The Impact of Gridded Weather Database on Soil Water Availability in Rice Crop Modeling

Caio Teodoro Menezes

Federal University of Goiás (UFG)

Derblai Casaroli (✉ derblai@ufg.br)

Federal University of Goiás: Universidade Federal de Goiás <https://orcid.org/0000-0001-8041-0066>

Alexandre Bryan Heinemann

Santo Antônio de Goiás, Goiás, Brazil

Vinicius Cintra Moschetti

Jundiaí, São Paulo, Brazil

Rafael Battisti

UFG, Goiânia

Research Article

Keywords: ORYZA (v3), Daily Gridded, Virtual weather station, NASA/POWER, Crop modeling

Posted Date: August 25th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-355960/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Theoretical and Applied Climatology on January 6th, 2022. See the published version at <https://doi.org/10.1007/s00704-021-03906-4>.

Abstract

In recent years, there has been an increase in studies suggesting that gridded weather database (GWD) is a suitable source for simulating crop yield. Brazil has low geospatial coverage by measured weather database (MWD). Based on that, this study aimed to compare two different GWD sources, Daily Gridded (DG) and NASA/POWER (NP), on the simulated yield of upland rice (UR) against the MWD input. The GWD and MWD were obtained for seven locations across UR Brazilian region, considering a period ranging from 1984 to 2016. GWD and MWD were used to estimate rice potential (Y_p) and attainable yield (Y_a), in clay soil and sandy soil, using ORYZA (v3) model. DG had the best performance for all variables. GWD-based yields had a reasonable performance. However, DG had a slightly better performance than NP in all conditions, DG-based yields showed RMSE values of 0.57, 0.71 and 0.52 for Y_p and Y_a in clay and sandy soil, whereas NP showed RMSE values of 0.86, 0.91 and 0.64. DG also showed higher R^2 and d values for yields assessed. Both GWD overestimated Y_a , these overestimations in DG-based yield were 3.54, 9.61, and 21.35% for Y_p and Y_a in clay and sandy soil respectively, in NP-based yield were 13.67, 18.45, 29.11%, showing that for both GWD-based yield increased as the soil type texture as well as water storage decreased. As a consequence, we do not recommend the use of precipitation data in daily time-step crop modeling.

Introduction

Mathematical models that simulate growth and yield of agricultural crops are important tools for decision making (Caetano and Casaroli 2017; Ferreira et al. 2019; Battisti et al. 2020; Caetano et al. 2021). The use of estimated weather data has been widely used, improving its degree of accuracy more and more (Jha et al. 2019; Dubrovsky et al. 2020). Thus, these data can be used in different applications, such as, to obtain yield gap in agrometeorological models (Santos et al. 2021; Paixão et al. 2021).

Useful and reliable weather datasets are a time-saving tool for crop modeling purposes. The absence of measured weather data (MWD) and a proper spatial resolution are impediments to predicting both, current and future effect of climate on crop yields. Therefore, gridded weather data (GWD) has become a welcomed alternative weather data source for agricultural purposes. GWD has the advantage of complete geospatial cover, thus becoming an option for regions where the MWD has a lower spatial distribution or incomplete data (Paixão et al. 2021).

There are two main sources of MWD in Brazil, the Brazilian National Institute of Meteorology (INMET), which provides MWD from automatic and conventional weather stations, and the Brazilian Water Agency (ANA), which has a rainfall station network across Brazilian territory. Both sources are available in low spatial resolution (Xavier et al. 2016). Central Brazil, the main region for grains and pulses farming (CONAB 2020), is situated in three river basins: Tocantins, Amazon, and Paraná. MWD resolutions are lower for Amazon and Tocantins river basins when compared with other parts of Brazil (Xavier et al. 2016). Currently, this is the greatest limitation for crop modeling in upland rice region (Heinemann et al. 2008).

Crop model simulation for application and agrometeorological purposes is made easier by GWD. It has been applied worldwide to simulate crop yields (Wart et al. 2013; Ruane et al. 2015; Mourtzinis et al. 2017; Jha et al. 2019; Müller et al. 2019; Paixão et al. 2021). Amongst worldwide available GWD for crop modeling purposes, it can be pointed out the NASA/POWER (NP). This dataset has a $0.5^\circ \times 0.5^\circ$ spatial resolution at a daily temporal resolution comprising most variables used in crop modeling. This dataset is updated weekly with a time series available since 1983.

In Brazil, there is an interpolated weather database denominated Daily Gridded (DG) (Xavier et al. 2016), which provides an agrometeorological daily and monthly grid data with two spatial resolution for the years 1980 until 2017. The highest resolution available – $0.1^\circ \times 0.1^\circ$, contains data for maximum and minimum air temperature, and the lowest resolution – $0.25^\circ \times 0.25^\circ$, contains data for the other variables: reference evapotranspiration, global solar radiation, rainfall, relative humidity, and wind speed. DG has been applied for wide variety of crops, such as soybean (Battisti and Sentelhas 2019; Junior and Sentelhas 2019; Teixeira et al. 2019), sugarcane (Monteiro et al. 2017; Cesconetto et al. 2018; Perin et al. 2019; Paixão et al. 2020), maize (Andrea et al. 2019), and bamboo (Battisti et al. 2019).

Nevertheless, there are still uncertainties related to the accuracy for its application in crop model simulation. GWD has been accurate in applications related to temperature, but uncertain in estimates related to precipitation. Those uncertainties may generate misleading water deficit and crop yield outcomes (Mourtzinis et al. 2017).

For this study, we selected ORYZA (v3) crop model (Li et al. 2017). The model predicts rice growth and yield as influenced by local environmental conditions, agronomic practices, and cultivar traits. Its strong ability to quantify the influence of soil water on rice growth and yield (Bouman and Laar 2006; Feng et al. 2007), allows the evaluation of rice cultivars' response under drought stress (Li et al. 2013). ORYZA (v3) has been used to quantify rice yield gaps at global, national, and regional scales (Boling et al. 2010; Laborte et al. 2012; Espe et al. 2016), to provide a cost-effective method for drought-tolerance screening (Li et al. 2013), and for determining drought-stress profiles for upland rice in Brazil under current (Heinemann et al. 2015) and future climates (Ramirez-Villegas et al. 2018).

Based on the current literature, we hypothesizes that GWD can be used for simulating rice yield using crop growth models in the areas of upland rice. This way, the study aims: (1) to compare measured weather data with two sources of gridded weather data (Daily Gridded and NASA/POWER); and (2) to evaluate the impact of measured and gridded weather data on the potential (Y_p) and attainable (Y_a) in upland rice yield, considering two soil types – clay and sandy.

Materials And Methods

Locations and weather databases

The measured weather data was obtained from seven locations (Table 1) across upland rice production region in Brazil. The data for six weather stations were obtained from the Brazilian Meteorological Service

(INMET) and one from the Brazilian Agricultural Research Corporation (EMBRAPA). Missing data in the MWD was filled out using the day of year from the climate normal. The percentage of missing data for each weather station is in Table 1. The locations are classified as tropical by Köppen-Geiger method, where Porto Nacional and Altamira are classified as Am climate type, and the other five locations as Aw (Alvares et al. 2013).

The two gridded weather databases (GWD) and its variables applied in this study are: i) Daily Gridded (Xavier et al. 2016), which has daily variables at $0.1^\circ \times 0.1^\circ$ for maximum and minimum air temperature, and $0.25^\circ \times 0.25^\circ$ spatial resolution for rainfall and solar radiation; and ii) NASA/POWER (National Aeronautics and Space Administration's POWER, 2020) which has daily variables, including maximum and minimum air temperature, rainfall and solar radiation at scale of $0.5^\circ \times 0.5^\circ$ horizontal resolution.

A historical daily weather data of maximum and minimum air temperature, incoming solar radiation ($\text{MJ m}^{-2} \text{ d}^{-1}$), and precipitation (mm d^{-1}) was obtained for the years 1984 until 2016 from both gridded datasets. The solar radiation of measured dataset was estimated by Angstrom-Prescott (Angstrom 1924; Prescott 1940) from sunshine hours because only weather station S7 contains a radiometer (Table 1).

Crop model simulations

The rice crop modeling ORYZA (v3) (Bouman et al. 2001) was used to assess the potential (Y_p) and attainable (Y_a) upland rice yield. ORYZA (v3) operates on a daily time-step; hence, it requires as input daily weather data such as rainfall, global solar radiation, and maximum and minimum air temperature. ORYZA (v3) is a process-based rice simulation model developed for a wide variety of applications in rice research (Bouman and Laar 2006; Li et al. 2013) e.g., puddled or non-puddled, with free or impeded drainage at a given depth in the soil profile. The 'PADDY' module is described in detail by Boling et al. (2007). The same study demonstrated that it is suitable for rainfed conditions.

The cultivar crop model parameters were taken from upland rice BRS Primavera cultivar. This is a short-cycle cultivar, with cycle 100 days from emergence to physiological maturity, and it is still used as cultivar check in the upland rice breeding program. Its parameters were obtained from a previous study (Heinemann et al. 2015) and they have been applied in many upland rice crop modeling applications in Brazil (Heinemann et al. 2015; Ramirez-Villegas et al. 2018; Heinemann et al. 2019). For all locations (see Table 1), the simulation start date was set 225 days prior to the sowing date in order to establish realistic soil water profiles and the sowing date was set in November 1st. Seed density was set at 200 seeds m^{-2} . A sandy and a clay soil type were used in the simulation. Classified as oxisol on the American classification system, both are predominant in the study region (Heinemann et al. 2015). Further parameters are described in Table S1.

Statistical analysis

The statistical indices used were mean error (ME), mean absolute error (MAE), root mean square error (RMSE), agreement index (d) (Willmot 1981), Pearson's correlation coefficient (r), and coefficient of

determination (R^2). This was applied to compare minimum e maximum air temperature, rainfall, solar radiation in the crop cycle as well as potential (Y_p) and attainable (Y_a) yield.

The sequence of graphics used in this study were empirical cumulative probability function (ECDF), scatter plot, regression lines, violin plot and line chart.

Results

Descriptive statistics to measured and gridded weather data

An empirical cumulative probability function (ECDF) was plotted (Fig. 1) to assess the four variables from the datasets in the crop cycle. ECDF for daily solar radiation (S_r) and maximum air temperature (T_{max}) in Daily Gridded (DG) data showed a higher similarity to the measured weather data (MWD) than NASA/POWER (NP) (Fig. 1a and c). NP only exhibits a higher similarity to MWD regarding minimum air temperature (T_{min}) (Fig. 1d). Both gridded weather data (GWD) had similar distributions to each other but with a higher degree of distinction to MWD (Fig. 1b). NP underestimates T_{max} in the range of 24.5°C to 35.8°C in about 95% of the times, which represents 6% of underestimation in this range; meanwhile DG T_{max} data is more similar to MWD for the ECDF (Fig. 1a). NP overestimated S_r in the range of 11 to 32 MJ m⁻² d⁻¹ in about 89% of the cases (Fig. 1d), it represents 5% of overestimation.

DG overestimates MWD rainfall in 73% for the cumulative events in the range of 0.1 to 13.36 mm; from 8% to 81% of measured events (Fig. 1c), it represents 84.6% of overestimation in this range. On average, this overestimation is about 163 mm cycle⁻¹. NP overestimates MWD rainfall in 73.7% of cumulative events in the range of 0.1 to 11.8 mm – from 5.4 to 79.1% of measured events (Fig. 1c), which results in a 95% of overestimation in this range. On average, this overestimation is about 160 mm cycle⁻¹. This trend takes place in both GWD from all assessed municipalities (Fig. 2).

In comparison, MWD precipitation starts after about 38% of the cumulative events. From 80 and 79% of the events onwards, for DG and NP, respectively, GWD underestimates extreme rainfall events (Fig. 1c). On average, it represents 27 and 31% of underestimation. This underestimation is about 165 and 196 mm cycle⁻¹ for DG and NP, respectively. DG underestimates total rainfall by 1.6 mm cycle⁻¹ (0.2%), whereas NP underestimates it by 36 mm cycle⁻¹ (4.5%).

Scatter plot with commonly employed indexes were also employed for comparison to other studies. DG data shows superior performance against NP data for all weather variables assessed (Fig. 3). T_{max} is the variable with best agreement and fit to MWD in DG dataset, showing coefficient of determination (R^2), root mean square error (RMSE) and agreement index (d) of 0.77, 1.34, and 0.93 against 0.46, 2.87, and 0.74 from NP, respectively (Fig. 3a and b). NP shows a tendency to underestimate T_{max} , thus exhibiting higher ME value than DG, -1.84 against -0.3, respectively (Fig. 1a and 3b).

T_{min} shows overall lower agreement with MWD from both GWD when compared to T_{max} . In this case, DG also shows better agreement and fit than NP, with R^2 , RMSE and d values of 0.6, 1.19, and 0.86

against 0.47, 1.49, and 0.83 from NP (Fig. 3c and d). Rainfall has the worst performance amongst all variables in both GWD when compared with MWD. DG shows R^2 , RMSE, and d values of 0.42, 10.7, and 0.75 whereas NP showed 0.11, 13.67, and 0.53, respectively.

Crop modeling outputs

All locations are assessed as an aggregate for statistical analysis (Fig. 4 and 5) and each location is assessed individually in Table 2. GWD show a better performance simulating Yp than Ya, but NP only exhibited better performance when all locations were aggregated (Table 2). Both GWD-based physiological maturity simulated days showed good agreement, with R^2 of 0.66 and 0.74, and RMSE of 1.93 and 2.85 for DG and NP respectively (Fig. 5).

Porto Nacional is the location with overall better performance for both GWD, R^2 of 0.65, 0.59 and 0.61, RMSE 0.29, 0.73 and 0.57 for each DG condition (Yp, Ya clay and Ya sandy) respectively, and R^2 of 0.17, 0.67 and 0.67 and RMSE of 0.63, 0.79 and 0.60 for each NP condition respectively. The worse agreement in NP Yp condition in relation to Ya condition also strengthens the fact of lower NP agreement in Tmax, Tmin and RS (Fig. 1 and 2) since these variables define Yp. Rio Branco is the city with worse performance in both GWD, except for DG Ya clay soil condition.

Discussion

Weather variables

This is the first study to evaluate the rainfall distribution pattern created by gridded weather database (GWD) considering its impact on soil condition in crop modeling. These patterns of GWD rainfall are of critical importance, prone to generate mistaken attainable yields. There are studies related to GWD rainfall estimation uncertainties in amount and distribution, such as Ruane et al. (2015), Xavier et al. (2016) and Mourtzinis et al. (2017). Battisti et al. (2019) showed that Daily Gridded (DG) underestimated rainfall by 21.17 mm cycle⁻¹ in the soybean crop, they also found good agreements between DG and MWD for Tmax and Tmin, reporting an r of 0.92 and RMSE of 0.72 for Tmax, and r of 0.87 and RMSE of 1.05 for Tmin.

The RMSE observed in DG rainfall data (Fig. 3) agrees with the ones reported by Xavier et al. (2016) for Amazon and Tocantins river basins, with values of 13.17 and 10.54, respectively. Lower ME values from DG agrees with previous statements, given that the dataset has a lower underestimation of total rainfall than NASA/POWER (NP). Both GWD has a good RS, agreement with MWD, DG has a slightly better agreement, showing R^2 , RMSE, and d of 0.65, 3.09, and 0.9, respectively, against 0.52, 4.09, and 0.84 from NP.

Simulated yield and phenology

Battisti et al. (2019) found a better performance at simulating attainable (Y_a) and potential (Y_p) yield using DG datasets for simulating soybean crop modeling. In our study, we found that DG has a better performance when simulating Y_p and Y_a than NP (Table 2, Fig. 4 and 5). Since potential yield depends only on maximum (T_{max}) and minimum (T_{min}) air temperature, as well as solar radiation (S_r), DG also showed better agreement than NP. Because Y_a has rainfall as input, the most uncertain variable, it is expected a reduction in its agreement to MWD ones.

The ORYZA (v3) crop model phenology is defined by air temperature and photoperiod (Bouman et al. 2001). Errors in GWD-based simulations for these variables are low enough not to cause high-magnitude errors for this estimation (Fig. 2 and 5). In this case the phenology is the same for all soil conditions, since drought don't change rice phenology in the simulation.

Weather variables play an important role in growth and crop yield (Sridevi and Chellamuthu 2015). Tropical rice has an optimal temperature range between 25 and 35°C. High temperatures, above 35°C (Hussain et al. 2019) for most cultivars or 36.6°C for the studied cultivar (Heinemann et al. 2015), have a negative impact on growth and pollination, leading to spikelet sterility. Likewise, low temperature, below 25°C, may cause delays in phenological stages (Hussain et al. 2019). Both situations decrease productivity.

Soil water availability effect on attainable yield

In our study, it is observed that GWD-based simulations have a closer result to Y_p , meanwhile GWD overestimates Y_a , when compared to MWD-based simulations (Fig. 4). GWD overestimates Y_a as a consequence of misestimating the daily rainfall amount and distribution (Fig. 1 and 3) (Wart et al. 2013; Mourtzinis et al. 2015; Xavier et al. 2016).

In general, this overestimation took place mainly in the distribution range of 0.1 to 12 mm, which accounts for around of 70% of the times observed (Fig. 3). On the other hand, extreme rainfall events ($> 50 \text{ mm d}^{-1}$) are underestimated by GWD (Fig. 1, 2 and 3). Therefore, gridded weather data ends up simulating an irrigation-sheet-like rainfall pattern in the range of 0.1 to 12mm, an overestimation by GWD based rainfall. This rainfall pattern is incompatible with measured weather data, resulting in increased productivity for all soils. As a consequence of soil characteristics, its overestimations increase as soil textures become coarser. It is possible to visualize this difference in the cumulative evapotranspiration (ETCUM) (Fig. 6).

Both regression lines from GWD and MWD showed that in both soil conditions, GWD overestimates MWD ETCUM, with higher overestimations in sandy soil condition (Fig. 4e and f). It is noticeable that cumulative evapotranspiration (ETCUM) from GWD tends to have increased overestimation in sandy soils, even though the accumulated rain is similar (Fig. 6e, f, i, and j). The violinplot (Fig. 4i and j) shows MWD ETCUM values ranging at lower values than GWD ones. The opposite pattern is seen in accumulated drainage (DRAIN) (Fig. 6k and l). In this regard, GWD DRAIN values are ranging in values lower than MWD ones, showing that irrigation-sheet-like rain pattern generated by GWD reduces total

drainage, a consequence of an underestimation of high-precipitation events and an overestimation of precipitation values in the range of 0.1 to 12 mm.

According to Bouman et al. (2001), when drought occurs in the ORYZA (v3) crop model, solar radiation absorption is negatively affected due to leaf rolling, reduced leaf expansion rate, and changed assimilate partitioning. Hence the reduced PARCUM in sandy soil conditions. Li et al. (2017) stated that photosynthesis for non-drought-tolerant cultivars linearly decreases as the soil water content decreases, the yield for drought-tolerant cultivars would be mildly impacted in a mild drought.

It is observed that DG and NP-based crop simulations overestimation YP in 3.5% and 13.7%, respectively. Meanwhile, for Ya in clay soil condition, there was 9.6% and 18.4% yield overestimation, with 67% and 75% of total events overestimating MWD, for DG and NP respectively. The widest differences between GWD and MWD-based simulations are observed for Ya in sandy soil condition, where there are 21.3% and 29.1% yield overestimation, with 84% and 85% of total events overestimating MWD for DG and NP datasets. These overestimation percentages are based on the mean errors and the average yield (Table 2, Fig. 4 and 7).

In Yp conditions, NP-based simulations overestimate MWD-based simulations in a wider productivity range, from 2.5 to 6.5 Mg ha⁻¹ (Fig. 7). NP overestimates MWD in about 20% of the simulations before DG, overestimating MWD ones in the range of 2.5 to 6.5 Mg ha⁻¹, whereas DG overestimates in the range of 3 to 5 Mg ha⁻¹ Fig. 7a. In Ya clay soil condition, both GWD overestimate MWD around the same range, from 2 to 4 Mg ha⁻¹ Fig. 7b. In Ya sandy soil condition as in occurred before, both GWD overestimate in about the same range of conditions, ranging from 1 to about 2.5 Mg ha⁻¹ Fig. 7c. It is noteworthy to point out that the graph curve is further away in relation to the last case (Fig. 7a, b, and c). Overestimation worsens as soil water storage capacity decreases.

In clay soil condition, only two DG GWD municipalities show higher values of attainable yield than MWD (Paracatu and Rio Branco), NP did not show any underestimation at clay soil condition, whereas both GWD did not show underestimation at sandy soil conditions (Table 2), which are less prone to happen. We also divided the 32 years' time series in two in order to assess the temporal impact on crop model (Table 3), its possible to notice that DG overestimate even more for all conditions in the second part of time series, in this period DG used more rain gauge in interpolations than the first period assessed (Xavier et al. 2016). NP overestimated more in the second period for Yp condition and underestimated for Ya in sandy soil condition (Table 3).

Similar yield trends are observed for all simulated conditions along 32 seasons (Fig. 7d, e, and f). Soybean simulations by Battisti et al. (2019), based on GWD and MWD, also found the same trends along sowing years in simulating attainable yield. These similar systematic errors may indicate model sensitivity on capturing climate effects on yield variability (Pirtiojja et al. 2015).

Ya was overestimated by the model running the GWD, similar to what has been found in other studies. Battisti et al. (2019), simulations of soybean yield with DG dataset, found the ME for Yp and Ya of 69 kg ha⁻¹ and 178 kg ha⁻¹, respectively. It was concluded that the DG data estimates soybean yield within acceptable error boundaries. However, this study used a soil with 0.162 cm³ cm⁻³ soil water storage, classified as high soil water storage (Jensen and Allen 2016), and set the root growth to hit maximum values of 120 cm, possibly smoothing the errors related to GWD precipitation rain pattern. Van Wart et al. (2015) found a ME for Ya of about 10% on maize, wheat, and rice simulations in several countries using the NP dataset. There is no study that analyzed the interaction between GWD rain pattern and soil water storage.

Limitations and future work

Although current findings point out an acceptable correspondence between GWD and MWD on crop model for rice, several limitations must be considered. GWD are known to be less accurate in regions where MWD are less available or do not exist (Wart et al. 2013; Xavier et al. 2016). These uncertainties increase as virtual weather stations get further away from measured weather stations (Van Wart et al. 2015). Second, GWD does not estimate daily MWD rainfall distribution with accuracy. Therefore, we suggest caution when using GWD as an input in crop models, especially in sandy soils, since they tend to have lower water storage, which is prone to worse attainable yield overestimation in upland rice. We assume this may happen in other crops as well, but there are no studies supporting this hypothesis. Third, studies comparing crop performance among different regions with distinct soil profiles may generate misleading results. Lastly, according to Li et al. (2017), ORYZA (v3) crop model has coupled nitrogen uptake with water uptake, hence affecting nitrogen studies with ORYZA (v3) that relies on GWD.

Conclusion

This study has shown that both Daily Gridded and NASA/POWER may be suitable weather data sources for simulating upland rice (UR) potential yield (Yp). Despite this accuracy, attainable yield (Ya) modeling should be avoided, especially short-cycle cultures on daily-step model simulations. These uncertainties are even higher at sandy soil conditions since it has lower soil water storage. Further studies using the GWD dataset for other crops in a wide range of locations and soils could improve its estimation capabilities while defining its acceptable use range.

Declarations

Acknowledgments

The authors would like to thank: The National Council for Scientific and Technological Development (CNPq) for the support to this study through a research fellowship for the first author.

Conflicts of interest

The authors declare no conflicts of interest.

Funding Statement

Not applicable.

Author's Contribution

All authors contributed equally to all sections of the manuscript. All authors read and approved the final manuscript.

Availability of data and material

Data can be made available if necessary.

Code availability

Not applicable

Ethics approval

Not applicable

Consent to participate

All authors agreed to participate in the manuscript.

Consent for publication

All authors agree with the publication of the manuscript in Theoretical and Applied Climatology.

References

1. Alvares CA, Stape JL, Sentelhas PC, de Moraes Gonçalves JL, Sparovek G (2013) Köppen's climate classification map for Brazil. *Meteorol Z* 22:711–728. DOI:10.1127/0941-2948/2013/0507
2. Andrea MSC, Dallacort R, Barbieri JD, Tieppo RC (2019) Impactos de Predições Climáticas Futuras sobre o Milho de Segunda Safra em Agrossistema em Região de Transição de Bioma no Estado do Mato Grosso. *Rev bras meteorol* [online]. 34:335–347. <https://doi.org/10.1590/0102-77863340241>
3. Angstrom A (1924) Solar and terrestrial radiation. *Q J R Meteorol Soc* 50: 121–125. <https://doi.org/10.1002/qj.49705021008>
4. Battisti R, Bender FD, Sentelhas PC (2019) Assessment of different gridded weather data for soybean yield simulations in Brazil. *Theor Appl Climatol*. 135:237–247. <https://doi.org/10.1007/s00704-018-2383-y>

5. Battisti R, Casaroli D, Júnior JA, Evangelista AWP, Mesquita M (2019) Agro-climatic zoning of bamboo as a support for crop farming in the central-north region of the Brazilian Savannah. *Pesqui Agropecu Trop* 49. <https://doi.org/10.1590/1983-40632019v4952794>
6. Battisti R, Ferreira MDP, Tavares EB, Knapp FM, Bender F, Casaroli D, Alves Junior J (2020) Rules for grown soybean-maize cropping system in Midwestern Brazil: Food production and economic profits. *Agricultural Systems* 182:102850. <https://doi.org/10.1016/j.agsy.2020.102850>
7. Battisti R, Sentelhas PC (2019) Characterizing Brazilian soybean-growing regions by water deficit patterns. *Field Crop Res* 240:95–105. <https://doi.org/10.1016/j.fcr.2019.06.007>
8. Bouman BAM, Kropff MJ, Tuong TP, Wopereis MCS, Ten Berge HFM, Van Laar HH (2001) *ORYZA2000: Modeling Lowland Rice*. Los Baños (Philippines): International Rice Research Institute, and Wageningen, Netherlands: Wageningen University and Research Centre. 235 pp
9. Bouman BAM, Van Laar HH (2006) Description and evaluation of rice growth model ORYZA 2000 under nitrogen limited conditions. *Agric Syst* 87:249–273. <https://doi.org/10.1016/j.agsy.2004.09.011>
10. Boling AA, Bouman BAM, Tuong TP, Murty MVR, Jatmiko SY (2007) Modelling the effect of groundwater depth on yield-increasing interventions in rainfed lowland rice in Central Java, Indonesia. *Agric Syst* 92:115–139. <https://doi.org/10.1016/j.agsy.2006.05.003>
11. Boling AA, Tuong TP, Keulen H, Bouman BAM, Suganda H, Spiertz JHJ (2010) Yield gap of rainfed rice in farmers' fields in Central Java, Indonesia. *Agric Syst* 103:307–315. <https://doi.org/10.1016/j.agsy.2010.02.003>
12. Caetano JM, Alves Júnior J, Casaroli D, Evangelista AWP (2021) Estimated productivity of sugarcane through the Agro-Ecological Zone method. *Revista Ceres* 68:1-9. <http://dx.doi.org/10.1590/0034-737x202168010001>
13. Caetano JM and Casaroli D (2017) Sugarcane yield estimation for climatic conditions in the state of Goiás. *Revista Ceres* 64:298-306. <http://dx.doi.org/10.1590/0034-737x201764030011>
14. Santos TG, Battisti R, Casaroli D, Alves Júnior J, Evangelista AWP (2021) Assessment of agricultural efficiency and yield gap for soybean in the Brazilian Central Cerrado biome. *Bragantia* 80:e1821. <http://dx.doi.org/10.1590/1678-4499.20200352>
15. Cesconetto LB, Pruski FF, Rodriguez RDG, Marcatti GE (2019) Potentiality of sugarcane expansion under irrigation conditions considering natural and potential water availability. *Agric Water Manag* 203:162–171. <https://doi.org/10.1016/j.agwat.2018.03.008>
16. CONAB (2019) Survey of crop season: rice. <http://www.conab.gov.br/conteudos.php?a=1253&>. Accessed 20 Jun 2019
17. Dubrovsky M, Huth R, Dabhi H (2020) Parametric gridded weather generator for use in present and future climates: focus on spatial temperature characteristics. *Theor Appl Climatol* 139:1031–1044. <https://doi.org/10.1007/s00704-019-03027-z>
18. Espe MB, Cassman KG, Yang H, Guilpart N, Grassini P, Wart J, Anders M, Beighley D, Harrell D, Linscombe S, McKenzie K, Mutters R, Wilson LT, Linquist BA (2016) Yield gap analysis of US rice

- production systems shows opportunities for improvement. *Field Crops Res* 196:276–283.
<https://doi.org/10.1016/j.fcr.2016.07.011>
19. Feng L, Bouman B, Tuong T, Cabangon R, Li Y, Lu G, Feng Y (2007) Exploring options to grow rice using less water in northern China using a modelling approach I. Field experiments and model evaluation. *Agric Water Manag* 88(1–3):1–13. <https://doi.org/10.1016/j.agwat.2006.10.006>
 20. Ferreira L, Alves Júnior J, Battisti R, Heinemann AB, Leite CV, Evangelista AWP, Casaroli D (2019) Assessment of economic returns by using a central pivot system to irrigate common beans during the rainfed season in Central Brazil. *Agric Water Manag* 224:105749.
<https://doi.org/10.1016/j.agwat.2019.105749>
 21. Heinemann AB, Dingkuhn M, Luquet D, Combres J, Chapman S (2007) Characterization of drought stress environments for upland rice and maize in central Brazil. *Euphytica* 162(3):395–410.
<https://doi.org/10.1007/s10681-007-9579-z>
 22. Heinemann AB, Barrios-Perez C, Ramirez-Villegas J, Arango-Londoño D, Bonilla-Findji O, Medeiros JC, Jarvis A (2015) Variation and impact of drought-stress patterns across upland rice target population of environments in Brazil. *J Exp Bot* 66:3625–3638. <https://doi.org/10.1093/jxb/erv126>
 23. Heinemann AB, Ramirez-Villegas J, Rebolledo MC, Neto GMFC, Castro AP (2019) Upland rice breeding led to increased drought sensitivity in Brazil. *F Crop Res* 231:57–67.
<https://doi.org/10.1016/j.fcr.2018.11.009>
 24. Heinemann AB, Ramirez-Villegas J, Stone LF, Silva AP, Matta DH, Diaz MEP (2020) The impact of El Niño Southern Oscillation on cropping season rainfall variability across Central Brazil. *Int J Climatol* 1:1–22. <https://doi.org/10.1002/joc.6684>
 25. Hussain S, Khaliq A, Ali B, Hussain S (2019) Temperature extremes: impact on rice growth and development. In: *Plant abiotic stress tolerance agronomic, molecular and biotechnological approaches*. Springer, Berlin, pp 153–171. <https://doi.org/10.9787/PBB.2018.6.1.1>
 26. INMET. The Brazilian National Institute of Meteorology (2019) Available at:
<http://www.inmet.gov.br/portal/> Accessed 20 may 2019
 27. Jha PK, Athanasiadis P, Gualdi S, Trabucco A, Mereu V, Shelia V, Hoogenboom G (2019) Evaluating the applicability of using daily forecasts from seasonal prediction systems (SPSSs) for agriculture: a case study of Nepal's Terai with the NCEP CFSv2. *Theor Appl Climatol* 135:1143–1156.
<https://doi.org/10.1007/s00704-018-2433-5>
 28. Laborte AG, de Bie KC, Smaling EM, Moya PF, Boling AA, Van Ittersum MK (2012) Rice yields and yield gaps in Southeast Asia: past trends and future outlook. *Eur J Agron* 36:9–20.
<https://doi.org/10.1016/j.eja.2011.08.005>
 29. Li T, Raman AK, Marcaida M III, Kumar A, Angeles O, Radanielson M (2013) Simulation of genotype performances across a larger number of environments for rice breeding using ORYZA2000. *Field Crops Res* 149:312–321. <https://doi.org/10.1016/j.fcr.2013.05.006>
 30. Li T, Angeles O, Marcaida M III, Manalo E, Manalili MP, Radanielson A, Mohanty S (2017) From ORYZA2000 to ORYZA (v3): an improved simulation model for rice in drought and nitrogen-deficient

- environments. *Agric For Meteorol* 237:246–256. <https://doi.org/10.1016/j.agrformet.2017.02.025>
31. Monteiro LA, Sentelhas PC, Pedra GU (2018) Assessment of NASA/ POWER satellite-based weather system for Brazilian conditions and its impact on sugarcane yield simulation. *Int J Climatol* 38: 1571–1581. <https://doi.org/10.1002/joc.5282>
 32. Mourtzinis S, Edreira JIR, Conley SP, Grassini P (2017) From grid to field: assessing quality of gridded weather data for agricultural applications. *Eur J Agron* 82:163–172. <https://doi.org/10.1016/j.eja.2016.10.013>
 33. Paixão JS, Casaroli D, Battisti R, Evangelista AWP, Alves Júnior J, Mesquita M (2020) Characterizing Sugarcane Production Areas Using Actual Yield and Edaphoclimatic Condition Data for the State of Goiás, Brazil. *Int J Plant Prod* 14:511–520. <https://doi.org/10.1007/s42106-020-00101-9>
 34. Paixão JS, Casaroli D, dos Anjos JCR, Alves Júnior J, Evangelista AWP, Dias HB, Battisti R (2021) Optimizing Sugarcane Planting Windows Using a Crop Simulation Model at the State Level. *Int J Plant Prod* (2021). <https://doi.org/10.1007/s42106-021-00134-8>
 35. Perin V, Sentelhas PC, Dias HB, Santos EA (2019) Sugarcane irrigation potential in Northwestern São Paulo, Brazil, by integrating Agrometeorological and GIS tools. *Agric Water Manag* 220:50–58. <https://doi.org/10.1016/j.agwat.2019.04.012>
 36. Pirttioja N, Carter TR, Fronzek S, et al (2015) Temperature and precipitation effects on wheat yield across a European transect: a crop model ensemble analysis using impact response surfaces. *Clim Res* 65:87–105. <https://doi.org/10.3354/cr01322>
 37. Prescott JA (1940) Evaporation from a water surface in relation to solar radiation. *Trans Roy Soc Sci Austr* 64:114–118. <https://doi.org/10.1155/2013/168048>
 38. Ramirez-Villegas J, Heinemann AB, Pereira de Castro A, Breseghello F, Navarro-Racines C, Li T, Rebolledo MC, Challinor AJ (2018) Breeding implications of drought stress under future climate for upland rice in Brazil. *Glob Chang Biol* 24:2035–2050. <https://doi.org/10.1111/gcb.14071>
 39. Ruane AC, Goldberg R, Chryssanthacopoulos J (2015) Climate forcing datasets for agricultural modeling: merged products for gap-filling and historical climate series estimation. *Agric For Meteorol* 200: 233–248. <https://doi.org/10.1016/j.agrformet.2014.09.016>
 40. Sparks A (2018) nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R. *J Open Source Softw* 3(30):1035. <https://doi.org/10.21105/joss.01035>
 41. Sridevi V, Chellamuthu V (2015) Impact of weather on rice-a review. *Int J Appl Res* 1:825–831
 42. Teixeira WR, Battisti R, Sentelhas PC, Moraes MF, Oliveira Junior A (2019) Uncertainty assessment of soya bean yield gaps using DSSAT-CSM-CROPGRO-Soybean calibrated by cultivar maturity groups. *J Agron Crop Sci* 205:533–544. <https://doi.org/10.1111/jac.12343>
 43. van Wart J, Grassini P, Cassman KG (2013) Impact of derived global weather data on simulated crop yields. *Glob Chang Biol* 19:3822– 3834. <https://doi.org/10.1111/gcb.12302>
 44. van Wart J, Grassini P, Yang H, Claessens L, Jarvis A, Cassman KG (2015) Creating long-term weather data from thin air for crop simulation modeling. *Agr Forest Meteorol* 209-210:49–58. <https://doi.org/10.1016/j.agrformet.2015.05.001>

45. Willmott CJ (1981) On the validation of models. Phys Geogr 2:184–194

46. Xavier AC, King CW, Scanlon BR (2015) Daily gridded meteorological variables in Brazil (1980–2013). Int J Climatol 36(6):2644–2659. <https://doi.org/10.1002/joc.4518>

Tables

Table 1 Brazilian measured weather data (MWD) locations

ID*	Location	State	Latitude (°)	Longitude (°)	Elevation (m ASML)	Weather station	**Missing data (%)
s1	Altamira	PA	-3.21	-52.21	74	INMET	1.68%
s2	Alto Parnaíba	MA	-9.1	-45.93	285	INMET	3.00%
s3	Jataí	GO	-17.91	-51.71	662	INMET	0.87%
s4	Paracatu	MG	-17.24	-46.88	712	INMET	2.01%
s5	Porto Nacional	TO	-10.71	-48.41	239	INMET	9.10%
s6	Rio Branco	AC	-9.96	-67.8	160	INMET	4.05%
s7	Santo Antônio de Goiás	GO	-16.28	-49.17	823	EMBRAPA	0.95%

*ID - weather station identification; **Percentage of missing data set considering maximum and minimum air temperature, precipitation, and solar radiation from 1984 to 2016

Table 2 Mean potential and attainable upland rice yield for both soil conditions simulated by ORYZAv3, using MWD and two GWD sources, Daily Gridded (DG) and NASA/POWER (NP), totalizing 32 crop seasons for seven locations in Brazil

Data	Cond	Site	Yield Mg ha ⁻¹		ME	MAE	RMSE	d	r	R ²
			MWD	GWD						
	Yp	Altamira	2.92	3.11	0.19	0.31	0.39	0.75	0.66	0.43
		Alto Parnaíba	3.76	4.05	0.29	0.38	0.44	0.72	0.65	0.42
		Jataí	4.40	5.04	0.63	0.63	0.69	0.65	0.83	0.69
		Paracatu	4.89	4.88	-0.01	0.34	0.60	0.74	0.50	0.25
DG		Porto Nacional	3.45	3.63	0.18	0.24	0.29	0.84	0.81	0.65
		Rio Branco	3.19	3.00	-0.19	0.25	0.34	0.62	0.41	0.17
		S. A. Goiás	5.04	4.95	-0.09	0.74	0.96	0.63	0.42	0.18
	Ya clay soil	Altamira	2.53	2.87	0.34	0.48	0.64	0.53	0.27	0.07
		Alto Parnaíba	2.15	2.33	0.18	0.45	0.58	0.90	0.83	0.68
		Jataí	3.00	3.69	0.69	0.79	0.91	0.67	0.68	0.46
		Paracatu	2.00	1.89	-0.12	0.50	0.85	0.84	0.71	0.51
		Porto Nacional	2.44	2.92	0.49	0.57	0.73	0.80	0.77	0.59
		Rio Branco	2.87	2.83	-0.04	0.23	0.30	0.73	0.53	0.28
		S. A. Goiás	3.22	3.41	0.19	0.63	0.78	0.74	0.58	0.33
	Ya sandy soil	Altamira	1.73	2.15	0.42	0.48	0.56	0.52	0.40	0.16
		Alto Parnaíba	1.25	1.49	0.24	0.38	0.46	0.82	0.76	0.57
		Jataí	1.69	2.27	0.58	0.59	0.66	0.61	0.74	0.54
		Paracatu	1.09	1.16	0.06	0.28	0.48	0.84	0.71	0.50
		Porto Nacional	1.47	1.93	0.46	0.50	0.57	0.76	0.78	0.61
		Rio Branco	1.84	2.06	0.22	0.31	0.36	0.68	0.60	0.36
		S. A. Goiás	1.75	2.12	0.37	0.45	0.53	0.71	0.69	0.48
	Yp	Altamira	2.92	3.22	0.30	0.38	0.47	0.53	0.37	0.13
		Alto Parnaíba	3.76	4.24	0.49	0.53	0.65	0.58	0.48	0.24
		Jataí	4.40	5.37	0.96	1.00	1.06	0.46	0.59	0.34

NP		Paracatu	4.89	5.51	0.62	0.83	1.02	0.51	0.34	0.11
		Porto Nacional	3.45	3.87	0.42	0.51	0.63	0.55	0.41	0.17
		Rio Branco	3.19	3.50	0.31	0.40	0.53	0.39	-0.05	0.00
		S. A. Goiás	5.04	5.72	0.68	1.09	1.30	0.47	0.19	0.03
	Ya clay soil	Altamira	2.53	2.80	0.27	0.38	0.46	0.82	0.77	0.59
		Alto Parnaíba	2.15	2.37	0.22	0.67	0.87	0.75	0.58	0.33
		Jataí	3.00	3.89	0.89	1.06	1.19	0.63	0.65	0.42
		Paracatu	2.00	2.20	0.20	0.61	0.90	0.84	0.72	0.52
		Porto Nacional	2.44	3.01	0.58	0.63	0.79	0.82	0.82	0.67
		Rio Branco	2.87	3.20	0.33	0.41	0.51	0.48	0.20	0.04
		S. A. Goiás	3.22	4.08	0.87	1.12	1.29	0.58	0.48	0.23
	Ya sandy soil	Altamira	1.73	2.03	0.30	0.32	0.41	0.75	0.79	0.63
		Alto Parnaíba	1.25	1.55	0.30	0.45	0.53	0.76	0.69	0.48
		Jataí	1.69	2.39	0.69	0.76	0.86	0.49	0.50	0.25
		Paracatu	1.09	1.41	0.32	0.50	0.63	0.76	0.67	0.45
		Porto Nacional	1.47	1.95	0.48	0.52	0.60	0.77	0.82	0.67
		Rio Branco	1.84	2.16	0.32	0.39	0.45	0.61	0.52	0.27
		S. A. Goiás	1.75	2.47	0.72	0.77	0.84	0.51	0.57	0.33

Data – datasets, Cond – soil condition, ME - mean error, MAE - mean absolute error, RMSE - root mean square error, d - agreement index, r - Pearson coefficient, R² - coefficient of determination, Y_p – potential yield, Y_a – attainable yield

Table 3 Attainable and potential yield (Y_a and Y_p) overestimation by GWD for both soil conditions simulated by ORYZAv3, using MWD and two GWD sources, Daily Gridded (DG) and NASA/POWER (NP). 32 crop seasons were assessed, as well as the first half (1984-1999) and the second half (2000-2015).

Time Series	Condition	Datasets Yield Mg ha ⁻¹			Datasets Overestimation	
		MWD	DG	NP	DG	NP
	Yp	3.95	4.09	4.49	3.54	13.67
1984 - 2015	Ya clay soil	2.60	2.85	3.08	9.61	18.45
	Ya sandy soil	1.55	1.88	1.99	21.35	29.11
	Yp	4.02	4.09	4.46	1.74	10.95
1984 - 1999	Ya clay soil	2.71	2.87	3.22	5.90	18.82
	Ya sandy soil	1.60	1.93	2.13	20.63	33.13
	Yp	3.87	4.08	4.51	5.43	16.54
2000 - 2015	Ya clay soil	2.48	2.82	2.93	13.71	18.15
	Ya sandy soil	1.48	1.82	1.85	22.97	25.00

MWD – Measured weather database; DG – Daily Gridded; NP – NASA/POWER; ME – Mean error

Figures

Figure 1

Empirical cumulative probability function (ECDF) plot for maximum and minimum air temperature (Tmax and Tmin) (a and b), Rainfall (c), and solar radiation (RS) (d) for all three datasets, measured weather database (MWD) in purple, and two gridded data (GWD) sources, Daily Gridded in green and NASA/POWER in orange. *Precipitation was cut off at 60 mm (c) to provide better visual effect.

Figure 2

Empirical cumulative probability function (ECDF) plot for all municipalities assessed considering only rainfall, measured weather database (MWD) in purple, and two gridded data (GWD) sources, Daily Gridded in green and NASA/POWER in orange. *Precipitation was cut off at 60 mm (c) to provide better visual effect

Figure 3

Relationship between gridded weather (GWD) and measured data (MWD) for the following variables: minimum and maximum air temperature (a, b, c and d), precipitation (e and f) and solar radiation (g and h), considering two sources of GWD, Daily Gridded (GD) with 1:1 green line (a, c, e and g) and NASA/POWER with 1:1 orange line (b, d, f and h). ME - mean error, MAE - mean absolute error, RMSE - root mean square error, d - agreement index, r - Pearson coefficient, R² - coefficient of determination

Figure 4

Relationship between upland rice yield simulated with MWD (x axis) and GWD (y axis), using two GWD sources, Daily Gridded (DG) with 1:1 green line (a, c and e) and NASA/POWER (NP) with 1:1 orange line (b,d and f) database, for potential (a and b) and attainable yield in clay soil condition (c and d) and attainable yield in sandy soil condition (e and f). ME - mean error, MAE - mean absolute error, RMSE - root mean square error, d - agreement index, r - Pearson coefficient, R² - coefficient of determination

Figure 5

Relationship between upland rice physiological maturity in days simulated with MWD (x axis) and GWD (y axis), using two GWD sources, Daily Gridded (DG) with 1:1 green line (a) and NASA/POWER (NP) with 1:1 orange line (b) database. ME - mean error, MAE - mean absolute error, RMSE - root mean square error, d - agreement index, r - Pearson coefficient, R² - coefficient of determination

Figure 6

Regression lines (a, b, c, d, e, f, g and h) and violin plots (i, j, k and l) assessing both clay and soil conditions for the following datasets and variables: Daily Gridded (DG) in green, NASA/POWER (NP) in orange and Measured Weather Database (MWD) in purple. TMAXC - Accumulative maximum air temperature from emergence (a and b), PARCUM - Photosynthetic radiation from emergence to maturity (c and d), ETCUM - Accumulative actual evapotranspiration in main field (e, f, i and j), DRAIN - Accumulative drainage in main field (g, h, k and l)

Figure 7

Empirical cumulative probability function (ECDF) (a, b and c) and line chart (d, e and f) showing the performance of measured data (MWD) in purple and gridded data (GWD), Daily Gridded (DG) in green and NASA/POWER (NP) in orange, considering three conditions, potential yield (Yp) (a and d), attainable yield in clay soil (b and e) and attainable yield in sandy soil (c and f). *line charts show the average yield values considering all locations along 32 crop seasons, whereas ECDF considers all values

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterial.docx](#)