Estimating breeding values for crossbred performance based on purebred data by accounting for dominance

Xinrou Huang
REG.NR.: 1024002

Major thesis Animal Breeding and Genomics (examcode: ABG80436)
August 2022

Supervisors/examiners: **Duenk Pascal**

Thesis: Animal Breeding and Genomics

Abstract

In genomic selection (GS), instead of using crossbred data as reference population, using purebred is more profitable and useful. However, purebred and crossbred are different while they have a different genetic correlation r_{PC}, mainly caused by the dominance effect. Therefore, the aim of the study is to investigate whether including dominance in a prediction model for crossbred performance based on purebred data can improve the model. The traits examined in the models were the body weight of broiler chicken at seven (BW7) and 35 days (BW35). Estimated breeding values (GEBVs) of purebred sires were calculated based on the average SNP effects. In additive models (MA), the average effects were gained through the models, and for dominance model (MAD), the average effects were computed trough estimated additive and dominance effects. To validate the GEBVs, a K-fold cross validation was used and the GEBVs were correlated with the mean performance of the crossbred offspring of the sires. The result showed that the accuracy did not differ much by involving dominance effect for both traits, with the highest mean accuracy for the dominance model for BW35 (0.3158). Overall, most of the replicates showed a superiority in the additive model. Even if the difference in accuracy between different models was not obvious, involving more non-additive genetic effect is still expected to improve the prediction models in the future.

Table of Content

Introduction

Nowadays, genomic selection (GS) is a common method used to estimate breeding values with a high accuracy and broad applicability (Goddard and Hayes, 2007). Genomic selection uses the genomic and phenotype data from a reference population to estimate the breeding value of the young selection candidates. By collecting those data, it is possible to establish a prediction model that captures the relationship between the phenotype and genotype. Thus, by using the genotype data of an individual, we can estimate its breeding values. This technique is especially profitable when the trait we are pursuing is sex limited or when the trait can only be measured later in life (Calus and Veerkamp, 2011). For example, when estimating breeding values for milk production of a bull, we usually need data from the female relatives of the bull. However, with genomic selection, we only need the genotype data of the bull to calculate its breeding value using the prediction model.

In order to improve performance of the commercial animals, crossbreeding is commonly used in modern farming (Pahmeyer and Britz, 2020). Crossbred animals used in livestock production can combine the beneficial properties of the purebred parental lines. In addition, crossbreeding allows breeders to take the advantages of heterosis (Zeng et al., 2013; Wakchaure et al., 2015). The performance of purebred parents and crossbred offspring are correlated, but different. This difference is measured by the purebred-crossbred genetic correlation (r_{pc}) and can be lower than one due to multiple aspects. Firstly, purebred animals are kept in an environment with higher level of hygiene, larger space and better equipment and food, while crossbred animals are kept in a more economical way. These differences can lead to gene-environment interaction (G×E). Secondly, the two populations also differ in their genetic background. When selecting based on the purebred parental line, response in their crossbred offspring is lower due to the r_{pc} lower than one. Crossbreds benefit from heterosis, which can be partly explained by the dominance effect, which is the interactions between alleles at the same locus (Falconer & Mackay, 1996; Xiang et al., 2016).

In crossbred breeding programs, establishing a prediction model based on purebred and crossbred data are both possible, however, in practice, genotype and phenotype data of the crossbred animals are not only difficult to collect but also expensive (Esfandyari, 2016). Therefore, training on purebred data is more ideal and practical.

The value of r_{pc}can be dependent on the dominance, epistasis and G×E effects, which is also the reason for $r_{pc}$to be lower than one. Genomic prediction models aimed at estimating breeding values for crossbred performance can therefore be improved by accounting for dominance. In the past, dominance effects were not always included in the model, mainly because of its difficulty in estimating (Esfandyari, 2016). However, with the single nucleotide polymorphism (SNP) technique, it is now possible to estimate these non-additive genetic effects (Varona et al., 2018). Studies on the genomic prediction of litter size of Landrace and Yorkshire pigs also showed an increase in prediction accuracy using models involving dominance effects (Esfandyari, 2016).

The aim of the study is to investigate to what extent accounting for dominance can improve a genomic prediction model for body weight in broiler chicken. To achieve this aim, I will compare accuracies, goodness of fit, and bias of genomic estimated breeding values (GEBV) from the additive model and dominance model.

Materials and Methods

The traits analysed in the model were the weight of the purebred broiler chicken at different ages: 7 days (early age) (BW7) and 35 days (BW35). At each age, the models were trained with only additive effect (MA) involved or with both additive and dominance effect (MAD), while only the genetics data of the parental purebred line was used. A K-fold cross validation was used for selecting the validation group so that every random group can serve as validation group once. The comparison was mainly built between MA and MAD to investigate the effect of including dominance in a genomic prediction model to estimate the genomic estimated breeding value (GEBV) for crossbred performance.

Materials

The data used in the analysis was provided by Cobb which was a breeding scenario consisting of three purebred boiler chicken lines (line A, line B and line C) and two crossbreds (line BC and line A(BC)) (Figure 1). Only line A was used to train the models, while for line BC only the genotype data was used as the allele frequencies of another parental line, and phenotypic data of line A(BC) was used to calculate the accuracy of the models.

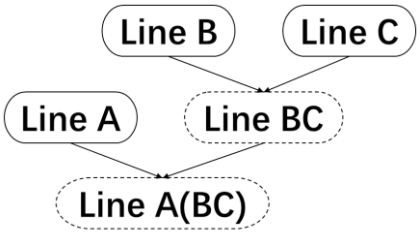


Figure 1 Breeding program for crossbred A(BC), with solid line box representing purebred lines and dotted box representing the crossbred lines.

The chickens were kept in five different pens and five different trials. The data involved 19148 genotypes and 15392 phenotypes data in total consisting of line A and line A(BC). For line A, 4688 phenotype data were available for BW7 and 4477 for BW35 at 50960 SNPs. After matching the phenotype and genotype data, for BW7, 4687 individuals had complete data, and for BW35, 4472 individuals had complete data. The data available were as below:

Table1. Phenotype and genotype data available for the five chicken lines

	Genotype	Phenotype		Data matching*	
		BW7	BW35	BW7	BW35
Line A	50960	4688	4472	4687	4472
Line BC	50960				
A(BC)		10602	10290		

**Data matching: the phenotype data that also have genotype data.*

The phenotype data included the weight of the chicken at 7 days and 35 days, sex, pen,

trial and parents

Data Cleaning

Phenotype data were cleaned by removing the outliers based on their standard deviations (four standard deviations from the mean) in each line. No individuals were removed for both traits after cleaning. And the genotype data were cleaned by removing the SNPs with the MAF (minor allele frequency) lower than 1% for line A and line BC. In line A, no outlier was found in the phenotype data while 3180 SNPs were removed with 47780 SNPs left.

Methods

In order to calculate GEBV of the purebred candidates for crossbred performance, R package 'BGLR' was used. Both the weight at 7 days and 35 days of the chicken were used as response variable to train the models (MA and MAD). The individuals used in the model for both MA and MAD were only purebred from line A. To improve the accuracy of the result, ten repetitions were conducted for each model.

Genomic prediction models

The additive model (MA) can be written as:

$$y_i = \mu + \Sigma x_{ij}\alpha_j + sex_i + pen_i + trial_i + m_i + e_i$$

In which y_i is the phenotypic value of line A (BW7 or BW35) of individual i, μ is the overall mean, x_{ij} is the copy number of a given allele of SNP j of individual i, coded 0, 1 and 2 for aa, Aa and AA, respectively, α_j is the random average effect for SNP j, sex_i is the sex of i (male/female), pen_i is the pen the individual i was kept, trial is the trial of i, m_i is the random maternal effect of individual i, e_i is the residual effect for animal i, and Σ denotes summation over all SNPs. The random effects (α_j, m_i) were assumed to follow a normal distribution.

The additive plus dominance model (MAD) can be written as:

$$y_i = \mu + \Sigma x_{ij}a_j + \Sigma z_{ij}d_j + sex_i + pen_i + trial_i + m_i + e_i$$

With Z_{ij} as the indicator variable for heterozygosity of individual i at SNP j ($z_{ij} = 0$ when homozygous at j and $z_{ij} = 1$ when heterozygous at j), a_j is the random additive effect for SNP j and d_j is the random dominance effect for SNP j both assumed following a normal distribution.

With the estimated additive and dominance effect we can calculate the genomic estimated breeding values (GEBV) for both purebred and crossbred.

In the models, the GEBV was calculated by multiplying the SNP effects and the genotypes. In the additive model, the SNP effects were directly the estimated average effects. In the dominance model, however, the SNP effects were calculated using the

formula:

$$\hat{\alpha}_{CB} = \hat{a} + (1 - 2p_{BC})\hat{d}$$

In which $\hat{a}$ is the estimated additive effect, $\hat{d}$ is the estimated dominance effect and p_{BC} is the population allele frequencies of the BC lines.

Genetic variances

Several genetic variances were involved and calculated in the models. For MA, the additive and maternal variance were involved in the models while MAD also involved the dominance variance. The genetic variances were calculated to identify how the effects contributed to the models.

In MA, the additive variance (σ_A^2) was computed from the additive SNP variances (σ_a^2) in the model:

$$\sigma_A^2 = \sum (2p_i q_i)\sigma_a^2$$

In which p_i represents the allele frequency at SNP i, and q_i is equal to (1- p_i).

In MAD, the additive and dominance variances (σ_D^2) can be calculated as (Vitezica et al., 2013):

$$\sigma_A^2 = \sum (2p_i q_i)\sigma_a^2 + \sum (2p_i q_i (q_i - p_i)^2)\sigma_d^2$$

$$\sigma_D^2 = \sum (2p_i q_i)^2 \sigma_d^2$$

In which σ_d^2 is the dominance SNP variances.

The proportions of the genetic variances were calculated by dividing the genetic variances by the total variance which is ($\sigma_A^2 + \sigma_M^2 + \sigma_e^2$) for MA and ($\sigma_A^2 + \sigma_M^2 + \sigma_D^2 + \sigma_e^2$) for MAD, while σ_M^2 is the maternal variances and σ_e^2 is the total residual variances:

Proportion of	Formula	
	MA	MAD
Additive variance (σ_A^2)	$\frac{\sigma_A^2}{\sigma_A^2 + \sigma_M^2 + \sigma_e^2}$	$\frac{\sigma_A^2}{\sigma_A^2 + \sigma_M^2 + \sigma_D^2 + \sigma_e^2}$
Maternal variance (σ_M^2)	$\frac{\sigma_M^2}{\sigma_A^2 + \sigma_M^2 + \sigma_e^2}$	$\frac{\sigma_M^2}{\sigma_A^2 + \sigma_M^2 + \sigma_D^2 + \sigma_e^2}$
Dominance variance (σ_D^2)	$\frac{\sigma_D^2}{\sigma_A^2 + \sigma_M^2 + \sigma_e^2}$	$\frac{\sigma_D^2}{\sigma_A^2 + \sigma_M^2 + \sigma_D^2 + \sigma_e^2}$

The heritabilities of the traits were calculated with same formula as for the proportion of additive variances with

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_M^2 + \sigma_D^2 + \sigma_e^2}$$

for MA and

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_M^2 + \sigma_D^2 + \sigma_e^2}$$

for MAD.

Goodness-of-fit

Goodness-of-fit was evaluated by the deviance information criterion (DIC) of the models. DIC is a method to calculate the goodness-of-fit based on the posterior mean And is well used in model comparisons (R. Meyer, 2014).

Validation

In order to evaluate the accuracy of the model, a K-fold cross validation was used. Based on the line information, 140 sires from line A had crossbred offspring. These sires were used as validation animals. The sires were equally distributed into ten folds with 14 sires each, and all the purebred offspring of the 14 sires would be grouped in each fold. In order to minimize the effect of certain sires in the group, each fold was treated as testing group once when the other nine folds were used to train the model. For each of the folds, the SNP effects and the deviance information criterion (DIC) will be stored.

The models were validated using the weighted correlation between the GEBV of the sires and their true breeding values (TBVs). The TBVs were gained from a study from Duenk et al. (2019b) on the same chickens, which were contributed from the mean performance of the CB offspring of each sire. The weighting was based on the number of offspring of the sires and the heritability of the traits as $w = \frac{\frac{1}{4}nh_{CB}^2}{1+\frac{1}{4}(n-1)h_{CB}^2}$, in which $h_{CB}^2$is the heritability of trait of the crossbred animals(0.18 for BW7 and 0.23 for BW35 (Duenk et al., 2019b)), and *n* represents the number of crossbred offspring.

Results

In order to compare the effect of involving the dominance effect in a genomic prediction model for body weight at different ages, the models were denoted as additive model at 7 days (MA_7), additive model at 35 days (MA_35), dominance model at 7 days (MAD_7) and dominance model at 35 days (MAD_35).

Heritability

A significant difference in heritability was found between different ages and models. At both ages, the additive model showed a higher heritability than the dominance model in all the repetitions (Figure 2). At 7 days, the heritability ranged from 0.274 to 0.301, where the heritability of MA (0.298) was significantly higher than MAD (0.278). At 35 days of age, the heritability ranged from 0.260 to 0.280, where the heritability of MA (0.275) was significantly higher than MAD (0.264).

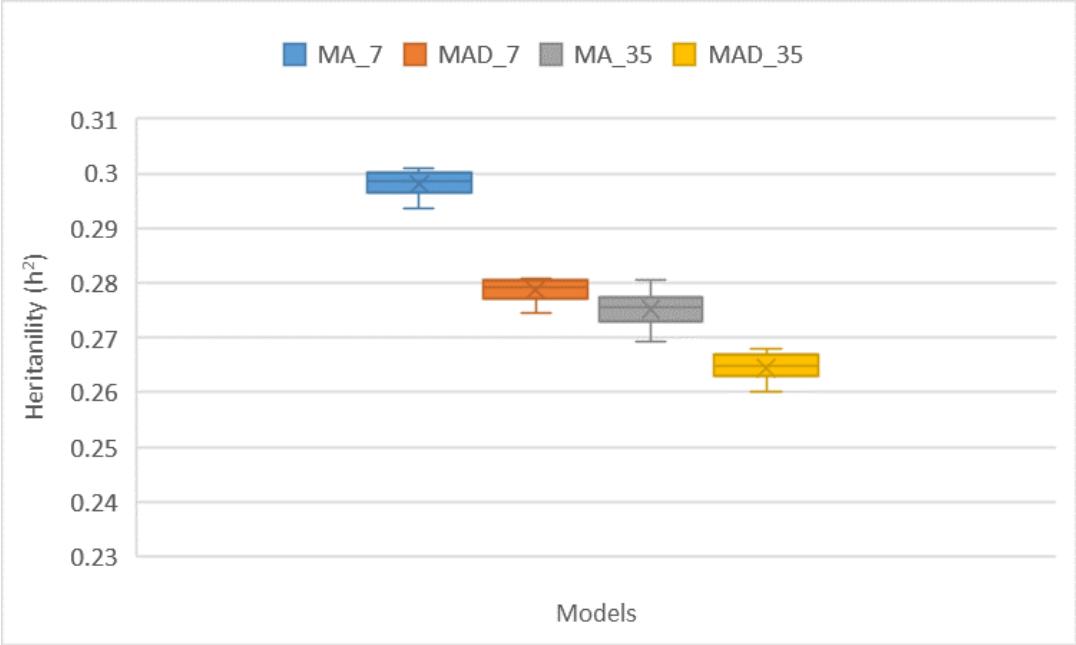
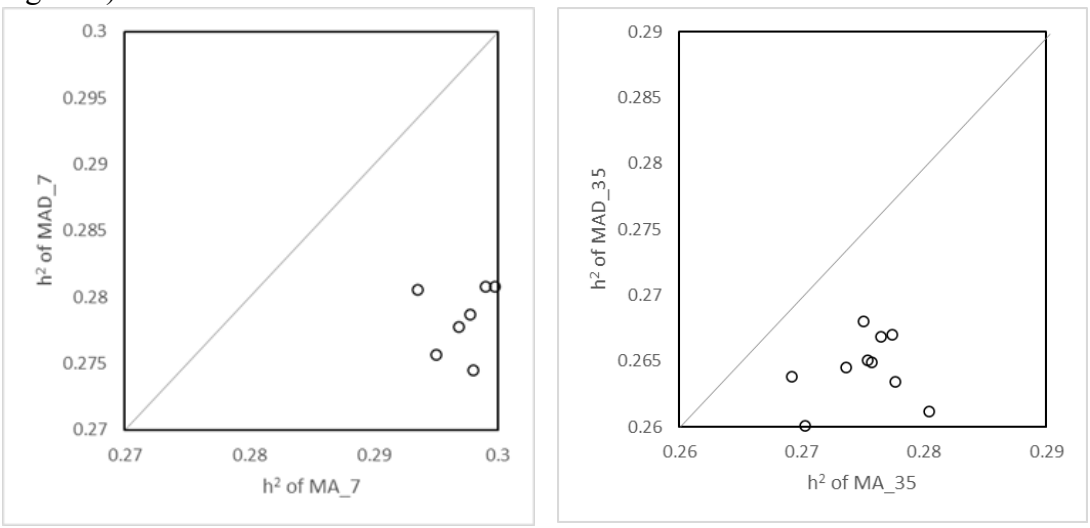


Figure 2. Heritabilities of different models. The y-axis is the heritability (h^2) of each model. The error bars represent for the upper and lower heritability values for each model.

For both traits, the heritability was higher in MA than in MAD for all the replications (Figure 3).



X(a) X(b)

Figure 3. Heritability of the models with MA on the x-axis and MAD on the y-axis. The dots represent for the value at every replication. X(a) is the heritabilities of the models at BW7 and X(b) is the heritabilities of the models at BW35.

Genetic variances

In the study, the additive, dominance and maternal variances were involved in the models (except for MA in which dominance effect was not considered). The proportion of σ_A^2 of the models were discussed in the heritability session as the proportion of σ_A^2 is equal to the h^2. Among the four models, the proportion of maternal variances (σ_M^2) were close to zero, while additive (σ_A^2) and dominance (σ_D^2) variances played an important

role (Figure 4). It is also shown that about 8% of the total variances was explained by σ_D^2, while for BW35 the proportion of σ_D^2 was slightly higher than BW7. The dominance models showed a higher overall genetics variances proportion compared to the additive models in which around 35% of the effect can be explained when involving both the additive and dominance variances (Figure 4).

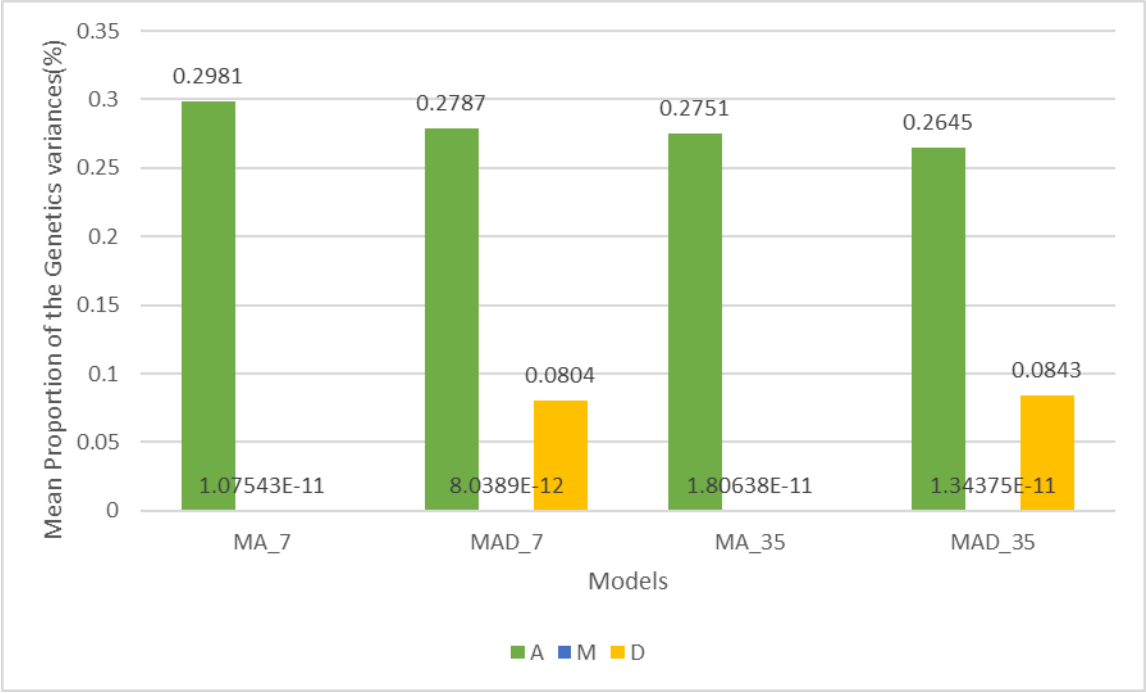


Figure 4. Mean of replicates of the proportions of different genetic variances in different models. "A" represents for the proportions of additive variances; "M" represents for the proportions of maternal variances; and "D" represents for the proportions of dominance variances

Accuracy of the GEBV

Overall, the accuracy of the prediction of the GEBV models were higher for BW35 than for BW7 (Figure 5). Comparing the mean accuracy at both ages, MA was superior to MAD. For BW7, the accuracy was 0.1009 for MA and 0.0884 for MAD. While for BW35, the difference between MA (0.3157) and MAD (0.3158) was much smaller.

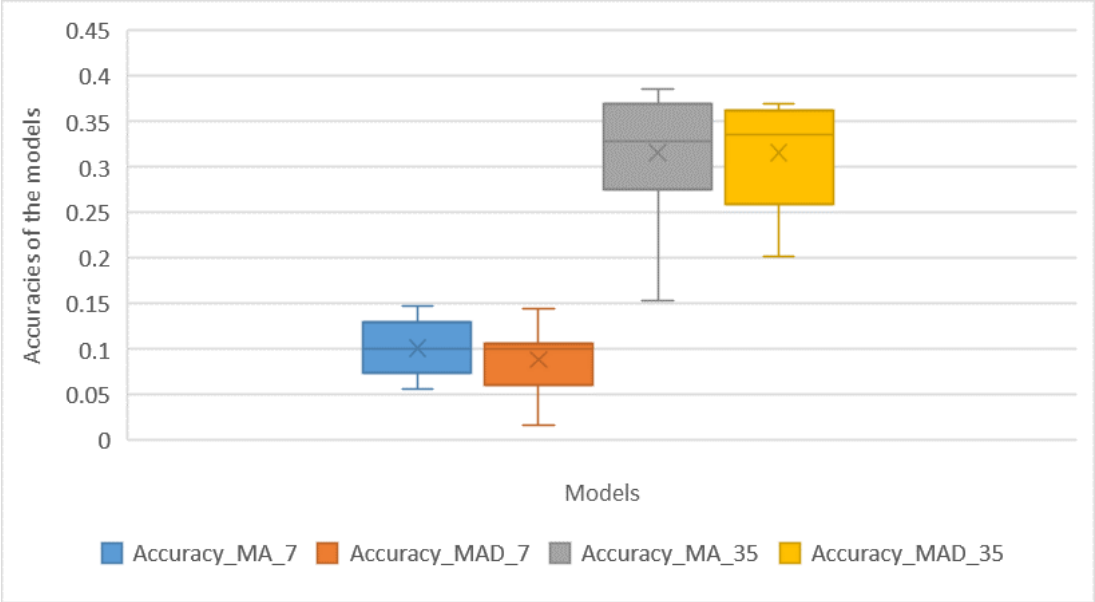
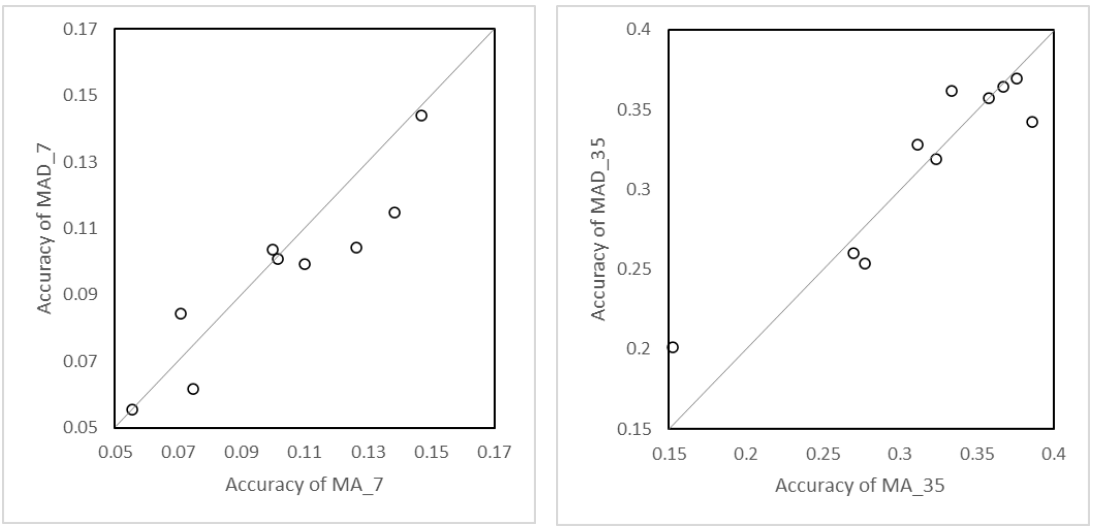


Figure 5. Accuracies of different models for chicken body weight at different ages. The x-axis are the different models, and the y-axis is the accuracy.

At 7 days, MA gave a higher accuracy than MAD (0.1009 over 0.0884). However, not all the replicates gave the same the result (Figure 6(a)). In two replicates, MAD was slightly better in accuracy than MA. For weight at 35 days, MA and MAD were almost equal in accuracy with MAD slightly better (0.3158 over 0.3157) while three out of ten replicates, MAD was slightly better than MA (Figure 6(b)).



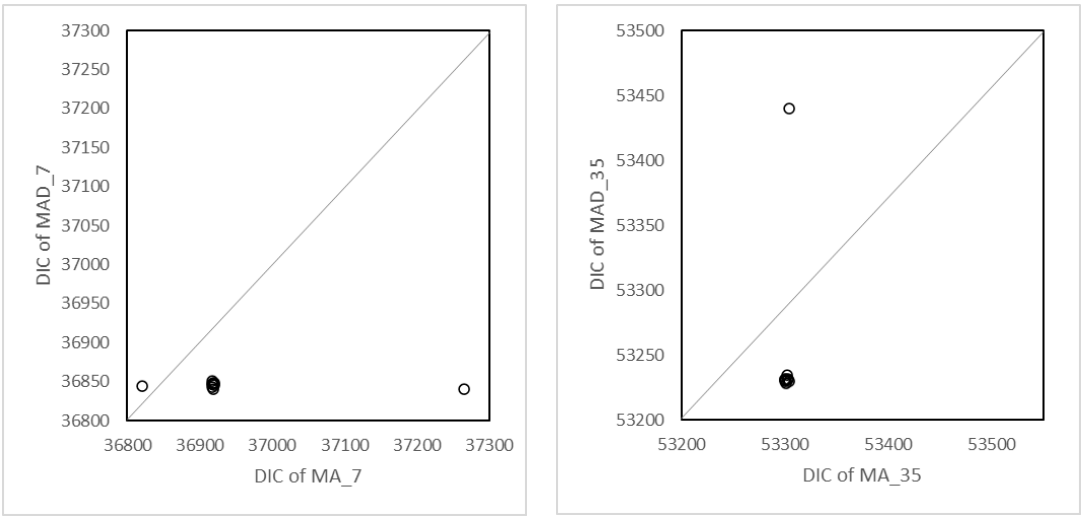
6(a) 6(b)

Figure 6. Accuracies of the models with MA on the x-axis and MAD on the y-axis. The dots represent for the value at every replication. 6(a) is the accuracy of the models for BW7 and 6(b) is the accuracy of the models for BW35.

Goodness-of-fit

At 7 days, only one replicate showed a higher goodness-of-fit of MAD over MA, while all the rest of the replicates indicated that MA is better. Especially at one dot (37264.6,

36840.7), the difference between the two models was the largest. For BW35, still only one replicate showed a higher goodness-of-fit of MAD over MA, while all the rest of the replicates indicated that MA is better. The same dot (53304.0, 53440.1) that MAD was higher in goodness-of-fit than MA, the difference between the teo models was the largest. Overall, MA fitted better than MAD and the models for BW7 were better than the models of BW35.



7(a) 7(b)

Figure 7. The deviance information criterion (DIC) of the models with MA on the x-axis and MAD on the y-axis. The dots represent for the value at every replication. 7(a) is the DIC of the models at 7 days (BW7) and 7(b) is the DIC of the models at 35 days (BW35).

Discussion

The aim of the study was to investigate to what extent accounting for dominance can improve a genomic prediction model for body weight in broiler chicken. The models were used to estimate the genomic estimated breeding values (GEBVs) of the selection candidates in the purebred parental line A. The GEBVs were correlated with the true breeding value of the selection candidates to figure out the accuracy of the models. The genetic variances were also gained from the models to study further the importance of dominance. The DIC were recorded to evaluate the goodness-of-fit of the models. Overall, MAD was not superior to MA even if dominance variances showed that dominance effects contributed to some extent in the models. The accuracy of the model for the weight at 7 days was significantly lower than the one at 35 days for both MA and MAD.

Heritability and other genetic variances

The estimated heritability, also shown as the proportion of additive variances (σ_A^2), was lower with MAD than with MA for both traits. This difference could be caused by the overestimation in σ_A^2, while maternal effect could be underestimated in both models.

Heritability (proportion of additive variances)

The heritability of body weight was expected to be lower for BW7 than for BW35 (Duenk et al, 2019a). However, the opposite was shown in our results, because the

heritability was higher for both MA and MAD at 7 days. Comparing to other papers related to broiler chickens, the heritability of MA for BW7 (0.298) was similar to the result in the paper of Chu et al. (2020) and Mebratie et al. (2019). For BW35, the heritability I estimated was 0.279 which was lower than the one reported by Chu et al. (2020) but higher than the one from Duenk et al. (2019a). For MAD, there is little investigation on the heritability of BW7 and BW35 of broiler chicken when involving dominance effect in the model.

Moreover, many reports indicated that weight at early age have a lower heritability than later weight (Adeleke et al., 2011; Kishore et al., 2002). There are also papers revealing that heritability of broiler chicken can decrease or fluctuate by aging (Adeyinka, 2006; Chu et al., 2020; Dana et al., 2011). This difference in the trend of heritabilities by aging can be influenced by the breed, method of estimation, sampling errors (SNPs/ individuals) and environmental error. When involving dominance effects, the proportions of σ_A^2 decreased by 1-2%, which could also indicate that some of the dominance variance is captured by the additive genetic effect in MA. Nguyen and Kiszlinger (2016) indicated in their study on several traits on cattle that when ignoring dominance effect, additive effect is likely to be overestimated.

Maternal variances

In the result, it is shown that the proportions of the maternal variances (σ_M^2) were close to zero while the maternal variances found in others' papers were higher than zero.

From a paper also on chicken showed that for the body weight at hatch, the proportion of σ_M^2 were 0.17 for MA and 0.15 for MAD, then decreased by the aging of the chicken (Jasouri et al., 2017). Another paper revealing the effect of egg size had shown a significant effect of egg size on the early weight of the chicken (Iqbal et al., 2016). The food, environment and the genetic makeup of the hen all play a role in hen egg size (Karell et al., 2008; Reed & Clark, 2011). Thus, the maternal effect is supposed to be higher than estimated. This difference can be caused by the lack of elements considered as maternal effects. In the models, maternal effect was referred by involving maternal effect as a random effect, while the age of the hen can also affect the initial weight of the offspring. In the future study, it can be useful to find out how to improve the way maternal effect is involved in the model.

In the heritability section, the heritabilityies for BW7 was higher than BW35, while in the other study on the same population, the condition was the opposite (Duenk et al, 2019a). Since the maternal effect is more expressed at the early age, it is possible that the maternal variances were counted as additive variances in the model.

Comparing between models, in MA, the proportions of σ_M^2 was higher than MAD for both BW7 and BW35. This result is similar to the one from a previous paper on another chicken breed for their weight at hatching time, however, for eight- and 12-weeks old chicken the proportions of σ_M^2 were the same for MA and MAD (Jasouri et al., 2017). In our result, the difference between MA and MAD for BW7 was also larger than the one for BW35.

Dominance variances

The dominance variances (σ_D^2) accounted for around 8% among the total variances with

BW35 slightly higher than BW7. This result is similar to the trend shown in the paper of Jasouri et al. (2017) while in my result the difference between ages was smaller, indicating that with aging of the chickens, the proportion of σ_D^2can increase (Le Rouzic et al., 2008).
The dominance effect can be underestimated or overestimated because of the different way the dominance variances were computed(Vitezica et al., 2013; Sun et al., 2014; Nguyen & Kiszlinger, 2016). Moreover, the effect of epistatic also contribute to the performance of the offspring (especially in the early weight of broiler chicken) but can be miss-treated as additive or dominance effect (Carlborg et al., 2003, Le Rouzic et al., 2008).

The accuracy of heritability and other genetic variances can be improved by improving the models considering more about the maternal effect in the model. Involving epistatic effect could be beneficial for the accuracy on estimating the genetic variances in the models (Jiang & Reif, 2015).

Accuracies & Goodness-of-fit

The accuracy of the model is one of the most important factors we need to look at when choosing a prediction model. In this study, purebred reference population was used to estimate GEBV for crossbred performance of purebred selection candidates. After involving the dominance effect, the model did not show an improvement in accuracy and goodness-of-fit.

Comparing between models

Even if the mean of accuracies of MAD for BW35 was higher than the one of MA, most of the accuracies of MA were higher than MAD for both BW7 and BW35, which was contrary to the expectation. Since the dominance variances can explain for 8% of among the total variances, it was expected that by involving the dominance effect in the model, the prediction model can be improved.
Another reason that an improvement of prediction accuracies was expected is that the correlation between the allele frequencies of line A and line BC was 0.57. As crossbred individuals benefit from heterosis, a variety in allele frequency between parental lines can also increase the contribution of dominance effects to the model (Wakchaure et al., 2015). Multiple papers have shown that after involving the dominance effect the prediction was more accurate than only accounting for the additive effect (Lee et al., 2008; Esfandyari, 2016; Liu et al., 2019). However, the condition was the opposite in my study, in which in most of the repetitions, MA had a higher accuracy for the crossbred performance. It was not a unique situation that after involving the dominance effect the accuracy of the model did not improve. In a study on genomic prediction for pureline layers, after involving the dominance effect, the predictability did not improve for several traits (Heidaritabar et al., 2016). Another study from Li et al. (2017) showed that when using crossbred as reference, the accuracies of MADs even decreased for all the seven traits measured. Other papers on poultry and pigs also showed that the accuracies between MA and MAD had little difference and variate between traits (Hidalgo et al., 2015; Liu at al., 2022). However, most of the papers on animal breeding

program still have shown that involving dominance effect can have a positive effect on predicting the performance of the crossbred offspring. One of the benefits on involving the dominance effects in the genomic prediction models for crossbred is also accounting for the allele frequency of another parental line.
In the result, the accuracies for BW35 were significantly higher than the ones for BW7. Several papers have shown that with age growing, prediction accuracy tends to fluctuate (Zhang et al., 2017; Gao et al., 2019; Teng et al., 2019). One of the reasons causing this difference between ages could be their difference in the genetic correlation (r_{PC}) between purebred and crossbred line. Based on the paper from Duenk et al. (2019a), which used the same group of broiler chickens as in the present study, r_{PC} between line A and line A(BC) was 0.64 to 0.80 for BW7 and 0.90 to 0.96 for BW35. Since non-additive effect can lead to a decrease in r_{PC}, the dominance and epistatic effect for BW35 may be weak (Duenk et al., 2021), which could be a reason of the similar accuracy between MA and MAD for BW35. Moreover, the heritability of the trait was also different for BW7 and BW35, which can affect the predictability of the model. As indicated in the paper of Duenk et al. (2019b), the heritability of body weight of broiler chicken at 35 days was higher than the one at 7 days.

Even if involving dominance effect was expected to improve the accuracy of the model, many papers listed above also have shown that the difference in accuracy between MA and MAD can differ by testing on different traits. With a low r_{PC} (compared to the one for BW35), the accuracies for models estimating for BW7 were supposed to be higher for MAD. In the models, there could be an underestimation for the dominance effect which made the accuracies of MA and MAD close to each other.

Methods to improve prediction models

Involve Epistatic effect. Since genomic selection has been applied on selecting breeding candidates, the accuracies of the prediction models have always been studied. We try to involve more factors that would influence the performance of the offspring. At the beginning, among the genetic effect, only the additive effect was involved in the model, while with the development of technology, dominance effect was also considered. A paper from Su et al. (2012) reported that by involving the epistatic effects the accuracy of prediction model improved, and after involving also the dominance effect, the accuracy increased more.
Although involving epistatic effect is expected to be beneficial for the prediction of the crossbred offspring, there are still room of improvement in understanding the interactions between genes and how to estimate the epistatic effect (Mackay, T. F., 2014).

Focusing on traits with lower r_{PC}.

A low r_{PC} can be caused by the express of dominance effect and an influence of G×E interaction. If the trait has a high r_{PC}, the impact of dominance effect on the offspring is supposed to be low, therefore, involving dominance effect would not have a great improvement to the accuracy of the model. While with a low r_{PC}, crossbred data can have a higher accuracy but is less related to the selection candidates (Van Grevenhof &

Van der Werf, 2015; Wientjes et al., 2020). Paper from Esfandyari et al. (2015) shown that using crossbred as reference population can improve the accuracy of the prediction model with an r_{PC} around 0.78. However, using crossbred is more difficult and less profitable than using purebred reference.

Consider about the G×E interaction. In this study, purebred parents and crossbred offspring were raised under the same environment, which is not realistic. Purebred parents are normally kept in an environment with a higher health standard, therefore when applying the model in a real farm, we cannot avoid that G×E interaction could affect the selection of breeding candidates. When this G×E interaction does not contribute much to decreasing the r_{PC}, there is no need to consider this effect. Otherwise, the purebred performance can be tested under the crossbred environment to avoid the influence of G×E interaction (Pascal et al., 2021).

Conclusion

While for the weight at 7 days (BW7) the mean accuracy of the additive model (MA) was higher, for 35 days (BW35) the mean accuracy did not differ much between the two models. Even if the maternal effect was also involved in the models, it was not shown in the results, which could be caused by the overestimation of the additive effect. The improvement in predictability of a model estimating crossbred performance by including dominance effect can be expected when r_{PC} is low. Therefore, it is important to improve also the accuracy of computing for r_{PC} while it varies between traits and lines. Although dominance effect did not show improvement in accuracy and goodness-of-fit of the models, we can still expect that by investigating on a more accurate way to compute dominance variances and involving maternal and dominance effect can be beneficial for genomic prediction models. With future technique developing, the price of genotyping will decrease more and allow more genotype and phenotype data available in training the models to get a higher accuracy.

References

Adeleke, M. A., Peters, S. O., Ozoje, M. O., Ikeobi, C. O. N., Bamgbose, A. M., & Adebambo, O. A. (2011). Genetic parameter estimates for body weight and linear body measurements in pure and crossbred progenies of Nigerian indigenous chickens. Livestock research for rural development, 23(1), 1-7.

Adeyinka, I. A., Oni, O. O., Nwagu, B. I., & Adeyinka, F. D. (2006). Genetic parameter estimates of body weights of naked neck broiler chickens. International journal of poultry science, 5(6), 589-592.

Carlborg, Ö., Kerje, S., Schütz, K., Jacobsson, L., Jensen, P., & Andersson, L. (2003). A global search reveals epistatic interaction between QTL for early growth in the chicken. Genome research, 13(3), 413-421.

Chu, T. T., Madsen, P., Norberg, E., Wang, L., Marois, D., Henshall, J., & Jensen, J.

(2020). Genetic analysis on body weight at different ages in broiler chicken raised in commercial environment. Journal of Animal Breeding and Genetics, 137(2), 245-259.

Dana, N., Vander Waaij, E. H., & Van Arendonk, J. A. (2011). Genetic and phenotypic parameter estimates for body weights and egg production in Horro chicken of Ethiopia. Tropical animal health and production, 43(1), 21-28.

Duenk, P., Calus, M. P., Wientjes, Y. C., Breen, V. P., Henshall, J. M., Hawken, R., & Bijma, P. (2019a). Estimating the purebred-crossbred genetic correlation of body weight in broiler chickens with pedigree or genomic relationships. Genetics Selection Evolution, 51(1), 1-11.

Duenk, P., Calus, M. P., Wientjes, Y. C., Breen, V. P., Henshall, J. M., Hawken, R., & Bijma, P. (2019b). Validation of genomic predictions for body weight in broilers using crossbred information and considering breed-of-origin of alleles. Genetics Selection Evolution, 51(1), 1-12.

Duenk, P., Bijma, P., Wientjes, Y. C., & Calus, M. P. (2021). Optimizing genomic selection for crossbred performance by model improvement and data collection. Journal of animal science, 99(8), skab205.

Esfandyari, H., Bijma, P., Henryon, M., Christensen, O. F., & Sørensen, A. C. (2016). Genomic prediction of crossbred performance based on purebred Landrace and Yorkshire data using a dominance model. Genetics Selection Evolution, 48(1), 1-9.

Esfandyari, H., Sørensen, A. C., & Bijma, P. (2015). A crossbred reference population can improve the response to genomic selection for crossbred performance. Genetics Selection Evolution, 47(1), 1-12.

Gao, N., Teng, J., Pan, R., Li, X., Ye, S., Li, J., ... & Zhang, Z. (2019). Accuracy of whole genome prediction with single-step GBLUP in a Chinese yellow-feathered chicken population. Livestock Science, 230, 103817.

Heidaritabar, M., Wolc, A., Arango, J., Zeng, J., Settar, P., Fulton, J. E., ... & Dekkers, J. C. (2016). Impact of fitting dominance and additive effects on accuracy of genomic prediction of breeding values in layers. Journal of Animal Breeding and Genetics, 133(5), 334-346.

Hidalgo, A. M., Zeng, J., Fernando, R. L., Lopes, M. S., & Dekkers, J. C. M. (2015). Evaluation of genomic prediction of purebreds for crossbred performance in pigs accounting for dominance effects. Exploiting genomic information on purebred and crossbred pigs, 133.

Iqbal, J., Khan, S. H., Mukhtar, N., Ahmed, T., & Pasha, R. A. (2016). Effects of egg size (weight) and age on hatching performance and chick quality of broiler breeder. Journal of applied animal research, 44(1), 54-64.

Jasouri, M., Zamani, P., & Alijani, S. (2017). Dominance genetic and maternal effects for genetic evaluation of egg production traits in dual-purpose chickens. British Poultry Science, 58(5), 498-505.

Jiang, Y., & Reif, J. C. (2015). Modeling epistasis in genomic selection. Genetics, 201(2), 759-768.

Kishore, P. V. L., Rao, G. N., Sharma, R. P., Praharaj, N. K., Gupta, B. R., & Satyanarayana, A. (2002). Inheritance of body weights in Synthetic broiler chickens. Indian Journal of Poultry Science, 37(2), 175-178.

Le Rouzic, A., Alvarez-Castro, J. M., & Carlborg, O. (2008). Dissection of the genetic architecture of body weight in chicken reveals the impact of epistasis on domestication traits. Genetics, 179(3), 1591-1599.

Lee, S. H., Van Der Werf, J. H., Hayes, B. J., Goddard, M. E., & Visscher, P. M. (2008). Predicting unobserved phenotypes for complex traits from whole-genome SNP data. PLoS genetics, 4(10), e1000231.

Li, Y., Hawken, R., Sapp, R., George, A., Lehnert, S. A., Henshall, J. M., & Reverter, A. (2017). Evaluation of non-additive genetic variation in feed-related traits of broiler chickens. Poultry Science, 96(3), 754-763.

Liu, T., Luo, C., Ma, J., Wang, Y., Shu, D., Qu, H., & Su, G. (2022). Including dominance effects in the prediction model through locus-specific weights on heterozygous genotypes can greatly improve genomic predictive abilities. Heredity, 128(3), 154-158.

Liu, Y., Xu, L., Wang, Z., Xu, L., Chen, Y., Zhang, L., ... & Li, J. (2019). Genomic prediction and association analysis with models including dominance effects for important traits in Chinese Simmental beef cattle. Animals, 9(12), 1055.

Mackay, T. F. (2014). Epistasis and quantitative traits: using model organisms to study gene–gene interactions. Nature Reviews Genetics, 15(1), 22-33.

Mebratie, W., Madsen, P., Hawken, R., Romé, H., Marois, D., Henshall, J., ... & Jensen, J. (2019). Genetic parameters for body weight and different definitions of residual feed intake in broiler chickens. Genetics Selection Evolution, 51(1), 1-12.

Meyer, R. (2014). Deviance information criterion (DIC). Wiley StatsRef: Statistics Reference Online, 1-6.

Nguyen, N. T., & Kiszlinger, H. N. (2016). Dominance effects in domestic populations. Acta Agraria Kaposvariensis, 20(1), 1-20.

Su, G., Christensen, O. F., Ostersen, T., Henryon, M., & Lund, M. S. (2012). Estimating additive and non-additive genetic variances and predicting genetic merits using genome-wide dense single nucleotide polymorphism markers.

Sun, C., VanRaden, P. M., Cole, J. B., & O'Connell, J. R. (2014). Improvement of prediction ability for genomic selection of dairy cattle by including dominance effects. PloS one, 9(8), e103934.

Teng, J., Gao, N., Zhang, H., Li, X., Li, J., Zhang, H., ... & Zhang, Z. (2019). Performance of whole genome prediction for growth traits in a crossbred chicken

population. Poultry science, 98(5), 1968-1975.

Van Grevenhof, I. E., & Van der Werf, J. H. (2015). Design of reference populations for genomic selection in crossbreeding programs. Genetics Selection Evolution, 47(1), 1-9.

Vitezica, Z. G., Varona, L., & Legarra, A. (2013). On the additive and dominant variance and covariance of individuals within the genomic selection scope. Genetics, 195(4), 1223-1230.

Wakchaure, R., Ganguly, S., Praveen, P. K., Sharma, S., Kumar, A., Mahajan, T., & Qadri, K. (2015). Importance of heterosis in animals: a review. International Journal of Advanced Engineering Technology and Innovative Science, 1(2), 1-5.

Wientjes, Y. C., Bijma, P., & Calus, M. P. (2020). Optimizing genomic reference populations to improve crossbred performance. Genetics Selection Evolution, 52(1), 1-18.

Zhang, Z., Xu, Z. Q., Luo, Y. Y., Zhang, H. B., Gao, N., He, J. L., ... & Zhang, X. Q. (2017). Whole genomic prediction of growth and carcass traits in a Chinese quality chicken population. Journal of animal science, 95(1), 72-80.