
Genotype by Environment Interactions in Livestock Farming



Francesco Tiezzi¹ and Christian Maltecca²

¹Department of Agriculture, Food, Environment and Forestry, University of Florence, Firenze, Italy

²Department of Animal Science, North Carolina State University, Raleigh, NC, USA

Article Outline

Glossary

Definition of the subject

Introduction

Expanding the Model

References

Glossary

Genotype by environment interaction The difference in response to environment changes due to different genotypes.

Breeding value Expected performance measured as deviation from the population mean of the progeny generated by a progenitor.

Phenotype Observable characteristics of an individual.

Resilience The ability to recover from stressful conditions.

Robustness The ability of not being perturbed by stressful conditions.

Tolerance The ability to cope with stressful conditions.

Plasticity The ability to change in response to environmental inputs.

Acclimation Increase of tolerance to stressful levels of environmental parameters.

Definition of the subject

Genotype by environment interaction, often referred to as “**G × E**,” is the phenomenon for

which the breeding value of an individual depends on the environmental conditions and the effect of an environmental factor depends on the individual’s genetic background. A breeding program that accounts for G×E allows, among other things, the development of breeds and lines that are particularly adapted to specific environmental and management conditions. Historically, animal breeders have neglected the phenomenon, and it has seldom been exploited in genetic evaluations. Conversely, plant breeders have made it familiar and developed breeding plans that take advantage of such interaction effects. The reasons for this difference are multiple but can mainly be reconducted to the following two: (1) The development of new plant germplasm is a process that is more affordable than the development of new animal germplasm. Such cost-effectiveness of plant breeding allows developing lines that will be used under a limited number of conditions. The development and maintenance of animal germplasm adapted to only marginal breeding systems will likely be uneconomical. (2) Animal husbandry allows managing livestock in a more precise manner than agronomy does with plants. For example, animals’ diets can be controlled with high precision, while rainfall and nutrient availability in the (open-field) soil can only be controlled to some extent. Better control of the environmental conditions lower the need to develop germplasm adapted to harsh conditions.

Introduction

Over the last seven decades, selective breeding has significantly contributed to increasing the productivity of animal systems. One significant step forward was taken when breeding values (BV) started to be calculated using mixed-model equation systems [1, 2, 3]. One of the main advantages of mixed models was the ability to disentangle the genetic from the environmental (systematic and random) effects. Consequently, individuals could be compared for their breeding

value even if they were reared under different environmental conditions. In formula notation:

$$P = G + E + G \times E \quad (1)$$

where P is the individual's phenotypic measure for a given trait of interest, G is the genetic component (the BV), E is the environmental component, and $G \times E$ is the interaction between the two components. In synthesis, mixed models were able to extract the heritable component G from P and provide an indicator for the BV of the individual. The G component is obtained *marginally* on the E component: an individual will have a deviation value from the rest of the population (the BV), regardless of the conditions under which it will be raised. The E component is the environmental deviation. This deviation is also *marginal* on the BV of the individual: a change in diet is assumed to change the performance of all the individuals in the same way, regardless of their genetic background and, therefore, their ability to react to the change. The $G \times E$ component is interpreted differently. Such component implies that genotypes will differ in response (P) according to the environmental conditions. The environmental state might determine the response depending on the genotype receiving it [4]. For example, a heritable component on tolerance to heat stress has been found on all livestock species ([5]; see later sections of this entry). The genetic control on tolerance to heat stress is attributable to the $G \times E$ component: the individual's BV depends on the temperature and humidity that the individual is exposed to. Such BV could be favorable under a comfortable environment but unfavorable under an uncomfortable environment, which means that the individual is a top performer under optimal conditions but is outperformed by other individuals under heat stress, that is, it is less tolerant than others and cannot acclimate to the less comfortable conditions. The same principle could apply to different diet and management systems. The *marginality* of the breeding value could be replaced with its *conditionality*, that is, the BV will depend on the environmental conditions. The inclusion of $G \times E$ in genetic evaluations and, therefore, breeding

programs implies a shift in the setup of the whole system. Since BV are no longer *marginal* on environmental conditions, this implies that such environmental conditions are adequately accounted for in the genetic evaluation and, therefore, in the breeding decisions.

This entry will expand the description of the concepts mentioned here: genotype by environment interaction, acclimation, tolerance, and environmental conditions. Examples will be provided that show the importance of accounting for $G \times E$ in livestock breeding and the pitfalls of its implementation.

The model mentioned above will be revisited and expanded with components that describe the genetic control of resilience in more detail. Next, the different definitions of environmental components and how these can be used in modeling $G \times E$ will be revised. Successively, the conceptual and practical description of the models that can predict breeding values for resilience will be described. A brief review of current literature on the topic will be provided, followed by some conceptual examples of the use of $G \times E$ in livestock breeding.

Expanding the Model

1. The Model Revisited

The model reported in formula (1) is widely accepted as a conceptual framework for the dissection of phenotypic variation in the two main components and their interaction. For this entry, that model will be expanded to better define the environmental component.

Following Jinks and Pooni [6] and Mulder et al. [7], E will be separated into two subcomponents, specifically a proper E component and an e component:

$$P = G + E_c + G \times E_c + e_r + G \times e_r \quad (2)$$

Using this parameterization:

- E_c is the controllable environmental component such as temperature in a barn or percentage of

protein in the feed. This can also be defined as the *macro-environmental* component.

- e_r is the random noncontrollable component of the environment, such as indirect social effects or the exposure to uncontrollable pathogens, termed *micro-environmental* component.

The genetic contribution needs to be redefined in three components: G , $G \times E_c$, and $G \times e_r$.

- G is the *marginal* deviation of an individual from the population genetic component mean so that its performance will not depend on the environmental conditions.
- $G \times E_c$ is the genetic component *conditional* on the *controllable* environmental component. If modeling can *predict* the E_c conditions in a given contemporary group, proper genetic material is allocated to optimize the use of resources. For example, genetic material can be tailored to expected temperature and humidity conditions, availability of energy in the diet, etc. This component expresses the macro-environmental plasticity, i.e., how a genotype shows a plastic phenotype given macro-environmental changes.
- $G \times e_r$ is the genetic component *conditional* on the *uncontrollable* environmental component. Despite the advances in technology, some environmental conditions will *not* be controlled (e_r). Still, e_r will affect the phenotype of farm animals and the $G \times e_r$ will imply a different ability of the individuals to overcome environmental changes. The ability of a genotype to show a plastic phenotype as defined by this component is also known as micro-environmental plasticity.

2. The Environmental Components

Formula (2) aims at dissecting the environmental component E in two parts, contingently on the possibility of controlling the factors that determine it. Ultimately our ability to control it defines whether the component is macro-environmental or micro-environmental. This has a clear zootechnical meaning. In livestock farming, the optimal

conditions for the animals' health, well-being, growth, and production are pursued. Over time, knowledge and technological inputs have shaped how farm animals are raised, with consequences on the systems' productivity and the farmers' welfare. The improvement of the farming conditions, broadly speaking, has been provided by controlling several factors. Among these are the control of the diet, ambient conditions (e.g., air temperature and humidity and CO_2 and NH_4 concentrations), and the control of the reproductive cycle. The evolution of the farming system has led to these factors being (at least potentially) controlled with accuracy and precision.

Nonetheless, such control is often lacking for several reasons that will not be covered in this entry. From our perspective, it will be more important to define the *controllable* and *not controllable* parts of the environmental variation [7].

As the domestication of livestock species occurred, the individual's diet was one of the first factors the farmers attempted to control. Nowadays, diets consist of rations formulated to maximize farm profitability and animal welfare. Rations are administered to individuals as a group or on an individual basis. The diet is controlled in its composition (e.g., energy content, protein content) and its mean of administration (e.g., solid or liquid). Often, the ration is also administered in a way that does not allow the animal to select its components.

Example: Environment and Diet Composition (ExA)

For example, pigs' rations are ground and/or pelleted so that the animals cannot choose between the corn and soybean that compose it. Therefore, the diet of a pig kept in indoor conditions is fully controllable, at least to the extent that the farm manager aimed at controlling. Let us assume that the ration of some growing pigs is formulated based on its energy and protein content, but its aminoacidic composition is not considered. A trial could be run where two diets are fed to groups of growing pigs, and their daily weight gain compared. The feed is also made with different batches of corn and soybean that vary in

their aminoacidic composition. Animals in each replicate of the study receive feed from different batches. The phenotypic variation between the groups in the trial could be reconducted to the controllable environmental variation component E_c because such variation was created in a *controlled* fashion. However, the different replicates received feed differing in aminoacidic composition, which inevitably affected their growth rate. The phenotypic variation between the replicates and within diet could be considered due to the *not controllable* environmental variation component e_r because it was not planned that the different batches received different diets for such parameters.

In this case, the E_c and e_r components could be directly attributed specific parameters that defined the diet, such as energy and protein content versus the content of specific amino acids. In reality, the same environmental parameter could show controllable and not controllable components.

Example: Environment and Sow Housing (ExB)

We could think of sows housed individually in farrowing pens, which in turn are allocated in different building rooms. The facility manager can monitor and control the temperature in each room with a dedicated thermostat located at the center of the room. However, each room has a window on only one side. The sows housed in the pens next to the window will experience higher temperatures during the day when the sunlight hits the window glass and lower temperatures at night since the windows cannot insulate as well as the walls do. Conversely, the sows in the other parts of the room will experience more stable conditions. If the breeder wanted to use the temperature as an environmental covariate in some heat tolerance study, a thermometer located in the middle of the room would only provide an average of the room conditions. But, ultimately, the pen-specific temperature affects the welfare of the sows, not the average room temperature. Using the average room temperature, the animal breeder would find some unexplained phenotypic variability, determined by the unaccounted environmental variance. While the average temperature in each room falls in the E_c component, being

controlled by the farm manager, the deviation between the single-pen temperature and the room average falls within the e_r component, given this is not monitored nor controlled by the farm manager.

The definition of E_c and e_r clearly implies a hierarchy of the components [9]: the micro-environmental component is nested within the macro-environmental component, both conceptually and statistically. The uncontrollable environmental changes in the micro-environment are meant to happen within the controllable macro-environment. This limitation is often convenient because it implies that the micro-environmental variation occurs within an individual and within a macro-environment, which makes its interpretation easier.

This second example (ExB) should be further expanded to mention the potential presence of covariance between E_c and e_r . Hypothetically, such covariance could be null: the average room temperature in a given day is not related to the range in temperature within the room, during that day, which results in the two effects being uncorrelated. However, one could also hypothesize that the range within the room (from closer to farther from the window) could depend on the average room temperature, having both these parameters a common source of variation which is the outside weather. For example, the sows closer to the windows (as compared to those farther from the windows) could suffer from cold stress more likely during a cold night than a warm day. Likewise, the same sows could suffer from heat stress more likely during a warm sunny day. Such covariance, inherent to the housing facility, could drive the environmental sources of variation to be correlated resulting in the two components being collinear. Such collinearity would be difficult to model and few attempts have been made [10, 8]. For the sake of simplicity, this covariance will be neglected in this entry; nonetheless it should be kept in mind of its potential presence. Ultimately, the dissection of E into E_c and e_r depends on the quality and completeness of the data as well as the modeling strategy (e.g., availability and use of average room temperature or pen-wise temperature). Consequently, the

covariance between the two terms will depend on the same factors.

The dissection of the environmental variance is an essential component for the optimal modeling of the $G \times E$. This concept will be further expanded in the following sections of this entry.

3. *The Breeding Value(s) in Relation to the Environmental Components*

Both formulas (1, 2) include a genetic component G . Such a component is marginal on the environment, which means that it will not depend on the environmental conditions experienced by an individual. When estimated breeding values are calculated and published, those are meant to be a deviation from the contemporary group (i.e., a set of individuals undergoing same macro-environmental conditions) average. Such definition acknowledges the presence of environmental variation (absorbed by the contemporary group effect) and assumes that such deviations due to G are the same under every condition. For example, dairy cows that are daughters of a given bull will show different yields in different farms. Still, the expectation of their deviation from the contemporary group will be the same, for example, 3 kilograms of milk above the group mates.

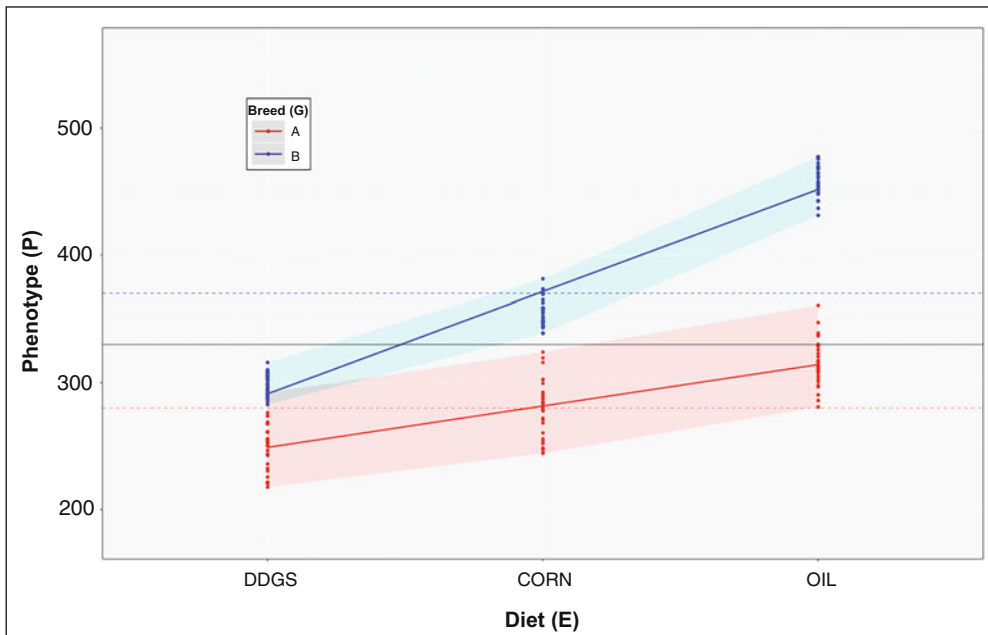
The $G \times E_c$ component, as defined in formula (2), is the most valuable for the animal breeder. Given the controllable environmental variation and the known parameters that describe it (e.g., energy content in the ration, average room temperature), the $G \times E_c$ component can be easily modeled with several statistical tools. In general, every model will yield a prediction of the BV *conditionally* on the environmental conditions, that is, depending on how such conditions vary. The deviation from the contemporary group mean, used as an example in the previous paragraph, now depends on the contemporary group itself. Practically, the dairy cow daughters of a given bull will show a deviation from the group mates that depends on the group itself: it will be (hypothetically) a +3 kilograms of milk in group A but a -0.5 kilograms of milk in group B.

The different performance due to $G \times E_c$ can probably be better understood going from the

individual to the breed level. Let us assume that the trial in example ExA (section 2) is run on pigs of different breeds, namely A and B. The growth rate of such pigs heavily depends on the breed, such that breed A pigs will show slower growth rates as compared to the breed B pigs. Being the average growth rate (of the pooled breeds) of 330 grams per day, the A and B pigs will show deviations of +40 and -50 grams per day, respectively. However, the market price of corn suddenly increases, and the farmer needs to find an alternative ingredient. While some ingredients could be more expensive, others will be more economical and the farmer decides to venture into the use of different feed ingredients hoping that, on average, the animals will show the same performance as if they were fed corn. The farmer creates therefore three rations: one with corn (CORN), one with distiller's dry grains (DDGS, lower energy content than corn), and one with vegetable oil (OIL, higher energy content than corn). The farmer feeds the three diets to both breeds present on the farm. A schematic representation of such experiment is reported in Fig. 1.

When the ration includes corn as the main energy sources, breed B pigs show a significantly faster growth rate, as pictured in the middle of the plot (CORN). When moved into a more energetic diet (OIL), both breeds show faster growth, but the advantage of breed B over breed A pigs becomes even larger. Probably, the selection history of breed B gives it a better growth potential which is not fully expressed on a corn diet but increases the gap between the two breeds when more energy is available. When the ration includes DDGS, both breeds slow down their growth and the difference between the breeds becomes smaller.

The farmer by testing the impact of breed, diet, and their interaction on the growth of the pigs has completed an experimental design that allows the dissection of the several components of formula (2). First, average growth of the individuals is represented by the horizontal black solid line. The G component is represented by the dashed red and blue lines, or actually by the difference (in response y) between the black line and each line. This difference is systematic across the diets



Genotype by Environment Interactions in Livestock Farming, Fig. 1 An idealized scenario for the dissection of GXE

(environments) and can be considered as the impact of the genetic background of the different breeds *marginally* on the environment. Likewise, the difference between the diets (not indicated by lines) can be considered as the E_c component, being purely environmental and *marginal* on the breed.

The $G \times E_c$ component is conceptually explained by 1) the fact that the difference between breeds changes depending on the diet, making such different *conditional* on the E_c (diet) components, and 2) the solid lines that interpolate the average of the breed by diet blocks, which show different slope. Such lines are considered reaction norms [4, 11], as they report what is the norm of the reaction of each genotype for going from an environment to another (increasing or decreasing dietary energy content). At the same time, the presence of $G \times E_c$ is proven by the fact the different diets show an impact on growth rate that depends on the breed. While it could be remunerative to replace DDGS with CORN or even with OIL for breed A, it will not be the case for breed B. Again, the effect of the diet (E_c) is *conditional* on the breed. Using a simple

example, it can be demonstrated how the three main components can be disentangled and interpreted at least from a zootechnical standpoint.

The $G \times e_r$ component, as introduced in (2), is probably the least intuitive of the components. For this component to be fully understood, the definition of e_r needs to be revisited in the modeling context. As mentioned in section 2 of this entry, the e_r component is the one that determines the phenotype due to random, uncontrollable environmental conditions. As such, these environmental conditions are unaccounted by the model, thus normally considered as the (statistical) residual error of the model. However, the presence of a component defined as $G \times e_r$ implies that the error will somehow be conditional on the genotype. In fact, the simplest definition of $G \times e_r$ could be that the genotype determines the residual error, that is, the phenotype uniformity. Going back to the example in Fig. 1, each point represents an individual from a given breed that was fed a given ration. It can be noted that the points for breed A show larger variation around the mean than the points from breed B, under any diet. This simple

example shows how the genotype can impact the dispersion of phenotypic measures beyond the variation due to the macro-environmental component E_c . Such dispersion is measured as the deviation of each phenotypic measure from the expectation of the group, that is, from the average of each breed by diet combination, which brings back the conceptual hierarchy mentioned in section 2 among the macro- and micro-environmental components. It should be mentioned though that the within-genotype phenotypic variation could also be determined by the within-family genomic variation, given by different gametic disequilibrium among the families [9]. However, for the sake of clarity, the within-family variation that is not accounted by explicit and controllable environmental variation should only be determined by the environmental components, in other words, the replicated observations of a genotype should be considered as replicates of a clone. While unrealistic, this will help the interpretation of the different components discussed in this entry. The $G \times e_r$ component expresses the ability of each genotype to withstand random environmental variations, that is, micro-environmental plasticity.

It should be stressed how the difference between macro- and micro-environmental plasticity depends on the availability and use of more or less precise environmental data. If the variables that define the macro-environment (and determine E_c) cannot be measured, the e_r component will absorb E_c and $G \times e_r$ (if modeled) will absorb $G \times E_c$. Using again Fig. 1 and assuming the ignorance about different diets, breed B would show larger within-genotype variation than the breed A. This difference in within-breed variance could be attributed to the breed-specific micro-environmental plasticity ($G \times e_r$). When diet information is made available though, part of the $G \times e_r$ will be absorbed by the macro-environmental component ($G \times E_c$). Again, the modeling of the two components implies a trade-off which depends mostly on the availability and use of the environmental covariates.

4. Variables Defining the Environmental Components

The environmental component of the models described requires descriptors that fully define the complexity of the environmental conditions. This is no different than the quality of genomic markers that are used in genomic selection. Markers are chosen to represent the physical variability of the genome together with their relevance for the traits of interest [12, 13]. At the same time, the information conveyed by the markers should not be redundant or biased to avoid model overparameterization or overfitting. Just as the genomic markers seldom account for the whole genetic variability (e.g., missing heritability), the environmental variables will not likely account for the entire environmental variation. This concept is central in modeling $G \times E$ and will be reintroduced at the end of this section.

Environmental variables can be either categorical or continuous [14, 15]. When modeling the diet, the different rations will (at least here) be considered as categorical variables describing discrete conditions. The assumption is that the variability between the diets is maximized while there is no (environmental) variation within the diets. Each diet is a standalone condition, which implies that each diet could have its peculiar conditions. In the examples ExA and ExB above (section 2), the categorical environmental variables would be the ration or the room. The environmental variables could also be continuous and describe conditions that vary within some interval. In this case, it is impossible to contrast different conditions straightforwardly, but rather the change in phenotype is described given a unit change in the environmental covariate. A single covariate only includes one dimension, but nonlinear modeling is possible. The change in the environmental covariate could lead to a different phenotype shift depending on the interval in the covariate itself (see quadratic trends). In the examples above, the continuous environmental variables could be the room or pen temperature or the energy and protein content in the ration. Several continuous variables that can describe categorical conditions lead to the possibility of having several continuous variables that describe the continuous variation. The room climatic conditions could be characterized by

daytime and nighttime temperatures, and the different rations are characterized by both energy and protein content. In this case, all the variables should be potentially considered in the model, but some will likely be more relevant than others. For example, heat stress for the sows (ExB) could result from a high daytime temperature or lack of night cooling. A good model should discern between the two to provide the basis of reasonable inference about the biology of the trait(s).

Nonetheless, the number of environmental variables could be large (hundreds, sometimes thousands), but their dimensionality could be limited. Again, this is no different from genomic markers that show low effective dimensionality because of linkage disequilibrium. In the same way, the environmental variables could have low effective dimensionality because of their intrinsic collinearity or the limited sample size. In the first case, daytime and nighttime temperatures could show some degree of collinearity due to common seasonal trends (ExB). In contrast, the energy and protein contents could also show collinearity because of some ration formulation constraints (ExA).

The kind of environmental covariate will strongly affect the choice of the model used to analyze the trait. Consequently, the nature of the environmental covariate will determine how the $G \times E_c$ component can be interpreted.

In the example ExB of the sows housed in the different pens of the room(s), the same environmental variable (temperature) had a portion that could be considered E_c and one that could be regarded as e_r . Again, the difference between the two components depended on the ability to be controlled in an experiment or monitored in the case of field data collection. Pen-specific thermometers would allow the E_c component to absorb e_r .

The breeder's goal is to design experiments or data collection protocols that maximize E_c and minimize e_r by having a more detailed description of the environmental conditions. While the purpose of this entry is not to emphasize the importance of precision livestock farming, it

should be evident how precise (but also accurate) data collection can significantly affect the quality of the results. Going back to the sows' example (B), more precision in defining the categorical environmental covariate would move the contemporary group to be defined as the pen rather than the room. This, in turn, could have the unwanted consequence of making data granular to a point that hamper the statistical learning for the model. Contiguous pen groups could have a low number of individuals allocated, and consequently, the standard errors for the group estimates would be larger. Suppose the pen temperature was used instead of the room temperature, in that case, it could happen that individuals were allocated in different intervals of temperature, with the experimental design potentially unbalanced (see section 9 of this entry). Ultimately, the optimal E_c and $G \times E_c$ effects modeling is a trade-off between goodness of fit and efficient parameterization.

5. *Modeling the Components and Interpretation of the Estimates: Part I*

Once the experiment or the field data collection is carried out, and the environmental variables are defined, the choice of the model will fall within a limited number of options. However, this should not be a trivial choice, and literature is vast in comparing statistical models given a fixed set of data [16, 17, 18].

The modeling of the E_c component is simple and is not different than, for example, an ANOVA aimed at dissecting the variance components following a complete randomized-blocks experimental design.

Categorical environmental variables can be included in the model as (fixed or random) cross-classified effects and contrasts can be obtained. Continuous variables can be included as (multiple) linear or polynomial covariates.

First, ANOVA could be used in the case of a balanced experimental design that includes specific breeds, lines, or families. In the simplest case, where categorical variables describe discreet environmental conditions, the model could be so defined as:

$$y = B + T + BT + e \quad (3)$$

where y is the phenotypes of interest, B is the breed component (or family, or genetic line, to be considered as modeling G), T is the treatment component (to be considered as modeling E_c), BT is the interaction between the two components (as modeling the $G \times E_c$), and e is the random residual component (which could absorb both the e_r and $G \times e_r$). In this case, all the components are disentangled using cross-classified effect, which include multiple levels of breed and treatments. When all the breeds are allocated over all the treatments, the phenotypic variance can be easily decomposed into the two main effects and their interaction. From a breeding standpoint, the breed estimates represent the germplasm's performance under average conditions. In contrast, the estimates for the BT component will describe the performance under specific conditions. At the same time, the solutions for the T treatments will represent the impact of such treatment over all the genetic material, while the solutions for the BT effect will represent how the treatments affect each breed specifically. If the design is not complete, the model will not be able to estimate the interaction component with the consequence of not predicting the performance of a given breed under a given treatment if this combination is not present in the design. The prediction of y to unobserved conditions could be obtained by replacing the categorical descriptors with continuous descriptors, that is, replacing the cross-classified effects with covariates. In this case, the model becomes:

$$y = B + t + Bt + e \quad (4)$$

where t is the treatment component as defined with a set of environmental covariates that describe and Bt is the interaction component modeled as a set of covariates for each breed. Here, the model would give regression coefficient estimates for all breeds (t) or specific for each breed (Bt). A practical example could be example A (section 3 and Fig. 1).

In comparison, the single regression coefficient of the t treatment component would show

the average change over the two breeds. Regardless, using the covariates would allow extrapolating each breed's performance to protein contents outside the formulated rations. While the prediction may have low accuracy, it is statistically possible and sets the basis for genomic predictions incorporating $G \times E_c$ effects.

Formula (4) does not allow to predict the performance of new, unobserved breeds (or genotypes) because such levels are not connected by any covariate. The use of pedigree or genomic information allows such prediction, similarly to how the environmental covariates estimate the E_c term. The model could therefore be reformulated as:

$$y = b + t + bt + e \quad (5)$$

where b is the breed component modeled as a set of covariates that describe (and link) the genotypes and bt is the interaction component, modeled as the product between genetic and environmental covariates. The latter term is the least common to find and should be interpreted as the interaction between covariates: the change in y given the unit change in b and the unit change in t . In practical breeding terms, each genomic marker could be assigned a specific regression coefficient for each environmental covariate and vice versa. The graphical representation of the predictions of this model necessarily requires the use of a three axes plot, also known as "surface plot." The advantage of this model is that the different genotypes and the different environments are not standalone items but are connected by covariates that allow transferring information from one level to another. In other words, to predict the performance of a given genotype, the information can be "borrowed" from different genotypes; likewise, the information can flow from a given environmental class to another. Ultimately, the interaction term informs the model of the particular outcome of a given genotype in a given environment.

There could be a fourth way to rearrange the same model, where the breed component is defined by a covariate, but the environmental effect is cross-classified. In formula:

$$y = b + T + bT + e \quad (6)$$

While all the terms are as described above, it should be noted that the interaction component bT is probably the most intuitive of all those described. Here, a genetic effect (e.g., genomic marker effect) is defined for each environment, being the last categorically defined. In other words, the biological model assumes that there could be a different genetic architecture for each of the environments considered, and this is explicitly modeled. This model allows to extrapolate the prediction to new genotypes but not to new environmental classes (Table 1).

6. Modeling the Components and Interpretation of the Estimates: Part II, E_c and $G \times E_c$

Selective breeding has taken strong advantage of the Best Linear Unbiased Prediction (BLUP) methodology implemented in the mixed models. These models allowed the estimation of both contemporary group (E_c) and additive genetic effects (G) simultaneously thanks to a covariance matrix that linked the different individuals (genotypes) based on the expected sharing of alleles. With the advent of inexpensive and high-throughput genotyping platforms, the genomic markers (mainly single nucleotide polymorphism markers) replaced the expectations of covariance with a more precise measure of shared inheritance of alleles. Conveniently, this genetic covariance matrix can be constructed considering markers

Genotype by Environment Interactions in Livestock Farming, Table 1 Recapitulation of the four models presented in formulas 3 to 6 and which models allow phenotype prediction across lines, environments, or both

	Across lines	Across environments	Across lines and environments
$y = B + T + bT + e$.	.	.
$y = B + t + bt + e$.	X	.
$y = b + T + bT + e$	X	.	.
$y = b + t + bt + e$	X	X	X

as covariates. In the case of its most common application, BLUP includes the animal additive genetic effect as cross-classified but connected by several covariates, either observed or unobserved. Such effect is usually fitted as random, where a parameter allows the shrinkage of the solutions to avoid overfitting and regulates the transmission of information through the covariance matrix.

The Baseline Model

The baseline model for genetic evaluations could be considered as follows:

$$y_{ijk} = T_0 + T_i + A_j + e_{ijk} \quad (7)$$

where y_{ijk} is the k th phenotypic record coming from the i th contemporary group, belonging to the j th individual, T_0 indicates the overall treatment effect, otherwise known as intercept in statistical terms, T_i is the cross-classified effect of the i th contemporary group, A_j is the cross-classified genetic effect of the j th individual, and e_{ijk} is the random residual deviation. Alternatively, the cross-classified contemporary group effect could be replaced by more covariates describing the environmental variation. The formula becomes:

$$y_{ijk} = T_0 + \sum_{l=1}^p t_l + A_j + e_{ijk} \quad (8)$$

where t_l is the l th covariate, the number of covariates can range from 1 to p . The presence of the covariate does not prevent the presence of the cross-classified effect, and for the sake of clarity, they will be kept separate in this entry.

These models do not include any interaction between the genotype and the environment effect but represent the most common models used in genetic evaluations.

The Multiple-Trait Model

The first model that should be considered for estimating GxE is the multiple-trait model (MTM). This category of models is widely used in animal breeding when more than one trait enters the selection index. Making use of the covariance existing among traits, this model

can produce breeding values for all individuals' traits, even if some traits are not recorded on the individual(s) of interest.

The MTM can estimate GxE by assuming that the phenotype belongs to a different trait depending on the (categorical) environmental condition of recording (different environments are different traits). Using the example of the pigs fed different diets (example A, section 2 of this entry), growth would be considered a different trait depending on the diet provided to the individuals (where the diet is the environment E_c). The model could be formulated as:

$$y_{ijk} = T_i + A_{ij} + e_{ijk} \quad (9)$$

where the term T_0 (the overall intercept) does not appear but is replaced by T_i , which is the trait-specific intercept. Likewise, the term A_j is replaced by a trait-specific genetic effect A_{ij} , that is, each individual is assigned a breeding value for the traits considered. The terms T_i and A_j are trait-specific but, in this case, also environment-specific. The possibility to transfer information across individuals is given by using a covariance structure imposed to the model via a matrix built on the pedigree or the genomic markers. Likewise, the possibility to transfer information across traits is given by a covariance structure that is imposed on the model via a variance-covariance matrix among the traits [19]. Considered together, the model allows transferring information across traits and individuals.

The MTM requires a categorical (discrete) definition of E_c , assuming that each environment is distinct and independent. The trait-specific intercept absorbs the difference between the categorical environments (i.e., E_c). The trait-specific breeding values absorb the variance of both the G and $G \times E_c$ components. The MTM does not provide solutions for the overall G component producing a breeding value estimation for each environment instead. However, since a non-null covariance is generally imposed on the calculation of the vectors of breeding values, the overall G component is blended into the single-environment components. In other words, the MTM cannot explicitly disentangle the G and $G \times E_c$ components.

The Random-Regression Model

A category of models widely used in estimating $G \times E_c$ are the random regression model (RRM), which were initially developed in livestock to model lactation curves [20]. An RRM involves defining the environmental component as a continuous gradient and fitting it in the model as a covariate. The general model is an expansion of the model in (6) and can be expressed as follows:

$$y_{ijk} = T_0 + t_i + A_{0j} + A_{1j} + e_{ijk} \quad (10)$$

As in (7), T_0 indicates the overall treatment effect, t_i the environmental effect E_c modeled as a covariate, A_{0j} is the j th-individual-wise intercept (G), A_{1j} is the j th-individual-wise effects over for the (nested) covariate ($G \times E_c$). The latter two terms could be counterintuitive but deserve attention.

First, it should be highlighted that T_0 and t_1 are population effects, as they are estimated for the whole sample of individuals included in the analysis. They express the value of the phenotype when all covariates have zero value (intercept) and the change in phenotype given a unit change in the covariate, respectively. Their interpretation is no different than the one from a regular regression model.

The terms A_{0j} and A_{1j} are individual effects, and they express the genetic value of the individual when the covariate has zero value and the change in genetic value for the individual given a unit change in the covariate. The term A_{0j} is conceptually (but not statistically) equivalent to the term A_j in formula (7), while the term A_{1j} is specific to the RRM. Such a term is, in fact, a regression coefficient that uses a covariate and is fitted as a nested random effect. Such a covariate is nested within the animal genetic effect and is fitted as random to limit overfitting. Biologically, the A_{1j} term represents the reaction norm as defined by Falconer [4]. Conceptually, the reaction norm shows the individual's reaction in the phenotype given a unit change in the environmental condition E_c .

The RRM requires a definition of the environmental component E_c as continuous. This

component is then absorbed by the model term t_1 . The G component is expressed by the term A_{0j} , whereas the $G \times E_c$ component is defined by the term A_{1j} . It should be noted that the RRM provides explicit modeling of the $G \times E_c$ component, which the MTM did not provide. The solutions for such a term express how an individual reacts to environmental changes compared to the population average reaction (t_1). Such expressions can be instrumental in the study of biology and the implementation of breeding programs. The individual-specific reaction indicates whether individuals are plastic or resilient to environmental changes. For a further comparison on the estimation of variance components for $G \times E_c$ reading Calus et al. [21] and Fikse et al. [18] is suggested.

The Reproducing Kernel Hilbert Spaces Regression Model

Both MTM and RRM provide an estimation of all G, E_c , and $G \times E_c$ components. Still, they have limitations in the number of environmental covariates that are included in the model. As mentioned in the previous sections, the number of environmental covariates describing E_c can be larger than 1, potentially assuming large values despite the limitation in their effective dimensionality given by collinearity. The RRM can explicitly model the genotype's reaction norms over each environmental covariate, but when the number of covariates becomes large (e.g., larger than 2), the model performance deteriorates in both the ability of estimating the model parameters as well as in the interpretability of the results. Ignoring the need to model the reaction norms over each covariate explicitly, multidimensional environmental data can be incorporated using Reproducing Kernel Hilbert Spaces (RKHS) regression. This class of models takes advantage of "kernels" that impose a covariance structure over the observations in a dataset. Animal breeders should be familiar with this type of model since BLUP solutions for breeding values are also based on a kernel (the relationship matrix among individuals). In fact, the BLUP animal model is a special case of RKHS [22].

Similar to using genomic markers (as covariates) to build a genomic kernel,

environmental covariates can be used to construct an environmental kernel. The latter can be used to let information flow among contemporary groups, in the same way that genomic marker information connects the individuals. The contemporary groups will be linked conditionally on the environmental covariates used: recalling the example (A) of the different diets fed to the pigs and assuming a larger number of diets tested, the different groups fed different diets will have different covariance whether they are linked based on energy and protein content alone or through a deeper characterization of the diets (e.g., fiber, amino acids, minerals).

Given a set of environmental covariates, the model can be formulated as:

$$y_{ijk} = T_0 + \sum_{l=1}^p t_l + A_{0j} + \sum_{l=1}^p A_{lj} + e_{ijk} \quad (11)$$

where T_0 and A_{0j} are as defined above, t_l is now a set p of fixed population effects based on covariates, and A_{lj} is now a set p of random individual effects based on these covariates. Because of high dimensionality (p) of A_{lj} and, potentially, also t_l ; such covariance can be modeled using the kernels constructed on the same set of covariates and the genomic markers.

7. Modeling the Components and Interpretation of the Estimates. Part III, e_p , and $G \times E_c$

The models presented in sections 5 and 6 assume the presence of a set of environmental covariates describing the environmental variation. Using the terminology introduced in section 2, these environmental covariates explain the macro-environmental changes that the individuals experience. In absence of such covariates, the animal breeder has the only option to model the variability that is found among the observations of a given genotype once all the other effects are accounted for. In its simplest modeling, the within-genotype variance is extracted from the residuals of any statistical analysis. Since all the models in section 6 include the residual term e_{ijk} , the necessary steps will involve calculating the

variance of the residuals for each individual. In fact, the residuals will contain the variation that is not accounted by the model and, potentially, absorbed by the G , E_c and $G \times E_c$ components. The $G \times e_r$ component will be considered as the differential ability, of each genotype, to determine the residual variation [9]. In other words, how the individual can maintain the uniformity of its phenotypes beyond the variation induced from the other components.

The analysis of the individual variance generated by the residuals could present some statistical challenges. First, if the residuals e_{ijk} are truly normally distributed, their squared values should be chi-square distributed. Since this could generate several issues in modeling, some data transformation procedures have been proposed [3]. Second, the repeated records of an individual could show some degree of autocorrelation, with consequent reduction of the measured variance, at least as expressed in (12). Therefore, covariance among the observations should be accounted for and some autoregressive structures among the residuals have been proposed [23].

Finally, e_{ijk} depends on the quality of the estimates for the other components. Lack in precision for such estimates would introduce noise in the residuals. While such noise cannot be removed with certainty, it can be definitely minimized with the appropriate modeling. Double hierarchical generalized linear modeling has been proposed as a class of structural models capable of estimating the components determining the expectation and the variance of all the factors in the model [24]. While such models present an enormous potential for modeling $G \times e_r$, they will not be further discussed in this entry but the reading of Rönnegård et al. [24, 25, 26] is encouraged.

In general, modeling the $G \times e_r$ component, in absence of environmental descriptors, is not a trivial task and needs particular attention.

8. Studies That Aimed at Estimating GxE for Economically Relevant Traits in Livestock

The literature of modeling $G \times E_c$ and $G \times e_r$ is vast, especially since the topic has gained more relevance during the last decade. The advent of

genomic selection has in fact made more feasible the estimation of these components given its ability to provide a stronger link between individuals, with consequent improved transfer of information across environments [15].

Dairy cattle is the livestock sector where this subject has been more frequently studied. One possible explanation for this is related to the large use of artificial insemination that allows the dissemination of germplasm across environments, with consequent strong genetic connectedness across environments through the paternal half-sib families. One of the most relevant applications of the study of $G \times E_c$ is when the environmental conditions are considered as the different countries, as was proposed by Fikse et al. [17], who studied across-country GxE for production traits in Guernsey cattle.

Farming system can also provide an alternative definition of the macro-environmental component. Within this space, organic farms vs pasture-based vs confined systems appear to be the first obvious distinction. Van Pelt et al. [27] studied macro-environmental plasticity of dairy cows for several traits when the environment was defined as the “use of grazing,” while Liu et al. [28], Pfeiffer et al. [29], and Shabalina et al. [30] used the organic certification as environmental condition, in Danish Holstein, Austrian Fleckvieh cattle, and German Holstein, respectively.

Similarly, Gerber et al. [31], Huquet et al. [32], Calus et al. [21], and Schmid et al. [33] studied macro-environmental plasticity for production traits in German Simmental, French Holstein, Dutch Holstein Friesian cattle, and German Brown Swiss cattle, defining the environmental conditions with different management parameters. Similarly, Santos et al. [34] and Mulim et al. [35, 36] studied production traits in Brazilian Holstein, and Bohlouli et al. [37] studied production traits as well as milk fat composition in German Holstein.

The genetic control of heat tolerance has largely justified the study of its $G \times E_c$. Ravagnolo et al. [38] conducted one of the first studies for tolerance to heat stress in US dairy cattle, Carrara et al. [39] and Negri et al. [40]

studied heat tolerance for production traits in Brazilian Holstein, while Ansari-Mahyari et al. [41] studied fertility in Iranian Holstein.

Nguyen et al. [42] and Cheruiyot et al. [43] worked toward the implementation of genomic selection for heat tolerance in Australian dairy cattle, so far being the only country that has implemented selection for heat tolerance in dairy cattle.

Fertility traits have also been subject of study. Strandberg et al. [44] studied fertility in UK Holstein, while Ismael et al. [45] worked on fertility traits in Danish Holstein, defining the environmental condition as the herd production level.

Finally, Lassen and Mark [46] studied conformation and workability traits defining the environmental conditions using herd housing.

The use of the RKHS model while popular in plants has been limited in livestock, but Tiezzi et al. [47] studied production traits in US Holstein cattle using this model, testing different sets of environmental covariates, from climatic, to herd management to geographical location.

Micro-environmental plasticity has not received the same attention. Initially, Clay et al. [48] were the first to study the genetic basis of the different within-family residual variance. Later on, Rönnegård et al. [49] studied micro-environmental plasticity for milk yield and somatic cell score in Swedish dairy cattle.

Environmental plasticity has also been studied in beef cattle. Paulo et al. [50] studied age at first calving in Italian Limousine and Charolais, Cardoso and Tempelman [51] studied post-weaning growth in Brazilian Angus using reaction norms on the contemporary group solutions, with the same model Silva et al. [52] studied reproductive traits in Nellore cattle and found non-null GxE for scrotal circumference and gestation length. As opposed to the study of the genetic basis of tolerance to heat stress in dairy cattle, Toghiani et al. [53] studied cold tolerance in composite beef breeds.

Again, micro-environmental plasticity has received less attention, but Neves et al. [54, 55] studied it in Nellore cattle and found non-null heritability estimates for birth weight and health scores.

Other studies on $G \times E_c$ can be found in swine. Silva et al. [15], Chen et al. [56], and Song et al. [57] studied productive and reproductive traits using the contemporary group solutions as environmental covariate. The necessity of improving heat tolerance of pigs has led to studies of $G \times E_c$ in this species. Zumbach et al. [58], Fragomeni et al. [59], Tiezzi et al. [60], and Usala et al. [61] studied the genetic components of heat tolerance on different traits.

Micro-environmental plasticity has also been studied in different minor livestock species. Mulder et al. [62] studied micro-environmental plasticity in broilers, Berghof et al. [63] studied layer chickens, Gutierrez et al. [64] studied it in mice, and Garreau et al. [65] studied it in rabbit.

Some methodological studies have to be mentioned. Since the definition of the environmental covariate is crucial in modeling $G \times E_c$, its choice will also impose some limitations in the potential outcome of the analysis. The covariate(s) could be estimated from the data itself using a two-step approach as in Silva et al. [15] in swine and Mota et al. [66] when studying tick resistance in Hereford and Bradford breeds. However, this has the limitation of being a “fixed” covariate, whose values are defined a priori when estimating the variance components for $G \times E_c$.

Su et al. [67] proposed a method where the environmental covariate could be directly estimated from the data, in a single pass, using MCMC sampling of the covariate value itself. This method has been applied several times [68, 69, 70] and appears to be promising in understanding the complex determination of the $G \times E_c$.

9. Effective Use of GxE in Livestock Breeding Programs

Livestock breeding programs have changed dramatically over the last century and further changes are coming as new technology and needs develop. While a relevant improvement was brought forward by the use of BLUP, new models will need to provide more precise breeding values. As an example, these improved breeding values could couple the G and the $G \times E_c$ components in order to tailor the breeding values

to specific management conditions [71, 72]. In support of this, just looking at the increases in global temperatures or the fluctuations of commodities prices illustrate that livestock, in less than 20 years, will be fed a different diet and will experience different ambient temperatures than today. These changes will be more or less gradual, but the animal breeder needs to be ready to provide precise breeding values in a timely fashion to improve animal welfare and farm productivity [73, 74, 75, 76, 77].

While the modeling of the macro- and micro-environmental components has been widely discussed in the previous sections of this entry, this section will focus on how these breeding values could be reported and published, what could be the limitations in their use, and how selection could take place once this information is made available.

How to Express and Report the Breeding Values

If the macro-environmental component is taken into account, breeding values can be defined as $G + G \times E_c$. These breeding values can in turn be generated using either MTM, RRM, or RKHS. These different implementations can provide different predictive abilities depending on the data structure and the quality and number of environmental covariates (see section 6 of this entry). MTM models will explicitly generate breeding values for each categorical environmental condition, facilitating their use [14]. In contrast, the RRM will generate breeding values for the intercept term (genetic value of an individual when the environmental covariate is zero) and for the slope term (change in genetic value of an individual per unit change in the environmental covariate, also known as reaction norm). Both these components will be centered to the population average. Consequently, breeding values for the slope term will be interpreted as a “resilience” parameter [17, 18].

If we assume that an environmental covariate is (or is correlated to) a measure of hypothetical stress, the model will estimate a population slope (t_1 in (10)), expressing the degree by which the population (on average) “reacts” to that stressor. RRM models will additionally produce individual

breeding values (A_{1j} in (10)) predicting how an individual will react to the environmental challenge, conditional on its genetic background. The breeding value will be adjusted by the population average, such that individuals with a negative value will have a stronger change when the environmental covariate increases (or decreases) in value, while individuals with positive values will show a weaker change. While in the former case, individuals will be considered more *plastic* (i.e., the environment more strongly shapes their phenotype), the latter case will identify the more *resilient* or *robust* individuals (i.e., whose phenotype is less affected by the environment). Importantly since we are talking about breeding values, plasticity resilience characteristics will be passed to the offspring, to the extent that this parameter is heritable. Whether these characteristics will be desirable or not in a livestock breeding scenario will be dependent largely on breeding objectives, as well as industry structure. The RRM will also generate breeding values for specific environmental conditions. Notably, while conditions might be defined by continuous covariates, in order to facilitate their use, it is likely that these will still be discretized in a second step.

In example ExA (also reported in Fig. 1), an RRM could be implemented using the energy content in the ration as environmental covariate, yet the breeding value could be extrapolated to discrete values of such environmental parameter to mimic an MTM. In this case, RRM could be advantageous compared to the MTM in situations in which the number of macro-environmental conditions is large and can be defined by some continuous covariate. In example ExB, the breeding values for heat tolerance could be expressed as those under hot, uncomfortable conditions simply by extrapolating the prediction under the high-temperature conditions.

RKHS are currently probably the least popular models to obtain breeding values for $G \times E_c$ in livestock. Nonetheless, their popularity is growing, and they can be advantageous when many environmental covariates are measurable [78]. Notably, RKHS can provide estimates of breeding value conditional on the environmental conditions. These predictions can be produced for all

combinations of environmental factors, even if not observed in the training dataset. Breeders could then optimize the use of germplasm under complex experimental designs or farming scenarios.

Caveats When Estimating Breeding Values for G×E

In the previous sections of this entry, a number of approaches used in estimating $G \times E$ have been described. Each of these models has pros and cons that are specific. All models described can perform poorly or produce misleading results, regardless of the statistical machinery employed. In most cases, these pitfalls are the result of hidden data structure not appropriately accounted for in the analysis. The estimation of the $G \times E_c$ component suffers the most when the experimental design (or the stratification in the field data collection) introduces artificial collinearity between the G and the E_c components. This collinearity will inevitably affect the estimation of the interaction among these components.

Macro-environmental Plasticity: Estimation of Variance Components

Considering the dataset pictured in Fig. 1, it can be safely assumed that the experimental design is balanced among the G and E_c components, that is, individuals from all breeds received all the diets, in equal numbers. In this case, a non-null estimation of the $G \times E_c$ component could be safely considered as robust. Nonetheless, a balanced design is necessary also for the proper estimation of the G and E_c components. The combination of breed and diet represents a “block,” that is, a group of individuals from the same breed (G) that received the same treatment (E_c). In a hypothetical unbalanced design, each breed might be fed a ration preferably, which is normally called *preferential treatment*. For instance, assume the least performing breed A was fed the DDGS ration and some of the CORN ration, while breed B was fed mostly the OIL ration and some of the CORN ration. While both breeds were fed the CORN diet and some “connection” was maintained, individuals from each breed were

preferentially treated. This could be the case where the farmer preferred to optimize the nutrition plan by tailoring it to the breed’s expected growth: the more “promising” breed was fed the more energetic diet, and vice versa. While this is unlikely to happen for an experimental design, it is actually frequently observed in studies from commercial livestock facilities.

The scenario pictured above would introduce collinearity between the genetic (e.g., breed) and environmental (e.g., diet) factors as a result of an artifact. This collinearity could be practically expressed if extracting a statistical unit from the OIL diet group will more probably be from breed B, while a statistical unit from the DDGS diet group will more probably be from breed A. In that case (again, the connection was maintained on the CORN diet) the estimation of the $G \times E_c$ component would be highly compromised, and the estimation of the G and E_c components would be compromised as well, regardless of the model employed (3) or (4). Specifically, in model (3) obtaining BT estimates for some blocks is not possible (e.g., breed A on OIL diet), which would become an evident limit of the model. In contrast, when using model (4), the breed-specific reaction norms coming from the element Bt would be estimated but with a large degree of error and the investigator would be required to interpret results extremely carefully.

Let us further hypothesize a less evident case of unbalanced design, in which all the breeds would receive all rations, but the number of individuals within each (breed by diet) block would be different. For example, the same number of individuals was initially allocated per group but the different mortality and morbidity for each block made them unbalanced in frequency, such that individuals from breed A could not thrive on the OIL diet while individuals from breed B could not thrive on the DDGS diet. Again, while the G and E_c components would be estimable with little distortion included, the $G \times E_c$ component would be compromised because of the collinearity mentioned introduced by the inherent structure of the experiment. The model (3) would provide BT estimates for the underrepresented blocks that

have larger standard errors. The model (4) would mask the bias in the estimates for the Bt slopes, since the most represented blocks would have a stronger impact than the other ones on the slope of the regression. Yet, such bias introduced in the Bt estimates would be hardly detectable, if not with a dedicated bootstrapping analysis or analysis of the within-block residuals. This latter case is probably the most frequent when breeding values for the $G \times E_c$ component is estimated even with robust methods. The model would produce variance components estimation as well as breeding values that do not show any particular issue, but the variance components for $G \times E_c$ would be inflated or deflated due to the artificial covariance between G and E_c . While the detection and estimation of such bias is beyond the scope of this entry, it is important to mention that bias would likely result in model overfitting and poor predictive ability.

Macro-environmental Plasticity: Prediction of Breeding Values

The ultimate goal of statistical modeling in animal breeding is to predict breeding values for selection candidates. If a $G \times E_c$ component is included, predictions are produced for a reasonable interval of environmental condition. Starting once more from the example in Fig. 1, breeding values will be obtained for individuals from both breeds over the three possible diets. As mentioned in the previous paragraph, such prediction would be obtained just like variance components assuming a valid design. In the case of unbalanced design when some breeds are not fed a specific diet breeding values prediction under those unobserved conditions would still be produced but would require a cautious interpretation. In this case the correct model would be conceptually similar to the one in (5), with covariates linking the categorical levels of breed/individual and diet so that the information can be transferred across breeds/individuals and across environmental conditions. In practical terms, MTM, RRM, and RKHS could all be used to obtain these predictions. However, in the absence of observations for breed A on the OIL diet, the estimation of the breed reaction norms from Bt (4) would be biased. This would lead to poor predictive ability.

In the case where all the breeds are fed all the diets but with different frequency per block, the model would provide better prediction of breeding values for each block. Yet, within-environment breeding value prediction would still be hampered by the experimental design, since bias would still be introduced in the solutions. In the example where morbidity and mortality were the factors that altered the structure of the design, it should be checked whether these factors occurred at random or were stratified over the breeds, diets, or both. In the case that they occurred as random, each block would present an equal number of “losses” (i.e., individuals whose phenotype could not be observed) and the designed would not be particularly altered. In the most unfortunate case where these losses might have occurred mostly within few blocks, and the breeder should proceed with extreme caution. If, for example, breed A showed more losses under the OIL diet because of excessive energy content and breed B showed more losses under diet DDGS for poor energy content, the experimental design would no longer provide a fair assessment of the performance of the two breeds over the three diets. Only the breed A individuals that acclimated to the high energy content were able to express their phenotype (no death or culling). For the same reason only the breed B individuals that acclimated a low energy content showed would have a phenotypic measure.

While in the previous section two extreme cases were presented, the animal breeder should always be cautious when performing $G \times E_c$ statistical analyses using field data, specifically concerning potential sources of bias in the design factors. Other sources of collinearity can be related to the distribution of genetic material in large breeding schemes, which is usually an issue when working with field data. Cao et al. [79] have for example shown out that such across-environment selection and sharing of germplasm can benefit genomic selection schemes that incorporate $G \times E_c$.

Micro-environmental Plasticity and Its Covariance with Macro-environmental Plasticity

In animal breeding, little attention is usually given to the study of micro-environmental plasticity,

consequently in this entry more relevance was given to macro-environmental plasticity. It should be considered that the difference between the two components is often arbitrary and rely for the most part on the ability to model this environmental component. The presence of one or more environmental covariates will allow to explain the within-genotype variation and reconvert it to estimates of one or more reaction norms. Yet, micro-environmental plasticity could be present and could be included in breeding programs.

Despite of the differences in data transformation and model used for this component, the use of breeding values for the $G \times e_r$ could be conceptually reduced to resilience, in a broad sense.

The breeding values for the $G \times e_r$ component will indicate how likely are the observations of an individual to be dispersed around an expected value (the breeding value for the G) component. Individuals with high resilience will also have highly stable and uniform phenotypes, and individuals with low resilience will also have unstable phenotypes. The random environmental variations will impact more individuals from the latter than the former group. These less resilient individuals will show more unpredictable variation because of their reduced ability to maintain homeostasis.

Uniformity is an important component in modern livestock farming systems, which are highly mechanized and need to rely on predictable outcomes. The predictability of the farm flow, meant as a whole from breeding to processing and packing, also allows better allocation of the resources. For example, uniform groups of individuals allow better care from the farm operators, examples can be found in the diet formulation of dairy cows or pen allocations in pigs. Therefore, there is an intrinsic value in having uniform individuals. At the same time, uniformity could be an indicator of robustness, since the individuals that show less variation could also be less perturbed by environmental changes. Last but not least, individuals with more uniform phenotypes could allow more precise culling or breeding decisions. We could make an example that culling depends on an independent threshold such that individuals are culled if not meeting the minimum

requirement and two cows, with same breeding value for the G component, are to be compared. Here, the cow with less uniform phenotypes would have higher chances of being culled because of the larger probability that the shown phenotype would fall below the threshold [9].

Breeding values for resilience (or uniformity, robustness) could be expressed simply as a gradient from a maximum to a minimum. These breeding values will probably gain more importance as livestock farming is becoming more “precise” and will need more predictable individual in order to optimize the inputs and forecast the outputs. Also, more uniform individuals will be needed as livestock farming operations become larger. As the monitoring of the individual welfare and performance will be performed by automated system, uniformity in phenotype (with particular relevance of behavior) will become essential for the successful management of the operation.

Future directions

In this entry a high-level conceptual framework of genotype by environment interactions and its importance in the selection and breeding process has been provided. While this has been a fertile area of research in crop science and plant breeding, the logistics and biological constraints have made $G \times E$ more difficult to study and less applicable in livestock farming. Recently, the development of methods to quantify genetic variability (through SNP panels and sequence information) and environmental variability (through increased technology present in the farm) has made the exploitation of $G \times E$ in livestock more promising. This confluence of technologies is happening at a critical time for agriculture in general, and in livestock specifically. Rising global temperature and extreme conditions are becoming more frequent and concerning. The understanding of genotype by environment interaction will become an increasingly crucial aspect of livestock practice both for granting uniformity of production in the face of more volatile conditions, as well in the ability to disseminate optimized germplasm in specific environments.

References

1. Hazel LN (1943) The genetic basis for constructing selection indexes. *Genetics* 28(6):476–490
2. Henderson CR, Quaas RL (1976) Multiple trait evaluation using relatives' records. *J Anim Sci* 43(6): 1188–1197
3. Hill WG (2014) Applications of population genetics to animal breeding, from Wright, Fisher and Lush to genomic prediction. *Genetics* 196(1):1–16
4. Falconer DS (1952) The problem of environment and selection. *Am Nat* 86(830):293–298
5. Misztal I (2016) Resilience and lessons from studies in genetics of heat stress. *J Anim Sci* 94:197–197
6. Jinks JL, Pooni HS (1988) The genetic basis of environmental sensitivity. *Proceedings of the Second International Conference on Quantitative Genetics*
7. Mulder HA, Rönnegård L, Fikse WF, Veerkamp RF, Strandberg E (2013) Estimation of genetic variance for macro- and micro-environmental sensitivity using double hierarchical generalized linear models. *Genet Sel Evol* 45(1):1–14
8. Berghof TV, Poppe M, Mulder HA (2019) Opportunities to improve resilience in animal breeding programs. *Front Genet* 9:692
9. Hill WG, Mulder HA (2010) Genetic analysis of environmental variation. *Genet Res* 92(5-6):381–395
10. Mulder HA, Bijma P, Hill WG (2008) Selection for uniformity in livestock by exploiting genetic heterogeneity of residual variance. *Genet Sel Evol* 40(1): 1–23
11. Finlay KW, Wilkinson GN (1963) The analysis of adaptation in a plant-breeding programme. *Aust J Agric Res* 14(6):742–754
12. Meuwissen TH (2009) Accuracy of breeding values of 'unrelated' individuals predicted by dense SNP genotyping. *Genet Sel Evol* 41(1):1–9
13. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF, Schenkel FS (2009) Invited review: reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci* 92(1):16–24
14. Windig JJ, Mulder HA, Bohthe-Wilhelmus DI, Veerkamp RF (2011) Simultaneous estimation of genotype by environment interaction accounting for discrete and continuous environmental descriptors in Irish dairy cattle. *J Dairy Sci* 94(6):3137–3147
15. Silva FF, Mulder HA, Knol EF, Lopes MS, Guimarães SEF, Lopes PS et al (2014) Sire evaluation for total number born in pigs using a genomic reaction norms approach. *J Anim Sci* 92(9):3825–3834
16. Hayes BJ, Daetwyler HD, Goddard ME (2016) Models for genome x environment interaction: examples in livestock. *Crop Sci* 56(2251):2259. <https://doi.org/10.2135/cropsci2015.07.0451>
17. Fikse F, Rekaya R, Weigel K (2003) Genotype by environment interaction for milk production traits in Guernsey cattle. *Interbull Bulletin* 27:9–9
18. Fikse WF, Rekaya R, Weigel KA (2003) Assessment of environmental descriptors for studying genotype by environment interaction. *Livest Prod Sci* 82(2-3):223–231
19. Mrode RA (2014) Linear models for the prediction of animal breeding values. *Cabi*.
20. Schaeffer LR (2004) Application of random regression models in animal breeding. *Livest Prod Sci* 86(1-3):35–45
21. Calus MPL, Groen AF, De Jong G (2002) Genotype x environment interaction for protein yield in Dutch dairy cattle as quantified by different models. *J Dairy Sci* 85(11):3115–3123
22. Gianola D, Van Kaam JB (2008) Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics* 178(4):2289–2303
23. Poppe M, Mulder HA, Veerkamp RF (2021) Validation of resilience indicators by estimating genetic correlations among daughter groups and with yield responses to a heat wave and disturbances at herd level. *J Dairy Sci* 104(7):8094–8106
24. Rönnegård L, Felleki M, Fikse F, Mulder HA, Strandberg E (2010) Genetic heterogeneity of residual variance—Estimation of variance components using double hierarchical generalized linear models. *Genet Sel Evol* 42:8
25. Rönnegård L, Valdar W (2011) Detecting major genetic loci controlling phenotypic variability in experimental crosses. *Genetics* 188:435–447
26. Rönnegård L, Valdar W (2012) Recent developments in statistical methods for detecting genetic loci affecting phenotypic variability. *BMC Genet* 13:63
27. van Pelt ML, de Jong G, Mulder HA (2016) Genotype by environment interaction for production, somatic cell score, workability, and conformation traits in Dutch Holstein-Friesian cows between farms with or without grazing. *J Dairy Sci*. <https://doi.org/10.3168/jds.2015-10555>
28. Liu A, Su G, Höglund J, Zhang Z, Thomasen J, Christiansen I, Wang Y, Kargo M (2019) Genotype by environment interaction for female fertility traits under conventional and organic production systems in Danish Holsteins. *J Dairy Sci* 102:8134–8147. <https://doi.org/10.3168/jds.2018-15482>
29. Pfeiffer C, Fuerst C, Schwarzenbacher H, Fuerst-Waltl B (2016) Genotype by environment interaction in organic and conventional production systems and their consequences for breeding objectives in Austrian Fleckvieh cattle. *Livest Sci* 185:50–55
30. Shabalina T, Yin T, May K, König S (2021) Proofs for genotype by environment interactions considering pedigree and genomic data from organic and conventional cow reference populations. *J Dairy Sci* 104(4): 4452–4466
31. Gerber A, Krogmeier D, Emmerling R, Götz KU (2018) Analysis of genotype by environment interaction for milk yield traits in first lactation of Simmental

- cattle. *J Anim Breed Genet* 125:382–389. <https://doi.org/10.1111/j.1439-0388.2008.00731.x>
32. Huquet B, Leclerc H, Ducrocq V (2012) Modelling and estimation of genotype by environment interactions for production traits in French dairy cattle. *Genet Sel Evol.* <https://doi.org/10.1186/1297-9686-44-35>
 33. Schmid M, Imortjust A, Emmerling R, Fuerst C, Hamann H, Bennewitz J (2021) Genotype-by-environment interactions at the trait level and total merit index level for milk production and functional traits in Brown Swiss cattle. *Animal* 15:100052. <https://doi.org/10.1016/j.animal.2020.100052>
 34. Santos JC, Henrique C, Malhado M, Cobuci JA, Paulo M, De Rezende G, Luiz P, Carneiro S (2020) Genotype-environment interaction for productive traits of Holstein cows in Brazil described by reaction norms. *Trop Anim Health Prod* 52:2425–2432
 35. Mulim HA, Fernando L, Pinto B, Valloto AA, Pedrosa VB (2021) Genotype by environment interaction for somatic cell score in Holstein cattle of southern Brazil via reaction norms. *Anim Biosci* 34:499–505
 36. Mulim HA, Luiz P, Carneiro S, Henrique C, Malhado M, Fernando L, Pinto B, Mourno GB, Valloto AA, Pedrosa VB (2021) Genotype by environment interaction for fat and protein yields via reaction norms in Holstein cattle of southern Brazil. *J Dairy Res* 88(16):22
 37. Bohlouli M, Yin T, Hammami H, Gengler N, König S (2021) Climate sensitivity of milk production traits and milk fatty acids in genotyped Holstein dairy cows. *J Dairy Sci* 104:6847–6860. <https://doi.org/10.3168/jds.2020-19411>
 38. Ravagnolo O, Misztal I, Hoogenboom G (2000) Genetic component of heat stress in dairy cattle, development of heat index function. *J Dairy Sci* 83(9): 2120–2125
 39. Carrara ER, Petrini J, Sulvian M, de Oliveira HR, Gregori AR, de Souza Iung LH, Miquilini M, Machado PF, Mourno GB (2021) Genetic parameters for milk yield and quality traits of Brazilian Holstein cows as a function of temperature and humidity index. *J Anim Breed Genet.* <https://doi.org/10.1111/jbg.12636>
 40. Negri R, Aguilar I, Feltes GL, Araujo Cobuci J (2021) Selection for test-day milk yield and thermotolerance in Brazilian Holstein Cattle. *Animals* 11(1):13
 41. Ansari-Mahyari S, Ojali MR, Forutan M, Riasi A, Brito LF (2019) Investigating the genetic architecture of conception and non-return rates in Holstein cattle under heat stress conditions. *Trop Anim Health Prod* 51(1847):853
 42. Nguyen TT, Bowman PJ, Haile-Mariam M, Pryce JE, Hayes BJ (2016) Genomic selection for tolerance to heat stress in Australian dairy cattle. *J Dairy Sci* 99(4): 2849–2862
 43. Cheruyiot EK, Nguyen TTT, Haile-Mariam M, Cocks BG, Abdelsayed M, Pryce JE (2020) Genotype-by-environment (temperature-humidity) interaction of milk production traits in Australian Holstein cattle. *J Dairy Sci* 103(2460):2476. <https://doi.org/10.3168/jds.2019-17609>
 44. Strandberg E, Brotherstone S, Wall E, Coffey MP (2009) Genotype by environment interaction for first-lactation female fertility traits in uK dairy cattle. *J Dairy Sci* 92:3437–3446. <https://doi.org/10.3168/jds.2008-1844>
 45. Ismael A, Strandberg E, Berglund B, Kargo M, Fogh A, Lövendahl P (2016) Genotype by environment interaction for activity-based estrus traits in relation to production level for Danish Holstein. *J Dairy Sci.* <https://doi.org/10.3168/jds.2016-11446>
 46. Lassen J, Mark T (2008) Short communication: Genotype by housing interaction for conformation and workability traits in Danish Holsteins. *J Dairy Sci* 91: 4424–4428. <https://doi.org/10.3168/jds.2008-1116>
 47. Tiezzi F, de Los Campos G, Gaddis KP, Maltecca C (2017) Genotype by environment (climate) interaction improves genomic prediction for production traits in US Holstein cattle. *J Dairy Sci* 100(3):2042–2056
 48. Clay JS, Vinson WE, White JM (1979) Heterogeneity of daughter variances of sires for milk yield. *J Dairy Sci* 62(6):985–989
 49. Rönnegård L, Felleki M, Fikse WF, Mulder HA, Strandberg E (2013) Variance component and breeding value estimation for genetic heterogeneity of residual variance in Swedish Holstein dairy cattle. *J Dairy Sci* 96(4):2627–2636
 50. Paulo M, De Rezende G, Henrique C, Malhado M, Bi S, Luiz P, Carneiro S, Adrián J, Bozzi R (2020) Genotype-environment interaction for age at first calving in Limousine and Charolais cattle raised in Italy, employing reaction norm model. *Livest Res Rural Dev* 232:103912. <https://doi.org/10.1016/j.livsci.2019.103912>
 51. Cardoso FF, Tempelman RJ (2012) Linear reaction norm models for genetic merit prediction of Angus cattle under genotype by environment interaction. *J Anim Sci* 90:2130–2141. <https://doi.org/10.2527/jas.2011-4333>
 52. de Lima Silva T, Carneiro PLS, Ambrosini DP, Lobo RB, Filho RM, Malhado CHM (2019) Genotype-environment interaction in the genetic variability analysis of reproductive traits in Nellore cattle. *Livest Sci* 230:103825. <https://doi.org/10.1016/j.livsci.2019.103825>
 53. Toghiani S, Hay E, Fragomeni B, Rekaya R, Roberts AJ (2020) Genotype by environment interaction in response to cold stress in a composite beef cattle breed. *Animal* 14(1576):587. <https://doi.org/10.1017/S1751731120000531>
 54. Neves HHR, Carvalheiro R, Roso VM, Queiroz SA (2011) Genetic variability of residual variance of production traits in Nellore beef cattle. *Livest Sci* 142(1-3):164–169
 55. Neves HH, Carvalheiro R, Queiroz SA (2012) Genetic and environmental heterogeneity of residual variance of weight traits in Nellore beef cattle. *Genet Sel Evol* 44(1):1–12

56. Chen S-Y, Freitas PH, Oliveira HR, Lázaro SF, Huang YJ, Howard JT et al (2021) Genotype-by-environment interactions for reproduction, body composition, and growth traits in maternal-line pigs based on single-step genomic reaction norms. *Genet Sel Evol* 53
57. Song H, Zhang Q, Misztal I, Ding X (2020) Genomic prediction of growth traits for pigs in the presence of genotype by environment interactions using single-step genomic reaction norm model. *J Anim Breed Genet* 137:523–534
58. Zumbach B, Misztal I, Tsuruta S, Sanchez JP, Azain M, Herring W, Holl J, Long T, Culbertson M (2008) Genetic components of heat stress in finishing pigs: development of a heat load function. *J Anim Sci* 86:2082–2088
59. Fragomeni BO, Lourenco DAL, Tsuruta S, Bradford HL, Gray KA, Huang Y, Misztal I (2016) Using single-step genomic best linear unbiased predictor to enhance the mitigation of seasonal losses due to heat stress in pigs. *J Anim Sci* 94:5004–5013
60. Tiezzi F, Brito LF, Howard J, Huang YJ, Gray K, Schwab C, Fix J, Maltecca C (2020) Genomics of Heat tolerance in reproductive performance investigated in four independent maternal lines of pigs. *Front Genet* 11:629
61. Usala M, Macciotta NP, Bergamaschi M, Maltecca C, Fix J, Schwab C, Shull C, Tiezzi F (2021) Genetic parameters for tolerance to heat stress in crossbred swine carcass traits. *Front Genet* 11:612815
62. Mulder HA, Hill WG, Vereijken A, Veerkamp RF (2009) Estimation of genetic variation in residual variance in female and male broiler chickens. *Animal* 3(12):1673–1680
63. Berghof TV, Bovenhuis H, Mulder HA (2019) Body weight deviations as indicator for resilience in layer chickens. *Front Genet* 10:1216
64. Gutiérrez JP, Nieto B, Piqueras P, Ibanez N, Salgado C (2006) Genetic parameters for canalisation analysis of litter size and litter weight traits at birth in mice. *Genet Sel Evol* 38(5):1–18
65. Garreau H, Bolet G, Larzul C, Robert-Granic C, Saleil G, SanCristobal M, Bodin L (2008) Results of four generations of a canalising selection for rabbit birth weight. *Livest Sci* 119(1-3):55–62
66. Mota RR, Lopes PS, Tempelman RJ, Silva FF, Aguilar I, Gomes CCG, Cardoso FF (2016) Genome-enabled prediction for tick resistance in Hereford and Braford beef cattle via reaction norm models. *J Anim Sci* 94:1834–1843. <https://doi.org/10.2527/jas.2015-0194>
67. Su G, Madsen P, Lund MS, Sorensen D, Korsgaard IR, Jensen J (2006) Bayesian analysis of the linear reaction norm model with unknown covariates. *J Anim Sci* 84(7):1651–1657
68. Shariati MM, Su G, Madsen P, Sorensen D (2007) Analysis of milk production traits in early lactation using a reaction norm model with unknown covariates. *J Dairy Sci* 90(12):5759–5766
69. Su G, Madsen P, Lund MS (2009) Reaction norm model with unknown environmental covariate to analyze heterosis by environment interaction. *J Dairy Sci* 92(5):2204–2213
70. Knap PW, Su G (2008) Genotype by environment interaction for litter size in pigs as quantified by reaction norms analysis. *Animal* 2(12):1742–1747
71. Mulder HA, Bijma P (2005) Effects of genotype x environment interaction on genetic gain in breeding programs. *J Anim Sci* 83(1):49–61
72. Mulder HA, Veerkamp RF, Ducro BJ, Van Arendonk JAM, Bijma P (2006) Optimization of dairy cattle breeding programs for different environments with genotype by environment interaction. *J Dairy Sci* 89(5):1740–1752
73. Rauw WM, Gomez-Raya L (2015) Genotype by environment interaction and breeding for robustness in livestock. *Front Genet* 6(1):15. <https://doi.org/10.3389/fgene.2015.00310>
74. Mulder HA (2016) Genomic selection improves response to selection in resilience by exploiting genotype by environment interactions. *Front Genet* 7(1):11. <https://doi.org/10.3389/fgene.2016.00178>
75. Renaudeau D, Collin A, Yahav S, de Basilio V, Gourdine JL, Collier RJ (2012) Adaptation to hot climate and strategies to alleviate heat stress in livestock production. *Animal* 6(707):728
76. Colditz IG, Hine BC (2016) Resilience in farm animals: biology, management, breeding and implications for animal welfare. *Anim Prod Sci* 56(12):1961–1983
77. Polsky L, von Keyserlingk MA (2017) Invited review: effects of heat stress on dairy cattle welfare. *J Dairy Sci* 100(11):8645–8657
78. Jarquin D, Crossa J, Lacaze X, Du Cheyron P, Daucourt J, Lorgeou J et al (2014) A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor Appl Genet* 127(3):595–607
79. Cao L, Liu H, Mulder HA, Henryon M, Thomasen JR, Kargo M, Sorensen AC (2020) Genomic breeding programs realize larger benefits by cooperation in the presence of genotype x environment interaction than conventional breeding programs. *Front Genet* 11:251