*Animal Science Papers and Reports* vol. 30 (2012) no. 4, 309-315
Institute of Genetics and Animal Breeding, Jastrzębiec, Poland

# Factors affecting genetic correlation estimates from dairy sires' genetic evaluations to assess genotype-environment interaction*

**Hugo H. Montaldo**[1,**], **Alejandra Pelcastre-Cruz**[2]

[1] Departamento de Genética y Bioestadística, Facultad de Medicina Veterinaria y Zootecnia,
Universidad Nacional Autónoma de México (UNAM),
Ciudad Universitaria, 04510, D.F., Mexico

[2] Departamento de Producción Agrícola y Animal, Universidad Autónoma Metropolitana,
Xochimilco, 04960, D.F., Mexico

*(Received March 26, 2012; accepted September 3, 2012)*

**Effects of trait heritability (0.05 or 0.25), effective daughters number (30 to 500), sires number (10 to 500), and sire selection (selecting or not the top 50% sires), were evaluated based upon standard error (SE) and bias of genetic correlations ($r_G$) between countries estimated from Calo's method ($r_G$) using simulated data. Calo's method is based on correlations between sire's predicted transmitting abilities (PTA) in two countries adjusted for reliabilities. Unselected sire's data analysis gave nearly unbiased $r_G$ in all cases, but selected sire's data analysis gave underestimates. Bias was from -0.34 to -0.05 for the 0.25 heritability trait (milk yield), and from -0.42 to -0.17 for the 0.05 heritability trait (functional). Underestimation of $r_G$ decreased with increased effective number of daughters (PTA's reliability), but was quite insensitive to number of sires. The SE of genetic correlations estimates decreased with increased PTA's reliability and sires number, and was higher for selected sires. Approximately 50 sires with PTA's reliabilities$\geq$0.97 on each country are required to obtain accurate (SE$\leq$0.02) and unbiased (bias$\leq$0.05|) $r_G$ with Calo's method using the best 50% selected sires. Many genetic correlation estimates between countries, already published using the Calo's method, may be underestimates, particularly for low heritability traits, and with low number of effective daughters in the importing country. Therefore, caution is required before interpreting the published $r_G$<1 as evidence for genotype-environment interaction.**

---

*Supported by the National Autonomous University of Mexico (UNAM) PAPIIT, grant IN 217509
**Corresponding author: montaldo@unam.mx

**KEY WORDS: bias /genetic correlation / genotype-environment interaction / international evaluation of bulls**

Genetic correlation [Robertson 1959] between environments has been suggested as a way to assess genotype-environment interaction (GEI). Genetic correlations <1 for a certain trait are interpreted as evidence of GEI [Falconer 1952, Montaldo 2001]. Mulder *et al.* [2006] suggested that genetic correlation less than 0.60 may indicate the need of developing independent breeding programs within each environment.

Accurate estimates of genetic correlations between countries for economically important traits in dairy cattle are required in order to develop efficient breeding programs worldwide [Mulder *et al.* 2006]. Estimates of genetic correlation ($r_G$) between countries are usually obtained using Multiple-Trait Across-Country Evaluations (MACE) methods which give nearly unbiased genetic correlation estimates with adequate data and pedigree information [Sigurdsson *et al.* 1996, Mark *et al.* 2005], but many published $r_G$ which involve countries lacking a solid dairy cattle breeding information infrastructure, are derived from correlations between predicted transmitting abilities (PTA) of sires in two countries [e.g. Vargas and Gamboa 2008, Montaldo *et al.* 2009], using the so called Calo's method [Calo *et al.* 1973]. Calo's method is based upon correlations between sire's predicted transmitting abilities (PTAs) in two environments, with corrections for reliabilities. Many of these estimates have been obtained using rather low effective numbers of daughters in the importing countries [e.g. Vargas and Gamboa 2008].

Underestimation of $r_G$ with Calo's method may be assumed due to sire selection in the exporting country based on normal theory [Cameron 1993], and it is known that the use of selected sire samples underestimate regression coefficients of importing on exporting breeding values, when these are based on PTA [Powell *et al.* 1994]. However, no studies have been published specifically to investigate how various factors may affect the statistic properties of $r_G$ as estimated using Calo's procedure.

The objective of this study was to evaluate the effects of heritability, accuracy of PTA, number of sires and sire selection on bias and accuracy of genetic correlations between countries estimated with the Calo's method, using simulated data.

## Material and methods

Simulated data were used. Genetic correlations were obtained from the observed correlation values between simulated sires' PTAs in two countries (i and j). PTA for one sire in countries i and j were obtained as: $PTA_i = \bar{y}_i b_i$ and $PTA_j = \bar{y}_j b_j$, where $\bar{y}_i$ and $\bar{y}_j$ are simulated averages of records for daughters of the sire from countries i and j using a sire model, with random and independent sire and error effects, sampled from independent normal standard distributions.

Appropriate standard deviations of sires ($\sigma'_s = (0.25h^2)^{0.5}$) and average error ($\sigma'_e = [(1-0.25h^2)/d]^{0.5}$), were used in the simulations, where $h^2$ is heritability and $d$ is the effective number of daughters. Weighting factors $b_i$ and $b_j$ were obtained as $d/(d+\alpha)$, where $\alpha$ is $(4-h^2)/h^2$ [Van Vleck 1993]. The effective number of daughters $d$ was randomly sampled from a Poisson distribution.

One thousand replicates were obtained for each combination formed by the number of sires (10, 50, 100 and 500), selection status of the sires (not selected or selected by keeping only the sires with PTA>0 values in the exporting country; approximately the top 50%), and numbers of effective daughters in the exporting and importing country (30 and 30, 500 and 30, and 500 and 500).

The situations considered are presented in Table 1. All options were evaluated for two traits according to the heritability value: milk yield with a heritability of 0.25 and a functional trait with a heritability of 0.05 (Tab. 1).

Genetic correlation between countries was estimated from the observed correlations among the PTA of each sire in two countries and their average reliabilities, according to the following formula [Calo *et al.* 1973]:

$$r_G = \frac{r_o}{\sqrt{arel_i \times arel_j}} \qquad (1)$$

where:

$r_G$ – estimated genetic correlation;

$r_o$ – estimated correlation among PTA;

$arel_i$ – average reliability of PTA from country i;

$arel_j$ – average reliability of PTA from country j.

Reliability is defined as the estimated $r^2_{T,PTA}$, where T is the true transmitting ability of the sire. The denominator of [1] is the expected correlation between evaluations if $r_G = 1$. The bias was estimated as $r_G$-1 and the accuracy was assumed to be inversely proportional to the standard error (SE) which was estimated as the standard deviation of the simulated $r_G$ values.

## Results and discussion

Results of the statistics from the simulated $r_G$ for a milk yield trait with heritability of 0.25, and for a functional trait with heritability of 0.05, are shown in Table 1. Calo's method, in absence of sire selection, gave nearly unbiased $r_G$ estimates (Tab. 1), which is not surprising, because the method is based on expected values [Calo *et al.* 1973] so it would be unbiased with unselected (random) samples.

Analysis of data with selected sires gave $r_G$ underestimated. Bias was from -0.34 to -0.05 for the milk yield and from -0.42 to -0.17 for the functional trait (Tab. 1). Underestimation of $r_G$ was lower with increased effective daughter number (PTA's reliability), but was quite unresponsive to number of sires (Tab 1). Larger

**Table 1.** Statistics for genetic correlations ($r_G$), estimated with the method of Calo from simulated genetic evaluations of dairy sires in two countries, for traits with a heritability of 0.25 or 0.05[1]

| Item | | | | Selection of best 50% | | | No selection | | |
|---|---|---|---|---|---|---|---|---|---|
| herita-bility | number of sires | $d_{exp}$[1] | $d_{imp}$[2] | $r_g$ | SE | bias | $r_G$ | SE | bias |
| 0.25 | 10 | 30 | 30 | 0.66 | 0.46 | -0.34 | 0.97 | 0.31 | -0.03 |
| 0.25 | 10 | 500 | 30 | 0.74 | 0.32 | -0.26 | 0.98 | 0.19 | -0.02 |
| 0.25 | 10 | 500 | 500 | 0.94 | 0.07 | -0.06 | 1.00 | 0.03 | 0.00 |
| 0.25 | 50 | 30 | 30 | 0.71 | 0.17 | -0.29 | 1.00 | 0.12 | 0.00 |
| 0.25 | 50 | 500 | 30 | 0.77 | 0.12 | -0.23 | 0.99 | 0.07 | -0.01 |
| 0.25 | 50 | 500 | 500 | 0.95 | 0.02 | -0.05 | 1.00 | 0.01 | 0.00 |
| 0.25 | 100 | 30 | 30 | 0.71 | 0.12 | -0.29 | 1.00 | 0.09 | 0.00 |
| 0.25 | 100 | 500 | 30 | 0.78 | 0.08 | -0.22 | 1.00 | 0.04 | 0.00 |
| 0.25 | 100 | 500 | 500 | 0.95 | 0.02 | -0.05 | 1.00 | 0.01 | 0.00 |
| 0.25 | 500 | 30 | 30 | 0.71 | 0.05 | -0.29 | 1.00 | 0.04 | 0.00 |
| 0.25 | 500 | 500 | 30 | 0.75 | 0.05 | -0.25 | 1.00 | 0.02 | 0.00 |
| 0.25 | 500 | 500 | 500 | 0.95 | 0.01 | -0.05 | 1.00 | 0.00 | 0.00 |
| 0.05 | 10 | 30 | 30 | 0.58 | 1.27 | -0.42 | 1.01 | 1.11 | 0.01 |
| 0.05 | 10 | 500 | 30 | 0.61 | 0.65 | -0.39 | 0.93 | 0.56 | -0.07 |
| 0.05 | 10 | 500 | 500 | 0.79 | 0.24 | -0.21 | 0.98 | 0.12 | -0.02 |
| 0.05 | 50 | 30 | 30 | 0.60 | 0.49 | -0.40 | 0.99 | 0.48 | -0.01 |
| 0.05 | 50 | 500 | 30 | 0.65 | 0.27 | -0.35 | 0.99 | 0.23 | -0.01 |
| 0.05 | 50 | 500 | 500 | 0.83 | 0.09 | -0.17 | 1.00 | 0.04 | 0.00 |
| 0.05 | 100 | 30 | 30 | 0.62 | 0.37 | -0.38 | 0.99 | 0.33 | -0.01 |
| 0.05 | 100 | 500 | 30 | 0.65 | 0.19 | -0.35 | 1.00 | 0.16 | 0.00 |
| 0.05 | 100 | 500 | 500 | 0.83 | 0.06 | -0.17 | 1.00 | 0.03 | 0.00 |
| 0.05 | 500 | 30 | 30 | 0.62 | 0.16 | -0.38 | 1.00 | 0.15 | 0.00 |
| 0.05 | 500 | 500 | 30 | 0.66 | 0.08 | -0.34 | 1.00 | 0.07 | 0.00 |
| 0.05 | 500 | 500 | 500 | 0.83 | 0.03 | -0.17 | 1.00 | 0.01 | 0.00 |

[1] based on 1000 replicates.

[2] $d_{exp}$= daughters' effective number for exporting country; $d_{exp}$= daughters' effective number for importing country. These daughter effective numbers are equivalent to a reliability of 0.67 for 30 effective daughters; 0.97 for 500 effective daughters and 0.99 with 1000 effective daughters for $h^2$=0.25; and to a reliability of 0.28 for 30 effective daughters, 0.86 for 500 effective daughters and 0.93 with 1000 effective daughters for $h^2$=0.05.

underestimation with lower reliability may explain the lower estimates for genetic correlations obtained with Calo's method for traits with lower compared to that for higher heritability traits [Powell *et al.* 1997, Rogers *et al.* 1998].

The SE of genetic correlations estimates decreased with increased effective number of daughters (PTA's reliability) and sires' number, and was higher for selected sires (Tab. 1). Additional simulations for both, selected and not selected sires, with the same reliabilities from each country, proved that bias and SE were exactly the same, even with different heritability values (figures not shown).

A situation with a low number of daughters in the importing country, which is the most common situation in dairy breeding [e.g. Vargas and Gamboa 2008], would increase considerably both the SE and the underestimation of $r_G$ (Tab. 1). Using sires with high reliability would give less underestimated $r_G$ values compared to those obtained from sires with low reliabilities using the Calo's method (Tab. 1). These results may at least partially explain the lower estimates for $r_G$ obtained with this method for traits of lower heritabilities (longevity, somatic cell counts, and clinical *mastitis*) compared to those for traits with higher heritabilities such as milk yield [Mrode and Swanson 1997, Powell *et al.* 1997, Rogers *et al.* 1998, Grignola and Schaeffer 2000].

Sire selection intensity used in this study for yield traits (0.8) is slightly lower than the average value (1.0) reported for yield traits in several countries for Holstein-Friesian sires [Powell *et al.* 2003], but within the observed range of 0.6 to 1.5 for these observed values. However, for functional traits, selection intensities achieved in dairy sires selection have been from slightly negative to 0.4 [Powell *et al.* 2003]; that is, lower than the value used in this study. Nevertheless, the importance of functional traits as selection criteria in dairy cattle is increasing worldwide [VanRaden 2004].

Montaldo *et al.* [2009] estimated $r_G$ for milk yield between Canada and Mexico at 0.77 with the method of Calo using 40 common sires with average reliabilities of 0.95 and 0.80, respectively. This $r_G$ estimate was compared to simulated $r_G$ obtained with the procedures used in this study, and the best of 50% selected sires, which was 0.83±0.10. These two values were similar, because 95% distribution of simulated data was between 0.66 and 0.96. Simulated $r_G$ with similar conditions, but with the best 25% of the sires selected gave $r_G$ of 0.74±0.12, and 95% of the distribution between 0.53 and 0.91. These results and another simulation results made for some of the conditions considered in Table 1 but with higher selection intensity (results not shown), illustrate further that a more intensive selection will increase underestimation of the genetic correlation. With small number of sires, the proportion selected for yield traits in dairy may be lower than 50%. In interpreting actual $r_G$ estimates with Calo's method, paternity identification errors may be an additional source of underestimation as found by Montaldo *et al.* [2009].

These results indicate that caution is required about the presence of GEI effects using this method, because the genetic correlations would be underestimated in many real life situations. The single most important controllable factor to attaining accurate and unbiased estimates of $r_G$, when the original sample of sires is selected, is to use only data from sires with high PTA's reliabilities in both countries.

The simulation model used in this study is simple and based on independent sire and error effects. Therefore, it may not consider all the factors that could influence $r_G$ between countries. However, the method used in this study may still be useful to identify the main effects that affect the bias and accuracy of $r_G$ estimates between countries using selected sire genetic evaluations. In MACE procedure all available daughters' data within each country with sire models are used and include the numerator relationship matrix between the sires, giving nearly unbiased estimates of the genetic correlations between countries with adequate data and pedigree information [Schaeffer 1985, Sigurdsson *et al.* 1996]. The inclusion of all data and genetic relationships through the numerator relationship matrix is required for unbiased estimates of genetic correlations between environments using the MACE approach [Mark *et al.* 2005]. However, bulls of interest in many countries are not always well connected between countries by pedigree links to allow accurate estimations and data information may be patchy. It is likely that while developed countries may have adequate data sets to perform these calculations, in some situations, Calo's method would be the only short-term practical option available for estimating $r_G$. One possibility to explore further is whether MACE estimates of $r_G$ or estimates obtained with complete data information analysis from each country [e.g. Ojango and Pollot 2002], may give underestimates of $r_G$ if pedigree information is shallow and data from other non-selected sires are incomplete.

Underestimation of $r_G$ coefficients due to sire selection is expected when this method is used, unless the reliabilities of the sire genetic evaluations are very high in both countries.

These results may be used both for interpreting correctly estimates of genetic correlations obtained from PTA in different environments and to design appropriate studies for assessing GEI using genetic evaluations form sires in two countries. It seems safer to interpret an $r_G$<0.60 rather than $r_G$<1.00, as evidence of GEI (Tab. 1), but SE will be reasonably low, only if a minimum number of sires and PTA's reliability is used, particularly for low heritability traits.

***Acknowledgement.*** *The MSc studies of A. Pelcastre-Cruz were founded by the National Council of Science and Technology of Mexico (CONACyT) with the scholarship number 14476.*

**References**

1. CALO L.L., MCDOWELL R.E., VANVLECK L.D., MILLER P.D.,1973 – Genetic aspects of beef production among Holstein-Friesians pedigree selected for milk production. ***Journal of Animal Science*** 37, 676-682.
2. CAMERON N.D., 1993 – Methodologies for estimation of genotype with environment interaction. ***Livestock Production Science*** 35, 237-249.
3. FALCONER D.S., 1952 – The problem of environment and selection. ***The American Naturalist*** 830, 293-298.
4. GRIGNOLA F., SCHAEFFER L.R., 2000 – Relationships between evaluations of Canadian and USA Holstein bulls for longevity and somatic cell score. ***Livestock Production Science*** 65, 161-165.
5. MARK T., MADSEN P., JENSEN J., FIKSE W.F., 2005 – Short communication: Difficulties in estimating across-country genetic correlations for weakly linked bull populations. ***Journal of Dairy Science*** 88, 3303-3305.
6. MONTALDO H.H., 2001 – Genotype-environment interactions in livestock breeding programs: a review. ***Interciencia*** 26, 229-235.

7. MONTALDO H.H., NÚŃEZ-SOTO S.G., RUÍZ-LÓPEZ F.J., CASTILLO-JUÁREZ H., 2009 – Selection response for milk production in conventional production systems in Mexico, using genetic evaluations of Holstein sires from Canada and the United States. ***Journal of Dairy Science*** 92, 5270-5275.
8. MRODE, R. A., SWANSON, G. J. T., 1997 – Association between somatic cell count progeny tests in the United States of America (USA), The Netherlands and the United Kingdom (UK). ***Proceedings of the British Society of Animal Science*** 1997, 15 (Abstract).
9. MULDER H.A., VEERKAMP R.F., DUCRO B.J., VAN ARENDONK J.A.M., BIJMA P., 2006 – Optimization of dairy cattle breeding programs for different environments with genotype by environment interaction. ***Journal of Dairy Science*** 89, 1740-175.
10. OJANGO J.M.K., POLLOTT G.E., 2002 –The relationship between Holstein bull breeding values for milk yield derived in both the UK and Kenya. ***Livestock Production Science*** 74, 1-12.
11. POWELL R.L., WIGGANS G.R., VANRADEN P. M., 1994 – Factors affecting calculation and use of conversion equations for genetic merit of dairy bulls. ***Journal of Dairy Science*** 77, 2679-2686.
12. POWELL R.L., VANRADEN P.M., WIGGANS G.R., 1997 – Relationship between United States and Canadian genetic evaluations of longevity and somatic cell score. ***Journal of Dairy Science*** 80, 1807-1812.
13. POWELL R.L., NORMAN H.D., SANDERS A.H., 2003 – Progeny testing and selection intensity for Holstein bulls in different countries. ***Journal of Dairy Science*** 86, 3386-3393.
14. ROBERTSON A. 1959 – The sampling variance of the genetic correlation coefficient. ***Biometrics*** 15, 469-485.
15. ROGERS G.W., BANOS G., NIELSEN U.S., PHILIPSSON J., 1998 – Genetic correlations among somatic cell scores, productive life, and type traits from the United States and udder health measures from Denmark and Sweden. ***Journal of Dairy Science*** 81, 1445-1453.
16. SCHAEFFER L.R., 1985 – Model for international evaluation of dairy sires. ***Livestock Production Science*** 12, 105-115.
17. SIGURDSSON A., BANOS G., PHILIPSSON J., 1996 – Estimation of genetic (co)variance components for international evaluation of dairy bulls. ***Acta Agriculturae Scandinavica,*** Section A, Animal Sciences 46, 129-136.
18. VANRADEN, P. M., 2004 – Invited review: Selection on net merit to improve lifetime profit. ***Journal of Dairy Science*** 87, 3125-3131.
19. VAN VLECK, L.D., 1993 – Selection index and introduction to mixed model methods. CRC Press, Boca Raton, CA.
20. VARGAS B., GAMBO A.G., 2008 – Genetic trends, genotype-environment interaction and inbreeding in Holstein and Jersey dairy cattle from Costa Rica. ***Técnica Pecuaria en México*** 46, 371-386.