http://dx.doi.org/10.1590/1678-4499.0324

Sample size in the estimation of correlation coefficients for corn hybrids in crops and accuracy levels

Marcos Toebe (1); Alberto Cargnelutti Filho (1*); Sidinei José Lopes (1); Cláudia Burin (1); Tatiani Reis da Silveira (1); Gabriele Casarotto (2)

(1) Universidade Federal de Santa Maria (UFSM), Departamento de Fitotecnia, Avenida Roraima, 1.000, 97105-900 Santa Maria (RS), Brasil.

(2) Universidade Federal do Rio Grande do Sul (UFRGS), Departamento de Agronomia, Avenida Paulo Gama, 110, 90040-060 Porto Alegre (RS), Brasil.

(*) Corresponding author: alberto.cargnelutti.filho@gmail.com

Received: *Sept. 19, 2014;* ***Accepted:*** *Oct. 8, 2014*

Basic Areas | Article

Abstract

This study determined the sample size necessary for the estimation of the Pearson linear correlation coefficients for single, triple and double corn hybrids in crops and accuracy levels. In 361, 373 and 416 plants, respectively, of the single, triple and double hybrids of the 2008/2009 crop and, in 1,777, 1,693 and 1,720 plants, respectively, of the single, triple and double hybrids of the 2009/2010 crop, twelve traits were measured: plant height, ear insertion height, ear weight, number of grain rows per ear, ear length and diameter, cob weight and diameter, weight of hundred grains, number of grains per ear, grain length and grain yield. Then, in each hybrid and crop, were estimated the correlation coefficients for the 66 pairs of traits and determined the sample size necessary to estimate the correlation coefficients in four accuracy levels [amplitudes of the confidence interval of 95% ($ACI_{95\%}$) of 0.15, 0.25, 0.35 and 0.45], by resampling with replacement. The sample size varies among hybrids, crops and pairs of traits. Larger sample size is required to estimate the correlation coefficient between weakly correlated traits and smaller sample size is needed to estimate the correlation coefficient between highly correlated traits. Independently of hybrid, crop and pairs of traits, 375, 195 and 120 plants are sufficient, respectively, to estimate the correlation coefficients with maximum $ACI_{95\%}$ of 0.25, 0.35 and 0.45.

Key words: *Zea mays* L, resampling, experimental design, linear relationships.

1. INTRODUCTION

Maize is the cereal with the largest volume of production worldwide estimated at 906.82 million tons for the crop of 2014/2015, in an area of 160.2 million hectares, Brazil is the world's third largest producer (FAO, 2014). Maize is used in food and feed and as industrial raw material, mainly due to the amount of reserves accumulated in grains (Fancelli & Dourado Neto, 2004). In this sense, the increase in corn productivity in recent decades has been assigned equally to management improvement and breeding (Duvick, 2005).

In breeding, plant selection can be performed directly or indirectly by studying the linear relationships between traits. For such analysis, it may be used the Pearson linear correlation coefficient (r), which measures the direction and intensity of the linear relationship between two random variables (Ferreira, 2009). The direction of the correlation may be either positive or negative, in the range of $-1 \le r \le 1$, wherein the intensity of the linear correlation is larger, the closer to $|1|$ is r. Additional studies as path analysis and canonical correlations have also been recommended for indirect selection of plants (Cruz & Regazzi, 1997).

For reliable results from linear relationships studies, correct sample size (number of plants) is required to be used for the estimation of correlation coefficients. These coefficients may be interpreted separately or used in further analysis, for example, path analysis and canonical correlations and therefore must be estimated accurately. Accordingly, Cargnelutti Filho et al. (2010) performed the sampling design for the estimation of linear correlation coefficients between traits of single, triple and double corn hybrids, based on data from a crop and for one accuracy level. The authors found that 300 plants were needed for the estimation of 91 pairs of traits, with maximum amplitude of 95% confidence interval ($ACI_{95\%}$) of 0.30, depending on the hybrid and pair of traits. Sample design studies for the estimation of Pearson correlation coefficients were also conducted in crambe (Cargnelutti Filho et al., 2011) and castor bean (Cargnelutti Filho et al., 2012). Still, Shieh (2010) evaluated the properties and the effect of sampling design on the linear correlation coefficient of Pearson and Bonett & Wright (2000) conducted sampling design studies for the estimation of Pearson, Kendall and Spearman correlation coefficients.

As noted above, sampling design studies for the estimation of correlation coefficients have been developed, including corn (Cargnelutti Filho et al., 2010). However, it is emphasized that the sample sizing, considering hybrids, crops and accuracy levels, is important to enable the reliable estimation of the correlation coefficients, widely used in studies of linear relationships. This study aimed to determine the sample size needed for the estimation of Pearson linear correlation coefficients for simple triple and double corn hybrids in crops and accuracy levels.

2. MATERIAL AND METHOD

Two experiments were conducted with corn (*Zea mays* L.), in the growing seasons of 2008/2009 (first experiment) and 2009/2010 (second experiment), in an area located in Santa Maria, Rio Grande do Sul State, Brazil (29°42'S, 53°49'W, 95 m altitude). In the first experiment, sown on 12/26/2008, four plots were sown with the single hybrid P32R21 four with the triple hybrid DKB566 and four with the double hybrid DKB747. In the second experiment, sown on 10/26/2009, sixteen plots were sown with the singe hybrid 30F53, sixteen with the triple hybrid DKB566 and sixteen with the double hybrid DKB747.

Each plot consisted of four rows of 6 m long, spaced at 0.80 m, with density adjusted to five plants per linear meter, representing a plant population of 62,500 plants ha^{-1}. Thus, each plot consisted of 120 plants, totaling 1,440 plants in the first experiment (3 hybrids × 4 plots/hybrid × 120 plants/plot) and 5,760 plants in the second experiment (3 hybrids × 16 plots/hybrid × 120 plants/plot). In each crop, plots of the single, triple and double hybrids were randomized in the experimental area. In both experiments, the basic fertilization was 750 kg ha^{-1} of the formula 3-24-18 (NPK) and the topdressing was 300 kg ha^{-1} of urea with 45% N. The other cultural practices were performed according to the recommendations for corn (Fancelli & Dourado Neto, 2004).

In the first experiment, we assessed 361, 373 and 416 plants, respectively, for single, triple and double hybrids and in the second experiment we evaluated 1,777, 1,693 and 1,720 plants, respectively, for single, triple and double hybrids. Plants evaluated contained the twelve traits described below. As a result, the final number of plants evaluated in each harvest differed between the single, triple and double hybrids. In each of the 6,340 plants, the following traits were measured: plant height at harvest (PH), ear insertion height (EIH), unhusked ear weight (EW), the number of grain rows per ear (NR), ear length (EL), ear diameter (ED), cob weight (CW), cob diameter (CD), weight of hundred grains (WHG), number of grains per ear (NGE), grain length (GL), calculated as the difference between the diameters of ear and cob divided by two, and grain yield (YIELD), in grams per plant. Next, for each hybrid in each experiment the Pearson linear correlation coefficient (r) was calculated for each of the 66 pairs of traits, and the significance of r checked by Student's t-test at 5% significance.

For each hybrid, in each experiment, 199 sample sizes were planned, with an initial sample size of ten plants and the others obtained with increment of five plants. Thus, the planned sample sizes were n = 10, 15, 20, ..., 1,000 plants. For each planned sample size, 1,000 resamples with replacement were obtained (Ferreira, 2009), the same number of resamples used in previous studies on sampling design to estimate the Pearson correlation coefficient (Cargnelutti Filho et al., 2010, 2012), and in each resample, the Pearson linear correlation coefficients (r) of the 66 pairs of traits were estimated. Thus, for each planned sample size 1,000 r estimates were obtained for each of the 66 pairs of traits. Based on these 1,000 estimates, we determined the 2.5% percentile, the mean, the 97.5% percentile and the range of the 95% confidence interval ($ACI_{95\%}$), by the difference between the 97.5% percentile and the 2.5% percentile.

To determine the sample size (number of plants) required for the estimation of r, from each of the 66 pairs of traits in each hybrid and experiment, it was initially set $ACI_{95\%}$ of r equal to 0.15 (higher accuracy), 0.25, 0.35 and 0.45 (lower accuracy). Then, we started with the initial sample size (n = 10 plants) and considered as adequate the sample size (n) the number of plants from which the $ACI_{95\%}$ of r was less than or equal to the maximum limit for each accuracy level (0.15, 0.25, 0.35 or 0.45). The correlation coefficients obtained with data from the first experiment and the sample size for $ACI_{95\%}$ of 0.30 (intermediate level of accuracy between 0.15 and 0.45) were presented by Cargnelutti Filho et al. (2010). Statistical analyses were run with the aid of the software R (R Development Core Team, 2014) and Microsoft Office Excel® software.

3. RESULTS AND DISCUSSION

The linear correlation coefficients (r) showed high variability among the 66 pairs of traits (Table 1). For hybrids evaluated in 2008/2009, r ranged from 0.02 (CD x GL) to 1.00 (EW x YIELD) for the single hybrid P32R21, from –0.20 (NR x WHG) to 1.00 (EW x YIELD) for the triple hybrid DKB566 and from –0.16 (NR x WHG) to 0.99 (EW x YIELD) for the double hybrid DKB747. In 2009/2010, the correlation coefficient ranged from –0.20 (EIH x EL) to 1.00 (EW x YIELD) for the single hybrid P32R21, from –0.12 (EIH x EL) to 1.00 (EW x YIELD) for the triple hybrid DKB566, and from –0.06 (NR x WHG) to 0.99 (EW x YIELD) for the double hybrid DKB747. Variability of

Table 1. Estimates of Pearson linear correlation coefficients[1] between the 66 pairs of traits of the single hybrid P32R21, the triple hybrid DKB566 and the double hybrid DKB747 in the 2008/2009 growing season and the single hybrid 30F53, the triple hybrid DKB566 and the double hybrid DKB747 in the 2009/2010 growing season

Trait[2]	PH	EIH	EW	NR	EL	ED	CW	CD	WHG	NGE	GL	YIELD
Single hybrid P32R21 (2008/2009) above the diagonal and triple hybrid DKB566 (2008/2009) below the diagonal[3]												
PH		0.61	0.40	0.21	0.38	0.35	0.40	0.29	0.29	0.33	0.24	0.38
EIH	0.74		0.26	0.09	0.24	0.24	0.20	0.14	0.22	0.23	0.20	0.26
EW	0.31	0.31		0.57	0.80	0.83	0.67	0.48	0.54	0.95	0.68	1.00
NR	0.09	0.17	0.42		0.38	0.70	0.23	0.41	0.11	0.65	0.57	0.59
EL	0.28	0.26	0.84	0.26		0.56	0.73	0.28	0.30	0.79	0.50	0.78
ED	0.26	0.31	0.83	0.57	0.66		0.47	0.55	0.53	0.79	0.84	0.83
CW	0.31	0.23	0.78	0.14	0.77	0.62		0.48	0.40	0.55	0.26	0.62
CD	0.19	0.16	0.54	0.43	0.44	0.68	0.61		0.40	0.40	0.02	0.46
WHG	0.27	0.18	0.48	–0.20	0.42	0.46	0.59	0.41		0.30	0.38	0.54
NGE	0.24	0.28	0.90	0.61	0.74	0.75	0.56	0.42	0.08		0.69	0.95
GL	0.22	0.30	0.73	0.46	0.57	0.86	0.40	0.21	0.33	0.70		0.70
YIELD	0.31	0.31	1.00	0.43	0.83	0.83	0.75	0.52	0.47	0.91	0.75	
Double hybrid DKB747 (2008/2009) above the diagonal and single hybrid 30F53 (2009/2010) below the diagonal												
PH		0.73	0.48	0.15	0.44	0.43	0.44	0.29	0.35	0.42	0.30	0.47
EIH	0.58		0.37	0.04	0.30	0.35	0.26	0.15	0.29	0.34	0.32	0.38
EW	0.42	–0.12		0.33	0.86	0.81	0.86	0.56	0.59	0.91	0.56	0.99
NR	0.31	0.01	0.53		0.20	0.54	0.20	0.46	–0.16	0.49	0.28	0.35
EL	0.34	–0.20	0.92	0.43		0.61	0.80	0.45	0.47	0.79	0.39	0.84
ED	0.39	–0.02	0.86	0.67	0.76		0.69	0.68	0.50	0.75	0.69	0.81
CW	0.42	-0.08	0.94	0.46	0.88	0.80		0.65	0.57	0.69	0.29	0.80
CD	0.36	0.04	0.68	0.54	0.58	0.73	0.70		0.35	0.47	–0.06	0.52
WHG	0.16	–0.14	0.49	-0.06	0.41	0.32	0.52	0.30		0.23	0.34	0.58
NGE	0.41	–0.09	0.94	0.62	0.88	0.85	0.83	0.64	0.19		0.56	0.92
GL	0.28	–0.05	0.72	0.55	0.65	0.88	0.62	0.31	0.23	0.73		0.60
YIELD	0.42	–0.13	1.00	0.54	0.91	0.86	0.92	0.67	0.48	0.94	0.72	
Triple hybrid DKB566 (2009/2010) above the diagonal and double hybrid DKB747 (2009/2010) below the diagonal												
PH		0.67	0.24	0.17	0.20	0.21	0.24	0.25	0.15	0.22	0.12	0.23
EIH	0.76		–0.01	0.02	–0.12	0.00	–0.02	0.03	-0.11	0.04	–0.02	–0.01
EW	0.29	0.15		0.42	0.79	0.86	0.85	0.64	0.55	0.91	0.72	1.00
NR	0.12	0.02	0.39		0.27	0.58	0.30	0.44	–0.03	0.54	0.48	0.43
EL	0.28	0.08	0.81	0.31		0.62	0.81	0.54	0.44	0.72	0.47	0.77
ED	0.26	0.15	0.82	0.59	0.62		0.69	0.67	0.42	0.83	0.89	0.86
CW	0.32	0.15	0.86	0.33	0.84	0.68		0.67	0.57	0.69	0.49	0.82
CD	0.31	0.15	0.70	0.50	0.61	0.77	0.74		0.36	0.59	0.26	0.63
WHG	0.23	0.11	0.47	–0.06	0.47	0.35	0.54	0.35		0.20	0.32	0.54
NGE	0.22	0.12	0.90	0.48	0.68	0.79	0.66	0.62	0.09		0.72	0.92
GL	0.09	0.08	0.57	0.40	0.34	0.76	0.31	0.18	0.20	0.60		0.73
YIELD	0.27	0.15	0.99	0.39	0.77	0.82	0.79	0.66	0.44	0.92	0.61	

(1) Value greater than |0.10|, |0.10| and |0.09|, respectively, for the single hybrid P32R21, the triple hybrid DKB566 and the double hybrid DKB747 of corn in 2008/2009 and greater than |0.04|, |0.04| and |0.04|, respectively, for the single hybrid 30F53, the triple hybrid DKB566 and the double hybrid DKB747 of corn in 2009/2010, is significant at 5% by t-test, with respectively 359, 371, 414, 1,775, 1,691 and 1,718 degrees of freedom.(2) PH: plant height at harvest; EIH: ear insertion height; EW: ear weight; NR: number of grain rows per ear; EL: ear length; ED: ear diameter; CW: cob weight; CD: cob diameter; WHG: weight of hundred grains; NGE: number of grains per ear; GL: grain length; YIELD: grain yield. (3) Correlation coefficients between the 66 pairs of traits in the single, triple and double hybrids in the experiment conducted in 2008/2009 have been previously presented by Cargnelutti Filho et al. (2010).

level and direction of the correlation coefficients were also observed in 91 pairs of traits in single, triple and double corn hybrids evaluated in a growing season (Cargnelutti Filho et al., 2010), in 210 pairs of traits in crambe (Cargnelutti Filho et al., 2011) and in 210 pairs of traits in castor bean hybrids (Cargnelutti Filho et al., 2012).

In general, weaker correlations between pairs of traits showed larger amplitudes between hybrids and crops, as can be observed for correlations between EIH and EW ($-0.12 \le r \le 0.37$), EIH and EL ($-0.20 \le r \le 0.30$), EIH and WHG ($-0.14 \le r \le 0.29$), EIH and NGE ($-0.09 \le r \le 0.34$), EIH and YIELD ($-0.13 \le r \le 0.38$), NR and WHG ($-0.20 \le r \le 0.11$) and between CD and GL ($-0.06 \le r \le 0.31$) (Table 1). Conversely, higher correlations between pairs of traits showed smaller fluctuations between hybrids and crops, as can be seen, for example, for the correlations between ED and NGE ($0.75 \le r \le 0.85$), EW and ED ($0.81 \le r \le 0.86$), ED and YIELD ($0.81 \le r \le 0.86$), EW and NGE ($0.90 \le r \le 0.95$), NGE and YIELD ($0.91 \le r \le 0.95$) and between EW and YIELD ($0.99 \le r \le 1.00$).

The sample size required for estimation of r, with ACI95% lower than or equal to 0.15, showed high variability among the 66 pairs of traits measured in single hybrid (10 plants ≤ n ≤ 890 plants), triple hybrid (10 plants ≤ n ≤ 990 plants) and double hybrid (10 plants ≤ n ≤ 800 plants) of the 2008/2009 crop and also in single hybrid (10 plants ≤ n > 1,000 plants), triple hybrid (10 plants ≤ n ≤ 825 plants) and double hybrid (10 plants ≤ n ≤ 880 plants) of 2009/2010 (Table 2). The larger sample sizes were needed for the estimation of the correlation coefficient between NR and EL (445 plants ≤ n ≤ 935 plants), NR and WHG (575 plants ≤ n ≤ 775 plants), PH and CD (565 plants ≤ n ≤ 800 plants), WHG and NGE (625 plants ≤ n ≤ 845 plants), EIH and CD (565 plants ≤ n ≤ 990 plants), CD and GL (590 plants ≤ n ≤ 880 plants) and between WHG and GL (585 plants ≤ n > 1,000 plants). These pairs of traits presented low correlation coefficients in all hybrids (single, triple and double) evaluated in 2008/2009 and 2009/2010 ($-0.20 \le r \le 0.43$) (Table 1).

Smaller sample sizes were required for estimation of r between traits with higher correlation, for example, for example, between EW and ED (40 plants ≤ n ≤ 100 plants), ED and YIELD (45 plants ≤ n ≤ 100 plants), EW and EL (25 plants ≤ n ≤ 120 plants) and between EL and YIELD (30 plants ≤ n ≤ 145 plants) (Table 2), which showed correlation coefficients in the range of $0.77 \le r \le 0.92$ (Table 1). Even smaller sample sizes were needed for the estimation of r between the traits EW and NGE (15 plants ≤ n ≤ 40 plants), NGE and YIELD (10 plants ≤ n ≤ 30 plants), and between EW and YIELD, which required only 10 plants to estimate the r, with $ACI_{95\%}$ less than or equal to 0.15, regardless of the hybrid and the crop (Table 2). These pairs of traits showed the highest correlations; the correlation between EW and NGE ($0.90 \le r \le 0.95$), NGE and YIELD ($0.91 \le r \le 0.95$) and between EW and YIELD ($0.99 \le r \le 1.00$) exhibited high values (Table 1). Thus, the higher the correlation between two traits, the smaller the sample size required for estimation of correlation, and vice versa, at a certain level of accuracy, as was already observed in previous studies (Bonett & Wright, 2000; Cargnelutti Filho et al., 2010, 2011, 2012; Shieh, 2010).

For conditions in which it is desired to estimate r for each of the 66 pairs of traits with maximum $ACI_{95\%}$ of 0.25, it would be necessary the measurement of 320, 365 and 295 plants, respectively, in the single hybrid P32R21, the triple hybrid DKB566 and the double hybrid DKB747 in 2008/09 and, 375, 315 and 310 plants, respectively, in the single hybrid 30F53, the triple hybrid DKB566 and the double hybrid DKB747 in 2009/2010 (Table 3). Thus, regardless of the hybrid, the harvest and the pair of traits, it would be recommended to measure 375 plants for estimating r with maximum $ACI_{95\%}$ of 0.25.

For estimation of r of each of the 66 pairs of traits with maximum $ACI_{95\%}$ of 0.35, it would be required the measurement of 165, 195 and 155 plants, respectively, in single, triple and double hybrids in 2008/2009 and, 190, 165 and 160 plants, respectively, in single, triple and double hybrids in 2009/2010 (Table 4). In turn, the measurement of 100, 110 and 100 plants, respectively, in single, triple and double hybrids in 2008/2009 and 120, 100 and 100 plants, respectively, in single, triple and double hybrids in 2009/2010 would be sufficient for the estimation of the correlation coefficient for each of the 66 pairs of traits with maximum $ACI_{95\%}$ of 0.45% (Table 5).

For the estimation of r in 91 pairs of traits in corn with maximum $ACI_{95\%}$ of 0.30, Cargnelutti Filho et al. (2010) recommended the measurement of up to 300 plants, depending on the hybrid and the pair of traits. In crambe, Cargnelutti Filho et al. (2011) found that for the estimation of r with maximum $ACI_{95\%}$ of 0.15%, the sample size ranged between 8 and 665 plants according to the pair of traits considered. Besides that, in 210 pairs of traits of two castor bean hybrids, Cargnelutti Filho et al. (2012) reported that 96 plants were sufficient for the estimation of r with maximum $ACI_{95\%}$ of 0.52. The authors also verified that for the estimation of r with $ACI_{95\%}$ of 0.20, the sample size varied between 10 and 661 plants, depending on the pair of traits.

In agreement with Bonett & Wright (2000) for the estimation of r through the Fisher confidence interval with $ACI_{95\%}$ of 0.10, it was necessary a sample size of n = 1,507 and n = 63 observations, respectively, for low (r = 0.10) and high (r = 0.90) correlation coefficients. The sample size of n = 168 and n = 13 observations would be sufficient, according to the authors, for estimation of

Table 2. Sample size (number of plants) for estimation of the Pearson correlation coefficient of 66 pairs of traits measured in the single hybrid P32R21, the triple hybrid DKB566 and the double hybrid DKB747 in the 2008/2009 growing season and in the single hybrid 30F53, the triple hybrid DKB566 and the double hybrid DKB747 in the 2009/2010 growing season for the range of the 95% confidence interval of 0.15

Trait[1]	PH	EIH	EW	NR	EL	ED	CW	CD	WHG	NGE	GL	YIELD
Single hybrid P32R21 (2008/2009) above the diagonal and triple hybrid DKB566 (2008/2009) below the diagonal												
PH		325	590	600	705	620	535	795	850	670	520	610
EIH	125		675	615	685	745	600	890	675	675	635	675
EW	615	600		280	105	60	290	475	370	15	180	10
NR	520	575	540		445	210	575	410	775	240	435	255
EL	660	520	70	935		290	165	790	640	125	405	135
ED	715	715	70	400	300		410	420	495	100	85	50
CW	570	605	125	565	150	250		385	565	410	670	365
CD	800	990	355	405	440	255	300		610	585	815	475
WHG	620	665	575	635	805	600	325	470		675	715	400
NGE	545	595	35	345	170	200	260	410	795		195	10
GL	675	575	135	585	410	75	355	590	750	255		180
YIELD	615	580	10	540	80	75	155	380	610	30	135	
Double hybrid DKB747 (2008/2009) above the diagonal and single hybrid 30F53 (2009/2010) below the diagonal												
PH		195	525	685	525	600	480	775	620	525	640	525
EIH	235		515	680	705	555	580	765	500	590	465	510
EW	510	550		520	55	100	60	390	300	40	420	10
NR	670	585	390		715	385	615	440	575	425	615	520
EL	650	510	25	610		350	70	500	395	115	635	65
ED	595	510	40	345	160		175	260	415	190	240	100
CW	515	570	30	420	40	110		240	285	200	635	105
CD	655	635	390	370	375	235	230		575	475	800	410
WHG	745	510	635	760	715	935	580	640		675	585	360
NGE	575	545	15	290	65	70	80	485	845		465	30
GL	595	520	145	530	305	80	250	865	>1.000	185		390
YIELD	505	550	10	390	30	50	40	400	675	15	145	
Triple hybrid DKB566 (2009/2010) above the diagonal and double hybrid DKB747 (2009/2010) below the diagonal												
PH		200	675	770	720	805	705	630	670	685	760	685
EIH	125		740	775	705	825	705	705	650	810	825	745
EW	585	590		520	120	45	95	300	470	25	140	10
NR	690	670	535		710	430	565	475	715	435	625	520
EL	630	640	95	655		365	85	465	565	200	495	130
ED	665	540	70	410	300		180	245	585	75	60	45
CW	560	590	85	580	70	175		280	490	200	360	135
CD	565	565	205	515	300	150	150		520	410	760	310
WHG	510	580	395	645	405	480	300	470		755	660	510
NGE	695	590	25	435	255	125	305	370	625		190	25
GL	745	560	350	675	605	160	590	880	625	305		140
YIELD	650	590	10	535	145	70	175	260	470	20	275	

[1] PH: plant height at harvest; EIH: ear insertion height; EW: ear weight; NR: number of grain rows per ear; EL: ear length; ED: ear diameter; CW: cob weight; CD: cob diameter; WHG: weight of hundred grains; NGE: number of grains per ear; GL: grain length; YIELD: grain yield.

Table 3. Sample size (number of plants) for estimation of the Pearson correlation coefficient of 66 pairs of traits measured in the single hybrid P32R21, the triple hybrid DKB566 and the double hybrid DKB747 in the 2008/2009 growing season and in the single hybrid 30F53, the triple hybrid DKB566 and the double hybrid DKB747 in the 2009/2010 growing season for the range of the 95% confidence interval of 0.25

Trait[1]	PH	EIH	EW	NR	EL	ED	CW	CD	WHG	NGE	GL	YIELD
Single hybrid P32R21 (2008/2009) above the diagonal and triple hybrid DKB566 (2008/2009) below the diagonal												
PH		120	225	255	240	240	200	290	315	225	190	225
EIH	55		245	230	255	290	245	320	250	245	235	245
EW	225	235		110	40	25	120	180	135	10	70	10
NR	225	225	230		175	90	215	170	290	90	165	100
EL	230	215	35	340		125	65	290	250	50	155	55

[1] PH: plant height at harvest; EIH: ear insertion height; EW: ear weight; NR: number of grain rows per ear; EL: ear length; ED: ear diameter; CW: cob weight; CD: cob diameter; WHG: weight of hundred grains; NGE: number of grains per ear; GL: grain length; YIELD: grain yield.

Table 3. Continued...

Trait[(1)]	PH	EIH	EW	NR	EL	ED	CW	CD	WHG	NGE	GL	YIELD
ED	265	265	35	140	115		160	155	205	40	35	25
CW	215	215	50	205	55	95		150	210	155	240	155
CD	285	365	130	165	175	100	115		220	215	305	180
WHG	225	255	210	215	275	225	115	175		255	270	135
NGE	215	215	15	130	65	75	95	180	285		75	10
GL	225	215	55	210	150	35	125	235	285	105		70
YIELD	225	225	10	230	35	40	65	130	215	15	55	
Double hybrid DKB747 (2008/2009) above the diagonal and single hybrid 30F53 (2009/2010) below the diagonal												
PH		80	200	255	200	235	185	275	220	200	260	200
EIH	95		185	250	245	205	210	290	190	205	190	185
EW	195	215		210	25	40	25	140	125	20	155	10
NR	250	220	145		250	145	225	160	205	175	230	210
EL	225	195	15	225		130	35	180	135	50	245	30
ED	215	215	20	125	60		75	95	155	70	95	40
CW	195	200	10	160	25	40		90	115	70	240	35
CD	225	240	140	160	130	75	85		225	170	295	160
WHG	270	190	230	305	275	340	205	250		245	220	130
NGE	195	210	10	120	25	30	35	175	295		165	15
GL	225	180	65	210	120	40	100	330	375	70		150
YIELD	195	215	10	145	15	25	15	160	230	10	65	
Triple hybrid DKB566 (2009/2010) above the diagonal and double hybrid DKB747 (2009/2010) below the diagonal												
PH		80	235	270	300	300	250	245	240	265	275	235
EIH	50		275	300	250	315	265	265	225	275	310	275
EW	230	240		190	45	25	35	120	165	15	60	10
NR	270	250	190		275	165	205	190	250	165	230	190
EL	230	235	40	255		140	35	175	200	75	190	50
ED	250	225	25	155	115		75	90	200	35	25	25
CW	210	245	35	230	30	70		110	190	80	135	40
CD	235	220	85	185	125	55	55		195	150	275	120
WHG	210	230	150	260	150	190	120	170		280	260	195
NGE	250	240	15	155	95	45	105	125	245		70	10
GL	255	235	125	270	210	70	220	310	230	130		55
YIELD	230	240	10	190	55	35	70	100	160	15	110	

[(1)] PH: plant height at harvest; EIH: ear insertion height; EW: ear weight; NR: number of grain rows per ear; EL: ear length; ED: ear diameter; CW: cob weight; CD: cob diameter; WHG: weight of hundred grains; NGE: number of grains per ear; GL: grain length; YIELD: grain yield.

Table 4. Sample size (number of plants) for estimation of the Pearson correlation coefficient of 66 pairs of traits measured in the single hybrid P32R21, the triple hybrid DKB566 and the double hybrid DKB747 in the 2008/2009 growing season and in the single hybrid 30F53, the triple hybrid DKB566 and the double hybrid DKB747 in the 2009/2010 growing season for the range of the 95% confidence interval of 0.35

Trait[(1)]	PH	EIH	EW	NR	EL	ED	CW	CD	WHG	NGE	GL	YIELD
Single hybrid P32R21 (2008/2009) above the diagonal and triple hybrid DKB566 (2008/2009) below the diagonal												
PH		70	120	135	130	110	105	150	160	125	105	120
EIH	30		125	125	140	140	125	165	125	140	120	125
EW	115	120		60	25	15	60	100	75	10	40	10
NR	120	125	125		85	50	120	95	135	55	90	60
EL	125	120	15	175		60	35	155	130	30	70	30
ED	140	135	25	80	60		85	90	100	25	25	15
CW	115	120	30	125	25	50		80	120	80	120	75
CD	155	195	70	85	95	50	60		125	120	160	100
WHG	125	130	110	130	155	115	65	90		120	135	80
NGE	115	115	10	75	35	45	55	90	160		40	10
GL	120	130	35	125	75	20	70	125	165	65		35
YIELD	115	120	10	125	20	25	40	70	130	10	35	

[(1)] PH: plant height at harvest; EIH: ear insertion height; EW: ear weight; NR: number of grain rows per ear; EL: ear length; ED: ear diameter; CW: cob weight; CD: cob diameter; WHG: weight of hundred grains; NGE: number of grains per ear; GL: grain length; YIELD: grain yield.

Table 4. Continued...

Trait[(1)]	PH	EIH	EW	NR	EL	ED	CW	CD	WHG	NGE	GL	YIELD
Double hybrid DKB747 (2008/2009) above the diagonal and single hybrid 30F53 (2009/2010) below the diagonal												
PH		45	95	130	105	125	95	155	125	125	125	95
EIH	50		100	130	125	100	125	145	100	120	100	100
EW	105	110		105	15	25	15	70	65	10	90	10
NR	140	110	80		155	75	120	80	110	100	125	105
EL	120	100	10	120		70	20	95	85	25	125	20
ED	110	110	15	70	40		40	50	90	40	55	25
CW	100	100	10	90	10	20		55	60	40	130	25
CD	130	130	70	75	65	35	40		125	95	155	80
WHG	140	110	130	160	130	180	110	120		125	115	65
NGE	105	110	10	60	15	20	20	95	170		90	10
GL	125	100	35	115	60	20	50	175	190	40		75
YIELD	105	110	10	80	10	15	10	90	130	10	35	
Triple hybrid DKB566 (2009/2010) above the diagonal and double hybrid DKB747 (2009/2010) below the diagonal												
PH		40	135	150	140	160	135	130	135	140	160	135
EIH	30		135	165	145	165	135	140	140	165	165	135
EW	110	115		95	30	15	20	65	90	10	35	10
NR	135	130	115		145	85	100	100	140	90	125	95
EL	115	115	25	125		75	20	95	110	45	100	30
ED	120	120	15	90	65		40	45	105	25	15	15
CW	110	110	20	115	20	40		55	100	40	80	25
CD	110	105	40	95	65	30	30		95	80	145	70
WHG	115	120	85	135	85	95	65	95		135	130	90
NGE	135	115	10	90	45	30	60	65	135		40	10
GL	130	130	70	135	120	40	120	160	125	70		35
YIELD	110	110	10	115	35	20	35	50	90	10	60	

[(1)] PH: plant height at harvest; EIH: ear insertion height; EW: ear weight; NR: number of grain rows per ear; EL: ear length; ED: ear diameter; CW: cob weight; CD: cob diameter; WHG: weight of hundred grains; NGE: number of grains per ear; GL: grain length; YIELD: grain yield.

Table 5. Sample size (number of plants) for estimation of the Pearson correlation coefficient of 66 pairs of traits measured in the single hybrid P32R21, the triple hybrid DKB566 and the double hybrid DKB747 in the 2008/2009 growing season and in the single hybrid 30F53, the triple hybrid DKB566 and the double hybrid DKB747 in the 2009/2010 growing season for the range of the 95% confidence interval of 0.45

Trait[(1)]	PH	EIH	EW	NR	EL	ED	CW	CD	WHG	NGE	GL	YIELD
Single hybrid P32R21 (2008/2009) above the diagonal and triple hybrid DKB566 (2008/2009) below the diagonal												
PH		40	75	70	80	70	65	85	95	75	70	75
EIH	20		80	80	80	85	75	95	90	85	75	80
EW	80	70		40	20	10	40	60	50	10	25	10
NR	75	70	75		60	30	75	55	90	40	60	40
EL	70	70	15	100		40	25	95	80	20	50	20
ED	90	85	15	50	35		50	50	60	20	15	10
CW	70	70	20	75	15	40		50	70	50	80	45
CD	90	110	45	55	60	30	40		75	70	100	60
WHG	75	80	75	75	90	75	40	60		80	90	50
NGE	65	65	10	45	25	30	30	55	90		30	10
GL	80	75	25	70	50	15	50	65	90	35		25
YIELD	80	75	10	75	15	15	20	45	75	10	25	
Double hybrid DKB747 (2008/2009) above the diagonal and single hybrid 30F53 (2009/2010) below the diagonal												
PH		30	70	85	65	70	70	95	70	70	75	70
EIH	30		70	85	80	70	80	95	65	75	65	60
EW	65	65		70	10	15	10	50	40	10	55	10
NR	85	70	50		90	50	75	55	70	60	75	65

[(1)] PH: plant height at harvest; EIH: ear insertion height; EW: ear weight; NR: number of grain rows per ear; EL: ear length; ED: ear diameter; CW: cob weight; CD: cob diameter; WHG: weight of hundred grains; NGE: number of grains per ear; GL: grain length; YIELD: grain yield.

Table 5. Continued...

Trait[1]	PH	EIH	EW	NR	EL	ED	CW	CD	WHG	NGE	GL	YIELD
EL	75	65	10	70		45	15	60	50	20	80	10
ED	75	70	10	45	20		25	30	55	25	35	20
CW	65	65	10	55	10	15		30	40	25	70	15
CD	70	80	35	55	45	25	25		65	60	100	55
WHG	85	65	70	95	90	105	65	80		80	80	45
NGE	65	70	10	40	10	15	10	55	100		55	10
GL	80	70	25	65	40	15	30	95	120	30		45
YIELD	65	65	10	50	10	10	10	35	70	10	25	
Triple hybrid DKB566 (2009/2010) above the diagonal and double hybrid DKB747 (2009/2010) below the diagonal												
PH		30	85	85	85	85	85	75	80	85	95	85
EIH	20		85	90	80	100	80	80	80	95	100	85
EW	75	75		60	15	10	10	45	60	10	25	10
NR	85	75	65		90	50	70	65	85	50	70	70
EL	75	75	15	80		50	15	60	65	30	65	20
ED	75	75	15	55	40		20	35	70	15	15	10
CW	70	75	15	80	10	20		35	55	25	45	15
CD	75	75	30	55	45	20	20		65	50	90	45
WHG	65	70	50	85	50	60	45	60		85	75	60
NGE	80	75	10	55	30	20	40	40	85		30	10
GL	85	75	40	85	70	25	75	100	75	40		25
YIELD	75	75	10	65	20	15	25	35	55	10	40	

[1] PH: plant height at harvest; EIH: ear insertion height; EW: ear weight; NR: number of grain rows per ear; EL: ear length; ED: ear diameter; CW: cob weight; CD: cob diameter; WHG: weight of hundred grains; NGE: number of grains per ear; GL: grain length; YIELD: grain yield.

these coefficients with $ACI_{95\%}$ of 0.30. According to Shieh (2010), the use of larger sample sizes reduced the bias and the root mean square error, associated with the estimates of r. The author found that higher root mean square errors are associated with small magnitude correlation coefficients and lower root mean square errors are associated with high magnitude correlation coefficients, either positive or negative.

In this study, regardless of the hybrid, the crop and the pair of traits, 375, 195 and 120 plants were sufficient, respectively, for the estimation of r with maximum $ACI_{95\%}$ of 0.25, 0.35 and 0.45 (Tables 3, 4 and 5). Thus, with an experiment with five treatments and four repetitions (20 plots in total) evaluating ten plants per plot (200 plants in total), r can be estimated with maximum $ACI_{95\%}$ of 0.35, provided that the effects of treatments and local control are suppressed. If, however, six plants per plot are evaluated (total 120 plants), it can be estimated the r of each pair of traits with maximum $ACI_{95\%}$ of 0.45, provided that the effects of treatments and local control are also suppressed. If a researcher seeks to estimate the r of 66 pairs of traits with maximum $ACI_{95\%}$ of 0.45 within each treatment, using four replications, the researcher must evaluate 30 plants per replication (120 plants per treatment), provided that the effect of local control is removed before the estimation of r.

4. CONCLUSION

The sample size varies among different hybrids, crops and pairs of traits. Larger sample size is needed to estimate the correlation coefficient between weakly correlated traits and smaller sample size is needed to estimate the correlation coefficient between highly correlated traits.

Independently of hybrid, crop and pairs of traits, 375, 195 and 120 plants are sufficient, respectively, to estimate the correlation coefficients with maximum $ACI_{95\%}$ of 0.25, 0.35 and 0.45.

ACKNOWLEDGEMENTS

To the National Council for Scientific and Technological Development (CNPq) and the Coordination for the Improvement of Higher Education Personnel (CAPES) for scholarships. To fellows and volunteers for their help in conducting the experiment and collecting data.

REFERENCES

Bonett, D. G., & Wright, T. A. (2000). Sample size requirements for estimating pearson, kendall and spearman correlations. Psychometrika, 65, 23-28. http://dx.doi.org/10.1007/BF02294183.

Cargnelutti, A. Fo., Toebe, M., Burin, C., Silveira, T. R., & Casarotto, G. (2010). Tamanho de amostra para estimação do coeficiente de correlação linear de Pearson entre caracteres de milho. Pesquisa Agropecuaria Brasileira, 45, 1363-1371. Retrieved from http://www.scielo.br/pdf/pab/v45n12/v45n12a05.pdf

Cargnelutti, A. Fo., Lopes, S. J., Toebe, M., Silveira, T. R., & Schwantes, I. A. (2011). Tamanho de amostra para estimação do coeficiente de correlação de Pearson entre caracteres de *Crambe abyssinica.* Revista Ciência Agronômica, 42, 149-158. http://dx.doi.org/10.1590/S1806-66902011000100019.

Cargnelutti, A. Fo., Lopes, S. J., Brum, B., Toebe, M., Silveira, T. R., & Casarotto, G. (2012). Tamanho de amostra para a estimação do coeficiente de correlação linear de Pearson entre caracteres de mamoneira. Semina. Ciências Agrárias, 33, 953-962. http://dx.doi.org/10.5433/1679-0359.2012v33n3p953.

Cruz, C. D., & Regazzi, A. J. (1997). Modelos biométricos aplicados ao melhoramento genético (2. ed.). Viçosa: UFV. 390 p.

Duvick, D. N. (2005). The contribution of breeding to yield advances in maize (*Zea mays* L.). Advances in Agronomy, 86, 83-145. http://dx.doi.org/10.1016/S0065-2113(05)86002-X.

Fancelli, A. L., & Dourado, D. No. (2004). Produção de milho (2. ed.). Guaíba: Agropecuária. 360 p.

Food and Agriculture Organization of the United Nations – FAO. (2014). Retrieved from http://statistics.amis-outlook.org/data/index.html#DOWNLOAD

Ferreira, D.F. (2009). Estatística básica. 2. ed. (664p). Lavras: UFLA.

R Development Core Team. (2014). R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Shieh, G. (2010). Estimation of the simple correlation coefficient. Behavior Research Methods, 42, 906-917. http://dx.doi.org/10.3758/BRM.42.4.906. PMid:21139158