Strategies to Combine Novel Traits across Countries: Example of Heat Stress

H.Hammami[1,2], J. Vandenplas[1,2], M.J. Carabaño[3], B. Logar[4], J. Bormann[5], C. Bertozzi[6] and N. Gengler[1]

[1]University of Liege, Gembloux Agro-Bio Tech, 5030 Gembloux, Belgium
[2]National Fund for Scientific Research, 1000 Brussels, Belgium
[3]Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, 28040 Madrid, Spain
[4]Agricultural Institute of Slovenia, 1000 Ljubljana, Slovenia
[5]Administration of Agricultural Technical Services, 1019 Luxembourg, Luxembourg
[6]Walloon Breeding Association, 5590 Ciney, Belgium

Abstract

Nowadays, novel traits are of great interest. However, phenotypes are siloed and mainly not shared. Heat stress is becoming problematic affecting animals' performances and their well-being. Heat stress tolerance as a novel trait is only addressed by isolated within-country research studies. Integration and combination of local and foreign information sources is needed for better accuracy genetic evaluations. Therefore, this study was aimed to test the potential combination of sources of external information towards the evaluation of heat stress tolerance of dairy cattle. Long-term cow performances linked to environmental descriptors (weather parameters as proxy to climate change) collected over 10 years under the temperate conditions of the Walloon Region of Belgium and the hotter and warm Mediterranean conditions of Andalusia and Castile-La-Mancha Spanish regions were available. A total of 1,604,775 milk, fat, and protein test-day (TD) records linked to average daily temperature humidity (THI) values for 3-day lag before each TD were considered. Under a first strategy considering free-access to raw-data (phenotype and pedigree), a joint evaluation was firstly run using reaction norm models where production traits were considered as function of THI. A Belgian and a Spanish evaluation were also run using the same model. An alternative strategy considering only access to external information (i.e. regression coefficients for additive genetic effects (â and their associated REL)) was tested. In this case, foreign â and their REL resulting from the Spanish evaluation were first converted to the Belgian trait and thereafter integrated in the Belgian evaluation using a Bayesian approach. Rank correlations between regression coefficients, â (of the 1,104 bulls having daughters only in Spain) estimated by Belgian evaluation and â estimated by the joint evaluation were moderate (<=0.70). Corresponding rank correlations between â estimated by joint and Bayesian evaluations were significantly higher (ranging from 0.967 to 0.998), indicating that the Bayesian evaluation integrating external information was in good concordance with the joint evaluation. Results from this study indicated that the integration of external information via the Bayesian approach has a good potential to improve the genetic evaluation of sparse and siloed novel traits.

Introduction

Today, national dairy cattle genetic evaluations concerned no more than 30 traditional traits related mostly to production, fertility and some functional traits. National estimated breeding value (EBV) for dairy breed sires are exchanged and combined through the Multi-trait across-country evaluations (MACE) performed by INTERBULL. Results for dairy breed sires for those traits are routinely communicated. This allows a better appreciation of the reliabilities (REL) of sires' EBV and permits greater genetic response when larger numbers of bulls are tested.

Dairy cattle breeders are affected by the global need for more sustainable, profitable, healthy and friendly-environmental food production. Additionally, the direct and

indirect risks of climate change impacting animals' performances and their well-being will add more pressure on the sustainability of the sector. Therefore, development of novel traits are needed to provide improved genetic and management tools.

Recent technologies and data capturing enable the collection of novel phenotypes. However, those novels traits are mainly collected on national small populations using local research projects and are not often genetically evaluated. Moreover, national datasets are siloed and not shared, affecting genetic evaluations accuracies and limiting the genetic gain. Under those conditions, genetic evaluations of exporting countries may be not fruitful to importing countries because of differences in selection criteria, genetic levels and ignorance of genotype by environment interactions.

Joint evaluations using combined raw data (phenotypes and pedigree) from exporting and importing countries are ideal when free access to all the information is guaranteed. However, this is not the case especially for novel traits due to their scarcity and also limitation of exchange, even when they are available. One alternative option could be based on the use of external information, i.e. EBV and associated REL provided from foreign evaluations, through its integration into a local genetic evaluation. Bayesian approaches where external information is used as priors for the local evaluation was found useful and promoted for genetic and genomic evaluations (e.g., Vandenplas *et al.*, 2013; Colinet *et al.*, 2013).

Climate change is an increasing concern worldwide. Heat tolerance as a novel trait linked to animal resilience and robustness under high weather and harsh conditions is only approached by research actions based on country-specific populations and production systems. Use of exchangeable raw or/and external information between countries to evaluate genetically cows' heat tolerance capacity across different environments, production systems and countries have to be challenged. Therefore, the main objective of this study was to ascertain ranking of bulls for

heat tolerance in a local "importing" country depending on information accessed from a foreign "exporting" country. The Walloon region of Belgium under temperate conditions, as "importing" country, and both Andalusia and Castile-La-Mancha regions of Spain, as "exporting" country under Mediterranean conditions, were selected as case study.

Materials and Methods

Cow performances and pedigree data were provided by the Walloon Breeding Association (AWE, Belgium) and the Confederation of Associations of Spanish Friesian (CONAFE, Spain). A total of 900,445 and 704,330 first-parity test day (TD) milk, fat, and protein yields from 113,282 and 81,752 Holstein daughters of 1,811 and 1,926 bulls in Belgium and Spain, respectively, were included. Both populations shared in common 369 bulls, each with more than 10 daughters. Those bulls have 56,265 and 36,212 daughters in Belgium and Spain, respectively.

Daily meteorological data (temperature and relative humidity) were provided by the consortium B-CGMS (CRA-W, IRM/KMI, ULg, VITO) for Belgium and AEMET for Spain. The daily THI values were computed as follows (NRC, 1971):

$$THI = (1.8 \times T_{db} + 32) - [(0.55 - 0.0055 \times RH) \times (1.8 \times T_{db} - 26.8)]$$

where T_{db} is the maximum daily temperature (°C) and RH is the minimum daily relative humidity (%).

Each TD record was merged with the average daily THI on 3 days prior to this TD, using the nearest weather station to the herd location based on minimum distance. The 3 day-lag was chosen following research of Bahmanova *et al.* (2008) and Hammami *et al.* (2008). The 95th percentiles of TD were recorded under mild to moderate heat stress (THI < 70) in Belgium. In Spain, the 75th percentiles of TD were measured under mild to moderate heat-stress days (THI > 70) and 95th percentiles under moderate to severe heat-stress days (THI >= 80).

The reaction norm model was used to perform the different genetic evaluations for milk, fat, and protein yield traits. The model was:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Q}_{hs}\mathbf{Wt} + \mathbf{Q}_{hs}(\mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{p}) + \mathbf{e}$$

where $\mathbf{y}$ was the vector of observations; $\mathbf{b}$ was the vector of the fixed effects including herd $\times$ TD, minor lactation stage (classes of 5 days in milk; DIM), and major lactation stage (classes of 73 DIM) $\times$ age at calving $\times$ season of calving effects; $\mathbf{t}$ was the fixed second-order Legendre polynomials coefficients designed to model the population mean effect of THI; $\mathbf{Q}_{hs}$ was the covariate matrix for second-order Legendre polynomials on standardized THI values in the interval [-1, 1]; $\mathbf{a}$ was the vector of additive genetic (AG) random regression coefficients; $\mathbf{p}$ was the vector of permanent environmental (PE) random regression coefficients; $\mathbf{e}$ was the vector of residuals; $\mathbf{X}$, $\mathbf{W}$, $\mathbf{Z}_1$, and $\mathbf{Z}_2$ were incidence matrices assigning observations to effects. Legendre polynomials were of order 2 with: $q_{hs0} = 1$, $q_{hs1} = x$, $q_{hs2} = 0.5*(3x^2 -1)$ where $x = -1 + 2 \times (THI - 32) / (83 - 32)$.

Four genetic evaluations were performed. A pedigree combining both Belgian and Spanish pedigree was used for all evaluations. First, a joint evaluation ($EVAL_J$) based on bivariate reaction norm model was performed using both Walloon and Spanish TD records. As detailed before, this case is an ideal one and is considered as the reference.

Second, local genetic evaluations for each population (i.e. $EVAL_{BEL}$ and $EVAL_{SPA}$) were performed using only the TD records related to this population. The REL associated to each of the 3 regression coefficients were computed as follows:

$$REL = 1 - PEV / \sigma^2_g$$

where σ^2_g is the additive genetic variance and PEV is the prediction error variance obtained by the inversion of the Left-hand-side of the mixed model equations.

Third, because $EVAL_J$ is an ideal case and is mostly not applicable, an alternative to integrate external information into an evaluation was tested. External information included regression coefficients for the additive genetic effect ($\hat{a}_{SPA}$) and their associated REL (REL_{SPA}) provided by $EVAL_{SPA}$. For each animal, $\hat{a}_{SPA}$ were converted to the Belgian trait as follows:

$$\hat{a}_{SPAc} = \mathbf{G}_{BELSPA} \, \mathbf{G}^{-1}_{SPA} \, \hat{a}_{SPA}$$

where $\hat{a}_{SPAc}$ is the vector of converted regression coefficients, $\mathbf{G}_{BELSPA}$ is the genetic additive covariance matrix between Belgium and Spain for the three regression coefficients and $\mathbf{G}^{-1}_{SPA}$ is the additive genetic (co)variance matrix for Spain.

For each of the three regressions coefficients, REL of the converted $\hat{a}_{SPA}$ (REL_{SPAc}) were estimated from all the REL_{SPA} resulted from $EVAL_{SPA}$ as follows (Powell *et al.*, 1994) as:

$$REL_{SPAc} = r^2_g * REL_{SPA}$$

where r^2_g is the genetic correlation coefficient between Belgium and Spain for the considered three regression coefficients.

The Bayesian evaluation ($EVAL_{BAYES}$) integrating external information, as defined above, into the Belgian evaluation was used as described by Vandenplas and Gengler (2012). The system of equations of the Belgian model integrating external information can be written as:

$$\begin{bmatrix} \mathbf{X'R^{-1}X} & \mathbf{X'R^{-1}Z} & \mathbf{X'R^{-1}Z} \\ \mathbf{Z'R^{-1}X} & \mathbf{Z'R^{-1}Z+G^{*-1}} & \mathbf{Z'R^{-1}Z} \\ \mathbf{Z'R^{-1}X} & \mathbf{Z'R^{-1}Z} & \mathbf{Z'R^{-1}Z+P^{-1}} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{p}} \end{bmatrix} = $$

$$\begin{bmatrix} \mathbf{X'R^{-1}y} \\ \mathbf{Z'R^{-1}y+G^{*-1}\mu_0} \\ \mathbf{Z'R^{-1}y} \end{bmatrix}$$

where the matrix G^{-1} is the inverse of the additive genetic (co)variance matrix and the matrix Λ is a block diagonal variance matrix with one block diagonal per individual (Quaas and Zhang, 2006).Each diagonal element i of Λ is equal to RE_{ij}/σ^2e for animals (j=1,2….,J) where RE_{ij} is the so-called record equivalents as defined by Misztal and Wiggans (1988) and σ^2e is the residual variance.

The four evaluations, i.e. $EVAL_J$, $EVAL_{BEL}$, $EVAL_{SPA}$ and $EVAL_{BAYES}$, were performed using the program blupf90 (Misztal, 2012) modified to integrate external information.

To appreciate the accuracy and precision of $EVAL_{BEL}$, $EVAL_{SPA}$ and $EVAL_{BAYES}$ for the three regression coefficients, spearman rank correlation coefficients between regression coefficients from $EVAL_J$ ($â_J$) and regressions coefficients from $EVAL_{BEL}$, $EVAL_{SPA}$ and $EVAL_{BAYES}$ were computed for bulls having only daughters in Spain and for bulls having daughters in both countries.

Results and Discussion

Under the ideal situation where raw data (phenotypes) are shared and free accessed, $EVAL_J$ provided estimates of genetic correlations. Moreover, the applied reaction norm models allowed good opportunities to identify the genetic merit of animals with low environmental sensitivity across THI trajectory (environments) without fixing a prior threshold point. Estimated genetic correlations between Belgium and Spain for milk, fat, and protein yields (Figure 1) were generally high across the whole THI trajectory (Figure 1). They ranged from 0.84 to 0.90 with a slight decrease

at extreme hot conditions, especially for fat and protein yields. These results indicate a possibility of genotype by environment interaction due to heat stress.
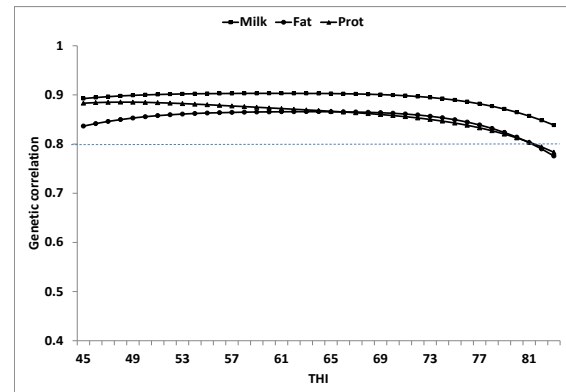

Figure 1. Genetic correlations for milk, fat, and protein yield between Belgium and Spain primiparous *Holsteins* across the THI trajectory.

Bohmanova *et al*. (2008) evaluated genotype by environment interactions due to heat stress within USA regions for milk yield of primaparous Hosteins. They found that heat tolerance EBV estimates were similar in cooler and hotter regions for only high-reliability sires.

In this study, the rank correlations for milk yield between regression coefficients ($â_{BEL}$ and $â_J$) for the 1,104 bulls having daughters only in Spain were moderate (<= 0.70 for all regression coefficients; Figure 2). Thus, the external information influenced the Belgian local evaluation only due to the relationships of Spanish bulls with animals in Belgium.

The estimates of $EVAL_{BAYES}$ showed a good concordance with those of $EVAL_J$ that pooled together raw data from both Belgium and Spain (Figure 2).
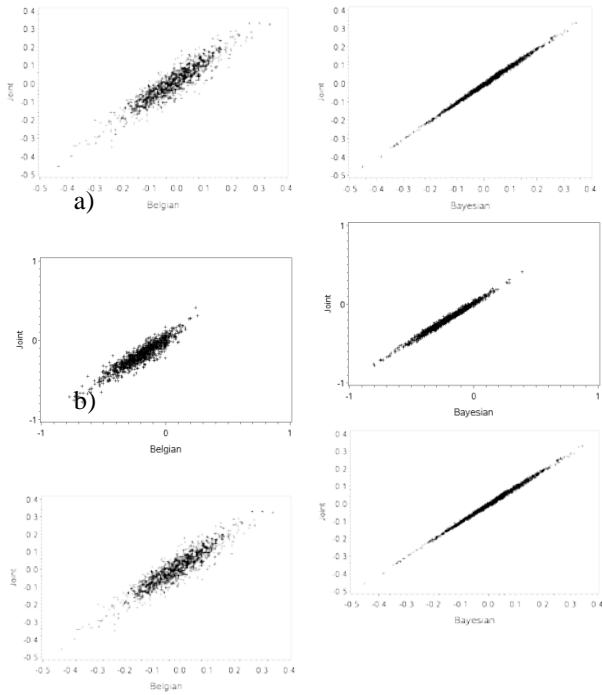
Figure 2. Rank correlations for milk yield between $\hat{a}_{BEL}$ and $\hat{a}_J$ (left) and correspondent correlations between $\hat{a}_{BAYES}$ and $\hat{a}_J$ (right) for a) intercept, b) first-order LP and c) second order LP for 1,104 bulls having daughters only in Spain.

Rank correlations for milk yield between regression coefficients ($\hat{a}_{BAYES}$ and $\hat{a}_J$) for the 1,104 bulls having only daughters in Spain were higher than corresponding correlations between $\hat{a}_{BEL}$ and $\hat{a}_J$.

The EVAL$_{BAYES}$ ranking of those foreign bulls was more similar to the ranking of EVAL$_J$, indicating that the integration of external information was efficient and resulted similarly as EVAL$_J$. These results corroborate with those reported by Quaas & Zhang (2006) and Vandenplas *et al*. (2013).

When comparing the different studied traits (Table 1), EVAL$_{BAYES}$ showed also very good concordance with the joint evaluation for fat and protein yields for bulls without daughters

in Belgium. Those bulls associated with external information ranked in EVAL$_{BAYES}$ very similar to EVAL$_J$.

Table 1. Spearman rank correlations for milk, fat, protein yields between regression coefficients for the additive genetic effect using joint evaluation ($\hat{a}_J$) and local separate ($\hat{a}_{BEL}$) and Bayesian evaluation ($\hat{a}_{BAYES}$) for the 1,104 bulls with only daughters in Spain.

Traits	Evaluation	Inter-cept	1st order LP	2nd order LP
Milk	$\hat{a}_{BEL}$	0.667	0.920	0.912
	$\hat{a}_{BAYES}$	0.978	0.992	0.997
Fat	$\hat{a}_{BEL}$	0.637	0.965	0.978
	$\hat{a}_{BAYES}$	0.980	0.994	0.989
Protein	$\hat{a}_{BEL}$	0.708	0.986	0.951
	$\hat{a}_{BAYES}$	0.977	0.998	0.996

Concerning the rankings of bulls with at least 10 daughters in both countries, the rank correlations between $\hat{a}_{BEL}$ and $\hat{a}_J$ for all traits (Table 2) were higher compared to correlations of bulls having only Spanish daughters when estimated by EVAL$_{BEL}$. The integration of external information for foreign bulls into the Belgian genetic evaluation via the Bayesian approach influenced very slightly the ranking of those bulls.

Table 2. Spearman rank correlations for milk, fat, protein yields between regression coefficients for the additive genetic effect using joint evaluation ($\hat{a}_J$) and local separate ($\hat{a}_{BEL}$) and Bayesian evaluation ($\hat{a}_{BAYES}$) for the 369 bulls with at least 10 daughters in Belgium and Spain.

Traits	Evaluation	Inter-cept	1st order LP	2nd order LP
Milk	$\hat{a}_{BEL}$	0.957	0.993	0.993
	$\hat{a}_{BAYES}$	0.967	0.993	0.996
Fat	$\hat{a}_{BEL}$	0.958	0.996	0.992
	$\hat{a}_{BAYES}$	0.981	0.996	0.992
Protein	$\hat{a}_{BEL}$	0.959	0.995	0.988
	$\hat{a}_{BAYES}$	0.978	0.998	0.996

Conclusions

The Bayesian approach integrated well external information obtained from an external evaluation, allowing an indirect recovering of non-accessed raw-data (phenotype and pedigree). Rankings of foreign bulls with or without daughters at local level of the Bayesian evaluation were similar to rankings resulting from EVAL$_J$. Therefore, the Bayesian approach showed a great interest to integrate external information when phenotypes are scarcely and/or not free accessed. The latter situation is still unfortunately charactering nowadays the various novel traits. The flexibility of Bayesian integration proved being well adapted to heat tolerance as novel trait. Currently, this approach was tested using as a case study, the Belgian Holsteins, raised under continental and temperate climate, and the Spanish populations, experiencing more heat and warm conditions under Mediterranean climate. Ongoing investigations are integrating several sources of external information provided by Luxembourg and Slovenia for better consolidation of the methods. Nevertheless, additional optimizations of the Bayesian integration are needed towards multi-trait use to avoid conversion step.

Acknowledgments

References

Bohmanova, J., Misztal, I., Tsuruta,, Norman, H.D. & Lawlor, T.J. 2008. Genotype by environment interaction due to heat stress. *J. Dairy Sci. 91,* 840-846.

Colinet, F., Vandenplas, J., Faux, P., Vanderick, S., Renaville, R., Bertozzi, C., Hubin, X. & Gengler, N. 2013. Wallon single step genomic evaluation system integrating local and MACE EBV. *Interbull Bulletin 47,* 203-210.

Hammami, H., Bormann, J., M'Hamdi, N., Montaldo, H.H. & Gengler, N. 2013. Evaluation of heat stress effects on production traits and somatic cell score of Holsteins in a temperate environment. *J. Dairy Sci. 96,* 1844-1855.

Quaas, R.L. & Zhang, Z. 2006. Multiple -breed genetic evaluation in the US beef cattle context: methodology. CD-ROM Commun. 24-12 in *Proc. 8thWorld Congr. Appli. Livest. Prod.,* Belo orizonte, Brazil.

Misztal, I. 2012. *BLUPF90 Family of Programs.* University of Georgia, Athens, GA, USA (available at http://nce.ads.uga.edu/~ignacy/newprograms.html; last accessed on 15 May 2014).

NRC. 1971. *A Guide to Environmental Research on Animals.* Natl. Acad. Sci., Washington, DC.

Powell, R.L., Wiggans, G.R. & VanRaden, P.M. 1994. Factors affecting calculation and use of conversion equations for genetic merit of dairy bulls. *J. Dairy Sci. 77,* 2679-2686.

Vandenplas, J. & Gengler, N. 2012. Comparison and improvements of different Bayesian procedures to integrate external information into genetic evaluations. *J. Dairy Sci. 95,* 1513-1526.

Vandenplas, J., Janssens, S., N. Buys, N. & Gengler, N. 2013. An integration of external information for foreign stallions into the Belgian genetic evaluation for jumping horses. *J. Anim. Breed. Genet. 130,* 209-217.