

Genotype-environment interactions in pig breeding

P.K. Mathur

Topigs Norsvin Research Center, PO Box 43, 6640 AA Beuningen, The Netherlands
pramod.mathur@topignorsvin.com

Summary

Genotype-environment interactions (GxE) should be expected to occur whenever more than one genotype and more than one environment are considered, although their magnitude may differ. It is important to make use of this natural phenomenon rather than ignoring it or considering it as a problem. There are several examples and estimates of GxE in pig breeding suggesting significant GxE with respect to heat stress, diseases, seasonal variations and variation in performance during different periods within farms. Reaction models are quite useful in estimating these GxE effects as genetic correlations between different levels of environmental stress. Several studies using these models have revealed lower genetic correlations between the intermediate environment and environments with higher or lower stress levels. These results suggest that indirect selection based on performance in an optimal environment, ignoring the GxE effects, would result in less than optimal performance in a stressful environment. At the same time, there are opportunities for enhancing performance under stressful environments. Most studies have revealed the availability of larger genetic variance in stressful environments, providing opportunities for selection within the environment and making faster genetic progress. Further, the reaction norm models also allow estimating breeding values of animals in non-stressful environments based on data from stressful environments for the selection of robust animals and sorting of terminal boars. Genomic information can improve the accuracy of these breeding values and enhance genetic gains. SNP-environment interactions and gene-environment interactions provide further opportunities for effective utilization of this phenomenon for future pig breeding.

Keywords: genotype-environment interactions, reaction norm models, heat stress, seasonality, diseases, challenge load

Introduction

In large scale pig breeding programs, the selection of breeding animals is typically carried out in high health nucleus environments while actual commercial production takes place in suboptimal environments with a variety of challenges due to diseases, climate, feeding and management related issues. In this context, genotype-environment interactions (GxE) between nucleus and commercial environments may play an important role in the performance of animals raised in commercial environments. Therefore, it becomes very important to account for GxE in estimation of breeding values of selection candidates in the nucleus. However, GxE effects are often ignored resulting in lower accuracy of breeding values in predicting performance in commercial farms.

A review of GxE in different livestock species and description of various terms used in this connection such as environmental sensitivity, plasticity, canalization, reaction norms and robustness is given by Rauw and Gomez-Raya (2015). In this review, the methods for

estimation of GxE and their usefulness are discussed. These are then followed by examples of GxE relating to different environmental conditions. Finally, some new developments and perspectives for the use of GxE in pig breeding programs are discussed.

Methods for estimation of GxE effects

In general, there are fewer studies regarding methods for the estimation of GxE for specific use in pig breeding compared to other livestock species and plants. Available studies have used different types of genotypes and environments using different assumptions in estimation methods. These have not only contributed to differences in the application of estimation methods but also to the overall attitude towards consideration or negligence of GxE in pig breeding programs. In a review on methods for estimating the magnitudes of GxE (Mathur, 2002) suggested that an important distinction should be made if the genotypes are considered as discrete classes of a few genotypes such as distinct breeds or lines usually considered as fixed effects or there are a large number of genotypes such as sires that are a sample from a population of genotypes usually considered as random effects. The same distinction also applies to environmental effects, whether they are a few classes such as the presence or absence of specific disease or a range of environmental effects such as the level of pathogen burden. In pig breeding, there are limited examples of estimation of GxE with respect to few breeds or lines. More commonly, sires or animals are considered as genotypes. Among the methods used, most common are bivariate and reaction norm models.

Bivariate models

These models are applied considering pairs of distinct classes of environment, e.g. heat stress or no heat stress, disease or healthy condition. The magnitude of GxE is then estimated as the genetic correlation between the environments (Falconer, 1952). The lower the correlation, the higher the interaction. However, quite often it is difficult to make a clear distinction between two environmental classes as the level of the environmental stress is often continuous in nature and reaction norm models are considered more appropriate.

Reaction norm models

In breeding value estimation, the aim is to estimate the genetic merit of each selection candidate with respect to the production environment. The GxE then refers to the expected reaction of the genotype to the level of environmental challenge. Random regression models are often suggested for genetic evaluations. In the context of GxE, those are often referred to as reaction norm models (De Jong and Bijma, 2002). These models include a random effect of the genotype (e.g. animal or sire) and a GxE interaction component as regression or slope of the genetic effect with respect to the environmental level.

Herrero-Medrano et al. (2015) compared three types of models: a) bivariate model considering two distinct classes, b) a reaction norm model accounting for environmental effects as a continuous variable, and c) a univariate model that does not include the GxE effect. It became clear that it is difficult to identify distinct boundaries for identifying periods high and low challenge, especially when herd-year-season (HYS) effects are considered as environmental range. In those situations, a reaction norm model is more appropriate. Further, considering the main goal to breed robust pigs that have a higher resilience to stressful

environments, it is more appropriate to use a reaction norm model compared to the univariate model. It was demonstrated by Silva et al. (2014) that the accuracy of breeding values in predicting progeny performance was higher using reaction norm model compared to univariate model. The use of genomic relationship matrix helps to increase the accuracy using a univariate model but genomic reaction norms further the accuracy even further.

A detailed description of the benefit of genomic reaction norm models is given by Mulder (2016). This is based on a simulation study comparing estimates of breeding values considering different sizes of reference populations, number of progeny per sire and covariance between intercept and slope of reaction norm models. The accuracy of genomic breeding values was higher using a genomic reaction norm model compared to the conventional pedigree based model. The difference was largest at the extremes of the environmental levels suggesting the utility of genotyping pigs from commercial environments as well. The study further highlights usefulness of reaction norm models for estimation and use of GxE interactions.

One of the requirements for the application of reaction norm models is that the genotype has to be tested in a large number of environments e.g. a sire's progeny needs to be tested under different temperatures, health status etc.. Considering an animal model, often the difficulty is that data on the animal's own performance in different environments may be limited. As an example, reproduction records of a sow may be available only for one or few parities that differ highly with respect to temperature. Knap and Su (2008) have demonstrated this difficulty with three sets of data and showed that accuracy of estimates increases with increase in data volume and with increase in environmental variation in the data. Therefore, the main requirement is that the reference population is spread across significantly different environments and large number of observations is available. Partly, this is already taking place due to increase in data collection and genotyping in commercial farms.

Comparison of bivariate and reaction norm models

Genetic correlations between environments can be estimated using bivariate models as well as reaction norm models. The reaction norm models have the advantage of their ability to estimate change in the genotypic effect per unit change in environmental effect considering continuous nature of the environmental effect. Genetic variances and heritability estimates often follow a parabolic shape with an increase in the environmental effects and the genetic correlations are lower among extreme environments and relatively higher among intermediate levels. In general, this reflects the natural phenomenon of GxE. The larger the differences between environments, the higher the likelihood of GxE. However, sometimes there are doubts whether the low genetic correlation is actually indicative of high GxE or it is just an artefact due to mathematical properties of the reaction norm model. This concern mainly arises as in part of the calculation of genetic variance for a specific environment, the variance due to slope is multiplied by the squared environmental value with linear reaction norms. If variance due to slope of the reaction norm is over estimated, then at the extremes, multiplying the variance with the squared environmental value could result in inflation of the estimate of the genetic variance, especially if the genetic variance is estimated for environmental levels beyond the range of observed data. Although this concern should not be valid if the slope, and covariance between intercept and slope are accurately estimated. Typically, there are fewer data records under extreme environmental conditions e.g. very high heat stress or very low

heat stress. These could result in such inaccuracies in the estimates of slope and covariance as demonstrated by Knap and Su (2008). Another possible reason for the inaccuracy of genetic variance could be due to polynomials used. For example, if a quadratic function is used in the mathematical model, while the data actually follow a cubic or some other function, then the estimates could deviate from the true situation. One way to verify this is to compare genetic correlations estimated from reaction norm models with bivariate models. This has been done in several studies (Rashidi et al., 2014; Herrero-Medrano et al., 2015; Sevillano et al., 2016) and the estimates were quite comparable.

Environmental effects

In commonly used models in pig breeding and also other livestock species, the term environment refers either to specific components of climate e.g. temperature, humidity, diseases, nutritional regimes etc. or a combination of these or similar effects. Here, the effects of heat stress, seasonal variation and disease challenges are described as examples of specific environmental effects.

Heat stress

GxE with respect to heat stress and finishing traits was investigated by (Zumbach et al., 2008). Here, sires were considered as genotypes and different months of increasing temperature were considered as environments. The genetic correlation between carcass weights in extreme cold months and months with maximum heat load was 0.02, between cold months and months with intermediate heat load was 0.52, and between months with intermediate heat load was 0.86 suggesting higher GxE with larger environmental differences.

GxE with respect to heat stress in two commercial lines was investigated by (Bloemhof et al., 2008) with the aim to breed for higher heat tolerance and to minimize unfavourable effects on reproduction traits. The study was based on the concept of upper critical temperature. According to this concept, pigs suffer from heat stress when temperature exceeds the upper critical temperature of the thermo-neutral zone. In a Dutch line, the upper critical temperature was 19.2°C for farrowing rate, 21.7°C for litter size and 19.6°C for total number of piglets born from first insemination. There was a decrease in farrowing rate by 1%, in litter size by 0.05 piglet and by 0.13 in total number of piglets born from first insemination per °C. However, in an international line there was a very small decrease such that even an upper critical temperature could not be estimated. This way, the two genotypes (Dutch line and international line) had differential expression over the two environments: a clear example of GxE between lines. In this case, the Dutch line is more suitable for countries with no heat stress while the international line is more suitable for countries with heat stress such as Spain and Portugal. This is an example of the usefulness of specific genotypes (e.g. specific lines) for specific environments. This phenomenon has been emphasised with several types of genotypes in poultry breeding as well (Mathur, 2003). These examples show that if GxE is ignored and only one line, e.g. the International line is used across all countries, it will result in suboptimal output for customers in the Netherlands. Similarly, using the Dutch line alone will be less beneficial for customer farms in Spain and Portugal affected by heat stress. Here, the GxE can be used as a resource by deciding to maintain specific lines that are most suited for the target environmental conditions.

Seasonal variation

An adverse effect of seasonal variation is often observed in reproductive performance of pigs. A common term is “seasonal infertility” which refers to anoestrus or a lower conception rate resulting in a reduced farrowing rate. The seasonal variation is related to the photoperiod or day length during different seasons of the year which mainly depends upon the geographical location. Sevillano et al. (2016) used cosine of the day of first insemination within a year to model photoperiod and estimated the rate of change in day length. As sows mainly respond to change in photoperiod, it was estimated using a cosine function of the photoperiod function. An important finding in this study was also that GxE exists for seasonal variation even after accounting for effect of temperature. As a next step, the combined effects of temperature and day length were grouped in to five classes ranging from non-stressful environments to severely stressful environments. The genetic correlations between environments decreased from non-stressful environments to environments with severe stress, suggesting increase in GxE with increases in environmental differences. The genetic correlation between neutral environment and other intermediate classes was high (≥ 0.90), indicating negligible GxE. However, the genetic correlation between the intermediate class and the class with highest stress was as low as 0.46 ± 0.13 , indicating substantial GxE and change in ranking of pigs between neutral and severe stress environments. In this case, use of reaction norm models will allow estimating two types of breeding values for each selection candidate. A breeding value for the slope, estimating the change in the breeding value per unit change in the day length due to GxE, and another for the intercept, considering no environmental effect. Using the day length function according to the farm location will allow ranking of all available terminal sires for that specific location. Those breeding values can be used to identify the sires that are most suitable for the location.

Diseases

In pig breeding programs, genetic evaluation and selection is mainly based on data from nucleus farms that have a specific pathogens free (SPF) or high health status while their progeny is expected to perform in commercial farms with lower or basic health status. In spite of the attention to biosecurity measures, several diseases occur in commercial farms. These diseases and their outbreaks lead to production losses and less than expected performance in commercial farms. In this context, GxE plays an important role and could be used in breeding for higher disease resilience (Mulder and Rashidi, 2017). In our study, farrowing records from 11 farms from the Netherlands and Canada were evaluated. The farm health status was classified as SPF, Basic (free from notifiable diseases but with variety of disease pathogens), Basic but free from APP (*Actinobacillus pleuropneumoniae*) and Basic but free from PRRS (Porcine Reproductive and Respiratory Syndrome). The genetic correlation between farms with SPF and non-SPF status ranged from 0.34 to 0.69, indicating large GxE effects. However, the estimates were also associated with large standard errors, ranging from 0.23 to 1.15, due to low connectedness between the farms. This represents a practical difficulty in obtaining data for such studies. Nevertheless, larger magnitudes of GxE between the farms with high health status e.g. nucleus farms and farms with low disease status e.g. commercial farms can be expected depending upon the biosecurity conditions and prevalent diseases. Efforts should be made obtain more data from commercial farms to get more accurate estimates of GxE. Those estimates of GxE could be used to identify lines with higher disease resilience and further to select sires within those lines that deliver best progeny performance

in commercial conditions.

One of the main difficulties in estimating the GxE with respect to diseases in pigs is the difficulty in having sufficient data on the level of disease challenge e.g. pathogen burden. Estimating the load of virus or bacteria requires collection of blood or tissue samples and rather expensive lab analysis which is hardly possible on a large number of pigs. On the other hand, a large number of performance records relating to growth, feed intake, reproduction etc. during different diseases are already available in the databases of breeding organizations. Therefore, some studies have used performance records as an alternative measure of level of disease challenge. The basic assumption is that diseases adversely affect performance. The higher the level of disease challenge, the lower is performance.

Lewis et al. (2009) used 30 day rolling averages of number of mummified piglets per litter for detecting outbreaks of PRRS. When this rolling average was above a 99% confidence threshold, the period was considered as a disease outbreak. This method was further refined using linear statistical models by Rashidi et al. (2014). Estimate of each herd-year-week effect and its standard error was used to test significant deviation from normal performance to detect disease outbreaks. The linear models approach (Rashidi et al., 2014) for the total number of pigs born had sensitivity of 78% compared to 21% by using weekly averages (Lewis et al., 2009). These studies were based on analysis of one trait at a time to detect disease outbreaks while quite often diseases like PRRS adversely affect several traits. In many such cases, one trait e.g. still birth is more affected in one farm while another trait e.g. number of pigs weaned is more affected in other farm. Hence, a combination of multiple traits could be more effective in detecting disease outbreaks and estimating GxE between disease and non-disease periods. Mathur et al. (2014a) described such a method. It was then tested and validated in several countries to detect outbreaks of PRRS and other diseases (Mathur et al., 2014b). Further, the method provides a measure of challenge load as a measure of the magnitude of the effect of any stressful environment. It can be used not only to differentiate between classes of disease and health but also as a covariate in reaction norm models to estimate the regression of the stress due to a combination of environmental effects.

A combination of multiple environmental effects

In commercial pig breeding, several authors have used estimates of herd-year-season (HYS) effects as a combination of climate, feed, management etc. e.g. (Knap and Su, 2008); (Silva et al., 2014). Herrero-Medrano et al. (2015) used the concept of challenge load (CL) (Mathur et al., 2014a) and herd-year-week (HYW) estimates of challenge load as measure of environmental level in reaction norm models. Ten classes of increasing CL were created. These classes were also used to account for the effect of heterogeneity of residual variances among the environmental levels. The genetic correlations between pairs of these classes ranged from 0.51 ± 0.06 for number of weaned pigs to 0.75 ± 0.03 for number born alive, due to GxE. Further, the genetic variances for these traits were significantly higher in the high challenge environment compared to those in low challenge environment. The correlations were 0.91 vs. 1.52 in number born alive, and 0.12 vs. 0.79 in number of lost pigs which was sum of stillborn and mummified pigs, respectively. Higher genetic variances under higher environmental stress should be expected as the fitness and survival traits such as number of stillborn and mummified pigs are not fully expressed in non-stressful environments. Given higher genetic variance, these GxE effects provide higher opportunity for selection and for

making more genetic progress in stressful environments.

Gene-environment interactions and gene networks

The GxE exists not only with respect to genotypes such as lines, sires or individuals but also with respect to specific, genes, genomic regions or SNPs. Genomic reaction norms can be used to estimate SNP-environment interactions. Silva et al. (2014) suggested this approach and identified SNPs that are common among multiple environment (HYS) levels and those that are different for total number of pigs born. Only 1% of the top 462 SNPs were common at all HYS levels. Even the effects of common SNPs differed remarkably across the HYS levels. Some SNPs had more favourable effects under unfavourable environment while effects of others were in opposite direction. Although the effects of individual SNPs were not very large there were indications of usefulness this SNP-environment interactions. These SNPs, either individually or as a combination, can be used to estimate genomic breeding values for specific environment and thereby to increase the response to selection in the target environment. The existence of SNP-environment interaction provides this unique opportunity.

In a study of total number of pigs born (TNB), Verardo et al. (2017) observed that SNPs relating to high HYS levels explained the higher variance than those relating to low HYS levels. Three LD blocks were identified relating to all, low and high HYS groups. Three important genes were identified: the *BMPRI1B* relating to the low HYS groups while *TGFB3* and *SLIT3* relating to the high HYS groups. All the three genes are known to play an important role in the ovulation process. *BMPRI1B* (bone morphogenetic receptor type 1B) gene has been related to the high prolificacy Booroola phenotype in sheep and reproductive performance of Iberian x Meishan F2 cross in pigs. The *TGFB3* (transforming growth factor beta 3) gene plays an important role in modulating theca cell function of pre- and postovulatory follicles in pigs. The *SLIT3* (Slit Guidance Ligand 3) could influence the process of follicle formation and oocyte survival in early ovarian developments in women. These results suggest that expression of a single trait might be controlled by different genes depending upon the environment and provide evidence of gene-environment interactions. In an earlier study, Mathur (2003) described gene-environment interactions in poultry with respect to major genes (naked-neck and dwarf). More evidences of GxE are expected in the future, with the increase in availability of genomic information and identifications of genes responsible for many biological processes. That will further increase our understanding of the interesting biological phenomenon of GxE and help in genetic selection.

Perspectives on estimation and use of GxE in breeding programs

In pig breeding there is a trend towards industry consolidation like many other industries. The breeding pigs for many countries around the world are produced by only a few breeding organisations. The environmental conditions in the target countries of a breeding organization can differ substantially. Therefore, large GxE can be expected between the nucleus and production environments. GxE can exist between farms but also between periods with and without stress within farms. The question then is, whether these breeding organisations should maintain different lines for different environments (e.g. countries or regions), if they should focus on selection for robustness in fewer lines or completely ignore the GxE and breed for maximizing productivity in the most favourable environment.

The decisions mainly depend upon the magnitudes of differences between environments, the magnitude of GxE for the trait of interest, availability of data and cost benefits of maintaining different lines. If the differences between the selection and target environments are large, such as high heat stress and disease challenge, larger GxE can be expected especially for maternal traits. This requires maintenance of one maternal line for countries and farms with heat stress and another for normal temperature conditions. Similar differences between environments can exist with respect to disease pressure, feed quality, housing and other management conditions. Then it could be practically difficult to maintain multiple lines. Therefore, another approach is to maintain fewer lines based on considerations other than GxE but breed for higher robustness or resilience in a specific line. However, there are very few examples of actual use of GxE in commercial pig breeding resulting into suboptimal genetic gains in stressful environments. The breeding organizations have the opportunity to account for GxE in genomic reaction norm models and obtain environment specific breeding values. These can be used to customize selection and sorting of sires to each of the target environment. Discovery for more genes affecting the gene-environment interactions should be expected in near future. These, in combinations with advanced genomic selection techniques, could further boost the genetic gains.

It is often argued that with the increasing trends towards large and highly technological commercial farms, there is no need to consider GxE because the environmental conditions will be better than in the traditional farms. This may be possible in some commercial farms. However, in many cases, it is either not possible or too expensive. As an example, air-conditioning in pig farms is possible but may not be cost effective. Similarly, in spite of increased attention on biosecurity, severe disease outbreaks occur for one reason or the other. Therefore, breeding of pigs that are more robust to environmental challenges, by effective use of GxE, should be given higher priority by pig breeding organisations. There should be a positive attitude towards GxE. Rather than considering it as a problem, it should be considered as a resource to enhance genetic gains in desired traits and environments.

List of References

- Bloemhof, S., E. Van der Waaij, J. Merks & E. Knol, 2008. Sow line differences in heat stress tolerance expressed in reproductive performance traits. *J. of Animal Sc.* 86(12): 3330-3337.
- De Jong, G. & P. Bijma, 2002. Selection and phenotypic plasticity in evolutionary biology and animal breeding. *Livestock Production Sc.* 78(3): 195-214.
- Falconer, D. S. 1952. The problem of environment and selection. *The American Naturalist* 86(830): 293-298.
- Herrero-Medrano, J. M., P. K. Mathur, J. ten Napel, H. Rashidi, P. Alexandri, E. F. Knol & H. A. Mulder, 2015. Estimation of genetic parameters and breeding values across challenged environments to select for robust pigs. *J. of Animal Sc.* 93(4): 1494-1502.
- Knap P. & G. Su, 2008. Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. *Animal.* 2(12): 1742-1747.
- Lewis, C., M. Torremorell, L. Galina-Pantoja & S. Bishop, 2009. Genetic parameters for performance traits in commercial sows estimated before and after an outbreak of porcine reproductive and respiratory syndrome. *J. of Animal Sc.* 87(3): 876-884.
- Mathur, P., 2002. Methods for estimation and use of genotype-environment interactions. In: *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production,*

- Montpellier, France, August, 2002. Session 18. p 1-8.
- Mathur, P., 2003. Genotype-environment interactions: problems associated with selection for increased production. In: W. M. Muir and S. E. Aggrey (editors.) *Poultry Genetics, Breeding and Biotechnology*. p 83–99. CABI Publishing, Oxon, UK.
- Mathur, P. K., J. M. Herrero-Medrano, P. Alexandri, E. F. Knol, J. ten Napel, H. Rashidi & H. A. Mulder, 2014a. Estimating challenge load due to disease outbreaks and other challenges using reproduction records of sows. *J. of Animal Sc.* 92(12): 5374-5381.
- Mathur, P. K., J. Herrero-Medrano, P. Alexandri, E. F. Knol, H. Rashidi, H. A. Mulder & J. ten Napel, 2014b. Genetic Selection for Disease Resistance and Tolerance in Pigs using Reproduction Records. In: *Proc. 10th World Congr. Genet. Appl. Livest. Prod.*, Vancouver, Canada
- Mulder, H. & H. Rashidi, 2017. Selection on resilience improves disease resistance and tolerance to infections. *J. of Animal Sc.* 95(8): 3346-3358.
- Mulder, H. A, 2016. Genomic selection improves response to selection in resilience by exploiting genotype by environment interactions. *Frontiers in genetics*, 7.
- Rashidi, H., H. A. Mulder, P. Mathur, J. A. van Arendonk, and E. F. Knol. 2014. Variation among sows in response to porcine reproductive and respiratory syndrome. *J. of Animal Sc.* 92(1): 95-105.
- Rauw, W. M. & L. Gomez-Raya. 2015. Genotype by environment interaction and breeding for robustness in livestock. *Frontiers in genetics* 6.
- Sevillano, C., H. Mulder, H. Rashidi, P. Mathur & E. Knol. 2016. Genetic variation for farrowing rate in pigs in response to change in photoperiod and ambient temperature. *J. of Animal Sc.* 94(8): 3185-3197.
- Silva, F. F., H. A. Mulder, E. F. Knol, M. S. Lopes, S. E. F. Guimarães, P. S. Lopes, P. K. Mathur, J. M. S. Viana & J. W. M. Bastiaansen, 2014. Sire evaluation for total number born in pigs using a genomic reaction norms approach. *J. of Animal Sc.* 92(9): 3825-3834.
- Verardo, L.L., Lopes M.S., Mathur P., Madsen O., Silva, F.F. Groenen, M.A. Knol, E.F. Lopes, P.S. & Guimarães, S.E., 2017. Gene networks for total number born in pigs across divergent environments. *Mammalian Genome*. Jun 2:1-10.
- Zumbach, B., I. Misztal, S. Tsuruta, J. P. Sanchez, M. Azain, W. Herring, J. Holl, T. Long & M. Culbertson, 2008. Genetic components of heat stress in finishing pigs: parameter estimation. *J. of Animal Sc.* 86(9): 2076-2081.