Spatial and temporal modelling for perennial crop variety selection trials

Joanne De Faveri

Doctor of Philosophy

School of Agriculture, Food and Wine
University of Adelaide
URRBRAE, SA 5064

September 2013

Contents

Abstract

This thesis involves the investigation and development of methods for analysing data from variety selection trials in perennial crops. This involves identifying best varieties from data collected at multiple times in field trials, often from multiple locations and involving multiple traits. For accurate variety predictions the methods for analysis of such data need to account for the spatial correlation typically present in field trials and the temporal correlation induced by the repeated measures nature of the data. The methods also need to model the variety effects over time. The methods presented are based on the linear mixed model and estimation is performed using residual maximum likelihood (REML).

Spatial analysis methods are applied to data from multiple harvest times for two perennial crop data sets. These analyses show that spatial correlation is evident and the spatial analysis methods improve model fit. Simulation studies also show the spatial analysis methods provide better predictions of variety effects (closer to the true effects).

As the data from perennial crop variety selection trials is measured over time there is also a need to account for the temporal correlation between measurements. Separable models are presented that model the spatial and temporal residual covariance structure. These methods are suitable for large numbers of harvests. Application to a multi-harvest lucerne breeding data set shows these models to be an improvement on historical analysis approaches.

At the genetic level the variety effects need to be modelled over time. Two approaches are presented. The first approach involves applying factor analytic models to variety by harvest effects and using clustering to aid in interpretation and selection. The second approach uses cubic smoothing spline random regression. These approaches are applied to data from two traits from a lucerne breeding trial and are shown to successfully model the variety by harvest effects and aid in selection. As data is usually obtained from multiple trials at different locations, the above approaches are extended to the multi-environment situation and applied to a multi-harvest, multi-environment lucerne data set.

While the separable spatio-temporal residual models show an improvement on analysing each harvest time separately, they are very restrictive in that they assume common spatial correlation parameters across harvests (or traits). The initial spatial analyses on the two multi-harvest perennial crop data sets reveal that spatial correlation often varies between harvests and between traits. A more suitable non-separable covariance model is investigated that allows for differing spatial correlation across time or traits. The approach is based on the Multivariate Autoregressive model, initially for spatial correlation in one direction. Subsequently the model is extended to the two directional row-column situation using the theory of Multivariate Conditional Autoregressive models. These models are applied to the lucerne multi-harvest and multi-trait data using code written in R, and are shown in most cases to be a significant improvement to the separable residual models previously investigated.

Declaration

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

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

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

SIGNED:

DATE:

Acknowledgements

I would like to sincerely thank my official supervisors Wayne Pitchford and Brian Cullis for their invaluable guidance and support throughout this PhD. I would also like to thank Ari Verbyla who started me on this journey and made sure I reached the end. All three have contributed greatly to the direction of this research and I have learnt so much from all of them. I am very appreciative of all of their efforts.

I would like to thank Robin Thompson for his ideas which instigated this PhD and for being so generous with his wisdom. I would also like to thank Shoba Venkatanagappa and the NSW DPI lucerne breeding team for providing me with the lucerne data set to work on. I also thank Guangdi Li for providing the chicory data. The data sets provided challenging motivation for the methods in this PhD and were a vital part of this research. I am also very grateful to David Butler and Arthur Gilmour for their support in implementing models in ASReml and ASReml-R.

I gratefully acknowledge the financial support from the NSW Department of Primary Industries, CSIRO and The University of Adelaide for providing a generous scholarship. I also acknowledge my employer Queensland Department of Agriculture, Fisheries and Forestry for providing me with extended leave without pay while completing the PhD.

I thank my colleagues at CMIS and DAFF for their support and patience. I especially wish to thank Mary, Joanna, Maree and Janet for their continued encouragement.

Finally I would like to thank my family, especially Ari and my two wonderful children Matthew and Lara, for their endless support, love and understanding throughout these years. Without you I would not have got this far.