

ESTIMATING GENETIC VARIATION AND COVARIATION FOR
FLYSTRIKE INCIDENCE IN AUSTRALIAN MERINO SHEEP
ON AN UNDERLYING NORMAL SCALE.

A. R. GILMOUR* and H. W. RAADSMAN**, AUSTRALIA

SUMMARY

We compare estimates of heritability, genetic and phenotypic correlation obtained using a generalized linear mixed model with those obtained by Henderson's method 3 and by restricted maximum likelihood for a binomial and normal trait.

INTRODUCTION

Flystrike is a serious problem to Australian wool producers. It occurs in hot humid weather and without prevention and remedial action by the grazier, causes significant production loss and death. The extent of the problem varies from year to year depending on the extent of summer rains. Considerable research has been directed at investigating the underlying genetic components of this problem and this paper reports an analysis of some of this data. The purpose of the paper is to review the methods we are using, rather than to report the results in detail. The technical details of the newer methods are presented elsewhere and are described here in conceptual terms only.

DESCRIPTION OF THE DATA

The data were collected from ewe yearlings representing 15 Merino genotypes maintained at Trangie Agricultural Research Centre. The flock was set up in 1975 and described by McGuirk *et al*, 1978. Young rams from the original sources are mated to ewes of the same genotype each year, in single sire mating groups. Sires are used in only one year. The sheep were classified into three birth/rearing classes: single/single, multiple/single and multiple/multiple.

In 1983, 310 ewe yearlings were subjected to a wetting regime in three groups, to induce fleece rot which predisposes to flystrike, and subsequently assessed for flystrike. Flystrike occurred naturally in 1984 and 315 ewe yearlings were assessed. A total of 87 sires were represented, nested within season and flocks. Many other traits were recorded including a wool assessment using a Fibre Fineness Distribution Analyser which gives a mean and standard deviation of fibre diameter based on 1000 individual fibres.

For the purposes of this paper, we estimated the heritabilities of flystrike and the standard deviation of fibre diameter (sdFD) and their genetic correlation. Flystrike was regarded as normally distributed for the incidence analysis and binomially distributed for the analysis of the underlying liability to flystrike incidence which was assumed normally distributed.

* Agricultural Research and Veterinary Centre,
Forest Rd, Orange, 2800, AUSTRALIA

** Agricultural Research Centre, Trangie, 2823, AUSTRALIA.

STATISTICAL METHODS AND RESULTS

a) Analyses involving flystrike incidence

Ordinary least squares was used to identify a suitable linear model. It was concluded that season, flock and wetting groups were significant sources of variation in both traits. Birth/rearing class and the flock by season interaction were significant for sdFD. The significance tests for flock, season and their interaction were based on the sires within flock by season line of the analysis of variance table. The model including all these factors was used for all results reported here.

The estimates of the sire and residual variance and covariance components shown below were obtained by Henderson's method 3 (Searle, 1971) and by restricted maximum likelihood (Harville, 1977).

Component	degrees of freedom	Variance		Covariance
		flystrike	sdFD	flystrike, sdFD
<u>Henderson's method 3</u>				
Sires	57	0.0145	0.0580	0.0083
Error	534	0.1430	0.3432	0.0180
<u>Restricted maximum likelihood</u>				
Sires	87	0.0141	0.0583	0.0084
Error	591	0.1434	0.3438	0.0179

These procedures give essentially the same results because the data is not severely unbalanced. Progeny per sire ranged from 3 to 12; 65 sires had from 6 to 9 progeny; the average was about 7.

The estimates of genetic parameters derived from these are

heritability flystrike	0.369 ± 0.163
heritability sdFD	0.579 ± 0.185
genetic correlation	0.285 ± 0.267
phenotypic correlation	0.105 ± 0.108

b) Analysis involving liability to flystrike

The method used here is the generalized linear mixed model described by Gilmour, Anderson and Rae (1985) and extensions of that method to multiple traits. The principle in this method is to assume an underlying normal liability to flystrike. The linear mixed model is assumed to relate to this liability which is predicted from the data by the expression

$$y = \bar{y} + (r-p)/z$$

where y is the predicted liability, the variable that is analysed,
 \bar{y} is the mean liability predicted by $x'b$,
 $p = \Phi(\bar{y})$ is the mean incidence, a function of \bar{y} ,
 r is the observed incidence,

$z = \phi(\bar{y})$ is a scale conversion factor, $dp/d\bar{y}$,
 x is the design vector for the observation and
 b is the vector of fixed effects.

The parameter we chose to use is the intraclass correlation,
 $t = \text{sire variance}/(\text{sire variance} + \text{error variance})$. The error variance
is 1. for liability on the standard normal scale.

On the incidence scale, the variance of the observations, $\text{var}(r)$,
is $p(1-p)$. In our mixed model, this is partly due to random sire
effects. A widely used result is that the intraclass correlation on the
incidence scale, t_r , is $tz^2/\{p(1-p)\}$. This implies that the part due to
sires is tz^2 and the residual is $p(1-p) - tz^2$. The mixed value
equations are then

$$\begin{bmatrix} X'WX & X'WZ \\ Z'WX & Z'WZ + G^{-1} \end{bmatrix} \begin{bmatrix} b \\ u \end{bmatrix} = \begin{bmatrix} X'Wy \\ Z'Wy \end{bmatrix}$$

where $W = DR^{-1}D$

D is $\text{diag}\{z_i\}$

R is $\text{diag}\{p_i(1-p_i) - z_i^2t\}$

G is $\text{diag}\{t\}$

X is the design matrix for fixed effects

Z is the design matrix for random effects.

These equations exactly parallel the normal case with R_n being
 $\text{diag}(1-t_n)$, G_n being $\text{diag}\{t_n\}$ and D_n being an identity. The extension
to multiple traits is then obvious. The two sets of equations are
combined and appropriate covariance terms added. The weight matrix for
the i^{th} data record is

$$\begin{bmatrix} z_i & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p_i(1-p_i) - z_i^2t & \text{cov} \\ \text{cov} & 1-t_n \end{bmatrix}^{-1} \begin{bmatrix} z_i & 0 \\ 0 & 1 \end{bmatrix}$$

where the error correlation is $c = \text{cov}/\sqrt{\{p_i(1-p_i) - z_i^2t\}\{1-t_n\}}$

Similarly, the diagonal block of G_m is

$$\begin{bmatrix} t & a \\ a & t_n \end{bmatrix} \text{ where } a = \text{genetic correlation } \sqrt{t t_n}.$$

This is a brief description of the mixed model equations. How
these are used to produce REML estimates of variance components is
discussed by Harville (1977) and Searle (1979).

The procedure was applied to the flystrike data and the estimate
of t for liability to flystrike was 0.194. A problem was encountered in
this analysis because three of the 30 season by flock classes had no
flystrike present. This implied a liability of $-\infty$ causing numerical
problems. This was overcome by changing the incidence of one sheep from
each class from zero to 1/4. Arbitrarily, the last record in each group
was changed.

For the estimation of the genetic correlation from the multiple
trait equations, the previously obtained results were taken as starting
values. The error and sire variances for sdFD were taken from the REML
analysis as was the error correlation between the traits. Whether this

is the best error correlation to use has not been considered yet. The genetic correlation was estimated as 0.2624 when t was held at 0.194. When t was allowed to move, it dropped to 0.190 and the genetic correlation increased to 0.265. The phenotypic correlation derived from these final values was 0.16.

DISCUSSION AND CONCLUSION

This paper demonstrates that the generalized linear mixed model can be used to obtain estimates of the heritability of liability for binomial traits when application of previously used procedures appears unsatisfactory. The estimate obtained is $4t = 0.777$ which is higher than that derived from the incidence analyses. Given the average incidence of flystrike over the two seasons was 0.228, the heritability of liability to flystrike obtained from the analysis of incidence is 0.59 (1.6x0.369). The higher value occurs because 26 of the 87 sire groups had zero incidence. While this is not uncommon with incidence data, it does raise doubts as to the reliability of any estimates obtained. The generalized linear mixed model would give more reliable results with data with either larger families or more families as discussed by Gilmour, Anderson and Rae (1985).

The two estimates of the genetic correlation are very similar. This agrees with the results reported by Mao (1976). The phenotypic correlations differ with the value involving the liability higher than that involving incidence. This agrees with the results of Mao (1976).

This work was supported by the Australian Wool Corporation.

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, pp 593-599.
- HARVILLE, D.A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *J. Am. Statist. Assoc.* 72, 320-340.
- MCGUIRK, B.J., ATKINS, K.D., KOWAL, Eva and THORNBERRY, K. (1978) Breeding for resistance to fleece rot and body strike - the Trangie programme. *Wool Technology and Sheep Breeding*. 26, pp 17-24.
- MAO, I.L. (1976). Correlation estimates for discontinuous data. *J. Dairy Sci.* 59, pp 2115-2121.
- SEARLE, S.R. (1971) *Linear models*. John Wiley & Sons. pp 443-451.
- SEARLE, S.R. (1971) *Notes on variance component estimation: A detailed account of maximum likelihood and kindred methodology*. Biometrics Unit, Warren Hall, Cornell University, Ithaca, New York. 14853.