Revista Brasileira de Milho e Sorgo

Brazilian Journal of Maize and Sorghum

ISSN 1980 - 6477

Journal homepage: www.abms.org.br/site/paginas

SAMPLE SIZE AND LINEAR RELATIONS IN TRAITS OF CUT AND GRAZING SORGHUM

Alberto Cargnelutti Filho[1] (✉), Valéria Escaio Bubans[1], Lucas Fillipin Osmari[1], Felipe Manfio Somavilla[1] and Vithória Morena Ortiz[1]

[1]Federal University of Santa Maria - Plant Science Department - Santa Maria (RS), Brazil.

E-mail: alberto.cargnelutti.filho@gmail.com, valeriabubans@hotmail.com, lucasfosmari@gmail.com, felipe-somavilla@hotmail.com, vithoria.ortiz159@gmail.com.

✉ Corresponding author

Abstract – The objectives of this work were to determine the sample size (number of plants) needed to estimate the mean of cut and grazing sorghum (*Sorghum bicolor* (L.) Moench), Nutribem cultivar, traits and investigate the linear relations among traits. At 86 days after sowing, 110 plants of sorghum were selected at random. The traits evaluated for each plant were: plant height, stem diameter, number of nodes, number of leaves, leaf fresh matter, stem fresh matter, shoot fresh matter, leaf dry matter, stem dry matter, and shoot dry matter. The sample size was determined to estimate the means of the traits, assuming estimation errors equal to 1% (higher precision), 2%, ..., and 20% (lower precision) of the mean. Scatter plots, correlation analysis, and path analysis investigated the relationship among traits. Fourteen plants were needed to estimate the means of plant height, stem diameter, number of nodes, and number of cut and grazing sorghum leaves, with a maximum error of 10% of the mean and a 95% confidence level. With the same precision, to estimate the means of leaf, stem, and shoot fresh and dry matter, 48 plants are needed. Plant height positively correlates with stem and shoots fresh and dry matter. The number of leaves has a positive linear relation with leaf fresh and dry matter.

Keywords: *Sorghum bicolor*, Sampling sufficiency, Experimental precision, Correlation, Path analysis.

TAMANHO DE AMOSTRA E RELAÇÕES LINEARES EM CARACTERES DE SORGO DE CORTE E PASTEJO

Resumo - Os objetivos deste trabalho foram determinar o tamanho de amostra (número de plantas) necessário para a estimação da média de caracteres de sorgo de corte e pastejo (*Sorghum bicolor* (L.) Moench), cultivar Nutribem, e investigar as relações lineares entre os caracteres. Aos 86 dias após a semeadura, foram selecionadas, aleatoriamente, 110 plantas de sorgo. Foram avaliados os seguintes caracteres: altura de planta, diâmetro de caule, número de nós, número de folhas, matéria fresca de folhas, matéria fresca de caule, matéria fresca de parte aérea, matéria seca de folhas, matéria seca de caule e matéria seca de parte aérea. Foi determinado o tamanho de amostra para a estimação da média dos caracteres, assumindo erros de estimação iguais a 1% (maior precisão), 2%, …, 20% (menor precisão) da média. Foi investigada a relação entre os caracteres por meio de diagramas de dispersão e análises de correlação e de trilha. Para a estimação da média da altura de planta, diâmetro de caule, número de nós e número de folhas de sorgo de corte e pastejo, com erro máximo de 10% da média e grau de confiança de 95%, são necessárias 14 plantas. Com essa mesma precisão, para a estimação da média das matérias fresca e seca de folhas, caule e parte aérea, são necessárias 48 plantas. A altura de planta tem relação linear positiva com as matérias fresca e seca de caule e de parte aérea. O número de folhas tem relação linear positiva com as matérias fresca e seca de folhas.

Palavras-chave: *Sorghum bicolor*, Suficiência amostral, Precisão experimental, Correlação, Análise de trilha.

How to cite

CARGNELUTTI FILHO, A.; BUBANS, V. E.; OSMARI, L. F.; SOMAVILLA, F. M.; ORTIZ, V. M. Sample size and linear relations in traits of cut and grazing sorghum. **Revista Brasileira de Milho e Sorgo**, v. 21, e1279, 2022.

DOI: https://doi.org/10.18512/rbms2022v21e1279

In experiments with crops, such as sorghum (*Sorghum bicolor* (L.) Moench), in its different types, classified according to their purpose of use (grain, silage, sweet, biomass, broom, and cut and grazing), when evaluating a trait it is common to find variability among plants, even among those subjected to the same treatment. Thus, it is essential to size the number of plants that must be evaluated to obtain reliable information on the trait under evaluation.

Sample sizing has been performed to estimate the mean traits of species of the Poaceae family, such as sorghum (Silva et al., 2005), maize (Toebe et al., 2014; Wartha et al., 2016), black oat (Cargnelutti Filho et al., 2015b), millet (Kleinpaul et al., 2017) and rye (Bandeira et al., 2018, 2019). These studies have found that, for estimating the mean with the same precision, there is variation in sample size (number of plants) among traits. Furthermore, for the same trait, there is variation in sample size among genotypes (Toebe et al., 2014; Wartha et al., 2016; Bandeira et al., 2018, 2019), sowing times (Bandeira et al., 2018, 2019), evaluation times (Cargnelutti Filho et al., 2015b; Kleinpaul et al., 2017; Bandeira et al., 2018) and agricultural years (Toebe et al., 2014).

In field experiments, it is common to measure several plant traits, which may or may not be linearly correlated. For example, in the case of sorghum (*Sorghum bicolor* (L.) Moench) cultivars used for cutting and grazing, such as the Nutribem cultivar, more significant amounts of fresh and dry matter are desired. Quantifying fresh and dry matter requires destructive sampling, so it is essential to know their relations with other traits, such as plant height, stem diameter, number of nodes, and number of leaves. Once correlations are proven, it is possible to estimate fresh and dry matter yields without needing to harvest plants.

The interrelations among traits can be visualized in scatter plots and investigated using Pearson's linear correlation coefficient (r) and path analysis. The coefficient r measures the intensity or degree of the linear relation between two random variables. The r sign expresses the association direction, while the intensity is represented by a numerical value between -1 and 1. In extreme situations, two traits may have perfect negative ($r = -1$) or perfect positive ($r = 1$) linear correlation, or lack of linear relation ($r = 0$) (Ferreira, 2009; Bussab & Morettin, 2017). Path analysis makes it possible to partition the correlation coefficients into direct and indirect effects of explanatory variables on the main variable and identify if there is a linear association between cause and effect (Cruz et al., 2012, 2014; Cruz, 2016).

Correlation and path analysis has been used in sorghum (Lombardi et al., 2015; Vendruscolo et al., 2016; Ceccon et al., 2017; Silva et al., 2017; Mengesha et al., 2019; Oliveira et al., 2021) and black oat (Cargnelutti Filho et al., 2015a). These studies have found critical linear associations among the traits used in plant breeding programs. Therefore, it is assumed that the inclusion of studies on sample sizing and association among sorghum traits aggregates essential information to support the planning of experiments with better precision and, consequently, with more excellent reliability in the results.

Thus, the objectives of this work were to determine the sample size (number of plants) needed to estimate the mean of cut and grazing sorghum, Nutribem cultivar, and traits and investigate the linear relations among traits.

Material and Methods

A uniformity trial (blank experiment) was conducted with cut and grazing sorghum, Nutribem cultivar, in Santa Maria, State of Rio Grande do Sul, Brazil (coordinates 29º42'S, 53º49'W, and 95 m altitudes), in an area of 8 m × 20 m (160 m²). The climate of this site is humid subtropical - Cfa, according to the Köppen-Geiger classification, with hot summers and no dry season (Alvares et al., 2013). The soil is *Argissolo Vermelho Distrófico Arênico* (Ultisol) (Santos et al., 2018). On October 28, 2020, 35 kg ha^{-1} of N, 135 kg ha^{-1} of P_2O_5, and 135 kg ha^{-1} of K_2O were incorporated into the soil, and sowing was carried out the broadcast, using 15 kg of seeds ha^{-1}, aiming at 450,000 plants ha^{-1}.

In the crop's flowering (reproductive stage), that is, at 86 days after sowing, a pilot sample of 110 plants was randomly harvested. The plants were separated into two parts, leaves and stem + inflorescence, considered in this study as leaves and stem, respectively, because some plants had no inflorescence yet. In each plant, the following traits were evaluated: plant height (PH, in cm), stem diameter (SD, in cm, measured with a caliper at 2 cm from the soil surface), number of nodes (NN), number of leaves (NL), leaf fresh matter (LFM, in g $plant^{-1}$), stem fresh matter (STFM, in g $plant^{-1}$), shoot fresh matter (SHFM = LFM + STFM, in g $plant^{-1}$), leaf dry matter (LDM, in g $plant^{-1}$), stem dry matter (STDM, in g $plant^{-1}$) and shoot dry matter (SHDM = LDM + STDM, in g $plant^{-1}$).

For the ten traits (PH, SD, NN, NL, LFM, STFM, SHFM, LDM, STDM, and SHDM), the following statistics were calculated: minimum; percentiles 1, 2.5, 25, 50 (median), 75, 97.5, and 99; maximum; amplitude; mean; variance; standard deviation; standard error of the mean; coefficient of variation; skewness; kurtosis; and p-value of the Kolmogorov-Smirnov normality test.

For each trait, based on the 110 plants, the sample size (n) was determined for estimation errors (semi-amplitudes of the confidence interval) fixed at 1%, 2%, ..., 20% of the mean (m), that is, 0.01×m (higher precision), 0.02×m, ..., 0.20×m (lower precision), with a 95% confidence level (1-α) by the expression $n = \left(\frac{t_{\alpha/2}\, s}{estimation\ error}\right)^2$ (Ferreira, 2009; Bussab & Morettin, 2017). The $t_{\alpha/2}$ is the critical value of the Student's t-distribution, whose area on the right is equal to α/2, that is, the value of t, such that $P(t>t_{\alpha/2})=\alpha/2$, with α=5% probability of error and n-1 degrees of freedom (n = 110 plants from the pilot sample and therefore 109 degrees of freedom in this study), and s is the estimate of the standard deviation.

After fixing n equal to 110 plants, the sample size used in the sampling, the estimation error was calculated as a percentage of the mean estimate (m) for each trait, using the expression: $estimation\ error = \frac{t_{\alpha/2}\, s}{\sqrt{n}\, m} \times 10$

Scatter plots were initially constructed to investigate the relations among the traits PH, SD, NN, NL, LFM, STFM, SHFM, LDM, STDM, and SHDM. Then, the matrix of Pearson's linear correlation coefficients (r) among the traits was determined, and Student's t-test was performed at a 5% significance level.

In the correlation matrix among the traits PH, SD, NN, and NL, multicollinearity was diagnosed and interpreted based on the condition number (CN), which is the ratio between the highest and lowest eigenvalue of the correlation matrix. Multicollinearity was classified as weak when CN ≤ 100, moderate to severe when 100 < CN < 1000, and severe when CN ≥ 1000 (Cruz et al., 2014; Cruz, 2016).

After that, path analysis of the main variables (LFM, STFM, SHFM, LDM, STDM, and SHDM) as a function of the explanatory variables (PH, SD, NN, and NL) was performed according to the methodology described in Cruz et al. (2012, 2014), totaling six path analysis. Finally, statistical analyses were conducted using Microsoft Office Excel® and Genes (Cruz, 2016).

Results and Discussion

For the data of plant height (PH), stem diameter (SD), number of nodes (NN), number of leaves (NL), leaf fresh matter (LFM), stem fresh matter (STFM), shoot fresh matter (SHFM), leaf dry matter (LDM), stem dry matter (STDM) and shoot dry matter (SHDM), the p-value of the Kolmogorov-Smirnov test ranged from 0.00 to 0.826, with mean of 0.431 (Table 1). The higher the p-value, the greater the adherence of the data to the normal distribution curve. Thus, assuming the significance level of 24.5%, it can be inferred that the assumption of normality was met for the traits PH, LFM, STFM, SHFM, LDM, STDM, and SHDM. For this test, the traits SD, NN, and NL showed low adherence to the normal distribution (p-value ≤ 0.023). However, the coefficients of skewness and kurtosis that do not differ from zero and three, respectively, and the mean's proximity to the median indicates that the data of the traits SD, NN, and NL showed slight distances from the standard distribution curve. Therefore, this data set is suitable for the studies of sample sizing based on Student's t-distribution and linear relations through correlation and path analysis.

The coefficient of variation (CV) for the traits PH, SD, NN, and NL ranged between 12.39% and 18.83%. On the other hand, the CV values for traits related to fresh and dry matter yields (LFM, STFM, SHFM, LDM, STDM, and SHDM) were relatively higher, ranging from 31.71% to 34.85%. (Table 1). Thus, for the same precision, a larger sample size is expected to estimate the means of the production traits (LFM, STFM, SHFM, LDM, STDM, and SHDM) compared to morphological traits (PH, SD, NN, and NL). Similar pattern was observed in black oat traits (Cargnelutti Filho et al., 2015b).

The sample sizes (number of plants) for estimating the mean, with estimation error (semi-amplitude of the 95% confidence interval) equal to 1% of the mean estimate (m), that is, 0.01×m (higher precision, in this study), ranged from 604 plants for PH to 4771 plants for LFM (Table 2). In Microsoft Office Excel®, these sizes are calculated by the following expressions: =ROUND.UP((((INVT(0.05;109)*19.7871)/(0.01*159.6636))^2);0) = 604 plants and = ROUND.UP((((INVT(0.05;109)*24.2730)/(0.01*69.6495))^2);0) = 4771 plants. Therefore, concerning plant height, it can be inferred, with 95% confidence, that the mean confidence interval (μ) obtained with 604 plants is μ±0.01m, that is, μ±1.597 cm because the mean height of the 110 plants sampled was 159.66 cm (Table 1). In the other extreme, the precision of μ±0.01m is obtained with 4771 plants for leaf fresh matter and, in this situation, the value would have been μ±0.696 grams because the mean leaf fresh matter of the 110 plants sampled was 69.65 grams. These results reveal that, for the same precision, the sample sizes vary among the traits, as observed in sorghum (Silva et al., 2005), maize (Toebe et al., 2014; Wartha et al., 2016), black oat (Cargnelutti Filho et al., 2015b), millet (Kleinpaul et al., 2017) and rye (Bandeira et al., 2018, 2019). It is also possible to infer that it is difficult to obtain estimates of the mean with this higher level of precision due to the high number of plants to be sampled.

With this estimation error of 1%, a smaller sample size was observed for the traits PH, SD, NN and NL ($604 \leq n \leq 1394$) when compared to the sample size required for the traits LFM, STFM, SHFM, LDM, STDM, and SHDM ($3951 \leq n \leq 4771$) (Table 2). These results are due to the higher coefficients of variation of the traits LFM, STFM, SHFM, LDM, STDM, and SHDM compared to those of the traits PH, SD, NN, and NL (Table 1). In black oat, a larger sample size is also necessary to estimate the means of production traits (fresh matter and dry matter)

DOI: https://doi.org/10.18512/rbms2022v21e1279

Table 1. Minimum, percentiles 1, 2.5, 25, 50 (median), 75, 97.5, and 99, maximum, amplitude, mean, variance, standard deviation, standard error, coefficient of variation (CV), skewness, kurtosis, and p-value of the Kolmogorov-Smirnov normality test of traits[(1)] measured in 110 plants of cut and grazing sorghum (*Sorghum bicolor* (L.) Moench), Nutribem cultivar.

Statistics	PH	SD	NN	NL	LFM	STFM	SHFM	LDM	STDM	SHDM
Minimum	109.00	1.00	6.00	3.00	25.19	73.18	102.12	6.43	10.27	17.00
Percentile 1	110.45	1.01	6.09	3.00	28.80	78.69	109.82	6.76	11.72	18.92
Percentile 2.5	120.80	1.17	7.00	3.73	30.16	95.14	128.56	7.55	14.17	22.70
Percentile 25	144.25	1.40	8.00	5.00	50.19	142.41	202.62	12.63	22.06	34.44
Median	163.00	1.60	9.00	5.00	70.47	198.23	273.76	17.52	31.51	47.78
Percentile 75	174.75	1.80	10.00	6.00	85.18	244.04	325.47	20.06	38.85	58.58
Percentile 97.5	190.55	2.20	11.00	7.00	115.23	318.06	426.95	27.93	51.56	79.18
Percentile 99	192.91	2.20	11.00	7.00	119.74	326.10	441.10	29.65	53.24	81.08
Maximum	197.00	2.50	11.00	7.00	123.45	327.05	445.72	30.82	58.40	89.22
Amplitude	88.00	1.50	5.00	4.00	98.26	253.87	343.60	24.39	48.13	72.22
Mean	159.66	1.60	8.91	5.25	69.65	198.02	267.67	16.82	31.04	47.86
Variance	391.53	0.08	1.33	0.98	589.18	4053.90	7205.30	32.65	109.37	241.30
Standard Deviation	19.79	0.29	1.15	0.99	24.27	63.67	84.88	5.71	10.46	15.53
Standard error	1.89	0.03	0.11	0.09	2.31	6.07	8.09	0.54	1.00	1.48
CV (%)	12.39	18.02	12.95	18.83	34.85	32.15	31.71	33.97	33.69	32.46
Skewness [(2)]	-0.42ns	0.44ns	-0.04ns	0.07ns	0.11ns	0.06ns	0.02ns	0.15ns	0.27ns	0.20ns
Kurtosis + 3 [(3)]	2.46ns	2.91ns	2.63ns	2.55ns	2.29ns	2.09*	2.19ns	2.51ns	2.43ns	2.55ns
p-value	0.245	0.023	0.005	0.000	0.809	0.625	0.551	0.826	0.603	0.627

[(1)] PH - plant height, in cm; SD - stem diameter, in cm; NN - number of nodes; NL - number of leaves; LFM - leaf fresh matter, in g plant^{-1}; STFM - stem fresh matter, in g plant^{-1}; SHFM - shoot fresh matter (SHFM = LFM + STFM), in g plant^{-1}; LDM - leaf dry matter, in g plant^{-1}; STDM - stem dry matter, in g plant^{-1}; and SHDM - shoot dry matter (SHDM = LDM + STDM), in g plant^{-1}. [(2)] * Skewness differs from zero, through the Student's t-test, at a 5% significance level. [(3)] * Kurtosis differs from three, through the Student's t-test, at a 5% significance level.

DOI: https://doi.org/10.18512/rbms2022v21e1279

Table 2. Sample size (number of plants) for estimating the mean of traits[(1)] of cut and grazing sorghum (*Sorghum bicolor* (L.) Moench), Nutribem cultivar, for estimation errors (semi-amplitudes of the confidence interval) equal to 1%, 2%, ..., 20% of the mean (m), that is, 0.01×m (higher precision), 0.02×m, ..., 0.20×m (lower precision), with a 95% confidence level (1-α). Estimation error, as a percentage of the mean estimate (m), based on the 110 plants sampled (Error, in %).

Error	PH	SD	NN	NL	LFM	STFM	SHFM	LDM	STDM	SHDM
1%	604	1275	659	1394	4771	4062	3951	4534	4459	4138
2%	151	319	165	349	1193	1016	988	1134	1115	1035
3%	68	142	74	155	531	452	439	504	496	460
4%	38	80	42	88	299	254	247	284	279	259
5%	25	51	27	56	191	163	159	182	179	166
6%	17	36	19	39	133	113	110	126	124	115
7%	13	27	14	29	98	83	81	93	91	85
8%	10	20	11	22	75	64	62	71	70	65
9%	8	16	9	18	59	51	49	56	56	52
10%	7	13	7	14	48	41	40	46	45	42
11%	5	11	6	12	40	34	33	38	37	35
12%	5	9	5	10	34	29	28	32	31	29
13%	4	8	4	9	29	25	24	27	27	25
14%	4	7	4	8	25	21	21	24	23	22
15%	3	6	3	7	22	19	18	21	20	19
16%	3	5	3	6	19	16	16	18	18	17
17%	3	5	3	5	17	15	14	16	16	15
18%	2	4	3	5	15	13	13	14	14	13
19%	2	4	2	4	14	12	11	13	13	12
20%	2	4	2	4	12	11	10	12	12	11
Error (%)	2.34	3.40	2.45	3.56	6.59	6.08	5.99	6.42	6.37	6.13

[(1)] PH - plant height, in cm; SD - stem diameter, in cm; NN - number of nodes; NL - number of leaves; LFM - leaf fresh matter, in g $plant^{-1}$; STFM - stem fresh matter, in g $plant^{-1}$; SHFM - shoot fresh matter (SHFM = LFM + STFM), in g $plant^{-1}$; LDM - leaf dry matter, in g $plant^{-1}$; STDM - stem dry matter, in g $plant^{-1}$; and SHDM - shoot dry matter (SHDM = LDM + STDM), in g $plant^{-1}$.

DOI: https://doi.org/10.18512/rbms2022v21e1279

compared to morphological traits (plant height, number of leaves, and number of tillers) (Cargnelutti Filho et al., 2015b).

When opting for allowing an estimation error of 20%, that is, 0.20×m (lower precision, in this study), and 95% confidence level, the number of plants to be sampled decreases to 2, 4, 2, 4, 12, 11, 10, 12, 12 and 11, respectively, for the traits PH, SD, NN, NL, LFM, STFM, SHFM, LDM, STDM, and SHDM (Table 2). An experiment would quickly evaluate this low number of plants (≤ 12). However, it would lead to less precision in estimating the means of the traits, that is, μ±31.93 cm, μ±0.32 cm, μ±1.78 nodes, μ±1.05 leaves, μ±13.93 grams, μ±39.60 grams, μ±53.53 grams, μ±3.36 grams, μ±6.21 grams, μ±9.57 grams, respectively, for the traits PH, SD, NN, NL, LFM, STFM, SHFM, LDM, STDM, and SHDM.

The results of this work allow the researcher to choose the sample size to estimate the mean of these traits with the desired precision. For example, if the option is a maximum estimation error of 10%, that is, 0.10×m, 48 plants would be sufficient to estimate the means of these ten traits (Table 2). Thus, when planning an experiment to be conducted in the field, in a completely randomized experimental design, to estimate the mean of each treatment with 10% precision, 48 plants per treatment should be evaluated. For example, if the experiment is planned with four replicates per treatment, 12 plants per replicate (48/4 = 12) would be sampled, that is, 12 plants per plot. Furthermore, if ten treatments were evaluated in the experiment, 480 plants (48 per treatment) would be sampled. It is worth pointing out that, for the traits PH, SD, NN, and NL, individual evaluations of the 12 plants in the plot are required, while for LFM, STFM, SHFM, LDM, STDM, and SHDM, the weighing of the 12 plants in the plot can be performed together.

If the option were to sample 110 plants, as used in this study, the estimation error would be 2.34, 3.40, 2.45, 3.56, 6.59, 6.08, 5.99, 6.42, 6.37, and 6.13% of the mean estimate (m), respectively, for the traits PH, SD, NN, NL, LFM, STFM, SHFM, LDM, STDM, and SHDM, that is, the maximum estimation error would be 6.59% (Table 2). This estimation error refers to the LFM trait and in Microsoft Office Excel® is obtained by the following expression: =((INVT(0.05;109)*24.2730)/(RAIZ(110)*69.6495))*100 = 6.59%.

Positive linear associations were found among the traits with Pearson's linear correlation coefficients (r) ranging from 0.332 to 0.989, with a mean of 0.718 (Table 3). Therefore, as the scores of a given trait increase, there is an increment in the scores of another trait. This linear pattern is crucial for identifying traits for indirect selection (Figure 1).

Leaf, stem, and shoot fresh matter and dry matter (LFM, STFM, SHFM, LDM, STDM, and SHDM) showed a higher degree of positive linear association (higher r values) with NL (0.663 ≤ r ≤ 0.826, with mean = 0.750) and decreasing linear association with the other traits in the following order: SD (0.613 ≤ r ≤ 0.690, with mean = 0.651), PH (0.460 ≤ r ≤ 0.765, with mean = 0.643) and NN (0.509 ≤ r ≤ 0.651, with mean = 0.602) (Table 3), which suggests that the number of leaves would be more strongly associated with the leaf, stem and shoot fresh and dry matter of sorghum. Therefore, path analysis is an appropriate complementary procedure to infer the genuine cause-effect relationships among the traits (Cruz et al., 2012, 2014). Furthermore, this analysis makes it possible to infer which explanatory trait (PH, SD, NN, and NL) directly affects the traits LFM, STFM, SHFM, LDM, STDM, and SHDM.

The multicollinearity diagnosis in the Pearson's linear correlation coefficient (r) matrix, among the explanatory variables PH, SD, NN, and NL, revealed a condition number of 5.905; that is, the matrix showed

DOI: https://doi.org/10.18512/rbms2022v21e1279

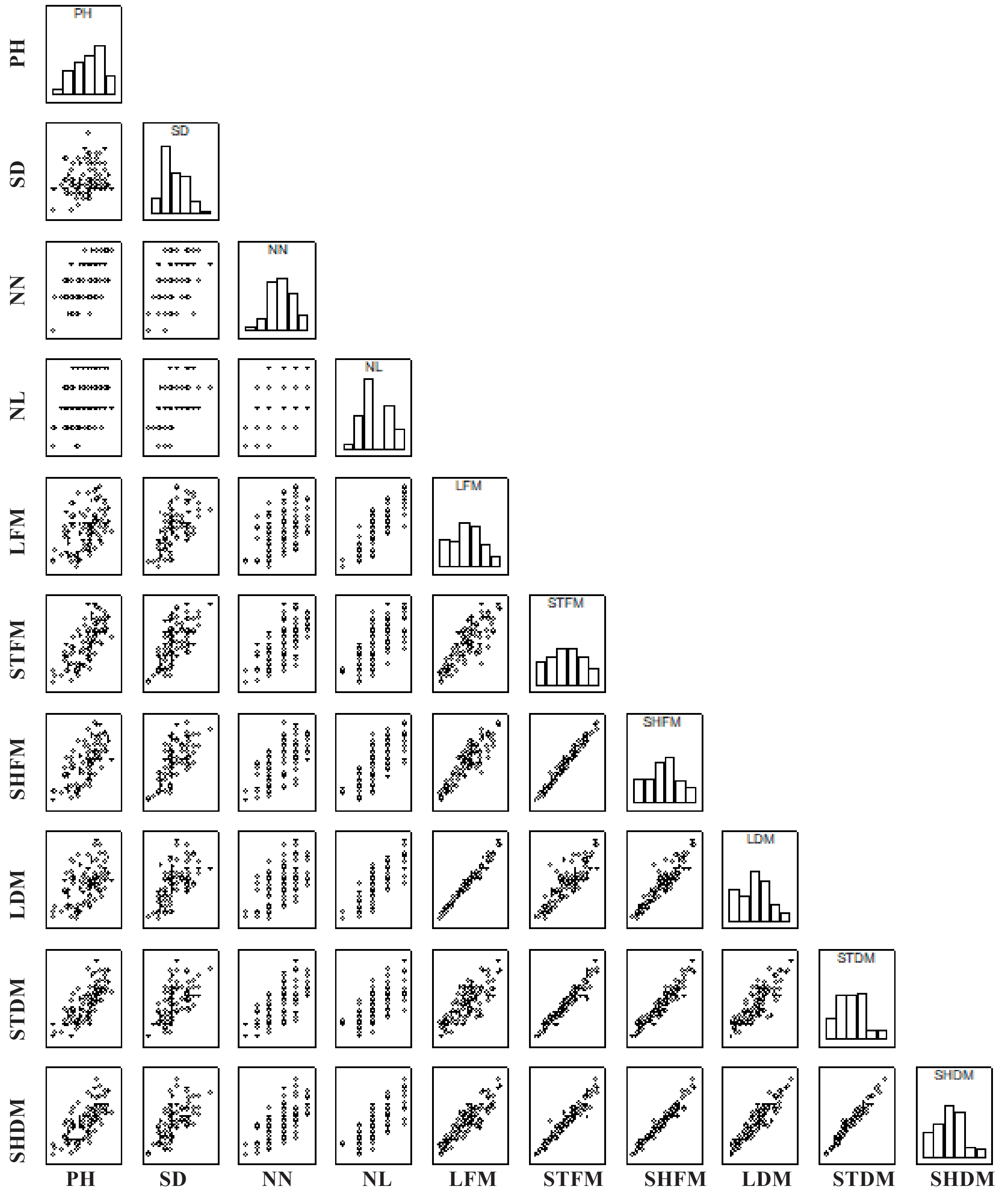

Figure 1. Matrix with frequency histogram (diagonally) and scatter plots (outside the diagonal) among plant height (PH), stem diameter (SD), number of nodes (NN), number of leaves (NL), leaf fresh matter (LFM), stem fresh matter (STFM), shoot fresh matter (SHFM = LFM + STFM), leaf dry matter (LDM), stem dry matter (STDM) and shoot dry matter (SHDM = LDM + STDM), measured at 86 days after sowing, in 110 plants of cut and grazing sorghum (*Sorghum bicolor* (L.) Moench), Nutribem cultivar.

Table 3. Pearson's linear correlation coefficients among the traits[(1)] measured in 110 plants of cut and grazing sorghum (*Sorghum bicolor* (L.) Moench), Nutribem cultivar.

	PH	SD	NN	NL	LFM	STFM	SHFM	LDM	STDM	SHDM
PH	1.000	0.332*	0.585*	0.412*	0.460*	0.742*	0.688*	0.503*	0.765*	0.700*
SD	0.332*	1.000	0.340*	0.552*	0.643*	0.675*	0.690*	0.639*	0.613*	0.648*
NN	0.585*	0.340*	1.000	0.462*	0.509*	0.637*	0.623*	0.551*	0.651*	0.641*
NL	0.412*	0.552*	0.462*	1.000	0.823*	0.686*	0.750*	0.826*	0.663*	0.750*
LFM	0.460*	0.643*	0.509*	0.823*	1.000	0.829*	0.908*	0.989*	0.786*	0.893*
STFM	0.742*	0.675*	0.637*	0.686*	0.829*	1.000	0.987*	0.856*	0.962*	0.963*
SHFM	0.688*	0.690*	0.623*	0.750*	0.908*	0.987*	1.000	0.925*	0.946*	0.977*
LDM	0.503*	0.639*	0.551*	0.826*	0.989*	0.856*	0.925*	1.000	0.831*	0.927*
STDM	0.765*	0.613*	0.651*	0.663*	0.786*	0.962*	0.946*	0.831*	1.000	0.979*
SHDM	0.700*	0.648*	0.641*	0.750*	0.893*	0.963*	0.977*	0.927*	0.979*	1.000

[(1)] PH - plant height, in cm; SD - stem diameter, in cm; NN - number of nodes; NL - number of leaves; LFM - leaf fresh matter, in g plant^{-1}; STFM - stem fresh matter, in g plant^{-1}; SHFM - shoot fresh matter (SHFM = LFM + STFM), in g plant^{-1}; LDM - leaf dry matter, in g plant^{-1}; STDM - stem dry matter, in g plant^{-1}; and SHDM - shoot dry matter (SHDM = LDM + STDM), in g plant^{-1}. * Significant at a 5% error probability by Student's t-test, with 108 degrees of freedom.

weak multicollinearity (Cruz et al., 2014; Cruz, 2016) (Table 4). Thus, it can be inferred that the adverse effect of multicollinearity is overcome and that the path analysis of the main traits LFM, STFM, SHFM, LDM, STDM, and SHDM of sorghum, as a function of the explanatory traits PH, SD, NN, and NL were performed under appropriate conditions (Cruz et al., 2012, 2014; Cruz, 2016).

SD showed a positive linear correlation ($0.613 \leq r \leq 0.690$) with the traits LFM, STFM, SHFM, LDM, STDM, and SHDM and direct effects ($0.232 \leq$ direct effect ≤ 0.342) of low magnitude. A similar linear association was also observed for NN ($0.509 \leq r \leq 0.651$), in addition to negligible direct effects ($0.103 \leq$ direct effect ≤ 0.171). Therefore, the associations of SD and NN with the traits LFM, STFM, SHFM, LDM, STDM, and SHDM are explained by indirect effects.

PH showed a positive linear correlation with LFM ($r = 0.460$) and with LDM ($r = 0.503$), and the direct effects on LFM (0.066) and LDM (0.101) were of low magnitude, with association explained by the more significant indirect effects of PH via NL on LFM (0.252) and LDM (0.244). However, there was a more significant association of PH with the traits STFM, SHFM, STDM, and SHDM ($0.688 \leq r \leq 0.765$) and direct effects ($0.347 \leq$ direct effect ≤ 0.478) with the same sign and of greater magnitude than indirect effects, confirming the cause-effect relationship between PH and the traits STFM, SHFM, STDM, and SHDM (Table 4). Thus, it can be inferred that taller plants have a more significant amount of stem and shoots fresh and dry matter.

NL showed a positive linear correlation with LFM ($r = 0.823$) and with LDM ($r = 0.826$), and the direct effects on LFM (0.611) and LDM (0.592) had

DOI: https://doi.org/10.18512/rbms2022v21e1279

Table 4. Pearson's correlation coefficients (r) and direct and indirect effects (path analysis) of the traits plant height (PH), stem diameter (SD), number of nodes (NN), and number of leaves (NL) on the traits leaf fresh matter (LFM), stem fresh matter (STFM), shoot fresh matter (SHFM = LFM + STFM), leaf dry matter (LDM), stem dry matter (STDM) and shoot dry matter (SHDM = LDM + STDM), measured at 86 days after sowing, in 110 plants of cut and grazing sorghum (*Sorghum bicolor* (L.) Moench), Nutribem cultivar.

Effect	Main variable					
	LFM	STFM	SHFM	LDM	STDM	SHDM
Direct of PH on	0.066	0.438	0.347	0.101	0.478	0.359
Indirect of PH via SD	0.083	0.114	0.109	0.077	0.087	0.087
Indirect of PH via NN	0.060	0.088	0.083	0.081	0.100	0.097
Indirect of PH via NL	0.252	0.102	0.148	0.244	0.100	0.157
Pearson's correlation (r)	0.460*	0.742*	0.688*	0.503*	0.765*	0.700*
Direct of SD on	0.249	0.342	0.327	0.232	0.262	0.262
Indirect of SD via PH	0.022	0.146	0.115	0.034	0.159	0.119
Indirect of SD via NN	0.035	0.051	0.048	0.047	0.058	0.056
Indirect of SD via NL	0.337	0.137	0.199	0.327	0.134	0.210
Pearson's correlation (r)	0.643*	0.675*	0.690*	0.639*	0.613*	0.648*
Direct of NN on	0.103	0.151	0.142	0.139	0.171	0.166
Indirect of NN via PH	0.039	0.256	0.203	0.059	0.279	0.210
Indirect of NN via SD	0.085	0.116	0.111	0.079	0.089	0.089
Indirect of NN via NL	0.283	0.115	0.167	0.274	0.112	0.176
Pearson's correlation (r)	0.509*	0.637*	0.623*	0.551*	0.651*	0.641*
Direct of NL on	0.611	0.248	0.361	0.592	0.242	0.381
Indirect of NL via PH	0.027	0.180	0.143	0.042	0.197	0.148
Indirect of NL via SD	0.138	0.189	0.181	0.128	0.145	0.144
Indirect of NL via NN	0.047	0.070	0.066	0.064	0.079	0.077
Pearson's correlation (r)	0.823*	0.686*	0.750*	0.826*	0.663*	0.750*
Coefficient of determination	0.746	0.821	0.824	0.765	0.798	0.813
Residual variable	0.254	0.179	0.176	0.235	0.202	0.187
Condition number	5.905	5.905	5.905	5.905	5.905	5.905

* Significant at 5% probability of error by Student's t-test, with 108 degrees of freedom.

the same sign and magnitude, confirming a cause-effect relationship (Table 4). An association of NL with the traits STFM, SHFM, STDM, and SHDM ($0.663 \leq r \leq 0.750$) was also observed, but the direct effects ($0.242 \leq$ direct effect ≤ 0.381) had a lower magnitude. Thus, plants with more leaves have more leaf fresh and dry matter.

Therefore, plant height and number of leaves can be used for indirect selection to increase sorghum plants' fresh and dry matter. The fact that it is not necessary to destroy the plants to measure the height and count the number of leaves is advantageous because it allows, if it is of interest, to keep the plants until the production of seeds. For direct selection, it would be necessary to destroy the plants for weighing the LFM, STFM, SHFM, LDM, STDM, and SHDM. In black oat, it was also found that the number of leaves per plant and plant height have a positive linear relation with fresh and dry matter and can be used for indirect selection (Cargnelutti Filho et al., 2015a). In sorghum (*Sorghum bicolor* (L.) Moench), cause-effect relationships among several traits have been found (Lombardi et al., 2015; Vendruscolo et al., 2016; Ceccon et al., 2017; Silva et al., 2017; Mengesha et al., 2019; Oliveira et al., 2021).

Conclusions

Fourteen plants are required to estimate the means of plant height, stem diameter, number of nodes, and number of leaves of cut and grazing sorghum (*Sorghum bicolor* (L.) Moench), Nutribem cultivar, with a maximum error of 10% of the mean and confidence level of 95%,. With this same precision, to estimate the means of leaf fresh matter, stem fresh matter, shoot fresh matter, leaf dry matter, stem dry matter, and shoot dry matter, 48 plants are required.

Plant height positively correlates with stem and shoots fresh and dry matter.

The number of leaves has a positive linear relation with leaf fresh and dry matter.

Acknowledgments

To the National Council for Scientific and Technological Development (CNPq - Processes 304652/2017-2; 159611/2019-9), the Coordination for the Improvement of Higher Education Personnel (CAPES) and the Rio Grande do Sul State Research Support Foundation (FAPERGS) for granting scholarships to the authors. To scholarship-holding students and volunteers for their assistance in data collection.

References

ALVARES, C. A.; STAPE, J. L.; SENTELHAS, P. C.; GONÇALVES, J. L. M.; SPAROVEK, G. Köppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, v. 22, n. 6, p. 711-728, 2013. DOI: https://doi.org/10.1127/0941-2948/2013/0507.

BANDEIRA, C. T.; CARGNELUTTI FILHO, A.; CHAVES, G. G.; NEU, I. M. M.; SILVEIRA, D. L.; PROCEDI, A. Suficiência amostral para estimar a média de caracteres produtivos de centeio. **Revista de Ciências Agrárias**, v. 42, n. 3, p. 751-760, 2019. DOI: https://doi.org/10.19084/rca.17759.

BANDEIRA, C. T.; CARGNELUTTI FILHO, A.; FOLLMANN, D. N.; BEM, C. M.; WARTHA, C. A.; THOMASI, R. M. Sample size to estimate the mean of morphological traits of rye cultivars in sowing dates and evaluation times. **Semina. Ciências Agrárias**, v. 39, n. 2, p. 521-532, 2018. DOI: https://doi.org/10.5433/1679-0359.2018v39n2p521.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística b***ásica*. São Paulo: Saraiva, 2017.

CARGNELUTTI FILHO, A.; TOEBE, M.; ALVES, B. M.; BURIN, C.; SANTOS, G. O.; FACCO, G.; NEU, I. M. M. Relações lineares entre caracteres de aveia preta. **Ciência Rural**, v. 45, n. 6, p. 985-992, 2015a. DOI: https://doi.org/10.1590/0103-8478cr20140500.

CARGNELUTTI FILHO, A.; TOEBE, M.; ALVES, B. M.; BURIN, C.; SANTOS, G. O.; FACCO, G.; NEU, I. M. M. Dimensionamento amostral para avaliar caracteres morfológicos e produtivos de aveia preta em épocas de avaliação. **Ciência Rural**, v. 45, n. 1, p. 9-13, 2015b. DOI: https://doi.org/10.1590/0103-8478cr20140504.

CECCON, G.; TEODORO, P. E.; SANTOS, A. Contribution of agronomic traits for sugar yield in sweet sorghum genotypes. **Revista Brasileira de Milho e Sorgo**, v. 16, n. 2, p. 232-239, 2017. DOI: https://doi.org/10.18512/1980-6477/rbms.v16n2p232-239.

CRUZ, C. D. Programa Genes: ampliado e integrado aos aplicativos R, Matlab e Selegen. **Acta Scientiarum Agronomy**, v. 38, n. 4, p. 547-552, 2016.

CRUZ, C. D.; CARNEIRO, P. C. S.; REGAZZI, A. J. **Modelos biométricos aplicados ao melhoramento genético**. 3. ed. Viçosa, MG: Universidade Federal de Viçosa, 2014.

CRUZ, C. D.; REGAZZI, A. J.; CARNEIRO, P. C. S. **Modelos biométricos aplicados ao melhoramento genético**. Viçosa, MG: Universidade Federal de Viçosa, 2012.

FERREIRA, D. F. **Estatística básica**. Lavras: Universidade Federal de Lavras, 2009.

KLEINPAUL, J. A.; CARGNELUTTI FILHO, A.; ALVES, B. M.; BURIN, C.; NEU, I. M. M.; SILVEIRA, D. L.; SIMÕES, F. M. Tamanho de amostra para estimação da média de caracteres de milheto em épocas de avaliação. **Revista Brasileira de Milho e Sorgo**, v. 16, n. 2, p. 251-262, 2017. DOI: https://doi.org/10.18512/1980-6477/rbms.v16n2p251-262.

LOMBARDI, G. M. R.; NUNES, J. A. R.; PARRELLA, R. A. C.; TEIXEIRA, D. H. L.; BRUZI, A. T.; DURÃES, N N. L.; FAGUNDES, T. G. Path analysis of agro-industrial traits in sweet sorghum. **Genetics and Molecular Research**, v. 14, n. 4, p. 16392-16402, 2015. DOI: http://dx.doi.org/10.4238/2015.

MENGESHA, G. H.; HAILEMARIAM, F. M.; MINDAYE, T. T.; LAKEW, B.; VERMA, R. P. S. Correlation and path analysis of yield, yield contributing and malt quality traits of Ethiopian sorghum (*Sorghum bicolor* (L.) Moench) genotypes. **African Journal of Plant Science**, v. 13, n. 8, p. 209-220, 2019. DOI: http://dx.doi.org/10.5897/AJPS2019.1788.

OLIVEIRA, T. C.; BARELLI, M. A. A.; SILVA, V. P.; AZEVEDO, R. F.; GONÇALVES, D. L.; SANTOS, P. R. J.; POLETINE, J. P.; GALBIATI, C.; TARDIN, F. D. Correlations between characters and path analysis in sweet sorghum (*Sorghum bicolor* (L.) Moench) genotypes for juice production. **Australian Journal of Crop Science**, v. 15, n. 2, p. 290-296, 2021. DOI: http://dx.doi.org/10.21475/ajcs.21.15.02.p3083.

SANTOS, H. G.; JACOMINE, P. K. T.; ANJOS, L. H. C.; OLIVEIRA, V. A.; LUMBRERAS, J. F.; COELHO, M. R.; ALMEIDA, J. A.; ARAÚJO FILHO, J. C.; OLIVEIRA, J. B.; CUNHA, T. J. F. **Sistema Brasileiro de Classificação de Solos.** 5. ed. rev. e ampl. Brasília, DF: Embrapa, 2018.

SILVA, K. J.; TEODORO, P. E.; MENEZES, C. B.; JÚLIO, M. P. M.; SOUZA, V. F.; SILVA, M. J.; PIMENTEL, L. D.; BORÉM, A. Contribution of morphoagronomic traits to grain yield and earliness in grain sorghum. **Genetics and Molecular Research**, v. 16, n. 2, gmr16029649, 2017. DOI: http://dx.doi.org/10.4238/gmr16029649.

SILVA, P. S. L.; BARBOSA, Z.; GONÇALVES, R. J. S.; SILVA, P. I. B.; NUNES, G. H. S. Sample size for the estimation of some sorghum traits. **Revista Brasileira de Milho e Sorgo**, v. 4, n. 2, p. 149-160, 2005. DOI: http://dx.doi.org/10.18512/1980-6477/rbms.v4n2p149-160.

TOEBE, M.; CARGNELUTTI FILHO, A.; BURIN, C.; CASAROTTO, G.; HAESBAERT, F. M. Tamanho de amostra para estimação da média e do coeficiente de variação em milho. **Pesquisa Agropecuária Brasileira**, v. 49, n. 11, p. 860-871, 2014. DOI: http://dx.doi.org/10.1590/S0100-204X2014001100005.

VENDRUSCOLO, T. P. S.; BARELLI, M. A. A.; CASTRILLON, M. A. S.; SILVA, R. S.; OLIVEIRA, F. T.; CORRÊA, C. L.; ZAGO, B. W.; TARDIN, F. D. Correlation and path analysis of biomass sorghum production. **Genetics and Molecular Research**, v. 15, n. 4, gmr15049086, 2016.

WARTHA, C. A.; CARGNELUTTI FILHO, A.; LÚCIO, A. D.; FOLLMANN, D. N.; KLEINPAUL, J. A.; SIMÕES, F. M. Sample sizes to estimate mean values for tassel traits in maize genotypes. **Genetics and Molecular Research**, v. 15, n. 4, gmr15049151, 2016. DOI: http://dx.doi.org/10.4238/gmr15049151.