ESTIMATION OF THE DEGREE OF CONNECTEDNESS BETWEEN TEST GROUPS OF STATION-TESTED BEEF BULLS

V.M. Roso1,2, F.S. Schenkel1 and S.P. Miller1

1CGIL, Department of Animal & Poultry Science, University of Guelph, Guelph, Ontario, Canada N1G 2W1
2GenSys Consultores Associados SC Ltda, Rua Guilherme Alves 170/304, Porto Alegre, Brazil, CEP 90680-000

INTRODUCTION

Connectedness among test groups (TG) is of interest in genetic evaluation of station-tested beef bulls, because comparisons of estimated breeding values (EBV) of bulls tested in different groups are made. The EBV of bulls from different TG are comparable due to use of appropriate methodology (BLUP) and genetic connectedness among groups. However, the accuracy of the comparisons depends upon the degree of connectedness among TG. With fewer genetic links between TG, comparison of bulls EBV in different TG is less accurate, even if the accuracy of EBV is high within the groups.

When genetic evaluation is under an animal model, connections occur through additive genetic relationships. Kennedy and Trus (1993) argued that the most appropriate measure of the connectedness is the average Prediction Error Variance (PEV) of differences in EBV between animals in different management units (e.g., TG), which are influenced by average genetic relationship between and within management units. However, computing this statistic is extremely time consuming and not feasible for routine application. Thus, Kennedy and Trus (1993) proposed to use the variance of estimated differences between management unit effects, which was highly correlated with the PEV of the differences between EBV in their simulation study. Mathur et al. (1999) also argued that the variance of estimated differences between management unit effects could be used as a measure of connectedness between two management units and proposed to calculate the connectedness rating (CR), defined as the correlation between estimated effects of two management units. Following Mathur et al. (1999), CR is less dependent on the size and structure of management units. For calculating CR, the authors used an iterative method, which allows them to obtain the inverse elements for some rows and columns (corresponding to TG in the MME, for example) of any large matrix for which a direct inverse is not possible. However, the computational effort to calculate CR is still substantial. Fries (1998) proposed the use of number of direct genetic links between TG (GLt) due to common sires and dams as a method for measuring degree of connectedness among TG. Obtaining GLt is computationally less demanding than the variance of estimated differences between management units and CR.

The objective of this study was to assess and compare the statistics CR and GLt for measuring the degree of connectedness among TG of station-tested beef bulls.
MATERIAL AND METHODS

Data were estimated weight gain (BEG) of 26,068 bulls tested in central evaluation stations in Ontario, Canada, from 1988 to 2000. The procedure used to obtain the BEG followed what is currently applied for genetic evaluation by Beef Improvement Ontario (BIO). Firstly, a fixed univariate linear regression of the weights \( (w_{ij}) \) (taken on average every 28 days \((j)\)) on the actual days on test \((d_{ij})\) for each bull \(i\) was estimated, using the model \( w_{ij} = a_i + \beta_i d_{ij} + e_{ij} \), where \(a_i\) and \(\beta_i\) are the intercept and linear regression coefficient of bull \(i\), respectively, and \(e_{ij}\) is the residual term. Secondly, the BEG was calculated multiplying \(\beta_i\) by the standard number of days on test (140 days). BEG was used as an observation in the follow genetic evaluation model:

\[
\text{BEG}_{iklm} = \sum_{n=1}^{14} b_n B_{in} + G_l + A_i + E_{iklm},
\]

where:

- \(b_n\) = linear regression coefficient of the \(n^{th}\) breed;
- \(B_{in}\) = breed composition of the \(i^{th}\) bull for the \(n^{th}\) breed;
- \(G_l\) = fixed effect of the \(l^{th}\) test group (\(l = 1\) to \(583\));
- \(A_i\) = random additive genetic effect of the \(i^{th}\) bull;
- \(E_{iklm}\) = random residual effect.

Random effects \(A\) and \(E\) were assumed independent with covariance matrices equal to \(A\sigma^2_A\), and \(I\sigma^2_E\), respectively. All available pedigree information was incorporated into the additive numerator relationship matrix \((A)\). PEV of EBV and estimated error variance of TG effects (EEV) were obtained using PEST (Groenevelt, 1990), assuming a heritability of 0.43, which was previously estimated for the same data set.

The degree of connectedness among TG was assessed using the following methods:

1) PEV of the difference between EBV of bulls from different TG (PEVd). The average PEVd of each TG was calculated and used as a measure of connectedness of the TG, following Kennedy and Trus (1993).

2) Connectedness rating, defined as the correlation between estimated effects of TG. The average CR of each TG was used as a measure of connectedness of the TG, following Mathur et al. (1999). CR between TG \(i\) and \(j\) was calculated by:

\[
\text{CR}_{ij} = \frac{\text{Cov}(\hat{G}_i, \hat{G}_j)}{\sqrt{\text{Var}(\hat{G}_i) \times \text{Var}(\hat{G}_j)}}
\]

3) Total number of direct genetic links between TG (GLt), defined as the links between TG due to common sires and dams. A small example of how to determine GLt is given bellow, where \(S\) stands for sire and \(D\) for dam. In this example, TG1 has 2 genetic links, via \(S1\) and \(1\) genetic link, via \(D1\), with TG2 and 2 genetic links, via \(S2\), with TG3, resulting in GLt equal to 5. TG2 has 2 genetic links, via \(S1\), and 1 genetic link, via \(D1\), with TG1, resulting in GLt equal to 3. TG3 has only two genetic links via \(S2\) with TG1, resulting in GLt equal to 2. TG4 has no genetic links with either of the TG, and is, therefore, disconnected, GLt equal to zero.
The PEVd was considered the most appropriate measure of connectedness and the alternative methods CR and GLt were compared with this statistic.

RESULTS AND DISCUSSION
The average of PEVd, CR, GLt, number of bulls per TG and number of sires per TG were 1583±41, 1.21±0.51, 702±690, 44±36 and 23±21, respectively. The total number of direct genetic links between TG was mainly due to sires (94.5%).

The observed relationship of PEVd with CR and GLt were curvilinear, as shown in Figure 1. By observation, small values of CR and GLt are associated with increasingly higher PEVd. The estimated linear correlation between CR and GLt was relatively high (0.86).

Figure 1. Observed relationship between PEVd and CR (left) and between PEVd and GLt (right) for groups of station-tested beef bulls

The observed average PEVd of each TG was modeled using two multiple regression models: (1) a quadratic-quadratic polynomial regression and a quadratic regression on number of sires and on the ratio bulls/sire. All effects were significant (P<0.01). The $R^2$ of the model was equal to 0.82 and CR accounted for 73% of total variation in PEVd. (2) a quadratic-quadratic polynomial regression on GLt, a linear regression on number of sires and a quadratic regression on the ratio bulls/sire. All effects were significant (P<0.01). The $R^2$ was 0.79 and GLt account for 76% of the total variation in PEVd. The estimated values of PEVd according
to CR and GLt showed that CR smaller than 1.0 and GLt smaller than 400 were associated with more rapidly increasing PEVd.

In the data set, on the basis of GLt, there was only one completely disconnected TG. Thus, to evaluate the effect of complete disconnectedness, 36 TG had sire and dam identifications modified to generate completely disconnected TG, covering a range of TG sizes from very small to large TG. The accuracy of bull’s EBV from disconnected TG would increase only with the size of the group, because there were no relationships among bulls within the created disconnected TG. Results indicate that increasing the size of disconnected groups reduces the PEV from 981 in a group with only 6 bulls to an asymptotical minimum value around 876, when 120 bulls are in the TG. Consequently, the average of PEVd between two disconnected TG with no relationship among bulls within the TG would be at least equal to 1752 (876+876). Relationship among bulls within a disconnected TG would increase the PEV of comparisons of EBV across TG. Therefore, connected TG with average PEVd greater or equal to 1752 would behave similarly to large disconnected TG of unrelated bulls with respect to PEVd.

Although disconnected TG had GLt equal to zero, the CR of those disconnected TG varied between 0.27 and 1.44, that is, a completely disconnected TG presented a large range of CR values. CR was highly correlated with the size of the TG (0.95). Therefore, on the basis of CR, it was not possible to differentiate completely disconnected TG from connected ones, because large disconnected TG had CR values similar to connected ones. EBV of bulls from completely disconnected TG should not be compared with EBV of bulls from other TG. The use of GLt allowed the identification of disconnected TG. GLt was computationally much less demanding than CR and can be easily routinely calculated. Thus, an assessment of the quality of the connectedness of a TG could potentially be obtained before beginning an evaluation by using the GLt, which is less dependent on the size of TG (correlation 0.81 vs 0.95) and it would not necessarily favor large TG, because relatively small TG could have large GLt and, consequently, low average PEVd.

**CONCLUSION**

CR is more dependent on the size of TG than GLt. The use of CR as a connectedness measure will cause larger TG to be favored. GLt, unlike CR, allows differentiation between completely disconnected TG from connected ones. GLt is computationally much less demanding than CR and can be routinely used with the aim of increasing the level of connectedness, obtaining more accurate comparison of EBV between TG.

**REFERENCES**