

Contents

Sponsors	2
Committees	3
Program	4
Abstracts – INVITED	12
Abstracts – CONTRIBUTED TALKS	23
Abstracts – CONTRIBUTED POSTERS	46
Author Index	55
Delegate List	56

Sponsors



Australasian Genstat User's Association Committee Inc.

Trevor Hancock:	Chairperson
Murray Hannah:	Vice Chairperson
Ray Correll:	Treasurer
Ari Verbyla:	Secretary
Roger Payne	Member
Jane Speijers	Member
Vivienne Doogan	Member
David Baird	Member
Jeff Wood	Member
Ruth Butler	Member

Conference Organising Committee

Trevor Hancock:	Convenor
Ari Verbyla:	Convenor Program Convenor
Ray Correll:	Treasurer
Roma Hardie:	Secretary
Julian Taylor:	Web Pages

Conference Program Committee

Ari Verbyla:	Convenor
Roger Payne	
David Baird	
Brian Cullis	
Jeff Wood	

Monday 4th December – Waite Campus, University of Adelaide

9:00 – 5:00	GenStat Advanced Skills Masterclass. Roger Payne, Darren Murray and Simon Harding
-------------	---

Tuesday 5th December

10:00	Bus Pickup – Granada Hotel
10:15	Bus Pickup – Waite Campus, Adelaide University
11:00	Bus Pickup – Adelaide Airport
12:30	Arrival at Whalers Inn Resort

12:00 – 5:00	Conference Registration – Flinders Room
1:00 – 2:00	Lunch & Welcome – Waterside Restaurant

Tuesday 5th December – Investigator Room

<i>Theme: Applications I Chair: Mario D'Antuono</i>	
2:00 – 2:25	Modelling the distribution of thrips caught in un-baited traps surrounding an odour baited trap. Ruth Butler ^a , David Teulon, Dale James & Melanie Davidson
2:25 – 2:50	Where's Smut? Vivienne Doogan ^a , Rosemary Kopittke & Tony Swain
2:50 – 3:15	Meta-Analyses of international trials on alternatives to METHYL BROMIDE, a major ozone depleting agent used for pre-planting soil fumigation. Debra Partington ^a , Murray Hannah & Ian Porter
3:15 – 3:40	Estimating mortality probabilities of Australian wild rabbits challenged with rabbit haemorrhagic disease using pre-challenge antibody cut-off values. Hwan-Jin Yoon ^a , Steve McPhee & Kym Butler
3:40 – 4:10	Afternoon Tea – Waterside Restaurant

<i>Theme: Applications II Chair: Ky Matthews</i>	
4:10 – 4:35	Statistics, the Law and the Thieving Fishers. Peter Johnstone ^a , Graeme Bremner & PJ Clarke
4:35 – 5:00	GenStat REML analyses for comparing sampling sites when measuring mohair quality on Angora goats. Kym Butler ^a & Bruce McGregor
5:00 – 5:25	Regression errors in “x” case study. David Saville ^a

6:30 – 11:00	Welcoming BBQ – Whalers Inn Resort (surrounding lawns) Live Acoustic from Brian Ruiz and Troy Loakes
--------------	---

^a Presenting Author

Wednesday 6th December – Investigator Room

<i>Theme: GenStat Session I Chair: Vivienne Doogan</i>	
8:15 – 8:30	Overview of Genstat 9th Edition. Roger Payne ^a
8:30 – 8:45	Genstat for Linux. Simon Harding ^a
8:45 – 9:00	Workbooks and other spreadsheet enhancements. David Baird ^a & Darren Murray
9:00 – 10:00	Invited: Double Hierarchical Generalized Linear Models. John Nelder ^a
10:00 – 10:30	Hierarchical Generalized Linear Models in practice. Roger Payne ^a
10:30 – 11:00	Morning Tea – Waterside Restaurant

<i>Theme: GenStat Session II Chair: Ruth Butler</i>	
11:00 – 11:25	GenBUGS – Running BUGS from GenStat. Darren Murray ^a
11:25 – 11:50	Hidden Markov Models. Roger LittleJohn ^a
11:50 – 12:05	Genstat Release 10. Roger Payne ^a
12:05 – 12:20	REML enhancements in Release 10. Sue Welham ^a
12:20 – 12:35	Custom Menus. Darren Murray ^a
12:35 – 1:00	Discussion
1:00 – 2:00	Lunch – Waterside Restaurant

^a Presenting Author

Wednesday 6th December – Investigator Room

<i>Theme: Poster Session Chair: Paul Eckermann</i>	
2:00 – 2:40	<p>Computations of some Scan statistics for detecting temporal clustering in animal diseases surveillance data. Mario D'Antuono^a & Chris Hawkins</p> <p>Detecting QTL for fruit quality traits in mango. Joanne De Faveri^a, Ian Bally, Natalie Dillon, Heather Smyth, Ralf Dietzgen</p> <p>GRAPES in Genstat – a procedure for monitoring sensory panel performance. Duncan Hedderley^a & Virginia Corrigan</p> <p>Spatial and temporal modelling of canopy temperature in wheat. Kathryn Lambkin^a, Ky Matthews & Scott Chapman</p> <p>Field assessment of faecal NIRS as a management tool for the northern cattle industry. David Reid^a & Desiree Jackson</p> <p>Use of Poisson Variation for Quality Assurance of Laboratory Worm Egg Counts. Andrew van Burgel^a</p> <p>Calculation of genetic parameters in a custard apple breeding program in south-east Queensland. Pat Pepper^a, Janet Giles, Michael MacBeth & Alan George</p>
<i>Biometrics Training Chair: Beverley Orchard</i>	
2:40 – 3:30	<p>GenStat Discovery in research methods training. Olena Kravchuk^a & Longbin Huang</p> <p>A magical mystery tour - demystifying basic statistical concepts, experimental design and analysis. Beverley Orchard^a, Lorraine Spohr & Sharon Nielsen</p> <p>Discussion</p>
3:30 – 4:00	Afternoon Tea – Waterside Restaurant

^a Presenting Author

Wednesday 6th December – Investigator Room

4:00 – 5:00	Australasian GenStat Users Association Inc. General Meeting
	<i>Special Invited Talk Chair: Chris Triggs</i>
5:30 – 6:30	An African Statistical Tale. Roger Stern ^a
7:00 – 9:00	Dinner – Waterside Restaurant

^a Presenting Author

Thursday 7th December: StatGen – Investigator Room

8:30 – 8:35	Introduction Ari Verbyla
-------------	------------------------------------

<i>Theme: SNPs Chair: Peter Baker</i>	
8:35 – 9:05	Invited: Dad's dilemma: Do I increase my daughters allowance now so she can have her genome sequenced or do I just wait for Xmas? David Mtichell ^a
9:05 – 9:55	Invited: Fitting statistical models to data with many more variables than observations. Harri Kiiveri ^a
9:55 – 10:45	Invited: Locating genetic links to disease using SNP chips. Ian Saunders ^a
10:45 – 11:15	Morning Tea – Waterside Restaurant

<i>Theme: Darts Chair: Wayne Pitchford</i>	
11:15 – 11:45	Invited: DArT genome scans: challenges and opportunities. Peter Wenzl ^a , Grzegorz Uszynski, Justin Bedo, Conrad Sanderson, Alison Smith, Cyril Cayla, Brian Cullis, Adam Kowalczyk, Eric Huttner & Andrzej Kilian
11:45 – 12:35	Invited: Normalisation and identification of polymorphic markers in DArT microarray data. Brian Cullis ^a , Alison Smith & Peter Wenzl
12:35 – 1:00	A whole of genome approach to QTL identification. Simon Diffey ^a
1:00 – 2:00	Lunch – Waterside Restaurant

^a Presenting Author

Thursday 7th December: StatGen – Investigator Room

<i>Theme: Statistical Genetics Chair: Peter Johnstone</i>	
2:00 – 2:25	Multi-trait QTL analysis using the complete linkage map. Ari Verbyla ^a , Brian Cullis, Christine Hackett, Adrian Newton & Bill Taylor
2:25 – 2:50	Estimating marker dosage and other challenges in linkage analysis and QTL mapping in sugarcane. Peter Baker ^a , Karen Aitken & Phillip Jackson
2:50 – 3:15	Divide and Conquer: partitioning genetic line effects into additive and non-additive effects in multi- environment trials. Helena Oakey ^a , Arunas P. Verbyla, Brian R. Cullis, Xianming Wei and Wayne S. Pitchford
3:15 – 3:45	Afternoon Tea – Waterside Restaurant

<i>Robin Thompson: A tribute Chair: Ari Verbyla</i>	
3:45 – 3:50	Introduction Ari Verbyla
3:50 – 4:20	Invited: Linear mixed models, REML and gene mapping. Peter Visscher ^a
4:20 – 4:50	Invited: REML, Rothamsted and Robin: A review. Roger Payne ^a
4:50 – 5:20	Invited: Colonial Collaborations – applications of REML for plant improvement. Brian Cullis ^a
5:20 – 5:25	Final Remarks Robin Thompson

6:30	Bus Departs from Whalers Inn to Currency Creek Winery
7:00 – 11:30	Conference Dinner, Currency Creek Winery Live Music from Geoff and his Technicolor Dream Coates

^a Presenting Author

Friday 8th December – Investigator Room

<i>Theme: Variety Trials Chair: Murray Hannah</i>	
8:30 – 8:55	Multi-environment QTL mixed models for drought stress adaptation in wheat. Ky Matthews ^a , Fred van Eeuwijk, Matthew Reynolds, C. Lynne McIntyre, Scott Chapman
8:55 – 9:20	On Efficiency of Barley Quality Testing in Australia. Katia Stefanova ^a , Alison Smith & Brian Cullis
9:20 – 9:45	Optimizing the barley crop variety testing program in Western Australia. Peter Clarke ^a
9:45 – 10:10	Evaluation of models for late-stage variety evaluation trials. Sue Welham ^a , Bev Gogel, Alison Smith & Brain Cullis
10:10 – 10:35	Models for Field Experiments with Repeated Ordinal Outcomes. Chris Triggs ^a , Arier Lee & John Andersen
10:35 – 11:00	Morning Tea – Waterside Restaurant

<i>Theme: Applications III Chair: Kym Butler</i>	
11:00 – 11:25	The distribution of tree diameters in native woodlands in Queensland. Christina Playford ^a & Steven Bray
11:25 – 11:50	Wambiana Grazing Trial: Impact of grazing strategies and variable rainfall on pasture composition. Angela Reid ^a , Yvette Everingham, Peter O'Reagain, John Bushell & Chris Holloway
11:50 – 12:15	Analysing zero-inflated fisheries counts using two-part conditional distributions. David Mayer ^a , Julie Robbins, Ian Halliday & Michelle Sellin

<i>Theme: Climate Change Chair: Julian Taylor</i>	
12:15 – 1:15	Invited: Climate Change – the near future. Roger Stern ^a
1:15 – 2:15	Lunch – Waterside Restaurant

2:30	Bus departs Whalers Inn Resort
4:00	Arrival at Adelaide Airport

^a Presenting Author

COLONIAL COLLABORATIONS - APPLICATIONS OF REML FOR PLANT IMPROVEMENT

Brian Cullis^a

Department of Primary Industries, Wagga Wagga, NSW, Australia.

On a visit to Australia in ca. 1990, Robin Thompson sowed the seeds for a collaboration that has, over the course of 15 years, revolutionised the approach to varietal selection in Australian plant improvement programs. It began, as always, with an airline boarding pass etched with a cryptic version of the average information algorithm for REML estimation of variance parameters. This, combined with the need to conduct a variance component analysis for a large unbalanced set of variety by environment data helped give rise to a computer program called AIREML. Of course this has evolved into the package now known as ASReml. ASReml has become the 'gold standard' for mixed models software in terms of the complexity of variance models that can be fitted, the ease with which quantities of interest can be predicted and the ability to handle large unbalanced data-sets. These features are underpinned by statistical research, the majority of which has been undertaken as collaborative work with Professor Thompson. The presentation will summarise some of this work, in particular key research aimed at improving genetic gain in plant improvement programs. The topics include multiplicative mixed models for variety by environment data, and the more general issues of prediction and inference in linear mixed models.

Email: brian.cullis@dpi.nsw.gov.au

^a Presenting Author

NORMALISATION AND IDENTIFICATION OF POLYMORPHIC MARKERS IN DART MICROARRAY DATA

Brian Cullis^{a,1}, Alison Smith¹ and Peter Wenzl²

¹*Department of Primary Industries, Wagga Wagga, NSW, Australia*

²*Diversity Arrays Technology P/L, Canberra, ACT*

DArT (Diversity Arrays Technology) is a high through-put wholegenome fingerprinting tool. Polymorphic markers are identified by hybridising fluorescently labelled representations of genomic DNA samples (referred to as target samples) to microarrays containing a large collection of clones, each typically spotted more than once per array. Each target sample is hybridised together with a reference DNA fragment (labelled with a different fluorescent dye) to measure the amount of DNA present in each spot on the array. The printing of arrays on glass slides is conducted using an automatic but complex process that will be described in detail. The key data obtained are the hybridisation intensities for both the target and reference channels for individual spots on each slide. Currently it is the log ratio of target to reference intensity that is used as the basis for analysis. Polymorphic clones are then identified as those showing ‘statistically significant’ differences across the targets tested.

Two important statistical issues for DArT data are methods for data normalisation and for identification of polymorphic clones. In this talk we present a mixed model approach for analysis that simultaneously accounts for nuisance variation between spots on each slide and identifies polymorphic clones. The approach is bivariate in nature, using the individual channel data rather than the ratio. A key component of the analysis is the identification of outliers. This is done at both a residual and clonal level. Our approach is based on the Alternative Outlier Model (AOM) for ordinary linear models in which outliers are assumed to arise from residuals with inflated variance. We generalise this concept for linear mixed models in which outliers may arise not only with respect to residual effects but also other random effects in the model. In the context of DArT data, the technique may first be used to identify outliers at the residual level as part of the data cleaning process. We then propose to identify clones that are outliers in the sense of having inflated clone by sample interaction variance as this may be indicative of polymorphism. In the context of AOM, effects with inflated variance may be identified using score tests. We propose an efficient resampling method for calculating a threshold for the test statistics.

In this talk we present the analysis of an example comprising the evaluation of 3456 clones across 93 slides, each hybridised with a different target sample.

Email: brian.cullis@dpi.nsw.gov.au

^a Presenting Author

FITTING STATISTICAL MODELS TO DATA WITH MANY MORE VARIABLES THAN OBSERVATIONS

Harri Kiiveri^a

CSIRO Mathematical & Information Sciences, WA, Australia.

In this talk I'll present a general method which can be used to fit a large class of statistical models to data with more variables than observations. The method involves simultaneous model fitting and variable selection and often produces (very) sparse well fitting models.

I'll give some illustrations from bioinformatics using gene expression and single nucleotide polymorphism (SNP) data sets with numbers of observations of the order of 100 and numbers of variables ranging from thousands to millions.

Email: harri.kiiveri@csiro.au

^a Presenting Author

DAD'S DILEMMA: DO I INCREASE MY DAUGHTERS ALLOWANCE NOW SO SHE CAN HAVE HER GENOME SEQUENCED OR DO I JUST WAIT FOR XMAS?

David Mitchell^a

CSIRO Mathematical & Information Sciences, NSW, Australia.

The last few years have seen a massive increase in the amount of genotyping information collected through the advent of new genotyping platforms. Doubling in information content every other year, it is now possible to reliably undertake whole genome association studies, in humans at least. On another front the cost of genome sequencing has decreased dramatically. Is there a logical end point to this technological development?

Where humans lead, other species follow. It is reasonable to assume that in short order many of our commercial crops and animals will soon have the same genotyping resources available as we have for humans. Will this change the way we undertake research? What will be the impact on “traditional” linkage studies and when should we stop what we do now and start doing something different?

Email: David.Mitchell@csiro.au

^a Presenting Author

DOUBLE HIERARCHICAL GENERALIZED LINEAR MODELSJohn A. Nelder^a*Imperial College, London, UK.*

GLMs are extended in three ways:

- (1) Joint modelling of both mean and dispersion
- (2) Random effects in the linear predictor for the mean model
- (3) Random effects in the linear predictor for the dispersion model

A single algorithm suffices to fit all models of the class, the analysis reducing to that of an interconnected set of GLMs. Survival models with random effects fall within the framework, also models with correlated random effects. The analysis does not require prior probabilities, quadrature or the use of the EM algorithm. Examples will be given of a wide range of applications, including new models for financial data and a new method of denoising pictures with missing pixels.

Email: j.nelder@imperial.ac.uk

^a Presenting Author

REML, ROTHAMSTED AND ROBIN: A REVIEWRoger Payne^a*VSN International, Hemel Hempstead, UK.*

The methodology of residual maximum likelihood (REML), developed with Desmond Patterson at Edinburgh in 1971, represents the first of Robin Thompson's major contribution to Statistics. REML provided a reliable and efficient framework for the analysis of linear models that contain several sources of random variation and whose component terms are not balanced (as they would be in ordinary analysis of variance). It also provided a precursor of Robin's future career at Rothamsted, in that it agreed with the Rothamsted-derived methodology of general balance where both were relevant, and extended the ideas to a wider range of situations.

This talk will review REML and its history, and discuss some of Robin's related contributions to Rothamsted, its statistics and GenStat.

Email: roger.payne@vsni.co.uk

^a Presenting Author

A STATISTICAL TOUR IN AFRICA

Roger Stern^a

Statistical Services Centre, University of Reading, UK.

Africa is on the move, and so is their statistics. A statistical tour is taken through a few African countries, starting in Kenya and Malawi and proceeding to Zambia, Uganda and Rwanda, mentioning other countries en route.

A “starter-pack” of good seeds and fertilizer was distributed to all the 2 million (or 3 million) rural families in Malawi, to help food security. Farmers are migrating from the South of Zambia, citing climate change as the reason while a local NGO thinks the main problem is unsustainable farming practices.

Improved statistics teaching in the University of Nairobi led to training support for government statistics in Kenya, and as well as hope for improved rights for donkeys. In Uganda, a broadened view of the scope of statistics may improve postgraduate research in agriculture, while the need for capacity-building in Rwanda’s national statistical system seems similar to recent work in neighbouring countries.

Email: r.d.stern@reading.ac.uk

^a Presenting Author

CLIMATE CHANGE – THE NEAR FUTURE

Roger Stern^a

Statistical Services Centre, University of Reading, UK.

Climate change is considered with a 10 to 20-year time horizon. This includes the time-scale for the Millennium Development Goals and is in our own lifetimes.

On this time-scale the climate modellers and donors agree that the study of climate change must proceed through an analysis of the existing climate variability. Hence it is important to analyse the existing historical records fully and understand how users could cope better with the existing vagaries and extremes.

Every country has a National Met Service and they are the main custodians of the historical climatic data. They need some support to process their data for the wide range of activities that depend on climate variability. Often the just simple statistics is needed, but there is also ample scope for modern modelling methods.

Email: r.d.stern@reading.ac.uk

^a Presenting Author

LOCATING GENETIC LINKS TO DISEASE USING SNP CHIPS

Ian Saunders^a

CSIRO Mathematical & Information Sciences, Adelaide, Australia.

The CSIRO Preventative Health Flagship has a major focus on the early detection of colorectal cancer (CRC). Included is a search for predisposing genes among CRC cases with familial association where the causal genes are unknown: 20% of all CRC. A preliminary sibling pair study using the Affymetrix GeneChip Human Mapping 50k Xba 240 Array is underway.

The high density of SNPs allows the allele sharing identical by descent (IBD) to be determined accurately from sib-pair data without parental genotypes. Using IBD status as the basis for testing allows the SNP haplotypes to vary between families without affecting the results. The analysis does not use test values at individual SNPs, but rather properties across the genome. Various properties have been studied, particularly those of runs of single-SNP test values above a critical level.

Simulations of these ensemble statistics determine critical values and the power of tests in the presence of single or multiple cancer susceptibility (CS) genes. The results indicate that a study of several hundred sib pairs will allow simultaneous location of two or three CS genes.

A preliminary study of 28 sib pairs known to contain genes related to Lynch Syndrome demonstrates the feasibility of the approach.

Email: Ian.Saunders@csiro.au

^a Presenting Author

LINEAR MIXED MODELS, REML AND GENE MAPPINGPeter M. Visscher^a*Queensland Institute of Medical Research, Brisbane, Australia.*

One of Robin Thompson's major contributions to statistical genetics has been the development and applications of methods to estimate variance components in linear mixed models with multiple random effects and complex covariance structures. In particular, the development and application of REML to estimate genetic (co)variances in large pedigreed populations has been widely adopted. A logical extension of the general framework of linear mixed models is to map genes affecting quantitative traits in pedigreed population using genetic linkage analysis. We discuss the development and application of (RE)ML variance component methods to map genes in outbred populations, and compare and contrast the approaches taken by statistical geneticists working in animal and human populations.

Email: peter.visscher@qimr.edu.au

^a Presenting Author

DART GENOME SCANS: CHALLENGES AND OPPORTUNITIES

Peter Wenzl^{a,1}, Grzegorz Uszynski¹, Justin Bedo², Conrad Sanderson², Alison Smith³, Cyril Cayla¹, Brian Cullis³, Adam Kovalczyk², Eric Huttner¹ and Andrzej Kilian¹

¹*Diversity Arrays Technology P/L, Canberra ACT, Australia,*

²*National ICT Australia Ltd., Canberra, ACT, Australia,*

³*Department of Primary Industries, Wagga Wagga, NSW, Australia.*

DArT (Diversity Arrays Technology) is a microarray-based genotyping technology that rapidly generates genome-wide genetic fingerprints. In the course of the last two years our laboratory has hybridised more than 40,000 microarrays and generated approximately ten million genetic data points. We have developed a laboratory information management system (DArTdb) and dedicated analytical software (DArTsoft) to manage and analyse the approximately two terabyte of microarray images generated in the process. Our primary focus has been on developing IT tools that fully automate the entire data-production process. The statistical analysis and mining of DArT data, however, is largely in its infancy. We will discuss key areas where a statistically more appropriate treatment of DArT data could enhance their information content and robustness and improve their usefulness for downstream genetic analyses. We will also highlight ways in which downstream applications such as genetic mapping and the identification of genotype-phenotype relationships might benefit from innovations in the analysis of high-density genetic data..

Email: peter@DiversityArrays.com

^a Presenting Author

ESTIMATING MARKER DOSAGE AND OTHER CHALLENGES IN LINKAGE ANALYSIS AND QTL MAPPING IN SUGARCANE

Peter Baker^{a,1,2}, Karen Aitken^{1,3}, Phillip Jackson^{1,4} and Christine Hackett⁵

¹ *Cooperative Research Centre for Sugar Industry Innovation through Biotechnology, University of Queensland, QLD, Australia,*

² *CSIRO Mathematical and Information Sciences, St. Lucia, QLD, Australia,*

³ *CSIRO Plant Industry, St. Lucia, QLD, Australia,*

⁴ *CSIRO Plant Industry, Aitkenvale, QLD, Australia,*

⁵ *Bioinformatics, Biomathematics and Statistics Scotland, Dundee, DD2 5DA, Scotland.*

Due to its complex genetics, sugarcane poses unique statistical challenges.

CSIRO in partnership with BSES through its involvement in the Cooperative Research Centre for Sugar Industry Innovation through Biotechnology is engaged in breeding programs for sugarcane by developing molecular markers and searching for quantitative trait loci (QTL) affecting important traits such as sucrose content and smut resistance. By finding markers associated with genes affecting a quantitative trait, breeders plan to employ marker assisted selection to speed up the breeding of new varieties.

Like some plants of agricultural importance, sugarcane is a polyploid in that it has more than two copies of each chromosome per cell. Unlike many such crops, it is also an autopolyploid in that it does not appear to exhibit preferential pairing. Given the added complexity of the genetics of autopolyploids, current QTL mapping methods for autopolyploids are not as well advanced nor as well tested as those for diploids. Currently, when dominant markers such as AFLPs are employed, the first step in autopolyploid linkage analysis or QTL mapping is to identify single dose markers which can then be treated as diploid for subsequent analysis. Current methods of assessing marker dosage appear to have limitations such as not allowing for multiple testing or measurement error. A Bayesian mixture model is outlined and compared to existing methods of classifying marker dosage for AFLPs. The relative information obtained for linkage of single dose markers to multi-dose markers is also presented. Implications for constructing marker maps and subsequent QTL analysis will be discussed.

Email: peter.baker@csiro.au

^a Presenting Author

GENSTAT REML ANALYSES FOR COMPARING SAMPLING SITES WHEN MEASURING MOHAIR QUALITY ON ANGORA GOATS

Kym Butler^a and Bruce McGregor

Departmentt of Primary Industries, Melbourne, Australia.

Using REML in GenStat to analyse data with more complex variance structure needs care. The relevant GenStat code to obtain the required output is not always obvious. This talk summarises a series of REML analyses for comparing sources of variability of key mohair quality attributes, at various sampling sites, on the fleeces of Angora goats. Average differences between sites, and how these differences vary between gender and flock, are obtained. Also differences amongst sites in between sire and between animal variability, and the covariance/correlation amongst sites at both the sire and individual goat level, are examined. The GenStat code to examine and fit models, and to obtain relevant parameter estimates with standard errors and confidence intervals, is discussed.

Email: kym.butler@dpi.vic.gov.au

^a Presenting Author

MODELLING THE DISTRIBUTION OF THRIPS CAUGHT IN UN-BAITED TRAPS SURROUNDING AN ODOUR BAITED TRAP

Ruth Butler^a, David Teulon, Dale James and Melanie Davidson

Crop & Food Research, Christchurch, New Zealand.

A field bioassay was undertaken to examine the distance over which an attractant chemical odour located at a trap can influence the number of thrips captured in nearby traps without the odour. Replicated arrays of water traps were used, consisting of a centre trap with or without odour surrounded by traps without odour at 0.5, 1, 2, 5, and 10m in eight directions for 48 hrs. The total number of thrips in each trap were counted. A sample of thrips from a subset of the traps were identified to species level, and the number of each species recorded. Initial Poisson analyses of the count data showed that thrips numbers in arrays with odour at the centre were substantially increased over numbers in arrays without the odour, and that the number of thrips in a trap varied with both the distance and direction from the centre trap. The distribution of thrips was examined more fully. First, the total number of Thrips caught in traps were modelled with a generalized non-linear model, using a non-standard logistic link to the distance of individual traps from the centre trap. This model was then modified by including an offset, to allow the number of thrips of each species to also be modelled. Thrips downwind of the centre traps were 10% above the background level for up to 5m from the trap for species attracted to the odour, whereas numbers in other directions had returned to close to the background by 1m from the trap.

Email: ButlerR@crop.cri.nz

^a Presenting Author

OPTIMIZING THE BARLEY CROP VARIETY TESTING PROGRAM IN WESTERN AUSTRALIA

Peter G. Clarke^a

Agriculture WA, Perth, Australia.

The problem addressed is where to site stage 4 barley trials and how many trials should be laid out each year.

The question of trial numbers is partly answered by calculating variance components from historical data and by a resampling exercise with varying trial numbers in the different agro-ecological zones.

The question of best sites is answered by using a mixture of multi-dimensional scaling to identify similar sites and hierarchical cluster analysis to group similar sites together.

Email: pclarke@agric.wa.gov.au

^a Presenting Author

A WHOLE OF GENOME APPROACH TO QTL IDENTIFICATION

Simon Diffey^a

Department of Primary Industries, Wagga Wagga, NSW, Australia

A QTL analysis requires two sources of data – phenotypic trait data and genotypic data. In a double haploid population genotypic information takes the form of breeding lines being scored according to the parental allele present at a set of molecular markers. These molecular markers are then assigned to linkage groups which generally represent chromosomes. The need for efficient designs and analyses of phenotypic trait data in a plant breeding context is generally well understood. The linear mixed model is a flexible tool for analysing phenotypic trait data as it allows a wide range of variance models and can accommodate missing and unbalanced data. The aim is to develop statistical models for plant breeding data that are better representations of reality and that account for both genetic and non-genetic sources of variation.

Genes or genomic regions that influence traits of interest are known as quantitative trait loci (QTL). It is of interest to find the number, genomic region, and the genetic effect of QTL. A whole of genome approach to QTL identification developed by Verbyla et al. (2006) is presented. This method has been shown to perform very well in a simulation study. In this method the statistical model developed for the phenotypic data is subsequently used in an algorithm that determines the best subset of marker information that explains the underlying genetic variation.

This approach has been implemented in the statistical software package R and uses the *samm* library for fitting the linear mixed model. This implementation allows a suitable model for the phenotypic trait data to be easily developed by considering sources of variation, heritability, outliers and model diagnostics before progressing to QTL identification. An example using data from a double haploid wheat population is presented.

Email: simon.diffey@dpi.nsw.gov.au

^a Presenting Author

WHERE'S SMUT?

Vivienne Doogan^a, Rosemary Kopittke and Tony Swain

Department of Primary Industries & Fisheries, Brisbane, QLD, Australia

Sugarcane smut, which is caused by a fungus, *Ustilago scitaminea*, was first found in Queensland near Childers in June 2006. Surveillance strategies designed to detect smut were developed to survey blocks of cane on the infected and surrounding properties. Due to operational issues these surveillance plans were based on two inspectors walking down an inter-row inspecting a double row of cane as the basic sampling unit.

To provide the sugarcane smut surveillance team with options, sampling strategies for detection of smut were designed for a range of confidence levels, prevalence levels and sensitivities of detection. These prevalence levels were all based on the detection of infected plants within the double-row sampling unit. However, the surveillance team wanted to be able to express their results on a per plant basis, rather than a double-row basis.

Simulations were conducted in GenStat to assess the effectiveness of inspection methods based on random and systematic examination of individual plants, single rows and double rows on a 'typical' block of sugarcane for a range of random and clustered infection patterns.

Results for simple random sampling of plants for inspection agreed with theory. For other surveillance schemes the probabilities of detection estimated from the simulations followed expected trends, but gave surprisingly low values for some configurations.

Results of the simulations will be presented and their practical implications discussed.

Email: Vivienne.Doogan@dpi.qld.gov.au

^a Presenting Author

STATISTICS, THE LAW AND THE THIEVING FISHERS

Peter Johnstone^{a,1}, Graeme Bremner² and P.J. Clarke²

¹*AgResearch Ltd, Mosgiel, New Zealand,*
²*Ministry of Fisheries, Dunedin, New Zealand.*

The purpose of the studies we will present was to apprehend and prosecute illegal fishers. The techniques we have used include standard model fitting and simulation studies. Although the results of these studies were convincing to us, the law and authorities were skeptical. We conclude that the disconnect between statistics and the law is incredulously large. We offer no solutions to encourage a reconnection.

Email: peter.johnstone@agresearch.co.nz

^a Presenting Author

HIDDEN MARKOV MODELS IN GENSTAT

Roger Littlejohn^a

AgResearch, Invermay, New Zealand

Hidden Markov models (HMMs) provide a useful class of models for a wide range of applications. There is assumed to be an underlying unobserved Markov chain, which moves through a set of states. For each state the observed process has a different distribution. Using the Baum-Welch algorithm it is possible to estimate the parameters of the observed process and the transition probabilities for the underlying process. I will discuss fitting HMMs in GenStat, with an application to sheep urination data.

Email: roger.littlejohn@agresearch.co.nz

MULTI-ENVIRONMENT QTL MIXED MODELS FOR DROUGHT STRESS ADAPTATION IN WHEAT

Ky Mathews^{a,1}, Fred van Eeuwijk², Matthew Reynolds³,
C. Lynne McIntyre¹ and Scott Chapman¹

¹*CSIRO, Plant Industry, Brisbane, Australia,*

²*Wageningen University Research, Wageningen, The Netherlands,*

³*ACIMMYT Mexico, Mexico City, Mexico.*

Mixed models are commonly used in plant breeding programs to understand genotype by environment interaction (G×E) in multi-environment trials. Variations in quantitative trait loci (QTL) expression across environments, QTL×E, underlies G×E. By incorporating molecular data into the mixed model framework we can model both the phenotypic (G×E) and genotypic (QTL×E) effects across environments. In complex traits, such as yield response to drought stress, major QTLs are unlikely and the investigation of smaller QTL×E effects is necessary.

The majority of commonly used QTL detection methods do not consider multiple environments and hence investigation of QTL×E effects is limited. Further, these methods are quite rigid and limited in their capacity to include explanatory covariates to account for real-life anomalies. We present a mixed modelling approach to detect QTL×E and demonstrate its flexibility to incorporate explanatory covariates.

A recombinant inbred spring wheat population was developed from parents with contrasting performance under drought stress, and evaluated in six trials in the northern region of Australia representing a range of drought stressed environments. One parent contained a rye translocation which had consequences for the molecular map. A covariate was used to accommodate this information. In addition, disease scores were measured on one trial and were also included as a covariate.

We present methodology aimed at the identification of QTLs that are stable across environments and suitable for targeting in a breeding program. We demonstrate the flexibility of this approach to accommodate the realities faced by plant breeding programs. We identified six putative QTLs with small effects and were able to relate their effects with environmental information.

Email: Ky.Mathews@csiro.au

^a Presenting Author

ANALYSING ZERO-INFLATED FISHERIES COUNTS USING TWO-PART CONDITIONAL DISTRIBUTIONS

David Mayer^a, Julie Robins, Ian Halliday and Michelle Sellin

Department of Primary Industries & Fisheries, Brisbane, QLD, Australia

The crucial role of freshwater flows into estuaries was investigated in a research project funded by the Coastal Zone CRC and Fisheries Research and Development Corporation. Stratified sampling, covering a number of flow events, was conducted in three central Queensland estuaries from 2002 to 2005. For each species, the count data from the two-minute spatial and temporal trawls typically include many zeros, and the distributions are significantly skewed. For our most common species (an anchovy), 44% of the samples had zero counts, and the maximum was 196.

This inflated zero-class violates the statistical assumptions of many standard analytical techniques. A more appropriate method is to use two-part conditional distributions – firstly, the Binomial to represent the proportion of zeros (simple presence or absence of each species in each trawl), and secondly, if present, a truncated distribution modelling the catch numbers (>0). For this second part there is a range of available distributions – researchers in ecology and entomology have typically used the Poisson or Negative Binomial for their discrete counts, whereas in meteorology, health statistics, fisheries and air-pollution research, continuous distributions such as the Gamma or log-Normal have been profitably employed.

Preliminary analyses showed the Gamma and the log-Normal to be superior and fairly equivalent, in terms of the distribution of residuals and stability of the estimates given this rather unbalanced data set. The zero-truncated Poisson was not suitable here, as it could not accommodate the more extreme skewness of these data.

For analyses, the log-Normal was used for the conditional part, and generalized linear model fitting was conducted in GenStat. The overall mean abundance for any given combination of model factors is the product of the Binomial proportion and the conditional mean. Taking *Penaeus merguensis* (banana prawns) as an example, standardized counts per trawl for each river, year and flow-condition were estimated. These were then related back to the freshwater flow amounts at these times and locations. The resultant patterns, and economic considerations of these, will be useful in determining water policy issues.

Email: david.mayer@dpi.qld.gov.au

^a Presenting Author

GENBUGS - RUNNING BUGS FROM GENSTATDarren Murray^a*VSN International, Hemel Hempstead, UK.*

This talk demonstrates how you can run a Bayesian analysis from GenStat. A new procedure is introduced that can be used to run WinBUGS or OpenBUGS from GenStat in batch mode using scripts. WinBUGS (Bayesian inference Using Gibbs Sampling, Spiegelhalter, Thomas, Best and Lund 2003) is an application that can be used for the Bayesian analysis of complex models using Markov chain Monte Carlo (MCMC) methods. WinBUGS is available for free at <http://www.mrc-bsu.cam.ac.uk/bugs/> and an open-source version of the core BUGS code (OpenBUGS) is also available at <http://mathstat.helsinki.fi/openbugs/>.

Spiegelhalter, D.J., Thomas, A., Best, N.G. and Lund, D. (2003). WinBUGS Version 1.4 Users Manual. MRC Biostatistics Unit, Cambridge.

Email: darren.murray@vsni.co.uk

^a Presenting Author

**DIVIDE AND CONQUER: PARTITIONING GENETIC LINE EFFECTS
INTO ADDITIVE AND NON-ADDITIVE EFFECTS IN MULTI-
ENVIRONMENT TRIALS**

Helena Oakey^{a,1}, Arunas Verbyla¹, Brain Cullis², Xianmin Wei and Wayne Pitchford¹

¹*School of Agriculture, Food and Wine, The University of Adelaide, SA, Australia,*

²*Department of Primary Industries, Wagga Wagga, NSW, Australia.*

A statistical approach to the analysis of multi-environment trials (METs) is presented, in which selection of best performing lines, best parents, and best combination of parents can be determined. The genetic effect of a line is partitioned into additive, dominance and residual non-additive effects. The dominance effects are estimated through the incorporation of the dominance relationship matrix, which is presented under varying levels of inbreeding. A computationally efficient way of fitting dominance effects is presented which partitions dominance effects into between family dominance and within family dominance line effects. The overall approach is applicable to inbred lines, hybrid lines and other general population structures where pedigree information is available.

Email: helena.oakey@adelaide.edu.au

^a Presenting Author

META-ANALYSES OF INTERNATIONAL TRIALS ON ALTERNATIVES TO METHYL BROMIDE, A MAJOR OZONE DEPLETING AGENT USED FOR PRE-PLANTING SOIL FUMIGATION

Debra Partington^a, Murray Hannah and Ian Porter

Department of Primary Industries, Melbourne, VIC, Australia.

Methyl Bromide has been widely used globally as a pre-planting soil fumigant in the horticultural industry. Due to its ozone depleting nature, effective alternatives have been sought. Whilst several alternatives have been identified, scepticism remains as to their efficacy, especially in some countries slow to embrace change or with trade interests to protect. In view of the importance of ozone depletion, a meta-analysis of international trials involving alternatives to methyl bromide was commissioned by the UN Methyl Bromide Technical Option Committee. This study was set up to summarise the state of knowledge on Methyl Bromide alternatives.

After considerable effort in constructing a database from relevant trials of adequate quality, meta-analyses were conducted in GenStat for yield data from several horticultural crops. These compared alternate chemical treatments, rates, methods of application, timings and management, in a range of environments. A number of acceptable alternative chemicals were identified.

We discuss some of the challenges faced in modelling such diverse data. These included

- Treatment classification definition,
- Modelling yield data with a wide array of measurement units and methods using a single model,
- Allowing for environmental factors like soil type, climate and level of pathogenic challenge, which could influence the relative performance of the chemical treatments, and
- Summarising and presenting the results from a complex model with simplicity.

Email: Debra.Partington@dpi.vic.gov.au

^a Presenting Author

HIERARCHICAL GENERALIZED LINEAR MODELS IN THE 9TH EDITION

Roger Payne^{a,1}, Youngjo Lee² and John Nelder³

¹*VSN International, Hemel Hempstead, UK,*

²*Seoul National University, Seoul, Korea,*

³*Imperial College, London, UK.*

Hierarchical generalized linear models (HGLMs) extend the familiar generalized linear models to situations where there may be more than one source of error variation. HGLMs include generalized linear mixed models (GLMMs) as a special case, but do not constrain the additional terms to follow a Normal distribution nor to have an identity link. They also allow for the modelling of the dispersion of the error terms, extending the ideas of Nelder & Pregibon (1987). The methodology provides improved estimation methods that reduce bias, and leads an efficient fitting algorithm which does not involve numerical integration.

The theory and practice of HGLMs has been greatly enhanced recently by the publication of a new expository book by Lee, Nelder & Pawitan (2006), and by the introduction of new procedures reimplementing the methodology for the 9th Edition of GenStat for Windows (Payne, Lee, Nelder & Noh 2006). This talk will illustrate the methodology, and its use in practice, with some of the worked examples from Lee, Nelder & Pawitan (2006) that are included with the 9th Edition.

Lee, Y., Nelder, J.A. & Pawitan, Y. (2006). Generalized Linear Models with Random Effects: Unified Analysis via H-likelihood. CRC Press.

Nelder, J.A. & Pregibon, D. (1987). An extended quasi-likelihood function. *Biometrika*, **74**, 221-232.

Payne, R.W., Lee, Y., Nelder, J.A. & Noh, M.S. (2006). Procedures HGANALYSE, HGDISEPLAY, HGDRANDOMMODEL, HGFIXEDMODEL, HGKEEP, HG PLOT, HGPREDICT and HGDRANDOMMODEL. GenStat Release 9 Reference Manual, Part 3 Procedure Library PL17, pp. 336-350.

Email: roger.payne@vsni.co.uk

^a Presenting Author

THE DISTRIBUTION OF TREE DIAMETERS IN NATIVE WOODLANDS IN QUEENSLAND

Christina Playford^a and Steven Bray

Department of Primary Industries & Fisheries, Rockhampton, QLD, Australia.

The monitoring of Queensland's native woodlands using TRAPS started in 1982 and currently monitors the life history of about 41000 woody plants located at a number of sites found throughout Queensland. The main objective in doing this was to develop an understanding of any changes in vegetation structure and population dynamics in native woodland. The data collected for each site included the size, location and species of individual plants at each recording date. Of particular interest here is the diameter of each plant. The diameter of each plant was calculated from the measured circumference taken at breast height (130cm from the ground). Plants with a circumference of less than 2 cm were not measured and were grouped into a single size class. In considering this the question of the distribution of tree diameters was raised and search of the literature undertaken. Subsequently a two part conditional distribution was considered and included a binomial distribution (small vs. large plants) and a gamma or exponential distribution (large plants).

Email: christina.playford@dpi.qld.gov.au

^a Presenting Author

WAMBIANA GRAZING TRIAL: IMPACT OF GRAZING STRATEGIES AND VARIABLE RAINFALL ON PASTURE COMPOSITION

Angela Reid^{a,1}, Yvette Everingham², Peter O'Reagain³,
John Bushell³ and Chris Holloway³

¹*Department of Primary Industries, Townsville, QLD, Australia,*

²*James Cook University, Townsville, QLD, Australia,*

³*Department of Primary Industries, Charters Towers, Australia.*

A commercial scale cattle grazing trial has been operational near Charters Towers (Queensland) since 1997 to assess the effect of different cattle grazing management strategies, under variable rainfall, on animal production and pasture composition.

The stocking rate strategies (heavy, light, variable, rotational resting and adjusting stocking level based on the Southern Oscillation Index (SOI)) were assigned to 10 paddocks (~100ha each) in a randomised block design. Intensive pasture sampling was conducted, in at least two 1 hectare sites within each of the three major soil types present in each paddock, to collect species frequency data. This paper will refer only to species composition data collected from 1998 to 2005. Rainfall was above average from 1998 to 2001 and well below average from 2002 to 2005.

Pasture species composition was distinctly different within each soil type (as evident in a biplot), thus data from each soil were analysed separately. Results will be presented for two of these soil types - texture contrast (TC) and clay. Principal components analyses indicated the main two drivers of pasture composition change were seasonal (rainfall) factors and grazing pressure, as represented by the first principal component (PC1) explaining up to 50 % of variation and the second principal component (PC2) explaining approximately 20% of variation respectively. Plots of the PC1 scores indicated a one year lag effect in species composition change when going from good to low rainfall years. For clay soils, PC2 indicated the heavy and SOI treatments were similar in pasture composition.

Repeated measures analysis of the PC1 scores showed no significant stocking rate by time interaction and no significant stocking rate effect ($p > 0.05$). The effect of time was highly significant ($p < 0.01$). For PC2, the significance of the stocking rate by time interaction was $p = 0.02$ for TC and $p = 0.10$ clay soils. For the TC soil, the effect of the SOI strategy was similar to light stocking in the earlier (wetter) years, but was similar to heavy stocking in the drier years. This has the management implication that the SOI strategy may work well in good rainfall years, but not in poor rainfall years.

Email: angela.reid@dpi.qld.gov.au

^a Presenting Author

REGRESSION ERRORS IN “X” CASE STUDYDave Saville^a*AgResearch Ltd, Lincoln, New Zealand.*

During my annual short courses in statistics, I discuss the linear regression model assumption that the “x values have no error”. Then, as a class exercise, each class member randomly selects data that is vertically distributed about the “true line” $y = 80 + 10x$ with a standard deviation (sd-y) of 20. To allow me to more intelligently comment upon the effect of violation of the “no errors in x” assumption, I recently decided to check out the behaviour of the linear regression model for 8 cases: “true slopes” of 0 and 10 in combination with standard deviations of the x-values of 0, 1, 2 and 4 (all with an sd-y of 20). In this talk I present the results from 10,000 simulations for each of the 8 cases, summarising the significance level/power for the test of the hypothesis “true slope=0”, the mean estimated slope, mean estimated variance and mean estimated elevation of the line at the mean x-value.

Email: dave.saville@agresearch.co.nz

^a Presenting Author

AN EFFICIENT RESAMPLING METHOD FOR OUTLIER DETECTION IN LINEAR MIXED MODELS

Alison Smith^{a,1}, Ari Verbyla² and Brian Cullis¹

¹*Department of Primary Industries, Wagga Wagga, NSW, Australia,*

²*School of Agriculture, Food and Wine, The University of Adelaide, SA, Australia.*

Identifying outliers in linear mixed model analysis is an important and difficult problem. Our approach is based on the Alternative Outlier Model (AOM) for ordinary linear models in which outliers are assumed to arise from residuals with inflated variance. We generalise this concept for linear mixed models in which outliers may arise not only with respect to residual effects but also other random effects in the model. Additionally the effects under scrutiny may be correlated. In the context of AOM, effects with inflated variance may be identified using score tests. We propose an efficient resampling method for calculating a threshold for the test statistics. The method is computationally less demanding than standard resampling procedures because the mixed model is fitted only once rather than in every resampling iteration. The procedure also provides a sample variance for each effect that can be used as the denominator in the score test rather than using expected information which may be difficult to compute. We present several simulation studies to show the performance of the method in terms of controlling Type I error rates.

Email: alison.smith@dpi.nsw.gov.au

^a Presenting Author

ON EFFICIENCY OF BARLEY QUALITY TESTING IN AUSTRALIA

Katia Stefanova^{a,1}, Alison Smith² and Brian Cullis²

¹*Department of Agriculture & Food, Perth, WA, Australia,*

²*Department of Primary Industries, Wagga Wagga, NSW, Australia.*

Barley quality traits associated with malting are of prime importance for variety selections and recommendations. In recent years in Australia there has been greater recognition of the need for appropriate experimental design and analysis for quality trait data.

The techniques described in Smith et al (2006) are now widely used for wheat quality testing. The accumulation of data derived from these experiments has enabled the investigation of efficiency and resource allocation for wheat quality testing in Australia.

There has been a major rationalisation of barley breeding in Australia recently, resulting in only three major programs. We have now assembled a barley quality data set that is based on a number of appropriately designed experiments from these programs.

In this presentation we describe the data, examine sources of variation (including variation between varieties, variety x environment interaction, variation between plots in the field and variation between samples in the laboratory), then examine the efficiency of testing schemes currently in use.

Email: kstefanova@agric.wa.gov.au

^a Presenting Author

MODELS FOR FIELD EXPERIMENTS WITH REPEATED ORDINAL OUTCOMES

Chris Triggs^{a,1}, Arier Lee¹ and John Anderson²

¹*University of Auckland, Auckland, New Zealand,*

²*Crop and Food Research, Pukekohe, New Zealand*

World wide, potato is considered one of the most important vegetable crops. Late blight caused by *Phytophthora infestans* is recognised as the most serious potato disease. A biennial field screening trial for resistance to late blight has been carried out at Pukekohe for over twenty years. Trials have been laid out as latinised row and column designs in a single rectangular array of plots, indexed by rows and columns. In each trial disease severity based on the percentage of affected foliage was repeatedly assessed on a 1-9 ordinal scale from the first sign of infection in each trial and at four to six subsequent occasions. Based on a McCullagh threshold model for ordinal responses the use of a model which takes into account repeated ordinal measurements and random effects arising from latinised row and column design is illustrated.

Email: triggs@stat.auckland.ac.nz

^a Presenting Author

MULTI-TRAIT QTL ANALYSIS USING THE COMPLETE LINKAGE MAP

Ari Verbyla^{a,1}, Brian Cullis², Christine Hackett³ and Adrian Newton⁴

¹*School of Agriculture, Food and Wine, The University of Adelaide, SA, Australia,*

²*NSW Department of Primary Industries, Wagga Wagga, NSW, Australia,*

³*Bioinformatics, Biomathematics and Statistics Scotland, Dundee, DD2 5DA, Scotland,*

⁴*Plant Pathology, Scottish Crop Research Institute, Dundee, DD2 5DA, Scotland.*

A split-plot experiment was conducted in the field in which treatments were applied to lines and parents of a barley doubled haploid population. The aim was to find and examine the possible differential expression of the QTL in relation to a control treatment. The analysis required a multi-trait (treatment) approach and was based on using the full linkage map simultaneously and an outlier detection method. A mixed model was used in which genetic, design and environmental effects were included. The QTL detection was based on a factor analytic or reduced rank formulation. Only pleiotropic QTL were detected using this approach but this included epistatic interactions.

Email: Ari.Verbyla@csiro.au

^a Presenting Author

EVALUATION OF MODELS FOR LATE-STAGE VARIETY EVALUATION TRIALS

Sue Welham^{a,1}, Bev Gogel², Alison Smith³ and Brian Cullis³

¹*Rothamsted Research, Harpenden, UK,*

²*School of Agriculture, Food and Wine, The University of Adelaide, SA, Australia,*

³*Department of Primary Industries, Wagga Wagga, NSW, Australia.*

The primary aim of crop variety evaluation programs is to reliably predict the future performance of potential new varieties relative to existing commercial varieties. This is achieved through multi-environment trials (METs), that is, series of field trials conducted across a range of geographic trial locations and possibly over several years. Varieties are tested for a range of agronomic traits with the aim of recommending superior varieties for commercial release. Usually, a program of METs is established to take varieties through from the initial selection of potential breeding lines to eventual commercial release and recommendation to farmers (late-stage variety evaluation trials). In this talk we are concerned with METs in the latter stages of this process. Many models have been proposed for MET data, which can be broadly classified as one-stage or two-stage analyses (Smith, Cullis & Thompson, 2005). In a one-stage analysis individual plot data from all trials is combined in a single analysis. In a two-stage analysis variety means are first obtained from the separate analysis of individual trials ({stage I}), and are then combined in an overall mixed model analysis ({stage II}). The stage II analysis may be unweighted or weighted to reflect the relative precision of variety means from each trial. In each case, a linear mixed model may be constructed to describe the structure of the data, and this model is usually fitted using REML estimation. Within this context, variety by environment effects are commonly modelled using either simple variance components or a factor analytic model. Within-trial variation may be modelled using either a spatial model or blocking factors corresponding to the design. This talk uses the characteristics of real MET datasets from Australia and the UK as a basis for a simulation study to evaluate these different mixed modelling approaches. Results and their implications for the analysis of late-stage variety testing programs are presented.

Smith, AB, Cullis, BR & Thompson R (2005) The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. *J. Agric. Sci.*, **143**, 1-14.

Email: sue.welham@bbsrc.ac.uk

^a Presenting Author

ESTIMATING MORTALITY PROBABILITIES OF AUSTRALIAN WILD RABBITS CHALLENGED WITH RABBIT HAEMORRHAGIC DISEASE USING PRE-CHALLENGE ANTIBODY CUT-OFF VALUES

Hwan-Jin Yoon^a, Steve McPhee and Kym Butler

Department of Primary Industries, Melbourne, Australia.

After the initial introduction of Rabbit Haemorrhagic Disease Virus (RHDV) into wild rabbits in Australia in 1995-96, the disease spread rapidly and became endemic in most rabbit prone areas. Understanding the epidemiology of introduced diseases and their impact on rabbit populations is necessary for optimising management decisions to reduce the damage they cause. The epidemiology of RHDV has been complicated due to benign forms calicivirus (bCV), co-circulating with RHDV and providing variable protection from RHDV. As yet, no specific virus has been isolated or described and so monitoring the antibody status of rabbit populations provides the only mechanism to indicate the presence of bCV and RHDV in the field. Quantitative measures of antibody values, using competition enzyme-linked immunosorbent assays (cELISA) and antibody isotyping ELISAs (isoELISAs), provided an indirect means to detect antibodies to bCV and more directly RHDV. To gain a better understanding of the epidemiology of RHDV, wild rabbits captured from four locations in Victoria were challenged with a lethal oral dose of RHDV and the pre-challenge antibody values were used to determine survival probability. Antibody cut-off values derived from the cELISA and the isoELISA's determined bCV or RHDV antibody class. Graphical exploratory data analysis and logistic regression modelling examined the relationship between the affect on the probability of survival relative to changes in antibody cut-off values describing bCV and RHDV antibody class.

Email: hwan-jin.yoon@dpi.vic.gov.au

^a Presenting Author

COMPUTATIONS OF SOME SCAN STATISTICS FOR DETECTING TEMPORAL CLUSTERING IN ANIMAL DISEASES SURVEILLANCE DATA

Mario D'Antuono^a and Chris Hawkins

Department of Agriculture and Food, Perth, WA, Australia.

As part of veterinary surveillance, there is an ongoing need to assess if several occurrences of an animal disease in a nominated period of time have exceeded the normal background level. If we can identify a “temporal cluster”, then this provides an indicator of the need for further investigation.

A recently published reference book discusses the use of a Scan test or statistic to test for temporal clusters in a time series of events retrospectively. The introduction of a Scan test via an approximate formula poses difficulties in understanding its derivation for non-statisticians and statisticians. It is even more of concern that these formulae can easily be applied without proper checks of the assumptions, and can calculate probability values outside the range 0 to 1.

An alternative approach is to use Monte Carlo simulation. It is practical to examine the distribution of expected clustering for observed data over time. By applying some simple probability models of the process that may have generated the given data, it is feasible to demonstrate clustering when and if it occurs.

We describe the computational processes and tools developed to assist in the application of Scan statistics to veterinary surveillance. In this way we avoid the problems of understanding and using formulae for approximating difficult test statistics. These computational processes will always produce P-values in the correct range, and also are more intuitive and informative because they give the user an insight into the processes that are used to evaluate the data.

Procedures to compute these statistics have been written to find an estimate of the probabilities of observing “k” events in a time-interval of width “w” for a number of scenarios and are available from <http://www.agric.wa.gov.au/biometrics>.

Email: mdantuono@agric.wa.gov.au

^a Presenting Author

DETECTING QTL FOR FRUIT QUALITY TRAITS IN MANGO

Joanne De Faveri^{a,1}, Ian Bally¹, Natalie Dillon¹, Heather Smyth² and Ralf Dietzgen²

¹*Department of Primary Industries & Fisheries, Mareeba, Australia,*

²*Department of Primary Industries & Fisheries, Brisbane, QLD, Australia.*

One of the aims of the recently commenced Queensland DPI&F Mango Genomics Project is to find molecular markers associated with genes affecting quantitative flavour and colour traits in mangoes. These markers may enable marker assisted selection to be implemented in the Australian mango breeding program.

Mangoes are an outcrossing diploid species with $2n=2x=40$ chromosomes (however there are some polyploid varieties). In this study related crosses of a number of mango varieties (each with >100 progeny) have been created and the progeny have been phenotyped for fruit quality traits. Expressed sequence tag (EST) libraries are being created and these ESTs along with SSRs & SNPs will be used to create a genetic linkage map to be used for mapping QTLs for flavour and colour traits.

Some of the statistical aspects to be considered in the linkage analysis & QTL mapping of this data will be presented.

Email: Joanne.DeFaveri@dpi.qld.gov.au

^a Presenting Author

GRAPES IN GENSTAT – A PROCEDURE FOR MONITORING SENSORY PANEL PERFORMANCE

Duncan Hedderley^a and Virginia Corrigan

Crop & Food Research, Palmerston North, New Zealand.

Panels of trained tasters are a useful tool for describing sensory differences between variants of foods (different batches, different formulations, different storage times). Ideally, each member of the panel will be able to discriminate between the samples using the attributes being scored; will be consistent in the scores they give the same samples on different occasions; and the panel will score the samples in a consistent fashion. Monitoring these things will help identify areas where the panel may need more training, or attributes which are not consistent enough to be useful. Schlich (1994) proposed a set of measures which can be calculated for each panellist on each attribute to address these questions, and graphed to compare panellists. We describe a Genstat version of this procedure.

Schlich, P (1994) GRAPES: a method and SAS® program for graphical representations of assessor performances, *Journal of Sensory Studies* **9**, 157-169

Email: hedderleyd@crop.cri.nz

^a Presenting Author

GENSTAT DISCOVERY IN RESEARCH METHODS TRAININGOlena Kravchuk^a and Longbin Huang*University of Queensland, Brisbane, QLD, Australia.*

The Discovery edition of GenStat has become popular in many developing countries since it was introduced in 2003. The edition contains the most acclaimed benefits of GenStat: diversity of statistical procedures, rigorousness of the analysis, and strong emphasis on design of experiments. This year, the School of Land and Food Sciences of the University of Queensland provided training in research methods to agricultural researchers in developing countries, partners of ACIAR. We have used the Discovery edition as an instrumental tool in the training course in Vietnam and have found a very favourable response from the participants towards the software. Knowing that the way Statistics is taught in the standard university statistical course in Vietnam is deficient in terms of the design and analysis of experiments, in the training course we aimed to demonstrate to the participants how much the quality of the data analysis improves when the analysis of an experiment is strongly linked to its design and research hypotheses. In our opinion, the way the contrast option is made available in the ANOVAs of common designs in the Discovery edition of GenStat adds educational value to the package. By teaching the participants how to use contrasts in GenStat, we not only taught them this feature of the package but introduced a better research practice to the researchers as they learned how to express their research hypotheses in the language of contrasts. Reflecting on our experience, we would suggest that using the edition in research methods training may help overcome many deficiencies in basic statistical university training in the developing world.

Email: o.kravchuk@uq.edu.au

^a Presenting Author

SPATIAL AND TEMPORAL MODELLING OF CANOPY TEMPERATURE IN WHEAT

Kathryn Lambkin^a, Ky Mathews and Scott Chapman

CSIRO Plant Industry, Brisbane, QLD, Australia

Canopy Temperature is an indicator of the water that a plant transpires and can be measured in both irrigated and drought conditions. It is a plot-level measurement taken in the field with a portable infra-red detector that measures point temperature. While spatial variation in soil conditions may affect the plot temperatures, the sampling procedure also introduces potential error. Large trials can take several hours to measure, during which time the ambient air temperature and wind may vary and confound results.

We model the design effects of row, column and replicate, fitting best spatial models to these data, also incorporating the measurement time effect. Canopy temperature was measured on several occasions between the vegetative and grain filling stages, allowing a repeated measures analysis of the data.

Email: kathryn.lambkin@csiro.au

^a Presenting Author

A MAGICAL MYSTERY TOUR - DEMYSTIFYING BASIC STATISTICAL CONCEPTS, EXPERIMENTAL DESIGN AND ANALYSIS

Beverley Orchard^{a,1}, Lorraine Spohr² and Sharon Nielsen³

¹*Department of Primary Industries, Wagga Wagga, NSW, Australia,*

²*Department of Primary Industries, Gosford, NSW, Australia,*

³*Department of Primary Industries, Orange, NSW, Australia*

Within NSW Department of Primary Industries (DPI), the demand for statistical consultancy and expertise continues to be strong. As part of the scientific process a scientist needs to communicate to a biometrician the nature of their experimental challenges and once the experiment has been conducted, to be able to carry out preliminary explorations of the data. Many research and technical staff recall the learning of statistical concepts as necessary but not ‘enjoyable or engaging’. Consequently, while there is a certain level of understanding of the validity of the concepts, there is an associated mystery or detachment which is evident when the concepts are discussed; this detachment is also reflected in scientific papers. The desire to state that treatments means are ‘greater’ or ‘less’ than some value regardless of confidence levels still prevails.

The process to ensure that the understanding of basic statistical concepts is ‘enjoyable and engaging’ is a well recognised challenge. In response, the biometrics unit of the NSW DPI has developed a short course which, not only improves the quantitative skills and understanding of basic statistical concepts for participants, but also increases their engagement with these basic concepts through active learning. This course provides the complete package: a two day course, an interactive CD and step-by-step practical notes for the use of statistical software including GenStat 9 for basic statistical analysis. Having been successfully conducted on eight occasions throughout NSW with encouraging assessments from participants, this course has achieved a balance of learning styles suitable for most adult learners in biological research field.

Email: beverley.orchard@dpi.nsw.gov.au

^a Presenting Author

CALCULATION OF GENETIC PARAMETERS IN A CUSTARD APPLE BREEDING PROGRAM IN SOUTH-EAST QUEENSLAND

Pat Pepper^a, Janet Giles, Michael Macbeth and Alan George

Department of Primary Industries & Fisheries, Brisbane, QLD, Australia.

As part of the custard apple breeding program at Maroochy Research Station, south-east Queensland, various combinations of 9 maternal and 18 paternal parents were crossed to produce 260 trees representing 49 crosses. There were between 1 and 32 trees of each cross. These trees were evaluated at three years of age by assessing usually 6 fruit per tree for fruit size and quality characteristics. Fruit data from parent trees were recorded in the same year, at which time parents were six years old.

Various models were used to estimate heritability, dominance, and maternal effects. These models included:

- (i) a maternal and paternal parent model with their non-genetic interaction,
- (ii) an individual tree model with non-genetic interaction of maternal and paternal effects and a non-genetic maternal effect,
- (iii) an individual tree model with non-genetic interaction of maternal and paternal effects with maternal effects absent from the model.

This last model recognised that the maternal component was close to zero for most traits, which was not surprising given the traits are far removed from maternal influence.

Data were analysed using the program ASREML. An age correction factor was included in the models to allow for the older age of parents at the time when data were recorded.

Because of the sparseness of the data, the estimates of the genetic parameters can only be indicative.

Email: pat.pepper@dpi.qld.gov.au

^a Presenting Author

FIELD ASSESSMENT OF FAECAL NIRS AS A MANAGEMENT TOOL FOR THE NORTHERN CATTLE INDUSTRY

David Reid^a and Desiree Jackson

Department of Primary Industries & Fisheries, Rockhampton, QLD, Australia.

Historically, producers have assessed animal condition to make decisions on supplementary feeding, paddock movements and selling of stock. The application of near infra-red spectroscopy (NIRS) technology enables producers to make an assessment of diet quality and thus make more timely and informed management decisions. The effective use of faecal NIRS requires producers to monitor pasture and animal condition on an ongoing basis.

An MLA-funded project assisted producers across Queensland to objectively record, preferably monthly, pasture and animal condition and submit faecal samples for NIRS prediction of pasture parameters such as percent crude protein (CP) and percent dry matter digestibility (DMD). Each paddock sampled was classified into a broad land type based on soil type (light/heavy) and pasture community.

To investigate the performance of the NIRS technology, box plots were initially used to look at relationships between the NIRS predictions and pasture and animal performance parameters. The impact of land type on the accuracy of NIRS predictions was also investigated.

Producers' visual assessment of whether the animals were losing, gaining or holding weight related favourably to the overall NIRS prediction of weight gain, although the predictions were variable and tended to overestimate gains. However, one needs to bear in mind that the calibration equation was based on a medium frame steer, weighing on average 300 kg; is confounded with land type, lactation status, and presence of browse; and is really an instantaneous prediction at the time of sampling. Predictions of pasture DMD and CP were also reasonably consistent with the producer's observation of animal performance, with DMD and CP decreasing as observed performance decreased.

The field data indicated the NIRS predictions were quite reasonable despite considerable variability. This variability may be due to many factors including land type, class of cattle, lactation status of cattle, type of pasture or browse grazed, bias in visual observation of animals, the lag between degrading pastures and the instantaneous NIRS prediction and the delayed effect of a rainfall event. These factors are being considered.

Email: David.Reid@dpi.qld.gov.au

^a Presenting Author

USE OF POISSON VARIATION FOR QUALITY ASSURANCE OF LABORATORY WORM EGG COUNTS

Andrew van Burgel^a

Department of Agriculture & Food, Albany, WA, Australia.

A new method is proposed for analysing the results of an annual quality assurance exercise done by parasitology laboratories. The exercise involves doing routine worm egg counts on multiple samples of sheep faeces. The samples sent to each lab are taken from bulk mixtures prepared at different densities of worm eggs, but the source of each sample is not revealed. Confidence intervals for the sample mean and variance ratio are calculated based on the variation of the Poisson distribution to identify outlying results that may indicate laboratories or operators with problems in their counting technique.

Email: avanburgel@agric.wa.gov.au

^a Presenting Author

Index of Presenting Authors

- Baker, P, 23
 Butler, K, 24
 Butler, R, 25
- Clarke, P. G., 26
 Cullis, B. R., 12, 13
- D'Antuono, M, 46
 De Faveri, J, 47
 Diffey, S, 27
 Doogan, V, 28
- Hedderley, D, 48
- Johnstone, P, 29
- Kiiveri, H, 14
- Kravchuk, O, 49
- Lambkin, K, 50
 Littlejohn, R, 30
- Mathews, K, 31
 Mayer, D, 32
 Mitchell, D, 15
 Murray, D, 33
- Nelder, J. A., 16
- Oakey, H, 34
 Orchard, B, 51
- Partington, D, 35
 Payne, R, 17, 36
 Pepper, P, 52
 Playford, C, 37
- Reid, A, 38
 Reid, D, 53
- Saunders, I, 20
 Saville, D, 39
 Smith, A, 40
 Stefanova, K, 41
 Stern, R, 18, 19
- Triggs, C, 42
- van Burgel, A, 54
 Verbyla, A, 43
 Visscher, P. M., 21
- Welham, S, 44
 Wenzl, P, 22
- Yoon, H, 45

Delegate List

Full Name		Email
Stewart	Andrews	stewart.andrews@vsni.co.uk
David	Baird	david.baird@agresearch.co.nz
Peter	Baker	peter.baker@csiro.au
Mary	Barnes	mary.barnes@csiro.au
Carey	Biggs	carey.biggs@vsni.co.uk
Gabriela	Borgognone	gabriela.borgognone@dpi.qld.gov.au
Dan	Burrell	daniel.burrell@dpi.qld.gov.au
Ruth	Butler	butlerR@crop.cri.nz
David	Butler	david.butler@dpi.qld.gov.au
Kym	Butler	kym.butler@dpi.vic.gov.au
Leigh	Callinan	leigh.callinan@dpi.vic.gov.au
Subhash	Chandra	subhash.chandra@dpi.vic.gov.au
Peter	Clarke	pclarke@agric.wa.gov.au
Damien	Collins	dc999@uow.edu.au
Ray	Correll	rho.environmentics@bigpond.com
Brian	Cullis	brian.cullis@dpi.nsw.gov.au
Mario	D'Antuono	mdantuono@agric.wa.gov.au
Joanne	DeFaveri	joanne.defaveri@dpi.qld.gov.au
Mick	Deland	deland.mick@saugov.sa.gov.au
Simon	Diffey	simon.diffey@dpi.nsw.gov.au
David	Donald	ddonald@bses.org.au
Vivienne	Doogan	vivienne.doogan@dpi.qld.gov.au
Paul	Eckermann	paul.eckermann@adelaide.edu.au
Robert	Forrester	bob.forrester@csiro.au
Peter	Geelan-Small	petergs@student.usyd.edu.au
Janet	Giles	janet.giles@dpi.qld.gov.au
Arthur	Gilmour	arthur.gilmour@dpi.nsw.gov.au
Trevor	Hancock	trevor.hancock@adelaide.edu.au
Murray	Hannah	murray.hannah@dpi.vic.gov.au
Simon	Harding	simon.harding@vsni.co.uk
Anne	Harris	anne.harris@dpi.nsw.gov.au
Kathy	Haskard	kathy.haskard@adelaide.edu.au
Duncan	Hedderley	hedderleyd@crop.cri.nz
Graham	Hepworth	hepworth@unimelb.edu.au
Veronica	Ingham	veronica.ingham@agriseach.com.au
Peter	Johnstone	peter.johnstone@agresearch.co.nz
Tim	Jones	timhjones@csiro.au
Ruth	Jordon-Som	rjordansom@nbpol.com.pg
Chandrapal	Kailasanathan	-
Harri	Kiiveri	harri.kiiveri@csiro.au
John	Koolaard	john.koolaard@agresearch.co.nz
Rosemary	Kopittke	rosemary.kopittke@dpi.qld.gov.au
Olena	Kravchuk	o.kravchuk@uq.edu.au
Kathryn	Lambkin	kathryn.lambkin@csiro.au
Alison	Lister	listera@lincoln.ac.nz
Roger	Littlejohn	roger.littlejohn@agresearch.co.nz
Ky	Mathews	ky.mathews@csiro.au
David	Mayer	david.mayer@dpi.qld.gov.au
Andrew	McLachlan	mclachlanA@crop.cri.nz
David	Mitchell	david.mitchell@csiro.au
Warren	Müller	warren.muller@csiro.au
Darren	Murray	darren.murray@vsni.co.uk
John	Nelder	j.nelder@imperial.ac.uk
Helen	Nicol	helennicol@iinet.net.au
Sharon	Nielsen	sharon.nielsen@dpi.nsw.gov.au

Sorn	Norng	sorn.norng@dpi.vic.gov.au
Helena	Oakey	helena.oakey@adelaide.edu.au
Maryann	O'Donnell	maryann@stats.net.au
Beverley	Orchard	beverley.orchard@dpi.nsw.gov.au
Debra	Partington	debra.partington@dpi.vic.gov.au
Roger	Payne	roger.payne@vsni.co.uk
Pat	Pepper	pat.pepper@dpi.qld.gov.au
Wayne	Pitchford	wayne.pitchford@adelaide.edu.au
Christina	Playford	christina.playford@dpi.qld.gov.au
Angela	Reid	angela.reid@dpi.qld.gov.au
David	Reid	david.reid@dpi.qld.gov.au
Ian	Saunders	ian.saunders@csiro.au
Dave	Saville	dave.saville@agresearch.co.nz
Jenny	Smith	jennifer.smith@dpi.vic.gov.au
Alison	Smith	alison.smith@dpi.nsw.gov.au
Lorraine	Spohr	lorraine.spohr@dpi.nsw.gov.au
Katia	Stefanova	kstefanova@agric.wa.gov.au
Roger	Stern	r.d.stern@reading.ac.uk
Joanne	Stringer	jstringer@bses.org.au
Julian	Taylor	julian.taylor@adelaide.edu.au
Robin	Thompson	robin.thompson@bbsrc.ac.uk
Fiona	Thomson	fiona.thomson@dpi.vic.gov.au
Thao	Tran	tratt006@students.unisa.edu.au
Chris	Triggs	triggs@stat.auckland.ac.nz
Andrew	Van Burgel	avanburgel@agric.wa.gov.au
Ari	Verbyla	ari.verbyla@adelaide.edu.au
Peter	Visscher	peter.visscher@qimr.edu.au
Andrew	Wallace	wallacea@crop.cri.nz
Sue	Welham	sue.welham@bbsrc.ac.uk
Peter	Wenzl	peter@diversityarrays.com
Hwan-Jin	Yoon	hwan-jin.yoon@dpi.vic.gov.au