An Acad Bras Cienc (2022) 94(1): e20200001 DOI 10.1590/0001-3765202220200001
Anais da Academia Brasileira de Ciências | *Annals of the Brazilian Academy of Sciences*
Printed ISSN 0001-3765 I Online ISSN 1678-2690
www.scielo.br/aabc | www.fb.com/aabcjournal

CELLULAR AND MOLECULAR BIOLOGY

A SAS code to estimate phenotypic-genotypic covariance and correlation matrices based on expected value of statistical designs to use in plant breeding

MEHDI RAHIMI & MATEO V. HERNANDEZ

Abstract: Phenotypic-genotypic covariance and correlation have been useful in crop and animal breeding programs. In the study of diversity of natural populations and different cultivars of plants that are examined based on statistical design, estimation of genotypic-phenotypic covariance through expected value of statistical designs mean square is hard and time-consuming when the number of studied traits is high. Moreover, the lack of a program in this field and manual calculations make the estimation more complicated. Therefore, in this study, one program was developed in SAS that can be used to calculate the genotypic-phenotypic covariance matrix through the first part of the program based on the expected value of applied statistical designs mean square. Then, based on the covariance matrix computed from the previous design model, their correlation matrix was calculated using the second part of the program based on the interactive matrix language (IML) of SAS. The phenotypic-genotypic covariance matrices of the 12 studied traits of rice are calculated based on this code. This program could compute phenotypic-genotypic covariance and correlation matrices based on the expected value of any statistical designs.

Key words: Computer program, covariance matrix, estimation, genetic correlation, plant breeding, statistical design.

INTRODUCTION

Diversity in plant species is extremely important in plant breeding because it provides the basis for effective selection of cultivars. The overall diversity within a population (phenotypic diversity) is due to the effects of genotype and environment (Govindaraj et al. 2015). Phenotypic changes are the visible variation in a trait within a population. This variation consists of two elements of environmental and genotypic variation and therefore its value differs in different environmental conditions. On the other hand, genotypic variation is related to genotypic difference between individuals within a population and is a major objective in plant breeding (Hermisson & Wagner 2004, Lewontin 2008).

Phenotypic variance (Vp) or observed variance is composed partly of genetic or heritable (Vg) and partly of non-heritable (Ve) variation. The ratio of the total genotypic variation to total phenotypic or observed variation is termed as coefficient of heritability in broad sense (Hill et al. 1998, Mather & Jinks 1977).

Phenotypic correlations between two traits can be influenced by inheritance, environment, or both. When the correlation is mainly genetic, genetic advancement is of more significance in

breeding programs. Genetic correlations indicate the amount of covariance of two similar genes or strong linkage in two different traits and the correlation of the environment is due to this fact that an environment can cause different variances in the both traits (Sultan 2000, Wolf et al. 1998).

One of the most important aims in plant breeding is to increase the yield per unit area. Finding suitable indices considering relationship between yield and important agricultural traits can play a significant role in selection plans for improving yield (Sölkner et al. 2008). In plant breeding, correlation between traits is also important because it measures amount and type of genetic and non-genetic relationship between two or more traits. Genotypic and phenotypic correlations between different traits may help the breeder in indirect selection for important traits through less complex traits that can be easily measured (Crossa et al. 2014, Stinchcombe et al. 2012).

In study of chilli (*Capsicum* spp.) based on morphological traits were used genetic and phenotypic correlations and by using these correlations, traits affecting chilli yield were identified through path analysis (Deepo et al. 2020). Also, in a study on seventy seven rice genotypes, phenotypic and genotypic correlations, genetic parameters and coefficients of genotypic and phenotypic variation were estimated by expected value of mean square of sources of variation for the traits and used to identify important traits (Parimala et al. 2020). In other study on wheat, genetic and phenotypic correlations were used to identify traits affecting yield, and through these correlations an effective step was taken to improve wheat yield (Kumari et al. 2020).

Perhaps the most important activity in all plant breeding programs is selection. Selection plans such as mass selection, progeny selection and recurrent selection are considered according to crop pollination method, gene action type and breeding purpose. The selecting action takes place in both pure and segregated populations (Acquaah 2009, Moreno-Gonzalez & Cubero 1993). Selection efficiency depends largely on the genetic diversity of the population and inheritance of the studied trait. The variation can be obtained from estimated variance components of a sample from total variance (Hallauer 2007). To achieve this purpose, one of the methods is to use evaluation of different traits of individuals or different genotypes based on repeated statistical designs and estimation of phenotypic, genotypic and environmental variance-covariance matrices through expected value of desired statistical designs. The phenotypic, genotypic and environmental correlation matrices are estimated through the above matrices (Roff 1997, Zeng et al. 1999). Many studies have shown that plant breeders have used phenotypic-genotypic variance-covariance and correlation for direct and indirect improvement of traits in different plants (Akhtar et al. 2011, Malik et al. 2005, Munir et al. 2007, Seyoum et al. 2012, Tripathi et al. 2011).

So far, no simple program has been available to estimate these matrices through the expected value of design. Therefore, the aim of this research was to develop a SAS program for estimating phenotypic, genotypic and environmental variance-covariance and correlation matrices through expected value of desired statistical designs.

MATERIALS AND METHODS

Formulas for combined analysis based on randomized complete block design (RCBD)

There are different designs to estimate phenotypic and genotypic covariance based on expected value of statistical designs such as completely randomized designs (CRD), randomized complete block design (RCBD), and split-plot designs in one or several

environments. Here, estimation of phenotypic and genotypic covariance is explained based on combined analysis for randomized complete block design and its formulas. However, based on the expected value of other designs, this covariance can also be calculated. In order to estimate phenotypic and genotypic variance of one trait, expected value of combined analysis was used according to Table I and the following relationships.

$$\sigma^2_{ge}(X) = \frac{MSge - MSe}{r} \quad (1)$$

$$\sigma^2_{g}(X) = \frac{MSg - MSge}{re} \quad (2)$$

$$\sigma^2_{p}(X) = \sigma^2_{g}(X) + \sigma^2_{ge}(X) + \sigma^2_{e} \quad (3)$$

W where σ^2_p is the phenotypic variance (Vp), σ^2_g, genotypic variance (Vg), σ^2_{ge}, genotype × environment interaction variance (Vge) and σ^2_e, environmental variance (Ve).

Moreover, phenotypic and genotypic covariance of two traits was calculated according to Table II and the following relationships based on the expected value of combined covariance analysis.

$$\sigma(xy)_{ge} = \frac{MPge - MPe}{r} \quad (4)$$

Table I. The combined variance analysis table and expected values of sources of variation for the trait x.

S.O.V.	DF.	SS_x	MS_x	$E(MS_x)$
Env	e-1	SSEnv.	MSEnv.	
Rep(Env)	e(r-1)	SSr	MSr	
Trt	t-1	SSg	MSg	
Trt×Env	(t-1)×(e-1)	SSge	MSge	
Error	e×(t-1)×(r-1)	SSe	MSe	

S.O.V.= sources of variation, DF.= degrees of freedom, SS_x=sum of squares of trait x, MS_x= mean squares of trait x, $E(MS_x)$= Expected value of MS, Env= environment, Rep=replication, Trt= treatment, e= number of environment, r= number of replication, t= number of treatment.

$$\sigma(xy)_{g} = \frac{MPg - MPge}{re} \quad (5)$$

$$\sigma(xy)_{p} = \sigma(xy)_{g} + \sigma(xy)_{ge} + \sigma(xy)_{e} \quad (6)$$

where $\sigma(xy)_p$ is the phenotypic covariance (COVp), $\sigma(xy)_g$, genotypic covariance (COVg), $\sigma(xy)_{ge}$, genotype × environment interaction covariance (COVge) and $\sigma(xy)_e$, environmental covariance (COVe). Combined variance analysis was performed for all the traits. If the effects of the treatment and treatment×environment interaction for all of them are significant, the traits are used to estimate phenotypic, genotypic and environmental variance-covariance matrices through expected value of the proposed design.

Development of a SAS code for phenotypic-genotypic covariance and correlations matrices

Here, we reported the development of a new SAS macro which computes phenotypic and genetic covariance as well as correlation matrices for several traits based on combined analysis (Supplementary Material-Table SI). Although this program is written for combined analysis of variance, it can be used for any statistical designs with some changes in the program. As

Table II. The combined covariance analysis table and expected value of sources of variation for the trait x and y.

S.O.V.	DF.	SP_{xy}	MP_{xy}	$E(MP_{xy})$
Env	e-1	SPEnv	MPEnv	
Rep(Env)	e(r-1)	SPr	MPr	
Trt	t-1	SPg	MPg	
Trt×Env	(t-1)×(e-1)	SPge	MPge	
Error	e×(t-1)×(r-1)	SPe	MPe	

S.O.V.= sources of variation, DF.= degrees of freedom, SP^{xy}= sum of the products for the trait x and y, $M_{Px}y$= mean of the products for the trait x and y, $E(MP_{xy})$=Expected value of MP, Env= environment, Rep=replication, Trt= treatment, e= number of environment, r= number of replication, t= number of treatment.

an example, this macro has been done based on a randomized complete block design (Table SII) and is presented with combined analysis of variance SAS macro. Thus, researchers, by comparing the program), could be able to modify this SAS macro based on their desired statistical designs (Table SI and SII).

General features of the program: an example

In this study, the data of 12 measured traits of 30 rice lines were used which were performed in a randomized complete block design with three replications in two separate experiments, i.e., normal and drought stress conditions (Table SIII). Users can bring data in CVS Excel format like sample data (Figure 1, Table SIII). General linear models were used for analyzing experimental design. In the INFILE section, path and name of data must be specified and changed based on user data (Figure 1). In the INPUT statement of the program, the variables namely ENV, REP, TRT and X1-Xn were internal to the program and showed the environment, replication, treatment, and number of traits (from one to n), respectively (Figure 1). Data input can be changed based on desired statistical designs and number of traits. In the phenotypic covariance and correlation matrices section, it should be specified that the number of traits for Var and Format statement such as Var x1-x12. Moreover, the 'Proc export' must specify the path for saving phenotypic covariance matrix and phenotypic correlations matrix (Figure 2).

The genotypic covariance and correlation matrices section is used to estimate the genotypic covariance and correlation matrices, whose class and model statement must be specified based on the desired statistical designs for proc GLM (Figure 3). In the Data DoF, the degrees of freedom are determined for the sources of variation based on type of statistical designs (Figure 3). In this section, some sources of variation should be added or decrease based on type of statistical designs used (Figure 3).

In the macro calculation section, the drop column should be changed based on the number of traits (Figure 4). In the next section, the true variance of the sources of variation is calculated based on the statistical designs used. In next part of this section, these variances need to be modified according to the desired statistical designs (Figure 4). After that in the IML section, the Read all var {} part must be changed according to the number of traits. Moreover, the TraitNames and Format parts should be changed according to number of traits. Finally, a path should be specified in the proc export to save the genetic covariance matrix and genetic correlations matrix.

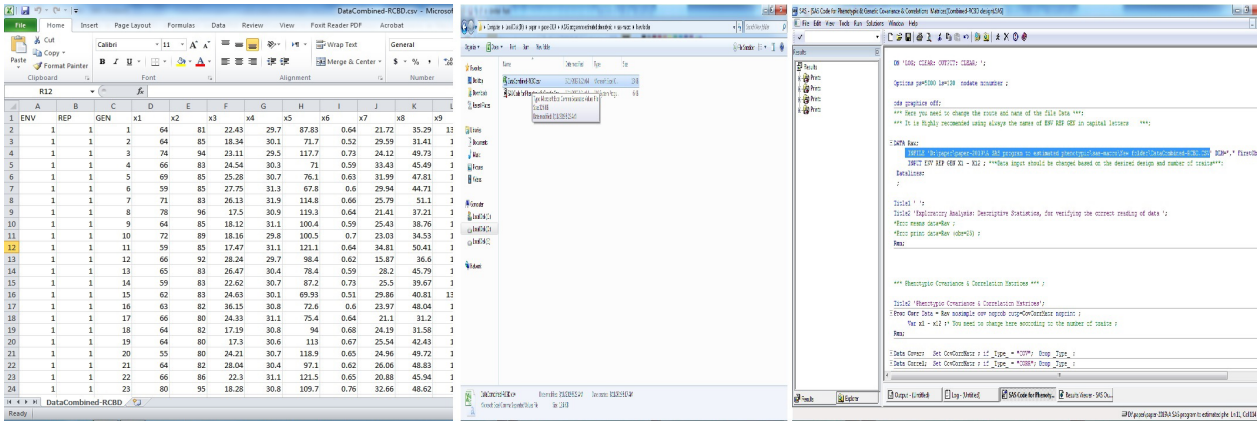


Figure 1. The prepared and saved data for use in the program.

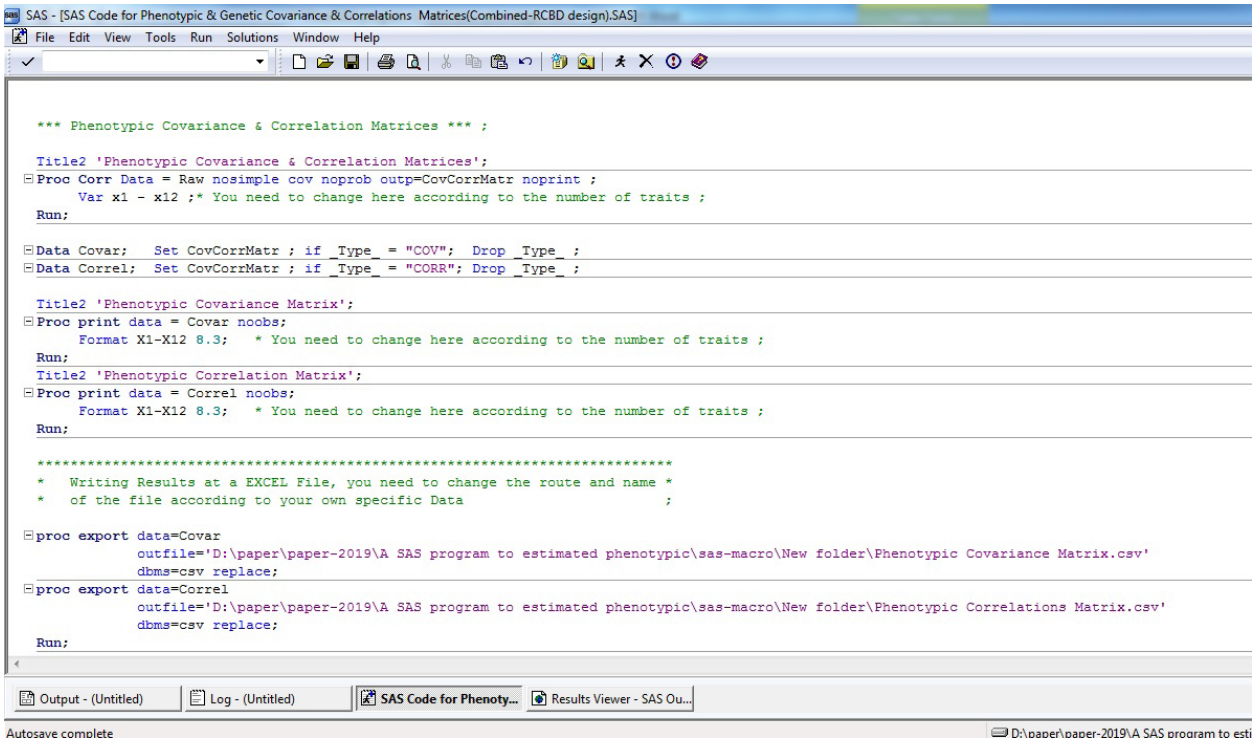


Figure 2. The phenotypic covariance and correlations matrix section.

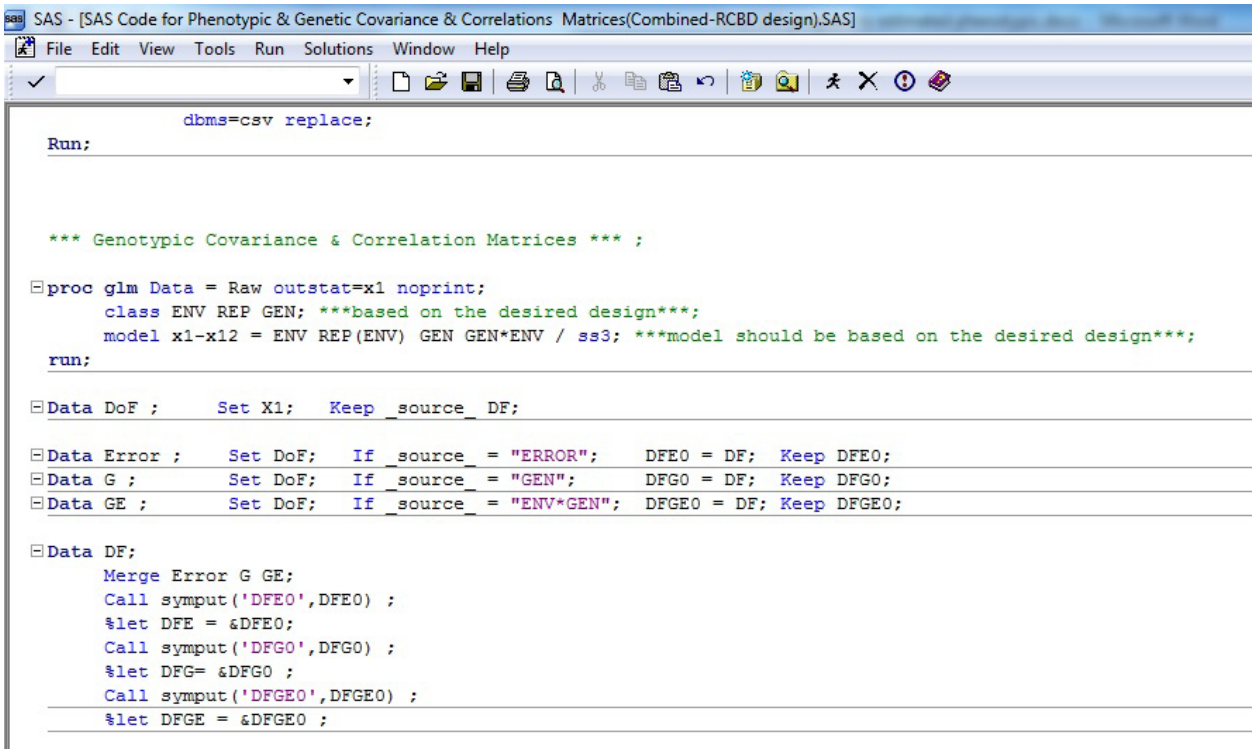


Figure 3. The model section of used statistical design.

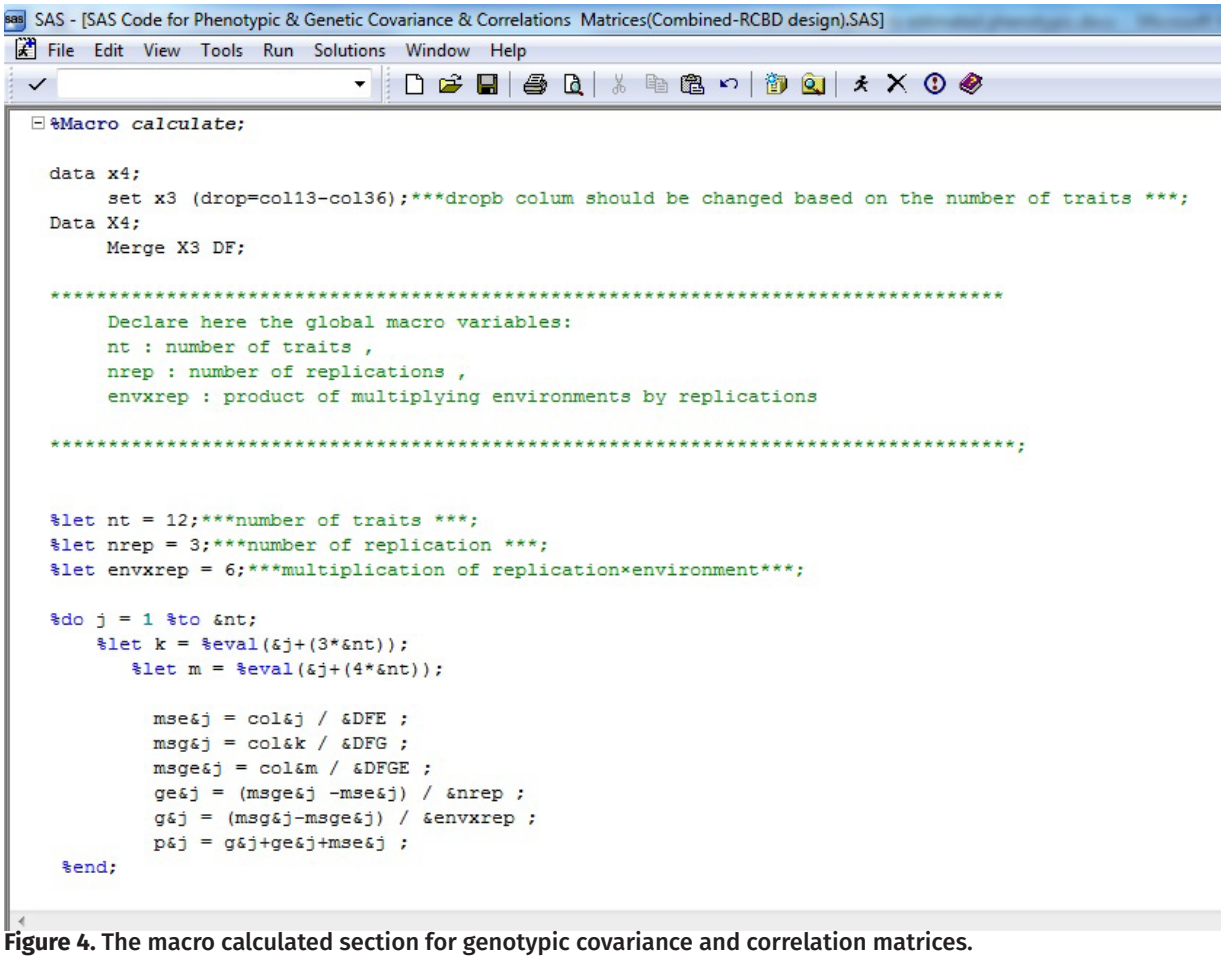


Figure 4. The macro calculated section for genotypic covariance and correlation matrices.

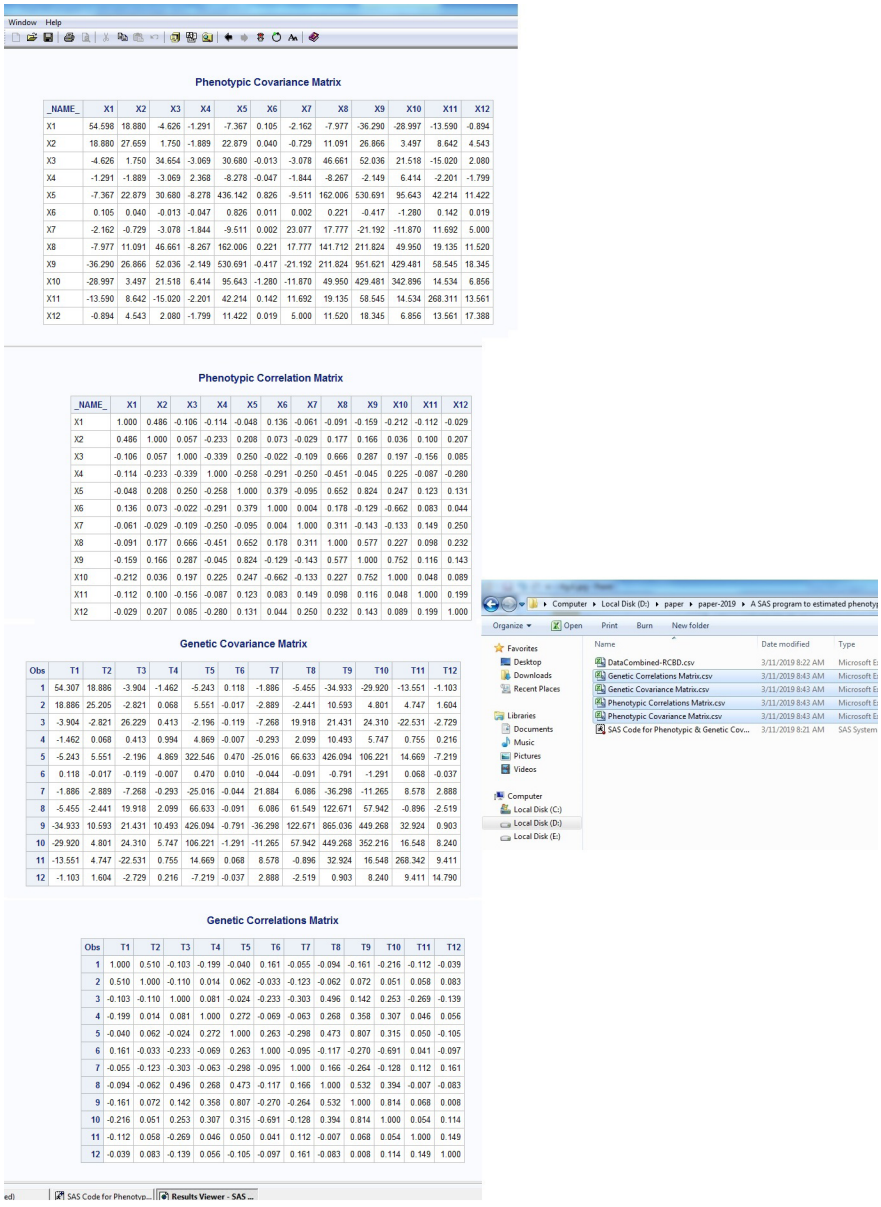
RESULTS

The SAS macro is shown for estimating variance-covariance matrix for 12 traits based on combined analysis. This recommendation can be changed for any number of traits as well as for any experimental design. This program stores the phenotypic and genotypic covariance and correlations matrices based on desired statistical designs and store it in a CVS Excel format for any number of traits measured in the path given to the program. The results are also shown in the result viewer or output section of the SAS program (Figure 5 and Table SIV to SVII). Researchers can use the information stored in Excel format for their breeding program. This program as well as data and output files are included in the supplemental data.

In first section of Figure 5, the phenotypic covariance matrices of the 12 studied traits are shown, and the same information is shown in Table SIV. In the next section of Figure 5, the correlation matrix of the 12 studied traits is shown and in Table SV, the phenotypic correlation matrix of 12 traits is stored in Excel format. Also, in the following sections of Figure 5, the genotypic covariance matrix and then the genotypic correlation matrix of the traits are shown. The genotypic covariance matrix and genotypic correlation matrix traits are stored in Table SVI and SVII in Excel format, respectively.

DISCUSSION

The phenotypic and genotypic correlation matrices are shown in Table SV and SVII,



Phenotypic Covariance Matrix

NAME	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	54.598	18.880	-4.626	-1.291	-7.367	0.105	-2.162	-7.977	-36.290	-28.997	-13.590	-0.894
X2	18.880	27.659	1.750	-1.889	22.879	0.040	-0.729	11.091	26.866	3.497	8.642	4.543
X3	-4.626	1.750	34.654	-3.069	30.680	-0.013	-3.078	46.661	52.036	21.518	-15.020	2.080
X4	-1.291	-1.889	-3.069	2.368	-8.278	-0.047	-1.844	-8.267	-2.149	6.414	-2.201	-1.799
X5	-7.367	22.879	30.680	-8.278	436.142	0.826	-9.511	162.006	530.691	95.643	42.214	11.422
X6	0.105	0.040	-0.013	-0.047	0.826	0.011	0.002	0.221	-0.417	-1.280	0.142	0.019
X7	-2.162	-0.729	-3.078	-1.844	-9.511	0.002	23.077	17.777	-21.192	-11.870	11.692	5.000
X8	-7.977	11.091	46.661	-8.267	162.006	0.221	17.777	141.712	211.824	49.950	19.135	11.520
X9	-36.290	26.866	52.036	-2.149	530.691	-0.417	-21.192	211.824	951.621	429.481	58.545	18.345
X10	-28.997	3.497	21.518	6.414	95.643	-1.280	-11.870	49.950	429.481	342.896	14.534	6.856
X11	-13.590	8.642	-15.020	-2.201	42.214	0.142	11.692	19.135	58.545	14.534	268.311	13.561
X12	-0.894	4.543	2.080	-1.799	11.422	0.019	5.000	11.520	18.345	6.856	13.561	17.388

Phenotypic Correlation Matrix

NAME	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12
X1	1.000	0.486	-0.106	-0.114	-0.048	0.136	-0.061	-0.091	-0.159	-0.212	-0.112	-0.029
X2	0.486	1.000	0.057	-0.233	0.208	0.073	-0.029	0.177	0.166	0.036	0.100	0.207
X3	-0.106	0.057	1.000	-0.339	0.250	-0.022	-0.109	0.666	0.287	0.197	-0.156	0.085
X4	-0.114	-0.233	-0.339	1.000	-0.258	-0.291	-0.250	-0.451	-0.045	0.225	-0.087	-0.280
X5	-0.048	0.208	0.250	-0.258	1.000	0.379	-0.095	0.652	0.824	0.247	0.123	0.131
X6	0.136	0.073	-0.022	-0.291	0.379	1.000	0.004	0.178	-0.129	-0.662	0.083	0.044
X7	-0.061	-0.029	-0.109	-0.250	-0.095	0.004	1.000	0.311	-0.143	-0.133	0.149	0.250
X8	-0.091	0.177	0.666	-0.451	0.652	0.178	0.311	1.000	0.577	0.227	0.098	0.232
X9	-0.159	0.166	0.287	-0.045	0.824	-0.129	-0.143	0.577	1.000	0.752	0.116	0.143
X10	-0.212	0.036	0.197	0.225	0.247	-0.662	-0.133	0.227	0.752	1.000	0.048	0.089
X11	-0.112	0.100	-0.156	-0.087	0.123	0.083	0.149	0.098	0.116	0.048	1.000	0.199
X12	-0.029	0.207	0.085	-0.280	0.131	0.044	0.250	0.232	0.143	0.089	0.199	1.000

Genetic Covariance Matrix

Obs	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
1	54.307	18.886	-3.904	-1.462	-5.243	0.118	-1.886	-5.455	-34.933	-29.920	-13.551	-1.103
2	18.886	25.205	-2.821	0.068	5.551	-0.017	-2.889	-2.441	10.593	4.801	4.747	1.604
3	-3.904	-2.821	26.229	0.413	-2.196	-0.119	-7.268	19.918	21.431	24.310	-22.531	-2.729
4	-1.462	0.068	0.413	0.994	4.869	-0.007	-0.293	2.099	10.493	5.747	0.755	0.216
5	-5.243	5.551	-2.196	4.869	322.546	0.470	-25.016	66.633	426.094	106.221	14.669	-7.219
6	0.118	-0.017	-0.119	-0.007	0.470	0.010	-0.044	-0.091	-0.791	-1.291	0.068	-0.037
7	-1.886	-2.889	-7.268	-0.293	-25.016	-0.044	21.884	6.086	-36.298	-11.265	8.578	2.888
8	-5.455	-2.441	19.918	2.099	66.633	-0.091	6.086	61.549	122.671	57.942	-0.896	-2.519
9	-34.933	10.593	21.431	10.493	426.094	-0.791	-36.298	122.671	865.036	449.268	32.924	0.903
10	-29.920	4.801	24.310	5.747	106.221	-1.291	-11.265	57.942	449.268	352.216	16.548	8.240
11	-13.551	4.747	-22.531	0.755	14.669	0.068	8.578	-0.896	32.924	16.548	268.342	9.411
12	-1.103	1.604	-2.729	0.216	-7.219	-0.037	2.888	-2.519	0.903	8.240	9.411	14.790

Genetic Correlations Matrix

Obs	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12
1	1.000	0.510	-0.103	-0.199	-0.040	0.161	-0.055	-0.094	-0.161	-0.216	-0.112	-0.039
2	0.510	1.000	-0.110	0.014	0.062	-0.033	-0.123	-0.062	0.072	0.051	0.058	0.083
3	-0.103	-0.110	1.000	0.081	-0.024	-0.233	-0.303	0.496	0.142	0.253	-0.269	-0.139
4	-0.199	0.014	0.081	1.000	0.272	-0.069	-0.063	0.268	0.358	0.307	0.046	0.056
5	-0.040	0.062	-0.024	0.272	1.000	0.263	-0.298	0.473	0.807	0.315	0.050	-0.105
6	0.161	-0.033	-0.233	-0.069	0.263	1.000	-0.095	-0.117	-0.270	-0.691	0.041	-0.097
7	-0.055	-0.123	-0.303	-0.063	-0.298	-0.095	1.000	0.166	-0.264	-0.128	0.112	0.161
8	-0.094	-0.062	0.496	0.268	0.473	-0.117	0.166	1.000	0.532	0.394	-0.007	-0.083
9	-0.161	0.072	0.142	0.358	0.807	-0.270	-0.264	0.532	1.000	0.814	0.068	0.008
10	-0.216	0.051	0.253	0.307	0.315	-0.691	-0.128	0.394	0.814	1.000	0.054	0.114
11	-0.112	0.058	-0.269	0.046	0.050	0.041	0.112	-0.007	0.068	0.054	1.000	0.149
12	-0.039	0.083	-0.139	0.056	-0.105	-0.097	0.161	-0.083	0.008	0.114	0.149	1.000

Figure 5. The output results.

respectively. These correlations can be used in correlational studies as well as path analysis (Anilkumar et al. 2019, Jhanavi et al. 2019, Mishra & Nandi 2018, Nirmal Raj & Gokulakrishnan 2018, Shivakumar et al. 2018) to identify important traits and used them in breeding programs. Also, the phenotypic and genotypic covariance matrices are shown in Table SIV and SVI, respectively that the variances are located in the diameter and covariance are placed outside the diameter. These variances can be used selection index (Almeida et al. 2019, Ghosh et al. 2018, Kour et al. 2018) studies as well as heritability of traits (Banik et al. 2018, Kumar et al. 2019, Raval et al. 2018).

In plant breeding programs, selection of traits based on genetic correlations is more beneficial because genotypic variance is passed on to the next generation. Heritability was also calculated based on the genotypic / phenotypic variance ratio. Traits that have higher heritability are easier to select. Evaluation of variability

components and inheritance of traits help plant breeders to improve crop. Breeders can use the knowledge of genetic variability available among and between crops as a guide to improving crop. Genetic and phenotypic correlations for plant breeding have been used in many researches in recent years. The SAS program for calculating genetic or phenotypic variance-covariance and genetic or phenotypic correlation can be a useful aid to plant and animal breeders and it will prevent mistakes in manual calculation.

CONCLUSION

The SAS program reported here was easy to use and the outputs were easy to understand and user-friendly. This program could compute phenotypic-genotypic covariance and correlation matrices based on the expected value of any statistical designs. The goodness and attraction of this program is that it doesn't need to know the language of the SAS program and the users can easily analyze data with this program. The program is not computationally intensive and should therefore run-on slower computers. Users are advised against making any changes to the program code based on your need and your statistical design.

Acknowledgments

The authors gratefully thank Dr. José Crossa for his comments and suggestions on this paper.

REFERENCES

ACQUAAH G. 2009. Principles of plant genetics and breeding. USA: J Wiley & Sons, 756 p.

AKHTAR N, NAZIR M, RABNAWAZ A, MAHMOOD T, SAFDAR M, ASIF M & REHMAN A. 2011. Estimation of heritability, correlation and path coefficient analysis in fine grain rice (*Oryza sativa* L.). J Anim Plant Sci 21: 660-664.

ALMEIDA GQD, SILVA JDO, RESENDE MDVD, MENEGUCI JLP & MATOS GR. 2019. Selection index via REML/BLUP for identifying superior banana genotypes in the central region of Goiás state, Brazil. Rev Ceres 66: 26-33.

ANILKUMAR G, UMESHA K, SHIVAPRIYA M, HALESH G, MARUTHIPRASAD B & DARSHAN G. 2019. Character association and path analysis for yield traits in coriander (*Coriandrum sativum* L.). Elec J Plant Breed 10: 224-229.

BANIK M, DEORE G, MANDAL AK & MHASE L. 2018. Genetic Variability and Heritability Studies in Chickpea (*Cicer arietinum* L.). Curr J Appl SciTech: 1-6.

CROSSA J, PEREZ P, HICKEY J, BURGUEÑO J, ORNELLA L, CERÓN-ROJAS J, ZHANG X, DREISIGACKER S, BABU R & LI Y. 2014. Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity 112: 48-60.

DEEPO DM, SARKER A, AKTER S, ISLAM MM, HASAN M & ZEBA N. 2020. Diversity and path analysis of chilli (*Capsicum* spp.) based on morphological traits in northern region of Bangladesh. Turkish J Agric Food SciTech 8: 179-185.

GHOSH T, ISLAM S, SHAHANAZ S, BISWAS S & TAREQ M. 2018. Genetic variability and selection index evaluation of some selected tomato lines for their yield and yield components. Bangladesh J Environ Sci 34: 73-78.

GOVINDARAJ M, VETRIVENTHAN M & SRINIVASAN M. 2015. Importance of genetic diversity assessment in crop plants and its recent advances: an overview of its analytical perspectives. Genet Res Int 2015: Article ID 431487, 14 p.

HALLAUER AR. 2007. History, contribution, and future of quantitative genetics in plant breeding: lessons from maize. Crop Sci 47: 4-19.

HERMISSON J & WAGNER GP. 2004. The population genetic theory of hidden variation and genetic robustness. Genet 168: 2271-2284.

HILL J, BECKER HC & TIGERSTEDT PM. 1998. Quantitative and ecological aspects of plant breeding. Netherlands: Springer Science & Business Media, 275 p.

JHANAVI D, PATIL H, RANJITHA B & JUSTIN P. 2019. Correlation and path analysis studies for growth, yield and quality traits in French bean (*Phaseolus vulgaris* L.). Environ Ecol 37: 22-26.

KOUR S, PRADHAN U, PATEL J & VAISHNAV P. 2018. Comparative study of selection indices based on different weights in forage sorghum [*Sorghum bicolor* (L.) Moench]. J Crop Weed 14: 17-23.

KUMAR N, PANDEY S, MISHRA S, MISHRA D & PANDEY V. 2019. Studies on heritability and genetic advance for the quantitative characters in Pea (*Pisum sativum* L.) in sodic condition. J Pharmacogn Phytochem 8: 310-312.

KUMARI P, DE N & KUMARI AKA. 2020. Genetic variability, correlation and path coefficient analysis for yield and quality traits in wheat (*Triticum aestivum* L.). Int J Curr Microbiol App Sci 9: 826-832.

LEWONTIN R. 2008. The genotype/phenotype distinction. Stanford Encyclopedia of Philosophy.

MALIK H, MALIK SI, HUSSAIN M, CHUGHTAI SR & JAVED HI. 2005. Genetic correlation among various quantitative characters in maize (*Zea mays* L.) hybrids. J Agric Soc Sci 3: 262-265.

MATHER K & JINKS JL. 1977. Introduction to biometrical genetics. Boston, USA: Springer Science & Business Media, 231 p.

MISHRA A & NANDI A. 2018. Correlation and path coefficient analysis for quality traits in tomato (*Solanum lycopersicon* L.). J Pharmacogn Phytochem 7: 1733-1738.

MORENO-GONZALEZ J & CUBERO J. 1993. Selection strategies and choice of breeding methods. In: Hayward MD, Bosemark NO, Romagosa I & Cerezo C (Eds), Plant Breeding: Principles and prospects. Netherlands: Springer Science+Business Media Dordrecht, p. 281-313.

MUNIR M, CHOWDHRY M & MALIK T. 2007. Correlation studies among yield and its components in bread wheat under drought conditions. Int J Agr Biol 9: 287-290.

NIRMAL RAJ R & GOKULAKRISHNAN J. 2018. Indirect selection for various yield attributing characters of maize hybrids across environments using correlation and path analysis. J Pharmacogn Phytochem 7: 1810-1812.

PARIMALA K, RAJU CS, PRASAD AH, KUMAR SS & REDDY SN. 2020. Studies on genetic parameters, correlation and path analysis in rice (*Oryza sativa* L.). J Pharmacogn Phytochem 9: 414-417.

RAVAL V, PATEL A, RATHOD S, SUMITA Z, JM V & CHAUDHARI B. 2018. Genetic variability, heritability and genetic advance studies in okra (*Abelmoschus esculentus* (L.) Moench). Int J Chem Stud 6: 3319-3321.

ROFF DA. 1997. Evolutionary quantitative genetics. USA: Springer Science & Business Media, 494 p.

SEYOUM M, ALAMEREW S & BANTTE K. 2012. Genetic variability, heritability, correlation coefficient and path analysis for yield and yield related traits in upland rice (*Oryza sativa* L.). J Plant Sci 7: 13-22.

SHIVAKUMAR M, RADHIKA K, SARLA N & VENKANNA V. 2018. Character association and path analysis for yield and its component characters in rice. Environ Ecol 36: 876-880.

SÖLKNER J, GRAUSGRUBER H, OKEYO AM, RUCKENBAUER P & WURZINGER M. 2008. Breeding objectives and the relative importance of traits in plant and animal breeding: a comparative review. Euphytica 161: 273-282.

STINCHCOMBE JR, KIRKPATRICK M & GROUP F-VTW. 2012. Genetics and evolution of function-valued traits: understanding environmentally responsive phenotypes. Trends Ecol Evol 27: 637-647.

SULTAN SE. 2000. Phenotypic plasticity for plant development, function and life history. Trends Plant Sci 5: 537-542.

TRIPATHI S, MARKER S, PANDEY P, JAISWAL K & TIWARI D. 2011. Relationship between some morphological and physiological traits with grain yield in bread wheat (*Triticum aestivum* L. em. Thell.). Trends Appl Sci Res 6: 1037-1045.

WOLF JB, BRODIE III ED, CHEVERUD JM, MOORE AJ & WADE MJ. 1998. Evolutionary consequences of indirect genetic effects. Trends Ecol Evol 13: 64-69.

ZENG Z-B, KAO C-H & BASTEN CJ. 1999. Estimating the genetic architecture of quantitative traits. Genet Res 74: 279-289.

SUPPLEMENTARY MATERIAL

Tables SI - SVII

How to cite

RAHIMI M & HERNANDEZ MV. 2022. A SAS code to estimate phenotypic-genotypic covariance and correlation matrices based on expected value of statistical designs to use in plant breeding. An Acad Bras Cienc 94: e20200001. DOI 10.1590/0001-3765202220200001.

Manuscript received on January 01, 2020;
accepted for publication on March 20, 2020

MEHDI RAHIMI[1]
https://orcid.org/0000-0001-5625-3275

MATEO V. HERNANDEZ[2]
https://orcid.org/0000-0002-0735-3242

[1]Graduate University of Advanced Technology, Institute of Science and High Technology and Environmental Sciences, Department of Biotechnology, Kerman Province, End of Haft Bagh-e-Alavi Highway Knowledge Paradise, Kerman, Iran

[2]Universidad Autónoma Chapingo, Department of Suelos, Carr. Federal México-Texcoco, Km 38.5, 56230 Texcoco, México

Correspondence to: **Mehdi Rahimi**
E-mail: mehdi83ra@yahoo.com

Author contributions

Rahimi M has designed and written the program and wrote the article. Hernandez MV developed and edited the program and also read and revised the manuscript. All authors have read and approved the manuscript.