
THE UNIVERSITY of ADELAIDE

# Improved Implementation of Scenario-Neutral Climate Impact Assessments for Water Resource Systems

Sam Anthony Culley

B. Eng Civil & Structural Engineering (Hons)

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Adelaide
Faculty of Engineering, Computer and Mathematical Sciences
School of Civil, Environmental and Mining Engineering

-April 2019-

# Table of Contents

**TABLE OF CONTENTS** **I**

**ABSTRACT** **III**

**STATEMENT OF ORIGINALITY** **V**

**ACKNOWLEDGEMENTS** **VI**

**LIST OF FIGURES** **VII**

**LIST OF TABLES** **XI**

**CHAPTER 1** **1**

1.1 Background on climate impact assessments 1

1.2 Research challenges for climate impact assessments 4

1.3 Overall Research Objectives 6

1.4 Thesis Organisation 7

**CHAPTER 2** **9**

Abstract 12

2.1 Introduction 13

2.2 Formalizing the Inverse Approach to Stochastic Generation 16

2.3 Case study 26

2.4 Investigation into the impact of increasing the efficiency of optimization 30

2.5 Ensuring the realism of hydrometeorological time series 35

2.6 Conclusions 45

2.7 Acknowledgements 47

2.8 References 47

**CHAPTER 3** **51**

**ABSTRACT** **54**

3.1 Introduction 55

3.2 Approach to identifying critical climate attributes 58

3.3 Case study, implementation and testing of proposed approach 68

3.4 Results and discussion 75

3.5 Summary and Conclusions 85

3.6 Acknowledgments 86

3.7 References 87

**CHAPTER 4** **95**

**ABSTRACT** **98**

4.1 Introduction 99

4.2 Pitfalls of time series generation for scenario-neutral studies 102

4.3 Methods 104

4.4 Results 116

4.5 Discussion 127

4.6 Conclusion 131

4.7 Acknowledgements 132

4.8 References 132

**CHAPTER 5** **138**

5.1 Research contribution 138

5.2 Limitations and future work 140

5.3 Final Recommendations 144

**REFERENCES** **146**

**APPENDIX A** **151**

**APPENDIX B** **164**

# Abstract

The uncertainty surrounding climate change presents a major challenge for the management of water resource systems, which are facing stresses to both supply and demand. To provide insight into how a system might perform under change, scenario-neutral climate impact assessments are being used increasingly to supplement 'scenario-led' climate projections. Scenario-neutral assessments stress-test a system against a range of potential climate changes, regardless of their plausibility, so that all major modes of system failure can be identified and characterised. This thesis focuses on overcoming several existing challenges with the implementation of scenario-neutral methods for complex systems. The specific aims of this research are to: (i) improve current methods of generating climate perturbed hydrometeorological time series to consider a wider set of changes, (ii) develop a method to identify the critical changes in climate conditions for inclusion in scenario-neutral climate impact assessments, and (iii) identify and demonstrate the key requirements of a scenario-neutral analysis, such that it will be consistent with the outcomes of a scenario-led analysis. These methods are demonstrated using two case studies: the Lake Como reservoir (Italy) for impact assessments, and Adelaide rainfall data (Australia) for time series generation.

For the first aim, this research advances the optimisation formulation that underpins an inverse approach to time series generation. This process uses formal optimisation techniques to identify the parameters of a stochastic weather generator that enable the generation of time series with desired climate attributes (statistics of climate variables). The advancements enable a greater number of climate attributes to be perturbed, while ensuring the realism of the time series. This allows scenario-neutral assessments to be implemented for more complex systems that require stress-testing beyond changes to means and seasonality of climate variables.

For the second aim, a method to identify the critical climate attributes for a system is proposed, which uses the partial mutual information algorithm to rank

a set of candidate attributes in order of significance. Critical attributes are then selected by considering the value of adding an additional attribute given its relative increase in the description of system performance. This allows the resulting scenario-neutral assessment to ensure that the modes of failure identified will be those to which a system is most vulnerable. Applied to the Lake Como reservoir, results show that an impact assessment using identified critical attributes such as frost days uncovers modes of flood prevention and irrigation supply failure not identified by the commonly used attributes average temperature and rainfall.

For the final aim, four key pitfalls in the scenario-neutral approach are identified and their effects are demonstrated using a set of diagnostics that compares implementations of the scenario-neutral approach with the results of a scenario-led analysis. Techniques for avoiding the pitfalls are also presented, building on the preceding advances in attribute identification and time series generation. Results show that when these techniques are applied, it is possible to reconcile scenario-neutral and scenario-led approaches. Collectively, this body of research improves the practical application of a scenario-neutral approach to better deliver on its underpinning principles.

# Statement of Originality

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

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.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

23-04-19

Sam Anthony Culley Date

# Acknowledgements

Firstly, I would like to thank my supervisors, Prof. Seth Westra, Prof. Holger Maier and Dr. Bree Bennett, for their guidance and support throughout my PhD. Their enthusiasm for research and problem solving has made every part of this project enjoyable, even when I was feeling overwhelmed and facing numerous challenges. I am very grateful to have had the opportunity to work with them these past four years, and am confident the skills I have developed as a result will be invaluable as I continue with my career.

I would also like to thank my colleagues and fellow PhD students in the hydrology and climate research group for their friendship and conversation. I have been grateful to have felt like part of a wider research group while undertaking this PhD. Thanks in particular go to Cameron McPhail and Robert Rundle, for always being willing to talk through problems. In addition, I would like to thank my friends I have lived with over the last four years, for helping me through the more stressful moments.

Last but not least, I would like to thank my family for their support and patience throughout. The last four years wouldn't have been possible without it.

# List of Figures

Figure 1-1 A scenario-neutral space, showing system response to percentages of historical precipitation and additive changes to historical temperature [*Culley et al.*, 2016]. .....3

Figure 1-2 The link between each research objective and the scenario-neutral approach to climate impact assessments.....7

Figure 2-1 Conceptual illustration of the inverse approach to stochastic generation. Each blue dot corresponds to a set of weather generator parameters that result in the generation of weather time series that have the set of climate attributes of one of the orange dots. Equal coverage of the climate attribute space ($a_1$, $a_2$) does not necessarily reflect equal coverage in the parameter space ($\theta_1$, $\theta_2$) of stochastic weather generators.....17

Figure 2-2 Generic steps in the optimization loop that underpins the inverse approach. .....19

Figure 2-3 The over-constrained optimization challenge: the grey surface indicates the feasible subspace, with equation $a_{(\text{ann. total rainfall})}=a_{(\text{no. wet days})}*a_{(\text{wet day amounts})}$. Historical conditions are shown at point (1) and the target is indicated by point (2). Point (3) indicates the shortest Euclidean distance between the target at point (2), and the feasible subspace. Point (4) indicates the shortest distance between the target and the feasible subspace, while ensuring zero error in the perturbed attribute. .....22

Figure 2-4 How penalty functions can change the fitness landscape to create a new minimum error solution during over-constrained optimization: (top panels) linear penalty term, and (bottom panels) quadratic penalty term.....25

Figure 2-5 Two-dimensional slices between the objective function value and both the Pwd, Pdd parameters (right) and the $\alpha$, $\beta$ parameters (left) from the four-parameter weather generator. The "unweighted" objective function is calculated for three attributes: total annual rainfall, number of wet days and 99$^{th}$ percentile rainfall. .....31

Figure 2-6 Optimization objective function values at each generation for two sets of parameter bounds: domain knowledge informed and uninformed. A target of historical conditions is searched for with both sets of bounds using an annual weather generator (left panel) and a seasonal weather generator (right panel).....35

Figure 2-7 Breakdown of the error in each attribute at the end of optimization for two target time series: zero change from the historical conditions (top) and a 30% decrease in total annual rainfall (bottom). The time series are formed using an annual weather generator and a linear penalty term where the scaling parameter $\lambda$ is varied.....37

Figure 2-8 Breakdown of the error in each attribute at the end of optimization with a varying scaling parameter $\lambda$ for two different penalties: a linear penalty term (top) and a quadratic penalty term (bottom). Time series with a requested 30% decrease in *Ptot* are simulated using a seasonal weather generator.....40

Figure 2-9 Breakdown of the error in each attribute at the end of optimization with two

"perturbed" attributes, *Ptot* and *nWet*. Time series are simulated using the seasonal weather generator for a target of historical conditions. Scaling parameters for the linear penalty term are changed separately for both attributes. ..... 42
Figure 2-10 A 4x4 scenario-neutral space made with the "unweighted" objective function (left) and the objective function with linear penalty terms (right). Targets are specified as percentage change from historical conditions, and the time series are made with a seasonal weather generator. The mean percentage error from historical conditions in the remaining "held" attributes is shown below each target. The red circle illustrates a point that requires the error in the "held" attributes to be doubled to reach the perturbed target. ..... 44
Figure 3-1 Main steps in proposed approach for selecting critical climate attributes (solid boxes) and the methods used to implement them (dashed boxes). ..... 60
Figure 3-2 Process for generating the climate time series that meet the candidate attribute targets, and resulting performance for critical attribute selection. ..... 63
Figure 3-3 Cumulative variance explained with each additional attribute in order of significance for the two performance criteria. ..... 76
Figure 3-4 PMI selection of the critical attributes and the next most significant attribute for the flood reliability criterion (left), and the irrigation deficit criterion (right). The residuals are plotted in black, and the MLPNN estimation is shown in orange. The function g in the axis labels represents the effect of the previous selected attributes on the attribute selected that iteration, A, and the outputs, P. ..... 80
Figure 3-5 Two scenario neutral spaces delineated by success and failure of an irrigation performance criteria with climate projections overlaid. One space is created using mean precipitation and temperature (top) and the other using the critical climate attributes (bottom). ..... 82
Figure 3-6 Two scenario neutral spaces delineated by success and failure of a flood performance criteria with climate projections overlaid. One space is created using mean precipitation and temperature (top) and the other using the critical climate attributes (bottom). ..... 84
Figure 4-1 Key steps in the generation of climate time series for scenario-neutral studies, common methods used and resulting pitfalls. ..... 103
Figure 4-2 The scenario-neutral framework, selected methods and pitfalls. Five implementations of the scenario-neutral approach are presented, to demonstrate the effect of four pitfalls. A comparison with a scenario-led approach is used as a diagnostic. .... 106
Figure 4-3 A scenario-neutral space for the irrigation deficit performance criteria for the benchmark, where perturbations to historical climate are presented for mean precipitation (Ptot) and the number of frost days (F0) in the x and y axes respectively. Changes from left to right show the temperature in June (TJun), and changes from top to bottom show the total rainfall in summer (PJJAtot). Climate projections are overlaid, with the color

inside the circles showing the scenario-led performance. ..... 117
Figure 4-4 A diagnostic for the benchmark methods. Fifty replicates of a stochastic weather generator are used to create a target that matches all candidate attributes (left panel) and matches the four critical attributes (right panel) of the projections, with the resulting spread in performance shown. The one-to-one line is shown in blue. ..... 118
Figure 4-5 A scenario-neutral space for the irrigation deficit performance criteria for pitfall 1, where perturbations to historical climate are presented for mean precipitation (*Ptot*) and mean temperature (*Tavg*) in the x and y axes respectively. Climate projections are overlaid, with the color inside the circles showing the scenario-led performance. ..... 120
Figure 4-6 A diagnostic for pitfall 1. Fifty replicates of a stochastic weather generator are used to create targets for the benchmark implementation (left), and targets that match only the *Ptot* and *Tavg* attributes of the projections (right panel), with the resulting spread in performance shown. The one-to-one line is shown in blue. ..... 121
Figure 4-7 A scenario-neutral space for the irrigation deficit performance criteria for pitfall 2, where perturbations to historical climate are presented for total annual precipitation (*Ptot*) and the total winter precipitation (*PDJFtot*) in the x and y axes respectively. Changes from left to right show the total rainfall in summer (*PJJAtot*). The temperature time series is constant throughout. Climate projections are overlaid, with the color inside the circles showing the scenario-led performance. ..... 122
Figure 4-8 A diagnostic for pitfall 2. Fifty replicates of a stochastic weather generator are used to create targets for the benchmark implementation (left panel), and targets that match only the three critical precipitation attributes of the projections (right panel), with the resulting spread in performance shown. The one-to-one line is shown in blue. ..... 123
Figure 4-9 A scenario-neutral space for the irrigation deficit performance criteria for pitfall 3, where an annual precipitation weather generator is used to create the climate time series. Perturbations to historical climate are presented for mean precipitation (*Ptot*) and the number of frost days (*F0*) in the *x* and *y* axes respectively. Changes from left to right show the temperature in June (*TJun*), and changes from top to bottom show the total rainfall in summer (*PJJAtot*). Climate projections are overlaid, with the color inside the circles showing the scenario-led performance. ..... 124
Figure 4-10 A diagnostic for pitfall 3. Fifty replicates of a stochastic weather generator are used to create a target that matches the four critical attributes of the projections, with the resulting spread in performance shown. The benchmark results are shown in the left panel, and pitfall 3 is shown in the right panel, where an annual precipitation weather generator is used to create the climate time series. The one-to-one line is shown in blue. ..... 125
Figure 4-11 A scenario-neutral space for the irrigation deficit performance criteria for pitfall 4, where perturbations to historical climate are presented for mean precipitation (*Ptot*) and the number of frost days (*F0*) in the *x* and *y* axes respectively. Changes from left to right

show the temperature in June (*TJun*), and changes from top to bottom show the total rainfall in summer (*PJJAtot*). No other attributes are set as targets. Climate projections are overlaid, with the color inside the circles showing the scenario-led performance. Some irrigation deficit values falls above the range of the color ramp (indicated by grey pixels). ..... 126

Figure 4-12 A diagnostic for pitfall 4. Fifty replicates of a stochastic weather generator are used to create a target that matches the four critical attributes of the projections, with the resulting spread in performance shown. The benchmark results are shown in the left panel, and pitfall 4 is shown in the right panel, where no other attributes are set as targets. The one-to-one line is shown in blue..... 127

# List of Tables

Table 2-1 List of "perturbed" (P) and "held" (H) attributes for annual and seasonal weather generator experiments. ..... 29
Table 2-2 Case study genetic algorithm parameters. ..... 32
Table 2-3 Third standard deviation bounds on the four parameters of the annual weather generator for an Australian data set. ..... 33
Table 2-4 Third standard deviation bounds on the eight parameters of the seasonal weather generator for an Australian data set. ..... 33
Table 3-1 Candidate climate attributes considered, where the attributes are defined as the average over the simulation period. ..... 71
Table 3-2 Expected change in attributes from historical baseline. ..... 72
Table 3-3 Order of critical attributes and the next most significant attribute (in italics) obtained using the PMI algorithm for the two performance criteria considered, as well as the corresponding values of cumulative variance explained (CVE). ..... 77
Table 4-1 Methods used in the scenario-neutral implementations. ..... 106
Table 4-2 RCP/GCM/RCM combinations and reference number for climate projections used in this study. ..... 108
Table 4-3 List of candidate attributes. ..... 111
Table 4-4 Range and bounds of candidate attributes. ..... 112
Table 4-5 Critical climate attributes from the candidate set. ..... 112
Table 4-6 Changes to benchmark methods to demonstrate pitfalls. ..... 115
Table 4-7 Critical attributes given a subset of candidate attributes. ..... 116

# Chapter 1

Globally, water resource systems are facing significant stress as a result of a changing climate [*Milly et al.*, 2008; *IPCC*, 2013]. In many cases, these systems were not designed with this change in mind. Performance of these systems is degrading, and in some cases will reach unacceptable conditions in the future [*Risbey*, 2011; *Paton et al.*, 2013; *IPCC*, 2014]. This necessitates the design of new management options, which given the uncertainty in climate change [*Maier et al.*, 2016] typically aim to be either robust to a wide range of potential changes [*Kasprzyk et al.*, 2013; *Herman et al.*, 2015; *Giuliani and Castelletti*, 2016], or adapt to changes in climate as they occur [*Haasnoot et al.*, 2013; *Kwakkel et al.*, 2016]. New systems also need to be designed with the same objective of accounting for uncertain future changes, to achieve successful levels of performance across their design life. In either case, the development of new management options requires information about how climate change will impact the water resource system in question. A series of techniques called climate impact assessments are used to gain this information, and examine how the performance of a water resource system might change in the future.

## 1.1 Background on climate impact assessments

Initially, the climate impact assessments presented in literature were 'scenario-led' (also known as 'top-down'), with a strong emphasis on the development of rigorous climate model projections. In the context of applications to water resource systems, these projections are provided in the form of climate time series typically leading up to the end of century, so that system performance in response to each projection can be determined at various time intervals [*Arnell*, 2004; *Brekke et al.*, 2009; *Vano et al.*, 2010; *Wilby and Dessai*, 2010; *Anghileri et al.*, 2011; *Giuliani and Castelletti*, 2016]. This snapshot of system performance can be used to characterise the success of different management options, but a full picture of the modes of system vulnerability is not provided. Therefore, the specific mechanisms of climate change that are affecting system

performance remain unknown.

'Scenario-neutral' (also known as 'bottom-up') impact assessments are increasingly being used in an effort to address the weaknesses of the scenario-led approach [*Brown and Wilby*, 2012]. Scenario-neutral approaches were first introduced as a stress-test against a range of changes in climate that extend beyond that range indicated by climate change projections [*Prudhomme et al.*, 2010; *Brown et al.*, 2012]. A decision-relevant definition of system performance, as well as a representative system model and any failure thresholds, are established through engaging with stakeholders [*Ekström et al.*, 2018], so that the performance in response to each climate time series created for the stress-test can then be determined. These time series are defined by the 'attributes'—specific statistics of a climate variable—that are perturbed (e.g. mean annual temperature, $90^{th}$ percentile precipitation). This information can then be presented in a scenario-neutral space, like Figure 1-1 [excerpt from *Culley et al.*, 2016]. In this example a reservoir system, aiming to manage performance in two competing objectives (irrigation deficit and flooded area), has its historical operating rules stress-tested against annual changes in precipitation volume and average temperature. The figure shows the change to those variables relative to historical conditions that would lead to failure in irrigation deficit, flooded area, or both objectives together.

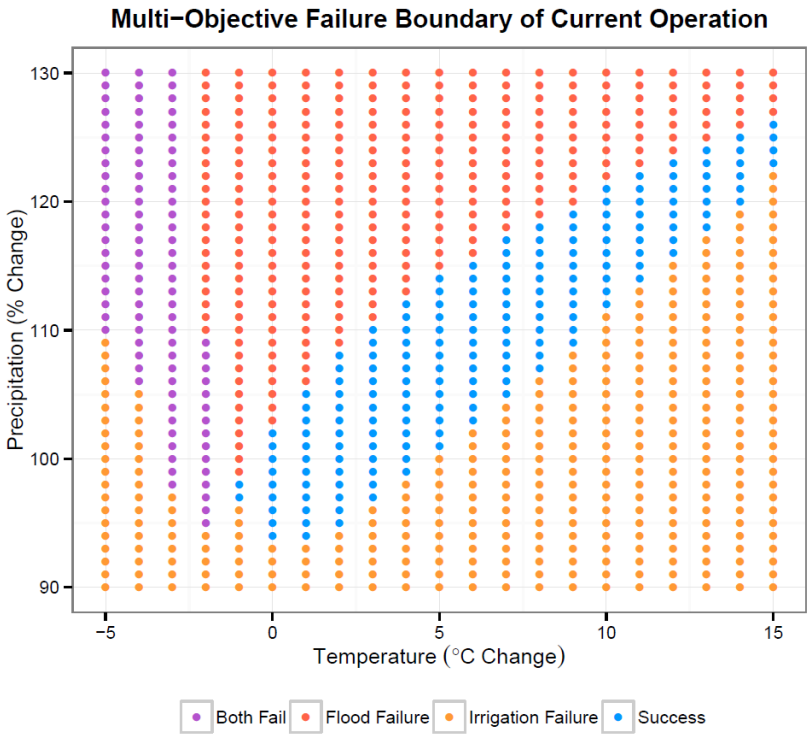


**Figure 1-1** A scenario-neutral space, showing system response to percentages of historical precipitation and additive changes to historical temperature [*Culley et al.*, 2016].

Provided a system is stress-tested against the climate attributes to which it is most sensitive, all the major modes of failure can then be identified. This allows for planning directly in response to this information. In the above example, *Culley et al.* [2016] continue the analysis to identify the optimal reservoir operation in response to each climate time series. This allows for an adaptive approach to be taken, and as the climate changes the reservoir operation can update accordingly [*Haasnoot et al.*, 2013; *Kwakkel et al.*, 2016]. Alternatively, a robust approach can be used, where such scenario-neutral spaces can be used to design the system operation that provides successful performance over the largest number of climate time series [*Brown et al.*, 2012; *Turner et al.*, 2014; *Whateley et al.*, 2014; *Taner et al.*, 2017].

In either case, information about the conditions under which current operation might fail is necessary. Scenario-neutral spaces define this threshold based on specific changes in climate. The likelihood of this failure occurring at a given point in time can then be investigated by considering any relevant climate

projections that indicate such changes in climate occurring [*Turner et al.*, 2014; *Whateley et al.*, 2014; *Culley et al.*, 2016]. Unlike a scenario-led approach, where a selection of climate projections are modelled directly, climate projections can now be interpreted with the context of the scenario-neutral space. This means in addition to a binary understanding of success or failure, an impact assessment can identify how close failure thresholds are to occurring, and which drivers of climate change lead to failure occurring.

## 1.2 Research challenges for scenario-neutral climate impact assessments

Although scenario-neutral approaches are increasingly being recognised as providing significant value to decision makers [*Ray et al.*, 2018], there are several key methodological challenges remaining for the development of scenario-neutral climate impact assessments.

Firstly, the techniques used to generate climate time series currently limit the climate attributes that can be explored in a scenario-neutral analysis. For example, the simplest method to create climate time series is to manually scale historical climate data [*Prudhomme et al.*, 2010]. However, this only allows for direct changes to the means of climate variables, which in most scenario-neutral climate impact assessments are precipitation and temperature [*Prudhomme et al.*, 2010; *Prudhomme et al.*, 2013; *Culley et al.*, 2016; *Wilcke and Bärring*, 2016; *Spence and Brown*, 2018]. This scaling technique can be improved to use seasonal scaling patterns, allowing change in the seasonality, but more complex attributes that measure the persistence and extremes of variables can still not be examined [*Wetterhall et al.*, 2011; *Keller et al.*, 2018]. To perturb these attributes, it is necessary to use stochastic weather generators [*Steinschneider and Brown*, 2013; *Wilby et al.*, 2014; *Bussi et al.*, 2016; *Guo et al.*, 2018]. Some approaches use weather generators to create different replicates of variability, but post-processing (e.g. quantile mapping [*Steinschneider and Brown*, 2013], or direct scaling [*Wilby et al.*, 2014]) is still required to meet intended perturbations. An approach that allows for perturbations to be met without any scaling is the inverse approach [*Guo et al.*, 2017; 2018]. This approach uses

formal optimisation techniques [*Maier et al.*, 2019] to identify the weather generator parameters that make time series with the required climate attributes. However, the approach has only been presented on a maximum of four attributes for one climate variable, and further development is needed to manage the generation of time series to meet perturbations in larger numbers of climate attributes.

Secondly, a key aspect of the scenario-neutral approach is that a system must be stress-tested against the changes to which it is most sensitive. In the wider scenario-neutral literature, the techniques for identifying these changes include global sensitivity analyses [*Gao et al.*, 2016], and scenario discovery [*Lempert et al.*, 2008; *Bryant and Lempert*, 2010; *Kasprzyk et al.*, 2013]. These techniques sample high-dimensional scenario-neutral spaces (including many non-climate factors) to identify the key drivers that affect system performance. However, a feature of these approaches is that the changes in drivers investigated are perturbations of single values, or scaling of time series. When climate time series are required to assess system performance, it becomes very difficult to perturb time series to meet a wide set of attributes, as discussed above. As a result, these methods cannot directly be applied to test a wide range of climate attributes for climate impact assessments. This is likely why many applications of scenario-neutral climate impact assessments that require time series do not test for sensitivity and default to using changes in mean annual precipitation and temperature as drivers [*Weiß*, 2011; *Wetterhall et al.*, 2011; *Singh et al.*, 2014; *Turner et al.*, 2014; *Whateley et al.*, 2014; *Bussi et al.*, 2016; *Culley et al.*, 2016]. However, given the complexity of water resource systems, not considering a wide range of attributes means it is likely that several key system climate drivers will therefore not be included in the analysis. This means that while a scenario-neutral space can be created and different management strategies can be compared, the system may not be tested against the changes in climate that could most affect system performance [*Brown and Wilby*, 2012; *Nazemi and Wheater*, 2014]. Consequently, there is a need for a formal method to identify the smallest number of climate attributes that have a significant impact on system performance, for the inclusions of climate time series in scenario-neutral

assessments.

A third challenge is that recently, applications of scenario-neutral climate impact assessments are comparing the performance indicated by a scenario-neutral space to the performance in response to climate projections, and finding that there is disagreement between the two approaches [*Taner et al.*, 2017; *Keller et al.*, 2018]. While there can be valid reasons for not relying on climate projection information for decision making, disagreement between the two approaches arises from a failure of scenario-neutral approaches to capture the key modes of change affecting system performance. To some extent, this is likely a result of the above mentioned limitations: the scenario-neutral climate impact assessments are not correctly perturbing the critical climate attributes for a system. As a result, identifying the cause of the difference between the scenario-neutral space and climate projections could lead to uncovering system vulnerabilities, and better deliver on the underpinning principle of the scenario-neutral approach to climate impact assessments. Consequently, there is a need to understand the source of any discrepancies between the two analyses, and be able to identify and avoid them.

## 1.3 Overall Research Objectives

The ambition of this thesis is to overcome a number of existing limitations with the implementation of scenario-neutral climate impact assessments for complex systems, and reconcile scenario-neutral and scenario-led approaches to provide a unified framework for climate impact assessments. Three specific research objectives have been identified, and the link between each research objective and the scenario-neutral approach as a whole is shown in Figure 1-2.

**Objective 1:** To formalise the inverse approach to stochastic time series generation, thereby improving the efficiency of the approach and ensuring the physical realism of the simulated time series (*Paper 1*):

**Objective 2:** To develop and evaluate an approach for identifying the smallest number of climate attributes that have a significant impact on system performance, for use in scenario-neutral impact assessments that require

hydrometeorological time series (*Paper 2*).

**Objective 3:** To present common pitfalls with the scenario-neutral approach that affect the validity of the results and demonstrate the effect of falling for each of these pitfalls (*Paper 3*).

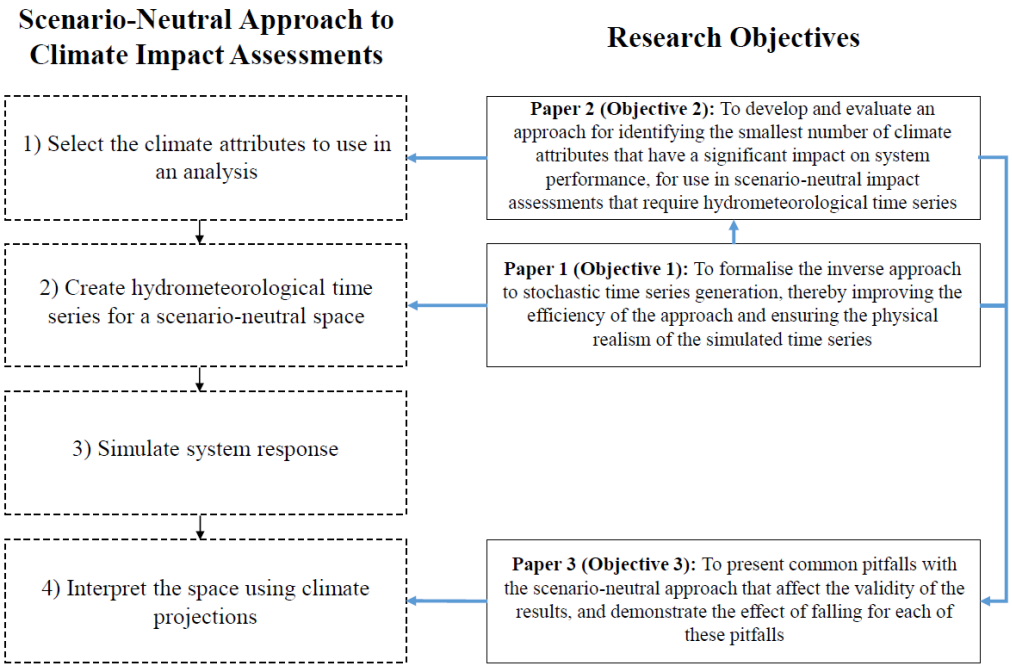


**Figure 1-2** The link between each research objective and the scenario-neutral approach to climate impact assessments

## 1.4 Thesis Organisation

This thesis contains five chapters with the main body of research presented in **Chapter 2** to **Chapter 4**. These chapters correspond to three journal papers. Chapter 2 (Paper 1) has been submitted to the Journal of Hydrology, Chapter 3 (Paper 2) has been submitted to Water Resources Research, and Chapter 4 (Paper 3) will also be submitted to Water Resources Research. The section, figure and table numbers have been modified in line with University guidelines but the manuscript material is otherwise unchanged.

**Chapter 2** presents an improvement to the inverse approach to stochastic time series creation that allows for more complex applications (Objective 1). This improvement, and an increase to the overall efficiency of the approach, are

demonstrated using climate data from Adelaide, Australia. The improved formulation is necessary for more complex applications of a scenario-neutral climate impact assessment and is therefore also used in Chapters 3 and 4.

**Chapter 3** presents an approach to identifying the critical climate attributes for use in scenario-neutral climate impact assessments (Objective 2). This approach is demonstrated on the Lake Como case study, and enables a scenario-neutral approach to uncover the key modes of failure for a system. This is used in Chapter 4, where a successfully validated implementation of the scenario-neutral approach is presented.

In **Chapter 4**, four key pitfalls for applications of the scenario-neutral approach are presented, and their effects are demonstrated (Objective 3). Some of these pitfalls address methodologies presented in Chapters 2 and 3, and others address wider limitations of the scenario-neutral approach identified in the literature.

Conclusions are provided in **Chapter 5**, which summarise: i) the research contributions, and ii) limitations and future research.

# Chapter 2

## *Generating realistic perturbed hydrometeorological time series to inform scenario-neutral climate impact assessments (Paper 1)*

Sam Culley, Bree Bennett, Seth Westra and Holger Maier

Journal of Hydrology, volume 576, pages 111-122

# Statement of Authorship

| | |
|---|---|
| Title of Paper | Generating realistic perturbed hydrometeorological time series to inform scenario-neutral climate impact assessments |
| Publication Status | ☐ Published ☐ Accepted for Publication<br>☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style |
| Publication Details | Submitted to the Journal of Hydrology, under review. |

## Principal Author

| | |
|---|---|
| Name of Principal Author (Candidate) | Sam Culley |
| Contribution to the Paper | Contributed to the conception and design of the project, performed analysis, interpreted the results, wrote manuscript and acted as corresponding author |
| Overall percentage (%) | 75% |
| 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 18/04/19 |

## 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 to 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 | Bree Bennett |
| Contribution to the Paper | Contributed to the conception and design of the project, analysis and interpretation of the research data, and editing the manuscript |
| Signature | Date 18/4/19 |

| | |
|---|---|
| Name of Co-Author | Seth Westra |
| Contribution to the Paper | Contributed to the conception and design of the project, analysis and interpretation of the research data, and editing the manuscript |
| Signature | Date 18-04-2019 |

| | | | |
|---|---|---|---|
| Name of Co-Author | Holger Maier | | |
| Contribution to the Paper | Contributed to the conception and design of the project; analysis and interpretation of the research data, and editing the manuscript | | |
| Signature | | Date | 18/4/19 |

# Abstract

Scenario-neutral approaches are used increasingly as a means of stress-testing climate-sensitive systems to a range of plausible future climate conditions. To ensure that these stress-tests are able to explore system vulnerability, it is necessary to generate hydrometeorological time series that represent all aspects of plausible future change (e.g. averages, seasonality, extremes). A promising approach to generating these time series is by inverting the stochastic weather generation problem to obtain weather time series that capture all the relevant statistical features of plausible future change. The objective of this paper is to formalize this "inverse" approach to weather generation, by both characterizing the process of optimizing weather generator parameters and proposing a numerically efficient solution that exploits prior knowledge and accounts for the complexity of the optimization landscape. The proposed approach also provides a structured way to ensure the physical realism of the generated weather time series, by using penalty-based objective functions to focus the optimization on the climate features deemed most relevant to the system being analyzed. A case study in Adelaide, Australia, is used to demonstrate specific implementations of this approach. The use of bounds on the weather generators dramatically decreases the time taken to create time series, and the use of penalties is shown to allow for change in some statistics to be prioritized, while still ensuring the realism of the time series.

## 2.1 Introduction

Scenario-neutral climate impact assessments are proving to be an effective way of assessing how a range of climate-sensitive systems might respond to plausible future climate changes. The scenario-neutral approach has been applied recently to flood protection, water supply and ecological systems [*Prudhomme et al.*, 2013; *Culley et al.*, 2016; *Poff et al.*, 2016], with these studies demonstrating that scenario-neutral approaches both lead to important insights into overall system sensitivities and vulnerabilities, and enable the identification of possible failure modes by determining how a system responds to step changes in climate [*Prudhomme et al.*, 2010]. These approaches are also increasingly being used to provide decision-theoretic information, describing conditions whereby one system configuration or design option is preferred to another [*Brown et al.*, 2012], or approximating the maximum operational adaptive capacity of the system [*Culley et al.*, 2016].

Although most scenario-neutral approaches have focused on changes in the mean state of climate variables [*Prudhomme et al.*, 2010; *Prudhomme et al.*, 2013; *Culley et al.*, 2016; *Wilcke and Bärring*, 2016; *Spence and Brown*, 2018], it is becoming increasingly apparent that critical system vulnerabilities may reside in other aspects of change—including variability, intermittency, extremes, seasonality and/or inter-annual persistence [*Steinschneider and Brown*, 2013; *Bussi et al.*, 2016; *Guo et al.*, 2017]. An important implication is that if key sensitives are not identified, major modes of system vulnerability may not be uncovered, thereby negating the stated benefit of scenario-neutral studies. This poses a deep challenge to the viability of scenario-neutral approaches: how should weather and hydrometeorological time series be generated to capture all possible aspects of future change?

The primary approach currently available to address this challenge within the scenario-neutral framework is through the use of stochastic weather generators, which contain sufficient flexibility to simulate a wide variety of possible future changes while maintaining the statistical features commonly associated with weather time series [*Steinschneider and Brown*, 2013; *Guo et al.*, 2018]. The

forward scaling approach presented by *Steinschneider and Brown* [2013] is capable of manipulating more complex measurements of precipitation, like persistence, by directly perturbing the parameters of a weather generator to generate baseline time series with different statistics. However, to produce the uniform perturbations to climate attributes required in a scenario-neutral assessment, this still requires some post-processing of time series, such as quantile mapping, to represent targeted changes in climate. Given scaling methods are still used, the range of attributes that can be perturbed is often limited to means and seasonality, by directly adjusting the baseline time series. To avoid the scaling process, *Guo et al.* [2018] provided the first structured attempt at inverting the stochastic generation problem, by varying the parameters of a stochastic generator through an optimization loop to simulate weather time series with pre-specified statistics or "attributes". The generation of time series using this approach was demonstrated for three climate variables (precipitation, temperature and evapotranspiration) [*Guo et al.*, 2017] and was benchmarked on a rainfall dataset from a catchment in South Australia [*Guo et al.*, 2018].

As scenario-neutral approaches are applied to increasingly complex systems, it becomes necessary to explore increasing numbers of hydrometeorological variables and statistics of those variables. For example, whereas *Culley et al.* [2016] focused on annual average rainfall and temperature in their case study on Lake Como flood management and irrigation requirements, it is likely that a thorough exploration of system vulnerability for this alpine lake would require exploration of attributes that affect features such as snow pack and snow melt rates, evaporation from the reservoir and evapotranspiration from the irrigation demand regions. These attributes could include winter precipitation amounts, the number of frost days in the year, growing season length, and so on. Stress testing the system to each individual change—and all the possible combinations of those changes—poses substantial numerical and computational barriers to the inversion problem. For example, the required runtimes indicated by *Guo et al.* [2018] for a simple application of three attributes (e.g. 8 hours for producing 100 simulated weather time series using 8 cores) suggest significant potential challenges for the widespread application of the inverse generation method, and

requires a structured approach for identifying opportunities for computational efficiencies. Consequently, there is a need to reduce the run times of the optimization loop that underpins the inversion process so that it can be applied to more complex systems within practical timeframes.

A further challenge is that, as the number of attributes to be perturbed increases, the likelihood of attempting to simulate infeasible changes will also increase. For example, consider the relationship between the attributes average annual rainfall, average rainfall intensity and average number of wet days. Given any two values of those attributes are held constant, there is only one value the third can take, and it is not numerically possible to simulate time series with further increases or decreases to that third attribute. This is particularly important when seeking to generate weather time series that capture specific changes, while seeking to match historical climate patterns in all other aspects to maintain physical realism. Consequently, there is a need to manage which attributes of a time series simulated as part of the inversion process achieve the requested change, in the event the requested change is infeasible.

The overarching objective of this paper is therefore to formalize the inversion problem, by focusing on two specific aims: (1) improving the computational efficiency of the optimization process; and (2) ensuring the physical realism of the simulated time series. The remainder of the paper proceeds as follows. Section 2.2 formally articulates the aims and details general approaches to meet them. Section 2.3 describes the case study of Adelaide, Australia, followed by Sections 2.4 and 2.5 that focus on the first and second aims, respectively, where the general approach is implemented for the case study, and is then tested to examine the impact of formalizing the inverse approach as presented in this paper. Conclusions are presented in Section 2.6.

## 2.2 Formalizing the Inverse Approach to Stochastic Generation

### 2.2.1 Overview of the Inverse Approach

*Guo et al.* [2018] presented the inverse approach as a technique to generate hydrometeorological time series that satisfy a set of target changes in specified climate attributes. In this context, 'attributes' are defined as statistics of particular hydrometeorological variables, such as the mean annual rainfall or number of wet days. The approach starts by setting targets $\mathbf{t}_j \in \mathbb{R}^n$, where *n* is the number of attributes considered, and $j = 1,...,m$ represents the number of target values of those attributes. The target changes may be represented as absolute values (e.g. simulating a time series with annual average rainfall of 960 mm), or alternatively they may be represented in terms of the percentage or absolute changes in attributes relative to historical climate (e.g. a 10% decrease in annual average rainfall, or 3°C increase in average annual temperature). The weather time series can be generated by changing only a single attribute at a time, or by simulating combinations of changes; for example *Guo et al.* [2018] simulated changes in two attributes over a regular grid.

Once the attribute targets are identified, the next step is to apply a formal optimization approach that involves modifying the parameter vector $\theta$ of some stochastic generator $g(\theta)$ that minimizes a measure between the relevant attributes $\mathbf{a}_j \in \mathbb{R}^n$ of the simulated weather time series and the target attributes ( $\mathbf{t}_j$ ). This is illustrated in Figure 2-1, whereby the target attributes are represented here in two dimensions ( $n = 2$ ) over a regular 7x7 grid (i.e. $j = 1,...,49$ ), in terms of a fraction or percentage change relative to a historical baseline. For each target, the inverse approach then adjusts parameter vector $\theta$ in order to achieve weather time series with desired attributes. This process is also represented mathematically as:

$$weather.ts_j = g\left(\arg\min_{\theta} O(\mathbf{a}_j, \mathbf{t}_j)\right) \quad \textbf{(1)}$$

where $O(\mathbf{a}_j, \mathbf{t}_j)$ describes a general measure of the difference between each

weather attribute and its target. For example, in the case of *Guo et al.* [2018], a simple Euclidean distance measure was used:

$$O(\mathbf{a}_j, \mathbf{t}_j) = \sqrt{\sum_{i=1}^{n} (a_{i,j} - t_{i,j})^2} \quad \textbf{(2)}$$

It is noted that the process of achieving the weather time series as described in Equation (1) is inherently iterative; namely the attributes $\mathbf{a}$ are calculated from the weather time series in the previous optimization step, until a stopping criterion is reached.

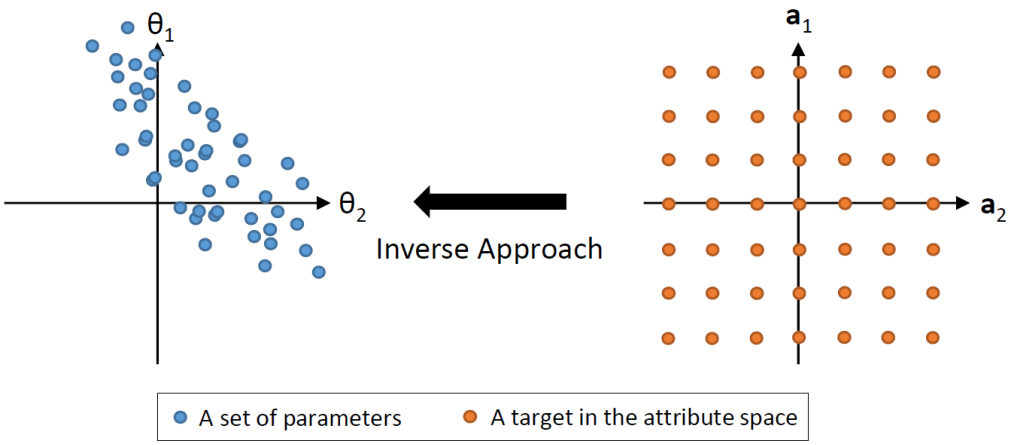


**Figure 2-1** Conceptual illustration of the inverse approach to stochastic generation. Each blue dot corresponds to a set of weather generator parameters that result in the generation of weather time series that have the set of climate attributes of one of the orange dots. Equal coverage of the climate attribute space ($a_1$, $a_2$) does not necessarily reflect equal coverage in the parameter space ($\theta_1$, $\theta_2$) of stochastic weather generators.

Although conceptually straightforward, there are two key challenges with the approach:

1. How to design an efficient optimization process that extends the inverse approach to high-dimensional spaces with high levels of accuracy and minimal runtimes (Section 2.2.2). In particular, computational issues were identified by *Guo et al.* [2018] as a significant challenge, and in its current form is likely to inhibit wider application of the inverse methodology.

2. How to ensure the realism of the weather time series (Section 2.2.3). A feature of the inverse approach is that any desired properties of the climate time series need to be included in the objective function. This provides an incentive to include a greater number of attributes in the objective function to maintain realism, increasing problem complexity, and the likelihood that an infeasible combination of attributes will be requested. A traditional Euclidean distance objective function does not provide a sufficiently robust approach for prioritizing some attributes above others, which is necessary when not all target changes can be met.

The following sections explore these two challenges in more detail.

### 2.2.2 Improving the computational efficiency of the optimization process

The generic steps in the optimization loop that underpins the inverse approach are shown in Figure 2-2, which consists of an iterative process for updating the parameters of a weather generator, $\theta_k$, until certain stopping criteria have been met. The approach to updating depends on the specific choice of optimization method (e.g. gradient descent versus stochastic searches), but all methods aim to improve the objective function value $O(\mathbf{a}_j, \mathbf{t}_j)$, which for this case consists of a measure of distance between the generated weather time series attributes and the target attributes. Potential stopping criteria for the optimization loop include the completion of a fixed number of iterations, stagnation in the optimization process or sufficiently small errors between the attribute values of the time series generated from the weather generator and the target attributes [*Zielinski et al.*, 2005].

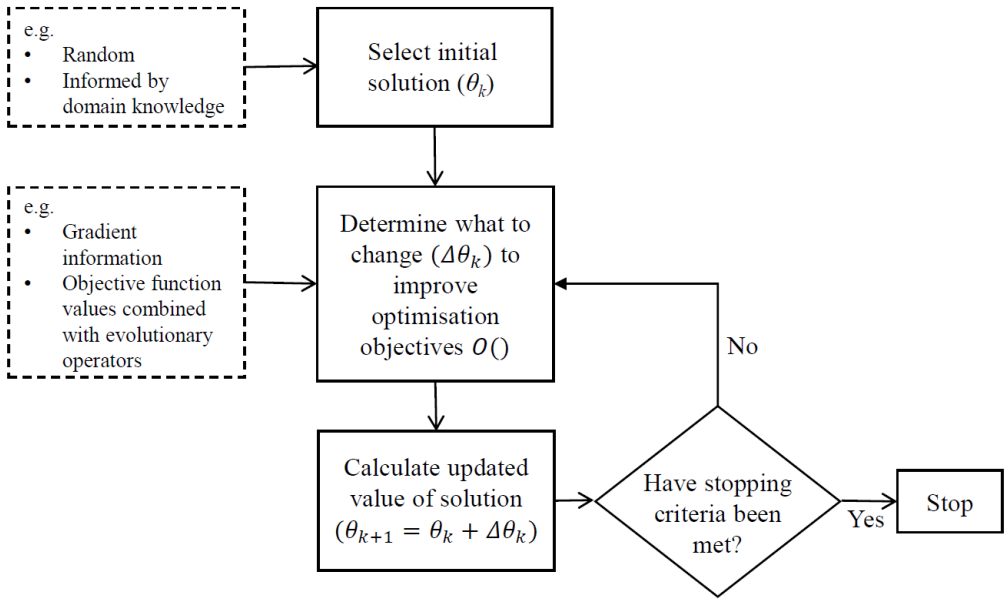


**Figure 2-2** Generic steps in the optimization loop that underpins the inverse approach.

The following two approaches can be used to improve the computational efficiency of the above processes: (i) selecting the optimization algorithm that is most suited to the characteristics of the optimization problem, and (ii) reducing the size of the search space as much as possible, without restricting the ability to identify the desired solutions.

For the first approach, it is necessary to diagnose the nature of the optimization fitness landscape—the relationship between the decision variables and the objective function—as this is critical for identifying the most efficient optimization algorithm for the class of problem to be tackled [*Maier et al.*, 2019]. For example, smooth fitness landscapes may enable computationally efficient hill climbing algorithms to find the global optimum (e.g. [*Nesterov*, 2007], whereas irregular fitness landscapes require stochastic methods [*Kingston et al.*, 2008]. In low dimensional problems, an enumeration methodology can be used to visualize the fitness landscape and examine its properties directly, whereas for higher-dimensional problems, the use of fitness landscape statistics that identify properties like the overall structure of the fitness landscape, any flat areas of the same function value, and the distance between good local optima and the global solution might be required [*Gibbs et al.*, 2011; *Malan and Engelbrecht*, 2013; *Maier et al.*, 2014; *Gibbs et al.*, 2015]. Given that the fitness landscape is defined

by the weather generators and attributes of interest, the most suitable optimization algorithm is likely to depend on the specific implementation of the inverse approach.

A challenge for the second approach is that, for all but the simplest problems, it is generally not possible to know *a priori* how the stochastic generation parameter vector $\theta$ maps into the attribute space (Figure 2-1). This makes it difficult to provide bounds on the weather generator parameters, which are necessary for some optimization algorithms, such that the bounds do not unintentionally prevent some requested time series from being generated. In order to address this issue, *Guo et al.* [2018] used very wide bounds on the weather generator parameters during the optimization process. However, this approach produces very large search spaces, which can result in significant increases in the computational effort associated with identifying the desired parameter values. An alternative approach used in this study is to refine the bounds of the weather generator parameters by assessing typical ranges of stochastic weather generator parameters applied to a broad set of current weather time series across a large geographic area, under the assumption that there are likely to be current weather "analogues" (e.g. weather time series in very warm and arid regions) that are representative of plausible future changes as a result of anthropogenic climate change.

### 2.2.3 Ensuring the physical realism of the simulated time series

The weather time series to be generated using the inverse approach are synthetic series and are thus not constrained by physical processes in the same manner as time series generated by weather and climate models. For example, if a target is to increase total annual rainfall by 15%, then it would be theoretically possible for the weather generator to produce time series whereby all the annual rainfall occurs within one season, or occurs uniformly across a whole year, or any other possible series that meets the total annual rainfall target.

The proposed conceptual approach for addressing this issue and ensuring physical realism is to generate time series that represent the proposed target

changes, but with all other aspects of the weather time series held at historical values. This is achieved by including a larger number of target attributes within the optimization process, by focusing on both "perturbed" attributes that represent the primary objective of the optimization, and "held" attributes that keep all other aspects of the weather time series as close to their historical values as possible. This substantially increases the complexity of the optimization problem, by increasing the number of attributes *n* that need to be considered as part of Equation 1, further highlighting the importance of reducing the size of the search space through other means, as discussed in Section 2.2.2.

Beyond the computational challenges associated with this increase in optimization complexity, there is a more fundamental problem: in many cases, setting a large number of both "perturbed" and "held" attributes will lead to requests for infeasible attribute combinations. Returning to the example of increasing annual average rainfall by 15%, we might seek to achieve this while holding the number of wet days, the amount per wet day and any other aspects of the annual rainfall time series at their historical values. However, this combination is not possible: increasing annual rainfall *can only be achieved* through either increasing the number of wet days, or the amount of rainfall per wet day, or some combination of the two.

The nature of the over-constrained optimization problem is illustrated in Figure 2-3, where we plot the feasible subspace of total annual rainfall, number of wet days and amounts per wet day in the three-dimensional space of possible attribute changes. This subspace is calculated assuming the total annual rainfall is the product of the number of wet days and the average wet day amount ($a_{(\text{ann. total rainfall})}=a_{(\text{no. wet days})}*a_{(\text{wet day amounts})}$). As an example, we seek to increase total annual rainfall (the "perturbed" attribute) by 15% from its historical value (leading from point 1, which represents the historical conditions, to point 2), which if the two other attributes are "held" at their historical values, is an infeasible target (i.e. point 2 does not lie within the feasible subspace).

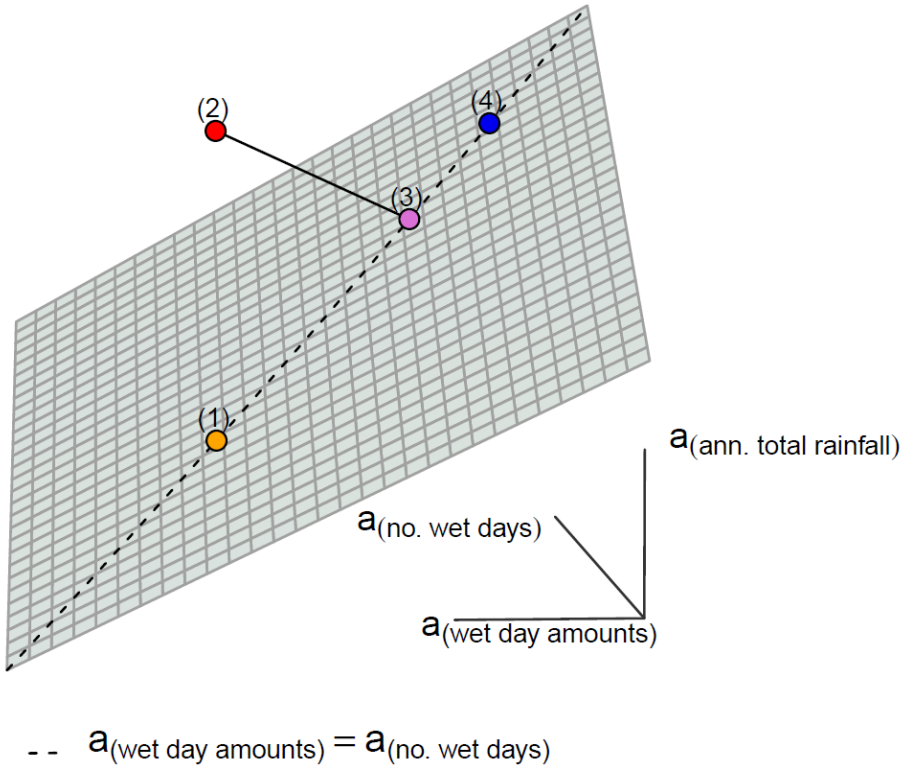


**Figure 2-3** The over-constrained optimization challenge: the grey surface indicates the feasible subspace, with equation $a_{(\text{ann. total rainfall})}=a_{(\text{no. wet days})}*a_{(\text{wet day amounts})}$. Historical conditions are shown at point (1) and the target is indicated by point (2). Point (3) indicates the shortest Euclidean distance between the target at point (2), and the feasible subspace. Point (4) indicates the shortest distance between the target and the feasible subspace, while ensuring zero error in the perturbed attribute.

If a Euclidean distance objective function (Eq. 2) is used in the optimization process (see Figure 2-2) the solution indicated by point (3) will be identified. Point (3) is the solution closest to the target that lies on the feasible subspace (i.e. point (3) is the orthogonal projection of point (2) onto the feasible subspace). However, this solution only increases the total annual rainfall by ~10%, and thus does not produce rainfall time series with the desired 15% change. It also leads to a 5% increase in the number of wet days and amounts per wet day. An alternative solution is to modify the objective function to place more emphasis on the "perturbed" attribute, and in so doing find a solution that produces the desired change in the "perturbed" attribute but keeps the remaining "held" attributes as close as possible to historical values. In this simple illustration, this might lead to total annual rainfall increasing by 15%, as originally sought, by

both increasing the number of wet days and amounts per wet day each by 7% (see point (4)).

As the objective function is used to measure how close the attributes of the simulated time series are to the intended targets, this also needs to manage any trade-offs between attributes. This suggests the use of penalties [*Coello Coello*, 2002] to modify the objective function to favor solutions with smaller errors in "perturbed" attributes. Here, the modification of the objective function (Eq. 2) is discussed with reference to two general penalty structures: a linear penalty structure that adds a linear term based on the error in the "perturbed" attributes (Eq. 3) and a quadratic structure that adds a squared term based on the error in the "perturbed" attributes (Eq. 4) (note that this equation can be rearranged as a weighted sum). The two modified objective functions are given by

$$O(\mathbf{a}_j, \mathbf{t}_j) = \sqrt{\sum_{i=1}^{n} \left(t_{i,j} - a_{i,j}\right)^2} + \sum_{k=1}^{p} \lambda_k * abs(t_{k,j} - a_{k,j}) \quad \textbf{(3)}$$

$$O(\mathbf{a}_j, \mathbf{t}_j) = \sqrt{\sum_{i=1}^{n} \left(t_{i,j} - a_{i,j}\right)^2 + \sum_{k=1}^{p} (\lambda_k - 1) * (t_{k,j} - a_{k,j})^2} \quad \textbf{(4)}$$

where $k = 1, ..., p$ represents the subset of $n$ "perturbed" attributes (i.e. $p \leq n$), $\lambda$ are the scaling parameters applied to the errors in the "perturbed" attributes. The remaining notation is consistent with Equation 2. Equations 3 and 4 reduce to the "unweighted" Euclidean distance objective function (Eq. 2) for $\lambda_k$=0 and 1, respectively.

The effect of these penalties and scaling parameter values can be illustrated with a continuation of the example in Figure 2-3, where we seek to increase total annual rainfall by 15% and hold the number of wet days and amounts per wet day at historical values. Here, instead of viewing the three-dimensional space of possible attribute changes, a two-dimensional slice through the space is shown in Figure 2-4, such that each panel displays the feasible solution subspace as a black line (this slice was represented by the dashed line in Figure 2-3). The over-constrained target is represented by a red point and the contours represent the

objective function value for each attribute combination in the 2D space (i.e. the fitness landscape). The minimum error solution, represented by the blue point, occurs where the smallest objective function value (i.e. 2D fitness landscape) intersects the feasible subspace (black line).

To demonstrate how the fitness landscape and minimum error solution change with different penalty structures and scaling parameters, $\lambda$, Figure 2-4 compares different $\lambda$ values for both a linear penalty structure (top panels) and a quadratic penalty structure (bottom panels). Moving left to right, the panels in Figure 2-4 illustrate the effect of increasing the scaling parameter, $\lambda$, on the "perturbed" attribute (annual total rainfall) in terms of the change in the fitness landscape and thus the location of the minimum error solution (blue point) for the two penalty structures. For the case where all attributes are equally weighted within the objective function (left most panels of Figure 2-4), the fitness landscape contours are circular and the identified minimum error solution increases the perturbed attribute by approximately 10%. As the scaling parameter, $\lambda$, is increased, the minimum error solution is moved along the feasible subspace line towards the solution with zero error in the "perturbed" attribute.

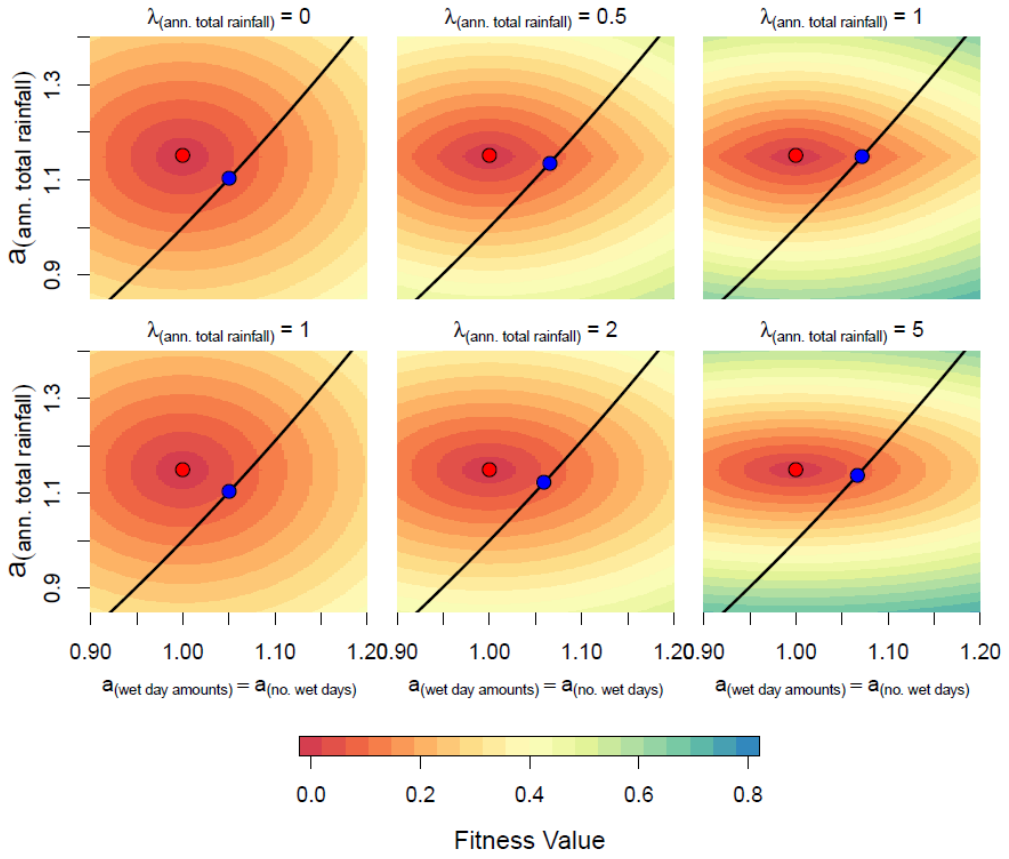


**Figure 2-4** How penalty functions can change the fitness landscape to create a new minimum error solution during over-constrained optimization: (top panels) linear penalty term, and (bottom panels) quadratic penalty term.

The rate at which the minimum error solution moves along the feasible line with change in a scaling parameter, λ, is dependent on the penalty structure. The linear penalty term (top panels) has the capacity to identify a solution with zero error in the "perturbed" attribute if the scaling parameter is sufficiently large. In contrast, the quadratic penalty term (bottom panels) exhibits asymptotic behaviour such that as the scaling parameter increases, the minimum error solution will get closer to the zero error solution in the "perturbed" attribute but will never intersect it. The choice of penalty is influenced by the problem application. For example, where it is important to meet the "perturbed" attribute target, the linear penalty term may be appropriate. However, if the "held" attributes also have a substantial impact on system performance, the quadratic penalty may be more appropriate.

It is noted that the illustrative example described in Figure 2-3 and Figure 2-4 is highly conceptual, and most widely used weather generators have much greater complexity to enable them to simulate the statistical features of realistic weather time series. The capacity to achieve specified target attribute combinations will be limited both by physical constraints (as illustrated in Figure 2-3 and Figure 2-4) and the ability of the weather generator to simulate the requisite combinations. For example, an annual Markov model would not be capable of simulating seasonal variability in various rainfall statistics, thereby leading to infeasible targets if the objective is to simulate seasonal variability. Conversely, overly complex weather generators would lead to a much higher-dimensional parameter set, $\theta$, as well as the need to constrain a larger number of attributes to ensure physical realism of the generated series, placing more burden on the optimization process. Care is therefore needed to ensure that a weather generator of appropriate complexity is selected to achieve the objectives of each investigation.

## 2.3 Case study

The issues and proposed solutions highlighted in the previous section are illustrated using rainfall data from a location in Adelaide, Australia. The region has a Mediterranean climate with an annual average rainfall of 532 mm. The rainfall for this region is highly seasonal with most rainfall occurring during winter (June, July and August) and spring (September, October and November) and the least rainfall occurring in the summer season (December, January and February). Historical rainfall time series for Adelaide (34.92°S, 138.62°E) were obtained from the Australian Water Availability Project (AWAP) dataset [*Raupach et al.*, 2012]. To minimise the influence of changing trends in rainfall, the period 1970 to 1999 was selected, since this period is relatively stationary.

Two stochastic daily weather generators are used in this study that follow the precipitation component of WGEN: (i) a simple four parameter model to aid in theoretical understanding of the optimization fitness landscape, and (ii) a more complex model to investigate how well the proposed developments work for more realistic applications [*Richardson and Wright,* 1984]. The simple weather

generator model used has only four parameters. Two parameters control the wet/dry sequence throughout the time series using a 1st order Markov chain. *Pdd* is the probability of a dry day given a dry day occurred previously, and *Pwd* is the probability of a dry day given a wet day occurred previously. For any two values of these parameters, the supplementary parameters *Pdw* and *Pww* are calculated, and sequences of wet and dry days for the length of the time series are obtained using a random number generator. The remaining two parameters control the amount of rainfall that occurs on wet days. These are the shape and rate parameters of the gamma distribution, $\alpha$ and $\beta$, from which each wet day rainfall amount is randomly sampled. Given the parameters do not vary throughout the year, this weather generator can only produce stochastic rainfall time series to meet a range of climate attributes measured at an annual level. It is therefore referred to throughout as the "annual" weather generator.

To perturb intra-annual attributes, a more complex weather generator is needed, with additional parameters and hence greater degrees of freedom to produce the required time series. The method used in this study is to extend the parameters of the simple model, where each of the original four parameters is specified as varying throughout the year. A harmonic model is used to control this variation, dictated by the mean, amplitude and phase angle of a harmonic (e.g. *Pdd* becomes *Pdd-m*, *Pdd-amp* and *Pdd-phase*) [*Richardson*, 1981]. This is the same approach *Richardson* [1981] used to create temperature and solar radiation time series; however, in this application the harmonic model is not creating the time series directly, but describing what values the parameters should take. The harmonic models are fixed to have 12 periods, allowing for each of the four annual WGEN parameters to take different monthly values throughout each year. This allows the perturbation of attributes at the seasonal level, and this model is referred to throughout as the "seasonal" weather generator.

Only eight parameters are used as decision variables for this seasonal weather generator, as the phase angle parameters are fixed at historical values, leaving just the mean and amplitude for *Pdd*, *Pwd*, $\alpha$ and $\beta$. The calibration process outlined by *Richardson* [1981] was used to determine the values of the four

phase angle parameters for the case study site of Adelaide. The *Pdd*, *Pwd*, $\alpha$ and $\beta$ phase angles were 0.355, 0.232, 3.53 and 2.46, respectively. This modification to the seasonal weather generator maintains the seasonal pattern in the generated time series such that most rainfall occurs in winter and spring, but still allows for the actual rainfall volume in each season to be perturbed separately. The trade-off with this new model is the large increase in computational effort required to find a solution given the increased search space.

The attribute sets used in the implementation of the inverse approach are listed in Table 2-1 for each weather generator type. Throughout this paper, *Ptot* is selected as the "perturbed" attribute, given its common usage in scenario-neutral impact studies, except in two instances designed to investigate applications for multiple "perturbed" attributes, where *nWet* is also selected. The remaining "held" attributes are included in the objective function for each simulation to ensure that these properties are maintained in the perturbed time series (as discussed in Section 2.2.3). Given the change in model complexity, different sets of attributes are specified for each weather generator. The seasonal model requires more "held" attributes, given the extra degrees of freedom provided by the parameters.

**Table 2-1** List of "perturbed" (P) and "held" (H) attributes for annual and seasonal weather generator experiments.

| Attribute | Description | Annual | Seasonal | Held Value |
|---|---|---|---|---|
| Ptot | total annual rainfall volume | P/H* | P/H* | 532.3 mm |
| nWet | annual number of wet days | P/H* | P/H* | 212.1 days |
| P99 | 99th percentile daily rainfall amount | H | H | 16.86 mm |
| P90 | 90th percentile daily rainfall amount | H | H | 4.636 mm |
| DSD | dry spell duration in days | H | H | 3.455 days |
| DJFtot | total rainfall volume in summer (DJF) | | H | 59.18 mm |
| WSR | ratio of total winter (JJA) to summer (DJF) rainfall | | H | 2.112 |

*This attribute is 'perturbed' or 'held', depending on the specific experiment.

## 2.4 Investigation into the impact of increasing the efficiency of optimization

The following section contains a specific implementation of the approach for increasing the efficiency of the optimization process proposed in Section 2.2.2. The optimization problem is analyzed for the case study, leading to the selection of an optimization algorithm that is suited to the fitness landscape (Section 2.4.1). The optimization process is then implemented on the case study with improvements to optimization efficiency due to restricting the bounds of the weather generator parameters (Section 2.4.2).

### 2.4.1 Selecting a suitable optimization algorithm

In order to determine the most appropriate optimization algorithm for the case study application, the nature of the fitness landscape is analyzed. This is done for the annual weather generator, as this enables an enumeration method to be used to generate the landscape. As a fitness landscape for a model with four parameters is five-dimensional, with the fifth dimension being the objective function value (i.e. the "fitness"), the complete fitness landscape is unable to be visualized. Consequently, in order to enable key aspects of the overall fitness landscape to be inspected, separate three-dimensional fitness landscapes are generated for the shape and rate parameters ($\alpha$ and $\beta$, respectively) and the probability of wet-dry and dry-dry parameters. The required fitness values are calculated using the "unweighted" objective function (Eq. 2) for three attributes all held at historical levels: *Ptot*, *nWet* and *P99*. As part of this process, the random seed used in the weather generator is held constant to maintain a set relationship between parameters and attribute values [*Guo et al.*, 2018].

The resulting fitness landscapes are shown in Figure 2-5. As can be seen, the fitness landscape in the left panel is smooth, as the gamma function from which rainfall amounts are sampled is continuous, so that changes to the gamma distribution parameters result directly in changes in rainfall volume (as a fixed random seed was used to eliminate stochastic "noise", as mentioned above). In contrast, the fitness landscape on the right is rough with many local optima. This is due to discrete changes in the response surface as a result of changing the

number of wet and dry days. Based on this finding, it is likely that irregular response surfaces will be a feature of Markov-based weather generators, including the higher-dimensional seasonal model also used in this paper (suggesting that there is no need to perform fitness landscape analysis for the more complex model). The diagnosis therefore suggests that stochastic search algorithms should be used for implementing the inverse approach, as hill-climbing methods are likely to get stuck in local optima and thus fail to find the best possible solution.

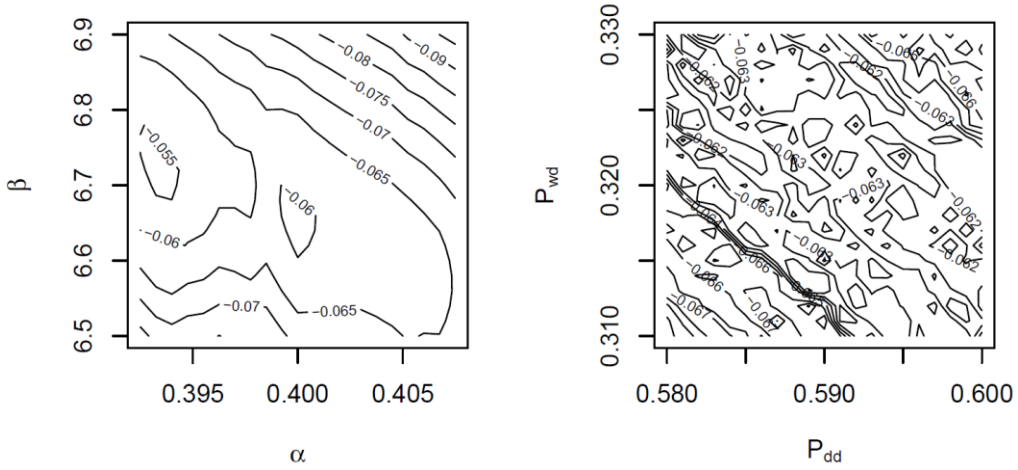


**Figure 2-5** Two-dimensional slices between the objective function value and both the Pwd, Pdd parameters (right) and the $\alpha$, $\beta$ parameters (left) from the four-parameter weather generator. The "unweighted" objective function is calculated for three attributes: total annual rainfall, number of wet days and 99$^{th}$ percentile rainfall.

There is a wide range of stochastic search algorithms that could be used for response surfaces such as that illustrated in Figure 2-5, including shuffled complex evolution, ant colony optimization and genetic algorithms [*Holland*, 1992; *Duan et al.*, 1993; *Dorigo et al.*, 1996]. In this study, a genetic algorithm is used [*Scrucca*, 2013], as this algorithm has been found to be effective in optimizing single objective functions with rough fitness landscapes. The parameter values for each operator of the genetic algorithm used are provided in Table 2-2. A population and number of generations of 200 were chosen to ensure solutions converged, and the remaining parameters were recommended by *Scrucca* [2013].

**Table 2-2** Case study genetic algorithm parameters.

| Operator | Value |
|---|---|
| Population | 200 |
| Mutation probability | 0.1 |
| Crossover probability | 0.8 |
| Crossover points | 1 |
| Generations | 200 |

### 2.4.2 Reducing the optimization search space

In order to determine appropriate bounds for the parameters of the weather generators for the case study location in Adelaide, both weather generators were calibrated to 2870 AWAP grid locations spanning all climatic regions of Australia (at a 50km resolution) using the approach set out in *Richardson* [1981]. Given the significant variability in rainfall time series across continental Australia (which spans tropical, temperate, alpine, Mediterranean, semi-arid and arid climates), this approach is likely to provide a reasonable proxy of the range of variability anticipated for the case study location as a result of future climate change. The actual domain knowledge informed parameter bounds were taken as the $0.3^{th}$ and $99.7^{th}$ percentile of the values of the 2870 rainfall time series and are summarized in Table 2-3 and Table 2-4.

**Table 2-3** Third standard deviation bounds on the four parameters of the annual weather generator for an Australian data set.

| **Parameter** | **Pdd** | **Pwd** | **α** | **β** |
|---|---|---|---|---|
| Domain bounds | 0.427 – 0.998 | 0.088 – 0.824 | 0.313 – 0.998 | 0.043 – 25.46 |
| Uninformed bounds | 0 – 1 | 0 – 1 | 0 – 10000 | 0 – 10000 |

**Table 2-4** Third standard deviation bounds on the eight parameters of the seasonal weather generator for an Australian data set.

| **Parameter** | **Pdd mean** | **Pdd amplitude** | **Pwd mean** | **Pwd amplitude** |
|---|---|---|---|---|
| Domain bounds | 0.38 – 0.99 | 0 – 0.36 | 0.09 – 0.73 | 0-0.32 |
| Uninformed bounds | 0 – 1 | 0 – 1 | 0 – 1 | 0 – 1 |
| **Parameter** | **α mean** | **α amplitude** | **β mean** | **β amplitude** |
| Domain bounds | 0.33 – 0.98 | 0 – 0.25 | 0.08 – 19.7 | 0.03 – 13.6 |
| Uninformed bounds | 0 – 10000 | 0 – 10000 | 0 – 10000 | 0 – 10000 |

To enable the benefits of reducing the size of the search space to be assessed, the inverse approach was used to generate time series with the set of attributes shown in Table 2-1. For this test, all attributes were set at a target of their historical levels (i.e. they were all "held" attributes). Both weather generators were used for time series generation, each with the domain knowledge informed parameters bounds and the wider, uniformed bounds used by *Guo et al.* [2018] (see Table 2-3 and Table 2-4). For the annual weather generator, the use of domain knowledge informed bounds was able to reduce the volume of the search space by seven orders of magnitude, whereas for the seasonal generator, the volume was reduced by fifteen orders of magnitude. All optimization runs were repeated 50 times from different random starting positions in the solution space to minimize the influence of the random search behavior of the genetic algorithm. In contrast, the weather generator seed was held constant for all simulations to ensure consistency in the fitness landscape, as mentioned previously.

Figure 2-6 compares the reduction in objective function value at each optimization generation when domain knowledge informed bounds and uninformed bounds are used to restrict the parameter values for the annual weather generator (left panel) and the seasonal weather generator (right panel). For the annual weather generator, the optimization with informed parameter bounds converges much more quickly and finds better solutions (i.e. three orders of magnitudes smaller) than the optimization with the uniformed bounds for the computational budget of 200 generations. The objective function error for the informed bounds experiment at generation 200 was 0.09, compared to 130 for the uninformed bounds experiment (left panel Figure 2-6). The benefits of using parameter informed bounds is more pronounced for the seasonal weather generator (right panel), such that at generation 200 the objective function errors are approximately five orders of magnitude larger when uninformed parameter bounds are used (56,212 compared to 0.202 for the informed bounds). This highlights the potential benefits of search space size reduction by using domain knowledge informed parameters in terms of increasing the computational efficiency of the inverse approach (and hence increasing the chances of finding better solutions), especially for higher-dimensional search spaces, such as those

associated with more complex weather generators.

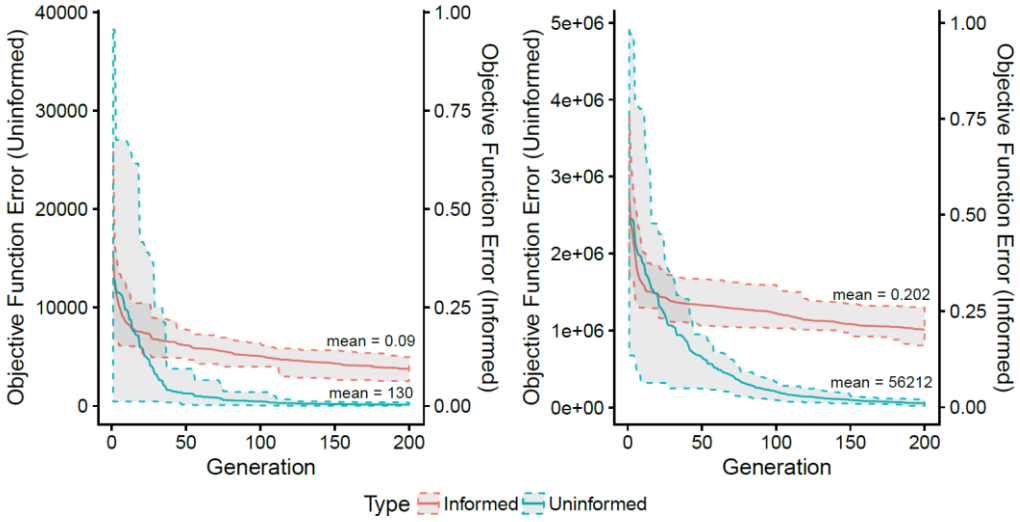


**Figure 2-6** Optimization objective function values at each generation for two sets of parameter bounds: domain knowledge informed and uninformed. A target of historical conditions is searched for with both sets of bounds using an annual weather generator (left panel) and a seasonal weather generator (right panel).

## 2.5 Ensuring the realism of hydrometeorological time series

To ensure realistic time series are generated by the inverse approach, the penalty structures presented in Section 2.2.3 are applied to the Adelaide case study. Section 2.5.1 tests how penalty functions work when creating targeted time series in different regions of a scenario-neutral space. The simple annual weather generator and the linear penalty function (Eq. 3) are used for this demonstration. Section 2.5.2 then compares the effect of the two penalty function structures, using the more complex seasonal weather generator. How well penalty functions can be used to focus on two "perturbed" attributes at once is investigated in Section 2.5.3. These results are specific to the weather generators, attributes and target time series used in this case study. Consequently, the process of examining how the results change with different penalty scaling parameters is something that should be repeated for each implementation of the inverse approach, to ensure the time series are created with the most appropriate trade-offs across the "perturbed" and "held" attributes.

### 2.5.1 Focusing on an attribute with two target perturbations

In order to determine how penalty functions perform in creating time series in different regions of a scenario-neutral space, two target time series are generated using the simple annual weather generator and a linear penalty term (Eq. 3). The "perturbed" attribute in both cases is *Ptot*, with the first target having no change in all selected attributes from historical levels (Table 2-1) and the second target having a 30% decrease in the total annual rainfall volume (*Ptot*) with no change in other attributes.

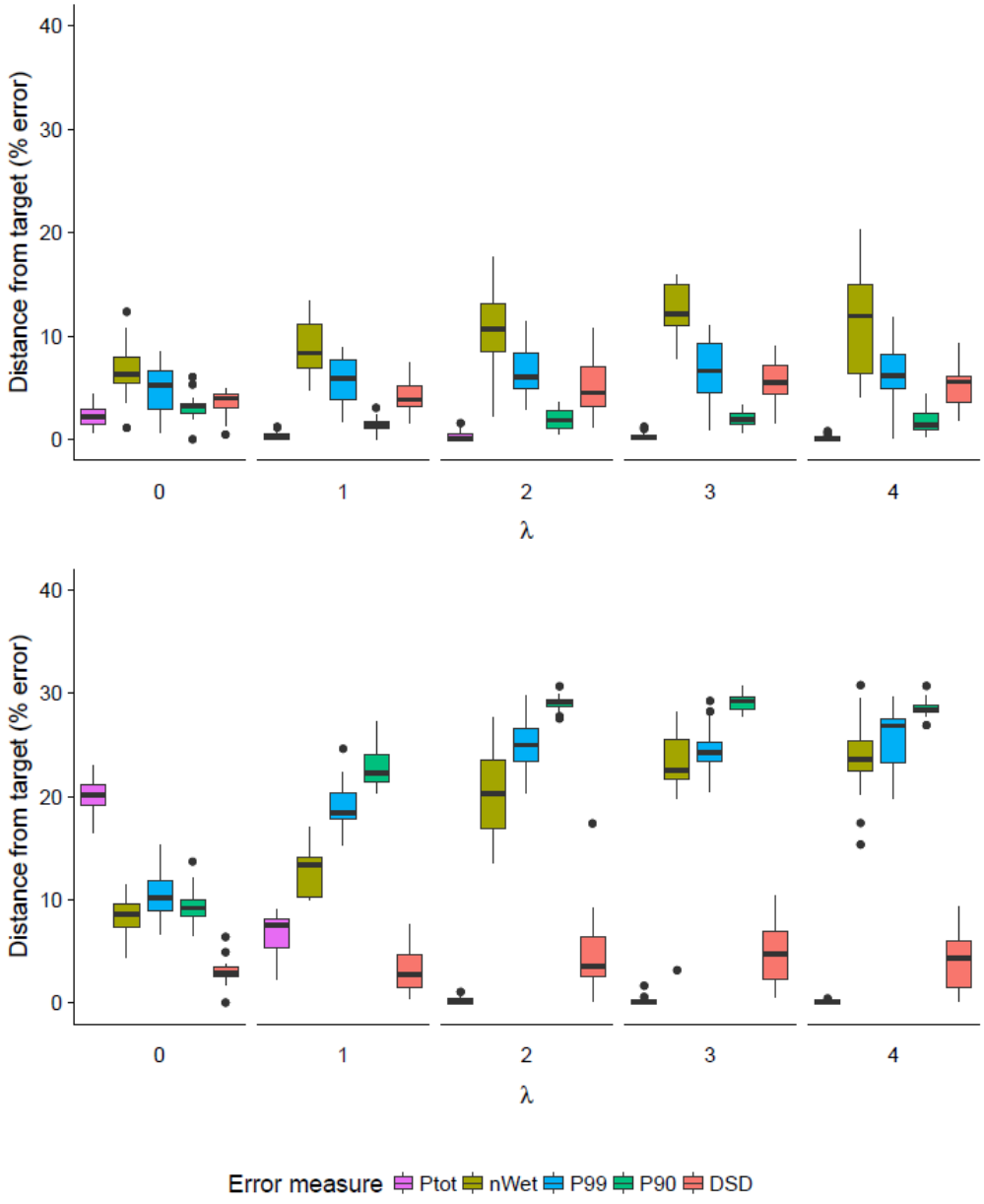


**Figure 2-7** Breakdown of the error in each attribute at the end of optimization for two target time series: zero change from the historical conditions (top) and a 30% decrease in total annual rainfall (bottom). The time series are formed using an annual weather generator and a linear penalty term where the scaling parameter λ is varied.

Figure 2-7 shows the distance from the target of each attribute as a percentage error across 50 optimisation seeds for varying λ values. For both targets the λ=0 cases show the error breakdown across the selected attributes using an "unweighted" objective function (i.e. equivalent to Eq. 2). For the historical

target when λ=0, the error is low and spread relatively evenly across each attribute (top panel). The error in this context arises because of structural deficiencies in the simple annual weather generator relative to the complex historical rainfall time series, so that the weather generator is not able to faithfully simulate all the historical values of the "perturbed" and "held" attributes. For the second target (a 30% decrease in *Ptot*), the error is much more varied across attributes and the *Ptot* attribute has 20% error, whereas the other attributes have less than 10% error (bottom panel). This again results from a lack of flexibility in the annual weather generator—it lacks the degrees of freedom to change *Ptot* alone.

To reduce the error in the attribute *Ptot*, its weight in the objective function needs to be increased. This trade-off in the error between the "perturbed" and the "held" attributes changes with increasing scaling parameter values, λ (Figure 2-7). Once λ=2, the error in *Ptot* is approximately zero; however, the error in three "held" attributes (*nWet*, *P90* and *P99*) has increased. The average dry spell duration (*DSD*) is the only attribute that does not increase its error, as it only depends on the wet/dry first order Markov chain. In contrast, the error in the number of wet days increases, which is likely because this attribute more directly affects the number of high rain days sampled in a year.

The above results demonstrate that the selection of λ can be used to manage the trade-off in error between *Ptot* and the "held" attributes. The decision as to which value of λ is most appropriate should be made on a case-by-case basis by considering the importance of errors in the "held" attributes relative to errors in the "perturbed" attributes. Note that the results indicate some targets require higher penalty scaling parameter values during optimization to make time series with zero error in the "perturbed" attribute.

### 2.5.2 Comparing two penalty function structures

In order to examine the differences between the two penalty structures proposed in Section 2.2.3, both penalty structures (Equations 3 and 4) are used to create the same target time series. The requested target time series corresponds to a

30% increase in *Ptot*, which is the "perturbed" attribute, with all other attributes held constant. The target time series are created using the seasonal weather generator for the desired attributes (Table 2-1), to see how the penalties perform with more attributes in the objective function. Figure 2-8 shows the breakdown of errors across each attribute for the linear penalty term (top) and the quadratic penalty (bottom). The trialed scaling parameter values differ between the investigated penalties: the linear penalty scaling parameter was varied from 0 to 4, and the quadratic penalty scaling parameter was varied from 5 to 25.

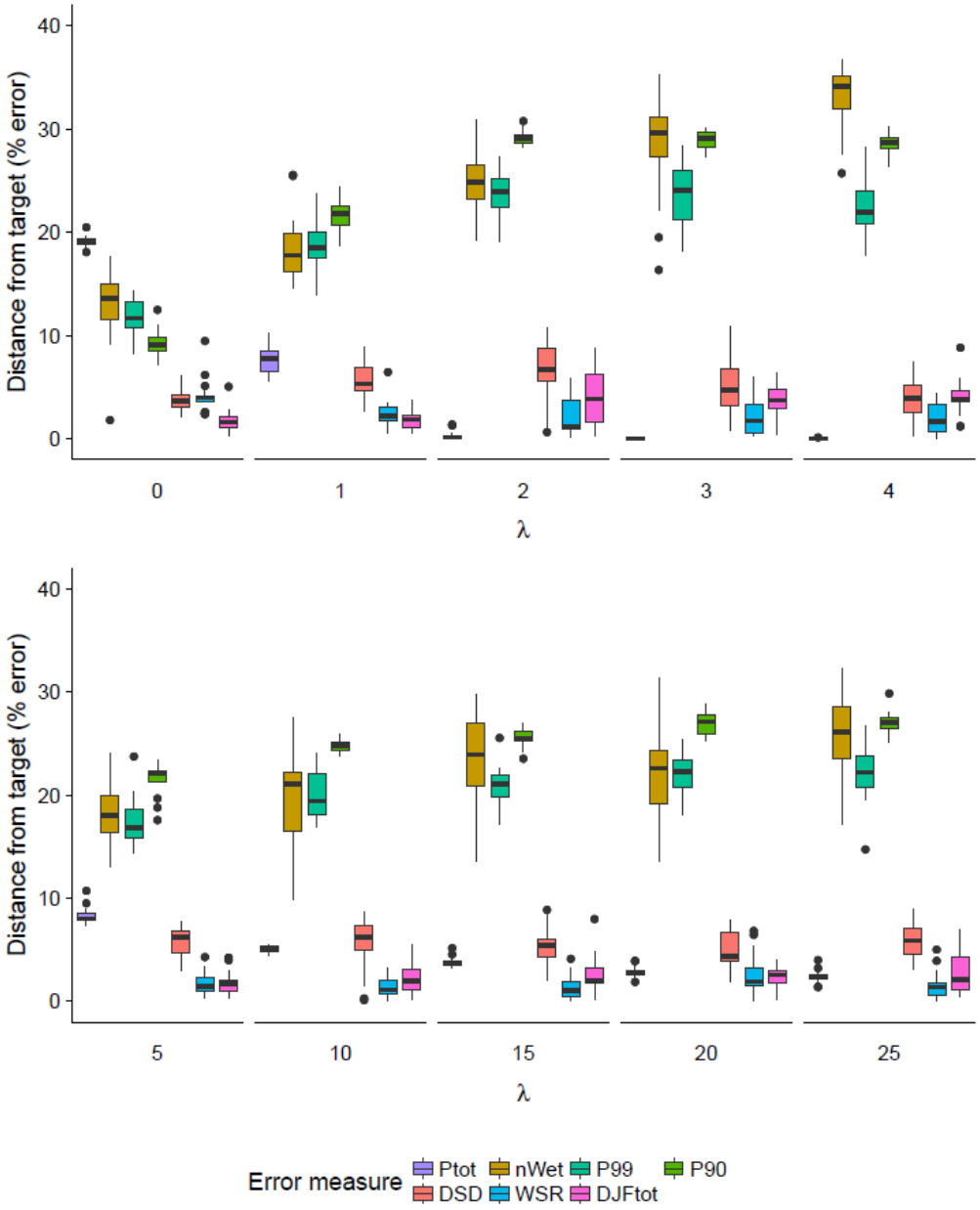


**Figure 2-8** Breakdown of the error in each attribute at the end of optimization with a varying scaling parameter λ for two different penalties: a linear penalty term (top) and a quadratic penalty term (bottom). Time series with a requested 30% decrease in *Ptot* are simulated using a seasonal weather generator.

As can be seen from Figure 2-8 (top panel), the linear penalty term performs in a similar manner to when it was used with the annual generator (Figure 2-7, bottom panel), despite the addition of two attributes. One difference is that the error in the *nWet* attribute is higher for the seasonal weather generator. It should

be noted that this distribution of error is a property of three separate elements of the optimization problem: the attributes chosen, the target set and the weather generator used. With more attributes in the objective function, attributes like *nWet* that had high error for the annual generator are now weighted relatively less and thus have higher error when the seasonal generator is used. However, despite the changes in error, a value of $\lambda$=2 is still enough to satisfy the "perturbed" attribute target.

For the quadratic penalty term, the error in the "perturbed" attribute does not reach zero, instead it approaches zero as the scaling parameter increases like in the example case shown in Figure 2-4 (Section 2.2.3). Note that larger scaling parameter values are used to reduce the error in the "perturbed" attribute, because the square root of the scaling parameter is taken in Equation 4 (Figure 2-8, bottom panel). As a result, there is less overall error in the simulated time series when averaged across the "perturbed" and "held" attributes. This is best demonstrated by examining the *nWet* attribute in the top and bottom panels of Figure 2-8. As the scaling parameter increases, the error in this attribute is around 5% less for the quadratic penalty term than it is with the linear penalty term, even though the "perturbed" attribute has near-zero error.

The decision behind which penalty structure to use in applications of the inverse approach should be made by considering the importance of error in the attributes. For example, if a target time series is set with the primary intention of reaching zero error in *Ptot*, and the other attributes are selected to make sure the stochastic time series stay similar to historical conditions, then the linear term penalty with high $\lambda$ value could be used. However, if the *Ptot* target does not need to be precisely simulated, and the other attributes have a strong bearing on system performance, the time series found using the quadratic penalty term might be more appropriate for analysis. As a result, for potential future applications it is likely that a process of trial-and-error would be needed to obtain an appropriate compromise in the trade-off in errors between attributes (and thus the penalty function and associated value of $\lambda$).

### 2.5.3 Focusing on two perturbed attributes

In order to determine if objective function penalties can be used to guide the error for multiple attributes, time series are created with two "perturbed" attributes. In addition to *Ptot*, these time series will be created with the number of wet days in the year (*nWet*) as a penalized attribute. Figure 2-9 shows the error breakdown for each attribute for the requested target of historical conditions using two "perturbed" attributes and varying the λ values for the linear penalty term (Eq. 3). Here, both *Ptot* and *nWet* are selected as "perturbed" attributes (Figure 2-7 and 2-8) as previous results demonstrated the difficulty in achieving low error in both *Ptot* and *nWet* simultaneously.

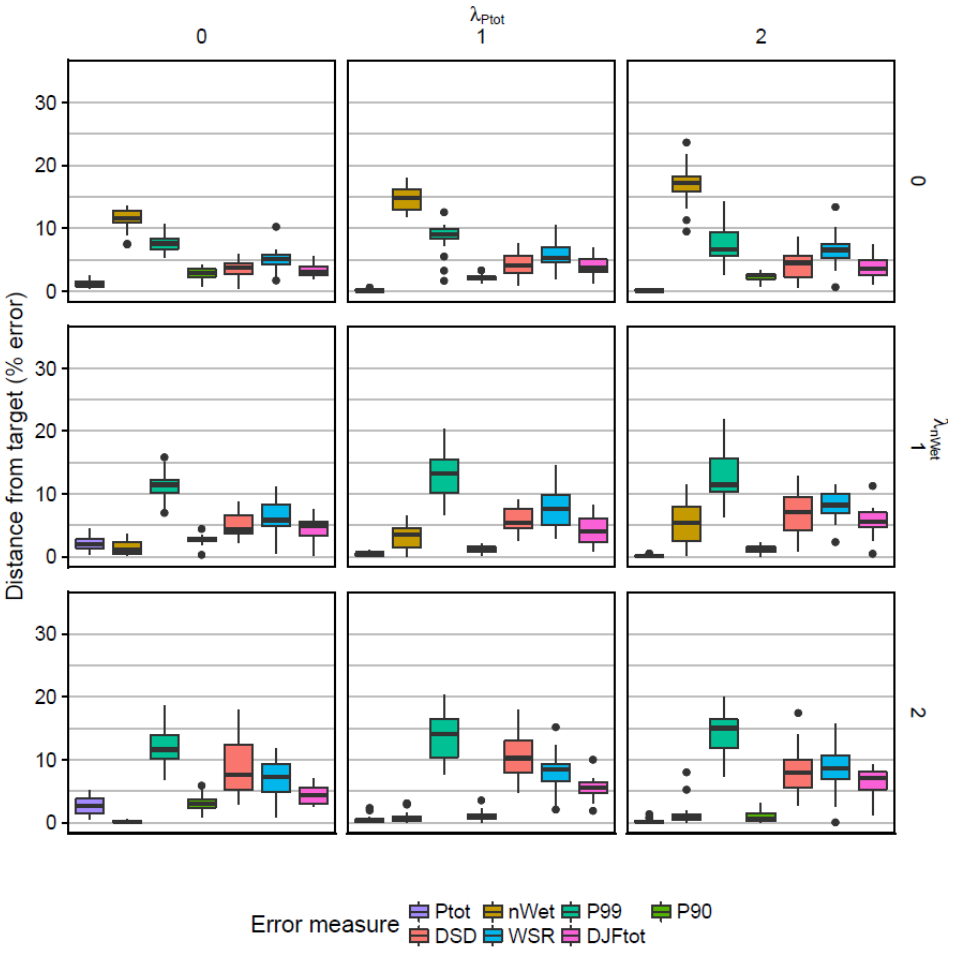


**Figure 2-9** Breakdown of the error in each attribute at the end of optimization with two "perturbed" attributes, *Ptot* and *nWet*. Time series are simulated using the seasonal weather generator for a target of historical conditions. Scaling parameters for the linear penalty term are changed separately for both attributes.

Figure 2-9 demonstrates that the seasonal weather generator can achieve near-zero error in both "perturbed" attributes, given two appropriate penalty scaling parameters. This is first seen when $\lambda_{Ptot}$=1 and $\lambda_{nWet}$=2. In this case, the penalty terms are enough to make the error in the five "held" attributes higher than they would be with only *Ptot* penalized (as in previous cases, their error stayed low while *nWet* increased). Further, these results show that the *nWet* attribute should be weighted twice as much as the *Ptot* attribute for both to achieve near-zero error from their target. Again, this ratio will be a property of the weather generator and requested targets.

To summarise the impact of penalty functions, we compare the difference in the creation of a 4x4 regular grid scenario-neutral space (16 target time series) using the "unweighted" objective function (Eq. 2) and the objective function with a linear penalty term (Eq. 3) with $\lambda_{Ptot}$=3 and $\lambda_{nWet}$=3 to ensure each target is met. The requested scenario-neutral space varies *Ptot* and *nWet* from 70 to 130% of their historical values. All other targets are "held" at historical conditions. Figure 2-10 compares the performance of the two optimization outcomes.

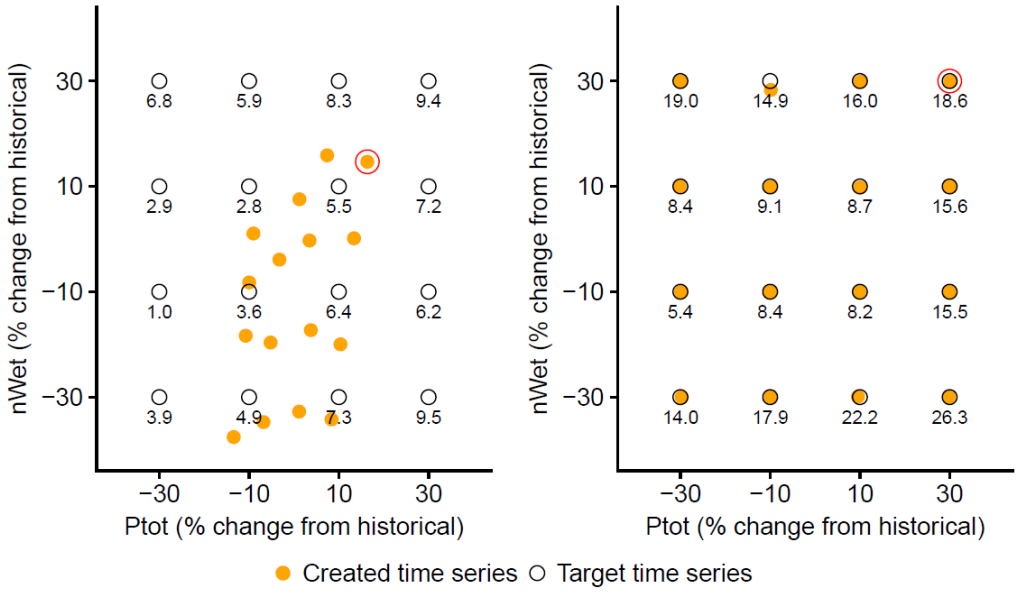


**Figure 2-10** A 4x4 scenario-neutral space made with the "unweighted" objective function (left) and the objective function with linear penalty terms (right). Targets are specified as percentage change from historical conditions, and the time series are made with a seasonal weather generator. The mean percentage error from historical conditions in the remaining "held" attributes is shown below each target. The red circle illustrates a point that requires the error in the "held" attributes to be doubled to reach the perturbed target.

When the "unweighted" objective function is over constrained, the two "perturbed" attributes cannot meet their targets (left panel Figure 2-10). This was also demonstrated in Section 2.5.2, where the error was spread across seven attributes when $\lambda$=0 (Figure 2-8, top panel). In Figure 2-10, this is seen as a clustering in the simulated targets, for the *Ptot* dimension in particular. The simulated targets are more varied in simulating *nWet*, achieving a 30% decrease but struggling to meet a 30% increase. This is due to the weather generator structure, where the number of wet days can be decreased with minimal impact on the rainfall volume extremes or seasonality as the wet/dry sequence is changed independently.

In contrast, the use of penalties in the objective function enables the generation of time series with attributes that match two target "perturbed" attributes for the majority of the scenario-neutral space (right panel Figure 2-10). However, this comes at the cost of increasing the error in the remaining "held" attributes, as

can be seen from the mean "held" attribute error shown below each target in Figure 2-10. Taking the example of the top right target (circled in red), when the unweighted objective function is used the "held" attributes are within 10% of historical levels on average, however, both "perturbed' attributes are ~15% away from their targets. In order for both "perturbed" attributes to reach their target, the error in the "held" attributes increases by a further 9% on average. Given the purpose of the "held" attributes is usually to ensure the realism of the time series, prioritization towards the "perturbed" attributes at the expense of the "held" attributes in most cases will be desirable.

## 2.6 Conclusions

The effectiveness of scenario-neutral approaches hinges on the ability to stress-test systems against plausible realizations of future climate. However, the range of changes in climate that can be examined is limited by the methods used to create the perturbed time series. Recently, the inverse approach has been presented as a method capable of producing perturbations to complex measures of hydrometeorological variables, by using formal optimization techniques with stochastic weather generators. Conceptually, this method can be applied to generate weather time series that represent not only changes in the averages, but also changes in the variability, intermittency, extremes, seasonality and/or inter-annual persistence. However, there are two key challenges to implementing the method: the large computational effort required to create the perturbed stochastic time series, and the difficulty in ensuring the realism of the time series. This paper presents approaches to overcome these challenges and improve the effectiveness of the inverse approach. Specific implementations were demonstrated using the case study of Adelaide, Australia, with a simple annual weather generator and a more complex seasonal weather generator.

As methods to increase the efficiency of an optimization process can be algorithm specific, a first step is to diagnose the nature of the optimization problem. For the weather generators used in the case study, the optimization fitness landscape was found to be irregular, due to the first-order Markov chain used to sequence wet and dry days. As a result, a genetic algorithm was selected,

and two sets of bounds on the decision variables were compared: one set of uninformed bounds, and one set based on the parameter values obtained when calibrating the weather generators to sites around Australia. Results demonstrated that the domain knowledge informed bounds increased the convergence of the optimization process by a significant amount and led to a reduction in fitness values by two orders of magnitude for the seasonal weather generator. This indicates that using domain knowledge of the weather generator parameters when employing the inverse approach can increase the efficiency of the approach, particularly with more complex weather generators.

The proposed approach for ensuring the realism of generated time series is to include attributes in the objective function that keep all properties of the time series near historical levels other than those that are being actively perturbed. This has the side effect of both making the problem more complex and creating infeasible target requests (e.g. increasing the total rainfall in the year without changing either the number of wet days or the average amount per day). The recommended solution is to add penalties to the objective function that prioritize meeting the "perturbed" attribute targets, while ensuring that the remaining attributes are "held" as close to their historical values as possible. Two penalty function structures were explored on multiple target perturbations, for various values of the penalty scaling parameter. When compared to an "unweighted" objective function, results show that the use of penalties is beneficial for creating realistic hydrometeorological time series for use in scenario-neutral spaces. Currently, the optimization approach is formulated to create step changes in climate attributes, as these time series are required for scenario-neutral spaces. Further work is required to extend this optimization formulation to creating transient time series, for use in other scenario-neutral impact assessments that do not generate scenario-neutral spaces.

Both these advances to the inverse approach allow for a greater range of perturbations to be made to historical climate records. This enables the stress-testing of systems against a broader range of climate attributes, and crucially, identifying system vulnerabilities in response to this change. This ultimately will

help ensure scenario-neutral analyses are able to identify system responses and potential failure modes to a much broader range of potential future climatic changes compared to traditional methods of time series perturbation.

## 2.7 Acknowledgements

Sam Culley was supported by an Australian Postgraduate Award. Bree Bennett was partially supported by a grant from the South Australian Goyder Institute for Water Research.

## 2.8 References

Brown, C., Y. Ghile, M. Laverty, and K. Li (2012), Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, Water Resources Research, 48(9), W09537.

Bussi, G., S. J. Dadson, C. Prudhomme, and P. G. Whitehead (2016), Modelling the future impacts of climate and land-use change on suspended sediment transport in the River Thames (UK), Journal of Hydrology, 542, 357-372.

Coello Coello, C. A. (2002), Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art, Computer Methods in Applied Mechanics and Engineering, 191(11), 1245-1287.

Culley, S., S. Noble, A. Yates, M. Timbs, S. Westra, H. Maier, M. Giuliani, and A. Castelletti (2016), A bottom-up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate, Water Resources Research, 52(9), 6751-6768.

Dorigo, M., V. Maniezzo, and A. Colorni (1996), Ant system: optimization by a colony of cooperating agents, Trans. Sys. Man Cyber. Part B, 26(1), 29-41.

Duan, Q. Y., V. K. Gupta, and S. Sorooshian (1993), Shuffled complex evolution approach for effective and efficient global minimization, Journal of Optimization Theory and Applications, 76(3), 501-521.

Gibbs, M. S., H. R. Maier, and G. C. Dandy (2011), Relationship between problem characteristics and the optimal number of genetic algorithm generations, Engineering Optimization, 43(4), 349-376.

Gibbs, M. S., H. R. Maier, and G. C. Dandy (2015), Using characteristics of the optimisation problem to determine the Genetic Algorithm population size when the number of evaluations is limited, Environ. Modell. Softw., 69, 226-239.

Guo, D., S. Westra, and H. R. Maier (2017), Use of a scenario-neutral approach to identify the key hydro-meteorological attributes that impact runoff from a natural catchment, Journal of Hydrology, 554, 317-330.

Guo, D., S. Westra, and H. R. Maier (2018), An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies, Journal of Hydrology, 556, 877-890.

Holland, J. H. (1992), Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence, 228 pp., MIT Press.

Kingston, G. B., G. C. Dandy, and H. R. Maier (2008), AI techniques for hydrological modeling and management, edited, pp. 67-69, Nova.

Maier, H. R., S. Razavi, Z. Kapelan, L. S. Matott, J. Kasprzyk, and B. A. Tolson (2019), Introductory overview: Optimization using evolutionary algorithms and other metaheuristics, Environ. Modell. Softw., 114, 195-213.

Maier, H. R., et al. (2014), Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions, Environ. Modell. Softw., 62(0), 271-299.

Malan, K. M., and A. P. Engelbrecht (2013), A survey of techniques for characterising fitness landscapes and some possible ways forward, Information Sciences, 241, 148-163.

Nesterov, Y. (2007), Gradient methods for minimizing composite objective function, edited, Citeseer.

Poff, N. L., et al. (2016), Sustainable water management under future uncertainty with eco-engineering decision scaling, Nature Clim. Change, 6(1), 25-34.

Prudhomme, C., A. L. Kay, S. Crooks, and N. Reynard (2013), Climate change and river flooding: Part 2 sensitivity characterisation for British catchments and example vulnerability assessments, Climatic Change, 119(3-4), 949-964.

Prudhomme, C., R. L. Wilby, S. Crooks, A. L. Kay, and N. S. Reynard (2010), Scenario-neutral approach to climate change impact studies: Application to flood risk, Journal of Hydrology, 390(3-4), 198-209.

Raupach, M., P. Briggs, V. Haverd, E. King, M. Paget, and C. Trudinger (2012), Australian Water Availability Project (AWAP): CSIRO Marine and Atmospheric Research, Canberra, Australia, edited.

Richardson, C. W. (1981), Stochastic simulation of daily precipitation, temperature, and solar radiation, Water resources research, 17(1), 182-190.

Richardson, C. W., and D. A. Wright (1984), WGEN: A model for generating daily weather variables.

Scrucca, L. (2013), {GA}: A Package for Genetic Algorithms in {R}, Journal of Statistical Software, 53(4), 1-37.

Spence, C., and C. Brown (2018), Decision Analytic Approach to Resolving Divergent Climate Assumptions in Water Resources Planning, Journal of Water Resources Planning and Management, 144(9), 04018054.

Steinschneider, S., and C. Brown (2013), A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, Water Resources Research, 49(11), 7205-7220.

Wilcke, R. A., and L. Bärring (2016), Selecting regional climate scenarios for impact modelling studies, Environ. Modell. Softw., 78, 191-201.

Zielinski, K., D. Peters, and R. Laur (2005), Stopping criteria for single-objective optimization, Proceedings of the Third International Conference on Computational Intelligence, Robotics and Autonomous Systems.

# Chapter 3

## *Identifying critical climate variables for use in scenario-neutral climate impact assessments based on climate-perturbed hydrometeorological time series (Paper 2)*

Sam Culley, Holger Maier, Seth Westra and Bree Bennett

Submitted to Water Resources Research

# Statement of Authorship

| | |
|---|---|
| Title of Paper | Identifying critical climate variables for use in scenario-neutral climate impact assessments based on climate-perturbed hydrometeorological time series |
| Publication Status | ☐ Published ☐ Accepted for Publication<br>☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style |
| Publication Details | Submitted to the journal Water Resources Research, under review. |

## Principal Author

| | |
|---|---|
| Name of Principal Author (Candidate) | Sam Culley |
| Contribution to the Paper | Contributed to the conception and design of the project, performed analysis, interpreted the 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 18/04/19 |

## 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 | Holger Maier |
| Contribution to the Paper | Contributed to the conception and design of the project, analysis and interpretation of the research data, and editing the manuscript |
| Signature | Date 18/4/19 |

| | |
|---|---|
| Name of Co-Author | Seth Westra |
| Contribution to the Paper | Contributed to the conception and design of the project, analysis and interpretation of the research data, and editing the manuscript |
| Signature | Date 18-04-2019 |

| | | | |
|---|---|---|---|
| Name of Co-Author | Bree Bennett | | |
| Contribution to the Paper | Contributed to the conception and design of the project, analysis and interpretation of the research data, and editing the manuscript | | |
| Signature | | Date | 18/ 4/19 |

# Abstract

Scenario-neutral climate impact assessments are being used increasingly to assess water resource system responses to possible climate changes. Essential to the success of such assessments is the simulation of changes in climate 'attributes' (statistics of climate variables) to which a system is most sensitive. However, in situations where system models require hydrometeorological time series as inputs, it can be challenging to simultaneously represent changes across a large number of climate attributes (e.g. the joint simulation of plausible changes in averages, seasonality, extremes and persistence of daily precipitation time series) within a scenario-neutral framework. This is likely to be the reason that existing studies have often defaulted to considering only changes to mean precipitation and temperature. However, this approach may miss important modes of system failure, thereby undermining one of the primary objectives of scenario-neutral studies which is to understand system sensitivity to plausible climate changes. To this end, we develop an approach that can identify 'critical' attributes with the greatest effect on system performance from a large number of plausible candidate attributes, enabling the generation of time series with attributes that correspond to key system vulnerabilities. The approach is tested on the regulated Lake Como reservoir in northern Italy considering two system performance criteria: irrigation deficit and flood reliability. Results show that system sensitivity can be adequately represented using only four critical attributes. Total annual rainfall and the number of frost days were shown to be critical for both performance criteria, with irrigation deficit also affected by autumn precipitation and December temperature, and flood reliability affected by June temperature and summer precipitation. The outcome of the scenario-neutral climate impact assessment using these attributes is shown to be very different to an analogous assessment using only mean precipitation and temperature, where for flood reliability, an assessment with mean precipitation and temperature shows 40% of climate projections indicating system failure, compared to only 15% when the identified critical climate attributes are used.

## 3.1 Introduction

Climate change can impact water resource systems through stresses to both supply and demand [*IPCC*, 2014]. These stresses are the result of changes to atmospheric variables such as precipitation and evapotranspiration, which influence water resource systems via a complex set of catchment-scale and system-level processes that in turn are dependent on the system's geography, configuration, operation and performance measures. Understanding the possible impacts of climate change on water resource systems therefore requires mapping changes in large-scale climate processes to changes in system performance, accounting for the unique features of each system [*Mastrandrea et al.*, 2010].

To this end, scenario-neutral climate impact assessments are being used increasingly to assess and convey the sensitivities of water resource systems to climate changes [*Prudhomme et al.*, 2010; *Brown and Wilby*, 2012]. These assessments work by 'stress testing' a system against a set of climatic time series that represent potential future climate conditions. These time series are then run through a system model, providing information on how the system responds to changes in climate conditions, and identifying critical performance thresholds and other decision-relevant information [*Prudhomme et al.*, 2010; *Brown et al.*, 2012]. This creates a scenario-neutral space (also referred to as an exposure space or a response surface) that maps system performance to changes in the statistics of climate variables, hereafter referred to as climate 'attributes'. The scenario-neutral space can be coupled with climate projections to understand the plausibility and possible timing of these changes [*Turner et al.*, 2014; *Taner et al.*, 2017], and to explore the robustness of different management strategies to the set of plausible future changes [*Brown et al.*, 2012; *Whateley et al.*, 2014; *Culley et al.*, 2016].

A key challenge for successfully executing a scenario-neutral assessment is to ensure that all key system sensitivities are identified [*Brown and Wilby*, 2012; *Nazemi and Wheater*, 2014], and as such a significant amount of research has been devoted to this. A prominent approach is scenario discovery [*Bryant and Lempert*, 2010], the aim of which is to identify potential future conditions

(scenarios) that have the biggest influence on a decision. This can be achieved using a number of approaches, such as the Patient Rule Induction Method or Classification and regression tree algorithms (PRIM and CART) [*Lempert et al.*, 2008]. A feature of studies in this field is that the scenario-neutral spaces are often generated by independently perturbing single values of a range of variables, or simply scaling a time series, many of which are not climate related (e.g. change in demand or costs) [*Kasprzyk et al.*, 2013; *Herman et al.*, 2014; *Shortridge and Guikema*, 2016; *Ray et al.*, 2018]. As a result, generation of the required scenario-neutral spaces can be achieved by sampling variables in a multi-dimensional space using methods such as Latin hypercube or SOBOL sampling, and then directly scaling inputs. This enables the consideration of scenario-neutral spaces that can be of quite high dimensionality (e.g. in the order of 10 dimensions).

However, the use of such high-dimensional scenario-neutral spaces becomes infeasible from a practical perspective for cases where many climate attributes of climate affected time series (rather than single values) are required to assess system performance. This is because the generation of perturbed time series of climate variables that simultaneously represent changes across a large number of climate attributes (e.g. particular combinations of changes in annual rainfall, extreme rainfall and dry-spell duration at specific times of the year) is extremely difficult [*Guo et al.*, 2018]. This difficulty increases with increasing dimensionality of the scenario-neutral space (i.e. with changes in a larger number of climate attributes that have to be satisfied by changes in the underlying hydrometeorological time series). Consequently, in order to be able to apply scenario-discovery methods to problems requiring perturbed time series of hydrometeorological variables as inputs to models used for the assessment of system performance, the dimensionality of the climate attributes to which these methods are applied has to be reduced as much as possible, while still including the climate attributes to which the decision of interest is most sensitive.

The difficulty of generating highly-dimensional scenario-neutral spaces based on perturbed hydrometeorological time series is likely to be the reason that most

time series-based scenario-neutral impact assessments have only considered changes in two climate attributes [*Weiß*, 2011; *Wetterhall et al.*, 2011; *Singh et al.*, 2014; *Turner et al.*, 2014; *Whateley et al.*, 2014; *Bussi et al.*, 2016; *Culley et al.*, 2016]. In addition, these attributes have typically been very simple, mainly corresponding to changes in mean precipitation and mean temperature, or in some cases changes to precipitation seasonality [*Prudhomme et al.*, 2013b; *Prudhomme et al.*, 2013a; *Kay et al.*, 2014] and shifts in peak flows [*Nazemi et al.*, 2013]. In these cases, the generation of perturbed time series with these attributes is much simpler than the generation of time series that correspond to changes in other climate attributes, such as extremes, intermittency and persistence.

The lack of consideration of a wider range of climate attributes in the generation of scenario-neutral spaces is likely to have a range of negative consequences, since if all important attributes are not identified, the results of a scenario-neutral impact study will be misleading. In particular, system failure modes may not be identified, and the relative performance of different decision alternatives under climate changes may be incorrect. Consequently, studies that have generated scenario-neutral spaces by perturbing time series of hydrometeorological time series suffer from the opposite problem to studies that have focused on scenario discovery, in that the dimensionality of the scenario-neutral space considered is likely to be too small, rather than too large.

Therefore, in cases where realistic climate perturbed time series with a wider range of changes in climate attributes (e.g. extremes, intermittencies and persistence) are required for performance assessment, there is a need to develop an approach that is able to identify the smallest number of climate attributes that have a significant impact on system performance. These are defined to be the ‘critical’ climate attributes for a system. Using the critical climate attributes will ensure that scenario-neutral analyses do not include unnecessary climate attributes that make it more difficult to generate time series, while including all attributes that have a significant impact on system performance thereby enabling all critical modes of system failure to be captured.

The objectives of this paper are therefore (i) to present an approach for identifying the most critical climate attributes for a given system, and (ii) to evaluate the benefits of using critical climate attributes in a scenario-neutral impact assessment. The approach is demonstrated and tested using the Lake Como system, a regulated lake in Northern Italy. Two performance criteria are considered—flood reliability and irrigation deficit—to further highlight the importance of tailoring the approach to each system objective [*Kasprzyk et al.*, 2013].

Details of the proposed approach can be found in Section 3.2, with a description of the case study and implementation of the proposed approach in Section 3.3. A description of further analysis designed to test how well the proposed approach performed is also presented in Section 3.3, as well as a demonstration of a decision scaling type impact assessment [*Brown et al.*, 2012] using the critical climate attributes. The results of implementing and testing the proposed approach are presented in Section 3.4, along with a discussion of its advantages and limitations. Conclusions are presented in Section 3.5.

## 3.2 Approach to identifying critical climate attributes

A systematic approach is proposed to identify 'critical' climate attributes that are most important for the analyzed system. The approach starts with the identification of a large set of 'candidate' attributes based on *a priori* knowledge of both the system dynamics and possible future climatic changes. Given the complexity of the response to projected climatic changes for most water resource systems, and the need for the candidate set to encompass all plausible drivers of change to system performance, the candidate attribute set in most cases is likely to be very large (potentially including statistics such as the average, extremes, seasonality, intermittency and inter-annual variability of multiple hydrometeorological variables such as precipitation, temperature and potential evapotranspiration). Working in such a high-dimensional attribute space would not only lead to computational challenges in simulating the scenario-neutral space, but also to difficulties in identifying failure boundaries and otherwise visualizing and interpreting such a high-dimensional space. For this reason, the

'critical' attributes should represent a reduced-dimension subset of the candidate attributes that are the most important for a particular system, and are also relatively independent of the other critical attributes—thereby maximizing the amount of 'new information' on system sensitivity provided by each critical attribute. The overall workflow for identifying these critical climate attributes is shown in Figure 3-1.

A core feature of the proposed approach is the use of the partial mutual information (PMI) algorithm introduced by *Sharma* [2000] to rank the set of candidate attributes [*Li et al*., 2015a, b] using a low-resolution sampling of the scenario-neutral space. This approach was selected because of its ability to represent non-linear effects as well as the ability to capture partial performance of each attribute given the previous attributes in the critical attribute set. The PMI provides a ranked list of attributes in order of importance. This is followed by calculation of the cumulative variance explained (CVE) to assess the added value of each additional attribute and can be used to identify a cut-off point for the critical attribute set.

Essential to the success of the approach is the ability to generate hydrometeorological time series that represent a wide range of potential future climate changes in the candidate attribute space, and which are used as inputs to the system model to simulate how system performance changes in response to possible climatic changes. This is achieved using the inverse approach of *Guo et al.* [2018], in which a formal optimization approach is used to identify the parameters in a weather generator that enables the generation of hydrometeorological time series that match the desired attribute values as closely as possible. The inverse approach is used both to provide the low-resolution sampling of the high-dimensional candidate attributes, as well as the more detailed simulation of the scenario-neutral space using the critical attributes. Further details of each of the steps in Figure 3-1 are given in the following sub-sections.

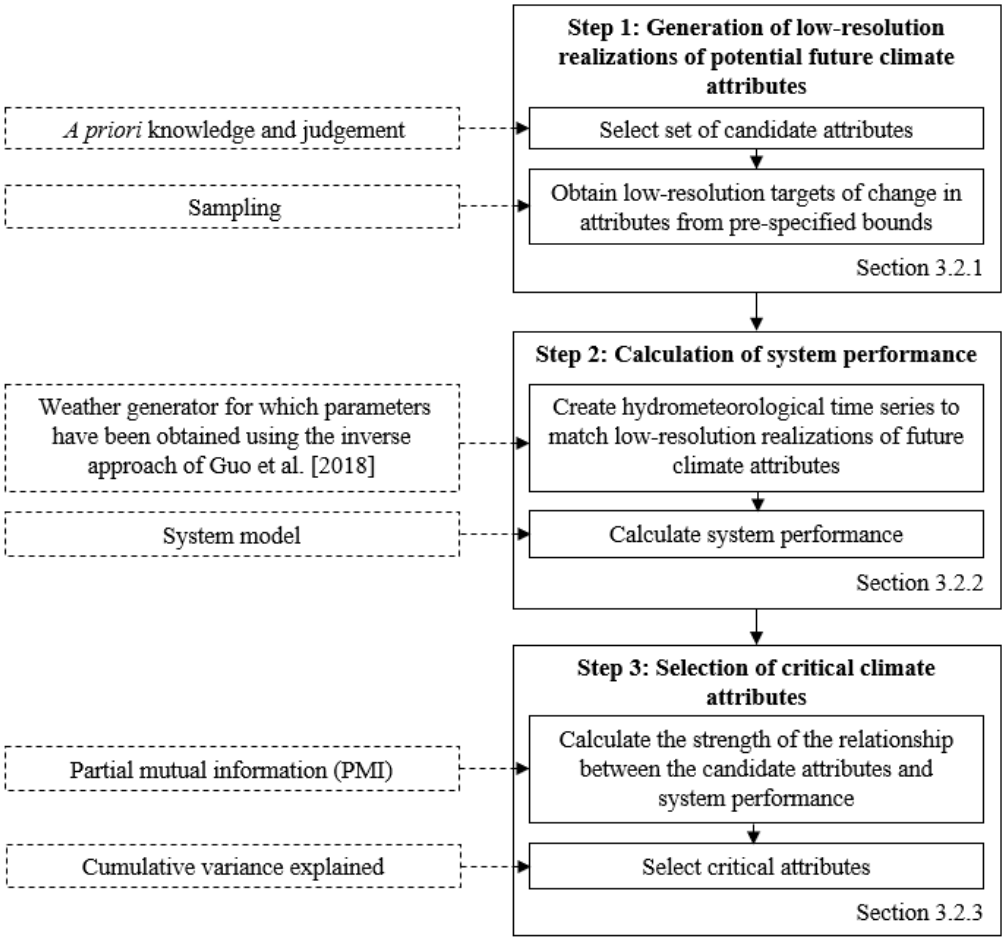


**Figure 3-1** Main steps in proposed approach for selecting critical climate attributes (solid boxes) and the methods used to implement them (dashed boxes).

### 3.2.1 Step 1: Generation of low-resolution realizations of potential future climate attributes

In this step, realizations of a wide range of potential future climate conditions that may influence system performance are generated at a low resolution. This requires a candidate set of attributes, from which different combinations are sampled from plausible ranges to represent potential future climate conditions to which the system of interest might be exposed. Details of these two sub-steps are given below.

#### *3.2.1.1 Select candidate attributes*

A key objective of the proposed approach is to ensure that scenario-neutral methods can identify all critical modes of system failure. As a result, it is

important to ensure that a wide range of climate attributes are identified *a priori*, from which the attributes that best describe system performance can be determined. This set is referred to as the 'candidate attributes' ($\boldsymbol{a} = a_1, \ldots, a_n$) and the resulting space of potential values of these attributes is referred to as the 'attribute space', $a \subseteq \mathbb{R}^n$ , where *n* is the number of candidate attributes.

While it is important to select a broad set of candidate attributes to ensure that all potential system failure modes are represented, as mentioned previously, this does increase the computational requirements of the next step in the process (Figure 3-1). Consequently, careful consideration should be given to the candidate attribute selection phase. It should be noted that as system performance is determined using a system model (see Section 3.2.2), the candidate attributes that can be considered are limited to those that can be used as system model inputs.

#### *3.2.1.2 Defining and sampling the candidate attribute space*

System performance across the entire candidate attribute space needs to be explored to understand the higher order interactions between attributes, and to determine which attributes most affect system performance. In order to define the candidate attribute space, bounds need to be placed on each of the candidate attributes. To ensure all possible drivers of system performance are considered, it is important that these bounds cover the full range of changes that might be expected as a result of future climate change. However, as the relative impact of a particular climate attribute on system performance can be a function of the selected range, care needs to be taken to ensure the selected range is plausible under projected future climate conditions. One potential method for achieving this is to extend the bounds just beyond those projected by relevant climate models and/or other available lines of evidence to allow for uncertainty in these models [*Brown and Wilby*, 2012].

In order to explore system performance across the entire candidate attribute space in a computationally efficient manner, a low-resolution representative sample of values from this space should be generated. This is referred to as the

set of low-resolution attribute 'targets', as the aim of subsequent steps is to generate hydrometeorological time series that have these attributes (Figure 3-1). As the number of attributes that could have an impact on system performance, and hence the space to be explored, is potentially quite large (e.g. $n > 10$), it is necessary to implement a sampling strategy that efficiently covers the attribute space. Methods like Latin hypercube sampling and improved distributed hypercube sampling have proved to be useful for this purpose [*Stein*, 1987; *Beachkofski and Grandhi*, 2002]. It is suggested, given large candidate sets ($n > 10$), that the number of samples taken, $s$, should be upwards of 10,000. This is sufficiently large enough to ensure selection accuracy using the PMI algorithm (Section 3.2.3) [*May et al.*, 2008; *Li et al.*, 2015b]. Consequently, the target attributes consist of the set $(A_{1,i,\ldots},A_{n,i})$, where $i = 1,\ldots,s$ and $s$ is the number of targets generated for the $n$ attributes.

### 3.2.2 Step 2: Calculation of system performance

In order to determine the climate attributes that have the largest impact on system performance, system performance must be determined for all low-resolution targets of candidate attributes generated in Section 3.2.1. This requires these target attributes to be converted to a form that can be used as inputs to a system model, which for many hydrological applications corresponds to hydrometeorological time series that have these attribute targets.

#### *3.2.2.1 Create hydrometeorological time series*

It is proposed to generate the hydrometeorological time series that match the low-resolution targets of the desired future climate attributes with the aid of a weather generator, resulting in a set of modelled targets of the desired climate attributes $(\hat{A}_{1,i,\ldots},\hat{A}_{n,i})$. In order to ensure the modelled attributes $(\hat{A}_{1,i,\ldots},\hat{A}_{n,i})$ are as similar as possible to the targets $(A_{1,i,\ldots},A_{n,i})$, it is proposed to use the inverse approach of *Guo et al. [2018]*. Implementation of this approach requires the use of an appropriate optimization algorithm, such as an evolutionary algorithm [*Maier et al.*, 2014; *Guo et al.*, 2018; *Maier et al.*, 2019], to identify the set of parameters of the stochastic weather generator, $\theta$, that minimizes an objective function that measures the difference between the values of $(A_{1,i,\ldots},A_{n,i})$

and ($\hat{A}_{1,i,\ldots},\hat{A}_{n,i}$), such as the root mean squared error (Figure 3-2).

The ability to meet the targets depends on both the weather generator and the set of candidate attributes used. Both of these factors are why it is necessary to preserve the modelled targets to calculate the influence of each attribute on the system (Figure 3-2). In relation to the former, many models for generating stochastic hydrometeorological time series are available, which generally differ in terms of their structure [*Richardson*, 1981; *Richardson and Wright*, 1984]. Simpler models may not have the degrees of freedom required to change the relevant attributes of a time series independently (e.g. winter precipitation and summer precipitation, when using a weather generator that does not simulate seasonal variability). Given the need to include a wide selection of attributes in the candidate set, it is recommended to use a more complex weather generator with the required degrees of freedom in its parameters. With regard to the latter, it is possible that attributes will be included in the candidate set that will co-vary with other attributes due to the structure of the weather generator used. This can result in an inability to meet samples formed by perturbing attributes independently. More information on challenges and possible solutions related to using the inverse approach to achieve specified target attributes is given in *Culley et al.* [2019].

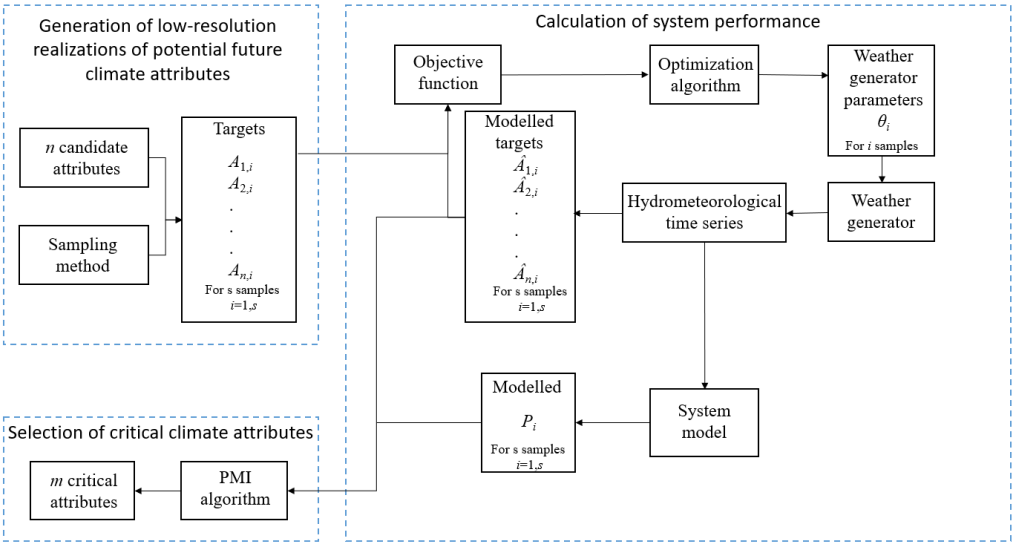


**Figure 3-2** Process for generating the climate time series that meet the candidate attribute targets, and resulting performance for critical attribute selection.

#### *3.2.2.2 Calculate system performance*

The model used to calculate system performance under plausible future climate conditions comprises either a single model or a coupled set of models that takes the generated hydrometeorological time series as inputs and simulates the resulting system performance as the output. This enables an assessment of how system performance changes under a range of plausible climate futures. It should be noted that the time step of the weather generator must match that of the system model so that system performance can be simulated for each climate time series.

The system model should be able to simulate the key processes of the investigated system, so that the strength of the simulated relationship between attributes and system performance is representative of the real system. This is a challenge that affects all scenario-neutral impact assessments, and it has recently been demonstrated that the choice of system model can lead to different conclusions about system sensitivity [*Guo et al.*, 2017b; *Broderick et al.*, 2019]. A further challenge affecting all scenario-neutral assessments is that models may be required to simulate system behavior outside the bounds of historical variability. Assessing system model performance under changed (or non-stationary) climate regimes is an active research endeavor, with possible approaches including differential split-sample testing and optimizing model performance for different climate conditions [*Klemes*, 1986; *Westra et al.*, 2014; *Fowler et al.*, 2016]. Given that the contribution of this paper is to improve a specific aspect of scenario neutral approaches, and not on assessing the utility of scenario-neutral approaches *per se*, the influence of the choice of system model is not examined.

### 3.2.3 Step 3: Selection of critical attributes

The final step is to reduce the *n* candidate attributes to the *m* critical attributes (i.e. the smallest number of attributes that have a significant impact on system performance) (Figure 3-2). As mentioned in the Introduction, this is critical in the case where the scenario-neutral space consists of attribute values that correspond to statistics of climate-perturbed hydrometeorological time series, as the inclusion of unnecessary attributes makes it more difficult to generate time

series that satisfy all required attributes simultaneously. Conversely, the exclusion of attributes that have a significant impact on system performance makes it more difficult to identify critical modes of system failure. As mentioned in Section 3.2.2.1, the method used for the selection of the *m* most critical attributes from the *n* candidate attributes has to not only be able to identify the relative influence of each candidate attribute on system performance, but also account for any redundancies (i.e. co-variations) in the candidate attributes, so that the smallest attribute set that has a significant impact on system performance can be identified.

#### *3.2.3.1 Determine relative significance of candidate attributes*

Given that co-variability in the candidate attributes needs to be taken into account in order to identify the smallest attribute set that has a significant impact on system performance [*Galelli et al.*, 2014], global sensitivity analysis methods based on sampling techniques such as SOBOL or eFAST [*Gao et al.*, 2016; *Pianosi et al.*, 2016; *Whateley and Brown*, 2016; *Guo et al.*, 2017a] are not suitable for this task. Instead, it is proposed to use a variant of the partial mutual information (PMI) algorithm [*Sharma*, 2000; *Li et al.*, 2015a] for this purpose, as it has been shown to be able to identify significant driving variables correctly for a range of linear and non-linear relationships, as well as their relative order of influence [*May et al.*, 2008; *Galelli et al.*, 2014; *Li et al.*, 2015a]. It is also able to take into account redundancy, or co-variability, between climate attributes. The use of this PMI algorithm will therefore enable the smallest number of climate attributes that have a significant impact on system performance to be identified, thereby reducing the dimensionality of the scenario-neutral space as much as possible, while still capturing all critical modes of system performance variation.

The PMI algorithm is used to rank candidate attributes in order of significance based on the mutual information between each of the *n* candidate attributes and system performance, after removing the effect of higher-ranked (i.e. more important) attributes. The mutual information (MI) between two variables is a measure of how much information about one variable is gained by an

observation of the other. For the set of targets, *s*, the MI between each attribute, $\hat{A}_k$, and the measure of system performance, *P*, can be approximated by the following equation:

$$I_{A_k,P} \approx \frac{1}{s}\sum_{i=1}^{s}\log\left[\frac{f\left(A_{k,i},P_i\right)}{f\left(A_{k,i}\right)f\left(P_i\right)}\right] \quad \textbf{(1)}$$

where *k* indexes the *n* candidate attributes (i.e. $k = 1, \ldots, n$), and *f* are the marginal or joint probability distributions functions of the attributes and system performance, which are typically estimated using kernel density estimators [*Galelli et al.*, 2014; *Li et al.*, 2015b]. The MI is non-negative and unbounded, unlike other dependence metrics (e.g. Pearson correlation, Spearman rank). It should be noted that as the generated hydrometeorological time series are used to assess system performance via the system model (Figure 3-1 and 3-2), the modelled targets ($\hat{A}_{1,i,\ldots},\hat{A}_{n,i}$) have to be used as part of the analysis to determine the *m* critical attributes with the aid of the PMI algorithm, rather than ($A_{1,i,\ldots},A_{n,i}$) (Figure 3-2).

As part of the PMI algorithm, the climate attribute with the highest MI value is selected first. In order to identify the climate attribute that has the next highest *additional* impact on system performance, the influence of the selected attribute needs to be removed. This is achieved by developing non-linear regression relationships between the already selected attribute and both the remaining attributes and system performance. The MI can then be calculated between the residuals of these models, which is referred to as the PMI. The required non-linear regression models can be developed using kernel based methods such as general regression neural networks [*Specht*, 1991], or kernel free methods such as multi-layer perceptron artificial neural networks (MLPANNs) [*Wu et al.*, 2014]. While kernel based methods are generally more computationally efficient, MLPs have been found to perform better if the data are highly non-Gaussian [*Li et al.*, 2015b]. The iterative process outlined above is repeated until all candidate attributes have been ranked.

### *3.2.3.2 Select critical attributes*

The rankings of the candidate attributes alone are not indicative of how many should be selected to create a scenario-neutral space. Further information is needed to select the *m* critical attributes from the ranked *n* candidate attributes. This can be achieved with the aid of the cumulative variance explained (CVE) in system performance, which is calculated for each selected attribute by:

$$CVE_j = 1 - \frac{SS_{err_j}}{SS_P} \quad \textbf{(2)}$$

Where $SS_P = \sum_{i=1}^{s} (P_i - \bar{P})^2$ is the total sum of squares of the system performance, *P*. For the first selected attribute (*j=1*), the sum of squares of the residuals of the non-linear regression between the first attribute and system performance is given by $SS_{err_1} = \sum_{i=1}^{s} (P_i - P_i)^2$, such that the CVE metric equates to an $R^2$ calculation. For each additional attribute, $SS_{err_j} = \sum_{i=1}^{s} \left( (P - g(r_j))_i - (P - g(r_j))_i \right)^2$ is the sum of squares of the residuals of the non-linear regression between the $j^{th}$ attribute and system performance, both of which have been updated by removing the effect of the set of *j-1* previously selected attributes, where $g(r_j)$ is the effect of the previous selected attributes, $r_j = a_{1,\ldots,a_{j-1}}$. This is similar to an analysis of variance metric (ANOVA), except that as opposed to being calculated for each attribute based on the initial sample, it is calculated iteratively after each non-linear regression transformation of the original sample.

The relative gain in this metric with each additional attribute can be used to decide how many attributes are critical for the system under consideration. This is achieved by visual inspection of the plot of cumulative variance explained and number of attributes (ranked in order of significance by the PMI algorithm) and exclusion of the attributes after which the CVE plateaus, leaving the *m* critical attributes. It should be noted that the absolute cumulative variance explained values at which the plateau occurs can provide an indication of whether all

critical attributes have been included in the candidate set. If a significant portion of variance in the performance is still unexplained, there are aspects of the climate time series investigated that are changing in a way affecting system performance, and they are not being measured by the candidate attributes.

## 3.3 Case study, implementation and testing of proposed approach

The case study of the regulated reservoir is used to demonstrate and test the approach outlined in Section 3.2. The critical climate attributes are identified for two competing system objectives (flood mitigation and irrigation supply), to determine whether different critical attributes are identified for each. Section 3.3.1 describes the case study models and data, and Section 3.3.2 describes the specific implementation of the steps described in Section 3.2. Section 3.3.3.1 presents further analysis used to assess how well the approach selects critical attributes. Finally, to demonstrate the utility of the proposed approach, Section 3.3.3.2 presents a comparison between the scenario-neutral spaces generated by considering the critical attributes with those generated by considering mean precipitation and temperature as the axes of the scenario-neutral space, which is the approach most commonly used in the literature.

### 3.3.1 Case study models and data

The Lake Como case study is located in the alpine region of northern Italy [*Anghileri et al.*, 2013; *Culley et al.*, 2016]. The maximum storage volume in the lake is 254 $Mm^3$ and provides water for one of the largest irrigation systems in Europe. There is a large snowmelt component to the inflows each year. The regulation of Lake Como is driven by two primary objectives: mitigating flooding and supplying irrigation needs, with an additional minimum environmental flow release constraint each day to ensure adequate conditions in the downstream Adda River. This study considers a simplified representation of the Lake Como system to illustrate the proposed approach, as a snowmelt component to the hydrological model means it is influenced by more complex attributes than annual means, and two competing objectives in the system model allows a comparison of how the identified critical climate attributes can differ.

The Lake Como catchment is modelled by a lumped HBV model [*Bergström and Singh*, 1995]. The hydrometeorological inputs to the model are daily precipitation and temperature time series. Evaporation is calculated internally and is dependent on temperature, as is the snowpack component that stores rainfall in the catchment as snow until a melting temperature is reached. The historical data include records of daily temperature, precipitation and streamflow into the lake across the baseline period (from 1965 to 1980) used to calibrate the hydrological model [see *Anghileri et al.*, 2011]. Catchment average precipitation data are used, derived from five gauges across the catchment, while the daily temperature data are from one site.

The reservoir system is modelled with a mass balance equation at a daily time step, in which there is a controlled release each day. The reservoir releases vary based on the time of year and height of the reservoir. In order to represent typical historical operational decisions over the baseline period, the releases are simulated using a radial basis function fitted to historical operating characteristics [*Giuliani et al.*, 2014; *Giuliani et al.*, 2015; *Culley et al.*, 2016]. As radial basis functions are dependent on reservoir height, they allow for some adaptation to climate variability both in and across different climate time series. The annual demand pattern is held constant throughout all climate time series; this does not reflect expected changes to demand as a result of changing climate regimes, but this does not diminish the value of the case study in meeting the primary objectives of this study.

In this application, the system model is configured to simulate two performance criteria calculated as annual average values: flood reliability and irrigation deficit. The flood reliability criterion is calculated as the fraction of days the reservoir height is below the flood threshold. For the irrigation deficit criterion, each daily release is compared to the irrigation demand for that day. If demand is met, the deficit is zero, otherwise the volume of the daily deficit is recorded in kL. The total annual deficit is then averaged over the number of years of simulation.

The twenty climate projections used in this analysis are from the EUROCODEX

project, where projections were created for Europe at a 50 km resolution (EUR 44) and a 12.5 km resolution (EUR 11), both of which are used in this study. The projections were bias-corrected using quantile mapping [*Boé et al.*, 2007; *Déqué*, 2007]. An increased focus on regional simulations in this project provided higher daily precipitation intensities that are not captured in the global climate model simulations. This project also showed more regional variation in projected change to selected ETCCDI indices [*ETCCDI*, 2013; *Jacob et al.*, 2014], such as the number of frost days in the year and average dry spell durations. Climate projections for the period of 2040-2060 were used in this study to determine the bounds of the attribute space (Sections 3.3.2.1).

### 3.3.2 Implementation of proposed approach

This section describes the implementation of the approach presented in Section 3.2. Sections 3.3.2.1 to 3.3.2.3 describe each of the three steps, respectively.

#### *3.3.2.1 Generation of realizations of potential future climate attributes*

The low-resolution realizations of potential future climate attributes are obtained by sampling from a pre-determined set of candidate attributes. This set of candidate attributes is restricted to the hydrometeorological variables temperature and precipitation, as these are the inputs to the system model for the case study. In selecting statistics of these two variables, our aim was to include a wide range of attributes, including those previously unexplored in scenario-neutral impact assessments. The ETCCDI list of extreme climate indices [*ETCCDI*, 2013] served as a starting point for this, providing attributes like the wet spell duration (WSD), 99$^{th}$ percentile rainfall (P99), and the number of frost days (F0). Several non-extreme attributes were also included, such as annual rainfall and average temperature, which are known to be important drivers in water resource systems. In practice, this process can also include discussions with stakeholders and incorporation of any existing expert knowledge. The full list of the fifteen candidate attributes ($n$=15) considered for Lake Como ($\boldsymbol{a} = a_{1,\ldots,}a_{15}$) is shown in Table 3-1.

**Table 3-1** Candidate climate attributes considered, where the attributes are defined as the average over the simulation period.

| Attribute name | Description | Units |
|---|---|---|
| Winter total rainfall (PDJF) | Annual summation of summer rainfall | mm |
| Summer total rainfall (PJJA) | Annual summation of winter rainfall | mm |
| Spring total rainfall (PMAM) | Annual summation of spring rainfall | mm |
| Autumn total rainfall (PSON) | Annual summation of autumn rainfall | mm |
| $99^{th}$ percentile rainfall (P99) | Volume of $99^{th}$ percentile rainfall event | mm |
| Average Wet Spell Duration (WSD) | Average length of consecutive wet days | mm |
| Annual rainfall volume (Ptot) | Annual summation of rainfall volume | mm |
| Number of wet days (nWet) | Annual count of wet days | Days |
| Average March temperature (TMar) | Average of daily temperature in March | ℃ |
| Average June temperature (TJun) | Average of daily temperature in June | ℃ |
| Average September temperature (TSep) | Average of daily temperature in September | ℃ |
| Average December temperature (TDec) | Average of daily temperature in December | ℃ |
| Average temperature (Tavg) | Annual average of daily temperature | ℃ |
| Annual temperature range (Trng) | Temperature difference between $5^{th}$ percentile and $95^{th}$ percentile day | ℃ |
| Frost days (F0) | Annual count of days with temperature less than 0℃ | Days |

Candidate attribute values were then calculated from the selected climate projection time series in order to determine the bounds of the attribute space. Each of the candidate attribute values was averaged over the period 2040-2060, and bounds were chosen to extend slightly beyond these values (Table 3-2). It should be noted that precipitation attributes are described by fraction changes from historical values, and temperature attributes are described by additive changes.

**Table 3-2** Expected change in attributes from historical baseline.

| | | Projections 2040-2060 | | Attribute bounds | |
|---|---|---|---|---|---|
| **Attribute** | **Target Type** | **Min** | **Max** | **Min** | **Max** |
| PDJF | fraction | 1.22 | 2.10 | 1.0 | 2.2 |
| PJJA | fraction | 0.43 | 1.20 | 0.4 | 1.2 |
| PMAM | fraction | 0.97 | 1.45 | 1.0 | 1.5 |
| PSON | fraction | 0.70 | 1.21 | 0.7 | 1.3 |
| P99 | fraction | 0.92 | 1.29 | 0.8 | 1.4 |
| WSD | fraction | 0.64 | 0.86 | 0.6 | 1.0 |
| Ptot | fraction | 0.89 | 1.30 | 0.8 | 1.4 |
| nWet | fraction | 0.89 | 1.03 | 0.8 | 1.1 |
| TMar | additive | 0.11 | 8.85 | 0 | 9 |
| TJun | additive | -0.43 | 8.41 | -1.0 | 8.5 |
| TSep | additive | 1.37 | 11.25 | 0 | 11.5 |
| TDec | additive | -1.53 | 8.25 | -2.0 | 8.5 |
| Tavg | additive | 1.31 | 6.14 | 0.0 | 6.5 |
| Trng | additive | -1.43 | 8.50 | -2.0 | 8.5 |
| F0 | additive | -59.8 | -12.5 | -70 | 0.0 |

A Latin hypercube sample (LHS) method was chosen to sample the candidate attribute space to generate the low-resolution scenario-neutral space. Fifteen thousand samples were taken from within the bounds to provide reasonable sampling of the scenario-neutral space, while considering the high computational cost of optimizing each climate time series.

#### *3.3.2.2 Calculation of system performance*

Climate time series of precipitation and temperature were first generated using the inverse approach for the 15,000 sampled targets of attributes. The inverse approach was implemented using the R package *fore*SIGHT [*Bennett et al.*, 2018] and required approximately 20,000 CPU hours to generate the samples. Each climate time series had a length of 21 years, which was selected to match the climate projection window of 2040-2060. The genetic algorithm used in *fore*SIGHT was used with default operators [*Scrucca*, 2013], and 300 generations. A single weather generator replicate was used for each target. The time series were then used as an input to the hydrological model to simulate daily flow. The Lake Como reservoir model then used these daily flow time series to

calculate the two performance criteria in response to each time series.

#### *3.3.2.3 Selection of critical attributes*

The PMI algorithm input variable selection was implemented to rank the set of candidate attributes (Section 3.2.3), using the modelled targets of candidate attributes calculated using each of these time series ($\hat{A}_{1,i},\ldots,\hat{A}_{15,i}$) as per Figure 3-2. The algorithm was implemented for the two performance criteria using the software developed by *Li et al.* [2015b]. A conventional Gaussian kernel was used with the Gaussian reference bandwidth [*May et al.*, 2008] to estimate the probability distributions functions for the partial mutual information calculation. A multi-layer perceptron artificial neural network (MLPANN) model was used as the non-linear regression technique to remove the influence of each selected attribute. This model is limited to one hidden layer, with a maximum of four nodes to avoid over-fitting [*Li et al.*, 2015b]. Both of these methods are recommended for use with highly non-linear data sets.

### 3.3.3 Methods to evaluate the utility of the proposed approach

#### *3.3.3.1 Selection of correct critical attributes*

In addition to implementing the proposed approach in order to identify the critical attributes for the case study, further analysis was performed to test how well the approach works. This was achieved by a test that investigates (i) how well the PMI algorithm performs in terms of ranking the candidate attributes in order of decreasing impact and (ii) how well the cumulative variance explained metric can be used to select critical attributes. The residuals obtained after removing the effect of each additional attribute (Section 3.2.3.1) were inspected visually, along with the relationship obtained from the corresponding MLPANN model. A decrease in the strength of the signal in the residuals as each attribute is removed indicates that the attributes are being ranked correctly. Such plots should be inspected for all selected critical attributes, plus the next most significant attribute. If the residuals for the next most significant attribute have no signal, there is no more information left to describe, and hence the correct number of critical attributes has been selected.

#### *3.3.3.2 Scenario-neutral climate impact assessments*

As mentioned at the beginning of Section 3.3, to demonstrate the utility of the proposed approach, two scenario-neutral impact assessments are presented: one with scenario-neutral spaces constructed using the critical attributes identified using the proposed approach and the other with a scenario-neutral space constructed using just mean precipitation and temperature as the axes of the scenario-neutral space. As part of the assessments, the scenario-neutral spaces were divided into regions of success and failure, as is done in decision scaling [*Brown et al.*, 2012], and the climate projections for the Lake Como region, which correspond to changes in each attribute in the year 2050, were overlaid. The proportions of climate projections that indicate unsuccessful performance for the two different scenario-neutral impact assessments (i.e. one using the critical attributes identified using the proposed approach as the axes of the scenario-neutral space and the other using mean precipitation and temperature) were then compared.

The acceptable thresholds for the performance criteria in this study were selected by considering simulated performance under current conditions and adding a small buffer. This is representative of safety factors that would have been used in the design of the Lake Como reservoir. For the irrigation deficit criterion, the selected threshold was a 10% increase in deficit compared to performance under current climate conditions (4224 kL deficit threshold compared to 3840 kL under current conditions). Similarly, the selected flood reliability threshold was 95%, which provides a buffer compared with the current reliability of 98%.

The stochastic time series for the scenario-neutral spaces were created using the technique outlined in Section 3.2.2. For the mean precipitation and temperature space, target changes were set using a high-resolution sample between the bounds in Table 3-2. For the critical attribute spaces, the first two dimensions were sampled in high-resolution between the bounds indicated in Table 3-2, and the third and fourth dimensions are explored using larger step changes (three or four intervals depending on the width of the bounds) to allow the space to be visualized in a lattice plot. The R package *fore*SIGHT is again used to

stochastically generate 21 years of daily precipitation and temperature data, with five weather generator replicates used for each target, requiring approximately 4500 CPU hours. The performance criteria were then calculated in response to these time series using the system model, and the average performance across the five replicates is shown for each target.

The computational cost of this high-resolution sampling of four dimensions is approximately a quarter of the 20,000 CPU hours required for the low-resolution sample of the 15 candidate attributes. However, if this high-resolution sample of attributes was extended to the remaining 11 attributes (using the same lattice plot structure), a lower-bound estimate of the CPU requirements would be 380,000 CPU years. This intractable computational cost, in addition to the visualisation challenges of the scenario-neutral space in 15 dimensions, demonstrates the need to take a low-resolution sample of the candidate attributes in order to then identify and use critical climate attributes for a scenario-neutral assessment.

## 3.4 Results and discussion

Section 3.4.1 presents the results for the approach detailed in Section 3.2. This includes the PMI calculations, as well as an analysis of how many critical attributes should be selected to capture system vulnerabilities for the different performance criteria. Section 3.4.2.1 presents the results of the additional analysis conducted to test how well the approach performs, and Section 3.4.2.2 presents the results of the decision scaling type analysis detailed in Section 3.3.3.2.

### 3.4.1 Selection of critical attributes

Figure 3-3 shows the cumulative variance explained (CVE) metric calculated for each of the candidate attributes, for each of the two performance criteria, while details of which attributes are critical and which attribute is the next most significant are given in Table 3-3. As can be seen from Figure 3-3, the CVE plateaus after four attributes for the flood reliability criterion. The most critical attribute for the flood reliability criterion is *F0* (Table 3-3), which makes sense from a physical perspective, as this attribute controls the number of days on

which temperatures are above and below zero degrees, and thus provides an indicator of the snow storage in the year. Historically, the largest inflows into the reservoir are in spring as the snow storage is released. Decreasing the number of frost days in the year means warming occurs earlier, changing the timing of the snowmelt. This causes more flooding events, given the reservoir operation was designed for historical conditions. *Ptot* is the next most significant attribute, which also makes physical sense, as annual rainfall largely dictates the amount of water in the reservoir. The next critical attribute is *PSON*, which, when the total annual rainfall is held constant, controls the seasonality in the shoulder seasons. Finally, *TDec* is selected, which, when the number of frost days is held constant, shifts the lowest temperatures into January and February and does not result in as much storage of snow in December. As the reservoir does not release water during December, this can cause the runoff from any high rainfall events to cause flooding, where previously this runoff would have been stored as snow. The addition of the next most significant attribute, *TMar*, increased the CVE value by only 0.001 (Table 3-3), and so this is not considered critical.

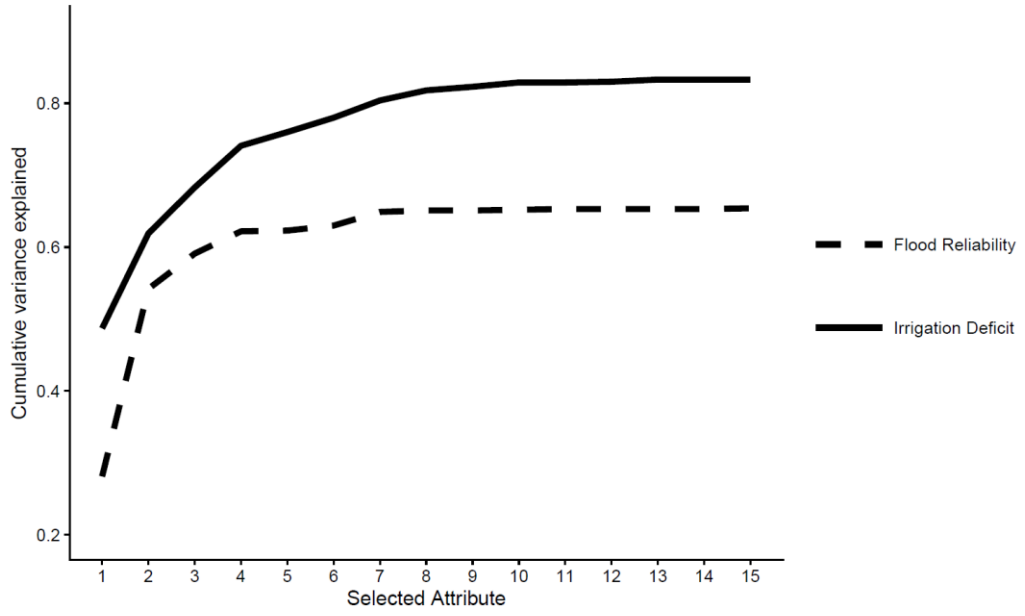


**Figure 3-3** Cumulative variance explained with each additional attribute in order of significance for the two performance criteria.

**Table 3-3** Order of critical attributes and the next most significant attribute (in italics) obtained using the PMI algorithm for the two performance criteria considered, as well as the corresponding values of cumulative variance explained (CVE).

| | **Flood Reliability** | | **Irrigation Deficit** | |
|---|---|---|---|---|
| **Order** | **Attribute** | **CVE** | **Attribute** | **CVE** |
| 1 | *F0* | 0.281 | *Ptot* | 0.487 |
| 2 | *Ptot* | 0.542 | *F0* | 0.619 |
| 3 | *PSON* | 0.591 | *TJun* | 0.683 |
| 4 | *TDec* | 0.622 | *PJJA* | 0.741 |
| | *TMar* | *0.623* | *TDec* | *0.760* |

For the irrigation deficit criterion, four attributes are selected as critical, as there is a distinct gradient change in the CVE value after the fourth attribute. It should be noted that the CVE value does not plateau until around eight attributes have been selected. However, the relative increase in information provided by these four additional variables is considered marginal, especially when traded-off against the additional difficulty of producing an 8-dimensional scenario-neutral space. As can be seen from Table 3-3, the most significant attribute for the average height criterion is *Ptot*, which is to be expected, as the irrigation deficit is primarily a function of the total amount of water in the system, which is mainly affected by total annual rainfall. The attribute having the second most significant impact on the average height criterion is *F0*, which affects the snowmelt in the year, as discussed above. This is a large source of annual inflow for the Lake Como reservoir that occurs as it becomes warmer, which is also when the irrigation demand increases. The next two critical attributes are *TJun* and *PJJA*, which both control the climate when the irrigation demand is highest, and affect the rainfall available and evaporation losses in that period, respectively. The next most significant attribute, *TDec*, impacts the timing of the snowmelt releases by shifting the days with lowest temperature. However, the increase in CVE by including this attribute is only 0.019 compared to the 0.058 increase when selecting the fourth attribute, and so it would not be worth the exponential increase in computation to include this fifth attribute in a scenario-neutral space (Table 3-3).

The CVE values in Figure 3-3 and Table 3-3 indicate that the total amount of variance explained by the critical attributes varies for the two performance criteria. Overall, the irrigation deficit criterion reaches 83% variance explained, compared to 65% for the flood reliability criterion. This suggests in both cases that there are aspects of the time series that are changing system performance that are not accounted for in the candidate attributes, and that this is more of an issue for the flood reliability criterion. Comparing the two performance criteria, the number of flooding events is subject to greater variability in the time series than the total irrigation deficit in the year, which is why it has a lower CVE value. This suggests that as system criteria become more affected by variability, there is a limit to the amount of system performance that can be predicted by a candidate set of attributes that mainly consists of annual and seasonal means.

In a real-life setting, this might necessitate the inclusion of a larger candidate set to ensure that all modes of system performance variability are accounted for in the scenario-neutral climate impact assessment, although this decision is somewhat subjective. The choice to include more attributes in a candidate set to better account for system performance also needs to be balanced by the usefulness of attributes as predictors for system performance. Given scenario-neutral spaces can be used to inform decision making, attributes that measure climate variability may not be useful to include in an assessment. However, the fact that even a broad set of 15 candidate attributes was not able to describe all of the variance in system performance highlights that the attributes considered in the vast majority of scenario neutral studies (i.e. average rainfall and average temperature) are unlikely to identify critical modes of system failure.

A comparison of the rankings of the critical attributes obtained for the different performance criteria provides a useful means of highlighting the utility and importance of using the PMI algorithm, combined with the CVE metric for identifying the climate attributes to which system performance is most sensitive. The results of the PMI analysis clearly demonstrate that different climate attributes are critical for different performance criteria, reinforcing the fundamental premise underpinning scenario-neutral climate impact assessments

that the axes of the scenario-neutral space need to be tailored to specific systems and performance measures.

### 3.4.2 Evaluating the utility of the proposed approach

#### *3.4.2.1 Testing for correct critical attributes*

The results of the test designed to evaluate the utility of the proposed approach are shown in Figure 3-4. For both the irrigation deficit and flood reliability criteria, the strength of the relationship in the residuals reduces with the addition of less significant attributes. At the fifth attribute, the residuals are close to noise, although there is still a slight relationship for the irrigation deficit criterion, which confirms the findings in Figure 3-3. This suggests that the PMI algorithm, combined with the CVE metric, were able to successfully select the critical attributes from the fifteen candidate attributes.

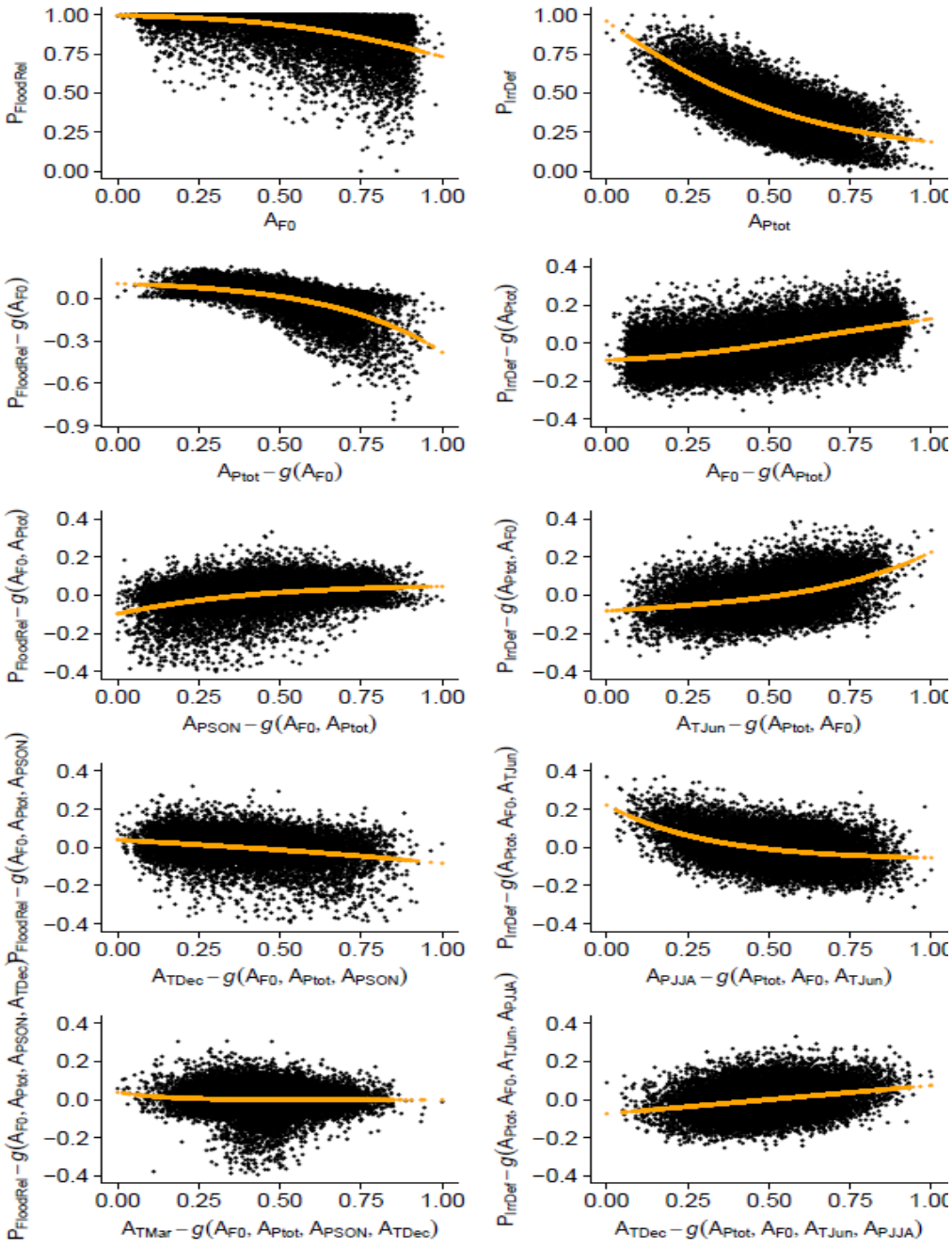


**Figure 3-4** PMI selection of the critical attributes and the next most significant attribute for the flood reliability criterion (left), and the irrigation deficit criterion (right). The residuals are plotted in black, and the MLPNN estimation is shown in orange. The function g in the axis labels represents the effect of the previous selected attributes on the attribute selected that iteration, A, and the outputs, P.

Figure 3-4 also demonstrates that a major difference between the two performance criteria across the 15,000 samples is that the flood reliability criterion is highly non-linear, whereas the irrigation deficit criterion is not (note

that both performance criteria are scaled). This is due to the proximity of the former to the upper bound of 1, caused by a number of time series that only have 0-2 flood events throughout the year. In contrast, there is irrigation deficit in every time series, and it varies much more linearly with the first selected attribute. This shows how the PMI algorithm performs well with both linear and highly non-linear relationships between attributes and system performance.

#### *3.4.2.2 Comparing scenario-neutral climate impact assessments*

For the irrigation deficit criterion there is a significant difference in the results obtained using defaults of mean precipitation and temperature as the axes of the scenario-neutral space, compared to using the critical attributes identified using the proposed approach (Figure 3-5). Using the scenario-neutral space of mean precipitation and temperature, eight of the twenty climate projections suggest that the irrigation deficit will be unacceptable due to decreases in *Ptot* and increases in *Tavg* (Figure 3-5a). In contrast, when considering the scenario-neutral space formed by the critical attributes, only three projections suggest failure of the system (projections 3, 5 and 6).

This difference is due to the fact that the scenario-neutral space formed by the critical attributes captures a more detailed range of system performance. For example, the majority of climate projections indicate a decrease in summer rainfall amounts, which increases the performance of the irrigation objective, given a fixed annual total. However, this information is not obtained when only changes in *Ptot* and *Tavg* are considered, as these do not correspond to key modes of system performance.

These results clearly show the utility of the proposed approach, as adoption of the commonly used *Ptot* and *Tavg* scenario-neutral space significantly overestimates future system failure in accordance with the irrigation deficit criterion. In addition to over-estimating the risk of future failure, consideration of the *Ptot* and *Tavg* scenario-neutral space is also likely to result in the adoption of ineffective adaptation strategies. For example, when only considering future changes in *Ptot* and *Tavg*, adaptation strategies designed to avoid system failure would be targeted at better responses to these drivers. In contrast, by gaining a

better understanding of the key modes of system failure identified with the aid of the proposed approach, adaptive strategies responding to severe decreases in summer rainfall and high increases in average June temperature can be developed.

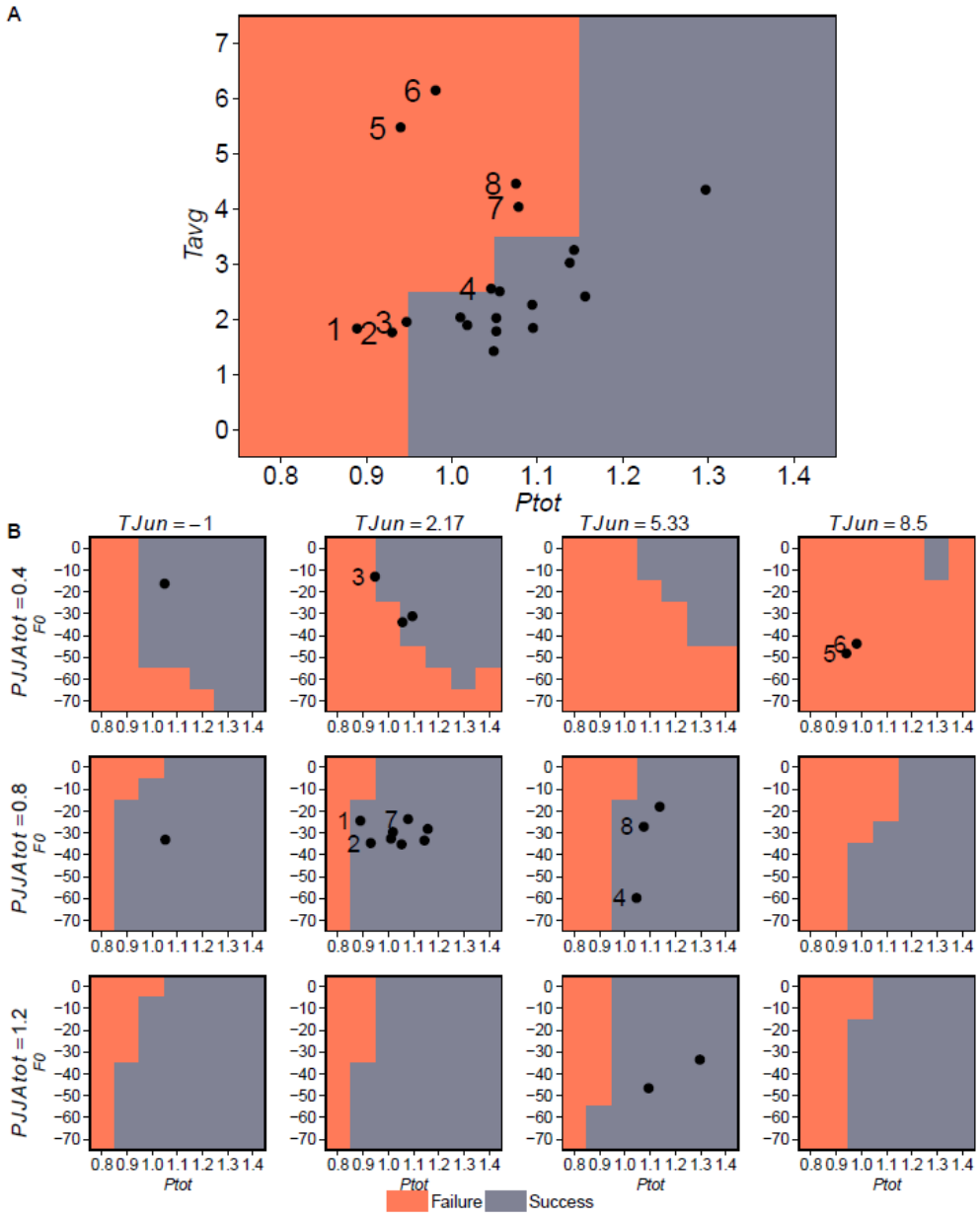


**Figure 3-5** Two scenario neutral spaces delineated by success and failure of an irrigation performance criteria with climate projections overlaid. One space is created using mean precipitation and temperature (top) and the other using the critical climate attributes (bottom).

As illustrated by Figure 3-6, the comparison between the *Ptot* and *Tavg* space and the critical climate attribute space for the flood reliability criterion also shows a difference between the number of failures obtained, although this difference is not as marked as that observed for the irrigation deficit criterion. In the *Ptot* and *Tavg* space, no projections indicate a large enough decrease in flood reliability to cause unacceptable performance, as this would require an increase in annual rainfall without an increase in average temperature. In the critical attribute space, only one projection predicts unacceptable performance (projection 1), which was the projection that was the closest to unacceptable performance in the *Ptot* and *Tavg* space. This shows that there can also be a risk of under-estimating failure when not using the critical climate attributes.

In addition, as was the case for the irrigation deficit criterion, adoption of the proposed approach was able to identify critical modes of system failure, facilitating the development of more effective adaptive strategies. For example, for the flood reliability criterion, application of the proposed approach was able to identify that high December temperatures can cause unacceptable levels of performance, which would not be possible if only *Ptot* and *Tavg* were considered. Identifying this information about system vulnerabilities is a key purpose of scenario-neutral impact assessments. These results indicate that these vulnerabilities can change with different system performance criteria and that these changes can only be identified by considering the climate attributes to which system performance is most sensitive.

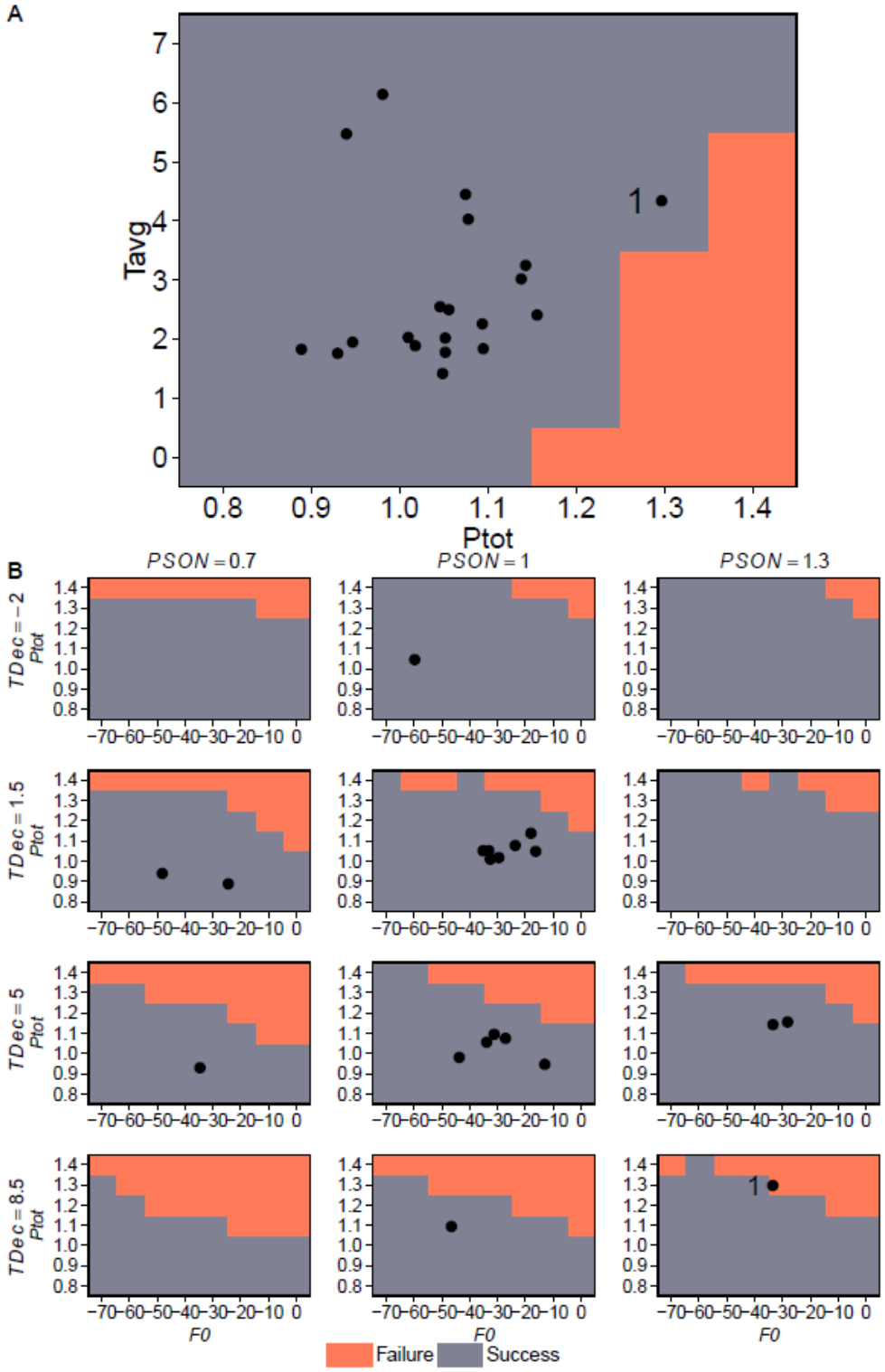


**Figure 3-6** Two scenario neutral spaces delineated by success and failure of a flood performance criteria with climate projections overlaid. One space is created using mean precipitation and temperature (top) and the other using the critical climate attributes (bottom).

## 3.5 Summary and Conclusions

This study presents a novel approach for selecting critical climate attributes for use in scenario-neutral impact assessments that require hydrometeorological time series. This is necessary, given than when large numbers of climate attributes are used to form high-dimensional scenario-neutral spaces, it is extremely difficult to generate the time series that perturb each attribute in the required way. It is because of this that most scenario-neutral studies that use hydrometeorological time series consider too few dimensions, often defaulting to examining means of precipitation and temperature. However, this makes it likely that for complex systems, the important climatic changes are not considered, and critical modes of failure can be missed. Therefore, there is a need to identify the smallest set of attributes that contains the most information about system performance. To keep the dimensionality low, this requires the consideration of both the relative importance of each attribute, and redundancy given selected attributes.

The proposed approach accomplishes this by generating synthetic hydrometeorological time series that have pre-specified values of a wide range of attributes, as determined by sampling the candidate attribute space using the inverse approach of *Guo et al.* [2018]. System performance for each of these time series is then calculated using a system model that converts the hydrometeorological time series into measures of system performance. Based on the time series and corresponding system performances, the PMI algorithm is then used to rank the candidate attributes in order of significance, which, crucially, considers any co-variability/redundancy given each selected attribute. The drop-off in additional information from each selected attribute, as measured through the CVE, can then be used to determine the critical set of attributes, such that as much of the system performance can be captured in as few dimensions as possible.

The approach was demonstrated on the Lake Como regulated reservoir system using two modelled performance criteria: irrigation deficit and flood reliability. The approach is shown to identify the smallest number of critical attributes for

each performance criterion, thereby enabling the high-resolution sampling of the critical climate scenario-neutral space, for use in scenario-neutral impact assessments like scenario-discovery and decision-scaling. This study found that one of the most commonly selected attributes in scenario-neutral studies, annual precipitation volume, was the most important attribute for only one objective: irrigation deficit. However, three additional attributes, the number of frost days, summer rainfall and average June temperature, were also needed to represent the irrigation deficit criterion. The flood reliability criterion, which was most sensitive to the number of frost days in the year, also required four attributes to describe most of the variance in system performance. This demonstrates the difficulty of knowing *a priori* which attributes are likely to be most important for a given system, and thus the need for a structured analysis for identifying the critical attributes for each system and performance criterion.

The approach is also shown to produce significantly difference outcomes when a scenario-neutral space uses critical climate attributes instead of the commonly selected annual average precipitation and temperature. Results indicate that the assessment of which projections cause acceptable and unacceptable levels of performance changes significantly for the irrigation objective, depending on the scenario-neutral space used. Further, the vulnerabilities due to climate projections are characterized as decreases in summer rainfall and increases in June temperature, in addition to changes in annual means. This suggests the use of critical climate attributes can allow more detailed planning in response to identified system vulnerabilities, which is the main purpose of scenario-neutral impact assessments.

## 3.6 Acknowledgments

The case study data used in this study are from Agenzia Regionale per la Protezione dell'Ambiente (http://ita.arpalombardia.it/ita/inde) and Consorzio dell'Adda (http://www.addaconsorzio.it/). The authors would like to thank ARPA and Eng. Bertoli from Consorzio dell'Adda for its provision. The authors would also like to thank M. Giuliani and A. Castelletti for their contribution to this research. Details of the climate time series data produced using the

*fore*SIGHT package (https://cran.r-project.org/web/packages/foreSIGHT/index.html), and case study performance data used in the results, are submitted as supporting information. Professional editor, Leticia Mooney, provided copyediting and proofreading services, according to the guidelines for editing research theses. Sam Culley was supported by an Australian Postgraduate Award.

## 3.7 References

Anghileri, D., F. Pianosi, and R. Soncini-Sessa (2011), A framework for the quantitative assessment of climate change impacts on water-related activities at the basin scale, Hydrol. Earth Syst. Sci., 15(6), 2025-2038.

Anghileri, D., A. Castelletti, F. Pianosi, R. Soncini-Sessa, and E. Weber (2013), Optimizing Watershed Management by Coordinated Operation of Storing Facilities, Journal of Water Resources Planning and Management, 139(5), 492-500.

Beachkofski, B., and R. Grandhi (2002), Improved distributed hypercube sampling, paper presented at 43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference.

Bennett, B., S. Culley, S. Westra, D. Guo, and H. R. Maier (2018), foreSIGHT: Systems Insights from Generation of Hydroclimatic Timeseries, edited, p. R package version 0.9.04.

Bergström, S., and V. Singh (1995), The HBV model, Computer models of watershed hydrology., 443-476.

Boé, J., L. Terray, F. Habets, and E. Martin (2007), Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies, International Journal of Climatology, 27(12), 1643-1655.

Broderick, C., C. Murphy, R. L. Wilby, T. Matthews, C. Prudhomme, and M. Adamson (2019), Using a Scenario-Neutral Framework to Avoid Potential Maladaptation to Future Flood Risk, Water Resources Research, 0(0).

Brown, C., and R. L. Wilby (2012), An alternate approach to assessing climate risks, Eos, Transactions American Geophysical Union, 93(41), 401-402.

Brown, C., Y. Ghile, M. Laverty, and K. Li (2012), Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, Water Resources Research, 48(9), W09537.

Bryant, B. P., and R. J. Lempert (2010), Thinking inside the box: A participatory, computer-assisted approach to scenario discovery, Technological Forecasting and Social Change, 77(1), 34-49.

Bussi, G., S. J. Dadson, C. Prudhomme, and P. G. Whitehead (2016), Modelling the future impacts of climate and land-use change on suspended sediment transport in the River Thames (UK), Journal of Hydrology, 542, 357-372.

Culley, S., B. Bennet, S. Westra, and H. Maier (2019), Generating realistic perturbed hydrometeorological time series to inform scenario-neutral climate impact assessments, Journal of Hydrology - Under review.

Culley, S., S. Noble, A. Yates, M. Timbs, S. Westra, H. Maier, M. Giuliani, and A. Castelletti (2016), A bottom-up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate, Water Resources Research, 52(9), 6751-6768.

Déqué, M. (2007), Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, Global and Planetary Change, 57(1–2), 16-26.

ETCCDI (2013), ETCCDI/CRD Climate Change Indicies, edited, Xuebin Zhang.

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 Resources Research, 52(3), 1820-1846.

Galelli, S., G. B. Humphrey, H. R. Maier, A. Castelletti, G. C. Dandy, and M. S. Gibbs (2014), An evaluation framework for input variable selection algorithms for environmental data-driven models, Environ. Modell. Softw., 62, 33-51.

Gao, L., B. A. Bryan, M. Nolan, J. D. Connor, X. Song, and G. Zhao (2016), Robust global sensitivity analysis under deep uncertainty via scenario analysis, Environ. Modell. Softw., 76, 154-166.

Giuliani, M., J. D. Herman, A. Castelletti, and P. Reed (2014), Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management, Water Resources Research, 50(4), 3355-3377.

Giuliani, M., A. Castelletti, F. Pianosi, E. Mason, and P. M. Reed (2015), Curses, Tradeoffs, and Scalable Management: Advancing Evolutionary Multiobjective Direct Policy Search to Improve Water Reservoir Operations, Journal of Water Resources Planning and Management, 04015050.

Guo, D., S. Westra, and H. R. Maier (2017a), Use of a scenario-neutral approach to identify the key hydro-meteorological attributes that impact runoff from a natural catchment, Journal of Hydrology, 554, 317-330.

Guo, D., S. Westra, and H. R. Maier (2017b), Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models, Water Resources Research, 53(1), 435-454.

Guo, D., S. Westra, and H. R. Maier (2018), An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies, Journal of Hydrology, 556, 877-890.

Herman, J. D., H. B. Zeff, P. M. Reed, and G. W. Characklis (2014), Beyond optimality: Multistakeholder robustness tradeoffs for regional water portfolio planning under deep uncertainty, Water Resources Research, 50(10), 7692-7713.

IPCC (2014), Summary for Policymakers. In: Climate Change 2014: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, IPCC.

Jacob, D., J. Petersen, B. Eggert, A. Alias, O. B. Christensen, L. M. Bouwer, A. Braun, A. Colette, M. Déqué, and G. Georgievski (2014), EURO-CORDEX: new high-resolution climate change projections for European impact research, Reg. Envir. Chang., 14(2), 563-578.

Kasprzyk, J. R., S. Nataraj, P. M. Reed, and R. J. Lempert (2013), Many objective robust decision making for complex environmental systems undergoing change, Environ. Modell. Softw., 42, 55-71.

Kay, A., S. Crooks, and N. Reynard (2014), Using response surfaces to estimate impacts of climate change on flood peaks: assessment of uncertainty, Hydrological Processes, 28(20), 5273-5287.

Klemes, V. (1986), Operational testing of hydrological simulation models, Hydrological Sciences Journal, 31(1), 13-24.

Lempert, R. J., B. P. Bryant, and S. C. Bankes (2008), Comparing algorithms for scenario discovery, RAND, Santa Monica, CA.

Li, X., H. R. Maier, and A. C. Zecchin (2015a), Improved PMI-based input variable selection approach for artificial neural network and other data driven environmental and water resource models, Environ. Modell. Softw., 65, 15-29.

Li, X., A. C. Zecchin, and H. R. Maier (2015b), Improving partial mutual information-based input variable selection by consideration of boundary issues associated with bandwidth estimation, Environ. Modell. Softw., 71, 78-96.

Maier, H. R., S. Razavi, Z. Kapelan, L. S. Matott, J. Kasprzyk, and B. A. Tolson (2019), Introductory overview: Optimization using evolutionary algorithms and other metaheuristics, Environ. Modell. Softw., 114, 195-213.

Maier, H. R., et al. (2014), Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions, Environ. Modell. Softw., 62(0), 271-299.

Mastrandrea, M., N. Heller, T. Root, and S. Schneider (2010), Bridging the gap: linking climate-impacts research with adaptation planning and management, Climatic Change, 100(1), 87-101.

May, R. J., H. R. Maier, G. C. Dandy, and T. M. K. G. Fernando (2008), Non-linear variable selection for artificial neural networks using partial mutual information, Environ. Modell. Softw., 23(10), 1312-1326.

Nazemi, A., and H. Wheater (2014), Assessing the vulnerability of water supply to changing streamflow conditions, Eos, Transactions American Geophysical Union, 95(32), 288-288.

Nazemi, A., H. S. Wheater, K. P. Chun, and A. Elshorbagy (2013), A stochastic reconstruction framework for analysis of water resource system vulnerability to climate-induced changes in river flow regime, Water Resources Research, 49(1), 291-305.

Pianosi, F., K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, and T. Wagener (2016), Sensitivity analysis of environmental models: A systematic review with practical workflow, Environ. Modell. Softw., 79, 214-232.

Prudhomme, C., A. L. Kay, S. Crooks, and N. Reynard (2013a), Climate change and river flooding: Part 2 sensitivity characterisation for British catchments and example vulnerability assessments, Climatic Change, 119(3-4), 949-964.

Prudhomme, C., S. Crooks, A. L. Kay, and N. Reynard (2013b), Climate change and river flooding: part 1 classifying the sensitivity of British catchments, Climatic Change, 119(3-4), 933-948.

Prudhomme, C., R. L. Wilby, S. Crooks, A. L. Kay, and N. S. Reynard (2010), Scenario-neutral approach to climate change impact studies: Application to flood risk, Journal of Hydrology, 390(3-4), 198-209.

Ray, P. A., L. Bonzanigo, S. Wi, Y.-C. E. Yang, P. Karki, L. E. García, D. J. Rodriguez, and C. M. Brown (2018), Multidimensional stress test for hydropower investments facing climate, geophysical and financial uncertainty, Global Environmental Change, 48, 168-181.

Richardson, C. W. (1981), Stochastic simulation of daily precipitation, temperature, and solar radiation, Water resources research, 17(1), 182-190.

Richardson, C. W., and D. A. Wright (1984), WGEN: A model for generating daily weather variables.

Scrucca, L. (2013), {GA}: A Package for Genetic Algorithms in {R}, Journal of Statistical Software, 53(4), 1-37.

Sharma, A. (2000), Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1 — A strategy for system predictor identification, Journal of Hydrology, 239(1), 232-239.

Shortridge, J. E., and S. D. Guikema (2016), Scenario Discovery with Multiple Criteria: An Evaluation of the Robust Decision-Making Framework for Climate Change Adaptation, Risk Analysis, 36(12), 2298-2312.

Singh, R., T. Wagener, R. Crane, M. Mann, and L. Ning (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(4), 3409-3427.

Specht, D. F. (1991), A general regression neural network, IEEE transactions on neural networks, 2(6), 568-576.

Stein, M. (1987), Large sample properties of simulations using Latin hypercube sampling, Technometrics, 29(2), 143-151.

Taner, M. Ü., P. Ray, and C. Brown (2017), Robustness-based evaluation of hydropower infrastructure design under climate change, Climate Risk Management, 18, 34-50.

Turner, S. W. D., D. Marlow, M. Ekström, B. G. Rhodes, U. Kularathna, and P. J. Jeffrey (2014), Linking climate projections to performance: A yield-based decision scaling assessment of a large urban water resources system, Water Resources Research, 50(4), 3553-3567.

Weiß, M. (2011), Future water availability in selected European catchments: a probabilistic assessment of seasonal flows under the IPCC A1B emission scenario using response surfaces, Natural Hazards and Earth System Sciences, 11(8), 2163-2171.

Westra, S., M. Thyer, M. Leonard, D. Kavetski, and M. Lambert (2014), A strategy for diagnosing and interpreting hydrological model nonstationarity, Water Resources Research, 50(6), 5090-5113.

Wetterhall, F., L. Graham, J. Andreasson, J. Rosberg, and W. Yang (2011), Using ensemble climate projections to assess probabilistic hydrological change in the Nordic region, Natural Hazards and Earth System Sciences, 11(8), 2295-2306.

Whateley, S., and C. Brown (2016), Assessing the relative effects of emissions, climate means, and variability on large water supply systems, Geophysical Research Letters, 43(21), 11,329-311,338.

Whateley, S., S. Steinschneider, and C. Brown (2014), A climate change range-based method for estimating robustness for water resources supply, Water Resources Research, 50(11), 8944-8961.

Wu, W., G. C. Dandy, and H. R. Maier (2014), Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling, Environ. Modell. Softw., 54, 108-127.

# Chapter 4

## *Scenario-neutral climate impact assessments: pitfalls, diagnostics and solutions (Paper 3)*

Sam Culley, Seth Westra, Bree Bennett and Holger Maier

# Statement of Authorship

| | |
|---|---|
| Title of Paper | Scenario-neutral climate impact assessments: pitfalls, diagnostics and solutions |
| Publication Status | ☐ Published ☐ Accepted for Publication<br>☐ Submitted for Publication ☑ Unpublished and Unsubmitted work written in manuscript style |
| Publication Details | Planned submission to the journal Water Resources Research. |

## Principal Author

| | |
|---|---|
| Name of Principal Author (Candidate) | Sam Culley |
| Contribution to the Paper | Contributed to the conception and design of the project, performed analysis, interpreted the results and wrote manuscript. |
| Overall percentage (%) | 75% |
| 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 18/04/19 |

## 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 | Contributed to the conception and design of the project, analysis and interpretation of the research data, and editing the manuscript |
| Signature | Date 18-04-2019 |

| | |
|---|---|
| Name of Co-Author | Bree Bennett |
| Contribution to the Paper | Contributed to the conception and design of the project, analysis and interpretation of the research data, and editing the manuscript |
| Signature | Date 18/4/19 |

| | | | |
|---|---|---|---|
| Name of Co-Author | Holger Maier | | |
| Contribution to the Paper | Contributed to the conception and design of the project; analysis and interpretation of the research data, and editing the manuscript | | |
| Signature | | Date | 18/4/19 |

# Abstract

Results from scenario-neutral climate impact assessments often do not align with those from a scenario-led analysis. Regardless of the validity or otherwise of scenario-led projections, it is argued here that differences between the approaches arise because modes of change that are potentially critical for system performance were not taken into account in the scenario-neutral analysis. This potentially undermines the scenario-neutral approach, as the identification of system vulnerabilities is its key motivation. This research presents four potential pitfalls associated with implementing the scenario-neutral approach, with the impact of each pitfall demonstrated by comparing to a benchmark implementation using state-of-the-art methods. The first two pitfalls lead to missing key modes of change relevant to a system through either defaulting to mean precipitation and temperature or not casting a wide enough net in an analytical approach. The remaining two pitfalls affect the generation of hydrometeorological time series, through either using insufficiently flexible perturbation methods to adequately stress-test a system or inadvertently changing other aspects of a time series that are not the focus of the analysis. For all cases, comparisons with projections from scenario-led methods represent a valuable diagnostic of the performance of scenario-neutral approaches, and directions are suggested for improving scenario-neutral analyses based on this diagnostic approach. The Lake Como reservoir is used as a case study, with results demonstrating that only by correctly implementing all the key requirements of the scenario-neutral methodology will the method identify all primary system vulnerabilities.

## 4.1 Introduction

Scenario-neutral impact assessments are proving to be an effective tool for supporting decision making under climate change uncertainty. Initially, *Prudhomme et al.* [2010] formalized the approach as a sensitivity analysis on a system impact model and provided a conceptual framework as an alternative to a scenario-led approach. This was followed by *Brown and Wilby* [2012], who provided a discussion of the key components of this alternate approach, and how it overcomes the limitations of a traditional scenario-led approach. Distilled to three main points, a scenario-neutral approach should: i) identify performance measures and resulting objectives for a system, ii) stress test a system under a wide range of climatic and non-climatic changes to identify hazards, and iii) assess the plausibility of those hazards occurring using model projections and other climatic lines of evidence [*Brown et al.*, 2012]. As part of the approach, a system is stress-tested against a number of climate variables (typically two), where system performance is shown in response to changes in climate from historical conditions via a scenario-neutral space (also referred to as a response surface, or an exposure space) [*Culley et al*., 2016; *Broderick et al*., 2019].

The scenario-neutral space serves as the basis for decision making because it is a simple visual representation of system performance that enables performance degradation and/or system failure boundaries to be characterized by changes in climate. It also extends the impact analysis beyond the range indicated by climate projections, reducing the reliance on their accuracy [*Raäisaänen*, 2007]. The bounds of the analysis are typically selected to extend slightly beyond the climate projections [*Culley et al*., 2019b], but if there is reason to believe that projections systematically underestimate or overestimate key modes of change, then this information can also be used to inform the bounds. Decision alternatives can be mapped on this space, showing how the outcome of different decisions perform across a range of climate conditions, allowing for the identification of tipping points for adaptation planning *[Haasnoot et al*., 2013; *Kwakkel et al*., 2016] or robustness calculations [*Whateley et al*., 2014; *Steinschneider et al*., 2015; *McPhail et al*., 2018]. Climate projections and other lines of evidence are then overlaid onto scenario-neutral spaces to assess the

likelihood any identified system vulnerabilities will occur. This information can be used to determine a system's maximum adaptive capacity [*Culley at al*., 2016] or to decide on a management strategy, where, for example, the strategy that is most robust to the set of projections is selected [*Whateley et al*., 2014].

Regardless of the decision-making technique, it is implied that system performance in response to the projections is adequately represented by the scenario-neutral space. However, this may not always be the case, with a recent study showing that system performance calculated directly in response to climate projections differs from the estimate given by a scenario-neutral space [*Taner et al*., 2017]. The potential discrepancy between scenario-neutral and scenario-led approaches calls into question whether the scenario-neutral approach adequately identifies the primary system vulnerabilities, and this in turn has the potential to undermine any actions resulting from scenario-neutral impact assessments.

Given that the same system model is used to estimate system performance, the difference in results between scenario-neutral and scenario-led approaches must be due to differences between the climate time series used. Scenario-led approaches select relevant climate projections for the system of interest and use system models to simulate the response [*Arnell*, 2004; *Brekke et al*., 2009; *Vano et al*., 2010; *Wilby and Dessai*, 2010; *Anghileri et al*., 2011]. Scenario-neutral approaches, however, identify climate drivers and perturb historical conditions to make time series with changes in those drivers. When these scenario-neutral time series are compared with a projection purportedly indicating the same changes, any differences in performance are therefore a result of differences in the two climate time series beyond those changes that are explicitly considered. Therefore, in order to understand any limitations of the scenario-neutral approach, there is a need to understand what aspect of the climate time series causes the difference in performance. It should be noted that the validity of the system model itself is also of concern, and has been demonstrated to affect the scenario-neutral spaces created [*Broderick et al*., 2019], but accounting for invalid system models will not reconcile differences between scenario-led and scenario-neutral assessments.

Several limitations with the scenario-neutral approach have been identified in literature, which could contribute to the disagreement in results. Firstly, it was stated in the initial applications of the scenario-neutral approach that the changes investigated should be those to which a system is most sensitive, and yet many studies default to changes in annual temperature and precipitation [*Weiß*, 2011; *Wetterhall et al*., 2011; *Turner et al*., 2014; *Whateley et al*., 2014; *Culley et al*., 2016]. There are many other statistical measurements of climate change that should be considered, which for the purposes of scenario-neutral analyses are referred to as climate 'attributes' (e.g. 99th percentile precipitation, measures of precipitation seasonality, the number of frost days in the year, the number of wet days). Some studies have included a formal sensitivity test in the scenario-neutral analysis [*Prudhomme et al*., 2013; *Ray et al*., 2018], but the number of climate attributes considered is still often restricted to a small number. Not stress-testing a system against the most sensitive climatic changes can cause modes of failure to be missed [*Culley et al*., 2019b].

An additional concern with the scenario-neutral approach is that the different techniques used to create the climate time series for scenario-neutral spaces can cause different system responses [*Keller et al*., 2018]. This is a concern because the analysis is used to indicate how a system will respond to specific climate conditions. Incorrectly linking system performance to changes in climate can lead to inaccurate performance maps and delineation of failure boundaries, undermining any decision reliant on the scenario-neutral spaces.

Although these limitations have been identified, their direct impact on the validity of the scenario-neutral approach has not been investigated. Therefore, there is a need for a formal diagnostic approach to identify when these limitations are present, and their likely effect. This will enable more successful applications of the scenario-neutral approach to climate impact assessments. The purpose of this paper is therefore to: i) identify common pitfalls with the scenario-neutral approach that may affect the validity of the results, and ii) demonstrate the effect of each pitfall. A diagnostic is presented which directly compares the results of a scenario-neutral analysis with a scenario-led analysis. A benchmark

implementation that uses state-of-the-art methods for undertaking a scenario-neutral approach is presented, to assess how similar scenario-led and scenario-neutral approaches can be. Alterations to these methods, representing some of the potential pitfalls associated with implementations of the scenario-neutral approach, are then presented and compared to the benchmark implementation, to demonstrate the impact of these pitfalls.

The remainder of the paper is organized as follows. Section 4.2 discusses the key requirements of generating time series to support scenario-neutral studies, and presents four potential pitfalls that can prevent the satisfaction of these requirements. Section 4.3 then presents the methods used to demonstrate the effect of each pitfall, with the results relative to a benchmark implementation presented in Section 4.4. A discussion of the utility of the diagnostic presented in this paper is given in Section 4.5, and conclusions are presented in Section 4.6.

## 4.2 Pitfalls of time series generation for scenario-neutral studies

In order to identify the potential pitfalls in the approach to generating climate time series for scenario-neutral studies, the two major steps in this approach are first presented. These are to: i) identify the critical climate attributes to be used in the analysis (where the critical climate attributes are defined to be the attributes that have significant influence on system performance), and ii) generate climate time series to systematically adjust the critical attributes as an input to the system stress test (Figure 4-1).

Two general methods are used for identifying the critical climate attributes (Step 1). The first method is to use *a priori* knowledge, typically from stakeholders, to directly select critical climate attributes without formally testing their importance on the system in question [*Culley et al.*, 2016; *Broderick et al.*, 2019]. This can lead to Pitfall 1, in which the analysis defaults to using mean precipitation and mean temperature as the critical climate attributes, as has been done in the majority of previous scenario-neutral studies [*Weiß*, 2011; *Wetterhall et al.*, 2011; *Turner et al.*, 2014; *Whateley et al.*, 2014; *Culley et al.*, 2016].

Although most water resource systems are sensitive to changes in these attributes, there is no guarantee that they are the only critical attributes, potentially leading to key modes of system failure being missed from the analysis. In contrast, the second method uses analytical techniques to identify the most sensitive attributes from a larger candidate set [*Ray et al.*, 2018; *Culley et al.*, 2019b]. This method still requires *a priori* understanding of likely system sensitivities to select a candidate set of attributes; however the use of a subsequent filtering stage means that the candidate set can be quite large (e.g. 15 attributes in [*Culley et al.*, 2019b]), so the likelihood of missing key system sensitivities is reduced. Nevertheless, the complexity of many water resource systems means that it is still possible to miss important attributes in the candidate set, and Pitfall 2 reflects the omission of important candidate attributes from the analysis.

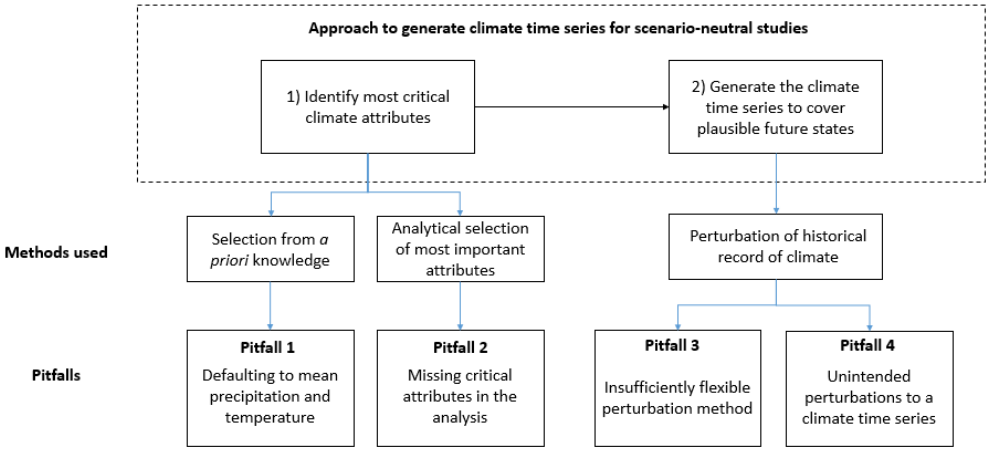


**Figure 4-1** Key steps in the generation of climate time series for scenario-neutral studies, common methods used and resulting pitfalls.

Once the critical attributes are identified, it is necessary to generate climate time series that reflect those critical attributes (Step 2) as inputs to a system model. Pitfall 3 can occur when the perturbation technique is not sufficiently flexible to perturb the most critical climate attributes. For example, the manual scaling of historical time series [*Prudhomme et al.*, 2010] generally only allows for perturbations to the means of climate variables, although some changes in seasonality can be achieved if monthly scaling factors are used. In contrast, the

use of stochastic weather generators enables modification of a much wider set of attributes [*Steinschneider and Brown,* 2013; *Guo et al.*, 2018; *Keller et al.*, 2018]; however depending on the choice of stochastic generator it may still not be possible to achieve certain combinations of attribute changes. Pitfall 4 represents the case where the perturbation technique used not only changes the *desired* attributes of the climate time series, but also inadvertently changes *other* attributes that have an effect on system performance. For example, calibrating a stochastic generator to a change in the mean may lead to unintended changes to other statistics (e.g. the seasonality, intermittency and/or extremes), which if not controlled for could lead to misleading results in the scenario-neutral analysis [*Culley et al.*, 2019a].

## 4.3 Methods

### 4.3.1 Overview

**In order to demonstrate the potential impact of the four key pitfalls of the scenario-neutral scenario-neutral approach outlined in Section 2, five alternative scenario-neutral implementations are presented (Figure 4-2,**

Table 4-1). Each implementation is demonstrated using the Lake Como reservoir in Italy, with an irrigation supply objective (Section 4.3.2). First, a benchmark implementation is presented, which uses state-of-the-art methods for the identification of the most critical climate attributes (benchmark method 1) and the generation of the climate perturbed time series (benchmark method 2). Benchmark method 1 consists of the method of *Culley et al.* [2019b], as it has been demonstrated to identify the critical climate attributes for a system. The method works by examining a wide range of climate attributes and then reducing these attributes to a subset of critical attributes using the partial mutual information (PMI) algorithm [*Sharma*, 2000]. Benchmark method 2 consists of the method of *Guo et al.* [2018], which presents an inverse approach to time series generation, capable of meeting the required perturbations in a wide range of climate attributes. This approach has been demonstrated to perturb climate time series across the scenario-neutral space while maintaining realism [*Culley*

*et al.*, 2019a] (Section 4.3.3).

**After establishing the benchmark implementation, four alternative 'pitfall' implementations are presented, each changing one aspect of the benchmark to demonstrate demonstrate the potential impact of each pitfall (Section 4.3.4). The first pitfall involves involves selecting mean precipitation and temperature as critical attributes. The second second pitfall is demonstrated by only considering precipitation attributes and holding holding other climate variables constant, reflecting the practice used in a number of of scenario-neutral studies where temperature is held constant [*Prudhomme et al.*, 2010; *al.*, 2010; *Broderick et al.*, 2019] as this variable is often considered of secondary importance importance to precipitation for many water resource systems. The third pitfall is demonstrated by using a less flexible stochastic weather generator as part of the inverse inverse approach to generating climate time series [*Guo et al.*, 2018] resulting in a limitation a limitation in the extent to which certain attributes can be perturbed. Lastly, the fourth fourth pitfall is demonstrated by only monitoring the changes in the critical climate climate attributes for the system when generating climate time series (**

Table 4-1), resulting in potentially unrealistic changes in other attributes that are not monitored.

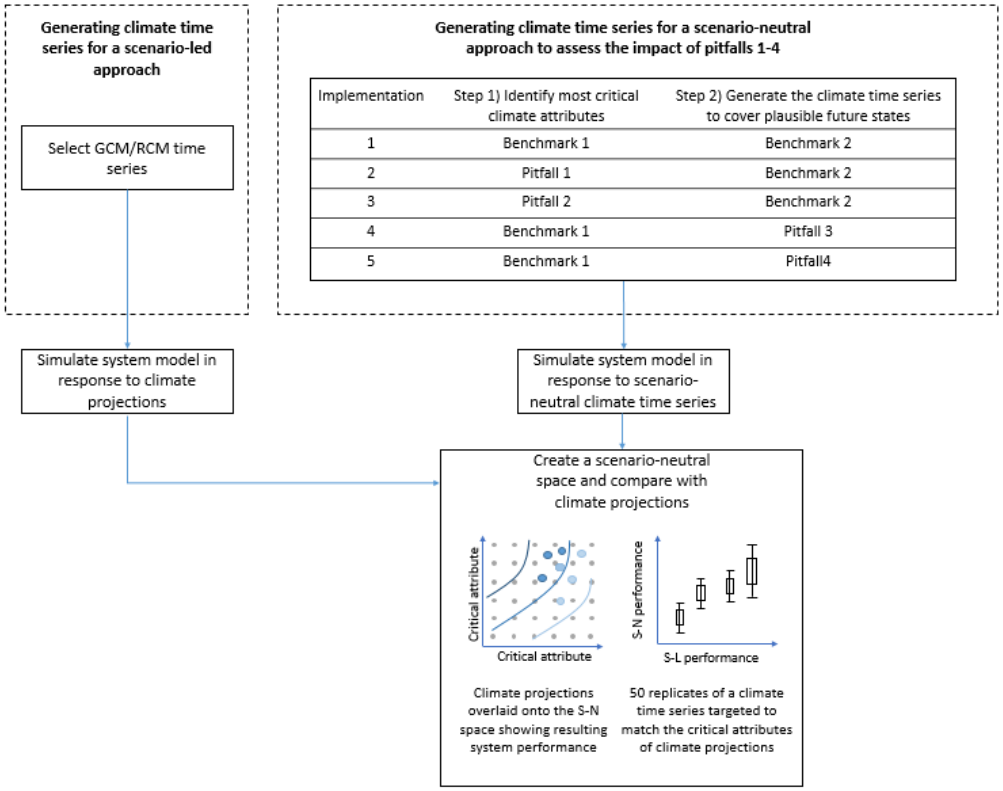


**Figure 4-2** The scenario-neutral framework, selected methods and pitfalls. Five implementations of the scenario-neutral approach are presented, to demonstrate the effect of four pitfalls. A comparison with a scenario-led approach is used as a diagnostic.

**Table 4-1** Methods used in the scenario-neutral implementations.

| | **Method** | **Reasons for selection** |
|---|---|---|
| Benchmark 1 | Analytical approach to reduce dimensions with credible candidate attribute set | Demonstrated to select critical attributes |
| Benchmark 2 | Inverse approach to time series generation | Demonstrated to perturb specific attributes for climate time series |
| Pitfall 1 | Stress test against mean annual precipitation and temperature | These attributes are the most common defaults |
| Pitfall 2 | Benchmark 1, but a limited candidate set of only precipitation attributes | Some studies keep other climate variables constant and perturb only precipitation |
| Pitfall 3 | Benchmark 2, but perturb time series with an insufficiently flexible annual precipitation weather generator | The choice of perturbation approach (e.g. scaling of historical time series or using a weather generator) can have a significant bearing on capacity to simulate changes to critical attributes |
| Pitfall 4 | Benchmark 2, but ignore secondary changes to climate attributes when creating scenario-neutral spaces | Perturbing time series to change critical attributes could lead to unintended changes to other aspects of the climate time series |

The recommended diagnostic approach for assessing the presence of one or more of the pitfalls is to compare the scenario-neutral time series with scenario-led time series (in this case based on downscaled GCM projections), as shown in Figure 4-2. This comparison can be achieved in two ways. The first is an overlay of the climate projections onto the scenario-neutral space, while also indicating system performance in response to climate projections. The second is to directly produce climate time series that match the critical attributes of the climate projections and compare across 50 replicates using the stochastic weather generators. Both these comparisons can determine if there are aspects of a climate time series (i.e. specific climate attributes) that affect system performance but are not included in the scenario-neutral analysis. For example, two precipitation time series that have the same annual total precipitation could provide different system performance when simulated, as a result of differences in seasonal patterns, intermittency or the magnitude and/or frequency of extreme events.

It is important to note the use of scenario-led projections as the recommended diagnostic does not require the assumption of the validity of those projections. Rather, a situation whereby a system responds differently to time series with the same changes in the attributes that are considered in the scenario neutral space (e.g. identical changes in annual average temperature and precipitation) indicates that the system is likely to be sensitive to a climate attribute that has not been considered by the scenario-neutral analysis. These sensitivities are intrinsic to the system (or at least the system model used for simulating system performance) rather than to the climate, so that the presence of errors or biases in climate projections does not negate the presence of that sensitivity.

### 4.3.2 Case study data and system models

The Lake Como case study comprises a reservoir in an alpine region of Northern Italy that supplies a major component of irrigation demand [*Anghileri et al.*, 2011; *Culley et al.*, 2016]. The operation of the reservoir is driven by two competing objectives: the prevention of floods for the nearby city of Como, and irrigation supply. There is a large snowmelt component to the annual inflow,

which historically arrives in late spring. The reservoir operation needs to store this inflow for the high summer demand while also avoiding flood events.

A baseline climate record of 1965-1980 is used in this study, with daily recordings of precipitation and temperature [see *Culley et al.*, 2016]. This is taken to be a stationary period not yet affected by climate change. Climate projections for the Lake Como area are available as part of the EUROCORDEX project [*Jacob et al.*, 2014], where ten different combinations of GCM and RCM are used to create climate projections for both RCP 4.5 and 8.5 (

Table 4-2). The time window of 2040-2060 is used in this study as a basis to calculate projected change for the case study relative to the stationary reference period.

**Table 4-2** RCP/GCM/RCM combinations and reference number for climate projections used in this study.

| Model Reference # | GCM | RCM | RCP |
|---|---|---|---|
| 1 | CM5 (CNRM CERFACS) | CCLM4 (CLMcom) | 4.5 |
| 2 | CM5 (CNRM CERFACS) | CCLM4 (CLMcom) | 8.5 |
| 3 | CM5 (CNRM CERFACS) | RCA4 | 4.5 |
| 4 | CM5 (CNRM CERFACS) | RCA4 | 8.5 |
| 5 | EARTH (ICEC) | CCLM4 (CLMcom) | 4.5 |
| 6 | EARTH (ICEC) | CCLM4 (CLMcom) | 8.5 |
| 7 | EARTH (ICEC) | HIRHAM5 (DMI) | 4.5 |
| 8 | EARTH (ICEC) | HIRHAM5 (DMI) | 8.5 |
| 9 | EARTH (ICEC) | RACMO22E (KNMI) | 4.5 |
| 10 | EARTH (ICEC) | RACMO22E (KNMI) | 8.5 |
| 11 | EARTH (ICEC) | RCA4 | 4.5 |
| 12 | EARTH (ICEC) | RCA4 | 8.5 |
| 13 | CanESM2 (CCCma) | RCA4 | 4.5 |
| 14 | CanESM2 (CCCma) | RCA4 | 8.5 |
| 15 | MIROC | RCA4 | 4.5 |
| 16 | MIROC | RCA4 | 8.5 |
| 17 | NCC | RCA4 | 4.5 |
| 18 | NCC | RCA4 | 8.5 |
| 19 | NOAA | RCA4 | 4.5 |
| 20 | NOAA | RCA4 | 8.5 |

The hydrological component is modelled by a lumped HBV model [*Bergström and Singh*, 1995], which takes daily precipitation and temperature as inputs,

while evaporation and snowmelt are calculated internally. The reservoir system is modelled by a mass storage balance equation, and its releases are controlled with a radial basis function [*Giuliani et al.*, 2014]. The release function takes the day of year and height of the reservoir as inputs to calculate the release, and has been calibrated to a historical operation period to match performance in both flood prevention and irrigation supply [*Culley et al.*, 2016]. The baseline climate period of 1965-1980 has been used to calibrate both the hydrological model [*Anghileri et al.*, 2011], and the reservoir storage operation [*Culley et al.*, 2016].

The performance criterion of irrigation deficit is used, which is calculated as an annual average over the simulation period, where the deficit is calculated by comparing the irrigation demand with the water release available for irrigation at a daily time step. A fixed seasonal demand pattern is applied to each year of simulation. The water available is calculated by subtracting a 5kL daily minimum environmental release from the total daily release. With the seasonal demand pattern, this performance criterion is sufficiently complex to be influenced by some seasonality of climate change, which is necessary to demonstrate how some pitfalls can be present for more complex applications. However, a potential limitation of this study is that the properties of the system model are unchanging (such as demand), which means it is assumed that the system model is accurately representing how the system will respond to different climate forcings.

### 4.3.3 State-of-the-art (benchmark) implementation of scenario-neutral approach

**As discussed in relation to**

Table 4-1, a state-of-the-art implementation of the scenario neutral approach is used as a benchmark against which the impact of the four pitfalls discussed in Section 4.2 can be assessed. Details of the methods for the identification of the most critical climate attributes and for the generation of the climate-perturbed hydro-meteorological time series are given in Sections 4.3.3.1 and 4.3.3.2.

#### *4.3.3.1 Identification of most critical climate attributes*

The partial mutual information (PMI) algorithm presented by *Culley et al.* [2019b] is used as the state-of-the-art benchmark approach to identifying the critical climate attributes, as it has been shown to identify the smallest number of attributes that have a significant impact on system performance. This method begins by specifying a range of possible 'candidate' climate attributes to which a system could be sensitive. A representative low-resolution sample of these candidate attributes is generated, with a climate time series generated to match each sample. With system performance calculated in response to each time series, the PMI algorithm [*Sharma*, 2000; *Li et al.*, 2015b; *Li et al.*, 2015a] is used to rank the candidate attributes in order of significance. The cumulative fraction of variance explained (CVE) in the irrigation deficit is calculated for each additional attribute to indicate the benefit of including that attribute in the critical set. Critical attributes are selected by considering the trade-off between an increase in the representation of key climatic drivers of system performance and the cost (not only in terms of computational load but also in terms of interpretability) of including an additional dimension in the high-resolution scenario-neutral spaces.

To implement the PMI approach, it is recommended that the candidate set is based on *a priori* system understanding, but should err on the side of including, rather than excluding, attributes to avoid the possibility of missing critical attributes. The set of candidate attributes used in this study is presented in Table 4-3, covering various statistical representations of precipitation and temperature—the inputs to the case study system model. Seven attributes represent annual-scale statistics, including means, extremes, and precipitation persistence. Eight seasonal attributes are also included, four each for precipitation and temperature. These attributes are included in anticipation of their effect on system performance, given the seasonal demand pattern the reservoir is attempting to supply.

**Table 4-3** List of candidate attributes.

| Attribute | Description | Units |
|---|---|---|
| *Ptot* | Total annual precipitation volume | mm |
| *nWet* | Number of wet days in the year | Days |
| *P99* | Volume of the 99$^{th}$ percentile precipitation event | mm |
| *avgWSD* | Average duration of consecutive wet days | Days |
| *PDJFtot* | Total December, January, February precipitation volume | mm |
| *PMAMtot* | Total March, April, May precipitation volume | mm |
| *PJJAtot* | Total June, July, August precipitation volume | mm |
| *PSONtot* | Total September, October, November precipitation volume | mm |
| *TMar* | Average temperature in March | °C |
| *TJun* | Average temperature in June | °C |
| *TSep* | Average temperature in September | °C |
| *TDec* | Average temperature in December | °C |
| *Tavg* | Annual average temperature | °C |
| *F0* | Number of days below zero in the year | days |
| *Trng* | Range between the annual 95$^{th}$ and 5$^{th}$ percentile temperature | °C |

Bounds are placed on the candidate attributes by considering the range indicated by projections from the 20 model combinations described in

Table 4-2, for the period 2040-2060. The calculated range is shown in Table 4-4, with precipitation attributes showing fractional change and temperature attributes showing additive change. The bounds for sampling are selected by extending slightly beyond this range. A Latin hypercube method is used to create 15,000 samples within the bounds, and *fore*SIGHT [*Bennett et al*., 2018], an R software package, is used to create each climate time series to best match the sampled attributes. The attributes of the created time series are used in the PMI algorithm along with the system model output in response to simulating each time series. More details on using *fore*SIGHT are given in Section 4.3.3.2.

**Table 4-4** Range and bounds of candidate attributes.

| Attribute | Range | Bounds |
|---|---|---|
| *Ptot* | 0.89 −1.30 | 0.8 − 1.4 |
| *nWet* | 0.91 − 1.03 | 0.8 − 1.1 |
| *P99* | 0.92 − 1.29 | 0.8 − 1.4 |
| *avgWSD* | 0.64 − 0.86 | 0.6 − 1.0 |
| *PDJFtot* | 1.22 − 2.10 | 1.0 − 2.2 |
| *PMAMtot* | 0.97 − 1.45 | 0.9 − 1.5 |
| *PJJAtot* | 0.43 − 1.20 | 0.4 − 1.2 |
| *PSONtot* | 0.70 − 1.21 | 0.7 − 1.3 |
| *TMar* | 0.11 − 8.85 | 0.0 − 9.0 |
| *TJun* | -0.42 − 8.41 | -1.0 − 8.5 |
| *TSep* | 1.37 − 11.25 | 0.0 − 11.5 |
| *TDec* | -1.53 − 8.25 | -2.0 − 8.5 |
| *Tavg* | 1.42 − 6.13 | 0.0 − 6.5 |
| *F0* | -59.8 − -13.09 | -70 − 0.0 |
| *Trng* | -1.43 − 8.51 | -2.0 − 9.0 |

The outcome of following the PMI-based process for the case study is shown in Table 4-5, which also provides the cumulative variance explained—a measure of the extent to which the set of attributes describes the full variability in sampled system model performance. The total annual precipitation (*Ptot*) is identified as the most significant climate attribute. The number of frost days (*F0*) comes second, with this attribute describing the days of temperature below zero, which affects the snow storage throughout the year, as well as the timing of the first major snowmelt release into the reservoir. The temperature in June (*TJun*) and the precipitation total in June, July, August (*PJJAtot*) are chosen next, showing the summer season has the biggest impact on irrigation deficit, which is likely due to the irrigation demand being highest in summer.

**Table 4-5** Critical climate attributes from the candidate set.

| Rank | Attribute | CVE |
|---|---|---|
| 1 | *Ptot* | 0.487 |
| 2 | *F0* | 0.619 |
| 3 | *TJun* | 0.683 |
| 4 | *PJJAtot* | 0.741 |

#### *4.3.3.2 Generation of climate-perturbed hydrometeorological time series*

Having identified the critical climate attributes for the scenario-neutral analysis,

the next step is to simulate climate time series that capture changes in those attributes. The inverse approach to time series creation using a seasonal weather generator is used as the second benchmark approach [*Guo et al.*, 2017a; 2018; *Culley et al.*, 2019a], as it has the ability to generate hydrometeorological time series that exhibit a wide range of desired climate attributes while maintaining the physical realism of the time series. This approach uses formal optimization techniques [*Maier et al.*, 2014; *Maier et al.*, 2019] to identify the set of stochastic weather generator parameters that provides a climate time series with the desired attributes. The optimization formulation aims to minimize the difference between a target set of attributes and the attributes measured from the time series created with the stochastic generator. This is typically accomplished using a Euclidean distance function [*Guo et al.*, 2017a; 2018].

The *fore*SIGHT software provides the stochastic models and the optimization formulation necessary to implement the inverse approach. Seasonal precipitation and temperature weather generators are used to perturb some of the seasonal attributes that might affect snowmelt dynamics. Crucially, this software enables the prioritization of some attributes over others when creating targeted perturbations in historical climate. This is because, to maintain physical realism of the generated time series, the optimization works by finding stochastic generator parameters and time series that match the targeted changes to attributes, while keeping all other aspects of the time series as close as possible to the historical record. This has the potential to lead to infeasible combinations—for example it is not possible to increase the total annual rainfall while keeping both the number of wet days and the average rainfall amount per wet day at historical levels. The approach taken by the *fore*SIGHT software is presented in *Culley et al.* [2019a], and uses a penalty functions to prioritize some attributes over others. This ensures that, wherever possible, the inverse approach identifies time series with the specified critical attribute combinations, while simultaneously ensuring that the remaining (non-critical) attributes are as close to historical conditions as possible.

As mentioned in Section 4.3.1, two approaches are taken to compare the

scenario-neutral approach with the scenario-led approach. The first is to overlay scenario-led time series onto a scenario-neutral space. To generate hydrometeorological time series for a scenario-neutral space, target changes in the critical attributes are identified at regular intervals between the bounds in Table 4-4 (Section 4.3.3.1), with all other candidate attributes held at historical levels. Each target is created with one optimization seed, and the default optimization operators of the *fore*SIGHT package are used. The scenario-neutral spaces are created using five weather generator replicates for each target and averaging the system response of each. All scenario-neutral time series created in this study are 21 years long, to match the length of climate projections used in the scenario-led analysis.

The second approach is to create scenario-neutral time series for the direct comparison with the scenario-led approach. To generate these hydrometeorological time series, the values of the critical attributes in the scenario-led time series are used as target perturbations, with all other attributes held at historical levels. The direct comparison with the scenario-led projections uses 50 weather generator replicates, to examine how the performance (irrigation deficit) varies as a function of random sampling variations. Performance across the 50 replicates is summarized using box-and-whisker plots, with outliers excluded beyond one and a half times the interquartile range. In the benchmark implementation only, a second set of targets is used for direct comparison with the scenario-led approach. This set of targets directly matches the values of all 15 candidate attributes in the scenario-led time series, rather than focusing solely on the critical attributes. This enables examination of the impact of using a smaller number of critical climate attributes to visualize a scenario-neutral space.

### 4.3.4 Implementations of scenario-neutral approach demonstrating impact of pitfalls

Details of the changes that have been made to the state-of-the-art benchmark methods in order to demonstrate the impact of each pitfall on the ability to generate scenarios that are commensurate with those produced using scenario-led approaches are given in Table 4-6.

**Table 4-6** Changes to benchmark methods to demonstrate pitfalls.

| **Step 1: Identify most critical climate attributes** | | | |
|---|---|---|---|
| | **Benchmark 1** | **Pitfall 1** | **Pitfall 2** |
| Candidate attribute selection | 15 for both precipitation and temperature | N/A | 8 precipitation attributes only |
| Critical attribute identification | Using PMI algorithm | Default to *Ptot, Tavg* | Equivalent to benchmark |
| **Step 2: Generate the climate time series to cover plausible future states** | | | |
| | **Benchmark 2** | **Pitfall 3** | **Pitfall 4** |
| Perturbation technique | Seasonal weather generator | Annual precipitation weather generator | Equivalent to benchmark |
| Monitored attributes | All remaining candidate attributes | Equivalent to benchmark | None |

Pitfall 1 was developed by using only *Ptot and Tavg* as the default attributes, rather than selecting a potentially larger pool of critical attributes based on a structured sensitivity analysis. To the extent possible, all other candidate attributes are held at historical levels, to ensure the time series created are still realistic.

Pitfall 2 was developed by using a candidate pool that considers only the eight precipitation-related attributes in Table 4-3, leaving the temperature time series at historical levels. The approach used in Section 4.3.3.1 was applied to this smaller candidate pool, leading to the critical attributes in Table 4-7. Here, *Ptot* was selected as the most critical attribute using the PMI algorithm, which is the same for the benchmark method (Table 4-5). However, the CVE metric indicates that this attribute alone now explains more of the variance in system performance across the new samples than it did when the temperature time series was also varying. In contrast, whereas *F0* was selected second in the benchmark implementation, the total winter rainfall (*PDJFtot*) is selected here. While this attribute does not control when the snow melts, it does control how much water is stored in snow given a fixed temperature series, and so might serve as a proxy for a similar phenomenon. The final critical attribute is the rainfall in summer (*PJJAtot*) which was also selected in the benchmark implementation.

**Table 4-7** Critical attributes given a subset of candidate attributes.

| Rank | Attribute | CVE |
|---|---|---|
| 1 | *Ptot* | 0.830 |
| 2 | *PDJFtot* | 0.862 |
| 3 | *PJJAtot* | 0.931 |

Pitfalls 3 and 4 both aim to recreate scenario-neutral time series with the same target perturbations as the benchmark implementation (i.e. same critical attributes), but with the different configurations of the *fore*SIGHT package. In particular, Pitfall 3 was developed by using an annual weather generator to perturb the precipitation time series, limiting the extent to which the inverse approach can achieve certain attribute combinations—particularly those with seasonal characteristics. Pitfall 4 was developed by no longer seeking to constrain non-critical attributes at their historical levels. This allows other candidate attributes to vary freely when generating the scenario-neutral climate time series, and can result highly unrealistic time series as discussed further in *Culley et al.* [2019a]. This pitfall only applies for applications of the scenario neutral methodology that are based on stochastic generation approaches, as simple scaling approaches would lead to predictable changes in the non-critical attributes.

## 4.4 Results

The following five sections present the results from each implementation of the scenario-neutral approach, and the diagnostic of comparing the scenario-neutral and scenario-led time series. The benchmark implementation is presented first (Section 4.4.1), to assess how similar scenario-led and scenario-neutral approaches can be. The four pitfalls are presented in the following four sections (Sections 4.4.2 – 4.4.5), to demonstrate the effect of each pitfall.

### 4.4.1 Benchmark methods

A comparison of the system performance from the scenario-neutral and scenario-led approaches is given in Figure 4-3. The twenty ‘scenario-led’ climate projections for the Lake Como case study are placed on the scenario-neutral space (colored dots, Figure 4-3) with the color inside the black circles indicating

the scenario-led irrigation deficit performance. Whereas the climate projections are placed precisely on the scenario-neutral space for critical attributes one and two, for critical attributes three and four the projections are laid onto the panel with attribute values that are closest to their projected value. Visual inspection suggests that there is some level of agreement between the two approaches as to what the system response to changes in the critical climate attributes will be. For example, it can be seen that the deficit is largest when *TJun* is highest and *PJJAtot* is lowest. This makes practical sense as the demand is highest in summer, and is not being met under these conditions. Further to this point, the irrigation deficit is shown to improve as the seasonality of rainfall changes to increase the amount of summer precipitation.

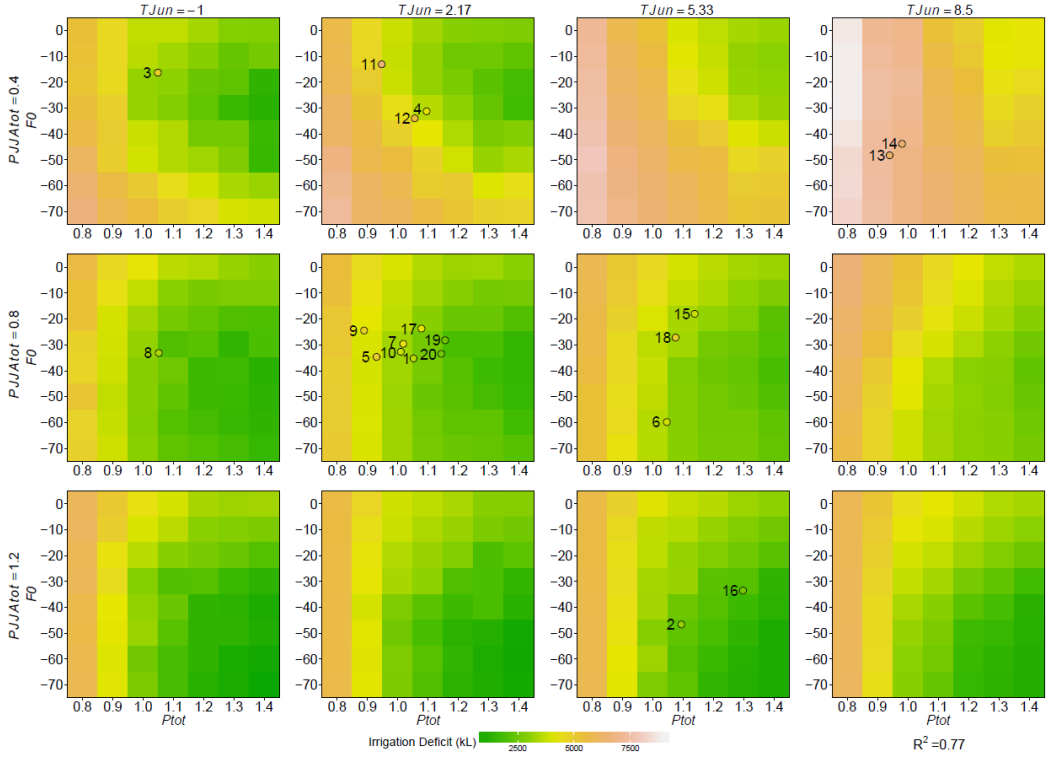


**Figure 4-3** A scenario-neutral space for the irrigation deficit performance criteria for the benchmark, where perturbations to historical climate are presented for mean precipitation (Ptot) and the number of frost days (F0) in the x and y axes respectively. Changes from left to right show the temperature in June (TJun), and changes from top to bottom show the total rainfall in summer (PJJAtot). Climate projections are overlaid, with the color inside the circles showing the scenario-led performance.

Figure 4-4 shows a more direct comparison between the performance in response

to each climate projection, with a box and whisker plot showing the irrigation deficit spread across fifty weather generator replicates. Two comparisons are shown: the time series generated with targets of change in all candidate attributes (15 attributes) is shown on the left panel, and the time series with change in just the critical attributes identified with the benchmark methods (four attributes) is shown on the right panel. Perturbing targets with all candidate attributes provides stochastic time series that match the scenario-led time series well, with an $R^2$ of 0.85. Differences in the scenario-neutral and scenario-led performance values indicate that the time series do not match exactly, which can also be seen by the spread in performance across each weather generator replicate. However, the nature of this error indicates that the differences between the time series are likely to be due to variability across weather generator replicates, instead of missing a key climate attribute.

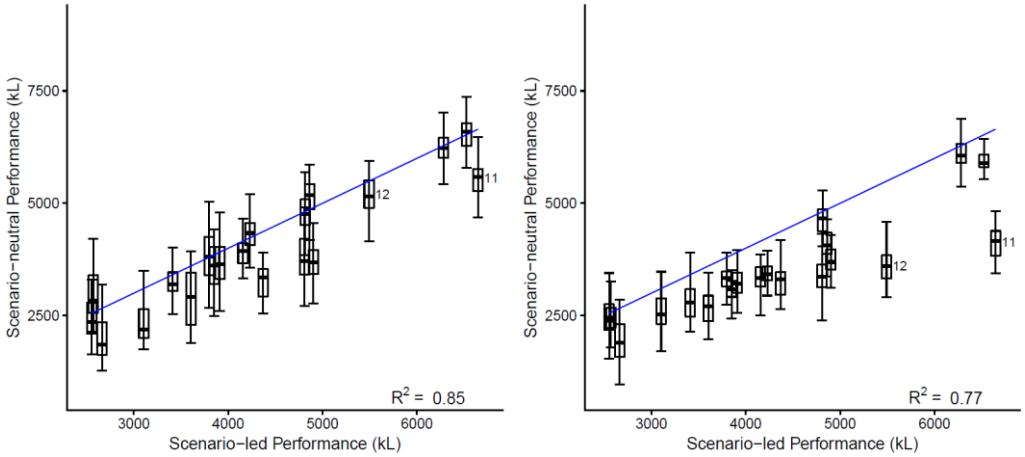


**Figure 4-4** A diagnostic for the benchmark methods. Fifty replicates of a stochastic weather generator are used to create a target that matches all candidate attributes (left panel) and matches the four critical attributes (right panel) of the projections, with the resulting spread in performance shown. The one-to-one line is shown in blue.

When only the four critical attributes are set as targets, and the other candidate attributes are held at historical levels, the overall match in performance is still strong ($R^2$ of 0.77), as evidenced by the right panel of Figure 4-4. The slight degradation the match in performance is mostly a result of two of the projections (11 and 12), where the irrigation deficit is slightly lower for the scenario-neutral time series compared to the scenario-led time series. Projections 11 and 12 come

from the same GCM/RCM combination (

Table 4-2), and have the largest decrease in summer precipitation, and when the weather generator makes a precipitation time series with only *Ptot* and *PJJAtot* as targets, the remaining seasons do not match the climate projection, causing an underestimation of irrigation deficit. These results indicate that the four critical climate attributes identified using state-of-the-art methods are a strong indicator of system performance.

### 4.4.2 Pitfall 1 – Defaulting to mean precipitation and temperature

The impact of pitfall 1 is shown in Figure 4-5, where *Ptot* and *Tavg* are used as default attributes for the scenario-neutral space. As can be seen, the match between the scenario-neutral and scenario-led results deteriorates compared to the benchmark implementation (Figure 4-3). This is mostly seen when considering the clustering of the mid-range deficit scenario-led time series (e.g. projections 3, 4, 11 and 12). This is in part because the scenario-neutral space for the demonstration of Pitfall 1 is two-dimensional, showing irrigation deficit change in response to *Ptot* and *Tavg* (Figure 4-5). The range of irrigation deficit captured by this scenario-neutral space is not as large, which can be seen when compared to deficit values in the top right panel of Figure 4-3. Given the severity of the irrigation deficit that can occur when *F0* and *PJJAtot* change, as is indicated by the color in the climate projection overlay, it is clear that an analysis with just this *Ptot* and *Tavg* space would not allow for a full understanding of likely system response to these conditions.

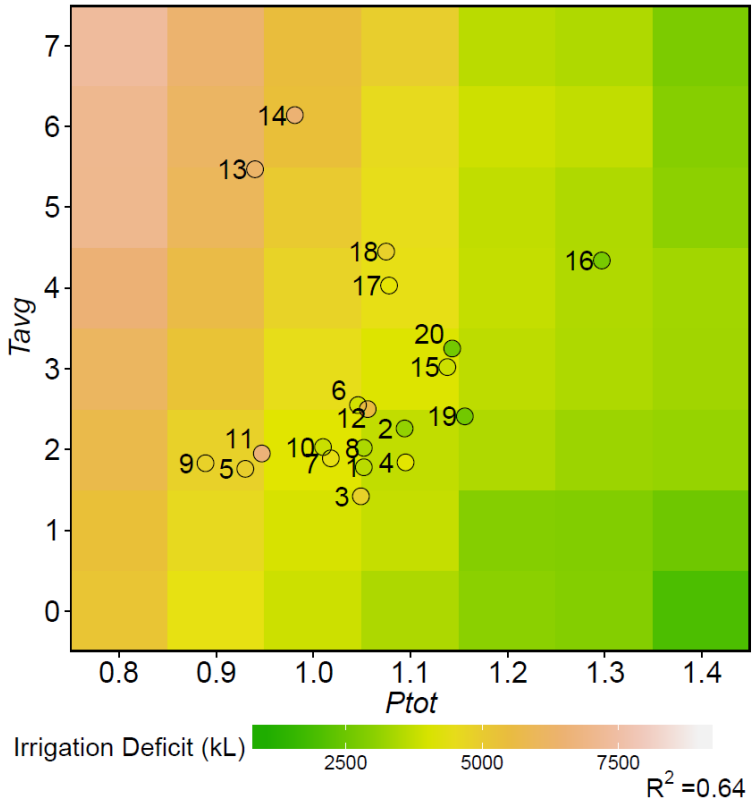


**Figure 4-5** A scenario-neutral space for the irrigation deficit performance criteria for pitfall 1, where perturbations to historical climate are presented for mean precipitation (*Ptot*) and mean temperature (*Tavg*) in the x and y axes respectively. Climate projections are overlaid, with the color inside the circles showing the scenario-led performance.

The mismatch in performance is shown more clearly in Figure 4-6, which shows comparisons between scenario-neutral and scenario-led results for the benchmark implementation on the left panel, and the time series generated with just mean precipitation and temperature as targeted changes on the right. Compared to the benchmark implementation, the irrigation deficit range in the scenario-neutral time series for pitfall 1 is less than the range in response to scenario-led time series. This is seen more clearly for the projections with lower irrigation deficit (e.g. projections 2, 19 and 20), where the scenario-neutral time series are overestimating deficit (Figure 4-6, right panel). This is due to several key attributes that affect system dynamics remaining constant; namely the seasonality throughout the year. Overall, it is evident that while these two attributes alone predict some system response, key mechanisms for the system are missing.

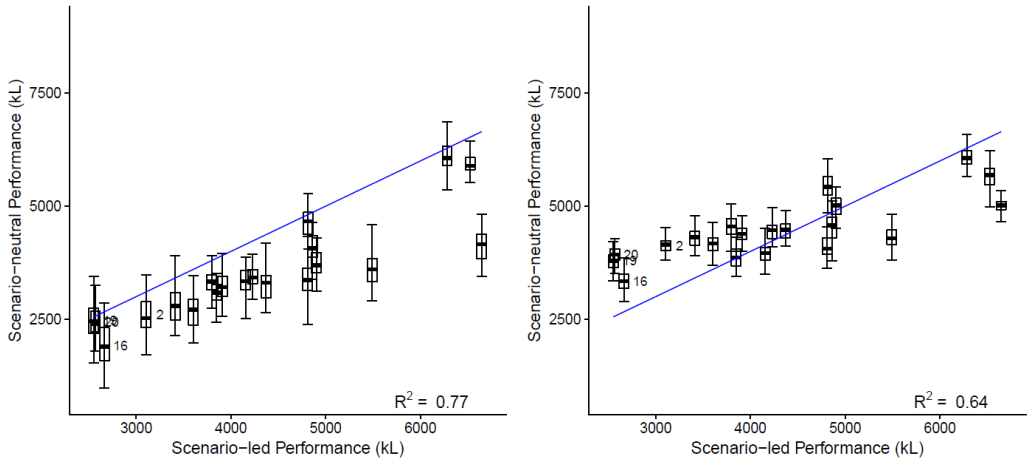


**Figure 4-6** A diagnostic for pitfall 1. Fifty replicates of a stochastic weather generator are used to create targets for the benchmark implementation (left), and targets that match only the *Ptot* and *Tavg* attributes of the projections (right panel), with the resulting spread in performance shown. The one-to-one line is shown in blue.

### 4.4.3 Pitfall 2 – Missing critical attributes in the analysis

The impact of pitfall 2, where the critical temperature attributes are excluded from the analysis, is shown in Figure 4-7. As can be seen, the range of irrigation deficit indicated by the scenario-neutral space is reduced when compared to the benchmark implementation (calculated range of 5306 kL for Figure 4-7 compared to 8199 kL for Figure 4-3). This produces a weaker match between the scenario-neutral and scenario-led analysis, where in particular the projections on the left panel of Figure 4-7 show a greater irrigation deficit than those in the scenario-neutral space. The dynamics of changing summer rainfall are still present in this scenario-neutral space; however, without the shifting temperature in summer, the deficit does not become as large. The comparison with climate projections supports this, where the projections with larger irrigation deficits do not match the scenario-neutral space.

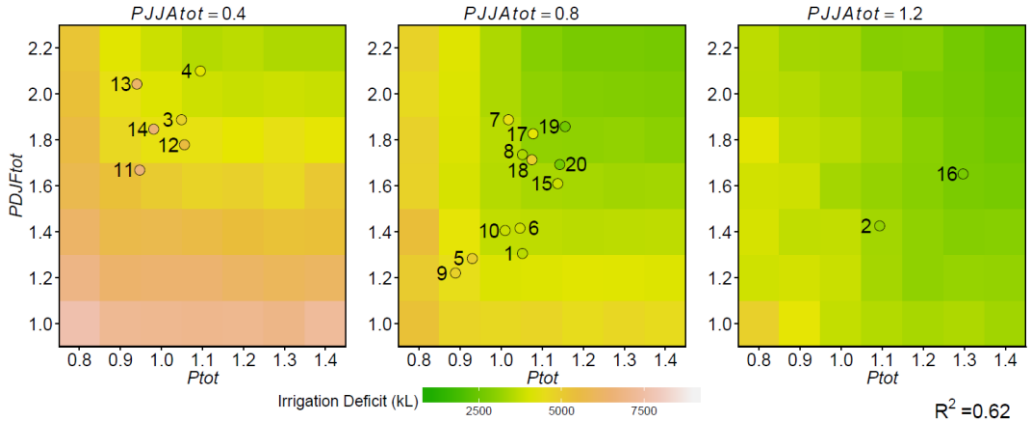


**Figure 4-7** A scenario-neutral space for the irrigation deficit performance criteria for pitfall 2, where perturbations to historical climate are presented for total annual precipitation (*Ptot*) and the total winter precipitation (*PDJFtot*) in the x and y axes respectively. Changes from left to right show the total rainfall in summer (*PJJAtot*). The temperature time series is constant throughout. Climate projections are overlaid, with the color inside the circles showing the scenario-led performance.

Figure 4-8 also shows the increased mismatch between the scenario-neutral and scenario-led time series when only critical precipitation attributes are used, particularly the under-estimation of the irrigation deficit for the high deficit scenarios (i.e., projections 13, 14 and 18). Despite this, many of the scenario-led time series with smaller irrigation deficit values match well. For these smaller irrigation deficit projections, considering perturbations to precipitation attributes alone is enough to capture the deficit behavior. However, for the high-deficit projections, even though the most critical attribute, *Ptot*, is correct, the temperature being held at historical conditions does not create a large enough irrigation deficit.

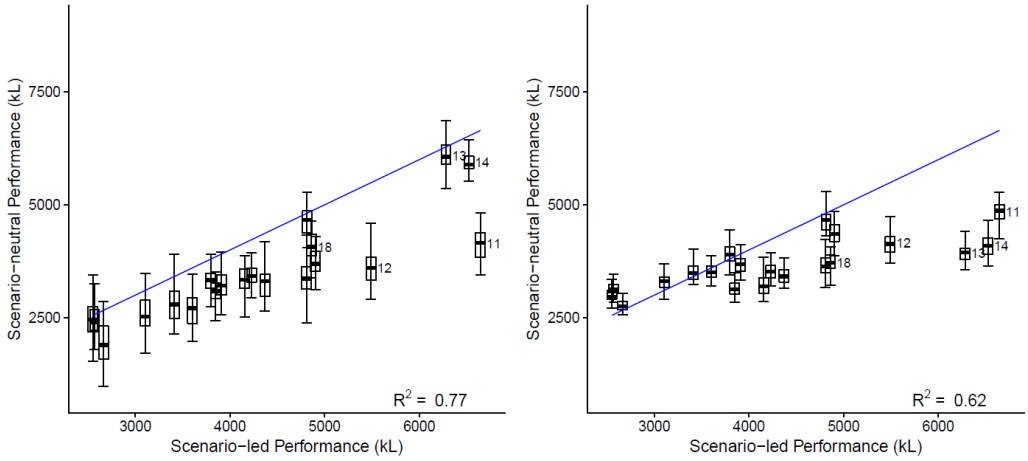


**Figure 4-8** A diagnostic for pitfall 2. Fifty replicates of a stochastic weather generator are used to create targets for the benchmark implementation (left panel), and targets that match only the three critical precipitation attributes of the projections (right panel), with the resulting spread in performance shown. The one-to-one line is shown in blue.

### 4.4.4 Pitfall 3 – Insufficiently flexible perturbation method

The impact of pitfall 3 can be seen in Figure 4-9. This pitfall limits the range of attributes that can be perturbed, as evidenced by the lack of change in the surface across each column of panels in Figure 4-9. In this implementation, the weather generator cannot make independent pertubations to both the annual preciptation (*Ptot*) and the summer precipation (*PJJAtot*). Given it is not possible to achieve both target attributes, the attribute that meets its target will be a function of the weather generator mechanics. In this case, *Ptot* is perturbed across each space, but the seasonal changes to summer rainfall are not successfully produced. This causes the low summer rainfall projections in particular to not match the panels they lie on (projections 3, 4, 11, 12). The error between the target critical attributes and the attributes of the scenario-neutral time series can be seen in Supplementary Material A.

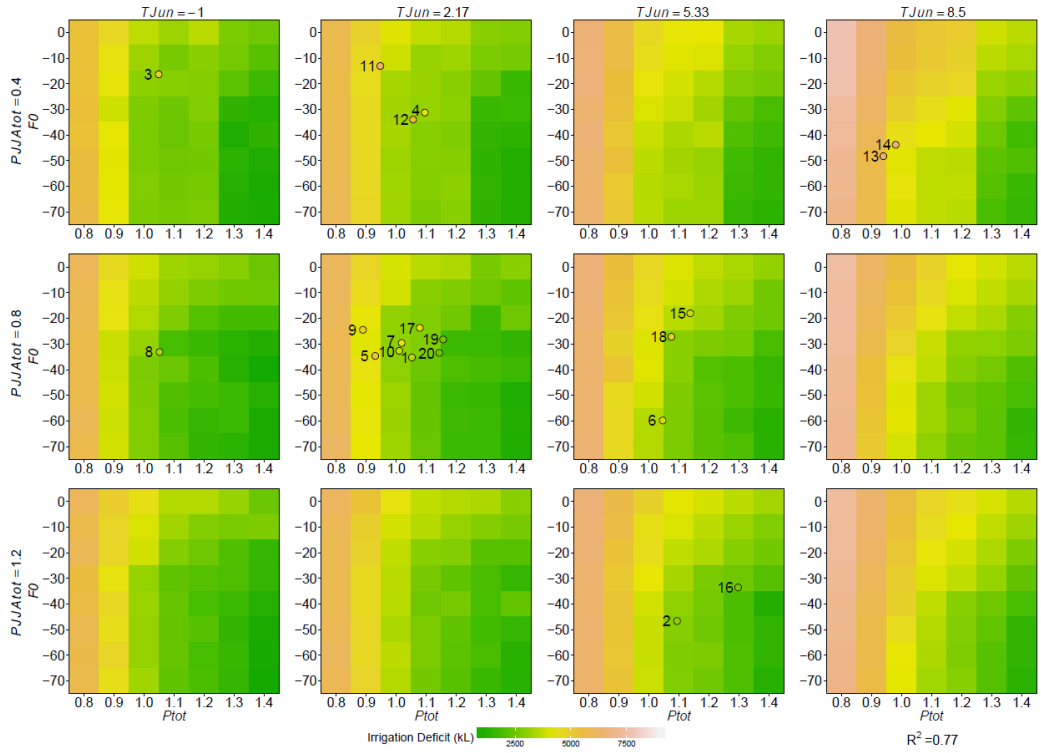


**Figure 4-9** A scenario-neutral space for the irrigation deficit performance criteria for pitfall 3, where an annual precipitation weather generator is used to create the climate time series. Perturbations to historical climate are presented for mean precipitation (*Ptot*) and the number of frost days (*F0*) in the *x* and *y* axes respectively. Changes from left to right show the temperature in June (*TJun*), and changes from top to bottom show the total rainfall in summer (*PJJAtot*). Climate projections are overlaid, with the color inside the circles showing the scenario-led performance.

Interestingly, the direct comparison between the performance arising from the scenario-neutral and scenario-led time series, however, indicates that the match with the scenario-led time series is similar to that of the benchmark implementation (Figure 4-10). The primary reason for this is that the first three of the four critical attributes are still being successfully perturbed, and it is only *PJJAtot* that sometimes does not meet its target. The requested targets for *PJJAtot* to match the scenario-led time series are not as extreme as the bounds used in the scenario-neutral space, and so the majority of scenario-neutral time series have similar medians to those of the benchmark implementation. It is only the high deficit climate projections, that have high decreases in *PJJAtot*, that do not match as well (e.g. projections 13, 14). As there are fewer high deficit

projections, the difference between this pitfall and the benchmark implementation is not as pronouced.

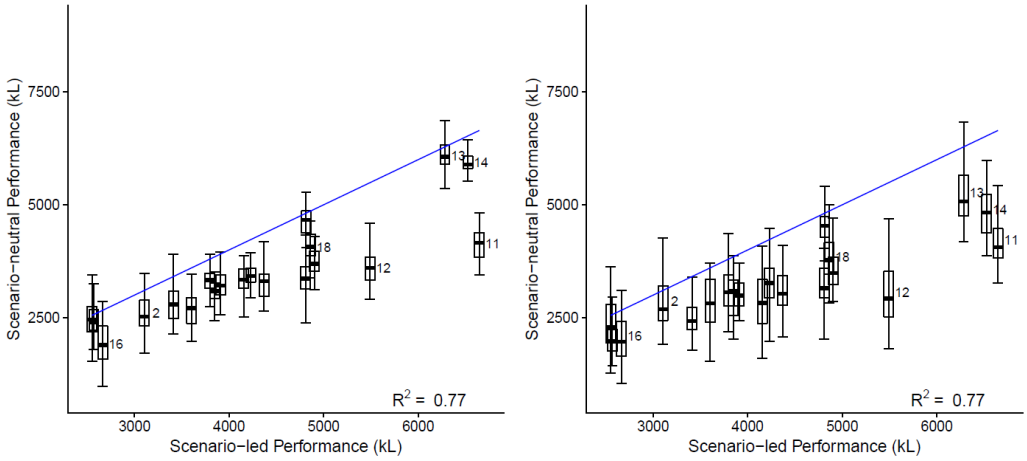


**Figure 4-10** A diagnostic for pitfall 3. Fifty replicates of a stochastic weather generator are used to create a target that matches the four critical attributes of the projections, with the resulting spread in performance shown. The benchmark results are shown in the left panel, and pitfall 3 is shown in the right panel, where an annual precipitation weather generator is used to create the climate time series. The one-to-one line is shown in blue.

### 4.4.5 Pitfall 4 – Unintended perturbations to a climate time series

Whereas the implementation of the inverse approach for the benchmark simulations and for pitfalls 1-3 involved perturbing a set of critical attributes while seeking to hold a wide range of other non-critical attributes at close to their historical level, in the case of Pitfall 4 the non-critical attributes were free to vary as part of the stochastic generation optimization process. The results are shown in Figure 4-11, and as can be seen there is no clear trend across some regions of the space. This is because while the time series meet the desired critical attributes, other aspects of the time series, like the other seasonal precipitation totals that were not part of the critical attribute set, vary from point to point. The error between the target non-critical attributes and the attributes of the time series can be seen in Supplementary Material B, and highlights the importance of carefully monitoring the stochastic generation process as part of any scenario-neutral analysis.

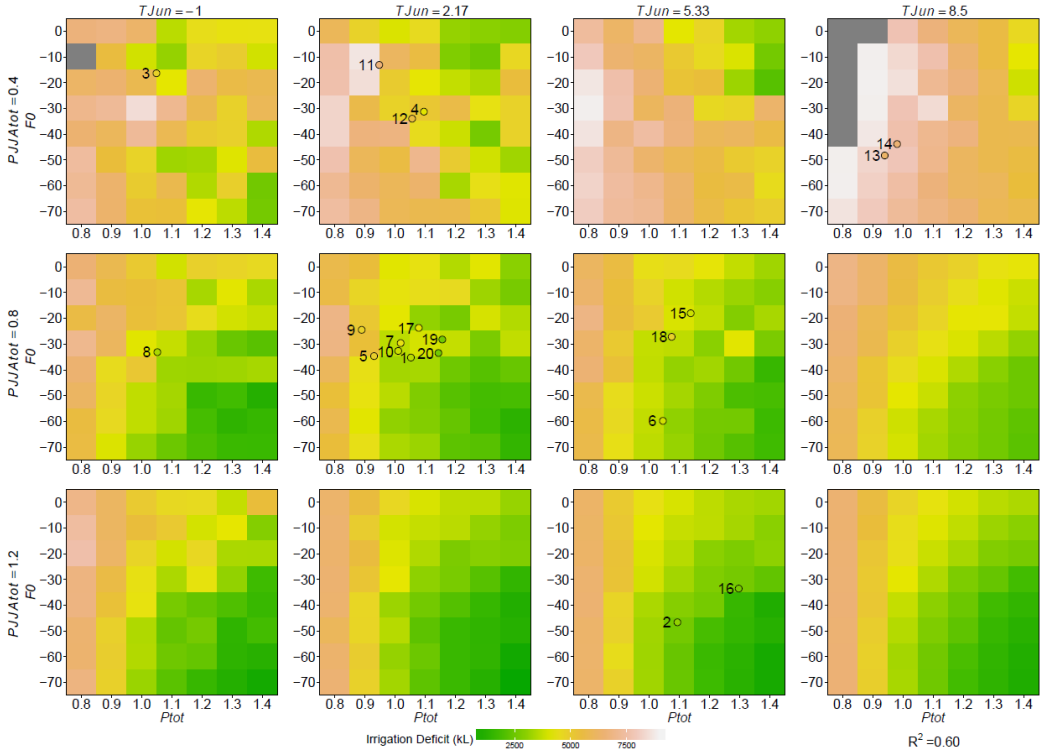


**Figure 4-11** A scenario-neutral space for the irrigation deficit performance criteria for pitfall 4, where perturbations to historical climate are presented for mean precipitation (*Ptot*) and the number of frost days (*F0*) in the *x* and *y* axes respectively. Changes from left to right show the temperature in June (*TJun*), and changes from top to bottom show the total rainfall in summer (*PJJAtot*). No other attributes are set as targets. Climate projections are overlaid, with the color inside the circles showing the scenario-led performance. Some irrigation deficit values falls above the range of the color ramp (indicated by grey pixels).

This can be seen more clearly when considering the spread in performance in response to fifty replicates of each point (Figure 4-12), which is much larger than the spread in performance in the benchmark implementation. This affects some projections more than others, where, like in the bottom right panels in Figure 4-11, the specific changes in attributes can create a more constrained optimization problem given the weather generators used (e.g. projection 11 compared to 12). When present, the large spread across replicates is in part due to the winter, spring and autumn precipitation totals that can change in each replicate, as only the *Ptot* and *PJJAtot* are used as targets. However, the largest cause of spread in performance results from the under-constrained temperature time series, where the seasonality of temperature can completely flip given that

only *F0* and *TJun* have been set as targeted changes.

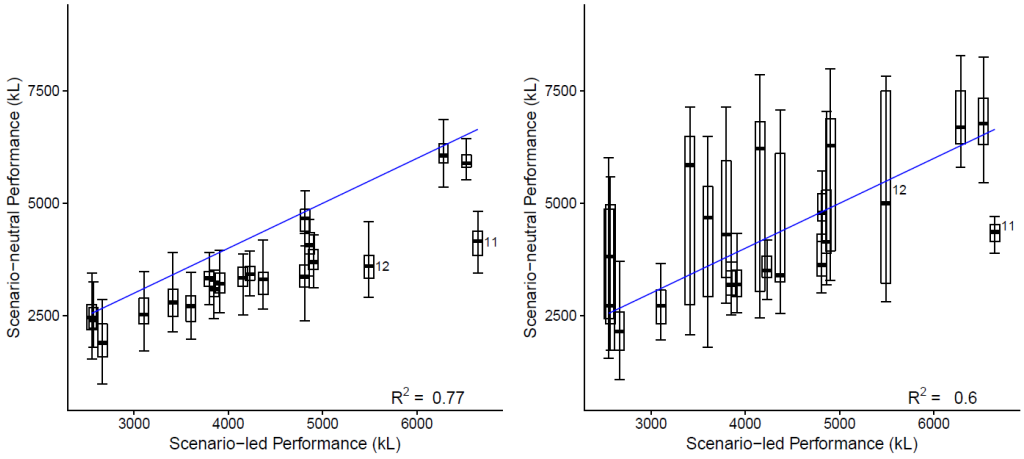


**Figure 4-12** A diagnostic for pitfall 4. Fifty replicates of a stochastic weather generator are used to create a target that matches the four critical attributes of the projections, with the resulting spread in performance shown. The benchmark results are shown in the left panel, and pitfall 4 is shown in the right panel, where no other attributes are set as targets. The one-to-one line is shown in blue.

## 4.5 Discussion

The difference in system performance obtained from scenario-led studies and scenario-neutral assessments is increasingly recognized in the literature [*Taner et al.*, 2017], and was verified here. In this section we ask: how do we know which set of assessments are more likely to be correct? And if there is a discrepancy, what is the most efficient way to find the likely cause and improve scenario-neutral performance?

### 4.5.1 Resolving conflicts between scenario-neutral and scenario-led projections

In the situation where there are significant differences between system performance obtained through scenario-neutral and scenario-led analyses, it is obvious to ask which of the two is most likely to be correct. It is argued here that neither is necessarily more correct than the other; rather the differences in system performance arise through differences in aspects of the time series not explicitly

represented by the scenario-neutral analysis.

Consider the example in which a scenario-neutral analysis is applied to changes in total annual precipitation and annual average temperature, and a scenario-neutral space is generated for a range of combinations of these attributes, holding all other attributes as close to historical conditions as possible. Under the assumption that the system model correctly represents the relevant system processes, the ensuing scenario-neutral space represents a valid and 'correct' system response to these simulated climatic changes. If a scenario-led climate projection (e.g. from a GCM) is superimposed onto the scenario-neutral space and exhibits a different performance to the scenario-neutral analysis for the same combination of total annual precipitation and annual average temperature, then given it is the same system model that is used, the only explanation is that some other aspect of the two time series must be different. For example, the scenario-led projection could also be suggesting a change in seasonality and number of wet days, whereas in the scenario-neutral analysis these are relatively unchanged.

To further explore the notion of the 'correctness' of both analyses, note that there are fundamental differences in the philosophical approach of scenario-neutral and scenario-led impact assessments. In particular, the former is essentially a sensitivity or 'what-if' analysis, exploring how a system might response to hypothetical future climates. In contrast, the latter seeks to use modelling statements of likely (or at least possible) future change, usually based on the best-available science at the time (e.g. projections from recent GCM/RCM combinations). Based on this, a discrepancy between the two approaches should lead to two key conclusions:

- The system exhibits sensitivity to attributes not considered in the scenario-neutral analysis. If this were not the case then there would be no mechanism by which the system model could produce different performances for the same changes to critical attributes.

- There is evidence that those attributes not considered in the scenario

neutral analysis could change in a future climate, at least to the extent that the models used to generate scenario-led projections can be trusted to simulate those attributes.

As a result of the above, it is natural to conclude that discrepancies between the two approaches highlight potential system vulnerabilities that could be important to the system, but were not accounted for by the scenario-neutral analysis. It is for this reason that the approach taken here is to use scenario-led projections as an evaluation metric for scenario-neutral analyses. And even if the scenario-led projections turn out to be incorrect (for example the projections exhibit significant biases in the relevant attributes), then at least the scenario-neutral method is alerted to possible sensitivities that may be worthy of further exploration.

Finally, as noted elsewhere in this paper, any estimates of changes to system performance as a function of climatic changes (regardless of whether the changes are generated by scenario-neutral or scenario-led approaches) are made with an assumption that the system model provides a reasonable representation of the system response to key climate forcings. This is a limitation of this study, as it has been demonstrated elsewhere that this should not be assumed to be the case [*Broderick et al.*, 2016; *Guo et al.*, 2017b]. It is therefore recommended that, as a core part of the scenario-neutral analysis, the assumptions and performance of system models are carefully scrutinized relative to the drivers of change being explored. A detailed review of approaches for model evaluation under changed climatic regimes is outside of the scope of this paper, but methods such as differential split-sample testing or optimizing model performance for different climate conditions represent several approaches to address this issue [*Westra et al.*, 2014; *Broderick et al.*, 2016; *Fowler et al.*, 2016].

### 4.5.2 Improving the performance of scenario-neutral assessments

The approach recommended here—namely a comparison of scenario-neutral and scenario-led assessments as a means of evaluating scenario-neutral performance—provides an indicator that one or several of the pitfalls identified

here has been fallen for. To this end, a systematic approach is recommended to further explore discrepancies and identify strategies for improving the scenario-neutral analysis.

Given that, as discussed above, a discrepancy between the two assessments is an indicator of differences in the climatic time series outside of the critical attributes being included in the scenario-neutral analysis, a logical start is to comprehensively review differences in all possible model attributes from the two approaches to assess potential directions for further exploration. An example of this calculation is shown in Supplementary Material C, where for each pitfall the difference between the candidate attributes of the scenario-neutral and scenario-led time series is shown. For pitfalls 1 and 2, this shows the selected critical attributes for each pitfall with low error, while there is higher error in the remaining candidate attributes, some of which are known to be important from the benchmark implementation. This process identifies if some attributes differ significantly between the two time series, but it does not identify if these attributes have a significant impact on system performance. To provide information on this latter question, the attribute will need to be added to a new candidate set, and the overall scenario-neutral analysis repeated. This process will work towards addressing pitfalls 1 and 2.

To further understand any possible modelling limitations (pitfalls 3 and 4), it is recommended to identify the climate model projections that depart most significantly from the scenario-neutral analysis, and assess whether there are any specific features associated with those projections. For example, in the benchmark implementation, system responses to scenario-neutral time series generated using the four critical climate attributes disagreed with climate projections 11 and 12. Both projections were unique in that they had a seasonal precipitation pattern that was not simulated by the weather generator used in this study when only summer was perturbed, therefore producing the largest error when generating scenario-neutral time series (Section 4.4.1). When using a different weather generator, or perturbation technique, the seasonality produced when perturbing *Ptot* and *PJJAtot* could be expected to change, and so would

the error. This shows the merit of the diagnostic technique coupled with an understanding of the evaluated system, as a disagreement between the scenario-led and scenario-neutral approaches can be attributed to a limitation in the stochastic weather generation models used.

## 4.6 Conclusion

Scenario-neutral impact assessments are being widely adopted as they expand upon a scenario-led analysis to uncover more information about a system's vulnerabilities. However, in practice, there are several ways in which a scenario-neutral analysis can be undermined, and thus produce misleading information on system sensitivity to potential climatic changes. Firstly, a system can be stress tested against climate attributes that are not those most critical for that system. Using these attributes it would still be possible to define modes of failure for a system; however, these modes may not describe the most important system sensitivities, and other more important modes of failure may be missed. Secondly, when generating climate time series for analysis, climate attributes that are not the focus of the study can also be perturbed. Should they still affect system performance, this creates a scenario-neutral space that does not best represent system response to the intended perturbed attributes.

A benchmark demonstration of an application of the scenario-neutral approach that is successful in identifying the changes in climate that most affect system performance is presented for the case study of Lake Como. A performance criterion of irrigation deficit is used, which is strongly affected by the seasonality of climate, given the demand pattern throughout the year. This benchmark implementation using state-of-the-art methods shows the scenario-neutral approach can agree with the results of a scenario-led approach, although with some loss in information due to only being able to visualize results in low dimensions. This paper then demonstrates how four pitfalls can cause key system dynamics to be missed, of which two can arise when specifying the attributes to be considered as part of the scenario-neutral analysis, and another two can arise when generating the time series to populate the scenario-neutral space. The method of illustrating this is to compare different implementations of a scenario-

neutral approach with climate projections that would be used in a scenario-led analysis. Identifying a difference in performance between a scenario-neutral time series developed, and the performance as indicated by a scenario-led time series under the same conditions, is used as an indicator that some aspect of the climate time series that is not captured in the scenario-neutral analysis is affecting system performance.

Once identified using the above diagnostic, these pitfalls can be avoided by using the methods first presented in the benchmark application. The validation with scenario-led time series can therefore be used to ensure the key motivation of a scenario-neutral climate impact assessment analysis is fulfilled, which is that key system vulnerabilities are identified. This allows the key concerns for a system to be addressed when using scenario-neutral approaches for adaptation planning and decision making.

## 4.7 Acknowledgements

The case study data used in this study are from Agenzia Regionale per la Protezione dell'Ambiente (http://ita.arpalombardia.it/ita/inde) and Consorzio dell'Adda (http://www.addaconsorzio.it/). The authors would like to thank ARPA and Eng. Bertoli from Consorzio dell'Adda for its provision. Details of the climate time series data produced using the *fore*SIGHT package (https://cran.r-project.org/web/packages/foreSIGHT/index.html), and case study performance data used in the results, are submitted as supporting information. Sam Culley was supported by an Australian Postgraduate Award.

## 4.8 References

Anghileri, D., F. Pianosi, and R. Soncini-Sessa (2011), A framework for the quantitative assessment of climate change impacts on water-related activities at the basin scale, Hydrol. Earth Syst. Sci., 15(6), 2025-2038.

Arnell, N. W. (2004), Climate change and global water resources: SRES emissions and socio-economic scenarios, Glob. Environ. Change-Human Policy Dimens., 14(1), 31-52.

Bennett, B., S. Culley, S. Westra, D. Guo, and H. R. Maier (2018), foreSIGHT: Systems Insights from Generation of Hydroclimatic Timeseries, edited, p. R package version 0.9.04.

Bergström, S., and V. Singh (1995), The HBV model, Computer models of watershed hydrology., 443-476.

Brekke, L. D., E. P. Maurer, J. D. Anderson, M. D. Dettinger, E. S. Townsley, A. Harrison, and T. Pruitt (2009), Assessing reservoir operations risk under climate change, Water Resources Research, 45(4), W04411.

Broderick, C., T. Matthews, R. L. Wilby, S. Bastola, and C. Murphy (2016), Transferability of hydrological models and ensemble averaging methods between contrasting climatic periods, Water Resources Research, 52(10), 8343-8373.

Broderick, C., C. Murphy, R. L. Wilby, T. Matthews, C. Prudhomme, and M. Adamson (2019), Using a Scenario-Neutral Framework to Avoid Potential Maladaptation to Future Flood Risk, Water Resources Research, 0(0).

Brown, C., and R. L. Wilby (2012), An alternate approach to assessing climate risks, Eos, Transactions American Geophysical Union, 93(41), 401-402.

Brown, C., Y. Ghile, M. Laverty, and K. Li (2012), Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, Water Resources Research, 48(9), W09537.

Culley, S., B. Bennett, S. Westra, and H. R. Maier (2019a), Generating realistic perturbed hydrometeorological time series to inform scenario-neutral climate impact assessments, Journal of Hydrology, 576, 111-122.

Culley, S., S. Westra, H. Maier, and B. Bennet (2019b), Identifying critical climate variables for use in scenario-neutral climate impact assessments based on climate-perturbed hydrometeorological time series Water Resources Research - Under review.

Culley, S., S. Noble, A. Yates, M. Timbs, S. Westra, H. Maier, M. Giuliani, and A. Castelletti (2016), A bottom-up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate, Water Resources Research, 52(9), 6751-6768.

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 Resources Research, 52(3), 1820-1846.

Giuliani, M., J. D. Herman, A. Castelletti, and P. Reed (2014), Many-objective reservoir policy identification and refinement to reduce policy inertia and myopia in water management, Water Resources Research, 50(4), 3355-3377.

Guo, D., S. Westra, and H. R. Maier (2017a), Use of a scenario-neutral approach to identify the key hydro-meteorological attributes that impact runoff from a natural catchment, Journal of Hydrology, 554, 317-330.

Guo, D., S. Westra, and H. Maier (2017b), Sensitivity of potential evapotranspiration to changes in climate variables for different Australian climatic zones.

Guo, D., S. Westra, and H. R. Maier (2018), An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies, Journal of Hydrology, 556, 877-890.

Haasnoot, M., J. H. Kwakkel, W. E. Walker, and J. ter Maat (2013), Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world, Global Environmental Change, 23(2), 485-498.

Jacob, D., J. Petersen, B. Eggert, A. Alias, O. B. Christensen, L. M. Bouwer, A. Braun, A. Colette, M. Déqué, and G. Georgievski (2014), EURO-CORDEX: new high-resolution climate change projections for European impact research, Reg. Envir. Chang., 14(2), 563-578.

Keller, L., O. Rössler, O. Martius, and R. Weingartner (2018), Comparison of scenario-neutral approaches for estimation of climate change impacts on flood characteristics, Hydrological Processes, 0(0).

Kwakkel, J. H., W. E. Walker, and M. Haasnoot (2016), Coping with the Wickedness of Public Policy Problems: Approaches for Decision Making under Deep Uncertainty, Journal of Water Resources Planning and Management, 01816001.

Li, X., A. C. Zecchin, and H. R. Maier (2015a), Improving partial mutual information-based input variable selection by consideration of boundary issues associated with bandwidth estimation, Environ. Modell. Softw., 71, 78-96.

Li, X., H. R. Maier, and A. C. Zecchin (2015b), Improved PMI-based input variable selection approach for artificial neural network and other data driven environmental and water resource models, Environ. Modell. Softw., 65, 15-29.

Maier, H. R., S. Razavi, Z. Kapelan, L. S. Matott, J. Kasprzyk, and B. A. Tolson (2019), Introductory overview: Optimization using evolutionary algorithms and other metaheuristics, Environ. Modell. Softw., 114, 195-213.

Maier, H. R., et al. (2014), Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions, Environ. Modell. Softw., 62(0), 271-299.

McPhail, C., H. Maier, J. Kwakkel, M. Giuliani, A. Castelletti, and S. Westra (2018), Robustness metrics: How are they calculated, when should they be used and why do they give different results?, Earth's Future, 6(2), 169-191.

Prudhomme, C., A. L. Kay, S. Crooks, and N. Reynard (2013), Climate change and river flooding: Part 2 sensitivity characterisation for British catchments and example vulnerability assessments, Climatic Change, 119(3-4), 949-964.

Prudhomme, C., R. L. Wilby, S. Crooks, A. L. Kay, and N. S. Reynard (2010), Scenario-neutral approach to climate change impact studies: Application to flood risk, Journal of Hydrology, 390(3-4), 198-209.

Raäisaänen, J. (2007), How reliable are climate models?, Tellus A: Dynamic Meteorology and Oceanography, 59(1), 2-29.

Ray, P. A., L. Bonzanigo, S. Wi, Y.-C. E. Yang, P. Karki, L. E. García, D. J. Rodriguez, and C. M. Brown (2018), Multidimensional stress test for hydropower investments facing climate, geophysical and financial uncertainty, Global Environmental Change, 48, 168-181.

Sharma, A. (2000), Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1 — A strategy for system predictor identification, Journal of Hydrology, 239(1), 232-239.

Steinschneider, S., and C. Brown (2013), A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, Water Resources Research, 49(11), 7205-7220.

Steinschneider, S., R. McCrary, S. Wi, K. Mulligan, L. Mearns, and C. Brown (2015), Expanded Decision-Scaling Framework to Select Robust Long-Term Water-System Plans under Hydroclimatic Uncertainties, Journal of Water Resources Planning and Management, 0(0), 04015023.

Taner, M. Ü., P. Ray, and C. Brown (2017), Robustness-based evaluation of hydropower infrastructure design under climate change, Climate Risk Management, 18, 34-50.

Turner, S. W. D., D. Marlow, M. Ekström, B. G. Rhodes, U. Kularathna, and P. J. Jeffrey (2014), Linking climate projections to performance: A yield-based decision scaling assessment of a large urban water resources system, Water Resources Research, 50(4), 3553-3567.

Vano, J. A., M. J. Scott, N. Voisin, C. O. Stockle, A. F. Hamlet, K. E. B. Mickelson, M. M. Elsner, and D. P. Lettenmaier (2010), Climate change impacts on water management and irrigated agriculture in the Yakima River Basin, Washington, USA, Climatic Change, 102(1-2), 287-317.

Weiß, M. (2011), Future water availability in selected European catchments: a probabilistic assessment of seasonal flows under the IPCC A1B emission scenario using response surfaces, Natural Hazards and Earth System Sciences, 11(8), 2163-2171.

Westra, S., M. Thyer, M. Leonard, D. Kavetski, and M. Lambert (2014), A strategy for diagnosing and interpreting hydrological model nonstationarity, Water Resources Research, 50(6), 5090-5113.

Wetterhall, F., L. Graham, J. Andreasson, J. Rosberg, and W. Yang (2011), Using ensemble climate projections to assess probabilistic hydrological change in the Nordic region, Natural Hazards and Earth System Sciences, 11(8), 2295-2306.

Whateley, S., S. Steinschneider, and C. Brown (2014), A climate change range-based method for estimating robustness for water resources supply, Water Resources Research, 50(11), 8944-8961.

Wilby, R. L., and S. Dessai (2010), Robust adaptation to climate change, Weather, 65(7), 180-185.

# Chapter 5

Scenario-neutral climate impact assessments are being used increasingly to aid decision making for water resource systems facing the uncertainty of climate change. With more frequent applications, several limitations with the approach are becoming apparent, including (i) the difficulty in generating climate time series to meet more complex perturbations in climate, (ii) the lack of a method to identify which changes in climate should be included in a scenario-neutral assessment to best describe system performance, and (iii) the analysis of a system given by a scenario-neutral space not aligning with the analysis given by climate projections.

This research furthers the practical implementation of scenario-neutral climate impact assessments by achieving the following aims: (i) to formalise the inverse approach to stochastic time series generation, thereby improving the efficiency of the approach and ensuring the physical realism of the simulated time series, (ii) to develop and evaluate an approach for identifying the smallest number of climate attributes that have a significant impact on system performance, for use in scenario-neutral impact assessments that require hydrometeorological time series, and (iii) to present common pitfalls with the scenario-neutral approach that affect the validity of the results and demonstrate the effect of falling for each of these pitfalls.

## 5.1 Research contribution

The overall contribution of this research is the improvement of applications of scenario-neutral climate impact assessments, so that they can satisfy the intended purpose of such assessments and successfully identify the vulnerabilities of a water resource system. The specific research contributions are as follows:

1. This research has formalised the underlying optimisation process that is used by the inverse approach to generate specific perturbations to statistics of climate time series (e.g. precipitation, temperature) [*Guo et*

*al.*, 2017; 2018]. The statistics of the climate variables (e.g. means, extremes and intermittency) are referred to as climate 'attributes'. Through the addition of penalty functions to the optimisation objective function, the new formulation provides a structured way to ensuring the realism of the time series, while allowing a greater number of attributes to be perturbed. Being able to perturb a wider range of climate attributes allows applications of the scenario-neutral approach to be successful when implemented on more complex systems. This is because making changes to climate time series beyond means and seasonality can be necessary to uncover key system sensitivities. By proposing a numerically efficient solution that exploits prior knowledge of weather generator parameters, the time taken to generate climate time series has been dramatically decreased, allowing for more applications in a practical setting.

2. This research proposes a formal approach to identifying the most critical climate attributes for a water resource system, out of a large selection of potential variables (e.g. rainfall, potential evapotranspiration) and statistics of those variables (e.g. means, extremes, intermittency, seasonality). The critical climate attributes are selected to be the smallest number of attributes that have a significant influence on system performance. This technique ensures that scenario-neutral climate impact assessments stress-test a system against the changes to which it is most sensitive. This will enable identification all the major modes of failure for a given water resource system. It has been demonstrated that for a single system, different definitions of system performance require different stress-tests, warranting the need for a formal approach, and not an over-reliance on *a priori* knowledge of a system. The attribute selection technique can also be used to expand uncertainty analysis techniques that extend beyond scenario-neutral climate impact assessments, but that still investigate climate scenarios (e.g. robust optimisation, scenario discovery) [*Haasnoot et al.*, 2013; *Kasprzyk et al.*, 2013; *Kwakkel et al.*, 2016]. Such studies regularly use means in

precipitation and temperature as the only climate indicators, and can instead consider more significant changes in climate to design new management strategies.

3. This research presents common pitfalls that can prevent the successful implementation of scenario-neutral climate impact assessments. These pitfalls address both limitations in the scenario-neutral approach identified in current literature [*Taner et al.*, 2017; *Keller et al.*, 2018], and potential limitations when using some of the new methods presented in this thesis. By presenting a diagnostic of comparing to a scenario-led analysis, these pitfalls can be readily identified. It was also demonstrated that by using the methods described in this thesis, the scenario-neutral approach to creating climate time series can align with the scenario-led approach. The validation with scenario-led projections can therefore be used to ensure the key motivation of a scenario-neutral climate impact assessment is fulfilled.

## 5.2 Limitations and future work

Limitations of this research, as well as recommended directions for future work, are discussed below.

### 5.2.1 Creation of climate time series

In Chapter 2, an increase in efficiency of the inverse approach is demonstrated. This is in part achieved by changing the optimisation formulation to include domain knowledge about the weather generator parameters, and decrease the bounds used to search, hence decreasing the time needed to converge on a solution [*Kingston et al.*, 2008; *Maier et al.*, 2014]. While beneficial, the process of acquiring domain knowledge needs to be repeated for each weather generator used in the future. It is also region specific, and implementation in other regions requires a large database of historical climate time series. As a result, further testing of the developed methods on a wider set of climate regimes is needed, as well as additional climate variables (beyond precipitation), to gain a better understanding of how beneficial the inclusion of domain knowledge is. Further,

while this increase in efficiency has allowed for the more complex research applications presented in this thesis, the computational time is still too large to be completed in more practical settings, particularly without the use of high performance computing. It is for this reason that simple scaling of historical time series is so readily adopted as a perturbation method. A more efficient implementation of the code, in a program language like C++ instead of R, is recommended to make the implementation window smaller with more limited computer resources.

Additionally, this research focussed only on the WGEN weather generators in its analysis. This means the improvement of focusing on perturbed attributes meeting their target changes at the expense held attributes has only been demonstrated for one set of weather generators. In order to understand how general the findings on trade-offs between perturbed and held climate attributes are, a wider selection of weather generators should be tested. Also, given the complexity of water resource systems, weather generators that can perturb attributes inter-annually, or at the sub-daily scale, should be incorporated into the inverse approach, to broaden the range of potential applications. As these weather generators become more complex, they will require more attributes in order to monitor the time series to ensure realism. This would require further testing of how the proposed solutions to managing the trade-offs between attributes work with greater numbers of climate attributes.

Currently, the optimisation is formulated to be a single objective problem, and it was found that it requires calibration for each set of attributes and weather generators. This finding suggests that the use of multi-objective optimisation would be a way to avoid this calibration, warranting the investigation of different optimisation algorithms [*Maier et al.*, 2014]. Alternatively, the realism of the time series can also be considered in a less direct way. Constraints during the optimisation process can be used to ensure that, while the trade-off between attributes is being searched, all the time series created by the weather generators remain realistic.

### 5.2.2 Identifying critical climate variables

In Chapter 3, a method for identifying the critical climate attributes was introduced. This method relies in part on the cumulative variance explained metric, which indicates what fraction of the variance in performance is being accounted for by each successive selected attribute. The set of performance values is generated by sampling across a representative range of candidate climate attributes. However, a finding from Chapter 2 is that the extent to which the created climate time series will match the requested targets is something that will vary depending on the weather generator model structure and candidate attributes selected. In particular, different weather generators can apply different baselines of variability to created perturbed time series that match a set of attributes, and the effect of this on the cumulative variance explained metric was not explored. To enhance the understanding of this process, the critical attribute selection from a candidate set can be explored across different weather generators, with a few different case studies. Further applications will allow a greater understanding of the cumulative variance explained metric, and how strongly it indicates that a scenario-neutral analysis is not accounting for a key system mechanism. This will also allow a more detailed examination of which candidate attributes should be considered when using the proposed approach to identify the critical climate attributes.

It was also demonstrated in Chapter 3 that four critical climate attributes were needed to describe significant changes in both system performance metrics. A simple analysis demonstrated that conducting a scenario-neutral impact assessment with these critical climate attributes suggested different outcomes when compared to the status quo approach of defaulting to using changes in mean precipitation and temperature in an analysis. Given the finding that scenario-neutral spaces should consist of more complex attributes, the implications on decision making when using these higher dimensional scenario-neutral spaces should be explored, by applications to scenario discovery [*Lempert et al.*, 2008; *Bryant and Lempert*, 2010; *Kasprzyk et al.*, 2013], adaptation pathways [*Haasnoot et al.*, 2013; *Kwakkel et al.*, 2016], robustness calculations [*Whateley et al.*, 2014; *Giuliani and Castelletti*, 2016; *McPhail et*

*al.*, 2018] and other scenario-neutral approaches to considering uncertainty.

It was also found that the four critical attributes were different for each performance objective in the Lake Como application. Further, this research generally suggests increasing the number of attributes included in a scenario-neutral analysis (beyond mean precipitation and temperature) is necessary. However, this creates a challenge for the visualisation of scenario-neutral spaces, as interpretability decreases with higher dimensions. This challenge is present for single-objective system, and is compounded when considering multi-objective systems, given that as shown in the case study different objectives may be sensitive to different critical attributes. The clear visualisation of system performance in response to changes in climate was one of the initial benefits over the scenario-led approaches. Additional techniques need to be developed to maintain the effective visualisation of system performance while including all critical attributes for a multi objective system. The different sets of critical attributes for multiple objectives also need to be incorporated into the above-mentioned techniques for calculations of robustness and development of adaptation strategies.

### 5.2.3 Avoiding pitfalls of the scenario-neutral approach

In Chapter 4, a comparison between a scenario-neutral and scenario-led analysis was presented for the case study of Lake Como, in order to validate the scenario-neutral approach. However, only one replicate of the climate model projections used in the scenario-led approach was available for comparison, compared to the 50 replicates used in the scenario-neutral analysis. This meant that while pitfalls surrounding the selection of critical attributes and how well the created time series matched the requested ones were discussed, the effect of natural variability on both scenario-neutral and scenario-led analyses was not. A comparison of the scenario-neutral approach with a larger set of climate projection data can be used to further distinguish the effects of variability in a weather generator from missing a key system attribute, when considering the mismatch between scenario-led and scenario-neutral analyses. This will enable applications of the scenario-neutral approach to have more confidence that the key system

vulnerabilities have been identified.

Finally, the techniques to successfully identify key system vulnerabilities presented in this thesis should be applied to more complex water resource systems, beyond the Lake Como system, to test the methodological refinements and ensure the recommendations to avoid pitfalls are sufficiently robust. A limitation of this case study is a lack of stakeholder engagement, which is a core part of the scenario-neutral approach. Applied to other water resource systems, the methodologies presented in this thesis should combine with stakeholder engagement to define decision-relevant system performance metrics and failure thresholds. Additionally, any known climate attributes a system is vulnerable to can be included as candidate attributes, and any knowledge on the range these attributes should be stress-tested over can be used. This ensures that key vulnerabilities identified by the methods presented in this research are meaningful to the system.

## 5.3 Final Recommendations

This thesis prompts the following recommendations for future implementations of scenario-neutral impact assessments.

The improved inverse approach to stochastic time series creation is the only method currently available that can provide direct control over both the perturbations to be made as part of an impact assessment, and aspects of a climate time series to be held constant and ensure realism. This approach should be implemented for any systems that require a sufficiently complex stress-test against a wider range of climate attributes than just annual means. Otherwise, the analysis will be limited to a subset of potential future changes in climate, undermining the utility of an impact assessment.

Identifying the changes in climate to which a system is most sensitive is a necessary part of a scenario-neutral impact assessment. This analysis should be enacted for each performance criteria in a system, given the likely differences in sensitivities. As a result, care should be taken to combine multiple objectives in a single scenario-neutral analysis. Each objective may have different

requirements for a scenario-neutral space.

Given the challenges in successfully implementing a scenario-neutral approach, the analysis should be directly compared with the results of a scenario-led analysis. Any discrepancy between the results can indicate that an aspect of a climate time series is affecting system performance and is not currently accounted for, and therefore will not be the focus of any solutions developed. It is therefore through reconciling the scenario-neutral and scenario-led approaches that the most complete understanding of a water resource system can be obtained.

# References

Anghileri, D., F. Pianosi, and R. Soncini-Sessa (2011), A framework for the quantitative assessment of climate change impacts on water-related activities at the basin scale, Hydrol. Earth Syst. Sci., 15(6), 2025-2038.

Arnell, N. W. (2004), Climate change and global water resources: SRES emissions and socio-economic scenarios, Glob. Environ. Change-Human Policy Dimens., 14(1), 31-52.

Brekke, L. D., E. P. Maurer, J. D. Anderson, M. D. Dettinger, E. S. Townsley, A. Harrison, and T. Pruitt (2009), Assessing reservoir operations risk under climate change, Water Resources Research, 45(4), W04411.

Brown, C., and R. L. Wilby (2012), An alternate approach to assessing climate risks, Eos, Transactions American Geophysical Union, 93(41), 401-402.

Brown, C., Y. Ghile, M. Laverty, and K. Li (2012), Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, Water Resources Research, 48(9), W09537.

Bryant, B. P., and R. J. Lempert (2010), Thinking inside the box: A participatory, computer-assisted approach to scenario discovery, Technological Forecasting and Social Change, 77(1), 34-49.

Bussi, G., S. J. Dadson, C. Prudhomme, and P. G. Whitehead (2016), Modelling the future impacts of climate and land-use change on suspended sediment transport in the River Thames (UK), Journal of Hydrology, 542, 357-372.

Culley, S., S. Noble, A. Yates, M. Timbs, S. Westra, H. Maier, M. Giuliani, and A. Castelletti (2016), A bottom- up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate, Water Resources Research, 52(9), 6751-6768.

Ekström, M., E. D. Gutmann, R. L. Wilby, M. R. Tye, and D. G. C. Kirono (2018), Robustness of hydroclimate metrics for climate change impact research, Wiley Interdisciplinary Reviews: Water, 5(4), e1288.

Gao, L., B. A. Bryan, M. Nolan, J. D. Connor, X. Song, and G. Zhao (2016), Robust global sensitivity analysis under deep uncertainty via scenario analysis, Environ. Modell. Softw., 76, 154-166.

Giuliani, M., and A. Castelletti (2016), Is robustness really robust? How different definitions of robustness impact decision-making under climate change, Climatic Change, 1-16.

Guo, D., S. Westra, and H. R. Maier (2017), Use of a scenario-neutral approach to identify the key hydro-meteorological attributes that impact runoff from a natural catchment, Journal of Hydrology, 554, 317-330.

Guo, D., S. Westra, and H. R. Maier (2018), An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies, Journal of Hydrology, 556, 877-890.

Haasnoot, M., J. H. Kwakkel, W. E. Walker, and J. ter Maat (2013), Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world, Global Environmental Change, 23(2), 485-498.

Herman, J. D., P. M. Reed, H. B. Zeff, and G. W. Characklis (2015), How should robustness be defined for water systems planning under change?, Journal of Water Resources Planning and Management, 141(10), 04015012.

IPCC (2013), Summary for Policymakers. 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, IPCC.

IPCC (2014), Summary for Policymakers. In: Climate Change 2014: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, IPCC.

Kasprzyk, J. R., S. Nataraj, P. M. Reed, and R. J. Lempert (2013), Many objective robust decision making for complex environmental systems undergoing change, Environ. Modell. Softw., 42, 55-71.

Keller, L., O. Rössler, O. Martius, and R. Weingartner (2018), Comparison of scenario-neutral approaches for estimation of climate change impacts on flood characteristics, Hydrological Processes, 0(0).

Kingston, G. B., G. C. Dandy, and H. R. Maier (2008), AI techniques for hydrological modeling and management, edited, pp. 67-69, Nova.

Kwakkel, J. H., W. E. Walker, and M. Haasnoot (2016), Coping with the Wickedness of Public Policy Problems: Approaches for Decision Making under Deep Uncertainty, Journal of Water Resources Planning and Management, 01816001.

Lempert, R. J., B. P. Bryant, and S. C. Bankes (2008), Comparing algorithms for scenario discovery, RAND, Santa Monica, CA.

Maier, H. R., J. H. A. Guillaume, H. van Delden, G. A. Riddell, M. Haasnoot, and J. H. Kwakkel (2016), An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together?, Environ. Modell. Softw., 81, 154-164.

Maier, H. R., S. Razavi, Z. Kapelan, L. S. Matott, J. Kasprzyk, and B. A. Tolson (2019), Introductory overview: Optimization using evolutionary algorithms and other metaheuristics, Environ. Modell. Softw., 114, 195-213.

Maier, H. R., et al. (2014), Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions, Environ. Modell. Softw., 62(0), 271-299.

McPhail, C., H. Maier, J. Kwakkel, M. Giuliani, A. Castelletti, and S. Westra (2018), Robustness metrics: How are they calculated, when should they be used and why do they give different results?, Earth's Future, 6(2), 169-191.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer (2008), Climate change - Stationarity is dead: Whither water management?, Science, 319(5863), 573-574.

Nazemi, A., and H. Wheater (2014), Assessing the vulnerability of water supply to changing streamflow conditions, Eos, Transactions American Geophysical Union, 95(32), 288-288.

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 Resources Research, 49(3),

1643-1667.

Prudhomme, C., A. L. Kay, S. Crooks, and N. Reynard (2013), Climate change and river flooding: Part 2 sensitivity characterisation for British catchments and example vulnerability assessments, Climatic Change, 119(3-4), 949-964.

Prudhomme, C., R. L. Wilby, S. Crooks, A. L. Kay, and N. S. Reynard (2010), Scenario-neutral approach to climate change impact studies: Application to flood risk, Journal of Hydrology, 390(3-4), 198-209.

Ray, P. A., M. Ü. Taner, K. E. Schlef, S. Wi, H. F. Khan, S. S. G. Freeman, and C. M. Brown (2018), Growth of the Decision Tree: Advances in Bottom-Up Climate Change Risk Management, JAWRA Journal of the American Water Resources Association, 0(0).

Risbey, J. S. (2011), Dangerous climate change and water resources in Australia, Reg. Envir. Chang., 11, S197-S203.

Singh, R., T. Wagener, R. Crane, M. Mann, and L. Ning (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(4), 3409-3427.

Spence, C., and C. Brown (2018), Decision Analytic Approach to Resolving Divergent Climate Assumptions in Water Resources Planning, Journal of Water Resources Planning and Management, 144(9), 04018054.

Steinschneider, S., and C. Brown (2013), A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments, Water Resources Research, 49(11), 7205-7220.

Taner, M. Ü., P. Ray, and C. Brown (2017), Robustness-based evaluation of hydropower infrastructure design under climate change, Climate Risk Management, 18, 34-50.

Turner, S. W. D., D. Marlow, M. Ekström, B. G. Rhodes, U. Kularathna, and P. J. Jeffrey (2014), Linking climate projections to performance: A yield-based decision scaling assessment of a large urban water resources system, Water Resources Research, 50(4), 3553-3567.

Vano, J. A., M. J. Scott, N. Voisin, C. O. Stockle, A. F. Hamlet, K. E. B. Mickelson, M. M. Elsner, and D. P. Lettenmaier (2010), Climate change impacts on water management and irrigated agriculture in the Yakima River Basin, Washington, USA, Climatic Change, 102(1-2), 287-317.

Weiß, M. (2011), Future water availability in selected European catchments: a probabilistic assessment of seasonal flows under the IPCC A1B emission scenario using response surfaces, Natural Hazards and Earth System Sciences, 11(8), 2163-2171.

Wetterhall, F., L. Graham, J. Andreasson, J. Rosberg, and W. Yang (2011), Using ensemble climate projections to assess probabilistic hydrological change in the Nordic region, Natural Hazards and Earth System Sciences, 11(8), 2295-2306.

Whateley, S., S. Steinschneider, and C. Brown (2014), A climate change range-based method for estimating robustness for water resources supply, Water Resources Research, 50(11), 8944-8961.

Wilby, R. L., and S. Dessai (2010), Robust adaptation to climate change, Weather, 65(7), 180-185.

Wilby, R. L., C. Dawson, C. Murphy, P. O'Connor, and E. Hawkins (2014), The Statistical DownScaling Model -Decision Centric (SDSM-DC): Conceptual basis and applications, Climate Research, 61.

Wilcke, R. A., and L. Bärring (2016), Selecting regional climate scenarios for impact modelling studies, Environ. Modell. Softw., 78, 191-201.

# Appendix A

## *Chapter 2 Journal Paper as published in Journal of Hydrology*

Culley, S., B. Bennett, S. Westra, and H. R. Maier (2019), Generating realistic perturbed hydrometeorological time series to inform scenario-neutral climate impact assessments, Journal of Hydrology, 576, 111-122.

Journal of Hydrology 576 (2019) 111–122

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

# Generating realistic perturbed hydrometeorological time series to inform scenario-neutral climate impact assessments

S. Culley*, B. Bennett, S. Westra, H.R. Maier

*School of Civil, Environmental and Mining Engineering, University of Adelaide, Adelaide, South Australia 5005, Australia*

ARTICLE INFO

This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Saman Razavi, Associate Editor

*Keywords:*
Scenario-neutral
Stochastic weather generation
Bottom-up
Inverse approach
Climate change impact assessment
Realistic perturbed hydrometeorological time series

ABSTRACT

Scenario-neutral approaches are used increasingly as a means of stress-testing climate-sensitive systems to a range of plausible future climate conditions. To ensure that these stress-tests are able to explore system vulnerability, it is necessary to generate hydrometeorological time series that represent all aspects of plausible future change (e.g. averages, seasonality, extremes). A promising approach to generating these time series is by inverting the stochastic weather generation problem to obtain weather time series that capture all the relevant statistical features of plausible future change. The objective of this paper is to formalize this "inverse" approach to weather generation, by both characterizing the process of optimizing weather generator parameters and proposing a numerically efficient solution that exploits prior knowledge and accounts for the complexity of the optimization landscape. The proposed approach also provides a structured way to ensure the physical realism of the generated weather time series, by using penalty-based objective functions to focus the optimization on the climate features deemed most relevant to the system being analyzed. A case study in Adelaide, Australia, is used to demonstrate specific implementations of this approach. The use of bounds on the weather generators dramatically decreases the time taken to create time series, and the use of penalties is shown to allow for change in some statistics to be prioritized, while still ensuring the realism of the time series.

## 1. Introduction

Scenario-neutral climate impact assessments are proving to be an effective way of assessing how a range of climate-sensitive systems might respond to plausible future climate changes. The scenario-neutral approach has been applied recently to flood protection, water supply and ecological systems (Culley et al., 2016; Poff et al., 2016; Prudhomme et al., 2013), with these studies demonstrating that scenario-neutral approaches both lead to important insights into overall system sensitivities and vulnerabilities, and enable the identification of possible failure modes by determining how a system responds to step changes in climate (Prudhomme et al., 2010). These approaches are also increasingly being used to provide decision-theoretic information, describing conditions whereby one system configuration or design option is preferred to another (Brown et al., 2012), or approximating the maximum operational adaptive capacity of the system (Culley et al., 2016).

Although most scenario-neutral approaches have focused on changes in the mean state of climate variables (Culley et al., 2016; Prudhomme et al., 2013; Prudhomme et al., 2010; Spence and Brown, 2018; Wilcke and Bärring, 2016), it is becoming increasingly apparent that critical system vulnerabilities may reside in other aspects of change—including variability, intermittency, extremes, seasonality and/or inter-annual persistence (Bussi et al., 2016; Guo et al., 2017; Steinschneider and Brown, 2013). An important implication is that if key sensitives are not identified, major modes of system vulnerability may not be uncovered, thereby negating the stated benefit of scenario-neutral studies. This poses a deep challenge to the viability of scenario-neutral approaches: how should weather and hydrometeorological time series be generated to capture all possible aspects of future change?

The primary approach currently available to address this challenge within the scenario-neutral framework is through the use of stochastic weather generators, which contain sufficient flexibility to simulate a wide variety of possible future changes while maintaining the statistical features commonly associated with weather time series (Guo et al., 2018; Steinschneider and Brown, 2013). The forward scaling approach presented by Steinschneider and Brown (2013) is capable of manipulating more complex measurements of precipitation, like persistence, by directly perturbing the parameters of a weather generator to generate baseline time series with different statistics. However, to produce the uniform perturbations to climate attributes required in a scenario-neutral assessment, this still requires some post-processing of time

* Corresponding author.
*E-mail address:* sam.culley@adelaide.edu.au (S. Culley).

https://doi.org/10.1016/j.jhydrol.2019.06.005
Received 26 September 2018; Received in revised form 8 April 2019; Accepted 4 June 2019
Available online 06 June 2019

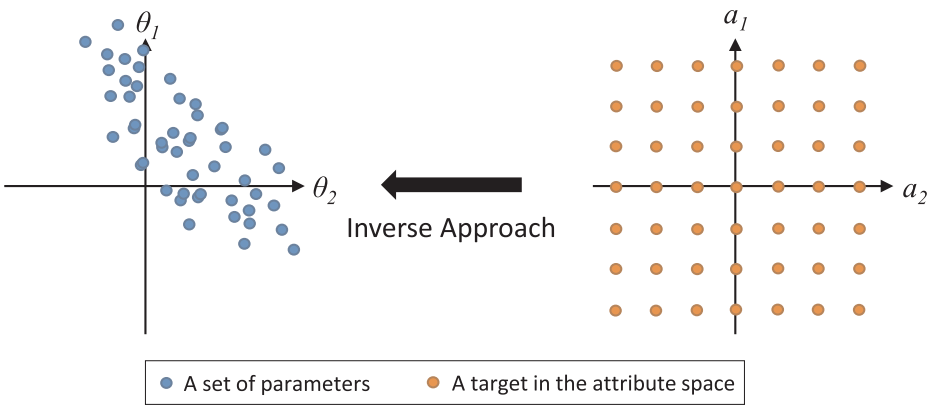


**Fig. 1.** Conceptual illustration of the inverse approach to stochastic generation. Each blue dot corresponds to a set of weather generator parameters that result in the generation of weather time series that have the set of climate attributes of one of the orange dots. Equal coverage of the climate attribute space ($a_1$, $a_2$) does not necessarily reflect equal coverage in the parameter space ($\theta_1$, $\theta_2$) of stochastic weather generators.

series, such as quantile mapping, to represent targeted changes in climate. Given scaling methods are still used, the range of attributes that can be perturbed is often limited to means and seasonality, by directly adjusting the baseline time series. To avoid the scaling process, Guo et al. (2018) provided the first structured attempt at inverting the stochastic generation problem, by varying the parameters of a stochastic generator through an optimization loop to simulate weather time series with pre-specified statistics or "attributes". The generation of time series using this approach was demonstrated for three climate variables (precipitation, temperature and evapotranspiration) (Guo et al., 2017) and was benchmarked on a rainfall dataset from a catchment in South Australia (Guo et al., 2018).

As scenario-neutral approaches are applied to increasingly complex systems, it becomes necessary to explore increasing numbers of hydrometeorological variables and statistics of those variables. For example, whereas Culley et al. (2016) focused on annual average rainfall and temperature in their case study on Lake Como flood management and irrigation requirements, it is likely that a thorough exploration of system vulnerability for this alpine lake would require exploration of attributes that affect features such as snow pack and snow melt rates, evaporation from the reservoir and evapotranspiration from the irrigation demand regions. These attributes could include winter precipitation amounts, the number of frost days in the year, growing season length, and so on. Stress testing the system to each individual change—and all the possible combinations of those changes—poses substantial numerical and computational barriers to the inversion problem. For example, the required runtimes indicated by Guo et al. (2018) for a simple application of three attributes (e.g. 8 h for producing 100 simulated weather time series using 8 cores) suggest significant potential challenges for the widespread application of the inverse generation method, and requires a structured approach for identifying opportunities for computational efficiencies. Consequently, there is a need to reduce the run times of the optimization loop that underpins the inversion process so that it can be applied to more complex systems within practical timeframes.

A further challenge is that, as the number of attributes to be perturbed increases, the likelihood of attempting to simulate infeasible changes will also increase. For example, consider the relationship between the attributes average annual rainfall, average rainfall intensity and average number of wet days. Given any two values of those attributes are held constant, there is only one value the third can take, and it is not numerically possible to simulate time series with further increases or decreases to that third attribute. This is particularly important when seeking to generate weather time series that capture specific changes, while seeking to match historical climate patterns in all other aspects to maintain physical realism. Consequently, there is a need to manage which attributes of a time series simulated as part of the inversion process achieve the requested change, in the event the requested change is infeasible.

The overarching objective of this paper is therefore to formalize the inversion problem, by focusing on two specific aims: (1) improving the computational efficiency of the optimization process; and (2) ensuring the physical realism of the simulated time series. The remainder of the paper proceeds as follows. Section 2 formally articulates the aims and details general approaches to meet them. Section 3 describes the case study of Adelaide, Australia, followed by Sections 4 and 5 that focus on the first and second aims, respectively, where the general approach is implemented for the case study, and is then tested to examine the impact of formalizing the inverse approach as presented in this paper. Conclusions are presented in Section 6.

## 2. Formalizing the inverse approach to stochastic generation

### 2.1. Overview of the inverse approach

Guo et al. (2018) presented the inverse approach as a technique to generate hydrometeorological time series that satisfy a set of target changes in specified climate attributes. In this context, 'attributes' are defined as statistics of particular hydrometeorological variables, such as the mean annual rainfall or number of wet days. The approach starts by setting targets $\boldsymbol{t}_j \in \mathbb{R}^n$, where $n$ is the number of attributes considered, and $j = 1, ...,m$ represents the number of target values of those attributes. The target changes may be represented as absolute values (e.g. simulating a time series with annual average rainfall of 960 mm), or alternatively they may be represented in terms of the percentage or absolute changes in attributes relative to historical climate (e.g. a 10% decrease in annual average rainfall, or 3 °C increase in average annual temperature). The weather time series can be generated by changing only a single attribute at a time, or by simulating combinations of changes; for example Guo et al. (2018) simulated changes in two attributes over a regular grid.

Once the attribute targets are identified, the next step is to apply a formal optimization approach that involves modifying the parameter vector $\theta$ of some stochastic generator $g(\theta)$ that minimizes a measure between the relevant attributes $\boldsymbol{a}_j \in \mathbb{R}^n$ of the simulated weather time series and the target attributes ($\boldsymbol{t}_j$). This is illustrated in Fig. 1, whereby the target attributes are represented here in two dimensions ($n = 2$) over a regular 7 × 7 grid (i.e. $j = 1, ...,49$), in terms of a fraction or percentage change relative to a historical baseline. For each target, the inverse approach then adjusts parameter vector $\theta$ in order to achieve weather time series with desired attributes. This process is also represented mathematically as:

$$weather.\ ts_j = g\left(\arg\min_{\theta} O(\boldsymbol{a}_j, \boldsymbol{t}_j)\right) \quad (1)$$

where $O(\boldsymbol{a}_j, \boldsymbol{t}_j)$ describes a general measure of the difference between each weather attribute and its target. For example, in the case of Guo et al. (2018), a simple Euclidean distance measure was used:

$$O(\boldsymbol{a}_j, \boldsymbol{t}_j) = \sqrt{\sum_{i=1}^{n} (a_{i,j} - t_{i,j})^2} \quad (2)$$

It is noted that the process of achieving the weather time series as described in Eq. (1) is inherently iterative; namely the attributes $\boldsymbol{a}$ are calculated from the weather time series in the previous optimization step, until a stopping criterion is reached.

Although conceptually straightforward, there are two key challenges with the approach:

1. How to design an efficient optimization process that extends the inverse approach to high-dimensional spaces with high levels of accuracy and minimal runtimes (Section 2.2). In particular, computational issues were identified by Guo et al. (2018) as a significant challenge, and in its current form is likely to inhibit wider application of the inverse methodology.
2. How to ensure the realism of the weather time series (Section 2.3). A feature of the inverse approach is that any desired properties of the climate time series need to be included in the objective function. This provides an incentive to include a greater number of attributes in the objective function to maintain realism, increasing problem complexity, and the likelihood that an infeasible combination of attributes will be requested. A traditional Euclidean distance objective function does not provide a sufficiently robust approach for prioritizing some attributes above others, which is necessary when not all target changes can be met.

The following sections explore these two challenges in more detail.

### 2.2. Improving the computational efficiency of the optimization process

The generic steps in the optimization loop that underpins the inverse approach are shown in Fig. 2, which consists of an iterative process for updating the parameters of a weather generator, $\theta_k$, until certain stopping criteria have been met. The approach to updating depends on the specific choice of optimization method (e.g. gradient descent versus stochastic searches), but all methods aim to improve the objective function value $O(\boldsymbol{a}_j, \boldsymbol{t}_j)$, which for this case consists of a measure of distance between the generated weather time series attributes and the target attributes. Potential stopping criteria for the optimization loop include the completion of a fixed number of iterations, stagnation in the optimization process or sufficiently small errors between the attribute values of the time series generated from the weather generator and the target attributes (Zielinski et al., 2005).

The following two approaches can be used to improve the computational efficiency of the above processes: (i) selecting the optimization algorithm that is most suited to the characteristics of the optimization problem, and (ii) reducing the size of the search space as much as possible, without restricting the ability to identify the desired solutions.

For the first approach, it is necessary to diagnose the nature of the optimization fitness landscape—the relationship between the decision variables and the objective function—as this is critical for identifying the most efficient optimization algorithm for the class of problem to be tackled (Maier et al., 2019). For example, smooth fitness landscapes may enable computationally efficient hill climbing algorithms to find the global optimum (e.g. Nesterov, 2007), whereas irregular fitness landscapes require stochastic methods (Kingston et al., 2008). In low dimensional problems, an enumeration methodology can be used to visualize the fitness landscape and examine its properties directly, whereas for higher-dimensional problems, the use of fitness landscape statistics that identify properties like the overall structure of the fitness landscape, any flat areas of the same function value, and the distance between good local optima and the global solution might be required (Gibbs et al., 2011; Gibbs et al., 2015; Maier et al., 2014; Malan and Engelbrecht, 2013). Given that the fitness landscape is defined by the weather generators and attributes of interest, the most suitable optimization algorithm is likely to depend on the specific implementation of the inverse approach.

A challenge for the second approach is that, for all but the simplest problems, it is generally not possible to know *a priori* how the stochastic generation parameter vector $\theta$ maps into the attribute space (Fig. 1). This makes it difficult to provide bounds on the weather generator parameters, which are necessary for some optimization algorithms, such that the bounds do not unintentionally prevent some requested time series from being generated. In order to address this issue, Guo et al. (2018) used very wide bounds on the weather generator parameters during the optimization process. However, this approach produces very large search spaces, which can result in significant increases in the computational effort associated with identifying the desired parameter values. An alternative approach used in this study is to refine the bounds of the weather generator parameters by assessing typical ranges of stochastic weather generator parameters applied to a broad set of current weather time series across a large geographic area, under the assumption that there are likely to be current weather "analogues" (e.g. weather time series in very warm and arid regions) that are representative of plausible future changes as a result of anthropogenic

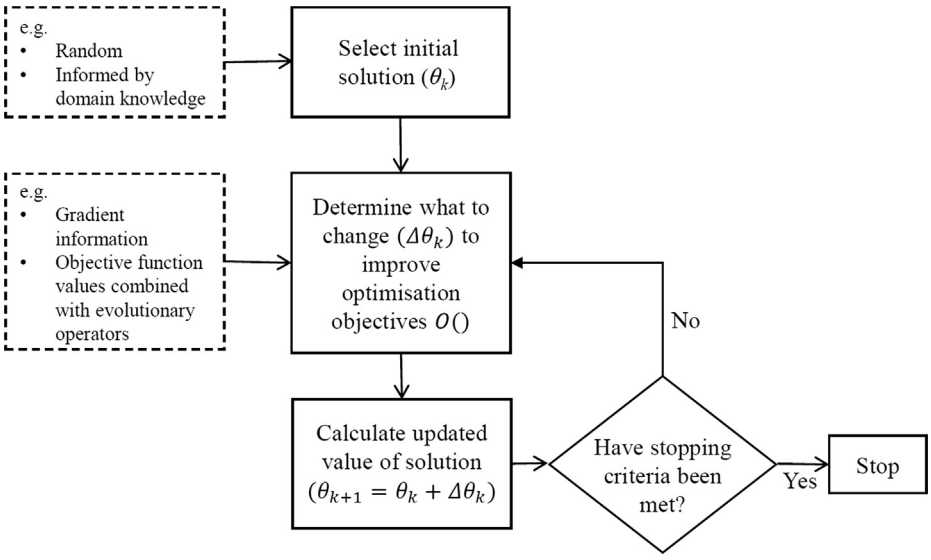


**Fig. 2.** Generic steps in the optimization loop that underpins the inverse approach.

climate change.

### 2.3. Ensuring the physical realism of the simulated time series

The weather time series to be generated using the inverse approach are synthetic series and are thus not constrained by physical processes in the same manner as time series generated by weather and climate models. For example, if a target is to increase total annual rainfall by 15%, then it would be theoretically possible for the weather generator to produce time series whereby all the annual rainfall occurs within one season, or occurs uniformly across a whole year, or any other possible series that meets the total annual rainfall target.

The proposed conceptual approach for addressing this issue and ensuring physical realism is to generate time series that represent the proposed target changes, but with all other aspects of the weather time series held at historical values. This is achieved by including a larger number of target attributes within the optimization process, by focusing on both "perturbed" attributes that represent the primary objective of the optimization, and "held" attributes that keep all other aspects of the weather time series as close to their historical values as possible. This substantially increases the complexity of the optimization problem, by increasing the number of attributes $n$ that need to be considered as part of Eq. (1), further highlighting the importance of reducing the size of the search space through other means, as discussed in Section 2.2.

Beyond the computational challenges associated with this increase in optimization complexity, there is a more fundamental problem: in many cases, setting a large number of both "perturbed" and "held" attributes will lead to requests for infeasible attribute combinations. Returning to the example of increasing annual average rainfall by 15%, we might seek to achieve this while holding the number of wet days, the amount per wet day and any other aspects of the annual rainfall time series at their historical values. However, this combination is not possible: increasing annual rainfall *can only be achieved* through either increasing the number of wet days, or the amount of rainfall per wet day, or some combination of the two.

The nature of the over-constrained optimization problem is illustrated in Fig. 3, where we plot the feasible subspace of total annual rainfall, number of wet days and amounts per wet day in the three-dimensional space of possible attribute changes. This subspace is

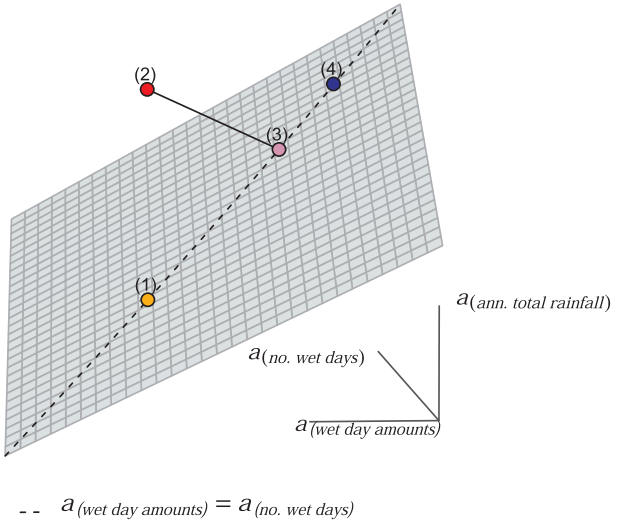


**Fig. 3.** The over-constrained optimization challenge: the grey surface indicates the feasible subspace, with equation $a_{(ann.\ total\ rainfall)} = a_{(no.\ wet\ days)} * a_{(wet\ day\ amounts)}$. Historical conditions are shown at point (1) and the target is indicated by point (2). Point (3) indicates the shortest Euclidean distance between the target at point (2), and the feasible subspace. Point (4) indicates the shortest distance between the target and the feasible subspace, while ensuring zero error in the perturbed attribute.

calculated assuming the total annual rainfall is the product of the number of wet days and the average wet day amount ($a_{(ann.\ total\ rainfall)} = a_{(no.\ wet\ days)} * a_{(wet\ day\ amounts)}$). As an example, we seek to increase total annual rainfall (the "perturbed" attribute) by 15% from its historical value (leading from point (1), which represents the historical conditions, to point (2)), which if the two other attributes are "held" at their historical values, is an infeasible target (i.e. point (2) does not lie within the feasible subspace).

If a Euclidean distance objective function (Eq. (2)) is used in the optimization process (see Fig. 2) the solution indicated by point (3) will be identified. Point (3) is the solution closest to the target that lies on the feasible subspace (i.e. point (3) is the orthogonal projection of point (2) onto the feasible subspace). However, this solution only increases the total annual rainfall by ~10%, and thus does not produce rainfall time series with the desired 15% change. It also leads to a 5% increase in the number of wet days and amounts per wet day. An alternative solution is to modify the objective function to place more emphasis on the "perturbed" attribute, and in so doing find a solution that produces the desired change in the "perturbed" attribute but keeps the remaining "held" attributes as close as possible to historical values. In this simple illustration, this might lead to total annual rainfall increasing by 15%, as originally sought, by both increasing the number of wet days and amounts per wet day each by 7% (see point (4)).

As the objective function is used to measure how close the attributes of the simulated time series are to the intended targets, this also needs to manage any trade-offs between attributes. This suggests the use of penalties (Coello Coello, 2002) to modify the objective function to favor solutions with smaller errors in "perturbed" attributes. Here, the modification of the objective function (Eq. (2)) is discussed with reference to two general penalty structures: a linear penalty structure that adds a linear term based on the error in the "perturbed" attributes (Eq. (3)) and a quadratic structure that adds a squared term based on the error in the "perturbed" attributes (Eq. (4)) (note that this equation can be rearranged as a weighted sum). The two modified objective functions are given by

$$O(\boldsymbol{a}_j, \boldsymbol{t}_j) = \sqrt{\sum_{i=1}^{n} (t_{i,j} - a_{i,j})^2} + \sum_{k=1}^{p} \lambda_k * abs(t_{k,j} - a_{k,j}) \quad (3)$$

$$O(\boldsymbol{a}_j, \boldsymbol{t}_j) = \sqrt{\sum_{i=1}^{n} (t_{i,j} - a_{i,j})^2 + \sum_{k=1}^{p} (\lambda_k - 1) * (t_{k,j} - a_{k,j})^2} \quad (4)$$

where $k = 1, ...,p$ represents the subset of $n$ "perturbed" attributes (i.e. $p \leqslant n$), $\lambda$ are the scaling parameters applied to the errors in the "perturbed" attributes. The remaining notation is consistent with Eq. (2). Eqs. (3) and (4) reduce to the "unweighted" Euclidean distance objective function (Eq. (2)) for $\lambda_k = 0$ and 1, respectively.

The effect of these penalties and scaling parameter values can be illustrated with a continuation of the example in Fig. 3, where we seek to increase total annual rainfall by 15% and hold the number of wet days and amounts per wet day at historical values. Here, instead of viewing the three-dimensional space of possible attribute changes, a two-dimensional slice through the space is shown in Fig. 4, such that each panel displays the feasible solution subspace as a black line (this slice was represented by the dashed line in Fig. 3). The over-constrained target is represented by a red point and the contours represent the objective function value for each attribute combination in the 2D space (i.e. the fitness landscape). The minimum error solution, represented by the blue point, occurs where the smallest objective function value (i.e. 2D fitness landscape) intersects the feasible subspace (black line).

To demonstrate how the fitness landscape and minimum error solution change with different penalty structures and scaling parameters, $\lambda$, Fig. 4 compares different $\lambda$ values for both a linear penalty structure (top panels) and a quadratic penalty structure (bottom panels). Moving left to right, the panels in Fig. 4 illustrate the effect of increasing the scaling parameter, $\lambda$, on the "perturbed" attribute (annual total rainfall)

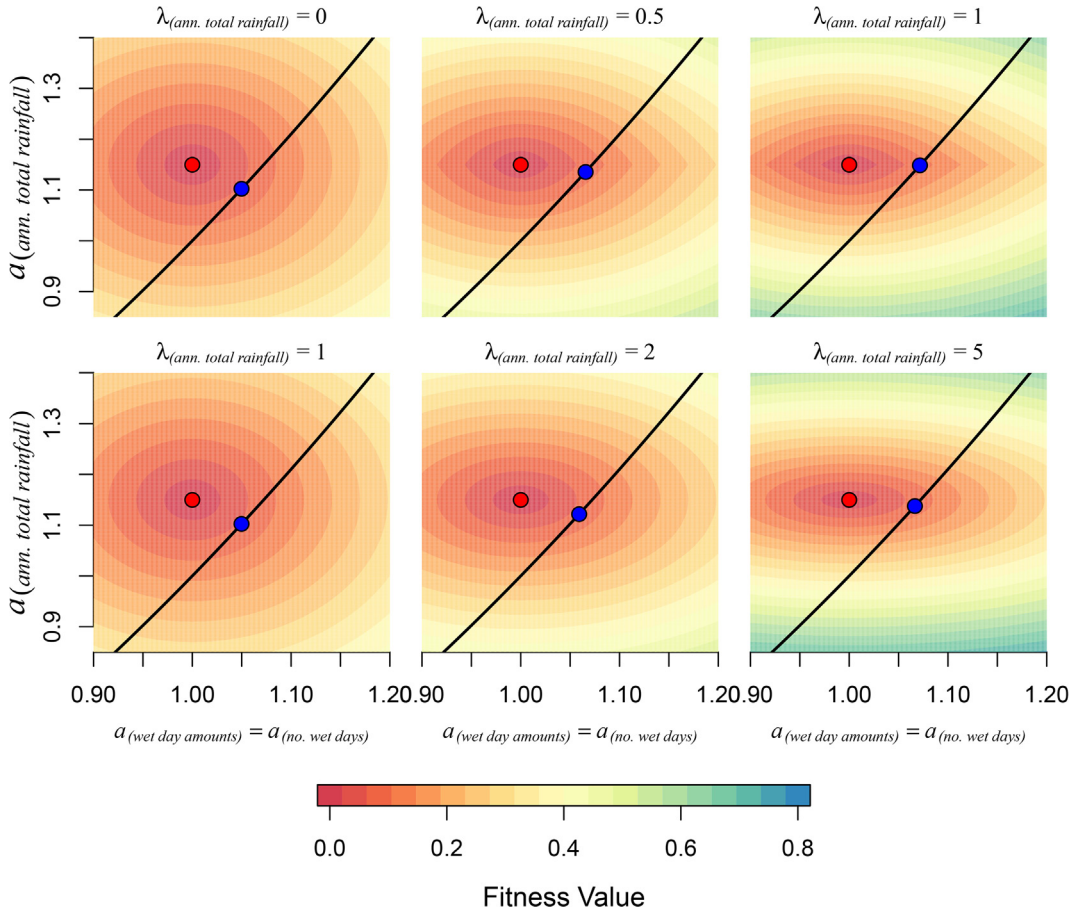


**Fig. 4.** How penalty functions can change the fitness landscape to create a new minimum error solution during over-constrained optimization: (top panels) linear penalty term, and (bottom panels) quadratic penalty term.

in terms of the change in the fitness landscape and thus the location of the minimum error solution (blue point) for the two penalty structures. For the case where all attributes are equally weighted within the objective function (left most panels of Fig. 4), the fitness landscape contours are circular and the identified minimum error solution increases the perturbed attribute by approximately 10%. As the scaling parameter, λ, is increased, the minimum error solution is moved along the feasible subspace line towards the solution with zero error in the "perturbed" attribute.

The rate at which the minimum error solution moves along the feasible line with change in a scaling parameter, λ, is dependent on the penalty structure. The linear penalty term (top panels) has the capacity to identify a solution with zero error in the "perturbed" attribute if the scaling parameter is sufficiently large. In contrast, the quadratic penalty term (bottom panels) exhibits asymptotic behaviour such that as the scaling parameter increases, the minimum error solution will get closer to the zero error solution in the "perturbed" attribute but will never intersect it. The choice of penalty is influenced by the problem application. For example, where it is important to meet the "perturbed" attribute target, the linear penalty term may be appropriate. However, if the "held" attributes also have a substantial impact on system performance, the quadratic penalty may be more appropriate.

It is noted that the illustrative example described in Figs. 3 and 4 is highly conceptual, and most widely used weather generators have much greater complexity to enable them to simulate the statistical features of realistic weather time series. The capacity to achieve specified target attribute combinations will be limited both by physical constraints (as illustrated in Figs. 3 and 4) and the ability of the weather generator to simulate the requisite combinations. For example, an annual Markov model would not be capable of simulating seasonal variability in various rainfall statistics, thereby leading to infeasible targets if the objective is to simulate seasonal variability. Conversely, overly complex weather generators would lead to a much higher-dimensional parameter set, $\theta$, as well as the need to constrain a larger number of attributes to ensure physical realism of the generated series, placing more burden on the optimization process. Care is therefore needed to ensure that a weather generator of appropriate complexity is selected to achieve the objectives of each investigation.

## 3. Case study

The issues and proposed solutions highlighted in the previous section are illustrated using rainfall data from a location in Adelaide, Australia. The region has a Mediterranean climate with an annual average rainfall of 532 mm. The rainfall for this region is highly seasonal with most rainfall occurring during winter (June, July and August) and spring (September, October and November) and the least rainfall occurring in the summer season (December, January and February). Historical rainfall time series for Adelaide (34.92°S, 138.62°E) were obtained from the Australian Water Availability Project (AWAP) dataset (Raupach et al., 2012). To minimise the influence of changing trends in rainfall, the period 1970–1999 was selected, since this period is relatively stationary.

Two stochastic daily weather generators are used in this study that follow the precipitation component of WGEN: (i) a simple four parameter model to aid in theoretical understanding of the optimization fitness landscape, and (ii) a more complex model to investigate how well the proposed developments work for more realistic applications

(Richardson and Wright, 1984). The simple weather generator model used has only four parameters. Two parameters control the wet/dry sequence throughout the time series using a 1st order Markov chain. *Pdd* is the probability of a dry day given a dry day occurred previously, and *Pwd* is the probability of a dry day given a wet day occurred previously. For any two values of these parameters, the supplementary parameters *Pdw* and *Pww* are calculated, and sequences of wet and dry days for the length of the time series are obtained using a random number generator. The remaining two parameters control the amount of rainfall that occurs on wet days. These are the shape and rate parameters of the gamma distribution, $\alpha$ and $\beta$, from which each wet day rainfall amount is randomly sampled. Given the parameters do not vary throughout the year, this weather generator can only produce stochastic rainfall time series to meet a range of climate attributes measured at an annual level. It is therefore referred to throughout as the "annual" weather generator.

To perturb intra-annual attributes, a more complex weather generator is needed, with additional parameters and hence greater degrees of freedom to produce the required time series. The method used in this study is to extend the parameters of the simple model, where each of the original four parameters is specified as varying throughout the year. A harmonic model is used to control this variation, dictated by the mean, amplitude and phase angle of a harmonic (e.g. *Pdd* becomes *Pdd-m*, *Pdd-amp* and *Pdd-phase*) (Richardson, 1981). This is the same approach Richardson (1981) used to create temperature and solar radiation time series; however, in this application the harmonic model is not creating the time series directly, but describing what values the parameters should take. The harmonic models are fixed to have 12 periods, allowing for each of the four annual WGEN parameters to take different monthly values throughout each year. This allows the perturbation of attributes at the seasonal level, and this model is referred to throughout as the "seasonal" weather generator.

Only eight parameters are used as decision variables for this seasonal weather generator, as the phase angle parameters are fixed at historical values, leaving just the mean and amplitude for *Pdd*, *Pwd*, $\alpha$ and $\beta$. The calibration process outlined by Richardson (1981) was used to determine the values of the four phase angle parameters for the case study site of Adelaide. The *Pdd*, *Pwd*, $\alpha$ and $\beta$ phase angles were 0.355, 0.232, 3.53 and 2.46, respectively. This modification to the seasonal weather generator maintains the seasonal pattern in the generated time series such that most rainfall occurs in winter and spring, but still allows for the actual rainfall volume in each season to be perturbed separately. The trade-off with this new model is the large increase in computational effort required to find a solution given the increased search space.

The attribute sets used in the implementation of the inverse approach are listed in Table 1 for each weather generator type. Throughout this paper, *Ptot* is selected as the "perturbed" attribute, given its common usage in scenario-neutral impact studies, except in two instances designed to investigate applications for multiple "perturbed" attributes, where *nWet* is also selected. The remaining "held" attributes are included in the objective function for each simulation to ensure that these properties are maintained in the perturbed time series (as discussed in Section 2.3). Given the change in model complexity, different sets of attributes are specified for each weather generator. The seasonal model requires more "held" attributes, given the extra degrees of freedom provided by the parameters.

## 4. Investigation into the impact of increasing the efficiency of optimization

The following section contains a specific implementation of the approach for increasing the efficiency of the optimization process proposed in Section 2.2. The optimization problem is analyzed for the case study, leading to the selection of an optimization algorithm that is suited to the fitness landscape (Section 4.1). The optimization process is then implemented on the case study with improvements to optimization efficiency due to restricting the bounds of the weather generator parameters (Section 4.2).

### 4.1. Selecting a suitable optimization algorithm

In order to determine the most appropriate optimization algorithm for the case study application, the nature of the fitness landscape is analyzed. This is done for the annual weather generator, as this enables an enumeration method to be used to generate the landscape. As a fitness landscape for a model with four parameters is five-dimensional, with the fifth dimension being the objective function value (i.e. the "fitness"), the complete fitness landscape is unable to be visualized. Consequently, in order to enable key aspects of the overall fitness landscape to be inspected, separate three-dimensional fitness landscapes are generated for the shape and rate parameters ($\alpha$ and $\beta$, respectively) and the probability of wet-dry and dry-dry parameters. The required fitness values are calculated using the "unweighted" objective function (Eq. (2)) for three attributes all held at historical levels: *Ptot*, *nWet* and *P99*. As part of this process, the random seed used in the weather generator is held constant to maintain a set relationship between parameters and attribute values (Guo et al., 2018).

The resulting fitness landscapes are shown in Fig. 5. As can be seen, the fitness landscape in the left panel is smooth, as the gamma function from which rainfall amounts are sampled is continuous, so that changes to the gamma distribution parameters result directly in changes in rainfall volume (as a fixed random seed was used to eliminate stochastic "noise", as mentioned above). In contrast, the fitness landscape on the right is rough with many local optima. This is due to discrete changes in the response surface as a result of changing the number of wet and dry days. Based on this finding, it is likely that irregular response surfaces will be a feature of Markov-based weather generators, including the higher-dimensional seasonal model also used in this paper (suggesting that there is no need to perform fitness landscape analysis for the more complex model). The diagnosis therefore suggests that stochastic search algorithms should be used for implementing the inverse approach, as hill-climbing methods are likely to get stuck in local optima and thus fail to find the best possible solution.

There is a wide range of stochastic search algorithms that could be used for response surfaces such as that illustrated in Fig. 5, including

**Table 1**
List of "perturbed" (P) and "held" (H) attributes for annual and seasonal weather generator experiments.

| Attribute | Description | Annual model | Seasonal model | Held Value |
|---|---|---|---|---|
| *Ptot* | total annual rainfall volume | P/H* | P/H* | 532.3 mm |
| *nWet* | annual number of wet days | P/H* | P/H* | 212.1 days |
| *P99* | 99th percentile daily rainfall amount | H | H | 16.86 mm |
| *P90* | 90th percentile daily rainfall amount | H | H | 4.636 mm |
| *DSD* | dry spell duration in days | H | H | 3.455 days |
| *DJFtot* | total rainfall volume in summer (DJF) | | H | 59.18 mm |
| *WSR* | ratio of total winter (JJA) to summer (DJF) rainfall | | H | 2.112 |

* This attribute is 'perturbed' or 'held', depending on the specific experiment.

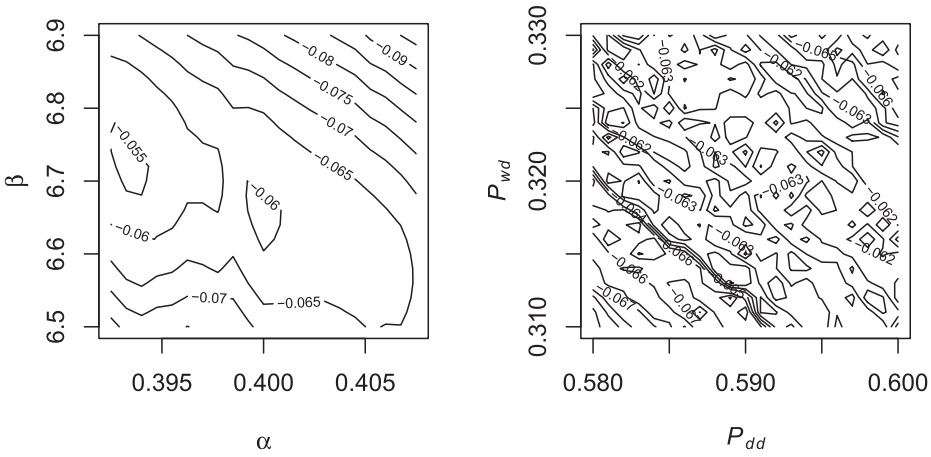


**Fig. 5.** Two-dimensional slices between the objective function value and both the Pwd, Pdd parameters (right) and the $\alpha$, $\beta$ parameters (left) from the four-parameter weather generator. The "unweighted" objective function is calculated for three attributes: total annual rainfall, number of wet days and 99th percentile rainfall.

**Table 2**
Case study genetic algorithm parameters.

| Operator | Value |
|---|---|
| Population | 200 |
| Mutation probability | 0.1 |
| Crossover probability | 0.8 |
| Crossover points | 1 |
| Generations | 200 |

shuffled complex evolution, ant colony optimization and genetic algorithms (Dorigo et al., 1996; Duan et al., 1993; Holland, 1992). In this study, a genetic algorithm is used (Scrucca, 2013), as this algorithm has been found to be effective in optimizing single objective functions with rough fitness landscapes. The parameter values for each operator of the genetic algorithm used are provided in Table 2. A population and number of generations of 200 were chosen to ensure solutions converged, and the remaining parameters were recommended by Scrucca (2013).

### 4.2. Reducing the optimization search space

In order to determine appropriate bounds for the parameters of the weather generators for the case study location in Adelaide, both weather generators were calibrated to 2870 AWAP grid locations spanning all climatic regions of Australia (at a 50 km resolution) using the approach set out in Richardson (1981). Given the significant variability in rainfall time series across continental Australia (which spans tropical, temperate, alpine, Mediterranean, semi-arid and arid climates), this approach is likely to provide a reasonable proxy of the range of variability anticipated for the case study location as a result of future climate change. The actual domain knowledge informed parameter bounds were taken as the 0.3th and 99.7th percentile of the values of the 2870 rainfall time series and are summarized in Tables 3 and 4.

To enable the benefits of reducing the size of the search space to be assessed, the inverse approach was used to generate time series with the set of attributes shown in Table 1. For this test, all attributes were set at a target of their historical levels (i.e. they were all "held" attributes). Both weather generators were used for time series generation, each with the domain knowledge informed parameters bounds and the wider, uniformed bounds used by Guo et al. (2018) (see Tables 3 and 4). For the annual weather generator, the use of domain knowledge informed bounds was able to reduce the volume of the search space by seven orders of magnitude, whereas for the seasonal generator, the volume was reduced by fifteen orders of magnitude. All optimization runs were repeated 50 times from different random starting positions in the solution space to minimize the influence of the random search behavior of the genetic algorithm. In contrast, the weather generator seed was held constant for all simulations to ensure consistency in the fitness landscape, as mentioned previously.

**Table 3**
Third standard deviation bounds on the four parameters of the annual weather generator for an Australian data set.

| Parameter | *Pdd* | *Pwd* | $\alpha$ | $\beta$ |
|---|---|---|---|---|
| Domain bounds | 0.427–0.998 | 0.088–0.824 | 0.313–0.998 | 0.043–25.46 |
| Uninformed bounds | 0–1 | 0–1 | 0–10,000 | 0–10,000 |

**Table 4**
Third standard deviation bounds on the eight parameters of the seasonal weather generator for an Australian data set.

| Parameter | *Pdd* mean | *Pdd* amplitude | *Pwd* mean | *Pwd* amplitude |
|---|---|---|---|---|
| Domain bounds | 0.38–0.99 | 0–0.36 | 0.09–0.73 | 0–0.32 |
| Uninformed bounds | 0–1 | 0–1 | 0–1 | 0–1 |
| Parameter | $\alpha$ mean | $\alpha$ amplitude | $\beta$ mean | $\beta$ amplitude |
| Domain bounds | 0.33–0.98 | 0–0.25 | 0.08–19.7 | 0.03–13.6 |
| Uninformed bounds | 0–10,000 | 0–10,000 | 0–10,000 | 0–10,000 |

Fig. 6 compares the reduction in objective function value at each optimization generation when domain knowledge informed bounds and uninformed bounds are used to restrict the parameter values for the annual weather generator (left panel) and the seasonal weather generator (right panel). For the annual weather generator, the optimization with informed parameter bounds converges much more quickly and finds better solutions (i.e. three orders of magnitudes smaller) than the optimization with the uniformed bounds for the computational budget of 200 generations. The objective function error for the informed bounds experiment at generation 200 was 0.09, compared to 130 for the uninformed bounds experiment (left panel Fig. 6). The benefits of using parameter informed bounds is more pronounced for the seasonal weather generator (right panel), such that at generation 200 the objective function errors are approximately five orders of magnitude larger when uninformed parameter bounds are used (56,212 compared to 0.202 for the informed bounds). This highlights the potential benefits of search space size reduction by using domain knowledge informed parameters in terms of increasing the computational efficiency of the inverse approach (and hence increasing the chances of finding better solutions), especially for higher-dimensional search spaces, such as those associated with more complex weather generators.

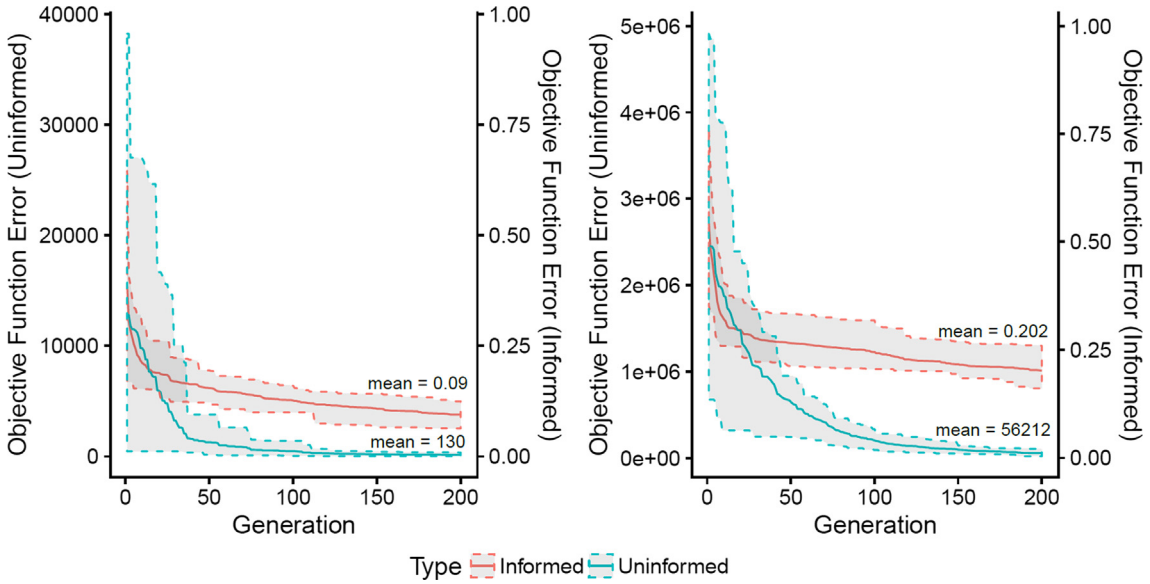


**Fig. 6.** Optimization objective function values at each generation for two sets of parameter bounds: domain knowledge informed and uninformed. A target of historical conditions is searched for with both sets of bounds using an annual weather generator (left panel) and a seasonal weather generator (right panel).

## 5. Ensuring the realism of hydrometeorological time series

To ensure realistic time series are generated by the inverse approach, the penalty structures presented in Section 2.3 are applied to the Adelaide case study. Section 5.1 tests how penalty functions work when creating targeted time series in different regions of a scenario-neutral space. The simple annual weather generator and the linear penalty function (Eq. (3)) are used for this demonstration. Section 5.2 then compares the effect of the two penalty function structures, using the more complex seasonal weather generator. How well penalty functions can be used to focus on two "perturbed" attributes at once is investigated in Section 5.3. These results are specific to the weather generators, attributes and target time series used in this case study. Consequently, the process of examining how the results change with different penalty scaling parameters is something that should be repeated for each implementation of the inverse approach, to ensure the time series are created with the most appropriate trade-offs across the "perturbed" and "held" attributes.

### 5.1. Focusing on an attribute with two target perturbations

In order to determine how penalty functions perform in creating time series in different regions of a scenario-neutral space, two target time series are generated using the simple annual weather generator and a linear penalty term (Eq. (3)). The "perturbed" attribute in both cases is *Ptot*, with the first target having no change in all selected attributes from historical levels (Table 1) and the second target having a 30% decrease in the total annual rainfall volume (*Ptot*) with no change in other attributes.

Fig. 7 shows the distance from the target of each attribute as a percentage error across 50 optimisation seeds for varying λ values. For both targets the λ = 0 cases show the error breakdown across the selected attributes using an "unweighted" objective function (i.e. equivalent to Eq. (2)). For the historical target when λ = 0, the error is low and spread relatively evenly across each attribute (top panel). The error in this context arises because of structural deficiencies in the simple annual weather generator relative to the complex historical rainfall time series, so that the weather generator is not able to faithfully simulate all the historical values of the "perturbed" and "held" attributes. For the second target (a 30% decrease in *Ptot*), the error is much more varied across attributes and the *Ptot* attribute has 20% error, whereas the other attributes have less than 10% error (bottom panel). This again results from a lack of flexibility in the annual weather generator—it lacks the degrees of freedom to change *Ptot* alone.

To reduce the error in the attribute *Ptot*, its weight in the objective function needs to be increased. This trade-off in the error between the "perturbed" and the "held" attributes changes with increasing scaling parameter values, λ (Fig. 7). Once λ = 2, the error in *Ptot* is approximately zero; however, the error in three "held" attributes (*nWet*, *P90* and *P99*) has increased. The average dry spell duration (*DSD*) is the only attribute that does not increase its error, as it only depends on the wet/dry first order Markov chain. In contrast, the error in the number of wet days increases, which is likely because this attribute more directly affects the number of high rain days sampled in a year.

The above results demonstrate that the selection of λ can be used to manage the trade-off in error between *Ptot* and the "held" attributes. The decision as to which value of λ is most appropriate should be made on a case-by-case basis by considering the importance of errors in the "held" attributes relative to errors in the "perturbed" attributes. Note that the results indicate some targets require higher penalty scaling parameter values during optimization to make time series with zero error in the "perturbed" attribute.

### 5.2. Comparing two penalty function structures

In order to examine the differences between the two penalty structures proposed in Section 2.3, both penalty structures (Eqs. (3) and (4)) are used to create the same target time series. The requested target time series corresponds to a 30% increase in *Ptot*, which is the "perturbed" attribute, with all other attributes held constant. The target time series are created using the seasonal weather generator for the desired attributes (Table 1), to see how the penalties perform with more attributes in the objective function. Fig. 8 shows the breakdown of errors across each attribute for the linear penalty term (top) and the quadratic penalty (bottom). The trialed scaling parameter values differ between the investigated penalties: the linear penalty scaling parameter was varied from 0 to 4, and the quadratic penalty scaling parameter was varied from 5 to 25.

As can be seen from Fig. 8 (top panel), the linear penalty term performs in a similar manner to when it was used with the annual generator (Fig. 7, bottom panel), despite the addition of two attributes. One difference is that the error in the *nWet* attribute is higher for the

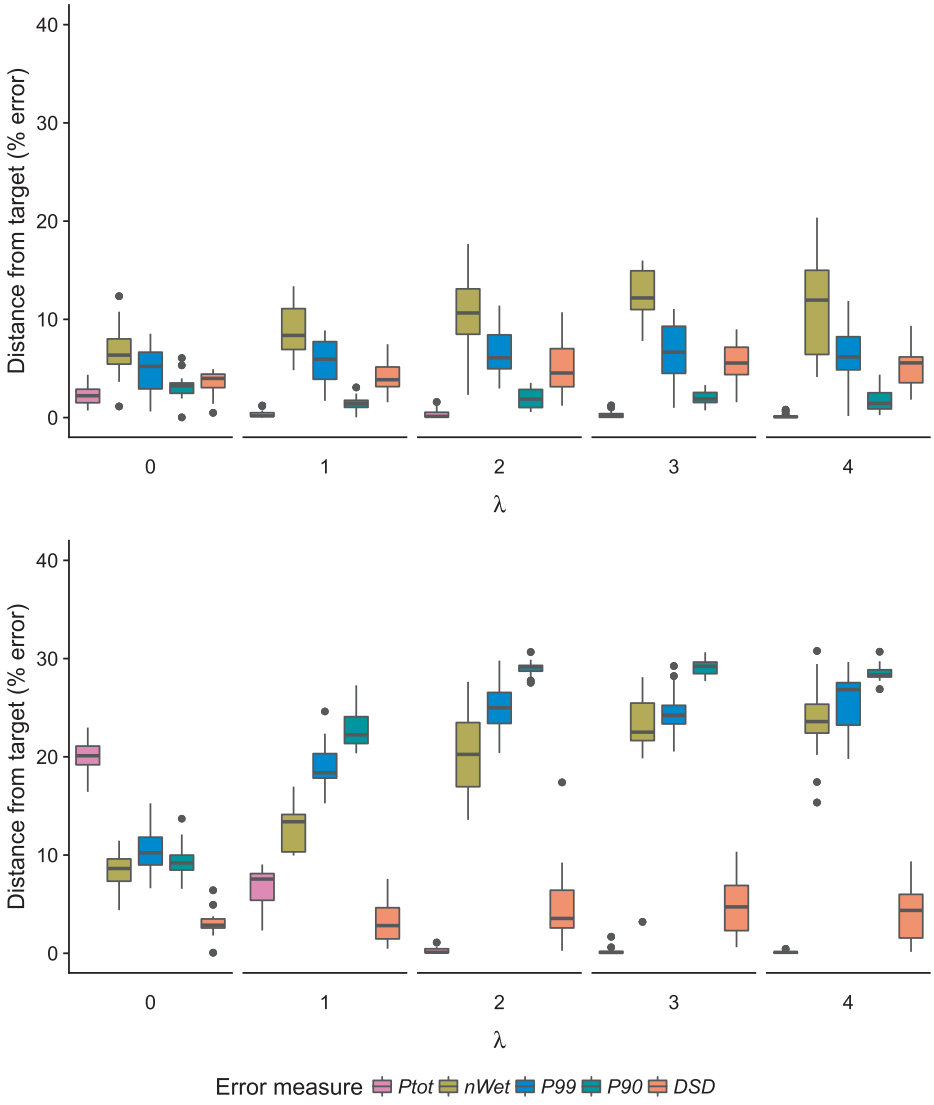


**Fig. 7.** Breakdown of the error in each attribute at the end of optimization for two target time series: zero change from the historical conditions (top) and a 30% decrease in total annual rainfall (bottom). The time series are formed using an annual weather generator and a linear penalty term where the scaling parameter λ is varied.

seasonal weather generator. It should be noted that this distribution of error is a property of three separate elements of the optimization problem: the attributes chosen, the target set and the weather generator used. With more attributes in the objective function, attributes like *nWet* that had high error for the annual generator are now weighted relatively less and thus have higher error when the seasonal generator is used. However, despite the changes in error, a value of $\lambda = 2$ is still enough to satisfy the "perturbed" attribute target.

For the quadratic penalty term, the error in the "perturbed" attribute does not reach zero, instead it approaches zero as the scaling parameter increases like in the example case shown in Fig. 4 (Section 2.3). Note that larger scaling parameter values are used to reduce the error in the "perturbed" attribute, because the square root of the scaling parameter is taken in Eq. (4) (Fig. 8, bottom panel). As a result, there is less overall error in the simulated time series when averaged across the "perturbed" and "held" attributes. This is best demonstrated by examining the *nWet* attribute in the top and bottom panels of Fig. 8. As the scaling parameter increases, the error in this attribute is around 5% less for the quadratic penalty term than it is with the linear penalty term, even though the "perturbed" attribute has near-zero error.

The decision behind which penalty structure to use in applications of the inverse approach should be made by considering the importance of error in the attributes. For example, if a target time series is set with the primary intention of reaching zero error in *Ptot*, and the other attributes are selected to make sure the stochastic time series stay similar to historical conditions, then the linear term penalty with high λ value could be used. However, if the *Ptot* target does not need to be precisely simulated, and the other attributes have a strong bearing on system performance, the time series found using the quadratic penalty term might be more appropriate for analysis. As a result, for potential future applications it is likely that a process of trial-and-error would be needed to obtain an appropriate compromise in the trade-off in errors between attributes (and thus the penalty function and associated value of λ).

### 5.3. Focusing on two perturbed attributes

In order to determine if objective function penalties can be used to guide the error for multiple attributes, time series are created with two "perturbed" attributes. In addition to *Ptot*, these time series will be created with the number of wet days in the year (*nWet*) as a penalized attribute. Fig. 9 shows the error breakdown for each attribute for the requested target of historical conditions using two "perturbed" attributes and varying the λ values for the linear penalty term (Eq. (3)). Here, both *Ptot* and *nWet* are selected as "perturbed" attributes (Figs. 7

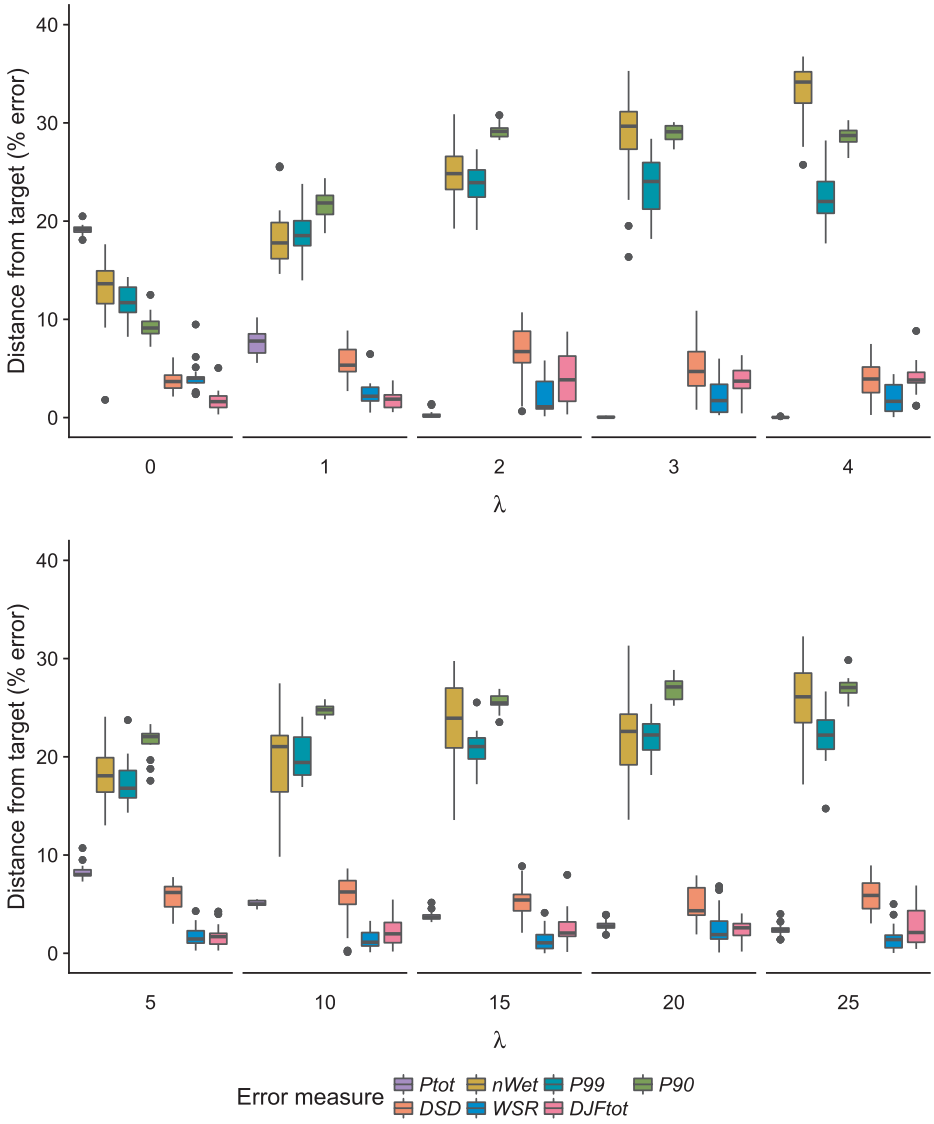


**Fig. 8.** Breakdown of the error in each attribute at the end of optimization with a varying scaling parameter λ for two different penalties: a linear penalty term (top) and a quadratic penalty term (bottom). Time series with a requested 30% decrease in *Ptot* are simulated using a seasonal weather generator.

and 8) as previous results demonstrated the difficulty in achieving low error in both *Ptot* and *nWet* simultaneously.

Fig. 9 demonstrates that the seasonal weather generator can achieve near-zero error in both "perturbed" attributes, given two appropriate penalty scaling parameters. This is first seen when $\lambda_{Ptot} = 1$ and $\lambda_{nWet} = 2$. In this case, the penalty terms are enough to make the error in the five "held" attributes higher than they would be with only *Ptot* penalized (as in previous cases, their error stayed low while *nWet* increased). Further, these results show that the *nWet* attribute should be weighted twice as much as the *Ptot* attribute for both to achieve near-zero error from their target. Again, this ratio will be a property of the weather generator and requested targets.

To summarise the impact of penalty functions, we compare the difference in the creation of a 4x4 regular grid scenario-neutral space (16 target time series) using the "unweighted" objective function (Eq. (2)) and the objective function with a linear penalty term (Eq. (3)) with $\lambda_{Ptot} = 3$ and $\lambda_{nWet} = 3$ to ensure each target is met. The requested scenario-neutral space varies *Ptot* and *nWet* from 70 to 130% of their historical values. All other targets are "held" at historical conditions. Fig. 10 compares the performance of the two optimization outcomes.

When the "unweighted" objective function is over constrained, the two "perturbed" attributes cannot meet their targets (left panel Fig. 10). This was also demonstrated in Section 5.2, where the error was spread across seven attributes when $\lambda = 0$ (Fig. 8, top panel). In Fig. 10, this is seen as a clustering in the simulated targets, for the *Ptot* dimension in particular. The simulated targets are more varied in simulating *nWet*, achieving a 30% decrease but struggling to meet a 30% increase. This is due to the weather generator structure, where the number of wet days can be decreased with minimal impact on the rainfall volume extremes or seasonality as the wet/dry sequence is changed independently.

In contrast, the use of penalties in the objective function enables the generation of time series with attributes that match two target "perturbed" attributes for the majority of the scenario-neutral space (right panel Fig. 10). However, this comes at the cost of increasing the error in the remaining "held" attributes, as can be seen from the mean "held" attribute error shown below each target in Fig. 10. Taking the example of the top right target (circled in red), when the unweighted objective function is used the "held" attributes are within 10% of historical levels on average, however, both "perturbed' attributes are ~15% away from their targets. In order for both "perturbed" attributes to reach their target, the error in the "held" attributes increases by a further 9% on average. Given the purpose of the "held" attributes is usually to ensure the realism of the time series, prioritization towards the "perturbed" attributes at the expense of the "held" attributes in most cases will be

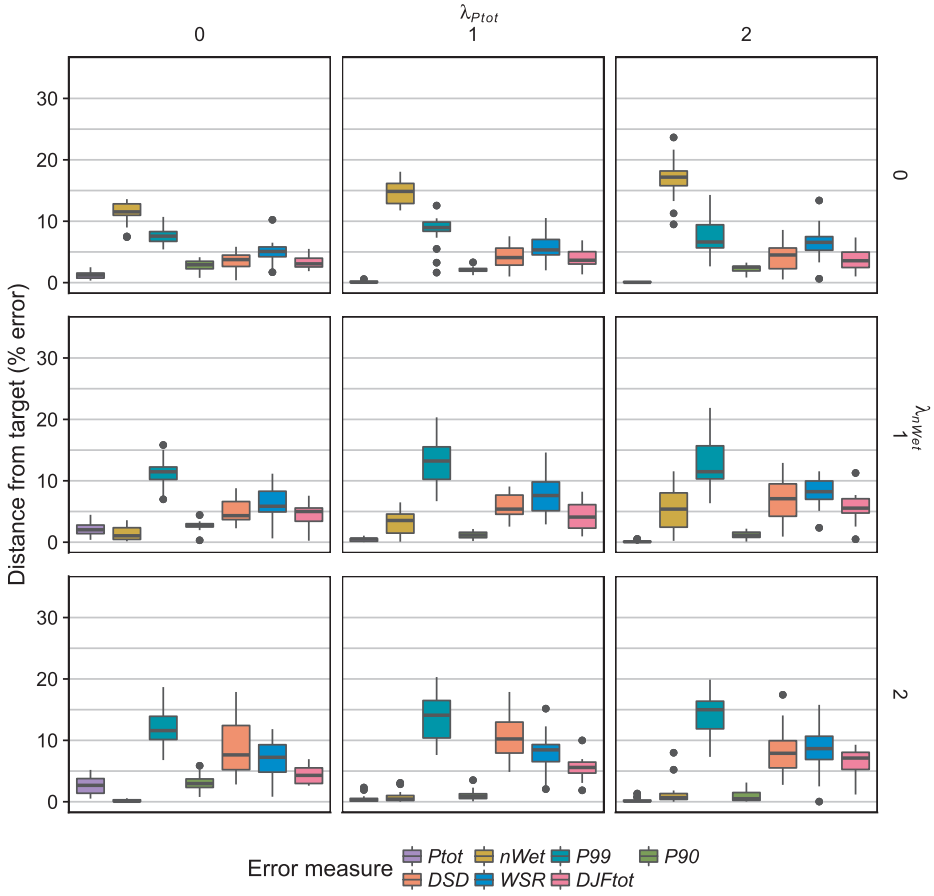


**Fig. 9.** Breakdown of the error in each attribute at the end of optimization with two "perturbed" attributes, *Ptot* and *nWet*. Time series are simulated using the seasonal weather generator for a target of historical conditions. Scaling parameters for the linear penalty term are changed separately for both attributes.

desirable.

## 6. Conclusions

The effectiveness of scenario-neutral approaches hinges on the ability to stress-test systems against plausible realizations of future climate. However, the range of changes in climate that can be examined is limited by the methods used to create the perturbed time series. Recently, the inverse approach has been presented as a method capable of producing perturbations to complex measures of hydrometeorological variables, by using formal optimization techniques with stochastic weather generators. Conceptually, this method can be applied to generate weather time series that represent not only changes in the averages, but also changes in the variability, intermittency, extremes, seasonality and/or inter-annual persistence. However, there are two key challenges to implementing the method: the large computational effort required to create the perturbed stochastic time series, and the difficulty in ensuring the realism of the time series. This paper

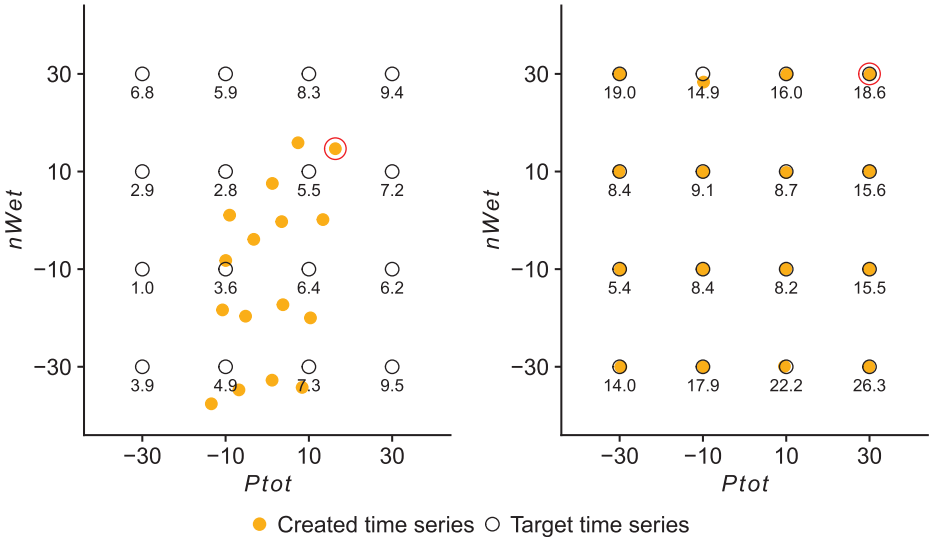


**Fig. 10.** A 4 × 4 scenario-neutral space made with the "unweighted" objective function (left) and the objective function with linear penalty terms (right). Targets are specified as percentage change from historical conditions, and the time series are made with a seasonal weather generator. The mean percentage error from historical conditions in the remaining "held" attributes is shown below each target. The red circle illustrates a point that requires the error in the "held" attributes to be doubled to reach the perturbed target.

presents approaches to overcome these challenges and improve the effectiveness of the inverse approach. Specific implementations were demonstrated using the case study of Adelaide, Australia, with a simple annual weather generator and a more complex seasonal weather generator.

As methods to increase the efficiency of an optimization process can be algorithm specific, a first step is to diagnose the nature of the optimization problem. For the weather generators used in the case study, the optimization fitness landscape was found to be irregular, due to the first-order Markov chain used to sequence wet and dry days. As a result, a genetic algorithm was selected, and two sets of bounds on the decision variables were compared: one set of uninformed bounds, and one set based on the parameter values obtained when calibrating the weather generators to sites around Australia. Results demonstrated that the domain knowledge informed bounds increased the convergence of the optimization process by a significant amount and led to a reduction in fitness values by two orders of magnitude for the seasonal weather generator. This indicates that using domain knowledge of the weather generator parameters when employing the inverse approach can increase the efficiency of the approach, particularly with more complex weather generators.

The proposed approach for ensuring the realism of generated time series is to include attributes in the objective function that keep all properties of the time series near historical levels other than those that are being actively perturbed. This has the side effect of both making the problem more complex and creating infeasible target requests (e.g. increasing the total rainfall in the year without changing either the number of wet days or the average amount per day). The recommended solution is to add penalties to the objective function that prioritize meeting the “perturbed” attribute targets, while ensuring that the remaining attributes are “held” as close to their historical values as possible. Two penalty function structures were explored on multiple target perturbations, for various values of the penalty scaling parameter. When compared to an “unweighted” objective function, results show that the use of penalties is beneficial for creating realistic hydrometeorological time series for use in scenario-neutral spaces. Currently, the optimization approach is formulated to create step changes in climate attributes, as these time series are required for scenario-neutral spaces. Further work is required to extend this optimization formulation to creating transient time series, for use in other scenario-neutral impact assessments that do not generate scenario-neutral spaces.

Both these advances to the inverse approach allow for a greater range of perturbations to be made to historical climate records. This enables the stress-testing of systems against a broader range of climate attributes, and crucially, identifying system vulnerabilities in response to this change. This ultimately will help ensure scenario-neutral analyses are able to identify system responses and potential failure modes to a much broader range of potential future climatic changes compared to traditional methods of time series perturbation.

## Acknowledgements

Sam Culley was supported by an Australian Postgraduate Award. Bree Bennett was partially supported by a grant from the South Australian Goyder Institute for Water Research.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2019.06.005.

## References

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.
Bussi, G., Dadson, S.J., Prudhomme, C., Whitehead, P.G., 2016. Modelling the future impacts of climate and land-use change on suspended sediment transport in the River Thames (UK). J. Hydrol. 542, 357–372.
Coello Coello, C.A., 2002. Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art. Comput. Methods Appl. Mech. Eng. 191 (11), 1245–1287.
Culley, S., et al., 2016. A bottom-up approach to identifying the maximum operational adaptive capacity of water resource systems to a changing climate. Water Resour. Res. 52 (9), 6751–6768.
Dorigo, M., Maniezzo, V., Colorni, A., 1996. Ant system: optimization by a colony of cooperating agents. Trans. Sys. Man Cyber. Part B 26 (1), 29–41.
Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76 (3), 501–521.
Gibbs, M.S., Maier, H.R., Dandy, G.C., 2011. Relationship between problem characteristics and the optimal number of genetic algorithm generations. Eng. Optim. 43 (4), 349–376.
Gibbs, M.S., Maier, H.R., Dandy, G.C., 2015. Using characteristics of the optimisation problem to determine the Genetic Algorithm population size when the number of evaluations is limited. Environ. Modell. Softw. 69, 226–239.
Guo, D., Westra, S., Maier, H.R., 2017. Use of a scenario-neutral approach to identify the key hydro-meteorological attributes that impact runoff from a natural catchment. J. Hydrol. 554, 317–330.
Guo, D., Westra, S., Maier, H.R., 2018. An inverse approach to perturb historical rainfall data for scenario-neutral climate impact studies. J. Hydrol. 556, 877–890.
Holland, J.H., 1992. Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence. MIT Press, pp. 228.
Kingston, G.B., Dandy, G.C., Maier, H.R., 2008. AI techniques for hydrological modeling and management. Nova 67–69.
Maier, H.R., et al., 2014. Evolutionary algorithms and other metaheuristics in water resources: current status, research challenges and future directions. Environ. Modell. Softw. 62, 271–299.
Maier, H.R., et al., 2019. Introductory overview: optimization using evolutionary algorithms and other metaheuristics. Environ. Modell. Softw. 114, 195–213.
Malan, K.M., Engelbrecht, A.P., 2013. A survey of techniques for characterising fitness landscapes and some possible ways forward. Inf. Sci. 241, 148–163.
Nesterov, Y., 2007. Gradient Methods for Minimizing Composite Objective Function. Citeseer.
Poff, N.L., et al., 2016. Sustainable water management under future uncertainty with eco-engineering decision scaling. Nat. Clim. Change 6 (1), 25–34.
Prudhomme, C., Kay, A.L., Crooks, S., Reynard, N., 2013. Climate change and river flooding: Part 2 sensitivity characterisation for British catchments and example vulnerability assessments. Clim. Change 119 (3–4), 949–964.
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.
Raupach, M., et al., 2012. Australian Water Availability Project (AWAP): CSIRO Marine and Atmospheric Research. Canberra, Australia.
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.
Scrucca, L., 2013. {GA}: A Package for Genetic Algorithms in {R}. J. Stat. Softw. 53 (4), 1–37.
Spence, C., Brown, C., 2018. Decision analytic approach to resolving divergent climate assumptions in water resources planning. J. Water Resour. Plann. Manage. 144 (9), 04018054.
Steinschneider, S., Brown, C., 2013. A semiparametric multivariate, multisite weather generator with low-frequency variability for use in climate risk assessments. Water Resour. Res. 49 (11), 7205–7220.
Wilcke, R.A., Bärring, L., 2016. Selecting regional climate scenarios for impact modelling studies. Environ. Modell. Softw. 78, 191–201.
Zielinski, K., Peters, D., Laur, R., 2005. Stopping criteria for single-objective optimization. In: Proceedings of the Third International Conference on Computational Intelligence, Robotics and Autonomous Systems.

# Appendix B

*Supplementary Material:*

*Scenario-neutral climate impact assessments: pitfalls, diagnostics and solutions (Paper 3)*

## *Supplementary Material A*

| Error from target, Panel PJJAtot=0.4, Tjun=-1 | | | | |
|---|---|---|---|---|
| **Attribute** | **Ptot** | **PJJAtot** | **Tjun** | **F0** |
| **Error Type** | **percentage** | **percentage** | **additive** | **additive** |
| 1 | 0.1 | 14.1 | 0.003 | 0.08 |
| 2 | 0.2 | 20.5 | 0.001 | 0.006 |
| 3 | 0.6 | 27 | 0.002 | 0.006 |
| 4 | 2.1 | 30.2 | 0 | 0.004 |
| 5 | 0.9 | 37.2 | 0.001 | 0.015 |
| 6 | 1.1 | 47.5 | 0.003 | 0.006 |
| 7 | 1 | 53.8 | 0.002 | 0.013 |
| 8 | 0.1 | 14.1 | 0.018 | 0.006 |
| 9 | 0.2 | 20.5 | 0.002 | 0.023 |
| 10 | 0.6 | 27 | 0.004 | 0.023 |
| 11 | 2.1 | 30.2 | 0.003 | 0.054 |
| 12 | 0.9 | 37.2 | 0.006 | 0.004 |
| 13 | 1.1 | 47.5 | 0.001 | 0.004 |
| 14 | 1 | 53.8 | 0.001 | 0.015 |
| 15 | 0.1 | 14.1 | 0.006 | 0.035 |
| 16 | 0.2 | 20.5 | 0.002 | 0.013 |
| 17 | 0.6 | 27 | 0.004 | 0.004 |
| 18 | 2.1 | 30.2 | 0.003 | 0.013 |
| 19 | 0.9 | 37.2 | 0.005 | 0.032 |
| 20 | 1.1 | 47.5 | 0.002 | 0.032 |
| 21 | 1 | 53.8 | 0.003 | 0.032 |
| 22 | 0.1 | 14.1 | 0.005 | 0.015 |
| 23 | 0.2 | 20.5 | 0.003 | 0.082 |
| 24 | 0.6 | 27 | 0.004 | 0.023 |
| 25 | 2.1 | 30.2 | 0 | 0.025 |
| 26 | 0.9 | 37.2 | 0.003 | 0.015 |
| 27 | 1.1 | 47.5 | 0.001 | 0.006 |
| 28 | 1 | 53.8 | 0.005 | 0.006 |
| 29 | 0.1 | 14.1 | 0.007 | 0.006 |
| 30 | 0.2 | 20.5 | 0 | 0.006 |
| 31 | 0.6 | 27 | 0.001 | 0.004 |
| 32 | 2.1 | 30.2 | 0.001 | 0.035 |
| 33 | 0.9 | 37.2 | 0.009 | 0.013 |
| 34 | 1.1 | 47.5 | 0.008 | 0.092 |
| 35 | 1 | 53.8 | 0.002 | 0.004 |
| 36 | 0.1 | 14.1 | 0.002 | 0.006 |
| 37 | 0.2 | 20.5 | 0.008 | 0.13 |
| 38 | 0.6 | 27 | 0.001 | 0.004 |
| 39 | 2.1 | 30.2 | 0 | 0.013 |

| | | | | |
|---|---|---|---|---|
| 40 | 0.9 | 37.2 | 0 | 0.023 |
| 41 | 1.1 | 47.5 | 0.008 | 0.025 |
| 42 | 1 | 53.8 | 0.001 | 0.006 |
| 43 | 0.1 | 14.1 | 0.001 | 0.004 |
| 44 | 0.2 | 20.5 | 0.001 | 0.006 |
| 45 | 0.6 | 27 | 0.002 | 0.013 |
| 46 | 2.1 | 30.2 | 0.007 | 0.035 |
| 47 | 0.9 | 37.2 | 0.003 | 0.004 |
| 48 | 1.1 | 47.5 | 0.003 | 0.051 |
| 49 | 1 | 53.8 | 0.002 | 0.044 |
| 50 | 0.1 | 14.1 | 0 | 0.006 |
| 51 | 0.2 | 20.5 | 0.01 | 0.044 |
| 52 | 0.6 | 27 | 0.005 | 0.015 |
| 53 | 2.1 | 30.2 | 0.002 | 0.006 |
| 54 | 0.9 | 37.2 | 0.001 | 0.006 |
| 55 | 1.1 | 47.5 | 0.003 | 0.006 |
| 56 | 1 | 53.8 | 0.002 | 0.006 |
| **Error from target, Panel PJJAtot=0.8, Tjun=-1** | | | | |
| **Attribute** | **Ptot** | **PJJAtot** | **Tjun** | **F0** |
| **Error Type** | **percentage** | **percentage** | **additive** | **additive** |
| 1 | 0.1 | 8 | 0 | 0.015 |
| 2 | 0.1 | 1.4 | 0.002 | 0.004 |
| 3 | 0 | 0 | 0.003 | 0.004 |
| 4 | 0.2 | 0.4 | 0.002 | 0.099 |
| 5 | 0.4 | 3.7 | 0.001 | 0.025 |
| 6 | 1 | 6.3 | 0.003 | 0.004 |
| 7 | 0.3 | 15.2 | 0.001 | 0.015 |
| 8 | 0.1 | 8 | 0.001 | 0.015 |
| 9 | 0.1 | 1.4 | 0.007 | 0.004 |
| 10 | 0 | 0 | 0.002 | 0.006 |
| 11 | 0.2 | 0.4 | 0.004 | 0.073 |
| 12 | 0.4 | 3.7 | 0.004 | 0.035 |
| 13 | 1 | 6.3 | 0.003 | 0.006 |
| 14 | 0.3 | 15.2 | 0.002 | 0.006 |
| 15 | 0.1 | 8 | 0.002 | 0.006 |
| 16 | 0.1 | 1.4 | 0.001 | 0.054 |
| 17 | 0 | 0 | 0.019 | 0.023 |
| 18 | 0.2 | 0.4 | 0.006 | 0.013 |
| 19 | 0.4 | 3.7 | 0.009 | 0.013 |
| 20 | 1 | 6.3 | 0.001 | 0.006 |
| 21 | 0.3 | 15.2 | 0.002 | 0.013 |
| 22 | 0.1 | 8 | 0.004 | 0.158 |
| 23 | 0.1 | 1.4 | 0.007 | 0.006 |
| 24 | 0 | 0 | 0.002 | 0.006 |

| | | | | |
|---|---|---|---|---|
| 25 | 0.2 | 0.4 | 0.001 | 0.004 |
| 26 | 0.4 | 3.7 | 0 | 0.004 |
| 27 | 1 | 6.3 | 0.002 | 0.013 |
| 28 | 0.3 | 15.2 | 0.001 | 0.004 |
| 29 | 0.1 | 8 | 0.013 | 0.08 |
| 30 | 0.1 | 1.4 | 0.014 | 0.015 |
| 31 | 0 | 0 | 0.004 | 0.025 |
| 32 | 0.2 | 0.4 | 0.001 | 0.006 |
| 33 | 0.4 | 3.7 | 0.008 | 0.063 |
| 34 | 1 | 6.3 | 0.004 | 0.035 |
| 35 | 0.3 | 15.2 | 0.008 | 0.092 |
| 36 | 0.1 | 8 | 0.001 | 0.006 |
| 37 | 0.1 | 1.4 | 0.001 | 0.015 |
| 38 | 0 | 0 | 0.006 | 0.025 |
| 39 | 0.2 | 0.4 | 0.009 | 0.006 |
| 40 | 0.4 | 3.7 | 0.006 | 0.004 |
| 41 | 1 | 6.3 | 0.001 | 0.006 |
| 42 | 0.3 | 15.2 | 0.003 | 0.025 |
| 43 | 0.1 | 8 | 0.001 | 0.006 |
| 44 | 0.1 | 1.4 | 0.001 | 0.006 |
| 45 | 0 | 0 | 0.002 | 0.015 |
| 46 | 0.2 | 0.4 | 0.001 | 0.013 |
| 47 | 0.4 | 3.7 | 0.003 | 0.013 |
| 48 | 1 | 6.3 | 0.01 | 0.006 |
| 49 | 0.3 | 15.2 | 0.003 | 0.08 |
| 50 | 0.1 | 8 | 0.002 | 0.015 |
| 51 | 0.1 | 1.4 | 0.002 | 0.004 |
| 52 | 0 | 0 | 0.002 | 0.004 |
| 53 | 0.2 | 0.4 | 0 | 0.015 |
| 54 | 0.4 | 3.7 | 0.004 | 0.013 |
| 55 | 1 | 6.3 | 0.001 | 0.006 |
| 56 | 0.3 | 15.2 | 0.001 | 0.004 |
| **Error from target, Panel PJJAtot=1.2, Tjun=-1** | | | | |
| **Attribute** | **Ptot** | **PJJAtot** | **Tjun** | **F0** |
| **Error Type** | **percentage** | **percentage** | **additive** | **additive** |
| 1 | 1.2 | 47 | 0.002 | 0.013 |
| 2 | 0.2 | 38.8 | 0.003 | 0.013 |
| 3 | 0.2 | 31.6 | 0.001 | 0.004 |
| 4 | 0.3 | 22.1 | 0.001 | 0.015 |
| 5 | 0.1 | 14.7 | 0.002 | 0.006 |
| 6 | 0 | 5.8 | 0.005 | 0.006 |
| 7 | 0 | 1.2 | 0.002 | 0.035 |
| 8 | 1.2 | 47 | 0.012 | 0.013 |
| 9 | 0.2 | 38.8 | 0.002 | 0.006 |

| | | | | |
|---|---|---|---|---|
| 10 | 0.2 | 31.6 | 0.001 | 0.025 |
| 11 | 0.3 | 22.1 | 0.002 | 0.004 |
| 12 | 0.1 | 14.7 | 0.001 | 0.015 |
| 13 | 0 | 5.8 | 0.001 | 0.006 |
| 14 | 0 | 1.2 | 0.003 | 0.025 |
| 15 | 1.2 | 47 | 0.001 | 0.006 |
| 16 | 0.2 | 38.8 | 0.004 | 0.004 |
| 17 | 0.2 | 31.6 | 0 | 0.004 |
| 18 | 0.3 | 22.1 | 0 | 0.015 |
| 19 | 0.1 | 14.7 | 0 | 0.006 |
| 20 | 0 | 5.8 | 0 | 0.194 |
| 21 | 0 | 1.2 | 0.001 | 0.006 |
| 22 | 1.2 | 47 | 0.003 | 0.006 |
| 23 | 0.2 | 38.8 | 0 | 0.025 |
| 24 | 0.2 | 31.6 | 0.003 | 0.054 |
| 25 | 0.3 | 22.1 | 0.006 | 0.025 |
| 26 | 0.1 | 14.7 | 0.015 | 0.023 |
| 27 | 0 | 5.8 | 0.011 | 0.006 |
| 28 | 0 | 1.2 | 0.002 | 0.006 |
| 29 | 1.2 | 47 | 0 | 0.006 |
| 30 | 0.2 | 38.8 | 0 | 0.082 |
| 31 | 0.2 | 31.6 | 0.003 | 0.004 |
| 32 | 0.3 | 22.1 | 0.01 | 0.004 |
| 33 | 0.1 | 14.7 | 0.001 | 0.063 |
| 34 | 0 | 5.8 | 0.009 | 0.092 |
| 35 | 0 | 1.2 | 0.002 | 0.015 |
| 36 | 1.2 | 47 | 0.002 | 0.006 |
| 37 | 0.2 | 38.8 | 0.004 | 0.054 |
| 38 | 0.2 | 31.6 | 0.001 | 0.006 |
| 39 | 0.3 | 22.1 | 0.002 | 0.044 |
| 40 | 0.1 | 14.7 | 0.001 | 0.015 |
| 41 | 0 | 5.8 | 0.005 | 0.013 |
| 42 | 0 | 1.2 | 0.002 | 0.006 |
| 43 | 1.2 | 47 | 0 | 0.013 |
| 44 | 0.2 | 38.8 | 0 | 0.035 |
| 45 | 0.2 | 31.6 | 0.009 | 0.006 |
| 46 | 0.3 | 22.1 | 0.007 | 0.006 |
| 47 | 0.1 | 14.7 | 0 | 0.073 |
| 48 | 0 | 5.8 | 0.005 | 0.015 |
| 49 | 0 | 1.2 | 0.003 | 0.013 |
| 50 | 1.2 | 47 | 0.006 | 0.063 |
| 51 | 0.2 | 38.8 | 0.007 | 0.025 |
| 52 | 0.2 | 31.6 | 0.001 | 0.015 |

| | | | | |
|---|---|---|---|---|
| 53 | 0.3 | 22.1 | 0 | 0.004 |
| 54 | 0.1 | 14.7 | 0.001 | 0.013 |
| 55 | 0 | 5.8 | 0.001 | 0.025 |
| 56 | 0 | 1.2 | 0.003 | 0.006 |

## *Supplementary Material B*

| | Error from target (historical target for non-critical attributes), Panel PJJAtot=0.4, Tjun=-1 | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Attribute** | **Ptot** | **nWet** | **P99** | **avgWSD** | **PDJFtot** | **PJJAtot** | **PMAMtot** | **PSONtot** | **TMar** | **TJun** | **TSep** | **TDec** | **Tavg** | **F0** | **Trng** |
| **Error Type** | **percent** | **percent** | **percent** | **percent** | **percent** | **percent** | **percent** | **percent** | **additive** | **additive** | **additive** | **additive** | **additive** | **additive** | **additive** |
| **1** | 0 | 22.9 | 11.5 | 42 | 89.5 | 0 | 12 | 50.9 | 3.491 | 0.001 | 3.491 | 6.512 | 1.693 | 0.006 | 0.756 |
| **2** | 0 | 21.2 | 3.7 | 35.3 | 129.4 | 0 | 27.8 | 48 | 7.133 | 0.001 | 0.02 | 6.588 | 1.709 | 0.006 | 0.729 |
| **3** | 0 | 16.1 | 1.8 | 30.5 | 172.4 | 0 | 38.6 | 42.4 | 7.102 | 0.001 | 0.09 | 6.719 | 1.761 | 0.006 | 0.451 |
| **4** | 0 | 9.3 | 8 | 27.2 | 218.9 | 0 | 48.9 | 38 | 0.145 | 0 | 7.233 | 6.348 | 1.674 | 0.006 | 0.891 |
| **5** | 0 | 7.7 | 16 | 23.4 | 260.1 | 0 | 57.7 | 29.9 | 3.343 | 0 | 3.654 | 6.401 | 1.656 | 0.006 | 0.894 |
| **6** | 0 | 7.9 | 23.3 | 23 | 302.9 | 0 | 67.9 | 23.8 | 2.977 | 0 | 4.082 | 6.442 | 1.678 | 0.006 | 0.844 |
| **7** | 0 | 4.8 | 22 | 2.2 | 359.2 | 0 | 68.9 | 16.3 | 3.285 | 0 | 3.792 | 6.426 | 1.654 | 0.006 | 0.708 |
| **8** | 0 | 22.9 | 11.5 | 42 | 89.5 | 0 | 12 | 50.9 | 15.056 | 0.039 | 6.944 | 7.765 | 2.209 | 0.013 | 1.672 |
| **9** | 0 | 21.2 | 3.7 | 35.3 | 129.4 | 0 | 27.8 | 48 | 2.964 | 0 | 5.144 | 7.574 | 2.214 | 0.006 | 1.268 |
| **10** | 0 | 16.1 | 1.8 | 30.5 | 172.4 | 0 | 38.6 | 42.4 | 0.191 | 0.002 | 8.196 | 7.693 | 2.346 | 0.006 | 0.679 |
| **11** | 0 | 9.3 | 8 | 27.2 | 218.9 | 0 | 48.9 | 38 | 0.54 | 0.001 | 7.282 | 7.211 | 2.087 | 0.006 | 1.657 |
| **12** | 0 | 7.7 | 16 | 23.4 | 260.1 | 0 | 57.7 | 29.9 | 7.756 | 0.001 | 0.178 | 7.103 | 1.956 | 0.006 | 2.461 |
| **13** | 0 | 7.9 | 23.3 | 23 | 302.9 | 0 | 67.9 | 23.8 | 7.969 | 0 | 0.102 | 7.752 | 2.249 | 0.006 | 1.376 |
| **14** | 0 | 4.8 | 22 | 2.2 | 359.2 | 0 | 68.9 | 16.3 | 4.374 | 0.001 | 3.792 | 7.572 | 2.265 | 0.006 | 1.026 |
| **15** | 0 | 22.9 | 11.5 | 42 | 89.5 | 0 | 12 | 50.9 | 4.591 | 0.002 | 4.363 | 8.512 | 2.716 | 0.015 | 1.779 |
| **16** | 0 | 21.2 | 3.7 | 35.3 | 129.4 | 0 | 27.8 | 48 | 8.257 | 0 | 1.122 | 8.987 | 2.91 | 0.006 | 1.461 |
| **17** | 0 | 16.1 | 1.8 | 30.5 | 172.4 | 0 | 38.6 | 42.4 | 8.616 | 0 | 0.687 | 8.849 | 2.81 | 0.006 | 1.539 |
| **18** | 0 | 9.3 | 8 | 27.2 | 218.9 | 0 | 48.9 | 38 | 4.627 | 0 | 4.486 | 8.598 | 2.721 | 0.006 | 1.789 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **19** | 0 | 7.7 | 16 | 23.4 | 260.1 | 0 | 57.7 | 29.9 | 11.751 | 0.002 | 2.301 | 9.229 | 2.955 | 0.006 | 1.067 |
| **20** | 0 | 7.9 | 23.3 | 23 | 302.9 | 0 | 67.9 | 23.8 | 12.074 | 0.013 | 2.988 | 8.875 | 2.756 | 0.035 | 1.864 |
| **21** | 0 | 4.8 | 22 | 2.2 | 359.2 | 0 | 68.9 | 16.3 | 11.606 | 0 | 2.59 | 8.454 | 2.669 | 0.006 | 1.936 |
| **22** | 0 | 22.9 | 11.5 | 42 | 89.5 | 0 | 12 | 50.9 | 8.606 | 0.002 | 1.561 | 9.811 | 3.3 | 0.006 | 1.887 |
| **23** | 0 | 21.2 | 3.7 | 35.3 | 129.4 | 0 | 27.8 | 48 | 12.222 | 0.001 | 2.227 | 9.556 | 3.169 | 0.006 | 2.17 |
| **24** | 0 | 16.1 | 1.8 | 30.5 | 172.4 | 0 | 38.6 | 42.4 | 19.184 | 0 | 9.428 | 9.695 | 3.158 | 0.006 | 2.386 |
| **25** | 0 | 9.3 | 8 | 27.2 | 218.9 | 0 | 48.9 | 38 | 15.445 | 0.007 | 5.837 | 9.302 | 3.016 | 0.013 | 1.961 |
| **26** | 0 | 7.7 | 16 | 23.4 | 260.1 | 0 | 57.7 | 29.9 | 12.167 | 0 | 2.079 | 9.817 | 3.283 | 0.006 | 1.743 |
| **27** | 0 | 7.9 | 23.3 | 23 | 302.9 | 0 | 67.9 | 23.8 | 12.259 | 0.002 | 2.15 | 9.632 | 3.258 | 0.006 | 1.999 |
| **28** | 0 | 4.8 | 22 | 2.2 | 359.2 | 0 | 68.9 | 16.3 | 15.935 | 0.001 | 5.966 | 9.506 | 3.117 | 0.006 | 2.487 |
| **29** | 0 | 22.9 | 11.5 | 42 | 89.5 | 0 | 12 | 50.9 | 5.876 | 0.008 | 4.916 | 10.214 | 3.584 | 0.006 | 2.75 |
| **30** | 0 | 21.2 | 3.7 | 35.3 | 129.4 | 0 | 27.8 | 48 | 8.622 | 0.027 | 1.902 | 10.083 | 3.457 | 0.004 | 2.967 |
| **31** | 0 | 16.1 | 1.8 | 30.5 | 172.4 | 0 | 38.6 | 42.4 | 15.693 | 0.001 | 5.263 | 9.977 | 3.368 | 0.006 | 2.926 |
| **32** | 0 | 9.3 | 8 | 27.2 | 218.9 | 0 | 48.9 | 38 | 12.229 | 0 | 1.39 | 10.474 | 3.634 | 0.006 | 2.352 |
| **33** | 0 | 7.7 | 16 | 23.4 | 260.1 | 0 | 57.7 | 29.9 | 12.112 | 0.002 | 1.855 | 9.716 | 3.273 | 0.006 | 3.449 |
| **34** | 0 | 7.9 | 23.3 | 23 | 302.9 | 0 | 67.9 | 23.8 | 12.733 | 0.001 | 1.837 | 10.644 | 3.707 | 0.006 | 2.461 |
| **35** | 0 | 4.8 | 22 | 2.2 | 359.2 | 0 | 68.9 | 16.3 | 9.479 | 0 | 1.413 | 10.366 | 3.611 | 0.006 | 2.534 |
| **36** | 0 | 22.9 | 11.5 | 42 | 89.5 | 0 | 12 | 50.9 | 9.168 | 0.001 | 1.884 | 10.495 | 3.686 | 0.006 | 3.37 |
| **37** | 0 | 21.2 | 3.7 | 35.3 | 129.4 | 0 | 27.8 | 48 | 2.714 | 0.008 | 9.64 | 11.876 | 4.388 | 0.004 | 1.535 |
| **38** | 0 | 16.1 | 1.8 | 30.5 | 172.4 | 0 | 38.6 | 42.4 | 6.193 | 0.003 | 5.151 | 10.73 | 3.896 | 0.006 | 2.782 |
| **39** | 0 | 9.3 | 8 | 27.2 | 218.9 | 0 | 48.9 | 38 | 2.872 | 0.007 | 9.084 | 11.306 | 4.152 | 0.006 | 2.204 |
| **40** | 0 | 7.7 | 16 | 23.4 | 260.1 | 0 | 57.7 | 29.9 | 6.342 | 0 | 5.119 | 10.966 | 3.905 | 0.006 | 2.817 |
| **41** | 0 | 7.9 | 23.3 | 23 | 302.9 | 0 | 67.9 | 23.8 | 9.836 | 0 | 1.179 | 10.606 | 3.677 | 0.006 | 3.478 |
| **42** | 0 | 4.8 | 22 | 2.2 | 359.2 | 0 | 68.9 | 16.3 | 15.384 | 0.001 | 4.816 | 10.188 | 3.445 | 0.006 | 4.046 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **43** | 0 | 22.9 | 11.5 | 42 | 89.5 | 0 | 12 | 50.9 | 9.683 | 0 | 2.063 | 11.235 | 4.04 | 0.006 | 3.37 |
| **44** | 0 | 21.2 | 3.7 | 35.3 | 129.4 | 0 | 27.8 | 48 | 9.491 | 0 | 2.096 | 11.219 | 4.04 | 0.006 | 3.377 |
| **45** | 0 | 16.1 | 1.8 | 30.5 | 172.4 | 0 | 38.6 | 42.4 | 3.399 | 0 | 8.833 | 11.496 | 4.259 | 0.006 | 2.795 |
| **46** | 0 | 9.3 | 8 | 27.2 | 218.9 | 0 | 48.9 | 38 | 12.811 | 0.001 | 1.243 | 10.868 | 3.923 | 0.006 | 3.489 |
| **47** | 0 | 7.7 | 16 | 23.4 | 260.1 | 0 | 57.7 | 29.9 | 3.454 | 0 | 8.122 | 11.048 | 3.999 | 0.006 | 3.589 |
| **48** | 0 | 7.9 | 23.3 | 23 | 302.9 | 0 | 67.9 | 23.8 | 12.417 | 0.016 | 1.097 | 10.893 | 3.795 | 0.006 | 4.102 |
| **49** | 0 | 4.8 | 22 | 2.2 | 359.2 | 0 | 68.9 | 16.3 | 3.287 | 0 | 8.095 | 10.808 | 3.927 | 0.006 | 3.596 |
| **50** | 0 | 22.9 | 11.5 | 42 | 89.5 | 0 | 12 | 50.9 | 13.665 | 0.002 | 0.908 | 12.637 | 4.655 | 0.006 | 2.938 |
| **51** | 0 | 21.2 | 3.7 | 35.3 | 129.4 | 0 | 27.8 | 48 | 3.796 | 0 | 8.389 | 11.672 | 4.299 | 0.006 | 3.772 |
| **52** | 0 | 16.1 | 1.8 | 30.5 | 172.4 | 0 | 38.6 | 42.4 | 7.225 | 0.011 | 5.843 | 12.376 | 4.692 | 0.006 | 2.648 |
| **53** | 0 | 9.3 | 8 | 27.2 | 218.9 | 0 | 48.9 | 38 | 13.647 | 0.041 | 0.685 | 12.56 | 4.663 | 0.006 | 2.97 |
| **54** | 0 | 7.7 | 16 | 23.4 | 260.1 | 0 | 57.7 | 29.9 | 6.826 | 0 | 5.705 | 12.078 | 4.508 | 0.006 | 3.079 |
| **55** | 0 | 7.9 | 23.3 | 23 | 302.9 | 0 | 67.9 | 23.8 | 3.642 | 0 | 8.768 | 12.043 | 4.516 | 0.006 | 3.129 |
| **56** | 0 | 4.8 | 22 | 2.2 | 359.2 | 0 | 68.9 | 16.3 | 3.767 | 0 | 8.924 | 12.005 | 4.552 | 0.006 | 3.003 |
| | **Error from target (historical target for non-critical attributes), Panel PJJAtot=0.8, Tjun=-1** | | | | | | | | | | | | | | |
| **Attribute** | **Ptot** | **nWet** | **P99** | **avgWSD** | **PDJFtot** | **PJJAtot** | **PMAMtot** | **PSONtot** | **TMar** | **TJun** | **TSep** | **TDec** | **Tavg** | **F0** | **Trng** |
| **Error Type** | **percent** | **percent** | **percent** | **percent** | **percent** | **percent** | **percent** | **percent** | **additive** | **additive** | **additive** | **additive** | **additive** | **additive** | **additive** |
| **1** | 0 | 19.5 | 9.7 | 46.3 | 0.1 | 0 | 27.4 | 23.2 | 3.433 | 0 | 3.269 | 6.271 | 1.543 | 0.006 | 1.536 |
| **2** | 0 | 13.3 | 4.9 | 37.5 | 29.6 | 0 | 11.7 | 15.4 | 3.415 | 0 | 3.623 | 6.409 | 1.65 | 0.006 | 1.092 |
| **3** | 0 | 9.8 | 0 | 34.3 | 66.2 | 0 | 7.2 | 13.5 | 3.256 | 0.001 | 3.981 | 6.777 | 1.826 | 0.006 | 0.133 |
| **4** | 0 | 12.4 | 7.4 | 37.1 | 99.2 | 0 | 22.6 | 7.1 | 0.423 | 0.001 | 7.448 | 6.443 | 1.705 | 0.006 | 0.668 |
| **5** | 0 | 5.8 | 12 | 29 | 141.4 | 0 | 52.2 | 16.4 | 6.689 | 0 | 0.209 | 6.489 | 1.633 | 0.006 | 1.081 |
| **6** | 0 | 2.6 | 14.5 | 22.4 | 179.3 | 0 | 70.2 | 14.4 | 7.129 | 0.01 | 0.082 | 6.647 | 1.723 | 0.004 | 0.434 |
| **7** | 0 | 2.7 | 22.9 | 16.9 | 220.8 | 0 | 88.3 | 14.1 | 6.881 | 0 | 0.234 | 6.76 | 1.733 | 0.006 | 0.608 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **8** | 0 | 19.5 | 9.7 | 46.3 | 0.1 | 0 | 27.4 | 23.2 | 7.792 | 0.001 | 0.421 | 7.732 | 2.259 | 0.006 | 1.2 |
| **9** | 0 | 13.3 | 4.9 | 37.5 | 29.6 | 0 | 11.7 | 15.4 | 4.018 | 0 | 4.223 | 7.69 | 2.301 | 0.006 | 0.972 |
| **10** | 0 | 9.8 | 0 | 34.3 | 66.2 | 0 | 7.2 | 13.5 | 7.979 | 0 | 0.546 | 8.07 | 2.428 | 0.006 | 0.423 |
| **11** | 0 | 12.4 | 7.4 | 37.1 | 99.2 | 0 | 22.6 | 7.1 | 7.868 | 0 | 0.356 | 7.685 | 2.233 | 0.006 | 1.055 |
| **12** | 0 | 5.8 | 12 | 29 | 141.4 | 0 | 52.2 | 16.4 | 0.213 | 0 | 8.147 | 7.857 | 2.37 | 0.006 | 0.634 |
| **13** | 0 | 2.6 | 14.5 | 22.4 | 179.3 | 0 | 70.2 | 14.4 | 8.097 | 0.003 | 0.352 | 8.151 | 2.456 | 0.006 | 0.523 |
| **14** | 0 | 2.7 | 22.9 | 16.9 | 220.8 | 0 | 88.3 | 14.1 | 4.361 | 0 | 4.147 | 7.947 | 2.406 | 0.006 | 0.644 |
| **15** | 0 | 19.5 | 9.7 | 46.3 | 0.1 | 0 | 27.4 | 23.2 | 1.206 | 0 | 7.872 | 8.516 | 2.71 | 0.006 | 1.879 |
| **16** | 0 | 13.3 | 4.9 | 37.5 | 29.6 | 0 | 11.7 | 15.4 | 11.974 | 0.001 | 2.739 | 8.97 | 2.836 | 0.006 | 1.547 |
| **17** | 0 | 9.8 | 0 | 34.3 | 66.2 | 0 | 7.2 | 13.5 | 8.355 | 0 | 1.146 | 9.169 | 2.956 | 0.006 | 1.115 |
| **18** | 0 | 12.4 | 7.4 | 37.1 | 99.2 | 0 | 22.6 | 7.1 | 12.189 | 0 | 2.539 | 9.318 | 3.026 | 0.006 | 0.841 |
| **19** | 0 | 5.8 | 12 | 29 | 141.4 | 0 | 52.2 | 16.4 | 8.131 | 0.002 | 1.523 | 9.266 | 2.983 | 0.006 | 0.778 |
| **20** | 0 | 2.6 | 14.5 | 22.4 | 179.3 | 0 | 70.2 | 14.4 | 8.536 | 0.002 | 0.831 | 8.975 | 2.845 | 0.006 | 1.333 |
| **21** | 0 | 2.7 | 22.9 | 16.9 | 220.8 | 0 | 88.3 | 14.1 | 8.232 | 0.002 | 1.278 | 9.073 | 2.919 | 0.006 | 1.002 |
| **22** | 0 | 19.5 | 9.7 | 46.3 | 0.1 | 0 | 27.4 | 23.2 | 12.605 | 0 | 2.384 | 9.97 | 3.339 | 0.006 | 1.739 |
| **23** | 0 | 13.3 | 4.9 | 37.5 | 29.6 | 0 | 11.7 | 15.4 | 12.186 | 0.001 | 2.183 | 9.636 | 3.196 | 0.006 | 2.17 |
| **24** | 0 | 9.8 | 0 | 34.3 | 66.2 | 0 | 7.2 | 13.5 | 5.24 | 0.002 | 5.653 | 10.523 | 3.698 | 0.006 | 0.632 |
| **25** | 0 | 12.4 | 7.4 | 37.1 | 99.2 | 0 | 22.6 | 7.1 | 8.608 | 0.001 | 1.577 | 9.81 | 3.319 | 0.006 | 1.67 |
| **26** | 0 | 5.8 | 12 | 29 | 141.4 | 0 | 52.2 | 16.4 | 8.432 | 0.005 | 1.824 | 9.763 | 3.329 | 0.006 | 1.799 |
| **27** | 0 | 2.6 | 14.5 | 22.4 | 179.3 | 0 | 70.2 | 14.4 | 5.427 | 0 | 4.975 | 10.035 | 3.456 | 0.006 | 1.438 |
| **28** | 0 | 2.7 | 22.9 | 16.9 | 220.8 | 0 | 88.3 | 14.1 | 5.345 | 0.01 | 5.455 | 10.065 | 3.537 | 0.006 | 1.025 |
| **29** | 0 | 19.5 | 9.7 | 46.3 | 0.1 | 0 | 27.4 | 23.2 | 12.347 | 0.001 | 1.937 | 10.18 | 3.481 | 0.006 | 2.714 |
| **30** | 0 | 13.3 | 4.9 | 37.5 | 29.6 | 0 | 11.7 | 15.4 | 5.795 | 0.001 | 4.575 | 9.916 | 3.365 | 0.006 | 3.038 |
| **31** | 0 | 9.8 | 0 | 34.3 | 66.2 | 0 | 7.2 | 13.5 | 15.834 | 0.001 | 5.012 | 10.538 | 3.587 | 0.006 | 2.61 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **32** | 0 | 12.4 | 7.4 | 37.1 | 99.2 | 0 | 22.6 | 7.1 | 9.391 | 0.005 | 1.797 | 10.693 | 3.778 | 0.015 | 2.105 |
| **33** | 0 | 5.8 | 12 | 29 | 141.4 | 0 | 52.2 | 16.4 | 12.133 | 0.023 | 1.054 | 10.579 | 3.733 | 0.044 | 2.208 |
| **34** | 0 | 2.6 | 14.5 | 22.4 | 179.3 | 0 | 70.2 | 14.4 | 11.927 | 0.008 | 2.121 | 9.452 | 3.136 | 0.054 | 3.419 |
| **35** | 0 | 2.7 | 22.9 | 16.9 | 220.8 | 0 | 88.3 | 14.1 | 16.024 | 0 | 5.565 | 10.176 | 3.436 | 0.006 | 3.037 |
| **36** | 0 | 19.5 | 9.7 | 46.3 | 0.1 | 0 | 27.4 | 23.2 | 12.237 | 0.001 | 1.325 | 10.51 | 3.67 | 0.006 | 3.194 |
| **37** | 0 | 13.3 | 4.9 | 37.5 | 29.6 | 0 | 11.7 | 15.4 | 6.484 | 0 | 4.631 | 10.587 | 3.717 | 0.006 | 3.109 |
| **38** | 0 | 9.8 | 0 | 34.3 | 66.2 | 0 | 7.2 | 13.5 | 8.928 | 0 | 2.741 | 11.326 | 4.066 | 0.006 | 2.447 |
| **39** | 0 | 12.4 | 7.4 | 37.1 | 99.2 | 0 | 22.6 | 7.1 | 12.767 | 0.001 | 1.685 | 10.757 | 3.754 | 0.006 | 3.271 |
| **40** | 0 | 5.8 | 12 | 29 | 141.4 | 0 | 52.2 | 16.4 | 2.897 | 0 | 8.773 | 10.951 | 3.971 | 0.006 | 2.657 |
| **41** | 0 | 2.6 | 14.5 | 22.4 | 179.3 | 0 | 70.2 | 14.4 | 6.289 | 0.002 | 5.596 | 11.261 | 4.083 | 0.015 | 2.365 |
| **42** | 0 | 2.7 | 22.9 | 16.9 | 220.8 | 0 | 88.3 | 14.1 | 6.327 | 0 | 4.717 | 10.435 | 3.681 | 0.006 | 3.232 |
| **43** | 0 | 19.5 | 9.7 | 46.3 | 0.1 | 0 | 27.4 | 23.2 | 9.026 | 0 | 2.16 | 10.987 | 3.901 | 0.006 | 3.849 |
| **44** | 0 | 13.3 | 4.9 | 37.5 | 29.6 | 0 | 11.7 | 15.4 | 9.707 | 0 | 1.672 | 11.059 | 3.927 | 0.006 | 3.694 |
| **45** | 0 | 9.8 | 0 | 34.3 | 66.2 | 0 | 7.2 | 13.5 | 9.815 | 0.006 | 2.566 | 11.893 | 4.38 | 0.006 | 2.589 |
| **46** | 0 | 12.4 | 7.4 | 37.1 | 99.2 | 0 | 22.6 | 7.1 | 9.968 | 0.001 | 1.987 | 11.411 | 4.149 | 0.006 | 3.088 |
| **47** | 0 | 5.8 | 12 | 29 | 141.4 | 0 | 52.2 | 16.4 | 6.367 | 0 | 5.224 | 11.153 | 4.034 | 0.006 | 3.58 |
| **48** | 0 | 2.6 | 14.5 | 22.4 | 179.3 | 0 | 70.2 | 14.4 | 6.763 | 0 | 5.376 | 11.571 | 4.245 | 0.006 | 2.905 |
| **49** | 0 | 2.7 | 22.9 | 16.9 | 220.8 | 0 | 88.3 | 14.1 | 9.92 | 0 | 1.935 | 11.162 | 4.013 | 0.006 | 3.546 |
| **50** | 0 | 19.5 | 9.7 | 46.3 | 0.1 | 0 | 27.4 | 23.2 | 10.255 | 0 | 2.386 | 12.371 | 4.597 | 0.006 | 3.158 |
| **51** | 0 | 13.3 | 4.9 | 37.5 | 29.6 | 0 | 11.7 | 15.4 | 7.22 | 0 | 5.67 | 12.425 | 4.693 | 0.006 | 2.665 |
| **52** | 0 | 9.8 | 0 | 34.3 | 66.2 | 0 | 7.2 | 13.5 | 6.978 | 0 | 5.73 | 12.184 | 4.556 | 0.006 | 3.095 |
| **53** | 0 | 12.4 | 7.4 | 37.1 | 99.2 | 0 | 22.6 | 7.1 | 6.847 | 0.001 | 5.376 | 11.877 | 4.397 | 0.006 | 3.474 |
| **54** | 0 | 5.8 | 12 | 29 | 141.4 | 0 | 52.2 | 16.4 | 7.493 | 0.002 | 5.711 | 12.35 | 4.77 | 0.006 | 1.778 |
| **55** | 0 | 2.6 | 14.5 | 22.4 | 179.3 | 0 | 70.2 | 14.4 | 6.994 | 0.001 | 5.854 | 12.475 | 4.645 | 0.006 | 2.716 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **56** | 0 | 2.7 | 22.9 | 16.9 | 220.8 | 0 | 88.3 | 14.1 | 7.36 | 0 | 5.728 | 11.994 | 4.541 | 0.006 | 2.851 |
| | **Error from target (historical target for non-critical attributes), Panel PJJAtot=1.2, Tjun=-1** | | | | | | | | | | | | | | |
| **Attribute** | **Ptot** | **nWet** | **P99** | **avgWSD** | **PDJFtot** | **PJJAtot** | **PMAMtot** | **PSONtot** | **TMar** | **TJun** | **TSep** | **TDec** | **Tavg** | **F0** | **Trng** |
| **Error Type** | **percent** | **percent** | **percent** | **percent** | **percent** | **percent** | **percent** | **percent** | **additive** | **additive** | **additive** | **additive** | **additive** | **additive** | **additive** |
| **1** | 0 | 19.2 | 10.5 | 31.1 | 59.5 | 0 | 47.3 | 25 | 3.088 | 0.001 | 3.982 | 6.639 | 1.709 | 0.006 | 0.411 |
| **2** | 0 | 14.4 | 4.3 | 37.4 | 38.8 | 0 | 35.9 | 9.7 | 7.107 | 0 | 0.176 | 6.928 | 1.883 | 0.006 | 0.005 |
| **3** | 0 | 12.6 | 7.8 | 40 | 7.3 | 0 | 21 | 2.2 | 3.405 | 0.001 | 3.521 | 6.42 | 1.651 | 0.006 | 0.655 |
| **4** | 0 | 10.5 | 11.7 | 36.5 | 21.4 | 0 | 0.3 | 1.3 | 3.431 | 0 | 3.626 | 6.445 | 1.655 | 0.006 | 0.822 |
| **5** | 0 | 10 | 17 | 35.8 | 51.5 | 0 | 15.6 | 9.2 | 3.342 | 0.002 | 3.933 | 6.737 | 1.819 | 0.006 | 0.041 |
| **6** | 0 | 2.2 | 19.2 | 26.2 | 85.3 | 0 | 24.3 | 20.7 | 0.421 | 0.004 | 7.538 | 6.517 | 1.736 | 0.006 | 0.519 |
| **7** | 0 | 1.6 | 23.7 | 27.3 | 118.6 | 0 | 43.2 | 24.1 | 14.447 | 0 | 7.645 | 6.763 | 1.645 | 0.006 | 1.001 |
| **8** | 0 | 19.2 | 10.5 | 31.1 | 59.5 | 0 | 47.3 | 25 | 7.718 | 0.002 | 0.672 | 8.057 | 2.426 | 0.006 | 0.474 |
| **9** | 0 | 14.4 | 4.3 | 37.4 | 38.8 | 0 | 35.9 | 9.7 | 10.911 | 0 | 2.958 | 7.712 | 2.206 | 0.006 | 1.445 |
| **10** | 0 | 10.6 | 7.4 | 37.3 | 8 | 0 | 19.7 | 2.9 | 7.765 | 0 | 0.209 | 7.742 | 2.264 | 0.006 | 1.007 |
| **11** | 0 | 10.5 | 11.7 | 36.5 | 21.4 | 0 | 0.3 | 1.3 | 7.879 | 0.001 | 0.473 | 7.857 | 2.336 | 0.006 | 0.626 |
| **12** | 0 | 10 | 17 | 35.8 | 51.5 | 0 | 15.6 | 9.2 | 3.91 | 0 | 4.296 | 7.648 | 2.27 | 0.006 | 0.985 |
| **13** | 0 | 2.2 | 19.2 | 26.2 | 85.3 | 0 | 24.3 | 20.7 | 7.436 | 0 | 0.318 | 7.621 | 2.178 | 0.006 | 1.301 |
| **14** | 0 | 1.6 | 23.7 | 27.3 | 118.6 | 0 | 43.2 | 24.1 | 0.21 | 0 | 8.349 | 7.869 | 2.396 | 0.006 | 0.67 |
| **15** | 0 | 19.2 | 10.5 | 31.1 | 59.5 | 0 | 47.3 | 25 | 11.563 | 0 | 2.622 | 8.679 | 2.678 | 0.006 | 1.806 |
| **16** | 0 | 14.4 | 4.3 | 37.4 | 38.8 | 0 | 35.9 | 9.7 | 11.665 | 0 | 2.315 | 9.239 | 2.907 | 0.006 | 0.895 |
| **17** | 0 | 10.6 | 7.4 | 37.3 | 8 | 0 | 19.7 | 2.9 | 8.424 | 0.001 | 0.682 | 8.709 | 2.731 | 0.006 | 1.735 |
| **18** | 0 | 10.5 | 11.7 | 36.5 | 21.4 | 0 | 0.3 | 1.3 | 0.911 | 0.001 | 8.435 | 8.836 | 2.889 | 0.006 | 1.139 |
| **19** | 0 | 10 | 17 | 35.8 | 51.5 | 0 | 15.6 | 9.2 | 15.518 | 0 | 6.436 | 8.986 | 2.804 | 0.006 | 1.55 |
| **20** | 0 | 2.2 | 19.2 | 26.2 | 85.3 | 0 | 24.3 | 20.7 | 8.233 | 0.002 | 0.478 | 8.345 | 2.558 | 0.006 | 2.275 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **21** | 0 | 1.6 | 23.7 | 27.3 | 118.6 | 0 | 43.2 | 24.1 | 8.231 | 0 | 0.968 | 8.856 | 2.801 | 0.006 | 1.577 |
| **22** | 0 | 19.2 | 10.5 | 31.1 | 59.5 | 0 | 47.3 | 25 | 15.116 | 0.001 | 5.897 | 9.247 | 2.838 | 0.006 | 3.334 |
| **23** | 0 | 14.4 | 4.3 | 37.4 | 38.8 | 0 | 35.9 | 9.7 | 8.697 | 0.001 | 1.469 | 9.9 | 3.31 | 0.006 | 1.722 |
| **24** | 0 | 10.6 | 7.4 | 37.3 | 8 | 0 | 19.7 | 2.9 | 8.587 | 0.001 | 1.718 | 10.081 | 3.407 | 0.006 | 1.668 |
| **25** | 0 | 10.5 | 11.7 | 36.5 | 21.4 | 0 | 0.3 | 1.3 | 8.546 | 0.001 | 1.612 | 9.781 | 3.282 | 0.006 | 1.93 |
| **26** | 0 | 10 | 17 | 35.8 | 51.5 | 0 | 15.6 | 9.2 | 16.252 | 0.004 | 5.819 | 10.256 | 3.426 | 0.006 | 1.561 |
| **27** | 0 | 2.2 | 19.2 | 26.2 | 85.3 | 0 | 24.3 | 20.7 | 12.201 | 0 | 2.128 | 9.758 | 3.241 | 0.006 | 2.064 |
| **28** | 0 | 1.6 | 23.7 | 27.3 | 118.6 | 0 | 43.2 | 24.1 | 8.526 | 0 | 1.235 | 9.223 | 3.037 | 0.006 | 2.613 |
| **29** | 0 | 19.2 | 10.5 | 31.1 | 59.5 | 0 | 47.3 | 25 | 15.601 | 0.001 | 5.116 | 10.463 | 3.419 | 0.006 | 2.876 |
| **30** | 0 | 14.4 | 4.3 | 37.4 | 38.8 | 0 | 35.9 | 9.7 | 12.062 | 0.001 | 1.894 | 10.182 | 3.393 | 0.006 | 3.028 |
| **31** | 0 | 10.6 | 7.4 | 37.3 | 8 | 0 | 19.7 | 2.9 | 12.6 | 0.005 | 1.721 | 10.64 | 3.645 | 0.006 | 1.989 |
| **32** | 0 | 10.5 | 11.7 | 36.5 | 21.4 | 0 | 0.3 | 1.3 | 8.942 | 0.003 | 1.252 | 9.943 | 3.372 | 0.025 | 3.265 |
| **33** | 0 | 10 | 17 | 35.8 | 51.5 | 0 | 15.6 | 9.2 | 8.829 | 0.001 | 1.666 | 10.146 | 3.469 | 0.006 | 2.879 |
| **34** | 0 | 2.2 | 19.2 | 26.2 | 85.3 | 0 | 24.3 | 20.7 | 8.259 | 0.008 | 1.828 | 9.732 | 3.266 | 0.006 | 2.937 |
| **35** | 0 | 1.6 | 23.7 | 27.3 | 118.6 | 0 | 43.2 | 24.1 | 11.69 | 0.007 | 1.307 | 10.185 | 3.462 | 0.006 | 2.945 |
| **36** | 0 | 19.2 | 10.5 | 31.1 | 59.5 | 0 | 47.3 | 25 | 16.242 | 0 | 4.89 | 11.155 | 3.918 | 0.006 | 2.947 |
| **37** | 0 | 14.4 | 4.3 | 37.4 | 38.8 | 0 | 35.9 | 9.7 | 2.852 | 0 | 8.499 | 10.895 | 3.926 | 0.006 | 2.939 |
| **38** | 0 | 10.6 | 7.4 | 37.3 | 8 | 0 | 19.7 | 2.9 | 12.529 | 0.02 | 2.409 | 10.057 | 3.329 | 0.004 | 4.206 |
| **39** | 0 | 10.5 | 11.7 | 36.5 | 21.4 | 0 | 0.3 | 1.3 | 11.954 | 0.034 | 1.769 | 9.733 | 3.226 | 0.023 | 4.724 |
| **40** | 0 | 10 | 17 | 35.8 | 51.5 | 0 | 15.6 | 9.2 | 15.89 | 0 | 4.946 | 10.668 | 3.674 | 0.006 | 3.482 |
| **41** | 0 | 2.2 | 19.2 | 26.2 | 85.3 | 0 | 24.3 | 20.7 | 5.907 | 0 | 6.05 | 11.458 | 4.221 | 0.006 | 1.98 |
| **42** | 0 | 1.6 | 23.7 | 27.3 | 118.6 | 0 | 43.2 | 24.1 | 12.62 | 0.014 | 1.501 | 10.633 | 3.709 | 0.025 | 3.041 |
| **43** | 0 | 19.2 | 10.5 | 31.1 | 59.5 | 0 | 47.3 | 25 | 9.999 | 0.005 | 1.963 | 11.382 | 4.168 | 0.006 | 3.177 |
| **44** | 0 | 14.4 | 4.3 | 37.4 | 38.8 | 0 | 35.9 | 9.7 | 16.287 | 0.001 | 4.557 | 11.834 | 4.176 | 0.006 | 3.065 |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **45** | 0 | 10.6 | 7.4 | 37.3 | 8 | 0 | 19.7 | 2.9 | 6.292 | 0.005 | 5.857 | 12.015 | 4.443 | 0.004 | 1.918 |
| **46** | 0 | 10.5 | 11.7 | 36.5 | 21.4 | 0 | 0.3 | 1.3 | 9.986 | 0 | 1.797 | 11.281 | 4.095 | 0.006 | 3.217 |
| **47** | 0 | 10 | 17 | 35.8 | 51.5 | 0 | 15.6 | 9.2 | 6.646 | 0 | 5.993 | 12.272 | 4.544 | 0.006 | 1.977 |
| **48** | 0 | 2.2 | 19.2 | 26.2 | 85.3 | 0 | 24.3 | 20.7 | 10.021 | 0.001 | 2.029 | 11.604 | 4.232 | 0.006 | 2.986 |
| **49** | 0 | 1.6 | 23.7 | 27.3 | 118.6 | 0 | 43.2 | 24.1 | 9.451 | 0.001 | 2.159 | 11.524 | 4.145 | 0.006 | 3.22 |
| **50** | 0 | 19.2 | 10.5 | 31.1 | 59.5 | 0 | 47.3 | 25 | 7.193 | 0 | 5.779 | 12.826 | 4.803 | 0.006 | 2.581 |
| **51** | 0 | 14.4 | 4.3 | 37.4 | 38.8 | 0 | 35.9 | 9.7 | 3.792 | 0 | 8.681 | 12.02 | 4.475 | 0.006 | 3.424 |
| **52** | 0 | 10.6 | 7.4 | 37.3 | 8 | 0 | 19.7 | 2.9 | 3.779 | 0 | 8.76 | 12.03 | 4.505 | 0.006 | 3.163 |
| **53** | 0 | 10.5 | 11.7 | 36.5 | 21.4 | 0 | 0.3 | 1.3 | 3.88 | 0 | 8.876 | 11.948 | 4.496 | 0.006 | 3.216 |
| **54** | 0 | 10 | 17 | 35.8 | 51.5 | 0 | 15.6 | 9.2 | 3.721 | 0 | 9.573 | 12.562 | 4.815 | 0.006 | 2.327 |
| **55** | 0 | 2.2 | 19.2 | 26.2 | 85.3 | 0 | 24.3 | 20.7 | 6.989 | 0 | 5.574 | 11.974 | 4.461 | 0.006 | 3.341 |
| **56** | 0 | 1.6 | 23.7 | 27.3 | 118.6 | 0 | 43.2 | 24.1 | 7.167 | 0 | 5.094 | 11.634 | 4.291 | 0.006 | 3.773 |

## *Supplementary Material C*

| | | Pitfall 1 absolute error between scenario-led time series and scenario-neutral time series | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Attribute** | **Ptot** | **nWet** | **P99** | **avgWSD** | **PDJFtot** | **PMAMtot** | **PJJAtot** | **PSONtot** | **TMar** | **TJun** | **TSep** | **TDec** | **Tavg** | **F0** | **Trng** |
| **Projection** | **Units** | **mm** | **days** | **mm** | **days** | **mm** | **mm** | **mm** | **mm** | **C** | **C** | **C** | **C** | **C** | **days** | **C** |
| **1** | **5th** | -3.205 | -12.11 | -4.72 | -0.772 | -38.723 | -2.55 | 33.655 | -28.823 | -2.034 | 4.104 | -3.926 | -3.803 | -0.001 | 34.376 | -0.56 |
| | **median** | -0.172 | 0.69 | -3.43 | -0.425 | -35.975 | 9.556 | 42.378 | -18.055 | -0.676 | 4.748 | -3.368 | -3.141 | 0 | 35.048 | -0.135 |
| | **95th** | 0.155 | 6.814 | -1.32 | -0.074 | -30.482 | 17.805 | 59.786 | -9.572 | 0.918 | 5.564 | -2.346 | -2.874 | 0.001 | 35.333 | 0.414 |
| **2** | **5th** | -1.9 | -15.99 | -2.71 | -0.832 | -53.833 | 4.776 | 26.473 | -8.828 | 3.627 | 1.067 | -2.894 | -8.036 | -0.001 | 43.164 | 1.201 |
| | **median** | -0.149 | -5.929 | -1.26 | -0.504 | -49.029 | 11.29 | 35.238 | 0.311 | 5.768 | 1.926 | -2.802 | -7.659 | 0 | 44.667 | 1.325 |
| | **95th** | 0.116 | 2.081 | 0.633 | 0.07 | -44.801 | 19.992 | 50.362 | 7.671 | 6.609 | 2.624 | -0.987 | -7.116 | 0 | 45.762 | 1.364 |
| **3** | **5th** | -1.664 | -19.57 | -7.86 | -0.766 | -125.39 | -57.907 | 184.949 | -33.977 | 4.909 | 5.183 | -5.837 | -1.913 | -0.001 | 15.831 | -4.931 |
| | **median** | -0.184 | -8.929 | -6.05 | -0.357 | -120.991 | -50.54 | 197.93 | -27.086 | 5.971 | 5.827 | -5.554 | -1.354 | 0 | 16.286 | -4.427 |
| | **95th** | 0.101 | 2.317 | -4.26 | -0.032 | -116.353 | -40.812 | 206.125 | -19.767 | 6.859 | 6.417 | -4.789 | -1.211 | 0.002 | 16.755 | -3.847 |
| **4** | **5th** | -1.343 | -16.93 | -7.81 | -0.598 | -151.448 | -62.119 | 204.068 | -22.449 | -3.553 | 2.428 | -6.948 | -4.536 | -0.001 | 29.926 | -2.785 |
| | **median** | -0.123 | -7.024 | -6.07 | -0.295 | -148.029 | -53.102 | 214.082 | -12.687 | -2.414 | 3.156 | -6.515 | -3.959 | 0 | 30.905 | -2.381 |
| | **95th** | 0.215 | 0.838 | -4.14 | 0.072 | -143.683 | -44.358 | 227.33 | -6.061 | -0.733 | 4.027 | -5.481 | -3.694 | 0 | 31.238 | -1.998 |
| **5** | **5th** | -0.073 | -5.238 | -0.57 | -0.72 | -46.863 | -77.777 | 64.042 | 37.34 | -0.354 | 4.787 | -1.376 | -7.318 | -0.001 | 33.995 | -1.881 |
| | **median** | 0.118 | 3.762 | 0.539 | -0.414 | -44.629 | -71.342 | 70.558 | 44.479 | 1.117 | 5.557 | -1.206 | -6.622 | 0 | 34.476 | -1.285 |
| | **95th** | 0.966 | 10.624 | 1.984 | -0.075 | -41.54 | -66.129 | 82.74 | 49.081 | 2.678 | 6.318 | -0.055 | -6.295 | 0.001 | 34.74 | -0.974 |
| **6** | **5th** | -1.523 | -5.481 | -3.05 | -0.674 | -56.463 | -45.784 | 91.856 | -24.994 | -1.562 | 1.728 | -2.626 | 0.603 | -0.001 | 53.255 | 1.418 |
| | **median** | -0.131 | 4.476 | -1.32 | -0.328 | -52.581 | -34.528 | 101.397 | -14.634 | 0.611 | 2.702 | -2.43 | 1.116 | 0 | 54.738 | 1.508 |
| | **95th** | 0.527 | 13.431 | 0.47 | 0.073 | -47.969 | -28.506 | 120.228 | -6.743 | 1.824 | 3.335 | -0.492 | 1.511 | 0.001 | 56.052 | 1.779 |
| **7** | **5th** | -1.387 | -7.169 | -6.31 | -0.761 | -128.006 | -26.795 | 116.259 | 5.91 | -0.501 | 4.206 | -4.138 | -3.486 | -0.002 | 28.186 | -0.426 |
| | **median** | -0.016 | 4.143 | -4.93 | -0.432 | -124.126 | -18.646 | 124.883 | 16.069 | 1.127 | 4.806 | -3.992 | -2.963 | 0 | 29.167 | 0.007 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **95th** | 0.722 | 10.926 | -3.28 | -0.095 | -119.58 | -11.627 | 137.428 | 23.851 | 2.243 | 5.664 | -2.723 | -2.5 | 0.001 | 29.598 | 0.364 |
| **8** | **5th** | -1.311 | -17.91 | -5.81 | -0.951 | -102.534 | -66.094 | 135.676 | -0.949 | 1.922 | 6.4 | -1.487 | -1.525 | -0.001 | 31.81 | 0.022 |
| | **median** | -0.084 | -6.429 | -4.17 | -0.504 | -98.538 | -56.535 | 145.667 | 9.025 | 3.864 | 7.024 | -1.438 | -1.04 | 0 | 32.667 | 0.52 |
| | **95th** | 0.205 | 2.762 | -2.56 | -0.215 | -93.014 | -49.976 | 160.437 | 17.048 | 4.807 | 7.873 | 0.003 | -0.459 | 0 | 33.048 | 0.851 |
| **9** | **5th** | -0.124 | -23.58 | 0.623 | -1.285 | -40.409 | -27.512 | 45.821 | -0.917 | -0.722 | 2.489 | -4.738 | -3.152 | -0.001 | 23.493 | -1.689 |
| | **median** | 0.103 | -12.6 | 1.997 | -0.888 | -38.53 | -21.309 | 52.557 | 6.933 | 0.339 | 3.091 | -4.175 | -2.561 | 0 | 24.238 | -1.368 |
| | **95th** | 0.719 | -5.524 | 3.2 | -0.523 | -35.842 | -17.026 | 63.247 | 11.782 | 1.775 | 3.719 | -3.372 | -2.156 | 0.001 | 24.476 | -0.859 |
| **10** | **5th** | -0.653 | -17.89 | -1.14 | -1.217 | -57.53 | -84.791 | 94.107 | 16.702 | -0.163 | 3.477 | -5.561 | -3.393 | -0.001 | 31.402 | -1.864 |
| | **median** | -0.017 | -8.119 | -0.22 | -0.829 | -54.524 | -74.761 | 103.644 | 25.448 | 1.666 | 4.242 | -5.491 | -2.835 | 0 | 32.143 | -1.374 |
| | **95th** | 0.324 | 0.067 | 1.727 | -0.343 | -49.953 | -69.543 | 117.653 | 33.252 | 2.804 | 5.134 | -3.751 | -2.3 | 0.001 | 32.598 | -1.007 |
| **11** | **5th** | -0.141 | -16.27 | -4.35 | -0.84 | -101.745 | -115.62 | 184.199 | 7.323 | 2.039 | 4.726 | -3.434 | -5.072 | -0.001 | 11.926 | -2.511 |
| | **median** | 0.091 | -6.238 | -3.62 | -0.449 | -99.682 | -108.018 | 194.003 | 14.414 | 3.351 | 5.606 | -2.845 | -4.597 | 0 | 12.714 | -1.937 |
| | **95th** | 0.923 | 2.229 | -2.28 | -0.022 | -95.567 | -102.746 | 202.418 | 20.854 | 4.691 | 6.428 | -1.833 | -4.145 | 0.001 | 13.269 | -1.534 |
| **12** | **5th** | -1.243 | -22.5 | -5.94 | -0.993 | -108.787 | -112.114 | 216.241 | -28.426 | 6.275 | 5.234 | -5.394 | -4.443 | -0.001 | 28 | -0.349 |
| | **median** | -0.096 | -10.74 | -4.13 | -0.596 | -104.441 | -100.194 | 224.295 | -20.743 | 8.364 | 6.104 | -5.191 | -4.114 | 0 | 29.357 | -0.282 |
| | **95th** | 0.101 | -0.69 | -2.59 | -0.053 | -101.004 | -93.045 | 239.008 | -12.181 | 9.306 | 6.739 | -3.536 | -3.633 | 0 | 31.126 | -0.131 |
| **13** | **5th** | -0.032 | -7.133 | -2.68 | -0.203 | -157.483 | -54.202 | 158.663 | 27.723 | 8.66 | 4.936 | -12.698 | -3.917 | -0.002 | 23.079 | -1.459 |
| | **median** | 0.093 | 4.024 | -1.34 | 0.14 | -155.187 | -47.359 | 166.761 | 35.676 | 9.915 | 6.454 | -11.268 | -3.472 | 0 | 25.429 | -0.954 |
| | **95th** | 0.729 | 10.607 | -0.39 | 0.55 | -150.965 | -43.104 | 179.248 | 41.167 | 12.624 | 7.45 | -11.108 | -3.092 | 0 | 27.598 | -0.471 |
| **14** | **5th** | -0.355 | 0.298 | -5.52 | -0.201 | -125.137 | -62.46 | 174.721 | -18.271 | 8.52 | 5.139 | -11.126 | -4.315 | -0.003 | 14.164 | -2.177 |
| | **median** | 0.032 | 9.619 | -4.02 | 0.086 | -122.004 | -51.223 | 184.203 | -9.356 | 9.631 | 6.657 | -10.709 | -3.922 | 0 | 16.714 | -1.461 |
| | **95th** | 0.889 | 15.179 | -2.38 | 0.424 | -119.39 | -46.158 | 198.573 | -3.071 | 11.521 | 7.767 | -10.347 | -3.505 | 0 | 18.052 | -1.053 |
| **15** | **5th** | -2.703 | -13.99 | -4.71 | -0.905 | -75.455 | -69.327 | 88.439 | 10.944 | 4.91 | 4.181 | -10.069 | -2.842 | -0.001 | 7.688 | -3.643 |
| | **median** | -0.164 | -3.667 | -2.73 | -0.51 | -70.441 | -58.489 | 106.495 | 21.505 | 6.292 | 5.016 | -9.268 | -2.39 | 0 | 9.238 | -3.212 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **95th** | 0.126 | 5.574 | -0.15 | -0.058 | -66.657 | -46.243 | 122.914 | 33.87 | 7.524 | 5.898 | -7.735 | -1.962 | 0.001 | 11.369 | -2.863 |
| **16** | **5th** | -2.056 | -21.22 | -5.32 | -0.725 | -66.012 | -57.489 | 68.337 | 3.288 | 7.941 | 4.565 | -10.31 | -9.591 | -0.002 | 16.143 | -1.435 |
| | **median** | -0.396 | -8.595 | -2.95 | -0.227 | -60.209 | -48.042 | 93.08 | 15.744 | 9.103 | 5.558 | -9.159 | -8.988 | 0 | 18.571 | -0.552 |
| | **95th** | 0.017 | 3.838 | 0.009 | 0.173 | -54.294 | -36.147 | 110.904 | 29.247 | 10.675 | 6.472 | -8.836 | -8.416 | 0.001 | 20.195 | -0.017 |
| **17** | **5th** | -2.217 | -9.286 | -2.41 | -0.222 | -113.445 | -101.07 | 172.688 | 3.038 | 2.8 | 6.941 | -9.4 | -3.558 | -0.002 | 8.155 | -2.625 |
| | **median** | -0.369 | 0.405 | -1.01 | 0.195 | -109.707 | -89.259 | 187.61 | 11.485 | 3.896 | 7.744 | -9.015 | -3.184 | 0 | 10.048 | -2.094 |
| | **95th** | 0.016 | 9.083 | 0.804 | 0.631 | -104.23 | -82.491 | 203.451 | 22.661 | 4.955 | 8.628 | -8.584 | -2.493 | 0.001 | 11.576 | -1.561 |
| **18** | **5th** | -1.89 | -12.97 | -3.62 | -0.18 | -96.319 | -121.249 | 168.379 | 10.757 | 7.202 | 5.177 | -11.097 | -6.032 | -0.002 | 8.81 | -2.77 |
| | **median** | -0.148 | -1.667 | -1.99 | 0.091 | -93.353 | -112.058 | 178.596 | 27.403 | 8.028 | 6.287 | -9.929 | -5.636 | 0 | 10.786 | -2.214 |
| | **95th** | 0.204 | 9.626 | -0.34 | 0.616 | -88.663 | -104.317 | 193.5 | 34.931 | 10.025 | 7.236 | -9.467 | -5.024 | 0.001 | 12.433 | -1.595 |
| **19** | **5th** | -2.286 | -3.919 | -3.68 | -0.245 | -112.196 | -29.457 | 138.441 | -32.291 | -0.867 | 4.195 | -5.27 | -6.968 | -0.001 | 23.307 | -2.061 |
| | **median** | -0.301 | 6.095 | -1.74 | 0.093 | -106.19 | -17.59 | 146.913 | -23.555 | 1.448 | 5.012 | -4.941 | -6.489 | 0 | 24.571 | -2.017 |
| | **95th** | 0.061 | 15.55 | 0.501 | 0.529 | -100.866 | -10.247 | 163.67 | -15.426 | 2.309 | 5.79 | -3.134 | -6.035 | 0 | 26.074 | -1.803 |
| **20** | **5th** | -1.71 | -6.179 | -5.27 | -0.181 | -87.672 | -34.955 | 121.723 | -43.919 | 3.304 | 5.583 | -10.044 | -7.39 | -0.002 | 21.831 | -1.96 |
| | **median** | -0.304 | 3 | -3.17 | 0.237 | -82.12 | -24.713 | 137.635 | -32.438 | 4.814 | 6.595 | -9.355 | -6.978 | 0 | 23.429 | -1.609 |
| | **95th** | 0.069 | 12.167 | -0.55 | 0.775 | -77.02 | -13.955 | 153.053 | -21.686 | 5.703 | 7.37 | -8.192 | -6.638 | 0.001 | 24.798 | -1.176 |
| | | **Pitfall 2 absolute error between scenario-led time series and scenario-neutral time series** | | | | | | | | | | | | | | |
| | **Attribute** | **Ptot** | **nWet** | **P99** | **avgWSD** | **PDJFtot** | **PMAMtot** | **PJJAtot** | **PSONtot** | **TMar** | **TJun** | **TSep** | **TDec** | **Tavg** | **F0** | **Trng** |
| **Projection** | **Units** | **mm** | **days** | **mm** | **days** | **mm** | **mm** | **mm** | **mm** | **C** | **C** | **C** | **C** | **C** | **days** | **C** |
| **1** | **5th** | -12.22 | -17.88 | -5.2 | -0.741 | -0.049 | -21.798 | -0.464 | -19.391 | 6.691 | 1.619 | -1.259 | -2.059 | 0.563 | -35.190 | 2.010 |
| | **median** | -0.11 | -3.952 | -2.8 | -0.247 | 0.012 | 1.735 | 0.047 | -3.227 | 6.691 | 1.619 | -1.259 | -2.059 | 0.563 | -35.190 | 2.010 |
| | **95th** | 0.494 | 8.914 | -0.03 | 0.241 | 0.731 | 18.065 | 1.345 | 19.874 | 6.691 | 1.619 | -1.259 | -2.059 | 0.563 | -35.190 | 2.010 |
| **2** | **5th** | -0.802 | -32.79 | -3.46 | -0.892 | -0.251 | -21.961 | -0.528 | -9.489 | 1.907 | 5.779 | -2.344 | 2.097 | 1.039 | -46.571 | 1.590 |
| | **median** | 0.031 | -12.36 | -0.9 | -0.374 | 0 | -1.011 | 0 | 1.308 | 1.907 | 5.779 | -2.344 | 2.097 | 1.039 | -46.571 | 1.590 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **95th** | 0.719 | -0.219 | 2.118 | 0.163 | 0.111 | 9.776 | 0.218 | 22.329 | 1.907 | 5.779 | -2.344 | 2.097 | 1.039 | -46.571 | 1.590 |
| **3** | **5th** | -80.29 | -30.08 | -11.3 | -0.278 | -0.307 | -86.563 | 0.218 | -63.805 | -0.177 | -0.430 | 0.671 | -3.699 | 0.204 | -16.286 | 5.359 |
| | **median** | -48.1 | -15.95 | -7.61 | 0.343 | 0.056 | -45.081 | 2.172 | -9.386 | -0.177 | -0.430 | 0.671 | -3.699 | 0.204 | -16.286 | 5.359 |
| | **95th** | -16.74 | -2.069 | -4.52 | 0.965 | 1.596 | -19.001 | 12.54 | 43.062 | -0.177 | -0.430 | 0.671 | -3.699 | 0.204 | -16.286 | 5.359 |
| **4** | **5th** | -69.11 | -27.89 | -10.8 | -0.218 | -0.136 | -88.241 | -0.033 | -35.275 | 8.494 | 3.314 | 1.776 | -1.234 | 0.623 | -31.190 | 4.393 |
| | **median** | -36.48 | -15.71 | -7.54 | 0.36 | 0.044 | -52.147 | 2.583 | 6.312 | 8.494 | 3.314 | 1.776 | -1.234 | 0.623 | -31.190 | 4.393 |
| | **95th** | -6.488 | -0.562 | -3.67 | 1.006 | 0.592 | -16.643 | 9.342 | 53.639 | 8.494 | 3.314 | 1.776 | -1.234 | 0.623 | -31.190 | 4.393 |
| **5** | **5th** | -29.29 | -18.35 | -2.83 | -0.599 | -0.071 | -117.644 | -0.152 | 38.373 | 5.062 | 0.804 | -3.795 | 1.367 | 0.538 | -34.619 | 3.091 |
| | **median** | -3.631 | -0.595 | -0.16 | -0.108 | 0.048 | -81.888 | 0.421 | 75.186 | 5.062 | 0.804 | -3.795 | 1.367 | 0.538 | -34.619 | 3.091 |
| | **95th** | 0.055 | 13.424 | 2.313 | 0.363 | 0.905 | -57.582 | 3.558 | 116.117 | 5.062 | 0.804 | -3.795 | 1.367 | 0.538 | -34.619 | 3.091 |
| **6** | **5th** | -54.46 | -8.86 | -4.15 | -0.307 | -0.05 | -80.77 | 0.035 | -23.751 | 7.629 | 5.497 | -2.735 | -6.457 | 1.328 | -59.762 | 1.500 |
| | **median** | -22.6 | 6.714 | -1.87 | 0.261 | 0.069 | -46.771 | 0.992 | 20.726 | 7.629 | 5.497 | -2.735 | -6.457 | 1.328 | -59.762 | 1.500 |
| | **95th** | -0.714 | 20.695 | 0.668 | 0.998 | 1.359 | -12 | 6.373 | 68.639 | 7.629 | 5.497 | -2.735 | -6.457 | 1.328 | -59.762 | 1.500 |
| **7** | **5th** | -0.854 | -21.77 | -8.51 | -0.689 | -0.154 | -21.337 | -0.15 | -3.701 | 5.539 | 1.835 | -1.053 | -2.433 | 0.675 | -29.571 | 2.146 |
| | **median** | -0.013 | -6.357 | -5.37 | -0.235 | 0.002 | -9.82 | 0 | 9.942 | 5.539 | 1.835 | -1.053 | -2.433 | 0.675 | -29.571 | 2.146 |
| | **95th** | 0.996 | 5.843 | -2.48 | 0.363 | 0.26 | 3.425 | 0.394 | 21.281 | 5.539 | 1.835 | -1.053 | -2.433 | 0.675 | -29.571 | 2.146 |
| **8** | **5th** | -51.14 | -25.8 | -7.6 | -0.616 | -0.132 | -92.127 | -0.001 | 0.432 | 3.219 | 0.065 | -3.710 | -4.472 | 0.801 | -33.095 | 2.031 |
| | **median** | -20.54 | -11.95 | -5.07 | -0.075 | 0.034 | -59.992 | 0.794 | 33.3 | 3.219 | 0.065 | -3.710 | -4.472 | 0.801 | -33.095 | 2.031 |
| | **95th** | -0.452 | 5.319 | -1.4 | 0.683 | 1.247 | -32.565 | 7.951 | 77.63 | 3.219 | 0.065 | -3.710 | -4.472 | 0.801 | -33.095 | 2.031 |
| **9** | **5th** | -8.87 | -32.03 | -1.44 | -1.122 | -0.108 | -59.556 | -0.088 | 8.738 | 5.951 | 3.466 | -0.454 | -2.777 | 0.616 | -24.429 | 3.278 |
| | **median** | -0.199 | -21 | 0.855 | -0.733 | 0.016 | -32.989 | 0.053 | 30.778 | 5.951 | 3.466 | -0.454 | -2.777 | 0.616 | -24.429 | 3.278 |
| | **95th** | 0.383 | -9.674 | 3.147 | -0.281 | 0.561 | -9.315 | 1.589 | 54.923 | 5.951 | 3.466 | -0.454 | -2.777 | 0.616 | -24.429 | 3.278 |
| **10** | **5th** | -56.98 | -24.2 | -2.82 | -0.78 | -0.139 | -118.241 | -0.03 | 19.733 | 5.332 | 2.763 | 0.360 | -2.623 | 0.809 | -32.619 | 3.885 |
| | **median** | -23.82 | -9.333 | -0.6 | -0.311 | 0.089 | -89.749 | 0.735 | 60.895 | 5.332 | 2.763 | 0.360 | -2.623 | 0.809 | -32.619 | 3.885 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **95th** | -0.736 | 6.25 | 2.196 | 0.411 | 1.019 | -64.793 | 8.356 | 103.014 | 5.332 | 2.763 | 0.360 | -2.623 | 0.809 | -32.619 | 3.885 |
| **11** | **5th** | -96.43 | -35.58 | -8.09 | -0.429 | -0.125 | -140.649 | -0.068 | -46.012 | 3.173 | 1.188 | -1.746 | -0.792 | 0.732 | -13.048 | 4.282 |
| | **median** | -62.11 | -19.45 | -5.3 | 0.352 | 0.212 | -114.882 | 4.723 | 46.982 | 3.173 | 1.188 | -1.746 | -0.792 | 0.732 | -13.048 | 4.282 |
| | **95th** | -15.65 | -4.467 | -2.08 | 1.205 | 2.32 | -49.95 | 18.314 | 101.61 | 3.173 | 1.188 | -1.746 | -0.792 | 0.732 | -13.048 | 4.282 |
| **12** | **5th** | -156.2 | -44.51 | -8.66 | -0.729 | -0.079 | -150.749 | 0.127 | -117.723 | -0.246 | 2.046 | 0.059 | -1.338 | 1.277 | -33.905 | 3.245 |
| | **median** | -94.58 | -23.14 | -6.18 | 0.332 | 0.231 | -109.718 | 4.245 | 5.511 | -0.246 | 2.046 | 0.059 | -1.338 | 1.277 | -33.905 | 3.245 |
| | **95th** | -28.88 | -9.033 | -3.14 | 1.272 | 1.726 | -27.208 | 15.546 | 101.711 | -0.246 | 2.046 | 0.059 | -1.338 | 1.277 | -33.905 | 3.245 |
| **13** | **5th** | -19.95 | -26.13 | -5.56 | -0.056 | -0.425 | -54.324 | -0.352 | -3.972 | 4.149 | 7.597 | 6.082 | -1.826 | 4.251 | -48.190 | 9.859 |
| | **median** | -0.067 | -9.571 | -2.11 | 0.439 | 0.034 | -32.782 | 0.204 | 30.219 | 4.149 | 7.597 | 6.082 | -1.826 | 4.251 | -48.190 | 9.859 |
| | **95th** | 2.808 | 4.538 | 0.447 | 1.224 | 1.329 | -9.08 | 7.518 | 49.666 | 4.149 | 7.597 | 6.082 | -1.826 | 4.251 | -48.190 | 9.859 |
| **14** | **5th** | -63.91 | -23.4 | -8.19 | 0.072 | -0.071 | -83.26 | 0.012 | -48.625 | 5.579 | 8.408 | 5.569 | -1.123 | 4.918 | -43.762 | 11.434 |
| | **median** | -29.24 | -4.381 | -4.68 | 0.621 | 0.107 | -44.86 | 2.592 | 12.74 | 5.579 | 8.408 | 5.569 | -1.123 | 4.918 | -43.762 | 11.434 |
| | **95th** | -6.574 | 12.155 | -2.32 | 1.405 | 1.29 | -19.694 | 13.97 | 56.217 | 5.579 | 8.408 | 5.569 | -1.123 | 4.918 | -43.762 | 11.434 |
| **15** | **5th** | -31.69 | -20.15 | -5.13 | -0.614 | -0.058 | -85.351 | -0.19 | 26.985 | 2.947 | 4.102 | 4.088 | -3.055 | 1.802 | -18.000 | 7.194 |
| | **median** | -4.956 | -5.881 | -2.53 | -0.147 | 0.042 | -59.042 | 0.15 | 51.112 | 2.947 | 4.102 | 4.088 | -3.055 | 1.802 | -18.000 | 7.194 |
| | **95th** | 0.219 | 9.162 | 0.811 | 0.24 | 0.543 | -42.986 | 2.105 | 76.759 | 2.947 | 4.102 | 4.088 | -3.055 | 1.802 | -18.000 | 7.194 |
| **16** | **5th** | -15.18 | -24.13 | -5.59 | -0.537 | -0.114 | -79.598 | -0.21 | 22.326 | 2.706 | 6.586 | 4.026 | 3.320 | 3.117 | -33.476 | 7.722 |
| | **median** | -0.538 | -9.738 | -2.25 | -0.167 | 0.013 | -52.135 | 0.046 | 51.739 | 2.706 | 6.586 | 4.026 | 3.320 | 3.117 | -33.476 | 7.722 |
| | **95th** | 0.145 | 5.017 | 1.09 | 0.362 | 0.411 | -22.86 | 1.011 | 76.691 | 2.706 | 6.586 | 4.026 | 3.320 | 3.117 | -33.476 | 7.722 |
| **17** | **5th** | -94.13 | -31.25 | -4.77 | -0.014 | -0.153 | -129.743 | 0.103 | -28.927 | 7.324 | 3.607 | 3.849 | -2.519 | 2.807 | -23.667 | 8.394 |
| | **median** | -57.97 | -9.048 | -1.92 | 0.786 | 0.069 | -93.402 | 2.008 | 28.797 | 7.324 | 3.607 | 3.849 | -2.519 | 2.807 | -23.667 | 8.394 |
| | **95th** | -3.443 | 6.574 | 1.339 | 1.268 | 1.503 | -55.042 | 13.939 | 109.591 | 7.324 | 3.607 | 3.849 | -2.519 | 2.807 | -23.667 | 8.394 |
| **18** | **5th** | -104 | -23.79 | -6.5 | 0.21 | -0.059 | -154.175 | -0.104 | 8.304 | 3.893 | 5.953 | 4.770 | -0.043 | 3.234 | -27.143 | 9.326 |
| | **median** | -70.36 | -7.786 | -3.41 | 0.885 | 0.116 | -126.865 | 1.895 | 53.351 | 3.893 | 5.953 | 4.770 | -0.043 | 3.234 | -27.143 | 9.326 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **95th** | -21.93 | 11.548 | -0.48 | 1.501 | 1.517 | -91.858 | 18.6 | 105.246 | 3.893 | 5.953 | 4.770 | -0.043 | 3.234 | -27.143 | 9.326 |
| **19** | **5th** | -43.94 | -15.56 | -5.13 | -0.098 | -0.153 | -38.684 | -0.052 | -29.39 | 6.547 | 2.839 | -0.169 | 1.108 | 1.190 | -28.190 | 4.949 |
| | **median** | -7.434 | -0.643 | -1.54 | 0.299 | 0.002 | -17.917 | 0.394 | 4.022 | 6.547 | 2.839 | -0.169 | 1.108 | 1.190 | -28.190 | 4.949 |
| | **95th** | -0.071 | 14.85 | 1.406 | 0.972 | 0.447 | 4.088 | 3.122 | 34.973 | 6.547 | 2.839 | -0.169 | 1.108 | 1.190 | -28.190 | 4.949 |
| **20** | **5th** | -63.36 | -15.83 | -6.31 | -0.022 | -0.074 | -63.342 | -0.033 | -51.286 | 4.973 | 3.194 | 4.181 | 1.478 | 2.033 | -33.429 | 6.141 |
| | **median** | -24.44 | 0.738 | -3.49 | 0.595 | 0.038 | -25.39 | 0.875 | 1.097 | 4.973 | 3.194 | 4.181 | 1.478 | 2.033 | -33.429 | 6.141 |
| | **95th** | -0.316 | 15.393 | -0.38 | 1.359 | 0.808 | -3.701 | 6.918 | 41.398 | 4.973 | 3.194 | 4.181 | 1.478 | 2.033 | -33.429 | 6.141 |
| | | **Pitfall 3 absolute error between scenario-led time series and scenario-neutral time series** | | | | | | | | | | | | | | |
| | **Attribute** | **Ptot** | **nWet** | **P99** | **avgWSD** | **PDJFtot** | **PMAMtot** | **PJJAtot** | **PSONtot** | **TMar** | **TJun** | **TSep** | **TDec** | **Tavg** | **F0** | **Trng** |
| **Projection** | **Units** | **mm** | **days** | **mm** | **days** | **mm** | **mm** | **mm** | **mm** | **C** | **C** | **C** | **C** | **C** | **days** | **C** |
| **1** | **5th** | -1.542 | -79.41 | -4.05 | -1.484 | 51.261 | -3.655 | -45.99 | -94.917 | -4.832 | -0.013 | 1.991 | 2.686 | 0.905 | -0.217 | -3.56 |
| | **median** | 0.044 | -49.83 | 1.435 | -1.145 | 78.25 | 25.75 | -26.464 | -69.257 | -4.2 | 0 | 2.615 | 3.722 | 1.178 | 0 | -2.698 |
| | **95th** | 6.417 | -7.16 | 6.345 | 0.693 | 97.005 | 42.455 | -5.706 | -56.003 | -3.36 | 0.018 | 3.718 | 4.427 | 1.451 | 0.048 | -1.971 |
| **2** | **5th** | -0.754 | -85.8 | -0.91 | -1.678 | 42.94 | 2.667 | -59.149 | -80.143 | 2.48 | -0.007 | 1.177 | -3.675 | 1.162 | -0.048 | -2.454 |
| | **median** | 0.438 | -57.69 | 3.509 | -1.176 | 68.985 | 28.218 | -40.244 | -56.058 | 3.661 | 0 | 2.183 | -3.096 | 1.46 | 0 | -1.575 |
| | **95th** | 12.83 | -15.74 | 8.811 | 0.189 | 95.911 | 47.446 | -7.496 | -32.854 | 4.445 | 0.017 | 3.491 | -2.327 | 1.693 | 0.095 | -0.591 |
| **3** | **5th** | -31.26 | -67.74 | -7.49 | -1.177 | -17.844 | -22.459 | 24.782 | -76.753 | 0.515 | -0.009 | 0.263 | 5.04 | 0.485 | -0.048 | -6.25 |
| | **median** | -3.48 | -44.91 | -3.64 | -0.739 | 2.275 | -5.274 | 46.352 | -51.666 | 1.182 | 0 | 0.699 | 5.764 | 0.722 | 0 | -5.6 |
| | **95th** | 0.465 | -6.933 | 0.709 | 0.475 | 22.254 | 15.156 | 61.275 | -26.757 | 1.819 | 0.014 | 1.561 | 6.525 | 0.969 | 0.121 | -4.494 |
| **4** | **5th** | -17.9 | -84.91 | -6.24 | -1.287 | -38.312 | -24.002 | 25.583 | -59.248 | -6.037 | -0.025 | -2.528 | 0.01 | 0.568 | -0.1 | -5.831 |
| | **median** | -4.686 | -41.48 | -3.43 | -0.696 | -17.875 | -3.745 | 51.811 | -37.567 | -4.972 | 0 | -1.696 | 0.785 | 0.963 | 0 | -4.884 |
| | **95th** | -0.004 | 3.026 | 2.969 | 0.445 | 11.62 | 20.085 | 67.628 | -18.716 | -4.188 | 0.011 | -0.534 | 1.658 | 1.227 | 0.048 | -4.191 |
| **5** | **5th** | -1.576 | -42.52 | -1.81 | -1.241 | 27.919 | -69.061 | -3.087 | -5.835 | -3.45 | -0.011 | 4.341 | 0.042 | 0.841 | -0.121 | -5.13 |
| | **median** | -0.004 | -20.64 | 1.531 | -0.856 | 41.364 | -53.026 | -0.048 | 11.855 | -2.71 | 0.001 | 5.234 | 0.93 | 1.074 | 0 | -3.953 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **95th** | 5.249 | 10.033 | 4.12 | 0.561 | 55.98 | -33.276 | 0.647 | 26.256 | -2.151 | 0.02 | 6.375 | 1.884 | 1.478 | 0.095 | -2.894 |
| **6** | **5th** | -1.432 | -43.26 | -1.74 | -1.108 | 27.643 | -21.82 | -1.211 | -62.233 | -3.146 | -0.006 | 2.032 | 5.512 | 1.14 | -0.026 | -2.975 |
| | **median** | -0.012 | -22.29 | 0.689 | -0.834 | 47.723 | -3.182 | 0.028 | -43.77 | -2.128 | 0 | 3.206 | 6.419 | 1.467 | 0 | -2.004 |
| | **95th** | 2.222 | 0.419 | 3.505 | -0.171 | 66.704 | 19.136 | 1.186 | -27.835 | -0.986 | 0.01 | 4.573 | 7.375 | 1.897 | 0.1 | -0.402 |
| **7** | **5th** | -6.53 | -39.52 | -5.39 | -1.279 | -34.264 | -3.307 | -0.664 | -29.888 | -4.164 | -0.01 | 1.174 | 2.512 | 0.61 | -0.074 | -3.323 |
| | **median** | -0.104 | -25 | -3.66 | -0.871 | -17.585 | 21.561 | 0.444 | -4.009 | -3.068 | 0 | 2.152 | 3.557 | 0.89 | 0 | -2.592 |
| | **95th** | 1.679 | 8.75 | -0.4 | 0.388 | 2.149 | 45.012 | 6.281 | 9.743 | -2.39 | 0.01 | 3.126 | 4.253 | 1.15 | 0.121 | -1.442 |
| **8** | **5th** | -8.652 | -62.42 | -5.14 | -1.504 | 2.586 | -32.853 | -0.431 | -30.925 | -1.878 | -0.008 | 4.362 | 6.09 | 0.332 | -0.074 | -3.934 |
| | **median** | -0.408 | -35.98 | -1.97 | -0.883 | 18.896 | -9.902 | 1.578 | -16.176 | -1.167 | 0 | 5.29 | 7.506 | 0.769 | 0 | -2.899 |
| | **95th** | 0.486 | -7.614 | 0.985 | -0.145 | 36.241 | 8.642 | 12.15 | 5.684 | -0.46 | 0.012 | 6.307 | 8.312 | 1.188 | 0.095 | -1.825 |
| **9** | **5th** | -0.715 | -61.67 | -0.16 | -1.862 | 21.487 | -30.561 | -5.619 | -39.243 | -3.368 | -0.012 | -0.902 | 0.927 | 0.387 | -0.048 | -4.639 |
| | **median** | 0.156 | -41.33 | 2.763 | -1.453 | 43.138 | -11.965 | -0.337 | -27.302 | -2.684 | 0 | 0.477 | 1.906 | 0.757 | 0 | -3.612 |
| | **95th** | 2.728 | -6.769 | 4.725 | -0.612 | 55.826 | 2.186 | 0.477 | -8.895 | -1.859 | 0.006 | 1.689 | 2.677 | 0.987 | 0.121 | -2.618 |
| **10** | **5th** | -6.741 | -56.82 | -2.58 | -1.679 | 25.75 | -58.985 | -1.241 | -15.058 | -2.963 | -0.031 | -0.901 | 1.617 | 0.305 | -0.074 | -5.377 |
| | **median** | -0.007 | -34.33 | 1.709 | -1.205 | 39.267 | -40.636 | -0.019 | -1.002 | -2.13 | 0 | 0.018 | 2.828 | 0.749 | 0 | -4.65 |
| | **95th** | 0.721 | 5.762 | 4.597 | -0.331 | 58.667 | -27.628 | 1.616 | 15.477 | -1.349 | 0.008 | 1.082 | 3.569 | 1.115 | 0.095 | -3.665 |
| **11** | **5th** | -16.58 | -87.35 | -5.57 | -1.477 | -13.8 | -92.535 | 39.099 | -21.697 | -2.665 | -0.011 | 1.974 | 0.334 | -0.062 | -0.048 | -4.907 |
| | **median** | -3.858 | -38.6 | -2.14 | -0.69 | 6.34 | -75.059 | 68.467 | -6.446 | -1.795 | 0.001 | 2.736 | 1.061 | 0.129 | 0 | -4.1 |
| | **95th** | -0.004 | 3.883 | 2.638 | 0.522 | 26.507 | -52.938 | 80.484 | 18.979 | -0.981 | 0.012 | 3.512 | 1.885 | 0.333 | 0.095 | -3.256 |
| **12** | **5th** | -16.18 | -89.62 | -5.31 | -1.487 | -5.608 | -80.108 | 47.707 | -63.505 | 2.236 | -0.011 | -0.038 | 1.515 | 0.049 | -0.048 | -4.811 |
| | **median** | -3.183 | -45.62 | -1.15 | -0.935 | 17.008 | -52.395 | 75.377 | -41.803 | 3.103 | 0 | 1.162 | 2.442 | 0.396 | 0 | -4.061 |
| | **95th** | 0.057 | -12.07 | 3.863 | 0.55 | 46.702 | -32.168 | 95.372 | -19.353 | 3.806 | 0.011 | 1.98 | 3.251 | 0.769 | 0.121 | -3.046 |
| **13** | **5th** | -27.13 | -69.11 | -3.59 | -0.853 | -66.058 | -36.491 | 23.639 | -12.811 | 1.656 | -0.005 | -7.396 | -0.674 | -1.578 | 0 | -9.583 |
| | **median** | -2.392 | -26 | -0.66 | -0.171 | -49.37 | -12.783 | 42.033 | 17.566 | 2.64 | 0.001 | -6.659 | 0.124 | -1.242 | 0 | -8.878 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **95th** | 0.575 | 17.388 | 4.419 | 0.907 | -24.698 | 6.323 | 58.943 | 33.395 | 3.809 | 0.01 | -5.852 | 0.683 | -0.985 | 0.048 | -8.031 |
| **14** | **5th** | -16.74 | -52.59 | -5.91 | -0.899 | -31.232 | -36.156 | 25.842 | -46.28 | 0.821 | -0.003 | -7.125 | -1.608 | -2.108 | 0 | -10.492 |
| | **median** | -2.748 | -21.38 | -2.61 | -0.227 | -10.137 | -15.364 | 51.249 | -30.656 | 1.822 | 0 | -6.345 | -1.14 | -1.798 | 0 | -9.836 |
| | **95th** | -0.032 | 20.726 | 2.3 | 1.27 | 15.58 | 5.702 | 64.611 | -6.52 | 3.119 | 0.002 | -5.67 | -0.412 | -1.46 | 0 | -8.914 |
| **15** | **5th** | -4.289 | -67.34 | -0.41 | -1.548 | 22.929 | -40.774 | -4.731 | -39.702 | -0.378 | -0.013 | -5.221 | 0.668 | -0.927 | -0.026 | -7.85 |
| | **median** | -0.002 | -47.69 | 1.552 | -1.169 | 45.255 | -23.679 | -0.018 | -18.657 | 0.383 | 0 | -4.45 | 1.218 | -0.583 | 0 | -6.804 |
| | **95th** | 2.043 | -22.82 | 4.345 | -0.708 | 70.345 | -7.757 | 0.479 | -2.322 | 1.21 | 0.011 | -3.708 | 1.716 | -0.363 | 0.048 | -6.238 |
| **16** | **5th** | -1.675 | -91.3 | 0.401 | -1.495 | 63.754 | -41.962 | -52.308 | -72.411 | 2.297 | -0.01 | -5.659 | -6.101 | -1.17 | -0.048 | -7.572 |
| | **median** | 0.164 | -57.24 | 3.519 | -1.091 | 87.136 | -6.331 | -31.12 | -46.266 | 2.936 | 0 | -4.935 | -5.525 | -0.923 | 0 | -7.029 |
| | **95th** | 8.96 | -28.52 | 8.601 | -0.239 | 115.544 | 15.674 | -5.059 | -25.604 | 3.878 | 0.003 | -4.058 | -4.862 | -0.644 | 0.048 | -6.258 |
| **17** | **5th** | -23.11 | -66.81 | -1.47 | -0.869 | -9.623 | -67.952 | 5.46 | -29.551 | -4.797 | -0.011 | -5.547 | 0.456 | -1.771 | -0.048 | -9.627 |
| | **median** | -3.242 | -36.79 | 1.928 | -0.385 | 17.463 | -48.242 | 29.696 | -7.132 | -4.04 | 0 | -3.945 | 1.46 | -1.466 | 0 | -8.749 |
| | **95th** | 0.014 | 0.138 | 6.124 | 0.984 | 37.906 | -16.748 | 45.253 | 12.156 | -3.041 | 0.005 | -2.631 | 2.315 | -1.144 | 0.048 | -7.837 |
| **18** | **5th** | -18.19 | -71.91 | -2.44 | -0.952 | 6.006 | -88.978 | 0.699 | -14.406 | 0.055 | -0.006 | -6.547 | -2.951 | -1.631 | -0.026 | -9.116 |
| | **median** | -3.574 | -34.91 | 0.749 | -0.336 | 30.728 | -62.41 | 24.934 | 3.059 | 1.068 | 0 | -5.496 | -2.213 | -1.342 | 0 | -8.518 |
| | **95th** | -0.003 | -5.433 | 6.027 | 0.485 | 57.99 | -40.321 | 42.569 | 25.626 | 2.181 | 0.008 | -4.625 | -1.627 | -1.11 | 0.095 | -7.726 |
| **19** | **5th** | -4.259 | -48.83 | -1 | -0.996 | -0.149 | 14.185 | -0.582 | -72.392 | -4.452 | -0.007 | -0.366 | -2.171 | 0.022 | -0.048 | -6.13 |
| | **median** | -0.015 | -29.02 | 1.642 | -0.552 | 15.183 | 33.95 | 0.289 | -51.077 | -3.704 | 0 | 0.757 | -1.074 | 0.303 | 0 | -5.443 |
| | **95th** | 0.475 | -4.114 | 4.085 | 0.316 | 39.189 | 51.676 | 2.414 | -32.164 | -2.816 | 0.019 | 1.854 | 0.026 | 0.722 | 0.048 | -4.245 |
| **20** | **5th** | -4.186 | -51.73 | -4.37 | -0.713 | 19.064 | 2.627 | -0.668 | -82.432 | -2.429 | -0.009 | -5.022 | -2.617 | -0.73 | -0.074 | -7.535 |
| | **median** | -0.026 | -30.95 | -0.76 | -0.312 | 37.382 | 24.183 | 0.03 | -58.55 | -1.557 | 0 | -3.797 | -1.539 | -0.381 | 0 | -6.71 |
| | **95th** | 1.536 | 4.948 | 2.611 | 0.537 | 53.736 | 45.314 | 0.679 | -41.763 | -0.699 | 0.007 | -2.574 | -0.716 | -0.071 | 0.048 | -6.061 |
| | | **Pitfall 4 absolute error between scenario-led time series and scenario-neutral time series** | | | | | | | | | | | | | | |
| | **Attribute** | **Ptot** | **nWet** | **P99** | **avgWSD** | **PDJFtot** | **PMAMtot** | **PJJAtot** | **PSONtot** | **TMar** | **TJun** | **TSep** | **TDec** | **Tavg** | **F0** | **Trng** |

| Projection | Units | mm | days | mm | days | mm | mm | mm | mm | C | C | C | C | C | days | C |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 5th | -0.059 | -28.01 | -5.57 | -0.953 | 4.467 | -37.185 | -0.028 | -51.372 | -4.996 | -0.002 | -13.649 | 4.167 | 1.531 | 0 | -1.825 |
| | median | 0.001 | -14.88 | -1.62 | -0.64 | 18.936 | 5.238 | 0.002 | -25.417 | 12.695 | 0 | -12.448 | 5.817 | 2.039 | 0 | -0.32 |
| | 95th | 0.062 | -2.717 | 1.238 | -0.227 | 41.011 | 40.557 | 0.025 | 20.355 | 14.847 | 0.001 | 5.644 | 7.118 | 2.476 | 0 | 0.696 |
| 2 | 5th | -0.102 | -35.74 | -3.22 | -1.09 | -17.204 | -40.255 | -0.041 | -64.782 | 1.849 | -0.001 | -10.231 | -3.002 | 1.525 | 0 | -1.619 |
| | median | -0.018 | -24 | 0.615 | -0.723 | 2.575 | -1.985 | -0.004 | -1.898 | 3.471 | 0 | 3.985 | -1.776 | 2.005 | 0 | -0.495 |
| | 95th | 0.053 | -1.002 | 4.301 | -0.025 | 14.04 | 74.059 | 0.026 | 35.684 | 17.015 | 0.022 | 5.867 | -0.329 | 2.448 | 0 | 0.843 |
| 3 | 5th | -0.067 | -43.83 | -8.42 | -0.756 | 63.167 | 11.128 | -0.024 | -173.02 | -0.048 | -0.001 | -16.578 | 6.271 | 1.136 | 0 | -4.129 |
| | median | 0.001 | -27.98 | -5.15 | -0.381 | 86.396 | 48.575 | 0 | -135.913 | 0.671 | 0 | 2.884 | 7.465 | 1.449 | 0 | -3.139 |
| | 95th | 0.08 | -10.88 | -2.08 | 0.097 | 101.872 | 96.304 | 0.025 | -103.313 | 20.084 | 0.002 | 3.978 | 8.368 | 1.779 | 0 | -1.893 |
| 4 | 5th | -0.069 | -37.82 | -8.04 | -0.742 | 52.059 | 13.93 | -0.025 | -158.597 | -6.832 | -0.005 | -16.281 | 1.712 | 1.418 | 0 | -3.872 |
| | median | -0.014 | -24.86 | -5.44 | -0.3 | 78.245 | 56.118 | 0.001 | -129.45 | 9.285 | 0 | -14.326 | 2.819 | 1.83 | 0 | -2.707 |
| | 95th | 0.109 | -12.12 | -1.68 | 0.051 | 93.581 | 90.241 | 0.028 | -93.455 | 11.869 | 0.005 | 2.58 | 3.897 | 2.274 | 0 | -1.016 |
| 5 | 5th | -0.091 | -31.79 | -2.04 | -1.017 | 20.858 | -90.499 | -0.019 | -19.783 | -3.771 | -0.004 | -11.067 | 1.513 | 1.437 | 0 | -3.215 |
| | median | -0.002 | -14.86 | 0.808 | -0.646 | 35.156 | -57.157 | 0.001 | 25.156 | 14.164 | 0 | -10.061 | 2.821 | 1.833 | 0 | -2.008 |
| | 95th | 0.056 | -1.598 | 3.591 | -0.222 | 52.495 | -17.802 | 0.021 | 59.792 | 15.746 | 0.002 | 8.338 | 3.75 | 2.348 | 0 | -0.786 |
| 6 | 5th | -0.104 | -22.9 | -4.34 | -0.769 | 33.426 | -49.602 | -0.031 | -97.679 | -4.26 | -0.003 | -8.912 | 6.455 | 1.8 | 0 | -1.612 |
| | median | 0.005 | -10.48 | -0.87 | -0.493 | 53.949 | -5.33 | 0.001 | -52.277 | -2.964 | 0 | 5.654 | 7.854 | 2.227 | 0 | -0.529 |
| | 95th | 0.091 | 5.767 | 2.44 | 0.04 | 74.15 | 36.769 | 0.035 | 1.787 | 11.781 | 0.002 | 7.556 | 9.641 | 2.666 | 0 | 0.616 |
| 7 | 5th | -0.088 | -25.62 | -8 | -0.915 | -16.643 | -16.938 | -0.023 | -79.99 | -4.706 | -0.001 | -13.773 | 4.146 | 1.37 | 0 | -1.345 |
| | median | -0.012 | -13.19 | -4.71 | -0.475 | 2.837 | 33.017 | -0.002 | -38.647 | -3.52 | 0 | 4.475 | 5.28 | 1.711 | 0 | -0.312 |
| | 95th | 0.045 | 3.729 | -0.89 | 0.058 | 22.934 | 79.111 | 0.022 | 21.767 | 14.323 | 0.005 | 6.084 | 6.073 | 2.04 | 0 | 0.948 |
| 8 | 5th | -0.082 | -39.32 | -7.53 | -0.994 | 32.249 | -33.505 | -0.034 | -98.922 | -1.946 | -0.004 | -11.042 | 7.773 | 0.987 | 0 | -2.671 |
| | median | -0.003 | -24.48 | -3.79 | -0.63 | 49.734 | 9.287 | 0.001 | -63.306 | 15.822 | 0 | -10.391 | 9.023 | 1.483 | 0 | -1.061 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **95th** | 0.079 | -1.664 | 0.044 | 0.06 | 67.864 | 45.798 | 0.027 | -2.989 | 17.574 | 0.004 | 8.365 | 10.256 | 1.959 | 0 | 0.569 |
| **9** | **5th** | -0.068 | -47.16 | -0.96 | -1.466 | 11.372 | -64.106 | -0.018 | -54.624 | -4.799 | -0.001 | -13.831 | 2.715 | 1.319 | 0 | -1.931 |
| | **median** | -0.016 | -31.98 | 1.433 | -1.153 | 25.115 | -16.214 | 0.002 | -12.807 | -3.976 | 0 | 4.064 | 4.006 | 1.741 | 0 | -0.629 |
| | **95th** | 0.061 | -17.98 | 4.983 | -0.828 | 40.211 | 22.667 | 0.017 | 33.901 | 13.202 | 0.002 | 5.192 | 4.865 | 2.062 | 0 | 0.516 |
| **10** | **5th** | -0.05 | -41.01 | -2.43 | -1.417 | 36.394 | -82.965 | -0.023 | -56.748 | -3.736 | -0.003 | -14.829 | 3.629 | 1.3 | 0 | -3.446 |
| | **median** | 0.005 | -28.33 | 0.489 | -0.997 | 54.557 | -37.591 | -0.006 | -18.744 | 4.78 | 0 | -6.101 | 4.854 | 1.692 | 0 | -2.215 |
| | **95th** | 0.087 | -12.3 | 3.653 | -0.533 | 71.625 | 3.713 | 0.02 | 39.789 | 15.052 | 0.006 | 4.441 | 6.296 | 2.198 | 0 | -0.817 |
| **11** | **5th** | -0.098 | -48.54 | -5.24 | -0.969 | 91.327 | -50.951 | -0.027 | -124.065 | -3.374 | -0.002 | 3.871 | 1.581 | 0.574 | 0 | -2.695 |
| | **median** | 0 | -33.71 | -3.27 | -0.518 | 112.434 | -9.987 | 0 | -95.024 | -2.749 | 0 | 5.075 | 2.546 | 0.85 | 0 | -1.445 |
| | **95th** | 0.06 | -14.49 | -0.16 | 0.006 | 128.801 | 18.508 | 0.027 | -74.407 | -2.002 | 0.002 | 6.556 | 3.492 | 1.177 | 0 | -0.176 |
| **12** | **5th** | -0.067 | -48.07 | -6.46 | -1.043 | 114.223 | -25.565 | -0.022 | -174.56 | 1.858 | -0.004 | -14.51 | 2.988 | 0.735 | 0 | -3.464 |
| | **median** | -0.012 | -30.86 | -3.46 | -0.635 | 136.818 | 11.36 | -0.002 | -149.443 | 10.504 | 0 | -6.148 | 4.114 | 1.221 | 0 | -1.865 |
| | **95th** | 0.04 | -16.08 | 0.726 | 0.185 | 157.053 | 46.639 | 0.022 | -120.978 | 20.79 | 0.003 | 4.569 | 5.473 | 1.51 | 0 | -0.855 |
| **13** | **5th** | -0.078 | -35.12 | -4.77 | -0.364 | 7.03 | 1.755 | -0.022 | -96.891 | 0.781 | -0.002 | -16.714 | -0.536 | -1.372 | 0 | -9.434 |
| | **median** | 0.004 | -17.93 | -1.69 | 0.001 | 25.327 | 45.811 | 0 | -69.49 | 3.844 | 0 | -6.137 | 0.77 | -0.969 | 0 | -8.38 |
| | **95th** | 0.063 | -4.674 | 3.137 | 0.48 | 46.18 | 76.86 | 0.023 | -36.623 | 13.532 | 0.002 | -3.085 | 2.233 | -0.527 | 0 | -7.325 |
| **14** | **5th** | -0.071 | -27.95 | -6.66 | -0.399 | 46.501 | -0.134 | -0.023 | -153.464 | 0.063 | -0.001 | -15.557 | -1.674 | -2.069 | 0 | -10.717 |
| | **median** | 0.001 | -13.1 | -3.57 | -0.02 | 73.626 | 49.3 | 0.001 | -121.077 | 2.82 | 0 | -7.283 | -0.977 | -1.682 | 0 | -9.529 |
| | **95th** | 0.084 | -1.829 | -0.33 | 0.316 | 92.346 | 89.186 | 0.024 | -83.741 | 11.64 | 0.002 | -4.182 | 0.075 | -1.38 | 0 | -8.651 |
| **15** | **5th** | -0.069 | -26.96 | -4.87 | -0.851 | 12.196 | -61.446 | -0.023 | -56.46 | -1.734 | -0.002 | -18.386 | 1.506 | -0.361 | 0 | -6.543 |
| | **median** | 0.001 | -16.14 | -1.27 | -0.562 | 28.308 | -20.368 | 0.001 | -9.538 | -0.793 | 0 | -2.086 | 2.744 | 0.113 | 0 | -4.575 |
| | **95th** | 0.08 | -0.8 | 1.988 | -0.166 | 54.313 | 25.342 | 0.032 | 35.049 | 15.673 | 0.001 | -0.286 | 3.57 | 0.352 | 0 | -3.734 |
| **16** | **5th** | -0.096 | -24.26 | -5.54 | -0.61 | 15.059 | -76.563 | -0.034 | -44.348 | 0.899 | -0.003 | -16.71 | -5.504 | -0.753 | 0 | -6.859 |
| | **median** | 0.009 | -11.67 | -2.24 | -0.269 | 27.762 | -40.893 | -0.007 | 2.768 | 6.773 | 0 | -8.728 | -4.406 | -0.441 | 0 | -5.663 |

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **95th** | 0.096 | 3.407 | 1.825 | 0.185 | 52.817 | 21.031 | 0.029 | 50.009 | 16.864 | 0.005 | -1.11 | -3.137 | 0.003 | 0 | -4.488 |
| **17** | **5th** | -0.063 | -30.94 | -3.33 | -0.29 | 52.05 | -28.801 | -0.03 | -143.16 | -6.135 | -0.002 | -17.965 | 2.291 | -0.991 | 0 | -7.521 |
| | **median** | 0.003 | -17.48 | -0.5 | 0.067 | 85.065 | 16.72 | -0.002 | -97.196 | -5.189 | 0 | -1.036 | 3.229 | -0.532 | 0 | -5.935 |
| | **95th** | 0.074 | -3.24 | 3.231 | 0.612 | 97.938 | 62.293 | 0.028 | -61.924 | 12.299 | 0.001 | 0.612 | 4.66 | -0.195 | 0 | -4.676 |
| **18** | **5th** | -0.08 | -32.55 | -4.37 | -0.367 | 71.583 | -58.902 | -0.022 | -102.599 | -1.187 | -0.001 | -17.472 | -2.451 | -1.182 | 0 | -8.245 |
| | **median** | -0.001 | -20.1 | -0.94 | 0.143 | 89.273 | -22.044 | -0.004 | -72.774 | 0.187 | 0 | -3.383 | -1.191 | -0.824 | 0 | -7.153 |
| | **95th** | 0.101 | -4.302 | 2.233 | 0.595 | 109.439 | 14.606 | 0.03 | -25.645 | 14.169 | 0.003 | -1.474 | -0.058 | -0.433 | 0 | -5.978 |
| **19** | **5th** | -0.112 | -19.9 | -4.26 | -0.324 | 20.986 | -5.961 | -0.026 | -132.802 | -5.39 | -0.002 | -14.663 | -0.296 | 0.767 | 0 | -4.299 |
| | **median** | 0.001 | -4.119 | -1.14 | 0.106 | 37.549 | 54.611 | -0.001 | -96.023 | -3.935 | 0 | 2.625 | 0.572 | 1.129 | 0 | -3.016 |
| | **95th** | 0.064 | 9.843 | 2.716 | 0.614 | 58.682 | 95.158 | 0.031 | -32.828 | 13.556 | 0.002 | 5.161 | 2.253 | 1.629 | 0 | -1.496 |
| **20** | **5th** | -0.099 | -18.79 | -6.41 | -0.175 | 33.645 | -11.639 | -0.041 | -132.39 | -3.113 | -0.003 | -18.558 | -0.63 | 0.077 | 0 | -5.748 |
| | **median** | -0.004 | -8.786 | -1.86 | 0.174 | 50.688 | 33.866 | -0.002 | -85.228 | 12.934 | 0 | -16.869 | 0.307 | 0.46 | 0 | -4.572 |
| | **95th** | 0.106 | 5.362 | 1.734 | 0.692 | 70.26 | 79.603 | 0.029 | -33.289 | 15.62 | 0.004 | -0.043 | 1.598 | 0.911 | 0 | -3.167 |