MODELLING THE IMPACT OF AGROMETEOROLOGICAL VARIABLES ON SOYBEAN YIELD IN THE MATO GROSSO DO SUL: 2000-2019

Lucas Eduardo de Oliveira Aparecido [1*],
Federal Institute of Mato Grosso do Sul (IFMS) - Navirai, Mato Grosso do Sul, Brazil.
* Correspondence author:
lucas-aparecido@outlook.com
https://orcid.org/0000-0002-4561-6760

Guilherme Botega Torsoni[1],
Federal Institute of Mato Grosso do Sul (IFMS) - Navirai, Mato Grosso do Sul, Brazil.
guilherme.torsoni@ifms.edu.br,
https://orcid.org/0000-0001-7178-2191

José Reinaldo da Silva Cabral de Moraes[1],
Federal Institute of Mato Grosso do Sul (IFMS) - Navirai, Mato Grosso do Sul, Brazil.
jose.moraes@ifms.edu.br,
https://orcid.org/0000-0002-8567-4893

Kamila Cunha de Meneses[2];
State University of Sao Paulo (FCAV/UNESP) - Jaboticabal, Sao Paulo, Brazil.
kamila.meneses@unesp.br.
https://orcid.org/0000-0001-9200-5260

João Antonio Lorençone[1];
Federal Institute of Mato Grosso do Sul (IFMS) - Navirai, Mato Grosso do Sul, Brazil.
joao.lorencone@gmail.com
https://orcid.org/0000-0002-1950-4853

Pedro Antonio Lorençone[1];
Federal Institute of Mato Grosso do Sul (IFMS) - Navirai, Mato Grosso do Sul, Brazil.
pedroantonio.lorencone@gmail.com
https://orcid.org/0000-0001-6831-3992

Abstract

The study of the soybean yield variability influenced by the climate contributes to the planning of strategies to mitigate its negative effects. Thus, our aim was to calibrate agrometeorological models for soybean yield forecast and identify the weather variables that most influence soybean yield. This study used historical series of climate and soybean yield data from soybean-producing locations in the Mato Grosso do Sul state, Brazil. The historical climate series was 20 years (2000-2019). The soybean production, yield, and planted area data of the localities were in the period from 2009-2018. Multiple Linear Regression analysis was the statistical tool used for data modeling. The models from the north and central regions forecast of anticipation of 2 months since the final data necessary to apply the model were EXC_{JANc} and P_{JANc}, respectively. The models calibrated for the southern region reported anticipation of one month since the final data necessary to apply the model was EXC_{FEVc}. The calibrated models used to forecast soybean yield as a function of climatic conditions have a high degree of significance ($p < 0.05$), high accuracy and errors lower. The models for the northern and central regions show a prevision of anticipation of 2 months before soybean harvest, a period that is essential for producers to be able to conduct pre- and post-harvest planning. The climate variable with the greatest negative influence ($r = -0.54$) on soybean yield in Mato Grosso do Sul state was water stress in December.

Keywords Crop modeling; Climate; Yield zoning; Spatial error model; *Glycine max* L.

1 Introduction

Worldwide production of soybeans was greater than 347 million tons in the harvest of 2017/18 when 126 million hectares were planted. The United States, Brazil, and Argentina producing 121million tons, 107 million tons, and 57 million tons, respectively (USDA, 2018). Brazil, in this harvest, produced 32.43% of soybeans worldwide, despite the large climate variability that occurs in production regions in the country (Sentelhas et al., 2015). The Mato Grosso do Sul State produced 7.35% of national production (CONAB, 2019).

Climate is one of the principal factors that cause reductions in soybean yield (Sentelhas et al., 2015). Soil water stress is the climatic variable that strongly limits crop yields (Battisti et al., 2017). The condition of soil water is a sensitive indicator of the future yield of grains (Martorano et al., 2009). Bonato et al. (1998) related that variation in meteorological factors in a region where soybeans are being cultivated will cause a reduction in crop growth, development, and production.

Soybeans reach their productive potential under appropriate climatic conditions, provided that no other limiting factors occur (Franke, 2000). Air temperature, solar radiation, soil moisture, and water stress are determinant meteorological factors in the efficiency of plant physiological processes (Bonato et al., 1998; Battisti et al., 2017).

Crop models are the best methods of quantitatively demonstrating the effects of climate on crop disease emergence, soybean yield and quality variation (Aparecido et al., 2018). Climatic factors are the main contributors to the occurrence and proliferation of plant diseases, however, these factors can be simulated from crop modeling (Rolim et al., 2008).

Studies like Fontana et al. (2001), Dourado Neto et al. (2004), and Martorano et al. (2012) showed that modelling is of fundamental importance for crop forecasting. However, studies on the forecast of the effects of climate variables on soybean development and yield in the state of Mato Grosso do Sul are still scarce in the literature.

The study of the soybean yield variability influenced by the climate is complex, however, it contributes to the planning of strategies to mitigate the negative effects caused by the climate in agricultural production. Thus, our aim was to calibrate agrometeorological models for soybean yield forecast and identify the weather variables that most influence soybean yield.

2 Materials and Methods

2.1 Locations and databases

The study used historical series of climate and soybean yield from soybean-producing locations in the Mato Grosso do Sul state, Brazil. The soybean production (number of sacks), yield (sacks ha^{-1}), and planted area (ha) data were obtained from the Association of Producers of Soybean, Corn, and other agricultural grains of the State of Mato Grosso do Sul - APROSOJA (www.aprosoja.com.br) website in the period from 2009-2019. We organized the data of the localities by the North, Center, and South regions of the state of Mato Grosso do Sul to create homogeneous groups based on their peculiarities (Fig. 1).

The daily air temperature (maximum, mean, and minimum, ° C) and daily precipitation (P, mm) data for 2000-2019 were obtained from the database of NASA Prediction of Worldwide Energy Resource (NASA POWER, 2019). Then the agrometeorological data were organized on a monthly scale.

2.2 Potential evapotranspiration and Climatological Water Balance

We calculated the potential evapotranspiration by the Camargo (1991) method, according to Eq. 1.

$$PET = 0.01 \times Q_o \times T_{mean} \times N \tag{1}$$

where Q_o is the extraterrestrial solar irradiance (mm day^{-1}); T_{mean} is the mean air temperature; N is the number of days of the month referred.

We estimated the climatological water balance of the localities studied by the method of Thornthwaite and Mather (1955) (Eqs 2-7). The available soil water capacity of 40 mm was used.

$$\text{if } (P-PET)_i < 0 = \begin{cases} Nac_i = Nac_{i-1} + (P-PET)_i \\ STO_I = AWCe^{\frac{(NAC_i)}{AWC}} \end{cases} \quad (2)$$

$$\text{if } (P-PET)_i \geq 0 = \begin{cases} STO_i = STO_{i-1} + (P-PET)_i \\ NAC_I = AWCe^{\frac{(STO_i)}{AWC}} \end{cases} \quad (3)$$

$$ALT_i = STO_i - STO_{i-1} \quad (4)$$

$$ALT_i = \begin{cases} P + |ALT_i| \text{ , } if\ ALT < 0 \\ PET_i \text{ , } \quad if\ ALT \geq 0 \end{cases} \quad (5)$$

$$DEF = PET - AET \quad (6)$$

$$SUR_i = \begin{cases} 0, & if\ AWC < 0 \\ (P-PET)_i - ALT_i, & if\ AWC = 0 \end{cases} \quad (7)$$

where AWC is available soil water capacity (mm); STO is soil water storage (mm); SUR is water surplus in the soil-plant-atmosphere system (mm); DEF is water deficiency in the soil-plant-atmosphere system (mm); NAC is the sum of rainfall – potential evapotranspiration; P is rainfall (mm); PET is potential evapotranspiration (mm); AET is actual evapotranspiration (mm); ALT is soil water storage of the current month - soil water storage of the preceding month (mm), and *i* is the monthly period.

2.3 Statistical analysis

The temporal variability of soybean production and yield were analyzed by planted area for the three regions that the studied localities were organized. The means of these variables were compared by the Scott-Knott test at the 5% probability level.

Multiple Linear Regression (MLR) analysis was the statistical tool used for data modeling (Eq. 8). The independent variables were the climatic variables: air temperature (ºC), rainfall (mm), potential evapotranspiration (mm), water deficit, and water excess (mm). The dependent variable in the model was soybean yield (sacks ha^{-1}). Innumerous models were generated for each region of Mato Grosso do Sul state (north, central, and south), so the model with the highest accuracy was selected for the regions.

$$Y = CL + aX_1 + bX_2 + cX_3 + dX_4 + eX_5 + \varepsilon \quad (8)$$

where Y is the soybean yield (sacks ha^{-1}) in the localities analyzed; a, b, c, d, and e are the model parameters (weights); X_1, X_2, X_3, X_4, and X_5 are the selected climatic variables, CL is the linear coefficient (constant term) and ε is random error.

The phenology of soybeans is shown in Fig. 2. We considered in the data modeling that soybean planting was in early October and harvesting occurred in late March of the following year according to a literature review. Therefore, the climatic data used in the prediction models were from October (OCT_P), November (NOV_P), and December (DEC_P) (soybean planting year), and January (JAN_C), February (FEB_C) and March (MAR_C) (soybean harvest year).

The estimation method employed was the minimum ordinary square (MOS), which minimizes the sum of the squared errors of the model (Draper and Smith, 1980), through a generalized reduced gradient (GRG_2) optimization system (Lasdon and Waren, 1982).

The assumptions tested to verify the adjustment of the model were: 1) collinearity analysis between explanatory variables (multicollinearity); 2) normality of the errors; and 3) homeostacity of the variables (Gujarati and Porter, 2011).

Pearson's correlation analysis (r) verified multicollinearity between the explanatory variables. Explanatory variables that demonstrated $r \geq 0.7$ were removed from the modeling. Collinearity of explanatory variables is a problem in the models, especially when the analysis of coefficient weights (elasticity or sensitivity) occurs (Gujarati and Porter, 2011). Also, we correlated climate variables with soybean production variables, so that we may identify which climate variables most influenced soybean cultivation in the studied localities. We used the Kolmogorov-Smirnov test to verify the normality of model errors.

After calibration of the models, we analyzed the sensitivity of the models (Gujarati and Porter, 2011). In this elasticity analysis, the angular coefficients (weights) of the independent variables were compared, therefore, the higher the weight of the climate variables, the more these variables influenced soybean production.

The models were calibrated using a routine from “Visual Basic for Applications” (VBA) from MS-Excel 2013. We used the following indices to select the best calibrated model for the regions: 1) Pearson correlation (r); 2) Adjusted coefficient of determination(R^2); 3) Wilmott Concordance (d); 4) Confidence Index (c) from Camargo and Sentelhas (1997); 5) Random error (Ea); 6) Systematic error (Es); 7) Maximum absolute error (ME); 8) Mean squared errors (MSE); 9) Root mean squared error (RMSE); 10) Mean absolute error (MAE); 11) Mean absolute percentage error (MAPE) (Eqs. 9 to 19). The regressions that presented the F test with a 5% probability, we selected these variables to verify a higher degree of confidence in the regressions.

$$r = \frac{\sum_{i=1}^{n}(Yobs_i - \overline{Yobs}) \times (Yest_i - \overline{Yest})}{\sqrt{\sum_{i=1}^{n}(Yobs_i - \overline{Yobs})^2} \times \sqrt{\sum_{i=1}^{n}(Yest_i - \overline{Yest})^2}} \tag{9}$$

$$R^2{}_{adjusted} = \left[1 - \frac{(1-R^2)\times(n-1)}{N-k-1}\right] \tag{10}$$

$$d = 1 - \frac{\sum_{i=1}^{N}(Yobs_1 - Yest_i)^2}{\sum_{i=1}^{N}(|Yest_i - \bar{Y}| + |Yobs_i - \bar{Y}|)} \tag{11}$$

$$c = r \cdot d \tag{12}$$

$$Ea = \sqrt{\frac{\sum_{i=1}^{N}(Yest_i - \bar{Y})^2}{N}} \tag{13}$$

$$Es = \sqrt{\frac{\sum_{i=1}^{N}(Yobs_i - \bar{Y})^2}{N}} \tag{14}$$

$$ME = \max(|Yobs_i - Yest_i|)_{i=1}^{n} \tag{15}$$

$$MSE = \frac{\sum_{i=1}^{N}(Yobs_i - Yest_1)^2}{N} \tag{16}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(Yobs_i - Yest_i)^2}{N}} \tag{17}$$

$$MAE = \frac{\sum_{i=1}^{N}|Yobs_i - Yest_i|}{N} \tag{18}$$

$$MAPE(\%) = \frac{\sum_{i=1}^{n}\left(\left|\frac{Yest_i - Yobs_i}{Yobs_i}\right| \times 100\right)}{N} \tag{19}$$

where $Yest_i$: interpolated variable; $Yobs_i$: observed variable; N: number of data, and k: number of independent variables in the regression.

We adopted for the performance interpretation of the confidence index of Camargo and Sentelhas (1997): > 0.85 = “Excellent”; 0.76 to 0.85 = “Very good”; 0.66 to 0.75 = “Good”; 0.61 to 0.65= “Average”; 0.51 to 0.60 = “Insatisfactory”; 0.41 to 0.50 = “Bad” and < 0.40 = “Terrible”.

3 Results and Discussion

There was high temporal variability in the agrometeorological elements studied in soybean-producing regions in the state of Mato Grosso do Sul (Fig. 3). The highest mean air temperatures(${}_{air}T$) occurred in October, December, and February in the northern, central, and southern regions of Mato Grosso do Sul (MS) state, respectively. In the northern region of the state occurred the highest ${}_{air}T$, with 27ºC. While, the lowest mean air temperatures of the regions occurred between June and July, where the lowest mean air temperature was in the southern region of MS of 19ºC. Also, the southern region of the state presented a high variation of mean air temperature between the regions. These results are within adequate air temperatures for soybean cultivation in MS. Similar results were found by Alvares et al. (2015).

The annual water deficit (WD) is more intense in the north of the State of Mato Grosso do Sul between May to October, with 140 mm y^{-1} (Fig. 3). In the south region, the WD was the lowest and occurred between August and September with an accumulated value of 29.76 mm y^{-1}. The WD of the central region occurred between July and August with an accumulated value of 39.71 mm y^{-1}. Fietz and Urchei (2002) reported similar results for WD when the evaluated the influence of WD on soybean cultivation in Mato Grosso do Sul.

In all regions, there was a significant increase in the production of soybeans from 2009 to 2018. For example, in the southern region, this value increased 267.13% during this period. This region presented an average production of 3,307,257.04 tons, while the central and northern regions had a production of 1,120,381.68 t and 971,569.03 t, respectively. The south of MS is the region with the largest area planted with soybeans, consequently, this region has a greater production. The growth and variation in production, area, and yield of soybeans between 2000 and 2018 are shown in Fig. 4.

The correlation between soybean yield and climatic variables for the State of MS shows distinct relationships (Fig. 5). In general, the largest direct correlations (+) were between water storage in December (ARM_{DEZ}) and real evapotranspiration in December and February (ETR_{DEZ} e ETR_{FEV}), these results showed that crop yield increased as ARM_{DEZ}, ETR_{DEZ}, and ETR_{FEV} increased. Thus plants have greater availability of water to conduct photosynthesis. It is important to emphasize that the variable with the lowest correlation (r = -0.02) with soybean yield was ETP_{NOV} (Fig. 5).

DEF_{DEZ} and DEF_{FEV} were the variables with the greatest negative correlations, with values between 0.54 and 0.41, respectively. Various authors have emphasized the negative influence of DEF in several crops, e.g., Martins et al. (2015) and Valeriano et al. (2018) for coffee crops, and Aparecido et al. (2018) for Annatto (*Bixa orellana* L.). DEF has a negative influence because it reduces the capacity for the evapotranspiration of plants, consequently, reduces net photosynthesis.

It is important to emphasize that in the selection process for the prediction variables for soybean yield, we applied the method of testing all possible combinations with up to four variables, which produced a total of 24,157 combinations of independent variables, from which we initially removed equations that showed multi-collinearity. The viable equations were ordered to reduce the MAPE and increase the adjusted R^2 ($p < 0.05$).

All the models calibrated to predict soybean yield was accurate and precise and had a low tendency (Table 1). The model calibrated for the north of Mato Grosso do Sul state yielded the following statistical indices: R = 0.4; R^2 = 0.38; d = 0.45; C = 0.18; Ea = 97; Es = 2.48; EAmax = 4.1; MSE = 7.06; RMSE = 2.66; MAE = 2.33; and MAPE = 4.63% (Table 2). A calibrated model with a MAPE of 5.197% (central region) was considered accurate since for average soybeans yield of 55 sacks ha^{-1}, there was a deviation of just ±2.80 sacks.ha^{-1}. Several authors who study crop modeling have reported that a model with MAPE below 6.063%, as found in the current study for the Central and South regions, is considered to have a low error for modeling using climate data (Moreto and Rolim, 2015).

The models calibrated for the regions of Mato Grosso do Sul are shown in Table 1. The models from the north and central regions show a prevision of anticipation of 2 months (59 days) since the final data necessary to apply the model were EXC_{JANc} and P_{JANc}, respectively. The models calibrated for the southern region reported anticipation of 1 month (31 days) since the final data necessary to apply the model was EXC_{FEVc}.

The variables selected to compose the prediction models were strictly related to water conditions since all model variables were water-based: P, ETP, ETR, and EXC. For the northern region, the variable with the greatest influence was $ETR_{JANc,}$ which represents the moment when this crop is in the initial phase of grain filling. The elasticity analysis of ETR_{JANc} demonstrates that it has a strong and direct relationship with soybean yield since its elasticity was +0.252 and significant at p<0.05. This elasticity indicates that there was an increase in 10% in ETR_{JANc} of soybean, this caused an increase of 2.252% in the crop yield (Table 1, Model [1]).

The spatial variation of predicted and real yield of soybeans in Mato Grosso do Sul is shown in Fig. 6. In the southern region real yield varied between 50.1 and 55 sacks ha^{-1}, while in the central region yield was above 55 sacks ha^{-1}, as observed in the localities of Ivinhema, Amaurilândia, and Batayporã (Fig. 6B). With high accuracy, these regression models were able to predict this spatial variation of soybean yield in Mato Grosso do Sul (Fig. 6B).

The deviation between the real and estimated yield of soybeans in Mato Grosso do Sul is observed in Fig. 6C. In 86% of the territory of Mato Grosso do Sul the models, as a function of climatic conditions, demonstrated deviations lower than 5 sacks ha^{-1}. In a few localities such as Costa Rica, Alcinópolis, Cassilândia, Camapuã, Maracaju, Bonito, and Eldorado, the models demonstrated deviations between 5 and 10 sacks ha^{-1}, however, these localities represent less than 10% of the total area of Mato Grosso do Sul. The performance of these models also underestimates soybean yield less than 54.5 sacks ha^{-1} (Fig. 7).

4 Conclusions and perspectives

The calibrated models used to forecast soybean yield as a function of climatic conditions have a high degree of significance, high accuracy, and errors lower.

The models for the northern and central regions show a prevision of anticipation of 2 months (59 days) before soybean harvest, a period that is essential for producers to be able to conduct pre- and post-harvest planning.

Water stress mainly in December (DEF_{DECp}) is the climate variable with the greatest negative influence on soybean yield in Mato Grosso do Sul state.

In the northern region of the state occur the highest air temperatures, of 27ºC. While, the lowest mean air temperatures of the regions occur between June and July, where the lowest mean air temperature is in the southern region of MS, with 19ºC. These results are within adequate air temperatures for soybean cultivation in MS.

The annual water deficit (WD) is more intense in the north of the State of Mato Grosso do Sul between May to October, with 140 mm y^{-1}.

Acknowledgements: We thank the Federal Institute of Mato Grosso Sul, Campus Naviraí, for funding this research.

Author's contribution: LEOA conceived of the project and together with GBT designed the study. GBT, JRSCM and KCM were responsible for collected the data and carried out the statistical analyses. JAL and PAL were responsible for the field work. All authors approved the fnal version of the manuscript.

Funding: This research was supported by the IFMS - Federal Institute of Education, Science and Technology of Mato Grosso do Sul - Campus of Naviraí, Naviraí, Brasil.

Compliance with ethical standards

Conflict of interest: The authors declare that they have no conflict of interest

References

Aparecido, L. E. D. O., Rolim, G. D. S., Moraes, J. R. D. S. C., Rocha, H. G., Lense, G. H. E., & Souza, P. S. (2018). Agroclimatic zoning for urucum crops in the state of Minas Gerais, Brazil. *Bragantia*, *77*(1), 193-200.

Alvares, C. A., Stape, J. L., Sentelhas, P. C., de Moraes, G., Leonardo, J., & Sparovek, G. (2013). Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, *22*(6), 711-728.

Bonato, E. R., Bertagnolli, P. F., Ignaczak, J. C., Tragnago, J. L., & Rubin, S. (1998). Performance of soybean cultivars in three sowing dates in Rio Grande do Sul, Brazil. *Pesquisa Agropecuaria Brasileira (Brazil), 33(6)*, 879-884.

CONAB (2019). Monitoring of the Brazilian harvest: grains, V.4, setembro 2019. Brasília: Conab, 2019. 158 p. https://www.conab.gov.br/info-agro/safras Acessed 22 December 2019.

Dourado Neto, D.; Sparovek, G.; Figueredo Júnior, L.G.M.; Fancelli, A.L.; Manfron, P.A.; Medeiros, S. L. P. (2004). Model for estimating the productivity of depleted corn grains based on soil water balance. *Revista Brasileira de Agrometeorologia*, Santa Maria - RS,12 (2), 359-367.

Fietz, C. R., & Urchei, M. A. (2002). Deficiência hídrica da cultura da soja na região de Dourados, MS. *Revista Brasileira de Engenharia Agrícola e Ambiental*, *6*(2), 262-265.

Fontana, D. C.; Berlato, M. A.; Lauschner, M. H.; Mello, R. W. (2001). Estimated soybean yield model in the State of Rio Grande do Sul. *Pesquisa Agropecuária Brasileira*, Brasília, 36(3), 399-403.

Franke, A. E. (2000). Need for supplemental irrigation in soybeans in the edaphoclimatic conditions of the Planalto Médio and Missões, RS. *Pesquisa Agropecuaria Brasileira (Brazil)*, 35(8) .

Lasdon, L. S.; Waren, A. D. (1982). GRG2 user´s guide. Depto of general Business, Shchoool of Business Administration, University of Texas, Austin, TX.

Martins, E.; Aparecido, L.E. O.; Santos, L.P.S.; Mendonça, J.M.A.; Souza, P.S. (2015). Influence of climatic conditions on the productivity and quality of coffee produced in the southern region of Minas Gerais. *Coffee Science*, 10, 499-506.

Martorano, L. G.; Bergamaschi, H.; Dalmago, G. A.; Faria, R. T.; Mielniczuk, J.; Comiran, F. (2009). Soil water status indicators with soybean under no-tillage and conventional tillage. *Revista Brasileira de Engenharia Agrícola e Ambiental*, Campina Grande, 13(4), 397–405.

Martorano, L.G., Bergamaschi, H.; Faria, R.T., Dalmago, G.A. (2012) Decision Strategies for Soil Water Estimations in Soybean Crops Subjected to No-Tillage and Conventional Systems, in Brazil. In: Manish Kumar (Org.). Problems, Perspectives and Challenges of Agricultural Water Management. *Croácia: InTech*,. 439-456.

Moreto, V. B.; Rolim, G. S. (2015). Estimation of annual yield and quality of "Valência" orange related to monthly water deficiencies. *Afr. J. Agric. Res*., 10 (6), 543-553.

Rolim, G. S.; Ribeiro, R. V.; Azevedo, F. A.; Camargo, M. B. P; Machado, E. C. (2008). Prediction of the number of fruits based on the amount of reproductive structures in orange trees. *Revista Brasileira de Fruticultura*, 30 (1), 48-53.

Valeriano, T. T. B.; Rolim, G.S.; Aparecido, L. E. O. (2017). A method to determine agro-climatic zones based on correlation and cluster analyses. *Theoretical and Applied Climatology*, *134*(3-4), 1355-1364.

Tables

Table 1. Calibrated models to estimate soybean supply in the state of Mato Grosso do Sul, as affected by climate control.

Regions	**Models**	**p-value**	**Forecasting**	
			Month	**Days**
NORTH	$Y = 0.029.\ P_{DEZp} – 0.245.\ ETR_{NOVp} – 0.252.\ ETR_{JANc} + 0.028.\ EXC_{JANc} + 82.461$	0.0001	2	59
CENTER	$Y = 0.0214.\ P_{JANc} – 0.290.\ ETP_{DEZp} - 0.012.\ ETR_{JANc} + 0.0183.\ EXC_{DEZp} + 63.91$	0.0004	2	59
SOUTH	$Y = 0.056.\ P_{DEZp} – 0.045.\ EXC_{NOVp} – 0.035.\ EXC_{JANc} + 0.053.\ EXC_{FEVc} + 39.817$	0.0031	1	31

Table 2. Statistical indices used to evaluate the accuracy, precision, and tendency of calibrated models used to forecast soybean yield in Mato Grosso do Sul

Statistical indices	Regions		
	CENTER	NORTH	SOUTH
R	0.55	0.4	0.84
R^2	0.452	0.38	0.689
d	0.72	0.45	0.91
C	0.39	0.18	0.77
Ea	1.82	0.97	2.31
Es	1.23	2.48	0.49
EAmax	4.6	4.1	5.3
MSE	4.8	7.06	5.59
RMSE	2.19	2.66	2.36
MAE	1.69	2.33	1.95
MAPE	5.19	4.63	6.06

Figures

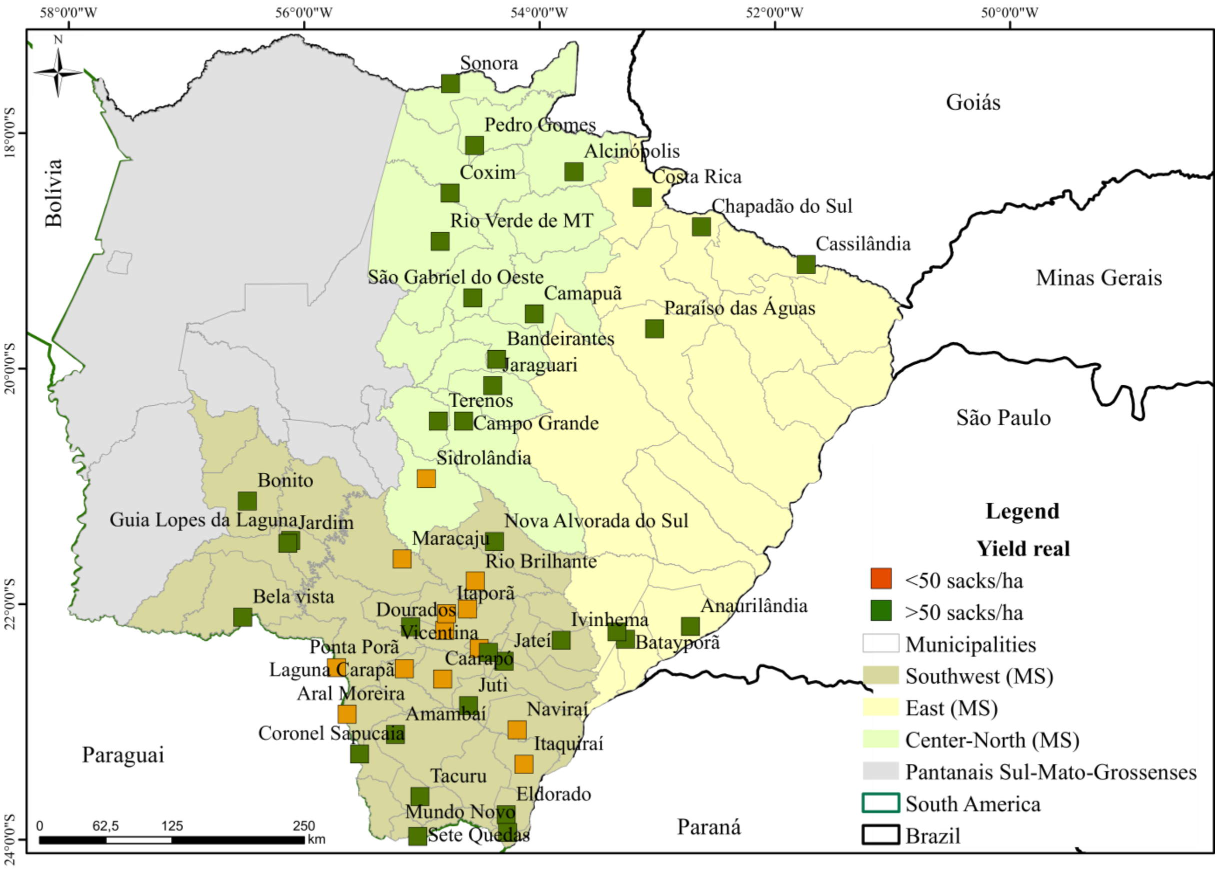
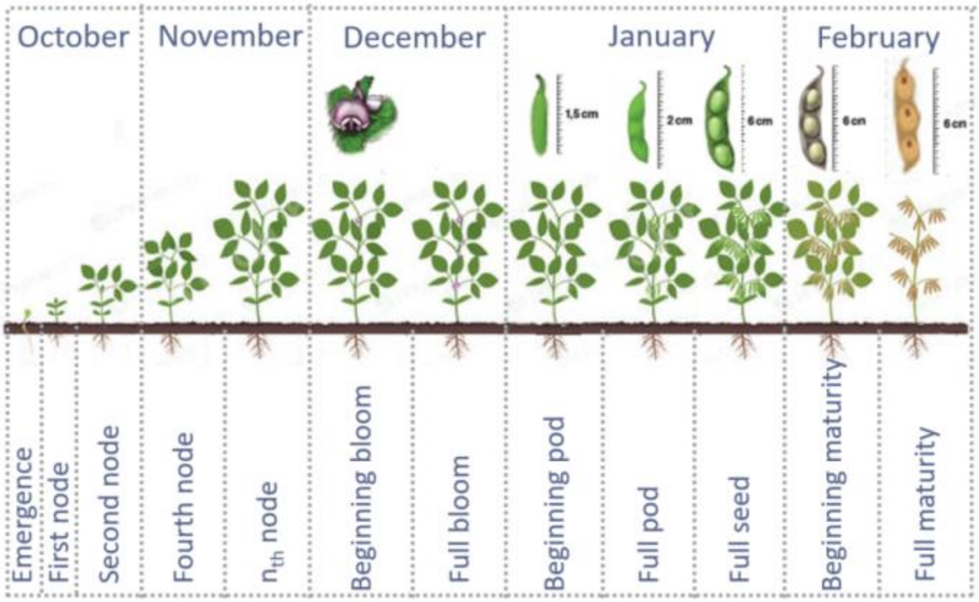


Fig. 1. The geographic location of soybean production regions in Mato Grosso do Sul, Brazil.

341 

342 **Fig. 2.** Phenology of planting and harvest of soybeans.

343

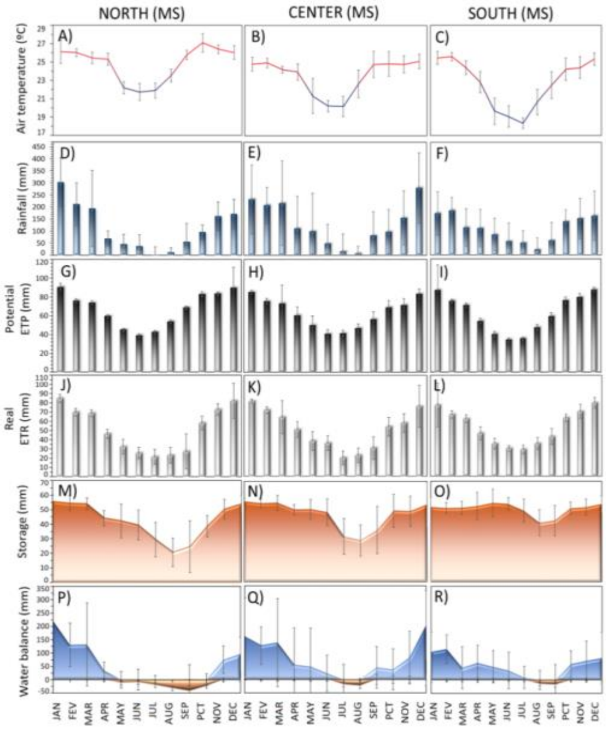


Fig. 3. Variation of climatic variables for the North, Central, and South regions of Mato Grosso do Sul.

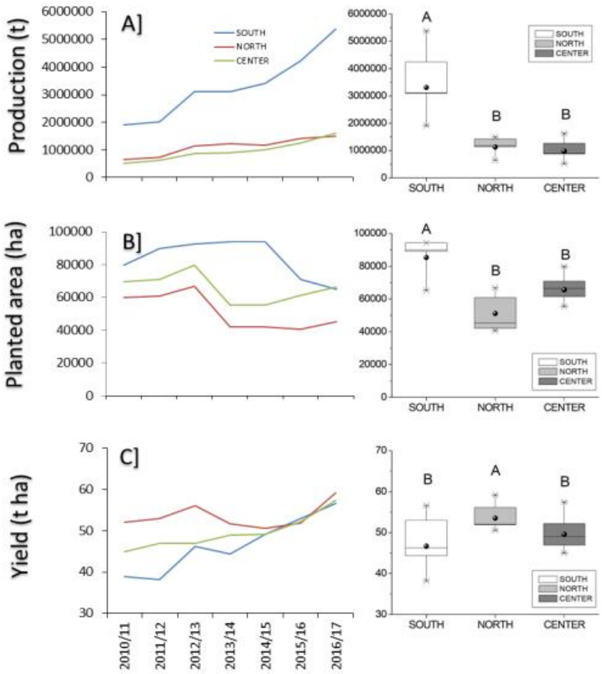
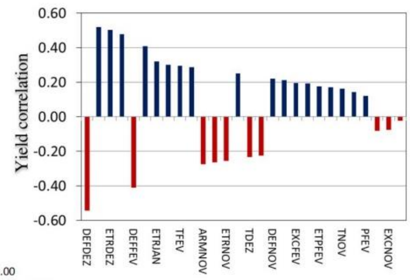


Fig. 4. Variation in production, area, and yield of soybean between 2008 and 2018 for the Northern, Central, and Southern regions of Mato Grosso do Sul, Brazil. Legend = Averages with identical capital letters do not significantly differ by the Scott-Knott test at 5 % probability.

	yield	TNOV	TDEZ	TJAN	TFEV	PNOV	PDEZ	PJAN	PFEV	ETPNOV	ETPDEZ	ETPJAN	ETPFEV	ETRNOV	ETRDEZ	ETRJAN	ETRFEV	ARMNOV	ARMDEZ	ARMJAN	ARMFEV	DEFNOV	DEFDEZ	DEFJAN	DEFFEV	EXCNOV	EXCDEZ	EXCJAN	EXCFEV
yield	1.00																												
TNOV	0.16	1.00																											
TDEZ	-0.23	0.81	1.00																										
TJAN	0.17	0.59	0.31	1.00																									
TFEV	0.30	0.56	0.35	0.56	1.00																								
PNOV	-0.08	-0.51	-0.42	-0.28	-0.72	1.00																							
PDEZ	0.30	-0.41	-0.62	-0.33	-0.53	0.71	1.00																						
PJAN	0.25	0.15	-0.16	-0.07	-0.22	0.20	0.52	1.00																					
PFEV	0.12	-0.27	-0.18	-0.07	0.01	0.18	0.18	-0.38	1.00																				
ETPNOV	-0.02	0.61	0.68	0.56	0.69	-0.45	-0.65	-0.45	0.00	1.00																			
ETPDEZ	-0.23	0.70	0.90	0.43	0.49	-0.40	-0.70	-0.41	-0.17	0.89	1.00																		
ETPJAN	0.14	0.58	0.32	1.00	0.57	-0.27	-0.35	-0.13	-0.08	0.60	0.47	1.00																	
ETPFEV	0.18	0.13	-0.11	0.01	0.32	-0.11	0.18	0.44	-0.10	-0.24	-0.26	-0.03	1.00																
ETRNOV	-0.26	-0.09	0.10	0.19	-0.28	0.56	0.03	-0.09	-0.07	0.15	0.23	0.22	-0.13	1.00															
ETRDEZ	0.50	0.18	-0.08	0.43	0.20	0.17	0.32	-0.08	0.29	0.26	0.06	0.44	-0.18	0.08	1.00														
ETRJAN	0.32	0.52	0.12	0.35	0.34	-0.06	0.35	0.53	-0.30	0.11	0.06	0.34	0.51	-0.13	0.21	1.00													
ETRFEV	0.48	-0.32	-0.63	0.00	-0.05	0.28	0.48	0.14	0.49	-0.32	-0.60	-0.04	0.22	-0.14	0.41	-0.05	1.00												
ARMNOV	-0.28	-0.50	-0.31	-0.22	-0.65	0.74	0.33	0.04	0.00	-0.40	-0.28	-0.21	-0.05	0.82	-0.15	-0.30	0.08	1.00											
ARMDEZ	0.52	-0.40	-0.71	0.00	-0.25	0.48	0.75	0.19	0.35	-0.42	-0.64	-0.02	-0.07	-0.06	0.70	0.06	0.75	0.13	1.00										
ARMJAN	0.21	0.26	0.00	-0.14	0.04	0.09	0.53	0.62	-0.29	-0.19	-0.15	-0.15	0.56	-0.20	-0.02	0.87	-0.08	-0.17	0.04	1.00									
ARMFEV	0.41	-0.35	-0.54	-0.02	-0.14	0.24	0.39	-0.07	0.66	-0.20	-0.48	-0.05	-0.19	-0.16	0.53	-0.27	0.88	0.01	0.78	-0.30	1.00								
DEFNOV	0.22	0.42	0.29	0.15	0.64	-0.77	-0.40	-0.17	0.07	0.43	0.29	0.14	-0.01	-0.83	0.08	0.18	-0.06	-0.98	-0.18	0.07	0.03	1.00							
DEFDEZ	-0.54	0.34	0.69	-0.04	0.18	-0.41	-0.73	-0.22	-0.34	0.42	0.64	-0.01	-0.04	0.10	-0.73	-0.12	-0.73	-0.08	-0.98	-0.09	-0.74	0.14	1.00						
DEFJAN	-0.26	-0.26	0.04	0.13	-0.07	-0.07	-0.54	-0.62	0.27	0.19	0.17	0.15	-0.55	0.25	0.00	-0.88	0.03	0.21	-0.08	-1.00	0.26	-0.12	0.12	1.00					
DEFFEV	-0.41	0.37	0.59	0.00	0.18	-0.32	-0.41	0.04	-0.53	0.23	0.50	0.03	0.18	0.08	-0.49	0.25	-0.92	-0.10	-0.79	0.30	-0.96	0.05	0.72	-0.25	1.00				
EXCNOV	-0.08	-0.59	-0.49	-0.35	-0.73	0.99	0.73	0.18	0.22	-0.51	-0.48	-0.34	-0.10	0.49	0.13	-0.09	0.30	0.72	0.50	0.09	0.27	-0.73	-0.43	-0.07	-0.34	1.00			
EXCDEZ	0.29	-0.44	-0.63	-0.36	-0.58	0.74	0.99	0.52	0.19	-0.67	-0.72	-0.39	0.15	0.06	0.28	0.29	0.51	0.38	0.74	0.49	0.41	-0.44	-0.71	-0.49	-0.45	0.76	1.00		
EXCJAN	0.19	0.15	-0.13	-0.08	-0.27	0.22	0.50	1.00	-0.38	-0.46	-0.39	-0.14	0.42	-0.07	-0.12	0.49	0.10	0.06	0.15	0.59	-0.10	-0.20	-0.18	-0.59	0.07	0.20	0.50	1.00	
EXCFEV	0.20	-0.30	-0.26	-0.06	0.05	0.16	0.22	-0.36	0.99	-0.02	-0.22	-0.07	-0.09	-0.11	0.35	-0.26	0.53	-0.03	0.43	-0.25	0.70	0.09	-0.42	0.23	-0.57	0.20	0.23	-0.37	1.00



354

355 **Fig. 5.** Effect of the climatic variables on soybean yield by Pearson correlation.

356

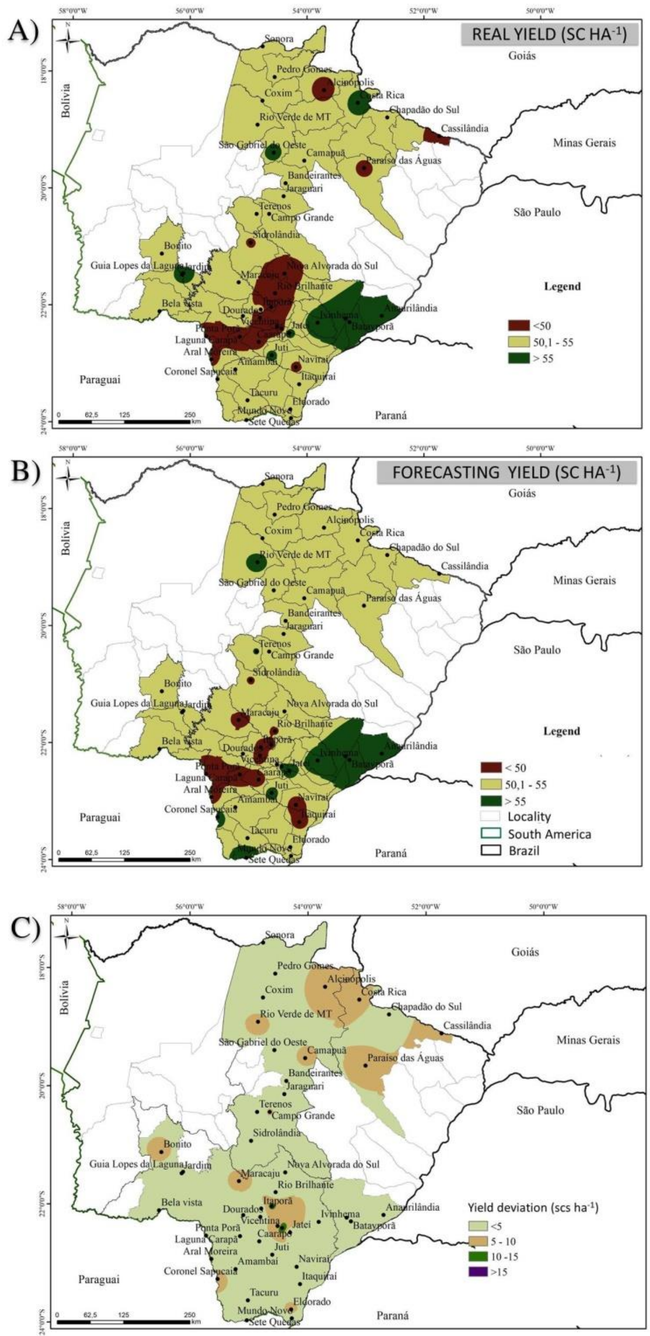
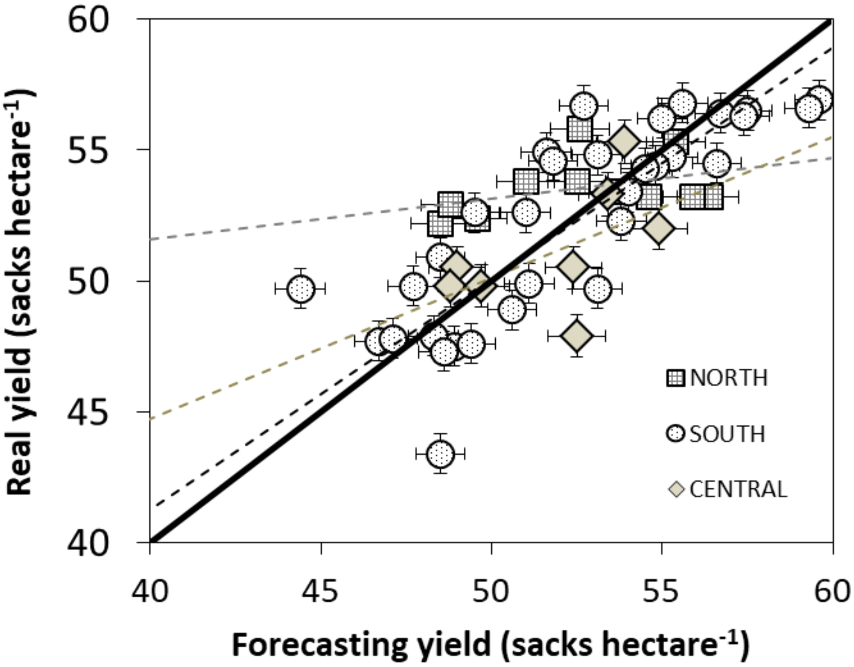


Fig. 6. Maps of real yield (A), forecasted yield (B), and the difference between real and forecasted yield (C) for the calibrated model in function of climate conditions in Mato Grosso do Sul, Brazil.



361

362 **Fig. 7.** Performance of the model of prediction of soybean yield in Mato Grosso do Sul, Brazil.