



REVISTA DA SOCIEDADE BRASILEIRA DE AGROMETEOROLOGIA

ISSN 2526-7043 www.sbagro.org.br DOI: http://dx.doi.org/10.31062/agrom.v32.e027580

# Soybean yield monitoring system for Brazil with FAO-AEZ crop model

**Rogério de Souza Nóia Júnior1(\*), José Lucas Safanelli2 , Lucas Fernandes de Souza3 and Durval Dourado Neto4**

 TUM School of Life Sciences, Digital Agriculture. Technical University of Munich, Freising, Germany. Email: rogeriosouzanoia@gmail.com Woodwell Climate Research Center. Falmouth MA, USA. E-mail: jsafanelli@woodwellclimate.org National Supply Company. Brasília, DF, Brazil. E-mail: lucas.fsouza@agricultura.gov.br "Luiz de Queiroz" College of Agriculture. University of São Paulo, Piracicaba, SP, Brazil. E-mail: ddourado@usp.br (\*)Corresponding author.

# ARTICLE INFO ABSTRACT

**Article history:** Received 1st February 2024 Accepted 16 July 2024

### **Index terms:**

crop yield monitoring food security *Glycine max* (L) Merr. yield anomaly

This study aimed at developing a soybean yield monitoring system based on a crop model. Previously calibrated and validated FAO Agro-Ecological Zone (FAO-AEZ) crop model was used to simulate soybean yields for 59 agroclimatic homogenous zones distributed across the Brazilian territory. The spatial performance of FAO-AEZ to simulate soybean yield at homogenous zones level was computed by comparing simulated median soybean yield with observed soybean yield data from national records after applying a trend correction on the time series of 2011 to 2020. To monitor the soybean crop in-season performance, we computed the soybean crop yield anomaly representing the in-season yield variation in relation to five previous cropping seasons with the same sowing dates. The in-season analysis was limited to the most recent cropping seasons reported by the national bulletins (2021 and 2022). The FAO-AEZ crop model performance between 2011 and 2020 reached an r2 of 0.59 and an RMSE of 402 kg ha-1 when compared to historical national statistics. In turn, the in-season analysis for 2022 revealed the model's capacity to anticipate signs of negative yield trends three months in advance compared to national bulletins. National yield estimations are possible at earlier months of the season further improving the prediction performance when approaching the end.

© 2024 SBAgro. All rights reserved.

# **Introduction**

Brazil is the largest producer and exporter of soybeans in the world, accounting for half of the soybean global market. The country raised US\$ 28 billion from soybean exports in 2020, which represents a third of the national agricultural sector exports (FAO, 2022). Variations in Brazilian soybean may affect the national economy and global food prices. Droughts and heat are the main

causes of the widespread historical crop yield failures in Brazil, which caused a shortfall of 10 million tons in 2012 (Nóia Júnior *et al.*, 2020). National yield prediction systems for soybeans in Brazil are important to minimize possible disruptions in food supply, helping agricultural commodity stakeholders and policymakers to plan mitigation strategies in advance.

Soybean crop yield monitoring and surveying in Brazil is a responsibility of the National Supply Company (CONAB), which is a governmental institute under the Ministry of Agriculture, Livestock, and Food Supply. The CONAB releases monthly agricultural bulletins during the soybean cropping season from October to July, with yield and production prospects mostly based on field surveys with observations made by crop experts together with market signals, remote sensing, and weather analysis from all soybean-producing regions. For the 2022 cropping season, for instance, CONAB's expectation in October 2021 was for national soybean production to exceed 140 million tons, but a historical drought in January of 2022 decreased the expectation to less than 125 million a few months later (CONAB, 2022). The soybean monitoring over the cropping season made by CONAB is resource-intensive and timeconsuming, which could be facilitated with the adoption of crop simulation models as an additional source of information.

Crop simulation models employed to predict soybean yield in Brazil have been the topic of several studies in the last years (Battisti *et al.*, 2017; Silva *et al.,* 2021). These studies helped to identify management practices to minimize the potential impacts of climate variability and climate change on soybean yield (Battisti & Sentelhas, 2017, 2019; Battisti *et al*., 2017; Nóia Júnior & Sentelhas, 2019; Silva et al., 2021), as well as to quantify its yield potential through yield gap analysis (Nóia Júnior & Sentelhas, 2020; Sentelhas *et al.*, 2015). Many crop simulation models, such as AQUACROP (Steduto *et al.,* 2009), DSSAT CROPGRO–Soybean (Boote *et al.*, 2003; Hoogenboom *et al.*, 2018), APSIM Soybean (Keating *et al.*, 2003), MONICA (Nendel *et al.*, 2011), and FAO Agro-ecological Zone (FAO-AEZ) (Doorenbos & Kassam, 1979; Kassam, 1977), were tested and demonstrated a good performance across a range of Brazilian growing conditions (Battisti *et al*., 2017; Sampaio *et al.*, 2020; Silva *et al.*, 2021). However, crop models usually require detailed information on initial conditions, soils, cultivars, and crop management at a field level, often hindering their use to simulate crop yield at broader scales such as the national level.

Some crop simulation models, such as the SIMPLE (Zhao *et al.,* 2019) and the FAO-AEZ require only commonly available inputs including daily weather data, crop management, and soil water holding parameters. These models include temperature and water, but no nutrient stress (Doorenbos & Kassam, 1979; Zhao *et al.*, 2019). In addition, they miss the waterlogging and plant disease effects on crop growth, as most of the existing crop simulation models (Rötter *et al.*, 2018). Thus, these simpler models should be applied to simulate crop yields only where waterlogging, pests, plant diseases, and nutrients are not major limiting factors.

With an average national soybean yield of 3.4 t ha-1 , Brazil is among the five countries with the highest

soybean yield in the world (FAO, 2022). In its production systems, pests, plant diseases, and nutrient losses are not the main factors behind its inter-annual yield variability observed across the country (Sentelhas *et al.*, 2015). Waterlogging can occur in different Brazilian regions, but it is more common during soybean harvest at the end of January (Lima *et al.*, 2019), when usually it no longer affects soybean yield but the quality of the harvested seed (Pasley *et al.*, 2020). Droughts and high temperatures are the main factors behind soybean yield inter-annual variability in Brazil (Sentelhas et al., 2015), which are stress factors to crop yield considered by crop models, such as the FAO-AEZ. Therefore, the objective of this study was to develop a soybean yield monitoring system in Brazil based on the simple and well-validated FAO-AEZ crop model.

# **Material and Methods**

# **Soybean yield anomalies and agro-climatic homogenous zones for Brazil**

A flowchart detailing the sequence of procedures adopted in this study is presented in Figure 1.

Soybean yield from 278 geographical micro-regions for 29 years from 1991 to 2019 (IBGE, 2022), representing the entire Brazilian soybean production were used to calculate yearly soybean yield anomaly. Yield anomalies  $(Y_{\text{ann}})$  were computed as the percent difference between observed  $(Y_{obs})$  and average  $(Y_{sys})$  trend-corrected yield divide by  $Y_{\text{ave}}$ :

$$
Y_{anm}(\%) = \frac{Y_{obs} - Y_{avg}}{Y_{avg}} \cdot 100
$$
 Equation (1)

Using yield anomalies, a hierarchical clustering analysis was performed across the geographical micro-regions and the 29 years. As a result, 59 soybean homogenous zones were defined based on their agro-climatic conditions, in a similar procedure as suggested for wheat in Brazil by Nóia Júnior *et al.* (2021). The 59 soybean homogenous zones were distributed in 21 states, including the Federal district of Brazil (Table 1). In our hierarchical clustering analysis, we restricted the inclusion of regions within the same group to only those that are contiguous. Given that the analysis focused on yield anomaly rather than absolute or mean yield levels, the paramount consideration lies in the interannual variation of soybean yield, predominantly influenced by regional interannual climatic fluctuations. Consequently, our classification of soybean homogenous zones primarily delineates these zones based on their interannual climatic variability during the soybean growing season. The soybean yield was trend corrected with a linear regression, as demonstrated by Guarin *et al.* (2020).

**Figure 1.** Flowchart of the procedures for soybean yield crop monitoring system in Brazil.



# **Locations, climate and soil data**

This study considered one representative location within each of the 59 soybean agro-climatic homogenous zones (Table 1). These 59 locations were defined based on the soybean planting intensity, and the point with the largest soybean planted area was selected as the most representative for each zone. From each location, daily weather data including maximum and minimum air temperature and rainfall were obtained from the gridded weather source of NASA POWER (Prediction of Worldwide Energy Resources, https://power.larc.nasa.gov/). The NASA POWER was found to be a suitable weather database for characterizing weather patterns and estimating soybean yield in Brazil (Battisti *et al*., 2019).

The predominant soil types of each site were selected using the soil data from IBGE (2022). Information about sand, clay, and silt contents and bulk density for each soil type were obtained from national-wide official soil surveys (BRASIL, 1981), and the soil water holding capacity of each soil was estimated using pedo-transfer functions developed by Reichert *et al.* (2009) (Table S1).

# **Crop simulation model**

This study simulated soybean yield by using the FAO – AEZ (Doorenbos & Kassam, 1979; Kassam, 1977). Calibration and validation of the FAO-AEZ to simulate soybean yield in Brazil was performed by Battisti *et al.* (2017). Battisti *et al.* (2017) calibrated and validated the FAO-AEZ using data from seven different locations distributed across Brazil in two cropping seasons (2014 and 2015) with several sowing dates (varying from October to January) and different regimes of water management. This model presented an average error of 640 kg ha<sup>-1</sup>, with a precision  $(r^2)$  of 77% and an accuracy (Willmott agreement index) of 92%. The genetic coefficient parameters for simulating soybean yield with FAO-AEZ are presented in Table 2.

For running the crop model, the initial soil water content was defined based on the water balance initiated six months before sowing, considering the prior crop as fallow. The FAO-AEZ crop model does not simulate soybean crop cycle duration, and thus we set its duration according to each region following average values of the national registry of commercial cultivars from the Ministry of **Table 1.** Agro-climatic homogenous zones for soybean in Brazil, representative locations and their associated geographical coordinates, water capacity and soybean cycle duration.





Agriculture, Livestock and Supply (MAPA, 2021).

yields. The percent yield anomaly (Equation 2) is therefore calculated for each replicate and further averaged:

# **In-season soybean crop monitoring**

To monitor the soybean crop in-season performance, we computed the soybean crop yield anomaly, which represents the in-season yield variation in relation to previous cropping seasons. This anomaly was calculated based on the relationship between the simulated soybean yield for the current cropping season and the simulated yield for the last five cropping seasons, both with the same sowing dates. The yield anomaly is calculated based on the last five cropping seasons because Brazilian public and private companies use this period as a reference (so the results are comparable). This index is calculated for 16 sowing dates from September to January, with weekly intervals, comprising almost the entire soybean sowing window in Brazil (except for the Northern state of Roraima). The in-season crop yield anomaly is computed with 15 days intervals within the current cropping season, which means that the first monitoring date will be carried out 15 days after the first sowing date in September.

To compute the in-season crop yield anomaly for a specified sowing date, we simulated soybean yield using in-season observed weather data from the sowing date to the monitoring date, completed until the maturity date with the weather data of the last five cropping seasons. This is the same approach used by Nóia Júnior *et al.* (2022), where more detailed explanations are provided. At the same time, we simulated soybean yield using weather data for the entire five last cropping seasons. This means that the crop model is run with 5 replicates representing the pair of the current cropping season yield with five completions and the last five historical cropping season

$$
Yield\;anomaly\; (%) = \frac{\sum_{h=1}^{5} 100 \left[ \left( \frac{Y_c}{Y_h} \right) - 1 \right]}{5} \text{ Equation (2)}
$$

After the end of the main sowing window period in January, the crop yield anomaly is spatially aggregated to the state level using the 5-year average production intensity levels, and temporally aggregated considering the state sowing progression curve. These aggregations allowed the interpretation of the final in-season soybean yield anomaly across the Brazilian territory and the comparison with national bulletins released by IBGE and CONAB. The historical sowing progress is not available before 2020 (in this period the company did not monitor the sowing date or did not make it public), therefore the in-season analysis was limited to the most recent cropping seasons reported by CONAB in 2021 and 2022.

# **National soybean yield and statistical analysis**

The statistical analyses were carried out using the R software (R Core Team, 2021). The spatial performance of FAO-AEZ to simulate soybean yield at homogenous zones level was computed by comparing simulated median soybean yield with observed soybean yield data from the IBGE after applying a trend-correction on the time series of 2011 to 2020 (IBGE, 2022b). For this, the simulated soybean yield was computed by multiplying the simulated yield anomaly by the observed average of soybean yield in the last 5 years for each agroclimatic homogenous zone. The performance of the FAO-AEZ to simulate historical

**Table 2.** Genetic coefficient parameters for simulating soybean yield with FAO-AEZ, according to Battisti et al. (2017).



national soybean yield was computed with the following statistical indices: bias (mean error), coefficient of determination ( $r^2$ ), root mean square error (RMSE), and mean absolute percentage error (MAPE).

# **Results**

# Soybean homogenous agro-climatic zones

The 59-soybean homogenous agro-climatic zones (Figure 2) were defined based on soybean yield anomalies as from 1991 to 2019. This analysis assumes that within a homogeneous zone, the soybean yield anomaly in a given varied from 2 to 23%, with a spatial variation simil year should be relatively similar among different locations. In 2012 the soybean anomaly was positive in practically all and the highest values in south-central Brazil. regions of northeastern Brazil (Figure 1b) but negative in most of southern Brazil (Figure 1d). In southern Brazil, crop simulation model for all homogenous agro-clin regions that presented negative anomalies in 2012 and positive anomalies in 2020 formed different groups (e.g., of 9.87% (Figure 4). The FAO-AEZ model overesting G2 and G1), compared to the ones with negative anomalies in both years (e.g., G4 and G5). Similar results are shown for the northeastern region in 2012 and 2016 (Figure 1b The observed yield from 2010 to 2020 is based of southern Brazil (Figure 1d). and 1c). As the soybean anomaly is relatively similar in different years within the agro-climatic homogeneous and Statistics (IBGE, 2022a) for each location of the anomalies in 2012 and positive and po zones, the monitoring of one representative location inside each one of these zones should be enough to represent the indices for model performance are the Coefficien regional patterns (NÓIA JÚNIOR *et al.*, 2021).

# **Performance of FAO-AEZ crop model to estimate observed regional soybean yield in Brazil**

The FAO-AEZ crop simulation historical model performance for soybean yield varied according to the soybean yield anomaly in most parts of the Northea

soybean agro-climatic homogenous zones of Brazil (Figure 3). The average bias of the soybean yield simulations varied from -318 to 236 kg ha<sup>-1</sup>, presenting the highest values in the northeast region of Brazil. For most of the locations, the bias presented positive values, indicating an overall underestimation of the crop model. The RMSE varied from 83 to 848 kg ha<sup>-1</sup>. The highest values of the RMSE were obtained in central Brazil, particularly in the states of Mato Grosso do Sul and São Paulo. Northern Brazil, such as north of Mato Grosso and Rondônia states, presented the smallest values of RMSE across the country. The MAPE varied from 2 to 23%, with a spatial variation similar to RMSE, where the lowest values were found in the north and the highest values in south-central Brazil.

Overall, the simulation of soybean yield with FAO-AEZ crop simulation model for all homogenous agro-climatic zones showed an  $r^2$  of 0.59, RMSE of 402.5 kg ha<sup>-1</sup>, and a MAPE of 9.87% (Figure 4). The FAO-AEZ model overestimates soybean yield below 2,500 kg ha<sup>-1</sup> and underestimates the soybean yield above 3,500 kg ha<sup>-1</sup>.

The observed yield from 2010 to 2020 is based on the reported yield from the Brazilian Institute of Geography and Statistics (IBGE, 2022a) for each location of the agroclimatic homogenous zones in Brazil. The statistical indices for model performance are the Coefficient of determination (r2 ), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

We tested the FAO-AEZ crop simulation model in 2016, observed regional soybean yield in Brazil and the south and extremely low of soybean yield in Brazil (Figure 5). The FAO-AEZ model captured the negative soybean yield anomaly in most parts of the Northeast, as

258 *et al.*, 2021). (e) 2020.**Figure 2. Soybean agro-climatic homogenous zones in Brazil.** (a) Soybean agro-climatic homogenous zones in Brazil based on interannual yield anomalies, and the soybean yield anomalies for northeastern in (b) 2012 and (c) 2016, and for southern in (d) 2012 and



**Figure 3.** Spatial performance of FAO-AEZ crop model to simulate soybean yield in Brazil. The statistical indices for model performance are the bias (a), Root Mean Squared Error (RMSE) (b), and Mean Absolute Percentage Error (MAPE) (c).



**Figure 4.** Observed versus simulated soybean yield with FAO-AEZ crop model.



well as in central regions of Brazil, such as central Mato horth and northeast of Brazil, and Grosso and north of Goiás. The model also captured a positive soybean anomaly in south-central Brazil.

**In-season soybean yield monitoring system in 2022**

Given the ability of the FAO-AEZ to simulate soybean yield, we extend the analysis to simulate in-season (Figure 6c). In May, with the soybean harvest compi soybean yield anomaly during the 2022 cropping season. This cropping season presented abnormal weather of more than 25% in three states in the southern re conditions, with droughts in the south and excessive rains in the north of Brazil, both from December to February. Our simulations with the FAO-AEZ model indicated in early January a positive soybean yield anomaly in the (Figure 6d). The FAO-AEZ model also indicated a

north and northeast of Brazil, and a negative anomaly in the south of Brazil (Figure 6a). In January, the simulations showed a soybean anomaly in south-central Brazil.<br>
29 Showed a soybean anomaly of almost -50% in the in the reported to 2010 to 2010 to 2010 to 2010 to 2010 to 2 southernmost state of Brazil, Rio Grande do Sul. The In-season soybean yield monitoring system in 2022 CONAB's official bulletins of January pointed to a drop in soybean yield in southern Brazil but of less than 10% (Figure 6c). In May, with the soybean harvest completed, the FAO-AEZ model indicated a soybean yield anomaly of more than 25% in three states in the southern region of Brazil, Mato Grosso do Sul, Paraná, and Rio Grande do Sul (Figure 6b). Similar results in southern Brazil were observed by the CONAB's bulletins after harvest, in May (Figure 6d). The FAO-AEZ model also indicated a yield

**Figure 5.** Comparison between (a) observed and (b) simulated soybean yield in 2016. The observed yield is from the Brazilian Institute of Geography and Statistics (IBGE, 2022a). Soybean yield was simulated with the FAO-AEZ crop model.



b) Simulated soybean yield anomaly in 2016



increase of up to 50% in the northeastern states of Brazil he 2022 cropping season with the FAO-AEZ mode (Figure 6b), which was not reported by CONAB (Figure

6d). The biweekly monitoring of the soybean yield during observed yield is from the Brazilian Institute of Geography and Statistics (IBGE, 2022a). 323 description in Portuguese for better engagement with Brazilian stakeholders. An in-season the 2022 cropping season with the FAO-AEZ model is available online at http://monitorasafra.gppesalq.agr. br/, with results and description in Portuguese for better

Figure 6. In-season monitoring system for soybean in Brazil during 2022 soybean cropping season. In-season soybean yield anomaly released by the National Company Supply (CONAB, 2022). 325 monitoring in January (a and c, three months before harvest) and in May (b and d, after harvest) simulated by FAO-AEZ crop model and









engagement with Brazilian stakeholders. An in-season state time-series visualization of yield anomaly is provided in Figure 7.

Biweekly simulations of the in-season soybean yield anomaly monitoring system shown in

## **Discussion**

The use of homogeneous zones relates to political and demographic census where the concept of small representative samples can represent a population (Schug *et al.*, 2021; Wardrop *et al.,* 2018; Zahnd *et al.*, 2019). In this study, this concept was applied to crop monitoring. With historical soybean yield anomaly from 1991 to 2019, we defined 59 soybean agro-climatic homogenous zones in Brazil. The concept of this analysis is that if a region presents similar year-to-year yield variability (i.e., neighboring counties have similar positive or negative yield anomalies), considering that edaphoclimatic factors representing the regional patterns are relatively identical. This approach has been applied to estimate and forecast national wheat yields in Brazil (Nóia Júnior *et al.*, 2021), and several other studies that quantify crop yield gaps and potential impacts of climate change on agriculture use a similar approach (Antolin *et al*., 2021; Asseng *et al.*, 2015; Marin *et al.*, 2016). However, the use of a single location to represent a larger region also has disadvantages and can generate erroneous estimates for some years, especially due to the spatial variability of extreme weather events that cannot be accounted for.

Crop simulation models and their improvement is a frequent topic in the scientific community (Maiorano *et al.*, 2017; Wang *et al.*, 2022, 2019). Despite this, crop models have proved to be a powerful tool for simulating yield and crop growth, particularly for estimating drought and heat impacts (Battisti *et al.*, 2017; Sampaio *et al.*, 2020). As high temperatures and drought are the main drivers of soybean year-to-year variability in Brazil, the FAO-AEZ crop model showed good overall performance  $(r^2 = 0.59)$ and RMSE =  $402 \text{ kg}$  ha<sup>-1</sup>, varying from 83 kg ha<sup>-1</sup> to 848 kg ha<sup>-1</sup> according to the region in Brazil) to be used as a soybean monitoring system in Brazil (Figures 3 and 4). In comparison, Battisti *et al.* (2020) simulating yield data from IBGE, demonstrated a mean error of 15% (RMSE), whereas our error stands at 402 kg ha<sup>-1</sup> or 11%. These results were obtained using climate data from NASA POWER, which could potentially be more accurate if local weather station data were utilized. However, such stations often suffer from significant missing data and are not evenly distributed geographically, complicating their nationalscale use. Similarly, if the study were based on a smaller regional scale, such as pixel-level analysis instead of using

Figure 7. Regional in-season monitoring for soybean yield anomaly in Brazil during 2022 cropping season. Soybean yield was simulated with the FAO-AEZ crop model. In-season soybean yield anomaly is represented by the red line, black line is the zero line, representing the expected yield (i.e., as the average of the last 5 years). The States are Bahia (BA), Goiás (GO), Maranhão (MA), Minas Gerais (MG), Mato Grosso do Sul (MS), Mato Grosso (MT), Pará (PA), Piauí (PI), Paraná (PR), Rondônia (RO), Rio Grande do Sul (RS), Santa Catarina (SC), São Paulo (SP) and Tocantins (TO).



homogeneous zones, these results might be more precise and better depict subnational variations in soybean yield.

The FAO-AEZ crop model satisfactorily reproduced the spatial distribution of soybean yield anomaly in the 2016 cropping season (Figure 5). In addition, it indicated early signs of extreme soybean yield loss during the 2022 cropping season in January, three months before closing the soybean harvest (Figure 6). In January 2022, official reports from CONAB indicated a good expectation of soybean production with a national estimate of 140 Mt. This number changed in May, with CONAB updating the national soybean production expectation to about 125 Mt.

In-season analysis for the 2022 soybean cropping season showed that the FAO-AEZ crop model overestimated soybean yield anomaly in northeastern Brazil (Figure 6) because excessive rainfall occurred in northeastern Brazil during the 2022 cropping season. The accumulated rainfall in January exceeded 500 mm, twice the normal for the region. This wet condition caused waterlogging and increased plant diseases, delaying soybean harvest, and causing anoxia and soybean seed rot (Dorigatti, 2021). Most of the existing crop simulation models, including the FAO-AEZ crop model, miss waterlogging and plant disease effects on crop growth, as well as the joint effects of many weather extremes (Rötter *et al.*, 2018). The improvement of agricultural monitoring systems performance depends on crop model improvements to include some missing processes. But not only crop simulation models depend on these improvements. Increasingly sophisticated machinelearning models need large, good-quality datasets to properly simulate crop yields (Paudel *et al.*, 2022).

Crop simulation models would also benefit from good quality agricultural big datasets. For example, a limitation of this study was the lack of information about the soybean sowing date in different Brazilian regions before the year 2020. As soybean yield in Brazil highly varies according to the sowing date (Nóia Júnior & Sentelhas, 2019), the performance of the model could have been better evaluated with accurate sowing progression data. In Brazil, CONAB started in 2021 to make available data on sowing date, phenology, and harvest progress over the cropping season at the state level. Even so, data with better spatial resolution would be important for testing crop modeling approaches. All these issues limit the improvement and implementation of crop monitoring and forecasting systems.

A soybean yield monitoring system at a regional resolution and with a crop simulation model may provide information on a spatial and temporal scale not available in the official forecasts (CONAB, 2022). The FAO-AEZ crop model only uses commonly available and public inputs including daily weather data, sowing date, and soil water holding parameters. With the FAO-AEZ crop model and weather data from NASA POWER (Power Project Team, 2022), sowing date from the general reports of CONAB, and soil water holding from national soil surveys (BRASIL, 1981), our soybean monitoring system could be updated on a daily basis to a target grid resolution of approximately a half degree, an improvement over the regional scale tested in this study.

It is important to mention that several studies have tested other methods to predict historical trends or forecast soybean yield in Brazil. Those methods rely mostly on machine learning with remote sensing or weather data, with variable prediction performance. Temporal analysis of vegetation indices obtained from satellite images can be used to explain a major part (60%) of the historical soybean yield variability with relatively moderate errors (Esquerdo et al., 2011; Liu & Kogan, 2002). More recently, the integration of both satellite information and weather data with machine learning has provided an improvement in the prediction capacity for forecasting in-season yield trends of a regional area in south Brazil, making possible the anticipation of potential crop shortages (Schwalbert et al., 2020). In addition, other prediction methods for soybean and other crops based on a time series of national reports employing machine learning were also tested and demonstrated satisfactory prediction performance (Abraham et al., 2020; Monteiro et al., 2022). However, the scaling of these methods over large geographical areas such as the Brazilian territory coupled with a routine that provides timely predictions for near-real-time monitoring may still be challenging due to computational and data restrictions. The approach employed in this study is based on a simple crop model running over representative locations of homogeneous zones and requiring minimal data input to facilitate national-wide monitoring and can be customized to other user needs or geographical extents.

# **Conclusion**

A soybean monitoring system based on a simple crop model has been introduced for Brazil. The inseason analysis for 2022 revealed the model's capacity to anticipate signs of negative yield trends three months in advance compared to national bulletins. National yield estimations are possible at earlier months of the season further improving the prediction performance when approaching the end.

There is the potential to combine this method with other crop simulation models or machine-learning in an ensemble approach to improve yield estimates. We also call for the importance of continuing crop model development and improvement, as well as the increase availability of agricultural big data for further testing any model types to reduce uncertainties. Crop monitoring systems will be increasingly important in a world with increasing variability of crop production due to climate effects.

# **Author Contributions**

All the authors contributed to the conception of the study. R. S. NÓIA JÚNIOR, J. L. SAFANELLI and L. F. DE SOUZA acquisition and analysis of data. R. S. NÓIA JÚNIOR and J. L. SAFANELLI wrote the first draft of the article and L. F. DE SOUZA and D. DOURADO NETO reviewed the article.

# **References**

ABRAHAM, E.R.; MENDES DOS REIS, J.G.; VENDRAMETTO, O.; OLIVEIRA COSTA NETO, P. L. de; CARLO TOLOI, R.; SOUZA, A.E. de; OLIVEIRA MORAIS, M. de. Time Series Prediction with Artificial Neural Networks: An Analysis Using Brazilian Soybean Production. **Agriculture**, v. 10, n. 10, p. 475. 2020. DOI: 10.3390/agriculture10100475

ANTOLIN, L. A.S.; HEINEMANN, A. B.; MARIN, F.R. Impact assessment of common bean availability in Brazil under climate change scenarios. **Agricultural Systems**, v. 191, p. 103174, 2021. DOI: 10.1016/j. agsy.2021.103174

ASSENG, S.; EWERT, F.; MARTRE, P.; RÖTTER, R. P.; LOBELL, D. B.; CAMMARANO, D.; KIMBALL, B. A.; OTTMAN, M. J.; WALL, G. W.; WHITE, J. W.; REYNOLDS, M. P.; ALDERMAN, P. D.; PRASAD, P. V. V.; AGGARWAL, P. K.; ANOTHAI, J.; BASSO, B.; BIERNATH, C.; CHALLINOR, A. J.; DE SANCTIS, G.; ZHU, Y. Rising temperatures reduce global wheat production. **Nature Climate Change**, v. 5, n. 2, p. 143–147, 2015. DOI:10.1038/nclimate2470

BATTISTI, R.; BENDER, F. D.; SENTELHAS, P. C. Assessment of different gridded weather data for soybean yield simulations in Brazil. **Theoretical and Applied Climatology**, v. 135, n. 1–2, p. 237–247, 2019. DOI: 10.1007/s00704-018-2383-y

BATTISTI, R.; FERREIRA, M. D. P.; TAVARES, É. B.; KNAPP, F. M.; BENDER, F. D.; CASAROLI, D.; ALVES JÚNIOR, J. Rules for grown soybean-maize cropping system in Midwestern Brazil: Food production and economic profits. **Agricultural Systems**, v. 182, p. 102850, 2020. DOI: 10.1016/j. agsy.2020.102850

BATTISTI, R.; SENTELHAS, P. C. Improvement of Soybean Resilience to Drought through Deep Root System in Brazil. **Agronomy Journal**, v. 109, n. 4, p. 1612–1622, 2017. DOI: 10.2134/agronj2017.01.0023

BATTISTI, R.; SENTELHAS, Pa. C.; BOOTE, K. J. Inter-comparison of performance of soybean crop simulation models and their ensemble in southern Brazil. **Field Crops Research**, v. 200, p. 28–37, 2017. DOI: 10.1016/j.fcr.2016.10.004

BATTISTI, R.; SENTELHAS, P. C.; BOOTE, K. J.; DE S. CÂMARA, G. M.; FARIAS, J. R. B.; BASSO, C. J. Assessment of soybean yield with altered water-related genetic improvement traits under climate change in Southern Brazil. **European Journal of Agronomy**, v. 83, p. 1–14, 2017. DOI: 10.1016/j.eja.2016.11.004

BATTISTI, R.; SENTELHAS, P. C. Characterizing Brazilian soybeangrowing regions by water deficit patterns. **Field Crops Research**, v. 240, p. 95–105, 2019. DOI: 10.1016/j.fcr.2019.06.007

BOOTE, K. J.; JONES, J. W.; BATCHELOR, W. D.; NAFZIGER, E. D.; MYERS, O. Genetic Coefficients in the CROPGRO–Soybean Model. **Agronomy Journal**, v. 95, n. 1, p. 32–51, 2003. DOI: 10.2134/agronj2003.3200

BRASIL. Ministério das Minas e Energia. Secretaria Geral. **Projeto RADAMBRASIL** [S. l.]: Levantamento de Recursos Naturais, 1981. vol. 25, 29, 31.

CONAB. **National Supply Company**: Agricultural information system. 2022. Available at: https://portaldeinformacoes.conab.gov.br/index. php/safras/safra-serie-historica. Accessed on: 1 Feb. 2022.

DOORENBOS, J.; KASSAM, A. H. **Yield response to water**. Rome: FAO

Irrigation and Drainage, Paper 33, 1979.

DORIGATTI, G. Excesso de chuvas atinge lavouras de soja na Bahia com alagamentos e plantas mortas. 2021. **Notícias Agrícolas**. Available at: https://www.noticiasagricolas.com.br/noticias/soja/304834-excessode-chuvas-atinge-lavouras-de-soja-na-bahia-com-alagamentos-eplantas-mortas.html#.YujkKHZBxaQ. Accessed on: 1 Feb. 2022.

ESQUERDO, J. C. D. M.; ZULLO JÚNIOR, J.; ANTUNES, J. F. G. Use of NDVI/ AVHRR time-series profiles for soybean crop monitoring in Brazil. **International Journal of Remote Sensing**, v. 32, n. 13, p. 3711–3727, 2011. DOI: 10.1080/01431161003764112

FAO. Food and Agriculture Organization of the United Nations. **FAOSTAT statistical database**. 2022. Available at: http://www.fao.org/ faostat/en/#home. Accessed on: 5 Jun. 2022.

GUARIN, J. R.; ASSENG, S.; MARTRE, P.; BLIZNYUK, N.. Testing a crop model with extreme low yields from historical district records. **Field Crops Research**, v. 249, p. 107269, 2020. DOI: 10.1016/j.fcr.2018.03.006

HOOGENBOOM, G.; PORTER, C.H.; SHELIA, V.; BOOTE, K.J.; SINGH, U.; WHITE, J.W.; HUNT, L.A.; OGOSHI, R.; LIZASO, J.I.; KOO, J.; ASSENG, S.; SINGELS, A.; MORENO, L.P.; JONES., J.W. **Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.7.2**. 2018.

IBGE. Brazilian Institute of Geography and Statistics. **Interactive soil maps**. 2022a. Available at: https://mapas.ibge.gov.br/tematicos/solos. Accessed on: 5 Jun. 2022.

IBGE. Brazilian Institute of Geography and Statistics. **Municipal agricultural research**. 2022b. Available at: https://sidra.ibge.gov.br/ pesquisa/pam/tabelas. Accessed on: 1 Feb. 2022.

KASSAM, A. H. **Calculation of Net Biomass Production and Yield of Crops**. Rome: Soil Resources, Management and Conservation Service, Land and Water Development Division, FAO, 1977.

KEATING, B.A; CARBERRY, P.S; HAMMER, G.L; PROBERT, M.E; ROBERTSON, M.J; HOLZWORTH, D; HUTH, N.I; HARGREAVES, J.N.G; MEINKE, H; HOCHMAN, Z; MCLEAN, G; VERBURG, K; SNOW, V; DIMES, J.P; SILBURN, M; WANG, E; BROWN, S; BRISTOW, K.L; ASSENG, S; SMITH, C.J. An overview of APSIM, a model designed for farming systems simulation. **European Journal of Agronomy**, v. 18, n. 3–4, p. 267–288, Jan. 2003. DOI: 10.1016/S1161-0301(02)00108-9

LIMA, M.; SILVA JUNIOR, C. A. da; RAUSCH, L.; GIBBS, H. K.; JOHANN, J. A. Demystifying sustainable soy in Brazil. **Land Use Policy**, v. 82, p. 349–352, 2019. DOI: 10.1016/j.landusepol.2018.12.016

LIU, W. T.; KOGAN, F. Monitoring Brazilian soybean production using NOAA/AVHRR based vegetation condition indices. **International Journal of Remote Sensing**, v. 23, n. 6, p. 1161–1179, 2002. DOI: 10.1080/01431160110076126

MAIORANO, A.; MARTRE, P.; ASSENG, S.; EWERT, F.; MÜLLER, C.; RÖTTER, R. P.; RUANE, A. C.; SEMENOV, M. A.; WALLACH, D.; WANG, E.; ALDERMAN, P. D.; KASSIE, B. T.; BIERNATH, C.; BASSO, B.; CAMMARANO, D.; CHALLINOR, A. J.; DOLTRA, J.; DUMONT, B.; REZAEI, E. E.; … ZHU, Y. Crop model improvement reduces the uncertainty of the response to temperature of multi-model ensembles. **Field Crops Research**, v. 202, p. 5–20, 2017. DOI: 10.1016/j.fcr.2016.05.001

MAPA. SISZARC: **Sistema de Zoneamento Agrícola de Risco Climático**. 2021. Available at: http://sistemasweb.agricultura.gov.br/siszarc/base. action.

MARIN, F. R.; MARTHA, G. B.; CASSMAN, K. G.; GRASSINI, P. Prospects for Increasing Sugarcane and Bioethanol Production on Existing Crop Area in Brazil. **BioScience**, v. 66, n. 4, p. 307–316, 2016. DOI: 10.1093/biosci/ biw009

MONTEIRO, L. A.; RAMOS, R. M.; BATTISTI, R.; SOARES, J. R.; OLIVEIRA, J. C.; FIGUEIREDO, G. K. D. A.; LAMPARELLI, R. A. C.; NENDEL, C.; LANA, M. A. Potential Use of Data-Driven Models to Estimate and Predict Soybean Yields at National Scale in Brazil. **International Journal of Plant Production**, 2022. DOI: 10.1007/s42106-022-00209-0

NENDEL, C.; BERG, M.; KERSEBAUM, K.C.; MIRSCHEL, W.; SPECKA, X.; WEGEHENKEL, M.; WENKEL, K.O.; WIELAND, R. The MONICA model: Testing predictability for crop growth, soil moisture and nitrogen dynamics. **Ecological Modelling**, v. 222, n. 9, p. 1614–1625, 2011. DOI: 10.1016/j.ecolmodel.2011.02.018.

NÓIA JÚNIOR, R. de S.; SAFANELLI, J. L.; SOUZA, L. F. de; MAULE, R. F.; BARRETO, A.; SANTO, L. N. S. dos; PAIXÃO, C. F. C. da; SOUZA, L. R. da S.; SILVA, A. R. da; DOURADO NETO, D. Sistema de monitoramento de produtividade de milho 2a safra no Mato Grosso. **Agrometeoros**, v. 30, 2022. DOI: 10.31062/agrom.v30.e027074

NÓIA JÚNIOR, R. de S.; FRAISSE, C. W.; KARREI, M. A. Z.; CERBARO, V. A.; PERONDI, D. Effects of the El Niño Southern Oscillation phenomenon and sowing dates on soybean yield and on the occurrence of extreme weather events in southern Brazil. **Agricultural and Forest Meteorology**, v. 290, p. 108038, 2020. DOI: 10.1016/j. agrformet.2020.108038

NÓIA JÚNIOR, R. de S.; MARTRE, P.; FINGER, R.; VAN DER VELDE, M.; BEN-ARI, T.; EWERT, F.; WEBBER, H.; RUANE, A. C.; ASSENG, S. Extreme lows of wheat production in Brazil. **Environmental Research Letters**, v. 16, n. 10, p. 104025, 1 Oct. 2021. DOI: 10.1088/1748-9326/ac26f3

NÓIA JÚNIOR, R. de S.; SENTELHAS, P. C. Soybean-maize succession in Brazil: Impacts of sowing dates on climate variability, yields and economic profitability. **European Journal of Agronomy**, v. 103, p. 140–151, 2019. DOI: 10.1016/j.eja.2018.12.008

NÓIA JÚNIOR, R. de S.; SENTELHAS, P. Cesar. Yield gap of the doublecrop system of main-season soybean with off-season maize in Brazil. **Crop and Pasture Science**, v. 71, n. 5, p. 445, 2020. DOI: 10.1071/ CP19372

PASLEY, H. R.; HUBER, I.; CASTELLANO, M. J.; ARCHONTOULIS, S. V. Modeling Flood-Induced Stress in Soybeans. **Frontiers in Plant Science**, v. 11, 2020. DOI: 10.3389/fpls.2020.00062.

PAUDEL, D.; BOOGAARD, H.; DE WIT, A.; VAN DER VELDE, M.; CLAVERIE, M.; NISINI, L.; JANSSEN, S.; OSINGA, S.; ATHANASIADIS, I. N. Machine learning for regional crop yield forecasting in Europe. **Field Crops Research**, v. 276, p. 108377, 2022. DOI: 10.1016/j.fcr.2021.108377

POWER PROJECT TEAM. **NASA Prediction of Worldwide Energy Resources**. 2022. Available at: https://power.larc.nasa.gov/. Accessed on: 5 Jun. 2022.

R CORE TEAM. **R**: A language and environment for statistical computing. 2021. Available at: https://www.r-project.org/.

REICHERT, J. M.; ALBUQUERQUE, J. A.; KAISER, D. R.; REINERT, D. J.; URACH, F. L.; CARLESSO, R. Estimation of water retention and availability in soils of Rio Grande do Sul. **Revista Brasileira de Ciência do Solo**, v. 33, n. 6, p. 1547–1560, Dec. 2009. DOI: 10.1590/S0100- 06832009000600004

RÖTTER, R. P.; HOFFMANN, M. P.; KOCH, M.; MÜLLER, C. Progress in modelling agricultural impacts of and adaptations to climate change. **Current Opinion in Plant Biology**, v. 45, p. 255–261, Oct. 2018. DOI: 10.1016/j.pbi.2018.05.009

SAMPAIO, L. S.; BATTISTI, R.; LANA, M. A.; BOOTE, K. J. Assessment of sowing dates and plant densities using CSM-CROPGRO-Soybean for soybean maturity groups in low latitude. **The Journal of Agricultural Science**, v. 158, n. 10, p. 819–832, 5 Dec. 2020. DOI: 10.1017/ S0021859621000204

SCHUG, F.; FRANTZ, D.; VAN DER LINDEN, S.; HOSTERT, P. Gridded population mapping for Germany based on building density, height and type from Earth Observation data using census disaggregation and bottom-up estimates. **PLOS ONE**, v. 16, n. 3, p. e0249044, 26 Mar. 2021.

DOI: 10.1371/journal.pone.0249044

SCHWALBERT, R. A.; AMADO, T.; CORASSA, G.; POTT, L. P.; PRASAD, P. V. V.; CIAMPITTI, I. A. Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. **Agricultural and Forest Meteorology**, v. 284, p. 107886, 2020. DOI: 10.1016/j.agrformet.2019.107886

SENTELHAS, P. C.; BATTISTI, R.; CÂMARA, G. M. S.; FARIAS, J. R. B.; HAMPF, A. C.; NENDEL, C. The soybean yield gap in Brazil – magnitude, causes and possible solutions for sustainable production. **The Journal of Agricultural Science**, v. 153, n. 8, p. 1394–1411, 24 Nov. 2015. DOI: 10.1017/S0021859615000313. Available at: https://www.cambridge.org/ core/product/identifier/S0021859615000313/type/journal\_article.

SILVA, E. H. F. M. da; ANTOLIN, L. A. S.; ZANON, A. J.; ANDRADE, A. S.; SOUZA, H. A. de; CARVALHO, K. dos S.; VIEIRA, N. A.; MARIN, F. R. Impact assessment of soybean yield and water productivity in Brazil due to climate change. **European Journal of Agronomy**, v. 129, p. 126329, 2021. DOI: 10.1016/j.eja.2021.126329

SILVA, E. H. F. M. da; BOOTE, K. J.; HOOGENBOOM, G.; GONÇALVES, A. O.; ANDRADE JUNIOR, A. S.; MARIN, F. R. Performance of the CSM-CROPGRO-soybean in simulating soybean growth and development and the soil water balance for a tropical environment. **Agricultural Water Management**, v. 252, p. 106929, Jun. 2021. DOI: 10.1016/j. agwat.2021.106929

STEDUTO, P.; HSIAO, T. C.; RAES, D.; FERERES, E. AquaCrop—The FAO Crop Model to Simulate Yield Response to Water: I. Concepts and Underlying Principles. **Agronomy Journal**, v. 101, n. 3, p. 426–437, 2009. DOI: 10.2134/agronj2008.0139s

WANG, E.; BROWN, H. E; REBETZKE, G. J; ZHAO, Z.; ZHENG, B.; CHAPMAN, S. C. Improving process-based crop models to better capture genotype×environment×management interactions. **Journal of Experimental Botany**, v. 70, n. 9, p. 2389–2401, 29 Apr. 2019. DOI: 10.1093/jxb/erz092

WANG, E.; HE, D.; WANG, J.; LILLEY, J. M.; CHRISTY, B.; HOFFMANN, M. P.; O'LEARY, G.; HATFIELD, J. L.; LEDDA, L.; DELIGIOS, P. A.; GRANT, B.; JING, Q.; NENDEL, C.; KAGE, H.; QIAN, B.; EYSHI REZAEI, E.; SMITH, W.; WEYMANN, W.; EWERT, F. How reliable are current crop models for simulating growth and seed yield of canola across global sites and under future climate change? **Climatic Change**, v. 172, n. 1–2, p. 20, 2022. DOI: 10.1007/s10584-022-03375-2

WARDROP, N. A.; JOCHEM, W. C.; BIRD, T. J.; CHAMBERLAIN, H. R.; CLARKE, D.; KERR, D.; BENGTSSON, L.; JURAN, S.; SEAMAN, V.; TATEM, A. J. Spatially disaggregated population estimates in the absence of national population and housing census data. **Proceedings of the National Academy of Sciences**, v. 115, n. 14, p. 3529–3537, 2018. DOI: 10.1073/pnas.1715305115

ZAHND, W. E.; ASKELSON, N.; VANDERPOOL, R. C.; STRADTMAN, L.; EDWARD, J.; FARRIS, P. E.; PETERMANN, V.; EBERTH, J. M. Challenges of using nationally representative, population-based surveys to assess rural cancer disparities. **Preventive Medicine**, v. 129, p. 105812, 2019. DOI: 10.1016/j.ypmed.2019.105812

ZHAO, C.; LIU, B.; XIAO, L.; HOOGENBOOM, G.; BOOTE, K. J.; KASSIE, B. T.; PAVAN, W.; SHELIA, V.; KIM, K. S.; HERNANDEZ-OCHOA, I. M; WALLACH, D.; PORTER, C. H.; STOCKLE, C. O.; ZHU, Y.; ASSENG, S. A SIMPLE crop model. **European Journal of Agronomy**, v. 104, p. 97–106, 2019. DOI: 10.1016/j.eja.2019.01.009

**CITATION** 

NÓIA JÚNIOR, R. S.; SAFANELLI, J. L.; SOUZA, L. F.; DOURADO NETO, D. Soybean yield monitoring system for Brazil with FAO-AEZ crop model. **Agrometeoros**, Passo Fundo, v.32, e027580, 2024.





ISSN 2526-7043 www.sbagro.org.br DOI: http://dx.doi.org/10.31062/agrom.v32.e027580

# Sistema de monitoramento de produtividade de soja para o Brasil com o modelo FAO-AEZ

**Rogério de Souza Nóia Júnior1(\*), José Lucas Safanelli2 , Lucas Fernandes de Souza3 e Durval Dourado Neto4**

 TUM School of Life Sciences, Digital Agriculture. Technical University of Munich, Freising, Germany. Email: rogeriosouzanoia@gmail.com Woodwell Climate Research Center. Falmouth MA, USA. E-mail: jsafanelli@woodwellclimate.org Companhia Nacional de Abastecimeno - CONAB. Brasília, DF. E-mail: lucas.fsouza@agricultura.gov.br Escola Superior de Agricultura "Luiz de Queiroz". Universidade de São Paulo, Piracicaba, SP. E-mail: ddourado@usp.br (\*)Autor para correspondência.

### INFORMAÇÕES RESUMO

### **História do artigo:**

Recebido em 01 de fevereiro de 2024 Aceito em 16 de julho de 2024

**Termos para indexação:**  monitoramento de produtividade de culturas segurança alimentar *Glycine max* (L.) Merr. anomalia de produtividade

# A variação anual na produtividade da soja no Brasil impacta diretamente a economia nacional e o abastecimento de alimentos. Este estudo propôs o desenvolvimento de um sistema de monitoramento da produtividade da soja baseado no modelo FAO Agro-Ecológico (FAO-AEZ), previamente validado e calibrado. O modelo foi utilizado para simular a produtividade da soja em 59 zonas homogêneas agroclimáticas em todo o Brasil. A avaliação do desempenho do modelo foi feita comparando as produtividades simuladas com dados observados de 2011 a 2020, corrigidos por tendências temporais. Além disso, foi realizada uma análise durante as safras de 2021 e 2022, monitorando a variação da produtividade em relação às cinco safras anteriores com as mesmas datas de semeadura. Os resultados indicaram um bom desempenho do modelo, com um coeficiente de determinação de 0,59 e um Erro Quadrático Médio de 402 kg ha-1 em comparação com os dados nacionais. Notavelmente, o modelo conseguiu antecipar sinais de tendências negativas na produtividade com três meses de antecedência em relação aos boletins nacionais. Este sistema de monitoramento utiliza dados meteorológicos diários e parâmetros simples de manejo de cultivo, oferecendo uma abordagem flexível e eficaz para estimativas de produtividade da soja, com potencial para adaptação a diferentes necessidades e contextos.

© 2024 SBAgro. Todos os direitos reservados.

NÓIA JÚNIOR, R. S.; SAFANELLI, J. L.; SOUZA, L. F.; DOURADO NETO, D. Soybean yield monitoring system REFERENCIAÇÃO NÓIA JÚNIOR, R. S.; SAFANELLI, J. L.; SOUZA, L. F.; DOURADO NETO, D. Soybean yield m<br>for Brazil with FAO-AEZ crop model. **Agrometeoros**, Passo Fundo, v.32, e027580, 2024.