

# Parameter estimation and selection efficiency under Bayesian and frequentist approaches in peach trials

**Julia Angelini**

Centro de Estudios Fotosintéticos y Bioquímicos

**Eugenia Belén Bortolotto**

CCT Rosario: CONICET Rosario

**Gabriela S Faviere**

CCT Rosario: CONICET Rosario

**Claudio F Pairoba**

UNR: Universidad Nacional de Rosario

**Gabriel H Valentini**

INTA: Instituto Nacional de Tecnología Agropecuaria

**Gerardo Domingo Lucio Cervigni** (✉ [cervigni@cefobi-conicet.gov.ar](mailto:cervigni@cefobi-conicet.gov.ar))

CCT Rosario: CONICET Rosario <https://orcid.org/0000-0001-7613-0328>

---

## Research Article

**Keywords:** BLUP, genotype-by-environment interaction, genetic gain, linear mixed model, multi-environment trials, peach breeding

**Posted Date:** April 7th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-377627/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

1 **Parameter estimation and selection efficiency under Bayesian and frequentist approaches in peach**  
2 **trials**

3 Julia Angelini<sup>a,#</sup>, Eugenia Belén Bortolotto<sup>a,#</sup>, Gabriela Soledad Faviere<sup>a,#</sup>, Claudio Fabián Pairoba<sup>b</sup>, Gabriel  
4 Hugo Valentini<sup>c,\*</sup>, Gerardo Domingo Lucio Cervigni<sup>a,\*</sup>

5 <sup>a</sup>Centro de Estudios Fotosintéticos y Bioquímicos (CEFOBI), Universidad Nacional de Rosario - CONICET,  
6 Suipacha 531, 2000 Rosario, Argentina.

7 <sup>b</sup>Secretaría de Ciencia y Tecnología, Universidad Nacional de Rosario, Argentina

8 <sup>c</sup>Estación Experimental Agropecuaria INTA San Pedro, Ruta 9, km 170, 2930 San Pedro, Argentina.

9 <sup>#</sup>These authors contributed equally to this work.

10 <sup>\*</sup>Corresponding authors.

11 TEL.:+54-3329-423-321(G. Valentini); ORCID: 0000-0002-5425-9400

12 TEL.: +54-341-437-1955 (G. Cervigni). ORCID: 0000-0001-7613-0328

13 E-mail addresses: valentini.gabriel@inta.gob.ar; cervigni@cefobi-conicet.gov.ar.

14 Postal address: Suipacha 531, (2000) – Rosario – Santa Fe – Argentina.

15

## 16 **Abstract**

17 Identification of stable and high-yielding genotypes is a real challenge in peach breeding, since genotype-by-  
18 environment interaction (GE) masks the performance of the materials. The aim of this work was to evaluate  
19 the effectiveness of parameter estimation and genotype selection solving the LMM under frequentist and  
20 Bayesian approaches. Fruit yield of 308 peach genotypes were assessed under different seasons and replication  
21 numbers arranged in a completely randomized design. Under the frequentist framework the restricted  
22 maximum likelihood method to estimate variance component and genotypic prediction was used. Different  
23 models considering environment, genotype and GE effects according to the likelihood ratio test and Akaike  
24 information criteria were compared. In the Bayesian approach, the mean and the variance components were  
25 assumed to be random variables having *a priori* non-informative distributions with known parameters.  
26 According the deviance information criteria the most suitable Bayesian model was selected. The full model  
27 was the most appropriate to calculate parameters and genotypic predictions, which were very similar in both  
28 approaches. Due to imbalance data, Cullis's method was the most appropriate to estimate heritability. It was  
29 calculated at 0.80, and selecting above 5% of the genotypes, the realized gain of 14.81 kg.tree<sup>1</sup> was attained.  
30 Genotypic frequentist and Bayesian predictions showed a positive correlation ( $r = 0.9991$ ;  $P = 0.0001$ ). Since  
31 the Bayesian method incorporates the credible interval for genetic parameters, genotypic Bayesian prediction  
32 would be a more useful tool than the frequentist approach and allowed the selection of 17 high-yielding and  
33 stable genotypes.

34 **Key words:** BLUP; genotype-by-environment interaction; genetic gain; linear mixed model;  
35 multienvironment trials; peach breeding.

## 36 **Introduction**

37 The peach tree [*Prunus persica* (L.) Batsch] belongs to the Rosaceae family and is native to China (Hedrick  
38 1917). It is one of the main fruit species grown in temperate and subtropical zones around the world, with  
39 relevant commercial, ornamental, and economic value. According to Faostat 2019 (<http://faostat.fao.org/>) the  
40 main peach and nectarine producing countries are China, Spain and Italy, while Argentina is the tenth largest  
41 producer in the world with 210,000 tons and the second in South America after Chile. The successful activity  
42 of many breeding programs worldwide has led to the release of more than 1000 new cultivars in the last

43 century (Mas-Goméz et al. 2021). The main peach-producing areas in Argentina are Mendoza, Rio Negro (in  
44 the so-called “High Valley”) and the northeastern of Buenos Aires province. San Pedro Agricultural  
45 Experimental Station of the National Institute of Agricultural Technology (INTA San Pedro) is located in this  
46 last region and has a wide peach and nectarine germplasm that is evaluated every season. Many of the peach  
47 varieties used in Argentina are initially selected in INTA San Pedro and the final evaluation is performed in  
48 each one of the producing regions. Despite the large number of cultivars obtained, exist a constant market and  
49 grower demand for new ones. In this way, breeders should consider different potential selection criteria such  
50 as performance, stability, adaptability and quality traits. The genetic basis of traits and the environment have  
51 a strong impact on phenotypic variation, from which useful information can be extracted. Such information  
52 includes determination of several parameter such as mean, variances, association between characters as well  
53 as heritability of each one of them. Genotype-by-environment interaction (GE) is a common phenomenon in  
54 multi-environment trial (MET). In peach, it can seriously affect stability production (Maulión et al. 2014;  
55 2016). The average fruit yield of genotypes evaluated in different environments could be a good parameter to  
56 select materials according to their performance and stability. However, phenotype is the result of the joint  
57 action between genes and environment. Depending on the nature of the GE, it could bring a serious problem  
58 for the breeder's work, because it can reduce the association between phenotypic and genotypic value. This is,  
59 GE may translate into inconsistent genotype behavior altering accession rank (crossover interactions - COI)  
60 or changes in the response without reordering (non-crossover interactions - NCOI) (Cruz and Regazzi 1997).  
61 Yield data from MET are frequently analyzed model in which the total variability is partitioned into the  
62 variance components due to genotypic, environmental and GE effects. Plant breeders have traditionally  
63 estimated variance and covariance components using a two-way analysis of variance (ANOVA) applying the  
64 moment-based estimators (Searle et al. 1992). Considering that the accuracy of genotypic value estimation is  
65 essential in a breeding program, it should be carried out by the most appropriate methods (Borges et al. 2010).  
66 The most popular approaches for estimating quantitative genetic parameters make use of the linear mixed  
67 models (LMM). They allow accounting for various confounding effects (Kruuk 2004; Wilson et al. 2010),  
68 achieving substantial progress in highlighting issues pertaining to fixed effects in quantitative genetic  
69 inferences (Wilson et al. 2010; Wolak et al. 2015). LMM are estimated by maximum likelihood or maximum

70 restricted likelihood (REML), and also Markov-Chain Monte Carlo (MCMC). While REML is associated with  
71 the frequentist approach, the MCMC is linked with the Bayesian framework. Under frequentist LMM, direct  
72 genotypic prediction using the best linear unbiased prediction (BLUP) (Henderson 1975, 1983; White and  
73 Hodge 1988) is more reliable than selection based on phenotype mean only. In this case, genotype can be  
74 assumed as a fixed (best linear unbiased estimation – BLUE) or random effect (BLUP) (Searle et al. 1992).  
75 According to Piepho and Möhring (2006), the most reasonable is to consider that the genotypes under  
76 evaluation are a random sample of a hypothetical population in which selection was performed. The Bayesian  
77 method has increased its attractiveness to analyze data in several fields of knowledge through the MCMC  
78 procedure, since it is efficient even with an arbitrary number of random effects, and provides accurate  
79 estimates of the parameters of interest. Bayesian models have a number of advantages over frequentist ones  
80 (Bolker 2008; Clark 2005; Gelman et al. 2013). Bayesian parameter estimates are more accurate when the  
81 sample size is small; the interpretation of the results is easier and more direct since they indicate the probability  
82 that a parameter assumes a certain value. It is also possible to include measures of uncertainty, missing data  
83 and different levels of variability; and finally, it allows to specify the parameter distribution based on *a priori*  
84 information (prior). Several MCMC algorithms are used to estimate the posterior distributions of parameters  
85 such as Metropolis, Gibbs and Hamiltonian. The Hamiltonian Monte Carlo (HMC) is a generalization of the  
86 Metropolis algorithm that works faster, especially in high dimensions, eliminating the random walk behavior  
87 endemic to Random Walk Metropolis and the Gibbs samplers.

88 Peach yield is a character whose main characteristic is its COI (Mauli3n et al. 2014; 2016; Angelini et al.  
89 2019), so that its statistical modeling is a real challenge. Measured and genetic parameters have been studied  
90 in reproductive characters (Souza et al. 1998a) and associated with peach fruits (Souza et al. 1998b) using  
91 LMM under the frequentist paradigm. However, this methodology was not used to model peach yield, or in  
92 the identification of high-yielding and stable genotypes. Moreover, to our knowledge, the Bayesian approach  
93 has not been used to estimate parameters in peach collection entries, so far. The objective of this work was to  
94 adjust frequentist and Bayesian LMM and compare their accuracy when genetic parameters are estimated, as  
95 well as to identify high-yielding and stable genotypes.

## 96 **Materials and methods**

97 Dataset

98 Fruit yield data set was obtained from 18 season evaluations (SE) involving a total of 308 genotypes, the  
99 amount of them evaluated in each SE is shown (Table 1). They include peach and nectarines for fresh market  
100 consumption with different origin and pedigree: The replication number, the fruits yield per genotype by SE,  
101 as well as the name and origin of each accession have been reported (Supplementary File 1). Furthermore, the  
102 known pedigree of 117 genotypes used is indicated (Supplementary File 2).

103 **Table 1**

104 Number of genotypes assessed in each season

SE1	SE2	SE3	SE4	SE5	SE6	SE7	SE8	SE9
76*	110*	110	103*	14*	112*	84*	27*	39*
SE10	SE11	SE12	SE13	SE14	SE15	SE16	SE17	SE18
68*	95*	160*	72*	177*	140*	147*	171*	164*

105 SE: season evaluation. SE1- SE8 (1991/92 - 1998/99); SE9 – SE11 (2000/01 - 2002/03) and SE12 – SE18  
106 (2005/06 - 2011/12); \* number of genotypes evaluated in each SE.

107  
108 Each SE corresponded to the evaluation of peach germplasm from INTA San Pedro located at 31° 41'12" S  
109 and 60° 47'32" W. Since all experiments were conducted in the same orchard but under different climatic  
110 conditions, each SE was considered different environment. Genotypes were arranged in a completely  
111 randomized design with a replication number between two and thirteen trees. Genotypes evaluated at least in  
112 three SE were included in the analysis. Each year the orchard received management recommended for the  
113 area; this included fungicide and insecticide sprays and pruning similar to the commercial orchards. To avoid  
114 overproduction and the consequent weakening of the plant, the amount of fruit was thinned before ripening,  
115 leaving one every 15 - 20 cm along the branch. Fruit yield was obtained for each replication by using a  
116 weighing machine and expressed in kg.tree<sup>-1</sup>.

117 Statistical analysis

118 Linear mixed model

119 The single-trait multi-environment statistical model is given by:

$$120 \quad y_{ijk} = \mu + G_i + E_j + GE_{ij} + \epsilon_{ijk}, \quad i = 1, \dots, n_G; j = 1, \dots, n_E; k = 1, \dots, n_R \quad (1)$$

121 where  $y_{ijk}$  is the yield from the k-th replication of the i-th genotype in the j-th environment,  $\mu$  is the overall  
122 mean,  $G_i$  is the effect of the i-th genotype,  $E_j$  is the effect of the j-th environment,  $GE_{ij}$  is the interaction

123 between the i-th genotype and the j-th environment, and  $\epsilon_{ijk}$  is the random error from the k-th replication of  
124 the i-th genotype in the j-th environment.

125 Genotype (G), environment (E) and GE were assumed as random (Hinkelmann and Kempthorne 2005).

126 Overall, apart from  $\mu$ , the following random-effect assumptions were made:  $G_i \sim N(0, \sigma_G^2)$ ,  $E_j \sim N(0, \sigma_E^2)$ ,

127  $GE_{ij} \sim N(0, \sigma_{GE}^2)$  and  $\epsilon_{ijk} \sim N(0, \sigma_\epsilon^2)$ .

128 Frequentist approach

129 The REML method was used to fit the mixed model in equation (1). To compare the fit of following models:

130  $M_{\text{null}}$  ( $M_N$ ): simply  $\mu$ ;  $M_{W/G}$ : only E;  $M_{W/GE}$ : G+E and  $M_{\text{full}}$  ( $M_F$ ): E+G+GE, the likelihood ratio test (LRT) and

131 Akaike information criteria (AIC) were used. The lowest value of AIC indicates the best model. LRT statistics

132 approximately follows a chi-square distribution with degrees of freedom equal to the number of additional

133 parameters in the more complex model. LRT and AIC have the following equations:  $LRT =$

134  $2(\log RL_+ - \log RL_-)$ ;  $AIC = -2 \log RL + 2d$ ; where  $RL$  represents the restricted maximum likelihood;

135  $RL_+$  and  $RL_-$  are the maximum values of the restricted likelihood of the more complex and simpler models,

136 respectively;  $d$  is the dimension of the model;  $n$  equals the number of observations valid for the REML

137 estimate.

138 Bayesian approach

139 Mean and variance components were assumed to be random variables having *a priori* distributions, with

140 known parameters. The most suitable prior information should be selected to fit the model. Several options

141 for non-informative priors for scale parameters in hierarchical models were reviewed by Gelman (2006), and

142 his recommendation includes the use of uniform, half-t, half-cauchy, and half-normal families of distributions.

143 The most suitable prior information should be selected to fit the model. Non-informative normal  $(0, 10^{10})$  for

144  $\mu$ , as well as eight standard deviation component distributions ( $P_{1-8}$ ) were tested:  $P_1$ : priors for the standard

145 deviation components  $\sigma_E$ ,  $\sigma_G$ ,  $\sigma_{GE}$  and  $\sigma_\epsilon$  follow inverse-gamma  $(10^{-3}, 10^{-3})$ ;  $P_2$ : priors for the standard

146 deviation components  $\sigma_E$ ,  $\sigma_G$ ,  $\sigma_{GE}$  and  $\sigma_\epsilon$  follow gamma  $(0.5, 0.5)$ ;  $P_3$ : priors for the standard deviation

147 components  $\sigma_E$ ,  $\sigma_G$ ,  $\sigma_{GE}$  and  $\sigma_\epsilon$  follow log-normal  $(0, 1)$ ;  $P_4$ : priors for the standard deviation components  $\sigma_E$ ,

148  $\sigma_G$ ,  $\sigma_{GE}$  and  $\sigma_\epsilon$  follow the positive part of the normal distribution denoted as half-normal (0,1) or equivalently  
 149 as normal(0,1)<sup>+</sup>;  $P_5$ : priors for the standard deviation components  $\sigma_E$ ,  $\sigma_G$ ,  $\sigma_{GE}$  and  $\sigma_\epsilon$  follow uniform (0,1000),  
 150  $P_6$ : priors for the standard deviation components  $\sigma_E$ ,  $\sigma_G$ ,  $\sigma_{GE}$  and  $\sigma_\epsilon$  follow the positive part of the normal  
 151 distribution denoted as half-normal (0,10) or equivalently as normal(0,10)<sup>+</sup>;  $P_7$ : priors for the standard  
 152 deviation components  $\sigma_E$ ,  $\sigma_G$ ,  $\sigma_{GE}$  and  $\sigma_\epsilon$  follow the positive part of the cauchy distribution denoted as half-  
 153 cauchy (0,5) or equivalently as cauchy (0,5)<sup>+</sup>;  $P_8$ : priors for the standard deviation components  $\sigma_E$ ,  $\sigma_G$ ,  $\sigma_{GE}$   
 154 and  $\sigma_\epsilon$  follow the positive part of the cauchy distribution denoted as half-cauchy (0, 2.5\*sd) or equivalently  
 155 as cauchy (0, 2.5\*sd)<sup>+</sup>, where sd is de standard deviation of the residuals of the frequentist model.  
 156 Model  $P_1$  to  $P_8$  were adjusted using four chains, each one comprising 4,000 iterations (the first 2,000 ones  
 157 were discarded) applying the HMC algorithm. Gelman-Rubin ( $\hat{R}$ ) statistics, which compare the variation  
 158 between and within chains (Gelman and Rubin, 1992), were used,  $\hat{R}$  values lower than 1.1 were good evidence  
 159 in favor of chain convergence. MCMC returns samples with some degree of autocorrelation, which varies  
 160 with the model and algorithm selected. Sampling quality was evaluated using the effective sample size. It is a  
 161 measure of the real sample size, discarding those values with correlation in successive simulations. Prediction  
 162 accuracy of the Bayesian model was established through leave-one-out cross-validation,  $\widehat{\text{elpd}}_{100}$ . The  
 163 difference of  $\widehat{\text{elpd}}_{100}$  between two models was expressed in a deviance information criteria (DIC) scale,  
 164 multiplying  $\widehat{\text{elpd}}_{100}$  by -2. A smaller value of DIC reflected a better suitability of the chosen priors and the  
 165 most suitable model among  $M_N$ ,  $M_{W/G}$ ,  $M_{W/GE}$  and  $M_F$ . Genotypic prediction with Bayesian (G-BP) and  
 166 frequentist (G-BLUP) models were obtained considering peach accessions unrelated, while the degree of  
 167 similarity between both predictions was measured using Spearman correlation.  
 168 Stable and high-yielding genotypes were ranked and selected based on their G-BP and predicted performance  
 169 ( $\hat{y}_i$ ) estimated as:  $\hat{y}_i = \hat{\mu} + \hat{G}_i + \bar{\bar{E}} + \overline{\widehat{GE}}_i$  (2), where  $\hat{\mu}$  is the overall mean estimate,  $\hat{G}_i$  is the genotypic  
 170 prediction of i-th genotype;  $\bar{\bar{E}}$  is the mean E prediction and  $\overline{\widehat{GE}}_i$  is the mean GE prediction of the i-th genotype.  
 171 Spearman's correlation to establish the association between  $\hat{y}_i$  and G-BP of selected genotypes was computed.  
 172 Heritability and genetic gain due to selection

173 Two broad-sense heritability measures were estimated:  $h_B^2 = \frac{\sigma_G^2}{\sigma_\epsilon^2 + \sigma_G^2 + \sigma_{GE}^2}$  (3) and, also according to Cullis et  
174 al. (2006):  $h_{BC}^2 = 1 - \bar{v}_{BLUP} / (2\sigma_G^2)$  (4), where  $\bar{v}_{BLUP}$  is the average variance of pair-wise differences  
175 between BLUPs. The genetic gain (GG) based on mean over replications was calculated for selection intensity  
176 of  $p$  is  $C \sigma_G h / \hat{\mu}$ , where  $C$  is a constant given by  $C = \frac{1}{p\sqrt{2\pi}} e^{-z_p^2/2}$  and  $z_p$  is the upper  $p$  quantile of standard  
177 normal distribution,  $h$  is the square root of the heritability, and  $\hat{\mu}$  is the mean response (Hinkelmann and  
178 Kempthorne 2005; Singh et al. 2012).  $C$  value is 2.06 for a 5% intensity of selection. The realized gain using  
179  $h_B^2$  or  $h_{BC}^2$  is: 1-  $RG_1 = h_B^2 \cdot S$  and 2-  $RG_2 = h_{BC}^2 \cdot S$ , where  $S$  is the selection differential (Falconer and Mackay  
180 1996).

181 Bayesian and frequentist analyzes were conducted with the R software (R Core Team, 2018). Bayesian  
182 inference was performed through the Stan language (Stan Development Team, 2017) using a script written in  
183 the lab available on request (angelini@cefobi-conicet.gov.ar), while the frequentist procedure was  
184 implemented using lme4 package.

## 185 **Results and discussion**

### 186 **Selection models**

187 Several Bayesian models with non-informative prior distribution according to DIC statistics were evaluated  
188 (Table 2). The smallest value of DIC reflects a better suitability of the chosen priors and model to the data  
189 (Gelman et al. 2004). Prior distribution above  $\mu$  and one different for the standard deviation components ( $P_1 -$   
190  $P_8$ ) were combined. The prior set  $P_3$  was the most appropriate since the lowest DIC values for each one was  
191 obtained (Table 2). In an ongoing breeding program over many years, information about distribution of  
192 estimated parameters is available. It can be included as informative prior to increase accuracy of inference  
193 estimated parameters (Omer et al. 2015). Since this is the first work of peach yield under Bayesian approaches,  
194 no prior information can be used. Therefore, genetic and non-genetic components were estimated using  $P_3$   
195 prior. Current trial estimates of variance components can be used as prior distributions for similar future trials.  
196 However, the posteriori distributions of the parameters naturally provide an update of the *a priori* distributions  
197 in the light of current datasets.

### 198 **Table 2**

199 DIC values for the selection of standard deviation component prior distribution for peach trials in eighteen  
 200 season at the INTA San Pedro, Argentina

Model	DIC values
P <sub>1</sub>	32817.50
P <sub>2</sub>	32824.30
P <sub>3</sub>	32808.70*
P <sub>4</sub>	32819.70
P <sub>5</sub>	32818.40
P <sub>6</sub>	32817.20
P <sub>7</sub>	32821.60
P <sub>8</sub>	32820.80

201 DIC: deviance information criteria; \*selected model with the lowest DIC.

202

203 To evaluate the fit of the models  $M_{null}$  ( $M_N$ ): simply  $\mu$ ;  $M_{W/G}$ : only E;  $M_{W/GE}$ : G+E and  $M_{full}$  ( $M_F$ ): E+G+GE,  
 204 LTR and AIC measures were considered for the frequentist model, while the DIC was used for the Bayesian  
 205 approach. Analyzing the measurements obtained with all the models (Table 3), it is evident that considering  
 206 only the general average ( $M_N$ ) was the least suitable, although it improved significantly after the environmental  
 207 effect (E) was included ( $M_{W/G}$ ). Since peach yield is sensitive to climatic variability because change in  
 208 temperature and precipitation can alter blooming and fruit development period, a significant effect of E on the  
 209 model is expected.

210 **Table 3**

211 Quality adjustment measures from complete and reduced models under Bayesian and frequentist approaches

Model	Bayesian approach		Frequentist approach	
	DIC		AIC	LRT (P)
$M_N$	34746.60		34741.33	-
$M_{W/G}$	34205.90		34246.73	497.20 (0.0000)
$M_{W/GE}$	33598.80		33848.12	400.61 (0.0000)
$M_F$	32808.70		33420.51	429.61 (0.0000)

212  $M_N$ : null model;  $M_{W/G}$  and  $M_{W/GE}$ : models without genotype or interaction effect; and  $M_F$ : full model, P:  
 213 significant at  $P \leq 0.05$ .

214

215 Bayesian and frequentist environmental effects (E-BP and E-BLUP) were similar (Table 4). According to the  
 216 magnitude of E-BP and E-BLUP the SE can be clustered (C) in C1, C2 and C3. The C1 grouped SE3, SE6,  
 217 SE16 and SE17 whose high and positive predictions would favor performance. On the contrary, C2 clustered  
 218 SE4, SE8, SE10, SE11 and SE15, which had a strong negative influence on yield, while C3 enclosed the  
 219 remaining SE1, SE2, SE5, SE7, SE9, SE12, SE13, SE14 and SE18 with moderate positive or negative effects  
 220 due to interval limit values. Mauli3n et al. (2014) have already indicated that the interaction between rain and

221 heat accumulation during fruit development period was associated with performance instability. Later, a  
222 portion of yield variation was explained by heat accumulation during fruit development period, rainfall during  
223 floral bud endo-dormancy and rainfall from floral bud endo- to ecodormancy (Maulión et al. 2016). The  
224 influence of climatic variables were reported previously by Erez and Couvillon (1987), and Lopez and DeJong  
225 (2007) among others, as essential for flowering and fruit development. Genotypes whose requirements do not  
226 match the climatic conditions of the region cannot be recommended for sustainable production. Angelini et  
227 al. (2019) showed that erratic behavior can be reduced by the release of cultivars adapted to a set of  
228 environments. Therefore, understanding of the nature of GE and, consequently, the matching of a genotype  
229 with appropriate locations to ensure a high and stable production is a major aim of any genetic program.

230 Inclusion of the G effect explains an important portion of data variation as shown by the lower AIC value and  
231 the significance of the LRT  $M_{W/GE}$  model (Table 3). This genotypic variability is caused partly because  
232 accessions of peach germplasm at INTA San Pedro were originated in different regions around the world with  
233 different pedigrees. The importance of genetic variability existing among materials should be emphasized.

234 Although selection of superior and high-yielding peach genotypes has been successful, new cultivars must be  
235 developed to attend the needs of both farmers and consumers. New cultivars and / or possible parents can be  
236 selected with diverse pedigree and origin. Finally, the  $M_F$  model had the lowest AIC, a significant LRT value  
237 and also the lowest DIC value (Table 3), representing the best model under frequentist and Bayesian  
238 frameworks to show the sensitivity to detect all variability sources. Since the presence of GE is a natural  
239 phenomenon in MET crops, including this effect in the final model to analyze agronomic data is a common  
240 fact, the full model in sorghum (Omer and Singh 2017), wheat (Mohammadi et al. 2015), cotton (Mora et al.  
241 2007), soybean (Volpato et al. 2019), among many others, was used. Although the GE characterization for  
242 main crops has been studied and reported extensively for many years, GE studies in peach started recently  
243 (Citadin et al. 2014; Maulión et al. 2014). Maulión et al. (2014) showed that peach yield is a character in  
244 which COI predominates over NCOI. These authors, using the Cruz and Castoldi (1991) test, determined that  
245 COI between pairs of SE ranged from 61 to 80%, while NCOI was estimated between 20 and 38%. Recently,  
246 Angelini et al. (2019) determined the most accurate tests about kind of crossover in a set of peach genotypes

247 were present. Therefore, full model were the most proper to estimate the genetic parameters and predict  
 248 genotypic values.

249 **Table 4**

250 Bayesian and frequentist environmental prediction of each season evaluation

SE	Bayesian approach			Frequentist approach
	E-BP	95 % Credible interval		E-BLUP
		Lower	Upper	
SE1	0.73	-1.31	2.89	0.73
SE2	1.65	-0.17	3.47	1.61
SE3	5.16	3.42	6.98	5.15
SE4	-7.18	-9.10	-5.29	-7.29
SE5	-0.09	-4.29	4.13	-0.21
SE6	8.33	6.54	10.05	8.34
SE7	0.83	-1.05	2.74	1.11
SE8	-4.53	-7.92	-1.17	-4.61
SE9	1.81	-1.00	4.65	1.87
SE10	-6.79	-9.04	-4.63	-6.80
SE11	-4.76	-6.59	-2.95	-4.81
SE12	-1.57	-3.44	0.27	-1.59
SE13	-1.21	-3.97	1.52	-1.13
SE14	0.80	-1.01	2.57	0.81
SE15	-4.62	-6.55	-2.72	-4.65
SE16	2.83	0.96	4.66	2.84
SE17	8.58	6.90	10.27	8.61
SE18	0.01	-1.74	1.76	0.01

251 SE: season evaluation; E-BP: environmental Bayesian prediction (posterior median); E-BLUP: environmental  
 252 frequentist prediction.

253

254 Estimates of variance components, heritability and genetic gain

255 The estimates of average, variance components, heritability and GG are shown (Table 5). Genetic and non-  
 256 genetic variance components estimated under frequentist and Bayesian models were very similar. For all  
 257 estimates, very narrow credible and confidence intervals for Bayesian and frequentist methodologies were  
 258 obtained, respectively.

259 **Table 5**

260 Parameters and measures obtained from fruit yield data under Bayesian and frequentist approaches

	Bayesian approach			Frequentist approach		
	Median	95 % Credible intervals		Estimation	95 % Confidence intervals	
		Lower	Upper		Lower	Upper
$\hat{\mu}$	12.46	11.97	12.95	12.46	10.06	14.84
$\sigma_e^2$	100.80	95.45	106.50	100.80	95.45	106.50
$\sigma_G^2$	27.14	20.07	36.00	27.98	20.79	36.84
$\sigma_E^2$	20.97	11.16	43.30	22.37	11.22	45.29
$\sigma_{GE}^2$	54.17	46.92	60.53	54.76	47.61	62.57
$h_n^2$	0.15	0.11	0.19	0.15	-	-
Measure obtained from Bayesian and frequentist estimates						
$CV_g$	41.86			42.45		
$CV_e$	80.46			80.44		
GG	0.13			0.33		
$\bar{v}_{BLUP}$	10.85			11.15		
$\bar{Y}_{BLUP}$	10.50			10.60		

$h_{BC}^2$	0.80	0.80
GGc	0.22	0.04
$S$	30.48	
$RG_1$	4.64	1.59
$RG_2$	14.08	8.48

$\hat{\mu}$ : overall mean;  $\sigma_\epsilon^2$ : residual variance;  $\sigma_G^2$ : genotypic variance;  $\sigma_E^2$ : environment variance;  $\sigma_{GE}^2$ : genotype-by-environment interaction variance;  $CV_g$ : genotypic coefficient of variation;  $CV_e$ : residual coefficient of variation;  $h_B^2$ : broad-sense heritability;  $\bar{v}_{BLUP}$ : average variance of pair-wise differences between BLUPs;  $\bar{Y}_{BLUP}$ : mean response;  $h_{BC}^2$ : corrected broad-sense heritability (Cullis et al., 2006); GG and GGc: genetic gain estimated using  $h_B^2$  and  $h_{BC}^2$ ;  $S$ : selection differential;  $RG_1$  and  $RG_2$  realized gain calculated with  $h_B^2$  and  $h_{BC}^2$ , respectively.

Variance components are those associated with the random effects of a model. Their knowledge is very useful in genetics and genetic breeding, since the type of genetic delineation and selection strategies that can be used depending on the information obtained from these components. The solution of the mixed model equations depends on the knowledge of the variance and covariance matrix, whose structure is known, but its components are often not. Currently, the standard method for estimating variance components is REML, developed by Patterson and Thompson (1971). REML is commonly used for estimating either the variance components or the solution of fixed and random effects (called BLUE and BLUP) model where phenotypic data are unbalanced or missing (Lu et al. 1999). Not only the estimated BLUE or BLUP depend on the components of variance and covariance, but also the estimates of genetic parameters such as heritability (Arnold et al., 1991).

Variance component estimates via REML can be affected by missing data in any character measured on a single three (Cappa and Cantet 2006). The authors proposed as an alternative the application of REML using the Bayesian approach through MCMC. Such procedures are used to determine the marginal posterior distributions of the parameters by algorithms converging to such marginal densities. The Bayesian procedure is an important tool in genetic evaluation since it considers the variability of all parameters in the model (Wright et al. 2000). In this sense, it is possible to characterize parameters through the mode, median or mean of the parameter's posterior distribution and obtain credible intervals for parameters and predicting genetic effects. As mentioned above, both approaches estimated parameter values and intervals of the same

magnitude. Under these conditions there would be no difference between Bayesian and frequentist approaches. Bayesian is superior to frequentist inference when the posterior distribution of a variance component is bimodal (Mathew et al. 2012). Sorensen and Gianola (2002) reported that when the mixed-model parameters are assigned non-informative distributions, Bayesian and frequentist inferences should be equivalent. On the contrary, the Bayesian approach showed a large reduction of error in estimate measures when the prior distribution is informative (Singh et al. 2015).

Spearman coefficient correlation between Bayesian and frequentist genotypic predictions was 0.9997 ( $P = 0.0001$ ). High correlation between the estimated predictions by REML/BLUP and independent chains (a variant of the methods of MCMC) were reported in maize (Mora and Arnhold 2006) cotton (Mora et al. 2007), and also between REML/BLUP and Gibbs algorithm in *Eucalyptus* (Soria et al. 1998; Mora and Perret, 2007) and *Pinus sylvestris* L. (Waldmann and Ericsson 2006). Previously, Blasco (2001) demonstrated that BLUP can be considered a Bayesian estimator, when this was constructed using a uniform and normal *a priori* distribution for environmental and genetic effects, respectively. This study identified a high correlation between REML/BLUP and the HMC algorithm of genotypic predictions of estimates. Therefore, independently from different considerations on the REML/BLUP and Bayesian algorithms used, estimates of both methods appear to be consistent in plant breeding.

According to G-BP and G-BLUP predictions the best 46 genotypes (about 15% of the selection index) were selected (Table 6). The high correlation (0.9997;  $P = 0.0001$ ) indicates a strong linear association between the predictions, consequently, no discrepancy among the selected genotypes was identified. An advantage of Bayesian procedures is the possibility of construing credible intervals of genetic parameters, including the G-BP values, which are obtained directly from the a posteriori distribution (Wright et al. 2000).

**Table 6**

Genotypes selected according to their Bayesian and frequentist genotypic predictions

Genotype	Origin	Bayesian			Frequentist BLUP
		Posterior median	95 % Credible interval		
			Lower	Upper	
Don Agustin	ARG	24.42	17.59	31.21	24.64
SP 10-32	ARG	12.72	5.58	19.84	12.89
Hermosillo	USA	12.70	8.04	17.46	12.77
SP 26-1	ARG	11.81	4.29	19.33	12.04
Tufth	USA	10.40	4.43	16.33	10.55
Eusebio	USA	9.87	4.79	14.93	9.93
M. Fortininer	USA	8.43	2.17	14.91	8.51
Pavia Verde	USA	8.31	1.54	15.05	8.44
SP 12-32	ARG	8.11	1.99	14.56	8.20
Sunprince	USA	7.67	1.88	13.44	7.73
YS-64	USA	7.61	0.39	14.90	7.66

Encore	USA	7.43	2.48	12.62	7.53
55RA15	USA	7.30	0.38	14.53	7.43
Fla 82-21	USA	6.93	-0.22	14.04	7.10
A 7-76	USA	6.61	-1.67	15.02	6.65
Lara	USA/ARG	6.57	1.40	11.65	6.52
Barceló	ARG	6.56	0.38	13.01	6.70
12ED36	USA	6.37	-0.83	13.66	6.37
84.351.029	ITA	6.25	-0.55	13.01	6.29
Pavia Blanco	ARG	6.14	-0.65	12.95	6.29
SP 16-31	ARG	6.11	-0.65	12.90	6.18
Flaminia	USA	6.05	-0.99	13.29	6.23
Carolina	USA	5.96	-2.09	14.45	6.07
Sunmist	USA	5.93	-1.14	13.10	6.04
SP 28-30	ARG	5.91	-1.90	13.75	5.95
Rey del Monte	URU	5.70	-0.08	11.59	5.73
SP 15-15	ARG	5.64	0.26	10.98	5.65
Summer Prince	USA	5.57	-0.53	11.61	5.63
M11	USA	5.54	-1.00	12.24	5.60
Southern Pearl	USA	5.46	-0.37	11.36	5.54
403-11	USA	5.44	-0.85	11.97	5.52
Sunright	USA	5.42	-2.99	14.00	5.57
Caldessi2010	ITA	5.36	-0.83	11.28	5.38
Don Carlos	ARG	5.19	0.64	9.73	5.19
Calred	USA	5.15	-0.59	10.72	5.16
Sunraycer	USA	5.12	-1.39	11.72	5.14
Texstar	USA	4.89	-0.95	10.69	4.90
Sundowner	USA	4.80	-1.79	11.32	4.72
Harbelle	USA	4.75	-1.47	11.08	4.79
Tropicsnow	USA	4.61	-0.42	9.69	4.65
Ark362	USA	4.57	-2.27	11.57	4.70
Caldessi2000	ITA	4.56	-2.28	11.30	4.60
Bowen	USA	4.51	-1.78	10.48	4.54
M. Sundance	USA	4.47	-1.95	10.77	4.47
Flordaking	USA	4.46	-1.94	10.83	4.39
Marhight	USA	4.34	-1.74	10.80	4.40
Selected mean		6.90	-0.48	13.38	6.97

309

310 Confidence intervals in the classical context can eventually be built by semi-parametric bootstrap techniques  
311 (Efron 1979). However, in the case of G-BLUP, Morris (2002) clarifies that bootstrap constantly  
312 underestimates data variation in finite samples. In this way, the G-BLUP will not be the best prediction. In the  
313 Bayesian perspective, the breeder may choose to select genotypes considering credibility limits obtained  
314 through a posteriori distribution for G-BP (Table 6). In this context, the genotypes that showed credibility  
315 intervals with negative values should not be selected because the range includes negative prediction values,  
316 indicating that the prediction values could be lower than the trial average. Seventeen from initial 52 genotypes  
317 would be selected (those marked in gray color, Table 6).

318 Heritability is the key parameter to study genetic changes in a breeding population undergoing selection  
319 among alternative breeding procedures (Cockerham 1963; Hill 1971). The first definition of heritability was  
320 proposed in the context of animal breeding, in which the basic unit of observation and selection is the  
321 individual animal. On the contrary, in plants there are many designs for crosses, and the observation unit can  
322 be diverse, from individual plants to mean genotypes tested in various environments according to some  
323 experimental design. According to Holland et al. (2002), this complicates both the definition and the estimate  
324 of heritability. An additional difficulty is that all heritable equations assume balanced data, while most  
325 characters exhibit some degree of imbalance. In addition, the standard definition of heritability considers that  
326 trials are analyzed by models that assume independent random effects for blocks, plots, plants, etc., whereas

327 analysis of field trials is often done by spatial models, which imply complex variance – covariance structures  
328 pertaining to observational units. In any case, some degree of disagreement in the estimated values is always  
329 expected, since the estimate of heritability depends on the variability of the population and the environment  
330 where the experiment was performed (Falconer and Mackay 1996). Therefore, there are several alternatives  
331 for unbalanced experiments and correlated genetic effects (Holland et al. 2002; Cullis et al. 2006; Helms and  
332 Hammond 2006). In this work, using the equation (3), the heritability yield was estimated at 0.15 (Table 5),  
333 which is extremely low, even for a very complex trait such as peach yield. Based on the best 17 genotypes  
334 and a selection index around 5%, the genetic gain would be 0.13 kg.tree<sup>-1</sup> (Table 5). Minvielle (1990)  
335 mentioned that a high estimated value of heritability indicates that the correlation between phenotype and  
336 genotype of the two individuals is also high, and an efficient phenotypic selection can be made. Therefore,  
337 low heritability values estimated in this study suggest that simple selection methods cannot be implemented.  
338 According to Cullis et al. (2006), calculation of broad-sense heritability based on mentioned equation (3), is  
339 expected to perform not so well in the case of strong unbalanced and genetic correlation due to the pedigree,  
340 in particular when response to selection is computed from them. Here, response to selection was not based on  
341 the accession pedigree, but our experiment is unbalanced and correlation between genetic effects must be  
342 assumed as existent given the kinship of the modern peach germplasm (Aranzana et al. 2003). Therefore, the  
343 broad-sense heritability estimates according to Cullis et al. (2006), equation (4), was five times greater than  
344 calculated with equation (3),  $h_{BC}^2 = 0.80$ , however, it represented an increase in GG<sub>C</sub> (0.13 to 0.22 kg.tree<sup>-1</sup>).  
345 Both realized gains, RG<sub>1</sub> and RG<sub>2</sub>, were greater than GG and GG<sub>C</sub>, especially RG<sub>2</sub>, because a large differential  
346 selection (S) was combined with a higher heritability (Table 5). To achieve efficient phenotypic selection,  
347 heritability and genetic gain are key parameters in a breeding program, given that they lead to changes on the  
348 crop genetic base, through crossings, selection, and inclusion of promising genotypes along with the  
349 elimination of those with poor performance. In this context, the interest lies in the prediction of superior  
350 genotypes. Selection index close to 5% should lead to relatively rapid genetic progress for yield, 14.81  
351 kg.tree<sup>-1</sup>. Selection based on performance should be considered in a broader context, e. g. flower density, flower  
352 per node, fruit set and size, among other traits, to achieve proper configuration traits to yield. At the same  
353 time, measurements and analysis of characteristics correlated (Nikolić et al. 2010; Cantín et al. 2010, Maulión

354 et al. 2016), simultaneous selection of characters (Tadeu et al. 2019) using different multivariate techniques,  
355 allow to analyze and have a broader and more exact view of the variability available in the germplasm.  
356 Selection indexes and different crossing and genomic selection strategies should also be included in the  
357 selection process. Develop superior genotypes and maintenance of genetic variability at the same time has  
358 always been a challenge for breeders.

#### 359 Genotype by environment analysis of peach genotype selected

360 The genotype prediction by environment (GE-BP) of 17 selected genotypes was estimated and the higher  
361 positive values are in grey (Table 7). Don Agustin had the maximum G-BP but it just showed a strong positive  
362 GE-BP only in SE4 and SE7 while Hermosillo showed the highest GE-BP in SE16, SP 26-1 in SE1, Tufth in  
363 SE2 and SE5, etc. On the other hand, although accessions SP 10-32, M. Fortininer, SP 12-32, 55RA15 and  
364 Barceló were selected among the best, they did not show an outstanding GE-PB in any SE. Since none of the  
365 materials kept their performance prediction in all SE, the instability of the character due to its strong GE is  
366 confirmed once more.

367 Because the genotype number evaluated in each environment was different, it is not possible to establish a  
368 rank based only on GE-BP to select the most stable and highly productive materials. Phenotypic predicted ( $\hat{y}_i$ )  
369 value according to equation (2) for all 17 genotypes selected was calculated (Table 8). Coefficient correlation  
370 between the rank of G-BP and  $\hat{y}_i$  showed a highly positive correlation:  $r = 0.85$  ( $P = 0.0001$ ). Small  
371 discrepancies between  $\hat{y}_i$  and G-BP order indicates that simultaneous selection for high yield and stability  
372 could be made considering only the G-BP, without the calculation of  $\hat{y}_i$  since it will simply delay the selection  
373 process. Previously, Maulión et al. (2014) showed that fruit yield was strongly correlated with superiority ( $P_i$ )  
374 (Lin and Binns 1988), reliability indexes ( $I_i$ ) (Eskridge, 1990), with non-parametric measurements for the  
375 genotype selection index (GSI) (Farshadfar 2008) and the rank sum (RS) (Farshadfar et al. 2011). However,  
376 Bayesian prediction can be used in unbalanced experiments or trials with missing data, but not the  $P_i$ ,  $I_i$ , GSI  
377 and RS measurements. On the other hand, Bayesian approaches allow incorporation of more information than  
378 fruit yield as well as more accurate estimates than those obtained by other methods.

380

**Table 7**

381

Genotype-by-environment Bayesian prediction (GE-BP) obtained for the 17 selected genotypes

	SE1	SE2	SE3	SE4	SE5	SE6	SE7	SE8	SE9	SE10	SE11	SE12	SE13	SE14	SE15	SE16	SE17	SE18
Don Agustin	4.62	-2.66	9.78	13.62	5.34	9.52	9.72			-1.34								
SP10-32	0.44	10.49	15.10	-0.52														
Hermosillo	7.27	5.21	-0.34	9.13		-0.05	-2.30			-6.29	-0.08	-7.54		7.91		20.19	-0.08	-7.51
SP 26-1	19.29		10.69	-6.06														
Tufth	-12.64	19.59	16.63	-7.54	11.56	-6.58												
Eusebio			-12.74	-6.72	-5.45	-9.32	2.83		0.94	0.28	-1.50	11.45		22.33		17.97		
M. Fortininer	-0.96	-1.70	13.73	3.59		1.85												
Pavia Verde	-2.60	-1.45	22.99	-2.31														
SP 12-32	7.93	19.39	5.74	-2.91		-14.41												
Encore	3.59	-2.96	3.83	1.20		-13.73				-2.58	-8.14	11.52		10.33		-7.84	11.21	8.25
Lara			-15.03	2.22		18.75	6.74	1.77	6.76	-0.14	-5.57						-2.09	-0.90
Don Carlos	1.44	-2.13	0.49	-2.01		-7.40	1.32			-2.35	-3.74	4.84		1.53	-5.40	2.69	-0.93	21.95
Sunprince										-3.99	-6.12	1.21	7.14	3.43	-6.20	-11.41	37.32	-6.43
Y5-64		-4.83	-6.44	-1.93		31.19	-2.89											
55RA15												0.52	-1.99	-2.67		10.83	13.7	-6.22
Barceló	10.3	12.13	5.2	-5.98							-8.46							
SP 15-15						15.5	-5.28	-4.56	16.85	3.76	-3.73	-1.43		-3.58	-3.78	-2.80		

382

SE: season evaluations.

383

**Table 8**

384

Predicted phenotypic value and Bayesian genotypic prediction rank of 17 selected genotypes

	G-BP	G-BP Rank	$\hat{y}_i$	$\hat{y}_i$ Rank
Don Agustin	24.42	1	43.29	1
SP 10-32	12.72	2	31.65	3
Hermosillo	12.70	3	27.79	5
SP 26-1	11.81	4	31.81	2
Tufth	10.40	5	27.80	4
Eusebio	9.87	6	20.21	15
M. Fortininer	8.43	7	25.93	6
Pavia Verde	8.31	8	25.02	8
SP 12-32	8.11	9	25.46	7
Sunprince	7.67	10	21.04	12
Y5-64	7.61	11	24.85	9
Encore	7.43	12	21.50	11
55RA15	7.30	13	23.70	10
Lara	6.57	14	20.43	14
Barceló	6.56	15	20.78	13
SP15-15	5.64	16	18.43	17
Don Carlos	5.19	17	18.67	16
Selected mean	14.80		30.98	

385 G-BP: genotypic Bayesian prediction;  $\hat{y}_i$ : Predicted phenotypic value.

## 386 **Conclusions**

387 Stable and high-yielding genotypes selection are always a tedious task for breeder, especially, when the  
388 genotype-by-environment interaction is present. In this work, we present information related to yield and yield  
389 stability in peach. Parameter and efficiency selection under frequentist and Bayesian linear mixed model were  
390 studied. The findings can be summarized as follows: 1- the full model was the best under both approaches, 2-  
391 near all parameters and genotypic prediction were very similar under frequentist and Bayesian methods,  
392 although, because the level of the imbalance data, Cullis method would be best suited to estimate the  
393 heritability, 3 - credible interval made the Bayesian prediction preferable to select high-yielding and stable  
394 genotype selection, and 4 - based on the presence of GE, genotype selection should be done in similar  
395 environments.

396 **Arcknowledgments** The authors appreciate funding support from Consejo Nacional de Investigaciones  
397 Científicas y Técnicas (CONICET). PIP N° 0411.

398 **Author contributions** GDLC and GHV are corresponding author. GDLC and JA wrote the paper and  
399 conducted the statistical analysis. GHV coordinated trials, performed the field measurements, generated  
400 phenotypes and revised the paper. EBB, GSF and CFP wrote and edited different versions of the manuscript.

401 **Author information** GDLC is a member of the Researcher Career of CONICET and JA, EBB and GSF are  
402 Ph.D students. GHV is member of Estación Experimental Agropecuaria INTA San Pedro and CFP of the  
403 Secretaría de Ciencia y Tecnología, Universidad Nacional de Rosario.

404 **Data availability** Data used during this study are included in this article and its supplemental files. Requests  
405 for additional information can be made to the authors.

406 **Animal Research** Not applicable.

407 **Consent to Participate** All the authors gave their consent to participation and for publication.

408 **Plant Reproducibility** Experiments were carried out at INTA San Pedro in compliance with the institutional  
409 and national guidelines. Field experiments were conducted in accordance with local law. Requests for  
410 additional information regarding genetic materials in this study can be made to the corresponding author GHV  
411 and will be considered without undue reservation.

412 **Clinical Trials Registration** Not applicable.

413 **Conflict of interest** The authors declare there to be no conflict of interest.

414 **Funding** This research was funded by Consejo Nacional de Investigaciones Científicas y Técnicas  
415 (CONICET): PIP N° 0411.

## 417 **References**

418 Angelini J, Faviere GS, Bortolotto EB, Arroyo L, Valentini GH, Cervigni GDL (2019) Biplot pattern  
419 interaction analysis and statistical test for crossover and non-crossover genotype-by-environment  
420 interaction in peach. *Sci Hort* 252: 298-309. <https://doi.org/10.1016/j.scienta.2019.03.024>

421 Aranzana MJ, Carbo J, Arús P (2003) Microsatellite variability in peach (*Prunus persica* L.) Batsch: cultivar  
422 identification, marker mutation, pedigree inferences and population structure. *Theor Appl Genet* 106:  
423 1341–1352. <https://doi.org/10.1007/s00122-002-1128-5>

424 Arnold JW, Bertrand JK, Benyshek LL, Ludwig C (1991) Estimates of genetic parameters for live animal  
425 ultrasound, actual carcass data, and growth traits in beef cattle. *J Anim Sci* 69:985–992.  
426 <https://doi.org/10.2527/1991.693985x>

427 Blasco A (2001) The Bayesian controversy in animal breeding. *Sci J Anim Sci* 79:2023.  
428 <https://doi.org/10.2527/2001.7982023x>

429 Bolker B (2008) *Ecological models and data in R*. Princeton University Press, New Jersey.

430 Borges V, Ferreira PV, Soares L, Santos GM, Santos AMM (2010) Seleção de clones de batata-doce pelo  
431 procedimento REML/BLUP. *Acta Sci Agron* 32:643-649. <https://doi.org/10.4025/actasciagron.v32i4.4837>

432 Cantín CM, Gogorcena Y, Moreno MA (2010) Phenotypic diversity and relationships of fruit quality traits in  
433 peach and nectarine [*Prunus persica* (L.) Batsch] breeding progenies. *Euphytica* 171:211–226.  
434 <https://doi.org/10.1007/s10681-009-0023-4>

435 Cappa EP, Cantet RJC (2006) Bayesian inference for normal multiple trait individual tree models with missing  
436 records via Full Conjugate Gibbs. *Can J For Res* 36:1276-1285. <https://doi.org/10.1139/x06-024>

437 Citadin I, Scariotto S, Sachet MRS, Rosa FJ, Raseira MCB, Wagner A (2014) Adaptability and stability of  
438 fruit set and production of peach trees in a subtropical climate. *Sci Agric* 71:133-138.  
439 <https://doi.org/10.1590/S0103-90162014000200007>

440 Clark JS (2005) Why environmental scientists are becoming Bayesians. *Ecol Lett* 8:2–14. <https://doi.org/10.1>  
441 1 1 Uj.1461-0248.2004.00702.x

442 Cockerham CC (1963) Estimation of genetic variances. In Hanson, W. D. and Robinson, H. F., eds. *Statistical*  
443 *genetics and plant breeding*. National Academy of Sciences, National Research Council, Washington,  
444 B.C.

445 Cruz CD, Castoldi FL (1991) Decomposição da interação genótipos x ambientes em partes simples e  
446 complexa. *Revista Ceres* 219:422-430.

447 Cruz CD, Regazzi AJ (1997) Interação genótipos x ambiente. In: *Modelos biométricos aplicados ao*  
448 *melhoramento genético*. (Ed) UFV Viçosa, Minas Gerais, Brasil.

449 Cullis BR, Smith AB, Coombes NE (2006) On the design of early generation variety trials with correlated  
450 data. *J Agr Biol Envir St* 11:381-393. <https://doi.org/10.1198/108571106X154443>

451 Efron B (1979) Bootstrap Methods: Another Look at the Jackknife. *Ann Stat* 7:1-26.  
452 <https://doi.org/10.1214/aos/1176344552>

453 Erez A, Couvillon GA (1987) Characterisation of the influence of moderate temperatures on rest completion  
454 in peach. *J Am Soc Hortic Sci* 112:677-680.

455 Eskridge KM (1990) Selection of stable cultivars using a safety-first rule. *Crop Sci* 30:369–374.  
456 <https://doi.org/10.2135/cropsci1990.0011183X003000020025x>

457 Falconer DS, Mackay FC (1996) *Introduction to Quantitative Genetics*. 4th Edition. Pearson Education  
458 Limited, Essex.

459 Farshadfar E (2008) Incorporation of AMMI stability value and grain yield in a single non-parametric index  
460 (GSI) in bread wheat. *Pak J Biol Sci* 11:1791–1796. <https://doi.org/10.3923/pjbs.2008.1791.1796>

461 Farshadfar E, Mahmodi N, Yaghotipoor A (2011) AMMI stability value and simultaneous estimation of yield  
462 and yield stability in bread wheat (*Triticum aestivum* L.). *Aust J Crop Sci* 5:1837–1844.

463 Faust M, Timon B (1995) Origin and dissemination of peach. *Horticult Rev* 17:331–  
464 379. <https://doi.org/10.1002/9780470650585.ch10>

465 Gelman A, Carlin J B, Stern HS, Rubin DB (2004) *Bayesian Data Analysis*, 2nd Edition. Chapman and Hall,  
466 New York.

467 Gelman A (2006) Prior distributions for variance parameters in hierarchical models. *Bayesian Anal* 1:515-  
468 534. <https://doi.org.10.1214/06-BA117A>.

469 Gelman A, Rubin D B (1992) Inference from Iterative Simulation Using Multiple Sequences. *Stat Sci* 7:457–  
470 511. <https://doi.org.10.1214/ss/1177011136>

471 Gelman A, Hwang J, Vehtari A (2013) Understanding predictive information criteria for Bayesian models.  
472 *Stat Comput* 24:997–1016. <https://doi.org.10.1007/s11222-013-9416-2>

473 Hedrick U P (1917) The peaches of New York. Report of the New York Agricultural Experimental Station of  
474 the year 1916. Lyon, Albany, New York

475 Helms TC, Hammond JJ (2006) Genetic Gain Equation with Correlated Genotype × Environment Effects.  
476 *Crop Sci* 46:1137. <https://doi.org.10.2135/CROPSCI2005.07-0212>

477 Henderson CR (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics*  
478 31:423–447.

479 Henderson CR (1983) Estimation of variances and covariances under multiple trait models. *J Dairy Sci*  
480 67:1581–1589. [https://doi.org/10.3168/jds.S0022-0302\(84\)81480-0](https://doi.org/10.3168/jds.S0022-0302(84)81480-0)

481 Hill WG (1971) Design and efficiency of selection experiments for estimation of genetic  
482 parameters. *Biometrics* 27:293–311. <https://doi.org/10.2307/2528996>

483 Holland JB, Nyquist WE, Cervantes-Martinez CT (2002). Estimating and interpreting heritability for plant  
484 breeding: an update. *Plant Breed Rev* 22:9–112. <https://doi.org/10.1002/9780470650202.ch2>

485 Hinkelmann K, Kempthorne O (2005) Design and Analysis of Experiments. Vol 2, Advanced Experimental  
486 Design, Wiley, Hoboken, New Jersey.

487 Kruuk LEB (2004) Estimating genetic parameters in natural populations using the 'animal model'. *Philos*  
488 *Trans R Soc Lond B Biol Sci* 359:873–890. <https://doi.org/10.1098/rstb.2003.1437>

489 Lin C S, Binns MR (1988) A superiority measure of cultivar performance for cultivar × location data. *Can J*  
490 *Plant Sci* 68:193–198. <https://doi.org/10.4141/cjps88-018>

491 Lu PX, White TL, Huber DA (1999) Estimating type B genetic correlations with unbalanced data and  
492 heterogeneous variances for half-sib experiments. *For Sci* 45:562-572.  
493 <https://doi.org/10.1093/forestscience/45.4.562>

494 Lopez G, DeJong TM (2007) Spring temperatures have a major effect on early stages of peach fruit growth. *J*  
495 *Hort Sci Biotech* 82, 507–512. <https://doi.org/10.1080/14620316.2007.11512266>

496 Mas-Gómez J, Cantín JC, Moreno M A, Prudencio AS, Gómez-Abajo M, Bianco L, Troglio M, Martínez-  
497 Gómez P, Rubio M, Martínez-García PJ (2021) Exploring Genome-Wide Diversity in the National Peach  
498 (*Prunus persica*) Germplasm Collection at CITA (Zaragoza, Spain). *Agronomy* 11:481-499

499 Mathew B, Bauer AM, Koistinen P, Reetz TC, León J, Sillanpää MJ (2012). Bayesian adaptive Markov chain  
500 Monte Carlo estimation of genetic parameters. *Heredity* 109:235–  
501 245. <https://doi.org/10.3390/agronomy11030481>

502 Maulión E, Arroyo L, Daorden ME, Valentini GH, Cervigni GDL (2016) Identification of peach accessions  
503 stability and adaptability in non-balanced trials through years. *Sci Hortic* 199:198–208.  
504 <https://doi.org/10.1016/j.scienta.2015.12.048>

505 Maulión E, Valentini G, Ornella L, Pairoba CF, Daorden ME, Cervigni GDL (2014) Study of statistic stability  
506 to select high-yielding and stable peach genotypes. *Sci Hortic* 175: 258–268.  
507 <https://doi.org/10.1016/j.scienta.2014.06.026>

508 Maulión E, Arroyo L, Daorden ME, Valentini GH, Cervigni GDL (2016) Performance profiling of *Prunus*  
509 *persica* (L.) Batsch collection and comprehensive association among fruit quality, agronomic and  
510 phenological traits. *Sci Hortic* 198: 385–397. <https://doi.org/10.1016/j.scienta.2015.11.017>

511 Minvielle F (1990) Principes d'amélioration génétique des animaux domestiques. INRA Paris et Université  
512 Laval, Quebec.

513 Mohammadi M, Sharifi P, Karimizadeh R, Alt Jafarby J, Khanzadeh H, Hosseinpour T, Poursiabidi M,  
514 Roustaii M, Hassanpour Hosni M, Mohammadi P (2015) Stability of grain yield of durum wheat genotypes  
515 by AMMI model. *Agriculture and Forestry* 61:181-193. <https://doi.org/10.17707/AgricultForest.61.3.18>

516 Mora F, Arnhold A (2006) Inferencia Bayesiana y metodología de modelos lineales mixtos aplicados al  
517 mejoramiento del maíz. *Cienc Inv Agr* 33:217-223.

518 Mora F, Perret S (2007). Aplicación de técnicas bayesianas en el análisis genético de árboles  
519 forestales. *Bosque* 28:198-206. <http://dx.doi.org/10.4067/S0717-92002007000300003>

520 Mora F, Pupim-Junior O, Scapim C A (2007) Predicción del efecto de cultivares de algodón en la presencia  
521 de interacción genotipo-ambiente. *Cienc Inv Agr* 34:11-19. [http://dx.doi.org/10.4067/S0718-](http://dx.doi.org/10.4067/S0718-16202007000100002)  
522 [16202007000100002](http://dx.doi.org/10.4067/S0718-16202007000100002)

523 Morris JS (2002) The BLUPs are not “best” when it comes to bootstrapping. *Stat Probabil Lett* 56:425-430.  
524 [https://doi.org/10.1016/S0167-7152\(02\)00041-X](https://doi.org/10.1016/S0167-7152(02)00041-X)

525 Nikolic D, Rakonjac V, Milatovic D, Fotiric M (2010) Multivariate analysis of vineyard peach [*Prunus persica*  
526 (L.) Batsch.] germplasm collection *Euphytica*. 171:227–234. <https://doi.org/10.1007/s10681-009-0032-3>

527 Omer SO, Abdalla AWH, Mohammed MH, Singh M (2015) Bayesian estimation of genotype-by-environment  
528 interaction in sorghum variety trials. *Commun Biometry Crop Sci* 10:82–95.

529 Omer SO, Singh M (2017) Comparing Bayesian and Frequentist Approaches for GGE Bi-plot Analysis in  
530 Multi-Environment Trials in Sorghum. *Eur J Exp Biol* 7:6-40. <https://doi.org/10.21767/2248-9215.100040>

531 Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal.  
532 *Biometrika* 58:545–554. <https://doi.org/10.2307/2334389>

533 Piepho H-P, Möhring J (2006) Selection in Cultivar Trials—Is It Ignorable? *Crop Sci*  
534 46:192. <https://doi.org/10.2135/cropsci2005.04-0038>

535 R Core Team, 2018. R: a language and environment for statistical computing. Vienna: R Foundation for  
536 Statistical Computing. Available from: <https://www.R-project.org/>

537 Searle SR, Casella G, McCulloch CE (1992) *Variance Components*. Wiley, New York,  
538 <https://doi.org/10.1002/9780470316856>

539 Singh M, Al-Yassin A, Omer SO (2015) Bayesian Estimation of Genotypes Means, Precision, and Genetic  
540 Gain Due to Selection from Routinely Used Barley Trials. *Crop Sci* 55:501-513.  
541 <https://doi.org/10.2135/cropsci2014.02.0111>

542 Singh M, van Ginkel M, Sarker A, Malhotra RS, Imtiaz M, Kumar S (2012) Increasing precision of even  
543 otherwise wellrun trials by capturing heterogeneity of plot error variances. *Agric Res* 1:285–294.  
544 <https://doi.org/10.1007/s40003-012-0024-8>

545 Sorensen D, Gianola D (2002) *Likelihood, Bayesian, and MCMC methods in quantitative genetics*. Springer,  
546 Berlin. <https://doi.org/10.1007/b98952>

- 547 Soria F, Basurco F, Toval G, Silió L, Rodríguez M C, Toro M (1998) An application of Bayesian techniques  
548 to the genetic evaluation of growth traits in *Eucalyptus globulus*. *Can J For Res* 28:1286–1294.  
549 <https://doi.org/10.1139/x98-099>
- 550 Souza VAB, Byrne DH, Taylor JF (1998a) Heritability, genetic and phenotypic correlations, and predicted  
551 selection response of several quantitative traits in peach. I. An analysis of several reproductive traits. *J*  
552 *Amer Soc Hort Sci* 123:598–603. <https://doi.org/10.21273/JASHS.123.4.598>
- 553 Souza VAB, Byrne DH, Taylor JF (1998b). Heritability, genetic and phenotypic correlations, and predicted  
554 selection response of several quantitative traits in peach. II. An analysis of several fruit traits. *J Amer Soc*  
555 *Hort Sci* 123:604–611. <https://doi.org/10.21273/JASHS.123.4.604>
- 556 Stan Development Team, 2017. Stan modeling language users guide and reference manual, version 2.17.0.  
557 Retrieved from <http://mc-stan.org/>
- 558 Tadeua MH, Pio R, Silva GN, Olmsteadc M, Cruzd CD, Machado de Souza FB, Bisia RB (2019) Methods  
559 for selecting peach cultivars in the tropics. *Sci Hortic* 252:252–259.  
560 <https://doi.org/10.1016/j.scienta.2019.01.016>
- 561 Volpato L, Alves RS, Teodoro PE, Vilela de Resende MD, Nascimento M, Nascimento ACC, Ludke  
562 WH, Lopes da Silva F, Borém A (2019). Multi-trait multi-environment models in the genetic selection of  
563 segregating soybean progeny. *PLoS ONE* 14, e0215315. <https://doi.org/10.1371/journal.pone.0215315>
- 564 Waldmann P, Ericsson T (2006) Comparison of REML and Gibbs sampling estimates of multi-trait genetic  
565 parameters in Scots pine. *Theor Appl Genet* 112:1441–1451. <https://doi.org/10.1007/s00122-006-0246-x>
- 566 White TL, Hodge GR (1988) Best linear prediction of breeding values in a forest tree improvement program.  
567 *Theor Appl Genet* 76:719-727. <https://doi.org/10.1007/BF00303518>
- 568 Wilson AJ, Réale D, Clements MN, Morrissey MM, Postma E, Walling CA, Kruuk LEB, Nussey DH  
569 (2010) An ecologist's guide to the animal model. *J Appl Ecol* 79:13–26. <https://doi.org/10.1111/j.1365-2656.2009.01639.x>
- 570
- 571 Wolak M E, Roff D A, Fairbairn DJ (2015) Are we underestimating the genetic variances of dimorphic  
572 traits? *Ecol Evol* 5:590–597. <https://doi.org/10.1002/ece3.1361>

573 Wright DR, Stern HS, Berger PJ (2000) Comparing traditional and Bayesian analyses of selection experiments  
574 in animal breeding. JABES 5:240-256. <https://doi.org/10.2307/1400533>

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1.docx](#)
- [SupplementaryTable2.docx](#)