

Echidna Mixed Models Software

A. R. Gilmour¹

¹ 11 Holman Way, Orange, NSW, 2800, Australia

Arthur.Gilmour@CargoVale.com.au (Corresponding Author)

Summary

Echidna is new software with the basic functionality of ASReml, but developed for free public use. It fits linear mixed models using the Average Information algorithm to estimate the variance parameters. It is targeted for use in plant and animal breeding contexts.

Keywords: ASReml, Echidna, mixed model software, average information

Introduction

ASReml (Gilmour et al., 2015) was written by Dr Arthur Gilmour in 1996 while employed by NSW Agriculture (DPI) in collaboration with Robin Thompson of Rothamsted Research (RR) and Brian Cullis also of DPI. DPI and RR decided to commercialize it with the release of ASReml 2 in 2002. DPI sold its interests to their commercial partner, VSN International, in 2011, after Dr Gilmour was retrenched in 2009. RR sold its interests to VSN, in 2014, after Robin Thompson retired.

ASReml quickly became widely used after its initial free release in 1996 and the core has remained stable over the past 20 years. The core was incorporated into Genstat as its REML procedure and is the basis for the asreml-r R software. The ASReml user interface has been extended over the years to allow a broader range of models to be fitted. However, commercialization has changed the development dynamic and restricted use of the product to those with deep pockets. Consequently, Dr Gilmour has undertaken to develop a new product, Echidna, similar to ASReml in functionality, as a platform for trialling new models and free for all to use.

Functionality of Echidna

Echidna is a new product written from scratch; the ASReml code is owned by VSN International. It is intended, in the first instance to fit the models commonly fitted in ASReml with similar syntax but with alternate internal procedures and revised output formats. It does not have the full functionality of ASReml but should have similar efficiency.

Echidna has been written to help scientists search out the incredible design of this created world in which we live. In the creation there is an vast amount of valuable genetic variation built in from the beginning which through careful breeding we can utilize to develop strains and varieties adapted to the great variety of production environments around the world.

The linear mixed model has proven to be a very powerful tool for selection of more productive genotypes. This is an international scientific endeavour and this software is offered as a tool for all to use to this end.

Echidna implements the functional syntax style first developed for `asreml-r` whereby the variance structure to be fitted is specified as a wrapper function in the specification of the model. For example, in a multivariate animal model, the base model is specified as, say,

```
Wwt Ywt ~ Trait us(Trait).nrm(Animal)
Residual units.us(Trait)
```

where `Wwt Ywt` are two response variables, `~` is read as 'is modelled as', `Trait` is a fixed effect fitting an intercept for each trait, `us(Trait).nrm(Animal)` specifies the random effects for the interaction of `Trait` and `Animal` and the variance structure for them which is a direct product of a 2x2 unstructured matrix (`us(.)`) and the numerator relationship matrix (`nrm(.)`), `Residual` is the keyword for specifying the residual, `units` specifies the rows (animals with data), assumed independent, and `us(Trait)` here specifies the 2x2 unstructured residual variance matrix.

Echidna uses Residual Maximum Likelihood and the Average Information algorithm (Gilmour et al., 1995) to estimate the variance structures and reports *inter-alia* the BLUPs and residuals from the model.

ASReml allows for an extensive range of variance structures used in plant and animal breeding and the more commonly used ones will be available in Echidna. Echidna is not intended to supplant ASReml, but rather to become a research platform where new procedures can be trialled and tested before being integrated into ASReml.

Stage of Development

Echidna now fits the most common models fitted in ASReml but some output options of ASReml are not yet available. Echidna currently fits traditional mixed models, the common spatial models used for multi-environment plant breeding trials, multivariate (pedigree based) animal models and GBLUP. The plan for development in 2018 involves the following stages

- Broaden the range of variance structures including Factor Analytic models
- Tabulation, Prediction, data and residual statistics
- Kenward Roger degrees of Freedom calculation
- Review of computational efficiency
- Genomic models
- (Double Hierarchical) Generalized Linear (Mixed) Models (PQL)
- Graphics

Basic documentation and examples are being prepared alongside development of the software.

New functionality should be implemented in 2019.

Availability

The latest builds of Echidna are available at www.EchidaMMS.org. The initial release will be for 64bit Windows. Unix and Mac versions should soon be available through partners.

Partners

The standalone program Echidna is being initially developed by Dr Gilmour. It is expected others will develop an R implementation and undertake research into improved computational methods to handle genomic models better. Any such improvements in Echidna relative to the performance of ASReml will be available for VSN to incorporate into ASReml, as they choose.

Support

Dr Gilmour welcomes support queries as they lead to software improvement, whether by making Echidna more robust, or by extending its functionality. As this is new software, users should initially expect to discover issues to be resolved. Large corporations with a major breeding program should continue to use ASReml for production purposes but are welcome to trial Echidna and report issues relating to their applications.

Acknowledgements

Dr Gilmour acknowledges the role of many mentors and collaborators in developing his skill base that has resulted in various programs including ASReml and Echidna. They include Professor Les Balaam (University of Sydney), Professor Robert Anderson (Massey University), Cameron Kirton, Brian Coote, Paul Nichols (Biometry Branch, NSW Agriculture), Dr Brian McGuirk, Dr Robert Banks, Dr Robin Thompson, Dr Karin Meyer, Dr Brian Cullis, Dr David Butler, Dr Roger Payne and many biometricians and fellow research scientists who have used the software. NSW Agriculture employed Dr Gilmour as a biometrician / agricultural research scientist for almost 42 years, funding his training and travel and providing an ideal work environment where he could develop software for use by colleagues. Since retrenchment, Dr Gilmour has received support from University of Wollongong, VSN International and several breeding companies. Jesus Christ provided the health, strength, opportunity, aptitude and interest to develop his skills.

List of References

- Gilmour, A.R., Anderson, R.D. and Rae, A.L. (1985). The analysis of binomial data by a generalized linear mixed model. *Biometrika*, 72: 593-599.
<http://dx.doi.org/10.1093/biomet/72.3.593>
- Gilmour, A.R., Luff, A.F., Fogarty, N.M. and Banks, R.G. (1994). Genetic parameters for ultrasonic fat depth and eye muscle measurements in live Poll Dorset sheep. *Australian Journal of Agricultural Research*, 45, 1281-91.

- Gilmour, A.R., Thompson, R. and Cullis, B.R. (1995). Average Information REML, an efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* 51, 1440-50.
- Gilmour, A.R., Cullis, B.R. and Verbyla, A. (1997). Accounting for Natural and Extraneous Variation in the Analysis of Field Experiments. *Journal of Agricultural, Biological and Environmental Statistics* 2, 269-293.
- Gilmour, A.R., (2000). Post Blocking Gone Too Far! Recovery of Information and Spatial Analysis in Field Experiments. *Biometrics* 56, 944-946.
- Gilmour, A.R., Cullis, B.R., Welham, S.J., Gogel, B, and Thompson, R. (2004) An efficient computing strategy for prediction in mixed linear models *Computational Statistics and Data Analysis* 44: 571-586.
- Gilmour, A.R. (2007) Mixed model regression mapping for QTL detection in experimental crosses. *Computational Statistics and Data Analysis* 51:3749-3764
- Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J. and Thompson, R. (2015). ASReml User Guide Release 4.1 Functional Specification, VSN International Ltd, Hemel Hempstead, HP1 1ES, UK www.vsnl.co.uk