Practical Implementation of the Scenario-Neutral Approach to Climate Impact Assessments for Hydrological Systems

by

Danlu Guo

Bachelor of Engineering Honours (Civil and Environmental)

Thesis submitted to The University of Adelaide,

Faculty of Engineering, Computer and Mathematical Sciences,

School of Civil, Environmental and Mining Engineering

in fulfilment of the requirements

for the degree of Doctor of Philosophy

Contents

Abstract

Understanding the impacts of climate change has particular significance for the future planning, design and operation of water resource systems. Scenario-neutral approaches are used increasingly to assess these possible impacts. These approaches allow water resource systems to be assessed independently of climate change projections, instead focusing on the sensitivity of specific systems to a large number of plausible climate change conditions. Once developed, these approaches can be used to better understand water resource system vulnerability, and provides a mechanism to incorporate multiple lines of evidence on possible future climatic changes into the climate impact assessment. The aim of this research therefore is to improve the practical implementation of scenario-neutral approaches by focusing on two key limitations: (1) limited capacity to generate a comprehensive climate exposure space to describe a large number of plausible future climate conditions; and (2) lack of understanding of how the physical process representation in rainfall-runoff models can affect future runoff projections.

The first limitation was addressed by developing an inverse approach to generate a climate exposure space that can represent a range of plausible future changes in rainfall and evapotranspiration. This is achieved by firstly identifying a set of climate variables (e.g. rainfall, temperature) and attributes of those variables (e.g. annual average, extremes) that might change in the future, and then modifying the parameters of a weather generator to perturb these variables and attributes within plausible ranges. This overcomes a long-standing problem in scenario-neutral studies, which have tended to focus only on a small subset of variables and attributes that might change in the future. The second limitation was addressed by examining the impact of alternative evapotranspiration representations within rainfall-runoff models, and assessing how this representation interacts with future evapotranspiration and runoff projections. The research showed that although the calibration and validation performance of alternative rainfall-runoff models may be similar under historical climate

conditions, the process representation can have a significant impact on future projections, highlighting the importance of model selection as part of the climate impact assessment process.

The outcomes of this research are demonstrated by implementing the enhanced scenario-neutral approach to a case study in Scott Creek catchment in South Australia. The results are used to show how different measures of runoff change as a function of different climate perturbations, and the climate variables most important for the system under plausible future climate conditions. This research therefore provides guidance for the future implementations of scenario-neutral framework, and thus greatly extends the applicability of this framework to a larger range of climate impact assessment problems.

Statement of Originality

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

Guo, D., Westra, S., Maier, H.R. 2016, An R Package for Modelling Actual, Potential and Reference Evapotranspiration, *Environmental Modelling and Software with Environment*. vol. 78, pp. 214-224. doi:10.1016/j.envsoft.2015.12.019.

Guo, D., Westra, S., Maier, H.R. 2016, An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies, *Journal of Hydrology*, Available online 22 March 2016, ISSN 0022-1694. doi:10.1016/j.jhydrol.2016.03.025.

Guo, D., Westra, S., and Maier, H. R. 2017. Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones, *Hydrology and Earth Systems Sciences*, 21, 2107-2126, doi:10.5194/hess-21-2107-2017, 2017.

Guo, D., Westra, S. & Maier, H. R. 2017. Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff

models. *Water Resources Research*, 53. 435-454, doi:10.1002/2016WR019627.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed: Date: 24/04/2017

Acknowledgements

First of all, I would like to express my deep appreciation to my supervisors, A./Prof. Seth Westra and Prof. Holger Maier, for their great effort in helping me through my PhD journey. Standing close to the finishing line, I am very proud of what we have achieved during these four years as a team. The pathway has not always been smooth, with countless challenges and unexpected incidents that could have been used as excuses to give up half-way. It would have been impossible for me to survive those hard times without the direction and encouragement from my supervisors. I am particularly grateful to Seth for his patient guidance on the technical aspects of my research, as well as on improving my communication skills. Moreover as an immature PhD candidate, the lessons I took from Seth on how to become an independent researcher have been, and will keep being, an invaluable treasure for my entire life. His words '*for science, it is important search for truth*' continuously remind me of the responsibility to deliver credible research to the community as a scientist. I am also grateful to Holger for his bright mind and deep insights, particularly on overseeing the big picture of our research and suggesting alternative approaches. Holger's generosity to share his extensive knowledge and research experience has greatly assisted us to derive practical solutions to a number of major difficulties we encountered. It has been an honourable pleasure to share the enthusiasm for research with Seth and Holger during my PhD, which motivates me to pursue a future career in research.

I would like to thank my mother, Yan Zhou, my father, Lei Guo, my partner, Xinyu Yang and other family and friends for their unconditional support and love during my candidature. My mother, who shared with me her life experience that *good things are worth the persistence*, has been a key motivation for me to keep this long-lasting commitment. My father and my partner are always standing behind me, and giving me the warmest encouragement at all times. Special thanks goes to my furry friend Tan Tan, who has always kept my morale positive with her cute (and sometimes annoying) disturbances when this research was being produced.

List of Figures

List of Tables

CHAPTER 1 Introduction

In modern society, water resource systems play a vital role in social, economic and environmental well-being. Water resource systems can be utilised for diverse objectives including water supply, irrigation, hydroelectricity and flood protection, and are commonly controlled by processes in the hydrological cycle such as precipitation, evapotranspiration and infiltration (IPCC, 2014, Dingman, 2015). These driving processes, when considered over the entire life of a water resource system, are likely to be influenced by potential changes in future climate, which can then impact the performance of the system in satisfying its objectives (Wilby and Dessai, 2010, Brown et al., 2012, Christensen and Lettenmaier, 2006, Wiley and Palmer, 2008). Some changes have already been observed in a range of climate features, such as the long-term averages and variability in temperature and precipitation (Westra et al., 2014a, Alexander et al., 2006, Easterling et al., 2000, Vincent et al., 2015), as well as extreme events such as droughts and floods (Petrow and Merz, 2009, Dai, 2011). Therefore, understanding the potential impacts of climate change is necessary for assessing the future capability for these systems, and is of particular significance for their future planning, design, operation and conservation practices.

To assess the potential climate change impact on water resource systems, information on future climatic and hydrological conditions is needed. Such information is based on our existing knowledge of relevant natural processes. Although this knowledge comes from an increasing number of sources of information that have been established from studies of climate science and hydrology (e.g. Chen et al., 2013, Ho et al., 2015, Ault et al., 2014, Vincent et al., 2015), the current use of this information for assessing climate change impact on water resource systems is far from comprehensive (Dessai and Hulme, 2004, Prudhomme et al., 2010). In most studies assessing the impact of climate change on water resource systems, a standard "top-down" or "scenario-led" framework has been employed. Under this framework, future projection(s) of the climate are first generated using global climate models (GCMs), followed by downscaling to obtain future regional climate conditions with regional climate models (RCMs) or statistical methods. The

resulting downscaled projections are then used to simulate future hydrological conditions through rainfall-runoff modelling (e.g. Crossman et al., 2013, Islam et al., 2014, Najafi and Moradkhani, 2015, Vaze and Teng, 2011), which can then provide recommendations for the future management of water resources (e.g. Crossman et al., 2013) as well as adaptation to extreme events under a changing climate (e.g. Kay et al., 2006).

Although the scenario-led approach has dominated the design of climate impact assessments for water resource systems, several limitations of this approach have been identified in recent literature, indicating that the relevance of scenario-led assessments to practical decision-making can be questionable (Brown et al., 2011, Yates et al., 2015, Prudhomme et al., 2010). As an important limitation, scenario-led assessments are often conducted in a one-size-fits-all manner, which relies heavily on the future climate conditions informed by global climate models (GCMs) and the associated regional climate models (RCMs), regardless of the specific water resource system under investigation (Brown et al., 2011; Prudhomme et al., 2010; Prudhomme et al., 2013). Different water resource systems are likely to be dominated by different physical processes across many spatial and temporal scales, and thus are influenced by different sets of key climatic and hydrologic variables and attributes (e.g. Jones et al. 2006, Donohue, Roderick and McVicar 2011). Therefore, the scenario-led studies may not adequately consider all plausible future conditions of the key hydro-climatic variables and attributes for individual water resource systems, while some climate conditions that are not considered may become tipping points that critically influence the performance of the specific systems (Prudhomme et al. 2013a, Poff et al. 2015, Brown et al. 2011).

An alternative way of assessing how the performance of water resource systems can be influenced by future climate change is the scenario-neutral approach. These approaches focus on assessing the sensitivity of the performance of individual water resources systems to a more comprehensive range of plausible climate change conditions, independent of climate change projections (Dessai and Hulme, 2004, Brown et al., 2012, Brown and Wilby,

2012, Nazemi and Wheater, 2014). The information generated from these approaches can be used to assess system vulnerability under alternative climate change scenarios, and to calculate climatic thresholds at which system performance begins to change abruptly (Brown et al., 2011, Poff et al., 2015). Furthermore, they can help to identify the important hydro-climatic variables and their attributes, as well as the critical states of these variables that affect the system under consideration. This latter feature is particularly useful as it allows a system-tailored study design centring on this particular water resource system and its key hydro-climatic variables. Such design can involve the selection of: (1) climate models that can predict these key variables with higher confidence (e.g. Flato et al., 2013, CSIRO and Bureau of Meteorology, 2015, Johnson et al., 2011); (2) strategies to downscale GCM-based projections that can better describe these key variables and attributes across different spatial and temporal scales (e.g. Johnson and Sharma, 2011); and (3) alternative sources of information (e.g. expert opinion and data from the paleo-climatic record) that can provide additional understanding of the possible variations in these key variables and attributes (e.g. Ault et al., 2014, Ho et al., 2015). Ultimately, this allows for the development of a more targeted set of projections for the water resource system of interest, by suggesting suitable models to predict the future changes in the key hydro-climatic variables and attributes for the system, as well as better approaches to link these changes to the water resource system (Vano et al., 2015, Steinschneider and Brown, 2013, Singh et al., 2014, Nazemi et al., 2013).

Practical implementations of scenario-neutral approaches vary with specific study objectives (e.g. to assess system vulnerability, to calculate critical climatic thresholds that affect system performance, or to identify the important hydro-climatic variables etc.). However, they generally include two core components, which are:

1) Defining the climate exposure space for a specific water resource system. This exposure space should consist of a range of plausible future changes in a number of hydro-climatic variables (e.g. rainfall,

temperature) and attributes of those variables (e.g. annual average, extremes) that are expected to impact catchment runoff;

2) Stressing the water resource system within the climate exposure space constructed in (1). This is often achieved via simulating the runoff responses to all the hydro-climatic conditions included in the exposure space using rainfall-runoff models.

Although the general principles underpinning scenario-neutral approaches have been well established (Dessai and Hulme, 2004, Nazemi and Wheater, 2014, Prudhomme et al., 2010, Brown et al., 2012), practical implementations are still scarce (Brown et al. 2012, Prudhomme et al. 2010, Prudhomme et al. 2013a, Poff et al. 2015, Kay, Crooks and Reynard 2014, Singh et al. 2014), within which several limitations have been identified, including:

1) **Design of the climate exposure space**

As a requirement of scenario-neutral approaches, a large number of plausible climate conditions should be used as part of the climate stress test of the water resource system under consideration. Ideally, this climate exposure space should include a number of hydro-climatic variables and attributes that are expected to change in the future, thus impacting the water resource system of interest (Prudhomme et al. 2010, Prudhomme et al. 2013a). However, current scenario-neutral studies have been mainly focused on a small number of hydro-climatic variables, with only changes in their long-term average values considered through applying annual or sub-annual change factors to historical data. Consequently, the resultant system sensitivity estimated may not show the full range of variability that can be expected in a greenhouse gas-enhanced climate (Prudhomme et al. 2013a, Steinschneider and Brown 2013).

2) **The choice of rainfall-runoff models to link climate exposure to the hydrological impact on specific water resource systems**

In current implementations of scenario-neutral approaches, there has been no formal selection process for the rainfall-runoff models to predict potential climate change on system hydrology. The selection of rainfall-runoff models has been largely based on their historical performance, without explicitly understanding their plausibility in representing relevant physical processes,

which can be particularly critical for runoff simulation under a changing climate (Kirchner, 2006). Within this issue, there is a particular lack of understanding of how climate impact assessments can be affected by alternative representation of evapotranspiration (ET) processes in different rainfall-runoff models. ET involves several complex sub-processes such as aerodynamic and advection processes, with varying influence of each sub-process with different local and hydro-climatic conditions such as aridity and evaporative surface (McMahon et al., 2013); furthermore, for modelling catchment water balance, an additional set of terrestrial hydrological processes also need to be considered to determine how precipitation is partitioned into ET, runoff and groundwater recharge (Oudin et al., 2005), which thus lead to the development of a number of alternative ET process representations (i.e. models). These varying process representations are often considered to have little impact on rainfall-runoff modelling under historical climate (Andréassian et al., 2004, Oudin et al., 2005b, Oudin et al., 2006). However, the role of ET process representation under a changing climate has not been investigated and thus remains unknown.

1.1 Research objectives

The overarching objective of this research is to improve the practical implementation of scenario-neutral climate impact assessment frameworks via addressing the abovementioned two limitations. This consists of the following specific objectives:

Objective 1: To improve the generation of the climate exposure space for scenario-neutral approaches by considering a range of plausible changes in a number hydro-climatic variables and their attributes, focusing on those that can influence rainfall and potential ET (PET), as they are the two key variables in linking large-scale climate change to local and regional water resources.

Objective 1.1: To develop a tool to estimate PET with alternative models and input data sets in a consistent manner. This tool allows PET estimation for

different case studies, as well as for different climate conditions included in the exposure space.

Objective 1.2: To understand the plausible changes in PET due to its driving climatic variables, as well as to identify the relative importance of these variables, under different hydro-climatic conditions. This understanding is particularly useful for the design of climate exposure spaces by informing which PET-related variables are important and should therefore be included in the exposure space.

Objective 1.3: To develop an approach to generate a climate exposure spaces consisting of a wide range of plausible changes in multiple rainfall attributes, including not only changes in annual average levels, but also shifts in variances, seasonal patterns and extremes.

Objective 2: To understand the role of ET process representations in rainfall-runoff modelling to propagate potential climate-induced changes to runoff projection, and thus to inform the selection of suitable rainfall-runoff models for representing ET under a changing climate.

Objective 3: To illustrate an implementation of the scenario-neutral approach with the abovementioned new knowledge and techniques on a case study water resource system, in order to identify the key climatic variables for the specific system.

1.2 Thesis overview

This thesis is organized into six chapters, where the main contributions are presented in Chapters 2 to 6. Each of these chapters is presented in the form of a technical paper. The way that these papers address the various objectives of the thesis is illustrated in Figure 1-1. Chapter 2 (Paper 1) has been published in Environmental Modelling & Software. Chapter 4 (Paper 3) has been published in the Journal of Hydrology. Chapter 5 (Paper 4) has been published in Water Resources Research. Chapters 3 (Paper 2) has been published in Hydrology and Earth System Sciences. Chapter 6 (Paper 5) provides an illustration of an integrated scenario-neutral approach using methods and

results obtained from previous chapters, which has been submitted to the Journal of Hydrology. Although the manuscripts have been reformatted in accordance with University guidelines, and sections renumbered for inclusion within this thesis, the material within these papers is otherwise presented herein as published. A copy for each of Paper 1, Paper 2, Paper 3 and Paper 4 "as published" is provided in Appendices 2B, 3B, 4B and 5B, respectively.

Chapter 2 introduces an R package to estimate actual, potential and reference ET using 17 well-known models. This package enables the implementation of alternative ET models with different data sets in a consistent manner (Objective 1.1). This tool facilitates the estimation of PET for different case studies, as well as for different climate conditions included in the exposure space, which were required in Chapters 3, 5 and 6.

Chapter 3 conducts a global sensitivity analysis to assess the implications of baseline climate conditions on the sensitivity of PET to a large range of plausible changes in four of its driving climatic variables, namely temperature (T), relative humidity (RH), solar radiation (R_s) and wind speed (u_z). The results obtained provide a comprehensive understanding of the plausible changes in PET due to these climate variables and the relative importance of these variables for affecting PET, across different hydro-climatic conditions (Objective 1.2). These results are particularly useful in selecting the hydro-climatic variables and attributes to consider in the design of the exposure space when implementing scenario-neutral approaches (Objective 1). Therefore, information on the relative importance of these climate variables was also utilised in the illustration of scenario-neutral approach in Chapter 6, for the design of the exposure space.

Chapter 4 proposes and demonstrates an optimization-based inverse approach to obtain hydro-climatic time series with uniform coverage across the possible ranges of change in multiple rainfall attributes (i.e. exposure space). This approach enables the generation of a wide range of plausible changes in rainfall (Objective 1.3), which is fundamental to the generation of the climate exposure spaces required for the implementation of scenario-neutral approaches (Objective 1). This approach was further adapted to

include a larger number of climatic variables and attributes (which are relevant to both rainfall and ET) to generate a more comprehensive climate exposure space in Chapter 6.

Chapter 5 explores the sensitivity of runoff to plausible future changes in PET using three conceptual rainfall-runoff models (GR4J, AWBM and IHACRES_CMD) and five catchments in climatologically different parts of Australia, with model performance evaluated via comparing the simulated ET with actual observations. This study improves our understanding of the role of ET process representations in rainfall-runoff modelling to propagate the potential climate-induced changes in runoff projection, and thus can be used to inform the selection of suitable rainfall-runoff models for representing ET in scenario-neutral climate impact assessments (Objective 2). Therefore, the model evaluation results were used to inform the selection of rainfall-runoff models in Chapter 6.

With these new understanding and approaches developed from the abovementioned chapters, **Chapter 6** illustrates a full implementation of the scenario-neutral approach to identify the key climatic variables for a specific water resource system (Objective 3).

Conclusions of the research within this thesis are provided in **Chapter 7**, which summarises: 1) the research contributions, 2) limitations and directions for further research.

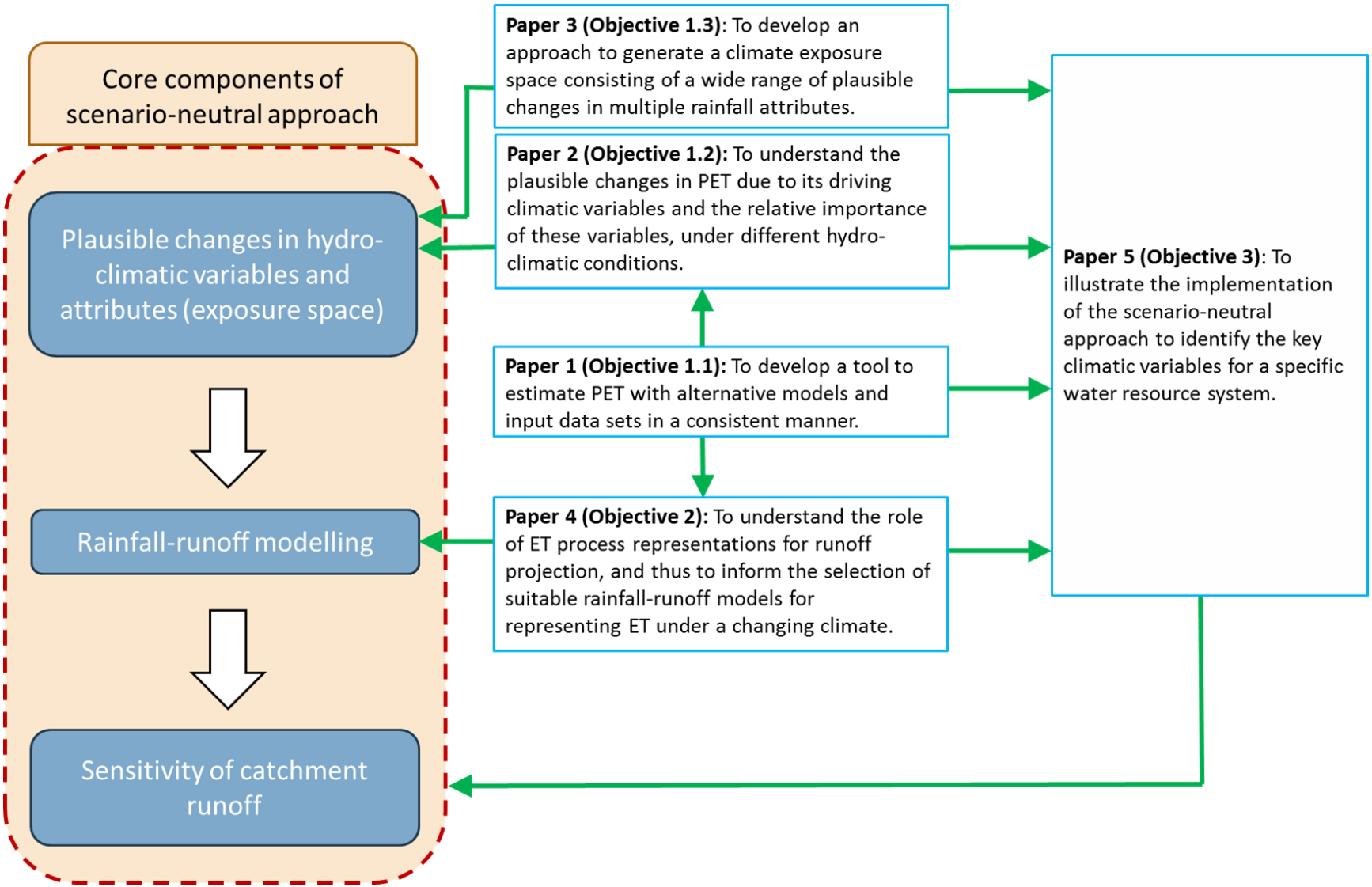


Figure 1-1: Linking of each of the papers to the objectives and the core components of the scenario-neutral approach.

CHAPTER 2 An R Package for Modelling Actual, Potential and Reference Evapotranspiration (Paper 1)

Statement of Authorship

itle of Paper	An R Package for Modelling Actual, Potential and Reference Evapotranspiration
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Guo, D., Westra, S. & Maier, H. R. (2016). An R package for modelling actual, potential and reference evapotranspiration. Environmental Modelling & Software, 78, 216-224.

Principal Author

Name of Principal Author (Candidate)	Danlu Guo
Contribution to the Paper	Designed the structures and features of R package, adapted models from literature to package, developed and maintained package to the latest version, wrote manuscript and acted as corresponding author.
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 19/12/2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Seth Westra
Contribution to the Paper	Suggested potential structures and features of R package, helped in literature and model interpretation, provided feedbacks during the evaluation of manuscript and the revision of journal paper.
Signature	Date 19/12/16

Name of Co-Author	Holger Maier
Contribution to the Paper	Motivated the development of package and journal paper, help to evaluate and edit the manuscript, as well as to develop responses to reviewers' comments.
Signature	Date 20/12/16

Please cut and paste additional co-author panels here as required.

Abstract

Evapotranspiration (ET) is a vital component of the hydrological cycle and there are a large number of alternative models for representing ET processes. However, implementing ET models in a consistent manner is difficult due to the significant diversity in process representations, assumptions, nomenclature, terminology, units and data requirements. An R package is therefore introduced to estimate actual, potential and reference ET using 17 well-known models. Data input is flexible, and customized data checking and pre-processing methods are provided. Results are presented as summary text and plots. Comparisons of alternative ET estimates can be visualized for multiple models, and alternative input data sets. The ET estimates also can be exported for further analysis, and used as input to rainfall-runoff models.

SOFTWARE AVAILABILITY

Description: Package *Evapotranspiration*

Developers: Danlu Guo, Seth Westra

Year First Available: 2014

E-mail: Danlu.guo@adelaide.edu.au

Website:http://cran.r-project.org/web/packages/Evapotranspiration/index.html

Hardware Requirement: General-purpose computer

Software Requirement: R version 2.10 or later

Programming Language: R

Licensing: This software is made freely available under the terms and conditions of the GNU General Public License

2.1 Introduction

Evapotranspiration (ET) is defined as the transfer of liquid water to the atmosphere as water vapor from bare soil and water bodies such as rivers and lakes (evaporation), as well as vegetated surfaces through plants' leaves (transpiration) (Allen et al., 1998, Dingman, 2015). ET is often one of the largest fluxes of water from catchments (Baumgartner et al., 1975), so that estimating its magnitude is critical for many applications. Factors that influence ET include: 1) the state of climate variables, such as temperature, relative humidity, wind speed and solar radiation, which influences the potential ET rate; 2) the water availability, which determines if actual evapotranspiration (AET) occurs at its potential rate (potential evapotranspiration, or PET) where sufficient water is present, or whether it occurs at a lower rate due to moisture limitations; and 3) the evaporative surface, with commonly modelled surfaces including natural catchments, 'reference' crops (ET_0), and open water bodies.

Understanding the dominant ET processes and quantifying ET rates provide useful information for diverse applications. For example, catchment management makes use of information on AET over the land surface, reservoir management requires information on open-water evaporation (e.g. McJannet et al., 2008), rainfall-runoff modelling often requires estimates of

catchment-averaged PET (e.g. Andréassian et al., 2004, Oudin et al., 2005a), and agricultural studies often require information on ET_0 (e.g. Shuttleworth and Wallace, 2009, Doorenbos, 1977). However, obtaining observations of these specific ET rates can be challenging. This is because the measurement of AET is difficult, typically involving sophisticated spatial and temporal scaling techniques from sap flow observations to represent the entire canopy, or using expensive micrometeorological eddy flux instrumentation that is generally not available for most practical applications; furthermore, PET is a conceptual quantity that cannot be 'measured' directly (Gasca-Tucker et al., 2007; Fisher et al., 2011). Therefore, these rates are usually estimated using models, so that the selection and implementation of ET process models becomes critical.

There are multiple models available for estimating ET rates. According to McMahon et al. (2013), alternative ET models can represent the same ET processes differently by: (1) placing emphasis on different sub-processes, such as mass transfer and energy balance processes; (2) focussing on the dominant processes that occur in different environments, including humid and arid climates; (3) having different requirements for inputting climate data and different interpretations of the constants' values; and (4) conforming to different hierarchies for handling missing data and adjusting biased estimates.

In order to provide better information on the selection of an appropriate model, guidance on ET model formulation and related issues was provided by McMahon et al. (2013). However, the implementation of these and other formulations is complicated by the significant diversity in process representations, assumptions, nomenclature, terminology, units and data requirements, which can make it difficult to implement the mathematical representations of these ET models, and can lead to coding inconsistencies and errors. This has a number of potentially negative implications on ET modelling studies, such as reducing confidence in the results presented, and providing difficulties for objectively comparing the results from different studies.

A practical aspect that can benefit from a more standardized approach to ET model implementation is the use of ensemble ET models. Applications of ensemble modelling can lead to a better understanding of ET model structural uncertainty (e.g. Beven and Freer, 2001; Duan et al., 2007; Kavetski and Fenicia, 2011; Velázquez et al., 2013), by:

1) assessing the impact of multiple ET models based on historical climate assumptions, to quantify PET and AET uncertainty (Tabari et al., 2013, Xu and Singh, 2000, Xu and Singh, 2002), and determine the effect of ET estimates on hydrologic modelling, water resource assessments (Horváth et al., 2010, Kannan et al., 2007, Oudin et al., 2005b, Yin and Brook, 1992, Rosenberry et al., 2007), ecological and agricultural studies (Fisher et al., 2011, Gasca-Tucker et al., 2007, Nichols et al., 2004), and;

2) assessing the impact of using multiple ET models under a changing climate, considering potential changes in both the ET-related processes and climate variables (Thompson et al., 2014, McKenney and Rosenberg, 1993, Kingston et al., 2009, Kay and Davies, 2008, Prudhomme and Williamson, 2013, Bormann, 2011, Donohue et al., 2010a).

To further support a range of ET modelling studies, there is a need to facilitate the implementation of different ET models in a convenient, consistent and efficient manner. There are some existing software packages focussing on specific ET modelling needs and aspects: such as the 'ET_0 Calculator' (Raes and Munoz, 2008) to calculate ET_0 using the FAO-56 Penman-Monteith model, the Fortran code 'Morton WREVAP' (McMahon et al., 2013) to implement the Morton ET models, and the R package 'SPEI' (Beguería et al., 2013), which includes multiple ET models and several drought indices to estimate the Standardized Precipitation-Evapotranspiration Index (SPEI). However, to our knowledge, there has not been a freely available tool which enables the implementation of a large number of alternative ET models in a consistent manner.

This paper presents an R software package to estimate ET from 17 alternative models: fifteen of the models are based on those summarize in McMahon et al. (2013), as well as the Jensen-Haise and the McGuinness-Bordne models, sourced from (Prudhomme and Williamson, 2013). These estimate a range of ET quantities (AET, PET and ET_0), take a range of climate processes and variables into account, and run at daily or monthly time-steps. Data input is flexible and data checking and pre-processing options are included. The availability of such a consistent software framework for implementing modelling approaches is important from the perspective of ensemble modelling, comparison among different models and data sets (for examples see Galelli et al., 2014, Dawson et al., 2007), as well as analysis of model and input uncertainty (Andrews et al., 2011, Leavesley et al., 2006, Clark et al., 2008).

The remainder of this paper is organized as follows. The package is described in Section 2.2, including the evapotranspiration models included, as well as the package structure and core functions. In Section 2.3, two different Australian catchments are used to demonstrate various features of the package including: (1) data pre-processing; (2) estimation of ET and producing summaries and plots of results; and (3) comparison of estimates with ensemble ET models and input data sets. In Section 2.4, some potential further analyses with the package and limitations are discussed, which are followed by the conclusions in Section 2.5.

2.2 The evapotranspiration package

2.2.1 Evapotranspiration models

The R package *Evapotranspiration* includes 17 models, which use one or several climate variables to estimate PET, AET and ET_0 at a single location using input data at sub-daily, daily and monthly resolutions. Although the models consist of different process representations, they are all based on the two fundamental components that drive ET:

1) Energy balance, which determines the latent heat of vaporization; and

2) Mass transfer, which influences the rate of movement of water vapor away from the evaporating surface.

The latent heat can be estimated considering the energy balance as:

$$\lambda E = R - H - G + A_d \quad \textbf{(2.1)}$$

where λ is the latent heat of vaporization, E is the rate of evapotranspiration, R is the net incoming radiation received at the soil/plant surfaces (which is determined by the total incoming solar radiation R_s), H is the sensible heat exchange with the atmosphere through convection (which is determined by the air temperature T), G is the heat exchange with the ground, and A_d is the net input of water advected energy, such as water inflow to a lake, which only applies for open-water bodies.

The mass transfer of water vapor is influenced by the vapor gradient (i.e. the difference between saturated and actual vapor pressure, which is related to relative humidity RH and temperature T) and wind speed u_z. Next to the evaporative surface, a thin non-turbulent layer of air provides resistance to evaporation flux, known as the aerodynamic resistance (Penman, 1948). For plant leaves, surface resistance is also important, as transpiration is regulated by the degree of stomatal opening in leaves (Monteith, 1991). Combining the energy balance and mass transfer components, the four key climate variables related to ET are T, RH, R_s and u_z (as illustrated in Figure 2-1).

Over the past decades, a large number of ET models have been developed by representing these processes in different ways. In this package, 17 of these models are included, which are based on different relationships among the ET processes and the four climate variables, and thus having different data requirements of climate variables and corresponding units (which are detailed in Table 2-1).

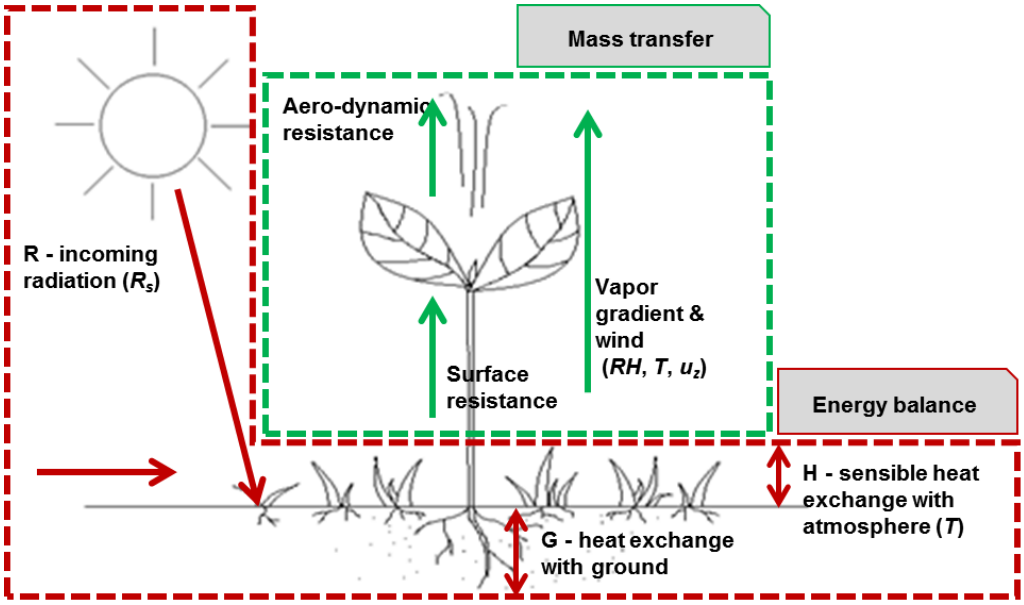


Figure 2-1: ET-related processes accounted for by the mass transfer and energy balance, with the relevant atmospheric variables in brackets: T = air temperature, R_s = incoming solar radiation, RH = relative humidity, u_z = wind speed.

The various models included in the package *Evapotranspiration* are detailed in Table 2-1. The PET and ET_0 models consider different sets of ET sub-processes and associated climate variables, including incoming radiation, vapor gradient, the heat exchanges with the atmosphere and the ground, advection processes and the surface resistance of vegetation (see the references in Table 2-1 for further details). The five AET models (i.e. Brutsaert-Strickler, Granger-Gray, Szilagyi-Jozsa, Morton CRAE and Morton CRWE) are all based on an observed complementary relationship (CR, first raised by Bouchet, 1963) between PET and AET, which states that as the evaporating surface dries, the decrease in AET is complemented by an equal increase in PET. The two Morton models (Morton, 1983b, Morton, 1983a) can estimate both the PET and AET explicitly at the equilibrium temperature (i.e. the temperature at the evaporating surface), by following the energy-balance and vapor transfer equations, respectively. Alternatively, the Brutsaert-Strickler, Granger-Gray and Szilagyi-Jozsa methods estimate AET by integrating the Penman and Priestley-Taylor models within the CR framework in different ways (Szilagyi, 2007, Granger and Gray, 1989, Brutsaert and Stricker, 1979). Note that these quantities are equivalent under special conditions: technically, when sufficient water is present, the rate of

PET and AET are equivalent to each other, and for a defined vegetated surface, the rate of PET and ET_0 are equivalent.

The equations for 15 ET models included in the package (all models except for Jensen-Haise and McGuinness-Bordne) are sourced from McMahon et al. (2013), which have all been verified with examples presented in their original paper. The availability of reliable verification is the key reason that we select the majority of ET models within this package from McMahon et al. (2013). For the other two structurally simple models, Jensen-Haise and McGuinness-Bordne, which are sourced from Prudhomme and Williamson (2013), there are no published examples of implementation available for verification. We have ensured that the equations are correct by verifying their formulae in a number of alternative references including Jensen and Haise (1963), Xu and Singh (2000) and Oudin et al. (2005b).

Table 2-1: Data requirements for different models. D = daily, M = monthly.

ET model name and corresponding function name in package	Time step	Climate input data required[1]					Quantity estimated		
		T_{max}, T_{min}	RH_{max}, RH_{min}	R_s	U_z	T_{dew}	*PET*	ET_0	*AET*
Penman 1948 (Penman, 1948) and Penman 1956 (Penman, 1956) *ET.Penman*	D	✓	✓	✓	✓		✓		✓(open water)
Penman-Monteith FAO-56 (Allen et al., 1998) and ASCE-EWRI (Allen et al., 2005) *ET.PenmanMonteith*	D	✓	✓	✓	✓			✓(short crop)	
Matt-Shuttleworth (Shuttleworth and Wallace, 2009) *ET.MattShuttleworth*	D	✓	✓	✓	✓			✓(well-watered)	
Priestley-Taylor (Priestley and Taylor, 1972) *ET.PriestleyTaylor*	D	✓	✓	✓			✓ (advection-free)		
PenPan[2](Rotstayn et al., 2006) *ET.PenPan*	D	✓	✓	✓	✓		✓		
Brutsaert-Strickler (Brutsaert and Stricker, 1979) *ET.BrutsaertStrickler*	D	✓	✓	✓	✓				✓ (areal)
Granger-Gray (Granger and Gray, 1989) *ET.GrangerGray*	D	✓	✓	✓	✓				✓ (areal)
Szilagyi-Jozsa (Szilagyi, 2007) *ET.SzilagyiJozsa*	D	✓	✓	✓	✓				✓
Makkink (De Bruin, 1981) *ET.Makkink*	D	✓		✓				✓	
Blaney-Criddle (Allen and Pruitt, 1986) *ET.BlaneyCriddle*	D	✓	✓	✓	✓			✓(well-watered)	
Turc (Turc, 1961) *ET.Turc*	D	✓		✓				✓	
Hargreaves-Samani (Hargreaves and Samani, 1985) *ET.HargreavesSamani*	D	✓						✓	
Chapman Australian[3](Chapman, 2003) *ET.ChapmanAustralian*	D	✓	✓	✓	✓		✓		
Jensen-Haise (Jensen and Haise, 1963, Prudhomme and Williamson, 2013, Xu and Singh, 2000) *ET.JensenHaise*	D	✓		✓			✓		
McGuiness-Bordne (Oudin et al., 2005b, Prudhomme and Williamson, 2013) *ET.McGuinnessBordne*	D	✓					✓		
Morton CRAE (Morton, 1983a) *ET.MortonCRAE*	M	✓		✓		✓	✓		✓
Morton CRWE (Morton, 1983b) *ET.MortonCRWE*	M	✓		✓		✓	✓		✓ (shallow lake)

***Note:**

1. T_{max}/T_{min} = maximum/minimum temperature (°C), R_s = incoming solar radiation ($MJ.m^{-2}$), RH_{max}/RH_{min} = maximum/minimum relative humidity (%), u_z = wind speed ($m.s^{-1}$), T_{dew} = dew point temperature (°C).

2. The original PenPan model estimates the actual evaporation from a Class-A Pan (i.e. a circular pan with diameter of 1.2 m and depth of 0.25 m, which is constructed of galvanised iron and supported on a

wooden frame at 30 to 50 mm above the ground). This rate of evaporation is closely related to the PET, so that is it possible to approximated PET from pan evaporation by adjustment using a pan coefficient (McMahon et al., 2013).

3. The original Chapman model (Chapman, 2003) uses only the measurements of Class-A Pan evaporation and is therefore fully empirical. However, in the *Evapotranspiration* package, it has been adapted to utilize the outputs of the PenPan model so it can be considered to capture the same set of ET sub-processes as the PenPan model.

2.2.2 Structure and core functions

The functions, data inputs and outputs, and graphical features of the package are summarized in Figure 2-2. The data pre-processing function *ReadInputs()* is developed for loading and processing sub-daily and daily raw climate data. The processed data are then ready to feed into the generic function *ET...()*, where each of the 17 different methods can be called by substituting the '…' by the function name (e.g. '*ET.Penman ()*' to call the Penman model). The function performs calculations for the relevant ET model and generates a calculation summary.

Having calculated the ET quantity, the function *ETPlot()* can then be called to plot the original estimates, as well as aggregations and averages at different time scales. Function *ETComparison()* facilitates comparison of results and visualization of uncertainties from using different models and/or different input data. Finally, *ETForcing()* enables the association between estimated ET and different climate variables to be plotted.

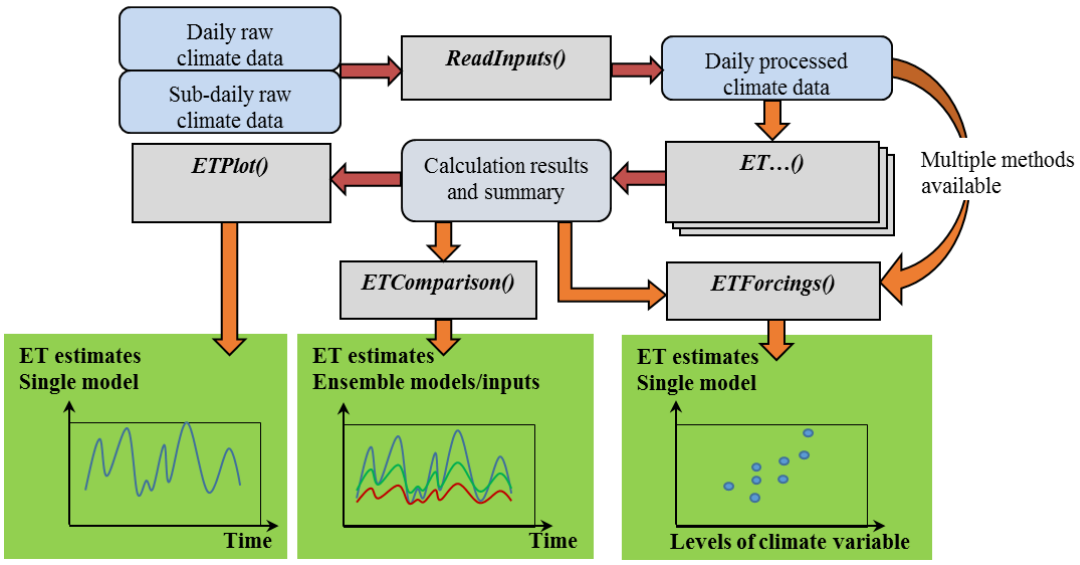


Figure 2-2: Schematic diagram of the features of the package *Evapotranspiration:* the blue boxes represent data or results that are produced and/or processed by the functions, represented in the grey boxes; the green boxes represent expected results.

Function *ReadInputs()* is designed for checking data availability, and identifies missing entries and errors from the input sub-daily or daily raw climate data. The availability of the date data (i.e. year, month and date) is checked first, since these data are compulsory for the function to read the time-series-like climate data. *ReadInputs*() then reads through the raw climate data presented, and reports all input variables that are available to use. Specific data requirements for the individual models (see Table 2-1) are checked prior to performing the calculations in function *ET…()*. A specific format of the input data is required in terms of variable names and units, as well as the input data file format, which is different for daily and sub-daily raw data (see Section 1.1 of the Appendix 2A, within which Table 2A.2 provides the detailed format requirements for the raw climate data). To assist users with preparing the raw input data, a summary of the relevant unit conversions is also provided in Table 2A.2 of the Appendix 2A.

Next, *ReadInputs()* checks for missing entries in each of the available climate variables, and the quality of the data is assessed against two user-defined threshold values for: (1) the maximum acceptable percentage of missing data; (2) the maximum acceptable duration of continuous missing data as a percentage of total data duration. If the data quality is not acceptable

(i.e. either of the percentage and/or duration of missing data has exceeded the user-defined threshold values), the program will be terminated with a warning message.

For data with acceptable quality but still containing some missing values, a warning is given with a default of assigning 'NA' for the missing values (which leads to 'NA's in the output estimates if they are used in *ET…()*). The user can also use the in-built gap-filling routine to interpolate for the missing values, with four alternative gap-filling methods (see Table 2A.3 in Appendix 2A for details) including:

1) Replacement with same-month average (adapted from Narapusetty et al., 2009);
2) Replacement with same-season average (adapted from Narapusetty et al., 2009);
3) Replacement with same day-of-the-year average (Narapusetty et al., 2009);
4) Interpolation between the two bounding values, which is only suitable for missing time increments in which values are available at adjacent increments (McMahon et al., 2013). When there is more than one consecutive missing entry, this interpolation fails, with a warning given.

The function also includes simple primary checks for abnormal values in each climate variable: for example, any temperature data greater than 100ºC are considered as abnormal. Warnings are issued for the abnormal values detected, and again, the users can choose if the abnormal values will be corrected in the function, using one of the four interpolation methods mentioned previously. Details of the four interpolation methods and definitions of abnormal values for each climate variable are presented in Tables 2A.4 and 2A.5 in Appendix 2A.

After completing the quality checks, all sub-daily raw data are aggregated to a daily time-step, as required by most ET models; such temporal

aggregation is not performed for raw climate data that are already available at a daily time-step.

As already discussed, *ET...()* is a generic function, which includes 17 different specific methods that are all named following the format of *ET.methodname()*, as detailed in Table 1-1. If a specific ET model is selected by user, the function first performs a specific check for the data requirement, which is different for each of the 17 models in *ET...()* (see Table 2-1 for details). If a certain input variable required by the ET model is not available, the function will search for whether there are alternative ways to estimate the missing variable from other available variables; however, if no alternative data or methods are available, the function will be terminated with a warning. The available methods to estimate missing input variables are summarized in Table 2A.3 in Appendix 2A.

In the case where a specific ET model is not specified (i.e. the generic function *ET()* is called directly instead of *ET.methodname()*), the first task *ET…()* performs is to estimate as many missing climate variables as possible. Then a default method to estimate ET is selected based on all the available input variables, which include both the original input variables presented, as well as the variables estimated from them. Wherever data are available, the ET model that has the highest data requirements (and thus provides the most detailed physically-based process representation) is selected as default. The detailed selection of default models for different data availability is in Table 2A.6 in Appendix 2A.

Besides the input climate data, a list of constants is also required in *ET…()*. The definitions and suggested values of all constants are summarized in Table 2A.7 in the Appendix 2A. A number of arguments are included in each ET model to allow additional user decisions in modelling ET. A common argument in all models is the choice of time-step for the output. The default time-step of the output ET estimates is daily for all models running at a daily time-step (i.e. all models except for Morton, as shown in Table 2-1), however, monthly and annual outputs can also be produced when specified; the two Morton models by default produce monthly output, while the user can also

choose to obtain annual output. For models with multiple versions (e.g. the Penman 1948 and 1956 models, which have different wind functions) and requiring additional user decisions (e.g. calculation options, assumptions) there are additional individual arguments to enable flexible choices among different pathways. The complete details on the use of constants and available arguments for different ET models are presented in Table 2A.8 in Appendix 2A.

Once being called with sufficient data, and provided all constants and arguments have been specified, function *ET...()* performs calculations for individual ET models. A user-friendly summary of the results is printed on the screen, which confirms the choice of model and sub-model, along with the corresponding versions, the quantities calculated, as well as options for alternative calculations and assumptions. A basic statistical summary of the entire output time-series is also presented (as illustrated with an example in Figure 2-3). The full results are stored as an R list file, as well as a csv file, which is automatically saved to the working directory. It contains both the calculation summary and the entire time series of the output, in which the ET estimates are organized in rows for different time increments.

A number of plotting tools are available to analyse the outputs. Function *ETPlot()* uses the estimated daily ET from individual ET models to generate aggregation plots and average plots at daily, monthly and annual time steps. Function *ETComparison()* produces comparison plots of different sets of ET estimates, to compare the outputs from (1) different ET models; (2) different versions of the same ET model (e.g. the 1948 and 1956 versions of the Penman model); (3) the same ET model with different calculation options, such as alternative approaches for data infilling and/or; (4) different sets of input climate data. For each quantity, three types of plots, including time series plots, non-exceedance probability plots and box plots, can be produced. Plots of uncertainty ranges can also be produced for daily estimates, monthly and annual aggregates and monthly and annual averages. Finally, the function *ETForcing()* is an additional plotting tool for visualizing the association between estimated ET and different climate variables within existing data.

2.3 Case studies

Two case studies have been used to demonstrate the core utilities of the package *Evapotranspiration*, using sub-daily climate data from meteorological sites at Adelaide (34.9290° S, 138.6010° E) and Alice Springs (23.7000° S, 133.8700° E) in Australia for the common period from 01/01/1989 to 30/03/2005.

2.3.1 Basic features: pre-processing input data, calculating and visualizing estimates

ReadInputs() is called first with the raw sub-daily data from the Adelaide case study, and the maximum percentage of acceptable number and duration of missing data set to 10% and 3%, respectively. The function displays a summary of data quality when checking through each input variable. The raw sub-daily data are then aggregated to a daily timescale. The missing values and abnormal values in each input variable are corrected with the corresponding averages from the same days of the year (i.e. day-of-the-year average). The processed data are then ready for the ET models to use.

The Penman open-water ET is estimated for the Adelaide case study using function *ET.Penman()*. The arguments are set so that (1) the time-step for calculation is daily; (2) the actual sunshine hours are used for calculating solar radiation; (3) the actual wind data are used; (4) the Penman 1948 wind function (Penman, 1948) is used to estimate the mass transfer component in the Penman model; and (5) the evaporative surface is open water (albedo = 0.08, roughness height = 0.001m). The calculated time series of Penman ET from *ET.Penman()* has been saved in an R data list, while output is printed to the screen, which confirms the choice of model and the selection of alternative calculation options, and also gives a basic statistical summary of the entire time-series of ET estimates.

Figure 2-3 is a screenshot of data processing and ET estimation with *ReadInputs()* and *ET.Penman()* for this case study.

```
> data <- ReadInputs(climatedata,constants,stopmissing=c(10,3), timestep = "subdaily",
+                    interp_missing = T, interp_abnormal = T,
+                    missing_method = "DoY average", abnormal_method = "DoY average")
The maximum acceptable percentage of missing data is 10 %
The maximum acceptable percentage of continuous missing data is 3 %
Warning: missing data of 'Tmax.daily'(daily maximum temperature), calculated from subdaily 'Temp.subdaily'
Warning: missing data of 'Tmin.daily'(daily minimum temperature), calculated from subdaily 'Temp.subdaily'
Warning: missing data of 'u2.subdaily', calculated from 'uz.subdaily
Number of missing values in uz.subdaily:  3
% missing data:  0.03 %
Maximum duration of missing data as percentage of total duration:  0.02 %
Warning: missing data of 'RHmax.daily'(daily maximum relative humidity), calculated from subdaily 'RH.subdaily'
Warning: missing data of 'RHmin.daily'(daily minimum relative humidity), calculated from subdaily 'RH.subdaily'
> Results <- ET.Penman(data, constants, solar="sunshine hours", wind="yes", windfunction_ver = "1948", alpha = 0.08, z0 = 0.001)
Penman Open-water Evaporation
Evaporative surface: water, albedo = 0.08 ; roughness height = 0.001 m
Sunshine hour data have been used for calculating incoming solar radiation
Wind data have been used for calculating the Penman evaporation. Penman 1948 wind function has been used.
Timestep: daily
Units: mm
1280 ET estimates obtained
Time duration: 2001-03-01 to 2004-08-31
Basic stats
Mean: 4.86
Max: 12.78
Min: 1.03
```

Figure 2-3: Example of a typical session of data processing with *ReadInputs()* and ET estimation with *ET.Penman()* for the Adelaide case study.

The plots of estimated daily ET and monthly averaged daily ET have been produced for the Adelaide case study using function *ETPlot()* (Figure 2-4). Although it is difficult to detect any trend from the highly fluctuating daily estimates (Figure 2-4a), there is a very strong seasonal pattern, displayed in the monthly average plot (Figure 2-4b). The ET peaks during the summer months, as would be expected due to the higher temperature and solar radiation during this time of the year.

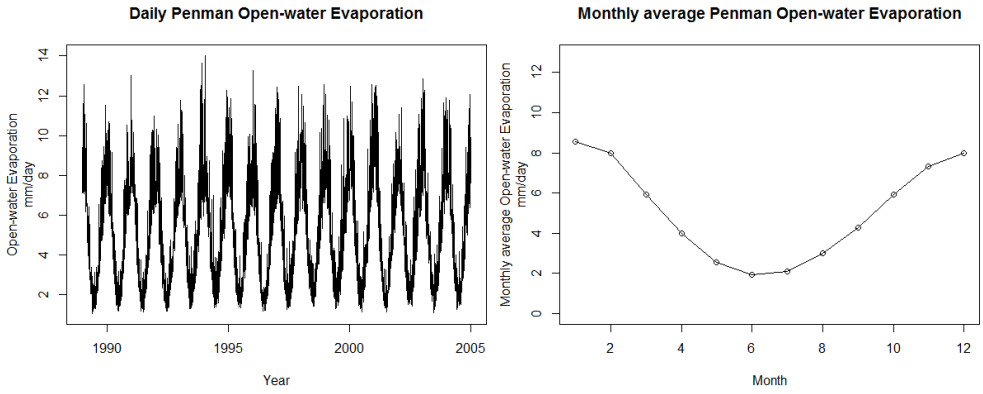


Figure 2-4: a) Daily estimates of Penman open-water ET (left panel); b) Monthly averaged daily Penman open-water ET (right panel) for the Adelaide case study, generated by *ETPlot()*.

2.3.2 Advanced features: analyses with ensemble models and different input data sets

The features of function *ETComparison()* are demonstrated for both the Adelaide and Alice Springs case studies. First, plots of the time series and the non-exceedance probabilities for monthly ET estimates have been produced to

compare estimates from the Penman-Monteith FAO56 and Priestley-Taylor models during 1989-1991 (Figure 2-5). From Figures 2-5a and 2-5b we observed that:

1) When comparing across the two case study sites, the inter-model differences in estimates are greater at Alice Springs, with the Priestley-Taylor model producing consistently lower estimates than the Penman-Monteith, and;
2) When comparing the seasonal patterns, the inter-model differences in estimates are most significant for the peak estimates, which occur in every summer (for example, at the start of 1990).

Figure 2-5c shows the distribution of the monthly estimates within the period and from the different models, which is consistent with previous observations: the ET estimates from the Priestley-Taylor model are consistently lower compared with those obtained using the Penman-Monteith model, with the greatest difference of approximately 100mm for the peak estimates at Alice Springs. These results reflect the structural differences in the two models, as the Penman-Monteith model explicitly takes the mass transfer for evapotranspiration into account, which is higher during summer periods and for arid and windy conditions (as reported in McKenney and Rosenberg, 1993; Yin et al., 2010), such as those experienced in Alice Springs.

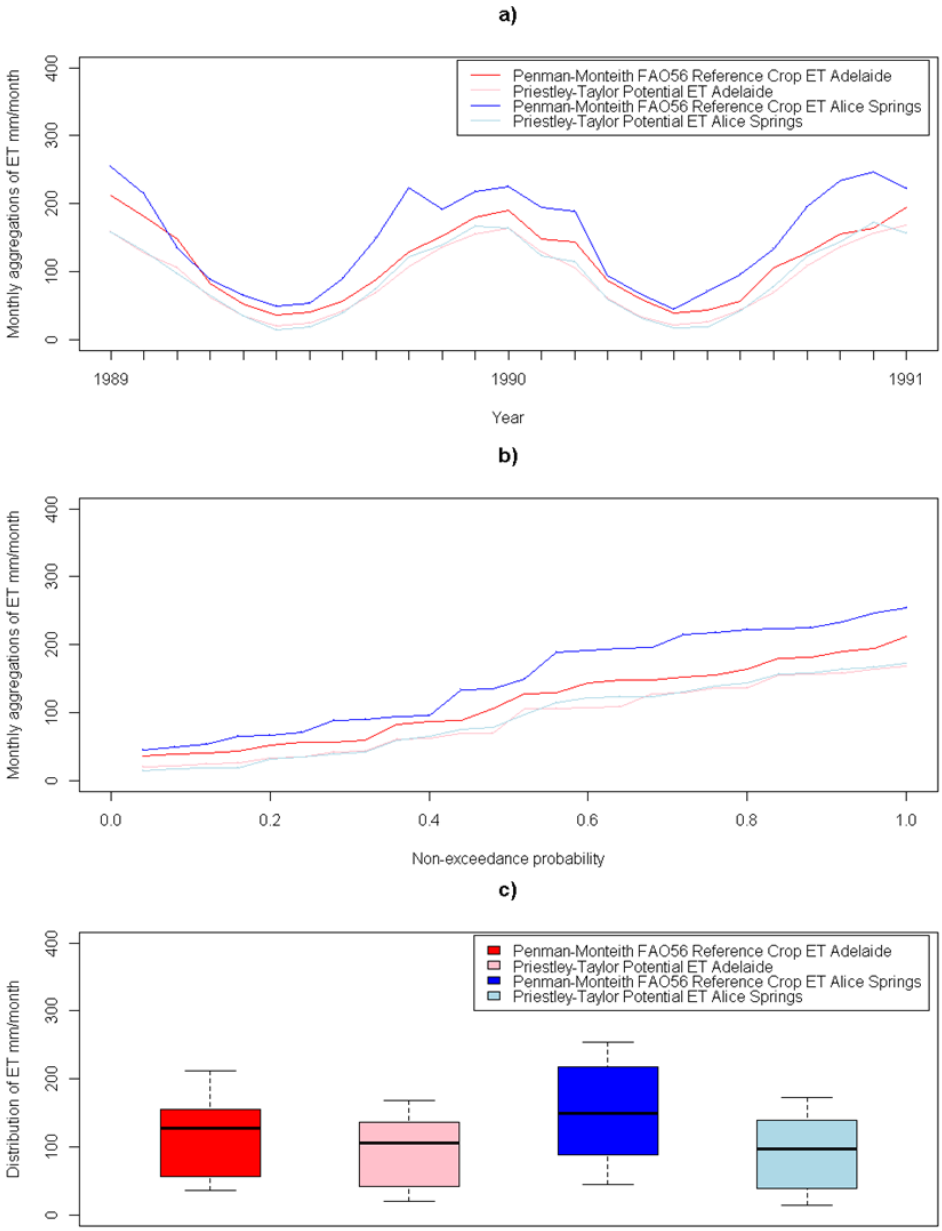


Figure 2-5: Comparison of monthly ET estimates from two models (Penman-Monteith FAO56 and Priestley-Taylor) and two locations (Adelaide and Alice Springs) using *ETComparison()* for a) time-series; b) non-exceedance probability, and; c) distribution.

Another application of *ETComparison()* is demonstrated in Figure 2-6, in which the effect of uncertainties in input climate data under climate changes are shown for the Adelaide case study, together with the model uncertainty. To maintain the simplicity and clarity of the example, we focus only on the potential uncertainties in the future temperature due to climate change, without considering the probability of individual changes or potential variations in other climate variables. We perturb the existing temperature data within a range of 0 to +8°C, which is considered to encompass all plausible

future changes in temperature in Australia by 2100 (Stocker et al., 2013). Within this range, 500 random samples have been drawn and the corresponding perturbations are applied to the historical time series of T_{max} and T_{min}, resulting in 500 sets of input climate data. These 500 sets were then fed into *ETComparison()* to generate the corresponding ET outputs using both the Penman-Monteith and Priestley-Taylor models. The resulting ranges of the monthly ET estimates for the period between 1989 and 1999 are shown in Figure 2-6. As can be seen, there is a greater range in ET estimates from the Penman-Monteith model. Since both ET models use the same temperature data as inputs, this indicates that temperature has a greater impact on the ET estimates obtained using the Penman-Monteith model—a pattern also observed in McKenney and Rosenberg (1993). This difference can be due to the structural differences between the two models: as the Penman-Monteith explicitly considers the mass transfer processes that are related to temperature, the importance of temperature is higher in the Penman-Monteith model.

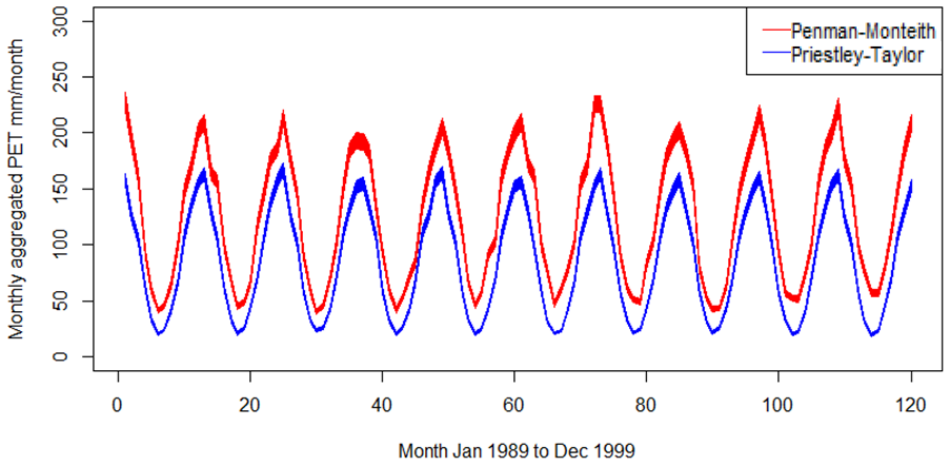


Figure 2-6: Uncertainties in monthly ET estimates from two models (Penman-Monteith FAO56 and Priestley-Taylor) at Adelaide, each executed for 500 sets of input data sampled with 0 to +8°C uncertainty in temperature, generated by *ETComparison()*.

It is worth mentioning that all climate variables other than temperature (i.e. *RH*, R_s and u_z) are kept at their current levels in this example, which is unrealistic under future conditions. Therefore, the results should only be considered as illustrative of the key feature of function *ETComparison()*, as a tool to compare ET estimates from multiple input data sets and ET models. In a formal assessment of the impact of climate-related input uncertainty, it is

necessary to consider the potential uncertainty in the full set of climate variables that influence ET (Whateley et al., 2014, Goyal, 2004).

2.4 Discussion

2.4.1 Further analyses with other software packages

The output from this package is formatted as time-series-like data in the *zoo* format (in which every data point is linked to a specific time point, see Zeileis et al., 2015), so it can be easily extracted and used as an input to other R-based software packages for a range of further analysis. Examples include using ET estimates as input to hydrologic models in *hydromad* (Andrews et al., 2011), and to investigate the sensitivity of ET estimates to changes in the input climate data using *sensitivity* (Pujol et al., 2014). Since the output is also saved to a csv file (as detailed in Section 2.2.2), it can also be imported to external software packages.

2.4.2 Limitations

Although the package provides features for checking missing values and errors in the input climate data, as well as interpolation methods for these problematic data, caution is required to minimize the risk of misuse. In developing this package, we have tested the data processing tools with our own test data sets, as well as a number of user-provided data sets, and we have ensured that the package runs free of errors with these existing data sets. However, since every data set is different, it is recommended that users should exercise their own quality-control procedure prior to using the package, to ensure that best-quality data are provided for ET estimation and the impact of data quality on the estimates is minimized.

Users should also be aware of the full assumptions and limitations prior to using any ET model in this package. Almost every ET model contains assumptions relating to the specific climate conditions under which the models apply. For example, some models assume that sub-processes related to ET are negligible, while other models are only calibrated to the climate of a specific region (a full list of assumptions and limitations for each individual

model is given in Table 2A.9 in Appendix 2A, which is summarized from the existing literature). These assumptions limit the models' ability to generalize to a wider range of climate zones, leading to varying performance of ET models under different climate settings (Rosenberry et al., 2007, Tabari et al., 2013). A further problem arises if the models are to be applied to estimate ET under climate change conditions, which can mean that existing ET processes and related climate variables are likely to be different to those for which the models are best suited, potentially causing deteriorating model performance (Prudhomme and Williamson, 2013, Thompson et al., 2014).

2.5 Summary and conclusions

This paper presents an R package *Evapotranspiration* for the estimation of actual, potential and reference crop ET using 17 models in a consistent, convenient and efficient manner. The pre-processing tool provides flexible methods for checking and processing raw input climate data, which are then fed into user-selected ET models. The presentation of results is in the form of both summary text and plots. Comparison between multiple ET models and input data sets is also supported. Estimates from the package can be conveniently extracted for further analysis, such as rainfall-runoff modelling and sensitivity analyses. It is hoped that this package will increase consistency in the results presented in ET studies, and increase our ability to investigate the impact of structural uncertainty in ET model formulations via the use of ensemble modelling.

2.6 Acknowledgements

The authors wish to thank Murray Peel and Thomas McMahon for their valuable comments on delineating the scope of the paper, and to Joseph Guillaume and an anonymous reviewer for their thoughtful comments on the manuscript.

Appendix 2A Supplementary to Chapter 2

1. *Preparing input data*

1.1 Name, unit and format requirements

The raw data to be used for function *ReadInputs()* should be saved in a csv file to be read in R, or directly as an R data frame, with every input variable in a column and every time step in a row. The input variables that can be processed within the package are listed in Table 2A.1, with corresponding naming and unit requirements. Some relevant unit conversion formulae are presented in Table 2A.2. Note that although the three variables defining the time of data records (i.e. year, month and day) are essential, it may not be necessary to supply all other input variables, as the specific data requirements vary across different ET models. Please refer to Table 2A.8 for details.

Table 2A.1. Supported input variables, and variable naming and unit requirements.

Input data variable	Required variable name	Units
Year	*Year*	-
Month	*Month*	-
Day	*Day*	-
Daily maximum temperature	*Tmax.daily*	°C
Daily minimum temperature	*Tmin.daily*	°C
Sub-daily temperature	*Temp.subdaily*	°C
Daily dew point temperature	*Tdew.daily*	°C
Sub-daily dew point temperature	*Tdew.subdaily*	°C
Daily maximum relative humidity	*RHmax.daily*	%
Daily minimum relative humidity	*RHmin.daily*	%
Sub-daily relative humidity	*RH.subdaily*	%
Daily incoming solar radiation	*Rs.daily*	$MJ.m^{-2}$
Sub-daily incoming solar radiation	*Rs.subdaily*	$MJ.m^{-2}$
Daily sunshine hours	*n.daily*	hour
Daily cloud cover	*Cd.daily*	Okta
Daily precipitation	*Precip.daily*	mm
Daily wind speed	*uz.daily*	$m.s^{-1}$
Daily wind speed at 2m	*u2.daily*	$m.s^{-1}$
Sub-daily wind speed	*uz.subdaily*	$m.s^{-1}$
Sub-daily wind speed at 2m	*u2.subdaily*	$m.s^{-1}$
Daily Class-A pan evaporation	*Epan.daily*	mm
Daily vapor pressure	*Vp.daily*	hPa
Sub-daily vapor pressure	*Vp.subdaily*	hPa

Table 2A.2. Unit conversion relevant to input variables.

Input data variable	Relevant unit conversions
Temperature	1 F = (F – 32) * 5/9 °C 1 K = K – 273.2 °C
Incoming solar radiation	1 $W.m^{-2}$= 10^{-6} $MJ.m^{-2}.s^{-1}$= 0.0036 $MJ.m^{-2}.h^{-1}$= 0.0864 $MJ.m^{-2}.day^{-1}$
Precipitation/evaporation	1 in = 25.4 mm
Wind speed	1 $km.h^{-1}$ = 0.2777 $m.s^{-1}$ 1 mph = 0.4470 $m.s^{-1}$ 1 $ft.s^{-1}$ = 0.3048 $m.s^{-1}$
vapor pressure	1 Pa = 0.01 hPa 1 bar = 100,000 Pa = 1000 hPa 1 atm = 101,325 Pa = 1013.25 hPa 1 mmHg = 133.3 Pa = 1.333 hPa

1.2 Alternative estimation methods for missing variables

In the situation where a climate variable is entirely missing, it may either be estimated via temporal aggregation from sub-daily data, or a specific model involving other climate variables. The methods available to estimate missing input variables are summarized in Table 2A.3.

Table 2A.3. Methods to estimate missing input variables from other variables, via either temporal aggregation or specific models.

Estimation via temporal aggregation		
Missing input variable	**Alternative input variable required**	
Daily maximum temperature	Sub-daily temperature	
Daily minimum temperature	Sub-daily temperature	
Daily maximum relative humidity	Sub-daily relative humidity	
Daily minimum relative humidity	Sub-daily relative humidity	
Daily incoming solar radiation	Sub-daily incoming solar radiation	
Daily wind speed	Sub-daily wind speed	
Daily wind speed at 2 meters	Sub-daily wind speed at 2 meters	
Daily dew point temperature	Sub-daily dew point temperature	
Estimation via models		
Missing input variable	**Alternative input variable required**	**Equation**
Daily dew point temperature (T_{dew}) in °C	Daily vapor pressure (v_a) in hPa	$T_{dew} = \frac{116.9 + 237.3\ln(v_a)}{16.78 - \ln(v_a)}$ (McJannet, 2008)
Sub-daily dew point temperature (T_{dew}) in °C	Sub-daily vapor pressure (v_a) in hPa	
Daily incoming solar radiation (R_s) in (three methods available)	Daily sunshine hours (n) in hours	$R_s = \left(a_s + b_s \frac{n}{N}\right) R_a$ where,

MJ.m^{-2}		a_s is the raction of extraterrestrial reaction reading earth on sunless days (dimensionless), which ideally require locally calibrated values (McMahon et al., 2013). See (Roderick, 1999) and also (McVicar et al., 2007) for regional examples for Australian applications. When no calibrated values are available, use: a_s = 0.25 (Allen et al., 1998) and; a_s = 0.23 for Australia (Roderick, 1999) $a_s + b_s$ is the fraction of extraterrestrial reaction reading earth on full-sun days (dimensionless), which ideally require locally calibrated values (McMahon et al., 2013). See (Roderick, 1999) and also (McVicar et al., 2007) for regional examples for Australian applications. When no calibrated values are available, use: b_s = 0.5 (Allen et al., 1998; Roderick, 1999) R_a is the extraterrestrial solar radiation in MJm^{-2}day^{-1}; *N* is the maximum possible duration of daylight hours in hours, $N = \frac{24}{\pi}\varpi_s$ And ϖ_s is the sunset hour angle in radians. (as in McMahon et al., 2013)
	Daily cloud covers (C_0) in Okta	$R_s = \left(a_s + b_s \frac{n}{N}\right) R_a$ $n = a_0 + b_0 C_0 + c_0 {C_0}^2 + d_0 {C_0}^3$ Where a_s, b_s, R_a and *N* are explained in the same way as above
	Monthly precipitation (P_j) in mm	$C_D = 1 + 0.5 log P_j + (log P_j)^2 \; where \; P_j \geq 1$ $C_D = 1 \; where P_j < 1$ (Linacre, 1993)
Daily wind speed at 2 meters (u_2) in m.s^{-1}	Daily wind speed (u_z) in m.s^{-1}	$u_2 = u_z \frac{\ln(\frac{z}{z_0})}{\ln(\frac{z}{z_0})}$ Where z_0 is the roughness height in meters (as in McMahon et al., 2013)

1.3 Definition for abnormal values

For quality control purposes, the package performs primary checks on the values of the climate data and generates warnings for any abnormal values detected. The 'abnormal' values for each climate variable are given in Table 2A.4.

Table 2A.4. Definitions of abnormal values for primary check of data quality.

Input data variable	Definition of abnormal values
Daily maximum temperature	> 100 °C
Daily minimum temperature	> Daily maximum temperature
Sub-daily temperature	> 100 °C
Daily dew point temperature	> 100 °C
Sub-daily dew point temperature	> 100 °C
Daily maximum relative humidity	> 100%
Daily minimum relative humidity	> Daily maximum relative humidity
Sub-daily relative humidity	> 100%
Daily incoming solar radiation	< 0 $MJ.m^{-2}$
Sub-daily incoming solar radiation	< 0 $MJ.m^{-2}$
Daily sunshine hours	< 0 hour
Daily cloud cover	< 0 Okta
Daily precipitation	< 0 mm
Daily wind speed	< 0 $m.s^{-1}$
Daily wind speed at 2m	< 0 $m.s^{-1}$
Sub-daily wind speed	< 0 $m.s^{-1}$
Sub-daily wind speed at 2m	< 0 $m.s^{-1}$
Daily Class-A pan evaporation	< 0 mm
Daily vapor pressure	< 0 hPa
Sub-daily vapor pressure	< 0 hPa

1.4 Interpolation methods available for missing/abnormal data entries

Wherever missing/abnormal entries are detected within the input climate data, the user can choose if these entries should be automatically filled/corrected. Four interpolation methods are available for gap-filling and correction of abnormal entries (Table 2A.5). Due to the scope of application of the package, only point-based temporal interpolation methods are included.

Table 2A.5. Interpolation methods available for gap-filling and correction of abnormal entries.

Interpolation method and corresponding arguments	Equation and explanation
Long-term mean of the same day of year (DoY) *"DoY average"*	The missing entry of variable *V*, at day *i* in year *j*, $V_{missing,ij}$, is approximated by the mean value of its values for the *i*th day from all other years. T is the total number of years data are available. $V_{missing,ij} = \frac{\sum_{y\in(1,2,\dots,j-1,j+1,\dots,T)} V_{avail.,iy}}{(T-1)}$ (Narapusetty et al., 2009)
Long-term mean of the same month *"monthly average"*	The missing entry of variable *V* at any day in month *k*, $V_{missing,k}$, is approximated by the mean value of all other available values from month *k*. $V_{missing,k} = \overline{V_{avail.,k}}$ (adapted from Narapusetty et al., 2009)

Long-term mean of the same season *"seasonal average"*	The missing entry of variable *V* at any day in season s $V_{missing,s}$, is approximated by the mean value of all other available values from season *s*. $V_{missing,s} = \overline{V_{avail.,s}}$ (adapted from Narapusetty et al., 2009)
Mean of the adjacent two data entries *"neighboring average"*	The missing entry of variable *V* at time *t* of the entire times series, is approximated by the average of its two adjacent entries. $V_{missing,t} = \frac{(V_{avail.,t-1} + V_{avail.,t+1})}{2}$ (McMahon et al., 2013)

2. Estimating ET

2.1 Default ET model when not specified

When the ET model selection is not specified by users, function *ET...()* determines the default model to use based on the availability of climate data presented. Wherever data are available, the more comprehensive, physically-based models are always preferred over the empirical models, as detailed in Table 2A.6.

Table 2A.6. Default ET models for different climate data availability.

Climate data available				**Default ET model**
T_{max},T_{min}	RH_{max},RH_{min}	R_s	U_z	
✓	✓	✓	✓	If short crop surface is specified in argument - > Penman-Monteith FAO56; If long crop surface is specified in argument - > Penman-Monteith ASCE-EWRI; If no surface is specified -> Penman.
✓	✓	✓		Priestley-Taylor
✓		✓		Makkink
✓				Hargreaves-Samani

2.2 Constants

The constants required to estimate ET consist of two parts: the universal constants are included in the data file '*defaultconstants.RData*' within the package, which should remain unchanged for most conditions; the case-specific constants should be prepared by individual users, which vary with case studies and are related to geographic locations and climatic/hydrologic conditions. Table 2A.7 provides details for all constants, as well as a summary of the default values for the universal constants, and recommended values for the case-specific constants.

Please note that not all constants are used for every single ET model – refer to Table 2A.8 for the list of required constants used for individual ET models. After obtaining all constants required, they should be compiled in a list variable named "*constants*" in R, which is ready to be used for the ET models.

Table 2A.7. Summary of constants required to estimate ET.

Universal constants		
Names	**Definitions**	**Default values**
lambda	Latent heat of vaporization ($MJ.kg^{-1}$)	2.45 at 20◦C
sigma	Stefan-Boltzmann constant ($MJ.m^{-2}.day^{-1}.◦K^{-4}$)	$4.903*10^{-9}$
Gsc	Solar constant ($MJ.m^{-2}.min^{-1}$)	0.0820
Roua	Mean density of air ($kg.m^{-3}$)	1.2 at 20◦C
Ca	Specific heat of air ($MJ.kg^{-1}.K^{-1}$)	0.001013
G	Soil heat flux ($MJ.m^{-2}$)	0, as this is assumed to be negligible for daily time-step (Allen et al., 1998)
alphaA	Albedo for Class-A pan (dimensionless)	0.14 (Rotstayn et al., 2006)
alphaPT	Priestley-Taylor coefficient (dimensionless)	Three default values depending on the model used: 1) 1.26 for Priestley-Taylor model (Priestley and Taylor, 1972) 2) 1.31 for Szilagyi-Jozsa model (Szilagyi, 2007) 3) 1.28 for Brutsaert-Strickler model (Brutsaert and Stricker, 1979)
ap	Constant in PenPan formula (dimensionless)	2.4
b0	Constant in Morton's procedure (dimensionless)	1 (Morton, 1983a)

b1	Constant in Morton's procedure ($W.m^{-2}$)	14 (Morton, 1983a)	
b2	Constant in Morton's procedure (dimensionless)	1.2 (Morton, 1983a)	
e0	Constant for Blaney-Criddle formula (dimensionless)	0.81917 (Frevert et al., 1983)	
e1	Constant for Blaney-Criddle formula (dimensionless)	-0.0040922 (Frevert et al., 1983)	
e2	Constant for Blaney-Criddle formula (dimensionless)	1.0705 (Frevert et al., 1983)	
e3	Constant for Blaney-Criddle formula (dimensionless)	0.065649 (Frevert et al., 1983)	
e4	Constant for Blaney-Criddle formula (dimensionless)	-0.0059864 (Frevert et al., 1983)	
e5	Constant for Blaney-Criddle formula (dimensionless)	-0.0005967 (Frevert et al., 1983)	
EpsilonMo	Land surface emissivity in Morton's procedure (dimensionless)	0.92 (Morton, 1983a)	
sigmaMo	Stefan-Boltzmann constant in Morton's procedure ($W.m^{-2}.K^{-4}$)	5.67e-08 (Morton, 1983a)	
Case-specific constants			
Names	**Definitions**	**Recommended values**	**Examples**
lat	Latitude (degree)	Use the actual latitude of case study	-34.9211 for Kent Town station, Adelaide
lat_rad	Latitude in radians (radians)	Use the actual latitude of case study	-0.6095 for Kent Town station, Adelaide
as	Fraction of extraterrestrial reaction reading earth on sunless days (dimensionless)	Ideally locally calibrated values should be used (McMahon et al., 2013). See (Roderick, 1999) and also (McVicar et al., 2007) for regional examples for Australian applications. When no calibrated values are available, use 0.25 (Allen et al., 1998).	0.23 for Australia (Roderick, 1999)
bs	*as* + *bs* is the fraction of extraterrestrial reaction reading earth on full-sun days (dimensionless)	Ideally locally calibrated values should be used (McMahon et al., 2013). See (Roderick, 1999) and also (McVicar et al., 2007) for regional examples for Australian applications. When no calibrated values are available, use 0.5 (Allen et al., 1998).	0.5 for Australia (Roderick, 1999)

Elev	Site elevation (m)	Use the actual elevation of study site	48 for Kent Town station, Adelaide
z	Wind instrument height (m)	Use the actual height of wind instrument or assume	Assumed to be 10
fz	Constant in Morton's procedure ($W.m^{-2}.mbar^{-1}$)	4 alternative values depending on the model used and the baseline temperature at study site: 1) 28.0 for CRAE model for $T >= 0◦C$ 2) 28.0*1.15 for CRAE model for $T < 0◦C$ 3) 25.0 for CRWE model for $T >= 0◦C$ 4) 28.75 for CRWE model for $T < 0◦C$ (Morton, 1983a, Morton, 1983b)	2 default values depending on the model used: 28.0 for CRAE model for Adelaide 25.0 for CRWE model for Adelaide
a_0	Constant for estimating sunshine hours from cloud cover data (dimensionless)	Depending on locations, see Chiew and McMahon (Chiew and McMahon, 1991) for Australian applications.	11.9 for Adelaide (Chiew and McMahon, 1991)
b_0	Constant for estimating sunshine hours from cloud cover data (dimensionless)	Depending on locations, see Chiew and McMahon (Chiew and McMahon, 1991) for Australian applications.	-0.15 for Adelaide (Chiew and McMahon, 1991)
c_0	Constant for estimating sunshine hours from cloud cover data (dimensionless)	Depending on locations, see Chiew and McMahon (Chiew and McMahon, 1991) for Australian applications.	-0.25 for Adelaide (Chiew and McMahon, 1991)
d_0	Constant for estimating sunshine hours from cloud cover data (dimensionless)	Depending on locations, see Chiew and McMahon (Chiew and McMahon, 1991) for Australian applications.	-0.0107 for Adelaide (Chiew and McMahon, 1991)
gammaps	Product of Psychrometric constant and atmospheric pressure as sea level ($mbar. ◦C^{-1}$)	2 alternative values depending on the model used and the baseline temperature at study site: 1) 0.66 for CRAE model for $T >= 0◦C$ 2) 0.66/1.15 for CRAE model for $T < 0◦C$ (Morton, 1983a)	0.66 for Adelaide
PA	Annual precipitation (mm)	Use the actual annual precipitation of case study	285.8 for Kent Town station, Adelaide
alphaMo	Constant in Morton's procedure (dimensionless)	2 alternative values depending on the baseline temperature at study site: 1) 17.27 when $T >= 0◦C$ 2) 21.88 when $T < 0◦C$ (Morton, 1983a)	17.27 for Adelaide
betaMo	Constant in Morton's procedure (◦C)	2 alternative values depending on the baseline temperature at study site: 1) 237.3 when $T >= 0◦C$	237.3 for Adelaide

		2) 265.5 when T < ◦C (Morton, 1983a)	
lambdaMo	Latent heat of vaporization in Morton's procedure (W.day.kg^{-1})	2 alternative values depending on the baseline temperature at study site: 1) 28.5 when T >= 0 ◦C 2) 28.5*1.15 when T < 0◦C (Morton, 1983a)	28.5 for Adelaide

2.3 ET models

The specifications of all ET models included in this package are summarized in Table 2A.8, including the names of corresponding functions, the available time steps for ET estimation, the ET quantities estimated, the requirements of input climate data and constants, and the actual equation and options available for calculation. The assumptions and known limitations of each ET model are summarized from existing literature and presented in Table 2A.9.

Table 2A.8. Specifications of available ET models within the package.

ET model name and corresponding function name in package	Time step	Quantity estimated				Climate input data required					Constants required[1]	Equation[2]	Calculation options with corresponding arguments
		PET	ET_0	AET	Pan ET	T_{max}, T_{min} (°C)	RH_{max}, RH_{min} (%)	R_s (MJ. m^{-2})	U_z (m. s^{-1})	T_{dew} (°C)			
Penman 1948 (Penman, 1948) and Penman 1956 (Penman, 1956) *ET.Penman*	D	✓		✓ (open water)		✓	✓	✓			*Elev, lambda (as λ in equation), lat_rad, Gsc, z, sigma;* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$ET = \frac{\Delta}{\Delta+\gamma}\frac{R_n}{\lambda}+\frac{\gamma}{\Delta+\gamma}E_\alpha$ Where Δ is the slope of vapor pressure curve in kPa°C^{-1}, $\Delta=\frac{4098[0.6108\exp(\frac{17.27*T_a}{T_a+237.3})]}{(T_a+237.3)^2}$ where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C; γ is the psychrometric constant in kPa°C^{-1}, $\gamma = 0.00163\frac{P}{\lambda}$ where P is the pressure at elevation z meters; R_n is the net incoming solar radiation at the evaporative surface in $MJm^{-2}day^{-1}$; E_a is the $E_\alpha = f(u)(v_a^* - v_a)$ in $mmday^{-1}$; $f(u) = 2.626 + 1.381u_2$ for Penman model version 1948, and; $f(u) = 1.313 + 1.381u_2$ for Penman model version 1956; v_a^* is the daily saturation vapor pressure in kPa, $v_a^* = \frac{v_a^*(T_{max})+v_a^*(T_{min})}{2}$ and $v_a^*(T_{max})$ and are the vapor pressures at temperatures T_{max} and T_{min} in °C; v_a is the mean daily actual vapor pressure in kPa, $v_a = 0.6108\exp\left[\frac{17.27T_d}{T_d+237.3}\right]$ and T_d is the dew point temperature in °C	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *wind* – if actual wind data is used in the model or alternative estimation method without wind data is used[3]. Default is to use wind data. *windfunction_ver* – if the 1948 version of the 1956 version of the Penman model is used. Default is 1948. *alpha* – user-defined albedo of the evaporative surface[4]. Default is 0.08 for open water. *z0* – user-defined roughness height of the evaporative surface. Default is 0.001 for open water.
Penman-Monteith FAO-56 (Allen et al., 1998) and ASCE-EWRI (Allen et al., 2005) *ET.PenmanMonteith*	D		✓ (short crop /tall crop			✓	✓	✓			*Elev, lambda (as λ in equation), lat_rad, Gsc, z, sigma, G;*	For FAO-56 model: $ET = \frac{0.408\Delta(R_n-G)+\gamma\frac{900}{T_a+273}u_2(v_a^*-v_a)}{\Delta+\gamma(1+0.34u_2)}\frac{R_{nw}}{\lambda}$ For ASCE model: $ET =$	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly

			)								In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$$\frac{0.408\Delta(R_n - G) + \gamma\frac{1600}{T_a + 273}u_2(v_a^* - v_a)}{\Delta + \gamma(1 + 0.38u_2)}$$ Where Δ is the slope of vapor pressure curve in kPa°C^{-1}, $\Delta = \frac{4098[0.6108\exp(\frac{17.27*T_a}{T_a+237.3})]}{(T_a+237.3)^2}$ where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C; R_n is the net incoming solar radiation at the evaporative surface in MJm^{-2}day^{-1}; γ is the psychrometric constant in kPa°C^{-1}, $\gamma = 0.00163\frac{P}{\lambda}$ where P is the pressure at elevation z meters; G is negligible for daily time step; v_a^* is the daily saturation vapor pressure in kPa, $v_a^* = \frac{v_a^*(T_{max})+v_a^*(T_{min})}{2}$ and $v_a^*(T_{max})$ and are the vapor pressures at temperatures T_{max} and T_{min} in °C; v_a is the mean daily actual vapor pressure in kPa, $v_a = 0.6108\exp\left[\frac{17.27T_d}{T_d+237.3}\right]$ and T_d is the dew point temperature in °C	from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *wind* – if actual wind data is used in the model or alternative estimation method without wind data is used[5]. Default is to use wind data. *crop* – user-defined crop type, can be either for short crop (results in use of FAO-56 version of this model) or tall crop (results in use of ASCE-EWRI version of this model). Default is for short crop.
Matt-Shuttleworth (Shuttleworth and Wallace, 2009) *ET.MattShuttleworth*	D		✓ (well-watered)			✓	✓	✓			*Elev, lambda (as λ in equation), lat_rad, Gsc, z, sigma, Roua (as* ρ_a in equation), *Ca (as* c_a *in equation);* In addition: *as, bs* – only If sunshine hours are	$$ET = \frac{1}{\lambda}\frac{\Delta R_n + \gamma\frac{\rho_a c_a u_2(VPD_2)}{r_c^{50}}(\frac{VPD_{50}}{VPD_2})}{\Delta + \gamma(1 + \frac{(r_s)_c u_2}{r_c^{50}})}$$ Where Δ is the slope of vapor pressure curve in kPa°C^{-1}, $\Delta = \frac{4098[0.6108\exp(\frac{17.27*T_a}{T_a+237.3})]}{(T_a+237.3)^2}$ where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C; γ is the psychrometric constant in kPa°C^{-1}, $\gamma = 0.00163\frac{P}{\lambda}$ where P is the pressure at elevation z meters; $(r_s)_c$ is the surface resistance of evaporative surface in sm^{-1}, which is a user-defined value through argument	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *alpha* – user-defined albedo of the evaporative surface. Default is 0.23 for short

											used to estimate solar radiation	*r_s*; $r_c^{50} = \frac{1}{0.41^2}\ln\left[\frac{(50-0.67h)}{(0.123h)}\right]\ln[\frac{(50-0.67h)}{(0.0123h)}]\frac{\ln[\frac{(2-0.08)}{0.0148}]}{\ln[\frac{(50-0.08)}{0.0148}]}$ and *h* is the crop height in meters; $\frac{VPD_{50}}{VPD_2} = \left(\frac{302(\Delta+\gamma)+70\gamma u_2}{208(\Delta+\gamma)+70\gamma u_2}\right) + \frac{1}{r_{clim}}[\left(\frac{302(\Delta+\gamma)+70\gamma u_2}{208(\Delta+\gamma)+70\gamma u_2}\right)\left(\frac{208}{u_2}\right) - (\frac{302}{u_2})]$ In which $r_{clim} = 86400\frac{\rho_a c_a (VPD)}{\Delta R_n}$; R_n is the net incoming solar radiation at the evaporative surface in $MJm^{-2}day^{-1}$	reference crop. *r_s* – user-defined surface resistance of the evaporative surface ($s.m^{-1}$). Default is 70 for short reference crop. *CH* – user-defined crop height (m) at the evaporative surface. Default is 0.12 for short reference crop.
Priestley-Taylor (Priestley and Taylor, 1972) *ET.PriestleyTaylor*	D	✓ (advection-free)				✓	✓	✓			*Elev, lambda (as λ in equation), lat_rad, Gsc, sigma, alphaPT (as α_{PT} in equation), G;* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$ET = \alpha_{PT} * [\frac{\Delta}{\Delta+\gamma}\frac{R_n}{\lambda} - \frac{G}{\lambda}]$ Where *Δ* is the slope of vapor pressure curve in kPa°C^{-1}, $\Delta = \frac{4098[0.6108\exp(\frac{17.27*T_a}{T_a+237.3})]}{(T_a+237.3)^2}$ where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C; *γ* is the psychrometric constant in kPa°C^{-1}, $\gamma = 0.00163\frac{P}{\lambda}$ where *P* is the pressure at elevation *z* meters; R_n is the net incoming solar radiation at the evaporative surface in $MJm^{-2}day^{-1}$; *G* is negligible for daily time step	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *alpha* – user-defined albedo of the evaporative surface. Default is 0.23 for short grass.
PenPan (Rotstayn et al., 2006) *ET.PenPan*	D	✓			✓	✓	✓	✓			*Elev, lambda (as λ in equation), lat_rad, Gsc, z, sigma, alpha, alpha (as α_p in*	$E_{PenPan} = \frac{\Delta}{\Delta+\alpha_p\gamma}\frac{R_{NPan}}{\lambda} + \frac{\alpha_p\gamma}{\Delta+\alpha_p\gamma}f_{Pan}(u)(v_a^* - v_a)$ Where *Δ* is the slope of vapor pressure curve in kPa°C^{-1}, $\Delta = \frac{4098[0.6108\exp(\frac{17.27*T_a}{T_a+237.3})]}{(T_a+237.3)^2}$ where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C;	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through

											equation), lat; In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	γ is the psychrometric constant in kPa°C^{-1}, $\gamma = 0.00163\frac{P}{\lambda}$ where P is the pressure at elevation z meters; R_{NPAN} is the net incoming solar radiation received at the Class-A Pan in MJm^{-2}day^{-1}; $f_{Pan}(u_2) = 1.201 + 1.621u_2$; v_a^* is the daily saturation vapor pressure in kPa, $v_a^* = \frac{v_a^*(T_{max})+v_a^*(T_{min})}{2}$ and $v_a^*(T_{max})$ and are the vapor pressures at temperatures T_{max} and T_{min} in °C; v_a is the mean daily actual vapor pressure in kPa, $v_a = 0.6108\exp\left[\frac{17.27T_d}{T_d+237.3}\right]$ and T_d is the dew point temperature in °C	estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *alpha* – user-defined albedo of the evaporative surface. Default is 0.23 for short grass. *est* – if the Pan evaporation or potential evaporation is estimated. Default is potential evaporation. *pan_coeff* – user-defined pan coefficient if potential evaporation is to be estimated. Default is 0.71 which is calibrated for Australia (McMahon et al., 2013). *overest* – if overestimation adjustment is made on the original results, which is divided by 1.078 (McMahon et al., 2013). Default is no adjustment.
Brutsaert-Strickler (Brutsaert and Stricker, 1979) *ET.BrutsaertStrickler*	D			✓ (are al)		✓	✓	✓			*Elev, lambda (as λ in equation), lat_rad, Gsc, z, sigma, alphaPT (as α_{PT} in equation);* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$ET = (2\alpha_{PT} - 1)\frac{\Delta}{\Delta+\gamma}\frac{R_n}{\lambda} - \frac{\gamma}{\Delta+\gamma}f(u_2)(v_a^* - v_a)$ Where Δ is the slope of vapor pressure curve in kPa°C^{-1}, $\Delta = \frac{4098[0.6108\exp(\frac{17.27*Ta}{Ta+237.3})]}{(T_a+237.3)^2}$ where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C; γ is the psychrometric constant in kPa°C^{-1}, $\gamma = 0.00163\frac{P}{\lambda}$ where P is the pressure at elevation z meters; R_n is the net incoming solar radiation at evaporative surface in MJm^{-2}day^{-1}; $f(u_2) = 2.626 + 1.381u_2$ v_a^* is the daily saturation vapor pressure in kPa, $v_a^* = \frac{v_a^*(T_{max})+v_a^*(T_{min})}{2}$ and $v_a^*(T_{max})$ and are the vapor	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *alpha* – user-defined albedo of the evaporative surface. Default is 0.23 for short grass.

												pressures at temperatures T_{max} and T_{min} in °C; v_a is the mean daily actual vapor pressure in kPa, $v_a = 0.6108 \exp\left[\frac{17.27T_d}{T_d+237.3}\right]$ and T_d is the dew point temperature in °C	
Granger-Gray (Granger and Gray, 1989) *ET.GrangerGray*	D			✓ (areal)		✓	✓	✓			*Elev, lambda (as λ in equation), lat_rad, Gsc, z, sigma, G;* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$ET = \frac{\Delta G_g}{\Delta G_g + \gamma}\frac{R_n - G}{\lambda} - \frac{\gamma G_g}{\Delta G_g + \gamma}E_a$ Where Δ is the slope of vapor pressure curve in kPa°C^{-1}, $\Delta = \frac{4098[0.6108\exp(\frac{17.27*T_a}{T_a+237.3})]}{(T_a+237.3)^2}$ where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C; γ is the psychrometric constant in kPa°C^{-1}, $\gamma = 0.00163\frac{P}{\lambda}$ where P is the pressure at elevation z meters; R_n is the net incoming solar radiation at evaporative surface in MJm^{-2}day^{-1}; G is negligible for daily time step; $G_g = \frac{1}{0.793+0.20e^{4.902Dp}} + 0.006D_p$ $D_p = \frac{E_a}{E_a + \frac{R_n - G}{\lambda}}$ E_a is the $E_\alpha = f(u)(v_a^* - v_a)$ in mmday^{-1}; $f(u) = 2.626 + 1.381u_2$ for Penman model version 1948, and; $f(u) = 1.313 + 1.381u_2$ for Penman model version 1956; v_a^* is the daily saturation vapor pressure in kPa, $v_a^* = \frac{v_a^*(T_{max})+v_a^*(T_{min})}{2}$ and $v_a^*(T_{max})$ and are the vapor pressures at temperatures T_{max} and T_{min} in °C; v_a is the mean daily actual vapor pressure in kPa, $v_a = 0.6108 \exp\left[\frac{17.27T_d}{T_d+237.3}\right]$ and T_d is the dew point temperature in °C	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *windfunction_ver* – if the 1948 version of the 1956 version of the Penman model is used. Default is 1948. *alpha* – user-defined albedo of the evaporative surface. Default is 0.23 for short grass.
Szilagyi-Jozsa (Szilagyi, 2007)	D			✓		✓	✓	✓			*Elev, lambda*	$ET =$	*ts* – time-step for ET

ET.SzilagyiJozsa											*(as λ in equation), lat_rad, Gsc, z, sigma, alphaPT;* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$2E_{PT}(T_e) - E_{Pen}$ Where, E_{PT} is Priestley-Taylor ET and E_{Pen} is Penman ET, T_e is the equilibrium temperature in °C, evaluated using $\frac{R_n}{\lambda E_{Pen}} = 1 + \frac{\gamma(T_e - T_a)}{(v_e^* - T_a)}$ Where v_e^* is the saturated vapor pressure at T_e; T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C.	estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *wind* – if actual wind data is used in the model or alternative estimation method without wind data is used *windfunction_ver* – if the 1948 version of the 1956 version of the Penman model is used. Default is 1948. *alpha* – user-defined albedo of the evaporative surface. Default is 0.23 for short grass. *z0* – user-defined roughness height of the evaporative surface (m). Default is 0.2 for short grass (McMahon et al., 2013).
Makkink (De Bruin, 1981) *ET.Makkink*	D		✓			✓		✓			*Elev, lambda (as λ in equation), lat_rad, Gsc;* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$ET = C_1\left(\frac{\Delta}{\Delta+\gamma}\frac{R_s}{2.45}\right) - C_2$ Where Δ is the slope of vapor pressure curve in kPa°C^{-1}, $\Delta = \frac{4098[0.6108\exp(\frac{17.27*T_a}{T_a+237.3})]}{(T_a+237.3)^2}$ where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C; γ is the psychrometric constant in kPa°C^{-1}, $\gamma = 0.00163\frac{P}{\lambda}$ where P is the pressure at elevation z meters; $C_1 = 0.61$ (dimensionless), $C_2 = 0.12$ mmday^{-1}	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours.
Blaney-Criddle (Allen and Pruitt,	D		✓			✓	✓	✓			*Elev, lambda*	$ET =$	*ts* – time-step for ET

1986) *ET.BlaneyCriddle*			(well-watered)								*(as λ in equation), lat_rad, Gsc, z, e0, e1, e2, e3, e4, e5;* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$\left(0.0043RH_{min} - \frac{n}{N} - 1.41\right) + b_{var}p_y$ Where, *N* is the maximum possible duration of daylight hours in hours, $N = \frac{24}{\pi}\varpi_s$ And ϖ_s is the sunset hour angle in radians; $b_{bar} = e_0 + e_1 RH_{min} + e_2\frac{n}{N} + e_3 u_2 + e_4 RH_{min}\frac{n}{N} + e_5 RH_{min} u_2$ $e_0 = 0.81917$ $e_1 = -0.0040922$ $e_2 = 1.0705$ $e_3 = 0.065649$ $e_4 = -0.0059684$ $e_5 = -0.005967$ p_y is the percentage of actual daytime hours for the day compared to the day-light hours for the entire year	estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *height* – if adjustment on the original results is made for the elevation of study site following: $ET_{adj} = ET\left[1 + 0.1\frac{Elev}{1000}\right]$ (Allen and Pruitt, 1986) Default is no adjustment.
Turc (Turc, 1961) *ET.Turc*	D		✓			✓		✓			*Elev, lambda (as λ in equation), lat_rad, Gsc;* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$ET = 0.013(23.88R_S + 50)\left(\frac{T_a}{T_a + 15}\right) + \left(1 + \frac{50 - RH}{70}\right)$ Where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C; *RH* is the average daily relative humidity calculated as (RH_{max} + RH_{min}) /2 in %.	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *humid* – if adjustment on the original results is made for low-humidity conditions where RH < 50, following: $ET_{adj} = ET\left[1 + \frac{50 - RH}{70}\right]$ (Alexandris et al., 2008) Default is no adjustment.
Hargreaves-Samani (Hargreaves and Samani, 1985) *ET.HargreavesSamani*	D		✓			✓					*Elev, lambda (as λ in equation), lat_rad, Gsc*	$ET = 0.0135C_{HS}\frac{R_a}{\lambda}(T_{max} - T_{min})^{0.5}(T_a + 17.8)$ Where, $C_{HS} = 0.00185(T_{max} - T_{min})^2 - 0.0433(T_{max} - T_{min}) +$	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily.

												0.4023; T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C.	
Chapman Australian (Chapman, 2001) *ET.ChapmanAustralian*	D	✓				✓	✓	✓			*Elev, lambda (as λ in equation), lat_rad, Gsc, lat, ap;* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$ET = (0.17 + 0.011Lat)E_{pan} + 10^{(0.66-0.211Lat)}$ Where Epan is the Class-A Pan evaporation in mmday^{-1}.	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *Penpan* – if Penpan model is used to estimate pan evaporation, or actual pan evaporation data is used. Default is to use Penpan model. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *alpha* – user-defined albedo of the evaporative surface. Default is 0.23 for short grass.
Jensen-Haise (Xu and Singh, 2000, Prudhomme and Williamson, 2013, Jensen and Haise, 1963) *ET.JensenHaise*	D	✓				✓		✓			*Elev, lambda (as λ in equation), lat_rad, Gsc;* In addition: *as, bs* – only If sunshine hours are used to estimate solar radiation	$ET = \frac{0.025}{\lambda}(R_s * (T_a + 3))$ Where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C.	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is daily. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours.
McGuiness-Bordne (Prudhomme and Williamson, 2013, Oudin et al., 2005b)	D	✓				✓					*Elev, lambda (as λ in equation),*	$ET = \frac{1}{68\lambda}(R_a * (T_a + 5))$	*ts* – time-step for ET estimation, can be daily, monthly or annual. Default is

ET.McGuinnessBordne											*lat_rad, Gsc*	Where T_a is the average daily temperature calculated as (T_{max} + T_{min}) /2 in °C; R_a is the extraterrestrial solar radiation in $MJm^{-2}day^{-1}$	daily.
Morton CRAE (Morton, 1983a) *ET.MortonCRAE*	M	✓		✓		✓		✓		✓	*lat_rad, epsilonMo (as ε_s in equation), fz (as f_z in quantion), b0 (as b_0 in quantion), b1, b2, lat, Elev, alphaMo, betaMo, sigmaMo (as σ in equation), Gsc, gammaps (as γp in equation), lambdaMo (as λ in equation)*	$$ET = \frac{1}{\lambda}(R_n - [\gamma p f_v + 4\varepsilon_s \sigma (T_e + 273)^3](T_e - T_a))$$ Where T_a is the mean air temperature in °C, as (T_{max} + T_{min}) /2; R_n is the net incoming solar radiation at the evaporative surface in Wm^{-2}; f_v is the vapor transfer coefficient in $Wm^{-2}mbar^{-1}$, $$f_v = (\frac{p_s}{p})^{0.5}\frac{f_z}{\xi}$$ Where p_s and p are the sea-level atmospheric pressure (in mbar) and at-site atmospheric pressure (in mbar) respectively; ξ is a stability factor (dimensionless) given as: $\xi = \frac{1}{0.28\left(1+\frac{v_D^*}{v_a^*}\right)+\frac{R_n\Delta}{\gamma p(\frac{p_s}{p})^{0.5}b_0 f_z(v_a^*-v_D^*)}}$ while ξ should always > 1; Δ is the slope of vapor pressure curve in $mbar°C^{-1}$ v_a^* is the saturation vapor pressure at air temperature (in mbar), v_D^* is the saturation vapor pressure at dew temperature T_{dew} (in mbar); T_e is the equilibrium temperature which can be estimated in °C, estimated iteratively by assuming a trial value T_e', and using: $\delta T_e = \frac{\frac{R_n}{f_v}+v_a^*-v_e^{*\prime}+\lambda_e(T_a-T_e')}{(\Delta_e'+\lambda_e)}$ and $$\lambda_e = \gamma p + \frac{4\varepsilon\sigma(T_e+273)^3}{f_v}$$	*ts* – time-step for ET estimation, can be monthly or annual. Default is monthly. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. est – ET quantity to be estimated, either for potential ET, wet areal ET or actual areal ET. Default is potential ET. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation). Default is using sunshine hours. *Tdew* – if actual Tdew data is used, or alternative estimation is made[6]. Default is to use data. alpha – only required when actual solar radiation data is used, to define albedo of the evaporative surface. Default is 0.23 for short grass.
Morton CRWE (Morton, 1983b) *ET.MortonCRWE*	M	✓		✓ (shallow lake)		✓		✓		✓	*lat_rad, epsilonMo (as ε_s in equation), fz (as f_z in quantion), b0 (as b_0 in*	Same equation as above (i.e. Morton CRAE) but with different values for constants – see Table 7 for details	*ts* – time-step for ET estimation, can be monthly or annual. Default is monthly. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine

											quantion), b1, b2, lat, Elev, alphaMo, betaMo, sigmaMo (as σ in equation), Gsc, gammaps (as γp in equation), lambdaMo (as λ in the equation)		hours, cloud cover or monthly precipitation). Default is using sunshine hours. est – ET quantity to be estimated, either for potential ET or shallow lake ET. Default is potential ET. *solar* – how solar radiation data is obtained (either directly from actual data, or through estimation using sunshine hours, cloud cover or monthly precipitation) *Tdew* – if actual Tdew data is used, or alternative estimation is made. Default is to use data. alpha – only required when actual solar radiation data is used, to define albedo of the evaporative surface. Default is 0.23 for short grass.

Note:

1. See Table D7 for interpretations and suggested values of the constants with corresponding names.

2. Only key steps for each ET model are summarised in the table, please refer to the corresponding reference for detailed calculations.

3. See Equation 33 in Valiantzas (2006) for the original equation.

4. See Table S3 in the supplementary material of McMahon et al. (2013) for a list of albedo values for different evaporative surfaces.

5. See Equation 39 in Valiantzas (2006) for the original equation.

6. See Table D3 in this document for the equation.

Table 2A.9. Assumptions and known limitations of each ET model.

ET model name	Assumptions	Known issues and limitations
Penman 1948/1956	Temperature of evaporative surface is unknown (McMahon et al., 2013). No heat storage or heat exchange with the ground, no advected energy, and hence the actual evaporation does not affect the overpassing air (Dingman, 2015).	Applies to most practical situations, but excludes extreme conditions, cases with high aerodynamic resistance, high humidity, and low temperature, which can lead to under-estimation by about 10% (Slatyer and McIlroy, 1961, Gao, 1988).
Penman-Monteith FAO-56/ASCE-EWRI	Temperature of evaporative surface is unknown (McMahon et al., 2013). G is negligible for daily time step (Allen et al., 1998). The FAO-56 model assumes an evaporative surface is covered with short grass of height 0.12m, with a surface resistance of 70s.m^{-1} and albedo of 0.23 (Allen et al., 1998). The ASCE-EWRI model assumes an evaporative surface is covered with short grass of height 0.5m, with a surface resistance of 45s.m^{-1} and albedo of 0.23 (Allen et al., 2005).	FAO-56 Reference Crop method should not be applied to the irrigation areas within semi-arid and windy regions, such as Australia (Shuttleworth and Wallace, 2009).
Matt-Shuttleworth	Evaporative surface is covered with short grass of height 0.12m, with a surface resistance of 70s.m^{-1} and albedo of 0.23 (Shuttleworth and Wallace, 2009). Evaporative surfaces are well-watered in semi-arid and windy areas (Shuttleworth and Wallace, 2009).	Only calibrated to irrigation (well-watered) areas that are semi-arid and windy (Shuttleworth and Wallace, 2009).
Priestley-Taylor	Evaporative surface is saturated (Priestley and Taylor, 1972).	Suitable for advection-free saturated surface. Performance can be improved by using seasonal-varying values of α_{PT} values depending on the seasons (Castellvi et al., 2001), evaporative surfaces as well as different observational periods (McMahon et al., 2013) Adjustments for vapour pressure deficit and available Energy should also be used to improve performance (Castellvi et al., 2001).
PenPan	Evaporative surface is a Class-A evaporation pan (Rotstayn et al., 2006).	The performance of the PenPan model in estimating Class-A pan evaporation are biased towards slightly higher values at lower evaporations (McMahon et al., 2013).
Brutsaert-Strickler	The actual ET and potential ET follow a Complementary Relationship (Brutsaert and Stricker, 1979).	The model sometimes generates negative ET values at daily time step (McMahon et al., 2013).
Granger-Gray	Evaporative surface is non-saturated lands (Granger and Gray, 1989).	No report in literature.
Szilagyi-Jozsa	The actual ET and potential ET follow a Complementary Relationship (Szilagyi, 2007).	Negative ET estimates can be obtained for days with very low net radiation (McMahon et al., 2013).
Makkink	Evaporative surface is covered by reference crop (De Bruin, 1981)	Calibrated only to cool climate conditions in the Netherlands (Xu and Singh, 2000).
Blaney-	Evaporative surface is covered with alfafa	This model should be used with caution in

Criddle	(Allen and Pruitt, 1986). Evaporative surface is within an adequately watered dry area where advection effects are strong (Allen and Pruitt, 1986, Yin and Brook, 1992).	equatorial regions (with relatively constant air temperature are relatively constant), small island and coastal areas (where air temperature is affected by sea temperature), high elevations (as environmental lapse rate induced low mean daily air temperature) and monsoonal and mid-latitude regions (as sunshine hours display large variety) (Doorenbos, 1977).
Turc	Evaporative surface is covered by reference crop(McMahon et al., 2013). Evaporative surface is within a humid region (Turc, 1961).	Adjustment may be used for non-humid conditions (*RH* < 50%) to improve performance (Alexandris et al., 2008).
Hargreaves-Samani	Evaporative surface is covered Alta fescue grass (Hargreaves and Samani, 1985).	Calibrated only to cool seasons in California (Xu and Singh, 2000).
Chapman Australian	Potential Evapotranspiration is well correlated with pan evaporation in Australia (Chapman, 2001).	Calibrated to Australia, so only applicable to Australia case studies (McMahon et al., 2013). Also only recommended if no other data than temperature are available to estimate ET (McMahon et al., 2013).
Jensen-Haise	Evaporative surface is within an adequately watered arid/semi-arid area (Jensen and Haise, 1963)..	Calibrated to arid/semi-arid regions in western US (Jensen and Haise, 1963).
McGuiness-Bordne	Evaporative surface is within a humid area (McGuinness and Bordne, 1972).	Calibrated only to Florida (Xu and Singh, 2000).
Morton CRAE	Vapour transfer is independent of wind speed (McMahon et al., 2013). The actual ET and potential ET follow a Complementary Relationship. Evaporative surface is a vegetated surface (Morton, 1983a).	Accurate measurements of humidity data are required (Morton, 1983a, McMahon et al., 2013). The model is not recommended for intervals of no longer than three days (Morton, 1983a). The model is not recommended for near edge conditions (e.g. edge of an oasis) (Morton, 1983a). Estimation for areal potential ET should only be used for large area (>1km^2) with unlimited water supply, while the potential ET should be used for a point, ideally a small irrigation area with unlimited water supply and surrounded by unirrigated area (Meteorology and Wang, 2001).
Morton CRWE	Vapour transfer is independent of wind speed (McMahon et al., 2013). The actual ET and potential ET follow a Complementary Relationship. Evaporative surface is a shallow lake-size water surface (Morton, 1983a).	The model is not recommended for deep lakes (>30m) (Morton, 1983a).

Appendix 2B Copy of Paper from Chapter 2

Guo, D., Westra, S., Maier, H.R. 2016, An R Package for Modelling Actual, Potential and Reference Evapotranspiration, *Environmental Modelling and Software with Environment*. vol. 78, pp. 214-224. doi:10.1016/j.envsoft.2015.12.019.

Environmental Modelling & Software 78 (2016) 216–224

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

An R package for modelling actual, potential and reference evapotranspiration

Danlu Guo*, Seth Westra, Holger R. Maier

School of Civil, Environmental & Mining Engineering, The University of Adelaide, Adelaide, 5005, Australia

ARTICLE INFO

Article history:
Received 1 June 2015
Received in revised form
31 December 2015
Accepted 31 December 2015
Available online 21 January 2016

Keywords:
Evapotranspiration (ET)
Evaporation
Ensemble modelling
R package
Evapotranspiration software

ABSTRACT

Evapotranspiration (ET) is a vital component of the hydrological cycle and there are a large number of alternative models for representing ET processes. However, implementing ET models in a consistent manner is difficult due to the significant diversity in process representations, assumptions, nomenclature, terminology, units and data requirements. An R package is therefore introduced to estimate actual, potential and reference ET using 17 well-known models. Data input is flexible, and customized data checking and pre-processing methods are provided. Results are presented as summary text and plots. Comparisons of alternative ET estimates can be visualized for multiple models, and alternative input data sets. The ET estimates also can be exported for further analysis, and used as input to rainfall-runoff models.

1. Introduction

Evapotranspiration (ET) is defined as the transfer of liquid water to the atmosphere as water vapor from bare soil and water bodies such as rivers and lakes (evaporation), as well as vegetated surfaces through plants' leaves (transpiration) (Allen et al., 1998; Dingman, 2015). ET is often one of the largest fluxes of water from catchments (Baumgartner et al., 1975), so that estimating its magnitude is critical for many applications. Factors that influence ET include: 1) the state of climate variables, such as temperature, relative humidity, wind speed and solar radiation, which influences the potential ET rate; 2) the water availability, which determines if actual evapotranspiration (AET) occurs at its potential rate (potential evapotranspiration, or PET) where sufficient water is present, or whether it occurs at a lower rate due to moisture limitations; and 3) the evaporative surface, with commonly modelled surfaces including natural catchments, 'reference' crops (ET_0), and open water bodies.

Understanding the dominant ET processes and quantifying ET rates provide useful information for diverse applications. For example, catchment management makes use of information on AET over the land surface, reservoir management requires information on open-water evaporation (e.g. McJannet et al., 2008), rainfall-runoff modelling often requires estimates of catchment-averaged PET (e.g. Andréassian et al., 2004; Oudin et al., 2005a,b), and agricultural studies often require information on ET_0 (e.g. Doorenbos, 1977; Shuttleworth and Wallace, 2009). However, obtaining observations of these specific ET rates can be challenging. This is because the measurement of AET is difficult, typically involving sophisticated spatial and temporal scaling techniques from sap flow observations to represent the entire canopy, or using expensive micrometeorological eddy flux instrumentation that is generally not available for most practical applications; furthermore, PET is a conceptual quantity that cannot be 'measured' directly (Gasca-Tucker et al., 2007; Fisher et al., 2011). Therefore, these rates are usually estimated using models, so that the selection and implementation of ET process models becomes critical.

There are multiple models available for estimating ET rates. According to McMahon et al. (2013), alternative ET models can represent the same ET processes differently by: (1) placing emphasis on different sub-processes, such as mass transfer and energy balance processes; (2) focussing on the dominant processes that occur in different environments, including humid and arid climates; (3) having different requirements for inputting climate data and different interpretations of the constants' values; and (4) conforming to different hierarchies for handling missing data and adjusting biased estimates.

In order to provide better information on the selection of an

* Corresponding author.
E-mail address: Danlu.Guo@adelaide.edu.au (D. Guo).

http://dx.doi.org/10.1016/j.envsoft.2015.12.019
1364-8152/

appropriate model, guidance on ET model formulation and related issues was provided by McMahon et al. (2013). However, the implementation of these and other formulations is complicated by the significant diversity in process representations, assumptions, nomenclature, terminology, units and data requirements, which can make it difficult to implement the mathematical representations of these ET models, and can lead to coding inconsistencies and errors. This has a number of potentially negative implications on ET modelling studies, such as reducing confidence in the results presented, and providing difficulties for objectively comparing the results from different studies.

A practical aspect that can benefit from a more standardized approach to ET model implementation is the use of ensemble ET models. Applications of ensemble modelling can lead to a better understanding of ET model structural uncertainty (e.g. Beven and Freer, 2001; Duan et al., 2007; Kavetski and Fenicia, 2011; Velázquez et al., 2012), by:

1) assessing the impact of multiple ET models based on historical climate assumptions, to quantify PET and AET uncertainty (Xu and Singh, 2000, 2002; Tabari et al., 2013), and determine the effect of ET estimates on hydrologic modelling, water resource assessments (Yin and Brook, 1992; Oudin et al., 2005a,b; Kannan et al., 2007; Rosenberry et al., 2007; Horváth et al., 2010), ecological and agricultural studies (Nichols et al., 2004; Gasca-Tucker et al., 2007; Fisher et al., 2011), and;
2) assessing the impact of using multiple ET models under a changing climate, considering potential changes in both the ET-related processes and climate variables (McKenney and Rosenberg, 1993; Kay and Davies, 2008; Kingston et al., 2009; Donohue et al., 2010; Bormann, 2011; Prudhomme and Williamson, 2013; Thompson et al., 2014).

To further support a range of ET modelling studies, there is a need to facilitate the implementation of different ET models in a convenient, consistent and efficient manner. There are some existing software packages focussing on specific ET modelling needs and aspects: such as the 'ET_0 Calculator' (Raes and Munoz, 2008) to calculate ET_0 using the FAO-56 Penman-Monteith model, the Fortran code 'Morton WREVAP' (McMahon et al., 2013) to implement the Morton ET models, and the R package 'SPEI' (Beguería et al., 2013), which includes multiple ET models and several drought indices to estimate the Standardized Precipitation-Evapotranspiration Index (SPEI). However, to our knowledge, there has not been a freely available tool which enables the implementation of a large number of alternative ET models in a consistent manner.

This paper presents an R software package to estimate ET from 17 alternative models: fifteen of the models are based on those summarize in McMahon et al. (2013), as well as the Jensen-Haise and the McGuinness-Bordne models, sourced from Prudhomme and Williamson (2013). These estimate a range of ET quantities (AET, PET and ET_0), take a range of climate processes and variables into account, and run at daily or monthly time-steps. Data input is flexible and data checking and pre-processing options are included. The availability of such a consistent software framework for implementing modelling approaches is important from the perspective of ensemble modelling, comparison among different models and data sets (for examples see Dawson et al., 2007; Galelli et al., 2014), as well as analysis of model and input uncertainty (Leavesley et al., 2006; Clark et al., 2008; Andrews et al., 2011).

The remainder of this paper is organized as follows. The package is described in Section 2, including the evapotranspiration models included, as well as the package structure and core functions. In Section 3, two different Australian catchments are used to demonstrate various features of the package including: (1) data pre-processing; (2) estimation of ET and producing summaries and plots of results; and (3) comparison of estimates with ensemble ET models and input data sets. In Section 4, some potential further analyses with the package and limitations are discussed, which are followed by the conclusions in Section 5.

2. The evapotranspiration package

2.1. Evapotranspiration models

The R package *Evapotranspiration* includes 17 models, which use one or several climate variables to estimate PET, AET and ET_0 at a single location using input data at sub-daily, daily and monthly resolutions. Although the models consist of different process representations, they are all based on the two fundamental components that drive ET:

1) Energy balance, which determines the latent heat of vaporization; and
2) Mass transfer, which influences the rate of movement of water vapor away from the evaporating surface.

The latent heat can be estimated considering the energy balance as:

$$\lambda E = R - H - G + A_d \quad (1)$$

where λ is the latent heat of vaporization, E is the rate of evapotranspiration, R is the net incoming radiation received at the soil/plant surfaces (which is determined by the total incoming solar radiation R_s), H is the sensible heat exchange with the atmosphere through convection (which is determined by the air temperature T), G is the heat exchange with the ground, and A_d is the net input of water advected energy, such as water inflow to a lake, which only applies for open-water bodies.

The mass transfer of water vapor is influenced by the vapor gradient (i.e. the difference between saturated and actual vapor pressure, which is related to relative humidity RH and temperature T) and wind speed u_z. Next to the evaporative surface, a thin non-turbulent layer of air provides resistance to evaporation flux, known as the aerodynamic resistance (Penman, 1948). For plant leaves, surface resistance is also important, as transpiration is regulated by the degree of stomatal opening in leaves (Monteith, 1991). Combining the energy balance and mass transfer components, the four key climate variables related to ET are T, RH, R_s and u_z

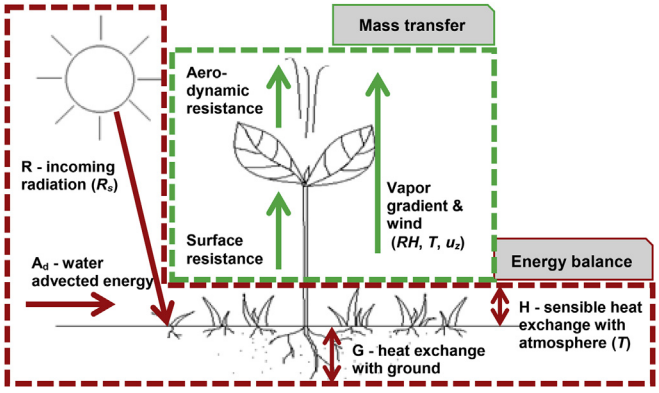


Fig. 1. ET-related processes accounted for by the mass transfer and energy balance, with the relevant atmospheric variables in brackets: T = air temperature, R_s = incoming solar radiation, RH = relative humidity, u_z = wind speed.

(as illustrated in Fig. 1).

Over the past decades, a large number of ET models have been developed by representing these processes in different ways. In this package, 17 of these models are included, which are based on different relationships among the ET processes and the four climate variables, and thus having different data requirements of climate variables and corresponding units (which are detailed in Table 1).

The various models included in the package *Evapotranspiration* are detailed in Table 1. The PET and ET_0 models consider different sets of ET sub-processes and associated climate variables, including incoming radiation, vapor gradient, the heat exchanges with the atmosphere and the ground, advection processes and the surface resistance of vegetation (see the references in Table 1 for further details). The five AET models (i.e. Brutsaert-Strickler, Granger-Gray, Szilagyi-Jozsa, Morton CRAE and Morton CRWE) are all based on an observed complementary relationship (CR, first raised by Bouchet, 1963) between PET and AET, which states that as the evaporating surface dries, the decrease in AET is complemented by an equal increase in PET. The two Morton models (Morton, 1983b, 1983a) can estimate both the PET and AET explicitly at the equilibrium temperature (i.e. the temperature at the evaporating surface), by following the energy-balance and vapor transfer equations, respectively. Alternatively, the Brutsaert-Strickler, Granger-Gray and Szilagyi-Jozsa methods estimate AET by integrating the Penman and Priestley-Taylor models within the CR framework in different ways (Brutsaert and Stricker, 1979; Granger and Gray, 1989; Szilagyi, 2007). Note that these quantities are equivalent under special conditions: technically, when sufficient water is present, the rate of PET and AET are equivalent to each other, and for a defined vegetated surface, the rate of PET and ET_0 are equivalent.

The equations for 15 ET models included in the package (all models except for Jensen-Haise and McGuinness-Bordne) are sourced from McMahon et al. (2013), which have all been verified with examples presented in their original paper. The availability of reliable verification is the key reason that we select the majority of ET models within this package from McMahon et al. (2013). For the other two structurally simple models, Jensen-Haise and McGuinness-Bordne which are sourced from Prudhomme and Williamson (2013), there are no published examples of implementation available for verification. We have ensured that the equations are correct by verifying their formulae in a number of alternative references including Jensen and Haise (1963), Xu and Singh (2000) and Oudin et al. (2005a).

2.2. Structure and core functions

The functions, data inputs and outputs, and graphical features of the package are summarized in Fig. 2. The data pre-processing function *ReadInputs()* is developed for loading and processing sub-daily and daily raw climate data. The processed data are then ready to feed into the generic function *ET…()*, where each of the 17 different methods can be called by substituting the '…' by the function name (e.g. '*ET.Penman ()*' to call the Penman model). The function performs calculations for the relevant ET model and generates a calculation summary.

Having calculated the ET quantity, the function *ETPlot()* can then be called to plot the original estimates, as well as aggregations and averages at different time scales. Function *ETComparison()* facilitates comparison of results and visualization of uncertainties from using different models and/or different input data. Finally, *ETForcing()* enables the association between estimated ET and different climate variables to be plotted.

Table 1
Data requirements for different models. D = daily, M = monthly.

ET model name and corresponding function name in package	Time step	Climate input data required[a]					Quantity estimated		
		T_{max}, T_{min}	RH_{max}, RH_{min}	R_s	U_z	T_{dew}	*PET*	ET_0	*AET*
Penman 1948 (Penman, 1948) and Penman 1956 (Penman, 1956) *ET.Penman*	D	✓	✓	✓	✓		✓		✓ (open water)
Penman-Monteith FAO-56 (Allen et al., 1998) and ASCE-EWRI (Allen et al., 2005) *ET.PenmanMonteith*	D	✓	✓	✓	✓			✓ (short crop)	
Matt-Shuttleworth (Shuttleworth and Wallace, 2009) *ET.MattShuttleworth*	D	✓	✓	✓	✓			✓ (well-watered)	
Priestley-Taylor (Priestley and Taylor, 1972) *ET.PriestleyTaylor*	D	✓	✓	✓			✓ (advection-free)		
PenPan[b] (Rotstayn et al., 2006) *ET.PenPan*	D	✓	✓	✓	✓		✓		
Brutsaert-Strickler (Brutsaert and Stricker, 1979) *ET.BrutsaertStrickler*	D	✓	✓	✓	✓				✓ (areal)
Granger-Gray (Granger and Gray, 1989) *ET.GrangerGray*	D	✓	✓	✓	✓				✓ (areal)
Szilagyi-Jozsa (Szilagyi, 2007) *ET.SzilagyiJozsa*	D	✓	✓	✓	✓				✓
Makkink (De Bruin, 1981)*ET.Makkink*	D	✓		✓				✓	
Blaney-Criddle (Allen and Pruitt, 1986) *ET.BlaneyCriddle*	D	✓	✓	✓	✓			✓ (well-watered)	
Turc (Turc, 1961) *ET.Turc*	D	✓		✓				✓	
Hargreaves-Samani (Hargreaves and Samani, 1985) *ET.HargreavesSamani*	D	✓						✓	
Chapman Australian[c] (Chapman, 2001) *ET.ChapmanAustralian*	D	✓	✓	✓	✓		✓		
Jensen-Haise (Jensen and Haise, 1963; Xu and Singh, 2000; Prudhomme and Williamson, 2013) *ET.JensenHaise*	D	✓		✓			✓		
McGuiness-Bordne (Oudin et al., 2005a; Prudhomme and Williamson, 2013) *ET.McGuinnessBordne*	D	✓					✓		
Morton CRAE (Morton, 1983a) *ET.MortonCRAE*	M	✓		✓		✓	✓		✓
Morton CRWE (Morton, 1983b) *ET.MortonCRWE*	M	✓		✓		✓	✓		✓ (shallow lake)

[a] T_{max}/T_{min} = maximum/minimum temperature (°C), R_s = incoming solar radiation ($MJ.m^{-2}$), RH_{max}/RH_{min} = maximum/minimum relative humidity (%), u_z = wind speed ($m.s^{-1}$), T_{dew} = dew point temperature (°C).

[b] The original PenPan model estimates the actual evaporation from a Class-A Pan (i.e. a circular pan with diameter of 1.2 m and depth of 0.25 m, which is constructed of galvanised iron and supported on a wooden frame at 30–50 mm above the ground). This rate of evaporation is closely related to the PET, so that is it possible to approximate PET from pan evaporation by adjustment using a pan coefficient (McMahon et al., 2013).

[c] The original Chapman model (Chapman, 2001) uses only the measurements of Class-A Pan evaporation and is therefore fully empirical. However, in the *Evapotranspiration* package, it has been adapted to utilize the outputs of the PenPan model so it can be considered to capture the same set of ET sub-processes as the PenPan model.

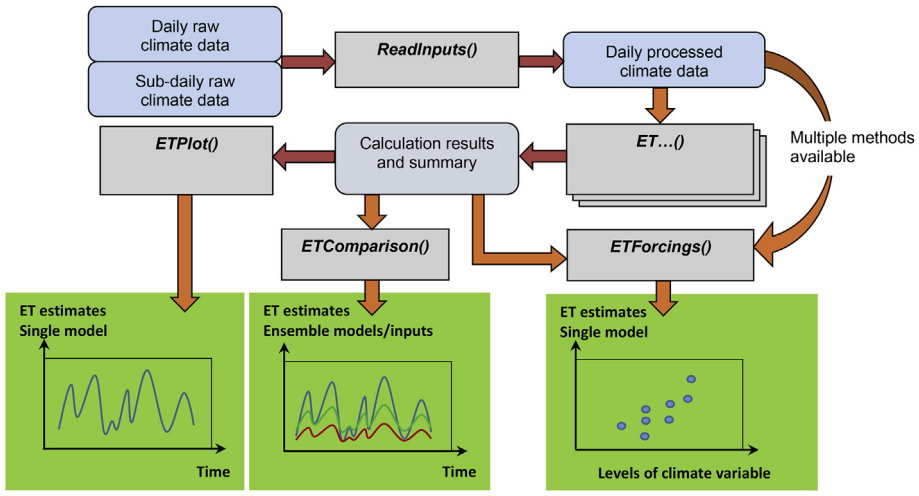


Fig. 2. Schematic diagram of the features of the package *Evapotranspiration:* the blue boxes represent data or results that are produced and/or processed by the functions, represented in the grey boxes; the green boxes represent expected results. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Function *ReadInputs()* is designed for checking data availability, and identifies missing entries and errors from the input sub-daily or daily raw climate data. The availability of the date data (i.e. year, month and date) is checked first, since these data are compulsory for the function to read the time-series-like climate data. *ReadInputs*() then reads through the raw climate data presented, and reports all input variables that are available to use. Specific data requirements for the individual models (see Table 1) are checked prior to performing the calculations in function *ET…()*. A specific format of the input data is required in terms of variable names and units, as well as the input data file format, which is different for daily and sub-daily raw data (see Section 1.1 of the supplementary material, within which Table 1 provides the detailed format requirements for the raw climate data). To assist users with preparing the raw input data, a summary of the relevant unit conversions is also provided in Table 2 of the supplementary material.

Next, *ReadInputs()* checks for missing entries in each of the available climate variables, and the quality of the data is assessed against two user-defined threshold values for: (1) the maximum acceptable percentage of missing data; (2) the maximum acceptable duration of continuous missing data as a percentage of total data duration. If the data quality is not acceptable (i.e. either of the percentage and/or duration of missing data has exceeded the user-defined threshold values), the program will be terminated with a warning message.

For data with acceptable quality but still containing some missing values, a warning is given with a default of assigning 'NA' for the missing values (which leads to 'NA's in the output estimates if they are used in *ET…()*). The user can also use the in-built gap-filling routine to interpolate for the missing values, with four alternative gap-filling methods (see Table 3 in the supplementary material for details) including:

1) Replacement with same-month average (adapted from Narapusetty et al., 2009);
2) Replacement with same-season average (adapted from Narapusetty et al., 2009);
3) Replacement with same day-of-the-year average (Narapusetty et al., 2009);
4) Interpolation between the two bounding values, which is only suitable for missing time increments in which values are available at adjacent increments (McMahon et al., 2013). When there is more than one consecutive missing entry, this interpolation fails, with a warning given.

The function also includes simple primary checks for abnormal values in each climate variable: for example, any temperature data greater than 100 °C are considered as abnormal. Warnings are issued for the abnormal values detected, and again, the users can choose if the abnormal values will be corrected in the function, using one of the four interpolation methods mentioned previously. Details of the four interpolation methods and definitions of abnormal values for each climate variable are presented in Tables 4 and 5 in the supplementary material.

After completing the quality checks, all sub-daily raw data are aggregated to a daily time-step, as required by most ET models; such temporal aggregation is not performed for raw climate data that are already available at a daily time-step.

As already discussed, *ET…()* is a generic function, which includes 17 different specific methods that are all named following the format of *ET.methodname()*, as detailed in Table 1. If a specific ET model is selected by user, the function first performs a specific check for the data requirement, which is different for each of the 17 models in *ET…()* (see Table 1 for details). If a certain input variable required by the ET model is not available, the function will search for whether there are alternative ways to estimate the missing variable from other available variables; however, if no alternative data or methods are available, the function will be terminated with a warning. The available methods to estimate missing input variables are summarized in Table 3 in the supplementary material.

In the case where a specific ET model is not specified (i.e. the generic function *ET()* is called directly instead of *ET.methodname()*), the first task *ET…()* performs is to estimate as many missing climate variables as possible. Then a default method to estimate ET is selected based on all the available input variables, which include both the original input variables presented, as well as the variables estimated from them. Wherever data are available, the ET model that has the highest data requirements (and thus provides the most detailed physically-based process representation) is selected as default. The detailed selection of default models for different data availability is in Table 6 of the supplementary material.

Besides the input climate data, a list of constants is also required in *ET…()*. The definitions and suggested values of all constants are

summarized in Table 7 in the supplementary material. A number of arguments are included in each ET model to allow additional user decisions in modelling ET. A common argument in all models is the choice of time-step for the output. The default time-step of the output ET estimates is daily for all models running at a daily time-step (i.e. all models except for Morton, as shown in Table 1), however, monthly and annual outputs can also be produced when specified; the two Morton models by default produce monthly output, while the user can also choose to obtain annual output. For models with multiple versions (e.g. the Penman, 1948, 1956 models, which have different wind functions) and requiring additional user decisions (e.g. calculation options, assumptions) there are additional individual arguments to enable flexible choices among different pathways. The complete details on the use of constants and available arguments for different ET models are presented in Table 8 in the supplementary material.

Once being called with sufficient data, and provided all constants and arguments have been specified, function *ET...()* performs calculations for individual ET models. A user-friendly summary of the results is printed on the screen, which confirms the choice of model and sub-model, along with the corresponding versions, the quantities calculated, as well as options for alternative calculations and assumptions. A basic statistical summary of the entire output time-series is also presented (as illustrated with an example in Fig. 3). The full results are stored as an R list file, as well as a csv file, which is automatically saved to the working directory. It contains both the calculation summary and the entire time series of the output, in which the ET estimates are organized in rows for different time increments.

A number of plotting tools are available to analyse the outputs. Function *ETPlot()* uses the estimated daily ET from individual ET models to generate aggregation plots and average plots at daily, monthly and annual time steps. Function *ETComparison()* produces comparison plots of different sets of ET estimates, to compare the outputs from (1) different ET models; (2) different versions of the same ET model (e.g. the 1948 and 1956 versions of the Penman model); (3) the same ET model with different calculation options, such as alternative approaches for data infilling and/or; (4) different sets of input climate data. For each quantity, three types of plots, including time series plots, non-exceedance probability plots and box plots, can be produced. Plots of uncertainty ranges can also be produced for daily estimates, monthly and annual aggregates and monthly and annual averages. Finally, the function *ETForcing()* is an additional plotting tool for visualizing the association between estimated ET and different climate variables within existing data.

3. Case studies

Two case studies have been used to demonstrate the core utilities of the package *Evapotranspiration*, using sub-daily climate data from meteorological sites at Adelaide (34.9290° S, 138.6010° E) and Alice Springs (23.7000° S, 133.8700° E) in Australia for the common period from 01/01/1989 to 30/03/2005.

3.1. Basic features: pre-processing input data, calculating and visualizing estimates

ReadInputs() is called first with the raw sub-daily data from the Adelaide case study, and the maximum percentage of acceptable number and duration of missing data set to 10% and 3%, respectively. The function displays a summary of data quality when checking through each input variable. The raw sub-daily data are then aggregated to a daily timescale. The missing values and abnormal values in each input variable are corrected with the corresponding averages from the same days of the year (i.e. day-of-the-year average). The processed data are then ready for the ET models to use.

The Penman open-water ET is estimated for the Adelaide case study using function *ET.Penman()*. The arguments are set so that (1) the time-step for calculation is daily; (2) the actual sunshine hours are used for calculating solar radiation; (3) the actual wind data are used; (4) the Penman 1948 wind function (Penman, 1948) is used to estimate the mass transfer component in the Penman model; and (5) the evaporative surface is open water (albedo = 0.08, roughness height = 0.001 m). The calculated time series of Penman ET from *ET.Penman()* has been saved in an R data list, while output is printed to the screen, which confirms the choice of model and the selection of alternative calculation options, and also gives a basic statistical summary of the entire time-series of ET estimates.

Fig. 3 is a screenshot of data processing and ET estimation with *ReadInputs()* and *ET.Penman()* for this case study.

The plots of estimated daily ET and monthly averaged daily ET have been produced for the Adelaide case study using function *ETPlot()* (Fig. 4). Although it is difficult to detect any trend from the highly fluctuating daily estimates (Fig. 4a), there is a very strong seasonal pattern, displayed in the monthly average plot (Fig. 4b). The ET peaks during the summer months, as would be expected due to the higher temperature and solar radiation during this time of the year.

```
> data <- ReadInputs(climatedata,constants,stopmissing=c(10,3), timestep = "subdaily",
+                    interp_missing = T, interp_abnormal = T,
+                    missing_method = "DoY average", abnormal_method = "DoY average")
The maximum acceptable percentage of missing data is 10 %
The maximum acceptable percentage of continuous missing data is 3 %
Warning: missing data of 'Tmax.daily'(daily maximum temperature), calculated from subdaily 'Temp.subdaily'
Warning: missing data of 'Tmin.daily'(daily minimum temperature), calculated from subdaily 'Temp.subdaily'
Warning: missing data of 'u2.subdaily', calculated from 'uz.subdaily
Number of missing values in uz.subdaily:  3
% missing data:  0.03 %
Maximum duration of missing data as percentage of total duration:  0.02 %
Warning: missing data of 'RHmax.daily'(daily maximum relative humidity), calculated from subdaily 'RH.subdaily'
Warning: missing data of 'RHmin.daily'(daily minimum relative humidity), calculated from subdaily 'RH.subdaily'
> Results <- ET.Penman(data, constants, solar="sunshine hours", wind="yes", windfunction_ver = "1948", alpha = 0.08, z0 = 0.001)
Penman Open-water Evaporation
Evaporative surface: water, albedo = 0.08 ; roughness height = 0.001 m
Sunshine hour data have been used for calculating incoming solar radiation
Wind data have been used for calculating the Penman evaporation. Penman 1948 wind function has been used.
Timestep: daily
Units: mm
1280 ET estimates obtained
Time duration: 2001-03-01 to 2004-08-31
Basic stats
Mean: 4.86
Max: 12.78
Min: 1.03
```

Fig. 3. Example of a typical session of data processing with *ReadInputs()* and ET estimation with *ET.Penman()* for the Adelaide case study.

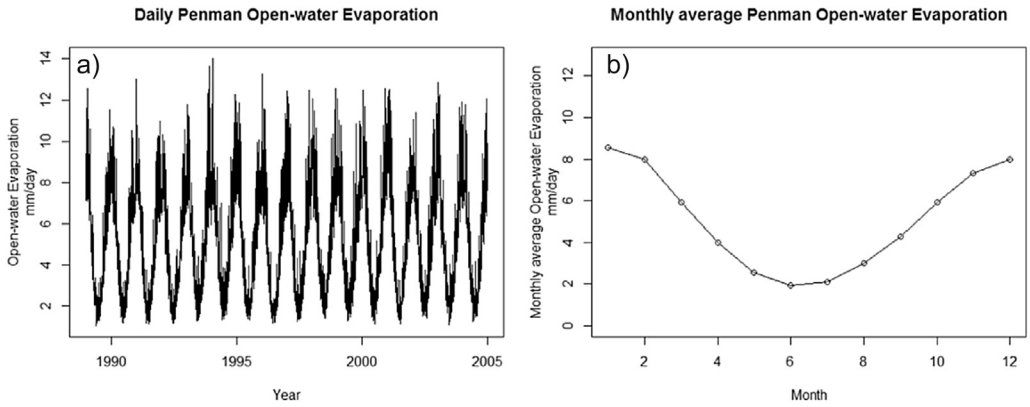


Fig. 4. a) Daily estimates of Penman open-water ET (left panel); b) Monthly averaged daily Penman open-water ET (right panel) for the Adelaide case study, generated by *ETPlot()*.

3.2. Advanced features: analyses with ensemble models and different input data sets

The features of function *ETComparison()* are demonstrated for both the Adelaide and Alice Springs case studies. First, plots of the time series and the non-exceedance probabilities for monthly ET estimates have been produced to compare estimates from the Penman-Monteith FAO56 and Priestley-Taylor models during 1989–1991 (Fig. 5). From Fig. 5a and b we observed that:

1) When comparing across the two case study sites, the inter-model differences in estimates are greater at Alice Springs, with the Priestley-Taylor model producing consistently lower estimates than the Penman-Monteith, and;
2) When comparing the seasonal patterns, the inter-model differences in estimates are most significant for the peak estimates, which occur in every summer (for example, at the start of 1990).

Fig. 5c shows the distribution of the monthly estimates within the period and from the different models, which is consistent with previous observations: the ET estimates from the Priestley-Taylor model are consistently lower compared with those obtained using the Penman-Monteith model, with the greatest difference of approximately 100 mm for the peak estimates at Alice Springs. These results reflect the structural differences in the two models, as the Penman-Monteith model explicitly takes the mass transfer for evapotranspiration into account, which is higher during summer periods and for arid and windy conditions (as reported in McKenney and Rosenberg, 1993; Yin et al., 2010), such as those experienced in Alice Springs.

Another application of *ETComparison()* is demonstrated in Fig. 6, in which the effect of uncertainties in input climate data under climate changes are shown for the Adelaide case study, together with the model uncertainty. To maintain the simplicity and clarity of the example, we focus only on the potential uncertainties in the future temperature due to climate change, without considering the probability of individual changes or potential variations in other climate variables. We perturb the existing temperature data within a range of 0 to +8 °C, which is considered to encompass all plausible future changes in temperature in Australia by 2100 (Stocker et al., 2013). Within this range, 500 random samples have been drawn and the corresponding perturbations are applied to the historical time series of T_{max} and T_{min}, resulting in 500 sets of input climate data. These 500 sets were then fed into *ETComparison()* to generate the corresponding ET outputs using both the Penman-Monteith and Priestley-Taylor models. The resulting ranges of the monthly ET estimates for the period between 1989 and 1999 are shown in Fig. 6. As can be seen, there is a greater range in ET estimates from the Penman-Monteith model. Since both ET models use the same temperature data as inputs, this indicates that temperature has a greater impact on the ET estimates obtained using the Penman-Monteith model—a pattern also observed in McKenney and Rosenberg (1993). This difference can be due to the structural differences between the two models: as the Penman-Monteith explicitly considers the mass transfer processes that are related to temperature, the importance of temperature is higher in the Penman-Monteith model.

It is worth mentioning that all climate variables other than temperature (i.e. *RH*, R_s and u_z) are kept at their current levels in this example, which is unrealistic under future conditions. Therefore, the results should only be considered as illustrative of the key feature of function *ETComparison()*, as a tool to compare ET estimates from multiple input data sets and ET models. In a formal assessment of the impact of climate-related input uncertainty, it is necessary to consider the potential uncertainty in the full set of climate variables that influence ET (Goyal, 2004; Whateley et al., 2014).

4. Discussion

4.1. Further analyses with other software packages

The output from this package is formatted as time-series-like data in the *zoo* format (in which every data point is linked to a specific time point, see Zeileis et al., 2015), so it can be easily extracted and used as an input to other R-based software packages for a range of further analyses. Examples include using ET estimates as input to hydrologic models in *hydromad* (Andrews et al., 2011), and to investigate the sensitivity of ET estimates to changes in the input climate data using *sensitivity* (Pujol et al., 2014). Since the output is also saved to a csv file (as detailed in Section 2.2), it can also be imported to external software packages.

4.2. Limitations

Although the package provides features for checking missing values and errors in the input climate data, as well as interpolation methods for these problematic data, caution is required to minimize the risk of misuse. In developing this package, we have tested the data processing tools with our own test data sets, as well as a number of user-provided data sets, and we have ensured that the package runs free of errors with these existing data sets. However, since every data set is different, it is recommended that users should exercise their own quality-control procedure prior to using

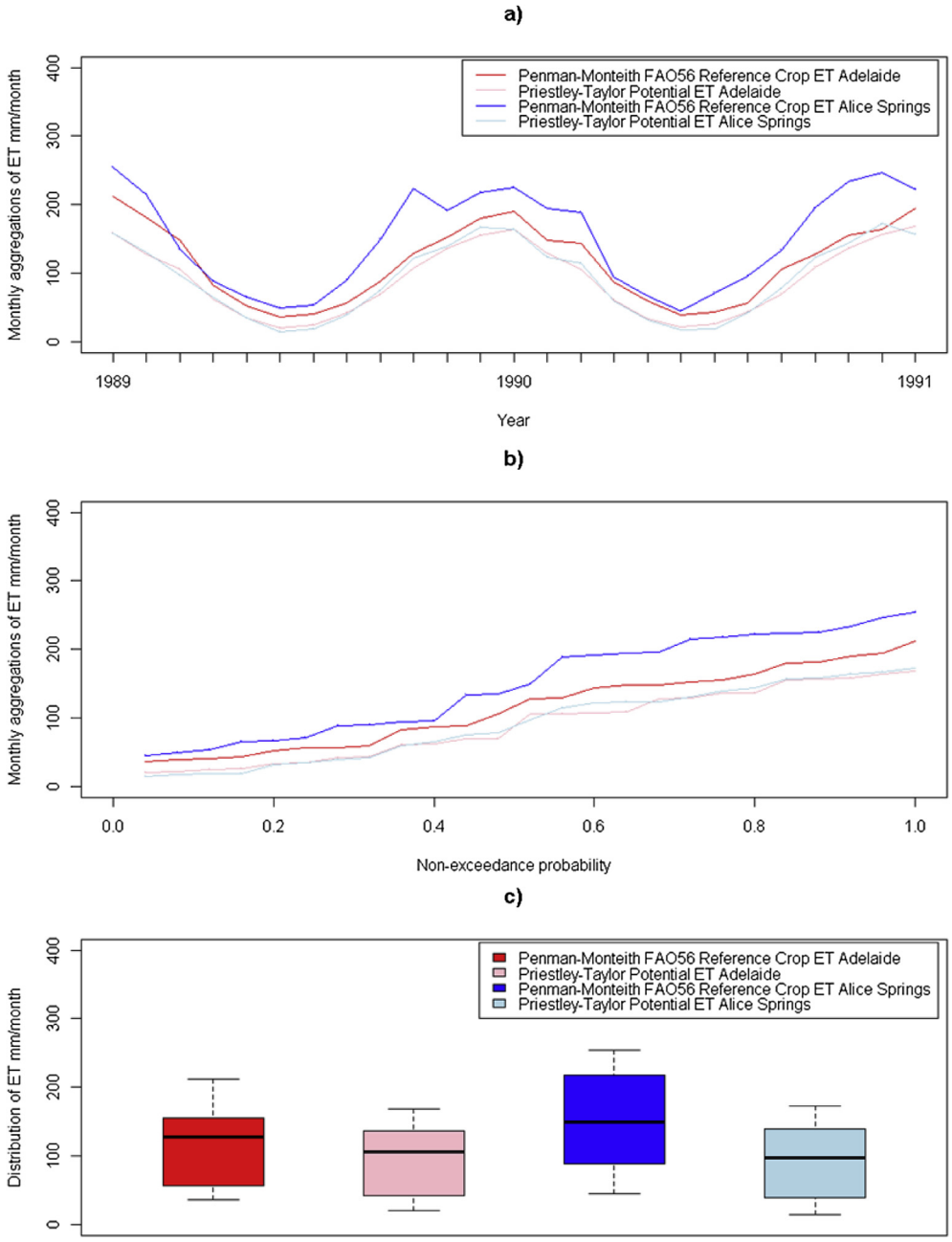


Fig. 5. Comparison of monthly ET estimates from two models (Penman-Monteith FAO56 and Priestley-Taylor) and two locations (Adelaide and Alice Springs) using *ETComparison()* for a) time-series; b) non-exceedance probability, and; c) distribution.

the package, to ensure that best-quality data are provided for ET estimation and the impact of data quality on the estimates is minimized.

Users should also be aware of the full assumptions and limitations prior to using any ET model in this package. Almost every ET model contains assumptions relating to the specific climate conditions under which the models apply. For example, some models assume that sub-processes related to ET are negligible, while other models are only calibrated to the climate of a specific region (a full list of assumptions and limitations for each individual model is given in Table 9 in the supplementary material, which is summarized from the existing literature). These assumptions limit the models' ability to generalize to a wider range of climate zones, leading to varying performance of ET models under different climate settings (Rosenberry et al., 2007; Tabari et al., 2013). A further problem arises if the models are to be applied to estimate ET under climate change conditions, which can mean that existing ET processes and related climate variables are likely to be different to those for which the models are best suited, potentially causing deteriorating model performance (Prudhomme and Williamson, 2013; Thompson et al., 2014).

5. Summary and conclusions

This paper presents an R package *Evapotranspiration* for the estimation of actual, potential and reference crop ET using 17 models in a consistent, convenient and efficient manner. The pre-processing tool provides flexible methods for checking and

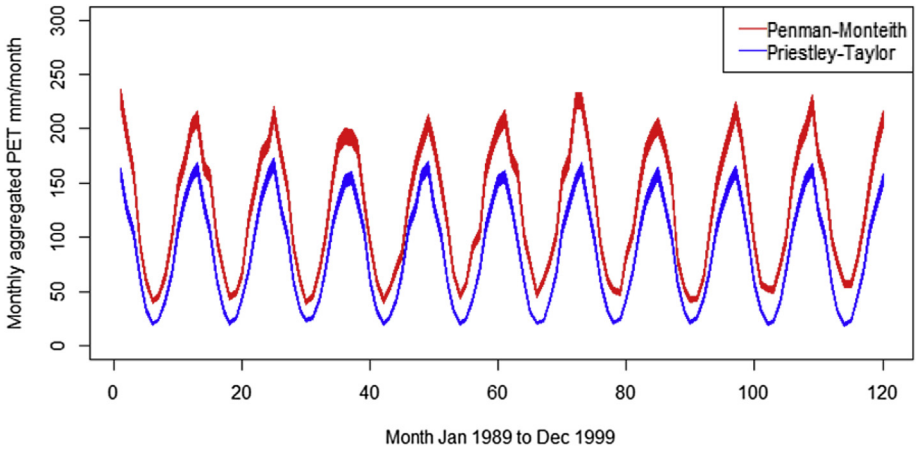


Fig. 6. Uncertainties in monthly ET estimates from two models (Penman-Monteith FAO56 and Priestley-Taylor) at Adelaide, each executed for 500 sets of input data sampled with 0 to +8 °C uncertainty in temperature, generated by *ETComparison()*.

processing raw input climate data, which are then fed into user-selected ET models. The presentation of results is in the form of both summary text and plots. Comparison between multiple ET models and input data sets is also supported. Estimates from the package can be conveniently extracted for further analysis, such as rainfall-runoff modelling and sensitivity analyses. It is hoped that this package will increase consistency in the results presented in ET studies, and increase our ability to investigate the impact of structural uncertainty in ET model formulations via the use of ensemble modelling.

Acknowledgements

The authors wish to thank Murray Peel and Thomas McMahon for their valuable comments on delineating the scope of the paper, and to Joseph Guillaume and an anonymous reviewer for their thoughtful comments on the manuscript.

Software availability

Description: Package *Evapotranspiration*.
Developers: Danlu Guo, Seth Westra.
Year First Available: 2014.
E-mail: Danlu.guo@adelaide.edu.au.
Website:http://cran.r-project.org/web/packages/Evapotranspiration/index.html.
Hardware Requirement: General-purpose computer.
Software Requirement: R version 2.10 or later.
Programming Language: R.

Licensing

This software is made freely available under the terms and conditions of the GNU General Public License.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.envsoft.2015.12.019.

References

Allen, R., Pruitt, W., 1986. Rational use of the FAO Blaney-Criddle formula. J. Irrigation Drainage Eng. 112 (2), 139–155.
Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements-FAO Irrigation and Drainage Paper 56, vol. 300. FAO, Rome, p. 6541.
Allen, R.G., Walter, I.A., Elliott, R.L., Howell, T.A., Itenfisu, D., Jensen, M.E., Snyder, R.L., 2005. The ASCE Standardized Reference Evapotranspiration Equation. ASCE Publications.
Andréassian, V., Perrin, C., Michel, C., 2004. Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models. J. Hydrology 286 (1–4), 19–35.
Andrews, F., Croke, B., Jakeman, A., 2011. An open software environment for hydrological model assessment and development. Environ. Model. Softw. 26 (10), 1171–1185.
Baumgartner, A., Reichel, E., Lee, R., 1975. The World Water Balance: Mean Annual Global, Continental and Maritime Precipitation, Evaporation and Run-off. Elsevier Scientific Publishing Company.
Beguería, S., Vicente-Serrano, S.M., Beguería, M.S., 2013. Package 'SPEI'.
Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrology 249 (1–4), 11–29.
Bormann, H., 2011. Sensitivity analysis of 18 different potential evapotranspiration models to observed climatic change at German climate stations. Clim. Change 104 (3–4), 729–753.
Bouchet, R.J., 1963. Evapotranspiration réelle et potentielle, signification climatique. In: Paper Presented at General Assembly Berkeley, Gentbrugge, Belgium.
Brutsaert, W., Stricker, H., 1979. An advection-aridity approach to estimate actual regional evapotranspiration. Water Resour. Res. 15 (2), 443–450.
Chapman, T., 2001. Estimation of evaporation in rainfall-runoff models. In: Ghassemi, F., Post, D., Sivapalan, M., Vertessy, R. (Eds.), MODSIM2001: Integrating Models for Natural Resources Management across Disciplines, Issues and Scales, MSSANZ, vol. 1, pp. 293–298.
Clark, M.P., Slater, A.G., Rupp, D.E., Woods, R.A., Vrugt, J.A., Gupta, H.V., Wagener, T., Hay, L.E., 2008. Framework for Understanding Structural Errors (FUSE): a modular framework to diagnose differences between hydrological models. Water Resour. Res. 44 (12).
Dawson, C.W., Abrahart, R.J., See, L.M., 2007. HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. Environ. Model. Softw. 22 (7), 1034–1052.
De Bruin, H., 1981. The determination of (reference crop) evapotranspiration from routine weather data. Evaporation Relat. Hydrology 25–37.
Dingman, S.L., 2015. Physical Hydrology. Waveland Press.
Donohue, R.J., McVicar, T.R., Roderick, M.L., 2010. Assessing the ability of potential evaporation formulations to capture the dynamics in evaporative demand within a changing climate. J. Hydrology 386 (1), 186–197.
Doorenbos, J., 1977. Guidelines for predicting crop water requirements. FAO Irrigation Drainage Pap. 24, 15–29.
Duan, Q., Ajami, N.K., Gao, X., Sorooshian, S., 2007. 'Multi-model ensemble hydrologic prediction using Bayesian model averaging. Adv. Water Resour. 30 (5), 1371–1386.
Fisher, J.B., Whittaker, R.J., Malhi, Y., 2011. ET come home: potential evapotranspiration in geographical ecology. Glob. Ecol. Biogeogr. 20 (1), 1–18.
Galelli, S., Humphrey, G.B., Maier, H.R., Castelletti, A., Dandy, G.C., Gibbs, M.S., 2014. An evaluation framework for input variable selection algorithms for environmental data-driven models. Environ. Model. Softw. 62 (0), 33–51.
Gasca-Tucker, D., Acreman, M., Agnew, C., Thompson, J., 2007. Estimating evaporation from a wet grassland. Hydrology Earth Syst. Sci. 11 (1).
Goyal, R.K., 2004. Sensitivity of evapotranspiration to global warming: a case study of arid zone of Rajasthan (India). Agric. Water Manag. 69 (1), 1–11.
Granger, R.J., Gray, D., 1989. 'Evaporation from natural nonsaturated surfaces. J. Hydrology 111 (1), 21–29.

Hargreaves, G.H., Samani, Z.A., 1985. 'Reference crop evapotranspiration from ambient air temperature. Am. Soc. Agric. Eng. no. fiche no. 85-2517.
Horváth, S., Szép, I.J., Makra, L., Mika, J., Pajtók-Tari, I., Utasi, Z., 2010. Effect of evapotranspiration parameterisation on the Palmer drought severity Index. Phys. Chem. Earth, Parts A/B/C 35 (1–2), 11–18.
Jensen, M.E., Haise, H.R., 1963. 'Estimating evapotranspiration from solar radiation', Proceedings of the American Society of Civil Engineers. J. Irrig. and Drain. Div. 89, 15–41.
Kannan, N., White, S.M., Worrall, F., Whelan, M.J., 2007. Sensitivity analysis and identification of the best evapotranspiration and runoff options for hydrological modelling in SWAT-2000. J. Hydrology 332 (3–4), 456–466.
Kavetski, D., Fenicia, F., 2011. Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. Water Resour. Res. 47 (11), W11511.
Kay, A.L., Davies, H.N., 2008. Calculating potential evaporation from climate model data: a source of uncertainty for hydrological climate change impacts. J. Hydrology 358 (3–4), 221–239.
Kingston, D.G., Todd, M.C., Taylor, R.G., Thompson, J.R., Arnell, N.W., 2009. Uncertainty in the estimation of potential evapotranspiration under climate change. Geophys. Res. Lett. 36 (20), L20403.
Leavesley, G.H., Markstrom, S.L., Viger, R.J., 2006. USGS modular modeling system (MMS)-precipitation-runoff modeling system (PRMS). Watershed models 159–177.
McJannet, D., Webster, I., Stenson, M., Sherman, B., 2008. Estimating Open Water Evaporation for the Murray-Darling Basin: a Report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project.
McKenney, M.S., Rosenberg, N.J., 1993. Sensitivity of some potential evapotranspiration estimation methods to climate change. Agric. For. Meteorology 64 (1–2), 81–110.
McMahon, T.A., Peel, M.C., Lowe, L., Srikanthan, R., McVicar, T.R., 2013. Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. Hydrol. Earth Syst. Sci. 17 (4), 1331–1363.
Monteith, J., 1991. 'Weather and water in the Sudano-Sahelian zone'. In: Sivakumar, M.V.K., Wallace, J.S., Renard, C., Giroux, C. (Eds.), Soil Water Balance in the Sudano-Sahelian Zone, Proceedings of the Niamey Workshop, Niamey, Niger, February 1991, Niamey, Niger, pp. 11–29. IAHS Publ. no. 199.
Morton, F.I., 1983a. Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. J. Hydrology 66 (1–4), 1–76.
Morton, F.I., 1983b. Operational estimates of lake evaporation. J. Hydrology 66 (1–4), 77–100.
Narapusetty, B., DelSole, T., Tippett, M.K., 2009. Optimal estimation of the climatological mean. J. Clim. 22 (18), 4845–4859.
Nichols, J., Eichinger, W., Cooper, D., Prueger, J., Hipps, L., Neale, C., Bawazir, A., 2004. Comparison of Evaporation Estimation Methods for a Riparian Area. Final report. UHR Technical Report, no. 436.
Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., Loumagne, C., 2005a. Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. J. Hydrology 303 (1–4), 290–306.
Oudin, L., Michel, C., Anctil, F., 2005b. Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 1—Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs? J. Hydrology 303 (1–4), 275–289.
Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. Proc. R. Soc. Lond. Ser. A. Math. Phys. Sci. 193 (1032), 120–145.
Penman, H.L., 1956. Evaporation: an introductory survey. Neth. J. Agric. Sci. 4, 9–29.
Priestley, C., Taylor, R., 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. Mon. Weather Rev. 100 (2), 81–92.
Prudhomme, C., Williamson, J., 2013. Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: a comparison of methods and associated uncertainty in future projections. Hydrology Earth Syst. Sci. 17 (4), 1365–1377.
Pujol, G., Iooss, B., Iooss, M.B., 2014. Package 'sensitivity'.
Raes, D., Munoz, G., 2008. The ETo calculator. Land water digital media Ser.
Rosenberry, D.O., Winter, T.C., Buso, D.C., Likens, G.E., 2007. Comparison of 15 evaporation methods applied to a small mountain lake in the northeastern USA. J. Hydrology 340 (3), 149–166.
Rotstayn, L.D., Roderick, M.L., Farquhar, G.D., 2006. A simple pan-evaporation model for analysis of climate simulations: evaluation over Australia. Geophys. Res. Lett. 33 (17).
Shuttleworth, W., Wallace, J., 2009. Calculating the water requirements of irrigated crops in Australia using the Matt-Shuttleworth approach. Trans. ASABE 52 (6), 1895–1906.
Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M., 2013. Climate change 2013: the physical science basis. In: Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5). Cambridge Univ Press, New York.
Szilagyi, J., 2007. On the inherent asymmetric nature of the complementary relationship of evaporation. Geophys. Res. Lett. 34 (2), L02405.
Tabari, H., Grismer, M.E., Trajkovic, S., 2013. Comparative analysis of 31 reference evapotranspiration methods under humid conditions. Irrigation Sci. 31 (2), 107–117.
Thompson, J.R., Green, A.J., Kingston, D.G., 2014. Potential evapotranspiration-related uncertainty in climate change impacts on river flow: an assessment for the Mekong River basin. J. Hydrology 510 (0), 259–279.
Turc, L., 1961. Estimation of irrigation water requirements, potential evapotranspiration: a simple climatic formula evolved up to date. Ann. Agron. 12 (1), 13–49.
Velázquez, J.A., Schmid, J., Ricard, S., Muerth, M.J., Gauvin St-Denis, B., Minville, M., Chaumont, D., Caya, D., Ludwig, R., Turcotte, R., 2012. An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. Hydrology Earth Syst. Sci. 17 (2013), 565–578.
Whateley, S., Steinschneider, S., Brown, C., 2014. A climate change range-based method for estimating robustness for water resources supply. Water Resour. Res. 50 (11), 8944–8961.
Xu, C.Y., Singh, V.P., 2000. Evaluation and generalization of radiation-based methods for calculating evaporation. Hydrol. Process. 14 (2), 339–349.
Xu, C.Y., Singh, V.P., 2002. Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. Water Resour. Manag. 16 (3), 197–219.
Yin, Y., Wu, S., Chen, G., Dai, E., 2010. Attribution analyses of potential evapotranspiration changes in China since the 1960s. Theor. Appl. Climatol. 101 (1–2), 19–28.
Yin, Z.-Y., Brook, G.A., 1992. Evapotranspiration in the Okefenokee Swamp watershed: a comparison of temperature-based and water balance methods. J. Hydrology 131 (1), 293–312.
Zeileis, A., Grothendieck, G., Ryan, J.A., Andrews, F., Zeileis, M.A., 2015. Package 'zoo'.

CHAPTER 3 Sensitivity of Potential Evapotranspiration to Changes in Climate Variables for Different Australian Climatic Zones (Paper 2)

Statement of Authorship

Title of Paper	Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones
Publication Status	☑ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Guo, D., Westra, S. & Maier, H. R. 2016. Sensitivity of potential evapotranspiration to changes in climate variables for different climatic zones. Hydrol. Earth Syst. Sci. Discuss., 2016, 1-43. Update: Guo, D., Westra, S., and Maier, H. R. 2017. Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones, Hydro. Earth Syst. Sci., 21, 2107-2126.

Principal Author

Name of Principal Author (Candidate)	Danlu Guo		
Contribution to the Paper	Designed experiments, conducted model simulations, analysed results, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/12/2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Seth Westra		
Contribution to the Paper	Suggested scope of study, helped to design experiments and interpret results, provided feedbacks on manuscript and responses to reviewers.		
Signature		Date	19/12/16

Name of Co-Author	Holger Maier		
Contribution to the Paper	Suggested scope of study, helped to evaluate the manuscript and the responses to reviewers.		
Signature		Date	20/12/16

Please cut and paste additional co-author panels here as required.

Abstract

Assessing the factors that have an impact on potential evapotranspiration (PET) sensitivity to changes in different climate variables is critical to understanding the possible implications of climatic changes on the catchment water balance. Using a global sensitivity analysis, this study assessed the implications of baseline climate conditions on the sensitivity of PET to a large range of plausible changes in temperature (T), relative humidity (RH), solar radiation (R_s) and wind speed (u_z). The analysis was conducted at 30 Australian locations representing different climatic zones, using the Penman-Monteith and Priestley-Taylor PET models. Results from both models suggest that the baseline climate can have a substantial impact on overall PET sensitivity. In particular, approximately two-fold greater changes in PET were observed in cool-climate energy-limited locations compared to other locations in Australia, indicating the potential for elevated water loss as a result of increasing actual evapotranspiration (AET) in these locations. The two PET models consistently indicated temperature to be the most important variable for PET, but showed large differences in the relative importance of the remaining climate variables. In particular, for the Penman-Monteith model wind and relative humidity were the second-most important variables for dry and humid catchments, respectively, whereas for the Priestley-Taylor model solar radiation was the second-most important variable, with greatest influence in warmer catchments. This information can be useful to inform the selection of suitable PET models to estimate future PET for different climate conditions, providing evidence on both the structural plausibility and input uncertainty for the alternative models.

3.1 Introduction

Assessing changes to evapotranspiration (ET) is critical in understanding the impacts of anthropogenic climate change on the catchment water balance. ET represents the dominant loss of water from catchments worldwide, with about 62% of global land-surface precipitation accounted for by ET (Dingman, 2015), and ET exceeding runoff in over 77% of the global land surface (Harrigan and Berghuijis, 2016). ET is affected by climate change through a cascade of processes that begins with the increasing concentration of greenhouse gases, followed by their attendant impacts on large-scale circulation and changes to the global distribution of energy and moisture. These large-scale processes lead to local-scale changes in the atmosphere, which in turn influence the catchment water balance through a set of terrestrial hydrological processes by which precipitation is converted into actual ET (AET), runoff and groundwater recharge (Oudin et al., 2005). Other factors that can potentially affect ET under a changing climate include changing land cover patterns (e.g. Liu et al., 2008), and the CO_2 fertilization effects that can limit the rate of plant transpiration under elevated levels of CO_2 (e.g. Prudhomme et al., 2014;Milly and Dunne, 2016).

Climate impact studies that investigate the influence of climate forcings on the catchment water balance are usually based on projections of future climate represented by climate variables such as temperature and solar radiation from general circulation models (GCMs), which are converted into potential ET (PET) using one or several PET models. The PET projections are combined with GCM projections of precipitation (P), which together can be used to directly estimate the water deficit (Taylor et al., 2013;Chang et al., 2016). Alternatively, rainfall-runoff models can be used to translate the changes in P and PET into changes in runoff (e.g. Akhtar et al., 2008;Chiew et al., 2009;Koedyk and Kingston, 2016), as well as associated information such as the impact on catchment streamflow (Wilby et al., 2006), water supply security (Paton et al., 2014, 2013) and flood risk (Bell et al., 2016). Therefore, to quantify the specific impact of changes in ET on the water balance, a good understanding of the sensitivity of PET to potential changes in its key

influencing climatic variables is required (Goyal, 2004;Tabari and Hosseinzadeh Talaee, 2014). This is particularly relevant given the recent focus on 'scenario-neutral' (or 'bottom-up') approaches to climate impact assessment (Brown et al., 2012;Prudhomme et al., 2010;Culley et al., 2016), which require the sensitivity of a given system to potential changes in climate forcings to be estimated (Prudhomme et al., 2013a;Steinschneider and Brown, 2013;Prudhomme et al., 2013b;Kay et al., 2014;Guo et al., 2016a).

Furthermore, the sensitivity of PET can provide critical evidence in relation to identifying models that are most appropriate for PET estimation under climate change conditions, which is particularly relevant to the ongoing debate on the potential trade-off between model complexity and reliability. Complex models such as the Penman-Monteith model are often recommended for their ability to better represent the physical processes that affect PET (McVicar et al., 2012;Donohue et al., 2010;Barella-Ortiz et al., 2013). For example, the Penman-Monteith model can account for the effects of wind, and thus can help explaining at least part of the observed decreases in pan evaporation with increases in temperature in many locations globally – the 'evaporation paradox' – due to the observed decreases in wind speed (Roderick et al., 2007;McVicar et al., 2008;Lu et al., 2016). However, simpler empirical models may also be preferable under some conditions, as they require a smaller number of input climate variables, which might be able to be projected with greater confidence with GCMs, and thus leading to greater confidence in the corresponding PET estimates (Kay and Davies, 2008;Ekström et al., 2007;Ravazzani et al., 2014). For example, there is reasonable confidence in projections of temperature and relative humidity in Australia for a given emission scenario, but less confidence in projections of wind due to sub-grid effects of orography and other land-surface features (Flato et al., 2013;CSIRO and Bureau of Meteorology, 2015). In these situations, models such as the Priestley-Taylor model that do not depend on wind may produce more reliable estimates of PET compared to the more complex Penman-Monteith model. Thus, the choice of climate variables to include in climate impact assessments must be informed both by the relative importance of each variable on projections of PET (e.g. Tabari and

Hosseinzadeh Talaee, 2014), and the likely confidence in the projections of each variable (e.g. Flato et al., 2013;Johnson and Sharma, 2009).

Sensitivity analysis methods have been employed in a number of recent studies to assess the overall sensitivity of PET estimated by the Penman-Monteith model to potential changes in climate, as well as to better understand the relative importance of different climate variables on overall PET sensitivity. For example, Goyal (2004) found that PET was most sensitive to perturbations in temperature, followed by solar radiation, wind speed and vapor pressure, at a single study site in an arid region in India. Tabari and Hosseinzadeh Talaee (2014) also looked at the sensitivity of PET to perturbations of historical climate data from eight meteorological stations representing four climate types in Iran, and concluded that the importance of wind speed and air temperature was lower while that of sunshine hours was higher for a humid location compared to an arid location. Gong et al. (2006) found that the differences in PET sensitivity across the upper, middle and lower regions of the Changjiang (Yangtze) basin in China were largely due to contrasting baseline wind speed patterns. However, most of these PET sensitivity analysis studies focused on a limited number of study sites and/or climatic zones, so that the specific causes for varying PET sensitivity at different locations, such as the roles of climatic and hydrological conditions, remain unclear. Consequently, it is difficult to extrapolate our existing knowledge of PET sensitivity and the relative importance of each climate variable to new locations, which is essential for assessing the water balance at regional scales.

To address the shortcomings of existing studies outlined above, this study aims to gain an understanding of (i) the sensitivity of PET estimates to changes in the key climatic variables which influence PET, and how these sensitivity estimates are affected by varying baseline hydrologic and climatic conditions at different locations; and (ii) the relative importance of these climatic variables for PET, and how this changes with the baseline hydrologic and climatic conditions at different locations. These aims were achieved by analyzing the responses of PET to perturbations in four of its driving climatic

variables, namely temperature (T), relative humidity (RH), solar radiation (R_s) and wind speed (u_z), at 30 study sites across Australia representing a range of climate zones. Both the Penman-Monteith and Priestley-Taylor models were used, as they represent different conceptualizations of the PET-related processes, with both models being widely used for climate impact assessments (Felix et al., 2013;Arnell, 1999;Gosling et al., 2011;Kay et al., 2009;Prudhomme and Williamson, 2013;Donohue et al., 2009). It is worth noting that the potential changes in one climate variable can be amplified or offset by changes in another variable (for examples see the discussions of 'evaporation paradox' in Lu et al., 2016;Roderick and Farquhar, 2002), which can affect the relative importance of each variable. To account for this effect, a global sensitivity analysis method was used, with similar methods being applied to account for the impact of joint variations in the input variables on the output from a variety of environmental models, ranging from conceptual rainfall-runoff models (e.g. Tang et al., 2007a;Tang et al., 2007c) to complex models which consider a number of surface-groundwater processes (e.g. Guillevic et al., 2002;van Griensven et al., 2006;Nossent et al., 2011). The results of the global sensitivity analysis in this study were presented in terms of both the range of potential changes in PET and relative sensitivity indices of each climate variable for PET, which were then used to elucidate the specific roles of varying baseline hydro-climatic conditions on influencing these sensitivity measures.

The subsequent sections of this paper are structured as follows. Section 3.2 introduces the data obtained from the 30 study sites required for the global sensitivity analysis. Section 3.3 describes the approach to the global sensitivity analysis of PET. Section 3.4 presents and discusses two sets of results which address the two study aims respectively: (i) the range of estimated changes in PET in response to potential changes in temperature, solar radiation, humidity and wind, and how this changes with location; and (ii) the relative importance of the four climate variables for estimating PET, and how this changes with location. The study is summarized and concluded in Section 3.5.

3.2 Data

To represent contrasting hydro-climatic conditions for assessing PET sensitivity, we selected case study locations within different Köppen classes in Australia. The original Köppen climate classification (Köppen et al., 1930;Köppen, 1931) provides a useful categorization of hydro-climatic conditions at specific locations, which is based on the long-term average levels and seasonal patterns of climatic and hydrologic variables, including temperature, relative humidity and rainfall. A 'modified Köppen classification' system has been adapted for Australia (as in Stern et al., 2000) and is now widely used in climatic and hydrologic studies to identify and categorize case study locations (e.g. Johnson and Sharma, 2009;Rustomji et al., 2009;Li et al., 2014;Guo et al., 2017).

As mentioned in the Introduction, both the Penman-Monteith and the Priestley-Taylor models were used to estimate PET for the global sensitivity analyses. The estimation of PET with these models relies on temperature, relative humidity, solar radiation and (for the Penman-Monteith model only) wind speed. In addition, the rainfall data were also obtained to assess the aridity of the different locations. We limited the selection of study sites to those with 10 or more years of continuous climate data with no more than 5 % missing records over the study period. This led to a final selection of 30 weather stations (Figure 3-1), with a consistent data period from 1 January 1995 to 31 December 2004. The data obtained at each site are detailed as below:

- **Daily maximum and minimum temperature (T in °C), maximum and minimum relative humidity (RH in %) and wind speed (u_z in m s^{-1}):** Data for each of these variables were obtained directly from each weather station.
- **Daily solar radiation (R_s in MJ m^{-2} day^{-1}):** Daily solar radiation was calculated from daily sunshine hour data (n in h) obtained from each

weather station, using the Ångström-Prescott equation as in McMahon et al. (2013).

- **Daily rainfall (mm/day)**: Daily rainfall data were obtained from a rain gauge at each weather station.

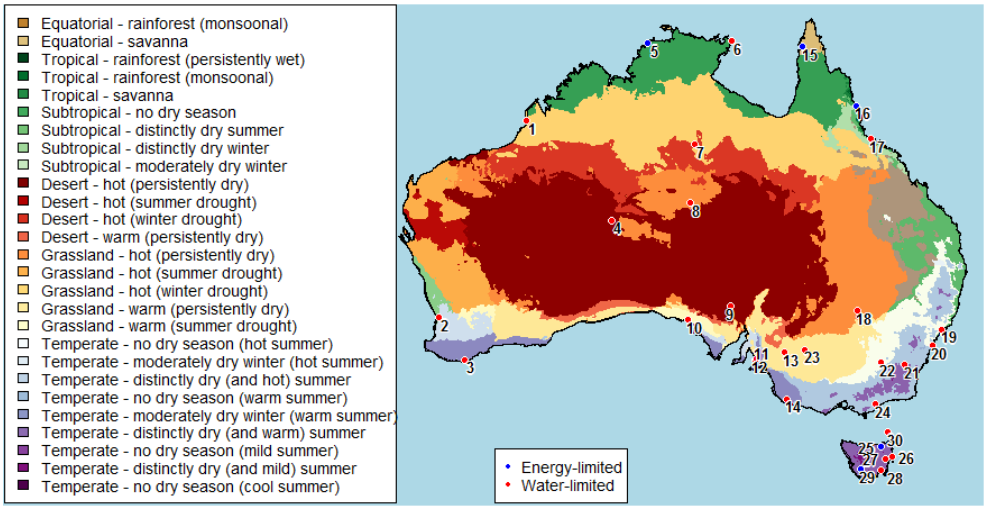


Figure 3-1: Locations of 30 Australian weather stations selected for analysis (see Table 3-1 for the full names of these weather stations), with reference to their corresponding climate classes derived following the modified Köppen classification (reproduced with data from Stern et al., 2000).

Table 3-1 shows the average values of the four PET-related climate variables, as well as the rainfall within the study period, at each of the 30 sites. As can be seen, there are large differences in the average values of each variable, highlighting large differences in the climatic conditions across the 30 sites. In addition, a quantity particularly relevant to ET processes is the long-term averaged ratio of PET to precipitation (PET/P), which describes whether a location is water-limited (PET/P >1) or energy-limited (PET/P < 1) (Gerrits et al., 2009;McVicar et al., 2010). This ratio was estimated for each site and is also shown in Table 3-1 (with the point colour in Figure 3-1 indicating whether the location is water-limited or energy-limited). The range of PET/P values indicates substantial variations in the water availability conditions at different study sites. Note that these ratios were based on the estimates of PET from the Penman-Monteith model. Although the use of Priestley-Taylor model resulted in different PET estimates at each site, the categorization of

water- and energy-limited catchments was generally consistent with those from Penman-Monteith, with different categories only shown at four out of the 30 study sites (sites 6, 19, 20 and 27).

Table 3-1: Names, locations and average climate conditions of the 30 weather stations over the study period (1995-2004).

No.	Study site name	Köppen class[1]	Lat (°S)	Long (°E)	Elev (m)	T (°C)	RH (%)	R_s (MJ m^{-2} day^{-1})	u_z (m s^{-1})	Annual P (mm)	Annual PET (mm)	Annual PET/P
1	Broome airport	13	-17.95	122.2	7.4	26.37	65.15	21.55	3.684	865	2003	2.317
2	Perth	8	-31.93	116.0	15.4	18.54	61.72	18.95	4.519	721	1751	2.429
3	Albany	4	-34.94	117.8	68	15.08	73.59	15.20	4.382	752	1126	1.498
4	Giles	24	-25.03	128.3	598	22.70	38.40	20.29	4.380	394	2344	5.947
5	Darwin	35	-12.42	130.9	30.4	27.42	69.27	20.33	3.393	1976	1864	0.944
6	Gove	35	-12.27	136.8	51.6	26.29	75.93	19.45	3.500	1607	1660	1.033
7	Tennant Creek	13	-19.64	134.2	375.7	25.73	37.21	21.64	4.759	539	2634	4.886
8	Alice Springs	15	-23.80	133.9	546	21.18	44.53	20.79	2.352	331	1822	5.503
9	Woomera	24	-31.16	136.8	166.6	19.41	46.57	19.40	5.057	151	2153	14.24
10	Ceduna	11	-32.13	133.7	15.3	16.92	62.04	18.20	5.450	266	1723	6.478
11	Adelaide airport	12	-34.95	138.5	2	16.37	63.04	16.91	4.213	454	1410	3.107
12	Adelaide (kent town)	12	-34.92	138.6	48	16.95	61.20	16.88	3.161	569	1372	2.409
13	Loxton	12	-34.44	140.6	30.1	16.50	59.41	17.59	3.250	255	1490	5.847
14	Mount Gambier	4	-37.75	140.8	63	13.45	72.77	14.91	4.460	731	1116	1.526
15	Weipa	41	-12.68	141.9	18	26.87	72.21	19.31	3.271	2154	1782	0.827
16	Cairns	36	-16.87	145.7	3	24.80	73.00	18.98	4.352	1985	1678	0.845
17	Townsville	35	-19.25	146.8	4.3	24.53	69.45	20.27	4.304	1099	1802	1.641
18	Cobar	15	-31.48	145.8	260	19.08	50.64	19.05	2.458	398	1565	3.936
19	Williamtown	9	-32.79	151.8	9	17.84	70.57	16.07	3.927	1145	1309	1.143
20	Sydney	9	-33.94	151.2	6	18.19	67.69	15.97	5.311	1017	1393	1.369
21	Canberra	6	-35.30	149.2	578.4	13.36	65.82	16.86	3.302	590	1226	2.078
22	Wagga Wagga	9	-35.16	147.5	212	15.77	61.78	17.48	3.288	552	1436	2.602
23	Mildura	12	-34.24	142.1	50	17.11	55.62	18.24	3.604	246	1645	6.681
24	East sale	6	-38.12	147.1	4.6	13.77	72.32	14.92	4.062	529	1093	2.067
25	Scottsdale	3	-41.17	147.5	197.5	13.19	70.55	14.23	2.921	931	912	0.980
26	Bicheno	3	-41.87	148.3	11	14.69	66.68	13.69	3.319	690	966	1.401
27	Lake Leake	3	-42.01	147.8	575	9.96	75.40	13.44	3.358	732	774	1.056
28	Hobart	3	-42.83	147.5	4	12.77	65.67	14.04	4.367	483	1097	2.273
29	Strathgordon village	3	-42.77	146.0	322	10.70	77.95	11.65	2.473	2626	699	0.266
30	Flinders Island	3	-40.09	148.0	9	13.54	73.59	14.34	6.399	654	1064	1.626

Note:

[1]**The Köppen classes are presented with their corresponding identifiers from Stern et al. (2000), as: 3. Temperate - no dry season (mild summer); 4. Temperate - distinctly dry (and warm) summer; 6. Temperate - no dry season (warm summer); 8. Temperate - moderately dry winter (hot summer); 9. Temperate - no dry season (hot summer); 11. Grassland - warm (summer drought); 12. Grassland - warm (persistently dry); 13. Grassland - hot (winter drought); 15. Grassland - hot (persistently dry); 24. Desert - hot (persistently dry); 35. Tropical - savanna; 36. Tropical - rainforest (monsoonal); 41 Equatorial - savanna.**

[2]***T*** **= temperature,** ***RH*** **= relative humidity,** R_s **= incoming solar radiation,** u_z **= wind speed, P = rainfall, PET = potential evapotranspiration calculated using the Penman-Monteith model.**

3.3 Method

3.3.1 Overview

A schematic of the approach followed in study is shown in Figure 3-2. As a required model input for the global sensitivity analysis, a large number of representative samples were first obtained for the four climate variables that influence PET (T, RH, R_s and u_z) at each study site, by perturbing the corresponding historical climate data (Section 3.3.2). The outputs of the global sensitivity analysis (i.e. the responses of PET) were estimated with the Penman-Monteith and Priestley-Taylor models (Section 3.3.3). To understand the PET sensitivity and the relative importance of the four climate variables in influencing PET and how these change with location, a global sensitivity analysis was conducted with the responses of PET to the climate perturbations (Section 3.3.4). This proceeded in two parts:

(1) To assess the sensitivity of PET to the climate variables, the range of percentage changes in PET in response to all the climate perturbations was estimated relative to the baseline PET at each location. To observe the impact of varying baseline hydro-climatic conditions, the ranges obtained from each PET model were also plotted against the baseline levels of each climate variable for all study sites.

(2) To assess the relative importance of each climate variable, the range of percentage responses in PET to all climate perturbations in (1) was first compared to the conditional range of percentage responses in PET

with holding each variable constant. This comparison enables an assessment of the relative impact of each variable on the potential responses of PET. An alternative presentation of the individual and interaction effects of the climate variables was achieved using the Sobol' method (Sobol' et al., 2007). Here, the total variance of PET was estimated based on different samples drawn from the perturbed ranges of each climate variable, and then partitioned into the individual contribution from each climate variable and their interactions (see Section 1 in Appendix 3A for details). The Sobol' first-order sensitivity indices were estimated and plotted against the baseline levels of each climate variable for all study sites to explore the role of varying baseline hydro-climatic conditions on the relative importance of each climatic variable for PET.

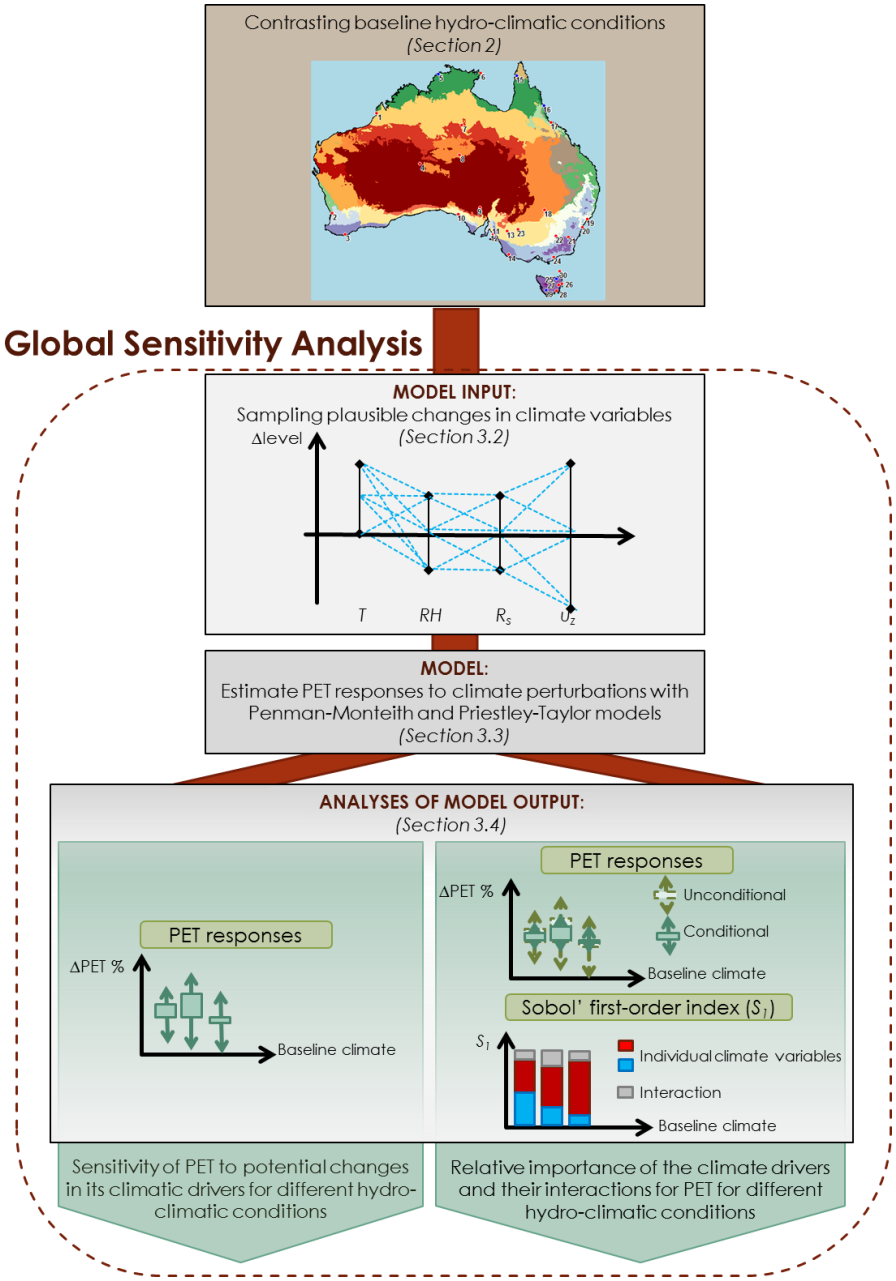


Figure 3-2: Schematic of the method used in this study.

3.3.2 Representing plausible changes in the climatic variables

As part of the global sensitivity analysis, a large number of representative combinations of the changes in the four climate variables (T, RH, R_s and u_z) were obtained. The upper and lower bounds for perturbing each climate variable were determined based on the uncertainty bounds of projections for 2100 for Australia (Stocker et al., 2013). The selected bounds are given in

Table 3-2, which are all slightly wider than those presented in Stocker et al. (2013) to encompass a comprehensive range of plausible future climate change scenarios. Within these bounds, samples were drawn for different combinations of changes in each climatic variable. Latin hypercube sampling (LHS) was used for this purpose due to its effectiveness in covering multi-dimensional input spaces (Osidele and Beck, 2001;Sieber and Uhlenbrook, 2005;Tang et al., 2007b).

Table 3-2: Plausible perturbation bounds for each climate variable relative to their current levels.

Climate variable	Perturbation range
T	0 to +8 °C
RH	-10 % to +10 %
R_s	-10 % to +10 %
u_z	-20 % to +20 %

Note: T = daily temperature, RH = daily relative humidity, R_s = daily incoming solar radiation, u_z = daily wind speed.

According to Nossent et al. (2011) and Zhang et al. (2015), the sample size was selected to ensure the convergence of the first- and total-order Sobol' sensitivity indices, which occurs when the width of the 95 % confidence intervals from 1000-fold bootstrap resampling of the each index is below 10 % of the corresponding mean obtained from bootstrapping. Specifically, we generated different sizes of LHS samples of climate perturbations with the historical climate data from one study site, from which the PET responses were estimated using the Penman-Monteith model. The 1000-fold bootstrap estimates for the Sobol' first- and total-order sensitivity indices for each climate variable were then derived (as in Eqn. 1.2 and 1.5 in Appendix 3A, respectively) for each sample size. It was observed that both the Sobol' indices began to converge when the sample size exceeded 5000, and this was therefore used as the LHS sample size for all the sensitivity experiments in this study. Based on this sample size, a total of 30000 Sobol' samples were compiled as required to estimate the first- and total-order indices (as detailed in Appendix 3A), which correspond to 30000 climate perturbations to be used to test PET sensitivity.

To generate time series of perturbed climate data, the 30000 joint perturbations to the four climate variables obtained by LHS were treated as

change factors, and applied to the time series of daily values of the corresponding historical data. Rather than using a single daily mean value of temperature and relative humidity, the two PET models used in this study require both the daily minimum and maximum values; therefore each pair of temperature variables and relative humidity variables was considered jointly and thus perturbed by the same amount for each day. In addition, to ensure physical plausibility of the perturbations, the daily maximum and minimum values of relative humidity were capped at a maximum of 100%.

3.3.3 Estimating PET responses to climate perturbation

To represent the responses in PET as a result of the climate perturbations, we used both the Penman-Monteith and Priestley-Taylor models, which provide contrasting process representations to estimate PET. The Penman-Monteith model is often referred to as a combinational model, as it combines the energy balance and mass transfer components of ET, and takes into account vegetation-dependent processes such as aerodynamic and surface resistances (Eqn. 2.1 in Appendix 3A). The model requires input of six climate variables, namely, T_{max}, T_{min}, RH_{max}, RH_{min}, R_s and u_z. The Priestley-Taylor model consists of a simpler structure, considering only the energy balance, without consideration of mass transfer or any impact from vegetation (Eqn. 3.1 in Appendix 3A). Therefore, the Priestley-Taylor model is also referred to as a radiation-based model. The model only requires five climate variables, including T_{max}, T_{min}, RH_{max}, RH_{min} and R_s.

To minimize the potential confounding effects of differences in vegetated surface, the evaporative surface was assumed to be reference crop for all study sites, so that it was possible to use the FAO-56 version of the Penman-Monteith model (Allen et al., 1998). The detailed formulations of the two PET models, as well as the relevant constants and assumptions, are included in McMahon et al. (2013). Both models were implemented using the R package *Evapotranspiration* (http://cran.r-project.org/web/packages/Evapotranspiration/index.html) (Guo et al., 2016b). From each model, two sets of estimated PET were obtained: (i) a single set of baseline (historical) PET data at each study site with the historical climate

data; (ii) 30000 sets of perturbed PET data at each study site corresponding to the 30000 sets of perturbed climate data obtained using LHS, as detailed in Section 3.3.2.

3.3.4 Analyses of PET sensitivity

To assess the overall sensitivity of PET to plausible climate change, we first estimated the annual average percentage changes in PET (relative to the baseline PET) using all climate perturbations at the 30 study sites, with estimates from both the Penmen-Monteith and Priestley-Taylor models. A closer investigation of how PET sensitivity varies with baseline climate was conducted by plotting the ranges of all monthly PET responses against the average levels of each climate variable, for all study sites and all months. The reason for the choice of monthly timescale is that for some study sites, the climate can vary substantially by season, so that an annual analysis might obscure important sub-annual effects.

To assess the relative importance of each climate variable for PET estimation from each model, we first compared the ranges of the two sets of PET changes, namely:

(1) The range of all potential changes in PET obtained from the entire 30000 sets of climate perturbations from LHS; and

(2) The conditional ranges of potential changes in PET assuming no change in one of the climate variables. This was obtained with using a subset of all climate perturbations used in (1), for which the changes in the specific conditioning climate variable were close to zero (within ±0.1 ◦C for T, and within ±0.1 % for the other three variables).

In this way any difference between (1) and (2) was purely contributed by the impact of changing the specific conditioning climate variable. To quantify and compare the relative importance of each climate variable, we then utilized the Sobol' method, which was implemented within the R package *sensitivity* (https://cran.r-project.org/web/packages/sensitivity/index.html). We estimated

the Sobol' first-order sensitivity indices (as in Eqn. 1.2, Appendix 3A) to assess the role of each individual climate variable for each PET model, at the 30 study sites. The sum of all interaction effects was also calculated for each location as the difference between the sum of all first-order indices and one (Eqn. 1.6, Appendix 3A). The Sobol' first-order indices were then plotted against the baseline levels of each climate variable at the 30 study sites, to assess how the relative importance changes with the baseline climatic conditions.

3.4 Results and discussion

3.4.1 Ranges of potential changes in PET in response to potential climate change for different climate zones

We start by assessing the potential changes in PET in response to the full set of climate perturbations at the 30 study sites at the annual timescale, using both the Penman-Monteith and Priestley-Taylor models. The results are presented in Table 3-3 in terms of the minimum, maximum and average changes of PET relative to the 1995-2004 baseline, in response to the 30000 sets of climate perturbation at each study site. The two models suggest similar average PET changes at most locations, with the average changes obtained from the Penman-Monteith model across all the locations (+13.38 %) being slightly higher than that for the Priestley-Taylor model (+10.91 %). Greater differences between the two models were observed when considering the ranges of changes. In particular, the minimum and maximum values (averaged across all the 30 sites) were -13.66 % and +47.09 % for the Penman-Monteith model, respectively, compared to -7.39 % and +34.47 % for the Priestley-Taylor model. This corresponds to a range for the Penman-Monteith model being approximately 45 % wider than that of the Priestley-Taylor model.

Table 3-3: Maximum, minimum and average of all possible changes in annual average PET in response to the full set of climate perturbations from the Penman-Monteith and Priestley-Taylor models at the 30 study sites (as % changes to baseline PET relative to the 1995-2004 baseline). The maximum and minimum changes from each model across all locations are shaded in grey.

No.	Study site name	Penman-Monteith	Priestley-Taylor

		Min.	Max.	Avg.	Min.	Max.	Avg.
1	Broome airport	-12.33	39.10	11.16	-9.61	33.75	9.59
2	Perth	-13.20	46.67	13.52	-7.98	34.17	10.62
3	Albany	-15.04	54.67	15.21	-7.28	35.49	11.63
4	Giles	-12.30	37.57	10.68	-7.73	25.83	7.27
5	Darwin	-12.73	39.10	10.92	-9.82	33.84	9.50
6	Gove	-13.10	41.34	11.53	-9.74	33.67	9.61
7	Tennant Creek	-12.28	36.45	10.21	-8.35	26.31	7.09
8	Alice Springs	-10.88	34.00	9.80	-8.00	27.41	7.92
9	Woomera	-12.84	43.48	12.73	-7.48	30.35	9.18
10	Ceduna	-13.97	49.61	14.39	-7.62	33.82	10.67
11	Adelaide airport	-14.47	49.80	14.17	-7.22	34.55	11.09
12	Adelaide (kent town)	-13.10	45.43	13.17	-7.15	33.70	10.78
13	Loxton	-12.55	44.05	12.96	-7.18	33.34	10.67
14	Mount Gambier	-15.33	57.97	16.00	-6.58	35.54	12.02
15	Weipa	-12.42	39.06	10.95	-9.66	32.98	9.36
16	Cairns	-14.80	44.74	12.08	-9.42	33.84	9.73
17	Townsville	-13.77	43.21	12.10	-9.43	34.26	9.90
18	Cobar	-10.62	37.49	11.36	-7.64	31.19	9.49
19	Williamtown	-13.64	47.99	13.68	-7.66	34.11	10.76
20	Sydney	-16.24	53.71	14.46	-7.61	35.24	10.98
21	Canberra	-12.41	46.17	13.85	-6.95	33.24	10.92
22	Wagga Wagga	-13.00	46.34	13.43	-7.09	33.27	10.74
23	Mildura	-12.61	44.50	13.05	-7.24	32.75	10.38
24	East sale	-14.43	53.82	15.34	-6.51	36.32	12.19
25	Scottsdale	-13.64	51.53	15.02	-5.42	40.00	13.47
26	Bicheno	-14.81	52.11	14.87	-4.91	46.38	15.68
27	Lake Leake	-16.06	60.36	16.45	-5.11	36.03	12.84
28	Hobart	-15.97	56.29	15.78	-4.57	50.36	17.77
29	Strathgordon village	-13.08	52.11	15.29	-4.66	33.83	12.35
30	Flinders Island	-18.05	64.07	17.15	-6.19	38.66	13.02
Average		**-13.66**	**47.09**	**13.38**	**-7.39**	**34.47**	**10.91**

For each PET model, the magnitudes of average potential changes in PET display substantial variation across the locations, with both models suggesting the lowest average changes at arid locations and highest average changes at humid locations, as was also observed in Table 3-3. Specifically, the Penman-Monteith model identified the highest average PET change at Flinders Island (+17.15 %), with the lowest average change at Alice Springs (+9.80 %). The Priestley-Taylor model identified the highest average change at Hobart (+17.77 %), with the lowest at Tennant Creek (+7.09 %).

To further investigate how potential change in PET varies with different climatic conditions, we now focus on the associations between the PET responses and the baseline levels of the four climate variables for each month of the year and across the 30 study sites. Starting with the Penman-Monteith model (Figure 3-3), it is clear that the PET response displays a clear association with the baseline levels of climate variables, with higher magnitude of responses for locations that are cooler (low T), more humid (high RH), and receiving less solar radiation (low R_s). The highest associations can be found with T (Figure 3-3a), with the monthly changes in PET ranging from -30.2% to +98.3 % for the lowest baseline T value of 5.0 °C, compared to a range of -13.3 % to +46.6 % for the highest baseline T of 30.3 °C. Similarly, the range of Penman-Monteith PET responses also shows clear decreases with baseline R_s (Figure 3-3c), and increases with baseline RH (Figure 3-3b). The baseline u_z (Figure 3-3d) levels show no obvious impact on the PET responses.

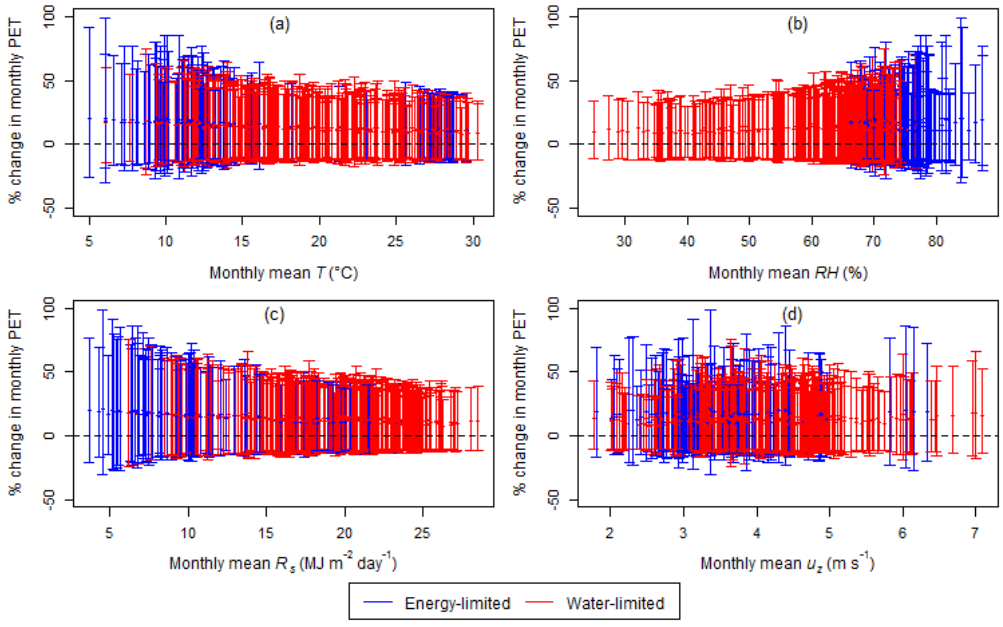


Figure 3-3: Ranges of monthly PET responses obtained from the Penman-Monteith model, plotted against the monthly baseline levels of (a) temperature, (b) relative humidity, (c) solar radiation and (d) wind speed at 30 study sites. Each vertical line represents the range of all potential changes in PET in response to the full set of climate perturbations for a single month at a single location, with the mean represented by the point on the line. The classification of energy- and water-limited months is based on the corresponding monthly PET/P ratios.

The potential responses in PET obtained from Priestley-Taylor was also investigated (Figure 3-4), and results are consistent with the results from the Penman-Monteith model, although the overall ranges of responses were smaller for each variable as anticipated from the results in Table 3-3. Interestingly, regardless of the choice of PET model, the range of PET responses at the monthly scale is larger than the range for the annual scale suggesting greater uncertainty at higher temporal resolutions.

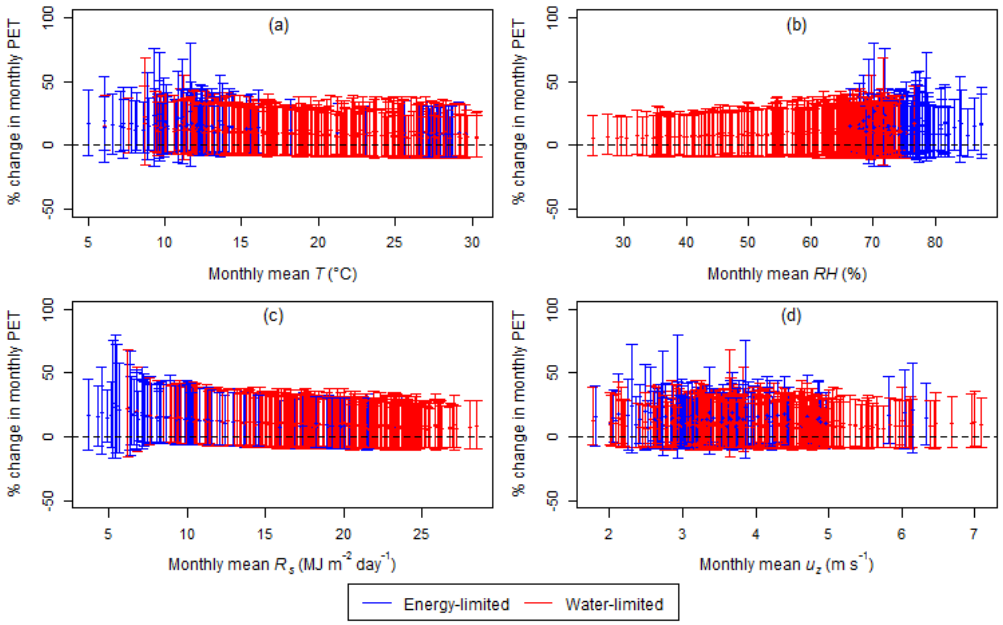


Figure 3-4: Range of monthly PET responses obtained from the Priestley-Taylor model, plotted against the monthly baseline levels of (a) temperature, (b) relative humidity, (c) solar radiation and (d) wind speed at 30 study sites. Each vertical line represents the range of all potential changes in PET in response to the full set of climate perturbations for a single month at a single location, with the mean represented by the point on the line. The classification of energy- and water-limited months is based on the corresponding monthly PET/P ratios.

In addition to assessing the impact of baseline climatic conditions, we are also interested in the role of baseline hydrological conditions (represented by the PET/P ratio at each study site) on the potential responses in PET. Since the hydrological conditions can vary substantially over the course of a year for each study site, for this analysis we focused on the PET/P ratios estimated on a monthly basis, and thus differ from the long-term PET/P ratios presented in Table 3-1. These results are also shown in Figures 3-3 and 3-4, with red-colored bars denoting water-limited conditions, and blue-colored bars denoting energy-limited conditions. These figures show that the magnitude of

potential responses in PET is generally larger under energy-limited conditions, regardless of the choice of PET model. In contrast, for water-limited conditions, the potential responses in PET only vary within approximately half of the entire range obtained from each PET model. However, when exploring the association with temperature (Figures 3-3a and 3-4a) in more detail, the magnitude of responses in PET is in fact lowest for energy-limited conditions during warm months (i.e. when T > 25 °C, corresponding to the monsoonal summer months in the northern parts of Australia), and highest for the energy-limited conditions during cool months (i.e. when T < 15 °C, corresponding to the wet winter months in southern Australia). This highlights that it is the atmospheric temperature, rather than the level of aridity, that appears to affect the potential responses in PET. This finding leads to a different interpretation to previous studies, which indicated that the dominant drivers of spatially varying PET include aridity (Tabari and Hosseinzadeh Talaee, 2014) and wind speed (Gong et al., 2006).

The above results also have potential implications on likely AET changes in a future climate. In particular, the above analysis shows that cool and humid regions and seasons appear to show the greatest potential responses in PET, and given that water is not expected to be limited for these cases, the ratio between AET and PET is also likely to be the greatest for these cases. As such, one might expect a greater change to AET occurring at the locations and during times of the year where PET is most sensitive to changes in climate.

As a potential limitation to the above analysis, some reliability issues of the Penman-Monteith model have been discussed in a recent study by Milly and Dunne (2016), which suggested that the Penman-Monteith model may overestimate the potential changes in PET in these energy-limited regions relative to a GCM-based AET benchmark. They concluded that the potential changes in ET would be better described by GCMs than 'off-line' PET models (such as the two models used in this study), as GCMs can explicitly consider more complex atmospheric processes, such as the interaction between CO_2 and stomatal conductance. Nevertheless, it should be noted that the current reliability of GCMs in simulating ET is also questionable, due to

the uncertainty in representing soil moisture and radiative energy at the evaporative surface (e.g. Seneviratne et al., 2013;Boé and Terray, 2008;Barella-Ortiz et al., 2013). In addition, due to the coarse scale of GCM output, downscaling is generally required to post-process output for use at local and regional scales, which often adds further bias and uncertainties to the GCM simulation and largely limits their applicability (e.g. Chen et al., 2012;Diaz-Nieto and Wilby, 2005). Therefore, although GCM results may be more suitable for large-scale assessments, catchment-scale climate impact assessments are likely to be informed by 'off-line' PET models for the foreseeable future. Consequently, the estimated potential changes in PET shown in this study will remain relevant for climate impact assessments conducted using these models.

3.4.2 Relative importance of climate variables affecting PET for different climate zones

We now explore the relative importance of each climate variable on overall PET sensitivity, by first visualizing the conditional responses of PET when holding each variable constant at its historical level while perturbing the remaining variables, and then comparing this to the unconditional estimates of all potential responses in PET (as shown in Figures 3-3 and Fig. 3-4). Figure 3-5 shows the ranges of the monthly unconditional responses in PET (dashed lines) and the ranges of the monthly responses conditioned on zero-change in each of T, RH, R_s and u_z (solid lines) for the Penman-Monteith model, plotted against the monthly baseline levels of the four climate variables at the 30 study sites.

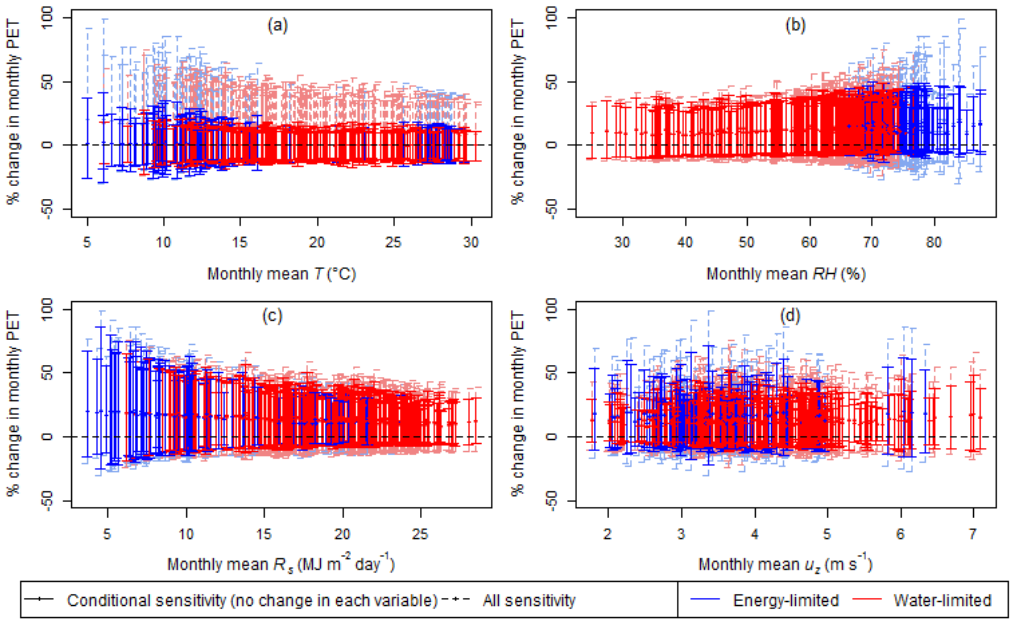


Figure 3-5: Range of monthly PET responses from the Penman-Monteith model, plotted against the monthly baseline levels of (a) temperature, (b) relative humidity, (c) solar radiation and (d) wind speed at 30 study sites. Each dashed (solid) line represents the range of all potential changes in PET in response to the full set of climate perturbations (conditioned on no-change in each climate variable) for a single month at a single location. The corresponding means are represented by the points on the lines. The classification of energy- and water-limited months is based on the corresponding monthly PET/P ratios.

The figure suggests that perturbations in *T* have the greatest impact on the potential changes in PET compared to other climate variables (Figure 3-5a), contributing to at least 45 % of the entire range of PET responses compared to the unconditional results. Humidity also plays a significant role, although only for higher humidity levels (contributing up to 57 % of the entire range of PET responses) with relatively minor influence for the less humid catchments (Figure 3-5b). In contrast, the role of solar radiation (Figure 3-5c) and wind (Figure 3-5d) is generally minor, with the range of unconditional responses being only slightly wider than the range of conditional responses.

A similar analysis was conducted for the Priestley-Taylor model (Figure 3-6), and shows somewhat different results compared to those obtained for the Penman-Monteith model. Consistent with Figure 3-5a, temperature has the greatest impact, but in this case contributes up to 85 % of the overall variability in PET responses (Figure 3-6a). As a result, the range of PET changes contributed by the remaining variables (i.e. conditional responses

with no-change in temperature) is much smaller. Unlike in Figure 3-5b, the role of relative humidity does not appear to increase significantly with increasing baseline humidity (Figure 3-6b) and in general contributes less than 33 % of the overall variability. The lower impact of *RH* on Priestley-Taylor PET compared to the impact on Penman-Monteith PET can be related to the structure of Priestley-Taylor model, which does not consider the aerodynamic processes, so that the impact of *RH* on PET through these processes is not accounted (see Eqn. 2.7, 2.15 and 2.16 in Appendix 3A). The role of solar radiation appears to be somewhat larger for high baseline solar radiation values (Figure 3-6c) and wind is shown to have no impact as expected, since wind is not an input into the Priestley-Taylor model (Figure 3-6d). However, it is worth noting that although the Priestley-Taylor model does not consider wind as an input variable, the range of unconditional responses of PET is slightly wider than the range of responses conditioned on no-change in wind. This is because the conditional responses were estimated with only a subset of all climate perturbations (Section 3.3.4), which may not consist of the entire range of perturbation in each of the other three climate variables.

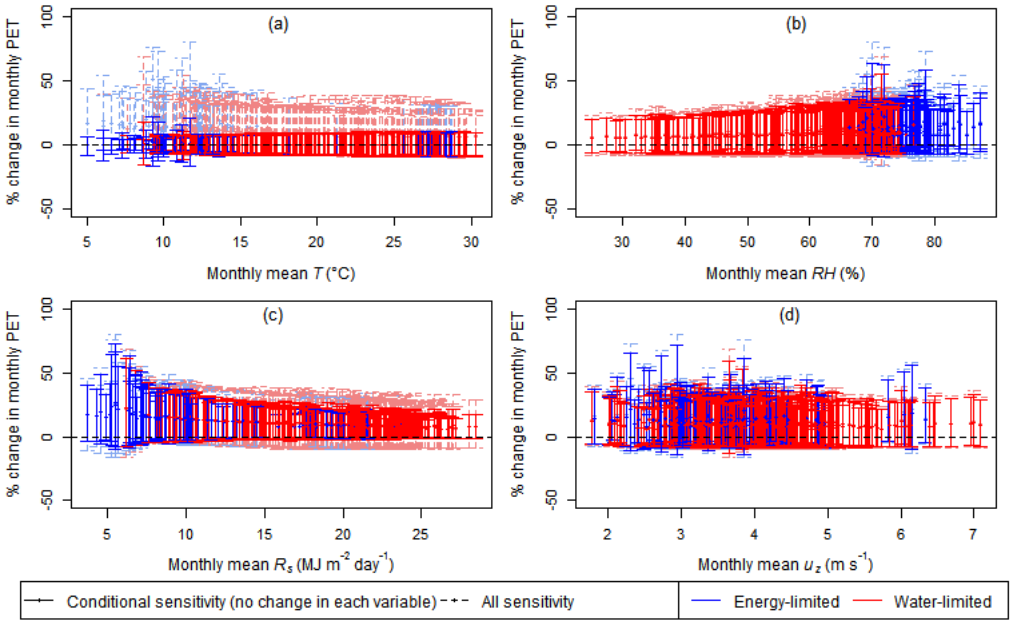


Figure 3-6: Range of monthly PET responses from the Priestley-Taylor model, plotted against the monthly baseline levels of (a) temperature, (b) relative humidity, (c) solar radiation and (d) wind speed at 30 study sites. Each dashed (solid) line represents the range of all potential change in PET in response to the full set of climate perturbations (conditioned on no-change in each

climate variable) for a single month at a single location. The corresponding means are represented by the points on the lines. The classification of energy- and water-limited months is based on the corresponding monthly PET/P ratios.

A more formal quantitative measure of the relative importance of each climate variable for PET is provided by the Sobol' indices. Figure 3-7 shows the Sobol' first-order indices of the Penman-Monteith PET to changes in the four climate variables at the annual scale, as well as their interactions. The first-order indices are plotted against the baseline levels of each climatic variable to observe the impact of baseline climate conditions. For presentation purposes, the baseline levels are represented by the rank of the baseline annual average value of each variable, rather than the absolute level of each climate variable across the 30 study sites. The Sobol' indices in the figure show that *T* is generally the most important variable for PET, with index values ranging from 0.46 to 0.62. Since the Sobol' indices suggest the partitioning of the total variance of PET, these results are consistent with Figure 3-5a, which suggests that perturbations in *T* contribute to at least 45 % of the variation in the estimated changes in PET. The role of wind and humidity in affecting the sensitivity values is also evident, with wind being the second-most important variable (with Sobol' indices up to 0.42) for sites with low baseline humidity, and humidity being the second-most important variable (with Sobol' indices up to 0.47) for sites that have high humidity (Figure 3-7b). Solar radiation is generally the variable with the lowest Sobol' indices, with the largest contributions (up to 18 %) can be observed for warm catchments (Figure 3-7a).

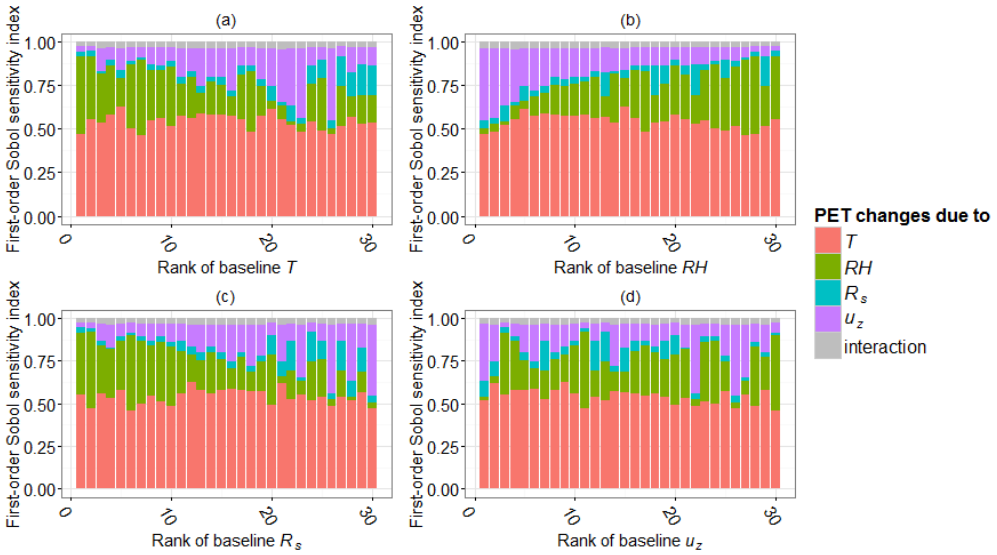


Figure 3-7: Sobol' first-order sensitivity indices of the Penman-Monteith model for changes in the four climate variables (colored) and their interaction effects (grey), plotted against the ranking of the average level of each climate variable at 30 study sites

The Sobol' sensitivity indices are also presented for the Priestley-Taylor model (Figure 3-8), and show substantial differences compared to those for the Penman-Monteith model. Temperature exhibits the largest sensitivity score in most cases, and ranges from 0.44 to 0.83. The relative role of temperature varies most clearly as a function of both the baseline temperature (Figure 3-8a) and the baseline solar radiation values (Figure 3-8c), with temperature being particularly important for low temperature and low solar radiation sites. As temperature and radiation increase, the relative role of solar radiation becomes more important, reaching Sobol' index values of up to 0.49. In contrast, the role of relative humidity is generally minor (with Sobol' indices within the range 0.03-0.1) and does not appear to vary as a function of baseline conditions. Finally, the role of wind is absent, given that this variable is not included as part of the Priestley-Taylor equation.

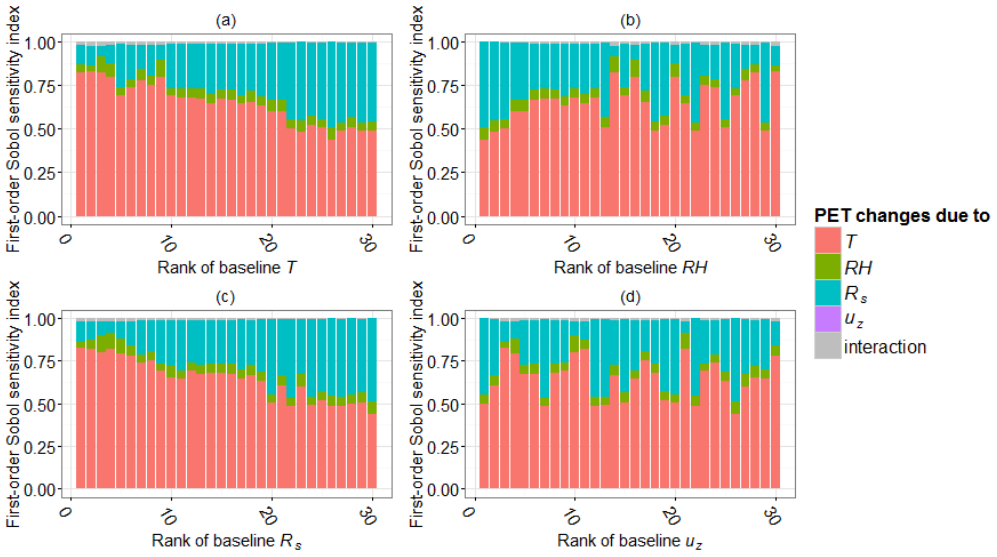


Figure 3-8: Sobol' first-order sensitivity indices of the Priestley-Taylor model for changes in the four climate variables (colored) and their interaction effects (grey), plotted against the ranking of the average level of each climate variable at 30 study sites

The differences between the Penman-Monteith and Priestley-Taylor models highlight the different physical assumptions underpinning the models, with aerodynamic processes being important for the Penman-Monteith model as indicated by the relative importance of *RH* and u_z for this model, whereas R_s has a critical role in the Priestley-Taylor model, which is closely linked to the emphasis of radiative energy as the energy source for ET in the model.

Finally, comparing Figure 3-7 and Figure 3-8, it is apparent that the interactions among the four climate variables on PET (shown as grey bars) are greater in the Penman-Monteith model compared to the Priestley-Taylor model. Specifically, these interactions contribute fractions of 0.03-0.04, and 0-0.02 of the total variance in PET for the Penman-Monteith and Priestley-Taylor models, respectively. The relative magnitude of the interaction effects in the two models can be again related to their structural differences: the higher interaction effects in Penman-Monteith can be a result of the larger number of variables in this model compared with those in the Priestley-Taylor model.

It is difficult to assess the consistency of these sensitivity results with existing literature, given the different methodologies and datasets used in other studies. Although most PET sensitivity studies used only the Penman-

Monteith PET model, there is still substantial discrepancy in results depending on the specific implementations of sensitivity analysis. For example, Gong et al. (2006) perturbed each of temperature, wind speed, relative humidity and solar radiation within ±20 % for the Changjiang basin in China, and observed that that relative humidity was generally the most important variable driving PET, followed by solar radiation, temperature and wind speed. This contrasted with our results from the Penman-Monteith model, which showed temperature as the most important variable and solar radiation as the least important variable for almost all the stations analyzed, and may be attributable to the different baseline climates as well as the perturbation ranges used for the sensitivity analysis between the two studies.

The results of our study were more consistent with Goyal (2004), who concluded that PET is most sensitive to potential changes in temperature for an arid region in India, by applying a ±20 % perturbation on each of temperature, solar radiation, wind speed and vapor pressure. In contrast, Tabari and Hosseinzadeh Talaee (2014) also used a ±20 % perturbation range, but on only three climate variables, namely temperature, wind speed and sunshine hours, for several climate regions in Iran. Their study concluded that the catchment aridity was a major determinant of the sensitivity to temperature, wind speed and humidity, whereas our analysis highlights the importance of baseline temperature and humidity, rather than the aridity (or water- or energy-limited status of the catchment) as a key driver.

PET sensitivity can further diversify by the choice of PET models, as illustrated in McKenney and Rosenberg (1993), in which the percentage changes in PET due to a +6 ◦C change can differ up to around 40 %, when estimated with eight alternative PET models. This lack of consistency in the relative importance of the climate variables for PET is not surprising given the findings of our study, as the results are strongly dependent on the design of the sensitivity analysis experiment, including the choice of study sites and study periods, the input climate variables considered, and the ways to perturb them (i.e. the choice of global or local perturbation and the ranges of perturbation in different input variables).

Nevertheless, the sensitivity results from this study suggest some distinct spatial patterns of the relative importance of different climate variables in Australia. Since the Penman-Monteith model is the most comprehensive physically-based PET model, the above regionalization of the PET sensitivity from this model can be used as a benchmark to identify the key climate variables for estimating PET under potential climate change. This information can be particularly useful to suggest the potential suitability of specific PET models for regional applications. For example, since the Penman-Monteith PET showed higher sensitivity to wind at dry locations (Figure 3-7b), it is expected that wind-dependent PET models (such as Penman and Penman-Monteith) would be more appropriate for predicting PET at these locations. In contrast, using simpler models that do not consider wind as an input (such as Priestley-Taylor) can be problematic for these locations. Although this study only examined two PET models, the results suggest that simpler empirical models are likely to ignore some potential dynamics and interactions within the climate variables, which makes them less preferred for PET estimation under changing climates.

Another particular issue in the selection of one or several PET models under a changing climate arises from considering the current reliability of available climate projections, as the models can show high levels of sensitivity to variables for which we currently do not have high-quality climate projections. For example, for a given emissions scenario, there is reasonable confidence in projections of temperature and relative humidity in Australia, but less confidence in projections of solar radiation and wind (Flato et al., 2013;CSIRO and Bureau of Meteorology, 2015). However the radiation-based Priestley-Taylor model can show high sensitivity to solar radiation, particularly for warm locations with high baseline solar radiation (Figures 3-8a and 3-8c), due to a particular emphasis on radiative energy and thus the empirical relationships between PET and solar radiation. Similarly, the Penman-Monteith model can exhibit higher sensitivity to wind for locations with low relative humidity (Figure 3-7b). Therefore, the use of GCM projections at these locations may lead to significant uncertainty in PET estimates due to the uncertainty in the driving variables.

3.5 Summary and conclusions

In this study, we used a global sensitivity analysis to investigate the sensitivity of PET and the relative importance four climatic variables which influence PET (T, RH, R_s and u_z) under plausible future changes in these variables. The sensitivity analysis was conducted at 30 Australian case study locations within different climate zones to understand the impact of varying baseline hydro-climatic conditions. For the sensitivity analysis, the historical climate data at each study site were first perturbed to represent a large number of plausible climate change conditions, and then the responses in PET were estimated with both the Penman-Monteith and Priestley-Taylor models, from which the sensitivity of PET was analysed. The key results are as follows:

- In general PET is most sensitive to potential changes in climate in regions with lower temperature, less solar radiation and greater humidity, where two-fold greater magnitude of changes in PET are expected compared to other locations in Australia.
- Within the plausible perturbations in T, RH, R_s and u_z, PET is generally most sensitive to T. The relative importance of the other climate variables varies substantially with the PET models. R_s has a dominant role in the radiation-based Priestley-Taylor model, highlighting the importance of radiative energy in the model. In contrast, the importance of RH and u_z are comparable for the Penman-Monteith model, whereas R_s has only little impact, reflecting the contribution of aerodynamic energy.
- The relative importance of climate variables in influencing PET depends very clearly on baseline climatic conditions. From Penman-Monteith, locations that are warmer, drier and receiving more solar radiation generally show greater sensitivity to u_z and lower sensitivity

to *RH*. For Priestley-Taylor, the importance of *T* increases while that of R_s decreases for cooler locations and locations receiving less solar radiation.

The global sensitivity analysis used in this study is a powerful tool for providing a comprehensive and consistent measure of PET sensitivity to different climatic variables, considering a wide range of possible changes in climate, across different models with different data requirements. However, we have identified space for improvements in further implementations. For example, the bounds of perturbation for each climate variable can have a substantial impact on PET sensitivity, and thus their selection requires careful justification (for example see Whateley et al., 2014;Shin et al., 2013). Therefore, alternative lines of evidence on possible changes in climate should be considered in setting these bounds: for example, the results of ensemble climate models (e.g. Collins et al., 2013), the impact of low-frequency climatic modes (e.g. Chen et al., 2013;Vincent et al., 2015), as well as findings from within paleoclimatology records (e.g. Ault et al., 2014;Ho et al., 2015).

The analysis in this study also lends itself to scenario-neutral analyses (Brown et al., 2012;Prudhomme et al., 2010), although the full implications on specific impacts of hydrological systems (e.g. flood risk, water supply, etc) would require the sensitivity analysis to be propagated to runoff via explicitly modelling the interaction between ET and rainfall-runoff processes (e.g. Garcia and Tague, 2015;Roy et al., 2016). Furthermore, potential changes to precipitation, which were not analyzed here but which can have a significant impact on future runoff, would need to be considered. Within this context, the incorporation of alternative lines of evidence can therefore not only be used to define the bounds of the perturbations, but can also be superimposed onto the exposure space (e.g. as in Prudhomme et al., 2013a;Culley et al., 2016) to provide insight into the likelihood of possible changes. The outcomes of our study can feed into such a scenario-neutral analysis by providing guidance on the variables that are likely to be most important for a particular location, as

well as providing insights on the potential implications of using alternative PET models on the overall sensitivity results.

Appendix 3A Supplementary to Chapter 3

1. *Sobol' sensitivity analysis (Sobol' et al., 2007)*

Sobol' is considered a variance-based method, in which the total variance in a model output due to changes in its inputs is estimated with a Monte-Carlo approach. To estimate the variances, a large number of samples is firstly drawn by varying all input variables at the same time, and then a Sobol' sequence is constructed by re-sampling from within these Monte-Carlo samples (Saltelli et al., 2010). According to Sobol' et al. (2007), to estimate the Sobol' first-order and total-order indices with a Monte-Carlo sample size of n consisting of p input variables, a Sobol' sequence with a total of $n.(p+2)$ samples is required, i.e. with $n.(p+2)$ model evaluations.

The total variance of model output is partitioned to the contribution of each individual input variable (i.e. first-order effects), as well as their interactions (i.e. higher-order effects), as follows (equation adapted from Zhang et al., 2015):

$$V_Y = \sum_{i=1}^{n} V_i \quad + \quad \sum_{i<j} V_{ij} + \sum_{i<j<k} V_{ijk} \ldots + V_{1,2\ldots,n} \tag{1.1}$$

Individual effects **Interactions**

The outputs from the Sobol' method comprise (equations adapted from Nossent et al., 2011):

1) First-order sensitivity index, which quantifies the individual contribution of each input variable to the total variance of the model's output:

$$S_i = \frac{V_i}{V_Y} \tag{1.2}$$

2) Second- and higher-order sensitivity indices, which quantify the contribution of interactions among two or more input variables to the total variance of the model's output:

For second-order: $S_{ij} = \frac{V_{ij}}{V_Y}$ (1.3)

For higher-order: $S_{ij\ldots n} = \frac{V_{ij\ldots n}}{V_Y}$ (1.4)

3) Total sensitivity index, which quantifies the contribution of each input variable, including its individual effect, as well as all its interactions with other input variables, to the total variance of the model's output:

$$S_{Ti} = S_i + \sum_{j \neq i} S_{ij} = 1 - \frac{V_{\sim i}}{V_Y} \quad (1.5)$$

From Eqns. 1.1 to 1.4, the sum of individual effects of all input variables and all their interactions equals one (adapted from Zhang et al., 2015):

$$1 = \underbrace{\sum_{i=1}^{n} S_i}_{\text{Individual effects}} + \underbrace{\sum_{i<j} S_{ij} + \sum_{i<j<k} S_{ijk} \ldots + S_{1,2\ldots,n}}_{\text{Interactions}} \quad (1.6)$$

2. *Penman-Monteith PET model (FAO-56) (as in McMahon et al., 2013)*

The Penman-Monteith PET model (FAO-56) is given as:

$$ET = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T_a + 273} u_2 (v_a^* - v_a)}{\Delta + \gamma(1 + 0.34u_2)} \quad (2.1)$$

The processes for estimating each of the variables in this equation are described in the following sections.

2.1 Estimating Δ in Eqn. 2.1

Δ is the slope of vapor pressure curve in kPa°C^{-1}, which is estimated by:

$$\Delta = \frac{4098[0.6108\exp(\frac{17.27 * T_a}{T_a + 237.3})]}{(T_a + 237.3)^2} \quad (2.2)$$

In Eqn. E2.2, T_a is the average daily temperature in °C, calculated as:

$$T_a = \frac{T_{max}+T_{min}}{2} \quad (2.3)$$

2.2 Estimating R_n in Eqn. 2.1

R_n is the net incoming solar radiation at the evaporative surface in MJ.m^{-2}.day^{-1}, which is estimated by:

$$R_n = R_{ns} - R_{nl} \quad (2.4)$$

In Eqn. 2.4, R_{ns} is the net shortwave solar radiation, estimated by:

$$R_{ns} = (1-\alpha)R_s \quad (2.5)$$

In Eqn. 2.5, α is the albedo at evaporative surface which is fixed at 0.23 in this equation, and R_s is the measured or estimated incoming solar radiation in MJ.m^{-2}.day^{-1}. R_{nl} is the net outgoing longwave radiation, estimated as:

$$R_{nl} = \sigma(0.34 - 0.14v_a^{0.5})\frac{(T_{max}+237.2)^4+(T_{min}+237.2)^4}{2}(1.35\frac{R_s}{R_{s0}} - 0.35) \quad (2.6)$$

In Eqn. 2.6: σ is is Stefan-Boltzmann constant = = 4.903*10^{-9} MJ.m^{-2}.day^{-1} °K^{-4}, v_a is the mean daily actual vapor pressure in kPa, R_{s0} is the clear-sky radiation in MJ.m^{-2}.day^{-1}. v_a and R_{s0} estimated by Eqns. 2.7 and 2.8, respectively:

$$v_a = \frac{v_a^*(T_{max})\frac{RH_{max}}{100}+v_a^*(T_{min})\frac{RH_{min}}{100}}{2} \quad (2.7)$$

$$R_{s0} = (0.75 + 2\times10^{-5}Elev)R_a \quad (2.8)$$

In Eqn. 2.8, *Elev* is the ground elevation above sea level at the measurement location, and R_a is the extraterrestrial solar radiation in MJ.m^{-2}.day^{-1}, estimated as:

$$R_a = \frac{1440}{\pi}G_{sc}d_r^2(\omega_s \sin(lat)\sin(\delta) + \cos(lat)\sin(lat)\sin(\omega_s)] \quad (2.9)$$

In Eqn. 2.9, G_{sc} is the solar constant = 0.0820 MJ.m^{-2}.min^{-1}, *lat* is the latitude in radiance, d_r is the inverse relative distance between Earth and Sun, δ is the solar declination in radians, and ω_s is the sunset hour angle in radians, The d_r, δ and ω_s are estimated as follows:

$$d_r^2 = 1 + 0.033\cos(\frac{2\pi}{365}DoY) \text{ with } DoY \text{ as the day of the year} \quad (2.10)$$

$$\delta = 0.409\sin(\frac{2\pi}{365}DoY - 1.39) \quad (2.11)$$

$$\omega_s = arcos[-\tan(lat)\tan(\delta)] \quad (2.12)$$

2.3 Estimating other variables in Eqn. 2.1

- G is negligible for daily time step.

- γ is the psychrometric constant in kPa°C^{-1}, estimated as:
$$\gamma = 0.00163\frac{P}{\lambda} \text{ where } P \text{ is the pressure at elevation } z \text{ meters} \quad (2.13)$$

- u_2 is the daily average wind speed measured at 2 meters in m.s^{-1}, which can be estimated from the measured wind speed at z meters as:
$$u_2 = u_z\frac{\ln(\frac{2}{z_0})}{\ln(\frac{z}{z_0})} \text{ where } z_0 \text{ is the roughness height in meters} \quad (2.14)$$

- (v_a^* - v_a) is the vapour pressure deficit in kPa, in which v_a is the mean daily actual vapor pressure in kPa, estimated as Eqn. 2.7; v_a^* is the daily saturation vapor pressure in kPa, estimated as:
$$v_a^* = \frac{v_a^*(T_{max}) + v_a^*(T_{min})}{2} \quad (2.15)$$

In Eqn. 2.15, $v_a^*(T_{max})$ and $v_a^*(T_{min})$ are the vapor pressures at temperatures T_{max} and T_{min} in °C are estimated with:

$$v_T^* = 0.6108\exp[\frac{17.27T}{T+237.3}] \quad (2.16)$$

3. Priestley-Taylor PET model (as in McMahon et al., 2013)

The Priestley-Taylor PET model is given as:

$$ET = \alpha_{PT} * [\frac{\Delta}{\Delta+\gamma}\frac{R_n}{\lambda} - \frac{G}{\lambda}] \quad (3.1)$$

where:

- α_{PT} is the albedo specifically used for the Priestley-Taylor model, since an evaporative surface of reference crop was assumed, this has a value of 1.12 which was for a similar surface of short grass (See Table S8 of the supplementary of McMahon et al., 2013),

- Δ is the slope of vapor pressure curve in $kPa°C^{-1}$, estimated as Eqn 2.2.

- γ is the psychrometric constant in $kPa°C^{-1}$, estimated as Eqn. 2.13.

- λ is the latent heat of vaporization, which is 2.45 $MJ.kg^{-1}$ at 20◦C.

- R_n is the net incoming solar radiation at the evaporative surface in $MJ.m^{-2}day^{-1}$, which is estimated in the same way as Eqn. 2.4.

- G is negligible for daily time step.

Appendix 3B Copy of Paper from Chapter 3

Guo, D., Westra, S., and Maier, H. R. 2017. Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones, *Hydrology and Earth Systems Sciences*, 21, 2107-2126, doi:10.5194/hess-21-2107-2017, 2017.

Hydrol. Earth Syst. Sci., 21, 2107–2126, 2017
www.hydrol-earth-syst-sci.net/21/2107/2017/
doi:10.5194/hess-21-2107-2017

Hydrology and Earth System Sciences
Open Access
EGU

Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones

Danlu Guo, Seth Westra, and Holger R. Maier

School of Civil, Environmental and Mining Engineering, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia

Correspondence to: Danlu Guo (danlu.guo@adelaide.edu.au)

Received: 26 August 2016 – Discussion started: 19 September 2016
Revised: 19 March 2017 – Accepted: 20 March 2017 – Published: 19 April 2017

Abstract. Assessing the factors that have an impact on potential evapotranspiration (PET) sensitivity to changes in different climate variables is critical to understanding the possible implications of climatic changes on the catchment water balance. Using a global sensitivity analysis, this study assessed the implications of baseline climate conditions on the sensitivity of PET to a large range of plausible changes in temperature (T), relative humidity (RH), solar radiation (R_s) and wind speed (u_z). The analysis was conducted at 30 Australian locations representing different climatic zones, using the Penman–Monteith and Priestley–Taylor PET models. Results from both models suggest that the baseline climate can have a substantial impact on overall PET sensitivity. In particular, approximately 2-fold greater changes in PET were observed in cool-climate energy-limited locations compared to other locations in Australia, indicating the potential for elevated water loss as a result of increasing actual evapotranspiration (AET) in these locations. The two PET models consistently indicated temperature to be the most important variable for PET, but showed large differences in the relative importance of the remaining climate variables. In particular for the Penman–Monteith model, wind and relative humidity were the second-most important variables for dry and humid catchments, respectively, whereas for the Priestley–Taylor model solar radiation was the second-most important variable, with the greatest influence in warmer catchments. This information can be useful to inform the selection of suitable PET models to estimate future PET for different climate conditions, providing evidence on both the structural plausibility and input uncertainty for the alternative models.

1 Introduction

Assessing changes to evapotranspiration (ET) is critical in understanding the impacts of anthropogenic climate change on the catchment water balance. ET represents the dominant loss of water from catchments worldwide, with about 62 % of global land-surface precipitation accounted for by ET (Dingman, 2015), and ET exceeding runoff in over 77 % of the global land surface (Harrigan and Berghuijis, 2016). ET is affected by climate change through a cascade of processes that begins with the increasing concentration of greenhouse gases, followed by their attendant impacts on large-scale circulation and changes to the global distribution of energy and moisture. These large-scale processes lead to local-scale changes in the atmosphere, which in turn influence the catchment water balance through a set of terrestrial hydrological processes by which precipitation is converted into actual ET (AET), runoff and groundwater recharge (Oudin et al., 2005). Other factors that can potentially affect ET under a changing climate include changing land cover patterns (e.g., Liu et al., 2008), and the CO_2 fertilization effects that can limit the rate of plant transpiration under elevated levels of CO_2 (e.g., Prudhomme et al., 2014; Milly and Dunne, 2016).

Climate impact studies that investigate the influence of climate forcings on the catchment water balance are usually based on projections of future climate represented by climate variables such as temperature and solar radiation from general circulation models (GCMs), which are converted into potential ET (PET) using one or several PET models. The PET projections are combined with GCM projections of precipitation (P), which together can be used to directly estimate the water deficit (Taylor et al., 2013; Chang et al.,

2016). Alternatively, rainfall–runoff models can be used to translate the changes in P and PET into changes in runoff (e.g., Akhtar et al., 2008; Chiew et al., 2009; Koedyk and Kingston, 2016), as well as associated information such as the impact on catchment streamflow (Wilby et al., 2006), water supply security (Paton et al., 2013, 2014) and flood risk (Bell et al., 2016). Therefore, to quantify the specific impact of changes in ET on the water balance, a good understanding of the sensitivity of PET to potential changes in its key influencing climatic variables is required (Goyal, 2004; Tabari and Hosseinzadeh Talaee, 2014). This is particularly relevant given the recent focus on "scenario-neutral" (or "bottom-up") approaches to climate impact assessment (Brown et al., 2012; Prudhomme et al., 2010; Culley et al., 2016), which require the sensitivity of a given system to potential changes in climate forcings to be estimated (Prudhomme et al., 2013a, b; Steinschneider and Brown, 2013; Kay et al., 2014; Guo et al., 2016a).

Furthermore, the sensitivity of PET can provide critical evidence in relation to identifying models that are most appropriate for PET estimation under climate change conditions, which is particularly relevant to the ongoing debate on the potential trade-off between model complexity and reliability. Complex models such as the Penman–Monteith model are often recommended for their ability to better represent the physical processes that affect PET (McVicar et al., 2012; Donohue et al., 2010; Barella-Ortiz et al., 2013). For example, the Penman–Monteith model can account for the effects of wind, and thus can help explain at least part of the observed decreases in pan-evaporation with increases in temperature in many locations globally – the "evaporation paradox" – due to the observed decreases in wind speed (Roderick et al., 2007; McVicar et al., 2008; Lu et al., 2016). However, simpler empirical models may also be preferable under some conditions, as they require a smaller number of input climate variables, which might be able to be projected with greater confidence with GCMs, and thus leading to greater confidence in the corresponding PET estimates (Kay and Davies, 2008; Ekström et al., 2007; Ravazzani et al., 2014). For example, there is reasonable confidence in projections of temperature and relative humidity in Australia for a given emission scenario, but less confidence in projections of wind due to sub-grid effects of orography and other land-surface features (Flato et al., 2013; CSIRO and Bureau of Meteorology, 2015). In these situations, models such as the Priestley–Taylor model that do not depend on wind may produce more reliable estimates of PET compared to the more complex Penman–Monteith model. Thus, the choice of climate variables to include in climate impact assessments must be informed both by the relative importance of each variable on projections of PET (e.g., Tabari and Hosseinzadeh Talaee, 2014), and the likely confidence in the projections of each variable (e.g., Flato et al., 2013; Johnson and Sharma, 2009).

Sensitivity analysis methods have been employed in a number of recent studies to assess the overall sensitivity of PET estimated by the Penman–Monteith model to potential changes in climate, as well as to better understand the relative importance of different climate variables on overall PET sensitivity. For example, Goyal (2004) found that PET was most sensitive to perturbations in temperature, followed by solar radiation, wind speed and vapor pressure, at a single study site in an arid region in India. Tabari and Hosseinzadeh Talaee (2014) also looked at the sensitivity of PET to perturbations of historical climate data from eight meteorological stations representing four climate types in Iran, and concluded that the importance of wind speed and air temperature was lower while that of sunshine hours was higher for a humid location compared to an arid location. Gong et al. (2006) found that the differences in PET sensitivity across the upper, middle and lower regions of the Changjiang (Yangtze) basin in China were largely due to contrasting baseline wind speed patterns. However, most of these PET sensitivity analysis studies focused on a limited number of study sites and/or climatic zones; therefore, the specific causes for varying PET sensitivity at different locations, such as the roles of climatic and hydrological conditions, remain unclear. Consequently, it is difficult to extrapolate our existing knowledge of PET sensitivity and the relative importance of each climate variable to new locations, which is essential for assessing the water balance at regional scales.

To address the shortcomings of existing studies outlined above, this study aims to gain an understanding of (i) the sensitivity of PET estimates to changes in the key climatic variables that influence PET, and how these sensitivity estimates are affected by varying baseline hydrologic and climatic conditions at different locations; and (ii) the relative importance of these climatic variables for PET, and how this changes with the baseline hydrologic and climatic conditions at different locations. These aims were achieved by analyzing the responses of PET to perturbations in four of its driving climatic variables, namely temperature (T), relative humidity (RH), solar radiation (R_s) and wind speed (u_z), at 30 study sites across Australia representing a range of climate zones. Both the Penman–Monteith and Priestley–Taylor models were used, as they represent different conceptualizations of the PET-related processes, with both models being widely used for climate impact assessments (Felix et al., 2013; Arnell, 1999; Gosling et al., 2011; Kay et al., 2009; Prudhomme and Williamson, 2013; Donohue et al., 2009). It is worth noting that the potential changes in one climate variable can be amplified or offset by changes in another variable (for examples see the discussions of "evaporation paradox" in Lu et al., 2016; Roderick and Farquhar, 2002), which can affect the relative importance of each variable. To account for this effect, a global sensitivity analysis method was used, with similar methods being applied to account for the impact of joint variations in the input variables on the output from a variety of environmental models, ranging from conceptual rainfall–runoff models (e.g., Tang et al., 2007a, c) to complex models that consider a number of surface–groundwater

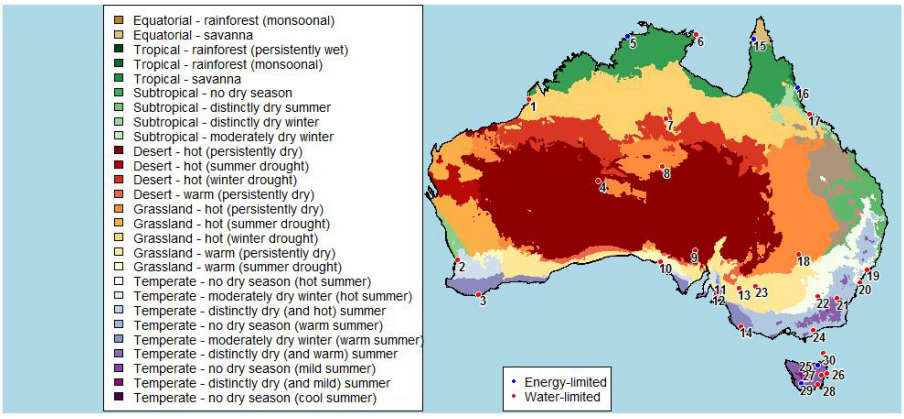


Figure 1. Locations of 30 Australian weather stations selected for analysis (see Table 1 for the full names of these weather stations), with reference to their corresponding climate classes derived following the modified Köppen classification (reproduced with data from Stern et al., 2000).

processes (e.g., Guillevic et al., 2002; van Griensven et al., 2006; Nossent et al., 2011). The results of the global sensitivity analysis in this study were presented in terms of both the range of potential changes in PET and relative sensitivity indices of each climate variable for PET, which were then used to elucidate the specific roles of varying baseline hydro-climatic conditions on influencing these sensitivity measures.

The subsequent sections of this paper are structured as follows. Section 2 introduces the data obtained from the 30 study sites required for the global sensitivity analysis. Section 3 describes the approach to the global sensitivity analysis of PET. Section 4 presents and discusses two sets of results that address the two study aims, respectively, (i) the range of estimated changes in PET in response to potential changes in temperature, solar radiation, humidity and wind, and how this changes with location, and (ii) the relative importance of the four climate variables for estimating PET, and how this changes with location. The study is summarized and concluded in Sect. 5.

2 Data

To represent contrasting hydro-climatic conditions for assessing PET sensitivity, we selected case study locations within different Köppen classes in Australia. The original Koöppen climate classification (Köppen et al., 1930; Köppen, 1931) provides a useful categorization of hydro-climatic conditions at specific locations, which is based on the long-term average levels and seasonal patterns of climatic and hydrologic variables, including temperature, relative humidity and rainfall. A "modified Köppen classification" system has been adapted for Australia (as in Stern et al., 2000) and is now widely used in climatic and hydrologic studies to identify and categorize case study locations (e.g., Johnson and Sharma, 2009; Rustomji et al., 2009; Li et al., 2014; Guo et al., 2017).

As mentioned in the Introduction, both the Penman–Monteith and the Priestley–Taylor models were used to estimate PET for the global sensitivity analyses. The estimation of PET with these models relies on temperature, relative humidity, solar radiation and (for the Penman–Monteith model only) wind speed. In addition, the rainfall data were also obtained to assess the aridity of the different locations. We limited the selection of study sites to those with 10 or more years of continuous climate data with no more than 5 % missing records over the study period. This led to a final selection of 30 weather stations (Fig. 1), with a consistent data period from 1 January 1995 to 31 December 2004. The data obtained at each site are detailed as below:

- Daily maximum and minimum temperature (T in °C), maximum and minimum relative humidity (RH in %) and wind speed (u_z in $\mathrm{m\,s^{-1}}$): data for each of these variables were obtained directly from each weather station.

- Daily solar radiation (R_s in $\mathrm{MJ\,m^{-2}\,day^{-1}}$): daily solar radiation was calculated from daily sunshine hour data (n in h) obtained from each weather station, using the Ångström–Prescott equation as in McMahon et al. (2013).

- Daily rainfall ($\mathrm{mm\,day^{-1}}$): daily rainfall data were obtained from a rain gauge at each weather station.

Table 1 shows the average values of the four PET-related climate variables, as well as the rainfall within the study period, at each of the 30 sites. As can be seen, there are large differences in the average values of each variable, highlighting large differences in the climatic conditions across the 30 sites. In addition, a quantity particularly relevant to ET processes is the long-term-averaged ratio of PET to precipitation (PET / P), which describes whether a location is water-limited (PET / $P > 1$) or energy-limited (PET / $P < 1$) (Ger-

Table 1. Names, locations and average climate conditions of the 30 weather stations over the study period (1995–2004).

No.	Study site name	Köppen class[a]	Lat (° S)	Long (° E)	Elev (m)	T^b (°C)	RH[2] (%)	R_s^b (MJ m^{-2} day^{-1})	u_z^b (m s^{-1})	Annual P^b (mm)	Annual PET^b (mm)	Annual PET / P^b
1	Broome airport	13	−17.95	122.2	7.4	26.37	65.15	21.55	3.684	865	2003	2.317
2	Perth	8	−31.93	116.0	15.4	18.54	61.72	18.95	4.519	721	1751	2.429
3	Albany	4	−34.94	117.8	68	15.08	73.59	15.20	4.382	752	1126	1.498
4	Giles	24	−25.03	128.3	598	22.70	38.40	20.29	4.380	394	2344	5.947
5	Darwin	35	−12.42	130.9	30.4	27.42	69.27	20.33	3.393	1976	1864	0.944
6	Gove	35	−12.27	136.8	51.6	26.29	75.93	19.45	3.500	1607	1660	1.033
7	Tennant Creek	13	−19.64	134.2	375.7	25.73	37.21	21.64	4.759	539	2634	4.886
8	Alice Springs	15	−23.80	133.9	546	21.18	44.53	20.79	2.352	331	1822	5.503
9	Woomera	24	−31.16	136.8	166.6	19.41	46.57	19.40	5.057	151	2153	14.24
10	Ceduna	11	−32.13	133.7	15.3	16.92	62.04	18.20	5.450	266	1723	6.478
11	Adelaide airport	12	−34.95	138.5	2	16.37	63.04	16.91	4.213	454	1410	3.107
12	Adelaide (kent town)	12	−34.92	138.6	48	16.95	61.20	16.88	3.161	569	1372	2.409
13	Loxton	12	−34.44	140.6	30.1	16.50	59.41	17.59	3.250	255	1490	5.847
14	Mount Gambier	4	−37.75	140.8	63	13.45	72.77	14.91	4.460	731	1116	1.526
15	Weipa	41	−12.68	141.9	18	26.87	72.21	19.31	3.271	2154	1782	0.827
16	Cairns	36	−16.87	145.7	3	24.80	73.00	18.98	4.352	1985	1678	0.845
17	Townsville	35	−19.25	146.8	4.3	24.53	69.45	20.27	4.304	1099	1802	1.641
18	Cobar	15	−31.48	145.8	260	19.08	50.64	19.05	2.458	398	1565	3.936
19	Williamtown	9	−32.79	151.8	9	17.84	70.57	16.07	3.927	1145	1309	1.143
20	Sydney	9	−33.94	151.2	6	18.19	67.69	15.97	5.311	1017	1393	1.369
21	Canberra	6	−35.30	149.2	578.4	13.36	65.82	16.86	3.302	590	1226	2.078
22	Wagga Wagga	9	−35.16	147.5	212	15.77	61.78	17.48	3.288	552	1436	2.602
23	Mildura	12	−34.24	142.1	50	17.11	55.62	18.24	3.604	246	1645	6.681
24	East sale	6	−38.12	147.1	4.6	13.77	72.32	14.92	4.062	529	1093	2.067
25	Scottsdale	3	−41.17	147.5	197.5	13.19	70.55	14.23	2.921	931	912	0.980
26	Bicheno	3	−41.87	148.3	11	14.69	66.68	13.69	3.319	690	966	1.401
27	Lake Leake	3	−42.01	147.8	575	9.96	75.40	13.44	3.358	732	774	1.056
28	Hobart	3	−42.83	147.5	4	12.77	65.67	14.04	4.367	483	1097	2.273
29	Strathgordon village	3	−42.77	146.0	322	10.70	77.95	11.65	2.473	2626	699	0.266
30	Flinders Island	3	−40.09	148.0	9	13.54	73.59	14.34	6.399	654	1064	1.626

Note: [a] The Köppen classes are presented with their corresponding identifiers from Stern et al. (2000), as (3) temperate – no dry season (mild summer); (4) temperate – distinctly dry (and warm) summer; (6) temperate – no dry season (warm summer); (8) temperate – moderately dry winter (hot summer); (9) temperate – no dry season (hot summer); (11) grassland – warm (summer drought); (12) grassland – warm (persistently dry); (13) grassland – hot (winter drought); (15) grassland – hot (persistently dry); (24) desert – hot (persistently dry); (35) tropical – savanna; (36) tropical – rainforest (monsoonal); (41) equatorial – savanna. [b] T is temperature, RH is relative humidity, R_s is incoming solar radiation, u_z is wind speed, P is rainfall, PET is potential evapotranspiration calculated using the Penman–Monteith model.

rits et al., 2009; McVicar et al., 2010). This ratio was estimated for each site and is also shown in Table 1 (with the point color in Fig. 1 indicating whether the location is water-limited or energy-limited). The range of PET / P values indicates substantial variations in the water availability conditions at different study sites. Note that these ratios were based on the estimates of PET from the Penman–Monteith model. Although the use of Priestley–Taylor model resulted in different PET estimates at each site, the categorization of water- and energy-limited catchments was generally consistent with those from Penman–Monteith, with different categories only shown at 4 out of the 30 study sites (sites 6, 19, 20 and 27).

3 Method

3.1 Overview

A schematic of the approach followed in study is shown in Fig. 2. As a required model input for the global sensitivity analysis, a large number of representative samples were first obtained for the four climate variables that influence PET (T, RH, R_s and u_z) at each study site, by perturbing the corresponding historical climate data (Sect. 3.2). The outputs of the global sensitivity analysis (i.e., the responses of PET) were estimated with the Penman–Monteith and Priestley–Taylor models (Sect. 3.3). To understand the PET sensitivity and the relative importance of the four climate variables in influencing PET and how these change with location, a global sensitivity analysis was conducted with the responses of PET to the climate perturbations (Sect. 3.4). This proceeded in two parts:

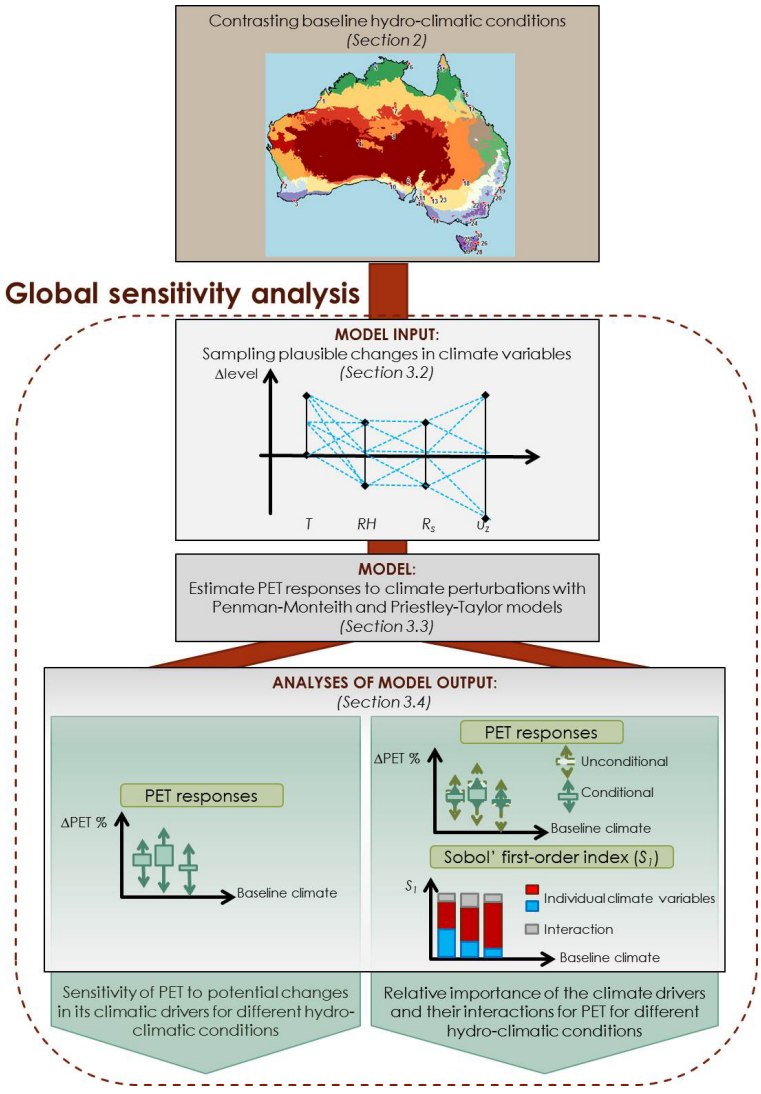


Figure 2. Schematic of the method used in this study.

1. To assess the sensitivity of PET to the climate variables, the range of percentage changes in PET in response to all the climate perturbations was estimated relative to the baseline PET at each location. To observe the impact of varying baseline hydro-climatic conditions, the ranges obtained from each PET model were also plotted against the baseline levels of each climate variable for all study sites.

2. To assess the relative importance of each climate variable, the range of percentage responses in PET to all climate perturbations in (1) was first compared to the conditional range of percentage responses in PET with holding each variable constant. This comparison enables an assessment of the relative impact of each variable on the potential responses of PET. An alternative presentation of the individual and interaction effects of the climate variables was achieved using the Sobol' method (Sobol' et al., 2007). Here, the total variance of PET was estimated based on different samples drawn from the perturbed ranges of each climate variable, and then partitioned into the individual contribution from each climate variable and their interactions (see Appendix A1 for details). The Sobol' first-order sensitivity indices were estimated and plotted against the baseline levels of each climate variable for all study sites to explore the role of varying baseline hydro-climatic conditions on the relative importance of each climatic variable for PET.

3.2 Representing plausible changes in the climatic variables

As part of the global sensitivity analysis, a large number of representative combinations of the changes in the four climate variables (T, RH, R_s and u_z) were obtained. The upper and lower bounds for perturbing each climate variable were determined based on the uncertainty bounds of projections for 2100 for Australia (Stocker et al., 2013). The selected bounds are given in Table 2, which are all slightly wider than those presented in Stocker et al. (2013) to encompass a comprehensive range of plausible future climate change scenarios. Within these bounds, samples were drawn for different combinations of changes in each climatic variable. Latin hypercube sampling (LHS) was used for this purpose due to its effectiveness in covering multi-dimensional input spaces (Osidele and Beck, 2001; Sieber and Uhlenbrook, 2005; Tang et al., 2007b).

According to Nossent et al. (2011) and Zhang et al. (2015), the sample size was selected to ensure the convergence of the first- and total-order Sobol' sensitivity indices, which occurs when the width of the 95 % confidence intervals from 1000-fold bootstrap re-sampling of the each index is below 10 % of the corresponding mean obtained from bootstrapping. Specifically, we generated different sizes of LHS samples of climate perturbations with the historical climate data from one study site, from which the PET responses were estimated using the Penman–Monteith model. The 1000-fold bootstrap estimates for the Sobol' first- and total-order sensitivity indices for each climate variable were then derived (as in Eqs. A2 and A5 in Appendix A1, respectively) for each sample size. It was observed that both the Sobol' indices began to converge when the sample size exceeded 5000, and this was therefore used as the LHS sample size for all the sensitivity experiments in this study. Based on this sample size, a total of 30 000 Sobol' samples were compiled as required to estimate the first- and total-order indices (as detailed in Appendix A1), which correspond to 30 000 climate perturbations to be used to test PET sensitivity.

To generate time series of perturbed climate data, the 30 000 joint perturbations to the four climate variables obtained by LHS were treated as change factors, and applied to the time series of daily values of the corresponding historical data. Rather than using a single daily mean value of temperature and relative humidity, the two PET models used in this study require both the daily minimum and maximum values; therefore, each pair of temperature variables and relative humidity variables was considered jointly and thus perturbed by the same amount for each day. In addition, to ensure physical plausibility of the perturbations, the daily maximum and minimum values of relative humidity were capped at a maximum of 100 %.

Table 2. Plausible perturbation bounds for each climate variable relative to their current levels.

Climate variable	Perturbation range
T	0 to +8 °C
RH	−10 to +10 %
R_s	−10 to +10 %
u_z	−20 to +20 %

Note: T is daily temperature, RH is daily relative humidity, R_s is daily incoming solar radiation, u_z is daily wind speed.

3.3 Estimating PET responses to climate perturbation

To represent the responses in PET as a result of the climate perturbations, we used both the Penman–Monteith and Priestley–Taylor models, which provide contrasting process representations to estimate PET. The Penman–Monteith model is often referred to as a combinational model, as it combines the energy balance and mass transfer components of ET, and takes into account vegetation-dependent processes such as aerodynamic and surface resistances (Eqs. A6 in Appendix A1). The model requires input of six climate variables, namely, T_{max}, T_{min}, RH_{max}, RH_{min}, R_s and u_z. The Priestley–Taylor model consists of a simpler structure, considering only the energy balance, without consideration of mass transfer or any impact from vegetation (Eq. A23 in Appendix A3). Therefore, the Priestley–Taylor model is also referred to as a radiation-based model. The model only requires five climate variables, including T_{max}, T_{min}, RH_{max}, RH_{min} and R_s.

To minimize the potential confounding effects of differences in vegetated surface, the evaporative surface was assumed to be the reference crop for all study sites; therefore, it was possible to use the FAO-56 version of the Penman–Monteith model (Allen et al., 1998). The detailed formulations of the two PET models, as well as the relevant constants and assumptions, are included in McMahon et al. (2013). Both models were implemented using the R package *Evapotranspiration* (http://cran.r-project.org/web/packages/Evapotranspiration/index.html) (Guo et al., 2016b). From each model, two sets of estimated PET were obtained: (i) a single set of baseline (historical) PET data at each study site with the historical climate data and (ii) 30 000 sets of perturbed PET data at each study site corresponding to the 30 000 sets of perturbed climate data obtained using LHS, as detailed in Sect. 3.2.

3.4 Analyses of PET sensitivity

To assess the overall sensitivity of PET to plausible climate change, we first estimated the annual average percentage changes in PET (relative to the baseline PET) using all

climate perturbations at the 30 study sites, with estimates from both the Penmen–Monteith and Priestley–Taylor models. A closer investigation of how PET sensitivity varies with baseline climate was conducted by plotting the ranges of all monthly PET responses against the average levels of each climate variable, for all study sites and all months. The reason for the choice of monthly timescale is that for some study sites, the climate can vary substantially by season; therefore, an annual analysis might obscure important sub-annual effects. To assess the relative importance of each climate variable for PET estimation from each model, we first compared the ranges of the two sets of PET changes, namely,

1. the range of all potential changes in PET obtained from the entire 30 000 sets of climate perturbations from LHS; and
2. the conditional ranges of potential changes in PET assuming no change in one of the climate variables. This was obtained with using a subset of all climate perturbations used in (1), for which the changes in the specific conditioning climate variable were close to zero (within $\pm 0.1\,^{\circ}\mathrm{C}$ for T, and within $\pm 0.1\,\%$ for the other three variables).

In this way any difference between (1) and (2) was purely contributed by the impact of changing the specific conditioning climate variable. To quantify and compare the relative importance of each climate variable, we then utilized the Sobol' method, which was implemented within the R package *sensitivity* (https://cran.r-project.org/web/packages/sensitivity/index.html). We estimated the Sobol' first-order sensitivity indices (as in Eq. A2, Appendix A1) to assess the role of each individual climate variable for each PET model, at the 30 study sites. The sum of all interaction effects was also calculated for each location as the difference between the sum of all first-order indices and one (Eq. A6, Appendix A1). The Sobol' first-order indices were then plotted against the baseline levels of each climate variable at the 30 study sites, to assess how the relative importance changes with the baseline climatic conditions.

4 Results and discussion

4.1 Ranges of potential changes in PET in response to potential climate change for different climate zones

We start by assessing the potential changes in PET in response to the full set of climate perturbations at the 30 study sites at the annual timescale, using both the Penman–Monteith and Priestley–Taylor models. The results are presented in Table 3 in terms of the minimum, maximum and average changes of PET relative to the 1995–2004 baseline, in response to the 30 000 sets of climate perturbation at each study site. The two models suggest similar average PET changes at most locations, with the average changes obtained from the Penman–Monteith model across all the locations ($+13.38\,\%$) being slightly higher than that for the Priestley–Taylor model ($+10.91\,\%$). Greater differences between the two models were observed when considering the ranges of changes. In particular, the minimum and maximum values (averaged across all the 30 sites) were -13.66 and $+47.09\,\%$ for the Penman–Monteith model, respectively, compared to -7.39 and $+34.47\,\%$ for the Priestley–Taylor model. This corresponds to a range for the Penman–Monteith model being approximately 45 % wider than that of the Priestley–Taylor model.

For each PET model, the magnitudes of average potential changes in PET display substantial variation across the locations, with both models suggesting the lowest average changes at arid locations and the highest average changes at humid locations, as was also observed in Table 3. Specifically, the Penman–Monteith model identified the highest average PET change at Flinders Island ($+17.15\,\%$), with the lowest average change at Alice Springs ($+9.80\,\%$). The Priestley–Taylor model identified the highest average change at Hobart ($+17.77\,\%$), with the lowest at Tennant Creek ($+7.09\,\%$).

To further investigate how potential change in PET varies with different climatic conditions, we now focus on the associations between the PET responses and the baseline levels of the four climate variables for each month of the year and across the 30 study sites. Starting with the Penman–Monteith model (Fig. 3), it is clear that the PET response displays a clear association with the baseline levels of climate variables, with higher magnitude of responses for locations that are cooler (low T), more humid (high RH), and receiving less solar radiation (low R_s). The highest associations can be found with T (Fig. 3a), with the monthly changes in PET ranging from -30.2 to $+98.3\,\%$ for the lowest baseline T value of $5.0\,^{\circ}\mathrm{C}$, compared to a range of -13.3 to $+46.6\,\%$ for the highest baseline T of $30.3\,^{\circ}\mathrm{C}$. Similarly, the range of Penman–Monteith PET responses also shows clear decreases with baseline R_s (Fig. 3c), and increases with baseline RH (Fig. 3b). The baseline u_z (Fig. 3d) levels show no obvious impact on the PET responses.

The potential responses in PET obtained from Priestley–Taylor was also investigated (Fig. 4), and results are consistent with the results from the Penman–Monteith model, although the overall ranges of responses were smaller for each variable as anticipated from the results in Table 3. Interestingly, regardless of the choice of PET model, the range of PET responses at the monthly scale is larger than the range for the annual scale suggesting greater uncertainty at higher temporal resolutions.

In addition to assessing the impact of baseline climatic conditions, we are also interested in the role of baseline hydrological conditions (represented by the PET / P ratio at each study site) on the potential responses in PET. Since the hydrological conditions can vary substantially over the course of a year for each study site, for this analysis we

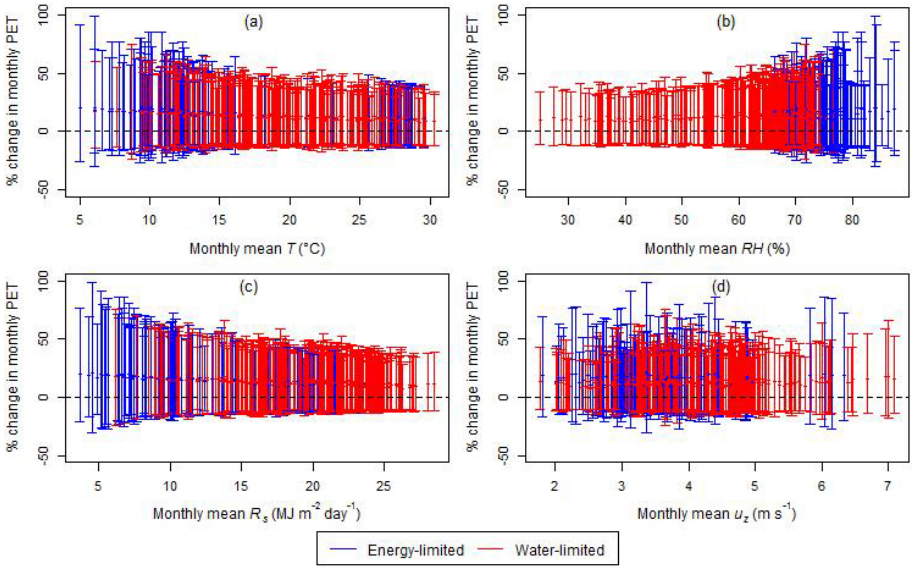


Figure 3. Ranges of monthly PET responses obtained from the Penman–Monteith model, plotted against the monthly baseline levels of **(a)** temperature, **(b)** relative humidity, **(c)** solar radiation and **(d)** wind speed at 30 study sites. Each vertical line represents the range of all potential changes in PET in response to the full set of climate perturbations for a single month at a single location, with the mean represented by the point on the line. The classification of energy- and water-limited months is based on the corresponding monthly PET / P ratios.

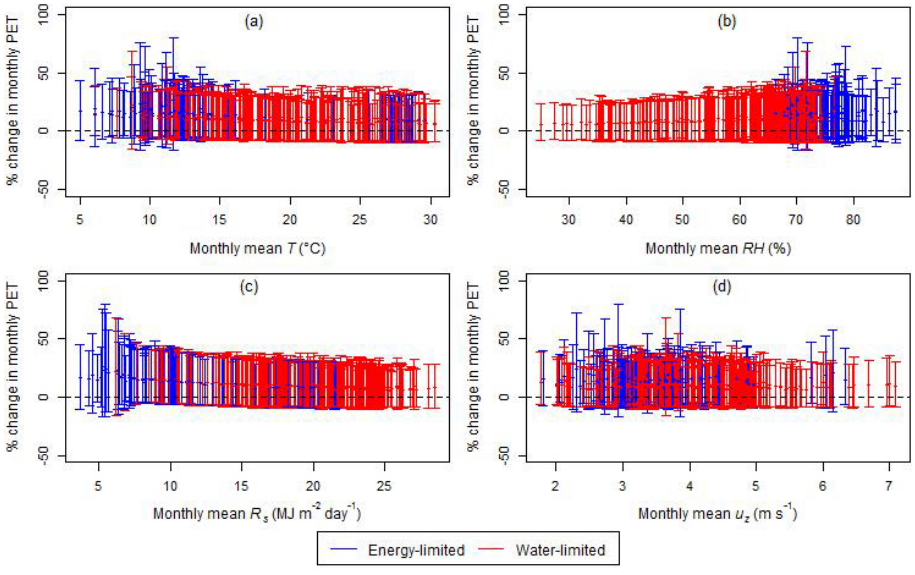


Figure 4. Range of monthly PET responses obtained from the Priestley–Taylor model, plotted against the monthly baseline levels of **(a)** temperature, **(b)** relative humidity, **(c)** solar radiation and **(d)** wind speed at 30 study sites. Each vertical line represents the range of all potential changes in PET in response to the full set of climate perturbations for a single month at a single location, with the mean represented by the point on the line. The classification of energy- and water-limited months is based on the corresponding monthly PET / P ratios.

focused on the PET / P ratios estimated on a monthly basis, and thus differ from the long-term PET / P ratios presented in Table 1. These results are also shown in Figs. 3 and 4, with red-colored bars denoting water-limited conditions, and blue-colored bars denoting energy-limited conditions. These figures show that the magnitude of potential responses in PET is generally larger under energy-limited conditions, regardless of the choice of PET model. In contrast, for water-limited conditions, the potential responses in PET only vary within approximately half of the entire range obtained from each PET model. However, when exploring the association with temperature (Figs. 3a and 4a) in more detail, the magnitude of responses in PET is in fact the lowest for energy-limited conditions during warm months (i.e., when

Table 3. Maximum, minimum and average of all possible changes in annual average PET in response to the full set of climate perturbations from the Penman–Monteith and Priestley–Taylor models at the 30 study sites (as % changes to baseline PET relative to the 1995–2004 baseline). The maximum and minimum changes from each model across all locations are in bold.

No.	Study site name	Penman–Monteith			Priestley–Taylor		
		Min.	Max.	Avg.	Min.	Max.	Avg.
1	Broome airport	−12.33	39.10	11.16	−9.61	33.75	9.59
2	Perth	−13.20	46.67	13.52	−7.98	34.17	10.62
3	Albany	−15.04	54.67	15.21	−7.28	35.49	11.63
4	Giles	−12.30	37.57	10.68	−7.73	25.83	7.27
5	Darwin	−12.73	39.10	10.92	−9.82	33.84	9.50
6	Gove	−13.10	41.34	11.53	−9.74	33.67	9.61
7	Tennant Creek	−12.28	36.45	10.21	−8.35	26.31	**7.09**
8	Alice Springs	−10.88	34.00	**9.80**	−8.00	27.41	7.92
9	Woomera	−12.84	43.48	12.73	−7.48	30.35	9.18
10	Ceduna	−13.97	49.61	14.39	−7.62	33.82	10.67
11	Adelaide airport	−14.47	49.80	14.17	−7.22	34.55	11.09
12	Adelaide (kent town)	−13.10	45.43	13.17	−7.15	33.70	10.78
13	Loxton	−12.55	44.05	12.96	−7.18	33.34	10.67
14	Mount Gambier	−15.33	57.97	16.00	−6.58	35.54	12.02
15	Weipa	−12.42	39.06	10.95	−9.66	32.98	9.36
16	Cairns	−14.80	44.74	12.08	−9.42	33.84	9.73
17	Townsville	−13.77	43.21	12.10	−9.43	34.26	9.90
18	Cobar	−10.62	37.49	11.36	−7.64	31.19	9.49
19	Williamtown	−13.64	47.99	13.68	−7.66	34.11	10.76
20	Sydney	−16.24	53.71	14.46	−7.61	35.24	10.98
21	Canberra	−12.41	46.17	13.85	−6.95	33.24	10.92
22	Wagga Wagga	−13.00	46.34	13.43	−7.09	33.27	10.74
23	Mildura	−12.61	44.50	13.05	−7.24	32.75	10.38
24	East sale	−14.43	53.82	15.34	−6.51	36.32	12.19
25	Scottsdale	−13.64	51.53	15.02	−5.42	40.00	13.47
26	Bicheno	−14.81	52.11	14.87	−4.91	46.38	15.68
27	Lake Leake	−16.06	60.36	16.45	−5.11	36.03	12.84
28	Hobart	−15.97	56.29	15.78	−4.57	50.36	**17.77**
29	Strathgordon village	−13.08	52.11	15.29	−4.66	33.83	12.35
30	Flinders Island	−18.05	64.07	**17.15**	−6.19	38.66	13.02
	Average	−13.66	47.09	13.38	−7.39	34.47	10.91

$T > 25\,^{\circ}\mathrm{C}$, corresponding to the monsoonal summer months in the northern parts of Australia), and the highest for the energy-limited conditions during cool months (i.e., when $T < 15\,^{\circ}\mathrm{C}$, corresponding to the wet winter months in southern Australia). This highlights the fact that it is the atmospheric temperature, rather than the level of aridity, which appears to affect the potential responses in PET. This finding leads to a different interpretation to previous studies, which indicated that the dominant drivers of spatially varying PET include aridity (Tabari and Hosseinzadeh Talaee, 2014) and wind speed (Gong et al., 2006).

The above results also have potential implications on likely AET changes in a future climate. In particular, the above analysis shows that cool and humid regions and seasons appear to show the greatest potential responses in PET, and given that water is not expected to be limited for these cases, the ratio between AET and PET is also likely to be the greatest for these cases. Therefore, one might expect a greater change to AET occurring at the locations and during times of the year where PET is most sensitive to changes in climate.

As a potential limitation to the above analysis, some reliability issues of the Penman–Monteith model have been discussed in a recent study by Milly and Dunne (2016), which suggested that the Penman–Monteith model may overestimate the potential changes in PET in these energy-limited regions relative to a GCM-based AET benchmark. They concluded that the potential changes in ET would be better described by GCMs than "off-line" PET models (such as the two models used in this study), as GCMs can explicitly consider more complex atmospheric processes, such as the interaction between CO_2 and stomatal conductance. Nevertheless, it should be noted that the current reliability of GCMs in simulating ET is also questionable, due to the uncertainty in

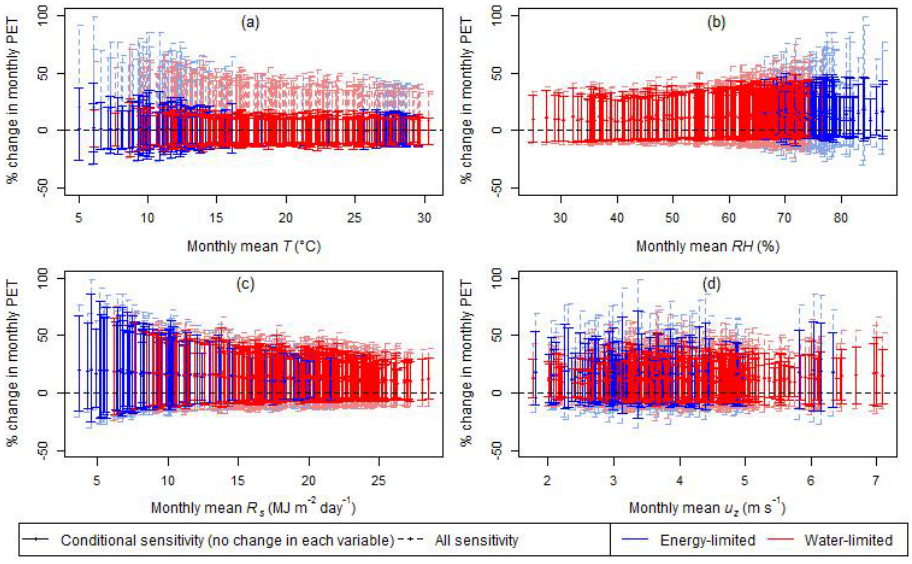


Figure 5. Range of monthly PET responses from the Penman–Monteith model, plotted against the monthly baseline levels of **(a)** temperature, **(b)** relative humidity, **(c)** solar radiation and **(d)** wind speed at 30 study sites. Each dashed (solid) line represents the range of all potential changes in PET in response to the full set of climate perturbations (conditioned on no change in each climate variable) for a single month at a single location. The corresponding means are represented by the points on the lines. The classification of energy- and water-limited months is based on the corresponding monthly PET / P ratios.

representing soil moisture and radiative energy at the evaporative surface (e.g., Seneviratne et al., 2013; Boé and Terray, 2008; Barella-Ortiz et al., 2013). In addition, due to the coarse scale of GCM output, downscaling is generally required to post-process output for use at local and regional scales, which often adds further bias and uncertainties to the GCM simulation and largely limits their applicability (e.g., Chen et al., 2012; Diaz-Nieto and Wilby, 2005). Therefore, although GCM results may be more suitable for large-scale assessments, catchment-scale climate impact assessments are likely to be informed by "off-line" PET models for the foreseeable future. Consequently, the estimated potential changes in PET shown in this study will remain relevant for climate impact assessments conducted using these models.

4.2 Relative importance of climate variables affecting PET for different climate zones

We now explore the relative importance of each climate variable on overall PET sensitivity, by first visualizing the conditional responses of PET when holding each variable constant at its historical level while perturbing the remaining variables, and then comparing this to the unconditional estimates of all potential responses in PET (as shown in Figs. 3 and 4). Figure 5 shows the ranges of the monthly unconditional responses in PET (dashed lines) and the ranges of the monthly responses conditioned on zero change in each of T, RH, R_s and u_z (solid lines) for the Penman–Monteith model, plotted against the monthly baseline levels of the four climate variables at the 30 study sites.

The figure suggests that perturbations in T have the greatest impact on the potential changes in PET compared to other climate variables (Fig. 5a), contributing to at least 45 % of the entire range of PET responses compared to the unconditional results. Humidity also plays a significant role, although only for higher humidity levels (contributing up to 57 % of the entire range of PET responses) with relatively minor influence for the less humid catchments (Fig. 5b). In contrast, the role of solar radiation (Fig. 5c) and wind (Fig. 5d) is generally minor, with the range of unconditional responses being only slightly wider than the range of conditional responses.

A similar analysis was conducted for the Priestley–Taylor model (Fig. 6), and shows somewhat different results compared to those obtained for the Penman–Monteith model. Consistent with Fig. 5a, temperature has the greatest impact, but in this case contributes up to 85 % of the overall variability in PET responses (Fig. 6a). As a result, the range of PET changes contributed by the remaining variables (i.e., conditional responses with no change in temperature) is much smaller. Unlike in Fig. 5b, the role of relative humidity does not appear to increase significantly with increasing baseline humidity (Fig. 6b) and in general contributes less than 33 % of the overall variability. The lower impact of RH on Priestley–Taylor PET compared to the impact on Penman–Monteith PET can be related to the structure of Priestley–Taylor model, which does not consider the aerodynamic processes; therefore, the impact of RH on PET through these processes is not accounted for (see Eqs. A13, A21 and A22 in Appendix A2). The role of solar radiation appears to be somewhat larger for high baseline solar radia-

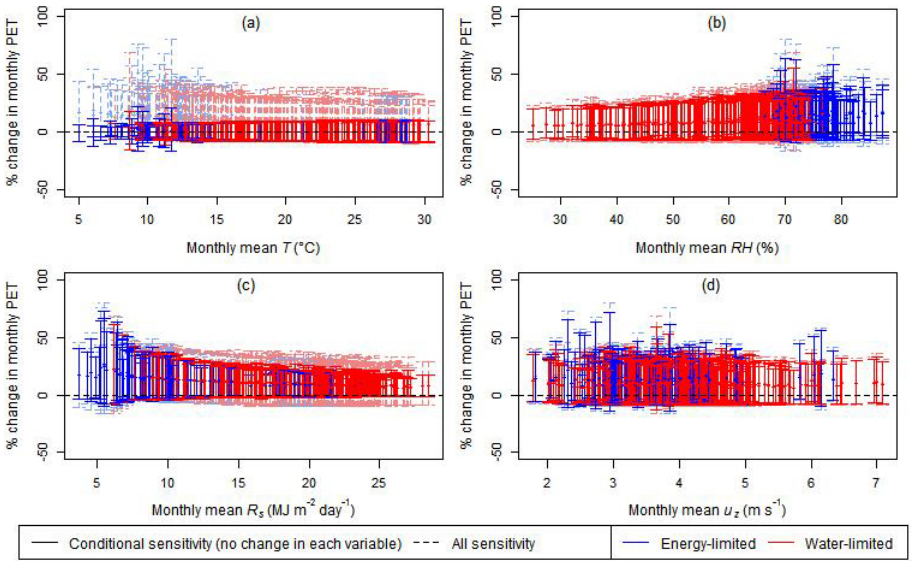


Figure 6. Range of monthly PET responses from the Priestley–Taylor model, plotted against the monthly baseline levels of **(a)** temperature, **(b)** relative humidity, **(c)** solar radiation and **(d)** wind speed at 30 study sites. Each dashed (solid) line represents the range of all potential change in PET in response to the full set of climate perturbations (conditioned on no change in each climate variable) for a single month at a single location. The corresponding means are represented by the points on the lines. The classification of energy- and water-limited months is based on the corresponding monthly PET / P ratios.

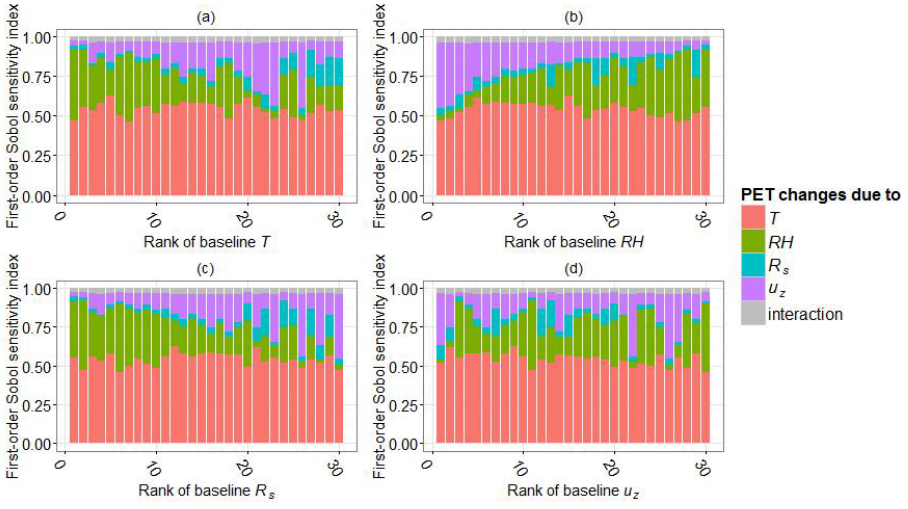


Figure 7. Sobol' first-order sensitivity indices of the Penman–Monteith model for changes in the four climate variables (colored) and their interaction effects (gray), plotted against the ranking of the average level of each climate variable at 30 study sites.

tion values (Fig. 6c) and wind is shown to have no impact as expected, since wind is not an input into the Priestley–Taylor model (Fig. 6d). However, it is worth noting that although the Priestley–Taylor model does not consider wind as an input variable, the range of unconditional responses of PET is slightly wider than the range of responses conditioned on no change in wind. This is because the conditional responses were estimated with only a subset of all climate perturbations (Sect. 3.4), which may not consist of the entire range of perturbation in each of the other three climate variables.

A more formal quantitative measure of the relative importance of each climate variable for PET is provided by the Sobol' indices. Figure 7 shows the Sobol' first-order indices of the Penman–Monteith PET to changes in the four climate variables at the annual scale, as well as their interactions. The first-order indices are plotted against the baseline levels of each climatic variable to observe the impact of baseline climate conditions. For presentation purposes, the baseline levels are represented by the rank of the baseline annual average value of each variable, rather than the absolute level of

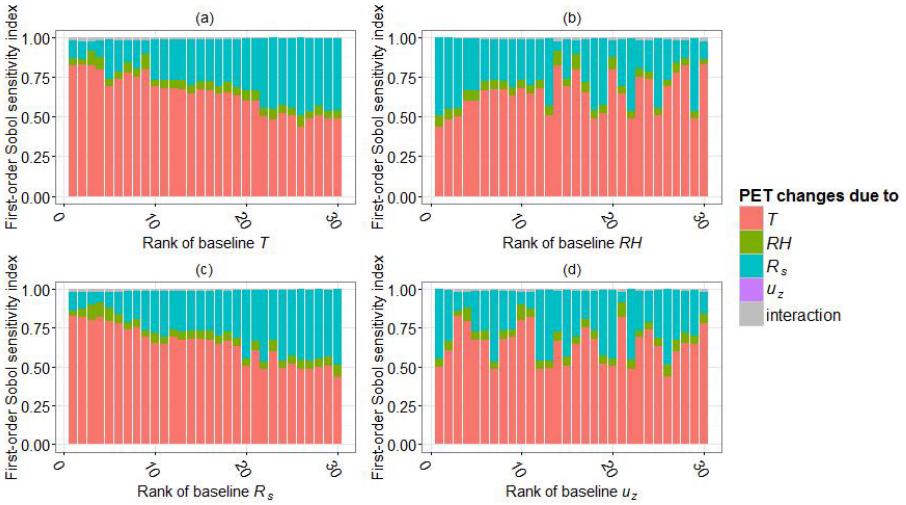


Figure 8. Sobol' first-order sensitivity indices of the Priestley–Taylor model for changes in the four climate variables (colored) and their interaction effects (gray), plotted against the ranking of the average level of each climate variable at 30 study sites.

each climate variable across the 30 study sites. The Sobol' indices in the figure show that T is generally the most important variable for PET, with index values ranging from 0.46 to 0.62. Since the Sobol' indices suggest the partitioning of the total variance of PET, these results are consistent with Fig. 5a, which suggests that perturbations in T contribute to at least 45 % of the variation in the estimated changes in PET. The role of wind and humidity in affecting the sensitivity values is also evident, with wind being the second-most important variable (with Sobol' indices up to 0.42) for sites with low baseline humidity, and humidity being the second-most important variable (with Sobol' indices up to 0.47) for sites that have high humidity (Fig. 7b). Solar radiation is generally the variable with the lowest Sobol' indices, where the largest contributions (up to 18 %) can be observed for warm catchments (Fig. 7a).

The Sobol' sensitivity indices are also presented for the Priestley–Taylor model (Fig. 8), and show substantial differences compared to those for the Penman–Monteith model. Temperature exhibits the largest sensitivity score in most cases, and ranges from 0.44 to 0.83. The relative role of temperature varies most clearly as a function of both the baseline temperature (Fig. 8a) and the baseline solar radiation values (Fig. 8c), with temperature being particularly important for low temperature and low solar radiation sites. As temperature and radiation increase, the relative role of solar radiation becomes more important, reaching Sobol' index values of up to 0.49. In contrast, the role of relative humidity is generally minor (with Sobol' indices within the range 0.03–0.1) and does not appear to vary as a function of baseline conditions. Finally, the role of wind is absent, given that this variable is not included as part of the Priestley–Taylor equation.

The differences between the Penman–Monteith and Priestley–Taylor models highlight the different physical assumptions underpinning the models, with aerodynamic processes being important for the Penman–Monteith model as indicated by the relative importance of RH and u_z for this model, whereas R_s has a critical role in the Priestley–Taylor model, which is closely linked to the emphasis of radiative energy as the energy source for ET in the model.

Finally, comparing Figs. 7 and 8, it is apparent that the interactions among the four climate variables on PET (shown as gray bars) are greater in the Penman–Monteith model compared to the Priestley–Taylor model. Specifically, these interactions contribute fractions of 0.03–0.04, and 0–0.02 of the total variance in PET for the Penman–Monteith and Priestley–Taylor models, respectively. The relative magnitude of the interaction effects in the two models can be again related to their structural differences; the higher interaction effects in Penman–Monteith can be a result of the larger number of variables in this model compared with those in the Priestley–Taylor model.

It is difficult to assess the consistency of these sensitivity results with existing literature, given the different methodologies and data sets used in other studies. Although most PET sensitivity studies used only the Penman–Monteith PET model, there is still substantial discrepancy in results depending on the specific implementations of sensitivity analysis. For example, Gong et al. (2006) perturbed each of temperature, wind speed, relative humidity and solar radiation within ±20 % for the Changjiang basin in China, and observed that relative humidity was generally the most important variable driving PET, followed by solar radiation, temperature and wind speed. This contrasted with our results from the Penman–Monteith model, which showed temperature as the most important variable and solar radiation as the least important variable for almost all the stations analyzed, and may be attributable to the different baseline climates as

well as the perturbation ranges used for the sensitivity analysis between the two studies.

The results of our study were more consistent with Goyal (2004), who concluded that PET is most sensitive to potential changes in temperature for an arid region in India, by applying a ±20 % perturbation on each of temperature, solar radiation, wind speed and vapor pressure. In contrast, Tabari and Hosseinzadeh Talaee (2014) also used a ±20 % perturbation range, but on only three climate variables, namely temperature, wind speed and sunshine hours, for several climate regions in Iran. Their study concluded that the catchment aridity was a major determinant of the sensitivity to temperature, wind speed and humidity, whereas our analysis highlights the importance of baseline temperature and humidity, rather than the aridity (or water- or energy-limited status of the catchment) as a key driver.

PET sensitivity can further diversify by the choice of PET models, as illustrated in McKenney and Rosenberg (1993), in which the percentage changes in PET due to a +6 °C change can differ up to around 40 %, when estimated with eight alternative PET models. This lack of consistency in the relative importance of the climate variables for PET is not surprising given the findings of our study, as the results are strongly dependent on the design of the sensitivity analysis experiment, including the choice of study sites and study periods, the input climate variables considered, and the ways to perturb them (i.e., the choice of global or local perturbation and the ranges of perturbation in different input variables).

Nevertheless, the sensitivity results from this study suggest some distinct spatial patterns of the relative importance of different climate variables in Australia. Since the Penman–Monteith model is the most comprehensive physically based PET model, the above regionalization of the PET sensitivity from this model can be used as a benchmark to identify the key climate variables for estimating PET under potential climate change. This information can be particularly useful to suggest the potential suitability of specific PET models for regional applications. For example, since the Penman–Monteith PET showed higher sensitivity to wind at dry locations (Fig. 7b), it is expected that wind-dependent PET models (such as Penman and Penman–Monteith) would be more appropriate for predicting PET at these locations. In contrast, using simpler models that do not consider wind as an input (such as Priestley–Taylor) can be problematic for these locations. Although this study only examined two PET models, the results suggest that simpler empirical models are likely to ignore some potential dynamics and interactions within the climate variables, which makes them less preferred for PET estimation under changing climates.

Another particular issue in the selection of one or several PET models under a changing climate arises from considering the current reliability of available climate projections, as the models can show high levels of sensitivity to variables for which we currently do not have high-quality climate projections. For example, for a given emissions scenario, there is reasonable confidence in projections of temperature and relative humidity in Australia, but less confidence in projections of solar radiation and wind (Flato et al., 2013; CSIRO and Bureau of Meteorology, 2015). However, the radiation-based Priestley–Taylor model can show high sensitivity to solar radiation, particularly for warm locations with high baseline solar radiation (Fig. 8a and c), due to a particular emphasis on radiative energy and thus the empirical relationships between PET and solar radiation. Similarly, the Penman–Monteith model can exhibit higher sensitivity to wind for locations with low relative humidity (Fig. 7b). Therefore, the use of GCM projections at these locations may lead to significant uncertainty in PET estimates due to the uncertainty in the driving variables.

5 Summary and conclusions

In this study, we used a global sensitivity analysis to investigate the sensitivity of PET and the relative importance four climatic variables which influence PET (T, RH, R_s and u_z) under plausible future changes in these variables. The sensitivity analysis was conducted at 30 Australian case study locations within different climate zones to understand the impact of varying baseline hydro-climatic conditions. For the sensitivity analysis, the historical climate data at each study site were first perturbed to represent a large number of plausible climate change conditions, and then the responses in PET were estimated with both the Penman–Monteith and Priestley–Taylor models, from which the sensitivity of PET was analyzed. The key results are as follows:

- In general PET is most sensitive to potential changes in climate in regions with lower temperature, less solar radiation and greater humidity, where 2-fold greater magnitude of changes in PET are expected compared to other locations in Australia.
- Within the plausible perturbations in T, RH, R_s and u_z, PET is generally most sensitive to T. The relative importance of the other climate variables varies substantially with the PET models. R_s has a dominant role in the radiation-based Priestley–Taylor model, highlighting the importance of radiative energy in the model. In contrast, the importance of RH and u_z are comparable for the Penman–Monteith model, whereas R_s has only little impact, reflecting the contribution of aerodynamic energy.
- The relative importance of climate variables in influencing PET depends very clearly on baseline climatic conditions. From Penman–Monteith, locations that are warmer, drier and receiving more solar radiation generally show greater sensitivity to u_z and lower sensitivity to RH. For Priestley–Taylor, the importance of T increases while that of R_s decreases for cooler locations and locations receiving less solar radiation.

The global sensitivity analysis used in this study is a powerful tool for providing a comprehensive and consistent measure of PET sensitivity to different climatic variables, considering a wide range of possible changes in climate, across different models with different data requirements. However, we have identified space for improvements in further implementations. For example, the bounds of perturbation for each climate variable can have a substantial impact on PET sensitivity, and thus their selection requires careful justification (for example see Whateley et al., 2014; Shin et al., 2013). Therefore, alternative lines of evidence on possible changes in climate should be considered in setting these bounds: for example, the results of ensemble climate models (e.g., Collins et al., 2013), the impact of low-frequency climatic modes (e.g., Chen et al., 2013; Vincent et al., 2015), as well as findings from within paleoclimatology records (e.g., Ault et al., 2014; Ho et al., 2015).

The analysis in this study also lends itself to scenario-neutral analyses (Brown et al., 2012; Prudhomme et al., 2010), although the full implications on specific impacts of hydrological systems (flood risk, water supply, etc.) would require the sensitivity analysis to be propagated to runoff via explicitly modeling the interaction between ET and rainfall–runoff processes (e.g., Garcia and Tague, 2015; Roy et al., 2017). Furthermore, potential changes to precipitation, which were not analyzed here but that can have a significant impact on future runoff, would need to be considered. Within this context, the incorporation of alternative lines of evidence can therefore not only be used to define the bounds of the perturbations, but also can be superimposed onto the exposure space (e.g., as in Prudhomme et al., 2013a; Culley et al., 2016) to provide insight into the likelihood of possible changes. The outcomes of our study can feed into such a scenario-neutral analysis by providing guidance on the variables that are likely to be most important for a particular location, as well as providing insights on the potential implications of using alternative PET models on the overall sensitivity results.

Data availability. The temperature, relative humidity, wind speed, sunshine hours and rainfall data for the five case study locations were obtained from the Climate Data Online project website, http://www.bom.gov.au/climate/data/ (Bureau of Meteorology, 2016).

Appendix A: Sobol' sensitivity analysis and PET models

A1 Sobol' sensitivity analysis (Sobol' et al., 2007)

Sobol' is considered a variance-based method, which requires estimation of the total variance in a model output due to changes in its inputs is estimated with a Monte Carlo approach. To estimate the variances, a large number of samples is firstly drawn by varying all input variables simultaneously, and then a Sobol' sequence is constructed by re-sampling from within these Monte Carlo samples (Saltelli et al., 2010). According to Sobol' et al. (2007), to estimate the Sobol' first-order and total-order indices with a Monte Carlo sample size of n consisting of p input variables, a Sobol' sequence with a total of $n \cdot (p+2)$ samples should be obtained, i.e., requiring $n \cdot (p+2)$ model evaluations.

Sobol' analysis partitions the total variance in model output to the contribution of each individual input variable (i.e., first-order effects), as well as their interactions (i.e., higher-order effects), as follows (equation adapted from Zhang et al., 2015):

$$V_Y = \underbrace{\sum_{i=1}^{n} V_i}_{\text{Individual effects}} + \sum_{i<j} V_{ij} + \underbrace{\sum_{i<j<k} V_{ijk} \ldots}_{\text{interactions}} + V_{1,2\ldots,n}. \quad \text{(A1)}$$

The outputs from Sobol' analysis include (equations adapted from Nossent et al., 2011)

1. first-order sensitivity index, which quantifies the individual contribution of each input variable to the total variance of the model's output:

$$S_i = \frac{V_i}{V_Y}; \quad \text{(A2)}$$

2. second- and higher-order sensitivity index, which quantifies the contribution of interactions among two or more input variables to the total variance of the model's output:

$$\text{For second} - \text{order}: \; S_{ij} = \frac{V_{ij}}{V_Y},$$
$$\text{For higher} - \text{order}: \; S_{ij\ldots n} = \frac{V_{ij\ldots n}}{V_Y};$$

3. total sensitivity index, which quantifies the total contribution of each input variable, including its individual effect as well as all its interactions with other input variables, to the total variance of the model's output:

$$S_{T_i} = S_i + \sum_{j \neq i} S_{ij} = 1 - \frac{V_{\sim i}}{V_Y}. \quad \text{(A3)}$$

From Eqs. (A1) to (A4), the sum of individual effects of all input variables and all their interactions equals one (adapted from Zhang et al., 2015):

$$1 = \underbrace{\sum_{i=1}^{n} S_i}_{\text{Individual} + \sum_{i<j} S_{ij} \text{ effects}} + \underbrace{\sum_{i<j<k} S_{ijk} \ldots}_{\text{interactions}} + S_{1,2\ldots,n}. \quad \text{(A4)}$$

A2 Penman–Monteith PET model (FAO-56) (as in McMahon et al., 2013)

The Penman–Monteith PET model (FAO-56) is given as

$$\text{ET} = \frac{0.408\Delta\,(R_\text{n} - G) + \gamma \frac{900}{T_\text{a}+273} u_2 \left(v_\text{a}^* - v_\text{a}\right)}{\Delta + \gamma\,(1 + 0.34 u_2)}. \quad \text{(A5)}$$

The process for estimating each of the variables in this equation are described in the following sections.

A2.1 Estimating Δ in Eq. (A7)

Δ is the slope of vapor pressure curve in kPa °C^{-1}, which is estimated by

$$\Delta = \frac{4098\left[0.6108 \exp\left(\frac{17.27 \cdot T_\text{a}}{T_\text{a}+237.3}\right)\right]}{(T_\text{a}+237.3)^2}. \quad \text{(A6)}$$

In Eq. (A8), T_a is the average daily temperature in °C, calculated as

$$T_\text{a} = \frac{T_\text{max} + T_\text{min}}{2}. \quad \text{(A7)}$$

A2.2 Estimating R_n in Eq. (A7)

R_n is the net incoming solar radiation at the evaporative surface in MJ m^{-2} day^{-1}, which is estimated by

$$R_\text{n} = R_\text{ns} - R_\text{nl}. \quad \text{(A8)}$$

In Eq. (A10), R_ns is the net shortwave solar radiation, estimated by

$$R_\text{ns} = (1-\alpha) R_\text{s}. \quad \text{(A9)}$$

In Eq. (A11), α is the albedo at evaporative surface which is fixed at 0.23 in this equation, and R_s is the measured or estimated incoming solar radiation in MJ m^{-2} day^{-1}. R_nl is the net outgoing longwave radiation, estimated as

$$R_\text{nl} = \sigma \left(0.34 - 0.14 v_\text{a}^{0.5}\right) \frac{(T_\text{max}+237.2)^4 + (T_\text{min}+237.2)^4}{2} \left(1.35 \frac{R_\text{s}}{R_{\text{s}_0}} - 0.35\right). \quad \text{(A10)}$$

In Eq. (A12): σ is Stefan–Boltzmann constant $= 4.903 \times 10^{-9}$ MJ m^{-2} day^{-1} °K^{-4}, v_a is the mean

daily actual vapor pressure in kilopascals, R_{s_0} is the clear-sky radiation in MJ m^{-2} day^{-1}. v_a and R_{s_0} estimated by Eqs. (A13) and (A14), respectively:

$$v_a = \frac{v_a^*(T_{max})\frac{RH_{max}}{100} + v_a^*(T_{min})\frac{RH_{min}}{100}}{2}, \quad (A11)$$

$$R_{s_0} = \left(0.75 + 2\times10^{-5}\text{Elev}\right) R_a. \quad (A12)$$

In Eq (A14), Elev is the ground elevation above sea level at the measurement location, and R_a is the extraterrestrial solar radiation in MJ m^{-2} day^{-1}, estimated as

$$R_a = \frac{1440}{\pi} G_{sc} d_r^2 (\omega_s \sin(\text{lat})\sin(\delta) + \cos(\text{lat})\sin(\text{lat})\sin(\omega_s)). \quad (A13)$$

In Eq. (A15), G_{sc} is the solar constant = 0.0820 MJ m^{-2} min^{-1}, lat is the latitude in radiance, d_r is the inverse relative distance between Earth and Sun, δ is the solar declination in radians and ω_s is the sunset hour angle in radians. The d_r, δ and ω_s are estimated as follows:

$$d_r^2 = 1 + 0.033\cos\left(\frac{2\pi}{365}\text{DoY}\right) \text{ with DoY as the day of the year,} \quad (A14)$$

$$\delta = 0.409\sin\left(\frac{2\pi}{365}\text{DoY} - 1.39\right), \quad (A15)$$

$$\omega_s = \arccos[-\tan(\text{lat})\tan(\delta)]. \quad (A16)$$

A2.3 Estimating other variables in Eq. (A7)

- G is negligible for daily time step.
- γ is the psychrometric constant in kPa °C^{-1}, estimated as

$$\gamma = 0.00163\frac{P}{\lambda} \text{ where } P \text{ is the pressure at elevation } z \text{ meters.} \quad (A17)$$

  u_2 is the daily average wind speed measured at 2 m in m s^{-1}, which can be estimated from the measured wind speed at z meters as

$$u_2 = u_z\frac{\ln\left(\frac{2}{z_0}\right)}{\ln\left(\frac{z}{z_0}\right)} \text{ where } z_0 \text{ is the roughness height in meters.} \quad (A18)$$

  $(v_a^* - v_a)$ is the vapor pressure deficit in kilopascals, in which v_a is the mean daily actual vapor pressure in kilopascals, estimated as Eq. (A13); v_a^* is the daily saturation vapor pressure in kilopascals, estimated as

$$v_a^* = \frac{v_a^*(T_{max}) + v_a^*(T_{min})}{2}. \quad (A19)$$

In Eq. (A21), $v_a^*(T_{max})$ and $v_a^*(T_{min})$ are the vapor pressures at temperatures T_{max} and T_{min} in °C are estimated with

$$v_T^* = 0.6108\exp\left[\frac{17.27T}{T+237.3}\right]. \quad (A20)$$

A3 Priestley–Taylor PET model (as in McMahon et al., 2013)

The Priestley–Taylor PET model is given as

$$\text{ET} = \alpha_{PT}\cdot\left[\frac{\Delta}{\Delta+\gamma}\frac{R_n}{\lambda} - \frac{G}{\lambda}\right] \quad (A21)$$

where

- α_{PT} is the albedo specifically used for the Priestley–Taylor model, since an evaporative surface of reference crop was assumed, this has a value of 1.12 which was for a similar surface of short grass (see Table S8 of the Supplement of McMahon et al., 2013);
- Δ is the slope of vapor pressure curve in kPa °C^{-1}, estimated as Eq. (A8);
- γ is the psychrometric constant in kPa °C^{-1}, estimated as Eq. (A18);
- λ is the latent heat of vaporization, which is 2.45 MJ kg^{-1} at 20 °C;
- R_n is the net incoming solar radiation at the evaporative surface in MJ m^{-2} day^{-1}, which is estimated in the same way as Eq. (A10);
- G is negligible for daily time step.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors wish to thank Giovanni Ravazzani and four anonymous reviewer for their thoughtful comments on the manuscript.

Edited by: N. Romano
Reviewed by: G. Ravazzani and three anonymous referees

References

Akhtar, M., Ahmad, N., and Booij, M. J.: The impact of climate change on the water resources of Hindukush–Karakorum–Himalaya region under different glacier coverage scenarios, J. Hydrol., 355, 148–163, doi:10.1016/j.jhydrol.2008.03.015, 2008.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration – Guidelines for computing crop water requirements, FAO Irrigation and drainage paper 56, FAO, Rome, 1998.

Arnell, N. W.: The effect of climate change on hydrological regimes in Europe: a continental perspective, Global Environ. Change, 9, 5–23, doi:10.1016/S0959-3780(98)00015-6, 1999.

Ault, T. R., Cole, J. E., Overpeck, J. T., Pederson, G. T., and Meko, D. M.: Assessing the Risk of Persistent Drought Using Climate Model Simulations and Paleoclimate Data, J. Climate, 27, 7529–7549, doi:10.1175/JCLI-D-12-00282.1, 2014.

Barella-Ortiz, A., Polcher, J., Tuzet, A., and Laval, K.: Potential evaporation estimation through an unstressed surface-energy balance and its sensitivity to climate change, Hydrol. Earth Syst. Sci., 17, 4625–4639, doi:10.5194/hess-17-4625-2013, 2013.

Bell, V. A., Kay, A. L., Davies, H. N., and Jones, R. G.: An assessment of the possible impacts of climate change on snow and peak river flows across Britain, Climatic Change, 136, 539–553, doi:10.1007/s10584-016-1637-x, 2016.

Boé, J. and Terray, L.: Uncertainties in summer evapotranspiration changes over Europe and implications for regional climate change, Geophys. Res. Lett., 35, L05702, doi:10.1029/2007GL032417, 2008.

Brown, C., Ghile, Y., Laverty, M., and Li, K.: Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, Water Resour. Res., 48, W09537, doi:10.1029/2011WR011212, 2012.

Bureau of Meteorology: Climate Data Online, available at: http://www.bom.gov.au/climate/data/ (last access: April 2017), 2016.

Chang, S., Graham, W. D., Hwang, S., and Muñoz-Carpena, R.: Sensitivity of future continental United States water deficit projections to general circulation models, the evapotranspiration estimation method, and the greenhouse gas emission scenario, Hydrol. Earth Syst. Sci., 20, 3245–3261, doi:10.5194/hess-20-3245-2016, 2016.

Chen, H., Xu, C.-Y., and Guo, S.: Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff, J. Hydrol., 434–435, 36–45, doi:10.1016/j.jhydrol.2012.02.040, 2012.

Chen, W., Lan, X., Wang, L., and Ma, Y.: The combined effects of the ENSO and the Arctic Oscillation on the winter climate anomalies in East Asia, Chinese Sci. Bull., 58, 1355–1362, doi:10.1007/s11434-012-5654-5, 2013.

Chiew, F. H. S., Teng, J., Vaze, J., Post, D. A., Perraud, J. M., Kirono, D. G. C., and Viney, N. R.: Estimating climate change impact on runoff across southeast Australia: Method, results, and implications of the modeling method, Water Resour. Res., 45, W10414, doi:10.1029/2008WR007338, 2009.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W. J., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A. J., and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 1029–1136, 2013.

CSIRO and Bureau of Meteorology: Climate Change in Australia Information for Australia's Natural Resource Management Regions, Technical Report, CSIRO and Bureau of Meteorology, Australia, 2015.

Culley, S., Noble, S., Yates, A., Timbs, M., Westra, S., Maier, H. R., Giuliani, M., and Castelletti, A.: A bottom-up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate, Water Resour. Res., 52, 6751–6768, doi:10.1002/2015WR018253, 2016.

Diaz-Nieto, J. and Wilby, R. L.: A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom, Climatic Change, 69, 245–268, 2005.

Dingman, S. L.: Physical Hydrology, 3rd Edn., Waveland Press, Long Grove, Illinois, USA, 2015.

Donohue, R. J., McVicar, T. R., and Roderick, M. L.: Generating Australian potential evaporation data suitable for assessing the dynamics in evaporative demand within a changing climate, CSIRO, Water for a Healthy Country National Research Flagship, 50 pp., 2009.

Donohue, R. J., McVicar, T. R., and Roderick, M. L.: Can dynamic vegetation information improve the accuracy of Budyko's hydrological model?, J. Hydrol., 390, 23–34, 2010.

Ekström, M., Jones, P. D., Fowler, H. J., Lenderink, G., Buishand, T. A., and Conway, D.: Regional climate model data used within the SWURVE project – 1: projected changes in seasonal patterns and estimation of PET, Hydrol. Earth Syst. Sci., 11, 1069–1083, doi:10.5194/hess-11-1069-2007, 2007.

Felix, T. P., Petra, D., Stephanie, E., and Martina, F.: Impact of climate change on renewable groundwater resources: assessing the benefits of avoided greenhouse gas emissions using selected CMIP5 climate projections, Environ. Res. Lett., 8, 024023, doi:10.1088/1748-9326/8/2/024023, 2013.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., and Eyring, V.: Evaluation of climate models, in: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, 741–866, 2013.

Garcia, E. S. and Tague, C. L.: Subsurface storage capacity influences climate–evapotranspiration interactions in three western United States catchments, Hydrol. Earth Syst. Sci., 19, 4845-4858, doi:10.5194/hess-19-4845-2015, 2015.

Gerrits, A., Savenije, H., Veling, E., and Pfister, L.: Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model, Water Resour. Res., 45, W04403, doi:10.1029/2008WR007308, 2009.

Gong, L., Xu, C.-Y., Chen, D., Halldin, S., and Chen, Y. D.: Sensitivity of the Penman–Monteith reference evapotranspiration to key climatic variables in the Changjiang (Yangtze River) basin, J. Hydrol., 329, 620–629, doi:10.1016/j.jhydrol.2006.03.027, 2006.

Gosling, S. N., Taylor, R. G., Arnell, N. W., and Todd, M. C.: A comparative analysis of projected impacts of climate change on river runoff from global and catchment-scale hydrological models, Hydrol. Earth Syst. Sci., 15, 279–294, ' doi:10.5194/hess-15-279-2011, 2011.

Goyal, R. K.: Sensitivity of evapotranspiration to global warming: a case study of arid zone of Rajasthan (India), Agr. Water Manage., 69, 1–11, doi:10.1016/j.agwat.2004.03.014, 2004.

Guillevic, P., Koster, R. D., Suarez, M. J., Bounoua, L., Collatz, G. J., Los, S. O., and Mahanama, S. P. P.: Influence of the Interannual Variability of Vegetation on the Surface Energy Balance – A Global Sensitivity Study, J. Hydrometeorol., 3, 617–629, doi:10.1175/1525-7541(2002)003<0617:IOTIVO>2.0.CO;2, 2002.

Guo, D., Westra, S., and Maier, H. R.: An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies, J. Hydrol., doi:10.1016/j.jhydrol.2016.03.025, in press, 2016a.

Guo, D., Westra, S., and Maier, H. R.: An R package for modelling actual, potential and reference evapotranspiration, Environ. Model. Softw., 78, 216–224, doi:10.1016/j.envsoft.2015.12.019, 2016b.

Guo, D., Westra, S., and Maier, H. R.: Impact of evapotranspiration process representation on runoff projections from conceptual rainfall–runoff models, Water Resour. Res., 53, 435–454, doi:10.1002/2016WR019627, 2017.

Harrigan, S. and Berghuijs, W.: The Mystery of Evaporation, Streams of Thought (Young Hydrologic Society), doi:10.5281/zenodo.57847, 2016.

Ho, M., Kiem, A. S., and Verdon-Kidd, D. C.: A paleoclimate rainfall reconstruction in the Murray-Darling Basin (MDB), Australia: 1. Evaluation of different paleoclimate archives, rainfall networks, and reconstruction techniques, Water Resour. Res., 51, 8362–8379, doi:10.1002/2015WR017058, 2015.

Johnson, F. and Sharma, A.: Measurement of GCM skill in predicting variables relevant for hydroclimatological assessments, J. Climate, 22, 4373–4382, 2009.

Kay, A. L. and Davies, H. N.: Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts, J. Hydrol., 358, 221–239, doi:10.1016/j.jhydrol.2008.06.005, 2008.

Kay, A. L., Davies, H. N., Bell, V. A., and Jones, R. G.: Comparison of uncertainty sources for climate change impacts: flood frequency in England, Climatic Change, 92, 41–63, doi:10.1007/s10584-008-9471-4, 2009.

Kay, A. L., Crooks, S. M., and Reynard, N. S.: Using response surfaces to estimate impacts of climate change on flood peaks: assessment of uncertainty, Hydrol. Process., 28, 5273–5287, doi:10.1002/hyp.10000, 2014.

Koedyk, L. P. and Kingston, D. G.: Potential evapotranspiration method influence on climate change impacts on river flow: a mid-latitude case study, Hydrol. Res., 47, 951–963, doi:10.2166/nh.2016.152, 2016.

Köppen, W. P.: Grundriss der Klimakunde, W. de Gruyter, Berlin, Germany, 1931.

Köppen, W. P., Geiger, R., Borchardt, W., Wegener, K., Wagner, A., Knoch, K., Sapper, K., Ward, R. D., Brooks, C. F., and Connor, A.: Handbuch der klimatologie, 1, Gebrüder Borntraeger, Berlin, Germany, 1930.

Li, L., Maier, H. R., Partington, D., Lambert, M. F., and Simmons, C. T.: Performance assessment and improvement of recursive digital baseflow filters for catchments with different physical characteristics and hydrological inputs, Environ. Model. Softw., 54, 39–52, doi:10.1016/j.envsoft.2013.12.011, 2014.

Liu, M., Tian, H., Chen, G., Ren, W., Zhang, C., and Liu, J.: Effects of Land-Use and Land-Cover Change on Evapotranspiration and Water Yield in China During 1900–2001, J. Am. Water Resour. Assoc., 44, 1193–1207, doi:10.1111/j.1752-1688.2008.00243.x, 2008.

Lu, X., Bai, H., and Mu, X.: Explaining the evaporation paradox in Jiangxi Province of China: Spatial distribution and temporal trends in potential evapotranspiration of Jiangxi Province from 1961 to 2013, Int. Soil Water Conserv. Res., 4, 45–51, doi:10.1016/j.iswcr.2016.02.004, 2016.

McKenney, M. S. and Rosenberg, N. J.: Sensitivity of some potential evapotranspiration estimation methods to climate change, Agr. Forest Meteorol., 64, 81–110, doi:10.1016/0168-1923(93)90095-Y, 1993.

McMahon, T. A., Peel, M. C., Lowe, L., Srikanthan, R., and McVicar, T. R.: Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis, Hydrol. Earth Syst. Sci., 17, 1331–1363, doi:10.5194/hess-17-1331-2013, 2013.

McVicar, T. R., Van Niel, T. G., Li, L. T., Roderick, M. L., Rayner, D. P., Ricciardulli, L., and Donohue, R. J.: Wind speed climatology and trends for Australia, 1975–2006: Capturing the stilling phenomenon and comparison with near-surface reanalysis output, Geophys. Res. Lett., 35, L20403, doi:10.1029/2008GL035627, 2008.

McVicar, T. R., Donohue, R. J., O'Grady, A. P., and Li, L.: The effects of climatic changes on plant physiological and catchment ecohydrological processes in the high-rainfall catchments of the Murray-Darling Basin: A scoping study, prepared for the Murray-Darling Basin Authority (MDBA) by the Commonwealth Scientific and Industrial Research Organization (CSIRO) Water for a Healthy Country National Research Flagship, MDBA, Canberra, ACT, Australia, 2010.

McVicar, T. R., Roderick, M. L., Donohue, R. J., Li, L. T., Van Niel, T. G., Thomas, A., Grieser, J., Jhajharia, D., Himri, Y., Mahowald, N. M., Mescherskaya, A. V., Kruger, A. C., Rehman, S., and Dinpashoh, Y.: Global review and synthesis of trends in observed terrestrial near-surface wind speeds: Implications for evaporation, J. Hydrol., 416–417, 182–205, doi:10.1016/j.jhydrol.2011.10.024, 2012.

Milly, P. C. D. and Dunne, K. A.: Potential evapotranspiration and continental drying, Nat.Clim. Change, 6, 946–949, doi:10.1038/nclimate3046, 2016.

Nossent, J., Elsen, P., and Bauwens, W.: Sobol'sensitivity analysis of a complex environmental model, Environ. Model. Softw., 26, 1515–1525, 2011.

Osidele, O. and Beck, M.: Identification of model structure for aquatic ecosystems usingregionalized sensitivity analysis, Water Sci. Technol., 43, 271–278, 2001.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall–runoff model: Part 2 – Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling, J. Hydrol., 303, 290–306, doi:10.1016/j.jhydrol.2004.08.026, 2005.

Paton, F. L., Maier, H. R., and Dandy, G. C.: Relative magnitudes of sources of uncertainty in assessing climate change impacts on water supply security for the southern Adelaide water supply system, Water Resour. Res., 49, 1643–1667, doi:10.1002/wrcr.20153, 2013.

Paton, F. L., Maier, H. R., and Dandy, G. C.: Including adaptation and mitigation responses to climate change in a multiobjective evolutionary algorithm framework for urban water supply systems incorporating GHG emissions, Water Resour. Res., 50, 6285–6304, doi:10.1002/2013WR015195, 2014.

Prudhomme, C. and Williamson, J.: Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: a comparison of methods and associated uncertainty in future projections, Hydrol. Earth Syst. Sci., 17, 1365–1377, doi:10.5194/hess-17-1365-2013, 2013.

Prudhomme, C., Wilby, R. L., Crooks, S., Kay, A. L., and Reynard, N. S.: Scenario-neutral approach to climate change impact studies: Application to flood risk, J. Hydrol., 390, 198–209, doi:10.1016/j.jhydrol.2010.06.043, 2010.

Prudhomme, C., Crooks, S., Kay, A., and Reynard, N.: Climate change and river flooding: part 1 classifying the sensitivity of British catchments, Climatic Change, 119, 933–948, doi:10.1007/s10584-013-0748-x, 2013a.

Prudhomme, C., Kay, A. L., Crooks, S., and Reynard, N.: Climate change and river flooding: Part 2 sensitivity characterisation for british catchments and example vulnerability assessments, Climatic Change, 119, 949–964, doi:10.1007/s10584-013-0726-3, 2013b.

Prudhomme, C., Giuntoli, I., Robinson, E. L., Clark, D. B., Arnell, N. W., Dankers, R., Fekete, B. M., Franssen, W., Gerten, D., Gosling, S. N., Hagemann, S., Hannah, D. M., Kim, H., Masaki, Y., Satoh, Y., Stacke, T., Wada, Y., and Wisser, D.: Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment, P. Natl. Acad. Sci., 111, 3262–3267, doi:10.1073/pnas.1222473110, 2014.

Ravazzani, G., Ghilardi, M., Mendlik, T., Gobiet, A., Corbari, C., and Mancini, M.: Investigation of Climate Change Impact on Water Resources for an Alpine Basin in Northern Italy: Implications for Evapotranspiration Modeling Complexity, PLOS ONE, 9, e109053, doi:10.1371/journal.pone.0109053, 2014.

Roderick, M. L. and Farquhar, G. D.: The Cause of Decreased Pan Evaporation over the Past 50 Years, Science, 298, 1410–1411, doi:10.1126/science.1075390, 2002.

Roderick, M. L., Rotstayn, L. D., Farquhar, G. D., and Hobbins, M. T.: On the attribution of changing pan evaporation, Geophys. Res. Lett., 34, L17403, doi:10.1029/2007GL031166, 2007.

Roy, T., Gupta, H. V., Serrat-Capdevila, A., and Valdes, J. B.: Using satellite-based evapotranspiration estimates to improve the structure of a simple conceptual rainfall–runoff model, Hydrol. Earth Syst. Sci., 21, 879–896, doi:10.5194/hess-21-879-2017, 2017.

Rustomji, P., Bennett, N., and Chiew, F.: Flood variability east of Australia's great dividing range, J. Hydrol., 374, 196–208, 2009.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S.: Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Compu. Phys. Commun., 181, 259–270, doi:10.1016/j.cpc.2009.09.018, 2010.

Seneviratne, S. I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A., Cheruy, F., Higgins, M. E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A., Dufresne, J.-L., Findell, K. L., Ghattas, J., Lawrence, D. M., Malyshev, S., Rummukainen, M., and Smith, B.: Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment, Geophys. Res. Lett., 40, 5212–5217, doi:10.1002/grl.50956, 2013.

Shin, M.-J., Guillaume, J. H. A., Croke, B. F. W., and Jakeman, A. J.: Addressing ten questions about conceptual rainfall–runoff models with global sensitivity analyses in R, J. Hydrol., 503, 135–152, doi:10.1016/j.jhydrol.2013.08.047, 2013.

Sieber, A. and Uhlenbrook, S.: Sensitivity analyses of a distributed catchment model to verify the model structure, J. Hydrol., 310, 216–235, doi:10.1016/j.jhydrol.2005.01.004, 2005.

Sobol', I. M., Tarantola, S., Gatelli, D., Kucherenko, S. S., and Mauntz, W.: Estimating the approximation error when fixing unessential factors in global sensitivity analysis, Reliab. Eng. Syst. Saf., 92, 957–960, doi:10.1016/j.ress.2006.07.001, 2007.

Steinschneider, S. and Brown, C.: A semiparametric multivariate, multi-site weather generator with low-frequency variability for use in climate risk assessments, Water Resour. Res., 49, 7205–7220, doi:10.1002/wrcr.20528, 2013.

Stern, H., De Hoedt, G., and Ernst, J.: Objective classification of Australian climates, Aust. Meteorol. Mag., 49, 87–96, 2000.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M.: Climate change 2013: The physical science basis, Intergovernmental Panel on Climate Change, in: Working Group I Contribution to the IPCC Fifth Assessment Report (AR5), Cambridge Univ Press, New York, 2013.

Tabari, H. and Hosseinzadeh Talaee, P.: Sensitivity of evapotranspiration to climatic change in different climates, Global Planet. Change, 115, 16–23, doi:10.1016/j.gloplacha.2014.01.006, 2014.

Tang, Y., Reed, P., van Werkhoven, K., and Wagener, T.: Advancing the identification and evaluation of distributed rainfall–runoff models using global sensitivity analysis, Water Resour. Res., 43, W06415, doi:10.1029/2006WR005813, 2007a.

Tang, Y., Reed, P., Wagener, T., and van Werkhoven, K.: Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation, Hydrol. Earth Syst. Sci., 11, 793–817, doi:10.5194/hess-11-793-2007, 2007b.

Tang, Y., Reed, P., Wagener, T., and van Werkhoven, K.: Comparing sensitivity analysis methods to advance lumped watershed model

identification and evaluation, Hydrol. Earth Syst. Sci., 11, 793–817, doi:10.5194/hess-11-793-2007, 2007c.

Taylor, I. H., Burke, E., McColl, L., Falloon, P. D., Harris, G. R., and McNeall, D.: The impact of climate mitigation on projections of future drought, Hydrol. Earth Syst. Sci., 17, 2339–2358, doi:10.5194/hess-17-2339-2013, 2013.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M., and Srinivasan, R.: A global sensitivity analysis tool for the parameters of multi-variable catchment models, J. Hydrol., 324, 10–23, doi:10.1016/j.jhydrol.2005.09.008, 2006.

Vincent, L. A., Zhang, X., Brown, R. D., Feng, Y., Mekis, E., Milewska, E. J., Wan, H., and Wang, X. L.: Observed Trends in Canada's Climate and Influence of Low-Frequency Variability Modes, J. Climate, 28, 4545–4560, doi:10.1175/JCLI-D-14-00697.1, 2015.

Whateley, S., Steinschneider, S., and Brown, C.: A climate change range-based method for estimating robustness for water resources supply, Water Resour. Res., 50, 8944–8961, 2014.

Wilby, R. L., Whitehead, P. G., Wade, A. J., Butterfield, D., Davis, R. J., and Watts, G.: Integrated modelling of climate change impacts on water resources and quality in a lowland catchment: River Kennet, UK, J. Hydrol., 330, 204–220, doi:10.1016/j.jhydrol.2006.04.033, 2006.

Zhang, X. Y., Trame, M. N., Lesko, L. J., and Schmidt, S.: Sobol Sensitivity Analysis: A Tool to Guide the Development and Evaluation of Systems Pharmacology Models, Pharmacometr. Syst. Pharmacol., 4, 69–79, doi:10.1002/psp4.6, 2015.

CHAPTER 4 An Inverse Approach to Perturb Historical Rainfall Data for Scenario-Neutral Climate Impact Studies (Paper 3)

Statement of Authorship

Title of Paper	An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Guo, D., Westra, S. & Maier, H. R. (2016). An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. Journal of Hydrology.

Principal Author

Name of Principal Author (Candidate)	Danlu Guo		
Contribution to the Paper	Designed experiments and conducted analysis, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/12/2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Seth Westra		
Contribution to the Paper	Suggested research gap and scope of research, helped to interpret data and results, provided feedbacks on manuscript and responses to reviewers' comments.		
Signature		Date	19/12/16

Name of Co-Author	Holger Maier		
Contribution to the Paper	Suggested scope of research and potential methodology, helped to interpret data and results, helped to evaluate manuscript and to develop responses to reviewers' comments.		
Signature		Date	20/12/16

Please cut and paste additional co-author panels here as required.

Abstract

Scenario-neutral approaches are being used increasingly for climate impact assessments, as they allow water resource system performance to be evaluated independently of climate change projections. An important element of these approaches is the generation of perturbed series of hydrometeorological variables that form the inputs to hydrologic and water resource assessment models, with most scenario-neutral studies to-date considering only shifts in the average and a limited number of other statistics of each climate variable. In this study, a stochastic generation approach is used to perturb not only the average of the relevant hydrometeorological variables, but also attributes such as the intermittency and extremes. An optimization-based inverse approach is developed to obtain hydrometeorological time series with uniform coverage across the possible ranges of rainfall attributes (referred to as the 'exposure space'). The approach is demonstrated on a widely used rainfall generator, WGEN, for a case study at Adelaide, Australia, and is shown to be capable of producing evenly-distributed samples over the exposure space. The inverse approach expands the applicability of the scenario-neutral approach in evaluating a water resource system's sensitivity to a wider range of plausible climate change scenarios.

4.1 Introduction

Scenario-neutral approaches are being used increasingly to assess the possible impact of climate change on the performance of water resources systems (Dessai and Hulme, 2004, Brown et al., 2012, Brown and Wilby, 2012, Nazemi and Wheater, 2014), as well as social and ecological systems (Gao et al., 2016, Poff et al., 2015). The information generated from these approaches can be used to assess system vulnerability under alternative climate change scenarios, and to calculate climatic thresholds at which system performance begins to change abruptly (Brown et al., 2011, Poff et al., 2015). Scenario-neutral approaches can also accommodate changes in climate projections without the need for additional analysis (Prudhomme et al., 2010), and can help to identify the important hydrometeorological variables, or particularly critical states of these variables that affect the system under consideration. The latter feature is particularly useful for selecting: (1) climate models; (2) strategies to generate future rainfall conditions from GCM-based projections (known as statistical downscaling); and (3) alternative 'lines of evidence' (e.g. expert opinion and data from the paleo-climatic record) that can provide useful information about these variables. Ultimately, this allows for the development of a more complete set of projections that describe how these variables might change in a greenhouse gas-enhanced climate (Vano et al., 2015, Steinschneider and Brown, 2013, Singh et al., 2014, Nazemi et al., 2013).

Central to the scenario-neutral approach is the analysis of system sensitivity to a range of hydrometeorological conditions. Such analyses involve exposing the system to perturbed hydrometeorological forcing data that reflect various hydrometeorological conditions that the system may confront in the future (referred to as the 'exposure space'). To this end, it is important to consider the possible variations not only in the average states of the relevant hydrometeorological variables, such as annual average rainfall and potential evapotranspiration (see Kay et al. 2014, Prudhomme et al. 2013a), but also their other attributes, including extremes, seasonality and interannual variability (Meselhe et al., 2009, Moody and Brown, 2013,

Prudhomme et al., 2010, Steinschneider and Brown, 2013). Indeed, assessments of historical and/or future changes to rainfall as a result of climate change have already indicated different changes to the averages (Collins et al., 2013), extremes (Westra et al., 2013, Alexander et al., 2006, Ajami et al., 2007, Westra et al., 2014a), temporal distribution (Rajah et al., 2014) and low-frequency variability (e.g. Johnson et al., 2011) of rainfall throughout the world. Similarly complex changes to other relevant hydrometeorological variables might also be expected, including potential evapotranspiration, and snowfall and melt.

One approach to generating perturbed hydrometeorological forcing data is by applying scaling factors to historical records of each of the relevant hydrometeorological variables. These factors can be applied at annual or monthly scales (Kay et al. 2014, Prudhomme et al. 2013a, Prudhomme et al. 2010, Paton et al., 2013, Singh et al. 2014), or different factors that can be applied across different quantiles in the entire distribution (Nazemi et al., 2013). Although such approaches might be viable for perturbing a small number of hydrometeorological variables and their attributes (i.e. low-dimensional exposure spaces), the capacity of these to represent the potentially complex variations in a wider range of variables and attributes (i.e. high-dimensional exposure spaces) is likely to be limited. Consequently, when using scaling factors to perturb historical data for climate impact assessments, the resultant projections may not show the full range of variability that can be expected in a greenhouse gas-enhanced climate (Prudhomme et al. 2013a, Prudhomme et al. 2010, Steinschneider and Brown 2013).

The use of stochastic generators has been proposed as an alternative to scaling factors to generate hydrometeorological data in a way that can account for a wider range of possible changes (Whateley et al., 2014). Some recent advances include the use of a multi-site weather generator that is capable of producing realistic time series of meteorological variables with shifts to the mean, standard deviation, extremes, daily-scale Markov transition probabilities and low-frequency (interannual) variability (for examples see Steinschneider and Brown, 2013, Wilby et al., 2014, Yates et al., 2015). This

is achieved through the perturbation of stochastic model parameters (including the transition probabilities and the autocorrelation coefficient) and the subsequent application of quantile correction, which, in combination, can be used to generate the high-dimensional exposure space. A challenge with this approach, however, is that it is difficult to assess *a priori* which parameters of the stochastic generator should be modified to produce time series at pre-specified points in the exposure space, potentially leading to insufficient exploration of the exposure space. This challenge arises both as a result of the non-linear mapping between the parameters of a stochastic generator and the statistics of the hydrometeorological variables, as well as due to the stochastic nature of the model, which means that a single parameter set will produce hydrometeorological data that span multiple points on the exposure space (Steinschneider and Brown, 2013).

In order to address the shortcomings of existing approaches in generating hydrometeorological data to form the exposure space, we introduce the concept and framework for an inverse approach with demonstration on a case study. The proposed inverse approach enables stochastic generators to be used to generate time series that uniformly span the desired range of the hydrometeorological variables and attributes of interest, and thus provides uniform coverage of the exposure space to serve the needs of scenario-neutral climate impact assessments. Although generally applicable to any parametric weather generator, this paper focuses on applying the method to rainfall time series for the following reasons:

1. Although stochastic generators have been used to generate a range of weather variables, including temperature, humidity, and wind (e.g. Semenov and Brooks, 1999, Racsko et al., 1991), the majority of applications have focused on the generation of rainfall data, due to their importance as an input to many water resource assessments (e.g. Jones and Thornton 1993, Chiew and McMahon 2002).

2. At daily or shorter timescales, rainfall is intermittent, highly skewed (with rainfall series typically exhibiting a large number of moderate rainfall days and a small number of very heavy rainfall days), and exhibits variability at seasonal, interannual and longer time scales (Dubrovský et al., 2000, Bastola et al., 2011b). As a result, rainfall is often regarded as a particularly challenging variable to simulate stochastically.

3. There has been a substantial amount of work on developing stochastic generation models to both generate replicates of historical rainfall data (Chen and Brissette, 2014, Furrer and Katz, 2008, Clark and Slater, 2006, Frost, 2004, Boughton and Droop, 2003, Beven, 1987, Langousis and Kaleris, 2014, Langousis et al., 2015), as well as downscaling GCM-based climate projections (Wilby et al., 2014, Yates et al., 2015, Allen and Pruitt, 1986, Bastola et al., 2011b, Fowler et al., 2007, Jones et al., 2011, Kay and Jones, 2012).

The remainder of this paper is structured as follows. In Section 4.2, we illustrate the alternative approaches that are currently available for generating an exposure space, including the historical scaling, forward and inverse approaches. This section also provides details of the proposed inverse approach. Section 4.3 introduces a case study and two stochastic generators that are used to illustrate both the proposed approach, as well as a simple forward approach that is used as a basis of comparison. The results are given in Section 4.4, followed by discussions in Section 4.5 and conclusions in Section 4.6.

4.2 Proposed inverse approach to exposure space generation

4.2.1 Rationale for an inverse approach to perturbing stochastic model parameters

As described in the introduction, a central feature of scenario-neutral approaches is the exploration of a water resource system's response to a range of different hydrometeorological conditions. This range of hydrological variables (e.g. rainfall, temperature, evapotranspiration) and the set of attributes of these variables (e.g. annual average, variance, seasonal differences, extremes) are collectively referred to as an 'exposure space', and represent the range of conditions of interest that a system may be exposed to under a future climate. For example, if a scenario-neutral approach was to be used to evaluate system sensitivity to changes in the average, variability and extremes of rainfall, then this would require generating a three-dimensional exposure space with each dimension representing one of the rainfall attributes.

Figure 4-1 illustrates the conceptual approaches that could be used to generate an exposure space **E**, which consists of the plausible future changes (represented as the gray shaded region, with the origin corresponding to no change) in various rainfall attributes of interest (represented by two axes A_1 and A_2, which refer to two generic rainfall attributes or groups of attributes). Two techniques are involved in the perturbation approaches – namely scaling of rainfall time series and stochastic rainfall generation (as shown in the two green squares). We use the term 'scaling' in the figure to collectively refer to perturbations that are directly applied to rainfall time series, through the use of change factors at annual, monthly or other time scales (Prudhomme et al., 2010, Prudhomme and Williamson, 2013, Kay et al., 2014), or more complex methods, such as quantile mapping (as used in Steinschneider and Brown, 2013). Consequently, the scaling technique can only modify rainfall intensity on wet days. The term 'stochastic generation' in the figure refers to indirect modification of the rainfall time series through changing the parameters of stochastic generators (as used in Steinschneider and Brown, 2013, Jones and

Page, 2001, Dubrovský et al., 2000). The parameter space $\boldsymbol{\Theta}$ consists of two axes of θ_1 and θ_2, which refer to two generic parameters or groups of parameters. The plausible ranges for all parameters are represented by the gray shaded region, while the origin represents the set of parameters corresponding to the historical rainfall condition.

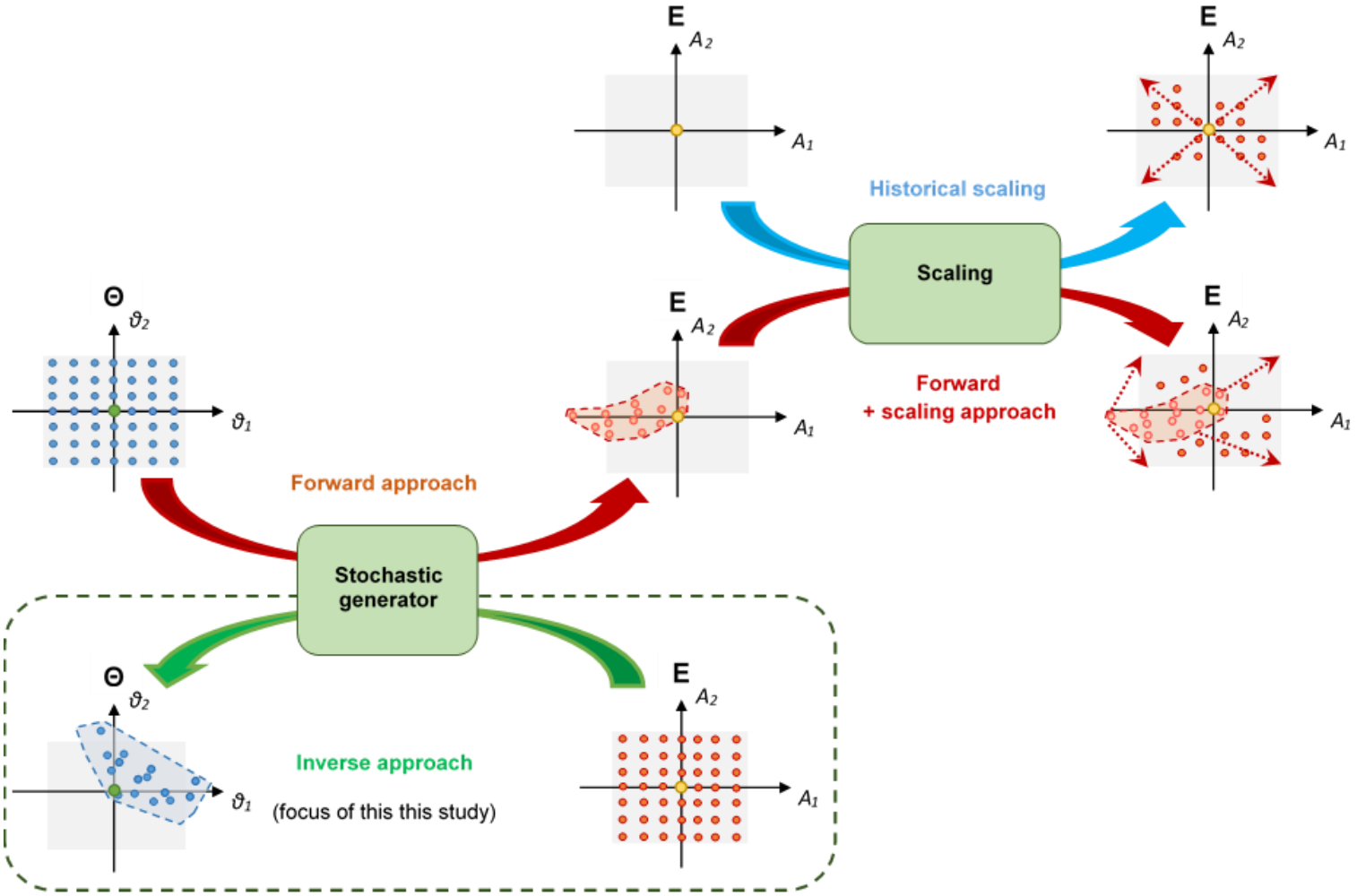


Figure 4-1: Outline of alternative approaches to developing an exposure space E for scenario-neutral climate impact assessments. Refer to main text for an explanation of each figure element. The focus of this study is on the approaches enclosed in the black dashed box.

The first approach, 'historical scaling' (as shown in the top-right corner of Figure 4-1) is analogous to the approach used by Prudhomme et al. (2010), Prudhomme et al. (2013), and Kay et al. (2014), in which additive and/or multiplicative scaling factors are applied directly to historical hydrometeorological time series to obtain the desired changes in the relevant variables (usually rainfall and potential evapotranspiration). Although conceptually simple, this approach is not capable of representing variations in the rainfall intermittency, such as the frequency and persistence of dry-/wet-day occurrence. Furthermore, it is difficult to apply this approach to higher-dimensional exposure spaces, since it becomes difficult to develop an approach to scale each attribute independently of the other attributes. Consequently, it can be difficult to sample some regions of the exposure space.

The remaining approaches use stochastic weather generators to obtain perturbed rainfall time series. The 'forward' approach (as illustrated in the middle of Figure 4-1) involves perturbing the parameters of stochastic generators over some pre-defined 'parameter space' to yield an exposure space (Dubrovský et al., 2000, Jones and Page, 2001). However, the non-linear mapping between the parameters of a stochastic generator and the attributes of the hydrometeorological variables means that it is unlikely that the full range of the desired exposure space will be covered. Conversely, some perturbations may lead to rainfall attributes with levels out of the defined plausible ranges of the exposure space. Consequently, further scaling may still be necessary after application of the forward approach. Steinschneider and Brown (2013) used this combined 'forward-plus-scaling' approach by firstly perturbing the parameters of a stochastic generator (including Markov chain transition probabilities and the autoregressive model for low-frequency variability) to obtain stochastic sequences without changing the historical rainfall intensity; the wet-day rainfall intensity in the stochastic sequences was subsequently quantile-mapped to yield a set of target daily rainfall series with desired levels of rainfall attributes. Although this approach is likely to provide a much better coverage of the exposure space, some portions of the exposure space may still remain poorly represented because of the difficulty in finding

parameters that will result in all combinations of the hydrometeorological attributes of interest.

The limitations of both the historical scaling and forward approaches motivate the 'inverse' approach proposed in this paper (bottom of Figure 4-1). Here, the desired values of the attributes of interest in the exposure space are the starting point for the analysis, followed by an optimization step to identify the stochastic generator parameters that produce stochastic sequences with these attributes. This approach provides control over the level of coverage of the exposure space, as required for the implementation of scenario-neutral approaches to climate impact assessments.

4.2.2 Overview of the inverse approach

To generate hydrometeorological time series with a range plausible attribute levels, the inverse approach is proposed as follows, which involves two primary steps:

1) **Identify a set of 'target' levels for each attribute included in the exposure space.** In order to achieve an even coverage of the exposure space, we first select the desired levels we would like to sample for each attribute included in the exposure space (referred to as 'target levels'). A number of different approaches can be used to select and combine the target levels (which produces individual 'target locations' in the exposure space), including sampling on a regular grid, or using more computationally efficient sampling methods, such as Latin hypercube sampling (Stein, 1987) or Hammersley sampling (Halton, 1960, Hammersley, 1960).

2) **Generate hydrometeorological time series that satisfy each target set of attributes.** For each target location of the exposure space, we combine

stochastic weather generation with a formal optimization approach to identify the best-fit parameter set for the stochastic generator. This parameter set should be capable of producing hydrometeorological time series with the levels of attributes corresponding to that particular target location, as detailed below.

During the optimization process, the decision variables are the parameters of the stochastic generator. The objective is to identify the parameters of the stochastic rainfall generator that minimize the difference between the values of the hydrometeorological attributes that correspond to the target location and those of the corresponding simulated values. The following objective function is proposed for minimization:

$$F_{obj,i} = \sqrt{\sum_{k=1}^{K}\{[\frac{(P_{s,i}^{k} - P_{his}^{k})}{P_{his}^{k}} - \frac{(\widehat{P_{i}^{k}} - P_{his}^{k})}{P_{his}^{k}}] * 100\%\}^2} \quad \textbf{(4.1)}$$

where *i = 1, 2, … n* for *n* target locations in the exposure space. For the k^{th} attribute of the hydrometeorological variable of interest (P^k), P_s^k represents the target level and $\hat{P}^k$ represents the simulated level from the stochastic generator. Since different attributes are likely to consist of different magnitudes, the difference between a target level and the simulated level is represented as a percentage change relative to its long-term averaged historical value (P_{his}^k) to ensure consistent scales across different attributes. The optimization problem can be solved using a variety of optimization algorithms, such as genetic algorithms (Holland, 1975) or shuffled complex evolution (Duan et al., 1993).

The optimization procedure proceeds as follows: (1) values of the parameters of the stochastic generator are perturbed based on the searching strategy of the selected optimization algorithm; (2) the corresponding time series of the desired hydrometeorological variables are generated; (3) the values of the attributes of interest are calculated; and (4) the objective function

is calculated in accordance with Eqn. 4.1, which drives the algorithm's searching behavior. Steps (1)–(4) are then repeated until the specified stopping criterion has been met, such as the completion of a pre-specified number of iterations or until the objective function value is sufficiently small (Maier et al., 2014). It is important to note that as part of the inverse approach, the random seed of the stochastic generator should be held constant to ensure that the optimization proceeds as efficiently as possible, as discussed further in the following section.

4.2.3 Random sampling issues of stochastic generators and the implications on the inverse approach

The stochastic component of the rainfall generator can produce substantial variations in the simulation of rainfall attributes, even with a single parameter set. This randomness can affect the efficiency of the optimization process used in the inverse approach. Essentially, every iteration of the optimization involves a comparison among multiple parameter sets in terms of their ability to generate the target locations in the exposure space. However, as a result of stochastic generation, a single parameter set can lead to multiple potential locations on the exposure space (Figure 4-2). This can then mislead the comparison and affect optimization efficiency, as changes made to parameters by the optimization algorithm in order to lead the search in one particular direction might actually have the opposite effect.

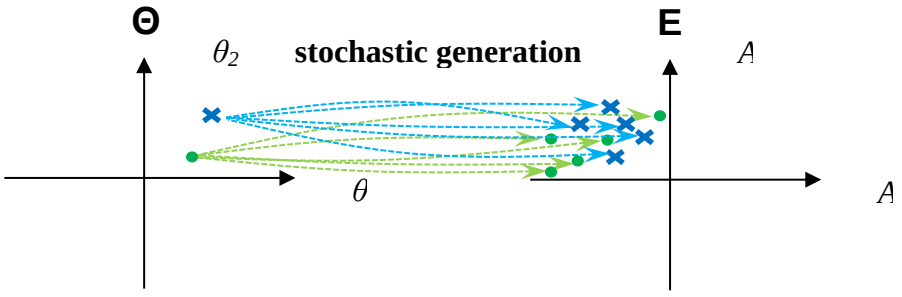


Figure 4-2: Using a single parameter set in the parameter space Θ can yield multiple points (shown by the dashed arrows) in the exposure space E as result of the stochastic nature of the weather generator. Different colors represent simulations from two different parameter sets.

To illustrate this issue, consider a simple optimization problem to find the best-fit parameters of a Gaussian distribution with the objective of getting a

'target' sample mean of $\bar{x} = 3$. Suppose that for one iteration the optimizer attempts to compare samples drawn from a simple Gaussian random generator ($X \sim N(\mu, \sigma)$) where the parameter μ is changed from 4.0 to 4.5, while holding σ at a constant value of 1. In the upper panel of Figure 4-3, we show 50 random values drawn with each parameter set. For this set of random values, the sample mean from $X \sim N(4.0, 1)$ is 4.2 compared with the sample mean from $X \sim N(4.5, 1)$, which is 4.0. Therefore, the resulted sample mean from $N(4.0, 1)$ is actually further away from the target sample mean of $\bar{x} = 3$ compared with $N(4.5, 1)$, so that the search direction of the optimizer may be misled. Although this variance can be reduced with a larger sample size or a longer simulation period, it can never be completely eliminated.

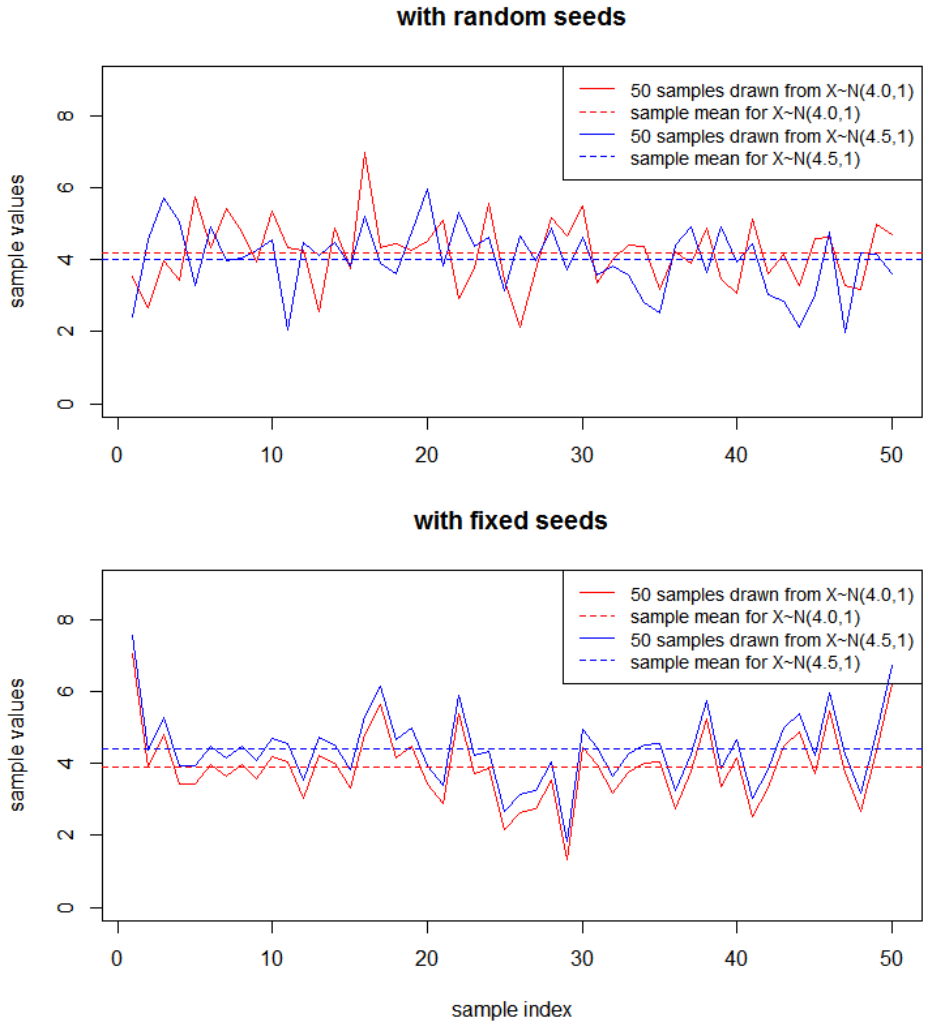


Figure 4-3: Implications of using random seed (upper panel) and fixed seed (lower panel) for comparing two parameter sets: *X*~*N*(4.0, 1) and *X*~*N*(4.5, 1), for simulating a sample mean of 3 from a Gaussian distribution.

To overcome this problem during optimization, the random number seed is held constant when producing the stochastic replicates. This ensures that any changes made to the parameters during the optimization process will lead the search in the desired direction. Using the same example, in the lower panel of Figure 4-3 we show 50 samples drawn from both X ~ *N*(*4.0, 1*) and X ~ *N*(*4.5, 1*) with the same random seed used for each pair of samples, resulting in samples means of 3.9 and 4.4 respectively, thereby indicating that *N*(*4.0, 1*) is better at producing a target sample mean of $\bar{x} = 3$. In this way, the stochastic generator is able to search through the correct directions on the parameter space to find parameters that converge toward the target rainfall attributes.

As discussed in Section 4.2.2, it is important to emphasize that the objective of the approach is to generate samples of hydrometeorological time series with specific levels for each attribute, rather than to identify the parameter sets that will produce those parameters in a population sense. Returning to the above example, the objective is to find a stochastic replicate with sample mean $\bar{x} = 3$, regardless of the values of the parameters (μ, σ) used to achieve this value. Consequently, once this goal has been met, the search can stop and the parameter values that were used to produce the stochastic time series corresponding to each target location can be discarded.

4.3 Case study

The proposed inverse approach is illustrated on rainfall data from a catchment in South Australia, using two stochastic rainfall generators: the Richardson model and the WGEN model. To provide a benchmark for the proposed inverse approach, its performance is also compared with that of a forward approach. The rainfall data, stochastic rainfall generators and the specific implementation of the forward and inverse approaches are described in this section.

4.3.1 Data

We used a rainfall time series from a gauge in the southern Mount Lofty Ranges close to Adelaide, South Australia, as a case study. The climate in this region is temperate, with most rainfall occurring in winter and spring (May to October). The mean annual rainfall was 913mm for the study period from 1989 to 2004. The daily rainfall data over this period have been used to represent the baseline (historical) rainfall conditions.

We used four rainfall attributes as the dimensions of the exposure space, with definitions and baseline values provided in Table 4-1. These attributes represent key features of rainfall patterns; namely, the average daily rainfall (*PD*), the wet day frequency (*WD*), a measure of the rainfall intermittency (*CDD*) and a measure of extreme rainfall (*Pex99*). These attributes have been commonly used to assess the performance of rainfall generators (Semenov, 2007, Kilsby et al., 2007, Hashmi et al., 2011, Fowler et al., 2007, Chen and Brissette, 2014), and are also closely related to several of the indices used for the detection and attribution of climate change, as described by the Expert Team on Climate Change Detection and Indices (ETCCDI; Klein Tank. et al., 2009).

Table 4-1: Definitions and annual average values of the four rainfall attributes that form the exposure space.

Rainfall attribute	Definition	Historical value
PD	Daily rainfall intensity averaged over all days	2.49 mm
WD	Average number of wet days	161 days
CDD	Average length of consecutive dry days	3.75 days
Pex99	The 99th percentile rainfall over wet days	37.6 mm

For each rainfall attribute we defined a plausible range for sampling (which defined the range of each dimension within the exposure space) of between 50% and 150% of the historical values. These bounds are wider than would be expected from most climate change projections (e.g. CSIRO and Bureau of Meteorology, 2015, Stocker et al., 2013), to encompass a large range of climate projections (for example from climate models) in the exposure space.

4.3.2 Stochastic rainfall generators

Two versions of the Richardson-type stochastic rainfall generator with different levels of complexity were used to generate the exposure space. We started with a simplified four-parameter model, which assumes uniform rainfall characteristics over the year. The advantage of this model is that it is possible to analytically derive the parameters that correspond to each target location in the exposure space. However, this simplified model uses a single value of each parameter throughout the year, and thus is unable to capture seasonal-scale variability. To highlight some practical issues with rainfall sampling, we then considered a more complex and widely used model—namely the WGEN (Richardson and Wright, 1984).

4.3.2.1 The four-parameter Richardson model

The simplified Richardson-type rainfall generator uses the following four parameters:

- The two parameters of the 1st order two-state Markov chain used for representing the transition probabilities of rainfall occurrence: p_{dd} (dry–dry probability) and p_{wd} (wet–dry probability), and
- The two parameters of a gamma distribution for representing the rainfall intensity on wet days: α (scale) and β (shape).

An approximate analytical expression relating two of the four output rainfall attributes (*PD* and *WD*) to the model parameters is given in Dubrovský et al. (2000) as:

$$PD = \frac{\alpha}{\beta} * \frac{WD}{365.25} \quad \textbf{(4.2)}$$

$$WD = 365.25 \times \frac{(1-p_{dd})}{(1-p_{dd}+p_{wd})} \quad \textbf{(4.3)}$$

These analytical expressions have been used when exploring the implications of random sampling issues on the inverse generation approach (Section 4.4.1.3).

4.3.2.2 The WGEN model

The WGEN model (Richardson and Wright, 1984) has the same structure as the simplified Richardson model, except that it uses a unique set of the four parameters for each month of the year, leading to a total of 48 parameters. This model has been used widely for climate impact studies, and is generally shown to capture most of the key features of daily rainfall series (Katz, 2002, Kim et al., 2007, Bastola et al., 2011b).

Since the proposed inverse approach involves optimization of the parameter values, a search space with low dimension (i.e. consisting of a small number of parameters as decision variables) is desired. To reduce the size of the parameter space in the inverse approach, the number of decision variables to be considered was reduced from 48 to eight by fitting harmonic functions to describe the seasonal variations of each parameter (Prudhomme et al. 2013a). The harmonic function takes the form of:

$$\beta(t) = \beta_0 + A(cos\frac{2\pi}{T}(t - \Phi)) \quad \textbf{(4.4)}$$

where $\beta(t)$ represents one of the four parameters during month (t = 1,…, T) with T = 12, β_0 represents the arithmetic mean of the parameter, A represents the amplitude and Φ corresponds to the month where the maximum occurs. It is worth mentioning that although parameter Φ can be varied as part of the optimization, the four-attribute exposure space in this case study was not designed to focus on shifts in rainfall seasonality (Section 4.3.1), so that Φ was held constant at its historically optimal value. To determine the value of Φ, we obtained the monthly estimates of p_{dd}, p_{wd}, α and β (based on the method in Richardson (1981)) using the historical rainfall data, and fitted a harmonic function to each parameter (Figure 4-4). The corresponding values of Φ were thus identified to be 2, 1, 8 and 1 (i.e., February, January, August and January) for the four parameters, respectively.

As a result, the optimization was performed on the mean (β_0) and amplitude (A) of each of the four model parameters, leading to an eight dimensional search space.

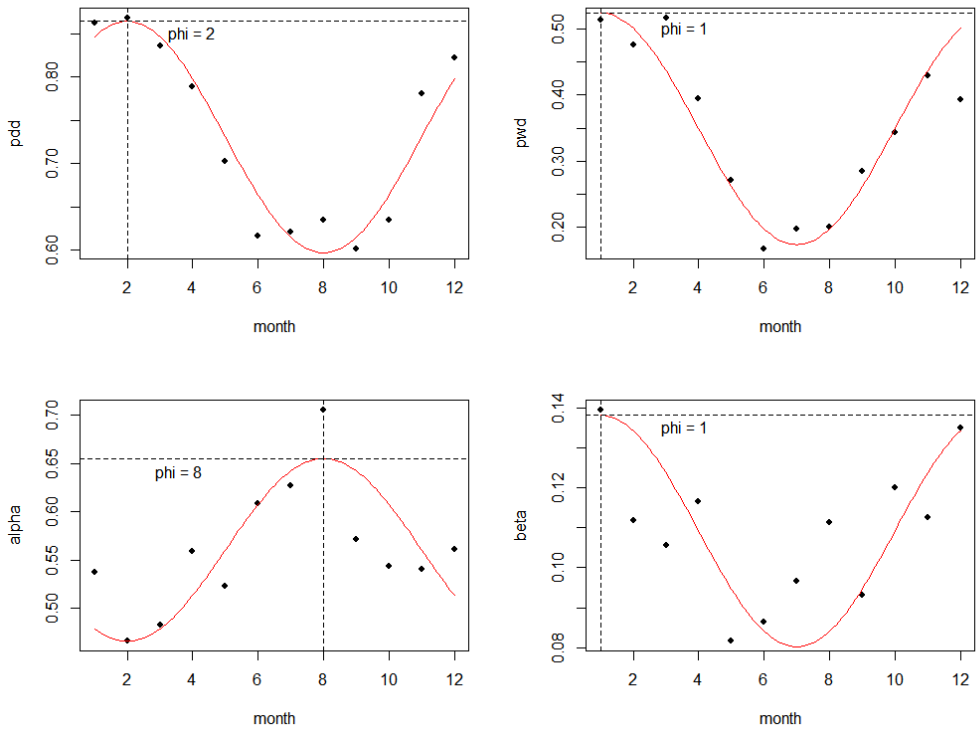


Figure 4-4: The monthly variations of p_{dd}, p_{wd}, α and β for WGEN to determine the values of Φ, obtained from existing rainfall data, black dots represent individual parameter values estimated for each month, whereas red curves show the fitted harmonic functions.

4.3.3 Sampling approach

As illustrated in Figure 4-1, application of the forward approach involves sampling the parameter space prior to using the stochastic model. Similarly, application of the inverse sampling approach involves the identification of target locations in the exposure space as the basis for optimization. One approach to sampling both the parameter space (in the forward approach) and exposure space (in the inverse approach) is to define a grid of evenly spaced points over the entire space. However, this can be inefficient, particularly for high-dimensional problems (for a large number of parameters/attributes in the exposure space in the forward/inverse approach). For example, if one wished to sample on a grid of width 10 for the parameter space of the four-parameter

Richardson model, then it would be necessary to evaluate a total of 10^4=10,000 separate parameter sets. This issue is particularly pertinent for the inverse approach, since optimization is required to find a parameter set that corresponds to each point in the exposure space. Therefore, to provide even coverage of the parameter or exposure space while keeping the sample sizes low, two structured sampling techniques have been employed, namely Latin hypercube sampling (LHS) and improved distributed hypercube sampling (IHS).

The objective of the analysis in this paper is to illustrate the inverse approach, by comparing its performance with the forward approach. Therefore and so for consistency, the objective of the sampling approach was to obtain 100 samples within the exposure space. For the forward approach, it is not known *a priori* whether a particular parameter set in the parameter space will yield a sample in our exposure space (i.e. within the plausible range of 50%–150% for each rainfall attribute, as defined in Section 4.3.1), so that the number of samples that need to be drawn from the parameter space is not known. To determine the total number of samples in a computationally efficient manner, we used the Latin hypercube sampling (LHS) method, which allows starting with a small sample size and adding new samples while keeping the previously generated ones. The LHS method involves sampling M variables with a desired sample size N by dividing the range of each variable into N equally probable intervals. N samples are then drawn so that any interval for each variable is only sampled once (Stein, 1987). To add n new sample points, the existing design is re-divided into $(N + n)$ intervals; the N old samples are kept which occupy N intervals, and then n new samples are drawn to fill the remaining n intervals.

Unlike the LHS method, the IHS method (Manteufel, 2001, Beachkofski and Grandhi, 2002) requires that the number of samples be specified *a priori*, but ensures more even coverage of the sampling space. This latter feature is attractive when sparsely sampling potentially high-dimensional spaces, and is therefore recommended to determine the target locations in the exposure space

for the inverse sampling approach. The IHS method is similar to the LHS method, with two additional objectives:

1) The average minimum distance between sample points equals the optimal distance d_{opt}. That is, if the span of each output variable is normalized to 1 so that the entire sample space is a hypercube of volume 1, then each sample point should cover an equal hypervolume (with dimension of M) within the entire space. This gives the optimal distance between sample points, i.e. $\boldsymbol{d_{opt}} = \sqrt[M]{\frac{1}{N_{Sample}}}$.
2) The coefficient of variance (*COV*) of all minima between each pair of sample points is close to zero.

4.3.4 Implementation of forward approach

As mentioned previously, the forward approach has been used to provide a benchmark against which the utility of the proposed inverse approach can be assessed. The approach involves the following steps:

1) The parameter space is constructed by selecting appropriate ranges of the parameters for the stochastic generators;
2) Parameter sets are drawn from the parameter space using a sampling strategy such as the LHS method described in Section 4.3.3;
3) The sampled parameter sets are used to generate time series of the hydrometeorological variables of interest (in this case, rainfall); and
4) The values of the attributes that define the exposure space are calculated for each of the generated hydrometeorological time series.

For the simple four-parameter Richardson model, the 'transition probability' parameters (p_{dd} and p_{wd}) both vary between 0 and 1. The two

'rainfall intensity' parameters (α and β) are for the gamma distribution and should be greater than 0 (Note that their values are mostly between 0 and 1 when calibrated to historical data; see Richardson and Wright, 1984). From a preliminary analysis for our case study, α and β values of 0.56 and 0.10 were obtained respectively, yielding rainfall time series with attributes that are close to the historical data. Therefore, although α and β do not physically have upper bounds and can take any value above 0, their ranges were set to be between 0 and 1 in the forward approach based on their historical values. The use of such a small range ensures that the parameter space surrounding the historical levels of the parameters is sufficiently sampled. For WGEN, the parameter ranges were defined in a similar way, so that the bounds of both the transition probability and rainfall intensity parameters were set to 0 and 1 for all months.

As mentioned previously, for both stochastic models, LHS was used to sample the parameter space. An initial Latin hypercube sample size of 100 parameter sets was used, and this was incremented until 100 rainfall time series were generated with attributes within the plausible bounds of the exposure space.

4.3.5 Implementation of proposed inverse approach

A general description of the inverse approach was provided in Section 4.2.2. The IHS method (Section 4.3.3) was used to determine the target locations for the optimization, which consist of 100 sets of combined levels of the four rainfall attributes that uniformly cover the exposure space.

For each target location, the best-fit parameter sets for both the four-parameter Richardson model and the WGEN model were identified using optimization. The shuffled complex evolution algorithm (Duan et al., 1993) was used as the optimization engine, due to its proven ability for solving complex optimization problems in hydrological studies (Gupta et al., 1999, Thyer et al., 1999, Wang et al., 2010). Based on the general formulation in Eqn. 4.1, the objective function to be minimized for both stochastic models was:

$$F_{obj,i} = \sqrt{\begin{array}{l}\{[\frac{(PD_{s,i}-PD_{his})}{PD_{his}} - \frac{(\widehat{PD_{\iota}}-PD_{his})}{PD_{his}}] * 100\}^2 + \{[\frac{(WD_{s,i}-WD_{his})}{WD_{his}} - \frac{(\widehat{WD_{\iota}}-WD_{his})}{WD_{his}}] * 100\}^2 + \\ \{[\frac{(CDD_{s,i}-CDD_{his})}{CDD_{his}} - \frac{(\widehat{CDD_{\iota}}-CDD_{his})}{CDD_{his}}] * 100\}^2 + \{[\frac{(Pex99_{s,i}-Pex99_{his})}{Pex99_{his}} - \frac{(\widehat{Pex99_{\iota}}-Pex99_{his})}{Pex99_{his}}] * 100\}^2\end{array}}$$

(4.5)

The constraints of the optimization consist of the plausible ranges of the parameters for both models. As mentioned in Section 4.3.4, the plausible range for the probability parameters (p_{dd}'s and p_{wd}'s) is between 0 and 1; for the intensity parameters (α's and β's), which do not have a physical upper limit, we defined the range to be between 0 to 10^4, which was wider than the range used for the forward approach (Section 4.3.4) to enable more extensive searching within the defined range.

For the WGEN model, since a harmonic function has been fitted to the monthly values of each of the probability and intensity parameters (Section 4.3.2.2), the actual decision variables for the optimization were the parameters of the harmonic functions (i.e. β_0 and A, which represent the mean and amplitude respectively, as in Eqn. 4.4). To ensure that the probability parameters were always within 0 and 1 while the intensity parameters were always within 0 and 10^4 during the optimization process, the values of the mean and amplitude for each of these parameters have been optimized sequentially. In the first step, the mean value of each parameter has been optimized with the amplitude kept as zero. Once the mean has been determined, a second optimization was conducted to estimate the amplitude. For example, if the mean of p_{dd} is found by the optimizer to be 0.3 in the first step, its amplitude must be constrained between 0 and 0.3 in the second step to avoid values of p_{dd} going beyond 0 and 1.

It should be noted that in determining the target locations, the IHS only checks the multi-dimensional uniformity of the overall distribution, without considering the physical interpretation for each individual target location. Therefore, it is important to ensure that each target location selected is physically plausible. For example, *PD* should always be less than *Pex99*, and *WD* should never exceed 365 days. For this study, these constraints were

automatically satisfied because a relatively small plausible range of 50%–150% was selected for each attribute. If the rainfall samples are required to show larger variances, it may be necessary to impose additional constraints in the optimization procedure to ensure the resultant samples remain physically plausible.

4.4 Results

4.4.1 The four-parameter Richardson model

4.4.1.1 Forward approach

The coverage of the exposure space obtained by applying the forward approach to the four-parameter Richardson model is shown in Figure 4-5, which shows high variances in some rainfall attributes. In particular, the generated *PD*, *CDD* and *Pex99* can go up to 15,000%, 6,000% and 80,000% of their corresponding historical values, respectively (Figure 4-5a), which are well outside the bounds of the exposure space. The sampled *WD* has lower variance with values up to only 226% of the historical values (since a year contains a maximum of 365 or 366 days), however, these values are still above the upper limit of the exposure space of 150%. The high variance leads to low sampling efficiency – to obtain 100 sets of combined levels of rainfall attributes within our exposure space, a total of 7635 LHS samples of parameter sets had to be generated (i.e. 98.7% samples were unacceptable and discarded). All 7635 sets of rainfall attributes are plotted in Figure 4-5a, with the 100 plausible samples shown in Figure 4-5b.

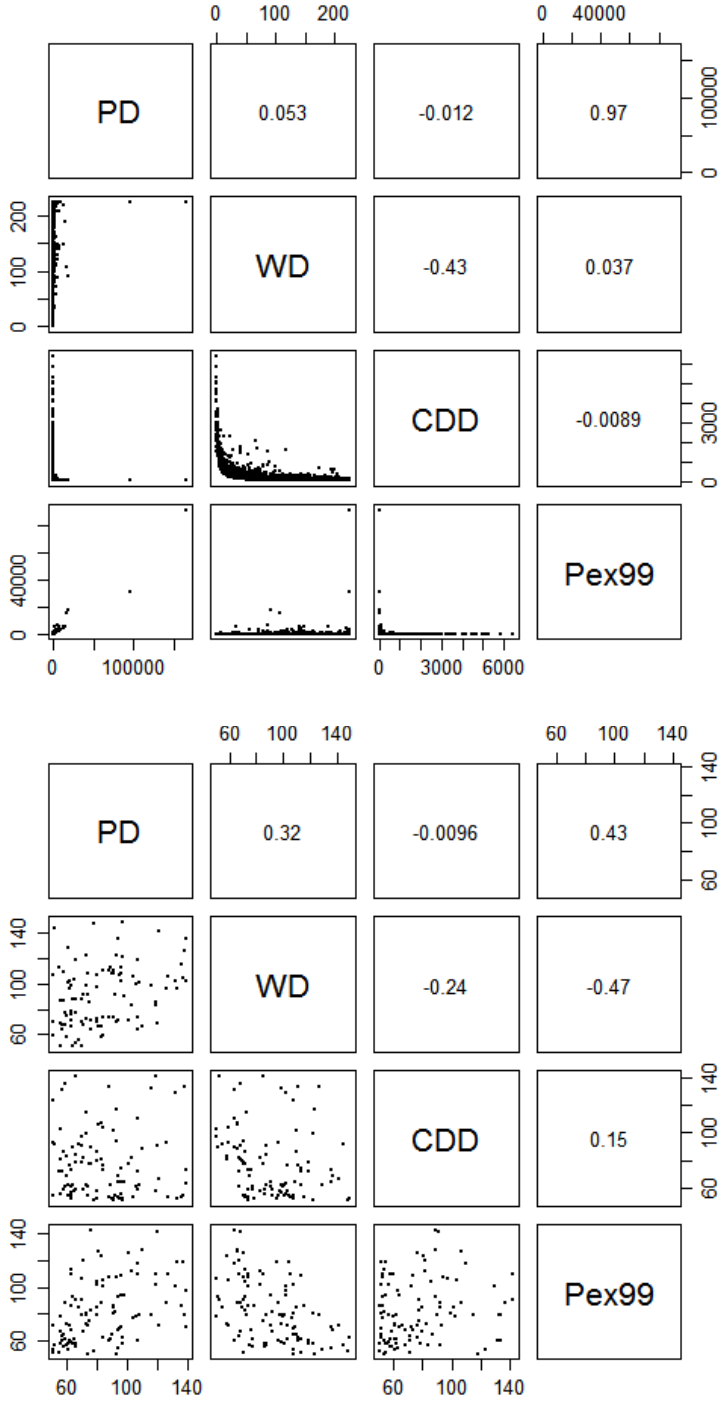


Figure 4-5: The four rainfall attributes described as a percentage relative to the historical values produced by the four-parameter Richardson model, by drawing LHS samples from the parameter space, for (a) all the 7635 samples, and (b) only the 100 samples that fall in the plausible range of between 50% and 150% for each rainfall attribute. The upper-right triangle displays pairwise correlations.

In addition to the issue of inefficient sampling, based on both a visual inspection of the coverage on the exposure space as well as consideration of

the correlation coefficients, it is clear that the coverage of the exposure space is uneven (Figure 4-5b). In particular, samples are clustered in small regions of the exposure space for each rainfall attribute, with other parts of the space receiving limited or no coverage. For example, the correlation between *PD* and *Pex99* is quite high, which results in better coverage over regions closer to the diagonal of the joint distribution of *PD* and *Pex99* than other regions.

The above problems with using the forward approach are most likely due to the non-linear translation from parameters to rainfall attributes through the stochastic generator, so that large variations in certain regions in parameters space result in small variations in exposure space and vice versa. This non-linearity will be further illustrated in the next section with the distribution of parameters identified through the inverse approach.

4.4.1.2 Inverse approach

Figure 6a shows the 100 target locations of desired rainfall attributes that have been determined using the IHS approach (Section 4.3.3). As can be seen, the IHS approach generates samples that appear to be uniformly distributed across the exposure space, with even coverage across each attribute and low cross-correlations between attributes.

The final set of combined levels of attributes corresponding to each of the 100 stochastically generated rainfall time series obtained using the inverse approach is presented in Figure 4-6b. As can be seen, the optimization-based approach is effective in producing the desired levels of rainfall attributes (i.e. target locations), with all of the 100 samples falling within the bounds of the exposure space and with relatively even coverage of the exposure space (Figure 4-6b). Therefore, the inverse approach delivers much better coverage of the exposure space than the forward approach (Figure 4-5).

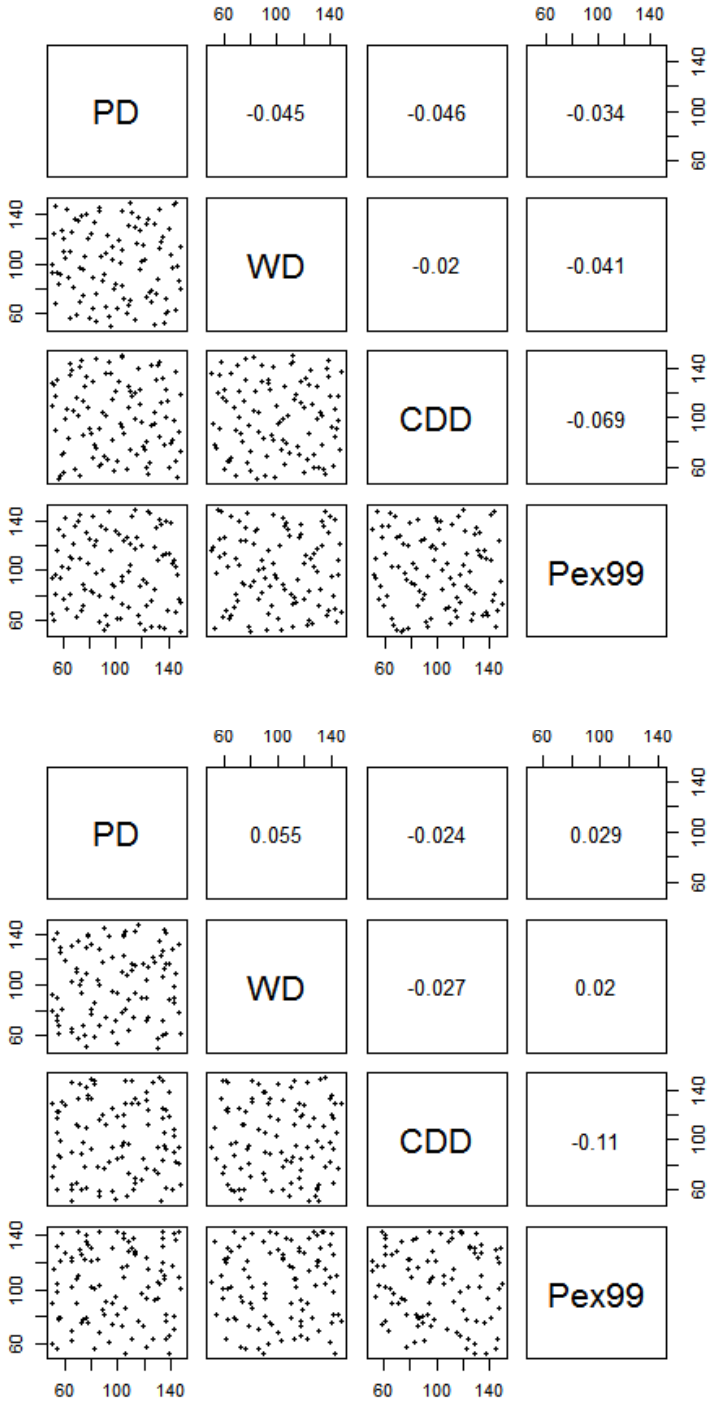


Figure 4-6: The four rainfall attributes described as a percentage relative to the historical values produced by the four-parameter Richardson model for (a) 100 selected IHS sample locations and (b) 100 corresponding optimized locations. The upper-right triangle displays pairwise correlations.

Figure 4-7 shows the values of the 100 parameter sets for the four-parameter Richardson model, identified via application of the inverse

approach, highlighting the non-linear mapping between parameter space and exposure space. This is best illustrated with the non-uniform distribution of the best-fit parameters, in contrast to the uniform distribution of the exposure space (Figure 4-7). Furthermore, the parameters have considerably different ranges compared with the *a priori* [0,1] ranges that were specified for the forward approach. For example, the values of p_{dd} generally vary within a narrower range of 0.5–0.9, whereas values of α are as high as 10. Therefore, the ranges of [0,1] defined for the four parameters in the forward approach (as detailed in Section 4.3.4) can significantly limit the resultant coverage of the exposure space. This also reflects the high degree of non-linearity in the mapping between the parameter values and the exposure space, as a small change in the exposure space may result in a large shift in parameter space.

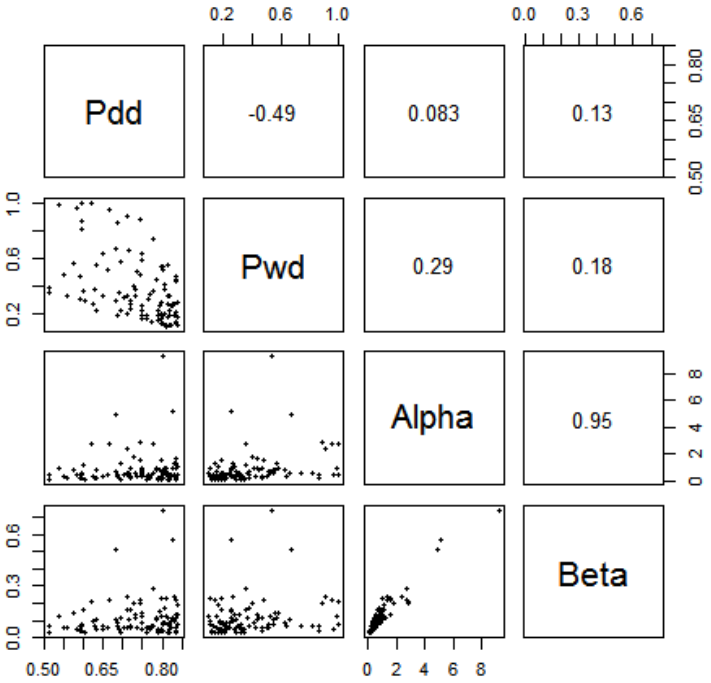


Figure 4-7: Parameters corresponding to the exposure space in Figure 4-6.

Interestingly, for the case study considered, although the inverse approach had the additional step of parameter optimization, the computational time required to obtain 100 samples was 32.6% shorter than for the forward approach. This is likely due to the large number of samples that were discarded in the forward approach.

4.4.1.3 Implications of random sampling on the inverse generation approach

In the above example, we fixed the random seed of the random number generator during the optimization process due to reasons discussed in Section 4.2.3. To illustrate the importance of this aspect of the optimization, we use the analytical expressions in Eqns. 4.2 and 4.3 to estimate the model parameters that will yield individual target locations from a grid consisting five evenly-spaced levels for each of *WD* and *PD* (50%, 75%, 100%, 125% and 150% of their historical values). These locations within the exposure space are given as green dots in Figure 4-8. We then generated 100 stochastic replicates from each of these parameter sets with different random seeds, which are shown as blue and red scatter about each of the target locations in Figure 4-8.

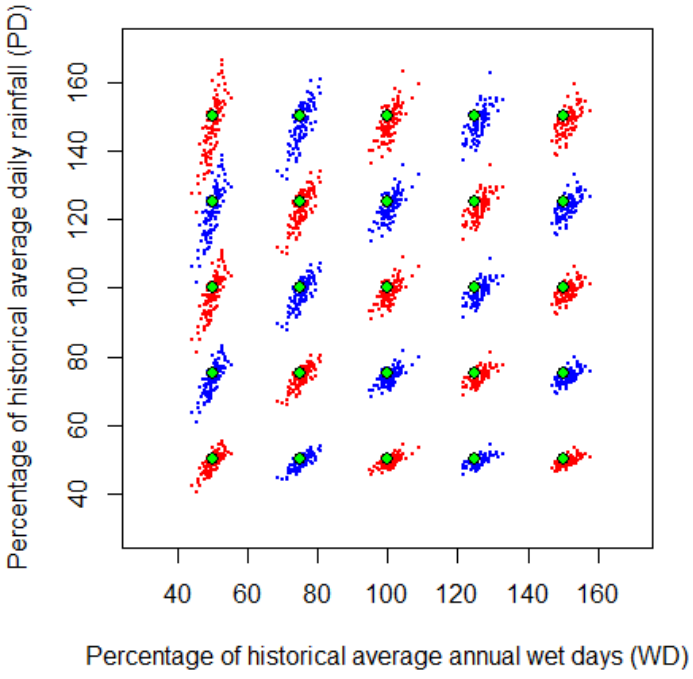


Figure 4-8: Stochastic behavior of the four-parameter Richardson model in simulating *WD* and *PD*, with 100 replicates for the analytically-solved parameter set corresponding to each single target location. Two separate colors (i.e. red and blue) were used to differentiate between adjacent target locations.

The stochastic nature of the model is clear for all target locations. For each parameter set, the 100 replicates of *WD* vary up to ±10% around their target

level, which is similar for all target levels of *WD* and *PD*. In contrast, the 100 replicates of *PD* are closer to the target level for smaller *PD* (e.g. up to ±10% around where the target level is 50%), while for larger *PD* target levels the spread among replicates increases substantially (e.g. up to ±40% around where the target level is 150%). Compared with the sampling resolution required in this study (shown in Figure 4-6a), the variability in Figure 4-8 is in fact much higher, which can adversely affect the capacity of the optimizer to find parameters that correspond to each target location, as discussed in Section 4.2.3.

4.4.2 The WGEN model

4.4.2.1 Forward approach

The coverage of the exposure space obtained by applying the forward approach to the WGEN model is shown in Figure 4-9. Similar to the results for the four-parameter model (Section 4.4.1.1), the forward approach shows low efficiency: to obtain 100 sets of rainfall attributes within the range of the exposure space, 1453 LHS samples of WGEN parameter sets were required (Figure 4-9a), which means that 93.1% of samples were discarded. With the 100 plausible sets in Figure 4-9b, the coverage of the exposure space is poor, which is also evident through the high pairwise correlations (such as between *PD* and *Pex99* and between *WD* and *CDD)*.

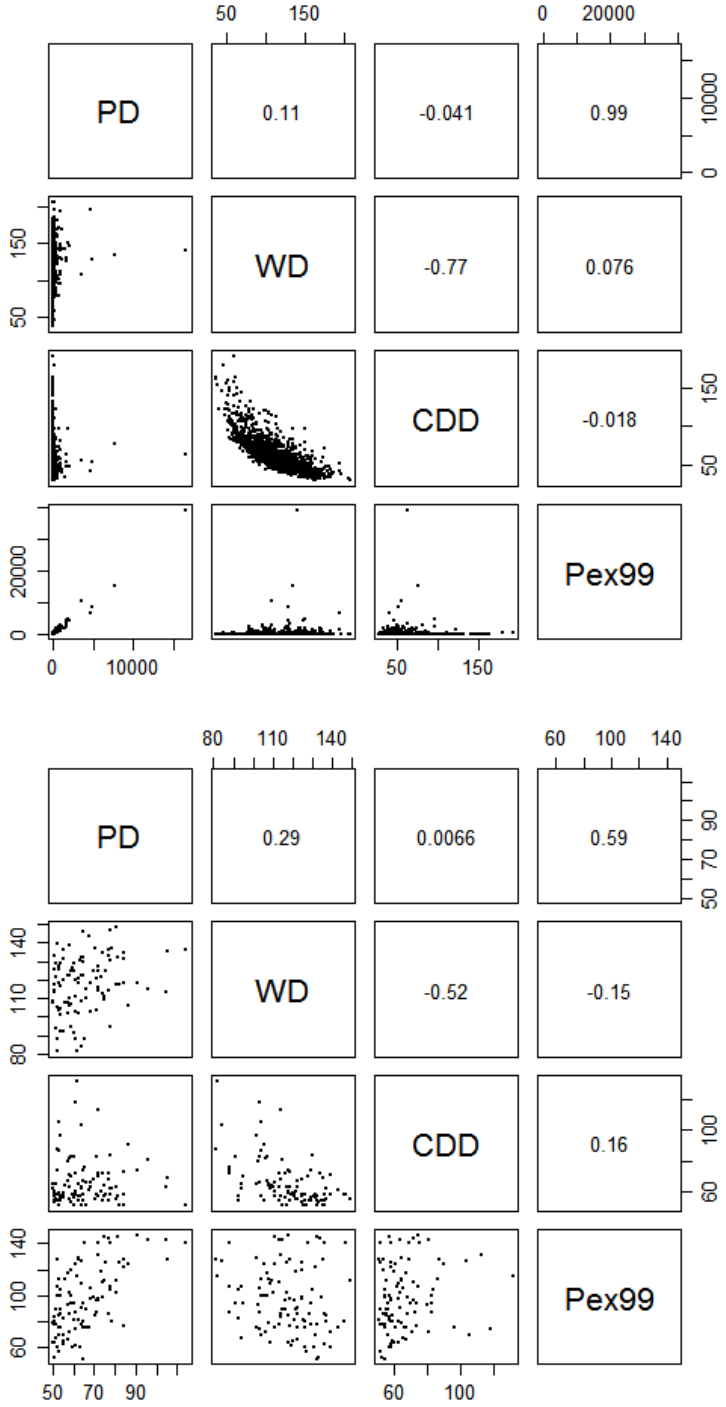


Figure 4-9: The four rainfall attributes described as a percentage relative to the historical values produced by the WGEN model, by drawing LHS samples from the parameter space, for (a) all 1453 samples, and; (b) only the 100 samples that fall in the plausible range of between 50% and 150% for each rainfall attribute. The upper-right triangle display pairwise correlations.

4.4.2.2 Inverse approach

To examine the performance of the inverse approach with WGEN, the 100 target locations which have been determined using the IHS approach (Section 4.3.3) are plotted in Figure 4-10a. The final optimized set of attributes corresponding to each of the 100 stochastically generated rainfall time series is presented in Figure 4-10b. The inverse approach is generally effective in evenly covering the exposure space and reproducing these target locations. In particular, this approach delivers much better coverage of the exposure space than the forward approach (Figure 4-9) in the following aspects:

1) All of the 100 samples are within the plausible output space defined in Table 4-1, suggesting effective control over the values of individual rainfall attributes; and
2) The joint distribution of multiple rainfall attributes is much more uniform across the exposure space, and the pairwise correlations between different attributes are reduced.

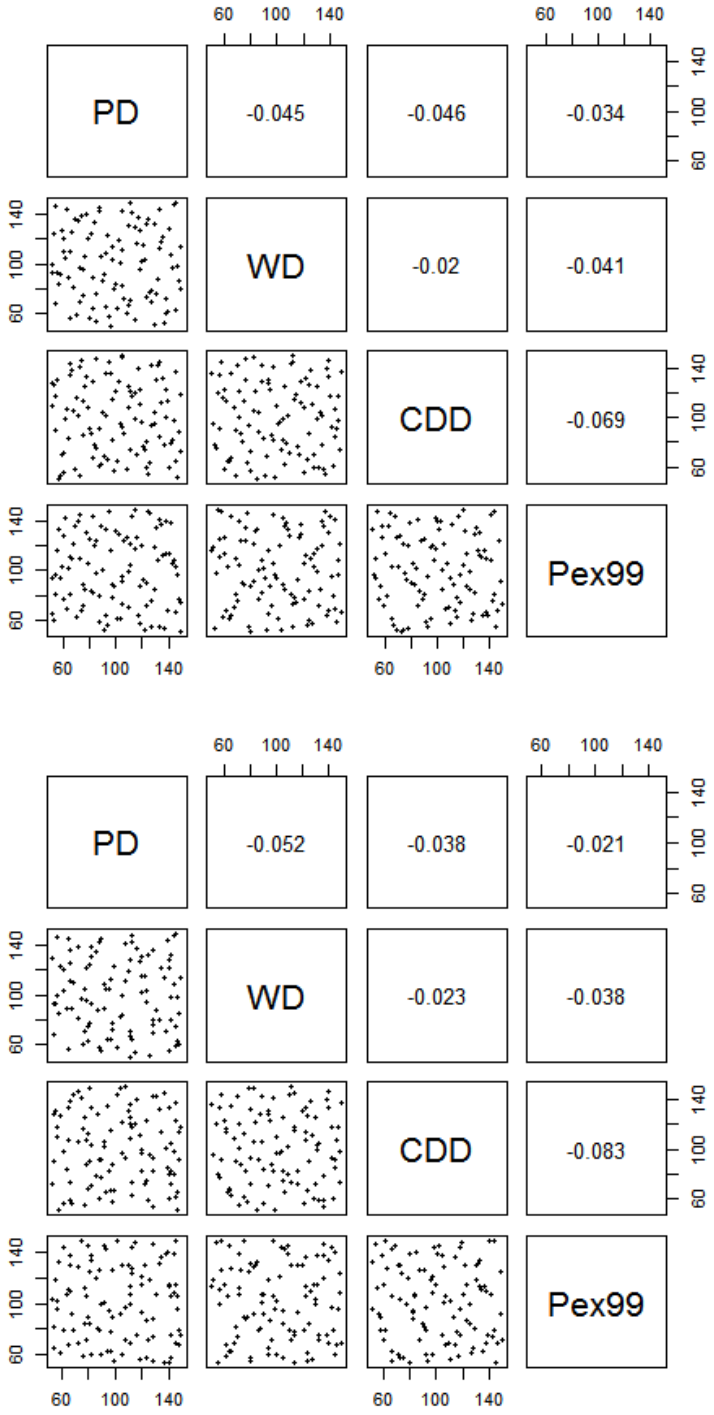


Figure 4-10: The four rainfall attributes described as a percentage relative to the historical values produced by the WGEN model for (a) 100 selected IHS sample locations and (b) 100 corresponding optimized locations. The upper-right triangle displays pairwise correlations.

As an alternative approach to assessing the relative uniformity of the sampling in the exposure space, the minimum distances between sample points in the exposure space are compared for both the forward (as orange

dots in Figure 4-11) and inverse approaches (as blue dots in Figure 4-11). The results show that the inverse sampling approach produces a more uniform coverage, as the minima between sample points are closer to the optimal distance, d_{opt} = 0.32 (see Section 4.3.3). Furthermore, the coefficient of variance (*COV*) of these minimum distances is also much lower (i.e. 0.52 for the inverse approach compared with 2.90 for the forward approach).

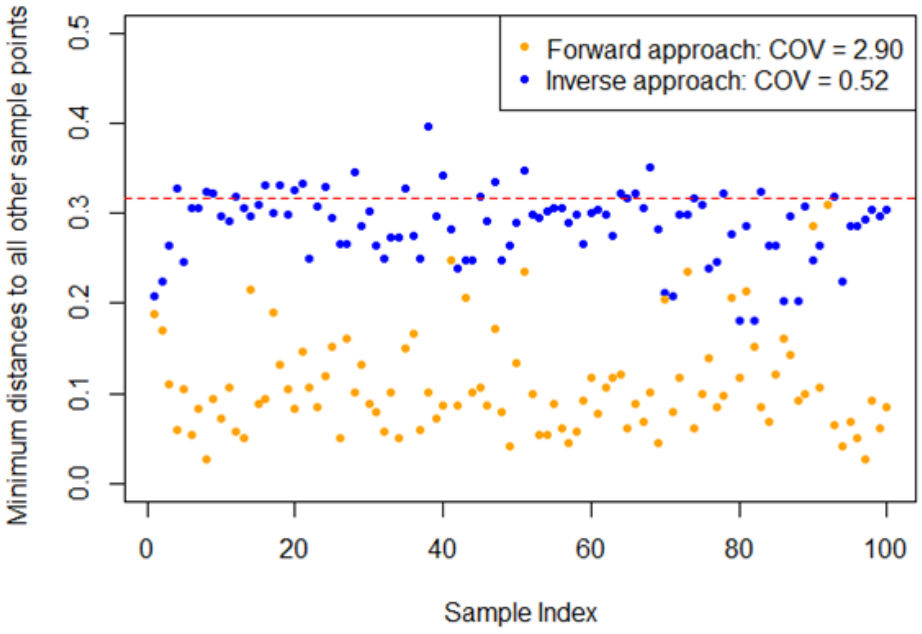


Figure 4-11: Minimum distances between sample points obtained from the forward and inverse approaches with the WGEN model. The dashed line shows the optimal distance with 100 plausible samples (d_{opt} = 0.32).

It is worth noting that to obtain 100 sample points on the exposure space with the WGEN model, the overall execution time required for implementing the inverse approach is 73% longer than that for the forward approach. This is most likely due to the difficulty in solving optimization problems with a larger number of parameters, as a result of the larger search space that has to be explored. However, the inverse approach ensures uniform coverage of the exposure space with the desired resolution, which is the key objective for constructing the exposure space. In contrast, the forward approach failed to obtain such coverage. Therefore, although associated with a higher computational expense, the proposed inverse approach is the only way of achieving the desired coverage of the exposure space.

4.5 Discussion

This study presented a framework for sampling various rainfall conditions to construct an exposure space for scenario-neutral climate impact assessments. Here, we discuss some practical considerations, as well as possible future adaptations of the framework.

4.5.1 Design of exposure space to represent more complex potential climate changes

The four rainfall attributes considered in the exposure space for this study (i.e. *PD*, *WD*, *CDD* and *Pex99*) are good descriptors of a range of changes of annual precipitation characteristics. However, there is also a range of other rainfall attributes that might be important when considering the impact of climate change, such as changes at seasonal or interannual timescales (e.g. Kwon et al., 2009, Johnson et al., 2011, Christensen et al., 2007). Furthermore, potential future variations in other climatic features, such as temperature, solar radiation and evapotranspiration, may also have a substantial impact on water resources (for examples see Prudhomme and Williamson 2013, Chiew and McMahon 2002).

The inverse approach presented here is sufficiently flexible to cater to all attributes in the exposure space that are of interest (provided they can be generated with an appropriate stochastic generator), although this comes at the expense of additional computational cost. Considering the trade-off between the flexibility of producing different climate patterns and computational effort, it is important to identify key hydrometeorological variables of interest, as well as their attributes, based on an understanding of the behavior of the system being analyzed.

Depending on the specific hydrometeorological variables involved, the format of the objective function may require modification from Eqn. 4.1, which was designed assuming multiplicative perturbations to attributes (e.g. changes expressed as a percentage of the historical value). For example, temperature changes are typically represented in an additive way (e.g.

increases in temperature by degrees Celsius; for examples see Chiew and McMahon 2002, Kingston et al. 2009), and this would require an adjustment to the objective function in Eqn. 4.1.

In this study, the boundary of the exposure space was set at 50-150% of the historical values of each attribute, which is sufficiently wide to incorporate a large number of possible changes in each of the rainfall attributes, while also using the same percentage changes across attributes to facilitate illustration. However, this framework can be easily adapted to incorporate tailored bounds for the exposure space, which should be carefully selected to suit the case study under consideration. In particular, if the bounds deviate too far from present conditions, a significant portion of samples will be unrealistic, even when extreme climate change impacts are considered. Conversely, if the bounds are too narrow, system response to some plausible climatic changes might not be considered (Whateley et al., 2014). Multiple sources of information could be considered in selecting these bounds, including GCM-based climate projections (e.g. Collins et al., 2013) of possible future climatic changes, and additional lines of evidence on possible changes to key variables, such as from long-term paleoclimatology reconstructions (e.g. Ho et al. 2015, Ault et al. 2014, Hansen and Sato 2012). In addition, it is worth specifying an exposure space with bounds that are wider than the range suggested from all currently available sources of information, so that additional climate change projections can be included in the analysis as they are developed (Steinschneider and Brown, 2013).

Finally, when determining the target locations consisting of different hydrometeorological attributes on the exposure space, it is desirable to ensure the physical realism of each individual location so that corresponding time series can be obtained with the aid of stochastic weather generators. This requires not only ensuring that the target levels of individual attribute are realistic (such as the constraints for the levels of *WD*, as discussed in Section 4.3.5), but also maintaining physically plausible relationships among multiple attributes (for example, a target location cannot consist a *WD* value of 100 days with a *CDD* value of 300 days, because this combination means that the

annual average wet day is 100 days while the annual average dry spell length is 300 days, which is physically unrealistic).

4.5.2 Stochastic generation of the exposure space

In this study, we used a sample size of 100 to represent different levels of changes for each individual attribute considered in the exposure space, with fixing the random seed across replicates to facilitate improved convergence during the optimization process. In this way, however, there is likely to be limited variability in between time series corresponding to different points on the exposure space.

This issue can be addressed in at least three ways:

1. Increase the sample size, and thus the coverage resolution in the exposure space. Increasing the exposure space resolution is likely to be particularly useful when the number of attributes increases, as this will lead to a corresponding increase in the dimensionality of the exposure space.
2. The length of each sample can be increased. Currently, the length is equal to the length of the historical data series (i.e. 15 years). However, it would be trivial to allow the simulation to run for longer periods of time to obtain greater stochastic variation. This will require the same number of optimized parameter sets, although because of the use of the same initial seed, there will still be significant similarities between individual samples.
3. The procedure can be repeated multiple times with different random seeds for each iteration, thereby generating multiple replicates. This would substantially increase the level of stochastic variability, although at the expense of additional computational time.

The appropriate sample size, length of each sample and number of replicates are likely to depend on the case study considered, as well as available computational resources.

4.5.3 Equi-finality in the optimization process

When using optimization to search for best-fit solutions, equi-finality issues are likely to arise (i.e. multiple solutions leading to the same results in the objective function, therefore they are not distinguishable during optimization, see Beven and Freer, 2001). This problem is further complicated within the proposed inverse approach, as the values of the objective function for optimization are based on results from stochastic models.

Equi-finality issues are likely to be greatest for low-dimensional exposure spaces, since higher-dimensional exposure spaces add constraints to the parameter space. For example, the chance that two contrasting combinations of P_{dd} and P_{wd} lead to the same combination of *WD* and *CDD* is much lower compared to that resulting in the same level of *WD* in isolation. Thus, increasing the number of attributes considered could have the additional advantage of reducing the number of feasible parameter sets to be considered.

It is worth noting that although equi-finality is likely to occur when the proposed inverse approach is implemented, the aim of the approach is to identify time series of outputs from the stochastic generator that result in desired values of the attributes included in the exposure space, and not to the identification of the resulting parameters in the stochastic generator, as discussed in Section 4.2.3. However, when different parameter sets lead to the same combination of attributes on the exposure space, the different time series of hydrometeorological variables which they produce can consist of varying degrees of physical realism. Therefore, checking the physical realism of the generated time series can potentially help to eliminate unrealistic parameter sets and thus resolve any equi-finality issues.

4.5.4 Computational efficiency and execution time

In our particular implementation of the proposed inverse approach, the computational time required to produce 100 evenly distributed samples is

around eight hours using an Intel Xeon E3 (2.60GHz, 8 Cores) processor with 32GB RAM for both the four-parameter model and the WGEN, suggesting a relatively high computational demand. However, in general, the computational effort required is dependent on a number of practical specifications, including the operating system, programming language and algorithm used. As the key focus of this study is to introduce and illustrate a new method, the above-mentioned specifications have not been optimized for computational efficiency. We have used the R package *lhs* (Carnell, 2012) for obtaining samples over the exposure space with the IHS and LHS methods, together with the shuffled complex evolution algorithm embedded in the R package *hydromad* (Andrews and Guillaume, 2013) for solving the optimal parameter values for the stochastic generator. We have also developed our own R-scripts to execute the four-parameter Richardson and the WGEN models. It is expected that an improved integration of these different modelling components with the aid of other programming languages, such as Fortran or C++, will further increase computational efficiency.

4.6 Conclusions

Generation of exposure spaces for scenario-neutral climate impact assessments should consider a range of potential variations in relevant hydrometeorological variables, including shifts in the average, intermittency, variability and extremes. The 'exposure space' describes the range of conditions of interest that a system may be exposed to under a future climate, and this paper presents and demonstrates an inverse approach to stochastically generating hydrometeorological time series to uniformly cover this exposure space.

The utility of the proposed inverse approach is benchmarked against a forward approach for rainfall generation for a South Australian catchment, using two Richardson-type stochastic rainfall generators of varying complexity. The results highlight the highly non-linear translation from parameter space to exposure space, and thus the need for the proposed inverse approach in order to obtain a relatively uniform coverage of the exposure

space. For both models, the inverse approach demonstrates better control of the sampling range, with 100% of samples falling within the exposure space. Furthermore, the uniformity of the coverage of the four-dimensional exposure space is substantially improved.

Several potential adaptations for future implementations of the framework have been discussed, which are: (1) the design of the exposure space to represent more complex changes in climate; (2) improvements to the way that stochastic samples in the exposure space are generated; (3) ways of reducing the effects of equifinality during the optimization process; and (4) methods for increasing computational efficiency. The flexibility of the proposed inverse approach enables consideration of all climate attributes of interest at the desired resolution, thereby expanding the applicability of the scenario-neutral approach to evaluating a water resource system's sensitivity to a wide range of plausible changes in climate.

4.7 Acknowledgements

The authors wish to thank Christel Prudhomme and an anonymous reviewer for their thoughtful comments on the manuscript. Seth Westra's time was supported by Australian Research Council Discovery project DP150100411.

Appendix 4A Supplementary to Chapter 4

Table 4A.1. Definition of terminology used in Chapter 4.

Terminology	Definition
exposure space	A large range of plausible hydrometeorological conditions that a system may confront in the future.
hydrometeorological attributes	The statistics calculated from the time-series of hydrometeorological variables (e.g. rainfall, temperature and potential evapotranspiration) which describe their average conditions, variations, extremes, as well as other features. For example, temperature is considered as a hydrometeorological variable, with potential attributes including annual average temperature, annual maximum temperature and seasonal average temperature, etc.
low-dimensional exposure space	An exposure space that consists of a small number of hydrometeorological variables and their attributes.
high-dimensional exposure space	An exposure space that consists of a large number of hydrometeorological variables and their attributes.
forward approach	The approach of perturbing the parameters of stochastic generators over some pre-defined 'parameter space' to yield an exposure space.
historical scaling	The approach to obtain a desired exposure by applying additive and/or multiplicative scaling factors directly to the historical time series of hydrometeorological variables (as in Kay et al., 2014; Prudhomme et al., 2010; Prudhomme and Williamson, 2013).
forward-plus-scaling approach	The approach that combines the use of the forward and scaling techniques to generate an exposure space. In this study, this specifically refers to the approach illustrated in Steinschneider and Brown (2013). In this approach, the parameters of a stochastic generator (namely, Markov chain transition probabilities and the autoregressive model for low-frequency variability) are first perturbed to obtain stochastic sequences without changing the historical rainfall intensity. The wet-day rainfall intensity in the stochastic sequences is subsequently quantile-mapped to yield a set of daily rainfall series with desired levels of rainfall attributes.
inverse approach	The approach aims to generate an exposure space by first selecting the desired values of the attributes of interest in the exposure space, followed by an optimization step to identify the stochastic generator parameters that produce stochastic sequences with these attributes.
target level	A desired level for a single meteorological attribute that is included in the exposure space.
target location	A desired combination of the target levels for each of the hydrometeorological attributes that are included in the exposure space.
best-fit parameters for the stochastic generator	The parameter set for the stochastic generator that produces hydrometeorological time series with the levels of attributes corresponding to a particular target location.

Appendix 4B Copy of Paper from Chapter 4

Guo, D., Westra, S., Maier, H.R. 2016, An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies, *Journal of Hydrology*, Available online 22 March 2016, ISSN 0022-1694. doi:10.1016/j.jhydrol.2016.03.025.

Journal of Hydrology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies

Danlu Guo *, Seth Westra, Holger R. Maier

School of Civil, Environmental and Mining Engineering, The University of Adelaide, North Terrace, Adelaide, SA 5000, Australia

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Scenario-neutral climate impact study
Stochastic generator
WGEN
Exposure space
Inverse approach
Optimization

SUMMARY

Scenario-neutral approaches are being used increasingly for climate impact assessments, as they allow water resource system performance to be evaluated independently of climate change projections. An important element of these approaches is the generation of perturbed series of hydrometeorological variables that form the inputs to hydrologic and water resource assessment models, with most scenario-neutral studies to-date considering only shifts in the average and a limited number of other statistics of each climate variable. In this study, a stochastic generation approach is used to perturb not only the average of the relevant hydrometeorological variables, but also attributes such as the intermittency and extremes. An optimization-based inverse approach is developed to obtain hydrometeorological time series with uniform coverage across the possible ranges of rainfall attributes (referred to as the 'exposure space'). The approach is demonstrated on a widely used rainfall generator, WGEN, for a case study at Adelaide, Australia, and is shown to be capable of producing evenly-distributed samples over the exposure space. The inverse approach expands the applicability of the scenario-neutral approach in evaluating a water resource system's sensitivity to a wider range of plausible climate change scenarios.

1. Introduction

Scenario-neutral approaches are being used increasingly to assess the possible impact of climate change on the performance of water resources systems (Brown et al., 2012; Brown and Wilby, 2012; Dessai and Hulme, 2004; Nazemi and Wheater, 2014), as well as social and ecological systems (Gao et al., 2016; Poff et al., 2015). The information generated from these approaches can be used to assess system vulnerability under alternative climate change scenarios, and to calculate climatic thresholds at which system performance begins to change abruptly (Brown et al., 2011; Poff et al., 2015). Scenario-neutral approaches can also accommodate changes in climate projections without the need for additional analysis (Prudhomme et al., 2010), and can help to identify the important hydrometeorological variables, or particularly critical states of these variables that affect the system under consideration. The latter feature is particularly useful for selecting: (1) climate models; (2) strategies to generate future rainfall conditions from GCM-based projections (known as statistical downscaling); and (3) alternative 'lines of evidence' (e.g. expert opinion and data from the paleo-climatic record) that can provide useful information about these variables. Ultimately, this allows for the development of a more complete set of projections that describe how these variables might change in a greenhouse gas-enhanced climate (Nazemi et al., 2013; Singh et al., 2014; Steinschneider and Brown, 2013; Vano et al., 2015).

Central to the scenario-neutral approach is the analysis of system sensitivity to a range of hydrometeorological conditions. Such analyses involve exposing the system to perturbed hydrometeorological forcing data that reflect various hydrometeorological conditions that the system may confront in the future (referred to as the 'exposure space'). To this end, it is important to consider the possible variations not only in the average states of the relevant hydrometeorological variables, such as annual average rainfall and potential evapotranspiration (see Kay et al., 2014; Prudhomme et al., 2013), but also their other attributes, including extremes, seasonality and interannual variability (Meselhe et al., 2009; Moody and Brown, 2013; Prudhomme et al., 2010; Steinschneider and Brown, 2013). Indeed, assessments of historical and/or future changes to rainfall as a result of climate change have already indicated different changes to the averages (Collins et al., 2013), extremes (Ajami et al., 2007; Alexander et al., 2006; Westra et al., 2013, 2014), temporal distribution (Rajah et al., 2014) and low-frequency variability (e.g. Johnson et al., 2011) of rainfall throughout the world. Similarly complex changes to other relevant hydrometeorological variables might also be expected, including potential evapotranspiration, and snowfall and melt.

* Corresponding author.
E-mail address: Danlu.Guo@Adelaide.edu.au (D. Guo).

http://dx.doi.org/10.1016/j.jhydrol.2016.03.025
0022-1694/

One approach to generating perturbed hydrometeorological forcing data is by applying scaling factors to historical records of each of the relevant hydrometeorological variables. These factors can be applied at annual or monthly scales (Kay et al., 2014; Paton et al., 2013; Prudhomme et al., 2013, 2010; Singh et al., 2014), or different factors that can be applied across different quantiles in the entire distribution (Nazemi et al., 2013). Although such approaches might be viable for perturbing a small number of hydrometeorological variables and their attributes (i.e. low-dimensional exposure spaces), the capacity of these to represent the potentially complex variations in a wider range of variables and attributes (i.e. high-dimensional exposure spaces) is likely to be limited. Consequently, when using scaling factors to perturb historical data for climate impact assessments, the resultant projections may not show the full range of variability that can be expected in a greenhouse gas-enhanced climate (Prudhomme et al., 2013, 2010; Steinschneider and Brown, 2013).

The use of stochastic generators has been proposed as an alternative to scaling factors to generate hydrometeorological data in a way that can account for a wider range of possible changes (Whateley et al., 2014). Some recent advances include the use of a multi-site weather generator that is capable of producing realistic time series of meteorological variables with shifts to the mean, standard deviation, extremes, daily-scale Markov transition probabilities and low-frequency (interannual) variability (for examples see Steinschneider and Brown, 2013; Wilby et al., 2014; Yates et al., 2015). This is achieved through the perturbation of stochastic model parameters (including the transition probabilities and the autocorrelation coefficient) and the subsequent application of quantile correction, which, in combination, can be used to generate the high-dimensional exposure space. A challenge with this approach, however, is that it is difficult to assess *a priori* which parameters of the stochastic generator should be modified to produce time series at pre-specified points in the exposure space, potentially leading to insufficient exploration of the exposure space. This challenge arises both as a result of the non-linear mapping between the parameters of a stochastic generator and the statistics of the hydrometeorological variables, as well as due to the stochastic nature of the model, which means that a single parameter set will produce hydrometeorological data that span multiple points on the exposure space (Steinschneider and Brown, 2013).

In order to address the shortcomings of existing approaches in generating hydrometeorological data to form the exposure space, we introduce the concept and framework for an inverse approach with demonstration on a case study. The proposed inverse approach enables stochastic generators to be used to generate time series that uniformly span the desired range of the hydrometeorological variables and attributes of interest, and thus provides uniform coverage of the exposure space to serve the needs of scenario-neutral climate impact assessments. Although generally applicable to any parametric weather generator, this paper focuses on applying the method to rainfall time series for the following reasons:

1. Although stochastic generators have been used to generate a range of weather variables, including temperature, humidity, and wind (e.g. Racsko et al., 1991; Semenov and Brooks, 1999), the majority of applications have focused on the generation of rainfall data, due to their importance as an input to many water resource assessments (e.g. Chiew and McMahon, 2002b; Jones and Thornton, 1993).
2. At daily or shorter timescales, rainfall is intermittent, highly skewed (with rainfall series typically exhibiting a large number of moderate rainfall days and a small number of very heavy rainfall days), and exhibits variability at seasonal, interannual and longer time scales (Bastola et al., 2011; Dubrovský et al., 2000). As a result, rainfall is often regarded as a particularly challenging variable to simulate stochastically.
3. There has been a substantial amount of work on developing stochastic generation models to both generate replicates of historical rainfall data (Beven, 1987; Boughton and Droop, 2003; Chen and Brissette, 2014; Clark and Slater, 2006; Frost, 2004; Furrer and Katz, 2008; Langousis and Kaleris, 2014; Langousis et al., 2015), as well as downscaling GCM-based climate projections (Allen and Pruitt, 1986; Bastola et al., 2011; Fowler et al., 2007; Jones et al., 2011; Kay and Jones, 2012; Wilby et al., 2014; Yates et al., 2015).

The remainder of this paper is structured as follows. In Section 2, we illustrate the alternative approaches that are currently available for generating an exposure space, including the historical scaling, forward and inverse approaches. This section also provides details of the proposed inverse approach. Section 3 introduces a case study and two stochastic generators that are used to illustrate both the proposed approach, as well as a simple forward approach that is used as a basis of comparison. The results are given in Section 4, followed by conclusions in Section 5.

2. Proposed inverse approach to exposure space generation

2.1. Rationale for an inverse approach to perturbing stochastic model parameters

As described in the introduction, a central feature of scenario-neutral approaches is the exploration of a water resource system's response to a range of different hydrometeorological conditions. This range of hydrological variables (e.g. rainfall, temperature, evapotranspiration) and the set of attributes of these variables (e.g. annual average, variance, seasonal differences, extremes) are collectively referred to as an 'exposure space', and represent the range of conditions of interest that a system may be exposed to under a future climate. For example, if a scenario-neutral approach was to be used to evaluate system sensitivity to changes in the average, variability and extremes of rainfall, then this would require generating a three-dimensional exposure space with each dimension representing one of the rainfall attributes.

Fig. 1 illustrates the conceptual approaches that could be used to generate an exposure space **E**, which consists of the plausible future changes (represented as the gray shaded region, with the origin corresponding to no change) in various rainfall attributes of interest (represented by two axes A_1 and A_2, which refer to two generic rainfall attributes or groups of attributes). Two techniques are involved in the perturbation approaches – namely scaling of rainfall time series and stochastic rainfall generation (as shown in the two green squares). We use the term 'scaling' in the figure to collectively refer to perturbations that are directly applied to rainfall time series, through the use of change factors at annual, monthly or other time scales (Kay et al., 2014; Prudhomme et al., 2010; Prudhomme and Williamson, 2013), or more complex methods, such as quantile mapping (as used in Steinschneider and Brown, 2013). Consequently, the scaling technique can only modify rainfall intensity on wet days. The term 'stochastic generation' in the figure refers to indirect modification of the rainfall time series through changing the parameters of stochastic generators (as used in Dubrovský et al., 2000; Jones and Page, 2001; Steinschneider and Brown, 2013). The parameter space Θ consists of two axes of θ_1 and θ_2, which refer to two generic parameters or groups of parameters. The plausible ranges for all parameters are represented by the gray shaded region, while the

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

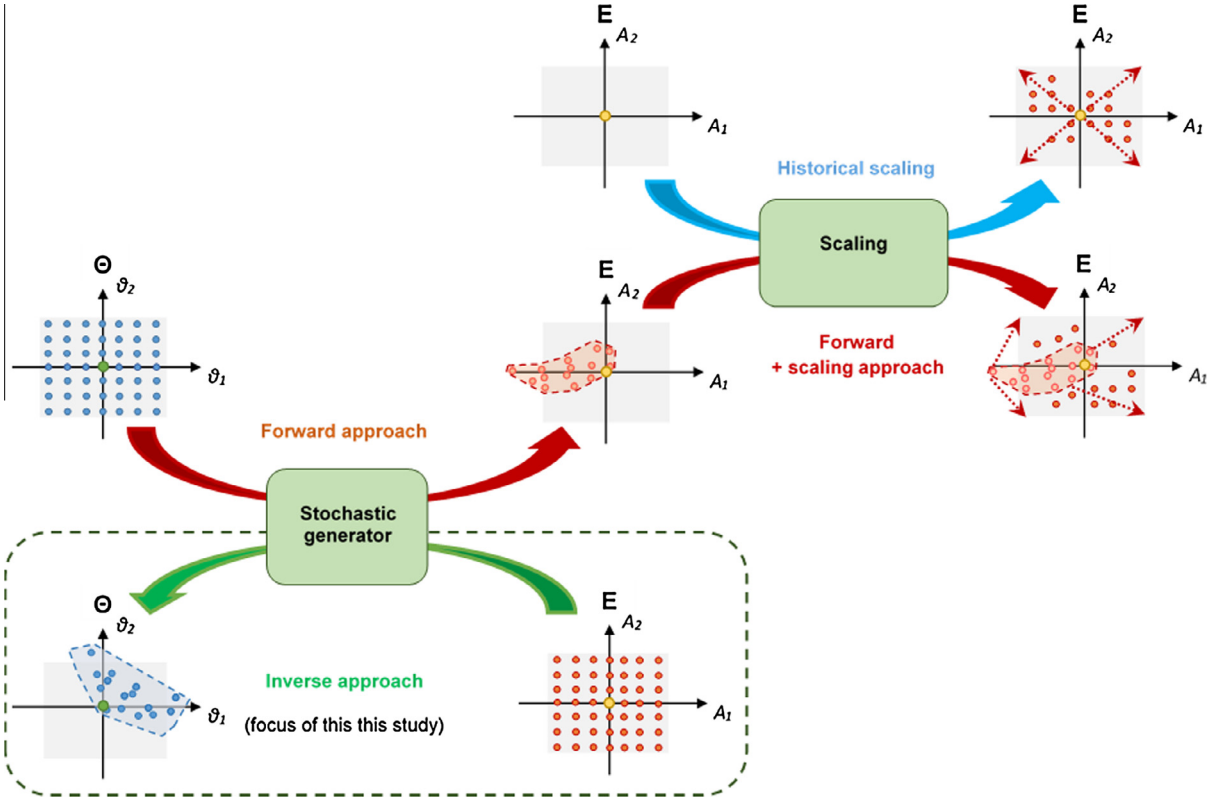


Fig. 1. Outline of alternative approaches to developing an exposure space **E** for scenario-neutral climate impact assessments. Refer to main text for an explanation of each figure element. The focus of this study is on the approaches enclosed in the black dashed box.

origin represents the set of parameters corresponding to the historical rainfall condition.

The first approach, 'historical scaling' (as shown in the top-right corner of Fig. 1) is analogous to the approach used by Prudhomme et al. (2010, 2013), and Kay et al. (2014), in which additive and/or multiplicative scaling factors are applied directly to historical hydrometeorological time series to obtain the desired changes in the relevant variables (usually rainfall and potential evapotranspiration). Although conceptually simple, this approach is not capable of representing variations in the rainfall intermittency, such as the frequency and persistence of dry-/wet-day occurrence. Furthermore, it is difficult to apply this approach to higher-dimensional exposure spaces, since it becomes difficult to develop an approach to scale each attribute independently of the other attributes. Consequently, it can be difficult to sample some regions of the exposure space.

The remaining approaches use stochastic weather generators to obtain perturbed rainfall time series. The 'forward' approach (as illustrated in the middle of Fig. 1) involves perturbing the parameters of stochastic generators over some pre-defined 'parameter space' to yield an exposure space (Dubrovský et al., 2000; Jones and Page, 2001). However, the non-linear mapping between the parameters of a stochastic generator and the attributes of the hydrometeorological variables means that it is unlikely that the full range of the desired exposure space will be covered. Conversely, some perturbations may lead to rainfall attributes with levels out of the defined plausible ranges of the exposure space. Consequently, further scaling may still be necessary after application of the forward approach. Steinschneider and Brown (2013) used this combined 'forward-plus-scaling' approach by firstly perturbing the parameters of a stochastic generator (including Markov chain transition probabilities and the autoregressive model for low-frequency variability) to obtain stochastic sequences without changing the historical rainfall intensity; the wet-day rainfall intensity in the stochastic sequences was subsequently quantile-mapped to yield a set of target daily rainfall series with desired levels of rainfall attributes. Although this approach is likely to provide a much better coverage of the exposure space, some portions of the exposure space may still remain poorly represented because of the difficulty in finding parameters that will result in all combinations of the hydrometeorological attributes of interest.

The limitations of both the historical scaling and forward approaches motivate the 'inverse' approach proposed in this paper (bottom of Fig. 1). Here, the desired values of the attributes of interest in the exposure space are the starting point for the analysis, followed by an optimization step to identify the stochastic generator parameters that produce stochastic sequences with these attributes. This approach provides control over the level of coverage of the exposure space, as required for the implementation of scenario-neutral approaches to climate impact assessments.

2.2. Overview of the inverse approach

To generate hydrometeorological time series with a range plausible attribute levels, the inverse approach is proposed as follows, which involves two primary steps:

(1) *Identify a set of 'target' levels for each attribute included in the exposure space.* In order to achieve an even coverage of the exposure space, we first select the desired levels we would like to sample for each attribute included in the exposure space (referred to as 'target levels'). A number of different approaches can be used to select and combine the target levels (which produce individual 'target locations' in the exposure space), including sampling on a regular grid, or using more computationally efficient sampling methods, such as Latin hypercube sampling (Stein, 1987) or Hammersley sampling (Halton, 1960; Hammersley, 1960).

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

(2) *Generate hydrometeorological time series that satisfy each target set of attributes.* For each target location of the exposure space, we combine stochastic weather generation with a formal optimization approach to identify the best-fit parameter set for the stochastic generator. This parameter set should be capable of producing hydrometeorological time series with the levels of attributes corresponding to that particular target location, as detailed below.

During the optimization process, the decision variables are the parameters of the stochastic generator. The objective is to identify the parameters of the stochastic rainfall generator that minimize the difference between the values of the hydrometeorological attributes that correspond to the target location and those of the corresponding simulated values. The following objective function is proposed for minimization:

$$F_{obj,i} = \sqrt{\sum_{k=1}^{K} \left\{ \left[\frac{(P_{s,i}^k - P_{his}^k)}{P_{his}^k} - \frac{(\widehat{P_i^k} - P_{his}^k)}{P_{his}^k} \right] * 100\% \right\}^2} \quad (1)$$

where $i = 1, 2, \ldots n$ for n target locations in the exposure space. For the kth attribute of the hydrometeorological variable of interest (P^k), P_s^k represents the target level and $\widehat{P}^k$ represents the simulated level from the stochastic generator. Since different attributes are likely to consist of different magnitudes, the difference between a target level and the simulated level is represented as a percentage change relative to its long-term averaged historical value (P_{his}^k) to ensure consistent scales across different attributes. The optimization problem can be solved using a variety of optimization algorithms, such as genetic algorithms (Holland, 1975) or shuffled complex evolution (Duan et al., 1993).

The optimization procedure proceeds as follows: (1) values of the parameters of the stochastic generator are perturbed based on the searching strategy of the selected optimization algorithm; (2) the corresponding time series of the desired hydrometeorological variables are generated; (3) the values of the attributes of interest are calculated; and (4) the objective function is calculated in accordance with Eq. (1), which drives the algorithm's searching behavior. Steps (1)–(4) are then repeated until the specified stopping criterion has been met, such as the completion of a pre-specified number of iterations or until the objective function value is sufficiently small (Maier et al., 2014). It is important to note that as part of the inverse approach, the random seed of the stochastic generator should be held constant to ensure that the optimization proceeds as efficiently as possible, as discussed further in the following section.

2.3. Random sampling issues of stochastic generators and the implications on the inverse approach

The stochastic component of the rainfall generator can produce substantial variations in the simulation of rainfall attributes, even with a single parameter set. This randomness can affect the efficiency of the optimization process used in the inverse approach. Essentially, every iteration of the optimization involves a comparison among multiple parameter sets in terms of their ability to generate the target locations in the exposure space. However, as a result of stochastic generation, a single parameter set can lead to multiple potential locations on the exposure space (Fig. 2). This can then mislead the comparison and affect optimization efficiency, as changes made to parameters by the optimization algorithm in order to lead the search in one particular direction might actually have the opposite effect.

To illustrate this issue, consider a simple optimization problem to find the best-fit parameters of a Gaussian distribution with the

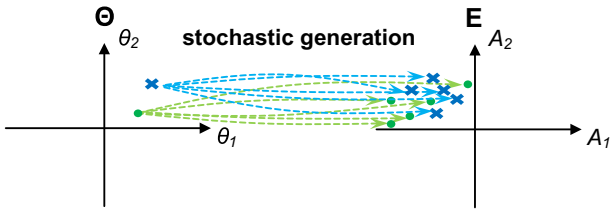


Fig. 2. Using a single parameter set in the parameter space Θ can yield multiple points (shown by the dashed arrows) in the exposure space **E** as result of the stochastic nature of the weather generator. Different colors represent simulations from two different parameter sets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

objective of getting a 'target' sample mean of $\bar{x} = 3$. Suppose that for one iteration the optimizer attempts to compare samples drawn from a simple Gaussian random generator ($X \sim N(\mu, \sigma)$) where the parameter μ is changed from 4.0 to 4.5, while holding σ at a constant value of 1. In the upper panel of Fig. 3, we show 50 random values drawn with each parameter set. For this set of random values, the sample mean from $X \sim N(4.0, 1)$ is 4.2 compared with the sample mean from $X \sim N(4.5, 1)$, which is 4.0. Therefore, the resulted sample mean from $N(4.0, 1)$ is actually further away from the target sample mean of $\bar{x} = 3$ compared with $N(4.5, 1)$, so that the search direction of the optimizer may be misled. Although this variance can be reduced with a larger sample size or a longer simulation period, it can never be completely eliminated.

To overcome this problem during optimization, the random number seed is held constant when producing the stochastic replicates. This ensures that any changes made to the parameters during the optimization process will lead the search in the desired direction. Using the same example, in the lower panel of Fig. 3 we show 50 samples drawn from both $X \sim N(4.0, 1)$ and $X \sim N(4.5, 1)$ with the same random seed used for each pair of samples, resulting in samples means of 3.9 and 4.4 respectively, thereby indicating that $N(4.0, 1)$ is better at producing a target sample mean of $\bar{x} = 3$. In this way, the stochastic generator is able to

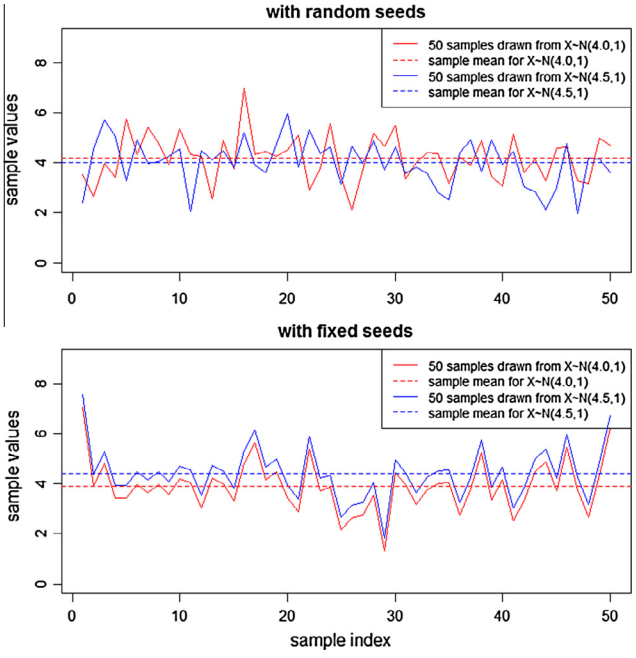


Fig. 3. Implications of using random seed (upper panel) and fixed seed (lower panel) for comparing two parameter sets: $X \sim N(4.0, 1)$ and $X \sim N(4.5, 1)$, for simulating a sample mean of 3 from a Gaussian distribution.

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

search through the correct directions on the parameter space to find parameters that converge toward the target rainfall attributes.

As discussed in Section 2.2, it is important to emphasize that the objective of the approach is to generate samples of hydrometeorological time series with specific levels for each attribute, rather than to identify the parameter sets that will produce those parameters in a population sense. Returning to the above example, the objective is to find a stochastic replicate with sample mean $\bar{x} = 3$, regardless of the values of the parameters (μ, σ) used to achieve this value. Consequently, once this goal has been met, the search can stop and the parameter values that were used to produce the stochastic time series corresponding to each target location can be discarded.

3. Case study

The proposed inverse approach is illustrated on rainfall data from a catchment in South Australia, using two stochastic rainfall generators: the Richardson model and the WGEN model. To provide a benchmark for the proposed inverse approach, its performance is also compared with that of a forward approach. The rainfall data, stochastic rainfall generators and the specific implementation of the forward and inverse approaches are described in this section.

3.1. Data

We used a rainfall time series from a gauge in the southern Mount Lofty Ranges close to Adelaide, South Australia, as a case study. The climate in this region is temperate, with most rainfall occurring in winter and spring (May to October). The mean annual rainfall was 913 mm for the study period from 1989 to 2004. The daily rainfall data over this period have been used to represent the baseline (historical) rainfall conditions.

We used four rainfall attributes as the dimensions of the exposure space, with definitions and baseline values provided in Table 1. These attributes represent key features of rainfall patterns; namely, the average daily rainfall (*PD*), the wet day frequency (*WD*), a measure of the rainfall intermittency (*CDD*) and a measure of extreme rainfall (*Pex99*). These attributes have been commonly used to assess the performance of rainfall generators (Chen and Brissette, 2014; Fowler et al., 2007; Hashmi et al., 2011; Kilsby et al., 2007; Semenov, 2007), and are also closely related to several of the indices used for the detection and attribution of climate change, as described by the Expert Team on Climate Change Detection and Indices (ETCCDI; Klein Tank et al., 2009).

For each rainfall attribute we defined a plausible range for sampling (which defined the range of each dimension within the exposure space) of between 50% and 150% of the corresponding historical value. These bounds were wider than would be expected from most climate change projections (e.g. CSIRO and Bureau of Meteorology, 2015; Stocker et al., 2013), to encompass a large range of climate projections (for example from climate models) in the exposure space.

Table 1
Definitions and annual average values of the four rainfall attributes that form the exposure space.

Rainfall attribute	Definition	Historical value
PD	Daily rainfall intensity averaged over all days	2.49 mm
WD	Average number of wet days	161 days
CDD	Average length of consecutive dry days	3.75 days
Pex99	The 99th percentile rainfall over wet days	37.6 mm

3.2. Stochastic rainfall generators

Two versions of the Richardson-type stochastic rainfall generator with different levels of complexity were used to generate the exposure space. We started with a simplified four-parameter model, which assumes uniform rainfall characteristics over the year. The advantage of this model is that it is possible to analytically derive the parameters that correspond to each target location in the exposure space. However, this simplified model uses a single value for each parameter throughout the year, and thus is unable to capture seasonal-scale variability. To highlight some practical issues with rainfall sampling, we then considered a more complex and widely used model—namely the WGEN (Richardson and Wright, 1984).

3.2.1. The four-parameter Richardson model

The simplified Richardson-type rainfall generator uses the following four parameters:

- The two parameters of the 1st order two-state Markov chain used for representing the transition probabilities of rainfall occurrence: p_{dd} (dry–dry probability) and p_{wd} (wet–dry probability), and
- The two parameters of a gamma distribution for representing the rainfall intensity on wet days: α (scale) and β (shape).

An approximate analytical expression relating two of the four output rainfall attributes (*PD* and *WD*) to the model parameters is given in Dubrovský et al. (2000) as:

$$PD = \frac{\alpha}{\beta} * \frac{WD}{365.25} \tag{2}$$

$$WD = 365.25 \times \frac{(1 - p_{dd})}{(1 - p_{dd} + p_{wd})} \tag{3}$$

These analytical expressions have been used when exploring the implications of random sampling issues on the inverse generation approach (Section 4.1.3).

3.2.2. The WGEN model

The WGEN model (Richardson and Wright, 1984) has the same structure as the simplified Richardson model, except that it uses a unique set of the four parameters for each month of the year, leading to a total of 48 parameters. This model has been used widely for climate impact studies, and is generally shown to capture most of the key features of daily rainfall series (Bastola et al., 2011; Katz, 2002; Kim et al., 2007).

Since the proposed inverse approach involves optimization of the parameter values, a search space with low dimension (i.e. consisting of a small number of parameters as decision variables) is desired. To reduce the size of the parameter space in the inverse approach, the number of decision variables to be considered was reduced from 48 to eight by fitting harmonic functions to describe the seasonal variations of each parameter (Prudhomme et al., 2013). The harmonic function takes the form of:

$$\beta(t) = \beta_0 + A\left(\cos\frac{2\pi}{T}(t - \Phi)\right) \tag{4}$$

where $\beta(t)$ represents one of the four parameters during month ($t = 1, \ldots, T$) with $T = 12$, β_0 represents the arithmetic mean of the parameter, A represents the amplitude and Φ corresponds to the month where the maximum occurs. It is worth mentioning that although parameter Φ can be varied as part of the optimization, the four-attribute exposure space in this case study was not designed to focus on shifts in rainfall seasonality (Section 3.1), so

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

that Φ was held constant at its historically optimal value. To determine the value of Φ, we obtained the monthly estimates of p_{dd}, p_{wd}, α and β (based on the method in Richardson (1981)) using the historical rainfall data, and fitted a harmonic function to each parameter (Fig. 4). The corresponding values of Φ were thus identified to be 2, 1, 8 and 1 (i.e., February, January, August and January) for the four parameters, respectively. As a result, the optimization was performed on the mean (β_0) and amplitude (A) of each of the four model parameters, leading to an eight dimensional search space.

3.3. Sampling approach

As illustrated in Fig. 1, application of the forward approach involves sampling the parameter space prior to using the stochastic model. Similarly, application of the inverse sampling approach involves the identification of target locations in the exposure space as the basis for optimization. One approach to sampling both the parameter space (in the forward approach) and exposure space (in the inverse approach) is to define a grid of evenly spaced points over the entire space. However, this can be inefficient, particularly for high-dimensional problems (for a large number of parameters/attributes in the exposure space in the forward/inverse approach). For example, if one wished to sample on a grid of width 10 for the parameter space of the four-parameter Richardson model, then it would be necessary to evaluate a total of $10^4 = 10{,}000$ separate parameter sets. This issue is particularly pertinent for the inverse approach, since optimization is required to find a parameter set that corresponds to each point in the exposure space. Therefore, to provide even coverage of the parameter or exposure space while keeping the sample sizes low, two structured sampling techniques have been employed, namely Latin hypercube sampling (LHS) and improved distributed hypercube sampling (IHS).

The objective of the analysis in this paper is to illustrate the inverse approach, by comparing its performance with the forward approach. Therefore for consistency, the objective of the sampling approach was to obtain 100 samples within the exposure space. For the forward approach, it is not known *a priori* whether a particular parameter set in the parameter space will yield a sample in our exposure space (i.e. within the plausible range of 50–150% for each rainfall attribute, as defined in Section 3.1), so that the number of samples that need to be drawn from the parameter space is not known. To determine the total number of samples in a computationally efficient manner, we used the Latin hypercube sampling (LHS) method, which allows starting with a small sample size and adding new samples while keeping the previously generated ones. The LHS method involves sampling M variables with a desired sample size N by dividing the range of each variable into N equally probable intervals. N samples are then drawn so that any interval for each variable is only sampled once (Stein, 1987). To add n new sample points, the existing design is re-divided into $(N + n)$ intervals; the N old samples are kept which occupy N intervals, and then n new samples are drawn to fill the remaining n intervals.

Unlike the LHS method, the IHS method (Manteufel, 2001; Beachkofski and Grandhi, 2002) requires that the number of samples be specified *a priori*, but ensures more even coverage of the sampling space. This latter feature is attractive when sparsely sampling potentially high-dimensional spaces, and is therefore recommended to determine the target locations in the exposure space for the inverse sampling approach. The IHS method is similar to the LHS method, with two additional objectives:

(1) The average minimum distance between sample points equals the optimal distance d_{opt}. That is, if the span of each output variable is normalized to 1 so that the entire sample space is a hypercube of volume 1, then each sample point should cover an equal hypervolume (with dimension of M) within the entire space. This gives the optimal distance between sample points, i.e. $d_{opt} = \sqrt[M]{\frac{1}{N_{Sample}}}$.

(2) The coefficient of variance (COV) of all minima between each pair of sample points is close to zero.

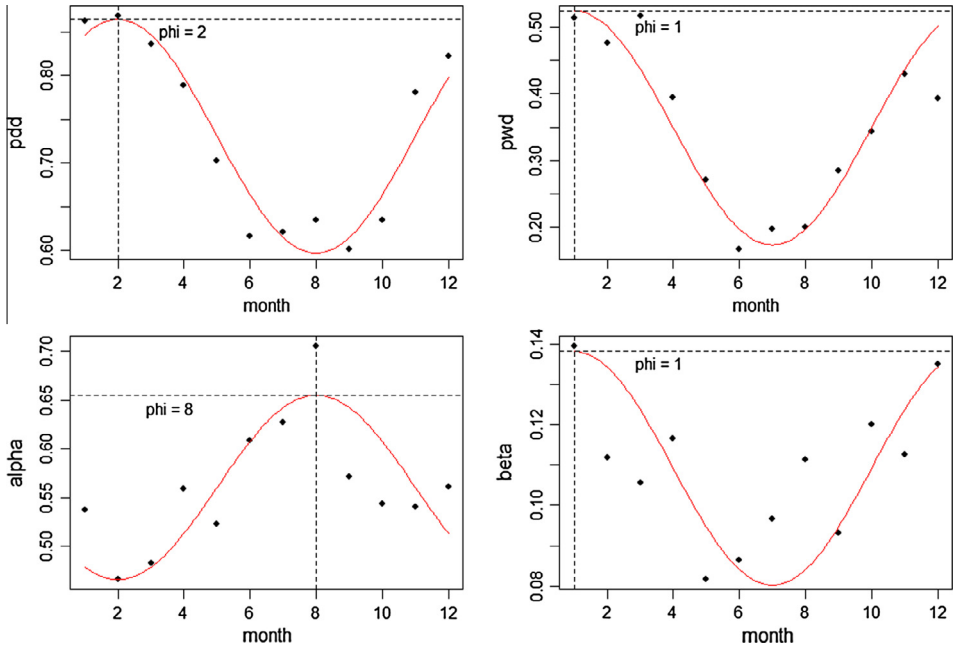


Fig. 4. The monthly variations of p_{dd}, p_{wd}, α and β for WGEN to determine the values of Φ, obtained from existing rainfall data, black dots represent individual parameter values estimated for each month, whereas red curves show the fitted harmonic functions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

3.4. Implementation of forward approach

As mentioned previously, the forward approach has been used to provide a benchmark against which the utility of the proposed inverse approach can be assessed. The approach involves the following steps:

(1) The parameter space is constructed by selecting appropriate ranges of the parameters for the stochastic generators;
(2) Parameter sets are drawn from the parameter space using a sampling strategy such as the LHS method described in Section 3.3;
(3) The sampled parameter sets are used to generate time series of the hydrometeorological variables of interest (in this case, rainfall); and
(4) The values of the attributes that define the exposure space are calculated for each of the generated hydrometeorological time series.

For the simple four-parameter Richardson model, the 'transition probability' parameters (p_{dd} and p_{wd}) both vary between 0 and 1. The two 'rainfall intensity' parameters (α and β) are for the gamma distribution and should be greater than 0 (Note that their values are mostly between 0 and 1 when calibrated to historical data; see Richardson and Wright, 1984). From a preliminary analysis for our case study, α and β values of 0.56 and 0.10 were obtained respectively, yielding rainfall time series with attributes that are close to the historical data. Therefore, although α and β do not physically have upper bounds and can take any value above 0, their ranges were set to be between 0 and 1 in the forward approach based on their historical values. The use of such a small range ensures that the parameter space surrounding the historical levels of the parameters is sufficiently sampled. For WGEN, the parameter ranges were defined in a similar way, so that the bounds of both the transition probability and rainfall intensity parameters were set to 0 and 1 for all months.

As mentioned previously, for both stochastic models, LHS was used to sample the parameter space. An initial Latin hypercube sample size of 100 parameter sets was used, and this was incremented until 100 rainfall time series were generated with attributes within the plausible bounds of the exposure space.

3.5. Implementation of proposed inverse approach

A general description of the inverse approach was provided in Section 2.2. The IHS method (Section 3.3) was used to determine the target locations for the optimization, which consist of 100 sets of combined levels of the four rainfall attributes that uniformly cover the exposure space.

For each target location, the best-fit parameter sets for both the four-parameter Richardson model and the WGEN model were identified using optimization. The shuffled complex evolution algorithm (Duan et al., 1993) was used as the optimization engine, due to its proven ability for solving complex optimization problems in hydrological studies (Gupta et al., 1999; Thyer et al., 1999; Wang et al., 2010). Based on the general formulation in Eq. (1), the objective function to be minimized for both stochastic models was:

$$F_{obj,i} = \sqrt{\begin{array}{l} \left\{\left[\frac{(PD_{s,i}-PD_{his})}{PD_{his}} - \frac{(\widehat{PD_i}-PD_{his})}{PD_{his}}\right]*100\right\}^2 + \left\{\left[\frac{(WD_{s,i}-WD_{his})}{WD_{his}} - \frac{(\widehat{WD_i}-WD_{his})}{WD_{his}}\right]*100\right\}^2 + \\ \left\{\left[\frac{(CDD_{s,i}-CDD_{his})}{CDD_{his}} - \frac{(\widehat{CDD_i}-CDD_{his})}{CDD_{his}}\right]*100\right\}^2 + \left\{\left[\frac{(Pex99_{s,i}-Pex99_{his})}{Pex99_{his}} - \frac{(\widehat{Pex99_i}-Pex99_{his})}{Pex99_{his}}\right]*100\right\}^2 \end{array}} \quad (5)$$

The constraints of the optimization consist of the plausible ranges of the parameters for both models. As mentioned in Section 3.4, the plausible range for the probability parameters (p_{dd}'s and p_{wd}'s) is between 0 and 1; for the intensity parameters (α's and β's), which do not have a physical upper limit, we defined the range to be between 0 to 10^4, which was wider than the range used for the forward approach (Section 3.4) to enable more extensive searching within the defined range.

For the WGEN model, since a harmonic function has been fitted to the monthly values of each of the probability and intensity parameters (Section 3.2.2), the actual decision variables for the optimization were the parameters of the harmonic functions (i.e. β_0 and A, which represent the mean and amplitude respectively, as in Eq. (4)). To ensure that the probability parameters were always within 0 and 1 while the intensity parameters were always within 0 and 10^4 during the optimization process, the values of the mean and amplitude for each of these parameters have been optimized sequentially. In the first step, the mean value of each parameter has been optimized with the amplitude kept as zero. Once the mean has been determined, a second optimization was conducted to estimate the amplitude. For example, if the mean of p_{dd} is found by the optimizer to be 0.3 in the first step, its amplitude must be constrained between 0 and 0.3 in the second step to avoid values of p_{dd} going beyond 0 and 1.

It should be noted that in determining the target locations, the IHS only checks the multi-dimensional uniformity of the overall distribution, without considering the physical interpretation for each individual target location. Therefore, it is important to ensure that each target location selected is physically plausible. For example, *PD* should always be less than *Pex99*, and *WD* should never exceed 365 days. For this study, these constraints were automatically satisfied because a relatively small plausible range of 50–150% was selected for each attribute. If the rainfall samples are required to show larger variances, it may be necessary to impose additional constraints in the optimization procedure to ensure the resultant samples remain physically plausible.

4. Results

4.1. The four-parameter Richardson model

4.1.1. Forward approach

The coverage of the exposure space obtained by applying the forward approach to the four-parameter Richardson model is shown in Fig. 5, which shows high variances in some rainfall attributes. In particular, the generated *PD*, *CDD* and *Pex99* can go up to 15,000%, 6000% and 80,000% of their corresponding historical values, respectively (Fig. 5a), which are well outside the bounds of the exposure space. The sampled *WD* has lower variance with values up to only 226% of the historical values (since a year contains a maximum of 365 or 366 days), however, these values are still above the upper limit of the exposure space of 150%. The high variance leads to low sampling efficiency – to obtain 100 sets of combined levels of rainfall attributes within our exposure space, a total of 7635 LHS samples of parameter sets had to be generated (i.e. 98.7% samples were unacceptable and discarded). All 7635 sets

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

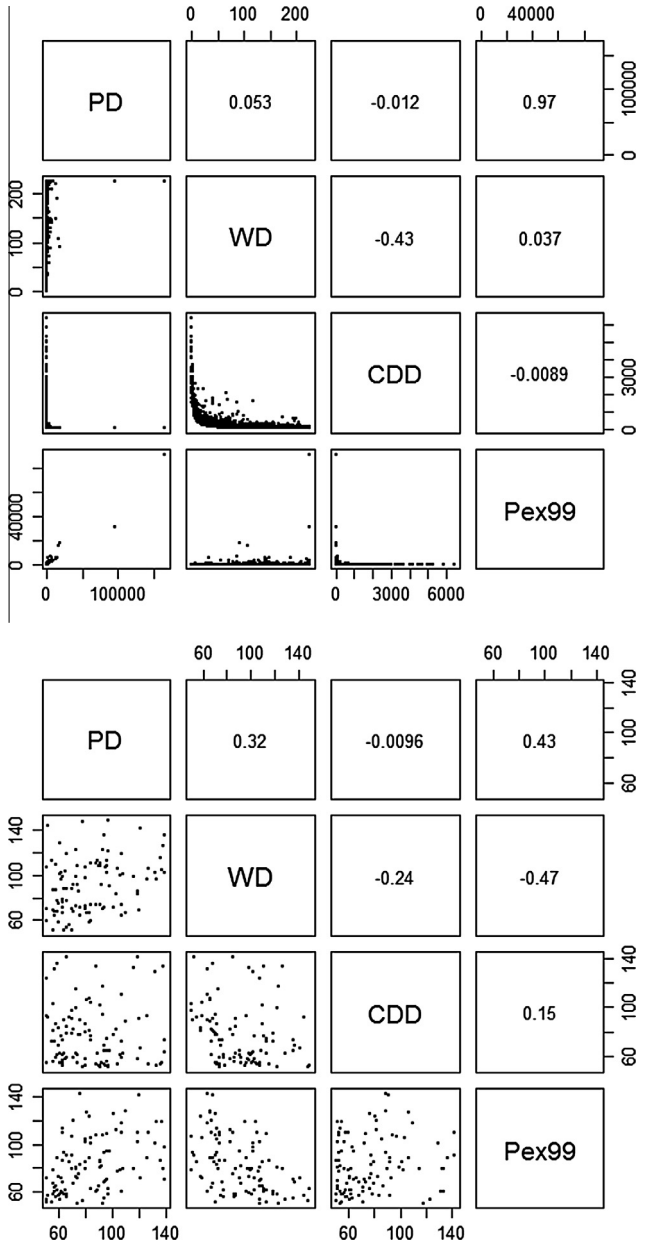


Fig. 5. The four rainfall attributes described as a percentage relative to the historical values produced by the four-parameter Richardson model, by drawing LHS samples from the parameter space, for (a) all the 7635 samples, and (b) only the 100 samples that fall in the plausible range of between 50% and 150% for each rainfall attribute. The upper-right triangle displays pairwise correlations.

of rainfall attributes are plotted in Fig. 5a, with the 100 plausible samples shown in Fig. 5b.

In addition to the issue of inefficient sampling, based on both a visual inspection of the coverage on the exposure space as well as consideration of the correlation coefficients, it is clear that the coverage of the exposure space is uneven (Fig. 5b). In particular, samples are clustered in small regions of the exposure space for each rainfall attribute, with other parts of the space receiving limited or no coverage. For example, the correlation between *PD* and *Pex99* is quite high, which results in better coverage over regions closer to the diagonal of the joint distribution of *PD* and *Pex99* than other regions.

The above problems with using the forward approach are most likely due to the non-linear translation from parameters to rainfall attributes through the stochastic generator, so that large variations in certain regions in parameters space result in small variations in exposure space and vice versa. This non-linearity will be further illustrated in the next section with the distribution of parameters identified through the inverse approach.

4.1.2. Inverse approach

Fig. 6a shows the 100 target locations of desired rainfall attributes that have been determined using the IHS approach (Section 3.3). As can be seen, the IHS approach generates samples that appear to be uniformly distributed across the exposure space, with even coverage across each attribute and low cross-correlations between attributes.

The final set of combined levels of attributes corresponding to each of the 100 stochastically generated rainfall time series obtained using the inverse approach is presented in Fig. 6b. As

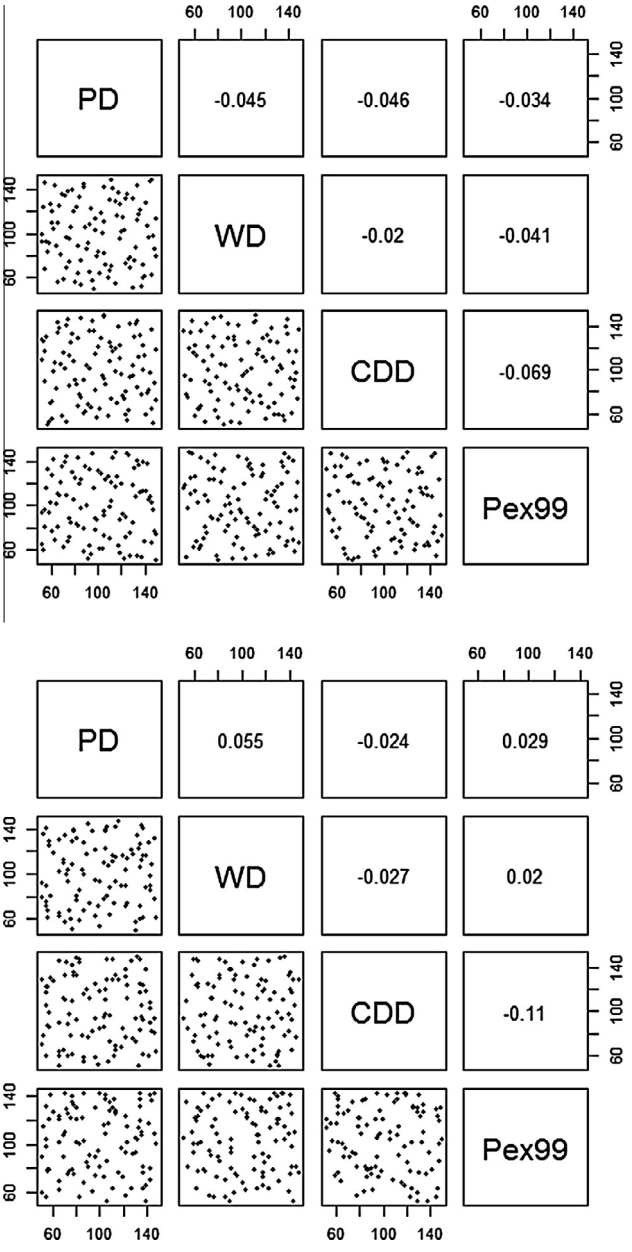


Fig. 6. The four rainfall attributes described as a percentage relative to the historical values produced by the four-parameter Richardson model for (a) 100 selected IHS sample locations and (b) 100 corresponding optimized locations. The upper-right triangle displays pairwise correlations.

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

can be seen, the optimization-based approach is effective in producing the desired levels of rainfall attributes (i.e. target locations), with all of the 100 samples falling within the bounds of the exposure space and with relatively even coverage of the exposure space (Fig. 6b). Therefore, the inverse approach delivers much better coverage of the exposure space than the forward approach (Fig. 5).

Fig. 7 shows the values of the 100 parameter sets for the four-parameter Richardson model, identified via application of the inverse approach, highlighting the non-linear mapping between parameter space and exposure space. This is best illustrated with the non-uniform distribution of the best-fit parameters, in contrast to the uniform distribution of the exposure space (Fig. 6). Furthermore, the parameters have considerably different ranges compared with the *a priori* [0,1] ranges that were specified for the forward approach. For example, the values of p_{dd} generally vary within a narrower range of 0.5–0.9, whereas values of α are as high as 10. Therefore, the ranges of [0,1] defined for the four parameters in the forward approach (as detailed in Section 3.4) can significantly limit the resultant coverage of the exposure space. This also reflects the high degree of non-linearity in the mapping between the parameter values and the exposure space, as a small change in the exposure space may result in a large shift in parameter space.

Interestingly, for the case study considered, although the inverse approach had the additional step of parameter optimization, the computational time required to obtain 100 samples was 32.6% shorter than for the forward approach. This is likely due to the large number of samples that were discarded in the forward approach.

4.1.3. Implications of random sampling on the inverse approach

In the above example, we fixed the random seed of the random number generator during the optimization process due to reasons discussed in Section 2.3. To illustrate the importance of this aspect of the optimization, we use the analytical expressions in Eqs. (2) and (3) to estimate the model parameters that will yield individual target locations from a grid consisting five evenly-spaced levels for each of *WD* and *PD* (50%, 75%, 100%, 125% and 150% of their historical values). These locations within the exposure space are given as green dots in Fig. 8. We then generated 100 stochastic replicates

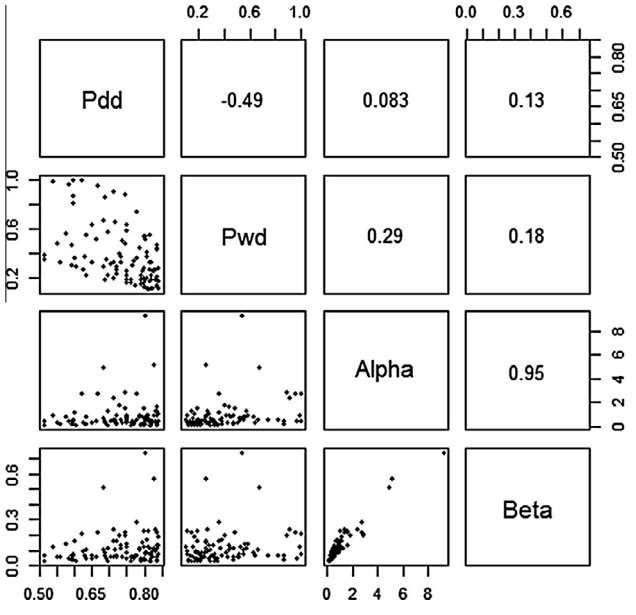


Fig. 7. Parameters corresponding to the exposure space in Fig. 6.

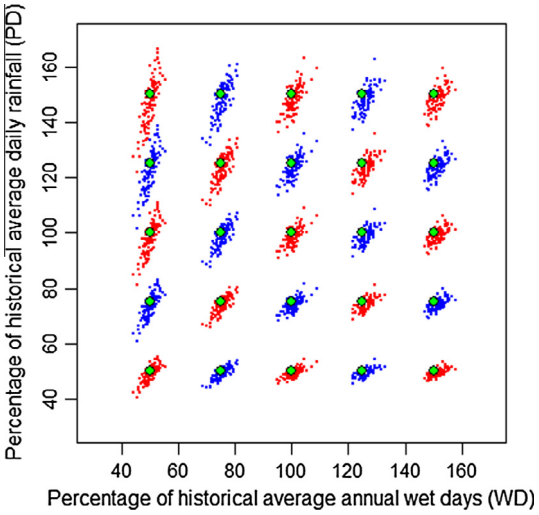


Fig. 8. Stochastic behavior of the four-parameter Richardson model in simulating *WD* and *PD*, with 100 replicates for the analytically-solved parameter set corresponding to each single target location. Two separate colors (i.e. red and blue) were used to differentiate between adjacent target locations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

from each of these parameter sets with different random seeds, which are shown as blue and red scatter about each of the target locations in Fig. 8.

The stochastic nature of the model is clear for all target locations. For each parameter set, the 100 replicates of *WD* vary up to ±10% around their target level, which is similar for all target levels of *WD* and *PD*. In contrast, the 100 replicates of *PD* are closer to the target level for smaller *PD* (e.g. up to ±10% around where the target level is 50%), while for larger *PD* target levels the spread among replicates increases substantially (e.g. up to ±40% around where the target level is 150%). Compared with the sampling resolution required in this study (shown in Fig. 6a), the variability in Fig. 8 is in fact much higher, which can adversely affect the capacity of the optimizer to find parameters that correspond to each target location, as discussed in Section 2.3.

4.2. The WGEN model

4.2.1. Forward approach

The coverage of the exposure space obtained by applying the forward approach to the WGEN model is shown in Fig. 9. Similar to the results for the four-parameter model (Section 4.1.1), the forward approach shows low efficiency: to obtain 100 sets of rainfall attributes within the range of the exposure space, 1453 LHS samples of WGEN parameter sets were required (Fig. 9a), which means that 93.1% of samples were discarded. With the 100 plausible sets in Fig. 9b, the coverage of the exposure space is poor, which is also evident through the high pairwise correlations (such as between *PD* and *Pex99* and between *WD* and *CDD*).

4.2.2. Inverse approach

To examine the performance of the inverse approach with WGEN, the 100 target locations which have been determined using the IHS approach (Section 3.3) are plotted in Fig. 10a. The final optimized set of attributes corresponding to each of the 100 stochastically generated rainfall time series is presented in Fig. 10b. The inverse approach is generally effective in evenly covering the exposure space and reproducing these target locations. In particular, this approach delivers much better coverage of the

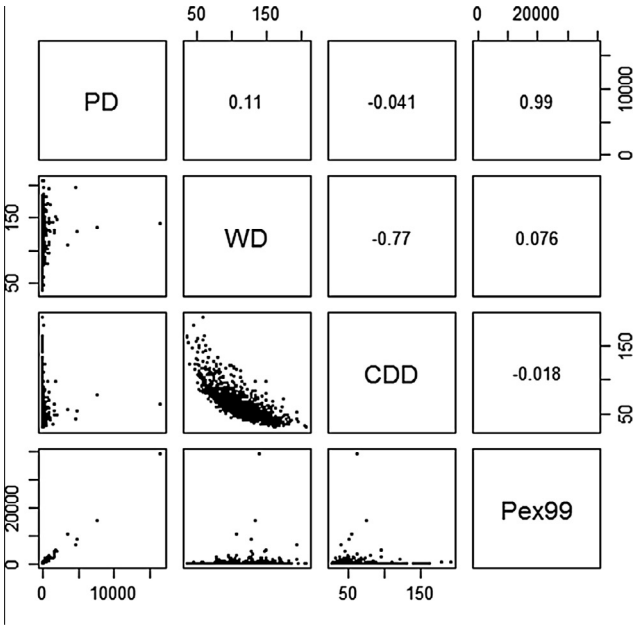


Fig. 9. The four rainfall attributes described as a percentage relative to the historical values produced by the WGEN model, by drawing LHS samples from the parameter space, for (a) all 1453 samples, and; (b) only the 100 samples that fall in the plausible range of between 50% and 150% for each rainfall attribute. The upper-right triangle display pairwise correlations.

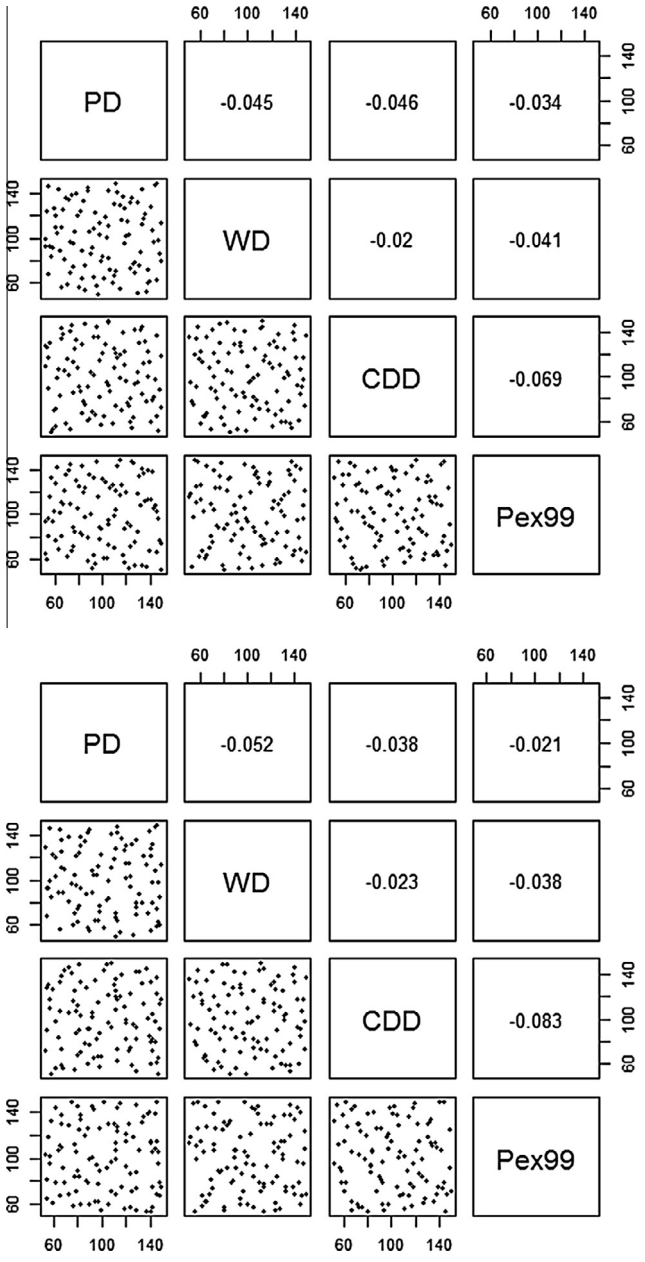


Fig. 10. The four rainfall attributes described as a percentage relative to the historical values produced by the WGEN model for (a) 100 selected IHS sample locations and (b) 100 corresponding optimized locations. The upper-right triangle displays pairwise correlations.

exposure space than the forward approach (Fig. 9) in the following aspects:

(1) All of the 100 samples are within the plausible output space defined in Table 1, suggesting effective control over the values of individual rainfall attributes; and
(2) The joint distribution of multiple rainfall attributes is much more uniform across the exposure space, and the pairwise correlations between different attributes are reduced.

As an alternative approach to assessing the relative uniformity of the sampling in the exposure space, the minimum distances between sample points in the exposure space are compared for both the forward (as orange dots in Fig. 11) and inverse approaches (as blue dots in Fig. 11). The results show that the inverse sampling approach produces a more uniform coverage, as the minima between sample points are closer to the optimal distance, $d_{opt} = 0.32$ (see Section 3.3). Furthermore, the coefficient of variance (*COV*) of these minimum distances is also much lower (i.e. 0.52 for the inverse approach compared with 2.90 for the forward approach).

It is worth noting that to obtain 100 sample points on the exposure space with the WGEN model, the overall execution time required for implementing the inverse approach is 73% longer than that for the forward approach. This is most likely due to the difficulty in solving optimization problems with a larger number of parameters, as a result of the larger search space that has to be explored. However, the inverse approach ensures uniform coverage of the exposure space with the desired resolution, which is the key objective for constructing the exposure space. In contrast,

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

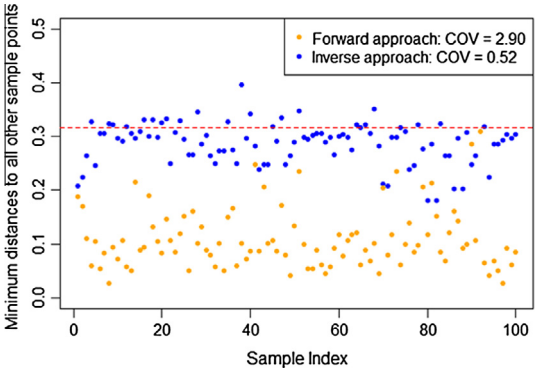


Fig. 11. Minimum distances between sample points obtained from the forward and inverse approaches with the WGEN model. The dashed line shows the optimal distance with 100 plausible samples (d_{opt} = 0.32).

the forward approach failed to obtain such coverage. Therefore, although associated with a higher computational expense, the proposed inverse approach is the only way of achieving the desired coverage of the exposure space.

5. Discussion

This study presented a framework for sampling various rainfall conditions to construct an exposure space for scenario-neutral climate impact assessments. Here, we discuss some practical considerations, as well as possible future adaptations of the framework.

5.1. Design of exposure space to represent more complex potential climate changes

The four rainfall attributes considered in the exposure space for this study (i.e. *PD*, *WD*, *CDD* and *Pex99*) are good descriptors of a range of changes of annual precipitation characteristics. However, there is also a range of other rainfall attributes that might be important when considering the impact of climate change, such as changes at seasonal or interannual timescales (e.g. Christensen et al., 2007; Johnson et al., 2011; Kwon et al., 2009). Furthermore, potential future variations in other climatic features, such as temperature, solar radiation and evapotranspiration, may also have a substantial impact on water resources (for examples see Chiew and McMahon, 2002a; Prudhomme and Williamson, 2013).

The inverse approach presented here is sufficiently flexible to cater to all attributes in the exposure space that are of interest (provided they can be generated with an appropriate stochastic generator), although this comes at the expense of additional computational cost. Considering the trade-off between the flexibility of producing different climate patterns and computational effort, it is important to identify key hydrometeorological variables of interest, as well as their attributes, based on an understanding of the behavior of the system being analyzed.

Depending on the specific hydrometeorological variables involved, the format of the objective function may require modification from Eq. (1), which was designed assuming multiplicative perturbations to attributes (e.g. changes expressed as a percentage of the historical value). For example, temperature changes are typically represented in an additive way (e.g. increases in temperature by degrees Celsius; for examples see Chiew and McMahon, 2002a; Kingston et al., 2009), and this would require an adjustment to the objective function in Eq. (1).

In this study, the boundary of the exposure space was set at 50–150% of the historical values of each attribute, which is sufficiently wide to incorporate a large number of possible changes in each of the rainfall attributes, while also using the same percentage changes across attributes to facilitate illustration. However, this framework can be easily adapted to incorporate tailored bounds for the exposure space, which should be carefully selected to suit the case study under consideration. In particular, if the bounds deviate too far from present conditions, a significant portion of samples will be unrealistic, even when extreme climate change impacts are considered. Conversely, if the bounds are too narrow, system response to some plausible climatic changes might not be considered (Whateley et al., 2014). Multiple sources of information could be considered in selecting these bounds, including GCM-based climate projections (e.g. Collins et al., 2013) of possible future climatic changes, and additional lines of evidence on possible changes to key variables, such as from long-term paleoclimatology reconstructions (e.g. Ault et al., 2014; Hansen and Sato, 2012; Ho et al., 2015). In addition, it is worth specifying an exposure space with bounds that are wider than the range suggested from all currently available sources of information, so that additional climate change projections can be included in the analysis as they are developed (Steinschneider and Brown, 2013).

Finally, when determining the target locations consisting of different hydrometeorological attributes on the exposure space, it is desirable to ensure the physical realism of each individual location so that corresponding time series can be obtained with the aid of stochastic weather generators. This requires not only ensuring that the target levels of individual attribute are realistic (such as the constraints for the levels of *WD*, as discussed in Section 3.5), but also maintaining physically plausible relationships among multiple attributes (for example, a target location cannot consist a *WD* value of 100 days with a *CDD* value of 300 days, because this combination means that the annual average wet day is 100 days while the annual average dry spell length is 300 days, which is physically unrealistic).

5.2. Stochastic generation of the exposure space

In this study, we used a sample size of 100 to represent different levels of changes for each individual attribute considered in the exposure space, with fixing the random seed across replicates to facilitate improved convergence during the optimization process. In this way, however, there is likely to be limited variability in between time series corresponding to different points on the exposure space, except for variations related to the target statistics.

This issue can be addressed in at least three ways:

1. Increase the sample size, and thus the coverage resolution in the exposure space. Increasing the exposure space resolution is likely to be particularly useful when the number of attributes increases, as this will lead to a corresponding increase in the dimensionality of the exposure space.
2. The length of each sample can be increased. Currently, the length is equal to the length of the historical data series (i.e. 15 years). However, it would be trivial to allow the simulation to run for longer periods of time to obtain greater stochastic variation. This will require the same number of optimized parameter sets, although because of the use of the same initial seed, there will still be significant similarities between individual samples.
3. The procedure can be repeated multiple times with different random seeds for each iteration, thereby generating multiple replicates. This would substantially increase the level of stochastic variability, although at the expense of additional computational time.

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

Table A.1
Definition of terminology used in this study.

Terminology	Definition
Exposure space	A large range of plausible hydrometeorological conditions that a system may confront in the future
Hydrometeorological attributes	The statistics calculated from the time-series of hydrometeorological variables (e.g. rainfall, temperature and potential evapotranspiration) which describe their average conditions, variations, extremes, as well as other features. For example, temperature is considered as a hydrometeorological variable, with potential attributes including annual average temperature, annual maximum temperature and seasonal average temperature, etc.
Low-dimensional exposure space	An exposure space that consists of a small number of hydrometeorological variables and their attributes
High-dimensional exposure space	An exposure space that consists of a large number of hydrometeorological variables and their attributes
Forward approach	The approach of perturbing the parameters of stochastic generators over some pre-defined 'parameter space' to yield an exposure space
Historical scaling	The approach to obtain a desired exposure by applying additive and/or multiplicative scaling factors directly to the historical time series of hydrometeorological variables (as in Kay et al., 2014; Prudhomme et al., 2010; Prudhomme and Williamson, 2013)
Forward-plus-scaling approach	The approach that combines the use of the forward and scaling techniques to generate an exposure space. In this study, this specifically refers to the approach illustrated in Steinschneider and Brown (2013). In this approach, the parameters of a stochastic generator (including Markov chain transition probabilities and the autoregressive model for low-frequency variability) are first perturbed to obtain stochastic sequences without changing the historical rainfall intensity. The wet-day rainfall intensity in the stochastic sequences is subsequently quantile-mapped to yield a set of daily rainfall series with desired levels of rainfall attributes
Inverse approach	The approach aims to generate an exposure space by first selecting the desired values of the attributes of interest in the exposure space, followed by an optimization step to identify the stochastic generator parameters that produce stochastic sequences with these attributes
Target level	A desired level for a single meteorological attribute that is included in the exposure space
Target location	A desired combination of the target levels for each of the hydrometeorological attributes that are included in the exposure space
Best-fit parameters for the stochastic generator	The parameter set for the stochastic generator that produces hydrometeorological time series with the levels of attributes corresponding to a particular target location

The appropriate sample size, length of each sample and number of replicates are likely to depend on the case study considered, as well as available computational resources.

5.3. Equi-finality in the optimization process

When using optimization to search for best-fit solutions, equi-finality issues are likely to arise (i.e. multiple solutions leading to the same results in the objective function, therefore they are not distinguishable during optimization, see Beven and Freer, 2001). This problem is further complicated within the proposed inverse approach, as the values of the objective function for optimization are based on results from stochastic models.

Equi-finality issues are likely to be greatest for low-dimensional exposure spaces, since higher-dimensional exposure spaces add constraints to the parameter space. For example, the chance that two contrasting combinations of P_{dd} and P_{wd} lead to the same combination of *WD* and *CDD* is much lower compared to that resulting in the same level of *WD* in isolation. Thus, increasing the number of attributes considered could have the additional advantage of reducing the number of feasible parameter sets to be considered.

It is worth noting that although equi-finality is likely to occur when the proposed inverse approach is implemented, the aim of the approach is to identify time series of outputs from the stochastic generator that result in desired values of the attributes included in the exposure space, and not to the identification of the resulting parameters in the stochastic generator, as discussed in Section 2.3. However, when different parameter sets lead to the same combination of attributes on the exposure space, the different time series of hydrometeorological variables which they produce can consist of varying degrees of physical realism. Therefore, checking the physical realism of the generated time series can potentially help to eliminate unrealistic parameter sets and thus resolve any equi-finality issues.

5.4. Computational efficiency and execution time

In our particular implementation of the proposed inverse approach, the computational time required to produce 100 evenly distributed samples is around eight hours using an Intel Xeon E3 (2.60 GHz, 8 Cores) processor with 32 GB RAM for both the four-parameter model and the WGEN, suggesting a relatively high computational demand. However, in general, the computational effort required is dependent on a number of practical specifications, including the operating system, programming language and algorithm used. As the key focus of this study is to introduce and illustrate a new method, the above-mentioned specifications have not been optimized for computational efficiency. We have used the R package *lhs* (Carnell, 2012) for obtaining samples over the exposure space with the IHS and LHS methods, together with the shuffled complex evolution algorithm embedded in the R package *hydromad* (Andrews and Guillaume, 2013) for solving the optimal parameter values for the stochastic generator. We have also developed our own R-scripts to execute the four-parameter Richardson and the WGEN models. It is expected that an improved integration of these different modeling components with the aid of other programming languages, such as Fortran or C++, will further increase computational efficiency.

6. Conclusions

Generation of exposure spaces for scenario-neutral climate impact assessments should consider a range of potential variations in relevant hydrometeorological variables, including shifts in the average, intermittency, variability and extremes. The 'exposure space' describes the range of conditions of interest that a system may be exposed to under a future climate, and this paper presents and demonstrates an inverse approach to stochastically generating hydrometeorological time series to uniformly cover this exposure space.

The utility of the proposed inverse approach is benchmarked against a forward approach for rainfall generation for a South Australian catchment, using two Richardson-type stochastic rainfall generators of varying complexity. The results highlight the highly non-linear translation from parameter space to exposure space, and thus the need for the proposed inverse approach in order to

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

obtain a relatively uniform coverage of the exposure space. For both models, the inverse approach demonstrates better control of the sampling range, with 100% of samples falling within the exposure space. Furthermore, the uniformity of the coverage of the four-dimensional exposure space is substantially improved.

Several potential adaptations for future implementations of the framework have been discussed, including: (1) the design of the exposure space to represent more complex changes in climate; (2) improvements to the way that stochastic samples in the exposure space are generated; (3) ways of reducing the effects of equifinality during the optimization process; and (4) methods for increasing computational efficiency. The flexibility of the proposed inverse approach enables consideration of all climate attributes of interest at the desired resolution, thereby expanding the applicability of the scenario-neutral approach to evaluating a water resource system's sensitivity to a wide range of plausible changes in climate.

Acknowledgements

The authors wish to thank Christel Prudhomme and an anonymous reviewer for their thoughtful comments on the manuscript. Seth Westra's time was supported by Australian Research Council Discovery project DP150100411.

Appendix A

See Table A.1.

References

Ajami, N.K., Duan, Q., Sorooshian, S., 2007. An integrated hydrologic Bayesian multimodel combination framework: confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resour. Res. 43 (1), W01403. http://dx.doi.org/10.1029/2005WR004745.

Alexander, L. et al., 2006. Global observed changes in daily climate extremes of temperature and precipitation. J. Geophys. Res.: Atmos. 111 (D5).

Allen, R., Pruitt, W., 1986. Rational use of the FAO Blaney–Criddle Formula. J. Irrig. Drain. Eng. 112 (2), 139–155. http://dx.doi.org/10.1061/(ASCE)0733-9437(1986) 112:2(139).

Andrews, F., Guillaume, J., 2013. Hydromad: Hydrological Model Assessment and Development. R package version 0.9-18.

Ault, T.R., Cole, J.E., Overpeck, J.T., Pederson, G.T., Meko, D.M., 2014. Assessing the risk of persistent drought using climate model simulations and paleoclimate data. J. Clim. 27 (20), 7529–7549. http://dx.doi.org/10.1175/JCLI-D-12-00282.1.

Bastola, S., Murphy, C., Sweeney, J., 2011. The sensitivity of fluvial flood risk in Irish catchments to the range of IPCC AR4 climate change scenarios. Sci. Total Environ. 409 (24), 5403–5415. http://dx.doi.org/10.1016/j.scitotenv.2011.08.042.

Beachkofski, B.K., Grandhi, R., 2002. Improved distributed hypercube sampling.

Beven, K., 1987. Towards the use of catchment geomorphology in flood frequency predictions. Earth Surf. Proc. Land. 12 (1), 69–82.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249 (1–4), 11–29. http://dx.doi.org/10.1016/S0022-1694(01)00421-8.

Boughton, W., Droop, O., 2003. Continuous simulation for design flood estimation—a review. Environ. Modell. Softw. 18 (4), 309–318. http://dx.doi.org/10.1016/S1364-8152(03)00004-5.

Brown, C., Ghile, Y., Laverty, M., Li, K., 2012. Decision scaling: linking bottom-up vulnerability analysis with climate projections in the water sector. Water Resour. Res. 48 (9), W09537. http://dx.doi.org/10.1029/2011WR011212.

Brown, C., Werick, W., Leger, W., Fay, D., 2011. A decision-analytic approach to managing climate risks: application to the upper great lakes1. JAWRA J. Am. Water Resour. Assoc. 47 (3), 524–534. http://dx.doi.org/10.1111/j.1752-1688.2011.00552.x.

Brown, C., Wilby, R.L., 2012. An alternate approach to assessing climate risks. Eos, Trans. Am. Geophys. Union 93 (41), 401–402.

Carnell, R., 2012. lhs: Latin Hypercube Samples. R package version 0.10.

Chen, J., Brissette, F.P., 2014. Comparison of five stochastic weather generators in simulating daily precipitation and temperature for the Loess Plateau of China. Int. J. Climatol. 34 (10), 3089–3105.

Chiew, F.H., McMahon, T.A., 2002a. Modelling the impacts of climate change on Australian streamflow. Hydrol. Process. 16 (6), 1235–1245.

Chiew, F.H.S., McMahon, T.A., 2002b. Modelling the impacts of climate change on Australian streamflow. Hydrol. Process. 16 (6), 1235–1245. http://dx.doi.org/10.1002/hyp.1059.

Christensen, J.H. et al., 2007. Regional climate projections. Climate Change, 2007: The Physical Science Basis. Contribution of Working group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. University Press, Cambridge, pp. 847–940, Chapter 11.

Clark, M.P., Slater, A.G., 2006. Probabilistic quantitative precipitation estimation in complex terrain. J. Hydrometeorol. 7 (1), 3–22. http://dx.doi.org/10.1175/JHM474.1.

Collins, M. et al., 2013. Long-term climate change: projections, commitments and irreversibility. In: Stocker, T.F. et al. (Eds.), Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 1029–1136. http://dx.doi.org/10.1017/CBO9781107415324.024.

CSIRO and Bureau of Meteorology, 2015. Climate Change in Australia Information for Australia's Natural Resource Management Regions: Technical Report. CSIRO and Bureau of Meteorology, Australia.

Dessai, S., Hulme, M., 2004. Does climate adaptation policy need probabilities? Clim. Policy 4 (2), 107–128.

Duan, Q., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76 (3), 501–521.

Dubrovský, M., Žalud, Z., Šťastná, M., 2000. Sensitivity of CERES-Maize yields to statistical structure of daily weather series. Clim. Change 46 (4), 447–472.

Fowler, H., Blenkinsop, S., Tebaldi, C., 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int. J. Climatol. 27 (12), 1547–1578.

Frost, A., 2004. Stochastic generation of point rainfall data at subdaily timescales: a comparison of DRIP and NSRP.

Furrer, E.M., Katz, R.W., 2008. Improving the simulation of extreme precipitation events by stochastic weather generators. Water Resour. Res. 44 (12).

Gao, L. et al., 2016. Robust global sensitivity analysis under deep uncertainty via scenario analysis. Environ. Modell. Softw. 76, 154–166. http://dx.doi.org/10.1016/j.envsoft.2015.11.001.

Gupta, H.V., Sorooshian, S., Yapo, P.O., 1999. Status of automatic calibration for hydrologic models: comparison with multilevel expert calibration. J. Hydrol. Eng.

Halton, J.H., 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Num. Math. 2 (1), 84–90. http://dx.doi.org/10.1007/BF01386213.

Hammersley, J.M., 1960. Monte Carlo methods for solving multivariable problems. Ann. N. Y. Acad. Sci. 86 (3), 844–874. http://dx.doi.org/10.1111/j.1749-6632.1960.tb42846.x.

Hansen, J.E., Sato, M., 2012. Paleoclimate implications for human-made climate change. In: Berger, A., Mesinger, F., Sijacki, D. (Eds.), Climate Change: Inferences from Paleoclimate and Regional Aspects. Springer Vienna, Vienna, pp. 21–47. http://dx.doi.org/10.1007/978-3-7091-0973-1_2.

Hashmi, M.Z., Shamseldin, A.Y., Melville, B.W., 2011. Comparison of SDSM and LARS-WG for simulation and downscaling of extreme precipitation events in a watershed. Stoch Environ. Res. Ris. Assess. 25 (4), 475–484.

Ho, M., Kiem, A.S., Verdon-Kidd, D.C., 2015. A paleoclimate rainfall reconstruction in the Murray-Darling Basin (MDB), Australia: 1. Evaluation of different paleoclimate archives, rainfall networks, and reconstruction techniques. Water Resour. Res. 51 (10), 8362–8379. http://dx.doi.org/10.1002/2015WR017058.

Holland, J.H., 1975. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. University of Michigan Press.

Johnson, F., Westra, S., Sharma, A., Pitman, A.J., 2011. An assessment of GCM skill in simulating persistence across multiple time scales. J. Clim. 24 (14), 3609–3623. http://dx.doi.org/10.1175/2011JCLI3732.1.

Jones, P., Harpham, C., Goodess, C., Kilsby, C., 2011. Perturbing a weather generator using change factors derived from regional climate model simulations. Nonlinear Process. Geophyss. 18 (4), 503–511.

Jones, P.G., Thornton, P.K., 1993. A rainfall generator for agricultural applications in the tropics. Agric. For. Meteorol. 63 (1), 1–19.

Jones, R., Page, C., 2001. Assessing the risk of climate change on the water resources of the Macquarie River Catchment. Integr. Models Nat. Resour. Manage. Across Discip., Issues Scales, 673–678.

Katz, R.W., 2002. Techniques for estimating uncertainty in climate change scenarios and impact studies. Clim. Res. 20 (2), 167–185.

Kay, A.L., Crooks, S.M., Reynard, N.S., 2014. Using response surfaces to estimate impacts of climate change on flood peaks: assessment of uncertainty. Hydrol. Process. 28 (20), 5273–5287. http://dx.doi.org/10.1002/hyp.10000.

Kay, A.L., Jones, R.G., 2012. Comparison of the use of alternative UKCP09 products for modelling the impacts of climate change on flood frequency. Clim. Change 114 (2), 211–230. http://dx.doi.org/10.1007/s10584-011-0395-z.

Kilsby, C. et al., 2007. A daily weather generator for use in climate change studies. Environ. Modell. Softw. 22 (12), 1705–1719.

Kim, B., Kim, H., Seoh, B., Kim, N., 2007. Impact of climate change on water resources in Yongdam Dam Basin, Korea. Stoch. Environ. Res. Ris. Assess. 21 (4), 355–373. http://dx.doi.org/10.1007/s00477-006-0070-5.

Kingston, D.G., Todd, M.C., Taylor, R.G., Thompson, J.R., Arnell, N.W., 2009. Uncertainty in the estimation of potential evapotranspiration under climate

change. Geophys. Res. Lett. 36 (20), L20403. http://dx.doi.org/10.1029/2009GL040267.
Klein Tank, A.M.G., Zwiers, F.W., Zhang, X., 2009. Guidelines on Analysis of Extremes in a Changing Climate in Support of Informed Decisions for Adaptation. World Meteorological Organisation.
Kwon, H.-H., Lall, U., Obeysekera, J., 2009. Simulation of daily rainfall scenarios with interannual and multidecadal climate cycles for South Florida. Stoch. Environ. Res. Ris. Assess. 23 (7), 879–896.
Langousis, A., Kaleris, V., 2014. Statistical framework to simulate daily rainfall series conditional on upper-air predictor variables. Water Resour. Res. 50 (5), 3907–3932. http://dx.doi.org/10.1002/2013WR014936.
Langousis, A., Mamalakis, A., Deidda, R., Marrocu, M., 2015. Assessing the relative effectiveness of statistical downscaling and distribution mapping in reproducing rainfall statistics based on climate model results. Water Resour. Res. http://dx.doi.org/10.1002/2015WR017556, n/a-n/a.
Maier, H. et al., 2014. Evolutionary algorithms and other metaheuristics in water resources: current status, research challenges and future directions. Environ. Modell. Softw. 62, 271–299.
Manteufel, R.D., 2001. Distributed hypercube sampling algorithm, Third AIAA non-deterministic approaches forum paper AIAA-2001-1673. In: 42nd structures, structural dynamics, and materials conference.
Meselhe, E., Habib, E., Oche, O., Gautam, S., 2009. Sensitivity of conceptual and physically based hydrologic models to temporal and spatial rainfall sampling. J. Hydrol. Eng. 14 (7), 711–720. http://dx.doi.org/10.1061/(ASCE)1084-0699(2009) 14:7(711).
Moody, P., Brown, C., 2013. Robustness indicators for evaluation under climate change: application to the upper Great Lakes. Water Resour. Res. 49 (6), 3576–3588.
Nazemi, A., Wheater, H.S., 2014. Assessing the vulnerability of water supply to changing streamflow conditions. Eos, Trans. Am. Geophys. Union 95 (32), 288-288.
Nazemi, A., Wheater, H.S., Chun, K.P., Elshorbagy, A., 2013. A stochastic reconstruction framework for analysis of water resource system vulnerability to climate-induced changes in river flow regime. Water Resour. Res. 49 (1), 291–305. http://dx.doi.org/10.1029/2012WR012755.
Paton, F., Maier, H., Dandy, G., 2013. Relative magnitudes of sources of uncertainty in assessing climate change impacts on water supply security for the southern Adelaide water supply system. Water Resour. Res. 49 (3), 1643–1667.
Poff, N.L. et al., 2015. Sustainable water management under future uncertainty with eco-engineering decision scaling. Nat. Clim. Change. http://dx.doi.org/10.1038/nclimate2765, advance online publication.
Prudhomme, C., Crooks, S., Kay, A., Reynard, N., 2013. Climate change and river flooding: part 1 classifying the sensitivity of British catchments. Clim. Change 119 (3–4), 933–948. http://dx.doi.org/10.1007/s10584-013-0748-x.
Prudhomme, C., Wilby, R.L., Crooks, S., Kay, A.L., Reynard, N.S., 2010. Scenario-neutral approach to climate change impact studies: application to flood risk. J. Hydrol. 390 (3–4), 198–209. http://dx.doi.org/10.1016/j.jhydrol.2010.06.043.
Prudhomme, C., Williamson, J., 2013. Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: a comparison of methods and associated uncertainty in future projections. Hydrol. Earth Syst. Sci. 17 (4), 1365–1377.
Racsko, P., Szeidl, L., Semenov, M., 1991. A serial approach to local stochastic weather models. Ecol. Model. 57 (1), 27–41.
Rajah, K. et al., 2014. Changes to the temporal distribution of daily precipitation. Geophys. Res. Lett. 41 (24), 8887–8894.
Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour. Res. 17 (1), 182–190.
Richardson, C.W., Wright, D.A., 1984. WGEN: A Model for Generating Daily Weather Variables. US Department of Agriculture, Agricultural Research Service Washington, DC, USA.
Semenov, M.A., 2007. Simulation of extreme weather events by a stochastic weather generator. Clim. Res. 35 (3), 203.
Semenov, M.A., Brooks, R.J., 1999. Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain. Clim. Res. 11 (2), 137–148.
Singh, R., Wagener, T., Crane, R., Mann, M.E., Ning, L., 2014. A vulnerability driven approach to identify adverse climate and land use change combinations for critical hydrologic indicator thresholds: application to a watershed in Pennsylvania, USA. Water Resour. Res. 50 (4), 3409–3427. http://dx.doi.org/10.1002/2013WR014988.
Stein, M., 1987. Large sample properties of simulations using Latin hypercube sampling. Technometrics 29 (2), 143–151.
Steinschneider, S., Brown, C., 2013. A semiparametric multivariate, multi-site weather generator with low-frequency variability for use in climate risk assessments. Water Resour. Res. http://dx.doi.org/10.1002/wrcr.20528, n/a-n/a.
Stocker, T.F. et al., 2013. Climate change 2013: The physical science basis. Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5). Cambridge Univ Press, New York.
Thyer, M., Kuczera, G., Bates, B.C., 1999. Probabilistic optimization for conceptual rainfall–runoff models: a comparison of the shuffled complex evolution and simulated annealing algorithms. Water Resour. Res. 35 (3), 767–773.
Vano, J.A., Kim, J.B., Rupp, D.E., Mote, P.W., 2015. Selecting climate change scenarios using impact-relevant sensitivities. Geophys. Res. Lett. 42 (13), 5516–5525.
Wang, Y.C., Yu, P.S., Yang, T.C., 2010. Comparison of genetic algorithms and shuffled complex evolution approach for calibrating distributed rainfall–runoff model. Hydrol. Process. 24 (8), 1015–1026.
Westra, S., Alexander, L.V., Zwiers, F.W., 2013. Global increasing trends in annual maximum daily precipitation. J. Clim. 26 (11), 3904–3918.
Westra, S. et al., 2014. Future changes to the intensity and frequency of short-duration extreme rainfall. Rev. Geophys. 52 (3), 522–555. http://dx.doi.org/10.1002/2014RG000464.
Whateley, S., Steinschneider, S., Brown, C., 2014. A climate change range-based method for estimating robustness for water resources supply. Water Resour. Res. 50 (11), 8944–8961.
Wilby, R.L., Dawson, C.W., Murphy, C., O'Connor, P., Hawkins, E., 2014. The Statistical DownScaling Model – Decision Centric (SDSM–DC): conceptual basis and applications. Clim. Res. 61 (3), 259–276.
Yates, D.N., Miller, K.A., Wilby, R.L., Kaatz, L., 2015. Decision-centric adaptation appraisal for water management across Colorado's Continental Divide. Clim. Risk Manage. 10, 35–50. http://dx.doi.org/10.1016/j.crm.2015.06.001.

Please cite this article in press as: Guo, D., et al. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. (2016), http://dx.doi.org/10.1016/j.jhydrol.2016.03.025

CHAPTER 5 Impact of Evapotranspiration Process Representation on Runoff Projections from Conceptual Rainfall-runoff Models (Paper 4)

Statement of Authorship

Title of Paper	Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models
Publication Status	☑ Published ☑ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Accepted for publication in Water Resources Research on 10/12/2016. Update: Guo, D., Westra, S. & Maier, H. R. 2017. Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models. Water Resources Research, 53. 435-454.

Principal Author

Name of Principal Author (Candidate)	Danlu Guo		
Contribution to the Paper	Designed experiments, conducted model simulations, analysed results, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	60%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	19/12/2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Seth Westra		
Contribution to the Paper	Suggested research gap and scope of study, helped to design experiments and interpret results, provided feedbacks on the manuscript and the responses to reviewers.		
Signature		Date	19/12/16

Name of Co-Author	Holger Maier		
Contribution to the Paper	Suggested scope of study, helped to evaluate the manuscript and the responses to reviewers.		
Signature		Date	20/12/16

Please cut and paste additional co-author panels here as required.

Abstract

Conceptual rainfall-runoff models are commonly used to estimate potential changes in runoff due to climate change. The development of these models has generally focused on reproducing runoff characteristics, with less scrutiny on other important processes such as the conversion from potential evapotranspiration (PET) to actual evapotranspiration (AET). This study uses three conceptual rainfall-runoff models (GR4J, AWBM and IHACRES_CMD) and five catchments in climatologically different regions of Australia to explore the role of ET process representation on the sensitivity of runoff to plausible future changes in PET. The changes in PET were simulated using the Penman-Monteith model and by perturbing each of the driving variables (temperature, solar radiation, humidity and wind) separately. Surprisingly, the results showed the potential of a more than seven-fold difference in runoff sensitivity per unit change in annual average PET, depending on both the rainfall-runoff model and the climate variable used to perturb PET. These differences were largely due to different ways used to convert PET to AET in the conceptual rainfall-runoff models, with particular dependencies on the daily wet/dry status, as well as the seasonal variations in store levels. By comparing the temporal patterns in simulated AET with eddy-covariance-based observations at two of the study locations, we highlighted some unrealistic behaviour in the simulated AET from AWBM. Such process-based evaluations are useful for scrutinizing the representation of physical processes in alternative conceptual rainfall-runoff models, which can be particularly useful for selecting models for projecting runoff under a changing climate.

5.1 Introduction

Climate change is expected to have significant implications on catchment-scale water resources, potentially affecting water security for municipal, agricultural, and industrial applications, environmental water quantity and quality, and flood hazard (CSIRO and Bureau of Meteorology, 2015; Hauser et al., 2009; IPCC, 2014; Turral et al., 2011). To assess these implications, rainfall-runoff models are commonly used to translate projected changes in atmospheric variables derived from large-scale general circulation models into regional or local runoff (for examples see Akhtar et al., 2008; Chiew et al., 2009b). This information can then be used to assess catchment yield (e.g. Haque et al., 2015), water quality (e.g. Crossman et al., 2013), water supply security (e.g. Christensen et al., 2004; Paton et al., 2013; 2014) and flood risk (e.g. Kay and Jones, 2012), amongst other variables.

Although various rainfall-runoff model classes have been developed with different levels of physical realism (Beven, 2011), most climate impact studies are based on the outputs of one or several calibrated conceptual rainfall-runoff models (Chiew et al., 2010; Islam et al., 2014; Najafi and Moradkhani, 2015; Vaze and Teng, 2011). The benefit of using conceptual rainfall-runoff models is that they are parsimonious and tend to perform well when calibrated to historical runoff data (Perrin et al., 2003; Wagener et al., 2003), particularly when compared to more complex physically-based and spatially distributed models, for which parameter estimation can be extremely challenging (Beven, 2001b; Butts et al., 2004; Goderniaux et al., 2009). However, since the performance of these conceptual rainfall-runoff models is largely assessed on runoff, in many cases the underlying physical processes such as evapotranspiration fluxes, runoff generation mechanisms and so on can be poorly represented (Li et al., 2015; Seibert and McDonnell, 2002; Seibert et al., 2003), with parameters being ‘falsely adjusted’ to maximize runoff performance without considering the realism of the internal physical processes (Andréassian et al., 2004; Beven, 2001a; Clark et al., 2016; Ewen et al., 2006; Minville et al., 2014) .

The surprisingly limited impact of physical process realism on the performance of conceptual rainfall-runoff models in simulating historical streamflow has been observed in multiple studies (Brouyère et al., 2004; Hartmann and Bárdossy, 2005; Kirchner, 2006; Vaze et al., 2010). This result has also been found in the specific context of evapotranspiration (ET) process representation, and reasonable performance of conceptual models has been obtained with the use of simplified methods for estimating potential ET (PET) that do not realistically consider the dynamics in all its driving climate variables (Oudin et al., 2005a), or do not sufficiently represent the spatial and temporal variations in ET (Andréassian et al., 2004; Chapman, 2003). Furthermore, reasonable conceptual rainfall-runoff model performance has been observed even when the ET input contains systematic and/or random errors (Oudin et al., 2006). A potential reason for these findings is that the calibration process of conceptual models can 'compensate' for reasonable levels of differences in ET estimates without significant impact on model performance in simulating runoff (Andréassian et al., 2004; McMahon et al., 2015). However, although such calibrated conceptual models may perform well when applied under conditions similar to the period used for calibration, the extent to which they can be applied to climatologically different conditions is less clear (Coron et al., 2012; Klemeš, 1986; Wagener et al., 2003; Westra et al., 2014b), as in this context it becomes much more critical to ensure that the model 'gets the right answers for the right reasons' (Kirchner, 2006). This is highlighted in a number of recent studies that found that although conceptual models with different structures can perform similarly when simulating historical runoff, they can produce very different estimates under a changing climate (Bae et al., 2011; Bastola et al., 2011a; Jiang et al., 2007; Mandoza et al., 2016; Velázquez et al., 2013).

Although the above studies suggest potential limitations in the performance of conceptual rainfall-runoff models under future climate, the specific role of ET representation within these models on runoff projections has not been widely investigated. Changes to ET-related processes are likely to be particularly important to overall catchment response, as actual ET (AET) volumes exceed runoff volumes for the majority of catchments worldwide

(Dingman, 2015). Furthermore, the potential changes in the climatic variables that can influence ET (e.g. temperature, solar radiation, humidity and wind) (IPCC, 2014) suggest that changes to ET-related processes could be extremely complex, with changes not only at annual timescales but also in terms of the daily and seasonal distributions of ET (as illustrated in Gong et al., 2006; Huo et al., 2013; Vicente-Serrano et al., 2014). However, the extent to which these potentially complex changes to ET can interact with the structure of conceptual rainfall-runoff models remains unknown, with these changes usually represented using seasonal or annual ET change factors in many climate impact studies (e.g. Chiew et al., 2009b; New et al., 2007; Teng et al., 2012).

The focus of this study is therefore to assess the impact of alternative ET process representations within conceptual rainfall-runoff models on runoff projections under perturbed climate conditions. This is achieved by focusing specifically on the following research questions:

1. To what extent can ET process representations within conceptual rainfall-runoff models impact runoff projections under plausible changes in the climate drivers of ET?
2. How do alternative conceptual rainfall-runoff model structures interact with the potentially complex changes to PET to produce projections of runoff and AET?
3. Is it possible to use AET observations to constrain which ET process representations within conceptual rainfall-runoff models are more realistic?

The above research questions were addressed with five case study locations from distinct climatic regions within Australia. Three lumped conceptual rainfall-runoff models—GR4J, AWBM and IHACRES_CMD—were used to represent contrasting ET process representations, and thus to

illustrate how potential changes in PET can propagate differently to runoff projections. PET was required as an input to all three rainfall-runoff models, and was estimated using the Penman-Monteith method. This method is considered to be the most comprehensive physically based PET model and is thus widely used for rainfall-runoff modelling and climate impact assessments (e.g. Arnell, 2004; Gosling et al., 2011; Kay et al., 2009). Use of the Penman-Monteith method enables an assessment of the effect of perturbing each of the driving climatic variables of PET (i.e. temperature, solar radiation, wind and humidity), and can provide information not only on annual average changes, but also on sub-annual (e.g. daily and seasonal) variations in PET sensitivity.

The remainder of this paper is structured as follows. The data used to address the three research questions are provided in the next section, together with an overview of the Penman-Monteith model and the three conceptual rainfall-runoff models. This is followed by a description of the methods employed to address each research question in Section 5.3. The results and their implications are discussed in Section 5.4, and a summary and conclusions are given in Section 5.5.

5.2 Data and models

5.2.1 Data

Figure 5-1 shows the names and locations of the five case study catchments, which are situated in climatologically different regions in Australia, as defined in the Australian Köppen climate classifications of Stern et al. (2000). As mentioned in the Introduction, three lumped conceptual rainfall-runoff models were used in this study, which required historical data for calibration. These data are, namely, catchment-average rainfall, PET and runoff at each location. The PET input data to these models were estimated using the Penman-Monteith model, which required data on temperature, relative humidity, solar radiation and wind speed.

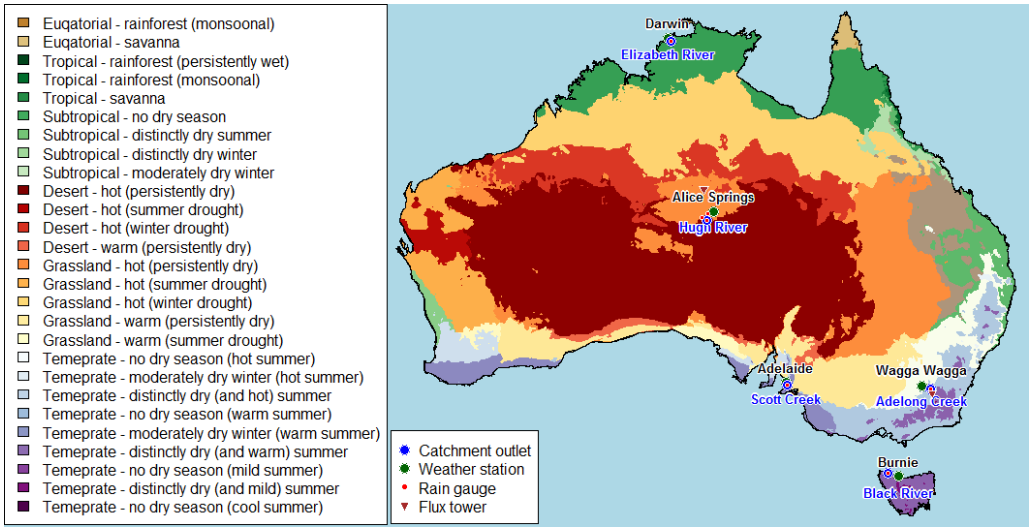


Figure 5-1: Locations for rain gauges, catchment outlets and weather stations from which data were obtained for calibration of rainfall-runoff models at the five case study catchments. The locations of two flux towers to obtain data for model evaluation at Alice Springs and Wagga Wagga are also shown. Coloring relates to Köppen climate classifications from *Stern et al.* (2000).

The source of each data variable is summarized below, with further details given in Table 5A.1 in Appendix 5A:

1. **Catchment runoff (ML/day)**: Daily runoff data were obtained from the gauging stations at the outlet of each catchment.
2. **Catchment-average rainfall (mm/day)**: In order to represent the catchment-average rainfall, daily rainfall data were obtained from a rain gauge within each of the four smaller catchments (Scott Creek, Black River, Elizabeth River and Adelong Creek). For the largest catchment (Hugh River), data from three rainfall gauges within the catchment were spatially averaged using the Thiessen polygon method.
3. **Daily maximum and minimum temperature (T_{max} *and* T_{min} in °C), maximum and minimum relative humidity (RH_{max} *and* RH_{min} in %) and wind speed (u_z in m/s):** Due to the limited availability of high-quality climate observations, data for each of these variables

were obtained from a weather station outside but nearby each catchment.

4. **Daily solar radiation (R_s in MJ/(m^2.day)):** Daily solar radiation was calculated from daily sunshine hour data (n in hrs) obtained from each weather station, using the Ångström-Prescott equation (McMahon et al., 2013) with constants for each location provided in Chiew and McMahon (1991).

Data for each variable were obtained for a consistent period from 1 January 1995 to 31 December 2003 at each location, with key statistics presented in Table 5-1. As can be seen, there are large differences in the average values of each variable, highlighting the potentially significant differences in dominant physical processes for ET and runoff production across these catchments. A quantity particularly relevant to ET processes is the long-term averaged ratio of PET to precipitation (PET/P), which describes whether a catchment is water-limited (PET/P >1) or energy-limited (PET/P < 1). The range of PET/P values indicates substantial variations in the water availability conditions at the five study sites.

Table 5-1: Average climate conditions at the five case study locations between 01/01/1995 and 31/12/2003.

Weather station	T_{max} (°C)	T_{min} (°C)	RH_{max} (%)	RH_{min} (%)	Rs (MJ/(m^2.day))	u_z (m/s)	Annual PET (mm)	Catchment (and area)	Annual P (mm)	Annual Q (mm)	PET/P
Adelaide	21.4	12.5	80.2	42.2	16.9	3.16	1372	Scott Creek (29 km^2)	892	133	1.54
Burnie	15.7	12.7	77.1	65.3	14.3	4.05	958	Black River (318.5 km^2)	1182	550	0.81
Darwin	31.2	23.7	88.5	50.1	20.3	3.39	1864	Elizabeth River (95.6 km^2)	1979	777	0.94
Alice Springs	28.5	13.8	65.6	23.5	20.8	2.35	1822	Hugh River (3324 km^2)	344	56.2	5.29
Wagga Wagga	21.6	9.99	83.2	40.3	17.5	3.29	1436	Adelong Creek (146.1 km^2)	799	195	1.80

***Note: T_{max} = maximum temperature; T_{min} = minimum temperature, R_s = incoming solar radiation, RH_{max} = maximum relative humidity, RH_{min} = minimum relative humidity, u_z = wind**

speed, PET = potential evapotranspiration calculated using the Penman-Monteith model, P = catchment-average precipitation, Q = average runoff at catchment outlet.

To determine the relative realism of different ET process representations, actual measurements of AET were used to benchmark the simulated AET from the three rainfall-runoff models. A challenge, however, is the limited availability of AET data for Australia. Lysimeter measurements are currently considered as the most accurate source of AET observations (Seneviratne et al., 2012; Wang and Dickinson, 2012); however, these data are most commonly available in irrigated regions (e.g. Bethune et al.; Northey et al., 2006; Zhang et al., 1999), with measurements from natural catchments being very scarce. The MODIS Land Product (http://daac.ornl.gov/MODIS/modis.shtml) also provides estimates of AET based on remotely sensed observations (Cleugh et al., 2007; Mu and Running, 2011) with a much wider spatial coverage; however, the AET data are only available on an 8-day interval.

The above constraints led us to select eddy covariance (EC)-based AET measurements, which are available through the OzFlux (TERN, 2012) network at 30 active sites across Australia (see: http://www.ozflux.org.au/). For our purpose of evaluating alternative ET representations, the AET observations should be obtained from nearby our case study catchments, as well as spanning a sufficient continuous time period. As a result of these constraints, only two OzFlux sites were selected, namely Alice Springs Mulga, which is close to the Hugh River catchment, and Tumbarumba, which is close to the Adelong Creek catchment (Cleverly, 2011; vanGorsel, 2013). Based on the EC data availability, the comparison between modelled and observed AET used a study period of between 3 September 2010 and 31 July 2014 for both catchments. The corresponding climatic and rainfall data within this period were also obtained from both flux sites to ensure internal consistency in the rainfall-runoff models.

5.2.2 Models

5.2.2.1 Penman-Monteith PET model

The Penman-Monteith model was adopted throughout this study, which combines the energy balance and mass transfer components of ET. To minimize the potential confounding effects of differences in vegetated surface, the evaporative surface was assumed to be reference crop for all study sites, for which the FAO-56 version of the Penman-Monteith model (Allen et al., 1998) was used. The FAO-56 Penman-Monteith model was implemented using the R package 'Evapotranspiration' (http://cran.r-project.org/web/packages/Evapotranspiration/index.html) (Guo et al., 2016b).

5.4.2.2 Three conceptual rainfall-runoff models

Contrasting ET process representations were considered by using three rainfall-runoff models (GR4J, AWBM and IHACRES_CMD), which all run on a daily time-step. The three models are all lumped conceptual models, although with very different soil moisture accounting (SMA) routines, which can affect the partitioning of precipitation into AET and runoff. To isolate the effect of different conceptual ET representations in the three models, the groundwater exchange was set to zero for each model. Furthermore, the routing components for all models were constrained to the GR4J routing model, so that differences between models can be attributed specifically to the SMA routines.

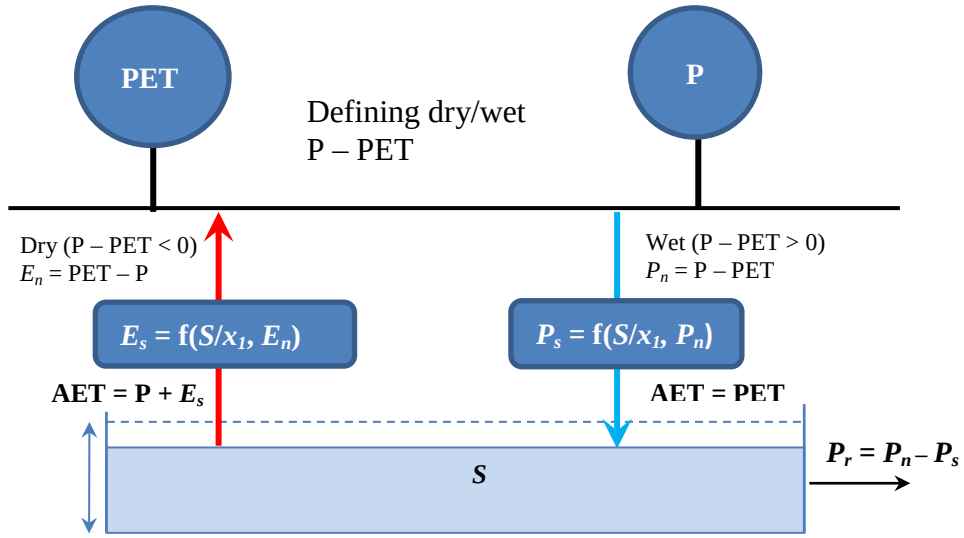
The first model, GR4J (Perrin et al., 2003), is a conceptual hydrologic model based on a SMA production store and an excess-rainfall routing store. The model requires two input variables, PET and precipitation (P). Interception is treated as a store with zero capacity, so that each day can be either 'wet' (P > PET), which produces a net precipitation P_n = P-PET, or 'dry' (P < PET) with a net evapotranspiration E_n = PET-P. For wet days a portion of the net precipitation P_n fills the production store. This portion is termed P_s and is a function of both P_n and the store level S relative to its capacity defined by parameter x_1. Runoff is produced from the total water available for routing, P_r, which is generated by the remaining net precipitation (P_n-P_s). The conversion of PET to AET from the production store is illustrated

in Figure 5-2a, which relies on the wet/dry status of a day. For wet days, AET is set to be equal to PET. For dry days, AET is modelled as the sum of the precipitation and the evapotranspiration from store E_s, which is modelled as a fraction of the net evapotranspiration E_n depending on the water level in the production store S relative to $x_{1.}$ Consequently, AET is always less than PET on dry days.

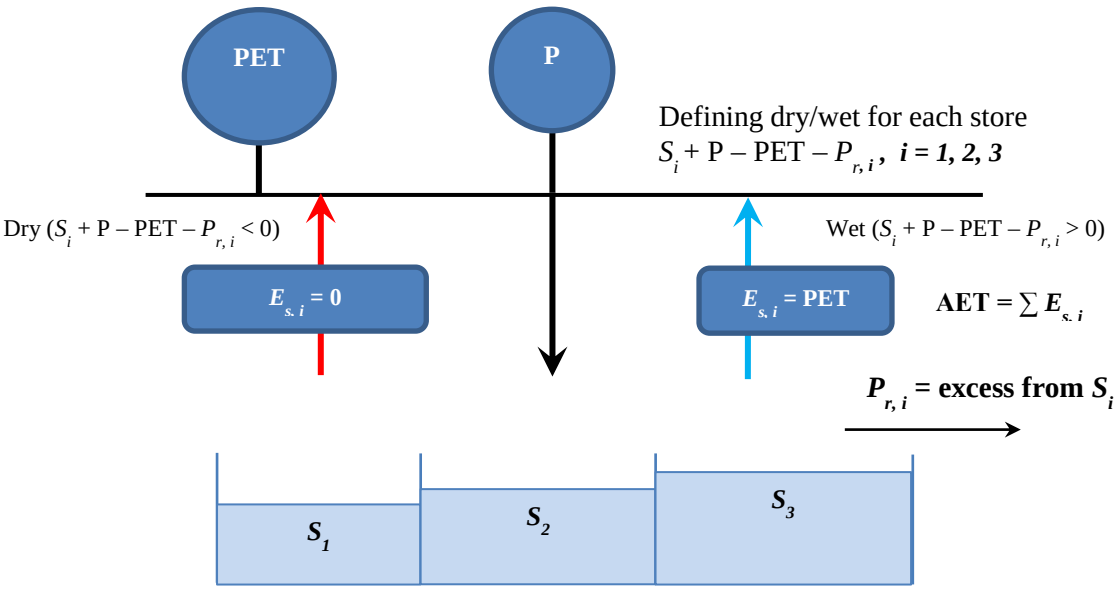
The schematic of AWBM is shown in Figure 5-2b. In contrast to GR4J, the production store in AWBM (Boughton, 2004) is represented by three individual stores (termed S_1, S_2 and S_3) with different capacities, and each store occupies a fixed fraction of the total catchment area, with the default values of 0.134, 0.433 and 0.433 used in this study. For each day precipitation P is added directly to each store without explicit consideration of interception. For each non-empty store, available runoff P_r is produced from any excess when the store is full, while evapotranspiration from the store, E_s, always equals PET until the store is completely empty (Boughton, 2009). Therefore, the overall AET, as the sum of E_s from all three stores, equals different proportions of PET, depending on which stores contain water. Specifically, the AET as a proportion of PET can be one of: 0 (when all three stores are empty), 0.134 (when both of S_2 and S_3 are empty), 0.433 (when both of S_1 and S_2 or both of S_1 and S_3 are empty), 0.567 (when only one of S_2 or S_3 is empty), 0.866 (when only S_1 is empty), or 1 (when all stores are non-empty). The level for each store S_i ($i = 1, 2, 3$), after accounting for P, PET and P_r from the specific store, equals $S_i + P - PET - P_{r,i}$. Therefore, on days with $P - PET > 0$, all stores are filled up (i.e. with increase in store levels), resulting in the overall AET = PET; when $P - PET < 0$, AET is always less than PET as determined by number of empty stores.

In IHACRES_CMD (Croke and Jakeman, 2004), the level of the store is represented by the catchment moisture deficit (CMD), which is the difference between the current level and the saturation level of the store. The schematic of the model is shown in Figure 5-2c. For each day, all precipitation P directly fills the store without explicit consideration of interception. The total water available for routing, P_r, is a proportion of P, which decreases with increasing

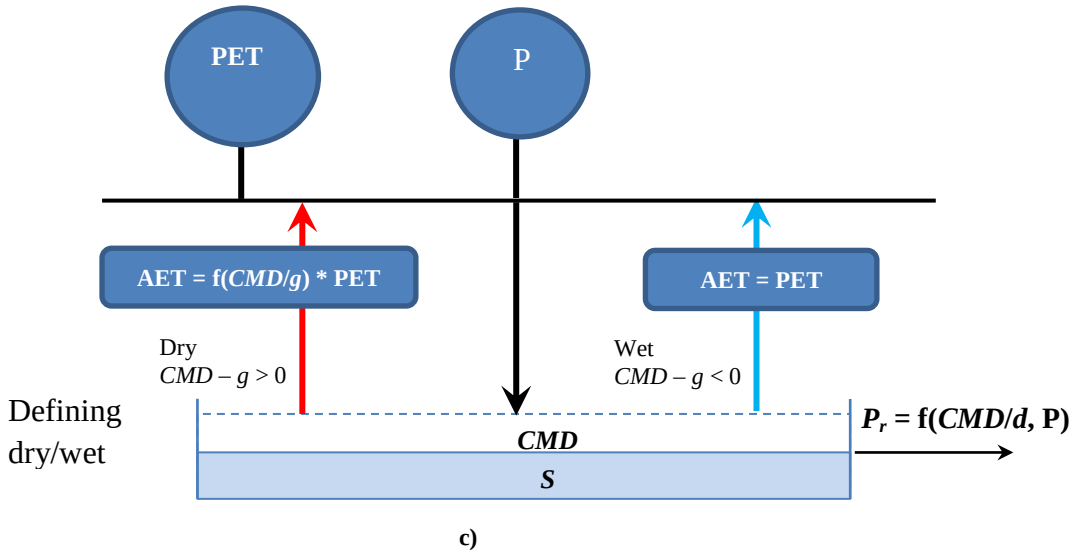
ratio of *CMD* relative to a flow-production threshold parameter *d*. AET is taken directly from the store (i.e. E_s = AET), which is estimated as a fraction of PET depending on whether *CMD* exceeds a threshold parameter *g*. When *CMD* > *g*, the fraction of AET to PET is always less than one, which decreases with increasing ratio of *CMD* to *g*, and approaches an asymptote of 0. In contrast, when *CMD* < *g*, AET = PET. Therefore, the conversion from PET to AET is independent of direct input of P, other than through its effect on changing the store levels.



a)



b)



c)

Figure 5-2: SMA routines in: a) GR4J; b) AWBM; and c) IHACRES_CMD (adapted from Perrin et al. (2005), Boughton (2004), Croke and Jakeman (2004) respectively). The red arrows represent dry days and the blue arrows represent wet days in each model.

The above rainfall-runoff models therefore provide examples of three contrasting methods of converting PET input data to AET: GR4J explicitly considers both the impact from store levels and the wet/dry status of a day, defined by individual rainfall events; AWBM simulates AET mostly as a function of rainfall which determines the wet/dry status, while the impact of instantaneous store levels is largely eliminated by the use of step functions to relate AET with the emptiness of each store; finally, IHACRES_CMD simulates AET as a function of only the store level and thus is relatively independent of rainfall for that day.

To simulate runoff, the PET data estimated for each case study were used as input to all three SMA models linked with the GR4J-routing model (implemented using the R package 'hydromad', available at: http://hydromad.catchment.org/ (Andrews and Guillaume, 2013)). For each case study, the three models were calibrated to the first six years of historical runoff data (from 1 January 1995 to 31 December 2000) at a daily time-scale, with the first year (1995) used as a warm-up period. The remaining three years of data (from 1 January 2001 to 31 December 2003) were used for validation. The objective function for calibration was the Nash-Sutcliffe coefficient of efficiency (NSE), which has been used widely in rainfall-runoff

modelling. In addition, the relative bias was also calculated to assess the bias in modelled runoff relative to observations.

The calibration and validation results are shown in Table 5-2. For each model, the calibration and validation performance varies by catchment, and the validation performance is on average slightly lower than the calibration performance. For each catchment, the values of each performance metric are generally similar across the three rainfall runoff models, although GR4J on average has the best performance and AWBM has the worst performance for the case study catchments.

Table 5-2: NSE for calibration and validation for the GR4J, AWBM and IHACRES_CMD models at the five case study sites, with relative bias shown in brackets.

SMA model name / Study site	GR4J		AWBM		IHACRES_CMD	
	Calibration	Validation	Calibration	Validation	Calibration	Validation
Adelaide	0.861 (0.011)	0.800 (0.025)	0.812 (-0.274)	0.804 (0.292)	0.856 (0.020)	0.809 (-0.028)
Burnie	0.862 (-0.010)	0.869 (-0.053)	0.799 (-0.089)	0.757 (0.018)	0.860 (0.008)	0.868 (-0.070)
Darwin	0.849 (0.052)	0.834 (-0.215)	0.825 (0.013)	0.823 (-0.148)	0.848 (0.048)	0.848 (-0.226)
Alice Springs	0.805 (-0.650)	0.581 (0.034)	0.778 (0.599)	0.569 (-0.737)	0.800 (-0.503)	0.585 (-0.912)
Wagga Wagga	0.748 (-0.072)	0.695 (0.083)	0.513 (-0.096)	0.618 (-0.298)	0.713 (-0.085)	0.642 (0.083)
Average NSE	**0.825**	**0.756**	**0.746**	**0.714**	**0.815**	**0.750**

5.3 Method

A schematic of the methodology is presented in Figure 5-3. The approach to addressing the three research questions described in the Introduction is detailed in Sections 5.3.1 to 5.3.3, respectively.

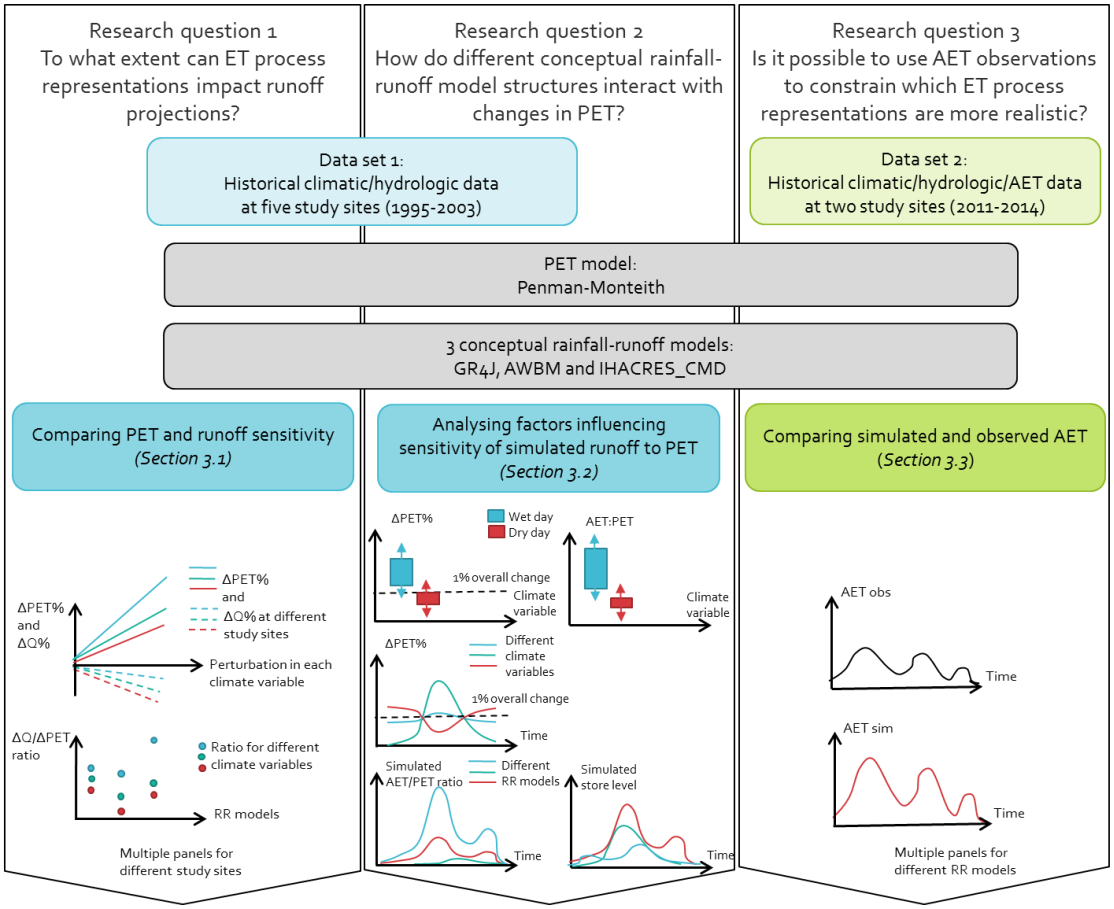


Figure 5-3: Schematic of methodology used to address the research questions posed in the Introduction.

5.3.1 *Impact of ET process representations on runoff projections*

To assess the impact of different ET process representations within conceptual rainfall-runoff models, we explored the sensitivity of both PET and runoff to changes in four climate variables related to PET, with differences in the ratio of sensitivity values between PET and runoff providing an indicator of the importance of ET process representation. The sensitivity curve method was employed to achieve this, as it has been widely applied in previous PET sensitivity studies due to its computational convenience and ease of interpretation (Goyal, 2004; McKenney and Rosenberg, 1993; Vicente-Serrano et al., 2014).

As required by the sensitivity curve method, perturbations were made to the four climatic variables which influence PET, namely temperature (T), relative humidity (RH), solar radiation (R_s) and wind speed (u_z), at each of the five study sites. This process yields multiple sets of perturbed climate data,

each obtained by perturbing one climate variable while keeping the others unchanged. It is worth noting that by perturbing each input variable independently, the sensitivity curve method does not take into account the potential joint-variations among the input variables due to correlations. However, according to a previous global sensitivity study across multiple locations in Australia, PET generally showed little sensitivity to these joint-variations among *T*, *RH*, R_s and u_z (Guo et al., 2017), so that it is reasonable to assume that this one-at-a-time climate perturbation is capable of capturing the majority of potential changes in PET due to the four climatic variables.

The bounds for perturbing each variable were selected to be slightly wider than the ranges of the projected changes in these variables by 2100 for Australia (Stocker et al., 2013) to encompass a comprehensive range of plausible future climate change scenarios (Table 5-3).

Table 5-3: Plausible perturbation ranges for each climate variable relative to their historical levels.

Climate variable to perturb	Perturbation range
T	0 to +8°C
RH	-10% to +10%
R_s	-10% to +10%
u_z	-20% to +20%

For each climate variable, the perturbation levels relative to its historical baseline consisted of four equidistant levels within the above-mentioned plausible ranges (i.e. +2, +4, +6 and +8°C for *T*, -10%, -5%, +5% and +10% for *RH* and R_s, and -20%, -10%, +10% and +20% for u_z), with each perturbation level being applied as a factor to the entire time series of the corresponding daily historical data. For *T* and *RH*, the Penman-Monteith model requires both the daily minimum and maximum values as input; therefore, the daily time-series of each pair of *T* and *RH* variables was considered jointly and thus perturbed by the same amount for each day. Since all perturbations for *T* were additive, whereas perturbations for R_s, *RH* and u_z were percentage changes, positive values of these variables were maintained during the perturbation. The perturbed RH data were also capped at 100% to avoid obtaining physically unrealistic humidity values. The sensitivity of PET

and runoff to different levels of perturbation in each of T, RH, R_s and u_z were then represented as the relative changes to their corresponding baseline historical data, as used in model calibration.

To compare the PET sensitivity with the runoff sensitivity from the three rainfall-runoff models, the ratio of the percentage change in annual average runoff for each percentage change in annual average PET—referred to as the runoff 'elasticity' to changes in PET (Chiew, 2006)—was also estimated.

5.3.2 The role of rainfall-runoff model structure on sensitivity to PET

The focus of this section is to better understand the effect of rainfall-runoff model structure on the sensitivity of runoff to changes in PET. As discussed in Section 5.2.2.2, the key structural differences of three rainfall-runoff models in the context of translating PET to AET and runoff are their manner of: (a) converting PET to AET on wet days and dry days; and (b) relating AET to the level of the soil moisture store(s). Therefore, to understand the causes of the impact of ET process representations on runoff projections, we investigated the impact of the following factors on the sensitivity of simulated runoff to PET:

1. Wet-day and dry-day patterns, to assess how rainfall-runoff models handle conversion from PET to AET on wet and dry days. Although the three rainfall-runoff models use different definitions of wet and dry days to determine how much of PET is converted to AET, for consistency we used a single definition of a wet day as one with rainfall greater than 1 mm for all models. The conversion from PET to AET was represented by the AET:PET ratio, which should be between 0 and 1 by definition. We compared the wet-/dry-day distributions of: (a) the daily PET responses to perturbations in each climate variable; and (b) the daily AET:PET ratios simulated from each rainfall-runoff model. We also investigated whether there were interactions between

them that can lead to different AET estimates for the same change in average PET.

2. Seasonal patterns, to assess how rainfall-runoff models relate AET to temporal variations in storage. Depending on the catchment, there is evidence that the production store can vary significantly by season (for example, see Westra et al. (2014b)). Therefore, we investigated the role of seasonality in storage on the simulated AET from rainfall-runoff models. We also investigated whether the responses of PET to each perturbed climate variable can vary seasonally and thus act synergistically or antagonistically to produce different AET simulated for the same total change in annual average PET. This was achieved by investigating the seasonal variations in: (a) the daily response of PET to perturbations in each individual climate variable; (b) the daily AET:PET ratios simulated in the three rainfall-runoff models; and (c) the daily storage levels simulated in the three rainfall-runoff models.

To provide a common basis for comparison, all the sensitivity results in this section were presented per 1% change in PET.

5.3.3 Relative realism of alternative ET process representations within conceptual rainfall-runoff models

To evaluate the relative realism of ET process representations within the three conceptual rainfall-runoff models, the AET simulated with the three models were compared with EC measurements of AET at Alice Springs and Wagga Wagga, as these were the only sites at which measured AET data were available. As a result of limitations in AET measurements at these sites (as discussed in Section 5.2.1), a qualitative approach was adopted for the comparison, in which the times-series of observed and simulated AET were

visually compared to evaluate each model's ability to represent the temporal variation of ET processes.

5.4 Results and discussions

5.4.1 Impact of ET process representations on runoff projections

The sensitivity of PET and runoff to changes in each climate variable are shown in Figuer 5-4, with each panel representing the sensitivity to a different climate variable (see Figures 5A.1 and 5A.2 in Appendix 5A for individual plots of PET and runoff sensitivity). The results show a large range of sensitivity values depending on (a) the perturbed climate variable, (b) the rainfall-runoff model and (c) the location. For example, PET is generally more sensitive to perturbations in T, with an 8°C increase in T leading to between a 20% and 32% increase in average PET depending on location (Figuer 5-4a). RH shows a negative relationship with PET, with a 10% increase in average RH leading to between a 2% and 13% decrease in average PET (Figure 5-4b). Sensitivity to R_s and u_z are consistently much smaller, with the increase in PET being between 3% and 5% for a 10% increase in average R_s (Figure 5-4c), and between 3% and 6% for a 20% increase in average u_z (Figure 5-4d).

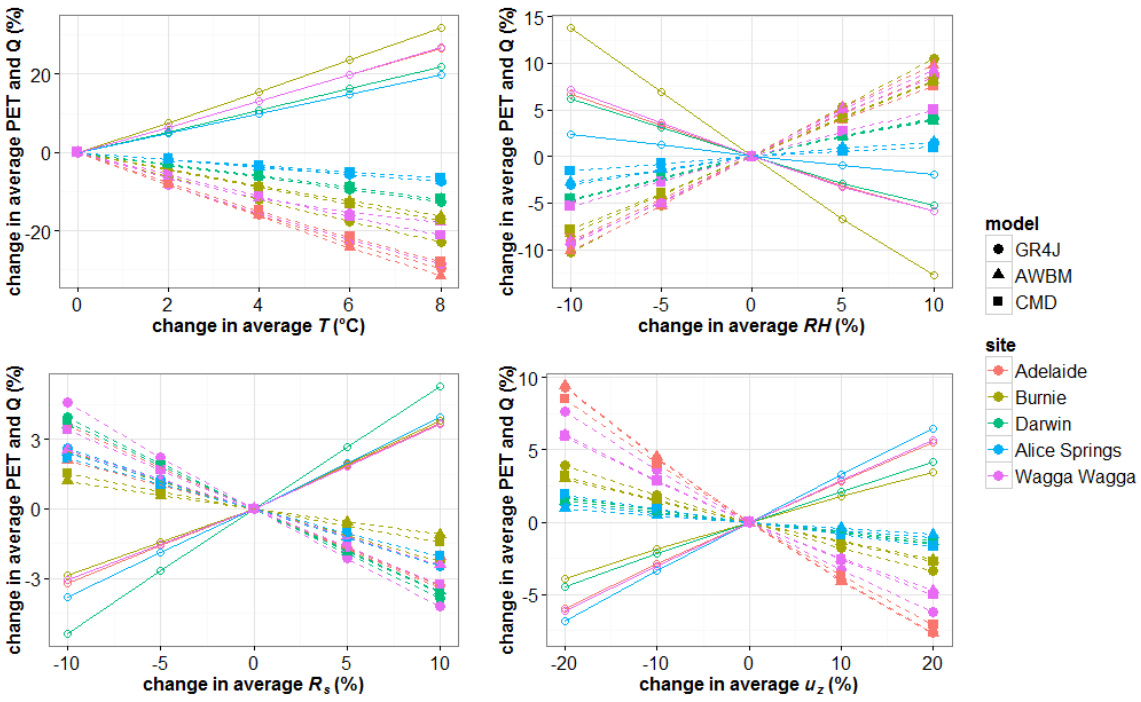


Figure 5-4: Sensitivity of PET and runoff to the four perturbed climate variables at the five study sites. The sensitivities for PET are represented by solid lines while for runoff dashed lines are used.

Also apparent from Figure 5-4 is that for a fixed climate perturbation, the runoff sensitivity always has an opposite sign to that of PET. However, the relative order of impact that the four climate variables have on runoff is consistent with that for PET. In particular, runoff generally shows a higher sensitivity to perturbations in *T*, for which an 8°C increase can cause between a 7% and 31% decrease in runoff depending on the location and rainfall-runoff model used (Figure 5-4a). *RH* also shows substantial sensitivity, leading to up to between a 3% and 10% increase in runoff for a 10% increase in *RH* (Figure 5-4b). R_s and u_z show much smaller impacts, with a 10% increase in R_s causing between a 1% and 4% decrease in runoff (Figure 5-4c), while a 20% increase in u_z can cause between a 1% and 8% decrease in runoff (Figure 5-4d). For both PET and runoff, the relationship between sensitivity values and perturbations in each climate variable are highly linear, suggesting that the results can be represented per unit change for both PET and runoff.

To assess the specific role of rainfall-runoff model choice on the sensitivity values in Figure 5-4, the runoff elasticity (i.e. the percentage change in annual average runoff for each percentage change in annual average

PET) simulated from three rainfall-runoff models is plotted for each study site in Figure 5-5. For the two most humid locations (Darwin and Burnie, which have the lowest long-term PET/P ratios as shown in Table 5-1), the elasticity values do not vary much across rainfall-runoff models and climate variables—this will be discussed further in Section 5.4.3.

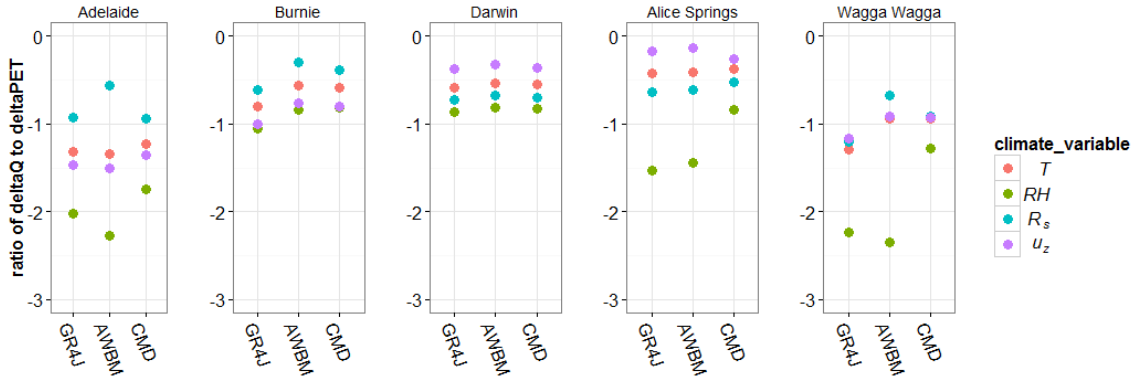


Figure 5-5: Runoff elasticity for different climate variables for each study site, as simulated from three rainfall-runoff models.

For the remaining locations (i.e. Adelaide, Alice Springs and Wagga Wagga), the choice of both the rainfall-runoff model and the climate variable being perturbed makes a substantial difference to the runoff elasticity. Although all models suggest that runoff shows higher elasticity when perturbing *RH* rather than other climate variables, GR4J and AWBM consistently show much greater elasticity due to *RH*, compared to IHACRES_CMD. Focusing on Alice Springs, GR4J suggests that for each 1% change in PET, perturbing *RH* leads to the greatest change in runoff of 1.5%, whereas perturbing u_z leads to the smallest runoff change of only 0.2%. This is equivalent to an over seven-fold difference in runoff elasticity attributable solely to perturbing different climate variables. A similar magnitude of difference in runoff elasticity is observed for AWBM. In contrast, IHACRES_CMD suggests that the runoff elasticity for perturbing all four climate variables is below 1%. This is quite surprising, as these conceptual rainfall-runoff models do not directly use the four climate variables as inputs, but instead use the PET estimated with the Penman-Monteith model. Consequently, the fact that the variation of runoff responses depends on the climate variable being perturbed clearly illustrates the impact of ET process

representations within individual rainfall-runoff models on propagating climate change signals to projected runoff.

To better understand these findings, in the next section we investigate the effect of rainfall-runoff model structure on the varying elasticity values shown above.

5.4.2 The role of rainfall-runoff model structure on sensitivity to PET

In the previous section, it was shown that the elasticity of runoff varied significantly depending on the variables used to perturb PET. Given that the results in Figure 5-5 are presented per unit change in annual average PET, it is anticipated that differences in elasticity values must be due to sub-annual variations that are different depending on the perturbing variable.

In this section we therefore assess the role of sub-annual variations in PET on runoff sensitivity, with a focus on daily and seasonal variations. For illustration purposes, we use the results from Alice Springs, which exhibited the largest differences in runoff elasticity among the climate variables (Figure 5-5), and is thus expected to illustrate the most significant impact of ET process representations. Analyses from other sites generally led to conclusions that are consistent with those from Alice Springs, and are available in Figure 5A.6 to Figure 5A.17 in Appendix 5A.

5.4.2.1 Interactions between PET sensitivity and AET simulated for wet and dry days

Figure 5-6a shows the wet/dry-day distribution of PET changes for a 1% change in annual average PET. Note that the same distribution of PET responses applies to GR4J, AWBM and IHACRES_CMD, since no rainfall-runoff modelling is yet involved. The figure indicates that for a fixed average change in PET, there are significant differences in the change for wet and dry days. Among all climate variables, *RH* causes the greatest difference in average dry-day and wet-day sensitivities of PET (0.743% and 3.04%, respectively). *T* and R_s also show lower sensitivity for dry days compared to wet days, but the differences are much smaller compared to those for *RH*

(with the two sensitivities becoming 0.989% and 1.01% for T, and are 0.917% and 1.27% for R_s, respectively). In contrast, u_z leads to higher sensitivities on dry days compared to wet days (with averages of 1.06% and 0.567%, respectively).

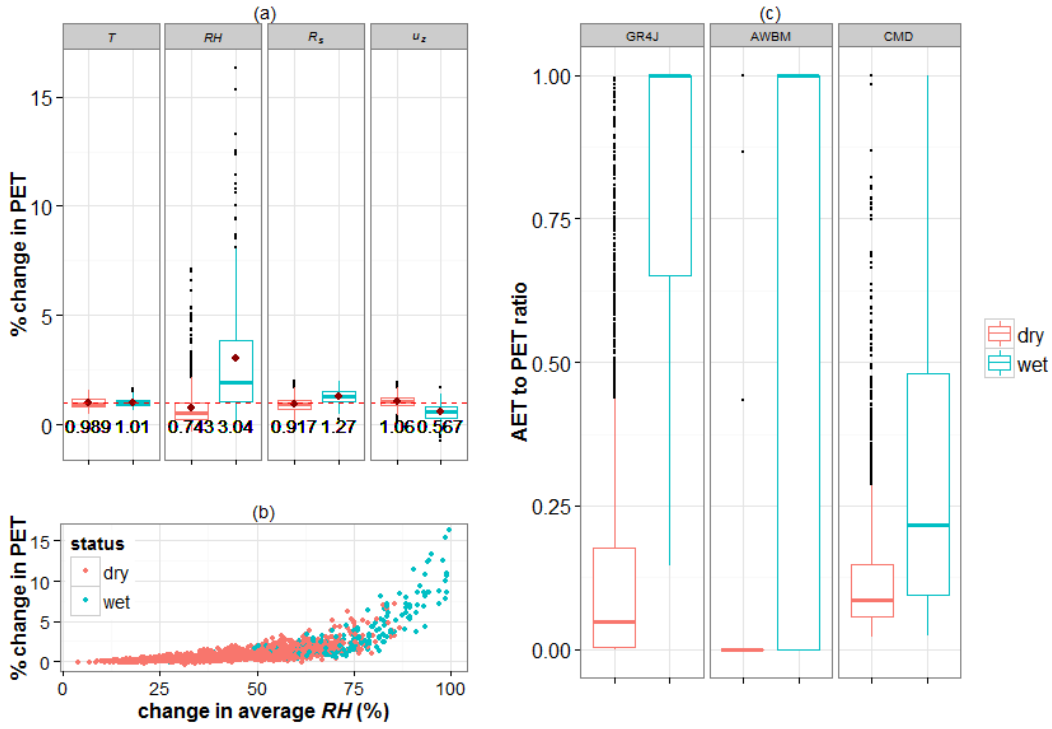


Figure 5-6: (a) Distribution of PET changes on wet and dry days, in response to perturbations in each climate variable (in each vertical panel) that result in an annual average PET change of 1% (indicated by the red dashed line), where a red dot represents the mean for each group; (b) all daily PET changes that lead to a 1% change in annual average PET from perturbing *RH*, against different baseline levels of *RH* for wet and dry days; and (c) AET:PET ratios on wet and dry days, simulated from AWBM, GR4J and IHACRES_CMD (in each vertical panel).

The substantially higher PET sensitivity to *RH* can be explained with the aid of Figure 5-6b, where all daily PET changes that led to a 1% change in annual average PET from perturbing *RH* (as the second panel in Figure 5-6a) are plotted against the corresponding baseline daily *RH*. The figure shows that: (a) wet days are typically associated with high baseline *RH*; (b) these high-*RH* wet days generally lead to greater change in PET. These illustrate that part of the higher PET sensitivity on wet days can be the result of using multiplicative perturbation for *RH*, as the average *RH* on wet days is around 1.5 times of that for dry-day *RH*, leading to greater perturbation of *RH* on wet

days than on dry days with any scaling factor. However, the use of a multiplicative perturbation method does not fully explain the four-fold difference in wet-/dry-day PET sensitivity in Figure 5-6a, which suggests that the structure of the Penman-Monteith model also contributes to greater PET changes during wet days compared to dry days.

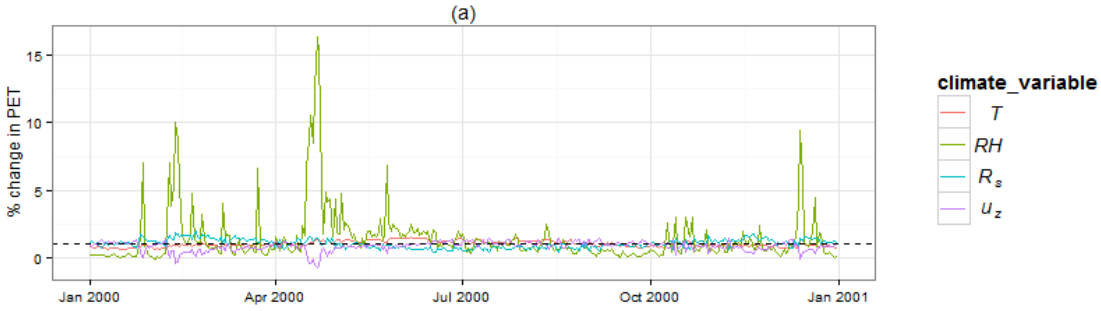
Having shown that the PET sensitivity differs from wet to dry days depending on the perturbing climate variable, we now investigate how PET is converted to AET differently on wet and dry days via different rainfall-runoff models. Figure 5-6c presents the distribution of daily AET:PET ratios on wet and dry days, simulated from the three rainfall-runoff models. All three models illustrate higher ratios of AET:PET during wet days compared to dry days, with GR4J and AWBM showing greater differences than IHACRES_CMD. This is because GR4J and AWBM both determine if AET = PET by checking if P – PET > 0, which produces a peak ratio of one for most wet days, while those below one correspond to days with P > 1 mm, but still not yet satisfying P – PET > 0. In contrast, IHACRES_CMD only uses store levels to determine AET, which is therefore largely independent of individual rainfall events, so that it does not lead to substantially higher AET on wet days. For dry days, the AET:PET ratios from GR4J and IHACRES_CMD show similar distributions, in contrast to those from AWBM. Although most dry-day ratios from GR4J and IHACRES_CMD are around zero, they can vary substantially with values spanning the entire range from zero to one, highlighting the direct impact of instantaneous store levels on AET in both models. In AWBM, this ratio is mostly zero with only a few anomalies, corresponding to dry days that have small rainfall amounts (0 < P < 1 mm), but still with water in some or all of the stores, leading to overall AET at discrete ratios of PET (i.e. 0, 0.433, 0.866 and 1).

Connecting these results with those in Figure 5-5, it is clear that for a fixed positive average change in PET, a greater decrease in runoff would be expected to occur for variables that cause the greatest change of PET during wet days, as a larger percentage of PET would be ‘lost’ to AET during wet days in both GR4J and AWBM. Based on this finding, one would expect the

highest runoff elasticity for both models for *RH*, followed by R_s, *T* and finally u_z. This is exactly the same order as was identified for Alice Springs in Figure 5-5.

5.4.2.2 Interactions between seasonal variations in PET sensitivity, simulated AET and storage

We now consider the role of seasonal patterns in the changes in PET as a result of perturbing each climate variable, and how they interact with simulated AET and storage. Figure 5-7a shows the seasonal distributions of the daily PET responses to perturbations in each climate variable that result in a 1% average change in PET at Alice Springs for the year 2000. The figure highlights substantial differences in the temporal distribution of PET responses, depending on the climate variable perturbed. For perturbations in *RH*, PET responses display high temporal variability: two clear peak periods are shown around February and May with the highest PET response exceeding 15 times average levels, while from July to October most responses are close to zero. In contrast, the variability in PET responses to the perturbations in *T*, R_s and u_z are much smaller and close to 1% for the entire year.



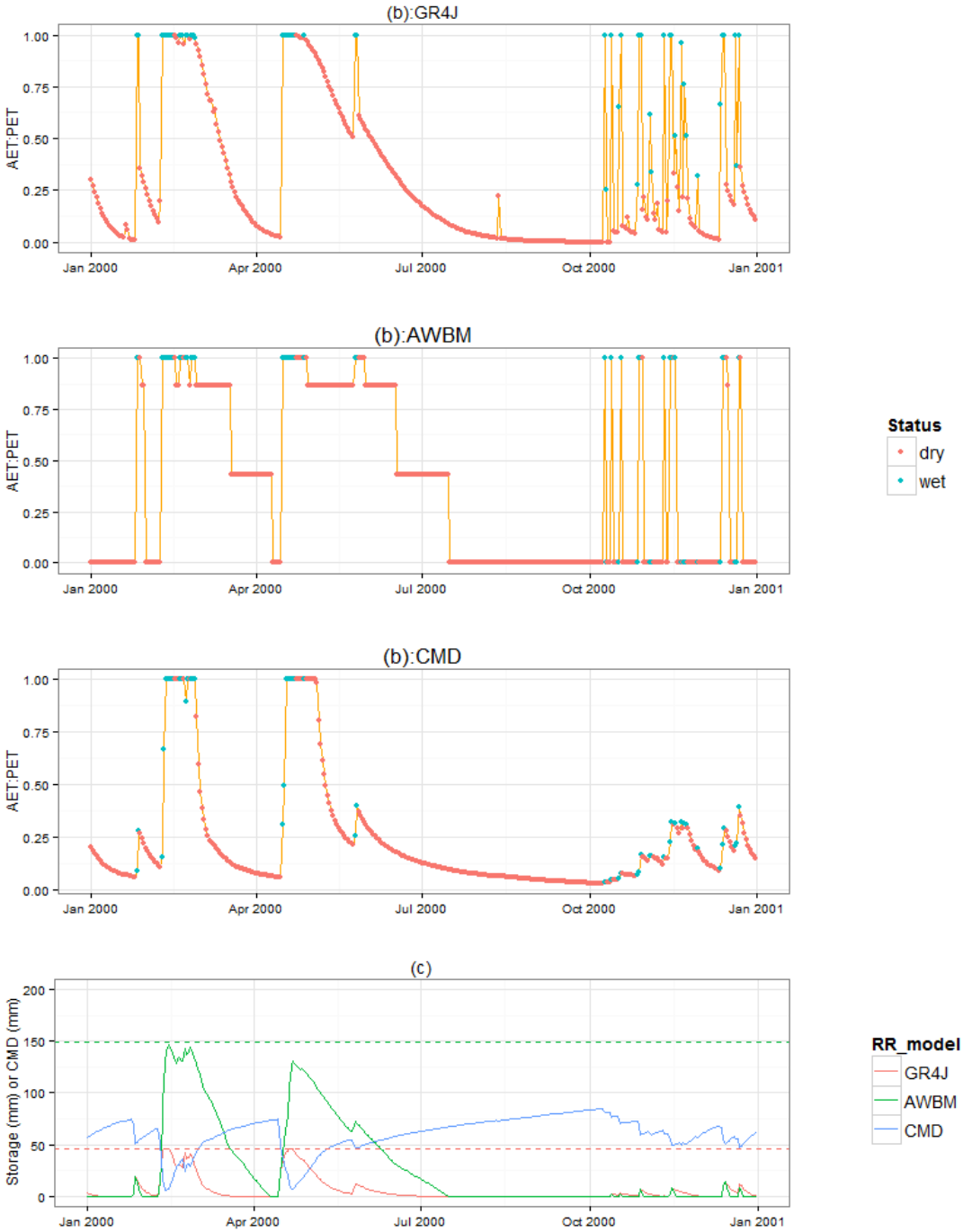


Figure 5-7: Seasonality of a) daily PET responses to perturbations in each climate variable that result in an average PET change of 1% (black dashed line); b) AET:PET ratios for wet and dry days from GR4J, AWBM and IHACRES_CMD; and c) levels of production stores from GR4J, AWBM (with dashed lines indicating store capacities) and IHACRES_CMD (as CMD levels). All results are from Alice Springs for the year 2000.

Having shown the different seasonal variations in the PET sensitivity depending on the perturbing climate variable, we now investigate the seasonal patterns in the conversion from PET to AET within the three rainfall-runoff

models. The simulated daily time-series of AET:PET ratios from the three models are shown in Figure 5-7b for Alice Springs for the year 2000.

Considering firstly the results from January to October, all models suggest two periods with peak AET:PET ratios occurring around February and May, respectively, which correspond to the simulated store levels from all models (Figure 5-7c), which also show two consistent peak periods. Following the peak periods, the models show contrasting temporal patterns in both AET and storage: from both GR4J and IHACRES_CMD, the AET:PET ratios show a smooth "recession period" from one to close to zero, while the store drains gradually at the same time. This is because both models limit the rate of AET when store levels decrease. In contrast, the role of the store level in simulated AET is less obvious for AWBM, as the AET:PET ratios drop in a 'stepwise' fashion (from 1 to levels of 0.866, 0.433 and then 0) following each peak period, corresponding to times when different stores become empty (as detailed in Section 5.2.2.2). In addition, compared to GR4J and IHACRES_CMD, the stores simulated in AWBM drain at a more constant rate (for example before April), because this model forces each store to keep evaporating at the rate of PET until empty.

Interestingly, after October, GR4J and AWBM both show instantaneous high AET:PET ratios on wet days, but returning to lower values once the rainfall ceases. This indicates the impact of individual rainfall events, which is independent of that of storage, since peak ratios can occur even when storage levels are consistently low. Such impact is not visible from IHACRES_CMD as it does not relate AET simulation to individual rainfall.

From these results, all three models show a fuller storage around both February and May. Since all models relate AET simulation to storage, a greater amount of PET is also converted to AET during this period, with associated implications on runoff volume. As these periods also correspond to when *RH* can cause the greatest changes in PET, we can expect that for a fixed positive average change in PET, perturbations in *RH* can lead to greater decreases in runoff from all three models. This, again, is consistent with the higher sensitivity of runoff to *RH* shown from all three models in Figure 5-5.

Combining results from Sections 5.4.2.1 and 5.4.2.2, we can see that:

1. PET shows substantially higher sensitivity to *RH* during wet days and around February and May;
2. Both GR4J and AWBM explicitly relate AET simulation to daily wet/dry status, leading to higher impact of changing PET on AET during wet days; and
3. All three models relate AET simulation to catchment storage, leading to higher impact of changing PET on AET when stores are fuller around February and May.

Relating the above results to observations for Alice Springs in Figure 5-5, it is clear that the higher elasticity values for *RH* compared to the other variables shown in all three models are likely to be attributable to these models having a higher contrast in the AET:PET ratios across different seasons where store levels vary substantially (Figure 5-7). Furthermore, the reason that GR4J and AWBM show even higher sensitivity to *RH* can be that these models having higher contrast in AET:PET ratios between wet day dry days across wet and dry days, which are associated with contrasting levels of *RH* (Figure 5-6).

In summary, in this section we show that there are clear daily and seasonal variations in the sensitivity of PET depending on which climate variable is perturbed. This can then interact with the way PET is converted to AET in the three models, which depends on whether a day is dry or wet, as well as the temporal variation in the production store. The interaction between these two mechanisms can lead to large variability in runoff sensitivity to different perturbing climate variables, even for the same change in annual average PET.

5.4.3 Relative realism of alternative ET process representations within conceptual rainfall-runoff models

Given the importance of rainfall-runoff model process representation in determining how much AET is converted from PET, we now investigate model realism by comparing observed time-series of AET with simulated AET from the three conceptual rainfall-runoff models at Alice Springs and Wagga Wagga (Figure 5-8).

For Alice Springs, the differences between model simulated AET clearly illustrate their contrasting ET process representations. GR4J illustrates both the effect of individual rainfall events, as well as change in store levels, with simulated AET peaking on individual wet days and gradually decreasing on the dry days after, which indicates drops in store levels. In AWBM, the effect of individual rainfall is also illustrated by the peak AET on wet days; however, the influence of retreating storage is not shown in AWBM, with AET either remaining at or near its peak value, or dropping suddenly to zero. This instantaneous drop is due to the 'discrete' function that AWBM uses to represent the relationship between AET and storage (as discussed in Section 5.2.2.2). In contrast to GR4J and AWBM, the simulated AET from IHACRES_CMD is not sensitive to individual rainfall events, as the occurrence of peak AET is not associated with wet days. This is because AET is simulated purely as a function of the moisture content in the store. In general, GR4J and IHACRES_CMD better resemble the temporal variation in the observed AET in Alice Spring, which indicates that the change in store levels is likely to be a key controlling mechanism for ET processes in this catchment.

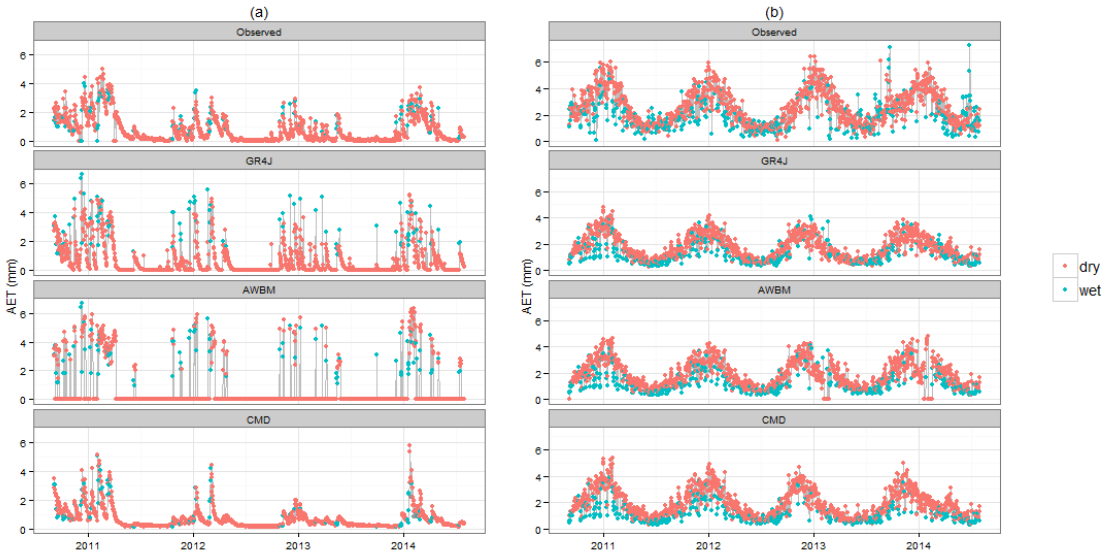


Figure 5-8: Comparison of the observed and simulated time-series of AET from GR4J, AWBM and IHACRES_CMD, at a) Alice Springs; and b) Wagga Wagga.

In contrast to Alice Springs, all three models show reasonable performance in producing the temporal patterns in the observed AET at Wagga Wagga, with only slight differences across models suggesting a lower impact of ET process representation. As an exception, at the start of 2013 and 2014, AWBM shows zero-AET periods, which are clearly different from observations. As illustrated in Figure 5-7, these periods correspond to times when all storages are empty, as a result of AWBM keeping each store evaporating at the PET rate (i.e. AET=PET).

The comparisons of AET simulation at the two study sites indicate that the impact of ET process representations is likely to be greater for drier catchments such as Alice Springs, where stores vary substantially on a seasonal basis, and experience frequent draining and rapid changes of levels in response to rainfall. In contrast, Wagga Wagga is a somewhat more humid catchment with lower seasonal variability (although still water limited, see Table 5-1). As such, it generally has fuller storage levels, which neither drain out completely nor rise rapidly with individual rainfall events. Under this condition, the simulated AET is likely to be less sensitive to different ways of defining the relationships between the AET:PET ratio and the storage level (e.g. GR4J and AWBM). Furthermore, with the consistently fuller stores leading to higher AET overall, the impact of different ET representations

across dry/wet status (e.g. GR4J and IHACRES_CMD) are likely to be limited. Therefore, model choice is likely to have less impact on ET simulations for catchments with relatively stable store levels. This can also be a possible explanation of the observations in Section 5.4.1, where humid locations, such as Darwin and Burnie, show fewer differences in runoff sensitivity to different climate variables across the three rainfall-runoff models.

It should be noted that although the EC-based AET measurements were used as a benchmark for the simulated AET from the three rainfall-runoff models, the comparison should be interpreted qualitatively rather than quantitatively, due to several limitations associated with the available data. In particular, both catchments used for this comparison were assumed to be covered with reference crop when estimating catchment PET for simulating AET (Section 5.2.2.1). However, the EC measurements were obtained from flux towers within tall canopy (Mulga woodland for Alice Springs and Eucalyptus forest for Wagga Wagga), which are likely to be valid only for a small region for the specific type of canopy around each flux tower, since EC measurements usually only cover spatial scales of hundreds of meters. This inconsistency can contribute to differences between the simulated and observed catchment AET. Furthermore, a number of studies have reported biases in AET measured by the EC method due to the failure to close the energy balance (e.g. Wang and Dickinson, 2012; Wilson et al., 2001). These biases appear to be particularly prominent for days with high relative humidity and wind (Ibrom et al., 2007), as well as during turbulent conditions at night or when precipitation or dew obscured the sensors (Meiresonne et al., 2003; Meyer et al., 2015; Wilson et al., 2001).

Despite the abovementioned limitations, the ET observations from the two sites illustrate contrasting temporal variations related to rainfall events and store levels, which can be explained by the unique characteristic of each catchment. Therefore, it is likely that these patterns shown in the data may nevertheless be useful to scrutinize rainfall-runoff models and assess which models simulate ET processes more realistically. Furthermore, the results also

highlight potential opportunities for more detailed process-based evaluation of rainfall-runoff models where data are available, which can provide valuable insights to support rainfall-runoff model selection complementary to the use of conventional runoff-centered performance assessment (e.g. calibration metrics in Table 5-2). These process-focused assessments are particularly important when evaluating the suitability of models for simulating catchment behavior under changed conditions, where the differences in the runoff projections from alternative rainfall-runoff models are likely to be greatest.

5.5 Summary and conclusions

In this study, we:

1. Assessed the impact of ET process representations within conceptual rainfall-runoff models on runoff projections under plausible changes in the climate drivers of ET;
2. Explored the interaction between conceptual rainfall-runoff model structure and the complex sub-annual changes in PET, and how this influences projections of runoff and AET; and
3. Evaluated the relative realism of alternative ET process representations within three conceptual rainfall-runoff models using AET observations.

We investigated five study sites across Australia with contrasting hydro-climatic conditions, at which the baseline PET was simulated using the Penman-Monteith model. Perturbation of PET was achieved by first perturbing the historical data of four climatic drivers of ET, namely temperature (T), relative humidity (RH), solar radiation (R_s) and wind (u_z), within plausible ranges of their future changes, and then estimating the corresponding PET with the Penman-Monteith model. The baseline and perturbed PET time series formed inputs to three structurally different conceptual rainfall-runoff models (GR4J, AWBM and IHACRES_CMD),

which were calibrated to historical runoff data at the five study sites. The calibrated rainfall-runoff models were used to simulate the baseline and perturbed runoff, and then to estimate runoff sensitivity to climate perturbations.

Our results illustrate that different ET process representations in conceptual rainfall-runoff models can have substantial impacts on the sensitivity of runoff projection under a changing climate. Specifically, it is possible to have an over seven-fold difference in runoff elasticity due to rainfall-runoff model choice and the climate variable that is perturbed. The significant differences in runoff elasticity were attributed to the different conceptualizations used to convert PET to AET on wet and dry days in the three conceptual rainfall-runoff models, as well as the contrasting relationships defined between AET and store levels.

Using eddy-covariance-based observations of AET, we also compared and evaluated the ET process representations within the three conceptual rainfall-runoff models at Alice Springs and Wagga Wagga. Results suggest that alternative ET process representations are likely to have a greater impact on drier catchments consisting of highly variable store levels, such as Alice Springs. As highlighted with the results from Alice Springs, AWBM poorly simulated the variations in observed AET, whereas GR4J and IHACRES_CMD performed better in simulating the temporal variation of AET as a result of changing store levels.

Although this study demonstrates that evapotranspiration process representation within conceptual rainfall-runoff models can substantially affect the results from climate impact assessments, quantitative sensitivity values are likely to vary on a case-by-case basis, depending on the catchment and the choice of PET and rainfall-runoff models. Therefore, specific studies are required for further exploration of the implications of modelling assumptions for a larger number of rainfall-runoff models, as well as their interactions on other case studies. Nevertheless, the results from this study highlight the importance of scrutinizing internal process realism within rainfall-runoff models as part of climate impact studies, and using alternative

sources of information on model performance in addition to the standard runoff-based metrics, such as by comparing simulated and observed AET or with the outputs of integrated surface-subsurface flow models (e.g. Li et al., 2015; Partington et al., 2013; Su et al., 2016; Vervoort et al., 2014)).

These findings are in line with a number of studies that suggest the need for an improved understanding of physical processes to better infer the potential changes in rainfall-runoff relationships under climate change, and thus to facilitate better modelling of future water resources (Fowler et al., 2016; Herman et al., 2013; Kirchner, 2006; Merz et al., 2011; Minville et al., 2014; Saft et al., 2016; Westra et al., 2014b). The study also provides a practical example of the multiple-hypothesis testing of rainfall-runoff models (Clark et al., 2011; Clark et al., 2016; Gupta et al., 2008; Harrigan et al., 2014), which suggests that models should be evaluated based on the validity of any individual hypothesis they describe. With the currently increasing availability of observations to support hypothesis testing (for example see Blöschl et al., 2016), this type of process-based model evaluation can greatly enhance our understanding on how alternative rainfall-runoff models can propagate climate change signals to hydrological impacts via different model conceptualizations and assumptions. As highlighted by the substantially different elasticity values found in this study, these process-based model evaluation studies will be particularly useful to inform the selection of rainfall-runoff models for climate impact studies, where good model performance under historical climate conditions provides only a partial guide as to how the model will perform under changed climate conditions.

5.6 Acknowledgements

The streamflow data used for the five catchments in this study were from the Australian Bureau of Meteorology's (BOM) Hydrologic Reference Station project website: www.bom.gov.au/hrs (Bureau of Meteorology, 2015). The rainfall, temperature, relative humidity, wind and sunshine hour data for the five weather stations were obtained from the Climate Data Online project website, http://www.bom.gov.au/climate/data/ (Bureau of Meteorology, 2016). The EC-based AET observations for Alice Springs and Wagga Wagga were obtained from the OzFlux website, at http://www.ozflux.org.au/ (Cleverly, 2011; vanGorsel, 2013). The authors would also like to thank the three anonymous reviewers, whose comments have helped to improve the quality of this paper significantly.

Appendix 5A Supplementary to Chapter 5

1. Introduction

Included in the supplementary material are:

- Details of data sources as supplements to those presented in Section 5.2.1 of the main text;

- Elasticity of PET and runoff, reproduced using the same results shown in Figure 5-4 of Chapter 5, but separating PET and runoff for better visualization; and

- The full analysis results for all five study sites, as supplement to those presented only for Alice Springs in Section 5.4.2 in the main text. To illustrate the impact of ET process representations, Section 5.4.2 used Alice Springs as an example, as it exhibited the greatest differences in runoff elasticity amongst the five case studies and thus indicating the most distinct impact of ET process representation. As shown in the supplementary material, similar analyses for the other four study sites (namely Adelaide, Burnie, Darwin and Wagga Wagga) generally lead to consistent conclusions.

2. Data and models – details of data sources

To calibrate three conceptual rainfall-runoff models for the five case study catchments, daily runoff data were obtained from the gauging stations at the outlet of each catchment, while the catchment-average rainfall was represented either with data from a single rain gauge within each catchment (for the four smaller catchments) or from multiple rain gauges using Thiessen polygon (Table 5A.1). The PET data for calibration were calculated using the Penman-Monteith model from climate data of temperature, relative humidity, solar radiation and wind. Due to the limited availability of high-quality observations, data for each of these climate variables were obtained from a weather station near each catchment.

Table 5A.1. Sources of data used for calibrating rainfall-runoff models

Weather station (BoM ID)	Catchment	Catchment outlet (BoM ID)	Rain gauge(s)
Adelaide (23090)	Scott Creek (29 km^2)	A5030502	23734
Burnie (91009)	Black River (318.5 km^2)	314213	91009
Darwin (14015)	Elizabeth River (95.6 km^2)	G8150018	14222
Alice Springs (15590)	Hugh River (3324 km^2)	G0050115	15562, 15564, 15642
Wagga Wagga (72150)	Adelong Creek (146.1 km^2)	410061	72000

3. Results and discussions – Elasticity of PET and Q

Figures 5A.1 and 5A.2 show the sensitivity of (a) PET and (b) runoff to changes in each climate variable, and corresponds to the results shown in Figure 5-4 in Chapter 5.

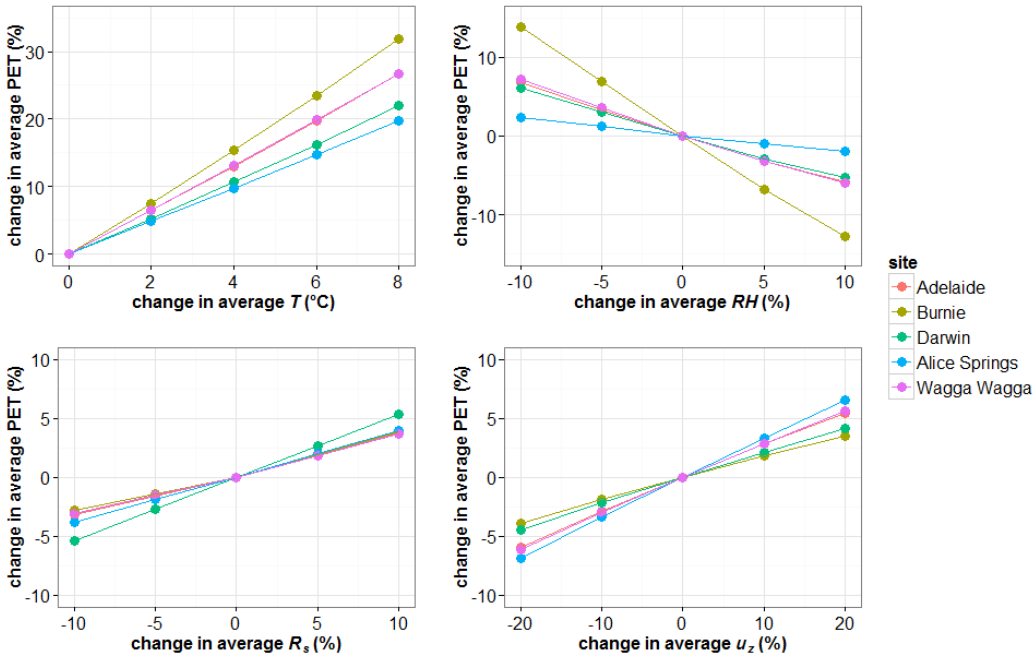


Figure 5A.1. Sensitivity of PET to the four perturbed climate variables at the five study sites.

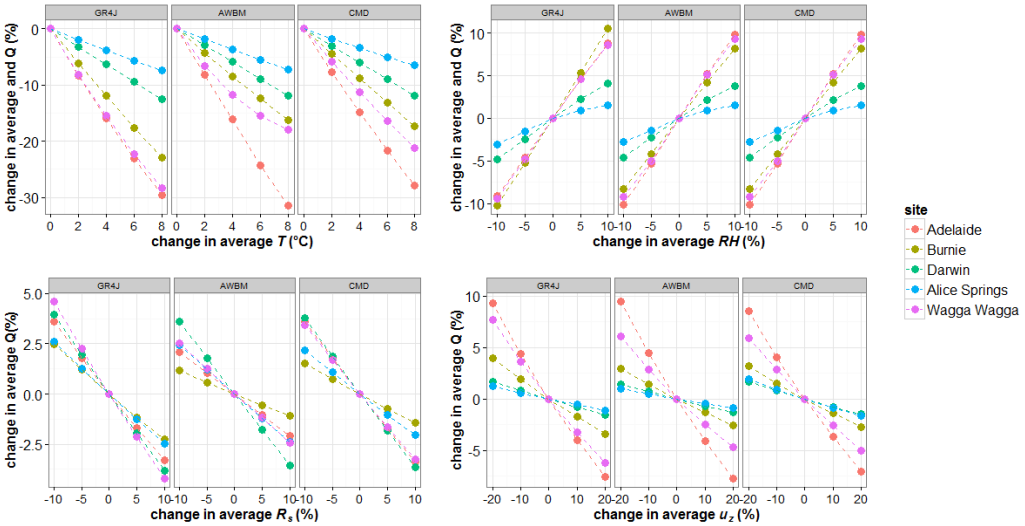


Figure 5A.2. Sensitivity of runoff to the four perturbed climate variables at the five study sites, with individual panels showing simulations from each of the three rainfall-runoff models.

4. Results and discussions – Interactions between PET sensitivity and process representations in rainfall-runoff models

Figure 5A.3 to Figure 5A.17 consist of the full results for all five study sites, corresponding to those presented in Figures 5-6 and 5-7 and discussed in Sections 5.4.2.1 and 5.4.2.2 of the main text, respectively. In addition, the absolute values of AET and PET for each site are presented to assist the understanding of AET:PET ratios.

As illustrated in Section 5.4.2 in the main text, amongst the five sites, Alice Springs exhibited the greatest differences in runoff elasticity depending on the climate variable being perturbed, thus indicating the most distinct impact of ET process representation. Therefore, Section 5.4.2 used Alice Springs as an example to illustrate the impact of ET process representations.

Figure 5A.6 to Figure 5A.17 illustrate the results for similar analyses for Adelaide, Burnie, Darwin and Wagga Wagga, which generally lead to consistent conclusions with those from Alice Springs.

Alice Springs

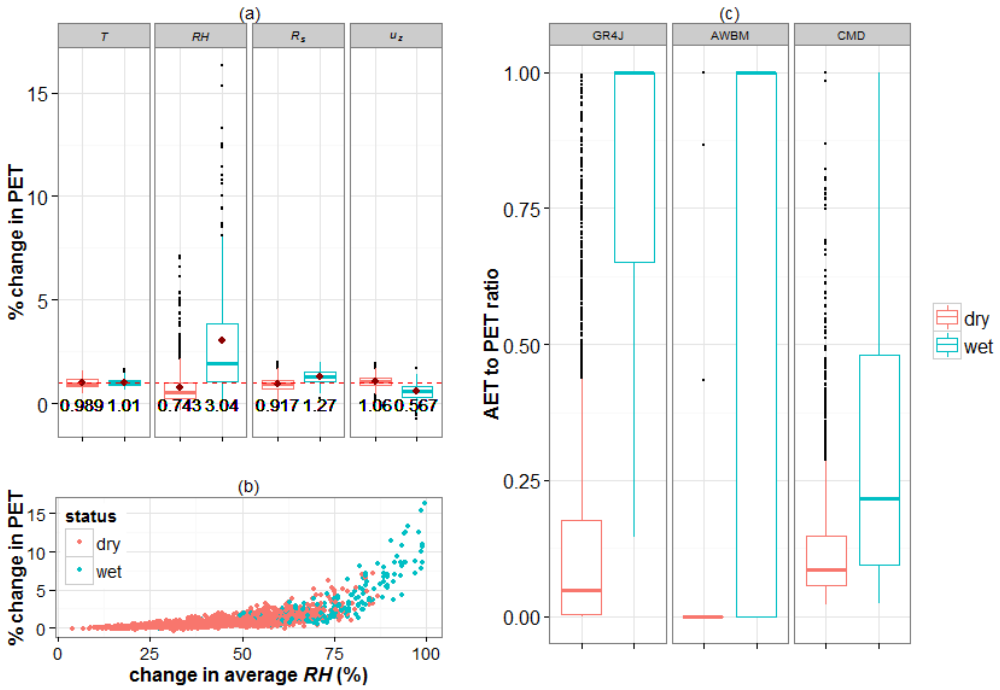


Figure 5A.3. Analyses of daily PET changes in Alice Springs, including: (a) Distribution of PET changes on wet and dry days, in response to perturbations in each climate variable (in each vertical panel) that result in an annual average PET change of 1% (indicated by the red dashed line), where a red dot represent the mean for each group; (b) scatterplot of PET changes against different baseline levels of RH for wet and dry days; and (c) AET:PET ratios on wet and dry days, simulated from AWBM, GR4J and IHACRES_CMD (in each vertical panel).

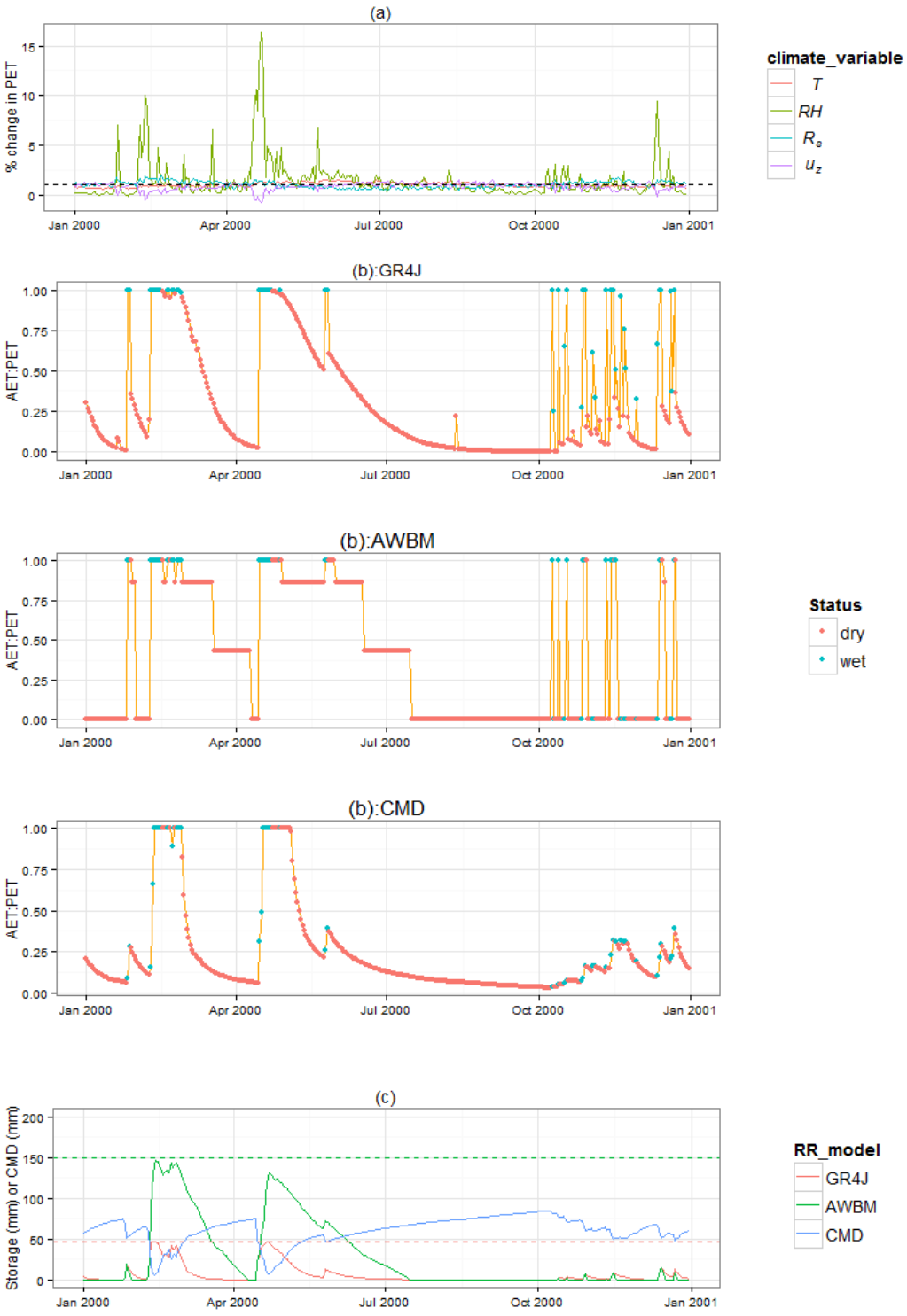


Figure 5A.4. Seasonality of daily a) PET responses to perturbations in each climate variable that result in an average PET change of 1% (black dashed line); b) AET:PET ratios for wet and dry days from GR4J, AWBM and IHACRES_CMD; and c) levels of production stores from GR4J, AWBM (with dashed lines indicating store capacities) and IHACRES_CMD (as CMD levels), at Alice Springs for 2000.

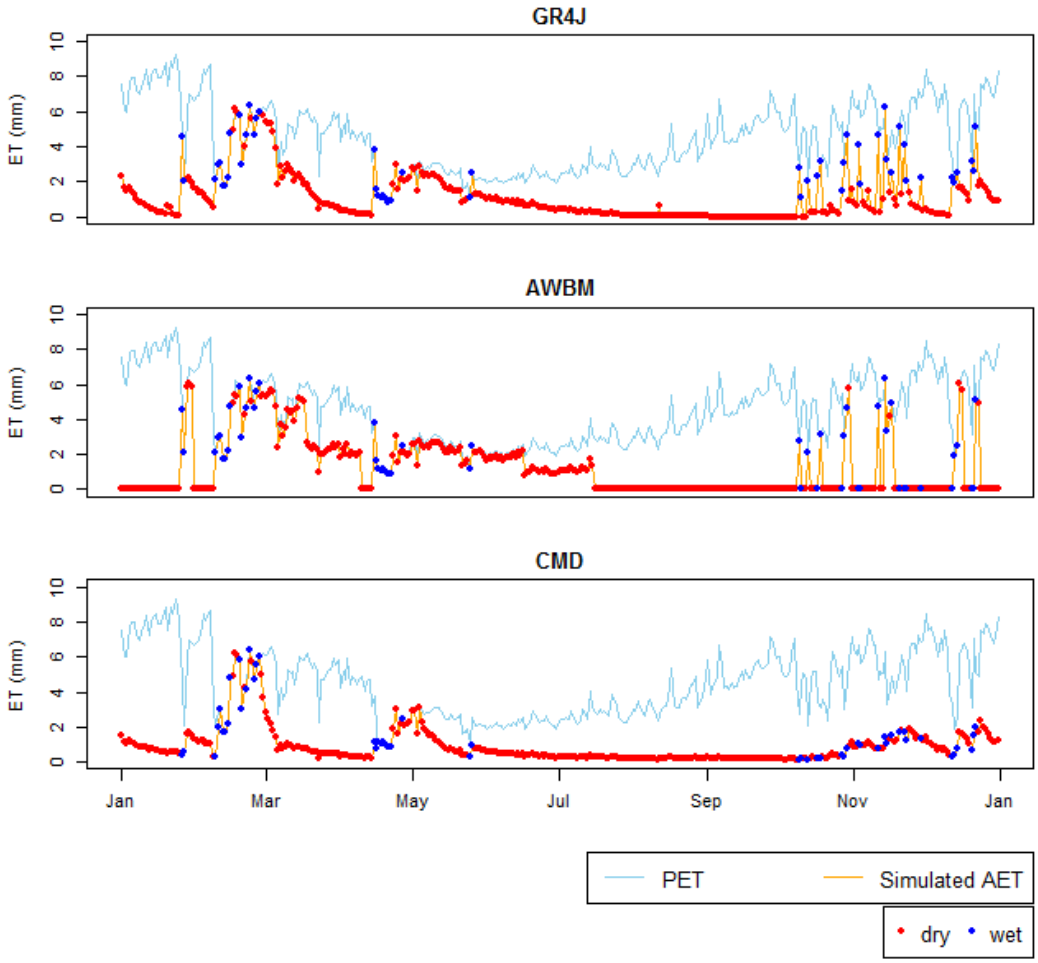


Figure 5A.5. Seasonality of the daily AET and PET for wet and dry days from GR4J, AWBM and IHACRES_CMD, at Alice Springs for 2000.

Adelaide

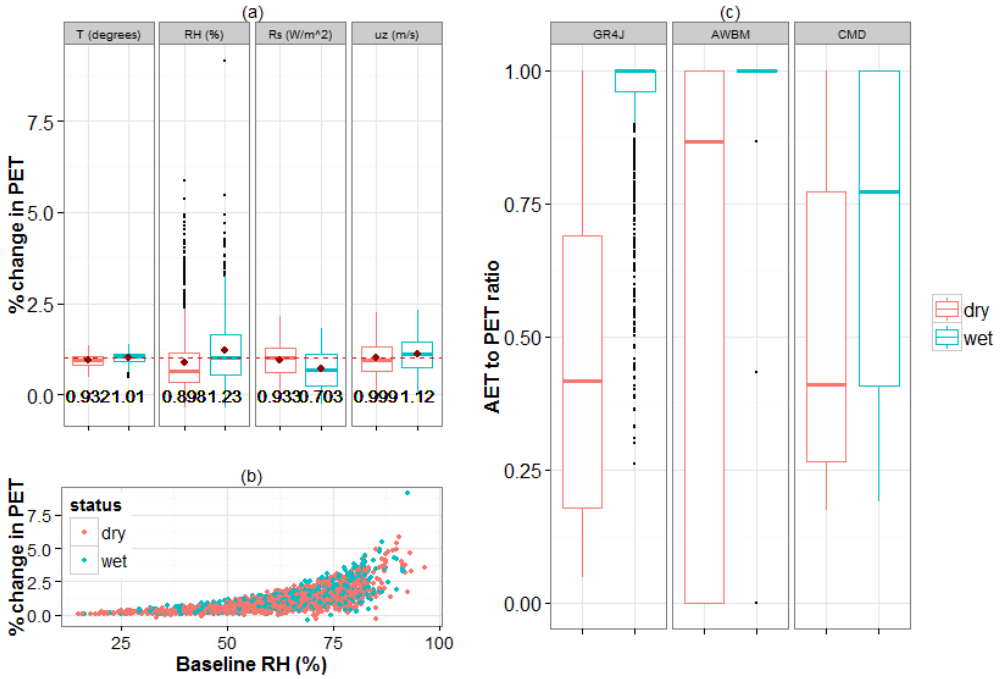


Figure 5A.6. Analyses of daily PET changes in Adelaide, including: (a) Distribution of PET changes on wet and dry days, in response to perturbations in each climate variable (in each vertical panel) that result in an annual average PET change of 1% (indicated by the red dashed line), where a red dot represent the mean for each group; (b) scatterplot of PET changes against different baseline levels of RH for wet and dry days; and (c) AET:PET ratios on wet and dry days, simulated from AWBM, GR4J and IHACRES_CMD (in each vertical panel).

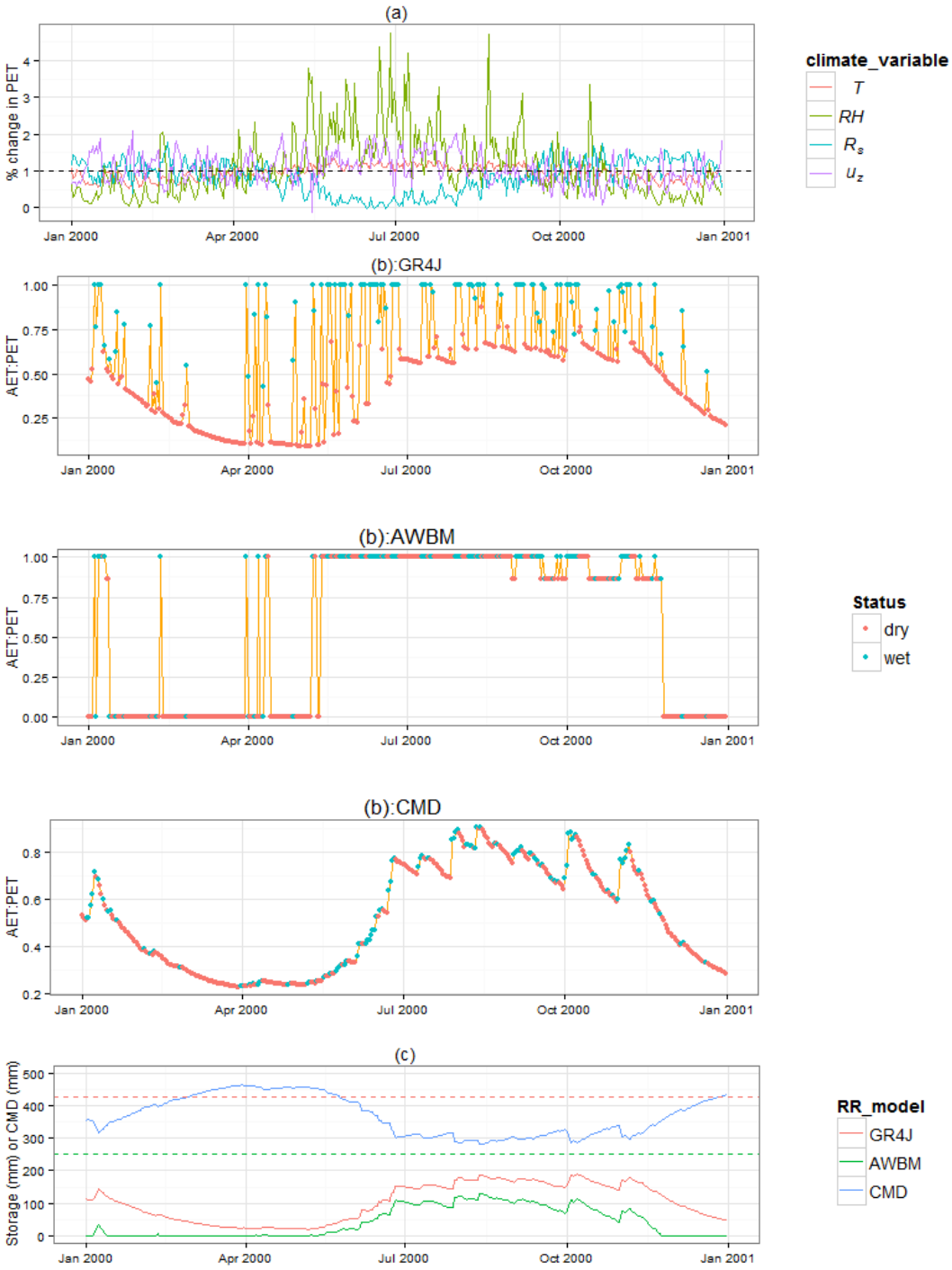


Figure 5A.7. Seasonality of daily a) PET responses to perturbations in each climate variable that result in an average PET change of 1% (black dashed line); b) AET:PET ratios for wet and dry days from GR4J, AWBM and IHACRES_CMD; and c) levels of production stores from GR4J, AWBM (with dashed lines indicating store capacities) and IHACRES_CMD (as CMD levels), at Adelaide for 2000.

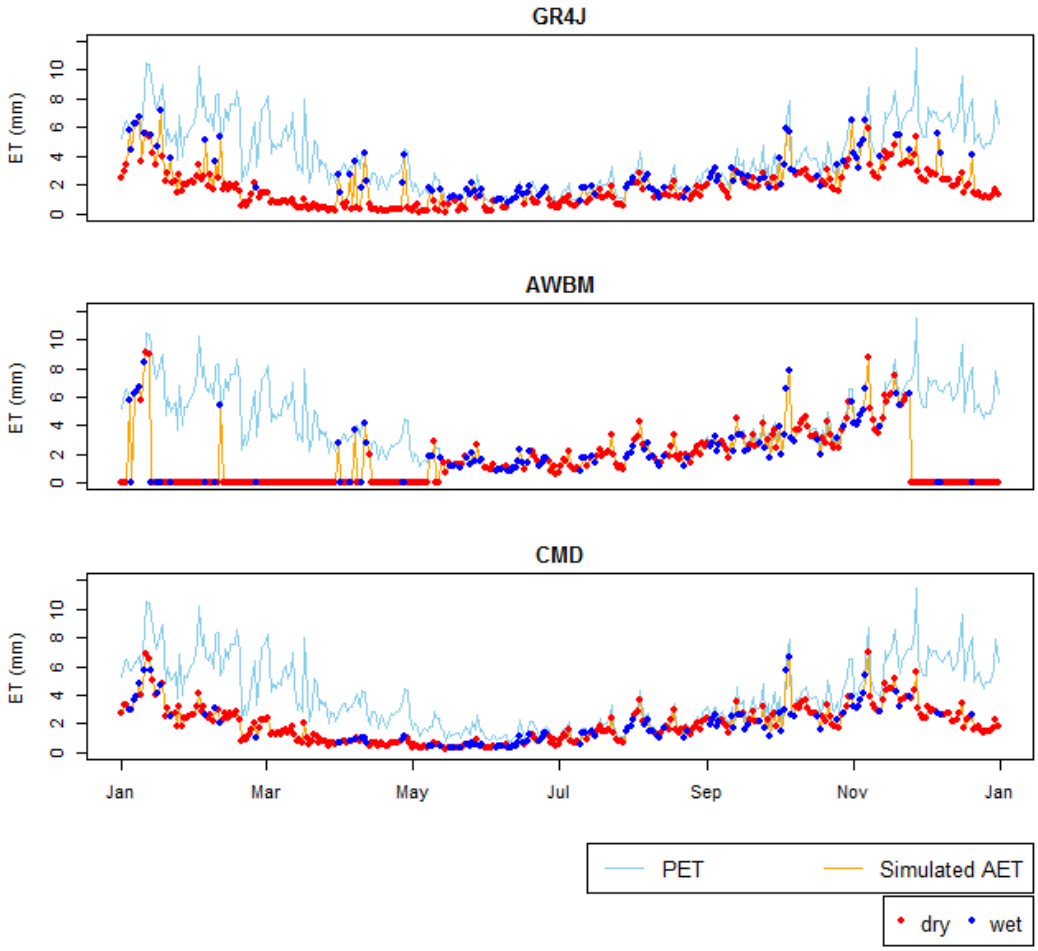


Figure 5A.8. Seasonality of the daily AET and PET for wet and dry days from GR4J, AWBM and IHACRES_CMD, at Adelaide for 2000.

Burnie

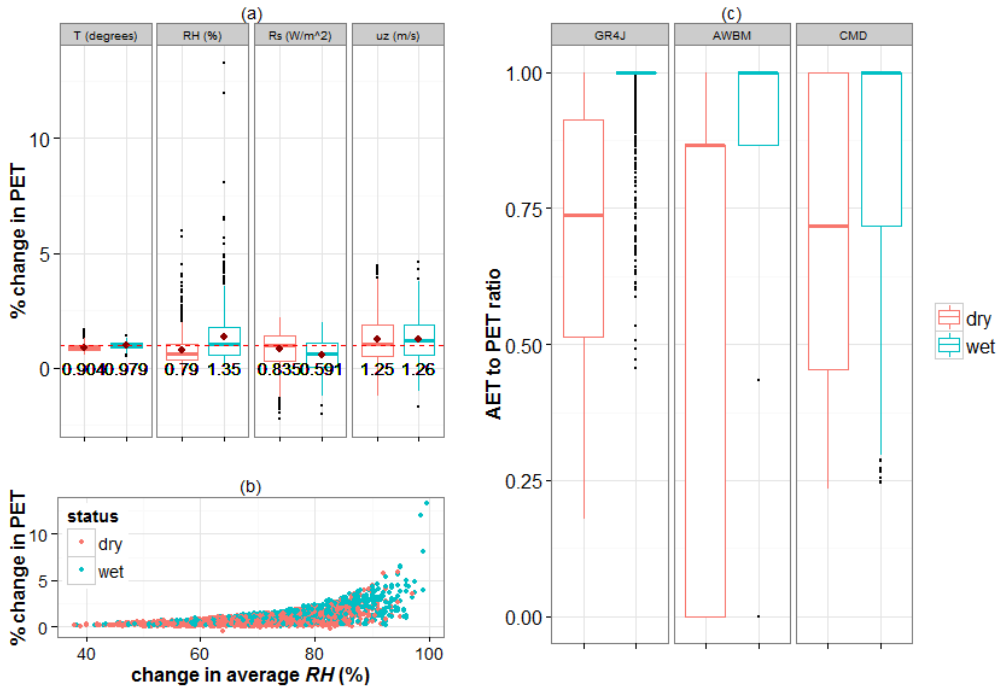


Figure 5A.9. Analyses of daily PET changes in Burnie, including: (a) Distribution of PET changes on wet and dry days, in response to perturbations in each climate variable (in each vertical panel) that result in an annual average PET change of 1% (indicated by the red dashed line), where a red dot represent the mean for each group; (b) scatterplot of PET changes against different baseline levels of RH for wet and dry days; and (c) AET:PET ratios on wet and dry days, simulated from AWBM, GR4J and IHACRES_CMD (in each vertical panel).

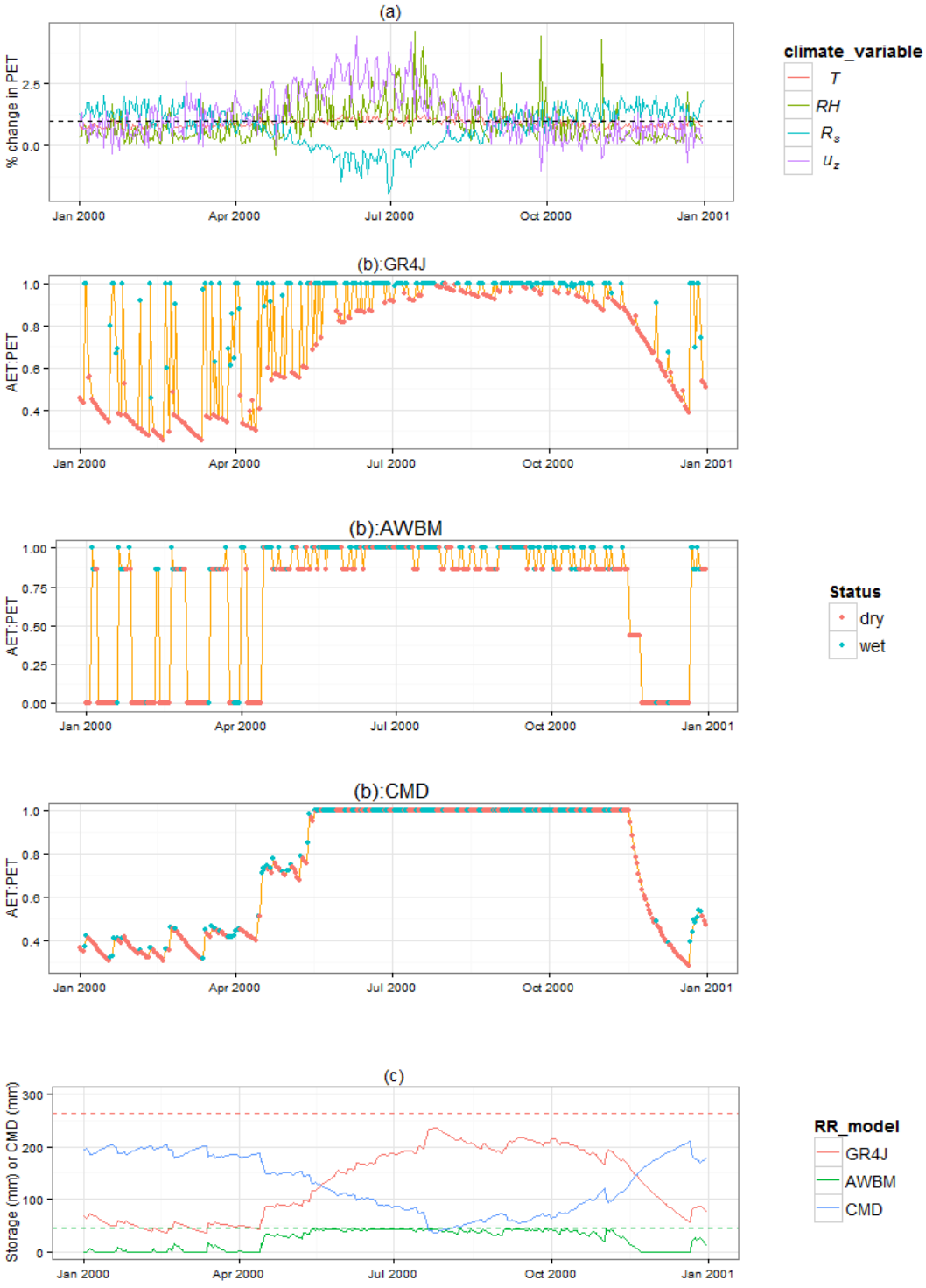


Figure 5A.10. Seasonality of a) daily PET responses to perturbations in each climate variable that result in an average PET change of 1% (black dashed line); b) AET:PET ratios for wet and dry days from GR4J, AWBM and IHACRES_CMD; and c) levels of production stores from GR4J, AWBM (with dashed lines indicating store capacities) and IHACRES_CMD (as CMD levels), at Burnie for 2000.

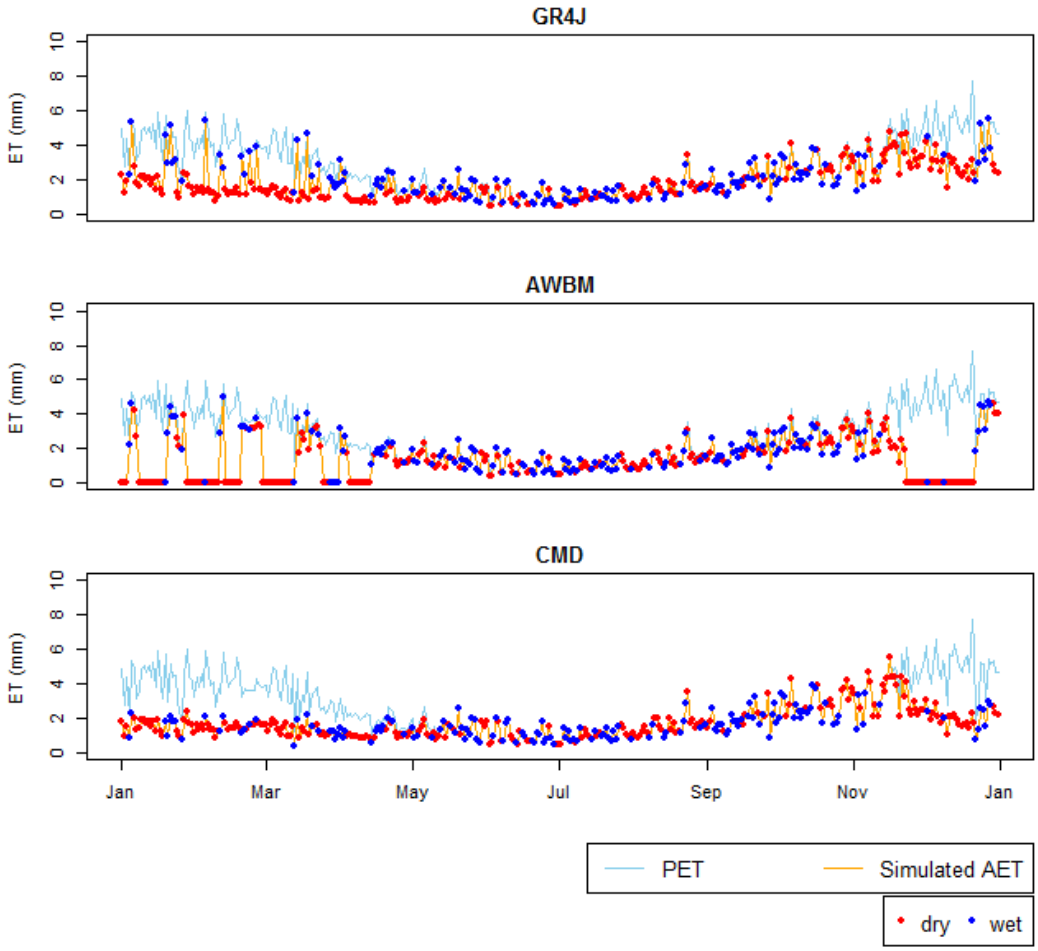


Figure 5A.11. Seasonality of the daily AET and PET for wet and dry days from GR4J, AWBM and IHACRES_CMD, at Burnie for 2000.

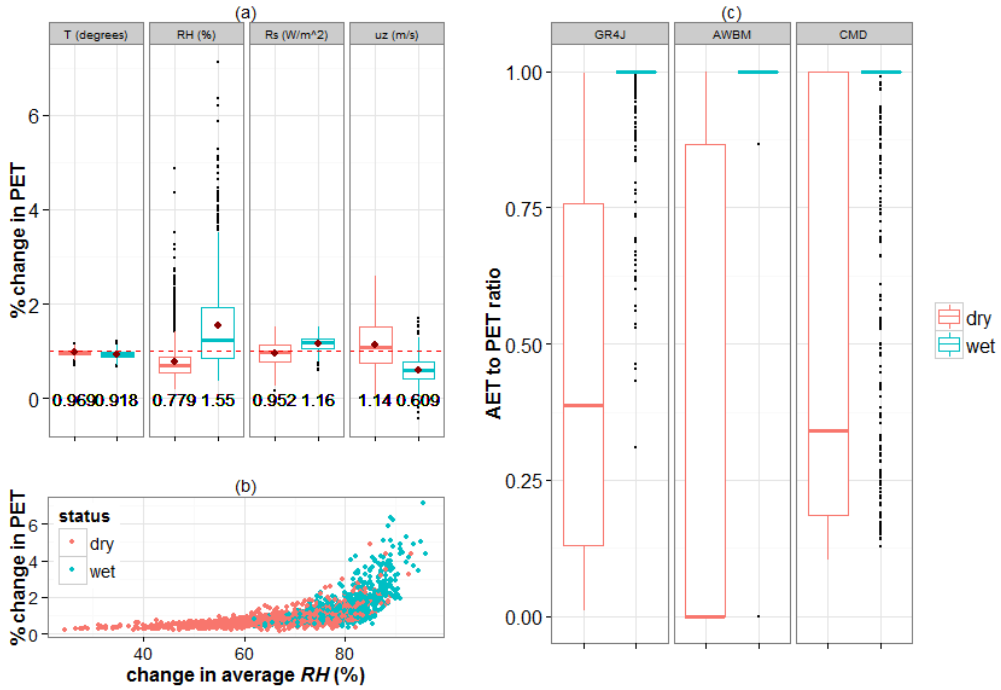


Figure 5A.12. Analyses of daily PET changes in Darwin, including: (a) Distribution of PET changes on wet and dry days, in response to perturbations in each climate variable (in each vertical panel) that result in an annual average PET change of 1% (indicated by the red dashed line), where a red dot represent the mean for each group; (b) scatterplot of PET changes against different baseline levels of RH for wet and dry days; and (c) AET:PET ratios on wet and dry days, simulated from AWBM, GR4J and IHACRES_CMD (in each vertical panel).

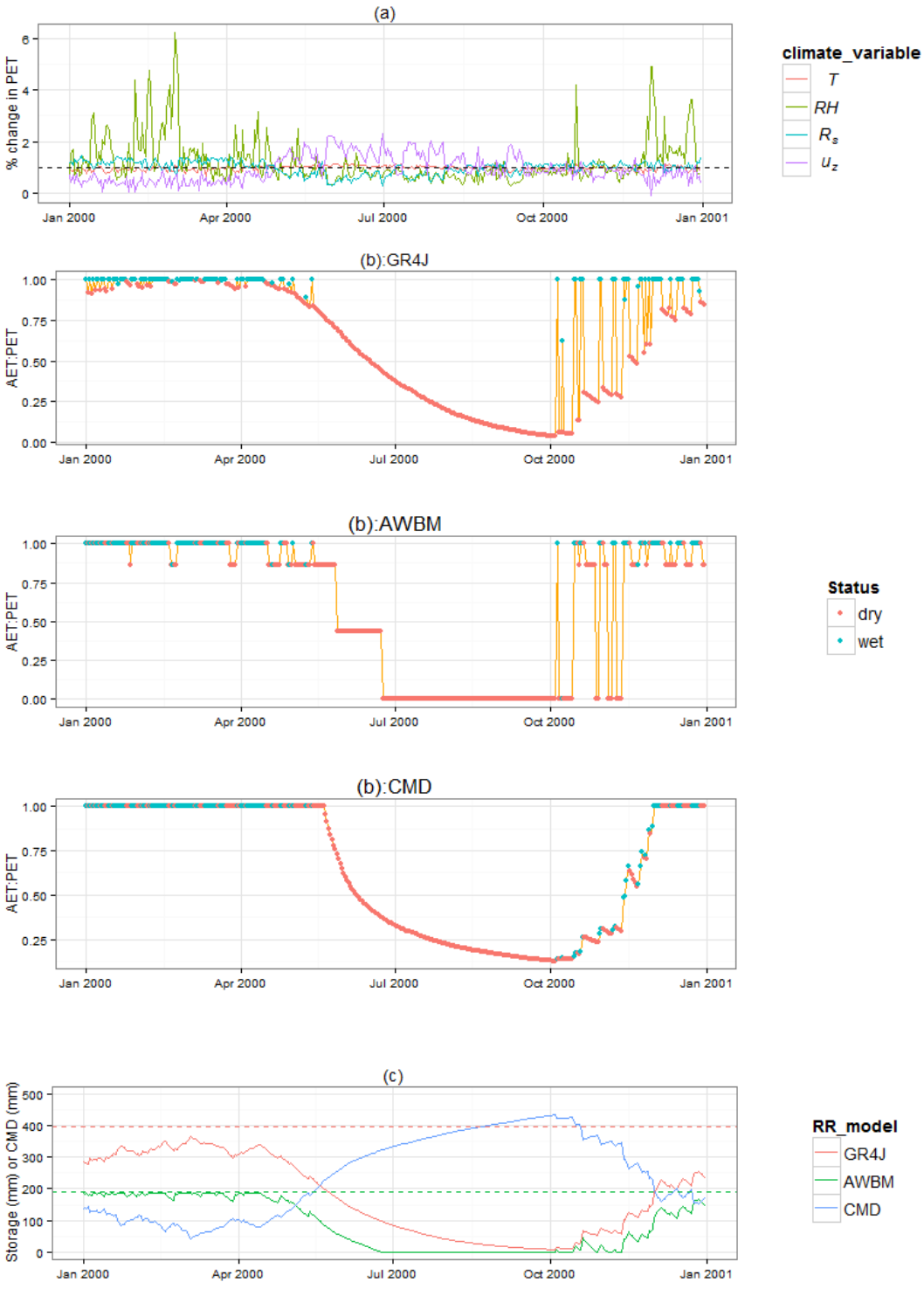


Figure 5A.13. Seasonality of a) daily PET responses to perturbations in each climate variable that result in an average PET change of 1% (black dashed line); b) AET:PET ratios for wet and dry days from GR4J, AWBM and IHACRES_CMD; and c) levels of production stores from GR4J, AWBM (with dashed lines indicating store capacities) and IHACRES_CMD (as CMD levels), at Darwin for 2000.

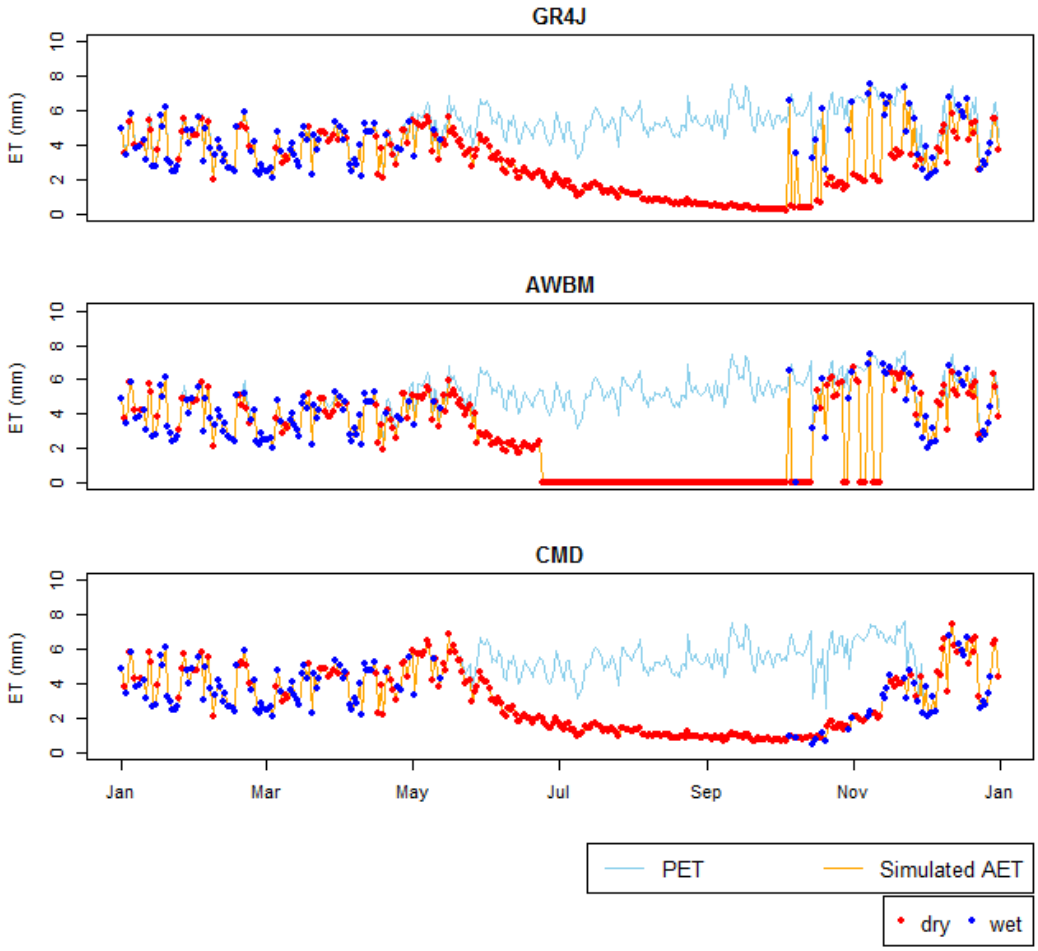


Figure 5A.14. Seasonality of the daily AET and PET for wet and dry days from GR4J, AWBM and IHACRES_CMD, at Darwin for 2000.

Wagga Wagga

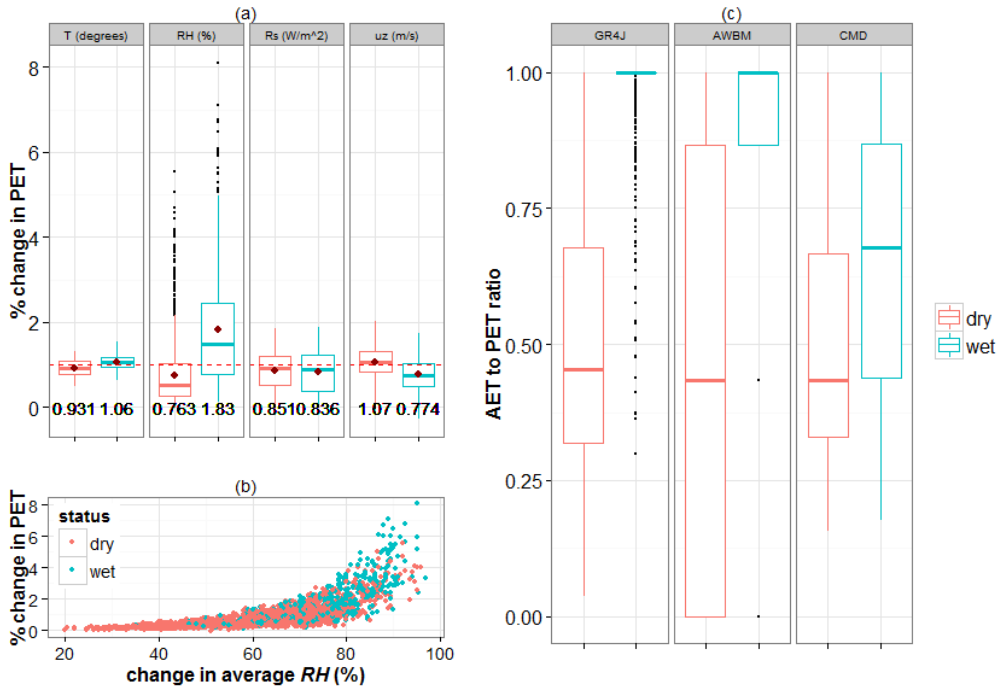


Figure 5A.15. Analyses of daily PET changes in Wagga Wagga, including: (a) Distribution of PET changes on wet and dry days, in response to perturbations in each climate variable (in each vertical panel) that result in an annual average PET change of 1% (indicated by the red dashed line), where a red dot represent the mean for each group; (b) scatterplot of PET changes against different baseline levels of RH for wet and dry days; and (c) AET:PET ratios on wet and dry days, simulated from AWBM, GR4J and IHACRES_CMD (in each vertical panel).

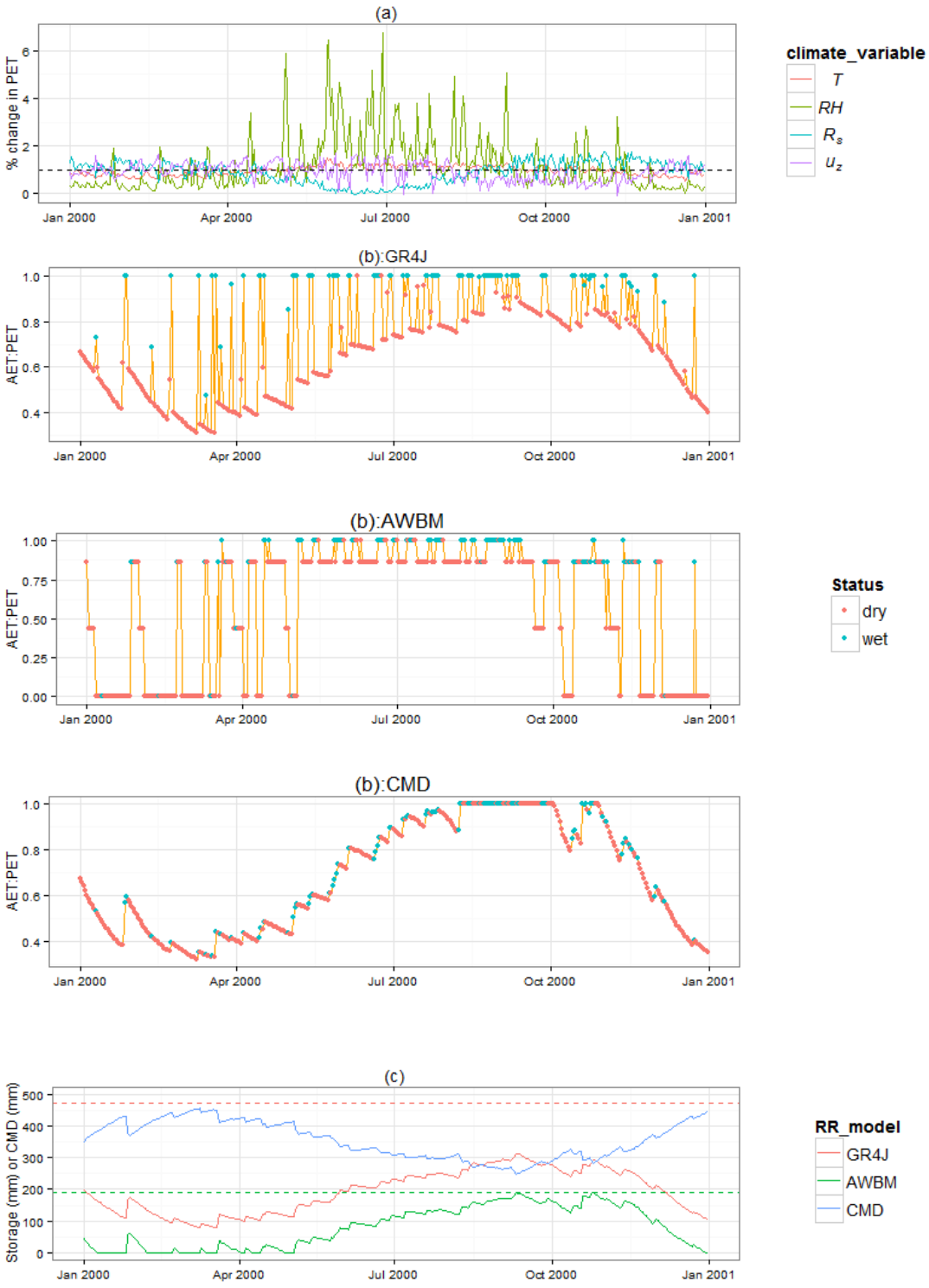


Figure 5A.16. Seasonality of a) daily PET responses to perturbations in each climate variable that result in an average PET change of 1% (black dashed line); b) AET:PET ratios for wet and dry days from GR4J, AWBM and IHACRES_CMD; and c) levels of production stores from GR4J, AWBM (with dashed lines indicating store capacities) and IHACRES_CMD (as CMD levels), at Wagga Wagga for 2000.

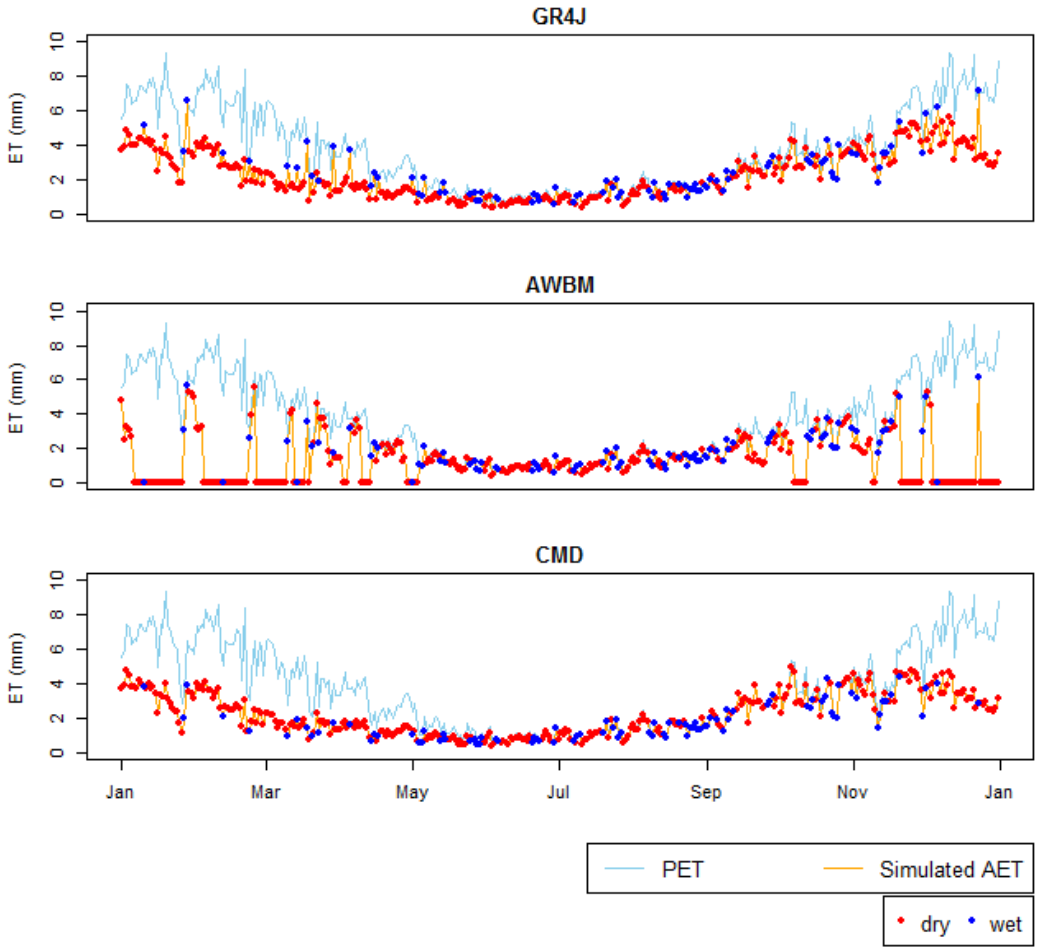


Figure 5A.17. Seasonality of the daily AET and PET for wet and dry days from GR4J, AWBM and IHACRES_CMD, at Wagga Wagga for 2000.

Appendix 5B Copy of Paper from Chapter 5

GUO, D., WESTRA, S. & MAIER, H. R. 2017. Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models. *Water Resources Research*, 53. 435-454, doi:10.1002/2016WR019627.

AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR019627

Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models

Danlu Guo [1], **Seth Westra** [1], **and Holger R. Maier** [1]

[1]School of Civil, Environmental and Mining Engineering, The University of Adelaide, North Terrace, Adelaide, South Australia, Australia

Key Points:
- Contrasting ET process representations can have substantial impact on runoff projections under a changing climate
- Conceptual rainfall-runoff models can interact with potentially complex changes to PET, causing contrasting runoff projections
- Comparing AET simulations with observations provides useful insights to evaluate the process representation within rainfall-runoff models

Supporting Information:
- Supporting Information S1

Correspondence to:
D. Guo,
Danlu.Guo@Adelaide.edu.au

Citation:
Guo, D., S. Westra, and H. R. Maier (2017), Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models, *Water Resour. Res.*, *53*, doi:10.1002/2016WR019627.

Received 9 AUG 2016
Accepted 9 DEC 2016
Accepted article online 22 DEC 2016

Abstract Conceptual rainfall-runoff models are commonly used to estimate potential changes in runoff due to climate change. The development of these models has generally focused on reproducing runoff characteristics, with less scrutiny on other important processes such as the conversion from potential evapotranspiration (PET) to actual evapotranspiration (AET). This study uses three conceptual rainfall-runoff models (GR4J, AWBM, and IHACRES_CMD) and five catchments in climatologically different regions of Australia to explore the role of ET process representation on the sensitivity of runoff to plausible future changes in PET. The changes in PET were simulated using the Penman-Monteith model and by perturbing each of the driving variables (temperature, solar radiation, humidity, and wind) separately. Surprisingly, the results showed the potential of a more than sevenfold difference in runoff sensitivity per unit change in annual average PET, depending on both the rainfall-runoff model and the climate variable used to perturb PET. These differences were largely due to different ways used to convert PET to AET in the conceptual rainfall-runoff models, with particular dependencies on the daily wet/dry status, as well as the seasonal variations in store levels. By comparing the temporal patterns in simulated AET with eddy-covariance-based observations at two of the study locations, we highlighted some unrealistic behavior in the simulated AET from AWBM. Such process-based evaluations are useful for scrutinizing the representation of physical processes in alternative conceptual rainfall-runoff models, which can be particularly useful for selecting models for projecting runoff under a changing climate.

1. Introduction

Climate change is expected to have significant implications on catchment-scale water resources, potentially affecting water security for municipal, agricultural, and industrial applications, environmental water quantity and quality, and flood hazard [*CSIRO and Bureau of Meteorology*, 2015; *Hauser et al.*, 2009; *IPCC*, 2014; *Turral et al.*, 2011]. To assess these implications, rainfall-runoff models are commonly used to translate projected changes in atmospheric variables derived from large-scale general circulation models into regional or local runoff [e.g., see *Akhtar et al.*, 2008; *Chiew et al.*, 2009]. This information can then be used to assess catchment yield [e.g., *Haque et al.*, 2015], water quality [e.g., *Crossman et al.*, 2013], water supply security [e.g., *Christensen et al.*, 2004; *Paton et al.*, 2013, 2014], and flood risk [e.g., *Kay and Jones*, 2012], amongst other variables.

Although various rainfall-runoff model classes have been developed with different levels of physical realism [*Beven*, 2011], most climate impact studies are based on the outputs of one or several calibrated conceptual rainfall-runoff models [*Chiew et al.*, 2010; *Islam et al.*, 2014; *Najafi and Moradkhani*, 2015; *Vaze and Teng*, 2011]. The benefit of using conceptual rainfall-runoff models is that they are parsimonious and tend to perform well when calibrated to historical runoff data [*Perrin et al.*, 2003; *Wagener et al.*, 2003], particularly when compared to more complex physically based and spatially distributed models, for which parameter estimation can be extremely challenging [*Beven*, 2001b; *Butts et al.*, 2004; *Goderniaux et al.*, 2009]. However, since the performance of these conceptual rainfall-runoff models is largely assessed on runoff, in many cases the underlying physical processes such as evapotranspiration fluxes, runoff generation mechanisms, and so on can be poorly represented [*Li et al.*, 2015; *Seibert and McDonnell*, 2002; *Seibert et al.*, 2003], with parameters being "falsely adjusted" to maximize runoff performance without considering the realism of the internal physical processes [*Andréassian et al.*, 2004; *Beven*, 2001a; *Clark et al.*, 2016; *Ewen et al.*, 2006; *Minville et al.*, 2014].

The surprisingly limited impact of physical process realism on the performance of conceptual rainfall-runoff models in simulating historical streamflow has been observed in multiple studies [*Brouyère et al*., 2004; *Hartmann and Bárdossy*, 2005; *Kirchner*, 2006; *Vaze et al*., 2010]. This result has also been found in the specific context of evapotranspiration (ET) process representation, and reasonable performance of conceptual models has been obtained with the use of simplified methods for estimating potential ET (PET) that do not realistically consider the dynamics in all its driving climate variables [*Oudin et al*., 2005], or do not sufficiently represent the spatial and temporal variations in ET [*Andréassian et al*., 2004; *Chapman*, 2003]. Furthermore, reasonable conceptual rainfall-runoff model performance has been observed even when the ET input contains systematic and/or random errors [*Oudin et al*., 2006]. A potential reason for these findings is that the calibration process of conceptual models can "compensate" for reasonable levels of differences in ET estimates without significant impact on model performance in simulating runoff [*Andréassian et al*., 2004; *McMahon*, 2015]. However, although such calibrated conceptual models may perform well when applied under conditions similar to the period used for calibration, the extent to which they can be applied to climatologically different conditions is less clear [*Coron et al*., 2012; *Klemeš*, 1986; *Wagener et al*., 2003; *Westra et al*., 2014], as in this context it becomes much more critical to ensure that the model "gets the right answers for the right reasons" [*Kirchner*, 2006]. This is highlighted in a number of recent studies that found that although conceptual models with different structures can perform similarly when simulating historical runoff, they can produce very different estimates under a changing climate [*Bae et al*., 2011; *Bastola et al*., 2011; *Jiang et al*., 2007; *Mendoza et al*., 2016; *Velázquez et al*., 2013].

Although the above studies suggest potential limitations in the performance of conceptual rainfall-runoff models under future climate, the specific role of ET representation within these models on runoff projections has not been widely investigated. Changes to ET-related processes are likely to be particularly important to overall catchment response, as actual ET (AET) volumes exceed runoff volumes for the majority of catchments worldwide [*Dingman*, 2015]. Furthermore, the potential changes in the climatic variables that can influence ET (e.g., temperature, solar radiation, humidity, and wind) [*IPCC*, 2014] suggest that changes to ET-related processes could be extremely complex, with changes not only at annual timescales but also in terms of the daily and seasonal distributions of ET (as illustrated in *Gong et al*. [2006], *Huo et al*. [2013], and *Vicente-Serrano et al*. [2014]). However, the extent to which these potentially complex changes to ET can interact with the structure of conceptual rainfall-runoff models remains unknown, with these changes usually represented using seasonal or annual ET change factors in many climate impact studies [e.g., *Chiew et al*., 2009; *New et al*., 2007; *Teng et al*., 2012].

The focus of this study is therefore to assess the impact of alternative ET process representations within conceptual rainfall-runoff models on runoff projections under perturbed climate conditions. This is achieved by focusing specifically on the following research questions:

1. To what extent can ET process representations within conceptual rainfall-runoff models impact runoff projections under plausible changes in the climate drivers of ET?
2. How do alternative conceptual rainfall-runoff model structures interact with the potentially complex changes to PET to produce projections of runoff and AET?
3. Is it possible to use AET observations to constrain which ET process representations within conceptual rainfall-runoff models are more realistic?

The above research questions were addressed with five case study locations from distinct climatic regions within Australia. Three lumped conceptual rainfall-runoff models—GR4J, AWBM, and IHACRES_CMD—were used to represent contrasting ET process representations, and thus to illustrate how potential changes in PET can propagate differently to runoff projections. PET was required as an input to all three rainfall-runoff models, and was estimated using the Penman-Monteith method. This method is considered to be the most comprehensive physically based PET model and is thus widely used for rainfall-runoff modeling and climate impact assessments [e.g., *Arnell*, 2004; *Gosling et al*., 2011; *Kay et al*., 2009]. Use of the Penman-Monteith method enables an assessment of the effect of perturbing each of the driving climatic variables of PET (i.e., temperature, solar radiation, wind, and humidity), and can provide information not only on annual average changes, but also on subannual (e.g., daily and seasonal) variations in PET sensitivity.

The remainder of this paper is structured as follows. The data used to address the three research questions are provided in the next section, together with an overview of the Penman-Monteith model and the three

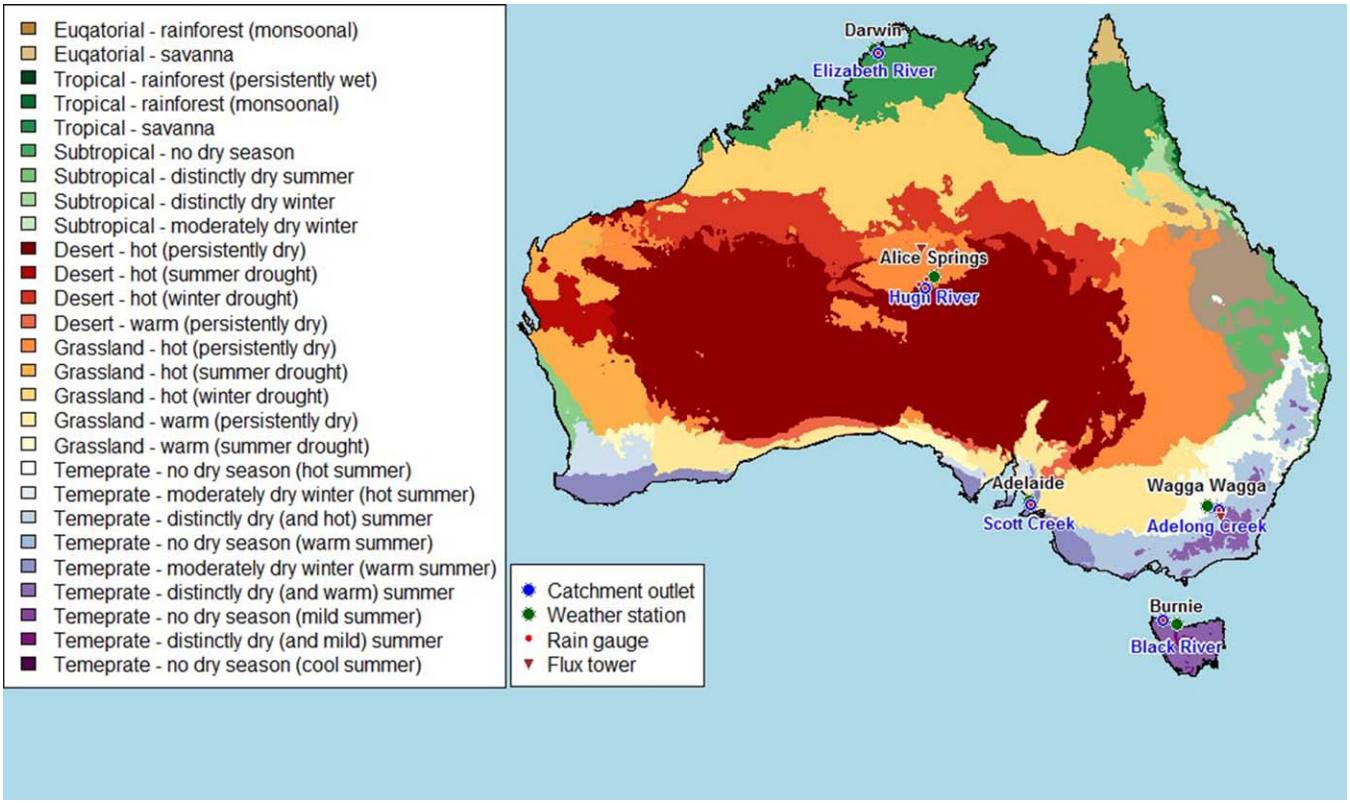


Figure 1. Locations for rain gauges, catchment outlets, and weather stations from which data were obtained for calibration of rainfall-runoff models at the five case study catchments. The locations of two flux towers to obtain data for model evaluation at Alice Springs and Wagga Wagga are also shown. Coloring relates to Köppen climate classifications from *Stern et al*. [2000].

conceptual rainfall-runoff models. This is followed by a description of the methods employed to address each research question in section 3. The results and their implications are discussed in section 4, and a summary and conclusions are given in section 5.

2. Data and Models

2.1. Data

Figure 1 shows the names and locations of the five case study catchments, which are situated in climatologically different regions in Australia, as defined in the Australian Köppen climate classifications of *Stern et al*. [2000]. As mentioned in section 1, three lumped conceptual rainfall-runoff models were used in this study, which required historical data for calibration. These data include catchment-average rainfall, PET, and runoff at each location. The PET input data to these models were estimated using the Penman-Monteith model, which required data on temperature, relative humidity, solar radiation, and wind speed.

The source of each data variable is summarized below, with further details given in supporting information Table S1:

1. Catchment runoff (ML/day): Daily runoff data were obtained from the gauging stations at the outlet of each catchment.
2. Catchment-average rainfall (mm/day): In order to represent the catchment-average rainfall, daily rainfall data were obtained from a rain gauge within each of the four smaller catchments (Scott Creek, Black River, Elizabeth River, and Adelong Creek). For the largest catchment (Hugh River), data from three rainfall gauges within the catchment were spatially averaged using the Thiessen polygon method.
3. Daily maximum and minimum temperature (T_{max} *and* T_{min} in °C), maximum and minimum relative humidity (RH_{max} *and* RH_{min} in %), and wind speed (u_z in m/s): Due to the limited availability of high-

Table 1. Average Climate Conditions at the Five Case Study Locations Between 1 January 1995 and 31 December 2003

Weather Station	T_{max} (°C)	T_{min} (°C)	RH_{max} (%)	RH_{min} (%)	Rs (MJ/(m^2.day))	u_z (m/s)	Annual PET (mm)	Catchment (and Area)	Annual P (mm)	Annual Q (mm)	PET/P
Adelaide	21.4	12.5	80.2	42.2	16.9	3.16	1372	Scott Creek (29 km^2)	892	133	1.54
Burnie	15.7	12.7	77.1	65.3	14.3	4.05	958	Black River (318.5 km^2)	1182	550	0.81
Darwin	31.2	23.7	88.5	50.1	20.3	3.39	1864	Elizabeth River (95.6 km^2)	1979	777	0.94
Alice Springs	28.5	13.8	65.6	23.5	20.8	2.35	1822	Hugh River (3324 km^2)	344	56.2	5.29
Wagga Wagga	21.6	9.99	83.2	40.3	17.5	3.29	1436	Adelong Creek (146.1 km^2)	799	195	1.80

[a]Note: Notation: T_{max} = maximum temperature; T_{min} = minimum temperature, R_s = incoming solar radiation, RH_{max} = maximum relative humidity, RH_{min} = minimum relative humidity, u_z = wind speed, PET = potential evapotranspiration calculated using the Penman-Monteith model, P = catchment-average precipitation, Q = average runoff at catchment outlet.

quality climate observations, data for each of these variables were obtained from a weather station outside but nearby each catchment.

4. Daily solar radiation (R_s in MJ/(m^2 day)): Daily solar radiation was calculated from daily sunshine hour data (n in h) obtained from each weather station, using the Ångström-Prescott equation [*McMahon et al*., 2013] with constants for each location provided in *Chiew and McMahon* [1991].

Data for each variable were obtained for a consistent period from 1 January 1995 to 31 December 2003 at each location, with key statistics presented in Table 1. As can be seen, there are large differences in the average values of each variable, highlighting the potentially significant differences in dominant physical processes for ET and runoff production across these catchments. A quantity particularly relevant to ET processes is the long-term averaged ratio of PET to precipitation (PET/P), which describes whether a catchment is water-limited (PET/P > 1) or energy-limited (PET/P < 1). The range of PET/P values indicates substantial variations in the water availability conditions at the five study sites.

To determine the relative realism of different ET process representations, actual measurements of AET were used to benchmark the simulated AET from the three rainfall-runoff models. A challenge, however, is the limited availability of AET data for Australia. Lysimeter measurements are currently considered as the most accurate source of AET observations [*Seneviratne et al*., 2012; *Wang and Dickinson*, 2012]; however, these data are most commonly available in irrigated regions [e.g., *Bethune et al*., 2001; *Northey et al*., 2006; *Zhang et al*., 1999], with measurements from natural catchments being very scarce. The MODIS Land Product (http://daac.ornl.gov/MODIS/modis.shtml) also provides estimates of AET based on remotely sensed observations [*Cleugh et al*., 2007; *Mu et al*., 2011] with a much wider spatial coverage; however, the AET data are only available on an 8 day interval.

The above constraints led us to select eddy covariance (EC)-based AET measurements, which are available through the OzFlux [*TERN*, 2012] network at 30 active sites across Australia (see: http://www.ozflux.org.au/). For our purpose of evaluating alternative ET representations, the AET observations should be obtained from nearby our case study catchments, as well as spanning a sufficient continuous time period. As a result of these constraints, only two OzFlux sites were selected, namely Alice Springs Mulga, which is close to the Hugh River catchment, and Tumbarumba, which is close to the Adelong Creek catchment [*Cleverly*, 2011; *vanGorsel*, 2013]. Based on the EC data availability, the comparison between modeled and observed AET used a study period of between 3 September 2010 and 31 July 2014 for both catchments. The corresponding climatic and rainfall data within this period were also obtained from both flux sites to ensure internal consistency in the rainfall-runoff models.

2.2. Models

2.2.1. Penman-Monteith PET Model

The Penman-Monteith model was adopted throughout this study, which combines the energy balance and mass transfer components of ET. To minimize the potential confounding effects of differences in vegetated surface, the evaporative surface was assumed to be reference crop for all study sites, for which the FAO-56 version of the Penman-Monteith model [*Allen et al*., 1998] was used. The FAO-56 Penman-Monteith model was implemented using the R package "Evapotranspiration" (http://cran.r-project.org/web/packages/Evapotranspiration/index.html) [*Guo et al*., 2016a].

2.2.2. Three Conceptual Rainfall-Runoff Models

Contrasting ET process representations were considered by using three rainfall-runoff models (GR4J, AWBM, and IHACRES_CMD), which all run on a daily time step. The three models are all lumped conceptual models, although with very different soil moisture accounting (SMA) routines, which can affect the partitioning of precipitation into AET and runoff. To isolate the effect of different conceptual ET representations in the three models, the groundwater exchange was set to zero for each model. Furthermore, the routing components for all models were constrained to the GR4J routing model, so that differences between models can be attributed specifically to the SMA routines.

The first model, GR4J [*Perrin et al*., 2003], is a conceptual hydrologic model based on an SMA production store and an excess-rainfall routing store. The model requires two input variables, PET and precipitation (P). Interception is treated as a store with zero capacity, so that each day can be either "wet" (P > PET), which produces a net precipitation P_n = P-PET, or "dry" (P < PET) with a net evapotranspiration E_n = PET-P. For wet days, a portion of the net precipitation P_n fills the production store. This portion is termed P_s and is a function of both P_n and the store level S relative to its capacity defined by parameter x_1. Runoff is produced from the total water available for routing, P_r, which is generated by the remaining net precipitation ($P_n - P_s$). The conversion of PET to AET from the production store is illustrated in Figure 2a, which relies on the wet/dry status of a day. For wet days, AET is set to be equal to PET. For dry days, AET is modeled as the sum of the precipitation and the evapotranspiration from store E_s, which is modeled as a fraction of the net evapotranspiration E_n depending on the water level in the production store S relative to x_1. Consequently, AET is always less than PET on dry days.

The schematic of AWBM is shown in Figure 2b. In contrast to GR4J, the production store in AWBM [*Boughton*, 2004] is represented by three individual stores (termed S_1, S_2, and S_3) with different capacities, and each store occupies a fixed fraction of the total catchment area, with the default values of 0.134, 0.433, and 0.433 used in this study. For each day, precipitation P is added directly to each store without explicit consideration of interception. For each nonempty store, available runoff P_r is produced from any excess when the store is full, while evapotranspiration from the store, E_s, always equals PET until the store is completely empty [*Boughton*, 2009]. Therefore, the overall AET, as the sum of E_s from all three stores, equals different proportions of PET, depending on which stores contain water. Specifically, the AET as a proportion of PET can be one of: 0 (when all three stores are empty), 0.134 (when both of S_2 and S_3 are empty), 0.433 (when both of S_1 and S_2 or both of S_1 and S_3 are empty), 0.567 (when only one of S_2 or S_3 is empty), 0.866 (when only S_1 is empty), or 1 (when all stores are nonempty). The level for each store S_i (i = 1, 2, 3), after accounting for P, PET, and P_r from the specific store, equals S_i + P − PET − $P_{r,\,i}$. Therefore, on days with P − PET > 0, all stores are filled up (i.e., with increase in store levels), resulting in the overall AET = PET; when P − PET < 0, AET is always less than PET as determined by number of empty stores.

In IHACRES_CMD [*Croke and Jakeman*, 2004], the level of the store is represented by the catchment moisture deficit (CMD), which is the difference between the current level and the saturation level of the store. The schematic of the model is shown in Figure 2c. For each day, all precipitation P directly fills the store without explicit consideration of interception. The total water available for routing, P_r, is a proportion of P, which decreases with increasing ratio of *CMD* relative to a flow-production threshold parameter *d*. AET is taken directly from the store (i.e., E_s = AET), which is estimated as a fraction of PET depending on whether *CMD* exceeds a threshold parameter *g*. When *CMD* > *g*, the fraction of AET to PET is always less than one, which decreases with increasing ratio of *CMD* to *g*, and approaches an asymptote of 0. In contrast, when *CMD* < *g*, AET = PET. Therefore, the conversion from PET to AET is independent of direct input of P, other than through its effect on changing the store levels.

The above rainfall-runoff models therefore provide examples of three contrasting methods of converting PET input data to AET: GR4J explicitly considers both the impact from store levels and the wet/dry status of a day, defined by individual rainfall events; AWBM simulates AET mostly as a function of rainfall which determines the wet/dry status, while the impact of instantaneous store levels is largely eliminated by the use of step functions to relate AET with the emptiness of each store; finally, IHACRES_CMD simulates AET as a function of only the store level and thus is relatively independent of rainfall for that day.

To simulate runoff, the PET data estimated for each case study were used as input to all three SMA models linked with the GR4J-routing model (implemented using the R package "hydromad," available at: http://

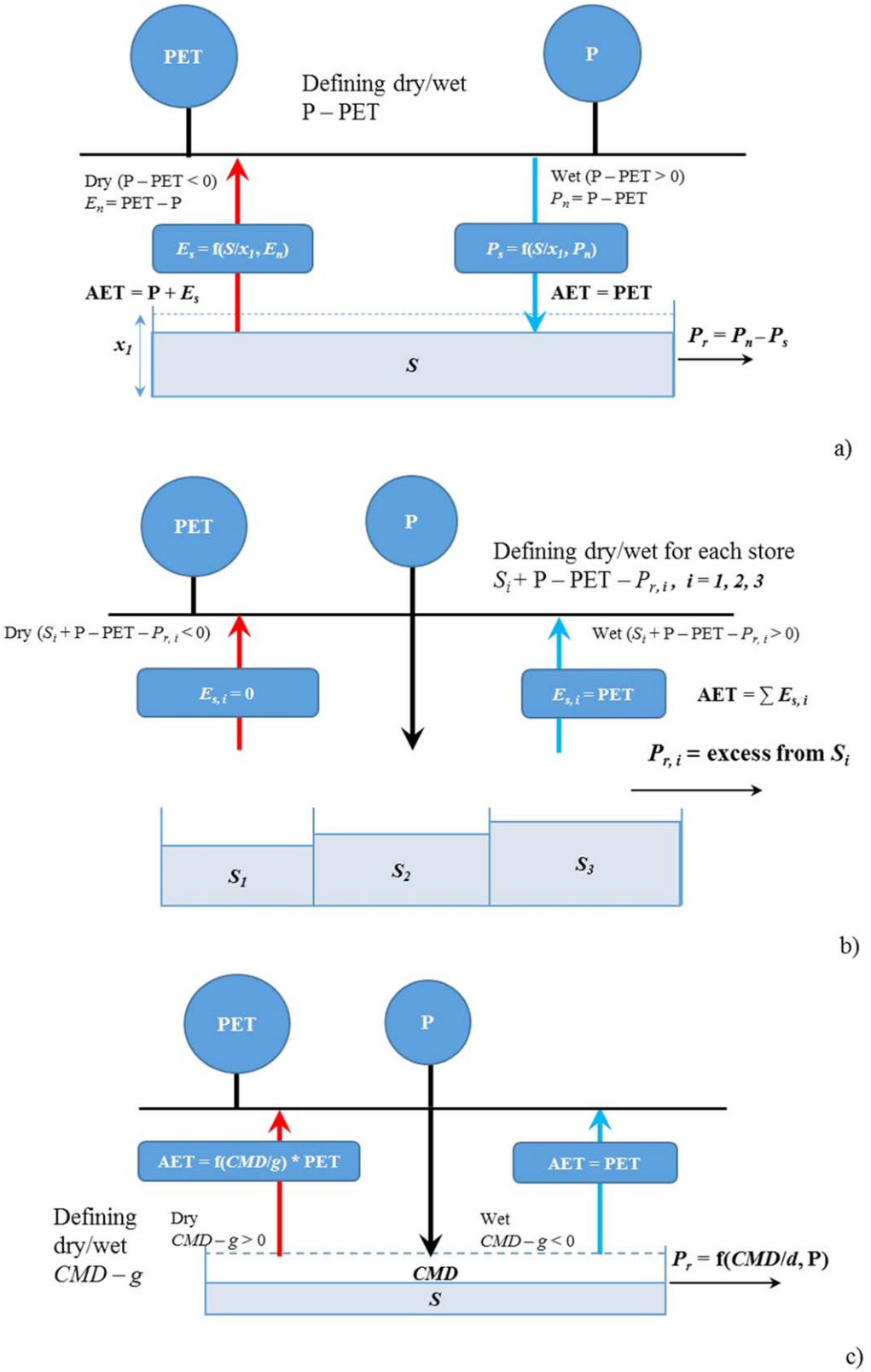


Figure 2. SMA routines in: (a) GR4J, (b) AWBM, and (c) IHACRES_CMD (adapted from *Perrin et al.* [2005], *Boughton* [2004], and *Croke and Jakeman* [2004], respectively). The red arrows represent dry days and the blue arrows represent wet days in each model.

hydromad.catchment.org/[*Andrews and Guillaume*, 2013]). For each case study, the three models were calibrated to the first 6 years of historical runoff data (from 1 January 1995 to 31 December 2000) at a daily timescale, with the first year (1995) used as a warm-up period. The remaining 3 years of data (from 1 January 2001 to 31 December 2003) were used for validation_ENREF_40. The objective function for calibration was the Nash-Sutcliffe coefficient of efficiency (NSE), which has been used widely in rainfall-runoff modeling. In addition, the relative bias was also calculated to assess the bias in modeled runoff relative to observations.

The calibration and validation results are shown in Table 2. For each model, the calibration and validation performance varies by catchment, and the validation performance is on average slightly lower than the calibration performance. For each catchment, the values of each performance metric are generally similar across the three rainfall-runoff models, although GR4J on average has the best performance and AWBM has the worst performance for the case study catchments.

3. Method

A schematic of the methodology is presented in Figure 3. The approach to addressing the three research questions described in section 1 is detailed in sections 3.1–3.3, respectively.

3.1. Impact of ET Process Representations on Runoff Projections

To assess the impact of different ET process representations within conceptual rainfall-runoff models, we explored the sensitivity of both PET and runoff to changes in four climate variables related to PET, with differences in the ratio of sensitivity values between PET and runoff providing an indicator of the importance of ET process representation. The sensitivity curve method was employed to achieve this, as it has been widely applied in previous PET sensitivity studies due to its computational convenience and ease of interpretation [*Goyal*, 2004; *McKenney and Rosenberg*, 1993; *Vicente-Serrano et al*., 2014].

As required by the sensitivity curve method, perturbations were made to the four climatic variables which influence PET, namely temperature (T), relative humidity (RH), solar radiation (R_s), and wind speed (u_z), at each of the five study sites. This process yields multiple sets of perturbed climate data, each obtained by perturbing one climate variable while keeping the others unchanged. It is worth noting that by perturbing each input variable independently, the sensitivity curve method does not take into account the potential joint variations among the input variables due to correlations. However, according to a previous global sensitivity study across multiple locations in Australia, PET generally showed little sensitivity to these joint variations among T, RH, R_s, and u_z (D. Guo et al., Sensitivity of potential evapotranspiration to changes in climate variables for different climatic zones, *Hydrology and Earth System Science Discussions*, doi:10.5194/hess-2016-441, in review, 2016), so that it is reasonable to assume that this one-at-a-time climate perturbation is capable of capturing the majority of potential changes in PET due to the four climatic variables.

The bounds for perturbing each variable were selected to be slightly wider than the ranges of the projected changes in these variables by 2100 for Australia [*Stocker et al*., 2013] to encompass a comprehensive range of plausible future climate change scenarios (Table 3).

Table 2. NSE for Calibration and Validation for the GR4J, AWBM, and IHACRES_CMD Models at the Five Case Study Sites, With Relative Bias Shown in Brackets

	SMA Model Name					
	GR4J		AWBM		IHACRES_CMD	
Study Site	Calibration	Validation	Calibration	Validation	Calibration	Validation
Adelaide	0.861 (0.011)	0.800 (0.025)	0.812 (−0.274)	0.804 (0.292)	0.856 (0.020)	0.809 (−0.028)
Burnie	0.862 (−0.010)	0.869 (−0.053)	0.799 (−0.089)	0.757 (0.018)	0.860 (0.008)	0.868 (−0.070)
Darwin	0.849 (0.052)	0.834 (−0.215)	0.825 (0.013)	0.823 (−0.148)	0.848 (0.048)	0.848 (−0.226)
Alice Springs	0.805 (−0.650)	0.581 (0.034)	0.778 (0.599)	0.569 (−0.737)	0.800 (−0.503)	0.585 (−0.912)
Wagga Wagga	0.748 (−0.072)	0.695 (0.083)	0.513 (−0.096)	0.618 (−0.298)	0.713 (−0.085)	0.642 (0.083)
Average NSE	0.825	0.756	0.746	0.714	0.815	0.750

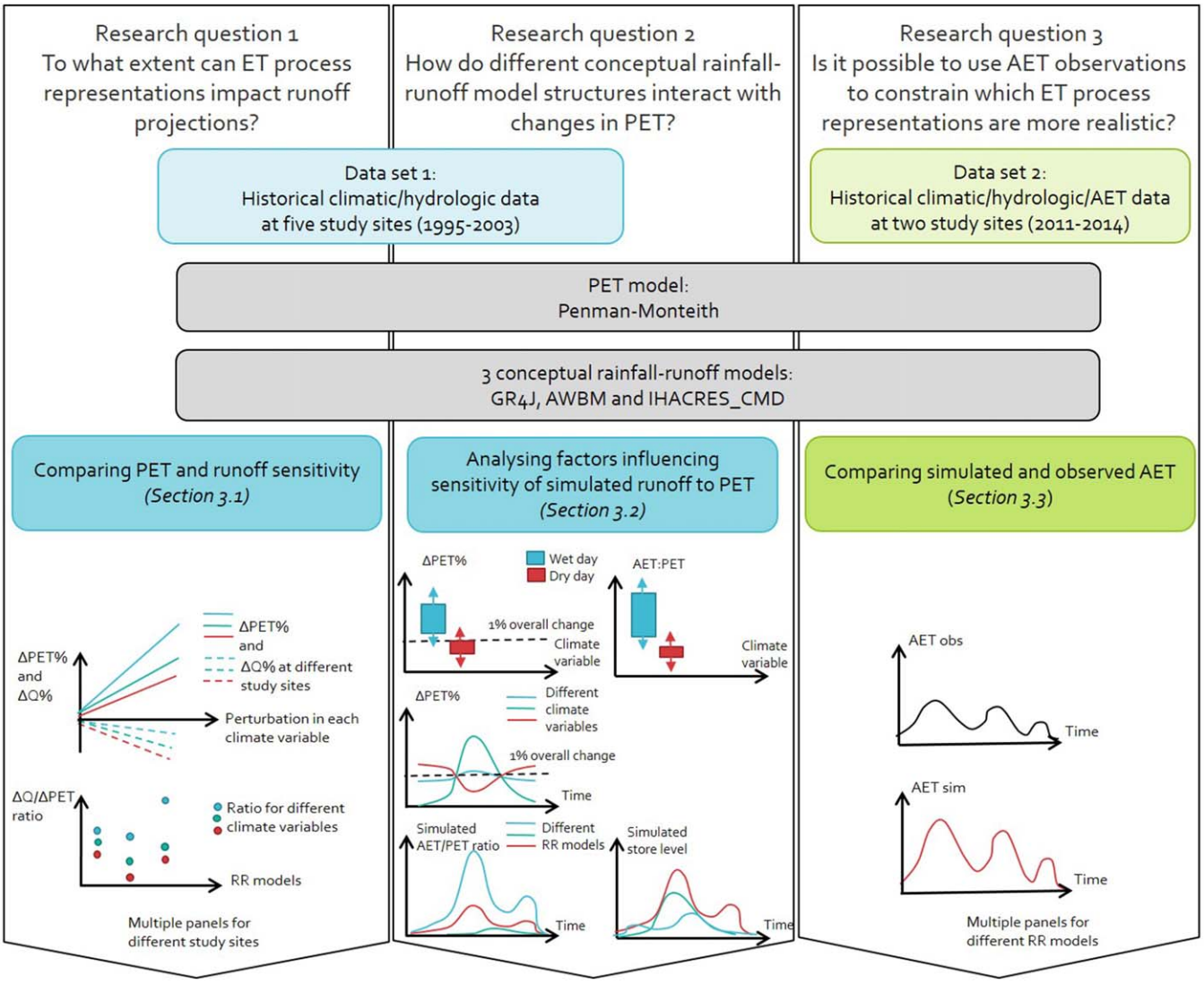


Figure 3. Schematic of methodology used to address the research questions posed in section 1.

For each climate variable, the perturbation levels relative to its historical baseline consisted of four equidistant levels within the above-mentioned plausible ranges (i.e., +2°C, +4°C, +6°C, and +8°C for *T*, 10%, −5%, +5%, and +10% for *RH* and R_s, and −20%, −10%, +10%, and +20% for u_z), with each perturbation level being applied as a factor to the entire time series of the corresponding daily historical data. For *T* and *RH*, the Penman-Monteith model requires both the daily minimum and maximum values as input; therefore, the daily time series of each pair of *T* and *RH* variables was considered jointly and thus perturbed by the same amount for each day. *Since all perturbations for T were additive, whereas perturbations for R_s, RH, and u_z were percentage changes, positive values of these variables were maintained during the perturbation. The perturbed RH data were also capped at 100% to avoid obtaining physically unrealistic humidity values.* The sensitivity of PET and runoff to different levels of perturbation in each of *T*, *RH*, R_s, and u_z were then represented as the relative changes to their corresponding baseline historical data, as used in model calibration.

Table 3. Plausible Perturbation Ranges for Each Climate Variable Relative to Their Historical Levels

Climate Variable to Perturb	Perturbation Range
T	0 to +8°C
RH	−10% to +10%
R_s	−10% to +10%
u_z	−20% to +20%

To compare the PET sensitivity with the runoff sensitivity from the three rainfall-runoff models, the ratio of the percentage change in annual average runoff for each percentage change in annual average PET—referred to as the runoff "elasticity" to changes in PET [*Chiew*, 2006]—was also estimated.

3.2. The Role of Rainfall-Runoff Model Structure on Sensitivity to PET

The focus of this section is to better understand the effect of rainfall-runoff model structure on the sensitivity of runoff to changes in PET. As discussed in section 2.2.2, the key structural differences of three rainfall-runoff models in the context of translating PET to AET and runoff are their manner of: (a) converting PET to AET on wet days and dry days and (b) relating AET to the level of the soil moisture store(s). Therefore, to understand the causes of the impact of ET process representations on runoff projections, we investigated the impact of the following factors on the sensitivity of simulated runoff to PET:

1. Wet-day and dry-day patterns, to assess how rainfall-runoff models handle conversion from PET to AET on wet and dry days. Although the three rainfall-runoff models use different definitions of wet and dry days to determine how much of PET is converted to AET, for consistency we used a single definition of a wet day as one with rainfall greater than 1 mm for all models. The conversion from PET to AET was represented by the AET:PET ratio, which should be between 0 and 1 by definition. We compared the wet-/dry-day distributions of: (a) the daily PET responses to perturbations in each climate variable and (b) the daily AET:PET ratios simulated from each rainfall-runoff model. We also investigated whether there were interactions between them that can lead to different AET estimates for the same change in average PET.
2. Seasonal patterns, to assess how rainfall-runoff models relate AET to temporal variations in storage. Depending on the catchment, there is evidence that the production store can vary significantly by season [e.g., see *Westra et al*., 2014]. Therefore, we investigated the role of seasonality in storage on the simulated AET from rainfall-runoff models. We also investigated whether the responses of PET to each perturbed climate variable can vary seasonally and thus act synergistically or antagonistically to produce different AET simulated for the same total change in annual average PET. This was achieved by investigating the seasonal variations in: (a) the daily response of PET to perturbations in each individual climate variable; (b) the daily AET:PET ratios simulated in the three rainfall-runoff models; and (c) the daily storage levels simulated in the three rainfall-runoff models.

To provide a common basis for comparison, all the sensitivity results in this section were presented per 1% change in PET.

3.3. Relative Realism of Alternative ET Process Representations Within Conceptual Rainfall-Runoff Models

To evaluate the relative realism of ET process representations within the three conceptual rainfall-runoff models, the AET simulated with the three models were compared with EC measurements of AET at Alice Springs and Wagga Wagga, as these were the only sites at which measured AET data were available. As a result of limitations in AET measurements at these sites (as discussed in section 2.1), a qualitative approach was adopted for the comparison, in which the time series of observed and simulated AET were visually compared to evaluate each model's ability to represent the temporal variation of ET processes.

4. Results and Discussion

4.1. Impact of ET Process Representations on Runoff Projections

The sensitivity of PET and runoff to changes in each climate variable is shown in Figure 4, with each plot representing the sensitivity to a different climate variable (see supporting information Figures S1 and S2 for individual plots of PET and runoff sensitivity). The results show a large range of sensitivity values depending on (a) the perturbed climate variable, (b) the rainfall-runoff model, and (c) the location. For example, PET is generally more sensitive to perturbations in T, with an 8°C increase in T leading to between a 20% and 32% increase in average PET depending on location (Figure 4a). RH shows a negative relationship with PET, with a 10% increase in average RH leading to between a 2% and 13% decrease in average PET (Figure 4b). Sensitivity to R_s and u_z is consistently much smaller, with the increase in PET being between 3% and 5% for a 10% increase in average R_s (Figure 4c) and between 3% and 6% for a 20% increase in average u_z (Figure 4d).

Also apparent from Figure 4 is that for a fixed climate perturbation, the runoff sensitivity always has an opposite sign to that of PET. However, the relative order of impact that the four climate variables have on runoff is consistent with that for PET. In particular, runoff generally shows a higher sensitivity to perturbations in T, for which an 8°C increase can cause between a 7% and 31% decrease in runoff

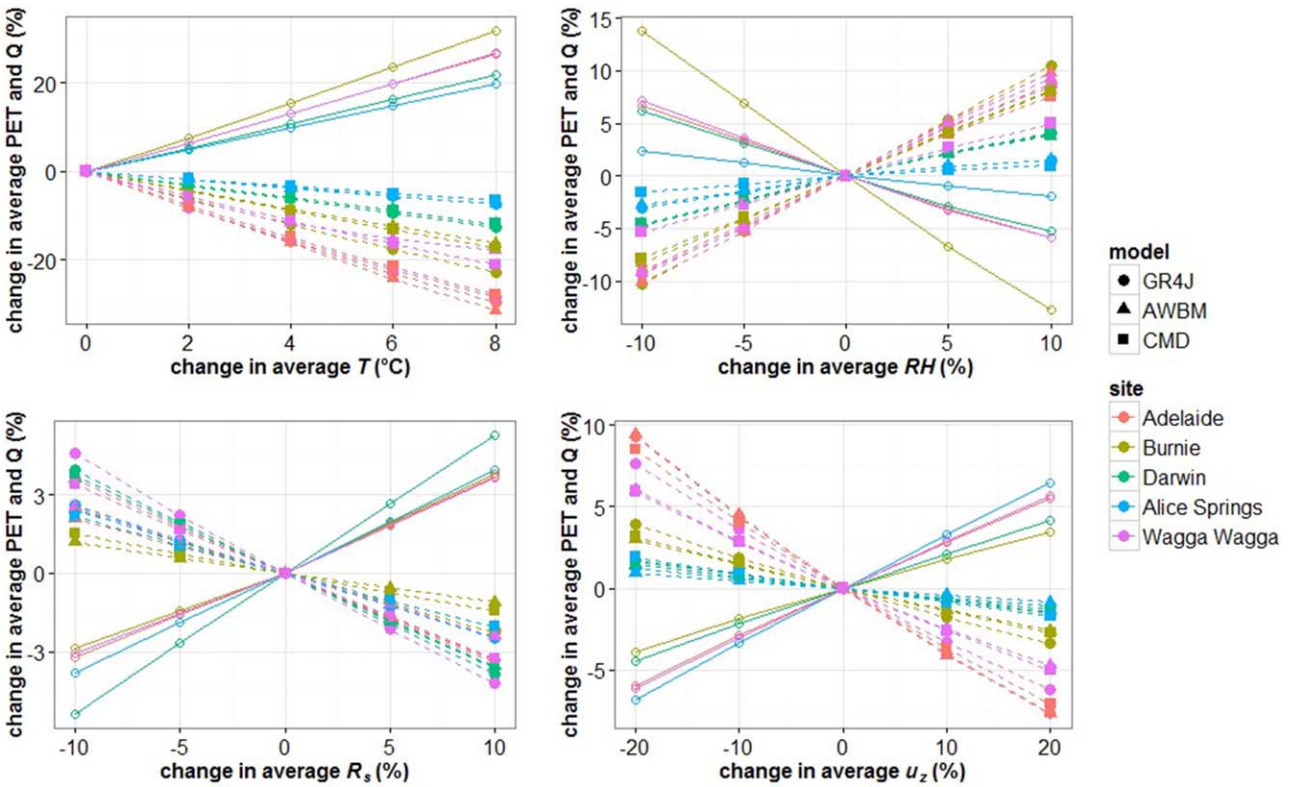


Figure 4. Sensitivity of PET and runoff to the four perturbed climate variables at the five study sites. The sensitivities for PET are represented by solid lines while for runoff dashed lines are used.

depending on the location and rainfall-runoff model used (Figure 4a). *RH* also shows substantial sensitivity, leading to up to between a 3% and 10% increase in runoff for a 10% increase in *RH* (Figure 4b). R_s and u_z show much smaller impacts, with a 10% increase in R_s causing between a 1% and 4% decrease in runoff (Figure 4c), while a 20% increase in u_z can cause between a 1% and 8% decrease in runoff (Figure 4d). For both PET and runoff, the relationship between sensitivity values and perturbations in each climate variable are highly linear, suggesting that the results can be represented per unit change for both PET and runoff.

To assess the specific role of rainfall-runoff model choice on the sensitivity values in Figure 4, the runoff elasticity (i.e., the percentage change in annual average runoff for each percentage change in annual average PET) simulated from three rainfall-runoff models is plotted for each study site in Figure 5. For the two most humid locations (Darwin and Burnie, which have the lowest long-term PET/P ratios as shown in Table 1), the elasticity values do not vary much across rainfall-runoff models and climate variables—this will be discussed further in section 4.3.

For the remaining locations (i.e., Adelaide, Alice Springs, and Wagga Wagga), the choice of both the rainfall-runoff model and the climate variable being perturbed makes a substantial difference to the runoff elasticity. Although all models suggest that runoff shows higher elasticity when perturbing *RH* rather than other climate variables, GR4J and AWBM consistently show much greater elasticity due to *RH*, compared to IHACRES_CMD. Focusing on Alice Springs, GR4J suggests that for each 1% change in PET, perturbing *RH* leads to the greatest change in runoff of 1.5%, whereas perturbing u_z leads to the smallest runoff change of only 0.2%. This is equivalent to an over sevenfold difference in runoff elasticity attributable solely to perturbing different climate variables. A similar magnitude of difference in runoff elasticity is observed for AWBM. In contrast, IHACRES_CMD suggests that the runoff elasticity for perturbing all four climate variables is below 1%. This is quite surprising, as these conceptual rainfall-runoff models do not directly use the four climate variables as inputs, but instead use the PET estimated with the Penman-Monteith model. Consequently, the fact that the variation of runoff responses depends on the climate variable being perturbed clearly

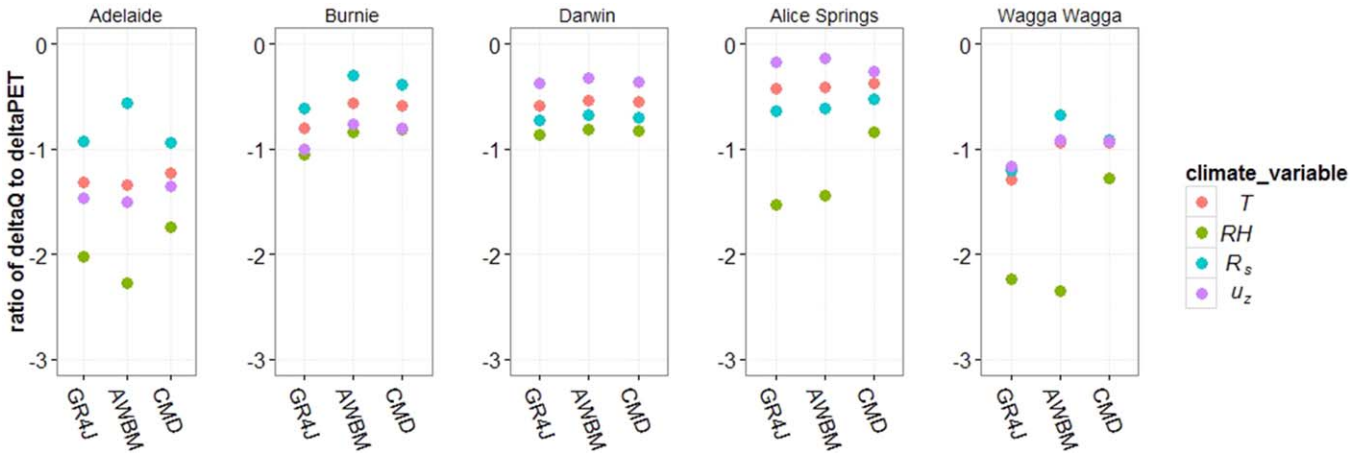


Figure 5. Runoff elasticity for different climate variables for each study site, as simulated from three rainfall-runoff models.

illustrates the impact of ET process representations within individual rainfall-runoff models on propagating climate change signals to projected runoff.

To better understand these findings, in the next section, we investigate the effect of rainfall-runoff model structure on the varying elasticity values shown above.

4.2. The Role of Rainfall-Runoff Model Structure on Sensitivity to PET

In the previous section, it was shown that the elasticity of runoff varied significantly depending on the variables used to perturb PET. Given that the results in Figure 5 are presented per unit change in annual average PET, it is anticipated that differences in elasticity values must be due to subannual variations that are different depending on the perturbing variable.

In this section, we therefore assess the role of subannual variations in PET on runoff sensitivity, with a focus on daily and seasonal variations. For illustration purposes, we use the results from Alice Springs, which exhibited the largest differences in runoff elasticity among the climate variables (Figure 5), and is thus expected to illustrate the most significant impact of ET process representations. Analyses from other sites generally led to conclusions that are consistent with those from Alice Springs, and are available in supporting information Figures S6–S17.

4.2.1. Interactions Between PET Sensitivity and AET Simulated for Wet and Dry Days

Figure 6a shows the wet/dry-day distribution of PET changes for a 1% change in annual average PET. Note that the same distribution of PET responses applies to GR4J, AWBM, and IHACRES_CMD, since no rainfall-runoff modeling is yet involved. The figure indicates that for a fixed average change in PET, there are significant differences in the change for wet and dry days. Among all climate variables, *RH* causes the greatest difference in average dry-day and wet-day sensitivities of PET (0.743% and 3.04%, respectively). *T* and R_s also show lower sensitivity for dry days compared to wet days, but the differences are much smaller compared to those for *RH* (with the two sensitivities becoming 0.989% and 1.01% for *T* and are 0.917% and 1.27% for R_s, respectively). In contrast, u_z leads to higher sensitivities on dry days compared to wet days (with averages of 1.06% and 0.567%, respectively).

The substantially higher PET sensitivity to *RH* can be explained with the aid of Figure 6b, where all daily PET changes that led to a 1% change in annual average PET from perturbing *RH* (as the second plot in Figure 6a) are plotted against the corresponding baseline daily *RH*. The figure shows that: (a) wet days are typically associated with high baseline *RH*; (b) these high-*RH* wet days generally lead to greater change in PET. These illustrate that part of the higher PET sensitivity on wet days can be the result of using multiplicative perturbation for *RH*, as the average *RH* on wet days is around 1.5 times of that for dry-day *RH*, leading to greater perturbation of *RH* on wet days than on dry days with any scaling factor. However, the use of a multiplicative perturbation method does not fully explain the fourfold difference in wet-/dry-day PET sensitivity in Figure 6a, which suggests that the structure of the Penman-Monteith model also contributes to greater PET changes during wet days compared to dry days.

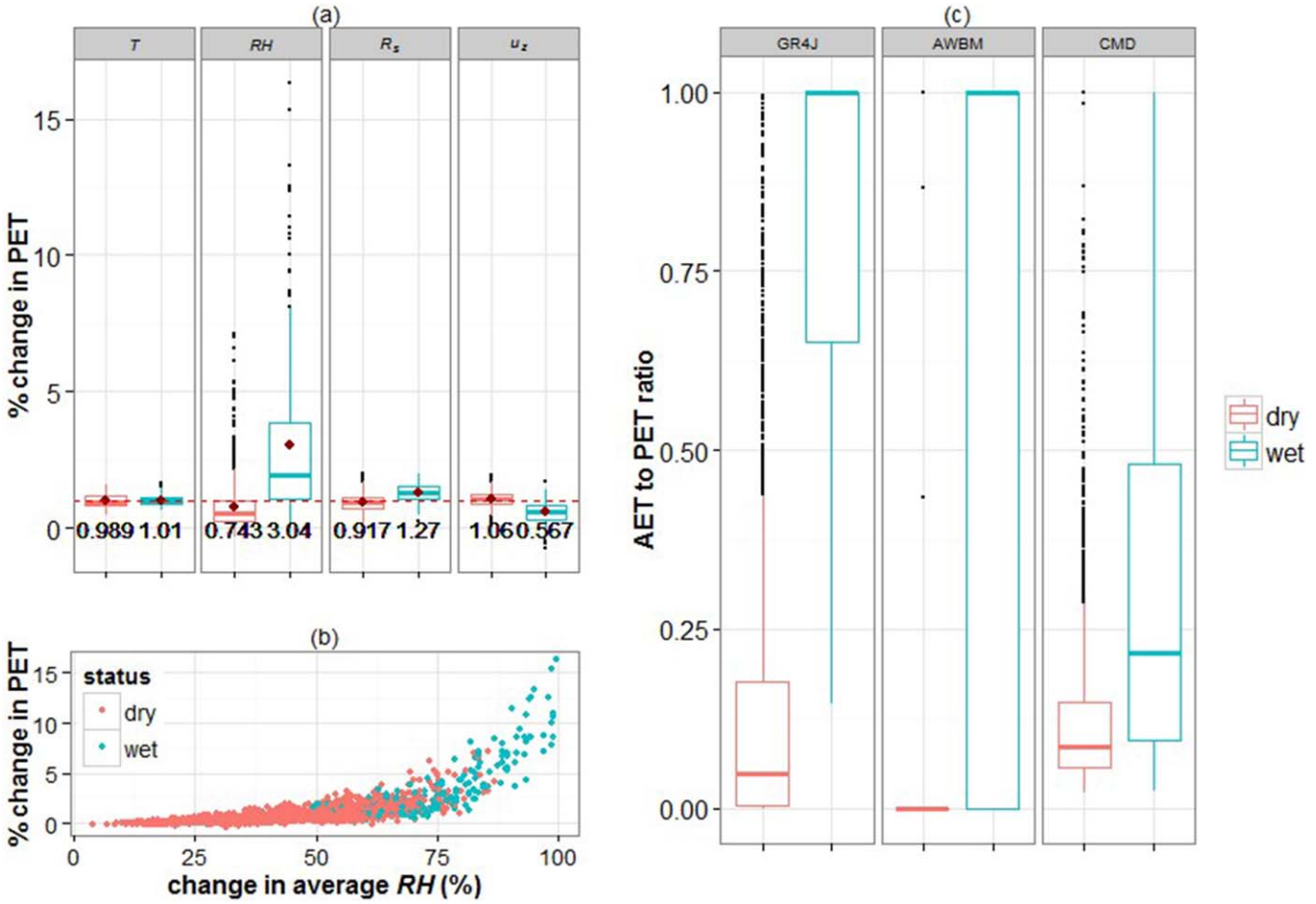


Figure 6. (a) Distribution of PET changes on wet and dry days, in response to perturbations in each climate variable (in each vertical plot) that result in an annual average PET change of 1% (indicated by the red dashed line), where a red dot represents the mean for each group; (b) all the daily PET changes that lead to a 1% change in annual average PET from perturbing RH, with different baseline levels of RH for wet and dry days; and (c) AET:PET ratios on wet and dry days, simulated from AWBM, GR4J, and IHACRES_CMD (in each vertical plot).

Having shown that the PET sensitivity differs from wet to dry days depending on the perturbing climate variable, we now investigate how PET is converted to AET differently on wet and dry days via different rainfall-runoff models. Figure 6c presents the distribution of daily AET:PET ratios on wet and dry days, simulated from the three rainfall-runoff models. All three models illustrate higher ratios of AET:PET during wet days compared to dry days, with GR4J and AWBM showing greater differences than IHACRES_CMD. This is because GR4J and AWBM both determine if AET = PET by checking if P – PET > 0, which produces a peak ratio of one for most wet days, while those below one correspond to days with P > 1 mm, but still not yet satisfying P – PET > 0. In contrast, IHACRES_CMD only uses store levels to determine AET, which is therefore largely independent of individual rainfall events, so that it does not lead to substantially higher AET on wet days. For dry days, the AET:PET ratios from GR4J and IHACRES_CMD show similar distributions, in contrast to those from AWBM. Although most dry-day ratios from GR4J and IHACRES_CMD are around zero, they can vary substantially with values spanning the entire range from 0 to 1, highlighting the direct impact of instantaneous store levels on AET in both models. In AWBM, this ratio is mostly zero with only a few anomalies, corresponding to dry days that have small rainfall amounts (0 < P < 1 mm), but still with water in some or all of the stores, leading to overall AET at discrete ratios of PET (i.e., 0, 0.433, 0.866, and 1).

Connecting these results with those in Figure 5, it is clear that for a fixed positive average change in PET, a greater decrease in runoff would be expected to occur for variables that cause the greatest change of PET during wet days, as a larger percentage of PET would be "lost" to AET during wet days in both GR4J and AWBM. Based on this finding, one would expect the highest runoff elasticity for both models for *RH*, followed by R_s, *T*, and finally u_z. This is exactly the same order as was identified for Alice Springs in Figure 5.

4.2.2. Interactions Between Seasonal Variations in PET Sensitivity, Simulated AET, and Storage

We now consider the role of seasonal patterns in the changes in PET as a result of perturbing each climate variable, and how they interact with simulated AET and storage. Figure 7a shows the seasonal distributions of the daily PET responses to perturbations in each climate variable that result in a 1% average change in PET at Alice Springs for the year 2000. The figure highlights substantial differences in the temporal distribution of PET responses, depending on the climate variable perturbed. For perturbations in *RH*, PET responses display high temporal variability: two clear peak periods are shown around February and May with the highest PET response exceeding 15 times average levels, while from July to October most responses are close to zero. In contrast, the variability in PET responses to the perturbations in T, R_s, and u_z are much smaller and close to 1% for the entire year.

Having shown the different seasonal variations in the PET sensitivity depending on the perturbing climate variable, we now investigate the seasonal patterns in the conversion from PET to AET within the three rainfall-runoff models. The simulated daily time series of AET:PET ratios from the three models are shown in Figure 7b for Alice Springs for the year 2000.

Considering first the results from January to October, all models suggest two periods with peak AET:PET ratios occurring around February and May, respectively, which correspond to the simulated store levels from all models (Figure 7c), which also show two consistent peak periods. Following the peak periods, the models show contrasting temporal patterns in both AET and storage: from both GR4J and IHACRES_CMD, the AET:PET ratios show a smooth "recession period" from 1 to close to 0, while the store drains gradually at the same time. This is because both models limit the rate of AET when store levels decrease. In contrast, the role of the store level in simulated AET is less obvious for AWBM, as the AET:PET ratios drop in a "stepwise" fashion (from 1 to levels of 0.866, 0.433, and then 0) following each peak period, corresponding to times when different stores become empty (as detailed in section 2.2.2). In addition, compared to GR4J and IHACRES_CMD, the stores simulated in AWBM drain at a more constant rate (e.g., before April), because this model forces each store to keep evaporating at the rate of PET until empty.

Interestingly, after October, GR4J and AWBM both show instantaneous high AET:PET ratios on wet days, but returning to lower values once the rainfall ceases. This indicates the impact of individual rainfall events, which is independent of that of storage, since peak ratios can occur even when storage levels are consistently low. Such impact is not visible from IHACRES_CMD as it does not relate AET simulation to individual rainfall.

From these results, all three models show a fuller storage around both February and May. Since all models relate AET simulation to storage, a greater amount of PET is also converted to AET during this period, with associated implications on runoff volume. As these periods also correspond to when *RH* can cause the greatest changes in PET, we can expect that for a fixed positive average change in PET, perturbations in *RH* can lead to greater decreases in runoff from all three models. This, again, is consistent with the higher sensitivity of runoff to *RH* shown from all three models in Figure 5.

Combining results from sections 4.2.1 and 4.2.2, we can see that:

1. PET shows substantially higher sensitivity to *RH* during wet days and around February and May.
2. Both GR4J and AWBM explicitly relate AET simulation to daily wet/dry status, leading to higher impact of changing PET on AET during wet days.
3. All three models relate AET simulation to catchment storage, leading to higher impact of changing PET on AET when stores are fuller around February and May.

Relating the above results to observations for Alice Springs in Figure 5, it is clear that the higher elasticity values for *RH* compared to the other variables shown in all three models are likely to be attributable to these models having a higher contrast in the AET:PET ratios across different seasons where store levels vary substantially (Figure 7). Furthermore, the reason that GR4J and AWBM show even higher sensitivity to *RH* can be that these models having higher contrast in AET:PET ratios between wet day dry days across wet and dry days, which are associated with contrasting levels of *RH* (Figure 6).

In summary, in this section, we show that there are clear daily and seasonal variations in the sensitivity of PET depending on which climate variable is perturbed. This can then interact with the way PET is converted to AET in the three models, which depends on whether a day is dry or wet, as well as the temporal

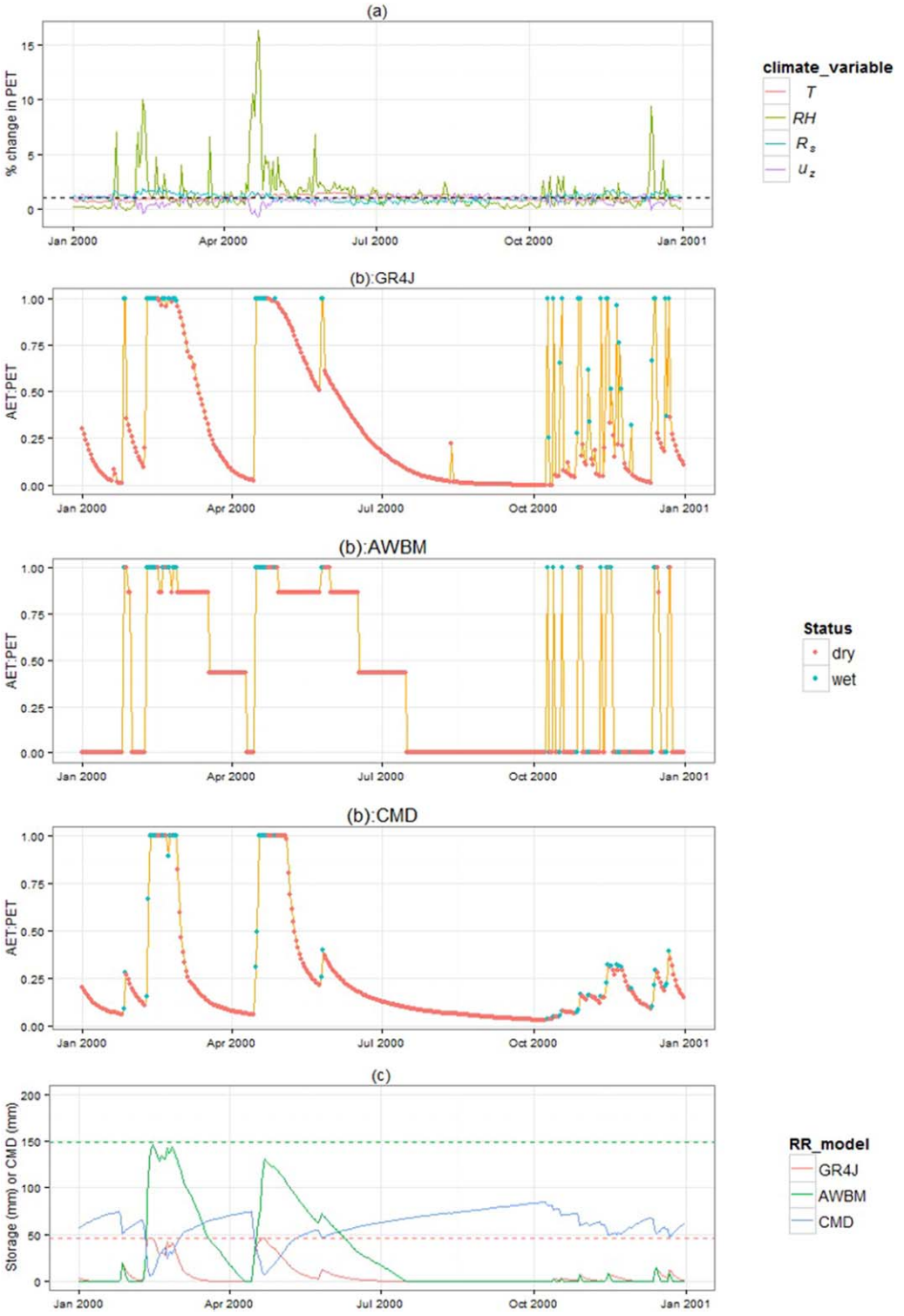


Figure 7. Seasonality of (a) daily PET responses to perturbations in each climate variable that result in an average PET change of 1% (black dashed line); (b) AET:PET ratios for wet and dry days from GR4J, AWBM, and IHACRES_CMD; and (c) levels of production stores from GR4J, AWBM (with dashed lines indicating store capacities), and IHACRES_CMD (as CMD levels). All results are from Alice Springs for the year 2000.

variation in the production store. The interaction between these two mechanisms can lead to large variability in runoff sensitivity to different perturbing climate variables, even for the same change in annual average PET.

4.3. Relative Realism of Alternative ET Process Representations Within Conceptual Rainfall-Runoff Models

Given the importance of rainfall-runoff model process representation in determining how much AET is converted from PET, we now investigate model realism by comparing observed time series of AET with simulated AET from the three conceptual rainfall-runoff models at Alice Springs and Wagga Wagga (Figure 8).

For Alice Springs, the differences between model simulated AET clearly illustrate their contrasting ET process representations. GR4J illustrates both the effect of individual rainfall events, as well as change in store levels, with simulated AET peaking on individual wet days and gradually decreasing on the dry days after, which indicates drops in store levels. In AWBM, the effect of individual rainfall is also illustrated by the peak AET on wet days; however, the influence of retreating storage is not shown in AWBM, with AET either remaining at or near its peak value, or dropping suddenly to zero. This instantaneous drop is due to the "discrete" function that AWBM uses to represent the relationship between AET and storage (as discussed in section 2.2.2). In contrast to GR4J and AWBM, the simulated AET from IHACRES_CMD is not sensitive to individual rainfall events, as the occurrence of peak AET is not associated with wet days. This is because AET is simulated purely as a function of the moisture content in the store. In general, GR4J and IHACRES_CMD better resemble the temporal variation in the observed AET in Alice Spring, which indicates that the change in store levels is likely to be a key controlling mechanism for ET processes in this catchment.

In contrast to Alice Springs, all three models show reasonable performance in producing the temporal patterns in the observed AET at Wagga Wagga, with only slight differences across models suggesting a lower impact of ET process representation. As an exception, at the start of 2013 and 2014, AWBM shows zero-AET periods, which are clearly different from observations. As illustrated in Figure 7, these periods correspond to times when all storages are empty, as a result of AWBM keeping each store evaporating at the PET rate (i.e., AET=PET).

The comparisons of AET simulation at the two study sites indicate that the impact of ET process representations is likely to be greater for drier catchments such as Alice Springs, where stores vary substantially on a seasonal basis, and experience frequent draining and rapid changes of levels in response to rainfall. In contrast, Wagga Wagga is a somewhat more humid catchment with lower seasonal variability (although still water-limited, see Table 1). As such, it generally has fuller storage levels, which neither drain out completely nor rise rapidly with individual rainfall events. Under this condition, the simulated AET is likely to be less sensitive to different ways of defining the relationships between the AET:PET ratio and the storage level (e.g., GR4J and AWBM). Furthermore, with the consistently fuller stores leading to higher AET overall, the impact of different ET representations across dry/wet status (e.g., GR4J and IHACRES_CMD) are likely to be limited. Therefore, model choice is likely to have less impact on ET simulations for catchments with relatively stable store levels. This can also be a possible explanation of the observations in section 4.1, where humid locations, such as Darwin and Burnie, show fewer differences in runoff sensitivity to different climate variables across the three rainfall-runoff models.

It should be noted that although the EC-based AET measurements were used as a benchmark for the simulated AET from the three rainfall-runoff models, the comparison should be interpreted qualitatively rather than quantitatively, due to several limitations associated with the available data. In particular, both catchments used for this comparison were assumed to be covered with reference crop when estimating catchment PET for simulating AET (section 2.2.1). However, the EC measurements were obtained from flux towers within tall canopy (Mulga woodland for Alice Springs and Eucalyptus forest for Wagga Wagga), which are likely to be valid only for a small region for the specific type of canopy around each flux tower, since EC measurements usually only cover spatial scales of hundreds of meters. This inconsistency can contribute to differences between the simulated and observed catchment AET. Furthermore, a number of studies have reported biases in AET measured by the EC method due to the failure to close the energy balance [e.g., *Wang and Dickinson*, 2012; *Wilson et al*., 2001]. These biases appear to be particularly prominent for days with high relative humidity and wind [*Ibrom et al*., 2007], as well as during turbulent conditions at night or when precipitation or dew obscured the sensors [*Meiresonne et al*., 2003; *Meyer et al*., 2015; *Wilson et al*., 2001].

Despite the abovementioned limitations, the ET observations from the two sites illustrate contrasting temporal variations related to rainfall events and store levels, which can be explained by the unique characteristic of each catchment. Therefore, it is likely that these patterns shown in the data may nevertheless be

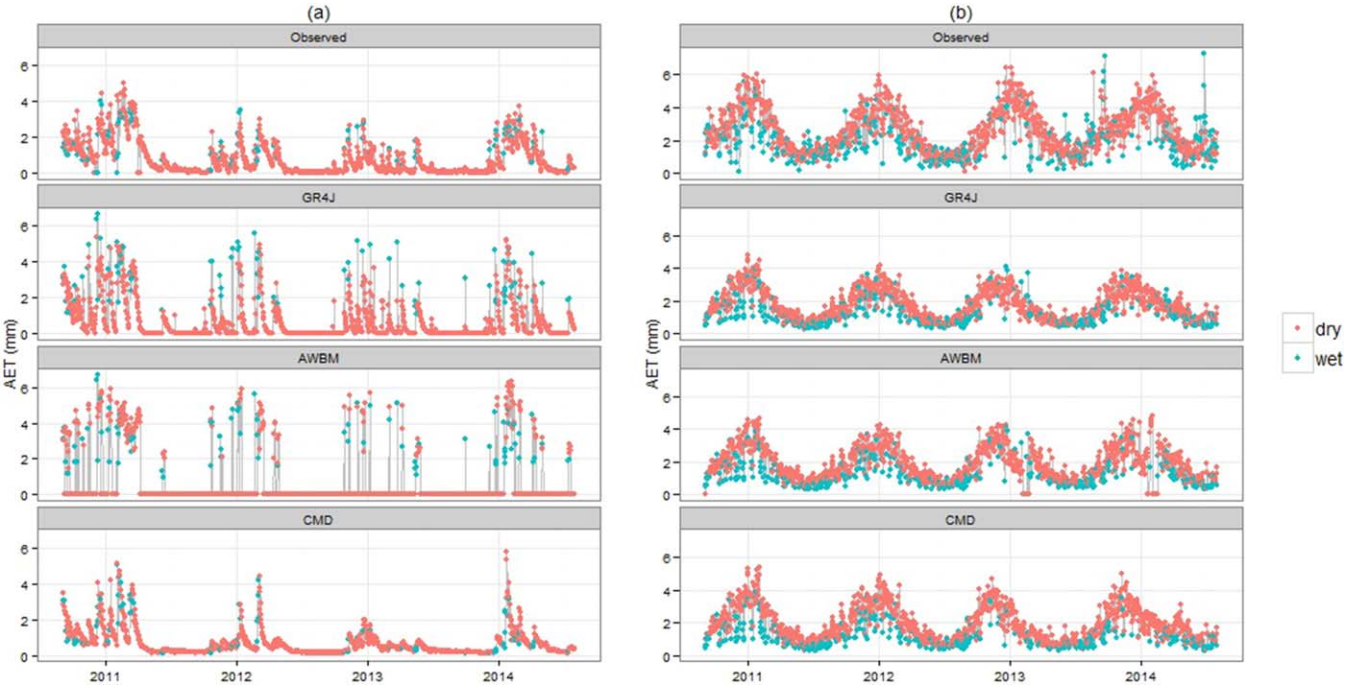


Figure 8. Comparison of the observed and simulated time series of AET from GR4J, AWBM, and IHACRES_CMD, at (a) Alice Springs and (b) Wagga Wagga.

useful to scrutinize rainfall-runoff models and assess which models simulate ET processes more realistically. Furthermore, the results also highlight potential opportunities for more detailed process-based evaluation of rainfall-runoff models where data are available, which can provide valuable insights to support rainfall-runoff model selection complementary to the use of conventional runoff-centered performance assessment (e.g., calibration metrics in Table 2). These process-focused assessments are particularly important when evaluating the suitability of models for simulating catchment behavior under changed conditions, where the differences in the runoff projections from alternative rainfall-runoff models are likely to be greatest.

5. Summary and Conclusions

In this study, we:

1. Assessed the impact of ET process representations within conceptual rainfall-runoff models on runoff projections under plausible changes in the climate drivers of ET.
2. Explored the interaction between conceptual rainfall-runoff model structure and the complex subannual changes in PET, and how this influences projections of runoff and AET.
3. Evaluated the relative realism of alternative ET process representations within three conceptual rainfall-runoff models using AET observations.

We investigated five study sites across Australia with contrasting hydro-climatic conditions, at which the baseline PET was simulated using the Penman-Monteith model. Perturbation of PET was achieved by first perturbing the historical data of four climatic drivers of ET, namely temperature (T), relative humidity (RH), solar radiation (R_s), and wind (u_z), within plausible ranges of their future changes, and then estimating the corresponding PET with the Penman-Monteith model. The baseline and perturbed PET time series formed inputs to three structurally different conceptual rainfall-runoff models (GR4J, AWBM, and IHACRES_CMD), which were calibrated to historical runoff data at the five study sites. The calibrated rainfall-runoff models were used to simulate the baseline and perturbed runoff, and then to estimate runoff sensitivity to climate perturbations.

Our results illustrate that different ET process representations in conceptual rainfall-runoff models can have substantial impacts on the sensitivity of runoff projection under a changing climate. Specifically, it is possible to have an over sevenfold difference in runoff elasticity due to rainfall-runoff model choice and the

climate variable that is perturbed. The significant differences in runoff elasticity were attributed to the different conceptualizations used to convert PET to AET on wet and dry days in the three conceptual rainfall-runoff models, as well as the contrasting relationships defined between AET and store levels.

Using eddy-covariance-based observations of AET, we also compared and evaluated the ET process representations within the three conceptual rainfall-runoff models at Alice Springs and Wagga Wagga. Results suggest that alternative ET process representations are likely to have a greater impact on drier catchments consisting of highly variable store levels, such as Alice Springs. As highlighted with the results from Alice Springs, AWBM poorly simulated the variations in observed AET, whereas GR4J and IHACRES_CMD performed better in simulating the temporal variation of AET as a result of changing store levels.

Although this study demonstrates that evapotranspiration process representation within conceptual rainfall-runoff models can substantially affect the results from climate impact assessments, quantitative sensitivity values are likely to vary on a case-by-case basis, depending on the catchment and the choice of PET and rainfall-runoff models. Therefore, specific studies are required for further exploration of the implications of modeling assumptions for a larger number of rainfall-runoff models, as well as their interactions on other case studies. Nevertheless, the results from this study highlight the importance of scrutinizing internal process realism within rainfall-runoff models as part of climate impact studies, and using alternative sources of information on model performance in addition to the standard runoff-based metrics, such as by comparing simulated and observed AET or with the outputs of integrated surface-subsurface flow models [e.g., *Li et al*., 2015; *Partington et al*., 2013; *Su et al*., 2016; *Vervoort et al*., 2014]).

These findings are in line with a number of studies that suggest the need for an improved understanding of physical processes to better infer the potential changes in rainfall-runoff relationships under climate change, and thus to facilitate better modeling of future water resources [*Fowler et al*., 2016; *Herman et al*., 2013; *Kirchner*, 2006; *Merz et al*., 2011; *Minville et al*., 2014; *Saft et al*., 2016; *Westra et al*., 2014]. The study also provides a practical example of the multiple-hypothesis testing of rainfall-runoff models [*Clark et al*., 2011, 2016; *Gupta et al*., 2008; *Harrigan et al*., 2014], which suggests that models should be evaluated based on the validity of any individual hypothesis they describe. With the currently increasing availability of observations to support hypothesis testing [e.g., see *Blöschl et al*., 2016], this type of process-based model evaluation can greatly enhance our understanding on how alternative rainfall-runoff models can propagate climate change signals to hydrological impacts via different model conceptualizations and assumptions. As highlighted by the substantially different elasticity values found in this study, these process-based model evaluation studies will be particularly useful to inform the selection of rainfall-runoff models for climate impact studies, where good model performance under historical climate conditions provides only a partial guide as to how the model will perform under changed climate conditions.

Acknowledgments
The streamflow data used for the five catchments in this study were from the Australian Bureau of Meteorology's (BOM) Hydrologic Reference Station project website: www.bom.gov.au/hrs [*Bureau of Meteorology*, 2015]. The rainfall, temperature, relative humidity, wind, and sunshine hour data for the five weather stations were obtained from the Climate Data Online project website, http://www.bom.gov.au/climate/data/ [*Bureau of Meteorology*, 2016]. The EC-based AET observations for Alice Springs and Wagga Wagga were obtained from the OzFlux website, at http://www.ozflux.org.au/ [*Cleverly*, 2011; *vanGorsel*, 2013]. The authors would also like to thank the three anonymous reviewers, whose comments have helped to improve the quality of this paper significantly.

References

Akhtar, M., N. Ahmad, and M. J. Booij (2008), The impact of climate change on the water resources of Hindukush-Karakorum-Himalaya region under different glacier coverage scenarios, *J. Hydrol.*, *355*(1–4), 148–163.
Allen, R. G., L. S. Pereira, D. Raes, and M. Smith (1998), Crop evapotranspiration–Guidelines for computing crop water requirements, *FAO Irrig. Drain. Pap. 56*, p. 300, Food Agric. Organ., Rome.
Andréassian, V., C. Perrin, and C. Michel (2004), Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models, *J. Hydrol.*, *286*(1–4), 19–35.
Andrews, F., and J. Guillaume (2013), *Hydromad: Hydrological Model Assessment and Development, R Package Version 0.9-18*. [Available at http://hydromad.catchment.org/.]
Arnell, N. W. (2004), Climate-change impacts on river flows in Britain: The UKCIPO2 scenarios, *Water Environ. J.*, *18*(2), 112–117.
Bae, D.-H., I.-W. Jung, and D. P. Lettenmaier (2011), Hydrologic uncertainties in climate change from IPCC AR4 GCM simulations of the Chungju Basin, Korea, *J. Hydrol.*, *401*(1–2), 90–105.
Bastola, S., C. Murphy, and J. Sweeney (2011), The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments, *Adv. Water Resour.*, *34*(5), 562–576.
Bethune, M., N. Austin, and S. Maher (2001), Quantifying the water budget of irrigated rice in the Shepparton Irrigation Region, Australia, *Irrig. Sci.*, *20*(3), 99–105.
Beven, K. J. (2001a), On hypothesis testing in hydrology, *Hydrol. Processes*, *15*(9), 1655–1657.
Beven, K. J. (2001b), How far can we go in distributed hydrological modelling?, *Hydrol. Earth Syst. Sci. Discuss.*, *5*(1), 1–12.
Beven, K. J. (2011), *Rainfall-Runoff Modelling: The Primer*, Wiley, 2nd ed., Chichester, U. K.
Blöschl, G., et al. (2016), The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: A hypothesis-driven observatory, *Hydrol. Earth Syst. Sci.*, *20*(1), 227–255.
Boughton, W. (2004), The Australian water balance model, *Environ. Model. Software*, *19*(10), 943–956.
Boughton, W. (2009), New approach to calibration of the AWBM for use on ungauged catchments, *J. Hydrol. Eng.*, *14*(6), 562–568.

Brouyère, S., G. Carabin, and A. Dassargues (2004), Climate change impacts on groundwater resources: Modelled deficits in a chalky aquifer, Geer basin, Belgium, *Hydrogeol. J.*, *12*(2), 123–134.
Bureau of Meteorology (2015), *Hydrologic Reference Stations*. [Available at http://www.bom.gov.au/water/hrs/.]
Bureau of Meteorology (2016), *Climate Data Online*. [Available at http://www.bom.gov.au/climate/data/.]
Butts, M. B., J. T. Payne, M. Kristensen, and H. Madsen (2004), An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation, *J. Hydrol.*, *298*(1–4), 242–266.
Chapman, T. G. (2003), Estimation of evaporation in rainfall-runoff models, in *MODSIM 2003 International Congress on Modelling and Simulation, July 2003*, edited by D. A. Post, Modelling and Simulation Society of Australia and New Zealand. [Available at http://www.mssanz.org.au.previewdns.com/MODSIM03/Volume_01/A03/04_Chapman.pdf.]
Chiew, F. H. S. (2006), Estimation of rainfall elasticity of streamflow in Australia, *Hydrol. Sci. J.*, *51*(4), 613–625.
Chiew, F. H. S., and T. A. McMahon (1991), The applicability of Morton's and Penman's evapotranspiration estimates in rainfall-runoff modelling, *Water Resour. Bull.*, *27*(4), 611–620.
Chiew, F. H. S., J. Teng, J. Vaze, D. A. Post, J. M. Perraud, D. G. C. Kirono, and N. R. Viney (2009), Estimating climate change impact on runoff across southeast Australia: Method, results, and implications of the modeling method, *Water Resour. Res.*, *45*, W10414, doi:10.1029/2008WR007338.
Chiew, F. H. S., D. G. C. Kirono, D. M. Kent, A. J. Frost, S. P. Charles, B. Timbal, K. C. Nguyen, and G. Fu (2010), Comparison of runoff modelled using rainfall from different downscaling methods for historical and future climates, *J. Hydrol.*, *387*(1–2), 10–23.
Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer (2004), The effects of climate change on the hydrology and water resources of the Colorado River Basin, *Clim. Change*, *62*(1), 337–363.
Clark, M. P., D. Kavetski, and F. Fenicia (2011), Pursuing the method of multiple working hypotheses for hydrological modeling, *Water Resour. Res.*, *47*, W09301, doi:10.1029/2010WR009827.
Clark, M. P., R. L. Wilby, E. D. Gutmann, J. A. Vano, S. Gangopadhyay, A. W. Wood, H. J. Fowler, C. Prudhomme, J. R. Arnold, and L. D. Brekke (2016), Characterizing uncertainty of the hydrologic impacts of climate change, *Curr. Clim. Change Rep.*, *2*(2), 55–64.
Cleugh, H. A., R. Leuning, Q. Mu, and S. W. Running (2007), Regional evaporation estimates from flux tower and MODIS satellite data, *Remote Sens. Environ.*, *106*(3), 285–304.
Cleverly, J. (2011), *Alice Springs Mulga OzFlux site OzFlux*, Australian and New Zealand Flux Research and Monitoring. [Available at hdl.102.100.100/14217.]
Coron, L., V. Andréassian, C. Perrin, J. Lerat, J. Vaze, M. Bourqui, and F. Hendrickx (2012), Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments, *Water Resour. Res.*, *48*, W05552, doi:10.1029/2011WR011721.
Croke, B. F. W., and A. J. Jakeman (2004), A catchment moisture deficit module for the IHACRES rainfall-runoff model, *Environ. Model. Software Environ. Data News*, *19*(1), 1–5.
Crossman, J., M. N. Futter, S. K. Oni, P. G. Whitehead, L. Jin, D. Butterfield, H. M. Baulch, and P. J. Dillon (2013), Impacts of climate change on hydrology and water quality: Future proofing management strategies in the Lake Simcoe watershed, Canada, *J. Great Lakes Res.*, *39*(1), 19–32.
CSIRO and Bureau of Meteorology (2015), Climate change in Australia information for Australia's Natural Resource Management Regions, technical report, Australia.
Dingman, S. L. (2015), *Physical Hydrology*: Third Edition, Waveland Press, Ill.
Ewen, J., G. O'Donnell, A. Burton, and E. O'Connell (2006), Errors and uncertainty in physically-based rainfall-runoff modelling of catchment change effects, *J. Hydrol.*, *330*(3–4), 641–650.
Fowler, K. J. A., M. C. Peel, A. W. Western, L. Zhang, and T. J. Peterson (2016), Simulating runoff under changing climatic conditions: Revisiting an apparent deficiency of conceptual rainfall-runoff models, *Water Resour. Res.*, *52*, 1820–1846, doi:10.1002/2015WR018068.
Goderniaux, P., S. Brouyère, H. J. Fowler, S. Blenkinsop, R. Therrien, P. Orban, and A. Dassargues (2009), Large scale surface-subsurface hydrological model to assess climate change impacts on groundwater reserves, *J. Hydrol.*, *373*(1–2), 122–138.
Gong, L., C.-y. Xu, D. Chen, S. Halldin, and Y. D. Chen (2006), Sensitivity of the Penman-Monteith reference evapotranspiration to key climatic variables in the Changjiang (Yangtze River) basin, *J. Hydrol.*, *329*(3–4), 620–629.
Gosling, S. N., R. G. Taylor, N. W. Arnell, and M. C. Todd (2011), A comparative analysis of projected impacts of climate change on river runoff from global and catchment-scale hydrological models, *Hydrol. Earth Syst. Sci.*, *15*(1), 279–294.
Goyal, R. K. (2004), Sensitivity of evapotranspiration to global warming: A case study of arid zone of Rajasthan (India), *Agric. Water Manage.*, *69*(1), 1–11.
Guo, D., S. Westra, and H. R. Maier (2016a), An R package for modelling actual, potential and reference evapotranspiration, *Environ. Model. Software*, *78*, 216–224.
Gupta, H. V., T. Wagener, and Y. Liu (2008), Reconciling theory with observations: Elements of a diagnostic approach to model evaluation, *Hydrol. Processes*, *22*(18), 3802–3813.
Haque, M. M., A. Rahman, D. Hagare, G. Kibria, and F. Karim (2015), Estimation of catchment yield and associated uncertainties due to climate change in a mountainous catchment in Australia, *Hydrol. Processes*, *29*(19), 4339–4349.
Harrigan, S., C. Murphy, J. Hall, R. L. Wilby, and J. Sweeney (2014), Attribution of detected changes in streamflow using multiple working hypotheses, *Hydrol. Earth Syst. Sci.*, *18*(5), 1935–1952.
Hartmann, G., and A. Bárdossy (2005), Investigation of the transferability of hydrological models and a method to improve model calibration, *Adv. Geosci.*, *5*, 83–87.
Hauser, R., S. Archer, P. Backlund, J. Hatfield, A. Janetos, D. Lettenmaier, and M. Walsh (2009), *The Effects of Climate Change on US Ecosystems*, US Global Change Research Program, Washington, D. C.
Herman, J. D., P. M. Reed, and T. Wagener (2013), Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior, *Water Resour. Res.*, *49*, 1400–1414.
Huo, Z., X. Dai, S. Feng, S. Kang, and G. Huang (2013), Effect of climate change on reference evapotranspiration and aridity index in arid region of China, *J. Hydrol.*, *492*, 24–34.
Ibrom, A., E. Dellwik, H. Flyvbjerg, N. O. Jensen, and K. Pilegaard (2007), Strong low-pass filtering effects on water vapour flux measurements with closed-path eddy correlation systems, *Agric. For. Meteorol.*, *147*(3–4), 140–156.
IPCC (2014), *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 1132 pp., IPCC, Cambridge, U. K.
Islam, S. A., M. A. Bari, and A. H. M. F. Anwar (2014), Hydrologic impact of climate change on Murray-Hotham catchment of Western Australia: A projection of rainfall-runoff for future water resources planning, *Hydrol. Earth Syst. Sci.*, *18*(9), 3591–3614.
Jiang, T., Y. D. Chen, C.-y. Xu, X. Chen, X. Chen, and V. P. Singh (2007), Comparison of hydrological impacts of climate change simulated by six hydrological models in the Dongjiang Basin, South China, *J. Hydrol.*, *336*(3–4), 316–333.

Kay, A. L., and R. G. Jones (2012), Comparison of the use of alternative UKCP09 products for modelling the impacts of climate change on flood frequency, *Clim. Change*, *114*(2), 211–230.
Kay, A. L., H. N. Davies, V. A. Bell, and R. G. Jones (2009), Comparison of uncertainty sources for climate change impacts: Flood frequency in England, *Clim. Change*, *92*(1–2), 41–63.
Kirchner, J. W. (2006), Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology, *Water Resour. Res.*, *42*, W03S04, doi:10.1029/2005WR004362.
Klemeš, V. (1986), Operational testing of hydrological simulation models, *Hydrol. Sci. J.*, *31*(1), 13–24.
Li, L., M. F. Lambert, H. R. Maier, D. Partington, and C. T. Simmons (2015), Assessment of the internal dynamics of the Australian Water Balance Model under different calibration regimes, *Environ. Model. Software*, *66*, 57–68.
McKenney, M. S., and N. J. Rosenberg (1993), Sensitivity of some potential evapotranspiration estimation methods to climate change, *Agric. For. Meteorol.*, *64*(1–2), 81–110.
McMahon, T. A., M. C. Peel, and D. J. Karoly (2015), Assessment of precipitation and temperature data from CMIP3 global climate models for hydrologic simulation, *Hydro. Earth Syst. Sci.*, *19*(1), 361–377.
McMahon, T. A., M. C. Peel, L. Lowe, R. Srikanthan, and T. R. McVicar (2013), Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: A pragmatic synthesis, *Hydrol. Earth Syst. Sci.*, *17*(4), 1331–1363.
Meiresonne, L., D. A. Sampson, A. S. Kowalski, I. A. Janssens, N. Nadezhdina, J. Cermák, J. Van Slycken, and R. Ceulemans (2003), Water flux estimates from a Belgian Scots pine stand: A comparison of different approaches, *J. Hydrol.*, *270*(3–4), 230–252.
Mendoza, P. A., M. P. Clark, N. Mizukami, E. D. Gutmann, J. R. Arnold, L. D. Brekke, and B. Rajagopalan (2016), How do hydrologic modeling decisions affect the portrayal of climate change impacts?, *Hydrol. Processes*, *30*(7), 1071–1095.
Merz, R., J. Parajka, and G. Blöschl (2011), Time stability of catchment model parameters: Implications for climate impact analyses, *Water Resour. Res.*, *47*, W02531, doi:10.1029/2010WR009505.
Meyer, W. S., E. Kondrlovà, and G. R. Koerber (2015), Evaporation of perennial semi-arid woodland in southeastern Australia is adapted for irregular but common dry periods, *Hydrol. Processes*, *29*(17), 3714–3726.
Minville, M., D. Cartier, C. Guay, L.-A. Leclaire, C. Audet, S. Le Digabel, and J. Merleau (2014), Improving process representation in conceptual hydrological model calibration using climate simulations, *Water Resour. Res.*, *50*, 5044–5073, doi:10.1002/2013WR013857.
Mu, Q., M. Zhao, and S. W. Running (2011), Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sens. Environ.*, *115*(8), 1781–1800.
Najafi, M. R., and H. Moradkhani (2015), Multi-model ensemble analysis of runoff extremes for climate change impact assessments, *J. Hydrol.*, *525*, 352–361.
Northey, J. E., E. W. Christen, J. E. Ayars, and J. Jankowski (2006), Occurrence and measurement of salinity stratification in shallow groundwater in the Murrumbidgee Irrigation Area, south-eastern Australia, *Agric. Water Manage.*, *81*(1–2), 23–40.
New, M., A. Lopez, S. Dessai, and R. Wilby (2007), Challenges in using probabilistic climate change information for impact assessments: An example from the water sector, *Philos. Trans. R. Soc. A*, *365*(1857), 2117–2131.
Oudin, L., C. Michel, and F. Anctil (2005), Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 1—Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs?, *J. Hydrol.*, *303*(1–4), 275–289.
Oudin, L., C. Perrin, T. Mathevet, V. Andréassian, and C. Michel (2006), Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models, *J. Hydrol.*, *320*(1–2), 62–83.
Partington, D., P. Brunner, S. Frei, C. T. Simmons, A. D. Werner, R. Therrien, H. R. Maier, G. C. Dandy, and J. H. Fleckenstein (2013), Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment, *Water Resour. Res.*, *49*, 5501–5519, doi:10.1002/wrcr.20405.
Paton, F. L., H. R. Maier, and G. C. Dandy (2013), Relative magnitudes of sources of uncertainty in assessing climate change impacts on water supply security for the southern Adelaide water supply system, *Water Resour. Res.*, *49*, 1643–1667, doi:10.1002/wrcr.20153.
Paton, F. L., H. R. Maier, and G. C. Dandy (2014), Including adaptation and mitigation responses to climate change in a multiobjective evolutionary algorithm framework for urban water supply systems incorporating GHG emissions, *Water Resour. Res.*, *50*, 6285–6304, doi:10.1002/2013WR015195.
Perrin, C., C. Michel, and V. Andréassian (2003), Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*, *279*(1), 275–289.
Saft, M., M. C. Peel, A. W. Western, J.-M. Perraud, and L. Zhang (2016), Bias in streamflow projections due to climate-induced shifts in catchment response, *Geophys. Res. Lett.*, *43*, 1574–1581, doi:10.1002/2015GL067326.
Seibert, J., and J. J. McDonnell (2002), On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resour. Res.*, *38*(11), 1241, doi:10.1029/2001WR000978.
Seibert, J., A. Rodhe, and K. Bishop (2003), Simulating interactions between saturated and unsaturated storage in a conceptual runoff model, *Hydrol. Processes*, *17*(2), 379–390.
Seneviratne, S. I., et al. (2012), Swiss prealpine Rietholzbach research catchment and lysimeter: 32 year time series and 2003 drought event, *Water Resour. Res.*, *48*, W06526, doi:10.1029/2011WR011749.
Stern, H., G. De Hoedt, and J. Ernst (2000), Objective classification of Australian climates, *Aust. Meteorol. Mag.*, *49*(2), 87–96.
Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (2013), Climate change 2013: The physical science basis, *Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5)*, Cambridge Univ. Press, New York.
Su, C.-H., J. F. Costelloe, T. J. Peterson, and A. W. Western (2016), On the structural limitations of recursive digital filters for base flow estimation, *Water Resour. Res.*, *52*, 4745–4764, doi:10.1002/2015WR018067.
TERN (2012), Coping with complex data streams: The OzFlux approach, *TERN e-NewsI*., October 2012. [Available at http://mymail.ezemsgs.com/em/message/email/view.php?id=959668&u=45618/.]
Teng, J., J. Vaze, F. H. S. Chiew, B. Wang, and J.-M. Perraud (2012), Estimating the relative uncertainties sourced from GCMs and hydrological models in modeling climate change impact on runoff, *J. Hydrometeorol.*, *13*(1), 122–139.
Turral, H., J. Burke, and J.-M. Faurès (2011), *Climate Change, Water and Food Security*, Food Agric. Organ., FAO water reports 36, Rome.
vanGorsel, E. (2013), Tumbarumba OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring. [Available at hdl:102.100.100/14241.]
Vaze, J., and J. Teng (2011), Future climate and runoff projections across New South Wales, Australia: Results and practical applications, *Hydrol. Processes*, *25*(1), 18–35.
Vaze, J., D. A. Post, F. H. S. Chiew, J. M. Perraud, N. R. Viney, and J. Teng (2010), Climate non-stationarity—Validity of calibrated rainfall-runoff models for use in climate change studies, *J. Hydrol.*, *394*(3–4), 447–457.

Velázquez, J. A., J. Schmid, S. Ricard, M. J. Muerth, B. Gauvin St-Denis, M. Minville, D. Chaumont, D. Caya, R. Ludwig, and R. Turcotte (2013), An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources, *Hydrol. Earth Syst. Sci.*, *17*(2), 565–578.

Vervoort, W. R., S. F. Miechels, F. F. van Ogtrop, and J. H. A. Guillaume (2014), Remotely sensed evapotranspiration to calibrate a lumped conceptual model: Pitfalls and opportunities, *J. Hydrol.*, *519*(Part D), 3223–3236.

Vicente-Serrano, S. M., C. Azorin-Molina, A. Sanchez-Lorenzo, J. Revuelto, E. Morán-Tejeda, J. I. López-Moreno, and F. Espejo (2014), Sensitivity of reference evapotranspiration to changes in meteorological parameters in Spain (1961–2011), *Water Resour. Res.*, *50*, 8458–8480, doi:10.1002/2014WR015427.

Wagener, T., N. McIntyre, M. J. Lees, H. S. Wheater, and H. V. Gupta (2003), Towards reduced uncertainty in conceptual rainfall-runoff modelling: Dynamic identifiability analysis, *Hydrol. Processes*, *17*(2), 455–476.

Wang, K., and R. E. Dickinson (2012), A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, *Rev. Geophys.*, *50*, RG2005, doi:10.1029/2011RG000373.

Westra, S., M. Thyer, M. Leonard, D. Kavetski, and M. Lambert (2014), A strategy for diagnosing and interpreting hydrological model nonstationarity, *Water Resour. Res.*, *50*, 5090–5113.

Wilson, K. B., P. J. Hanson, P. J. Mulholland, D. D. Baldocchi, and S. D. Wullschleger (2001), A comparison of methods for determining forest evapotranspiration and its components: Sap-flow, soil water budget, eddy covariance and catchment water balance, *Agric. For. Meteorol.*, *106*(2), 153–168.

Zhang, L., W. R. Dawes, P. G. Slavich, W. S. Meyer, P. J. Thorburn, D. J. Smith, and G. R. Walker (1999), Growth and ground water uptake responses of lucerne to changes in groundwater levels and salinity: Lysimeter, isotope and modelling studies, *Agric. Water Manage.*, *39*(2–3), 265–282.

CHAPTER 6 Use of a Scenario-Neutral Approach to Identify the Key Climate Attributes that Impact Runoff from a Natural Catchment (Paper 5)

Statement of Authorship

Title of Paper	Use of a scenario-neutral approach to identify the key climate attributes that impact runoff from a natural catchment
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☑ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	N/A Update: submitted to Journal of Hydrology, in review.

Principal Author

Name of Principal Author (Candidate)	Danlu Guo
Contribution to the Paper	Designed experiments and conducted analysis, wrote manuscript.
Overall percentage (%)	85%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 19/12/2016

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Seth Westra
Contribution to the Paper	Suggested the scope and contribution of study, helped to interpret results, provided feedbacks for manuscript.
Signature	Date 19/12/16

Name of Co-Author	Holger Maier
Contribution to the Paper	Suggested the scope and contribution of study, helped to evaluate manuscript.
Signature	Date 20/12/16

Please cut and paste additional co-author panels here as required.

Abstract

Scenario-neutral approaches are being used increasingly for assessing the potential impact of climate change on water resource systems, as these approaches allow the performance of these systems to be evaluated independently of climate change projections. However, practical implementations of these approaches are still scarce, with a key limitation being the ability to generate a range of plausible future time series of hydro-meteorological data. In this study we apply a recently developed inverse approach to stochastic generation to support the scenario-neutral analysis, and thus identify the key hydro-meteorological variables to which the system is most sensitive. The stochastic generator simulates 'perturbed' hydro-meteorological time series which represent plausible future changes in (1) the average, extremes and seasonal patterns of rainfall; and (2) the average values of temperature (T_a), relative humidity (RH) and wind speed (u_z) as variables that drive PET. The perturbed hydro-meteorological time series were then fed through a conceptual rainfall-runoff model to simulate the potential changes in runoff as a function of changes in the climate variables using both correlation and Sobol' sensitivity analyses. The method was applied to a case study catchment in South Australia, and the results showed that the most important climate attributes for runoff were winter rainfall followed by the annual average rainfall, while the PET-related climate variables show comparatively little importance. The high importance of winter rainfall can be related to the winter-dominated nature of both the rainfall and runoff regimes in this catchment. The approach illustrated in this study can greatly enhance our understanding of the key climate attributes and the key processes that are likely to drive catchment runoff under a changing climate, thus enabling the design of tailored climate impact assessments to specific water resource systems.

6.1 Introduction

Scenario-neutral approaches are increasingly being used to assess the potential impact of climate change on the performance of water resources systems (Dessai and Hulme, 2004, Brown et al., 2012, Brown and Wilby, 2012, Nazemi and Wheater, 2014). Scenario-neutral approaches generally involve the stress testing of specific water resource systems against plausible future climate conditions, which can thus provide useful information for assessing system vulnerability under alternative climate change scenarios, and for defining climatic thresholds at which system performance begins to change abruptly (Brown et al., 2011, Poff et al., 2015). Another unique feature of scenario-neutral approaches is their ability to identify the hydro-meteorological variables that have the greatest impact on the specific water resource system under consideration. This feature is particularly useful for developing a system-tailored study design that can involve the selection of: (1) climate models that illustrate higher performance in predicting these key variables (e.g. Flato et al., 2013, CSIRO and Bureau of Meteorology, 2015, Johnson et al., 2011); (2) strategies to improve the generation of fine-scale future rainfall and climate conditions from GCM-based projections (known as statistical downscaling) (e.g. Johnson and Sharma, 2011); and (3) alternative 'lines of evidence' (e.g. expert opinion and data from the paleo-climatic record) that can provide additional understanding of the possible variations in these key variables (e.g. Ault et al., 2014, Ho et al., 2015). Ultimately, scenario-neutral approaches can support the development of a tailored set of projections for the specific water resource system of interest, by suggesting suitable models to predict the future changes in the key hydro-climatic variables and attributes for the system, as well as better approaches to link these changes to the water resource system (Vano et al., 2015, Steinschneider and Brown, 2013, Singh et al., 2014, Nazemi et al., 2013).

Although the general principles underpinning scenario-neutral approaches have been well established (Dessai and Hulme, 2004, Nazemi and Wheater, 2014, Prudhomme et al., 2010, Brown et al., 2012), practical implementations have only appeared in the literature relatively recently (Brown et al. 2012,

Prudhomme et al. 2010, Prudhomme et al. 2013a, Poff et al. 2015, Kay et al. 2014, Singh et al. 2014, Culley et al. 2016). A key challenge in the implementation of scenario-neutral approaches is the generation of a set of plausible climate conditions (referred to as the 'exposure space'; see Culley et al, 2016) to which a system might be exposed in the future. Ideally, this exposure space should consider a range of possible variations not only in the average states of the relevant hydro-meteorological variables, such as annual average rainfall and potential evapotranspiration (see Kay et al. 2014, Prudhomme et al. 2013a), but also a number of other attributes for each variable, including extremes, seasonality and inter-annual variability (Meselhe et al., 2009, Moody and Brown, 2013, Prudhomme et al., 2010, Steinschneider and Brown, 2013). In this way, the sensitivity of water resource systems can be tested against a comprehensive range of potential climate changes that can be expected in a greenhouse gas-enhanced climate (Prudhomme et al. 2013a, Steinschneider and Brown 2013).

Despite the benefits of considering a wide range of hydro-meteorological variables and a variety of their attributes, in most previous scenario-neutral studies climate exposure spaces have generally been produced with perturbations on a small number of hydro-meteorological variables. In particular, most studies have relied on perturbing annual and/or month average rainfall and potential evapotranspiration through the use of simple scaling factors (Kay et al. 2014, Prudhomme et al. 2013a, Prudhomme et al. 2010, Paton et al., 2013, Singh et al. 2014). To expand the applicability of scenario-neutral approaches to investigate the implications of changes not only to the averages but also to other attributes of the climate variables of interest, the use of stochastic generators has been proposed. This has been illustrated in Whateley et al. (2014), in which the parameters of a weather generator were perturbed, followed by quantile correction of the generated time series, to achieve a set of pre-specified 'target' levels of climate statistics. A challenge with this approach, however, arises with the difficulty to assess *a priori* which parameters of the stochastic generator should be modified to produce hydro-meteorological time series that represent the different target

statistics, which potentially leads to insufficient exploration of the exposure space.

A potential way to address this issue was proposed by Guo et al. (2016a), who developed an 'inverse' approach to enable stochastic generation of time series that represent pre-specified changes of each hydro-meteorological variable of interest. In this way, the inverse approach enables a range of plausible future climate conditions to be explored, and thus can be used to provide uniform coverage of the exposure space and to serve the needs of comprehensive scenario-neutral climate impact assessments. Although the approach is applicable to perturb any hydro-meteorological variable(s) that can be generated by parametric stochastic generators, Guo et al. (2016a) focused only on perturbing four attributes of rainfall (i.e. average daily rainfall, annual wet days, average dry-spell length and the 99^{th} percentile of daily rainfall). Furthermore, the approach has not yet been utilized as part of a formal application of the scenario-neutral approach. This study aims therefore to extend the applicability of the inverse approach of Guo et al. (2016a) to a wider set of climate variables, and to apply the perturbed hydro-meteorological time series to test the climate sensitivity of a water resource system. This enables the identification of the key hydro-meteorological variables and attributes that the system is particularly sensitive to as part of the scenario-neutral framework.

In this study, implementation of the scenario-neutral approach was illustrated with a natural catchment located in South Australia, and we aimed to identify the key climate attributes that impact catchment runoff by testing the sensitivity of runoff to a range of plausible future changes in hydro-meteorological conditions. To better represent these changes, we considered a large number of plausible changes in both the rainfall and potential evapotranspiration (PET), as they are the two key variables through which potential changes in climate can influence catchment runoff. In achieving this, we perturbed climate attributes including average annual rainfall, extreme rainfall and the key seasonal patterns, as well as the driving variables of PET—namely temperature, relative humidity and wind speed. The perturbed

time series of climate variables were translated to PET using the Penman-Monteith model. The perturbed rainfall and PET time series were then used to simulate catchment runoff with a lumped conceptual rainfall-runoff model, GR4J (Perrin et al., 2003) that has been extensively tested in this catchment (Westra et al, 2014b). The relationships between runoff responses and perturbations in the rainfall attributes and PET-related climate variables were then assessed using both correlation analysis and Sobol' sensitivity analysis (Sobol' et al., 2007). These results allowed us to identify the key climate attributes that impact runoff from the case study catchment.

The subsequent sections of this paper are structured as follows. Section 6.2 introduces the case study catchment and data used to illustrate the scenario-neutral approach. Section 6.3 describes the methodology used to implement the scenario-neutral approach to identify the key climate attributes for the case study catchment, which details the generation of climate exposure spaces with the inverse approach, followed by a climate stress test on the catchment runoff. Section 6.4 presents and discusses two sets of results regarding (i) the performance of the inverse approach in terms of its ability to generate the desired climate exposure space; and (ii) the key climate attributes identified to have high impact on catchment runoff. The study is summarized and concluded in Section 6.5.

6.2 Case study and data

To illustrate the implementation of the scenario-neutral approach aimed at understanding the climate sensitivity of a water resource system, we focused on a case study for the Scott Creek catchment (29 km^2) in the southern Mount Lofty Ranges close to Adelaide, South Australia. Scott Creek is a sub-catchment of the Onkaparinga Catchment and has a Mediterranean climate. The mean annual rainfall and PET are 892 mm and 1372 mm, respectively, for the study period from 1995 to 2004. The catchment has a winter-dominated rainfall regime, with over 75% of the rainfall occurring within the period from July to September (winter and early spring) in an average year. Furthermore, the distribution of runoff is highly skewed, with approximately

30% of the total flow volume contributed from the top 1% of flow days (Westra et al., 2014b).

As part of the scenario-neutral approach, a number of relevant climate attributes should be selected to construct the climate exposure space (Prudhomme et al. 2013a). For this case study catchment, we focused on six climate attributes to describe a range of plausible changes in rainfall and PET, as defined in Table 6-1. To represent various rainfall patterns, we focused on variations in the average wet-day rainfall (*PD*) as a measure of average rainfall, and the 99th percentile of wet-day rainfall (*Pex99*) as a measure of extreme rainfall. In addition, winter rainfall (*PJJA*) was considered, since the catchment has a winter-dominated rainfall regime. The potential changes in daily average temperature (T_a), relative humidity (*RH*) and wind speed (u_z) were included in the exposure space, since they were identified as the key driving variables for PET at the case study location (Guo et al., 2017). The average values of each attribute calculated over the study period were used as the baseline, and are also provided in Table 6-1.

Table 6-1: Definitions and baseline values of the six climate attributes included in the exposure space.

Climate attribute	Definition	Average baseline value (1995-2004)
PD	Daily rainfall intensity averaged over all wet days	6.38 mm/day
Pex99	99th percentile of daily rainfall over wet days	40.0 mm/day
PJJA	Daily rainfall intensity averaged over winter (June, July and August)	7.28 mm/day
T_a	Daily average temperature	17.0 ◦C
RH	Daily average relative humidity	61.2 %
u_z	Daily average wind speed	3.16 m.s^{-1}

As mentioned in the Introduction, GR4J was used to simulate the response of catchment runoff to various hydro-meteorological conditions included in the climate exposure space, as it has been shown to provide a good representation of underlying physical processes for the case study catchment (Guo et al., 2017). To calibrate the model, historical rainfall and runoff data were obtained at the case study catchment, and the PET data were estimated

using the Penman-Monteith model with historical climate data of temperature, relative humidity, solar radiation and wind speed, using the R package *Evapotranspiration* (http://cran.r-project.org/web/packages/Evapotranspiration/index.html) (Guo et al., 2016b). Since over 80% of the catchment area is covered by grass (Evans and Jakeman, 1998), the evaporative surface for the entire catchment was assumed to be reference crop, enabling the FAO-56 version of the Penman-Monteith model (Allen et al., 1998) to be used. The data sources are detailed as below:

1. **Catchment-average rainfall (mm)**: Since the catchment is relatively small, daily rainfall data were obtained from a rain gauge within the catchment and assumed to represent the catchment-average rainfall.
2. **Daily maximum and minimum temperature (T_{max} *and* T_{min} in °C), maximum and minimum relative humidity (RH_{max} *and* RH_{min} in %) and wind speed (u_z in m/s):** Due to the limited availability of high-quality climate data, data on each of these variables were obtained directly from the Kent Town weather station, which is located approximately 20 km from the catchment outlet.
3. **Daily solar radiation (R_s in MJ/(m^2.day)):** Daily solar radiation was calculated from daily sunshine hour data (n in hrs) obtained from the Kent Town weather station, using the Ångström-Prescott equation (McMahon et al., 2013), with constants provided in Chiew and McMahon (1991).
4. **Catchment runoff (ML/day)**: Daily runoff data were obtained from the gauging stations at the outlet of the Scott Creek catchment.

With the abovementioned data, GR4J was calibrated following a 70:30 split-sample calibration (Klemeš, 1986), which led to a calibration period

from 1995 to 2001, and a validation period between 2002 and 2004. The model illustrated satisfactory performance over both periods for calibration (with a NSE of 0.873 and a relative bias of 0.007) and validation (with a NSE of 0.855 and a relative bias of 0.009). In addition to calibrating the rainfall-runoff model, the climate data from the combined calibration and validation period were used as baseline to produce climate perturbations for the climate exposure space, and then to simulate the corresponding changes in runoff with the calibrated rainfall-runoff model.

To assess the sensitivity of runoff with the scenario-neutral approach, we focused on five key attributes representing different features of runoff for the Scott Creek catchment; namely the average daily runoff for all days (*Qavg*), a measure of peak flow as the 99th percentile of daily runoff (*Q99*), a measure of low flow as the 10th percentile of daily runoff (*Q10*), and the average daily runoff for winter and spring, which contribute to majority of the annual flow (*QJJA* and *QSON*). The definition of each runoff attribute and the corresponding baseline value over the study period is shown in Table 6-2.

Table 6-2: Definitions and baseline values of the five runoff attributes selected to represent the runoff response for the case study catchment.

Runoff attribute	Definition	Average baseline value
Qavg	Daily average runoff over all days	0.367 mm
Q99	99th percentile of daily runoff over all days	5.014 mm
Q10	10th percentile of daily runoff over all days	0.006 mm
QJJA	Daily runoff averaged over winter (June, July and August)	0.912 mm
QSON	Daily runoff averaged over spring (September, October and November)	6.3 mm

6.3 Methods

6.3.1 Overview

A schematic of the approach followed in this study is shown in Figure 6-1. We first generated the climate exposure space, consisting of a large number of combinations of plausible changes in the six selected climate attributes related to rainfall and PET (as summarized in Table 6-1, which are visualized as three arbitrary attributes—***A***, ***B*** and ***C***—in Figure 6-1). This was achieved by

extending the inverse approach from Guo et al. (2016a), which allowed perturbation of hydro-meteorological time series using a stochastic generator (Section 6.3.2). With the exposure space we then conducted a climate stress test for the runoff from the case study catchment. We first used GR4J to simulate the responses of the five runoff attributes (as summarized in Table 6-2) to all climate perturbations obtained from the inverse approach. The runoff responses were then assessed to identify the key climate attributes for catchment runoff (Section 6.3.3), using two separate analyses:

1) The Spearman's rank coefficient correlations were first estimated between the runoff attributes and the climate attributes included in the climate exposure space, to detect any association between each runoff attribute and each climate attribute; and
2) A Sobol' sensitivity analysis (Sobol' et al., 2007, see Appendix 6A.1. for details) was then used to assess the relative importance of each climate attribute for each runoff attribute, with the aid of the Sobol' first-order and total-order sensitivity indices. For each runoff attribute, the first-order index for each climate attribute represents the portion of total variance in the runoff attribute that is contributed solely by this particular climate attribute, whereas the total-order index represents the total effects of this climate attribute plus all its interactions with other attributes.

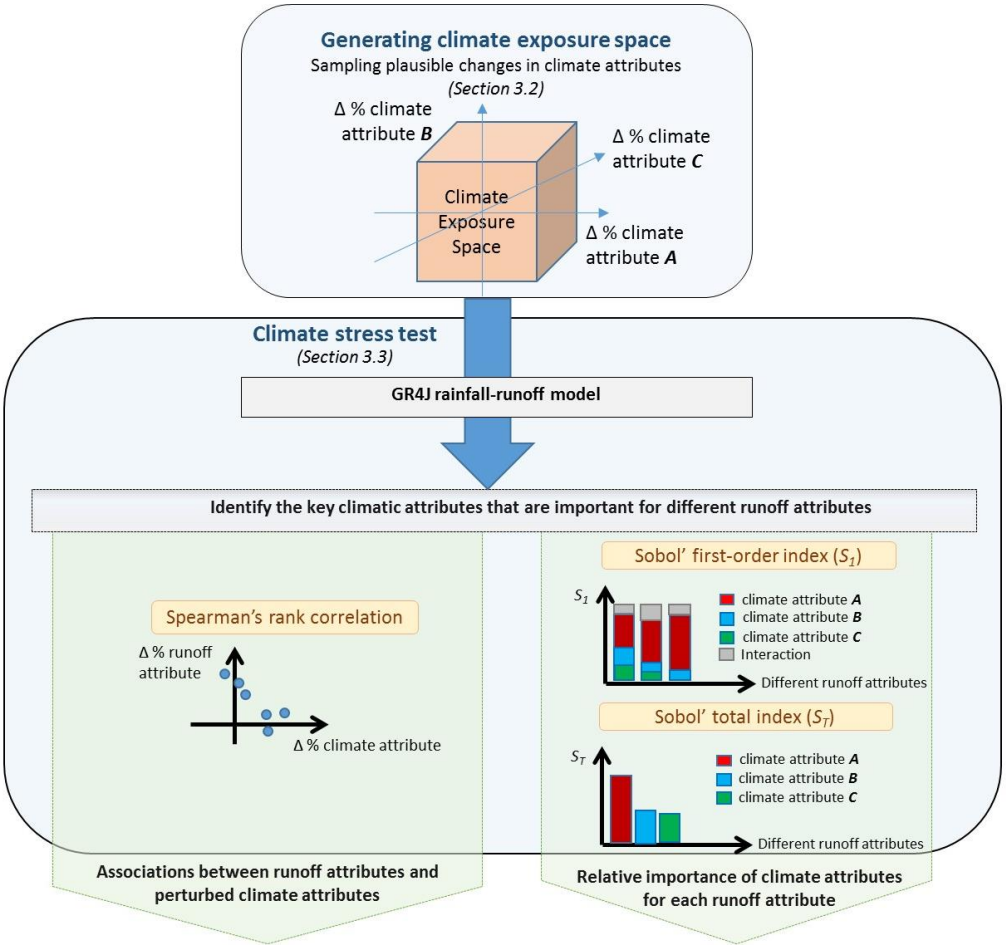


Figure 6-1: Schematic of the method used in this study. Note that although the dimension of the climate exposure space was six (i.e. consisting of six climate attributes), it is represented in three dimensions corresponding to three arbitrary subsets *A*, *B* and *C* for ease of visualization in this figure.

6.3.2 Generation of climate exposure space

To generate the climate exposure space, which is made up of various plausible changes in the six climate attributes (Table 6-1), we followed the inverse approach proposed in Guo et al. (2016a), with an extension to enable consideration of the three additional climate variables that affect PET (i.e. T_a, RH and u_z, as summarised in Table 6-1). The inverse approach involves two key steps:

(1) Identification of a large number of combinations of various 'target levels' for each climate attribute to include in the exposure space; and

(2) Generation of hydro-meteorological time series that correspond to each target location within the exposure space.

Further details on each of these steps are provided in the sections below.

6.3.2.1 Selection of the combinations of target levels in the exposure space

The first step in identifying the target values of each attribute is to specify the sampling bounds, which define the ranges within which to perturb each climate attribute. These were determined based on recent projections of future climate conditions for the case study location by the year 2090 (CSIRO and Bureau of Meteorology, 2015, IPCC, 2014) (Table 6-3). The bounds are slightly wider than suggested from most climate change projections, to ensure that a comprehensive range of plausible future conditions is included in the exposure space.

Table 6-3: Sampling bounds of the six climate attributes included in the exposure space.

Climate attribute	Sampling bounds
PD	-30 – +10%
Pex99	0 – +30%
PJJA	-40 – 0%
T_a	+2 – 6 ◦C
RH	-10 – +5%
u_z	-10 – 0%

Within the plausible bounds, a number of samples were drawn to define different combinations of target levels within the climate exposure space. Latin hypercube sampling (LHS) was used as the sampling method, due to its effectiveness in covering multi-dimensional sampling spaces (Osidele and Beck 2001, Sieber and Uhlenbrook 2005, Tang et al. 2007b). It is worth noting that the two analyses to assess runoff responses to these climate perturbations (i.e. the correlation analysis and the Sobol' analysis, as mentioned in Section 6.3.1) required different numbers of samples to be included in the climate exposure space. For the correlation analysis, we used a sample size of 1000, which was considered to be sufficient to illustrate

correlation patterns between the runoff attributes and the climate attributes. For the Sobol' sensitivity analysis, the sample size should be sufficient to ensure the convergence of the estimates of the Sobol' first-order and total-order sensitivity indices (Nossent et al., 2011, Zhang et al., 2015). Zhang et al. (2015) further defined the condition for convergence as the number of samples for which, for the most sensitive input variable, the width of 95 % confidence intervals from bootstrap resampling of both of its first-order and total-order indices are below 10 % of the corresponding means obtained from bootstrapping.

To determine the sample size within the climate exposure space for the Sobol' analysis, we generated different numbers of LHS samples, which were then translated to the corresponding numbers of hydro-meteorological time series with the approach detailed in Section 6.3.2.2. To check for convergence, for each sample size we derived the 1000-fold bootstrap estimates for both the first- and total-order Sobol' indices (as per Eqns. 1.2 and 1.5 in Appendix 6A.1, respectively) for the most sensitive climate attribute. It was observed that convergence for all Sobol' indices occurred when the LHS sample size exceeded 2400, which was therefore used as the LHS sample size for the Sobol' sensitivity analysis. Note that to estimate Sobol' first- and total-order indices, these 2400 LHS samples are required to be re-sampled to form a Sobol' sequence with 19200 samples, as detailed in Appendix 6A.1.

6.3.2.2 Optimization-based generation of hydro-meteorological time series to achieve target levels

To generate hydro-meteorological time series corresponding to each target location in the exposure space, the stochastic weather generator WGEN (Richardson, 1981) was used. This generator has been widely applied for climate impact studies, and has generally been shown to capture most of the key features of daily rainfall series (Katz, 2002, Kim et al., 2007, Bastola et al., 2011a).

WGEN uses a Markov chain model to stochastically generate daily rainfall occurrence, and a parametric distribution model to simulate rainfall intensity on wet days. The original WGEN model consists of individual monthly models for occurrence and intensity, whereas in this study, seasonal models were used to minimize the number of model parameters. For other climate variables, such as temperature, the model divides every year into 26 intervals, and generates each climate variable according to its specific distributions on wet/dry days in each interval. Therefore, to generate the climate attributes included in the exposure space, the following WGEN parameters were required:

- 16 parameters for rainfall, consisting of four parameters for each season: two parameters of the 1st order two-state Markov chain used for representing the transition probabilities of rainfall occurrence: p_{dd} (dry–dry probability) and p_{wd} (wet–dry probability), and the rate and shape parameters (α and β) for the Gamma distribution of rainfall intensity on wet days;
- 312 parameters to define the distributions of the three PET-related climate variables, namely: one parameter for the mean and one parameter for the standard deviation for each of the three variables, for each of the dry and wet days and during each of the 26 intervals; and
- Nine parameters defining the correlation structures between the three PET-related climate variables, namely: six parameters for the lag-1 cross-correlations and three parameters for the lag-0 cross-correlations. Note that these correlations are for the residuals of the three climate variables after accounting for their individual distributions for each of the 26 intervals.

In order to generate hydro-meteorological time series corresponding to each target location (determined through the sampling process outlined in Section 6.3.2.1), the best-fit parameter set for the WGEN model was identified using optimization. A genetic algorithm (GA) was used as the optimization engine, due to its proven efficiency, particularly for solving high-dimensional optimization problems in hydrological studies (Ndiritu and Daniell, 2001, Cheng et al., 2002, Shafii and De Smedt, 2009, Gibbs et al., 2012).

The objective function to be minimized was:

$$F_{obj,i} = \sum \sqrt{\left[\frac{\left(ATT_{s,i}^{j} - ATT_{his}\right)}{ATT_{his}} - \frac{\left(\widehat{ATT_{s,i}^{j}} - ATT_{his}\right)}{ATT_{his}}\right] * 100} \quad \textbf{(6.1)}$$

In Eqn. 6.1, $j = 1, 2, \ldots 6$ for each of the six climate attributes considered in the exposure space (Table 6-1), and $i = 1, 2, \ldots n$ for n combinations of target levels in the exposure space, which corresponds to the sizes of the two exposure spaces we constructed for the correlation and Sobol' analyses (1000 and 19200 respectively, as mentioned in Section 6.3.2.1). For the j^{th} climate attribute in the exposure space (ATT^{j}), $ATT_{s,i}^{j}$ represents the i^{th} target value and $\widehat{ATT_{s,i}^{j}}$ represents the corresponding simulated value from the stochastic generator. Since different attributes are likely to have different magnitudes, the difference between a target level and the corresponding simulated level is represented as a percentage change relative to its baseline value (ATT_{his}^{j}) (Table 6-1) to ensure consistent scales across attributes during the optimization process.

When simulating the rainfall time series, since only changes in rainfall intensity were included in the exposure space (Table 6-1), the probability parameters (p_{dd}'s and p_{wd}'s) that determine the rainfall occurrences were fixed at historical levels during the optimization process. To simulate the rainfall intensity for each season, both non-negative Gamma parameters (α's and β's) required were varied between 10^{-2} and 10 during the optimization process. This range was identified from preliminary experiments as the optimal range

for 1) quick convergence of the optimization algorithm; and 2) generation of realistic daily rainfall intensity values.

For temperature, relative humidity and wind speed, since only changes in average conditions are of interest for the exposure space (Table 6-1), during the optimization only the mean values for wet days and dry days for each attribute were varied within the corresponding plausible ranges, as defined in Table 6-3. To minimize the number of parameters to optimize, all other parameters, namely the standard deviation and the correlations among residuals, were fixed at corresponding historical values.

The optimization for each target location proceeded as follows. A wet-/dry-day sequence was first generated with the historical rainfall occurrence parameters of WGEN. Following this, four sequential optimization steps were conducted to search for the remaining WGEN parameters, aiming to fit specific levels of different attributes. These steps consisted of:

(1) Optimizing for the best α and β parameters from the gamma distribution, which yielded the target levels of both *PD* and *Pex99* simultaneously. It is worth mentioning that the winter β used for each iteration was determined with an additional layer of optimization to achieve the target level of *PJJA*, while all other α and β values were sampled randomly;

(2) Optimizing for the best mean wet- and dry-day temperature for WGEN, which yielded the target level of T_a;

(3) Optimizing for the best mean wet- and dry-day relative humidity for WGEN, which yielded the target level of *RH*; and

(4) Optimizing for the best mean wet- and dry-day wind speed for WGEN, which yielded the target level of u_z.

The convergence criterion was set to a value of 0.1 for the objective function (Eqn. 6.1). In addition, as suggested in Guo et al. (2016a), during

these optimization processes, the random seed of the stochastic generator was held constant. This was due to the consideration that the stochastic generator can introduce random behavior to the generated hydro-meteorological time series, which can mislead the optimization algorithm and thereby slow down the optimization process. Therefore, fixing the random seed ensured that any changes in objective function values from one iteration of the optimization process to the next were solely due to changes in the optimization decision variables (i.e. the model parameters), rather than a combination of these changes and any randomness introduced by the stochastic generator.

Using the best-fit parameters of WGEN obtained from the optimization, the 1000 sets (for the Spearman correlation analysis) and the 19200 sets (for the Sobol' sensitivity analysis) of perturbed hydro-meteorological time series were generated. The perturbed daily time series of rainfall from WGEN were used as a direct input to the calibrated GR4J model. The perturbed daily time series for the three PET-related variables (temperature, relative humidity and wind speed) were used to estimate daily PET with the Penman-Monteith model, and then fed into the calibrated GR4J. It is worth noting that the Penman-Monteith model requires both daily maximum and minimum time series for temperature and relative humidity as inputs. Therefore, the daily perturbations to each pair of temperature and relative humidity variables were determined by the changes in the corresponding perturbations in average temperature and relative humidity relative to the baseline levels. To ensure physical plausibility of the perturbations, the daily maximum and minimum values of relative humidity were capped at 100%. In addition, since solar radiation was not a climate attribute included in the exposure space, the historical solar radiation time series were used for estimating all sets of perturbed PET time series.

6.3.3 Climate stress test on catchment runoff

With the perturbed rainfall and PET time series obtained from Section 6.3.2, the calibrated GR4J model (Section 6.2) was run to simulate the corresponding catchment runoff. To identify the key climate attributes that influence catchment runoff, we first used scatter plots to visualize the

association between the runoff responses and the perturbations in each climate attribute included in the exposure space using the 1000 samples. The Spearman's rank correlation coefficient between each runoff attribute and each climate attribute was also calculated to indicate the strength of the relationship between attributes. We then conducted Sobol' analysis with the exposure space with 19200 samples and the corresponding modelled runoff responses, to assess the relative importance of each climate attribute in influencing each runoff attribute. Both the Sobol' first- and total-order sensitivity indices for each runoff attribute were estimated (as Eqns. 1.2 and 1.5 in Appendix 6A.1., respectively), which indicates how much the total variance of the runoff was due to the individual contribution of each individual climate attribute and how much of that was due to the interactions among multiple climate attributes.

6.4 Results

6.4.1 Performance of inverse approach

The performance of the inverse approach in terms of its ability to generate the desired climate exposure space (Section 6.3.2) is shown in Figure 6-2, as pairs plots of the sampled changes relative to the baseline level for each of the six climate attributes (as specified in Table 6-1). We illustrate this with our first exposure space with 1000 samples, as it consists of fewer samples than that used for the Sobol' analysis (19200 samples), and therefore allows for clearer visualization. As can be seen, the inverse approach is effective in producing the desired target locations, with all samples falling within the bounds of the exposure space (as defined in Table 6-3). In addition, the coverage of the climate exposure space is relatively uniform, with low correlations between each pair of climate attributes. Therefore, the exposure spaces generated with the inverse approach enable us to assess the sensitivity of catchment runoff to a wide range of plausible climate conditions.

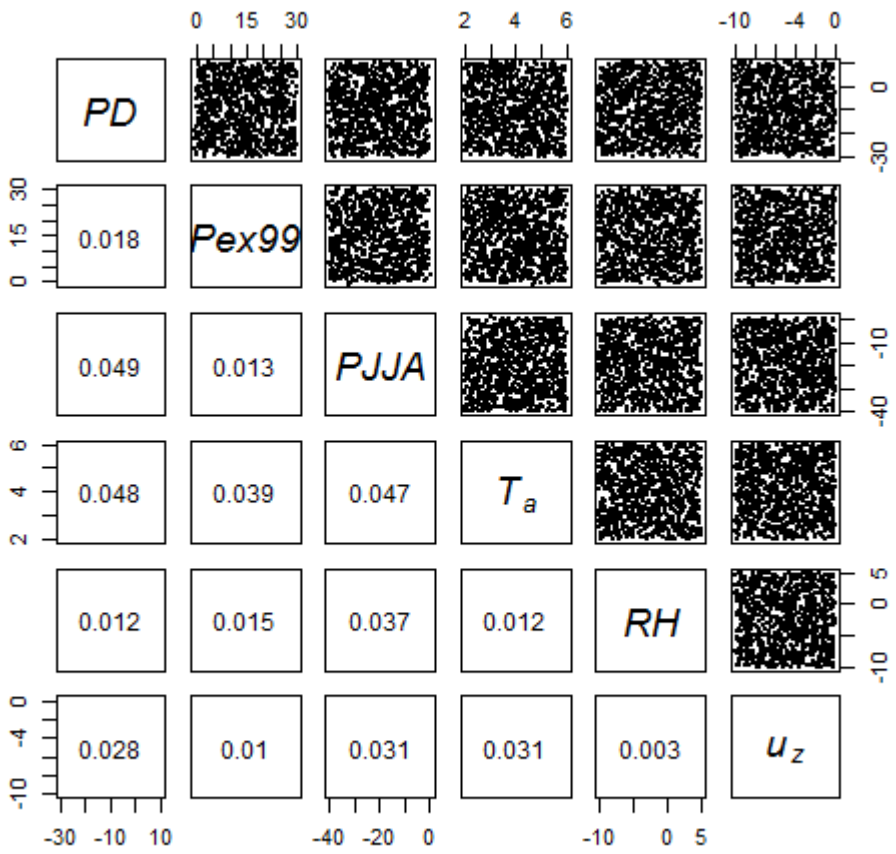


Figure 6-2: Climate exposure space consisting of 1000 LHS samples generated with the inverse approach, described as a percentage change for *PD*, *Pex99*, *PJJA*, *RH* and u_z, and absolute changes for T_a (in ◦C), relative to the corresponding baseline levels (Table 6-1). The lower-right triangle displays pairwise Spearman's correlation coefficients.

6.4.2 Key climate attributes for catchment runoff

6.4.2.1 Spearman's rank correlation coefficients

The responses of the five runoff attributes to the 1000 sets of perturbed climate attributes within the exposure space are plotted in Figure 6-3, as percentage changes relative to the corresponding baseline levels. The correlation between each runoff attribute and each climate attribute is summarized in Table 6-4 in terms of Spearman's rank correlation coefficients (ρ), with shaded cells indicating correlations that are significant at a 0.05 level ($p < 0.05$). Overall, the runoff attributes show strongest correlations with

PJJA, followed by the *PD*. Also, common to all runoff attributes is that the three climate attributes related to PET (*T*, *RH* and u_z) generally have very low correlations with the runoff attributes, with the absolute magnitudes of most ρ less than 0.1.

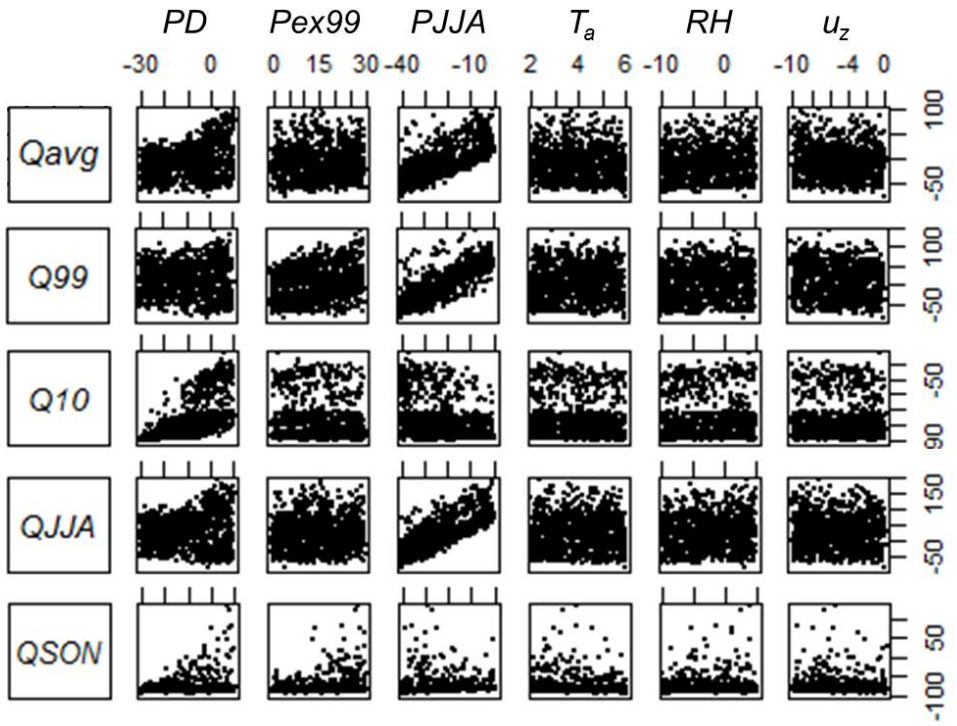


Figure 6-3: Responses of the five runoff attributes (as percentage changes relative to the baseline levels) to 1000 perturbations of the six climate attributes considered in the exposure space.

Table 6-4: Spearman's rank correlation coefficients (ρ) between the five runoff attributes and the six climate attributes considered in the exposure space, as plotted in Figure 6-3. Shaded cells indicate that correlations are significant at the 0.05 level.

	PD	***Pex99***	***PJJA***	***T_a***	***RH***	***u_z***
Qavg	0.251	0.050	0.718	-0.048	0.079	-0.066
Q99	-0.097	0.228	0.819	-0.034	0.037	-0.078
Q10	0.802	-0.151	-0.430	-0.092	0.062	0.020
QJJA	0.106	0.039	0.784	-0.012	0.060	-0.064
QSON	0.143	0.112	0.472	-0.204	0.060	-0.065

Table 6-4 shows that the Spearman correlation coefficients vary across different runoff attributes. However, a common pattern is observed within the

runoff attributes that are related to high flows, as they all display strong associations with winter rainfall. Specifically, *Qavg*, *Q99* and *QJJA* all show high correlations with *PJJA* (with ρ values of 0.718, 0.819 and 0.784, respectively, which are all significant at the 0.05 level), with strong near-linear relationships illustrated in Figure 6-3. Following *PJJA*, the second most influencing climate attribute is *PD* for *Qavg* and *QJJA* (with correlation coefficients of 0.251 and 0.106, respectively, significant at the 0.05 level), whereas *Q99* shows a significant correlation with *Pex99* (ρ = 0.228). This indicates that it is the seasonality of the rainfall, rather than the total annual rainfall *per se*, that appears to be the strongest driver for most of the runoff attributes considered.

In contrast to the above finding, as an indicator of low flow, *Q10* illustrates a somewhat opposite pattern to those shown for *Qavg*, *Q99* and *QJJA*, as it displays the highest correlation with *PD* (ρ = 0.802), followed by a moderate negative correlation with *PJJA* (ρ = -0.430). Another contrasting response is observed in *QSON*, which does not show a particularly high correlation with any rainfall attribute. The strongest correlation is with *PJJA*, with a correlation coefficient of only 0.472.

It is worth noting that *Q10* shows negative correlations with *PJJA* and *Pex99*, which is likely to be explained by the method of climate perturbation: as multiple variables are perturbed at the same time during the inverse approach, for a specific average rainfall intensity (*PD*), an increase in winter/extreme rainfall (*PJJA*/*Pex99*) has to be achieved by reducing the low rainfall intensity and/or rainfall during the drier seasons, which is likely to lead to a decrease in low flow (*Q10*).

6.4.2.2 Sobol' sensitivity analysis

Having assessed the correlations between the five runoff attributes and the six climate attributes included in the exposure space, we now quantify the relative importance of each climate attribute for the runoff attributes by using the Sobol' indices. The Sobol' first-order indices of each climate attribute, as well as their interactions, are presented in Figure 6-4, and are plotted against

the five different runoff attributes. The first-order indices for each runoff attribute sum up to one, and describe the partitioning of total variance to the contribution from each individual climate attribute, as well as the total effect of their interactions.

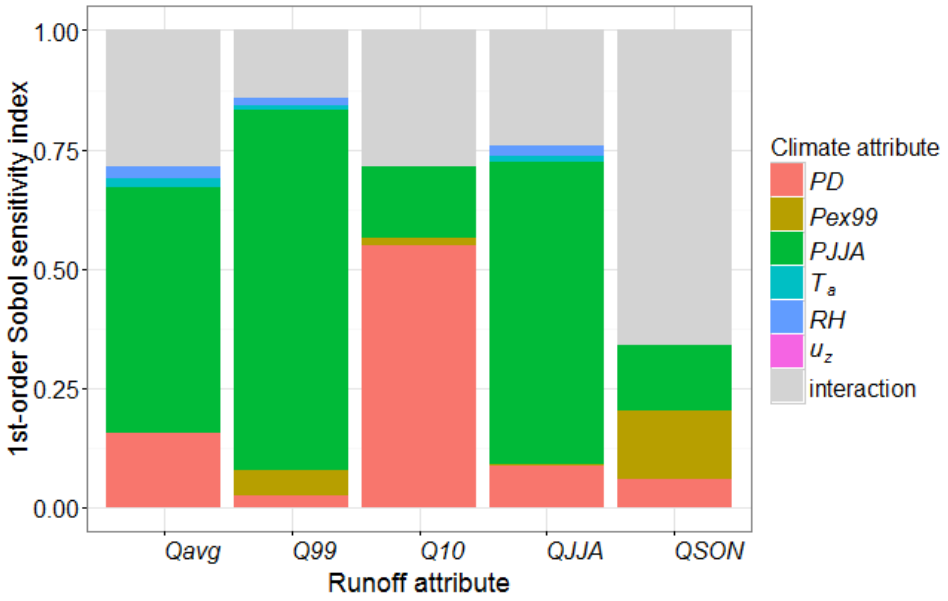


Figure 6-4: Sobol' first-order sensitivity indices of the five runoff attributes to changes in the six climate attributes (colored) and their interaction effects (grey).

Consistent with the correlation results in Section 6.4.2.1, the first-order indices generally suggest that the two most important climate attributes for runoff are *PJJA*, followed by *PD*. Again, the runoff attributes that are related to high flows show similar sensitivity patterns. Specifically, for *Qavg*, *Q99* and *QJJA* the first-order indices of *PJJA* always exceed 0.5 (0.513, 0.757 and 0.634, respectively), indicating that more than half of the variation in the responses of each of these runoff attributes is contributed by perturbations in *PJJA*. Following *PJJA*, *PD* is the second most important variable for *Qavg* and *QJJA* (with index values of 0.161 and 0.086, respectively), while *Pex99* is the second most important variable for *Q99* (with an index value of 0.052).

In contrast to *Qavg*, *Q99* and *QJJA*, the first-order indices for *Q10* suggest that *PD* is the dominant variable (with an index value of 0.548), whereas for *QSON*, the importance of *PD*, *Pex99* and *PJJA* are of similar magnitude (with index values of 0.058, 0.143 and 0.146, respectively). Lastly, *T*, *RH* and u_z all

show first-order indices below 0.03 for all runoff attributes, which again indicates low impact of PET-related changes on runoff.

In addition to providing confirmation of the results obtained using the correlation analysis (Section 6.4.2.1), the Sobol' analysis also provides information on the interactions among the six climate attributes (Eqn. 1.6 in Appendix 6A.1, represented with grey bars in Figure 6-4). As can be seen, the two runoff attributes that are least affected by these interactions are *Q99* and *QJJA*, with both interaction terms having values below 0.25. In contrast, the interactions are slightly higher for *Qavg* (0.285) and *Q10* (0.284), and significantly higher for *QSON* (0.660), with more than 60% of the total variance contributed by interactions, rather than changes in individual climate attributes.

The interaction effects are also reflected in Figure 6-5, which shows the Sobol' total-order sensitivity indices of the six climate attributes for each runoff attribute. The magnitude of the total-order index for each climate attribute represents the sum of both the individual contribution from that climate attribute and all its interaction with other attributes. Comparing these total-order indices for each climate attribute to the corresponding first-order indices in Figure 6-4, it is clear that the most distinct differences are observed for *QSON*, which displays much higher total-order indices to *PD*, *Pex99* and *PJJA* (0.749, 0.328 and 0.588, respectively) compared to the first-order indices (as 0.058, 0.143 and 0.146, respectively). Consistent with Figure 6-4, this highlights the importance of the interactions among the three climate attributes for *QSON*, in which the relative importance of the interaction effects exceed that of any individual attribute.

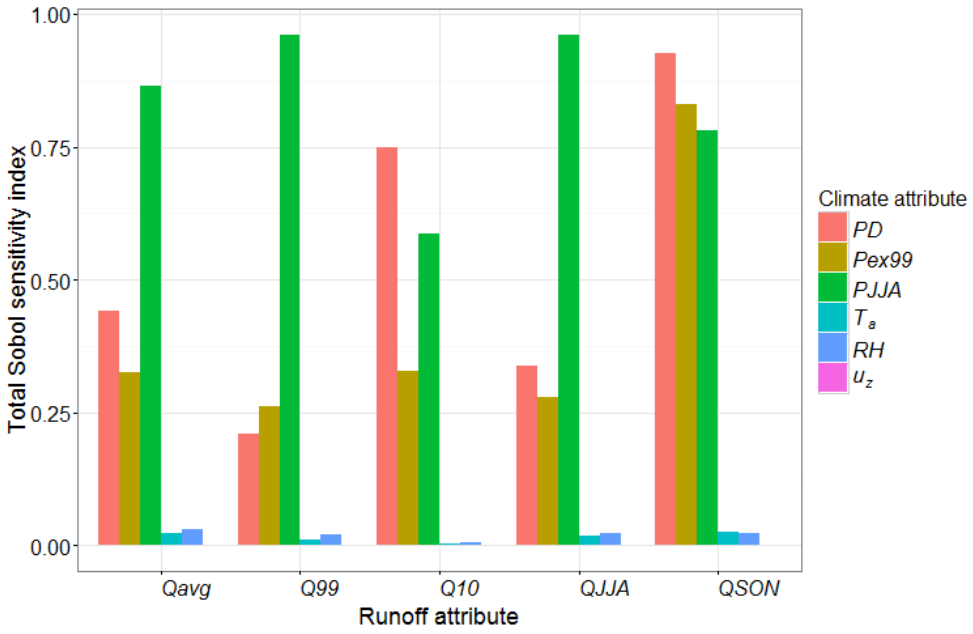


Figure 6-5: Sobol' total-order sensitivity indices of the five runoff attributes to changes in the six climate attributes.

6.5 Discussion

This study sought to expand the scope of scenario-neutral studies to consider possible changes in a broader range of variables and attributes of those variables, to account for not only mean changes in key atmospheric drivers but also the extremes and seasonality. The results from the correlation and Sobol' sensitivity analyses presented in Section 6.4 highlight the relative importance of the various climate attributes considered for simulated runoff generation for the Scott Creek catchment, which clearly reflect the unique catchment characteristics. As illustrated in both analyses, winter rainfall (*PJJA*) was found to be a key climate attribute for this catchment, and has a substantial impact on the average and extreme runoff (*Qavg* and *Q99*) as well as the winter runoff (*QJJA*). The high sensitivity of average runoff to changes in winter rainfall clearly illustrates the winter-dominated nature for both rainfall and runoff in the Scott Creek catchment, which is consistent with previous literature (Westra et al., 2014b, Westra et al., 2014c). The high sensitivity of winter runoff to winter rainfall also indicates that a large portion of rainfall is converted to runoff during winter, leading to a near-linear runoff response to rainfall. Interestingly, the extreme runoff was found to be much

more sensitive to the winter rainfall than to the extreme rainfall (*Pex99*), highlighting that accumulations of continuous rainfall events over winter are likely to have higher impacts on runoff compared to individual extreme rainfall events. In contrast to the abovementioned runoff attributes that are related to the high flows, the low flow (*Q10*) is dominantly driven by average rainfall (*PD*), which reflects the heavy-tailed nature of the distribution of daily runoff.

The results also suggest that potential changes in PET generally have a low impact on runoff from the Scott Creek catchment. This is likely related to the water-limited nature of the catchment, for which the long-term average rainfall is substantially lower than the long-term average PET (McVicar et al., 2010). From a water-balance perspective, since PET represents the upper limit of actual ET (AET), the AET from such a water-limited catchment is likely to be lower than PET for the majority of the year, which leads to low sensitivity of AET to any changes in PET. For this reason, the limited sensitivity to changes in PET might be expected to hold for catchments located in similar climate zones to Scott Creek, but is unlikely to hold in general. In particular, the AET for energy-limited catchments (for which long-term average rainfall exceeds long-term average PET) is more likely to be affected by changes in PET, since water is generally available for evaporation in these catchments (Kay et al., 2013, McMahon et al., 2015). Furthermore, these energy-limited catchments are also likely to experience greater changes in PET due to changes in their driving climate variables (see Guo et al.. 2017, which reported approximately two-fold PET sensitivity in cool-climate energy-limiated locations in Australia, compared to other locations). This highlights the value in the scenario-neutral analysis as a means of identifying the key attributes that would be expected to affect hydrological response for any particular catchment.

The above results also highlight the importance of considering various attributes for each climate variable, such as sub-annual features and extremes, in the design of climate exposure spaces for scenario-neutral approaches. For example, the results suggest that winter rainfall has a much greater impact on

high flows than average rainfall. This is particularly pertinent for Scott Creek catchment, as projections for climate change suggest potentially much stronger seasonal changes compared to annual average changes, with higher projected declines for rainfall during the wet winter months compared to all other seasons (CSIRO and Bureau of Meteorology, 2015, IPCC, 2014). Therefore, considering winter rainfall in the exposure space can warrant a more comprehensive assessment of the potential responses in catchment runoff, compared with if only changes in average rainfall are considered, as is done in most previous scenario-neutral studies (e.g. Kay et al. 2014, Prudhomme et al. 2013a, Prudhomme et al. 2010). Furthermore, inclusion of sub-annual attributes is particularly important when considering catchments with highly seasonal variations in rainfall and streamflow regimes, as they are likely to be driven by different hydrological processes during different seasons and thus respond differently to changes in hydro-meteorological conditions across seasons (e.g. Sorg et al., 2012, Chang and Jung, 2010, Barnett et al., 2005).

It is worth highlighting a limitation within the study results, which further emphasizes the significance of considering sub-annual variations in constructing climate exposure spaces. As highlighted in Figures 6-4 and 6-5, spring runoff (*QSON*) shows substantial sensitivity to interactions among the climate attributes, which exceeds that for any individual attribute. An obvious interpretation of these results is that spring runoff is driven by more complex processes that involve non-linear interactions of the climate attributes, under the assumption that all key climate attributes that influence spring runoff have been included in the climate exposure space. However, an alternative explanation is provided in Figure 6-6, which is an expanded version of Figure 6-3, in which the corresponding *PSON* obtained from each of the 1000 generated rainfall time series used in the correlation analyses is also shown. It is clear that *QSON* shows a strong correlation with *PSON*. This indicates that the changes in *QSON* are likely to be predominantly driven by *PSON*, which changes as a result of perturbing other climate attributes with the stochastic generator. However, since *PSON* was not considered in our design of the climate exposure space and was not sampled together with the other climate

attributes, these impacts from varying *PSON* were not recognized by the Sobol' analyses, but were instead explained as interactions. This example illustrates that the exclusion of relevant climate attributes in the exposure space can lead to misinterpretation of the sensitivity results obtained from scenario-neutral approaches.

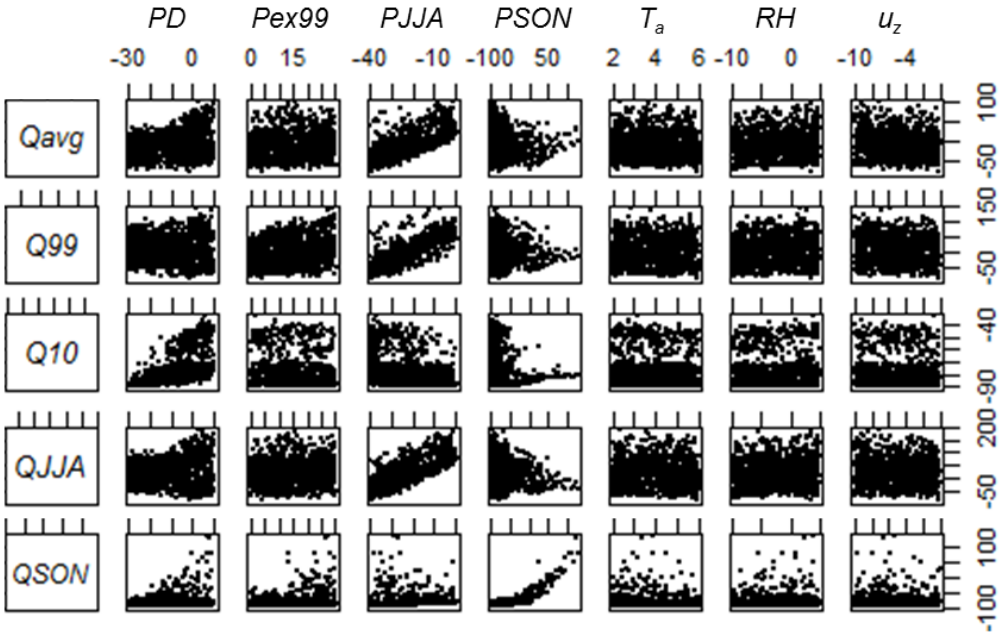


Figure 6-6: Responses of the five runoff attributes (as percentage changes relative to the historical levels) to 1000 perturbations of the six climate attributes considered in the exposure space, as well as to *PSON*, which was modified as a result of perturbing the six attributes.

The abovementioned limitation is partly related to the current capacity of the inverse approach, for which it is challenging to generate climate exposure spaces with high dimensions, especially when a number of attributes of one climate variable are considered. For example, in this study, we have considered three attributes for the single variable of rainfall (i.e. the average and 99th percentile of rainfall intensity, and winter average values). For testing the sensitivity of the catchment to various plausible climate conditions, different combinations of perturbations in these attributes were required, which were obtained by the implementing the staged optimization processes outlined in Section 6.3.2.2. It can be expected that the larger the number of attributes to consider for a specific climate variable, the more constraints will be imposed on the optimization, which would increase the difficulty and

computational effort associated with the successful implementation of the inverse approach. This highlights the need to improve the inverse approach when dealing with such cases.

6.6 Conclusions

In this study, we illustrated a formal implementation of the scenario-neutral approach, to identify the key climatic attributes influencing different runoff attributes from the Scott Creek catchment in South Australia. To achieve this, we first extended the applicability of a recently developed inverse approach to enable stochastic generation of climate exposure space, which consists of perturbed time series of a larger number of climate variables, namely: (1) rainfall to represent plausible future changes in its average conditions, extremes and seasonal patterns; and (2) three climatic variables related to PET (i.e., temperature, relative humidity, and wind speed) to represent plausible changes in their mean values. With these perturbed hydro-meteorological time series, we simulated potential changes in catchment runoff with a conceptual rainfall-runoff model, GR4J. By investigating the relationships between runoff responses and perturbations in each climate attribute using both correlation and Sobol' sensitivity analyses, we identified the key climate attributes that influence different runoff attributes from the catchment considered.

The results from this study show that different runoff attributes are dominated by changes in the rainfall attributes, while the PET-related variables have a relatively minor effect. Specifically:

- The runoff attributes that are related to high-flow, namely average runoff, extreme runoff and winter runoff, show substantial sensitivity to winter rainfall, followed by average rainfall;
- The runoff attribute that is related to low-flow is mainly affected by changes in average rainfall; and
- Spring runoff does not show high sensitivity to any individual rainfall attribute considered in the climate exposure space, but displays a high sensitivity to the interactions among these attributes.

The key climate attributes highlighted for different runoff attributes are closely linked to the unique characteristics of the case study catchment, which exhibits significant seasonal variations in not only the rainfall and runoff, but also the key processes involved in converting rainfall to runoff. These results also highlight the need to design individual implementations of scenario-neutral approaches for specific case studies, as different water resource systems are likely to be driven by different climate attributes and key processes (e.g. Merz and Blöschl, 2009, van der Kamp et al., 2003, Gaál et al., 2012), which can lead to identification of different sets of key climate attributes for each individual system.

In the context of scenario-neutral approaches, the information on key climate attributes can greatly enhance our understanding of the processes that are likely to impact different runoff attributes for a specific water resource system under a changing climate. Such understanding allows us to potentially tailor the design of climate impact assessments for the specific system by focusing on the key climate processes and variables that the system is particularly sensitive to.

Appendix 6A Supplementary to Chapter 6

Sobol' sensitivity analysis (Sobol' et al., 2007)

Sobol' is considered a variance-based method, in which the total variance in a model output due to changes in its inputs is estimated with a Monte-Carlo approach. To estimate the variances, a large number of samples is firstly drawn by varying all input variables at the same time, and then a Sobol' sequence is constructed by re-sampling from within these Monte-Carlo samples (Saltelli et al., 2010). According to Sobol' et al. (2007), to estimate the Sobol' first-order and total-order indices with a Monte-Carlo sample size of n consisting of p input variables, a Sobol' sequence with a total of $n.(p+2)$ samples is required, i.e. with $n.(p+2)$ model evaluations.

The total variance of model output is partitioned to the contribution of each individual input variable (i.e. first-order effects), as well as their interactions (i.e. higher-order effects), as follows (equation adapted from Zhang et al., 2015):

$$V_Y = \sum_{i=1}^{n} V_i \quad + \quad \sum_{i<j} V_{ij} + \sum_{i<j<k} V_{ijk} \ldots + V_{1,2\ldots,n} \tag{1.1}$$

Individual effects **Interactions**

The outputs from the Sobol' method comprise (equations adapted from Nossent et al., 2011):

4) First-order sensitivity index, which quantifies the individual contribution of each input variable to the total variance of the model's output:

$$S_i = \frac{V_i}{V_Y} \tag{1.2}$$

5) Second- and higher-order sensitivity indices, which quantify the contribution of interactions among two or more input variables to the total variance of the model's output:

For second-order: $S_{ij} = \frac{V_{ij}}{V_Y}$ (1.3)

For higher-order: $S_{ij\ldots n} = \frac{V_{ij\ldots n}}{V_Y}$ (1.4)

6) Total sensitivity index, which quantifies the contribution of each input variable, including its individual effect, as well as all its interactions with other input variables, to the total variance of the model's output:

$$S_{Ti} = S_i + \sum_{j \neq i} S_{ij} = 1 - \frac{V_{\sim i}}{V_Y} \quad (1.5)$$

From Eqns. 1.1 to 1.4, the sum of individual effects of all input variables and all their interactions equals one (adapted from Zhang et al., 2015):

$$1 = \sum_{i=1}^{n} S_i \quad + \quad \sum_{i<j} S_{ij} + \sum_{i<j<k} S_{ijk} \ldots + S_{1,2\ldots,n} \quad (1.6)$$

Individual effects **Interactions**

CHAPTER 7 Conclusions

Understanding the impacts of climate change has particular significance for the future planning, design and operation of water resource systems. Scenario-neutral approaches are used increasingly to assess these possible impacts. These approaches allow water resource systems to be assessed independently of climate change projections, instead focusing on the sensitivity of specific systems to a large number of plausible climate change conditions. Once developed, these approaches can be used to better understand water resource system vulnerability, and provides a mechanism to incorporate multiple lines of evidence on possible future climatic changes into the climate impact assessment. This research improves the practical implementation of scenario-neutral approaches by focusing on two key limitations: (1) limited capacity to generate a comprehensive climate exposure space to describe a large number of plausible future climate conditions; and (2) lack of understanding of how the physical process representation in rainfall-runoff models, especially for the evapotranspiration processes, can affect future runoff projections. The contributions of this thesis that seek to address these two limitations are summarised briefly below.

7.1 Research contribution

The overall contribution of this research is the improvement of the practical implementation of scenario-neutral approaches. Therefore, this research provides recommendations for the future implementations of scenario-neutral framework, and thus greatly extends the applicability of this framework to a larger range of climate impact assessment problems.

The specific research contributions are as follows:

1. This research has led to an enhanced methodology for the generation of the climate exposure space to represent a range of plausible future changes in a number of climate variables (e.g. rainfall, temperature, relative humidity, wind speed) and attributes of those variables (e.g. mean, extremes, variability) (Chapters 3, 4 and 6). Specifically,

Chapter 3 identifies the key climate variables that can have big impact on evapotranspiration across different climate zones, which informs the choice of evapotranspiration-related climate variables to include in the climate exposure space. Chapter 4 introduces an inverse approach to generating a range of plausible changes in rainfall, including average conditions as well as extremes and seasonal patterns. This approach is further extended in Chapter 6 which considers a combined set of climate variables and attributes that are capable of describing potential changes in both rainfall and evapotranspiration, for constructing a more comprehensive climate exposure space. By achieving these outcomes, this research overcomes a long-standing problem in scenario-neutral studies, which have tended to focus on only a small subset of variables and attributes that might change in the future.

2. This research improves the understanding of the role of physical process representations in rainfall-runoff models in providing runoff projections. The focus of this research was on the evapotranspiration processes, which has received relatively little attention in existing literature compared to rainfall processes (Chapter 5). The research showed that although the calibration and validation performance of alternative rainfall-runoff models may be similar under historical climate conditions, the contrasting representation of evapotranspiration processes can have a significant impact on future evapotranspiration and runoff projections. These findings highlight that alternative rainfall-runoff models should be carefully assessed based on their process representation to inform model selection in climate impact assessments.
3. This research integrates the abovementioned new knowledge and techniques to implement the enhanced scenario-neutral approach to identify the key climatic variables for a natural catchment (Chapter 6). This approach can be transferrable to identify the key climate attributes for other water resource systems, which can greatly enhance our understanding of the processes that are likely to impact different runoff attributes for specific systems under a changing climate. Such

understanding allows us to design tailored climate impact assessments for any specific water resource system by focusing on the key climate processes and variables that show greater impacts on the system.

7.2 Limitations and future work

The limitations of this research and the associated future work recommended are discussed below:

7.2.1 Evapotranspiration process representation

In Chapters 3, 5 and 6, the potential impact of climate change on evapotranspiration is represented by considering a large number of plausible future conditions of the climate variables that are relevant to evapotranspiration. However, the potential dynamics in some other factors influencing evapotranspiration are not considered, including: 1) vegetation cover which determines the proportion of incoming radiative energy received at the evaporative surface (i.e. albedo) (e.g. Feddema et al., 2005, Liu et al., 2008); 2) plant-related processes, such as the impact of increasing CO_2 (e.g. Milly and Dunne, 2016, Bell et al., 2011, Prudhomme et al., 2014). These factors can greatly impact the rate of transpiration from vegetated surfaces, therefore, they should be considered for conducting a more comprehensive assessment of the climate change impacts on evapotranspiration.

To facilitate more comprehensive estimation of evapotranspiration under a changing climate, the impacts of vegetation can be explicitly modelled. This can be achieved by integrating evapotranspiration modelling with land-surface models which are capable to describe the influences of plant cover on the driving energy for evapotranspiration (Allen et al., 2011, Feddema et al., 2005), and/or climate models which consider the feedbacks between CO_2 and plant dynamics (Milly and Dunne, 2016, Bell et al., 2011), with both approaches involve more complex models and are thus associated with increasing requirements for data and simulation work. Further exploration of the impact of climate on evapotranspiration is warranted given the finding that potential changes in evapotranspiration can vary substantially by baseline

climate (Chapter 3), and can also have a much greater impact on runoff than previously understood (Chapter 5).

7.2.2 Rainfall-runoff model selection

In Chapter 5, three rainfall-runoff models with different structures are employed to understand how the contrasting ways of representing physical processes can impact the simulated runoff responses to potential climate change. However, this analysis is limited to the use of lumped conceptual models, which are currently the main class of rainfall-runoff models used in climate impact studies (e.g. Chiew et al., 2010, Islam et al., 2014, Najafi and Moradkhani, 2015, Vaze and Teng, 2011). However, physically-based and distributed models have not been considered, and might be able to provide a more in-depth understanding on the effects of physical process representation.

To enhance the understanding of the impact of different process representations within alternative rainfall-runoff models, the comparison of rainfall-runoff models can be extended to a larger number of structurally different rainfall-runoff models. In particular, since this study focused on lumped conceptual models, significant insight can be gained by the inclusion of more complex models, such as models with more explicit representation of specific hydrological processes (e.g. base flow and inter-flow), and/or more detailed consideration of catchment heterogeneity (e.g. land cover types, grid cells and vertical layers) (Harrigan et al., 2014, Clark et al., 2011). These further investigations are warranted given the finding in Chapter 5, which highlighted that performance on projecting runoff in historical climate does not guarantee the sufficiency for future impact assessments, and therefore more effort is needed for conducting model inter-comparison focusing on their ability of representing the physical processes rather than traditional calibration metrics. In addition, to support the inclusion of these models for inter-comparison studies, additional effort is also required for the collection and quality control of relevant hydrological data, since physically-based rainfall-runoff models are generally more data-intensive (e.g. Blöschl et al., 2016).

7.2.3 Implementation of inverse approach to generate climate exposure space

In Chapters 4 and 6, the inverse approach has been implemented to generate a climate exposure space for scenario-neutral studies. However, these implementations are limited in the number of climate attributes considered in the climate exposure space. This this is due to two factors. Firstly, as pilot exploration of stochastically generating exposure space in the two studies, the stochastic weather generators employed were all based on the first-order Markov rainfall occurrence model and Gamma rainfall intensity model (Richardson, 1981), which had relatively simple structures and thus lower computational demand. However, these simple structures can limit the flexibility to explore more plausible changes on a larger range of climate variables and attributes. Secondly, the current capacity of the inverse approach is not yet sufficient to warrant exposure spaces with higher dimensions. This is especially an issue when a number of attributes of one climate variable are considered, as more constraints need to be imposed during the optimization process, thus leading to an increase in the difficulty and computational effort for finding the best-fit parameters of the stochastic generator. As illustrated in Chapter 6, insufficient exploration of the exposure space can be particularly problematic when some key climate features for the system were not considered, which led to misinterpretation of the results obtained from scenario-neutral approaches.

The flexibility of the generation of the climate exposure space can be improved by considering alternative stochastic generators, which include: 1) more complex and alternative structures of stochastic generators, such as other forms of Richardson-type models with higher-order Markov models to describe rainfall dependency and/or more complex distributions models to include more patterns in low-frequency variability and extremes of rainfall (e.g. Wilks, 1999, Mehrotra and Sharma, 2007); 2) models with alternative structures, such as alternative renewal models that determine rainfall occurrence based on distributions of wet and dry spells (e.g. Semenov and Barrow, 1997, Heneker et al., 2001). In addition, the optimization process can

also be improved to reduce the computational demand, so that a larger number of climate attributes can be included in the climate exposure space. This can be addressed by considering alternative searching behaviour over the exposure space, implying different weightings constraint the target level of each climate attribute (as opposed to the uniform weighting as implemented in Chapters 4 and 6), and/or employing more efficient programming languages such as Fortran or C++ (as opposed to R used in Chapters 4 and 6).

References

AJAMI, N. K., DUAN, Q. & SOROOSHIAN, S. 2007. An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resources Research,* 43**,** W01403.

AKHTAR, M., AHMAD, N. & BOOIJ, M. J. 2008. The impact of climate change on the water resources of Hindukush–Karakorum–Himalaya region under different glacier coverage scenarios. *Journal of Hydrology,* 355**,** 148-163.

ALEXANDER, L., ZHANG, X., PETERSON, T., CAESAR, J., GLEASON, B., KLEIN TANK., A. M. G., HAYLOCK, M., COLLINS, D., TREWIN, B. & RAHIMZADEH, F. 2006. Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research: Atmospheres (1984–2012),* 111.

ALEXANDRIS, S., STRICEVIC, R. & PETKOVIC, S. 2008. Comparative analysis of reference evapotranspiration from the surface of rainfed grass in central Serbia, calculated by six empirical methods against the Penman-Monteith formula. *European Water,* 21**,** 17-28.

ALLEN, R., IRMAK, A., TREZZA, R., HENDRICKX, J. M. H., BASTIAANSSEN, W. & KJAERSGAARD, J. 2011. Satellite-based ET estimation in agriculture using SEBAL and METRIC. *Hydrological Processes,* 25**,** 4011-4027.

ALLEN, R. & PRUITT, W. 1986. Rational Use of The FAO Blaney-Criddle Formula. *Journal of Irrigation and Drainage Engineering,* 112**,** 139-155.

ALLEN, R. G., PEREIRA, L. S., RAES, D. & SMITH, M. 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *FAO, Rome,* 300**,** 6541.

ALLEN, R. G., WALTER, I. A., ELLIOTT, R. L., HOWELL, T. A., ITENFISU, D., JENSEN, M. E. & SNYDER, R. L. 2005. *The ASCE Standardized Reference Evapotranspiration Equation,* ASCE Publications.

ANDRÉASSIAN, V., PERRIN, C. & MICHEL, C. 2004. Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models. *Journal of Hydrology,* 286**,** 19-35.

ANDREWS, F., CROKE, B. & JAKEMAN, A. 2011. An open software environment for hydrological model assessment and development. *Environmental Modelling & Software,* 26**,** 1171-1185.

ANDREWS, F. & GUILLAUME, J. 2013. hydromad: Hydrological Model Assessment and Development. R package version 0.9-18.

ARNELL, N. W. 2004, Climate-change impacts on river flows in Britain: the UKCIPO2 scenarios, *Water and Environment Journal,* 18(2), 112-117.

ARNELL, N. W. 1999. The effect of climate change on hydrological regimes in Europe: a continental perspective. *Global Environmental Change,* 9**,** 5-23.

AULT, T. R., COLE, J. E., OVERPECK, J. T., PEDERSON, G. T. & MEKO, D. M. 2014. Assessing the Risk of Persistent Drought Using Climate Model Simulations and Paleoclimate Data. *Journal of Climate,* 27**,** 7529-7549.

BAE, D.-H., JUNG, I.-W. & LETTENMAIER, D. P. 2011, Hydrologic uncertainties in climate change from IPCC AR4 GCM simulations of the Chungju Basin, Korea, *Journal of Hydrology*, 401(1–2), 90-105.

BARELLA-ORTIZ, A., POLCHER, J., TUZET, A. & LAVAL, K. 2013. Potential evaporation estimation through an unstressed surface-energy balance and its sensitivity to climate change. *Hydrol. Earth Syst. Sci.*, 17**,** 4625-4639.

BARNETT, T. P., ADAM, J. C. & LETTENMAIER, D. P. 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, 438**,** 303-309.

BASTOLA, S., MURPHY, C. & SWEENEY, J. 2011a. The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments, Advances in Water Resources, 34(5), 562-576.

BASTOLA, S., MURPHY, C. & SWEENEY, J. 2011b. The sensitivity of fluvial flood risk in Irish catchments to the range of IPCC AR4 climate change scenarios. *Science of The Total Environment*, 409**,** 5403-5415.

BAUMGARTNER, A., REICHEL, E. & LEE, R. 1975. *The world water balance: mean annual global, continental and maritime precipitation, evaporation and run-off*, Elsevier scientific publishing company.

BEACHKOFSKI, B. K. & GRANDHI, R. 2002. Improved distributed hypercube sampling.

BEGUERÍA, S., VICENTE-SERRANO, S. M. & BEGUERÍA, M. S. 2013. Package 'SPEI'.

BELL, V. A., GEDNEY, N., KAY, A. L., SMITH, R. N. B., JONES, R. G. & MOORE, R. J. 2011. Estimating Potential Evaporation from Vegetated Surfaces for Water Management Impact Assessments Using Climate Model Output. *Journal of Hydrometeorology*, 12**,** 1127-1136.

BELL, V. A., KAY, A. L., DAVIES, H. N. & JONES, R. G. 2016. An assessment of the possible impacts of climate change on snow and peak river flows across Britain. *Climatic Change*, 136**,** 539-553.

BETHUNE, M., AUSTIN, N. & MAHER, S. 2001. Quantifying the water budget of irrigated rice in the Shepparton Irrigation Region, Australia, *Irrigation Science*, 20(3), 99-105.

BEVEN, K. 1987. Towards the use of catchment geomorphology in flood frequency predictions. *Earth Surface Processes and Landforms*, 12**,** 69-82.

BEVEN, K. 2001a. On hypothesis testing in hydrology, *Hydrological processes*, 15(9), 1655-1657.

BEVEN, K. 2001b. How far can we go in distributed hydrological modelling?, *Hydrology and Earth System Sciences Discussions*, 5(1), 1-12.

BEVEN, K. 2011. *Rainfall-Runoff Modelling: The Primer: Second Edition.* Wiley. Chichester, UK.

BEVEN, K. & FREER, J. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249**,** 11-29.

BLÖSCHL, G., BLASCHKE, A. P., BROER, M., BUCHER, C., CARR, G., CHEN, X., EDER, A., EXNER-KITTRIDGE, M., FARNLEITNER,

A., FLORES-OROZCO, A., HAAS, P., HOGAN, P., KAZEMI AMIRI, A., OISMÜLLER, M., PARAJKA, J., SILASARI, R., STADLER, P., STRAUSS, P., VREUGDENHIL, M., WAGNER, W. & ZESSNER, M. 2016. The Hydrological Open Air Laboratory (HOAL) in Petzenkirchen: a hypothesis-driven observatory. *Hydrol. Earth Syst. Sci.,* 20**,** 227-255.

BOÉ, J. & TERRAY, L. 2008. Uncertainties in summer evapotranspiration changes over Europe and implications for regional climate change. *Geophysical Research Letters,* 35**,** n/a-n/a.

BORMANN, H. 2011. Sensitivity analysis of 18 different potential evapotranspiration models to observed climatic change at German climate stations. *Climatic change,* 104**,** 729-753.

BOUCHET, R. J. 1963. Evapotranspiration réelle et potentielle, signification climatique. *General Assembly Berkeley.* Gentbrugge, Belgium: Int. Assoc. Sci. Hydrol.

BOUGHTON, W. 2004. The Australian water balance model, *Environmental Modelling & Software,* 19(10), 943-956.

BOUGHTON, W. 2009. New approach to calibration of the AWBM for use on ungauged catchments, *Journal Of Hydrologic Engineering,* 14(6), 562-568.

BOUGHTON, W. & DROOP, O. 2003. Continuous simulation for design flood estimation—a review. *Environmental Modelling & Software,* 18**,** 309-318.

BROUYÈRE, S., CARABIN, G. & DASSARGUES, A. 2004. Climate change impacts on groundwater resources: modelled deficits in a chalky aquifer, Geer basin, Belgium, *Hydrogeology Journal,* 12(2), 123-134.

BROWN, C., GHILE, Y., LAVERTY, M. & LI, K. 2012. Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector. *Water Resources Research,* 48**,** W09537.

BROWN, C., WERICK, W., LEGER, W. & FAY, D. 2011. A Decision-Analytic Approach to Managing Climate Risks: Application to the Upper Great Lakes1. *JAWRA Journal of the American Water Resources Association,* 47**,** 524-534.

BROWN, C. & WILBY, R. L. 2012. An alternate approach to assessing climate risks. *Eos, Transactions American Geophysical Union,* 93**,** 401-402.

BRUTSAERT, W. & STRICKER, H. 1979. An advection-aridity approach to estimate actual regional evapotranspiration. *Water Resources Research,* 15**,** 443-450.

BUREAU OF METEOROLOGY, 2015. Hydrologic Reference Stations. http://www.bom.gov.au/water/hrs/.

BUREAU OF METEOROLOGY, 2016. Climate Data Online. http://www.bom.gov.au/climate/data/

BUTTS, M. B., PAYNE, J. T., KRISTENSEN, M. & MADSEN, H. 2004. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. Journal of Hydrology, 298, 242-266.

CARNELL, R. 2012. lhs: Latin Hypercube Samples. R package version 0.10.

CASTELLVI, F., STOCKLE, C., PEREZ, P. & IBANEZ, M. 2001. Comparison of methods for applying the Priestley–Taylor equation at a regional scale. *Hydrological Processes,* 15**,** 1609-1620.

CHANG, H. & JUNG, I.-W. 2010. Spatial and temporal changes in runoff caused by climate change in a complex large river basin in Oregon. *Journal of Hydrology,* 388**,** 186-207.

CHANG, S., GRAHAM, W. D., HWANG, S. & MUÑOZ-CARPENA, R. 2016. Sensitivity of future continental United States water deficit projections to general circulation models, the evapotranspiration estimation method, and the greenhouse gas emission scenario. *Hydrol. Earth Syst. Sci.,* 20**,** 3245-3261.

CHAPMAN, T. Estimation of evaporation in rainfall-runoff models. *In:* GHASSEMI, F., POST, D., SIVAPALAN, M. & VERTESSY, R., eds. MODSIM2001: Integrating models for Natural Resources Management across Disciplines, Issues and Scales, 2003. MSSANZ, 293-298.

CHEN, H., XU, C.-Y. & GUO, S. 2012. Comparison and evaluation of multiple GCMs, statistical downscaling and hydrological models in the study of climate change impacts on runoff. *Journal of Hydrology,* 434–435**,** 36-45.

CHEN, J. & BRISSETTE, F. P. 2014. Comparison of five stochastic weather generators in simulating daily precipitation and temperature for the Loess Plateau of China. *International Journal of Climatology,* 34**,** 3089-3105.

CHEN, W., LAN, X., WANG, L. & MA, Y. 2013. The combined effects of the ENSO and the Arctic Oscillation on the winter climate anomalies in East Asia. *Chinese Science Bulletin,* 58**,** 1355-1362.

CHENG, C. T., OU, C. P. & CHAU, K. W. 2002. Combining a fuzzy optimal model with a genetic algorithm to solve multi-objective rainfall–runoff model calibration. *Journal of Hydrology,* 268**,** 72-86.

CHIEW, F. H. S. 2006. Estimation of rainfall elasticity of streamflow in Australia, *Hydrological Sciences Journal*, 51(4), 613-625.

CHIEW, F. H. & MCMAHON, T. A. 1991. The applicability of Morton's and Penman's evapotranspiration estimates in rainfall-runoff modelling. *JAWRA Journal of the American Water Resources Association,* 27**,** 611-620.

CHIEW, F. H. & MCMAHON, T. A. 2002. Modelling the impacts of climate change on Australian streamflow. *Hydrological Processes,* 16**,** 1235-1245.

CHIEW, F. H. S., KIRONO, D. G. C., KENT, D. M., FROST, A. J., CHARLES, S. P., TIMBAL, B., NGUYEN, K. C. & FU, G. 2010. Comparison of runoff modelled using rainfall from different downscaling methods for historical and future climates. *Journal of Hydrology,* 387**,** 10-23.

CHIEW, F. H. S., TENG, J., VAZE, J. & KIRONO, D. G. C. 2009a. Influence of global climate model selection on runoff impact assessment. *Journal of Hydrology,* 379**,** 172-180.

CHIEW, F. H. S., TENG, J., VAZE, J., POST, D. A., PERRAUD, J. M., KIRONO, D. G. C. & VINEY, N. R. 2009b. Estimating climate change impact on runoff across southeast Australia: Method, results,

and implications of the modeling method. *Water Resources Research,* 45**,** W10414.

CHRISTENSEN, J. H., HEWITSON, B., BUSUIOC, A., CHEN, A., GAO, X., HELD, R., JONES, R., KOLLI, R. K., KWON, W. & LAPRISE, R. 2007. Regional climate projections. *Climate Change, 2007: The Physical Science Basis. Contribution of Working group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, University Press, Cambridge, Chapter 11***,** 847-940.

CHRISTENSEN, N. & LETTENMAIER, D. P. 2006. A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River basin. *Hydrol. Earth Syst. Sci. Discuss.*, 3**,** 3727-3770.

CHRISTENSEN, N. S., WOOD, A. W., VOISIN, N., LETTENMAIER, D. P. & PALMER, R. N., 2004., The Effects of Climate Change on the Hydrology and Water Resources of the Colorado River Basin, *Climatic Change*, 62(1), 337-363.

CLARK, M. P., KAVETSKI, D. & FENICIA, F. 2011. Pursuing the method of multiple working hypotheses for hydrological modeling. *Water Resources Research,* 47**,** W09301.

CLARK, M. P. & SLATER, A. G. 2006. Probabilistic Quantitative Precipitation Estimation in Complex Terrain. *Journal of Hydrometeorology,* 7**,** 3-22.

CLARK, M. P., SLATER, A. G., RUPP, D. E., WOODS, R. A., VRUGT, J. A., GUPTA, H. V., WAGENER, T. & HAY, L. E. 2008. Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research,* 44.

CLARK, M. P., WILBY, R. L., GUTMANN, E. D., VANO, J. A., GANGOPADHYAY, S., WOOD, A. W., FOWLER, H. J., PRUDHOMME, C., ARNOLD, J. R. & BREKKE, L. D. 2016. Characterizing Uncertainty of the Hydrologic Impacts of Climate Change. *Current Climate Change Reports,* 2**,** 55-64.

CLEUGH, H. A., LEUNING, R., Mu, Q. & RUNNING, S. W. 2007. Regional evaporation estimates from flux tower and MODIS satellite data, *Remote Sensing of Environment*, 106(3), 285-304.

CLEVERLY, J. 2011. Alice Springs Mulga OzFlux site OzFlux: Australian and New Zealand Flux Research and Monitoring. hdl: 102.100.100/14217.

COLLINS, M., KNUTTI, R., ARBLASTER, J., DUFRESNE, J.-L., FICHEFET, T., FRIEDLINGSTEIN, P., GAO, X., GUTOWSKI, W. J., JOHNS, T., KRINNER, G., SHONGWE, M., TEBALDI, C., WEAVER, A. J. & WEHNER, M. 2013. Long-term Climate Change: Projections, Commitments and Irreversibility. *In:* STOCKER, T. F., QIN, D., PLATTNER, G.-K., TIGNOR, M., ALLEN, S. K., BOSCHUNG, J., NAUELS, A., XIA, Y., BEX, V. & MIDGLEY, P. M. (eds.) *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.

CORON, L., ANDRÉASSIAN, V., PERRIN, C., LERAT, J., VAZE, J., BOURQUI, M. & HENDRICKX, F. 2012. Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments, *Water Resources Research*, 48(5), n/a-n/a.CROKE, B. F. W., and A. J. JAKEMAN (2004), A catchment moisture deficit module for the IHACRES rainfall-runoff model, Environmental Modelling and Software with Environment Data News, 19(1), 1-5.

CROSSMAN, J., FUTTER, M. N., ONI, S. K, WHITEHEAD, P. G., JIN, L., BUTTERFIELD, D., BAULCH, H. M. & DILLON, P. J. 2013. Impacts of climate change on hydrology and water quality: Future proofing management strategies in the Lake Simcoe watershed, Canada, *Journal of Great Lakes Research*, 39(1), 19-32.

CSIRO AND BUREAU OF METEOROLOGY 2015. Climate Change in Australia Information for Australia's Natural Resource Management Regions: Technical Report. Australia: CSIRO and Bureau of Meteorology.

CULLEY, S., NOBLE, S., YATES, A., TIMBS, M., WESTRA, S., MAIER, H. R., GIULIANI, M. & CASTELLETTI, A. 2016. A bottom-up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate. *Water Resources Research*, 52**,** 6751-6768.

DAI, A. 2011. Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900–2008. *Journal of Geophysical Research: Atmospheres*, 116**,** D12115.

DAWSON, C. W., ABRAHART, R. J. & SEE, L. M. 2007. HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts. *Environmental Modelling & Software*, 22**,** 1034-1052.

DE BRUIN, H. 1981. The determination of (reference crop) evapotranspiration from routine weather data. *Evaporation in relation to hydrology***,** 25-37.

DESSAI, S. & HULME, M. 2004. Does climate adaptation policy need probabilities? *Climate Policy*, 4**,** 107-128.

DIAZ-NIETO, J. & WILBY, R. L. 2005. A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. *Climatic Change*, 69**,** 245-268.

DINGMAN, S. L. 2015. *Physical Hydrology: Third Edition*, Waveland Press.

DONOHUE, R. J., MCVICAR, T. R. & RODERICK, M. L. 2009. Generating Australian potential evaporation data suitable for assessing the dynamics in evaporative demand within a changing climate.

DONOHUE, R. J., MCVICAR, T. R. & RODERICK, M. L. 2010a. Assessing the ability of potential evaporation formulations to capture the dynamics in evaporative demand within a changing climate. *Journal of Hydrology*, 386**,** 186-197.

DONOHUE, R. J., MCVICAR, T. R. & RODERICK, M. L. 2010b. Can dynamic vegetation information improve the accuracy of Budyko's hydrological model? *Journal of hydrology*, 390**,** 23-34.

DONOHUE, R. J., RODERICK, M. L. & MCVICAR, T. R. 2011. Assessing the differences in sensitivities of runoff to changes in climatic conditions across a large basin. *Journal of Hydrology,* 406**,** 234-244.

DOORENBOS, J. 1977. Guidelines for predicting crop water requirements. *FAO irrigation and drainage paper,* 24**,** 15-29.

DUAN, Q., AJAMI, N. K., GAO, X. & SOROOSHIAN, S. 2007. Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources,* 30**,** 1371-1386.

DUAN, Q., GUPTA, V. K. & SOROOSHIAN, S. 1993. Shuffled complex evolution approach for effective and efficient global minimization. *Journal of optimization theory and applications,* 76**,** 501-521.

DUBROVSKÝ, M., ŽALUD, Z. & ŠŤASTNÁ, M. 2000. Sensitivity of CERES-Maize yields to statistical structure of daily weather series. *Climatic Change,* 46**,** 447-472.

EASTERLING, D. R., MEEHL, G. A., PARMESAN, C., CHANGNON, S. A., KARL, T. R. & MEARNS, L. O. 2000. Climate Extremes: Observations, Modeling, and Impacts. *Science,* 289**,** 2068-2074.

EKSTRÖM, M., JONES, P., FOWLER, H., LENDERINK, G., BUISHAND, T. & CONWAY, D. 2007. Regional climate model data used within the SWURVE project? 1: projected changes in seasonal patterns and estimation of PET. *Hydrology and Earth System Sciences Discussions,* 11**,** 1069-1083.

EVANS, J. P. & JAKEMAN, A. J. 1998. Development of a simple, catchment-scale, rainfall-evapotranspiration-runoff model. *Environmental Modelling & Software,* 13**,** 385-393.

EWEN, J., O'DONNELL, G., BURTON, A. & O'CONNELL, E. 2006. Errors and uncertainty in physically-based rainfall-runoff modelling of catchment change effects, *Journal of Hydrology*, 330(3–4), 641-650.

FEDDEMA, J. J., OLESON, K. W., BONAN, G. B., MEARNS, L. O., BUJA, L. E., MEEHL, G. A. & WASHINGTON, W. M. 2005. The Importance of Land-Cover Change in Simulating Future Climates. *Science,* 310**,** 1674-1678.

FELIX, T. P., PETRA, D., STEPHANIE, E. & MARTINA, F. 2013. Impact of climate change on renewable groundwater resources: assessing the benefits of avoided greenhouse gas emissions using selected CMIP5 climate projections. *Environmental Research Letters,* 8**,** 024023.

FISHER, J. B., WHITTAKER, R. J. & MALHI, Y. 2011. ET come home: potential evapotranspiration in geographical ecology. *Global Ecology and Biogeography,* 20**,** 1-18.

FLATO, G., MAROTZKE, J., ABIODUN, B., BRACONNOT, P., CHOU, S. C., COLLINS, W., COX, P., DRIOUECH, F., EMORI, S. & EYRING, V. 2013. Evaluation of climate models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press.

FOWLER, H., BLENKINSOP, S. & TEBALDI, C. 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology,* 27**,** 1547-1578.

FOWLER, K. J. A., PEEL, M. C., WESTERN, A. W., ZHANG, L. & PETERSON T. J. 2016. Simulating runoff under changing climatic conditions: Revisiting an apparent deficiency of conceptual rainfall-runoff models, *Water Resources Research*, 52(3), 1820-1846.

FREVERT, D. K., HILL, R. W. & BRAATEN, B. C. 1983. Estimation of FAO evapotranspiration coefficients. *Journal of irrigation and drainage engineering*, 109**,** 265-270.

FROST, A. 2004. *Stochastic generation of point rainfall data at subdaily timescales: a comparison of DRIP and NSRP*.

FURRER, E. M. & KATZ, R. W. 2008. Improving the simulation of extreme precipitation events by stochastic weather generators. *Water Resources Research,* 44.

GAÁL, L., SZOLGAY, J., KOHNOVÁ, S., PARAJKA, J., MERZ, R., VIGLIONE, A. & BLÖSCHL, G. 2012. Flood timescales: Understanding the interplay of climate and catchment processes through comparative hydrology. *Water Resources Research,* 48**,** n/a-n/a.

GALELLI, S., HUMPHREY, G. B., MAIER, H. R., CASTELLETTI, A., DANDY, G. C. & GIBBS, M. S. 2014. An evaluation framework for input variable selection algorithms for environmental data-driven models. *Environmental Modelling & Software,* 62**,** 33-51.

GAO, L., BRYAN, B. A., NOLAN, M., CONNOR, J. D., SONG, X. & ZHAO, G. 2016. Robust global sensitivity analysis under deep uncertainty via scenario analysis. *Environmental Modelling & Software,* 76**,** 154-166.

GAO, W. 1988. Applications of solutions to non-linear energy budget equations. *Agricultural and Forest Meteorology,* 43**,** 121-145.

GARCIA, E. S. & TAGUE, C. L. 2015. Subsurface storage capacity influences climate–evapotranspiration interactions in three western United States catchments. *Hydrol. Earth Syst. Sci.,* 19**,** 4845-4858.

GASCA-TUCKER, D., ACREMAN, M., AGNEW, C. & THOMPSON, J. 2007. Estimating evaporation from a wet grassland. *Hydrology & Earth System Sciences*, 11.

GERRITS, A., SAVENIJE, H., VELING, E. & PFISTER, L. 2009. Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model. *Water Resources Research,* 45.

GIBBS, M. S., MAIER, H. R. & DANDY, G. C. 2012. A generic framework for regression regionalization in ungauged catchments. *Environmental Modelling & Software,* 27–28**,** 1-14.

GODERNIAUX, P., BROUYÈRE, S., FOWLER, H. J., BLENKINSOP, S., THERRIEN, R., ORBAN, P. & DASSARGUES, A. 2009. Large scale surface–subsurface hydrological model to assess climate change impacts on groundwater reserves. *Journal of Hydrology,* 373**,** 122-138.

GONG, L., XU, C.-Y., CHEN, D., HALLDIN, S. & CHEN, Y. D. 2006. Sensitivity of the Penman–Monteith reference evapotranspiration to key climatic variables in the Changjiang (Yangtze River) basin. *Journal of Hydrology*, 329**,** 620-629.

GOSLING, S. N., TAYLOR, R. G., ARNELL, N. W. & TODD, M. C. 2011. A comparative analysis of projected impacts of climate change on

river runoff from global and catchment-scale hydrological models. *Hydrol. Earth Syst. Sci.,* 15**,** 279-294.

GOYAL, R. K. 2004. Sensitivity of evapotranspiration to global warming: a case study of arid zone of Rajasthan (India). *Agricultural Water Management,* 69**,** 1-11.

GRANGER, R. J. & GRAY, D. 1989. Evaporation from natural nonsaturated surfaces. *Journal of Hydrology,* 111**,** 21-29.

GUILLEVIC, P., KOSTER, R. D., SUAREZ, M. J., BOUNOUA, L., COLLATZ, G. J., LOS, S. O. & MAHANAMA, S. P. P. 2002. Influence of the Interannual Variability of Vegetation on the Surface Energy Balance—A Global Sensitivity Study. *Journal of Hydrometeorology,* 3**,** 617-629.

GUO, D., WESTRA, S. & MAIER, H. R. 2016a. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. *Journal of Hydrology*. Available online 22 March 2016, ISSN 0022-1694. doi:10.1016/j.jhydrol.2016.03.025.

GUO, D., WESTRA, S. & MAIER, H. R. 2016b. An R package for modelling actual, potential and reference evapotranspiration. *Environmental Modelling & Software,* 78**,** 216-224. doi:10.1016/j.envsoft.2015.12.019.

GUO, D., WESTRA, S. & MAIER, H. R. 2017. Sensitivity of potential evapotranspiration to changes in climate variables for different climatic zones. *Hydrol. Earth Syst. Sci.,* 21, 2107-2126, doi:10.5194/hess-21-2107-2017, 2017.

GUO, D., WESTRA, S. & MAIER, H. R. 2017. Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models. *Water Resources Research*, 53. doi:10.1002/2016WR019627.

GUPTA, H. V., SOROOSHIAN, S. & YAPO, P. O. 1999. Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *Journal of Hydrologic Engineering*.

GUPTA, H. V., WAGENER, T. & LIU, Y. 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation, *Hydrological Processes*, 22(18), 3802-3813.

HAUSER, R., ARCHER, S., BACKLUND, P., HATFIELD, J., JANETOS, A., LETTENMAIER, D. & WALSH M. 2009. The effects of climate change on US Ecosystems, US Global Change Research Program, Washington, D.C., USA

HERMAN, J. D., REED, P. M. & WAGENER, T. 2013. Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior, *Water Resources Research*, 49(3), 1400-1414.

HAQUE, M. M., RAHMAN, A., HAGARE, D., KIBRIA, G. & KARIM, F. 2015. Estimation of catchment yield and associated uncertainties due to climate change in a mountainous catchment in Australia, *Hydrological Processes*, 29(19), 4339-4349.

HARGREAVES, G. H. & SAMANI, Z. A. 1985. Reference crop evapotranspiration from ambient air temperature. *American Society of Agricultural Engineers*.

HARRIGAN, S., MURPHY, C., HALL, J., WILBY, R. L. & SWEENEY, J. 2014. Attribution of detected changes in streamflow using multiple working hypotheses. *Hydrol. Earth Syst. Sci.,* 18**,** 1935-1952.

HARTMANN, G. & BÁRDOSSY A. 2005. Investigation of the transferability of hydrological models and a method to improve model calibration, *Advances in Geosciences*, 5, 83-87.

HALTON, J. H. 1960. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik,* 2**,** 84-90.

HASHMI, M. Z., SHAMSELDIN, A. Y. & MELVILLE, B. W. 2011. Comparison of SDSM and LARS-WG for simulation and downscaling of extreme precipitation events in a watershed. *Stochastic Environmental Research and Risk Assessment,* 25**,** 475-484.

HAMMERSLEY, J. M. 1960. Monte Carlo methods for solving multivariable problems. *Annals of the New York Academy of Sciences,* 86**,** 844-874.

HANSEN, J. E. & SATO, M. 2012. Paleoclimate Implications for Human-Made Climate Change. *In:* BERGER, A., MESINGER, F. & SIJACKI, D. (eds.) *Climate Change: Inferences from Paleoclimate and Regional Aspects.* Vienna: Springer Vienna.

HENEKER, T. M., LAMBERT, M. F. & KUCZERA, G. 2001. A point rainfall model for risk-based design. *Journal of hydrology,* 247**,** 54-71.

HO, M., KIEM, A. S. & VERDON-KIDD, D. C. 2015b. A paleoclimate rainfall reconstruction in the Murray-Darling Basin (MDB), Australia: 2. Assessing hydroclimatic risk using paleoclimate records of wet and dry epochs. *Water Resources Research*, 51, 8380–8396

HOLLAND, J. H. 1975. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*, University of Michigan Press.

HORVÁTH, S., SZÉP, I. J., MAKRA, L., MIKA, J., PAJTÓK-TARI, I. & UTASI, Z. 2010. Effect of evapotranspiration parameterisation on the Palmer Drought Severity Index. *Physics and Chemistry of the Earth, Parts A/B/C,* 35**,** 11-18.

HUO, Z., DAI, X., FENG, S., KANG, S. & HUANG, G. 2013. Effect of climate change on reference evapotranspiration and aridity index in arid region of China, *Journal of Hydrology*, 492, 24-34.

IBROM, A., DELLWIK, E., FLYVBJERG, H., JENSEN, N. O. & Pilegaard, K. 2007. Strong low-pass filtering effects on water vapour flux measurements with closed-path eddy correlation systems, *Agricultural and Forest Meteorology,* 147(3–4), 140-156.

IPCC 2014. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. *In:* FIELD, C. B., BARROS, V. R., DOKKE, D. J., MASTRANDREA, M. D., BILIR, T. E., CHATTERJEE, M., EBI, K. L., ESTRADA, Y. O., GENOVA, R. C., GIRMA, B., KISSEL, E. S., LEVY, A. N., MACCRACKEN, S., MASTRANDREA, P. R. & L.L., W. (eds.). Cambridge, United Kingdom and New York, NY, USA.

ISLAM, S. A., BARI, M. A. & ANWAR, A. H. M. F. 2014. Hydrologic impact of climate change on Murray–Hotham catchment of Western

Australia: a projection of rainfall–runoff for future water resources planning. *Hydrol. Earth Syst. Sci.,* 18**,** 3591-3614.

JENSEN, M. E. & HAISE, H. R. 1963. Estimating evapotranspiration from solar radiation. *Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division,* 89**,** 15-41.

JIANG, T., CHEN, Y. D., XU, C.-Y., CHEN, X., CHEN, X. & SINGH V. P. 2007. Comparison of hydrological impacts of climate change simulated by six hydrological models in the Dongjiang Basin, South China, *Journal of Hydrology*, 336(3–4), 316-333.

JOHNSON, F. & SHARMA, A. 2009. Measurement of GCM skill in predicting variables relevant for hydroclimatological assessments. *Journal of Climate,* 22**,** 4373-4382.

JOHNSON, F. & SHARMA, A. 2011. Accounting for interannual variability: A comparison of options for water resources climate change impact assessments. Water Resources Research, 47, n/a-n/a.

JOHNSON, F., WESTRA, S., SHARMA, A. & PITMAN, A. J. 2011. An Assessment of GCM Skill in Simulating Persistence across Multiple Time Scales. *Journal of Climate,* 24**,** 3609-3623.

JONES, P., HARPHAM, C., GOODESS, C. & KILSBY, C. 2011. Perturbing a Weather Generator using change factors derived from Regional Climate Model simulations. *Nonlinear Processes in Geophysics,* 18**,** 503-511.

JONES, P. G. & THORNTON, P. K. 1993. A rainfall generator for agricultural applications in the tropics. *Agricultural and forest meteorology,* 63**,** 1-19.

JONES, R. & PAGE, C. 2001. Assessing the risk of climate change on the water resources of the Macquarie River Catchment. Integrating Models for Natural Resources Management across Disciplines, issues and scales, 673-678.

JONES, R. N., CHIEW, F. H. S., BOUGHTON, W. C. & ZHANG, L. 2006. Estimating the sensitivity of mean annual runoff to climate change using selected hydrological models. *Advances in Water Resources,* 29**,** 1419-1429.

KANNAN, N., WHITE, S. M., WORRALL, F. & WHELAN, M. J. 2007. Sensitivity analysis and identification of the best evapotranspiration and runoff options for hydrological modelling in SWAT-2000. *Journal of Hydrology,* 332**,** 456-466.

KATZ, R. W. 2002. Techniques for estimating uncertainty in climate change scenarios and impact studies. *Climate Research,* 20**,** 167-185.

KAVETSKI, D. & FENICIA, F. 2011. Elements of a flexible approach for conceptual hydrological modeling: 2. Application and experimental insights. *Water Resources Research,* 47**,** W11511.

KAY, A. L., BELL, V. A., BLYTH, E. M., CROOKS, S. M., DAVIES, H. N. & REYNARD, N. S. 2013. A hydrological perspective on evaporation: historical trends and future projections in Britain. *Journal of Water and Climate Change,* 4**,** 193-208.

KAY, A. L., CROOKS, S. M. & REYNARD, N. S. 2014. Using response surfaces to estimate impacts of climate change on flood peaks: assessment of uncertainty. *Hydrological Processes,* 28**,** 5273-5287.

KAY, A. L. & DAVIES, H. N. 2008. Calculating potential evaporation from climate model data: A source of uncertainty for hydrological climate change impacts. *Journal of Hydrology,* 358**,** 221-239.

KAY, A. L., DAVIES, H. N., BELL, V. A. & JONES, R. G. 2009. Comparison of uncertainty sources for climate change impacts: flood frequency in England. *Climatic Change,* 92**,** 41-63.

KAY, A. L. & JONES, R. G. 2012. Comparison of the use of alternative UKCP09 products for modelling the impacts of climate change on flood frequency. *Climatic Change,* 114**,** 211-230.

KILSBY, C., JONES, P., BURTON, A., FORD, A., FOWLER, H., HARPHAM, C., JAMES, P., SMITH, A. & WILBY, R. 2007. A daily weather generator for use in climate change studies. *Environmental Modelling & Software,* 22**,** 1705-1719.

KIM, B., KIM, H., SEOH, B. & KIM, N. 2007. Impact of climate change on water resources in Yongdam Dam Basin, Korea. *Stochastic Environmental Research and Risk Assessment,* 21**,** 355-373.

KINGSTON, D. G., TODD, M. C., TAYLOR, R. G., THOMPSON, J. R. & ARNELL, N. W. 2009. Uncertainty in the estimation of potential evapotranspiration under climate change. *Geophysical Research Letters,* 36**,** L20403.

KIRCHNER, J. W. 2006. Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research,* 42**,** n/a-n/a.

KLEIN TANK., A. M. G., ZWIERS, F. W. & ZHANG, X. 2009. Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. *Climate Data and Monitoring.* World Meteorological Organisation.

KLEMEŠ, V. 1986. Operational testing of hydrological simulation models. *Hydrological Sciences Journal,* 31**,** 13-24.

KOEDYK, L. P. & KINGSTON, D. G. 2016. Potential evapotranspiration method influence on climate change impacts on river flow: a mid-latitude case study. *Hydrology Research*.

KÖPPEN, W., GEIGER, R., BORCHARDT, W., WEGENER, K., WAGNER, A., KNOCH, K., SAPPER, K., WARD, R. D., BROOKS, C. F. & CONNOR, A. 1930. *Handbuch der klimatologie,* Gebrüder Borntraeger Berlin, Germany.

KÖPPEN, W. P. 1931. Grundriss der klimakunde.

KWON, H.-H., LALL, U. & OBEYSEKERA, J. 2009. Simulation of daily rainfall scenarios with interannual and multidecadal climate cycles for South Florida. *Stochastic Environmental Research and Risk Assessment,* 23**,** 879-896.

LANGOUSIS, A. & KALERIS, V. 2014. Statistical framework to simulate daily rainfall series conditional on upper-air predictor variables. *Water Resources Research,* 50**,** 3907-3932.

LANGOUSIS, A., MAMALAKIS, A., DEIDDA, R. & MARROCU, M. 2015. Assessing the relative effectiveness of statistical downscaling and distribution mapping in reproducing rainfall statistics based on climate model results. *Water Resources Research***,** n/a-n/a.

LEAVESLEY, G. H., MARKSTROM, S. L. & VIGER, R. J. 2006. USGS modular modeling system (MMS)-precipitation-runoff modeling system (PRMS). *Watershed models***,** 159-177.

LI, L., LAMBERT, M. F., MAIER, H. R., PARTINGTON, D. & SIMMONS, C. T. 2015. Assessment of the internal dynamics of the Australian Water Balance Model under different calibration regimes, Environmental Modelling & Software, 66, 57-68.

LI, L., MAIER, H. R., PARTINGTON, D., LAMBERT, M. F. & SIMMONS, C. T. 2014. Performance assessment and improvement of recursive digital baseflow filters for catchments with different physical characteristics and hydrological inputs. *Environmental Modelling & Software,* 54**,** 39-52.

LINACRE, E. T. 1993. Data-sparse estimation of lake evaporation, using a simplified Penman equation. *Agricultural and Forest Meteorology,* 64**,** 237-256.

LIU, M., TIAN, H., CHEN, G., REN, W., ZHANG, C. & LIU, J. 2008. Effects of Land-Use and Land-Cover Change on Evapotranspiration and Water Yield in China During 1900-20001. *JAWRA Journal of the American Water Resources Association,* 44**,** 1193-1207.

LU, X., BAI, H. & MU, X. 2016. Explaining the evaporation paradox in Jiangxi Province of China: Spatial distribution and temporal trends in potential evapotranspiration of Jiangxi Province from 1961 to 2013. *International Soil and Water Conservation Research,* 4**,** 45-51.

MAIER, H., KAPELAN, Z., KASPRZYK, J., KOLLAT, J., MATOTT, L., CUNHA, M., DANDY, G., GIBBS, M., KEEDWELL, E. & MARCHI, A. 2014. Evolutionary algorithms and other metaheuristics in water resources: current status, research challenges and future directions. *Environmental Modelling & Software,* 62**,** 271-299.

MANTEUFEL, R. D. Distributed hypercube sampling algorithm. Third AIAA non-deterministic approaches forum paper AIAA-2001-1673, 42nd structures, structural dynamics, and materials conference, 2001.

MCJANNET, D., WEBSTER, I., STENSON, M. & SHERMAN, B. 2008. *Estimating open water evaporation for the Murray-Darling Basin: A report to the Australian Government from the CSIRO Murray-Darling Basin Sustainable Yields Project*.

MCKENNEY, M. S. & ROSENBERG, N. J. 1993. Sensitivity of some potential evapotranspiration estimation methods to climate change. *Agricultural and Forest Meteorology,* 64**,** 81-110.

MCMAHON, T. A., PEEL, M. C. & KAROLY, D. J. 2015. Assessment of precipitation and temperature data from CMIP3 global climate models for hydrologic simulation. *Hydrol. Earth Syst. Sci.,* 19**,** 361-377.

MCMAHON, T. A., PEEL, M. C., LOWE, L., SRIKANTHAN, R. & MCVICAR, T. R. 2013. Estimating actual, potential, reference crop and pan evaporation using standard meteorological data: a pragmatic synthesis. *Hydrol. Earth Syst. Sci.,* 17**,** 1331-1363.

MCVICAR, T. R., DONOHUE, R. J., O'GRADY, A. P. & LI, L. 2010. The effects of climatic changes on plant physiological and catchment ecohydrological processes in the high-rainfall catchments of the Murray-Darling Basin: A scoping study. *Prepared for the Murray-Darling Basin Authority (MDBA) by the Commonwealth Scientific and*

Industrial Research Organization (CSIRO) Water for a Healthy Country National Research Flagship, MDBA, Canberra, ACT, Australia.

MCVICAR, T. R., RODERICK, M. L., DONOHUE, R. J., LI, L. T., VAN NIEL, T. G., THOMAS, A., GRIESER, J., JHAJHARIA, D., HIMRI, Y., MAHOWALD, N. M., MESCHERSKAYA, A. V., KRUGER, A. C., REHMAN, S. & DINPASHOH, Y. 2012. Global review and synthesis of trends in observed terrestrial near-surface wind speeds: Implications for evaporation. *Journal of Hydrology*, 416–417**,** 182-205.

MCVICAR, T. R., VAN NIEL, T. G., LI, L., HUTCHINSON, M. F., MU, X. & LIU, Z. 2007. Spatially distributing monthly reference evapotranspiration and pan evaporation considering topographic influences. *Journal of hydrology*, 338**,** 196-220.

MCVICAR, T. R., VAN NIEL, T. G., LI, L. T., RODERICK, M. L., RAYNER, D. P., RICCIARDULLI, L. & DONOHUE, R. J. 2008. Wind speed climatology and trends for Australia, 1975–2006: Capturing the stilling phenomenon and comparison with near-surface reanalysis output. *Geophysical Research Letters*, 35.

MEHROTRA, R. & SHARMA, A. 2007. A semi-parametric model for stochastic generation of multi-site daily rainfall exhibiting low-frequency variability. *Journal of Hydrology*, 335**,** 180-193.

MEIRESONNE, L., SAMPSON, D. A., KOWALSKI, A. S., JANSSENS, I. A., NADEZHDINA, N., CERMÁK, J., VAN SLYCKEN, J. & CEULEMANS, R. 2003. Water flux estimates from a Belgian Scots pine stand: a comparison of different approaches, *Journal of Hydrology*, 270(3–4), 230-252.

MENDOZA, P. A., CLARK, M. P., MIZUKAMI, N., GUTMANN, E. D., ARNOLD, J. R., BREKKE, L. D. & RAJAGOPALAN, B. 2016. How do hydrologic modeling decisions affect the portrayal of climate change impacts? *Hydrological Processes*, 30**,** 1071-1095.

MERZ, R. & BLÖSCHL, G. 2009. A regional analysis of event runoff coefficients with respect to climate and catchment characteristics in Austria. *Water Resources Research*, 45**,** n/a-n/a.

MERZ, R., PARAJKA, J., & BLÖSCHL, G. 2011. Time stability of catchment model parameters: Implications for climate impact analyses, *Water Resources Research*, 47(2), n/a-n/a.

MESELHE, E., HABIB, E., OCHE, O. & GAUTAM, S. 2009. Sensitivity of Conceptual and Physically Based Hydrologic Models to Temporal and Spatial Rainfall Sampling. *Journal of Hydrologic Engineering*, 14**,** 711-720.

METEOROLOGY, A. B. O. & WANG, Q. 2001. *Climatic atlas of Australia: evapotranspiration*, Bureau of Meteorology.

MEYER, W. S., KONDRLOVÀ, E. & KOERBER G. R. 2015. Evaporation of perennial semi-arid woodland in southeastern Australia is adapted for irregular but common dry periods, *Hydrological Processes*, 29(17), 3714-3726.

MILLY, P. C. D. & DUNNE, K. A. 2016. Potential evapotranspiration and continental drying. *Nature Clim. Change*, advance online publication.

MINVILLE, M., CARTIER, D., GUAY, C., LECLAIRE, L.-A., AUDET, C., LE DIGABEL, S., & MERLEAU, J. 2014. Improving process representation in conceptual hydrological model calibration using climate simulations, *Water Resources Research*, 50(6), 5044-5073.

MOODY, P. & BROWN, C. 2013. Robustness indicators for evaluation under climate change: Application to the upper Great Lakes. *Water Resources Research,* 49**,** 3576-3588.

MORTON, F. I. 1983a. Operational estimates of areal evapotranspiration and their significance to the science and practice of hydrology. *Journal of Hydrology,* 66**,** 1-76.

MORTON, F. I. 1983b. Operational estimates of lake evaporation. *Journal of Hydrology,* 66**,** 77-100.

MU, Q., ZHAO, M., & RUNNING, S. W. 2011. Improvements to a MODIS global terrestrial evapotranspiration algorithm, *Remote Sensing of Environment*, 115(8), 1781-1800.

NAJAFI, M. R. & MORADKHANI, H. 2015. Multi-model ensemble analysis of runoff extremes for climate change impact assessments. *Journal of Hydrology,* 525**,** 352-361.

NARAPUSETTY, B., DELSOLE, T. & TIPPETT, M. K. 2009. Optimal Estimation of the Climatological Mean. *Journal of Climate,* 22**,** 4845-4859.

NAZEMI, A. & WHEATER, H. S. 2014. Assessing the vulnerability of water supply to changing streamflow conditions. *Eos, Transactions American Geophysical Union,* 95**,** 288-288.

NAZEMI, A., WHEATER, H. S., CHUN, K. P. & ELSHORBAGY, A. 2013. A stochastic reconstruction framework for analysis of water resource system vulnerability to climate-induced changes in river flow regime. *Water Resources Research,* 49**,** 291-305.

NDIRITU, J. G. & DANIELL, T. M. 2001. An improved genetic algorithm for rainfall-runoff model calibration and function optimization. *Mathematical and Computer Modelling,* 33**,** 695-706.

NEW, M., LOPEZ, A., DESSAI, S. & WILBY, R. 2007. Challenges in using probabilistic climate change information for impact assessments: an example from the water sector. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 365**,** 2117-2131.

NICHOLS, J., EICHINGER, W., COOPER, D., PRUEGER, J., HIPPS, L., NEALE, C. & BAWAZIR, A. 2004. Comparison of evaporation estimation methods for a riparian area. *Final report. UHR Technical Report*.

NORTHEY, J. E., CHRISTEN, E. W., AYARS, J. E. & JANKOWSKI, J. 2006. Occurrence and measurement of salinity stratification in shallow groundwater in the Murrumbidgee Irrigation Area, south-eastern Australia, *Agricultural Water Management*, 81(1–2), 23-40.

NOSSENT, J., ELSEN, P. & BAUWENS, W. 2011. Sobol'sensitivity analysis of a complex environmental model. *Environmental Modelling & Software,* 26**,** 1515-1525.

OSIDELE, O. & BECK, M. 2001. Identification of model structure for aquatic ecosystems usingregionalized sensitivity analysis. *Water Science & Technology,* 43**,** 271-278.

OUDIN, L., MICHEL, C. & ANCTIL, F. 2005a. Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 1—Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs?, *Journal of Hydrology*, 303(1–4), 275-289.

OUDIN, L., HERVIEU, F., MICHEL, C., PERRIN, C., ANDRÉASSIAN, V., ANCTIL, F. & LOUMAGNE, C. 2005b. Which potential evapotranspiration input for a lumped rainfall–runoff model?: Part 2—Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling. *Journal of Hydrology,* 303**,** 290-306.

OUDIN, L., PERRIN, C., MATHEVET, T., ANDRÉASSIAN, V. & MICHEL, C. 2006. Impact of biased and randomly corrupted inputs on the efficiency and the parameters of watershed models. *Journal of Hydrology,* 320**,** 62-83.

PARTINGTON, D., BRUNNER, P., FREI, S., SIMMONS, C. T., WERNER, A. D., THERRIEN, R., MAIER, H. R., DANDY, G. C. & FLECKENSTEIN, J. H. 2013. Interpreting streamflow generation mechanisms from integrated surface-subsurface flow models of a riparian wetland and catchment, *Water Resources Research*, 49(9), 5501-5519.

PATON, F., MAIER, H. & DANDY, G. 2013a. Relative magnitudes of sources of uncertainty in assessing climate change impacts on water supply security for the southern Adelaide water supply system. *Water Resources Research,* 49**,** 1643-1667.

PATON, F. L., MAIER, H. R. & DANDY, G. C. 2014. Including adaptation and mitigation responses to climate change in a multiobjective evolutionary algorithm framework for urban water supply systems incorporating GHG emissions. *Water Resources Research,* 50**,** 6285-6304.

PENMAN, H. L. 1948. Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences,* 193**,** 120-145.

PENMAN, H. L. 1956. Evaporation: An introductory survey. *Netherlands Journal of Agricultural Science,* 4**,** 9-29.

PERRIN, C., MICHEL, C. & ANDRÉASSIAN, V. 2003. Improvement of a parsimonious model for streamflow simulation. *Journal of hydrology,* 279**,** 275-289.

PETROW, T. & MERZ, B. 2009. Trends in flood magnitude, frequency and seasonality in Germany in the period 1951–2002. *Journal of Hydrology,* 371**,** 129-141.

POFF, N. L., BROWN, C. M., GRANTHAM, T. E., MATTHEWS, J. H., PALMER, M. A., SPENCE, C. M., WILBY, R. L., HAASNOOT, M., MENDOZA, G. F., DOMINIQUE, K. C. & BAEZA, A. 2015. Sustainable water management under future uncertainty with eco-engineering decision scaling. *Nature Clim. Change,* advance online publication.

PRIESTLEY, C. & TAYLOR, R. 1972. On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly weather review,* 100**,** 81-92.

PRUDHOMME, C., CROOKS, S., KAY, A. & REYNARD, N. 2013a. Climate change and river flooding: part 1 classifying the sensitivity of British catchments. *Climatic Change,* 119**,** 933-948.

PRUDHOMME, C., GIUNTOLI, I., ROBINSON, E. L., CLARK, D. B., ARNELL, N. W., DANKERS, R., FEKETE, B. M., FRANSSEN, W., GERTEN, D., GOSLING, S. N., HAGEMANN, S., HANNAH, D. M., KIM, H., MASAKI, Y., SATOH, Y., STACKE, T., WADA, Y. & WISSER, D. 2014. Hydrological droughts in the 21st century, hotspots and uncertainties from a global multimodel ensemble experiment. *Proceedings of the National Academy of Sciences,* 111**,** 3262-3267.

PRUDHOMME, C., KAY, A. L., CROOKS, S. & REYNARD, N. 2013b. Climate change and river flooding: Part 2 sensitivity characterisation for british catchments and example vulnerability assessments. *Climatic Change,* 119**,** 949-964.

PRUDHOMME, C., REYNARD, N. & CROOKS, S. 2002. Downscaling of global climate models for flood frequency analysis: where are we now? *Hydrological Processes,* 16**,** 1137-1150.

PRUDHOMME, C., WILBY, R. L., CROOKS, S., KAY, A. L. & REYNARD, N. S. 2010. Scenario-neutral approach to climate change impact studies: Application to flood risk. *Journal of Hydrology,* 390**,** 198-209.

PRUDHOMME, C. & WILLIAMSON, J. 2013. Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: a comparison of methods and associated uncertainty in future projections. *Hydrology and Earth System Sciences,* 17**,** 1365-1377.

PUJOL, G., IOOSS, B. & IOOSS, M. B. 2014. Package 'sensitivity'.

RACSKO, P., SZEIDL, L. & SEMENOV, M. 1991. A serial approach to local stochastic weather models. *Ecological modelling,* 57**,** 27-41.

RAES, D. & MUNOZ, G. 2008. The ETo Calculator. *Land and water digital media series*.

RAJAH, K., O'LEARY, T., TURNER, A., PETRAKIS, G., LEONARD, M. & WESTRA, S. 2014. Changes to the temporal distribution of daily precipitation. *Geophysical Research Letters,* 41**,** 8887-8894.

RAVAZZANI, G., GHILARDI, M., MENDLIK, T., GOBIET, A., CORBARI, C. & MANCINI, M. 2014. Investigation of Climate Change Impact on Water Resources for an Alpine Basin in Northern Italy: Implications for Evapotranspiration Modeling Complexity. *PLOS ONE,* 9**,** e109053.

RICHARDSON, C. W. 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resources Research,* 17**,** 182-190.

RICHARDSON, C. W. & WRIGHT, D. A. 1984. *WGEN: A model for generating daily weather variables*, US Department of Agriculture, Agricultural Research Service Washington, DC, USA.

RODERICK, M. L. 1999. Estimating the diffuse component from daily and monthly measurements of global radiation. *Agricultural and Forest Meteorology,* 95**,** 169-185.

RODERICK, M. L. & FARQUHAR, G. D. 2002. The Cause of Decreased Pan Evaporation over the Past 50 Years. *Science,* 298**,** 1410-1411.

RODERICK, M. L., ROTSTAYN, L. D., FARQUHAR, G. D. & HOBBINS, M. T. 2007. On the attribution of changing pan evaporation. *Geophysical research letters,* 34.

ROSENBERRY, D. O., WINTER, T. C., BUSO, D. C. & LIKENS, G. E. 2007. Comparison of 15 evaporation methods applied to a small mountain lake in the northeastern USA. *Journal of Hydrology,* 340**,** 149-166.

ROTSTAYN, L. D., RODERICK, M. L. & FARQUHAR, G. D. 2006. A simple pan-evaporation model for analysis of climate simulations: Evaluation over Australia. *Geophysical Research Letters,* 33.

ROY, T., GUPTA, H. V., SERRAT-CAPDEVILA, A. & VALDES, J. B. 2016. Using Satellite-Based Evapotranspiration Estimates to Improve the Structure of a Simple Conceptual Rainfall-Runoff Model. *Hydrol. Earth Syst. Sci. Discuss.,* 2016**,** 1-28.

RUSTOMJI, P., BENNETT, N. & CHIEW, F. 2009. Flood variability east of Australia's great dividing range. *Journal of Hydrology,* 374**,** 196-208.

SAFT, M., PEEL, M. C., WESTERN, A. W., PERRAUD, J.-M. & ZHANG L. 2016, Bias in streamflow projections due to climate-induced shifts in catchment response, *Geophysical Research Letters*, 43(4), 1574-1581.

SALTELLI, A., ANNONI, P., AZZINI, I., CAMPOLONGO, F., RATTO, M. & TARANTOLA, S. 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications,* 181**,** 259-270.

SEIBERT, J. & MCDONNELL, J. J. 2002. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resources Research*, 38(11), 23-21-23-14.

SEIBERT, J., RODHE, A., & BISHOP, K. 2003. Simulating interactions between saturated and unsaturated storage in a conceptual runoff model, *Hydrological Processes*, 17(2), 379-390.

SEMENOV, M. & BARROW, E. 1997. USE OF A STOCHASTIC WEATHER GENERATOR IN THE DEVELOPMENT OF CLIMATE CHANGE SCENARIOS. *Climatic Change,* 35**,** 397-414.

SEMENOV, M. A. 2007. Simulation of extreme weather events by a stochastic weather generator. *Climate Research,* 35**,** 203.

SEMENOV, M. A. & BROOKS, R. J. 1999. Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain. *Climate Research,* 11**,** 137-148.

SENEVIRATNE, S. I., LEHNER, I., GURTZ, J., TEULING, A. J., LANG, H., MOSER, U., GREBNER, D., MENZEL, L., SCHROFF, K., VITVAR, T. & ZAPPA, M. 2012. Swiss prealpine Rietholzbach research catchment and lysimeter: 32 year time series and 2003 drought event. *Water Resources Research,* 48, n/a-n/a.

SENEVIRATNE, S. I., WILHELM, M., STANELLE, T., VAN DEN HURK, B., HAGEMANN, S., BERG, A., CHERUY, F., HIGGINS, M. E., MEIER, A., BROVKIN, V., CLAUSSEN, M., DUCHARNE, A., DUFRESNE, J.-L., FINDELL, K. L., GHATTAS, J., LAWRENCE, D. M., MALYSHEV, S., RUMMUKAINEN, M. & SMITH, B. 2013. Impact of soil moisture-climate feedbacks on CMIP5 projections: First

results from the GLACE-CMIP5 experiment. *Geophysical Research Letters,* 40**,** 5212-5217.

SHAFII, M. & DE SMEDT, F. 2009. Multi-objective calibration of a distributed hydrological model (WetSpa) using a genetic algorithm. *Hydrol. Earth Syst. Sci.,* 13**,** 2137-2149.

SHIN, M.-J., GUILLAUME, J. H. A., CROKE, B. F. W. & JAKEMAN, A. J. 2013. Addressing ten questions about conceptual rainfall–runoff models with global sensitivity analyses in R. *Journal of Hydrology,* 503**,** 135-152.

SHUTTLEWORTH, W. & WALLACE, J. 2009. Calculating the water requirements of irrigated crops in Australia using the Matt-Shuttleworth approach. *Transactions of the ASABE,* 52**,** 1895-1906.

SIEBER, A. & UHLENBROOK, S. 2005. Sensitivity analyses of a distributed catchment model to verify the model structure. *Journal of Hydrology,* 310**,** 216-235.

SINGH, R., WAGENER, T., CRANE, R., MANN, M. E. & NING, L. 2014. A vulnerability driven approach to identify adverse climate and land use change combinations for critical hydrologic indicator thresholds: Application to a watershed in Pennsylvania, USA. *Water Resources Research,* 50**,** 3409-3427.

SLATYER, R. & MCLLROY, I. 1961. Practical Microclimatology, 300 pp., Commonwealth Scientific and Industrial Research Organization Australia. UNESCO, Paris, France.

SOBOL', I. M., TARANTOLA, S., GATELLI, D., KUCHERENKO, S. S. & MAUNTZ, W. 2007. Estimating the approximation error when fixing unessential factors in global sensitivity analysis. *Reliability Engineering & System Safety,* 92**,** 957-960.

SORG, A., BOLCH, T., STOFFEL, M., SOLOMINA, O. & BENISTON, M. 2012. Climate change impacts on glaciers and runoff in Tien Shan (Central Asia). *Nature Clim. Change,* 2**,** 725-731.

STEIN, M. 1987. Large sample properties of simulations using Latin hypercube sampling. *Technometrics,* 29**,** 143-151.

STEINSCHNEIDER, S. & BROWN, C. 2013. A semiparametric multivariate, multi-site weather generator with low-frequency variability for use in climate risk assessments. *Water Resources Research***,** n/a-n/a.

STERN, H., DE HOEDT, G. & ERNST, J. 2000. Objective classification of Australian climates. *Australian Meteorological Magazine,* 49**,** 87-96.

STOCKER, T. F., QIN, D., PLATTNER, G.-K., TIGNOR, M., ALLEN, S. K., BOSCHUNG, J., NAUELS, A., XIA, Y., BEX, V. & MIDGLEY, P. M. 2013. Climate change 2013: The physical science basis. *Intergovernmental Panel on Climate Change, Working Group I Contribution to the IPCC Fifth Assessment Report (AR5)(Cambridge Univ Press, New York).*

SZILAGYI, J. 2007. On the inherent asymmetric nature of the complementary relationship of evaporation. *Geophysical Research Letters,* 34**,** L02405.

SU, C.-H., COSTELLOE, J. F., PETERSON, T. J., & WESTERN, A. W. 2016. On the structural limitations of recursive digital filters for base flow estimation, *Water Resources Research,* 52(6), 4745-4764.

TABARI, H., GRISMER, M. E. & TRAJKOVIC, S. 2013. Comparative analysis of 31 reference evapotranspiration methods under humid conditions. *Irrigation Science,* 31**,** 107-117.

TABARI, H. & HOSSEINZADEH TALAEE, P. 2014. Sensitivity of evapotranspiration to climatic change in different climates. *Global and Planetary Change,* 115**,** 16-23.

TANG, Y., REED, P., VAN WERKHOVEN, K. & WAGENER, T. 2007a. Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis. *Water Resources Research,* 43**,** n/a-n/a.

TANG, Y., REED, P., WAGENER, T. & VAN WERKHOVEN, K. 2007b. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrology and Earth System Sciences Discussions,* 11**,** 793-817.

TAYLOR, I. H., BURKE, E., MCCOLL, L., FALLOON, P. D., HARRIS, G. R. & MCNEALL, D. 2013. The impact of climate mitigation on projections of future drought. *Hydrol. Earth Syst. Sci.,* 17**,** 2339-2358.

TENG, J., VAZE, J., CHIEW, F. H. S., WANG, B. & PERRAUD, J.-M. 2012. Estimating the Relative Uncertainties Sourced from GCMs and Hydrological Models in Modeling Climate Change Impact on Runoff. *Journal of Hydrometeorology,* 13**,** 122-139.

TERN. 2012. Coping with complex data streams: the OzFlux approach, TERN e-Newsletter October 2012. http://mymail.ezemsgs.com/em/message/email/view.php?id=959668&u=45618.

THOMPSON, J. R., GREEN, A. J. & KINGSTON, D. G. 2014. Potential evapotranspiration-related uncertainty in climate change impacts on river flow: An assessment for the Mekong River basin. *Journal of Hydrology,* 510**,** 259-279.

THYER, M., KUCZERA, G. & BATES, B. C. 1999. Probabilistic optimization for conceptual rainfall-runoff models: A comparison of the shuffled complex evolution and simulated annealing algorithms. *Water Resources Research,* 35**,** 767-773.

TURC, L. 1961. Estimation of irrigation water requirements, potential evapotranspiration: a simple climatic formula evolved up to date. *Ann. Agron,* 12**,** 13-49.

TURRAL, H., BURKE, J. & FAURÈS, J.-M. 2011. Climate change, water and food security, FAO water reports 36, FAO, Rome.

VALIANTZAS, J. D. 2006. Simplified versions for the Penman evaporation equation using routine weather data. *Journal of Hydrology,* 331**,** 690-702.

VANGORSEL, E. 2013. Tumbarumba OzFlux tower site OzFlux: Australian and New Zealand Flux Research and Monitoring hdl: 102.100.100/14241.

VAN DER KAMP, G., HAYASHI, M. & GALLÉN, D. 2003. Comparing the hydrology of grassed and cultivated catchments in the semi-arid Canadian prairies. *Hydrological Processes,* 17**,** 559-575.

VAN GRIENSVEN, A., MEIXNER, T., GRUNWALD, S., BISHOP, T., DILUZIO, M. & SRINIVASAN, R. 2006. A global sensitivity

analysis tool for the parameters of multi-variable catchment models. *Journal of Hydrology*, 324**,** 10-23.

VANO, J. A., KIM, J. B., RUPP, D. E. & MOTE, P. W. 2015. Selecting climate change scenarios using impact-relevant sensitivities. *Geophysical Research Letters*, 42**,** 5516-5525.

VAZE, J., POST, D. A., CHIEW, F. H. S., PERRAUD, J. M., VINEY, N. R. & TENG, J. 2010. Climate non-stationarity – Validity of calibrated rainfall–runoff models for use in climate change studies, *Journal of Hydrology*, 394(3–4), 447-457.

VAZE, J. & TENG, J. 2011. Future climate and runoff projections across New South Wales, Australia: results and practical applications. *Hydrological Processes*, 25**,** 18-35.

VELÁZQUEZ, J. A., SCHMID, J., RICARD, S., MUERTH, M. J., GAUVIN ST-DENIS, B., MINVILLE, M., CHAUMONT, D., CAYA, D., LUDWIG, R. & TURCOTTE, R. 2013. An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrology & Earth System Sciences*, 17**,** 565-578.

VERVOORT, W. R., MIECHELS, S. F., VAN OGTROP, F. F. & GUILLAUME, J. H. A. 2014. Remotely sensed evapotranspiration to calibrate a lumped conceptual model: Pitfalls and opportunities, *Journal of Hydrology*, 519, Part D, 3223-3236.

VICENTE-SERRANO, S. M., AZORIN-MOLINA, C., SANCHEZ-LORENZO, A., REVUELTO, J., MORÁN-TEJEDA, E., LÓPEZ-MORENO, J. I. & ESPEJO, F. 2014. Sensitivity of reference evapotranspiration to changes in meteorological parameters in Spain (1961–2011), *Water Resources Research*, 50(11), 8458-8480

VINCENT, L. A., ZHANG, X., BROWN, R. D., FENG, Y., MEKIS, E., MILEWSKA, E. J., WAN, H. & WANG, X. L. 2015. Observed Trends in Canada's Climate and Influence of Low-Frequency Variability Modes. *Journal of Climate*, 28**,** 4545-4560.

WAGENER, T., MCINTYRE, N., LEES, M. J., WHEATER, H. S. & GUPTA, H. V. 2003. Towards reduced uncertainty in conceptual rainfall-runoff modelling: dynamic identifiability analysis, *Hydrological Processes*, 17(2), 455-476.

WANG, K. & DICKINSON, R. E. 2012. A review of global terrestrial evapotranspiration: Observation, modeling, climatology, and climatic variability, *Reviews of Geophysics*, 50(2).

WANG, Y. C., YU, P. S. & YANG, T. C. 2010. Comparison of genetic algorithms and shuffled complex evolution approach for calibrating distributed rainfall–runoff model. *Hydrological processes*, 24**,** 1015-1026.

WESTRA, S., ALEXANDER, L. V. & ZWIERS, F. W. 2013. Global increasing trends in annual maximum daily precipitation. *Journal of Climate*, 26**,** 3904-3918.

WESTRA, S., FOWLER, H. J., EVANS, J. P., ALEXANDER, L. V., BERG, P., JOHNSON, F., KENDON, E. J., LENDERINK, G. & ROBERTS, N. M. 2014a. Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, 52**,** 522-555.

WESTRA, S., THYER, M., LEONARD, M., KAVETSKI, D. & LAMBERT, M. 2014b. A strategy for diagnosing and interpreting hydrological model nonstationarity. *Water Resources Research,* 50**,** 5090-5113.

WESTRA, S., THYER, M., LEONARD, M. & LAMBERT, M. 2014c. Impacts of Climate Change on Surface Water in the Onkaparinga Catchment–Volume 3: Impact of Climate Change on Runoff in the Onkaparinga Catchment. *Goyder Institute for Water Research Technical Report Series*.

WHATELEY, S., STEINSCHNEIDER, S. & BROWN, C. 2014. A climate change range-based method for estimating robustness for water resources supply. *Water Resources Research,* 50**,** 8944-8961.

WILBY, R. L., DAWSON, C. W., MURPHY, C., O' CONNOR, P. & HAWKINS, E. 2014. The Statistical DownScaling Model - Decision Centric (SDSM-DC): conceptual basis and applications. *Climate Research,* 61**,** 259-276.

WILBY, R. L. & DESSAI, S. 2010. Robust adaptation to climate change. *Weather,* 65**,** 180-185.

WILBY, R. L., WHITEHEAD, P. G., WADE, A. J., BUTTERFIELD, D., DAVIS, R. J. & WATTS, G. 2006. Integrated modelling of climate change impacts on water resources and quality in a lowland catchment: River Kennet, UK. *Journal of Hydrology,* 330**,** 204-220.

WILCKE, R. A. I. & BÄRRING, L. 2016. Selecting regional climate scenarios for impact modelling studies. *Environmental Modelling & Software,* 78**,** 191-201.

WILEY, M. & PALMER, R. 2008. Estimating the Impacts and Uncertainty of Climate Change on a Municipal Water Supply System. *Journal of Water Resources Planning and Management,* 134**,** 239-246.

WILKS, D. S. 1999. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agricultural and Forest Meteorology,* 93**,** 153-169.

WILSON, K. B., HANSON, P. J., MULHOLLAND, P. J., BALDOCCHI, D. D. &WULLSCHLEGER, S. D. 2001. A comparison of methods for determining forest evapotranspiration and its components: sap-flow, soil water budget, eddy covariance and catchment water balance, *Agricultural and Forest Meteorology,* 106(2), 153-168.

XU, C. Y. & SINGH, V. P. 2000. Evaluation and generalization of radiation-based methods for calculating evaporation. *Hydrological Processes,* 14**,** 339-349.

XU, C. Y. & SINGH, V. P. 2002. Cross comparison of empirical equations for calculating potential evapotranspiration with data from Switzerland. *Water Resources Management,* 16**,** 197-219.

YATES, D. N., MILLER, K. A., WILBY, R. L. & KAATZ, L. 2015. Decision-centric adaptation appraisal for water management across Colorado's Continental Divide. *Climate Risk Management,* 10**,** 35-50.

YIN, Y., WU, S., CHEN, G. & DAI, E. 2010. Attribution analyses of potential evapotranspiration changes in China since the 1960s. *Theoretical and Applied Climatology,* 101**,** 19-28.

YIN, Z.-Y. & BROOK, G. A. 1992. Evapotranspiration in the Okefenokee Swamp watershed: a comparison of temperature-based and water balance methods. *Journal of Hydrology,* 131**,** 293-312.

ZEILEIS, A., GROTHENDIECK, G., RYAN, J. A., ANDREWS, F. & ZEILEIS, M. A. 2015. Package ‘zoo’.

ZHANG, L., DAWES, W. R., SLAVICH, P. G., MEYER, W. S., THORBURN, P. J., SMITH, D. J. & WALKER, G. R. 1999. Growth and ground water uptake responses of lucerne to changes in groundwater levels and salinity: lysimeter, isotope and modelling studies, *Agricultural Water Management*, 39(2–3), 265-282.

ZHANG, X. Y., TRAME, M. N., LESKO, L. J. & SCHMIDT, S. 2015. Sobol Sensitivity Analysis: A Tool to Guide the Development and Evaluation of Systems Pharmacology Models. *CPT: Pharmacometrics & Systems Pharmacology*, 4**,** 69-79.