



# Consideration of the appropriate variation sources of the statistical model and their impacts on plant breeding

Rita Carolina de Melo<sup>1\*</sup>, Jefferson Luís Meirelles Coimbra<sup>1</sup>, Nicole Trevisani<sup>1</sup>, Altamir Frederico Guidolin<sup>1</sup> and Sydney Antonio Frehner Kavalco<sup>2</sup>

<sup>1</sup>Departamento de Agronomia, Instituto de Melhoramento e Genética Molecular, Universidade do Estado de Santa Catarina, Avenida Camões, 2090, 88520-000, Lages, Santa Catarina, Brazil. <sup>2</sup>Centro de Pesquisa para Agricultura Familiar, Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina, Chapecó, Santa Catarina, Brazil. \*Author for correspondence. E-mail: rita\_carol\_mel@hotmail.com

**ABSTRACT.** The present work has aimed to assess the consideration of the appropriate variation sources of the statistical model and their impacts on the conclusions plant breeding. The Value for Cultivation and Use test was conducted to assess three common locations (Lages, Ponte Serrada, and Canoinhas) and four non-common locations (Chapecó, Guatambu, Urussanga, and Campos Novos). The grain yields of six bean genotypes were evaluated in order to represent the imbalance between the common and non-common locations. The statistical analysis considered two situations: *i*) union of the location factors and cultivation years, with a single variation source called *environment* and *ii*) decomposition of the mean square values of the two factors, location and year. According to the simplified analysis (environmental variation source), the F test for the genotype factor was highly significant ( $p = 0.0006$ ). On the other hand, the hypothesis test for the genotype factor was not significant ( $p = 0.7370$ ) when the decomposition of mean squares was used. The simplified analysis presents some erroneous points, such as the use of a mean residue to test the hypothesis of the genotype factor, since this factor is composed of several sources of variation, and there is no exact F test. However, approximate F tests can be obtained by constructing linear combinations of average squares. This fact notes the relevance of considering the appropriate sources of variation within the statistical model, with a direct impact on the conclusions and recommendations of cultivars with superior performance.

**Keywords:** VCU trials; variance analysis; mathematical expectations.

Received on October 3, 2017.  
Accepted on December 15, 2017.

## Introduction

In Brazil, the cultivation of beans (*Phaseolus vulgaris* L.) is carried out by small, medium and large agricultural producers. This classification can be applied for area extension as well as the level of investment applied to the cultivation of this important agricultural product (IBGE, 2015). However, the national average grain yield is low (approximately 1,000 kg ha<sup>-1</sup>) and varies greatly by region and agricultural year (CONAB, 2016). In the State of Santa Catarina, the performance of genotypes recommended for cultivation can vary by more than 100% (1,925 to 3,885 kg ha<sup>-1</sup>).

A number of trials must be carried out in various cultivation locations and/or several agricultural years (Dias, Pitombeira, Teófilo, & Barbosa, 2009; Schmildt, Nascimento, Cruz, & Oliveira, 2011; Nassir & Ariyo, 2011; Silva et al., 2013; Boukid et al., 2017; Carvalho, Damasceno-Silva, Rocha, & Oliveira, 2017) in order to meet the demand for more productive and stable cultivars. This requirement is prescribed by the Ministry of Agriculture, Livestock and Supply (MAPA), the federal agency responsible for indicating the minimum requirements for registration at the National Cultivars Registry (RNC). The determination of the Value for Cultivation and Use (VCU) is one of the most important requirements for the protection and/or registration of new genetically superior constitutions. Briefly, in the case of bean cultivation, the minimum requirements for the determination of VCU establish that researchers must evaluate lineages in at least three locations (of significance for the culture), in a minimum period of two agricultural years.

However, the data imbalance is a common feature in these tests, when a historical series of years is considered. Certain genotypes start to show relative importance compared to others. Under such a condition, they are not conducted in certain locations in the following year (Yan, 2015). Therefore, the

locations vary from year to year, and only a certain number of locations that are common to all years remain, generating unbalanced data (Ignaczak & Silva, 1978).

The factorial analysis for a set of trials is easy to apply and widely informative, with balanced data. However, it can lead researchers to great misunderstanding in the presence of imbalanced data (Wechsler, 1998). Under such a condition (imbalanced factorials), the orthogonality is destroyed, and the calculations of the sums of squares become much more complex (Wechsler, 1998). In addition, the phenotypic value may not be an accurate estimate of the true genetic value (Borges et al., 2009). Several strategies are widely used for the imbalanced data analysis, including the approach in which the factors location and year are put together and the synthesis over *environments* is more frequently used (Bertoldo et al., 2009; Gupta et al., 2013). However, summarizing over years or locations is not always appropriate when the standards of each year and each location are complex and variable. In addition, this approach disregards the interactions between location, genotype and year and requires a subjective synthesis of the results (Yan, 2015).

Alternatively, the variance of the factors with imbalanced treatments (non-common locations) can be decomposed into three components: i) variance between the common treatments, ii) variance between the non-common treatments, and iii) variance between common and non-common treatments (Ignaczak & Silva, 1978). The assessment of these components using the analysis of variance with imbalanced data can contribute to the true estimation of the genetic value and affect the decisions of a plant breeder. Therefore, the present work has aimed to verify the consideration of the appropriate variation sources within the statistical model in Value for Cultivation and Use test and the impacts on the conclusions related to plant breeding.

## Material and methods

### Value for Cultivation and Use (VCU) test

The analyses considered six common bean genotypes in all locations: BRS Campeiro, IPR Uirapuru, CHP 01-238, FTs 1, LP 09-40, and LP 09-181. The trials were conducted in seven locations in the State of Santa Catarina during two agricultural years (2012/13 and 2013/14). They came from an experimental network constituted by the Agricultural Research and Rural Extension Company of Santa Catarina (EPAGRI), through the Value for Cultivation and Use of Bean (VCU) test, together with the Universidade do Estado de Santa Catarina (UDESC):

- i) 3 Locations common to (LC) years: Lages (LG), Ponte Serrada (PS), and Canoinhas (CA);
- ii) 4 Locations Non-common to (LNC) years: Chapecó (CH), Guatambu (GT), Urussanga (UR), and Campos Novos (CN).

Thus, the genotypes were evaluated in 10 different environments: Canoinhas/2013, Canoinhas/2014, Lages/2013, Lages/2014, Ponte Serrada/2013, Ponte Serrada/2014, Chapecó/2013, Urussanga/2013, Campos Novos/2014, and Guatambu/2014.

### Experimental Design

The Ministry of Agriculture, Livestock and Supply (MAPA) provides the minimum requirements for launching bean cultivars, including the arrangement of the experiment in a randomized block design with four replicates. The experimental unit was composed of four lines of four meters in length, with a spacing of 0.50 m between rows, and a density of 15 seeds per linear meter. The useful area consisted of the two central lines with a 0.50 m border. The trait grain yield was measured in kilograms per hectare ( $\text{kg ha}^{-1}$ ) of each plot.

### Data Statistical Analysis

The data were analyzed considering the following.

- i) Union of the local factors and cultivation years, with a single variation source called environment:

$$Y_{ijk} = \mu + B_i + G_j + AM_k + G^*AM_{jk} + e_{ijk}$$

where:  $Y_{ijk}$  is the average of the variable response of the genotype  $j$ , in block  $i$ , environment  $k$ ;  $\mu$  refers to a general mean effect;  $B_i$  is the fixed effect of the block  $i$ ;  $G_j$  is the fixed effect of the genotype  $j$ ;  $AM_k$  is the fixed effect of the environment  $k$ ;  $G^*AM_{jk}$  is the interaction of the  $j$ th level of the factor genotype and  $k$ th level of the factor environment; and  $e_{ijk}$  refers to the residue effect.

ii) Decomposition of the mean squares for the location and year factors:

$$Y_{ijkl} = \mu + B_i + L_j + A_k + G_l + (LA)_{jk} + (LG)_{jl} + (AG)_{kl} + (LAG)_{jkl} + e_{ijkl}$$

where:  $Y_{ijkl}$  is the average of the variable response of the genotype  $l$ , in block  $i$ , in location  $j$  and in the year  $k$ ;  $\mu$  refers to a general mean effect;  $B_i$  is the fixed effect of the block  $i$ ;  $L_j$  is the fixed effect of the location  $j$ ;  $A_k$  is the fixed effect of the year  $k$ ;  $G_l$  is the fixed effect of the genotype  $l$ ;  $(LA)_{jk}$  is the fixed effect of the interaction between location  $j$  and year  $k$ ;  $(LG)_{jl}$  is the fixed effect of the interaction between location  $j$  and genotype  $l$ ;  $(AG)_{kl}$  is the fixed effect of the interaction between the year  $k$  and genotype  $l$ ;  $(LAG)_{jkl}$  is the fixed effect of the interaction between location  $j$  and year  $k$  and genotype  $l$ ; and  $e_{ijkl}$  refers to the residue effect.

In addition, due to data imbalance (non-common locations), a decomposition of imbalanced factors was adopted, as proposed by Ignaczak and Silva (1978), in which the variance of the factors with imbalanced treatments and their interactions with the other factors were decomposed into three components: *i*) variance between the common treatments; *ii*) variance between the non-common treatments; and *iii*) variance between common and non-common treatments.

The expectations of the mean squares were used to verify the mean square of the appropriate residue for the analysis of each factor and the respective approximate F test. Contrasts of interest between the locations were performed within the variance of *Common Locations* and *Non-common Locations*. All the analyses were performed using the SAS software (Statistical Analysis System) and the *Proc glm* command (Littell, Milliken, Stroup, Wolfinger, & Shabenberger, 2006).

## Results and discussion

### Union of the location and cultivation year factors, with a single variation source called environment

The genotypes candidates for the launching of cultivars are usually evaluated in different *environments*. In the present situation, or the introduced model, the genotypes were evaluated in 10 different environments. Under such conditions, the environmental factor was significant and showed a significant mean square (283875748), which is considered the greatest magnitude among all the main effects of the analysis of variance (Table 1). This result can be explained in three ways, which may or may not occur simultaneously: *i*) significant variance of the effect between the years, *ii*) significant variance of the effect between the locations, and *iii*) significant variance of the effect between years and locations.

**Table 1.** Analysis of variance for grain yield (kg ha<sup>-1</sup>) from a group of ten trials using bean genotypes evaluated in the Value for Cultivation and Use (VCU) test. The environmental variation is composed of the union of the location and agricultural year factors. UDESC-IMEGEM, Lages, Santa Catarina, 2017.

Cause of variation	Degrees of freedom	Average squared	P-value
Block	3	342571	0.1535
Environment	9	283875748	0.0006
Genotype	5	881941	< 0.0001
Genotype x environment	45	384798	0.0008
Experimental error	177	192894	
	R <sup>2</sup> = 0.89	CV = 16.44%	$\bar{Y} = 2,671 \text{ kg ha}^{-1}$

Several studies have demonstrated that the environmental factor strongly affects grain yield (Bertoldo et al., 2009; Coimbra et al., 2009), which was corroborated in this study. The great challenge of plant breeding is elucidating the participation of this factor in the genotypic performance in order to adjust the genetic constitutions to the environment (Vencovsky, Ramalho, & Toledo, 2012). However, the alternation of genotypes and locations in breeding cycles is often observed, especially in annual and fast-cycle species, such as beans (Pereira et al., 2016). Therefore, locations can affect the performance of certain genetic constitutions, while years may also present superior effects. Under these conditions, breeders cannot attribute real causes to the greatest or lowest genotypic performance, which directly affects their decisions. Finally, will the genotypic performance be stable in the next year and/or in other regions (Gupta et al., 2013).

Table 1 demonstrates that the F test for the factor genotype was highly significant ( $p = 0.0006$ ). In other words, the six genotypes evaluated probably show a different agronomic performance when assessed or

tested in these environments. This is an important fact, considering the explicit need for variant genotypes that meet the specific and regional demands of a genetic breeding program (Bertoldo et al., 2009). The evaluated genotypes also respond immediately in different proportions to changes in the environments. This reveals different grain yield standards that may affect the changes in the classification of these genotypes, according to the location and agricultural year assessed (Table 1).

At first, the analysis of variance considering the union of the location and year factors is informative. However, detailed information can be obtained by decomposing the causes of variation for the imbalance when the experiments are joined with only part of the genotypes common to all the locations (Ignaczak & Silva, 1978). Therefore, the variance of the environment (9 degrees of freedom) can be decomposed: i) variance of the years (1 degree of freedom), ii) variance of the locations (6 degrees of freedom), and iii) variance of the interaction between years and common locations (2 degrees of freedom) (Table 2). This decomposition allowed for the understanding of the effect of the imbalanced factors on the analysis of variance and the ability to infer the relevant causes of variation.

### Decomposition of the mean squares for the location and year factors

The decomposition of the mean squares and their subsequent consideration in the mathematical model yielded different results, compared to the analysis of variance using the union of the location and year factors. It also offered additional information on the effects of the factors and their interactions on the behavior of genotypes, as well as their influence on plant breeding.

Undoubtedly, the great discrepancy between the two situations proposed in this paper refers to the effect of genotypes on grain yield. In the present model, the genotype effect (G) showed no significant difference with the decomposition of the mean squares (Table 2). It disagrees with what is observed in the analysis of variance, considering the union of the location and year factors.

Likewise, the effect of the year on grain yield was not significant (Table 2). The genotypes showed a stable behavior over the years. Comparing to the first proposed analysis (union of the location and year factors), the researcher would be induced to believe that years and locations present simultaneous activity on grain yield.

**Table 2.** Decomposition of the variance analysis for the grain yield ( $\text{kg ha}^{-1}$ ) of a group of ten experiments of bean genotypes evaluated by the Value for Cultivation and Use (VCU) test. UDESC-IMEGEM, Lages, Santa Catarina, 2017.

Causes of variation	Degrees of freedom	Average squared	Percentage
Block (B)	3	342571 <sup>ns</sup>	-
Genotype (G)	5	881941 <sup>ns</sup>	-
Years (A) <sub>p/LC</sub>	1	4763306 <sup>ns</sup>	-
Locations (L)	(6)	41762639 <sup>*</sup>	-
Common (LC)	2	47049643 <sup>*</sup>	38%
Non-common (LNC)	3	19483210 <sup>*</sup>	23%
LC x LNC	1	98026914 <sup>*</sup>	39%
A x LC	2	74514 <sup>ns</sup>	-
G x A <sub>p/LC</sub>	5	867749 <sup>*</sup>	-
G x L	(30)	358477 <sup>ns</sup>	-
G x LC	10	361466 <sup>ns</sup>	34%
G x LNC	15	408868 <sup>ns</sup>	57%
G x (LC vs. LNC)	5	201330 <sup>ns</sup>	9%
G x A x LC	10	222284 <sup>ns</sup>	-
Average experimental error	177	192894	
Total	239		

\*Significant at 5% error probability by the F test.

Therefore, it is noted that factor decomposition provides very important additional information. The year is undoubtedly the most unpredictable factor in the experimental research (Lobell & Gourdjii, 2012; Assefa et al., 2015), and since grain yield is a strictly quantitative trait, the year effect may be a hindrance to the work of breeders. With such information, researchers may perform a more efficient selection, since the years of research were enough to discriminate the genotypes.

Contrarily, the effect of locations (6 degrees of freedom) was more significant in the analysis of variance (QM = 41762639) compared to the other main effects. The decomposition of this factor, with its respective degrees of freedom, into *Common Locations (LC)*, *Non-common Locations (LNC)* and their interaction (*LC x*

LNC) also showed a significant effect. LC presents the greatest relative importance, accounting for 38% of the local variance, to the detriment of the LNCs, which contributed only 23% of the variance. In other words, the most significant effect of the location factor was closely related to the common locations (LG, PS, and CA) in the analysis of variance. The LC x LNC interaction, in turn, presented the greatest mean squared value of the analysis of variance (98026914) and a contribution of 39% in the variance, mainly due to the LCs (Table 2).

These results suggest that plant breeding must meet the specific needs of each agricultural region by considering the characteristics of each municipality and the intrinsic variations of each environment (Bertoldo et al., 2009). Other very useful information for plant breeding is to know in which locations breeders can conduct and evaluate their genetic constitutions in such a way to avoid strong phenotypic interference caused by the environmental effect (Coimbra et al., 2009). Therefore, it is evident that the common locations have greater power to discriminate genotypes, due to their greater relative importance in the total of the variance of locations, to the detriment of the non-common locations, which may mask the genetic effects and lead to inconsistent genotype classifications.

Regarding the effect of the interaction between the factors  $G \times A_{p/LC}$ , a significant difference was also observed at a 5% probability of error. When the genotypes were evaluated in common locations over the years, they revealed a different behavior for grain yield. Again, this fact can be explained by the significant contribution of the LCs in the analysis of variance (Table 2). The effect of the genotype x environment interaction describes the different behavior of the genotypes in the occurrence of contrasting environments. Its effect can make the different genotypes assessed provide a huge diversity of standards and results (Piepho, Herndl, Pötsch, & Bahn, 2017). Contrarily, the interaction between the genotypes and locations ( $G \times L$ ) was not significant (Table 2). Only the main effect of location contributed significantly to grain yield, while the genotypes were not discriminated, either in common or non-common locations.

One of the main objectives of a breeding program is the recommendation of new genetic constitutions fit to the cultivation environments (Vencovsky et al., 2012). The analyses proposed in this work reveal that biased interpretations of the genotype effect can be caused by the variation between the common and non-common locations (imbalanced factor), whose decomposition was not considered in the first analysis and that changed the residue estimation. However, a question arises: which analysis proposes the true variation causes for the hypothesis tests of this trial? The comparative analysis of the results reveals that the analysis considering the union of the location and year factors presents some erroneous aspects.

The first incongruent aspect in the analysis of variance that considers the union of the location and year factors refers to the denominator for the preparation of the hypothesis test. The F test is calculated based on the mean square of the residue (192894) for all factors. However, the inspection of the mathematical expectations of the mean squares ( $E(MS)$ ), related to the unit effects, allowed verifying the residue most appropriate to the hypothesis test, according to the inferences that the experiment aims to derive (Silva, 1999; Coimbra et al., 2006). The analysis of the mathematical expectations of the mean squares ( $E(MS)$ ) corresponds to what is expected to occur with the populational average of the response variable that is focused on. The genotype factor, for example, is composed of several variation sources, such as the following interactions: *genotype x year x location*, *genotype x year* and *genotype x location*. Therefore, it cannot be equated with the expectations of the mean squares of the residue. Similar to the genotype factor, the effects of years (A) and locations (L), as well as their decompositions (common and non-common), also lack a mathematical expectation that provides an exact F test.

In such cases, it is necessary to construct linear combinations of mean squares to obtain the respective hypothesis test. Therefore, the hypothesis test based on the use of the total residue – as in the analysis considering the union of the location and year factors – can provide biased estimates that do not agree with the true genetic value. In other words, in some situations, it is necessary to compose the expectations of the mean squares in order to obtain an exact F test.

Another erroneous aspect in the analysis that considers the union of the location and year factors refers to the categorization of the experimental factors. The statistical model must correctly present the trial structure (Silva, 1999). The biometric models proposed do not usually distinguish the effects of two categories of experimental factors: i) *Treatment factors*, whose levels are randomly attributed to elementary units, under the control of the researcher and ii) *Intrinsic factors*, whose levels are determined by the units themselves (Silva, 1999). The nature of the trials for cultivar recommendation (VCU-RNC) demands

candidate genotypes to be evaluated in different locations and years. In the first analysis, these factors were considered solely as a variation source called *environment*. However, the environmental factor is considered an *intrinsic* factor, since it cannot be replicated in the experiment, and its levels are determined by the units themselves. For example, it is not possible to replicate the combination of the Canoinhas/2013 environment, since it includes permanent characteristics of the location Canoinhas and of the year 2013 concerning the general aspects of climate, soil type and rainfall distribution, which should ideally remain constant but vary unpredictably.

Other relevant information provided by the decomposition of the imbalanced factors refers to the comparison between the environments. The use of multiple comparison tests may not be adequate to investigate the superiority of one environment over others, in analyses using imbalanced data. Therefore, the variance of the LC x LNC interaction can be explored based on comparisons of interest and with the use of contrasts (Table 3).

**Table 3.** Univariate contrasts for the trait grain yield ( $\text{kg ha}^{-1}$ ), considering the effects of *Common locations (LC)* - Lages (LG), Ponte Serrada (PS), and Canoinhas (CA) and *Non-common locations (LNC)* - Chapecó (CH), Guatambu (GT), Urussanga (UR), and Campos Novos (CN). UDESC-IMEGEM, Lages, Santa Catarina, 2017.

Contrast	Degrees of freedom	Average square	Percentage
<i>Locations</i>	6	41762639 <sup>*</sup>	-
<i>Common (LC)</i>	(2)	47049643 <sup>*</sup>	-
C <sub>1</sub> : LG vs. PS + CA	1	77996928 <sup>*</sup>	83%
C <sub>2</sub> : PS vs. CA	1	16102359 <sup>*</sup>	17%
<i>Non-common (LNC)</i>	(3)	19483210 <sup>*</sup>	-
C <sub>1</sub> : CN vs. CH + GT + UR	1	4011764 <sup>*</sup>	7%
C <sub>2</sub> : CH vs. GT + UR	1	30850619 <sup>*</sup>	53%
C <sub>3</sub> : GT vs. UR	1	23587258 <sup>*</sup>	40%

<sup>\*</sup>Significant at 5% error probability by the F test.

The flexibility of the contrast technique allows for comparing the effects of the desired variation on the genotypes, considering the non-common locations, in compliance with the criteria adopted by the commissions that manage the launch of cultivars. Therefore, inferences can be set based on a probability of error, rather than solely comparing a superior mean performance of 5% in relation to the controls (Silva, 2014).

According to the contrast analysis – in relation to the common locations – it can be observed that LG refers to the greatest variation for the period and genotypes considered, with 83% of the variation of the common locations, to the detriment of the comparison between PS and CA, only 17%. Regarding the non-common locations, CH shows greater variation compared to the other three non-common locations, namely, 53% of the variation. A similar example can be attributed to GT, which, compared to UR, revealed 40% of the total variation of the non-common locations. These locations may reveal a low ability to distinguish genotypes. Thus, researchers should be cautious when such locations are the determinants for the selection of certain genetic constitutions (Table 3).

In general, the analysis of trials for cultivar recommendation is often composed of a series of imbalanced data. In such a condition, since orthogonality is missed, the calculations of the sums of squares become much more complex (Wechsler, 1998), and the phenotypic value may not be a faithful estimate of the true genetic value. Researchers constantly perform a superficial statistical analysis and suppress the environments that are not replicated over the performance of the Value for Cultivation and Use test (Pereira et al., 2010). In other occasions, researchers multiply the number of locations and agricultural years, in a simple and incorrect way, considering a single source of variation called *environment* in their statistical model. Therefore, part of the information is lost, and the experiments are not analyzed as previously planned.

The decomposition of the mean squares can be an advantageous alternative for analyzing the complexity of imbalanced data, by maximizing all possible interest inferences in the network of trials. The analysis and interpretation of the results, respecting the due breaks of the degrees of freedom, shows that the treatment factor genotypes presented no significant differences by the F test when tested by the composition of the appropriate mean squares. This highlights the relevance of the consideration of the appropriate sources of variation in the mathematical model, which directly affects the conclusions and recommendations of cultivars with superior performances, actually proven by cultivar recommendation tests.

## Conclusion

The consideration of the appropriate sources of variation in the statistical model in experiments of Value for Cultivation and Use test directly affects the conclusions related to plant breeding. The decomposition of the imbalanced factors is an advantageous alternative for the detailed understanding of the relevant causes of variation.

## Acknowledgements

The authors would like to acknowledge the Universidade do Estado de Santa Catarina (UDESC, Brazil) (University of the State of Santa Catarina), Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina (EPAGRI, Brazil) (Agricultural Research and Rural Extension Company of Santa Catarina), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brazil) (National Council for Scientific and Technological Development), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil) (Coordination for the Improvement of Higher Education Personnel), and the Fundação de Apoio à Pesquisa Científica e Tecnológica do Estado de Santa Catarina (FAPESC, Brazil) (Foundation for Support to Scientific and Technological Research of the State of Santa Catarina) for the scholarships and financial support for the development of this research.

## References

- Assefa, T., Wu, J., Beebe, S. E., Rao, I. M., Marcomin, D., & Claude, R. J. (2015). Improving adaptation to drought stress in small red common bean: phenotypic differences and predicted genotypic effects on grain yield, yield components and harvest index. *Euphytica*, *203*(3), 477-489. DOI: 10.1007/s10681-014-1242-x.
- Bertoldo, J. G., Coimbra, J. L. M., Nodari, B. O., Guidolin, A. F., Hemp, S., Barili, L. D., ... Rozzeto, D. S. (2009). Stratification of the state of Santa Catarina in macro-environments for bean cultivation. *Crop Breeding and Applied Biotechnology*, *9*(4), 335-343. DOI: 10.12702/1984-7033.v09n04a08.
- Borges, V., Soares, A. A., Resende, M. D. V., Reis, M. S., Cornélio, V. M. O., & Soares, P. C. (2009). Progresso genético do programa de melhoramento de arroz de terras altas de minas gerais utilizando modelos mistos. *Revista Brasileira de Biometria*, *27*(3), 478-490.
- Boukid, F., Prandi, B., Sforza, S., Sayar, R., Seo, Y. W., Mejri, M., & Yacoubi, I. (2017). Understanding the effects of genotype, growing year, and breeding on tunisian durum wheat allergenicity. 1. The Baker's Asthma case. *Journal of Agricultural and Food Chemistry*, *65*(28), 5831-5836. DOI: 10.1021/acs.jafc.7b02040.
- Carvalho, L. C. B., Damasceno-Silva, K. J., Rocha, M. M., & Oliveira, G. C. X. (2017). Genotype x environment interaction in cowpea by mixed models. *Revista Ciência Agronômica*, *48*(5), 872-878. DOI: 10.5935/1806-6690.20170103.
- Coimbra, J. L. M., Souza, V. Q., Kopp, M. M., Silva, J. G. C., Oliveira, A. C., & Carvalho, F. I. F. (2006). Esperanças matemáticas dos quadrados médios: uma análise essencial. *Ciência Rural*, *36*(6), 1730-1738. DOI: 10.1590/S0103-84782006000600010.
- Coimbra, J. L. M., Bertoldo, J. G., Elias, H. T., Hemp, S., Vale, N. M. d., Toaldo, D., ... Kopp, M. M. (2009). Mineração da interação genótipo x ambiente em *Phaseolus vulgaris* L. para o Estado de Santa Catarina. *Ciência Rural*, *39*(2), 355-363.
- Companhia Nacional de Abastecimento. [CONAB]. (2016). *Acompanhamento da safra brasileira de grãos. v.4, safra 2016/17, Segundo Levantamento*. Brasília, DF: CONAB.
- Dias, F. T. C., Pitombeira, J. B., Teófilo, E. M., & Barbosa, F. S. (2009). Adaptabilidade e estabilidade fenotípica para o caráter rendimento de grãos em cultivares de soja para o Estado do Ceará. *Revista Ciência Agronômica*, *40*(1), 129-134.
- Gupta, S. K., Rathore, A., Yadav, O. P., Rai, K. N., Khairwal, I. S., Rajpurohit, B. S., & Das, R. R. (2013). Identifying mega-environments and essential test-locations for pearl millet cultivar selection in India. *Crop Science*, *53*(6), 2444-2453. DOI: 10.2135/cropsci2013.01.0053
- Instituto Brasileiro de Geografia e estatística [IBGE]. (2015). Pesquisa Mensal de Previsão e Acompanhamento das Safras Agrícolas no Ano Civil. *Levantamento Sistemático da Produção Agrícola*, *29*(10), 1-79.

- Ignaczak, J. C., & Silva, J. G. C. (1978). Análise conjunta de grupo de experimentos com alguns locais e tratamentos não comuns. *Pesquisa Agropecuária Brasileira*, 13(3), 56-66.
- Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Shabenberger, O. (2006). *SAS® for mixed models*. Cary, NC: SAS Institute.
- Lobell, D. B., & Gourdji, S. M. (2012). The influence of climate change on global crop productivity. *Plant Physiology*, 160(4), 1686-1697. DOI: 10.1104/pp.112.208298
- Nassir, A. L., & Ariyo, O. J. (2011). Genotype x Environment Interaction and Yield-Stability Analyses of Rice Grown in Tropical Inland Swamp. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca*, 39(1), 220-225. DOI: 10.15835/nbha3915591
- Pereira, H. S., Melo, L. C., Faria, L. C., Peloso, M. J. D., Díaz, J. L. C., & Wendland, A. (2010). Indicação de cultivares de feijoeiro-comum baseada na avaliação conjunta de diferentes épocas de semeadura. *Pesquisa Agropecuária Brasileira*, 45(6), 571-578. DOI: 10.1590/S0100-204X2010000600006
- Pereira, T. C. V., Schmit, R., Haveroth, E. J., Melo, R. C., Coimbra, J. L. M., Guidolin, A. F., & Backes, R. L. (2016). Reflexo da interação genótipo x ambiente sobre o melhoramento genético de feijão. *Ciência Rural*, 46(3), 411-417. DOI: 10.1590/0103-8478cr20130998
- Piepho, H. P., Herndl, M., Pötsch, E. M., & Bahn, M. (2017). Designing an experiment with quantitative treatment factors to study the effects of climate change. *Journal of Agronomy and Crop Science*, 203(6), 584-592. DOI: 10.1111/jac.12225
- Schmidt, E. R., Nascimento, A. L., Cruz, C. D., & Oliveira, J. A. R. (2011). Avaliação de metodologias de adaptabilidade e estabilidade de cultivares milho. *Acta Scientiarum. Agronomy*, 33(1), 51-58. DOI: 10.4025/actasciagron.v33i1.5817
- Silva, J. G. C. (1999). A consideração da estrutura das unidades em inferências derivadas do experimento. *Pesquisa Agropecuária Brasileira*, 34(6), 911-925. DOI: 10.1590/S0100-204X1999000600001
- Silva, J. G. C. (2014). Análise crítica do processo de lançamento de cultivares. *Revista da Estatística UFOP*, 3(2), 16-21.
- Silva, T. R. C., Amaral Júnior, A. T., Gonçalves, L. S. A., Candido, L. S., Vittorazzi, C., & Scapim, C. A. (2013). Agronomic performance of popcorn genotypes in Northern and Northwestern Rio de Janeiro State. *Acta Scientiarum. Agronomy*, 35(1), 57-63. DOI: 10.4025/actasciagron.v35i1.15694
- Vencovsky, R., Ramalho, M. A. P., & Toledo, F. H. R. B. (2012). Contribution and perspectives of quantitative genetics to plant breeding in Brazil. *Crop Breeding and Applied Biotechnology*, 12(Spec), 7-14. DOI: 10.1590/S1984-70332012000500002
- Wechsler, F. S. (1998). Fatoriais fixos desbalanceados: uma análise mal compreendida. *Pesquisa Agropecuária Brasileira*, 33(3), 231-262.
- Yan, W. (2015). Mega-environment analysis and test location evaluation based on unbalanced multiyear data. *Crop Science*, 55(1), 113-122. DOI: 10.2135/cropsci2014.03.0203.