Scientia Agricola

DOI: http://doi.org/10.1590/1678-992X-2020-0361

ISSN 1678-992X

Research Article

Row–Col and Bayesian approach seeking to improve the predictive capacity and selection of passion fruit

André Oliveira Souza[1], Alexandre Pio Viana[2], Fabyano Fonseca e Silva[3], Camila Ferreira Azevedo[4], Flavia Alves da Silva[2]*, Fernando Higino Lima e Silva[5]

[1]Instituto Federal do Espírito Santo, Rod. BR 482, km 47 – 29148-015 – Alegre, ES – Brasil.
[2]Universidade Estadual do Norte Fluminense Darcy Ribeiro/ CCTA – Lab. de Melhoramento Genético Vegetal, Av. Alberto Lamego, 2000 – 28013-602 – Campos dos Goytacazes, RJ – Brasil.
[3]Universidade Federal de Viçosa – Depto. de Zootecnia, Av. Peter Henry Rolfs, s/ n – 36570-000 – Viçosa, MG – Brasil.
[4]Universidade Federal de Viçosa – Depto. de Estatística – Viçosa, MG – Brasil.
[5]Instituto Federal de Educação, Ciência e Tecnologia de Goiás, km 01 – 75901-970 – Rio Verde, GO – Brasil.
*Corresponding author <flavia_uems@hotmail.com>

Edited by: Thomas Kumke

Received November 13, 2020
Accepted March 11, 2021

ABSTRACT: Methods for genetic improvement of semi–perennial species, such as passion fruit, often involve large areas, unbalanced data, and lack of observations. Some strategies can be applied to solve these problems. In this work, different models and approaches were tested to improve the precision of estimates of genetic evaluation models for several characteristics of the passion fruit. A randomized block design (RBD) model was compared to *a posteriori* correction, adding two factors to the model (post–hoc blocking Row–Col). These models were also combined with the frequentist and Bayesian approaches to identify which combination yields the most accurate results. These approaches are part of a strategic plan in a perennial plant breeding program to select promising genitors of passion to compose the next selection cycle. For Bayesian, we tested two *priors*, defining different values for the distribution parameters of effect variances of the model. We also performed a cross–validation test to choose *a priori* values and compare the frequentist and Bayesian approaches using the root mean square error (RMSE) and the correlation between the predicted and observed values, called Predictive capacity of the model (PC). The model with the post–hoc blocking Row–Col design captured the spatial variability for productivity and number of fruits, directly affecting the experimental precision. Both approaches applied to the models showed a similar performance, with predictive capacity and selective efficiency leading to the selection of the same individuals.

Keywords: REML, *posteriori*, experimental precision

Introduction

The passion fruit is indigenous to Tropical America, with more than 150 species native to Brazil (Bernacci et al., 2015). *Passiflora edulis* is the most marketable passion fruit species in Brazil, accounting for more than 95 % of the fruit farms in the country and also the most planted species worldwide (Meletti et al., 2011). Passion fruit production in the state of Rio de Janeiro is below the national average, mainly due to the lack of improved varieties adapted to edaphoclimatic conditions; however, breeding programs are working to improve this condition (Cavalcante et al., 2019). In passion fruit, breeding methods face challenges similar to other perennial species. The genetic evaluation of individuals under selection requires a large experimental site, due to the planting spacing causing a large spatial variation, compromising data precision and/or unbalancing experiments.

The post–hoc blocking Row–Col design is a method to control the local effect, which is described as an efficient for genetic improvement experiments on passion fruit (Silva et al., 2016; Machado et al., 2020). This technique consists of superimposing a structure of rows and columns to the original design and the effects of blocks and treatments, which allows greater accuracy in the model parameter estimates, redirecting part of the model error to these effects (Gezan et al., 2006).

The Bayesian approach is another method for more accurate estimates, especially few observations are available and designs are unbalanced (Silva et al., 2013; Silva et al., 2018; Silva et al., 2020). This approach allows to use of *a priori* distributions incorporated into the model, which can be an advantage. Compared to the frequentist restricted maximum likelihood/best linear unbiased predictor REML/BLUP approach, normally used in genetic improvement programs for estimation/prediction. If *a priori* distribution chosen is non–informative or the number of observations describes the data very well, the Bayesian approach tends to converge to the maximum likelihood estimation (Resende et al., 2014).

Thus, techniques to speed up the breeding steps, which usually take many years, can be an advantage in terms of saving resources. Thus, this study combined alternative methods, such as the post–hoc blocking Row–Col design and Bayesian approach, to improve accuracy of estimate components of variance, heritability, and selection gain in observations carried out in experiments of genetic improvement of passion fruit under recurrent intrapopulation selection.

Materials and Methods

Observations

We collected data from 81 progenies of full siblings of passion fruit *Passiflora edulis* Sims, after the third cycle of recurrent selection of the genetic improvement program of passion fruit developed in the State of Rio de Janeiro. The progenies were obtained from 27 half–sibling progenies selected by Silva et al. (2009).

The study was carried out in northwestern Rio de Janeiro State, Brazil (21°40′ S, 42°04′ W, altitude 76 m a.s.l.) aiming to generate cultivars adapted to similar local conditions. We used a randomized block design (RBD) with two replicates and the experimental unit consisted of five plants. The experimental site, despite only two blocks, comprised 8505 m^2 and some unevenness was observed in the field that could not be captured by the block, but it was captured by a model incorporating a post-hoc blocking Row-Col design.

Agronomic characteristics

The phenotyping of progenies was carried out for the following characteristics, according to Silva et al. (2017): number of fruits (NF), obtained by counting the number of total fruits in each plot; total production (PROD in kg per plant) expressed in kg per plot, obtained by the quantity harvested during the experiment; fruit mass (FM in g) expressed in grams, obtained by the arithmetic mean of the mass of 15 fruits sampled per plot; fruit width (FW in mm); fruit length (FL in mm); percentage of fruit pulp (FP), obtained by the ratio between pulp mass and total fruit mass; peel thickness (PT in mm), determined in the median portion of the sliced fruits in the direction of greater diameter; total soluble solids content (TSS in °Brix), obtained through a portable digital refractometer, with a reading range from 0 to 95° Brix degrees, in pulp juice aliquots.

Statistical Analyses

The model used was the linear mixed model:

$$y = Z_1 p + Z_2 b + Z_3 r + Z_4 c + \varepsilon$$

where: y is the vector of the phenotypic values of the trait; p [1, ..., 81] is the parametric vector of random effects of progeny associated to vector y by the incidence matrix known Z_1, assuming that $p \sim N(0, I \otimes \sigma_p^2)$; b [1, 2] is the parametric vector of random effects of blocks associated to vector y by the incidence matrix known Z_2, assuming $b \sim N(0, I \otimes \sigma_b^2)$; r [1, ..., 13] is the parametric vector of random effects of rows associated to vector y by the incidence matrix known Z_3, assuming that $r \sim N(0, I \otimes \sigma_r^2)$; c [1, ..., 18] is the parametric vector of random effects of columns associated to vector y by the incidence matrix known Z_4, assuming that $c \sim N(0, I \otimes \sigma_c^2)$; ε is the vector of random residual effects, assuming that $\varepsilon \sim N(0, I \otimes \sigma_\varepsilon^2)$. We assumed that the (co)variance matrices follow an inverted Wishart distribution, which was used as *a priori* to model the variance-covariance matrix.

Next, we tested a new model, removing the row and column effects from the previous described model. These two models were named complete or Row-Col and reduced or RBD (only randomized block design). The models were compared using the likelihood ratio test (LRT) using the Chi-square test with two degrees of freedom and a 5 % probability level. After testing these two models using the frequentist approach, we also tested them with the Bayesian approach. As we could choose *a priori* in the Bayesian approach, before comparing the results of the frequentist approach with the Bayesian, we tested two *a priori* approaches for the Bayesian approach. Two sets of values for the parameters of the distributions (inverted Wishart) of the variances were tested, denominated *priors* to Bayes I and *priors* to Bayes II (Table 1) and then we chose the model that contained *a priori* with the best fit to compare with the frequentist.

The first set of values for variance distributions (Bayes I) considers the default of package described by Hadfield (2010) not very informative, with wide variance. The second set of *priori* choices (Bayes II) was justified, as suggested by Hadfield (2010), with the insertion of the alpha.mu and alpha.v extension parameters, now based on the results from the frequentist itself. This implies that we expected a result at least close to the frequentist. The model (*a priori*) in the Bayesian approach was chosen considering the smallest square root of the mean squared error, based on the lowest Deviance Information Criterion (DIC) value between the models. The models were considered equal for $|\Delta| < 2$ between the DIC values of the models (Spiegelhalter et al., 2002).

For Bayes I and Bayes II, we used the Monte Carlo method based on Markov Chains (MCMC), according to Hadfield (2010), using the *MCMCglmm::MCMglmm* package::function in *R* language. We determined 1,000,000 iterations (*nitt*), disregarding the first 100,000 (*burn-in*) and performing a 1:3 (*thin*) sampling, totaling a chain with 300,000 iterations to obtain the variance components (*a posteriori* distribution). The Markov Chain convergence was tested by the Geweke criterion (Geweke, 1991) according to Cowles and Carlin (1996) by using the *coda::geweke.diag* package::function (Plummer et al. 2006) in *R* language. The model following the frequentist approach was also adjusted using *R* language (R Core Team, v.4.3.1).

The predictive capacity of the model was also considered in a cross-validation test with five-folds using 80 % of data for training and 20 % for validation, randomly sampled in each fold. Thus, we considered the mean correlation of the predicted dependent variables and those observed from the folds of cross-validation. This same model choice approach was repeated when the frequentist and Bayesian models were compared, performing a new cross-validation test, and comparing the models by their predictive capacity and the mean squared error in the model results.

Table 1 – A priori to values of parameters of distribution for the variance components.

A priori	υ	v	alpha.mu	Alpha.v	Classification	μ
Bayes I	1	0.002	0	0	Gamma Inverse (GI)	N (0.10^8)
Bayes II	1	1	0	25^2	Half–Cauchy (HC)	N (0.10^8)

Genetic parameters

Some genetic parameters were estimated in both approaches, using the complete or reduced model, according to their fit to each variable. If the complete model was significantly different from the reduced model for any variable, the complete model was used to estimate the genetic parameters. Heritability ($\hat{h}^2$) was estimated by:

$$\hat{h}^2 = \frac{\hat{\sigma}^2_{progeny}}{\hat{\sigma}^2_{progeny} + \hat{\sigma}^2_{block} + \hat{\sigma}^2_{column} + \hat{\sigma}^2_{row} + \hat{\sigma}^2_{error}}$$

where: $\hat{\sigma}^2_{error}$ is the residual variance estimate; $\hat{\sigma}^2_{block}$ is the variance estimate due to the block factor; $\hat{\sigma}^2_{column}$ is the variance estimate due to the column factor; $\hat{\sigma}^2_{row}$ is the variance estimate due to the row factor; $\hat{\sigma}^2_{progeny}$ is the estimate of the genotypic variance.

Confidence intervals for heritability were also obtained for the frequentist approach (CI for $\hat{h}^2$) and for the confidence interval for heritability for the Bayesian approach. A ranking was made selecting 30 individuals in which the selection gain was obtained by the delta method, based on:

$$G_s = 100.\left(\frac{\acute{X}_{30} - \acute{X}_G}{\acute{X}_G}\right)$$

where: selection gain is relative to the overall mean ($\acute{X}_G$) and the mean of the first 30 (thirty) selected progenies ($\acute{X}_{30}$). The scripts used in this work are available at https://github.com/CAIOAGRO0/scripts_from_papers/blob/main/Row-Col_Bayesian_passion.rar

Results and Discussion

The complete model, which incorporates a post-hoc blocking Row-Col design by frequentist approach, showed a significant difference for the factors (Row-Col) tested by the likelihood ratio test (LRT) via Chi-Square only for the variables with the small number of observations, that is, NF and PROD (Statistics D of 16.70 and 17.10 respectively) (Table 2). Possibly, as both variables result from the total sum of the plot and are not sampled in fruits with replicates, a variance captured by Row-Col may not be necessary when more observations occur in the other variables. The comparison of the models using the Bayesian approach showed that the Row-Col model presented statistical significance for the variables NF, PROD, FP, PT, and TSS by difference between the observed DIC in models considering Row-Col and without considering ($|\Delta| < 2$ = –44.90, –42.35, –2.16, 2.70 and –8.25, respectively). This result is possibly due to the choice of an appropriate informative *a priori*, which provided more accurate estimates allowing a better fit (Table 2). As mentioned in the methods, two *priors* were tested for the Bayesian approach and only *a priori* with the best fit was considered. *A priori* Bayes II was chosen because it presented more accurate metrics in relation to *a priori* Bayes I.

Table 2 – Likelihood ratio test (LRT) for comparison between the randomized block design and post–hoc blocking Row–Col models using frequentist and Bayesian approach.

Variable	Model	Frequantist		Bayesian	
		Deviance	Statistics (D)	DIC	Δ
NF	*Row–Col*	1935.20	16.70^{*}	1898.68	-44.90^{d}
	DBC	1951.90		1943.58	
PROD	*Row–Col*	1296.50	17.10^{*}	1266.68	-42.35^{d}
	DBC	1313.40		1309.03	
FM	*Row–Col*	1553.6	0.70^{ns}	1550.31	0.81^{i}
	DBC	1554.3		1549.44	
FL	*Row–Col*	1602.00	1.00^{ns}	1036.75	1.68^{i}
	DBC	1603.00		1038.43	
FD	*Row–Col*	970.10	0.65^{ns}	942.73	1.38^{i}
	DBC	969.45		941.35	
FP	*Row–Col*	1026.5	0.83^{ns}	1019.72	-2.16^{d}
	DBC	1027.3		1021.88	
PT	*Row–Col*	496.87	0.00^{ns}	480.03	2.70^{d}
	DBC	496.87		477.33	
TSS	*Row–Col*	523.24	4.36^{ns}	496.38	-8.25^{d}
	DBC	527.60		504.63	

NF = Number of fruits; PROD = Total production; FM = Fruit mass; FL = Fruit length; FD = Fruit width; FP = Fruit pulp; PT = Peel thickness; TSS = Total Soluble Solids; DIC = Deviance Information Criterion; Difference between model DICs (Δ) = i similar models and d different models; D = Likelihood ratio test; *significant at 5 % probability level; ns = not significant.

Machado et al. (2020) evaluated 135 progenies of full siblings of passion fruit and observed superior performance of randomized block design + Row-Col, compared to the only randomized block design. Kempton et al. (1994), Silva et al. (2016), and Machado et al. (2020) found similar results for traits fruit mass, number of fruits, and TSS content. According to these authors, the randomized block design + Row-Col is a low-cost technique that improves experimental precision, favoring thus selection of superior genotypes. This technique less costly especially when compared to the Bayesian approach, which requires time and financial resources involving the renting a cloud server or maintenance on local server in addition electricity costs.

Adding *a posteriori* factors to a model may not seem an adequate approach, since this was possibly not planned before the deployment of the experiment. Overall, the randomized block design is effective when variability within the replicates is relatively small, which is rare when a large number of genotypes is evaluated, as in the case of this study and many other perennial cultures.

Another point to consider is the size experiment, which easily increases for large plants and it also increases the probability of a factor to be an error source and ignored in the model. For example, the literature does not have many reports indicating that farming can also influence the number of samples in trials with genetic improvement programs for perennial and fruit species. The problem in these experiments is the time needed to carry out the farming tasks (from planting to harvesting) or the evaluation of all the fruits of the plot. Depending on the labor force available, the time required to complete an activity may be enough to cause differences between measurements. In this study, the post-hoc blocking Row-Col model was capable of capturing some experimental errors, due to a large number of treatments, presenting promising results, and were selected to estimate the components of variance and prediction of the genetic values.

To compare frequentist and Bayesian approaches, a cross validation (5-folds) was used to assess model accuracy. In general, the metrics of predictive capacity (PC) and predictive efficiency (RMSE) showed similar values for both approaches (Table 3), indicating that both tested approaches have the same generalization capacity, regardless of the design. We expected to find a greater discrepancy between the approaches, as noted by Silva et al. (2020), but apparently a simpler approach was satisfactory. The breeder decides whether the precision obtained by the Bayesian approach corresponds to the investment of processing time and resources for the analysis. In this study, as the differences in predictive capacity were close to 0.03, when they existed, they still did not have a major effect, at least in the short term.

Table 3 – Estimates of the metrics obtained by cross-validation for model fit by frequentist and Bayesiana approaches in models considering Row–Cow design when significative.

Variable	Model	PC	RMSE
NF	Frequentist*	0.91	54.43
	Bayesian*	0.90	61.33
PROD	Frequentist*	0.89	7.98
	Bayesian*	0.89	8.19
FM	Frequentist	0.76	22.82
	Bayesian	0.73	23.58
FL	Frequentist	0.86	4.22
	Bayesian	0.85	4.29
FD	Frequentist	0.88	2.97
	Bayesian	0.88	3.00
PT	Frequentist	0.87	0.74
	Bayesian	0.87	0.75
FP	Frequentist	0.77	4.43
	Bayesian*	0.81	4.12
TSS	Frequentist	0.87	0.77
	Bayesian*	0.90	0.71

NF = Number of fruits; PROD = Total production; FM = Fruit mass; FL = Fruit length; FD = Fruit width; PT = Peel thickness; FP = Fruit pulp; TSS = Total Soluble Solids; *model considering Row–Cow design; PC = Predictive capacity of the model; RMSE = Square root of the mean squared error of the model.

The Bayesian approach tends to converge with the frequentist approach REML/BLUP even when little informative *prior* is provided and when the data satisfactorily represents a population. Thus, this similar performance between the two approaches may not compensate for the choice of the Bayesian approach, as observed by Roh et al. (2004). The Bayesian approach can always provide more accurate estimates with shorter credibility intervals (equivalent to confidence intervals for frequentist), providing basis for better decision-making. Other authors, such as Alijani et al. (2012), decided that the superior performance for Bayesian methodology compared to the frequentist was worth the investment. Silva et al. (2020) evaluated 17 full-sibling families of guava and observed the same results. These authors emphasize that the Bayesian methodology was pronounced in situations with small data sets and/or incomplete data.

Similarity in the predictive capacity of the models shows an advantage that the breeder may have using the Bayesian approach. Here, we do not use *a priori* based on data from previous experiments. However, based only on the results of a frequentist model, we managed to have a similar and even better predictive capacity for some variables, such as FP (0.77 for frequentist and 0.81 Bayesian) and TSS (0.87 for frequenters and 0.90 Bayesian, Table 3). Further investments in processing time confer confidence to the breeder in their planning due to more accurate estimates and the large data volume with the Bayesian approach increase accuracy of the predictive capacity of the model.

Estimates of some genetic parameters were performed for the population. Heritability and selection gain showed satisfactory values for the variable productivity and number of fruits (Table 4). Variances for the effects were slightly smaller using the Bayesian approach (Compare Table 4 and Table 5). This result

Table 4 – Estimation of genetic parameters of 81 passion fruit progenies of full siblings of the third cycle of recurrent selection via mixed models restricted maximum likelihood/best linear unbiased predictor.

Var.	$\hat{\sigma}^2_{progeny}$	$\hat{\sigma}^2_{block}$	$\hat{\sigma}^2_{row}$	$\hat{\sigma}^2_{col}$	$\hat{\sigma}^2_{error}$	$\hat{h}^2$	IC– $\hat{h}^2$	G_s
NF	4410.00	0.00	2156.00	0.00	4651.00	0.39	0.22; 0.56	34.12
PROD	66.28	1.44	44.74	0.00	99.78	0.31	0.14; 0.48	28.17
FM	200.96	87.37	–	–	665.67	0.21	0.01; 0.41	5.38
FL	18.59	7.99	–	–	25.81	0.35	0.13; 0.58	4.09
FD	12.00	1.88	–	–	13.90	0.43	0.25; 0.62	3.57
FP	8.16	2.52	–	–	25.54	0.23	0.03; 0.42	4.47
PT	0.58	0.01	–	–	0.81	0.42	0.24; 0.60	–7.31
TSS	0.71	0.09	–	–	0.95	0.41	0.22; 0.59	4.50

NF = Number of fruits; PROD = Total production; FM = Fruit mass; FL = Fruit length; FD = Fruit width; PT = Peel thickness; FP = Fruit pulp; TSS = Total Soluble Solids; $\hat{\sigma}^2_{progeny}$ is the estimate of genetic variance; $\hat{\sigma}^2_{error}$ estimate of residual variance; $\hat{\sigma}^2_{block}$ estimate of variance due to the block factor; $\hat{\sigma}^2_{col}$ estimate of variance due to column factor; $\hat{\sigma}^2_{row}$ estimate of variance due to row factor; $\hat{h}^2$ heritability; (G_s) selection gain expressed in %; (IC– $\hat{h}^2$) confidence interval for heritability at 95 % probability level.

Table 5 – Estimation of genetic parameters of 81 passion fruit progenies of the third cycle of recurrent selection via Bayesian Inference.

Var.	$\hat{\sigma}^2_{progeny}$	$\hat{\sigma}^2_{block}$	$\hat{\sigma}^2_{row}$	$\hat{\sigma}^2_{col}$	$\hat{\sigma}^2_{error}$	$\hat{h}^2$	G_s	HPD
NF	4178.00	2873.00	2075.00	147.10	4975.00	0.35	31.98	0.15; 0.53
PROD	67.27	1559.00	53.07	5.61	102.20	0.21	27.00	0.00; 0.39
FM	180.00	3658.00	–	–	702.20	0.14	4.61	0.00; 0.31
FL	19.14	1847.00	–	–	26.45	0.18	4.02	0.00; 0.43
FD	12.41	2183.00	–	–	14.21	0.22	5.53	0.00; 0.50
FP	8.20	1339.00	2.00	1.39	24.89	0.11	4.21	0.00; 0.31
PT	0.60	1067.00	–	–	0.84	0.19	–7.16	0.00; 0.48
TSS	0.67	868.60	0.06	0.18	0.88	0.14	4.01	0.00; 0.41

NF = Number of fruits; PROD = Total production; FM = Fruit mass; FL = Fruit length; FD = Fruit width; FP = Fruit pulp; PT = Peel thickness; TSS = Total Soluble Solids; $\hat{\sigma}^2_{progeny}$ is the estimate of the genetic variance; $\hat{\sigma}^2_{error}$ estimate of residual variance; $\hat{\sigma}^2_{block}$ estimate of variance due to the block factor; $\hat{\sigma}^2_{col}$ estimate of variance due to column factor; $\hat{\sigma}^2_{row}$ estimate of variance due to row factor; $\hat{h}^2$ heritability obtained by the a posteriori distribution; G_s = selection gain, expressed in %; (HPD – $\hat{h}^2$) highest posterior density is the credibility interval for heritability at the 95 % probability level.

highlights the existence of genetic variability and selective potential between passion fruit progenies. The confidence intervals (CI) (Table 4) for heritability have lower amplitudes than the Highest Posterior Density (HPD) (Table 5). This difference in the amplitude of HPD and CI, except *a posteriori* for NF, is because the *a posteriori* distribution for heritability is bimodal, and the HPD amplitude may be lower as it improves *a priori*, making it more informative with results from previous experiments.

Heritability estimates ranged from 0.21 to 0.43 in the frequentist approach. The highest values were obtained for the variables FD, PT, TSS, and NF (0.43, 0.42, 0.41, and 0.39 respectively). These estimates presented values within the expected range for these traits, considering they are controlled by a large number of genes and are highly influenced by the environment (Resende et al., 2014). Cavalcante et al. (2019) found similar estimates for the same crop. Viana et al. (2004) evaluated the same variables in passion fruit and found higher heritability estimates for all variables except TSS, indicating a favorable situation for selection. Smaller estimates may be due to the selection cycle. In this study, we used data from the third recurrent selection cycle and the plants are more homogeneous by increasing the frequency of favorable alleles for the desirable traits.

The variables of greatest interest, NF and PROD, had the highest selection gains, 34 % and 28 %, respectively. We can observe a positive gain for all variables, except for PT, which showed a negative gain, according to expectations. The PT variable showed a negative gain of –7 %, which is interesting as a decrease of the expression of this variable is expected. Selection gains for variables FM (5 %), LDF (4 %), TDF (3 %), FP (4 %), and TSS (4 %) presented low estimates, possibly due to the successive selection cycles; however, these results proved the efficiency of the continuous selection in a recurrent program for passion fruit.

Thus, we proceeded with the selection of the best individuals in a ranking, selecting the 30 best individuals. The promising progenies were kept in both approaches, only their ordering was changed (Tables 6 and 7).

Table 6 – Ranking of the first 30 passion fruit progenies from full siblings of the third cycle of recurrent selection obtained from the mean value added to the predicted genotypic value (u + g) through a frequentist approach using restricted maximum likelihood/best linear unbiased predictor for each variable.

Rank	NF	PROD	FM	FL	FD	FP	PT	TSS
1	56	26	69	68	3	81	7	6
2	26	56	5	8	8	19	54	75
3	1	47	3	44	27	54	26	38
4	11	11	8	27	42	68	27	53
5	24	1	68	5	48	6	76	69
6	47	24	42	41	41	17	37	68
7	68	68	47	69	69	18	18	44
8	38	70	72	40	5	41	50	35
9	16	16	75	37	78	28	29	63
10	70	38	51	7	60	36	28	47
11	9	31	78	23	68	50	14	52
12	23	9	2	52	2	30	23	51
13	34	22	41	10	72	45	6	56
14	10	69	6	73	52	65	36	66
15	5	48	33	13	53	7	81	79
16	31	23	7	42	75	16	70	11
17	22	10	53	53	50	35	12	5
18	60	60	73	48	24	34	34	20
19	48	5	1	32	9	40	19	25
20	15	15	24	29	59	67	41	48
21	19	34	52	35	11	23	43	8
22	12	19	44	75	47	57	64	50
23	69	27	29	78	71	9	8	43
24	79	78	59	4	23	25	16	21
25	27	50	48	54	33	29	68	33
26	80	79	27	72	61	43	55	39
27	67	36	31	24	7	64	65	2
28	28	80	35	59	22	69	69	60
29	54	75	60	80	13	20	62	65
30	45	41	9	50	51	79	59	41
$\acute{X}_G$	160.97	22.18	167.47	82.35	74.75	42.25	7.93	14.20
$\acute{X}_{30}$	215.89	28.43	176.48	85.71	77.42	44.14	7.35	14.84

NF = Number of fruits; PROD = Total production; FM = Fruit mass; FL = Fruit length; FD = Fruit width; PT = Peel thickness; FP = Fruit pulp; TSS = Total Soluble Solids; ($\acute{X}_G$) is the overall mean; ($\acute{X}_{30}$) mean of the first 30 (thirty) selected progenies.

Table 7 – Ranking of the first 30 passion fruit progenies from full siblings of the third cycle of recurrent selection obtained from the mean value added to the predicted genotypic value (u + g) using the Bayesian approach for each variable.

Rank	NF	PROD	FM	FL	FD	FP	PT	TSS
1	56	26	69	68	3	19	7	6
2	26	56	5	8	8	68	54	44
3	1	47	3	44	27	81	26	53
4	11	11	8	27	42	54	27	35
5	47	1	68	5	48	6	76	75
6	24	24	42	41	41	17	37	38
7	68	70	47	69	69	41	18	68
8	38	68	72	40	5	18	50	63
9	70	16	75	37	78	28	29	69
10	16	38	51	7	60	30	28	47
11	9	31	78	23	68	36	14	66
12	23	9	2	52	2	65	23	5
13	34	22	41	10	72	45	6	52
14	10	69	6	73	52	50	36	51
15	22	48	33	13	53	16	81	8
16	31	23	7	42	75	35	70	56
17	5	10	53	53	50	40	12	20
18	60	19	73	48	24	7	34	50
19	48	34	1	32	9	67	19	79
20	19	27	24	29	59	57	41	25
21	12	60	52	35	11	34	43	11
22	15	5	44	75	47	23	64	48
23	69	15	29	78	71	79	8	21
24	79	79	59	4	23	43	16	39
25	27	36	48	72	33	69	68	43
26	80	50	27	54	61	9	55	2
27	67	78	31	59	7	29	65	41
28	45	80	35	24	22	25	69	59
29	36	75	60	80	13	26	62	60
30	54	44	9	50	51	20	59	17
$\acute{X}_G$	160.60	22.03	167.30	82.33	74.76	42.30	7.96	14.14
$\acute{X}_{30}$	211.96	27.98	175.02	85.65	77.70	44.08	7.39	14.72

NF = Number of fruits; PROD = Total production; FM = Fruit mass; FL = Fruit length; FD = Fruit width; PT = Peel thickness; FP = Fruit pulp; TSS = Total Soluble Solids; ($\acute{X}_G$) is the overall mean, ($\acute{X}_{30}$) mean of the first 30 (thirty) selected progenies.

Conclusion

The model with the post-hoc blocking Row-Col design captured the spatial variability for productivity and number of fruits traits, influencing directly the experimental precision. It indicates that the technique can be recommended for selection in the genetic improvement of passion fruit, mainly for variables with a small number of observations and large experiments.

Both approaches applied to the models showed similar performance, with predictive capacity and selective efficiency leading to the selection of the same individuals. As both approaches tended to converge, the advantage of Bayesian inference described in the literature with little or unbalanced data was not necessary for this experiment.

Acknowledgments

This study was financed in part by the Coordination for the Improvement of Higher Level Personnel – Brazil (CAPES) – Finance Code 001, and Carlos Chagas Filho Foundation for Research Support in the State of Rio de Janeiro (FAPERJ) – E-26/010.001454/2019.

Authors' Contributions

Conceptualization: Souza, A.O.; Viana, A.P.; **Data acquisition:** Silva, F.H.L.; **Data analysis:** Silva, F.H.L.; **Design of methodology:** Silva, F.F.; Azevedo, C.F.; **Formal analysis:** Silva, F.F.; Azevedo, C.F.; **Writing and editing:** Souza, A.O.; Silva, F.A.; Viana, A.P.

References

Alijani, S.; Jasouri, M.; Pirany, N.; Kia, H.D. 2012. Estimation of variance components for some production traits of Iranian Holstein dairy cattle using Bayesian and AI-REML methods. Pakistan Veterinary Journal 32: 562-566.

Bernacci, L.C.; Cervi, A.C.; Milward-De-Azevedo, M.A.; Nunes, T.S.; Imig, D.C.; Mezzonato, A.C. 2015. Passifloraceae. In: List of species of flora of Brazil. Botanical Garden of Rio de Janeiro = Passifloraceae. In: Lista de espécies da Flora do Brasil. Jardim Botânico do Rio de Janeiro. Available at: http://floradobrasil.jbrj.gov.br/jabot/floradobrasil/FB182 (Accessed Jan 21, 2020) (in Portuguese).

Cavalcante, N.R.; Viana, A.P.; Almeida Filho, J.E.; Pereira, M.G.; Ambrósio, M.; Santos, E.A.; Ribeiro, R.M.; Rodrigues, D.L.; Sousa, C.M.B. 2019. Novel selection strategy for half-sib families of sour passion fruit Passiflora edulis (*Passifloraceae*) under recurrent selection. Genetics and Molecular Research 18: gmr18305. https://doi.org/10.4238/gmr18305

Cowles, M.K.; Carlin, B.P 1996. Markov chain Monte Carlo convergence diagnostics: a comparative review Journal of the American Statistical Association 91: 883-904. https://doi.org/10.2307/2291683

Geweke, J. 1991. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. Federal Reserve Bank of Minneapolis, Minneapolis, MN, USA.

Gezan, S.A.; Huber, D.A.; White, T.L. 2006. Post hoc blocking to improve heritability and precision of best linear unbiased genetic predictions. Canadian Journal of Forest Research 36: 2141-2147.

Hadfield, J.D. 2010. MCMCglmm: MCMC methods for multi-response GLMMs in R. Journal of Statistical Software 33: 1-22. https://doi.org/10.1002/ana.22635

Kempton, R.A.; Seraphin, J.C.; Sword, A.M. 1994. Statistical analysis of two-dimensional variation in variety yield trials. Journal of Agricultural Science 122: 335-342.

Machado, R.F.C.; Krause, W.; Silva, C.A.; Fachi, L.R.; Silva, F.H.L.; Viana, A.P. 2020 Implications of the post hoc blocking row-col technique on the intrapopulational improvement of the passion fruit. Euphytica 216: 62. https://doi.org/10.1007/s10681-020-02595-w

Meletti, L.M.M. 2011. Advances in passion fruit culture in Brazil. Revista Brasileira de Fruticultura. 33: 83-91. https://doi.org/10.1590/0100-29452019155

Plummer, M.; Best, N.; Cowles, K.; Vines, K. 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6: 7-11.

Resende, M.D.V.; Silva, F.F.; Azevedo, C.F. 2014. Mathematical, Biometric and Computational Statistics: Mixed, Multivariate, Categorical and Generalized Models (REML/BLUP), Bayesian Inference, Random Regression, Genomic Selection, QTLGWAS, Spatial and Temporal Statistics, Competition, Survival = Estatística Matemática, Biométrica e Computacional: Modelos Mistos, Multivariados, Categóricos e Generalizados (REML/BLUP), Inferência Bayesiana, Regressão Aleatória, Seleção Genômica, QTL-GWAS, Estatística Espacial e Temporal, Competição, Sobrevivência. Suprema Editora, Visconde do Rio Branco, MG, Brazil (in Portuguese).

Roh, S.H.; Kim, B.W.; Kim, H.S.; Min, H.S.; Yoon, H.B.; Lee, D.H.; Jeon, J.T.; Lee, J.G. 2004. Comparison between REML and Bayesian via Gibbs sampling algorithm with a mixed animal model to estimate genetic parameters for carcass traits in Hanwoo (Korean native cattle). Journal of Animal Science and Technology 5: 719-728.

Silva, F.A.; Viana, A.P.; Corrêa, C.C.G.; Carvalho, B.M.; Sousa, C.M.B.; Amaral, B.D.; Ambrósio, M.; Glória, L.S. 2020. Impact of bayesian inference on the selection of Psidium guajava. Scientific Reports 10: 1-9.

Silva, J.A.; Sanches, A.; Carla, A.; Andrade, B. 2018. Bayesian approach, traditional method, and mixed models for multienvironment trials of soybean. Pesquisa Agropecuária Brasileira 53: 1093-1100. https://doi.org/10.1590/S0100-204X2018001000002

Silva, F.H.L.; Viana, A.P.; Freitas, J.C.O.; Santos, E.A.; Rodrigues, D.L.; Amaral Junior, A.T. 2017. Prediction of genetic gains by selection indexes and REML/BLUP methodology in a population of sour passion fruit under recurrent selection. Acta Scientiarum-Agronomy 39: 183-190.

Silva, F.H.L.; Muñoz, P.R.; Vincent, C.I. 2016. Generating relevant information for breeding *Passiflora edulis*: genetic parameters and population structure. Euphytica 208: 609-619. https://doi.org/10.1007/s10681-015-1616-8

Silva, F.F.; Viana, J.M.; Faria, V.R.; Resende, M.D. 2013. Bayesian inference of mixed models in quantitative genetics of crop species. Theoretical and Applied Genetics 7: 1749-1761. https://doi.org/10.1007/s00122-013-2089-6

Silva, M.G.M.; Viana, A.P.; Gonçalves, G.M.; Amaral Júnior, A.T.; Pereria, M.G. 2009. Recurrent intra-population selection in yellow passion fruit: Alternative to capitalize on genetic gains. Ciência e Agrotecnologia 33: 170-176 (in Portuguese, with abstract in English).

Spiegelhalter, D.J.; Best, N.; Carlin, B.; Van der Linde, A. 2002. Bayesian measures of model complexity and fit. Journal of The Royal Statistical Society 64: 583-640.

Viana, A.P.; Pereira, T.N.S.; Pereira, M.G.; Amaral Júnior, A.T.; Souza, M.M.; Maldonado, J.F.M. 2004. Genetic parameters in yellow passion fruit populations. Revista Ceres 51: 545-555.