Artificial neural networks based on segmented model for adaptability and stability evaluation of soybean genotypes

Wagner Faria Barbosa[1], Thais do Prado Hashimoto[2], Éder Matsuo[3], Ana Carolina Campana Nascimento[1], Camila Ferreira Azevedo[1], André Ricardo Gomes Bezerra[4], Moysés Nascimento[1*]

[1]Universidade Federal de Viçosa - Campus Viçosa, Departamento de Estatística, 36570-900, Viçosa, MG, Brasil
[2]Universidade Federal de Viçosa - Campus Viçosa, Departamento de Biologia Geral, 36570-900, Viçosa, MG, Brasil
[3]Universidade Federal de Viçosa - Campus Rio Paranaíba, Instituto de Ciências Exatas e Tecnológicas, 38810-000, Rio Paranaíba, MG, Brasil
[4]Limagrain Brasil S.A., Jataí, GO, Brasil

*Corresponding author: moysesnascim@ufv.br

Abstract

Unlike models based on simple linear regressions, segmented models can better assess the adaptability and stability of genotypes, demonstrating a nonlinear response pattern over environmental variation. However, these methods can be under statistical limitations, such as the Type Error II increase and biased estimates. Therefore, this work aimed to transpose the concepts of adaptability and stability from the statistical analysis of a segmented model to the discriminatory potential of an artificial neural network (ANN) and use it to classify soybean (*Glycine max* (L.) Merr.) genotypes. An ANN training was carried out with the grain yield of 7,200 soybean genotypes simulated in 15 different environments. The ANN topology chosen was the one that had less than 1% of error in the testing phase with 1,800 simulated genotypes. A total of 9,000 simulated soybean genotypes were previously arranged in 18 different classes, which represented the combination of nine classes of adaptability by the Verma and collaborators (VCM) method and two classes of stability (invariability concept) by the Finlay & Wilkinson (FW) method. Finally, the grain production of ten real soybean genotypes was inputted into the ANN-trained model, and the classification regarding adaptability and stability was obtained. There was 90% agreement between the ANN and VCM analyses regarding the adaptability classification and 20% regarding stability. With the methods presented in this work, it was demonstrated that the potential of using ANNs to assess the adaptability of genotypes is strong. In addition, since stability was introduced in the ANN as a different concept from that used to classify the genotypes by the statistical method, such classification needs to be reviewed and further improved.

Keywords: *Glycine max*; artificial intelligence, genotypes × environments interaction; data simulation, bioinformatics.
Abbreviations: ANN_artificial neural networks; MET_multi-environmental trials; VCM_Verma et al. (1978); FW_Finlay and Wilkinson (1963); MSE_Mean square Error; ER_Eberhart and Russell (1966); NID_Normally Independently Distributed.

Introduction

The use of methods based on machine learning in agronomy, such as artificial neural networks (ANN), is increasing (Sousa et al., 2022). Briefly, ANN are models that work as a network of biological neurons capable of processing a large amount of data using self-learning (Haykin, 2009). Compared with a statistical framework, ANN have the advantage of not requiring *prior* assumptions about the model, which allows their adjustment to the most diversified problems (Rosado et al., 2022).

Therefore, it is unsurprising that ANN models have been used in breeding programs to predict the genetic values of animals and plants (Abdollahi-Arpanahi et al., 2020; Rosado et al., 2020). Still, in the context of plant breeding, the use of ANN models has also become an interesting approach to dealing with the interaction between genotypes and environments in multi-environmental trials (MET) (Alves et al., 2019; Nascimento et al., 2013). For example, Nascimento et al. (2013) proposed a methodology of adaptability and phenotypic stability based on the training of an ANN,

considering the methodology of Eberhart and Russell (ER). The authors chose the ER method, which is widely used because of its simplicity and efficiency in analyzing MET (Janick, 2003).

Although attractive, the ER is based on the fit of simple linear regression models (Cruz et al., 2012). Therefore, it does not allow to study the potential nonlinear pattern of genotype responses throughout the environmental variation (Nascimento et al., 2020) and makes ER-based ANN models equally deficient in assessing the adaptability and stability of genotypes. On the other hand, segmented regression models like the one proposed by Verma et al. (1978) (VCM) can distinctly evaluate the performance of genotypes in unfavorable and favorable environments. This allows such model to find the "ideal" genotype, which should present low sensitivity to adverse conditions and increasing yield as the environment improves, besides high stability.

As mentioned in previous works, the use of ANN models to assess adaptability and stability is preferred to avoid the

statistical limitations of the simpler methods, such as biases in the estimates of regression coefficients and the increase of Type Error II (Nascimento et al., 2013; Teodoro et al., 2015). Thus, by joining the improvement aspects of the VCM method with an ANN approach, a new method can be established to remove the limitations described above. In addition, as proposed by Nascimento et al. (2013), the ANN approach can also include an adapted method based on the Finlay and Wilkinson (1963) method (FW) to assess stability, which is based on the invariant classification of a given genotype after data linearization.

Therefore, because of the issues raised above, this study (1) proposed the development of an ANN-based segmented model to evaluate the adaptability and stability of genotypes and (2) the application of such model for the classification of soybean (Glycine max (L.) Merr.) genotypes using real data. Finally, the soybean genotypes' classification regarding their adaptability and stability could be compared with that from regular statistical analysis, i.e., that upon which the neural network was based.

Results

Analysis of variance

The analysis of variance indicated that the soybean genotypes presented distinguished performances in the face of different environmental conditions, which is attested by the significant interaction (P ≤ 0.01) between genotypes and environments. The existence of genetic variability between genotypes for grain production was also verified (Table 2).

Neural network construction

Since the genotypes should be analyzed in unfavorable and favorable environments, the variance analyses were also performed considering these two conditions to implement the proposal. In these separate analyses, the MSE for unfavorable and favorable environments were 50,141.65 (degree of freedom (df) = 108) and 43,471.95 (df = 162). These results were used to implement the normally distributed random deviations into the grain production and test the angular coefficients of the estimated ER models of the simulated genotypes (these analyses are not shown since they are not required to interpret the results but to provide estimates for data simulation).

With the simulated data set, ANN models were trained and tested. A model with 15 neurons in the single hidden layer and 0.94% errors in the testing phase, which converged after 528 iterations, was selected, and subsequently used to classify the real soybean genotypes regarding adaptability and stability.

Phenotypic adaptability and stability analysis

Out of the ten soybean genotypes, none was considered "ideal" among those that presented productivity above the overall mean of the environments (4,074.78 $\mathrm{kg\ ha^{-1}}$) (Table 3). The best scenarios were found for the genotypes CZ 26B36 IPRO, DM 66I68 RSF IPRO, and ST 644 IPRO, which were classified as having general adaptability and high stability/invariability by the ANN output. In addition, the M 6210 IPRO genotype was considered exclusively responsive to favorable environments by the VCM method and the ANN (Table 3).

The ANN showed 90% agreement with the VCM method to discriminate the adaptability of soybean genotypes, however, only 20% agreement on stability (Table 3).

Discussion

The significance observed between the variance component related to genotype, and environmental effect indicates differential genotypic responses to environmental effects. The edaphoclimatic conditions of each environment (Figure 1), situated in different biomes and present differences in latitude, longitude, and climatic effects, such as rainfall and temperature, can be used to corroborate this result. Under such a situation, studies of adaptability and stability become necessary to detail the behavior of each genotype within the different environments evaluated, as Cruz et al. (2012) pointed out.

The use of methods based on Computational Intelligence, such as artificial neural networks (ANN), is increasing in agronomy (González-Camacho et al., 2016; Sousa et al., 2020; Kujawa et al., 2021, Costa et al., 2022). Specifically, the first proposal to study the phenotypic adaptability and stability used Eberhart and Russell (ER) methodology as prior information to construct an ANN and recommend genotypes (Nascimento et al., 2013). This approach was successfully used to evaluate the phenotypic adaptability and stability in cowpea genotypes (Teodoro et al., 2015), soybean (Alves et al., 2019; Oda et al., 2022), and cotton (Carvalho et al., 2015) avoids statistical limitations of the simpler methods, such as biases in the estimates of regression coefficients and the increase of Type Error II provides by small sample sizes (Teodoro et al., 2015). However, different from the ANN proposed by Nascimento et al. (2013), the proposed methodology in this manuscript allows to study the pattern nonlinear of genotype responses to environmental variation, that is, to find the "ideal" genotype defined by Verma et al. (1978). Considering the soybean data set, none evaluated considered genotype presented low sensitivity to adverse conditions and increasing yield as the environment improves, that is, the behavior of an "ideal" genotype. Specifically, three of the five genotypes with productivity above the overall mean (CZ 26B36 IPRO, DM 66I68 RSF IPRO, M 6210 IPRO, M 6410 IPRO, ST 644 IPRO) were classified as general adaptability. This result is expected since the evaluated genotype set is composed of released cultivars recommended for several environments.

It is interesting to highlight that although sophisticated methodologies, for example, GGE Biplot (Yan et al., 2000), AMMI (Gauch, 2006), and Centroid methods (Nascimento et al., 2009a, b) are presented in the literature, none of them can find the "ideal" genotype.

The low agreement between the results obtained by ANN and the VCM method in terms of the stability parameter can be explained by the difference in the concept of this parameter used in each of these approaches, one based on invariance and the other based on regression deviations. Precisely, the VCM method mirrors the concept of stability from its auxiliary method, i.e., the ER method, which is applied separately in unfavorable and favorable environments. Its stability concept is based on the predictability of genotype behavior (simplified in Table 1). On the other hand, this stability concept differs from that used by ANN, which was based on the invariance of genotype behavior after linearization of the data, adapting the FW method. However, a comparison with the FW method is not feasible as it is not a bi-segmented regression. The presented results showed that the proposed methodology that uses Computational intelligence with a segment regression (Verma et al., 1978) can recommend

Table 1. Adaptability and stability classes of genotypes based on, respectively, the arrangement of the β_{1i} and $\sigma_{d_i}^2$ parametric values according to the Verma et al. (1978) method.

Adaptability Classes	Parametric values of β_{1i}		Adaptability[1]
	Unfavorable	Favorable	
1	$\beta_{1i} = 1$	$\beta_{1i} = 1$	Overall
2	$\beta_{1i} < 1$	$\beta_{1i} < 1$	Specific for unfavorable environments
3	$\beta_{1i} > 1$	$\beta_{1i} > 1$	Specific for favorable environments
4	$\beta_{1i} = 1$	$\beta_{1i} < 1$	Not recommended
5	$\beta_{1i} < 1$	$\beta_{1i} > 1$	Ideal
6	$\beta_{1i} > 1$	$\beta_{1i} = 1$	Not recommended
7	$\beta_{1i} = 1$	$\beta_{1i} > 1$	Specific for favorable environments
8	$\beta_{1i} < 1$	$\beta_{1i} = 1$	Specific for unfavorable environments
9	$\beta_{1i} > 1$	$\beta_{1i} < 1$	Not recommended
Stability Classes	Parametric values of $\sigma_{d_i}^2$		Stability[1]
	Unfavorable	Favorable	
1	$\sigma_{d_i}^2 = 0$	$\sigma_{d_i}^2 = 0$	High
2	$\sigma_{d_i}^2 = 0$	$\sigma_{d_i}^2 > 0$	Low
3	$\sigma_{d_i}^2 > 0$	$\sigma_{d_i}^2 = 0$	Low
4	$\sigma_{d_i}^2 > 0$	$\sigma_{d_i}^2 > 0$	Low

[1]Simplified adaptability and stability classes of the genotypes throughout both unfavorable and favorable environments.
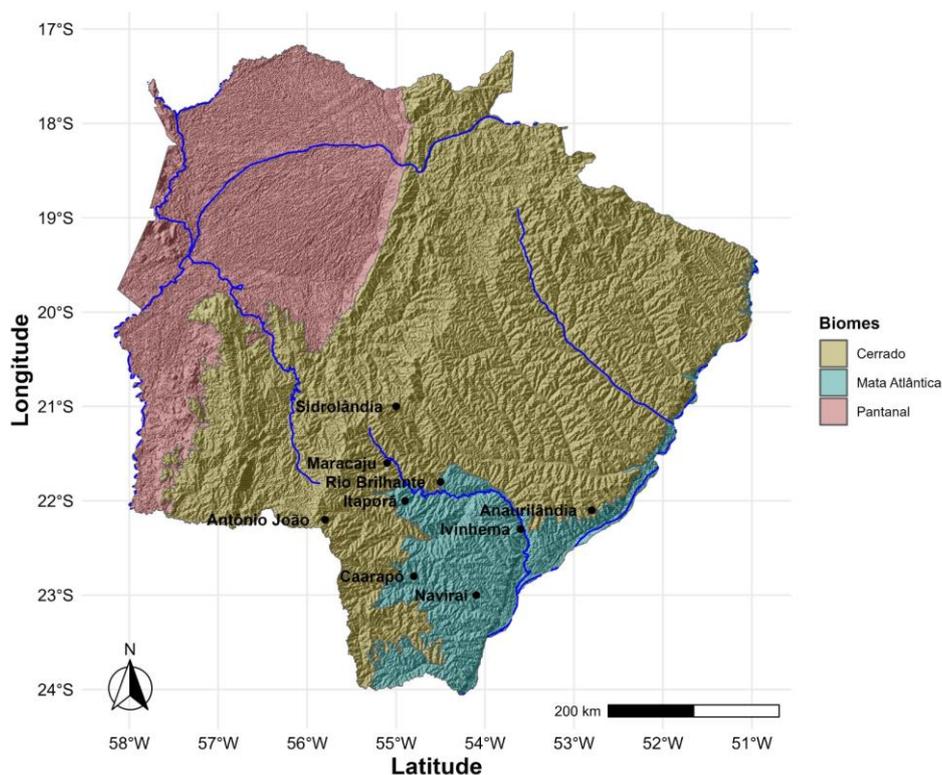

Figure 1. Location of the state of Mato Grosso do Sul in South America, with its respective biomes (Pantanal, Cerrado, and Mata Atlântica).

Table 2. Variance analysis of the real soybean genotypes' grain yield (kg ha^{-1}) (*G. max*).

FV	GL	SQ	QM	F
Blocks/Environments	30	4252146.55	141738.22	
Genotypes	9	8535633.04	948403.67	6.56**
Environments	14	90127512.44	6437679.46	45.42**
Genotypes × Environments	126	18211847.33	144538.47	3.13**
Error	270	12457753.65	46139.83	
TOTAL	449	133584893.00		

**Significant at 1% probability by F test. The coefficient of variation was 5.27 %.

Table 3. Mean grain yield and classification regarding adaptability and stability of 10 soybean genotypes (*G. max*) evaluated in 15 environments in the State of Mato Grosso do Sul by the methods of Verma et al. (1978) and artificial neural networks (ANN).

Genotype	Grain yield (kg ha^{-1})[1]	Verma, Chahal & Murty		ANN	
		Adaptabilty[2]	Stability[3]	Adaptabilty[2]	Stability[4]
		(Unf. | Fav.)	(Both)	(Unf. | Fav.)	(Both)
AS 3680 IPRO	3993.02	(=1 | =1)	High	(=1 | =1)	High
BRASMAX GARRA IPRO	3942.01	(>1 | =1)	Low	(>1 | =1)	High
BMX POTÊNCIA RR	3919.56	(=1 | =1)	Low	(=1 | =1)	High
BS 2606 IPRO	3958.68	(<1 | =1)	Low	(=1 | =1)	Low
CZ 26B36 IPRO	4075.17	(=1 | =1)	Low	(=1 | =1)	High
DM 66I68 RSF IPRO	4301.00	(=1 | =1)	Low	(=1 | =1)	High
M 6210 IPRO	4152.15	(=1 | >1)	Low	(=1 | >1)	High
M 6410 IPRO	4269.23	(=1 | <1)	Low	(=1 | <1)	High
ST 644 IPRO	4196.45	(=1 | =1)	Low	(=1 | =1)	High
TEC 7022 IPRO	3940.54	(=1 | =1)	Low	(=1 | =1)	High
Overall mean	4074.78				

[1]Average of the genotype above the overall mean of the experiment is in bold. [2]Adaptability based on the method of Verma, Chahal & Murty (1978) in terms of β_{1i} values, respectively, for unfavorable (left) and favorable (right) environments. [3]Stability based on reclassification for both environments according to Table 1. [4]Stability based on invariability after data linearization (adapted from Finlay & Wilkinson method).

genotypes considering potential nonlinear patterns of genotype responses to environmental variation. Additionally, it is interesting to highlight that using neural networks to assess phenotypic adaptability and stability allows simulating genotypes based on different methodologies. In this way, it is possible to create networks that classify genotypes based on different concepts according to the researcher's interest.

Material and methods

Firstly, it is worth mentioning that to establish the ANN based on VCM and FW methods as proposed in this work; real data must be present, followed by statistical analysis in which some estimated parameters were used to simulate the training and testing data used by the ANN models.

Experimental data
The data used here were obtained from experiments carried out by the Phytotechnics sector of the Fundação MS para Pesquisa e Difusão de Tecnologias Agropecuárias with ten soybean cultivars named AS 3680 IPRO, BRASMAX GARRA IPRO (63i64RSF IPRO), BMX POTÊNCIA RR, BS 2606 IPRO, CZ 26B36 IPRO, DM 66I68 RSF IPRO, M 6210 IPRO, M 6410 IPRO, ST 644 IPRO, and TEC 7022 IPRO. The cultivars were planted in the crop year 2020/2021 in 15 different experimental areas distributed among nine municipalities (Anaurilândia (22°08'S; 52°45'W; 370 msnm), Antônio João (22°10'S; 55°46'W; 630 msnm), Caarapó (22°45'S; 54°47'W; 390 msnm), Itaporã (22°03'S; 54°55'W; 400 msnm), Ivinhema (22°20'S; 53°39'W; 370 msnm), Maracaju (21°38'S; 55°06'W; 360 msnm), Navirai (22°59'S; 54°06'W; 370 msnm), Rio Brilhante (21°50'S; 54°32'W; 310 msnm) and Sidrolândia (21°00'S; 54°59'W; 450 msnm)) of the State of Mato Grosso do Sul, Brazil. The experiments were carried out under randomized blocks (3 replicates), and the experimental unit consisted of four 5.0 m long rows spaced 0.5 m from each other. The useful area of each plot was 4.0 m^2, with the two central rows being harvested, discounting 0.50 m of the border at the ends to obtain the production per plot. With this result, it was obtained the yield productivity in kg ha^{-1}.

Analyses of variance
The soybean yield (kg ha^{-1}) data were submitted to a joint analysis of variance in the software Genes (Cruz, 2013). The model adopted for the analysis was $Y_{ijk} = \mu + R/E_{k(j)} + G_i + E_j + GE_{ij} + \xi_{ijk}$, where Y_{ijk} is the phenotypic mean; μ is the general mean; $R/E_{k(j)}$ is the effect of the kth repetition (block) in the jth environment; G_i is the fixed effect of the ith genotype; E_j is the effect of the jth environment Normally Independently Distributed (NID); GE_{ij} is the effect of the interaction of the ith genotype in the jth environment NID with mean equal to 0 and variance denoted by σ_{ge}^2; and ξ_{ijk} is the experimental error NID with mean equal to 0 and variance denoted by σ_e^2. In addition, variance analyses considering the same model as above-mentioned were carried out with the data, which were previously split into according to unfavorable and favorable environments (more details are given below).

Segmented model for adaptability evaluation
The VCM method characterizes the adaptability and stability of genotypes by the interpretation of the angular coefficients and the regression deviation of two simple linear regressions, which are estimated for each genotype in unfavorable and favorable environments. Then, initially, it is necessary to recognize the unfavorable and favorable environments through the environmental index (I_j) given by $I_j = \frac{1}{g}\sum_i^g Y_{ij} - \frac{1}{ga}\sum_i^g \sum_j^a Y_{ij}$. Once the environments are identified, the response of each genotype to environmental variations can be analyzed within two distinct groups of environments (unfavorable and favorable) through a simple regression model, which, in this case, is based on the ER method. The statistical model considered by these authors is defined by $Y_{ij} = \beta_{0i} + \beta_{1i}I_j + \psi_{ij}$; where Y_{ij} is the mean of the ith genotype in the jth environment; β_{0i} is the regression coefficient that measures the response of the ith genotype throughout the environments; I_j is the environmental index; and ψ_{ij} is the random effect, which is decomposed as follows $\psi_{ij} = \delta_{ij} + \bar{\varepsilon}_{ij}$, where δ_{ij} is the regression deviation and $\bar{\varepsilon}_{ij}$ is the mean experimental error. The estimators of the adaptability and stability parameters are defined, respectively, by $\hat{\beta}_{1i} = \frac{\sum_j Y_{ij}}{\sum_j I_j^2}$ and $\hat{\sigma}_{d\,i}^2 = \frac{MSD_i - MSE}{r}$,

where MSD_i is the mean square deviation of the ith genotype, MSE is the mean square of the error and, r is the number of repetitions. In this case, since the ER method is applied in two groups of environments, the j number of environments has size a_u and a_f, denoting, respectively, the number of unfavorable and favorable environments. In addition, thirty-six classes of genotypes are possible to be generated from the arrangement [i.e., (3 × 3) × (2 × 2)] among parametric values of β_{1i} and $\sigma_{d\,i}^2$ (Table 1), which are tested under the respective hypotheses: $H_{0_{(\beta_1)}}$: $\beta_{1i} = 1$ versus $H_{1_{(\beta_1)}}$: $\beta_{1i} \neq 1$ (Student's t-test) and $H_{0_{(\sigma_d^2)}}$: $\sigma_{di}^2 = 0$ versus $H_{1_{(\sigma_d^2)}}$: $\sigma_{di}^2 > 0$ (F-test).

Artificial neural networks based on segmented model

Initially, aiming to expand the data set for the training and testing of the network, yield data were simulated based on the information from the experiment data under study. Therefore, the values of I_j were firstly estimated for each environment, and then, from its sign (positive or negative), two sets of environments were defined, one containing six unfavorable environments (i.e., $a_u = 6$ with negative values) and the other containing nine favorable ones (i.e., $a_f = 9$ with positive values). Posteriorly, 500 vectors containing values of Y_{ij} were simulated by the application of the ER model considering each possible parametric value of β_{1i} (i.e., $\beta_{1i} < 1$, $\beta_{1i} = 1$ and $\beta_{1i} > 1$) for each set of the previously categorized environments; i.e., vectors representing grain yield in unfavorable environments had six values of Y_{ij} simulated from the three possible values of β_{1i} and, vectors representing grain yield in favorable environments had nine values of Y_{ij} simulated from the three possible values of β_{1i}. Finally, the groups of vectors were arranged in nine classes by concatenating the vectors (without changing the order) obtained from each value of β_{1i} for both sets of favorable and favorable environments. This procedure created a set of 4,500 simulated genotypes representing their production behavior throughout 15 environments.

The parametric values used to simulate each value of Y_{ij} according to the model $Y_{ij} = \beta_{0i} + \beta_{1i}I_j + \psi_{ij}$ were $\beta_{0i} = \bar{X}_G$ (general average of the grain yield data of the real soybean genotypes); β_{1i} = a random value generated from a uniform distribution with the parameters a and b (i.e., $U[a; b]$), where $U[0.90; 1.10]$, $U[0.00; 0.89]$ and $U[1.11; 2.00]$ were used to the respective classes $\beta_{1i} = 1$, $\beta_{1i} < 1$ ou $\beta_{1i} > 1$ and, ψ_{ij} = a random value of a Normal distribution $N(0, \hat{\sigma}^2)$ for the ith genotype in the jth environment, where $\hat{\sigma}^2$ was $\hat{\sigma}_u^2$, whether it was the estimated MSE from the variance analysis carried out with only unfavorable environments and, $\hat{\sigma}^2$ was $\hat{\sigma}_f^2$, whether it was the estimated MSE from the variance analysis carried out with only favorable environments. In addition, simulated Y_{ij} values were included in the dataset of the 4,500 genotypes only if, a bilateral t test, under the hypotheses H_0: $\beta_{1i} = 1$ versus H_1: $\beta_{1i} \neq 1$, indicated the belonging of $\hat{\beta}_{1i}$ in the possible three classes of the parameter β_{1i}.

The ANN, used this set of 4,500 simulated soybean genotypes to classify the real genotypes according to the nine classes of adaptability coming from Table 1. However, since simulated values of Y_{ij} identifying the four classes of stability in (Table 1) may overlap each other when considering $\sigma_{d_i}^2 = 0$ and $\sigma_{d_i}^2 > 0$, the phenotypic stability classification by the ANN was carried out via an adapted

concept of stability, which is the invariance of results after data linearization. Thus, the simulated values of Y_{ij} of the 4,500 genotypes were submitted to a logarithmic transformation producing a linearization in which regression deviations are expected to be zero (Nascimento et al., 2013). These linearized Y_{ij} values were added to the set of 4,500 previously simulated genotypes (totaling 9,000 genotypes) so that an adapted concept of invariance of genotype behavior could be inserted, as proposed by Finlay and Wilkinson (1963). Therefore, the ANN could classify a genotype of high stability/invariability if, after linearization, the classification matched the origin class before linearization. In contrast, if the classification was another, the genotype was considered of low stability/invariability (Nascimento et al., 2013).

Since linearization was performed simultaneously in Y_{ij} values representing the total of the 15 environments for each simulated genotype, only two classes (high and low) of adapted stability/invariability could be generated (i.e., not for each set of unfavorable and favorable environments); therefore, the ANN classification output was compared to those two simplified stability classes of Table 1 and 18 classes of genotypes were then generated by data simulation. Finally, from the 9,000 simulated genotypes, 80% (i.e., 400 per class) were randomly chosen and used for the training of ANN models, and the remaining 20% (i.e., 100 per class) were used for the testing phase.

The ANN models were built by single-layer backpropagation neural networks (Hastie et al., 2009), in which the variables Z_i are weighted functions of the input variables X_i. The outputs Y_k are modeled as functions of these combinations. The sigmoid activation function was used in the single-layer, and the "softmax" function was used as the output function. The estimation of network parameters (weights) was performed by minimizing the sum of squares of the errors using the gradient descent algorithm.

For the training of the ANN models with the simulated dataset, auxiliary parameters were also defined as the learning rate ($L = 0.0005$), and the maximum number of iterations ($Iter_{max.} = 5,000$), the initialization interval of weights [-0.0002; 0.0002] and the number of hidden layer neurons that varied from 4 to 15 in each ANN convergence attempt. These auxiliary parameters were used in an algorithm programmed in the R software using the *nnet* function of the *nnet* package (Venables and Ripley, 2002) to find an ANN configuration whose maximum classification error in the testing phase was 1%. After choosing the ANN model with the smallest error, the set of real soybean genotypes was submitted to it for classification in terms of adaptability and stability. The R codes used in this manuscript are available at https://github.com/licaeufv.

Conclusion

An ANN based on a segmented model as the VCM model was powerful for classifying soybean genotypes regarding their adaptability and, possibly, can help breeders interpret data from the behavior of any cultivar in the face of environmental variations considering adapted ANN models for each situation. In addition, since stability was introduced in the ANN as a different concept from that used to classify the genotypes by the (VCM) statistical method, such classification needs to be reviewed and further improved.

Authors' contributions
All authors contributed equally to the manuscript, critically revised it, and approved its final version.

References

Abdollahi-Arpanahi R, Gianola D, Peñagaricano F (2020) Deep learning versus parametric and ensemble methods for genomic prediction of complex phenotypes. Genet Sel Evol. 52: 12.

Alves GF, Nogueira JPG, Machado R, Ferreira S da C, Nascimento M, Matsuo E (2019) Stability of the hypocotyl length of soybean cultivars using neural networks and traditional methods. Cienc Rural. 49 (3): e20180300.

Carvalho LP de, Teodoro PE, Barroso LMA, Farias FJC, Morello C de L, Nascimento M (2018) Artificial neural networks classify cotton genotypes for fiber length. Crop Breed Appl Biot. 18: 200–204.

Costa WG, Valadares CB, Silva Junior AC, Arcanjo ES, Almeida HC, Cruz CD, Carvalho IR, Silva JAG, Azevedo CF, Nascimento M, Nascimento ACC (2022) Genetic diversity and interaction between the maintainers of commercial soybean cultivars using self-organizing maps. Crop Sci. 1: 1-13.

Cruz, CD (2013) Genes: a software package for analysis in experimental statistics and quantitative genetics. Acta Sci-Agron. 35, 271–276.

Cruz CD, Regazzi AJ, Carneiro PC (2012) Métodos biométricos aplicados ao melhoramento genético. 1ed. Editora UFV, Viçosa, Minas Gerais.

Janick J (2003) Plant Breeding Reviews, Volume 24, Part 1: Long-term Selection: Maize, 24th ed. Wiley, New York.

Eberhart SA, Russell WA (1966) Stability parameters for comparing varieties. Crop Sci. 6:36–40

Finlay KW, Wilkinson GN (1963) The analysis of adaptation in a plant-breeding programme. Aust J Agr Res. 14:742–754.

Gauch HG (2006) Statistical analysis of yield trials by AMMI and GGE. Crop Sci. 46:1488–1500.

González-Camacho JM, Crossa J, Pérez-Rodríguez P, Ornella L, Gianola D (2016) Genome-enabled prediction using probabilistic neural network classifiers. BMC Genomics. 17: 208.

Hastie T, Tibshirani R, Friedman JH (2009) The elements of statistical learning: data mining, inference, and prediction, 2nd ed. Springer, New York.

Haykin S (2009) Neural networks and learning machines. 3ed. Prentice Hall. Hoboken, New Jersey.

Janick J (2003) Plant Breeding Reviews, Volume 24, Part 1: Long-term Selection: Maize, 24th ed. Wiley, New York.

Kujawa S, Niedbała G (2021) Artificial Neural Networks in Agriculture. Agriculture.11: 497.

Nascimento M, Cruz CD, Campana, ACM, Tomaz RS, Salgado CC, Ferreira RP (2009a). Alteração no método centroide de avaliação da adaptabilidade genotípica. Pesqui Agropecu Bras. 44: 263-269.

Nascimento M, Ferreira A, Campana ACM, Salgado CC, Cruz CC (2009b) Multiple Centroid Methodology to analyze genotype adaptability. Crop Breed Appl Biot. 9: 8-16.

Nascimento M. Nascimento ACC, Silva FF, Teodoro PE, Azevedo CF, Oliveira TRA de, Amaral-Junior AT, Cruz CD, Farias FJC, Carvalho LP (2020) Bayesian segmented regression model for adaptability and stability evaluation of cotton genotypes. Euphytica. 216: 30.

Nascimento M, Peternelli LA, Cruz CD, Nascimento ACC, Ferreira R de P, Bhering LL, Salgado CC (2013) Artificial neural networks for adaptability and stability evaluation in alfalfa genotypes. Crop Breed Appl Biot. 13: 152–156.

Oda MC, Sediyama T, Cruz CD, Nascimento M, Matsuo E (2022) Adaptability and yield stability of soybean genotypes by mean Eberhart and Russell methods, artificial neural networks and centroid. Agron Sci Biot. 8: 1-13.

Rosado RDS, Cruz CD, Barili LD, Carneiro JE de S, Carneiro PC, Carneiro VQ, Silva, JT da, Nascimento M (2020) Artificial Neural Networks in the Prediction of Genetic Merit to Flowering Traits in Bean Cultivars. Agriculture. 10: 638.

Rosado RDS, Penso GA, Serafini GAD, Magalhães dos Santos CE, Picoli EA de T, Cruz CD, Barreto CAV, Nascimento M, Cecon PR (2022) Artificial neural network as an alternative for peach fruit mass prediction by non-destructive method. Sci Hort. 299: 111014.

Sousa IC de, Nascimento M, Sant'anna I de C, Caixeta ET, Azevedo CF, Cruz, CD, Silva FL da, Alkimim ER, Nascimento, ACC, Serão NVL (2022) Marker effects and heritability estimates using additive-dominance genomic architectures via artificial neural networks in Coffea canephora. PLOS ONE. 17: e0262055.

Sousa IC de, Nascimento M, Silva GN, Nascimento ACC, Cruz CD, Almeida DP de, Pestana, KN, Azevedo CF, Zambolim L, Caixeta ET (2020) Genomic prediction of leaf rust resistance to Arabica coffee using machine learning algorithms. Sci Agr. 78: e20200021

Teodoro PE, Barroso LMA, Nascimento M, Torres FE, Sagrilo E, Santos AD, Ribeiro LP (2015) Redes neurais artificiais para identificar genótipos de feijão-caupi semiprostrado com alta adaptabilidade e estabilidade fenotípicas. Pesqui Agropecu Bras. 50: 1054-1060.

Venables WN, Ripley BD (2002). Modern applied statistics with S-Plus. 4ed. Springer. New York.

Verma MM, Chahal GS, Murty BR (1978) Limitations of conventional regression analysis a proposed modification. Theor Appl Genet. 53: 89–91.

Yan W, Hunt LA, Sheng Q, Szlavniscs Z (2000) Cultivar evaluation and mega-environment investigation based on the GGE Biplot. Crop Sci. 40:597–605.