

## The lasso as a tool for model selection

Daniel I. Brown<sup>1</sup>, Byron J.T. Morgan<sup>1</sup> and Ian T. Jolliffe<sup>2</sup>

<sup>1</sup> Institute of Mathematics, Statistics and Actuarial Science, University of Kent, UK

<sup>2</sup> School of Engineering, Computer Science and Mathematics, University of Exeter, UK

Mark-recapture-recovery models are becoming increasingly complex, which makes the task of model selection all the more difficult. There are a number of approaches available, in both the classical and Bayesian contexts, for dealing with the complications associated with age structure and time variation. Finding the best model without these approaches can be difficult due to the potentially large model space associated with a large number of covariates. A method used for estimation in linear models, when there are a large number of parameters, is the lasso (least absolute shrinkage and selection operator) [1]. The lasso results in some coefficients, within the model, being set to zero, thus giving more interpretable models. This method can be applied in the ring-recovery context when there is time variation in the survival parameter, and a number of different covariates are put forward to explain the variation. The use of the lasso is demonstrated on white stork (*Ciconia ciconia*) data and compared to a previous study. Blackbird (*Turdus Merula*) data with a large number of weather covariates are also analysed.

## References

- [1] Tibshirani R (1996) Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society, Series B 58:267-288