

Multicollinearity in canonical correlation analysis in maize

B.M. Alves, A. Cargnelutti Filho and C. Burin

Departamento de Fitotecnia, Universidade Federal de Santa Maria,
Santa Maria, RS, Brasil

Corresponding author: A. Cargnelutti Filho
E-mail: alberto.cargnelutti.filho@gmail.com

Genet. Mol. Res. 16 (1): gmr16019546

Received November 21, 2016

Accepted February 16, 2017

Published March 30, 2017

DOI <http://dx.doi.org/10.4238/gmr16019546>

Copyright © 2017 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

ABSTRACT. The objective of this study was to evaluate the effects of multicollinearity under two methods of canonical correlation analysis (with and without elimination of variables) in maize (*Zea mays* L.) crop. Seventy-six maize genotypes were evaluated in three experiments, conducted in a randomized block design with three replications, during the 2009/2010 crop season. Eleven agronomic variables (number of days from sowing until female flowering, number of days from sowing until male flowering, plant height, ear insertion height, ear placement, number of plants, number of ears, ear index, ear weight, grain yield, and one thousand grain weight), 12 protein-nutritional variables (crude protein, lysine, methionine, cysteine, threonine, tryptophan, valine, isoleucine, leucine, phenylalanine, histidine, and arginine), and 6 energetic-nutritional variables (apparent metabolizable energy, apparent metabolizable energy corrected for nitrogen, ether extract, crude fiber, starch, and amylose) were measured. A phenotypic correlation matrix was first generated among the 29 variables for each of the experiments. A multicollinearity diagnosis was later performed within each group of variables using methodologies such as variance inflation factor and condition number. Canonical correlation analysis

was then performed, with and without the elimination of variables, among groups of agronomic and protein-nutritional, and agronomic and energetic-nutritional variables. The canonical correlation analysis in the presence of multicollinearity (without elimination of variables) overestimates the variability of canonical coefficients. The elimination of variables is an efficient method to circumvent multicollinearity in canonical correlation analysis.

Key words: *Zea mays* L.; Linear relations among variables; Variance inflation factor; Condition number; Variable elimination

INTRODUCTION

According to the Food and Agriculture Organization of the United Nations, maize (*Zea mays* L.) is one of the world's most widely produced grains (FAO, 2016). Brazil ranks first with 66.98 million tons of grains produced in the 2015/2016 harvest season. In Brazil, maize cultivation covers an area of 15.92 million hectares, with an average yield of 4207 kg/ha (CONAB - Companhia Nacional de Abastecimento, 2016). It is widely used as protein and energy source in animal feed, with a large part of the production destined for animal and human consumption. Thus, the productive and nutritional characterization of the grains is of considerable importance.

Modern and competitive agriculture drives maize-breeding programs to focus on improving the nutritional composition of the grains, in addition to increasing its yield. However, according to a study carried out by Duvick (2005), the selection of more productive plants presents an inverse relationship with the contents of some nutritional components of interest.

The study of linear relationships between characters is important for the genetic improvement of plants through indirect selection. This study is performed when the characters of interest to be improved have low heritability or are difficult to measure; when early selection of plants or their lineages is desired; and when the simultaneous selection of more than one character was desired (Cruz and Regazzi, 1997). Studies using multivariate analyses are important in the genetic improvement of plants because they enable an early selection of the researcher's characters of interest, thus reducing fieldwork, time, and financial resources. Therefore, it is necessary to check how the characters of interest are related.

The complex linear associations among groups of agronomic and nutritional variables in maize make it difficult to classify the individual importance of each variable within the group or among groups of variables. The information is multidimensional, owing to antagonism and interactions among these variables and other abiotic factors, making it difficult to select superior genotypes. Therefore, canonical correlation analysis is a multivariate statistical procedure that enables verification of the linear relationship between two groups or sets of variables (X and Y) (Hair et al., 2009).

Multivariate analysis presents assumptions to be verified, including multicollinearity. Multicollinearity observed in the correlation matrix is variable. If the variables are highly correlated (with high degree of multicollinearity), then the inferences may be erroneous or unreliable. This strong association among variables is a problem that can affect parameter estimation (Gujarati and Porter, 2011).

The canonical correlation analysis presents multicollinearity as an assumption

required for the results to be reliable. The multicollinearity diagnosis can be performed by methodologies such as variance inflation factor (VIF) and condition number (CN) (Gujarati and Porter, 2011).

Toebe and Cargnelutti Filho (2013a,b) studied the effects of multicollinearity on path analysis in maize crop. They concluded that multicollinearity has a greater effect on the estimation of direct effects in path analyses than multivariate non-normality. They also confirmed that the traditional path analysis, with a high degree of multicollinearity in the correlation matrix, generates path coefficient estimates without biological significance. Path analysis under multicollinearity has already been studied in pepper (Carvalho et al., 1999), maize (Carvalho et al., 2001), and soybean crops (Bizeti et al., 2004), and the consequences of multicollinearity have already been verified in canola crop (Coimbra et al., 2005).

The elimination of variables to circumvent the multicollinearity problem in canonical correlation analysis has already been used in castor beans (Brum et al., 2011) and maize crops (Alves et al., 2016a,b).

Alves et al. (2016a,b) studied the linear dependence among groups of phenological, morphological, productive, protein-nutritional, and energetic-nutritional characteristics in maize crop. For the maize crop, it was found that there were effects of multicollinearity in the canonical correlation analysis. Thus, the objective of this study was to evaluate the effects of multicollinearity under two forms of canonical correlation analysis (with and without elimination of variables) in maize crop.

MATERIAL AND METHODS

The data used in this study were obtained from three experiments conducted with maize crop in the experimental area of the Department of Plant Sciences, Federal University of Santa Maria, Rio Grande do Sul State (latitude 29°42'S, longitude 53°49'W, and 95 m in altitude), during the 2009/2010 crop season. In the first experiment, 36 early maturing maize genotypes were evaluated; in the second experiment, 22 super-early maturing maize genotypes were evaluated; and in the third experiment, 18 transgenic maize genotypes were evaluated. The genotypes belonged to the Network of Evaluation of Maize Cultivars of Rio Grande do Sul, coordinated by the State Agricultural Research Foundation (Fundação Estadual de Pesquisa Agropecuária - FEPAGRO).

The experiments were conducted in randomized blocks with three replicates. The plots were composed of two rows of 5 m length, with 0.80 m spacing between rows and 0.20 m spacing between plants in each row. On October 26, 2009, maize seeds were sown manually, with a base fertilization application consisting of 37.5 kg/ha nitrogen (N), 150 kg/ha phosphorus (P_2O_5), and 150 kg/ha potassium (K_2O). Between November 1 and 3, 2009, the seedlings emerged. The population was adjusted by thinning the seedlings to achieve 62,500 plants/ha. Approximately 200 kg/ha nitrogen was applied to the cover crops in three applications, when the plants had three, five, and ten leaves. On March 15, 2010, the maize ears were harvested manually.

In each replicate of each genotype, the agronomic variables such as the number of days from sowing until male flowering (MF), the number of days from sowing until female flowering (FF) (considered when 50% of plants in the plots had male and female flowering), plant height in the harvest (PH, cm), ear insertion height at harvest (EH, cm), relative ear placement ($EP = EH/PH$) (measured in all plants of the plot), number of plants (NP, plants/

ha), number of ears (NE, ears/ha), ear index (EI = NE/NP), ear weight (EW, t/ha), grain yield at 13% humidity (GY, t/ha), and one thousand grain weight (TGW, g) were measured. Next, a sample of 500 g maize kernels was taken from each plot. It was then packed in a paper bag and dried in a forced air circulation oven until it reached 10% humidity. After drying, the grains were ground in micro-mill (MA-630, Marconi) to obtain a sample with a grain size between 0.30 and 0.50 mm. Each ground sample was packed in a hermetically sealed package to determine the nutritional analyses of the grains.

With the samples from each replicate of each genotype, the protein-nutritional variables, such as crude protein (CP), lysine (Lys), methionine (Met), cysteine (Cys), threonine (Thr), tryptophan (Trp), valine (Val), isoleucine (Ile), leucine (Leu), phenylalanine (Phe), histidine (His), and arginine (Arg); and the energetic-nutritional variables, such as apparent metabolizable energy (AME, kcal/kg), apparent metabolizable energy corrected for nitrogen (AMEn, kcal/kg), ether extract (EE), crude fiber (CF), starch (ST), and amylose (AML), were determined in raw matter percentage (%RM). The evaluations were performed by means of near-infrared reflectance spectroscopy (NIRS - Near Infrared Reflectance Spectroscopy), using calibration adjustment by the CEAN 010 analytical procedure of Adisseo Brasil S.A. Because the NIRS equipment was not calibrated to determine amylose content in maize grains, this variable was determined according to the iodometric technique described by Martinez and Cuevas-Perez (1989). This technique makes use of dissolution, gelatinization, acidification, and addition of iodine solution to form a complex with the starch, which can then be read using a spectrophotometer at 620 nm.

In each experiment, analysis of variance was performed, and the assumptions of the mathematical model were tested for each variable. The normality of the errors was verified by the Kolmogorov-Smirnov test, and the homogeneity of residual variances was verified using the Bartlett test. The phenotypic correlation matrix was first estimated among the 29 variables. A multicollinearity diagnosis was later performed on the correlation matrix of each group of variables (agronomic, protein-nutritional, and energetic-nutritional variables).

The methods used to verify the degree of multicollinearity of the phenotypic correlation matrix were the VIF that provides the VIF for each variable, and the CN that provides CN for the group of variables. The VIF is an indicator of how much the variance of the coefficient is inflated compared to the non-significant correlations with any other variables in the model. VIF values below 10 are considered adequate, indicating the absence of multicollinearity, whereas VIF values above 10 indicate a high degree of multicollinearity among variables (Hair et al., 2009). The CN is the ratio of the largest to the smallest eigenvalue of the correlation matrix. According to the criteria established by Montgomery and Peck (1982), there are three degrees of multicollinearity namely severe, when the number of conditions is greater than 1000; moderate to strong, when the number of conditions is between 100 and 1000; and weak, when the number of conditions is less than 100. VIF and CN were used as decision criteria for the elimination of highly correlated variables.

In each experiment, the canonical correlation analysis was performed, with and without eliminating variables, according to the procedures described by Cruz and Regazzi (1997). In the method without elimination of variables, the groups were correlated as agronomic (MF, FF, PH, EH, EP, NP, NE, EI, EW, GY, and TGW) and protein-nutritional (CP, Lys, Met, Cys, Thr, Trp, Val, Ile, Leu, Phe, His, and Arg), and agronomic (MF, FF, PH, EH, EP, NP, NE, EI, EW, GY, and TGW) and energetic-nutritional variables (AME, AMEn, EE, CF, ST, and AML). In the method with elimination of variables, the groups were correlated as agronomic (MF,

FF, PH, EH, EP, NP, NE, EI, EW, GY, and TGW) and protein-nutritional (CP, Lys, Met, Cys, Thr, Trp, Val, Ile, Leu, Phe, His, and Arg), and agronomic (MF, FF, PH, EH, EP, NP, NE, EI, EW, GY, and TGW) and energetic-nutritional variables (AME, AMEn, EE, CF, ST, and AML).

The associations among groups of variables under the two methods of analysis (with and without elimination of variables) were presented through canonical pairs, accompanied by respective canonical coefficients. To evaluate the significance of the canonical correlation between groups of variables, the chi-square test was used at 5% significance (Cruz and Regazzi, 1997). Statistical analyses were performed using the GENES software (Cruz, 2013) and the Microsoft Office Excel® application.

RESULTS AND DISCUSSION

Based on analysis of variance, it was found that the genotypes have genetic variability, except starch (ST) for early maturing genotypes and for the transgenic class genotypes, the number of plants (NP), number of ears (NE), ears index (EI), methionine (Met), and starch (ST). The assumption of normality of errors was observed in 25, 28, and 28 of the 29 variables in early maturing, super-early maturing, and transgenic genotypes, respectively. The homogeneity of residual variances was observed in 19, 25, and 26 of the 29 variables in early maturing, super-early maturing, and transgenic genotypes, respectively. The high percentage of compliance with assumptions lends credibility to analysis of variance.

It is important to verify the degree of multicollinearity in each group of variables to obtain a safe estimate and appropriate biological interpretation of the evaluation of the linear association between two groups of variables (Cruz and Regazzi, 1997). In the multicollinearity diagnosis based on the VIF in early maturing, super-early maturing, and transgenic genotypes, the agronomic, protein-nutritional, and energetic-nutritional variables showed a high degree of multicollinearity without elimination of variables ($VIF > 10$) (Tables 1, 2, and 3).

On the other hand, with the elimination of variables, there was a significant reduction in the degree of multicollinearity ($VIF < 10$) of the variables within each group (agronomic, protein-nutritional, and energetic-nutritional) in all three experiments, except for ear height (EH), which had $VIF = 11.5$ in transgenic genotypes (Tables 4, 5, and 6).

In early maturing genotypes, the CN values, without elimination of variables, for the groups of agronomic, protein-nutritional, and energetic-nutritional variables were 196,42, 805, and 1019, respectively. The elimination of highly correlated variables within each group reduced this value to 36 for agronomic variables, 32 for protein-nutritional variables, and 6 for energetic-nutritional variables. In super-early maturing genotypes, the CN, without elimination of variables, was 146,612 for the group of agronomic variables, 2138 for protein-nutritional variables, and 1599 for energetic-nutritional variables. With the elimination of variables, the CN was reduced to 48 for agronomic variables, 49 for protein-nutritional variables, and 35 for energetic-nutritional variables. In transgenic genotypes, without elimination of variables, the CN was 16,566 for agronomic variables, 3900 for protein-nutritional variables, and 515 for energetic-nutritional variables. The elimination of variables reduced the CN to 60, 33, and 14 for agronomic, protein-nutritional, and energetic-nutritional variables, respectively. The elimination of highly correlated variables in all the groups resulted in reduction of CN values ($CN \leq 100$), indicating weak multicollinearity among variables.

The approach of elimination of variables to circumvent multicollinearity problems in multivariate analyses, such as canonical correlation analysis and path analysis, has already

Table 1. Canonical coefficients of canonical pairs between the agronomic and protein-nutritional, and agronomic and energetic-nutritional variables of 36 early maturing maize genotypes (the canonical correlation analysis was performed without the elimination of variables).

Variables ⁽¹⁾	Canonical pairs											VIF ⁽²⁾
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	
	Agronomic											
MF	0.19	-0.46	0.21	-0.45	0.49	0.46	2.35	0.13	-0.53	0.28	-1.32	8.6
FF	-0.95	0.56	-0.52	0.65	-0.04	0.74	-2.27	0.68	1.83	-0.42	2.58	18.3
PH	-2.37	-0.14	-6.92	1.14	6.09	6.24	-4.93	-6.43	1.83	3.99	0.31	216.0
EH	4.96	0.23	17.48	-1.22	-13.61	-12.80	11.49	14.13	-4.02	-9.63	-0.92	1122.1
EP	-3.80	-0.37	-14.28	0.69	11.73	9.94	-9.75	-11.88	3.13	7.51	1.32	759.5
NP	-1.73	2.76	3.66	2.15	2.21	5.14	-1.56	4.28	0.01	7.63	2.68	146.0
NE	4.26	-9.80	-13.61	-7.57	-8.08	-17.46	7.22	-15.70	-2.42	-27.38	-7.20	1832.8
EI	-3.27	8.01	12.03	6.84	7.02	17.10	-6.30	14.97	2.08	22.64	7.45	1444.0
EW	-2.18	-4.86	1.05	-0.78	1.06	-8.46	4.10	1.95	-3.63	-2.68	-7.95	207.0
GY	1.35	6.24	-1.19	0.63	-0.78	7.12	-4.55	-2.07	4.90	2.73	7.48	206.3
TGW	0.83	-0.39	-0.51	0.59	-0.06	-0.02	0.43	0.21	-0.58	0.07	0.10	2.0
	Protein-nutritional											
CP	0.18	0.19	0.18	0.02	0.18	-0.26	0.05	1.36	0.15	0.37	-0.17	2.2
Lys	-0.85	-0.12	0.31	1.03	-0.79	-1.08	-1.01	0.34	0.11	0.16	0.47	7.0
Met	0.27	0.05	-0.90	-0.23	0.67	-0.83	0.94	0.30	-0.16	0.27	-1.34	5.2
Cys	0.66	-0.24	-0.45	-0.21	2.17	1.26	-0.36	-0.37	0.94	-0.99	-0.71	9.9
Thr	-0.02	-0.20	-0.40	-2.00	1.00	0.56	-0.15	-0.69	2.70	0.14	-2.24	18.3
Trp	-0.57	-0.31	-0.17	-1.14	0.17	0.28	1.37	0.77	-1.43	-0.03	1.31	8.1
Val	-2.99	1.19	-0.75	-0.09	2.68	-2.03	-0.49	-1.25	1.32	-0.64	-0.14	33.0
Ile	1.63	-1.17	-1.00	0.84	-2.36	-0.86	-1.40	1.82	-1.06	-1.56	0.92	22.5
Leu	-0.36	-0.69	-0.78	-1.50	0.71	-0.48	-1.36	-0.62	-1.26	-0.74	-2.43	15.2
Phe	0.77	0.63	2.35	3.22	-2.08	3.47	1.50	0.65	0.19	3.76	3.07	60.5
His	-0.58	-0.15	0.82	0.93	-2.39	-0.74	1.91	0.45	0.73	-0.35	-1.57	15.8
Arg	1.98	-0.14	0.81	-0.91	0.53	0.10	-0.83	-1.86	-2.13	-0.13	2.13	19.9
r	0.92 ^{ns}	0.86 ^{ns}	0.78 ^{ns}	0.69 ^{ns}	0.62 ^{ns}	0.56 ^{ns}	0.42 ^{ns}	0.38 ^{ns}	0.30 ^{ns}	0.13 ^{ns}	0.08 ^{ns}	
	Agronomic											
MF	0.32	-1.27	0.11	-0.91	0.83	0.74	-	-	-	-	-	8.6
FF	-1.34	1.79	-0.97	0.84	-0.73	-1.01	-	-	-	-	-	18.3
PH	-4.66	5.36	-4.02	0.47	-1.75	-8.32	-	-	-	-	-	216.0
EH	10.17	-13.48	7.62	-2.12	4.87	18.99	-	-	-	-	-	1122.1
EP	-7.87	11.03	-5.88	2.58	-4.00	-15.65	-	-	-	-	-	759.5
NP	-2.02	4.96	-6.58	-0.56	7.82	-1.58	-	-	-	-	-	146.0
NE	7.10	-17.05	23.66	0.92	-26.70	7.29	-	-	-	-	-	1832.8
EI	-5.75	14.30	-20.92	-1.74	23.75	-7.23	-	-	-	-	-	1444.0
EW	-1.37	-9.29	6.37	3.48	0.66	3.61	-	-	-	-	-	207.0
GY	-0.03	10.11	-5.81	-2.61	-1.10	-2.96	-	-	-	-	-	206.3
TGW	0.58	-0.27	-0.02	-0.03	-0.15	0.01	-	-	-	-	-	2.0
	Energetic-nutritional											
AME	8.43	-6.36	-1.22	-3.74	-3.12	-8.48	-	-	-	-	-	208.7
AMEn	-7.18	5.95	1.36	3.88	2.82	8.87	-	-	-	-	-	190.6
EE	-1.25	0.19	-0.34	0.96	0.81	-0.01	-	-	-	-	-	3.3
CF	-0.19	-0.40	0.30	-0.39	0.78	0.52	-	-	-	-	-	1.3
ST	0.33	0.43	-0.96	0.00	0.25	-0.49	-	-	-	-	-	1.5
AML	0.20	0.26	0.63	0.44	0.32	-0.64	-	-	-	-	-	1.2
r	0.76 ^{ns}	0.71 ^{ns}	0.63 ^{ns}	0.51 ^{ns}	0.36 ^{ns}	0.33 ^{ns}						

⁽¹⁾Agronomic: MF = number of days from sowing until male flowering; FF = number of days from sowing until female flowering; PH = plant height, in cm; EH = ear insertion height, in cm, and EP = ear placement (EP = EH/PH); NP = number of plants, in plants/ha; NE = number of ears, in ear/ha; EI = ear index; EW = ear weight, in t/ha; GY = grain yield, in t/ha and TGW = one thousand grains weight, in g; protein-nutritional: CP = crude protein, in %RM; Lys = lysine, in %RM; Met = methionine, in %RM; Cys = cysteine, in %RM; Thr = threonine, in %RM; Trp = tryptophan, in %RM; Val = valine, in %RM; Ile = isoleucine, in %RM; Leu = leucine, in %RM; Phe = phenylalanine, in %RM; His = histidine, in %RM, and Arg = arginine, in %RM; energetic-nutritional: AME = apparent metabolizable energy, in kcal/kg, AMEn = apparent metabolizable energy corrected for nitrogen, in kcal/kg, EE = ether extract, in %RM, CF = crude fiber, in %RM, ST = starch, in %RM, and AML = amylose, in %RM. r = canonical correlation. ^{ns}The chi-square test at 5% error probability was non-significant. ⁽²⁾VIF = variance inflation factor.

Table 2. Canonical coefficients of canonical pairs between the agronomic and protein-nutritional, and agronomic and energetic-nutritional variables of 22 super-early maturing maize genotypes (the canonical correlation analysis was performed without the elimination of variables).

Variables ⁽¹⁾	Canonical pairs											VIP ⁽²⁾
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	
	Agronomic											
MF	-0.45	-0.38	0.35	-1.95	0.57	1.18	-0.73	0.74	2.01	0.01	-1.59	13.7
FF	1.51	2.80	-0.77	2.06	-0.34	-1.93	0.07	0.45	-1.34	-0.59	1.31	22.9
PH	-1.96	-5.48	-0.64	4.11	2.65	-3.30	-0.83	5.52	12.15	-3.33	4.43	278.7
EH	3.09	9.50	0.87	-6.99	-5.87	6.53	0.93	-12.50	-26.36	8.25	-9.88	1243.9
EP	-2.23	-7.58	0.05	6.22	4.28	-4.30	-0.59	9.26	19.19	-6.56	8.08	700.5
NP	3.86	-7.18	-1.09	9.11	-2.49	0.19	-7.05	-9.70	-1.19	18.70	7.40	706.4
NE	-16.45	30.48	5.17	-36.66	8.75	-0.50	32.47	41.02	2.80	-74.93	-32.06	12034.3
EI	17.92	-28.65	-5.59	36.33	-8.39	0.59	-31.11	-38.51	-3.91	72.21	30.28	11161.7
EW	-0.70	-0.83	0.51	-3.29	5.99	2.72	0.82	2.15	-6.24	2.03	-0.58	104.2
GY	0.18	1.56	0.67	3.41	-5.74	-3.85	-1.50	-2.83	6.51	-2.38	1.45	122.8
TGW	0.25	0.57	-0.46	0.13	0.29	0.34	0.37	-0.13	0.93	-0.35	-0.54	2.3
	Protein-nutritional											
CP	0.47	-0.60	-0.11	0.54	0.17	0.03	-0.54	0.44	-0.43	-0.69	-0.85	3.1
Lys	0.61	1.51	0.15	-1.39	0.55	-0.23	1.08	-1.98	0.33	1.31	-1.04	13.1
Met	0.08	0.77	0.88	1.22	-0.04	0.87	-1.23	0.76	-0.56	1.05	-1.19	10.9
Cys	0.00	1.90	0.07	-0.45	-0.76	0.89	-0.20	0.43	-2.58	-0.08	0.69	19.6
Thr	-2.49	-2.13	0.42	-0.89	0.25	-0.02	-1.84	3.19	-0.66	1.07	-1.18	28.6
Trp	1.17	-0.38	0.68	0.36	-1.00	1.43	-2.29	0.04	-0.37	0.94	0.32	11.7
Val	-3.72	-5.38	-0.81	3.93	-1.84	1.65	-0.77	-1.06	-3.69	-3.45	-1.04	93.3
Ile	3.12	2.98	-3.10	-0.92	0.85	-3.48	0.03	-1.53	1.53	1.62	2.69	65.0
Leu	1.30	-1.41	-2.68	-1.58	2.89	1.04	-3.45	1.24	-0.16	0.74	0.86	39.5
Phe	-1.85	2.84	5.53	-0.11	-4.83	-0.85	7.68	-0.17	4.10	-1.31	-3.57	157.9
His	-1.71	1.42	1.36	-0.50	0.55	-0.96	-2.17	-1.87	2.26	0.97	-1.57	26.9
Arg	3.11	-1.12	-2.93	-0.06	3.35	0.43	2.97	0.95	0.06	-2.03	5.19	74.6
r	1.00*	1.00*	0.99 ^{ns}	0.93 ^{ns}	0.90 ^{ns}	0.83 ^{ns}	0.79 ^{ns}	0.74 ^{ns}	0.48 ^{ns}	0.16 ^{ns}	0.11 ^{ns}	
	Agronomic											
MF	-1.02	0.01	1.97	-0.68	-0.46	0.60	-	-	-	-	-	13.7
FF	0.64	-0.33	-0.92	0.42	0.59	-0.24	-	-	-	-	-	22.9
PH	0.68	9.11	3.94	9.75	-2.08	4.03	-	-	-	-	-	278.7
EH	-0.54	-19.31	-10.54	-20.20	5.58	-7.97	-	-	-	-	-	1243.9
EP	0.95	14.66	7.87	15.73	-3.77	5.65	-	-	-	-	-	700.5
NP	7.92	-3.60	-17.84	3.82	0.72	14.80	-	-	-	-	-	706.4
NE	-31.63	14.14	72.69	-18.69	-5.15	-61.74	-	-	-	-	-	12034.3
EI	29.64	-13.83	-70.08	17.31	5.90	60.27	-	-	-	-	-	11161.7
EW	1.32	-0.46	1.31	-1.87	2.01	-0.95	-	-	-	-	-	104.2
GY	-0.51	0.58	-1.86	2.75	-2.20	-0.12	-	-	-	-	-	122.8
TGW	-0.15	0.61	0.71	-0.77	0.08	0.41	-	-	-	-	-	2.3
	Energetic-nutritional											
AME	-0.75	-6.23	-1.73	3.58	3.06	14.43	-	-	-	-	-	272.9
AMEn	1.47	6.75	-0.08	-3.71	-1.41	-11.11	-	-	-	-	-	186.8
EE	0.26	-0.63	1.80	-0.57	-1.26	-3.15	-	-	-	-	-	15.5
CF	0.00	-0.69	-0.38	0.06	0.73	-0.34	-	-	-	-	-	1.3
ST	-0.64	0.23	0.62	-0.42	0.38	-0.46	-	-	-	-	-	1.4
AML	-0.23	-0.15	0.68	1.00	0.32	-0.68	-	-	-	-	-	2.1
r	0.97*	0.94 ^{ns}	0.80 ^{ns}	0.77 ^{ns}	0.73 ^{ns}	0.46 ^{ns}						

For abbreviations, see Table 1. *Significant by the chi-square test at 5% error probability. ^{ns}Non-significant.

been considered previously in canonical correlation analysis for castor beans (Brum et al., 2011) and maize (Alves et al., 2016a,b) crops, and in path analysis for chili (Carvalho et al., 1999), maize (Carvalho et al., 2001; Toebe and Cargnelutti Filho, 2013a,b), and soybean (Bizeti et al., 2004; Nogueira et al., 2012) crops.

In the canonical correlation analysis for early maturing genotypes, without elimination of variables, canonical correlations were not significant for all canonical pairs (Table 1). However, with the elimination of variables, the first canonical correlation between agronomic and protein-nutritional variables ($r = 0.79$) was significant. Among the agronomic and energetic-nutritional variables, there was no significant canonical correlation (Table 4). Therefore, a difference was observed in the results with and without elimination of variables.

Table 3. Canonical coefficients of canonical pairs between the agronomic and protein-nutritional, and agronomic and energetic-nutritional variables of 18 transgenic maize genotypes (the canonical correlation analysis was performed without the elimination of variables).

Variables ⁽¹⁾	Canonical pairs											VIF ⁽²⁾
	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	11th	
	Agronomic											
MF	-0.48	1.80	1.35	0.51	1.85	0.64	0.99	2.74	0.52	-1.08	0.61	19.7
FF	1.31	-0.64	-1.00	-1.60	-0.80	-0.34	1.01	-2.00	0.79	-0.47	1.07	13.5
PH	-2.27	-1.95	-0.38	2.32	0.80	11.50	-0.17	-5.40	3.59	-2.94	4.57	219.0
EH	3.72	4.11	-0.52	-4.21	-1.56	-26.65	2.59	15.99	-8.70	9.18	-10.94	1303.7
EP	-2.43	-3.94	0.31	3.00	0.72	16.30	-3.83	-11.90	5.19	-6.09	5.46	546.7
NP	5.72	0.45	0.25	0.14	0.93	4.50	10.30	-0.63	-8.98	7.08	-1.88	295.0
NE	-13.27	-1.13	-0.35	1.16	-3.22	-10.64	-23.19	0.98	21.64	-17.93	4.87	1654.5
EI	12.62	0.74	0.65	-1.06	3.25	9.66	21.31	-0.44	-18.53	16.20	-4.19	1342.6
EW	1.43	-0.62	0.16	1.77	0.47	0.28	-1.67	2.96	-1.92	0.16	1.52	23.4
GY	-1.45	0.46	0.57	-2.42	0.09	-0.01	2.55	-2.10	2.27	-1.11	-0.24	25.8
TGW	-0.33	0.47	-0.38	-0.18	-0.22	-1.27	-0.26	-1.26	-0.03	0.66	-0.14	4.3
	Protein-nutritional											
CP	-0.04	0.38	0.27	1.08	-0.32	0.28	-0.11	-0.29	-0.44	-0.89	-0.07	2.9
Lys	-0.45	1.21	-0.18	-0.37	-0.21	-0.37	-1.89	-2.17	-2.10	0.20	1.43	17.2
Met	-1.04	0.11	-1.14	0.83	-1.15	-0.59	1.39	-0.06	1.50	-0.47	-0.22	10.3
Cys	2.42	-0.23	-0.53	-0.48	-1.70	-1.13	2.31	1.60	1.43	0.07	-0.52	24.4
Thr	-3.82	0.90	-3.60	0.26	-4.21	0.89	4.81	5.50	0.84	-1.40	-2.69	124.8
Trp	2.14	-0.58	0.04	-0.39	1.11	0.24	0.37	-0.93	0.27	0.45	-0.09	11.8
Val	3.28	-2.51	-2.34	-2.65	-1.22	2.91	1.82	1.76	2.95	-0.49	-1.97	92.1
Ile	-0.46	-0.38	2.03	1.93	0.80	-1.37	-1.95	-4.04	-2.11	1.96	1.19	42.3
Leu	-1.25	-0.43	-0.63	-0.89	-2.01	2.50	3.53	0.77	1.42	-0.47	-1.59	33.7
Phe	1.54	2.35	2.69	1.11	5.09	-2.27	-6.78	-0.80	-3.41	0.60	5.20	182.9
His	-6.84	1.88	-4.08	1.26	-0.91	-0.32	2.13	-0.30	-2.13	-1.55	1.33	82.8
Arg	3.64	-1.53	6.19	-0.92	3.85	-0.81	-3.84	-1.15	2.21	1.87	-2.22	104.9
r	1.00*	1.00*	1.00*	1.00*	1.00*	1.00 ^{ns}	0.82 ^{ns}	0.72 ^{ns}	0.67 ^{ns}	0.47 ^{ns}	0.11 ^{ns}	
	Agronomic											
MF	1.81	2.95	0.03	1.18	-1.13	1.21	-	-	-	-	-	19.7
FF	-0.11	-1.20	-0.51	-0.88	0.24	0.81	-	-	-	-	-	13.5
PH	-4.01	2.90	-1.08	-1.39	-7.74	0.63	-	-	-	-	-	219.0
EH	7.61	-4.04	5.31	-2.95	18.38	-2.15	-	-	-	-	-	1303.7
EP	-5.18	-0.03	-3.45	1.33	-11.40	0.83	-	-	-	-	-	546.7
NP	1.98	3.34	0.59	-6.18	-1.29	-11.24	-	-	-	-	-	295.0
NE	-4.78	-7.58	-2.49	14.71	3.02	27.39	-	-	-	-	-	1654.5
EI	4.38	7.28	2.04	-13.35	-2.75	-24.10	-	-	-	-	-	1342.6
EW	-0.82	2.63	-1.38	-2.01	-0.55	-0.50	-	-	-	-	-	23.4
GY	1.38	-1.57	0.94	2.52	-0.72	1.36	-	-	-	-	-	25.8
TGW	0.31	-0.85	0.35	0.33	1.48	-0.44	-	-	-	-	-	4.3
	Energetic-nutritional											
AME	-0.67	5.63	1.92	-2.27	5.58	-4.00	-	-	-	-	-	88.2
AMEn	-0.10	-5.12	-1.07	3.59	-4.53	3.23	-	-	-	-	-	71.2
EE	0.67	-0.45	0.29	-1.14	-0.38	1.34	-	-	-	-	-	4.0
CF	-0.54	0.89	0.89	-0.06	-0.26	0.12	-	-	-	-	-	2.0
ST	1.03	0.05	-0.73	0.36	-0.62	-0.24	-	-	-	-	-	2.2
AML	0.25	-0.97	0.11	-0.33	-0.20	-0.58	-	-	-	-	-	1.5
r	1.00*	0.98 ^{ns}	0.94 ^{ns}	0.81 ^{ns}	0.63 ^{ns}	0.51 ^{ns}						

For abbreviations, see Table 1. *Significant by the chi-square test at 5% error probability. ^{ns}Non-significant.

Without the elimination of variables, all groups were independent; however, with the elimination of variables, the groups of agronomic and protein-nutritional variables presented linear dependence.

In super-early maturing genotypes, without elimination of variables, the canonical correlations were significant ($r = 1.00$) for the first two canonical pairs between agronomic and protein-nutritional variables. Among agronomic and energetic-nutritional variables, the first canonical pair showed a significant canonical correlation ($r = 0.97$) (Table 2). With the elimination of variables, the canonical correlation of the first canonical pair was significant ($r = 0.93$ and $r = 0.92$) between agronomic and protein-nutritional variables, and between

Table 4. Canonical coefficients of canonical pairs between the agronomic and protein-nutritional, and agronomic and energetic-nutritional variables of 36 early maturing maize genotypes (the canonical correlation analysis was performed with the elimination of variables).

Variables ⁽¹⁾	Canonical pairs						VIF ⁽²⁾
	1st	2nd	3rd	4th	5th	6th	
	Agronomic						
MF	-	-	-	-	-	-	-
FF	0.76	0.04	0.36	0.26	-0.78	1.56	3.9
PH	0.38	0.18	-0.56	-1.03	0.03	0.21	1.7
EH	-0.10	0.45	-0.39	0.84	-0.18	-0.06	1.6
EP	-	-	-	-	-	-	-
NP	0.11	-0.53	0.16	-0.19	0.59	0.11	1.7
NE	1.12	0.96	0.35	-0.22	-2.34	0.12	7.8
EI	-	-	-	-	-	-	-
EW	-	-	-	-	-	-	-
GY	-1.05	-1.09	-0.54	0.36	0.66	1.31	5.3
TGW	-0.37	0.82	0.82	-0.28	-0.15	-0.39	1.8
	Protein-nutritional						
CP	-0.41	0.10	-0.33	0.95	-0.23	-0.52	1.5
Lys	0.34	-0.07	-0.89	-0.14	1.75	0.27	4.1
Met	-0.84	0.60	0.94	0.47	0.05	-0.39	2.3
Cys	-0.29	1.25	-0.49	-0.79	-0.85	0.92	4.1
Thr	-	-	-	-	-	-	-
Trp	0.88	-0.81	0.85	1.38	-0.98	0.53	5.3
Val	-	-	-	-	-	-	-
Ile	-	-	-	-	-	-	-
Leu	0.62	-0.11	-0.02	-1.01	0.33	-1.37	3.4
Phe	-	-	-	-	-	-	-
His	-	-	-	-	-	-	-
Arg	-	-	-	-	-	-	-
r	0.79*	0.70 ^{ns}	0.48 ^{ns}	0.41 ^{ns}	0.21 ^{ns}	0.17 ^{ns}	
	Agronomic						
MF	-	-	-	-	-	-	-
FF	-1.06	-0.09	-0.69	-1.17	-0.08	-	3.9
PH	-1.11	0.07	0.00	0.61	-0.06	-	1.7
EH	0.75	0.46	-0.83	-0.01	-0.15	-	1.6
EP	-	-	-	-	-	-	-
NP	0.26	-0.07	0.49	0.46	0.06	-	1.7
NE	-0.85	-0.35	-2.01	-0.09	0.97	-	7.8
EI	-	-	-	-	-	-	-
EW	-	-	-	-	-	-	-
GY	-0.23	-0.53	0.67	-1.11	-1.15	-	5.3
TGW	0.30	0.30	-0.03	0.02	1.27	-	1.8
	Energetic-nutritional						
AME	-	-	-	-	-	-	-
AMEn	0.70	0.51	-0.17	0.23	0.83	-	1.5
EE	-0.47	0.35	-0.60	-0.38	-1.00	-	1.8
CF	-0.22	-0.28	-0.60	0.84	0.13	-	1.2
ST	0.15	0.54	0.55	0.19	-0.87	-	1.4
AML	0.64	-0.42	-0.46	-0.43	-0.45	-	1.2
r	0.66 ^{ns}	0.58 ^{ns}	0.35 ^{ns}	0.25 ^{ns}	0.04 ^{ns}		

For abbreviations, see Table 1. *Significant by the chi-square test at 5% error probability. ^{ns}Non-significant.

agronomic and energetic-nutritional variables, respectively (Table 5). Among agronomic and protein-nutritional variables, there was a reduction in the number of canonical pairs with the elimination of variables. On the other hand, between agronomic and energetic-nutritional variables, there was no reduction in the number of canonical pairs with the elimination of variables. However, the magnitude of the canonical correlations was lower with the elimination of variables between agronomic and protein-nutritional variables, and between agronomic and energetic-nutritional variables.

In transgenic genotypes, without elimination of variables, the first five canonical pairs presented significant canonical correlations ($r = 1.00$) for all pairs between agronomic and

Table 5. Canonical coefficients of canonical pairs between the agronomic and protein-nutritional, and agronomic and energetic-nutritional variables of 22 super-early maturing maize genotypes (the canonical correlation analysis was performed with the elimination of variables).

Variables ⁽¹⁾	Canonical pairs						VIF ⁽²⁾
	1st	2nd	3rd	4th	5th	6th	
	Agronomic						
MF	-	-	-	-	-	-	-
FF	0.69	1.52	-0.17	-0.62	2.07	0.94	8.4
PH	0.25	-0.43	0.51	-0.59	0.13	-1.06	2.2
EH	-0.42	0.61	-0.64	0.69	-1.19	-0.16	3.6
EP	-	-	-	-	-	-	-
NP	0.18	-0.07	0.37	0.28	-0.79	-0.02	1.5
NE	-0.14	1.28	-1.87	-1.28	0.85	0.21	7.7
EI	-	-	-	-	-	-	-
EW	-	-	-	-	-	-	-
GY	-0.45	0.55	1.69	0.68	0.43	0.37	4.7
TGW	0.58	0.06	-0.21	-0.84	-0.15	0.56	1.5
	Protein-nutritional						
CP	-0.45	0.54	-0.68	-0.61	0.66	0.59	2.1
Lys	-0.04	-1.44	0.44	-0.52	-0.80	1.39	5.1
Met	0.58	1.36	-0.12	-0.73	-0.38	-0.30	3.0
Cys	1.26	0.23	1.89	1.28	0.53	0.07	7.1
Thr	-	-	-	-	-	-	-
Trp	-0.86	0.91	-1.69	1.25	-0.70	-0.26	6.6
Val	-	-	-	-	-	-	-
Ile	-	-	-	-	-	-	-
Leu	0.26	-1.38	-0.63	-0.83	0.91	-0.98	4.8
Phe	-	-	-	-	-	-	-
His	-	-	-	-	-	-	-
Arg	-	-	-	-	-	-	-
r	0.93*	0.80 ^{ns}	0.64 ^{ns}	0.57 ^{ns}	0.45 ^{ns}	0.21 ^{ns}	-
	Agronomic						
MF	-	-	-	-	-	-	-
FF	0.29	0.01	0.29	0.76	-1.06	-	8.4
PH	-0.22	0.22	0.16	-0.88	-0.90	-	2.2
EH	0.33	-0.16	0.60	0.42	0.88	-	3.6
EP	-	-	-	-	-	-	-
NP	0.47	0.21	-0.83	-0.28	-0.14	-	1.5
NE	-0.99	0.29	0.51	1.68	-1.41	-	7.7
EI	-	-	-	-	-	-	-
EW	-	-	-	-	-	-	-
GY	0.61	-0.30	0.03	-1.37	1.01	-	4.7
TGW	0.04	0.98	0.14	0.57	-0.06	-	1.5
	Energetic-nutritional						
AME	-	-	-	-	-	-	-
AMEn	-0.35	2.39	0.40	-1.38	0.16	-	7.9
EE	1.23	-1.65	-0.13	1.24	-0.59	-	6.1
CF	0.05	-0.41	0.74	-0.50	-0.54	-	1.3
ST	-0.47	0.03	0.53	0.73	-0.21	-	1.1
AML	0.00	-0.99	0.70	0.24	0.69	-	2.0
r	0.92*	0.80 ^{ns}	0.69 ^{ns}	0.48 ^{ns}	0.22 ^{ns}	-	-

For abbreviations, see Table 1. *Significant by the chi-square test at 5% error probability. ^{ns}Non-significant.

protein-nutritional variables. The canonical correlation of the first canonical pair was significant ($r = 1.00$) between agronomic and energetic-nutritional variables (Table 3). Nevertheless, with the elimination of variables, canonical correlations were not significant for all canonical pairs between agronomic and protein-nutritional variables, and between agronomic and energetic-nutritional variables (Table 6). The occurrence of at least one canonical pair with significant canonical correlation shows the existence of linear dependence between the groups of variables under study.

In early maturing, super-early maturing, and transgenic genotypes, the magnitude of the canonical correlation for each canonical pair was always greater without elimination of

Table 6. Canonical coefficients of canonical pairs between the agronomic and protein-nutritional, and agronomic and energetic-nutritional variables of 18 transgenic maize genotypes (the canonical correlation analysis was performed with the elimination of variables).

Variables ⁽¹⁾	Canonical pairs						VIF ⁽²⁾
	1st	2nd	3rd	4th	5th	6th	
	Agronomic						
MF	-	-	-	-	-	-	-
FF	-0.24	-0.21	1.13	-1.14	0.57	-2.11	7.8
PH	-0.27	-0.36	-1.76	-1.34	0.33	-0.80	7.2
EH	-0.29	-0.23	0.84	1.73	-1.07	2.36	11.5
EP	-	-	-	-	-	-	-
NP	-0.21	0.09	0.04	0.68	-0.31	-0.51	1.3
NE	-0.46	-0.43	-0.30	-0.20	0.62	-0.47	2.2
EI	-	-	-	-	-	-	-
EW	-	-	-	-	-	-	-
GY	0.40	-1.05	0.59	-0.53	0.43	-0.94	3.0
TGW	-0.31	0.68	0.47	0.06	0.57	0.38	1.6
	Protein-nutritional						
CP	-0.21	0.38	0.19	1.37	-0.51	0.03	2.4
Lys	1.51	1.00	0.59	-0.13	-0.30	-0.54	4.0
Met	-1.08	0.28	-0.42	0.34	0.89	-0.91	3.2
Cys	-1.03	-0.21	0.96	-0.56	1.24	-0.59	4.2
Thr	-	-	-	-	-	-	-
Trp	-0.89	-0.29	-0.27	-1.16	-0.67	0.81	3.4
Val	-	-	-	-	-	-	-
Ile	-	-	-	-	-	-	-
Leu	0.90	-0.16	-0.91	0.24	-0.02	1.48	3.9
Phe	-	-	-	-	-	-	-
His	-	-	-	-	-	-	-
Arg	-	-	-	-	-	-	-
r	0.94 ^{ns}	0.88 ^{ns}	0.62 ^{ns}	0.51 ^{ns}	0.36 ^{ns}	0.17 ^{ns}	
	Agronomic						
MF	-	-	-	-	-	-	-
FF	-0.83	0.06	-1.70	1.11	1.53	-	7.8
PH	2.16	-0.86	-0.75	1.05	0.24	-	7.2
EH	-1.05	1.53	1.66	-1.77	-1.09	-	11.5
EP	-	-	-	-	-	-	-
NP	-0.25	-0.27	-0.08	0.22	0.05	-	1.3
NE	-0.06	-0.16	-0.79	0.04	0.03	-	2.2
EI	-	-	-	-	-	-	-
EW	-	-	-	-	-	-	-
GY	-0.13	-0.23	-0.29	0.34	1.56	-	3.0
TGW	-0.40	0.04	0.55	0.84	-0.52	-	1.6
	Energetic-nutritional						
AME	-	-	-	-	-	-	-
AMEn	1.03	-0.84	1.09	0.09	-0.82	-	3.6
EE	-0.27	1.36	-1.15	0.34	-0.14	-	3.4
CF	1.08	0.21	-0.12	0.52	0.40	-	1.7
ST	-1.00	-0.04	0.63	0.36	0.58	-	1.9
AML	-0.24	0.51	0.34	-0.96	-0.03	-	1.4
r	0.92 ^{ns}	0.85 ^{ns}	0.59 ^{ns}	0.37 ^{ns}	0.21 ^{ns}		

For abbreviations, see Table 1. ^{ns}The chi-square test at 5% error probability was non-significant.

variables than with the elimination of variables between groups of agronomic and protein-nutritional variables, and between agronomic and energetic-nutritional variables (Tables 1, 2, 3, 4, 5, and 6). This shows that the presence of multicollinearity (without elimination of variables) affects the estimation of canonical correlations.

In early maturing genotypes, the only canonical pair with significant canonical correlation, with elimination of variables, shows linear dependence between the groups of agronomic and protein-nutritional variables, and shows that the higher the number of ears (NE), the higher the tryptophan content (Trp) in the grains (Table 4).

In super-early maturing genotypes, with and without elimination of variables, the results highlight a linear dependence between the groups of variables. However, the relationships between canonical pairs, with and without elimination of variables, are distinct. Without elimination of variables, the canonical coefficients presented wide variability, making adequate inferences impossible (Table 2). With elimination of variables, the variability of the canonical coefficients was reduced, allowing biological inferences (Table 5). Among the agronomic and protein-nutritional variables, the canonical pair with significant canonical correlation shows that the more the number of days from sowing until female flowering (FF), the higher the cysteine (Cys) content in grains (Table 5). In the case of agronomic and energetic-nutritional variables, the canonical pair with significant canonical correlation shows that the lower the number of ears (NE), the higher the ether extract (EE) content in grains (Table 5). A study by Alves et al. (2016a,b) found a linear dependence between the group of phenological and protein-nutritional variables and between the groups of productive and energetic-nutritional variables, and the same associations among the variables. In the present study, the phenological, morphological, and productive variables were grouped in a group of agronomic variables. Thus, the relationships were confirmed.

In transgenic genotypes, without elimination of variables, a wide variability was observed in the estimation of canonical coefficients, and it was not possible to perform adequate associations between the groups of agronomic and protein-nutritional, and agronomic and energetic-nutritional variables (Table 3). With the elimination of variables, the canonical pairs presented non-significant canonical correlations, showing that the groups were independent, i.e., it was not possible to make associations between groups of variables (Table 6).

In general, it was observed that in early maturing, super-early maturing, and transgenic genotypes, the canonical correlations without elimination of variables were greater than those with elimination of variables. In the presence of multicollinearity, the estimated canonical coefficients can assume absurd values without any coherence with the biological phenomenon under study. In the presence of multicollinearity, the super estimation has already been seen in path analysis for soybean (Bizeti et al., 2004), canola (Coimbra et al., 2005), and maize (Toebe and Cargnelutti Filho, 2013a,b). The analysis of canonical correlation in the presence of multicollinearity in maize crop was performed by Ceccon et al. (2016), who verified linear dependence between the group of primary and secondary components of production. However, owing to non-compliance with the multicollinearity assumption, the canonical correlations and the estimated canonical coefficients can generate misunderstandings in the interpretation of the results.

The results obtained in the present study show that the analysis of canonical correlation in the presence of multicollinearity (without elimination of variables) presents a wide variability of canonical coefficients, resulting in overestimation of the canonical coefficients and providing unreliable estimates. This makes the interpretation of the relationships between the groups of agronomic and protein-nutritional, and between agronomic and energetic-nutritional variables in early maturing, super-early maturing, and transgenic genotypes difficult. Therefore, it can be inferred that the elimination of variables is a solution to circumvent the effects of multicollinearity in canonical correlation analysis and make accurate inferences in the genetic improvement of maize plants, which can enhance the nutritional quality of grains through indirect selection.

CONCLUSIONS

The canonical correlation analysis in the presence of multicollinearity (without elimination of variables) overestimates the variability of canonical coefficients.

The elimination of variables is an efficient method to circumvent multicollinearity in canonical correlation analysis.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Processes #401372/2015-4 and #305057/2013-8) for granting scholarships to the authors. Mr. Washington Neves of the Adisseo Brasil Company kindly conducted the laboratory tests. Student fellows and volunteers assisted in data collection.

REFERENCES

- Alves BM, Cargnelutti Filho A, Toebe M and Burin C (2016a). Linear relations among phenological, morphological, productive and protein-nutritional traits in early maturing and super-early maturing maize genotypes. *J. Cereal Sci.* 70: 229-239. <http://dx.doi.org/10.1016/j.jcs.2016.06.013>
- Alves BM, Cargnelutti Filho A, Burin C and Toebe M (2016b). Correlações canônicas entre caracteres agrônômicos e nutricionais proteicos e energéticos em genótipos de milho. *Rev. Bras. Milho Sorgo* 15: 171-185. <http://dx.doi.org/10.18512/1980-6477/rbms.v15n2p171-185>
- Bizeti HS, Carvalho CGP, Souza JRP and Destro D (2004). Path analysis under multicollinearity in soybean. *Braz. Arch. Biol. Technol.* 47: 669-676. <http://dx.doi.org/10.1590/S1516-89132004000500001>
- Brum B, Lopes SJ, Storck L, Lúcio AD, et al. (2011). Correlações canônicas entre variáveis de semente, plântula, planta e produção de grãos em mamoneira. *Cienc. Rural* 41: 404-411. <http://dx.doi.org/10.1590/S0103-84782011000300007>
- Carvalho CGP, Oliveira VR, Cruz CD and Casali VWD (1999). Análise de trilha sob multicolinearidade em pimentão. *Pesqui. Agropecu. Bras.* 34: 603-613. <http://dx.doi.org/10.1590/S0100-204X1999000400011>
- Carvalho CGP, Borsato R, Cruz CD and Viana JMS (2001). Path analysis under multicollinearity in S0 x S0 maize hybrids. *Crop Breed. Appl. Biotechnol.* 1: 263-270. <http://dx.doi.org/10.13082/1984-7033.v01n03a06>
- Cecon G, Santos A, Teodoro PE and Junior CAS (2016). Relationships between primary and secondary yield components of a maize population after 13 stratified mass selection cycles. *J. Agron.* 15: 33-38. <http://dx.doi.org/10.3923/ja.2016.33.38>
- Coimbra JLM, Benin G, Vieira EA, Oliveira AC, et al. (2005). Consequências da multicolinearidade sobre a análise de trilha em canola. *Cienc. Rural* 35: 347-352. <http://dx.doi.org/10.1590/S0103-84782005000200015>
- CONAB - Companhia Nacional de Abastecimento (2016). Acompanhamento da safra brasileira: grãos, décimo segundo levantamento. Conab, Brasília. Available at [http://www.conab.gov.br/OlalaCMS/uploads/arquivos/16_09_09_15_18_32_boletim_12_setembro.pdf]. Accessed October 23, 2016.
- Cruz CD (2013). GENES - a software package for analysis in experimental statistics and quantitative genetics. *Acta Sci. Agron.* 35: 271-276. <http://dx.doi.org/10.4025/actasciagron.v35i3.21251>
- Cruz CD and Regazzi AJ (1997). Modelos biométricos aplicados ao melhoramento genético. UFV, Viçosa.
- Duvick DN (2005). The contribution of breeding to yield advances in maize (*Zea mays* L.). *Adv. Agron.* 86: 83-145.
- FAO (2016). FAOSTAT. Available at [<http://faostat3.fao.org/browse/Q/QC/E>]. Accessed October 23, 2016.
- Gujarati DN and Porter DC (2011). Econometria básica. 5th edn. Porto Alegre: AMGH.
- Hair JF, Black WC, Babin BJ, Anderson RE, et al. (2009). Análise multivariada de dados. 6th edn. Bookman, Porto Alegre.
- Martinez C and Cuevas-Perez F (1989). Evaluación de localidad culinária y molineradel arroz. 3rd edn. CIAT, Cali.
- Montgomery DC and Peck EA (1982). Introduction to linear regression analysis. John Wiley and Sons, New York.

- Nogueira APO, Sedyama T, Sousa LB, Hamawaki OT, et al. (2012). Análise de trilha e correlações entre caracteres em soja cultivada em duas épocas de semeadura. *Biosci. J.* 28: 877-888.
- Toebe M and Cargnelutti Filho A (2013a). Multicollinearity in path analysis of maize (*Zea mays* L.). *J. Cereal Sci.* 57: 453-462. <http://dx.doi.org/10.1016/j.jcs.2013.01.014>
- Toebe M and Cargnelutti Filho A (2013b). Não normalidade multivariada e multicolinearidade na análise de trilha em milho. *Pesqui. Agropecu. Bras.* 48: 466-477. <http://dx.doi.org/10.1590/S0100-204X2013000500002>