Theoretical and Applied Climatology (2019) 135:237–247
https://doi.org/10.1007/s00704-018-2383-y

ORIGINAL PAPER

Assessment of different gridded weather data for soybean yield simulations in Brazil

R. Battisti[1] · F. D. Bender[2] · P. C. Sentelhas[2]

Received: 3 August 2017 / Accepted: 11 January 2018 / Published online: 18 January 2018

Abstract

A high-density, well-distributed, and consistent historical weather data series is of major importance for agricultural planning and climatic risk evaluation. A possible option for regions where weather station network is irregular is the use of gridded weather data (GWD), which can be downloaded online from different sources. Based on that, the aim of this study was to assess the suitability of two GWD, AgMERRA and XAVIER, by comparing them with measured weather data (MWD) for estimating soybean yield in Brazil. The GWD and MWD were obtained for 24 locations across Brazil, considering the period between 1980 and 2010. These data were used to estimate soybean yield with DSSAT-CROPGRO-Soybean model. The comparison of MWD with GWD resulted in a good agreement between climate variables, except for solar radiation. The crop simulations with GWD and MWD resulted in a good agreement for vegetative and reproductive phases. Soybean potential yield (*Yp*) simulated with AgMERRA and XAVIER had a high correlation ($r > 0.88$) when compared to the estimates with MWD, with the RMSE of about 400 kg ha^{-1}. For attainable yield (*Ya*), estimates with XAVIER resulted in a RMSE of 700 kg ha^{-1} against 864 kg ha^{-1} from AgMERRA, both compared to the simulations using MWD. Even with these differences in *Ya* simulations, both GWD can be considered suitable for simulating soybean growth, development, and yield in Brazil; however, with XAVIER GWD presenting a better performance for weather and crop variables assessed.

Keywords AgMERRA · XAVIER · Crop model · Potential yield · Attainable yield · Crop phenology

1 Introduction

The food demand is increasing around the world, which is associated with both population and per capita consumption increase and dietary changes for food insecurity. Under this scenario, soybean crop has an important role for being a grain rich in protein and oil, serving for human consumption and animal feed (EMBRAPA 2015). The increase of soybean demand can be supplied by increasing the growth area or crop yield (Sentelhas et al. 2015; Davis et al. 2016). The increase of growth area is limited in many countries, where yield improvement is the only way to produce more food. In Brazil, there is land available for agricultural expansion of about 65 Mha, but mainly in regions with degraded pastures and restrictive climates (Monteiro and Sentelhas 2017).

For increasing crop yield, it is essential to understand the best management strategies to improve crop resilience (Battisti et al. 2017a; Do Rio et al. 2015), mainly in the zones of higher climatic risk (Heinemann et al. 2016; Zanon et al. 2016). These analyses can be done by using crop simulation models, estimating yields and yield gaps based on simulated potential (*Yp*) and attainable (*Ya*) yields, together with actual yield (*Yr*) (Van Ittersum et al. 2013; Sentelhas et al. 2015) and integrating soil, weather, and cultivar effects on crop growth and development (Nendel et al. 2014; Battisti et al. 2017b). For these analyses, it is crucial to have a consistent historical weather data series with at least 30 years (WMO 1989) and with suitable spatial distribution to represent consistently the temporal and spatial climatic variability (Grassini et al. 2015; Ruane et al. 2015).

In Brazil, suitable temporal weather data and spatial weather station density are limited, mainly in the Midwest and North regions of the country, to where soybean crop is advancing. In

✉ R. Battisti
battisti@ufg.br

[1] College of Agronomy, Federal University of Goiás, Goiânia, GO, Brazil

[2] Department of Biosystems Engineering, ESALQ, University of São Paulo, Piracicaba, SP, Brazil

these regions, there are few weather stations with series no longer than 15 years (Xavier et al. 2015) and with gaps (Battisti et al. 2017a), which make their use very restrict to agrometeorological studies. One of the ways to solve these problems is the use of gridded weather data (GWD), which are based on the interpolation of observed ground-based data and merged satellite products (Xavier et al. 2015; Ruane et al. 2015).

GWD have a high potential to be used for agricultural applications, mainly for planning purposes, through the crop simulation models. In this way, Bai et al. (2010) showed that a combination of NASA satellite solar radiation with ground-station temperature is an excellent option for filling weather data gaps, allowing the simulation of maize potential yield in China. Bandaru et al. (2017) used four gridded weather databases to simulate the biomass of a short-rotation woody crop in Columbia Plateau, highlighting the importance of accounting for uncertainties in biomass estimates by GWD. Mourtzinis et al. (2017), on the other hand, observed that the use of weather data interpolated from existing meteorological stations for a region with high density of weather stations was better than using two different gridded weather databases (Daymet and PRISM) to simulate maize yield in U.S. Corn Belt.

Considering the lack of historical weather data series in the Brazilian soybean regions and the availability of a large number of daily gridded weather databases around the world (Xavier et al. 2015; Ruane et al. 2015), the main objective of this study was to assess the suitability of two gridded weather databases: AgMERRA, obtained from a reanalysis of satellite and observed weather data (Ruane et al. 2015), and XAVIER, obtained from the interpolation of several ground-based station data available for Brazil (Xavier et al. 2015); in comparison with observed weather data to simulate soybean development and yield in Brazil.

2 Material and methods

2.1 Locations and weather databases

The weather databases employed in this study were obtained from 24 locations in the main soybean regions across Brazil, which represent the areas with the highest soybean production intensity (Fig. 1). These regions encompass different climates, from subtropical humid in the south to tropical with intense dry season during the winter in the north. Details about the climate characteristics of these locations can be found in Alvares et al. (2013). The daily weather data comprise solar radiation, maximum and minimum air temperature, rainfall, mean relative humidity, and wind speed. The period used for soybean growth simulations was between 1980 and 2010, being the same for the three databases, totalizing 30 growing seasons.

Measured weather data (MWD) were obtained from Brazilian Meteorological Service (INMET) (21 locations) and Agronomic Institute of Paraná (IAPAR) (three locations). The homogeneity test of MWD followed the approach used by Xavier et al. (2015). In these weather databases, 20% of missing weather data was found. For rainfall, missing data were replaced by measured data from the closest rainfall station from Brazilian Water Agency (ANA), which has a dense rainfall station network covering all Brazilian territory. Missing air temperature and relative humidity data were replaced by estimates from linear relationships with nearby weather stations, choosing those with the same climate classification (Alvares et al., 2013). Solar radiation was estimated using Angstrom equation (Angstrom, 1924), which has input sunshine hours, or by the method proposed by Hargreaves and Samani (1982), which uses maximum and minimum air temperature as inputs (Allen et al. 1998; Pereira et al. 2002). Missing wind speed data were filled out with daily historical average values for each day.

The daily GWD were obtained from two publicly available weather databases. The first weather database was developed based on interpolated ground-based stations available for Brazil, aiming to grid data of precipitation and reference evapotranspiration, obtained from Xavier et al. (2015) (XAVIER). XAVIER used 2890 rain gauges and 735 weather stations (260 conventional and 475 automatics) for their interpolations, testing different methods to that and using cross-validation to evaluate the efficiency of GWD in the resolution of 0.25° for latitude and longitude.

The second gridded weather database was the AgMERRA (Ruane et al. 2015), which is one of the global weather database used by AgMIP (The Agricultural Model Intercomparison and Improvement Project) (Rosenzweig et al. 2013). AgMERRA uses reanalysis of climate forcing datasets (provided by MERRA, MERRA-Land, CRU, WM, GPCC, TRMM, CMORPH, PERSIANN, NASA/GEWEX SRB), with a higher spatial resolution for agricultural application. The satellite datasets were adjusted based on measured monthly data, from 737 weather stations in agricultural regions around the world, using smart interpolation, which considers elevation to create daily gridded weather database with a spatial resolution of 0.25°.

2.2 Crop simulation model

The soybean development and yield simulations were performed using the CSM-CROPGRO-Soybean model (Boote et al. 2003), present in the software Decision Support

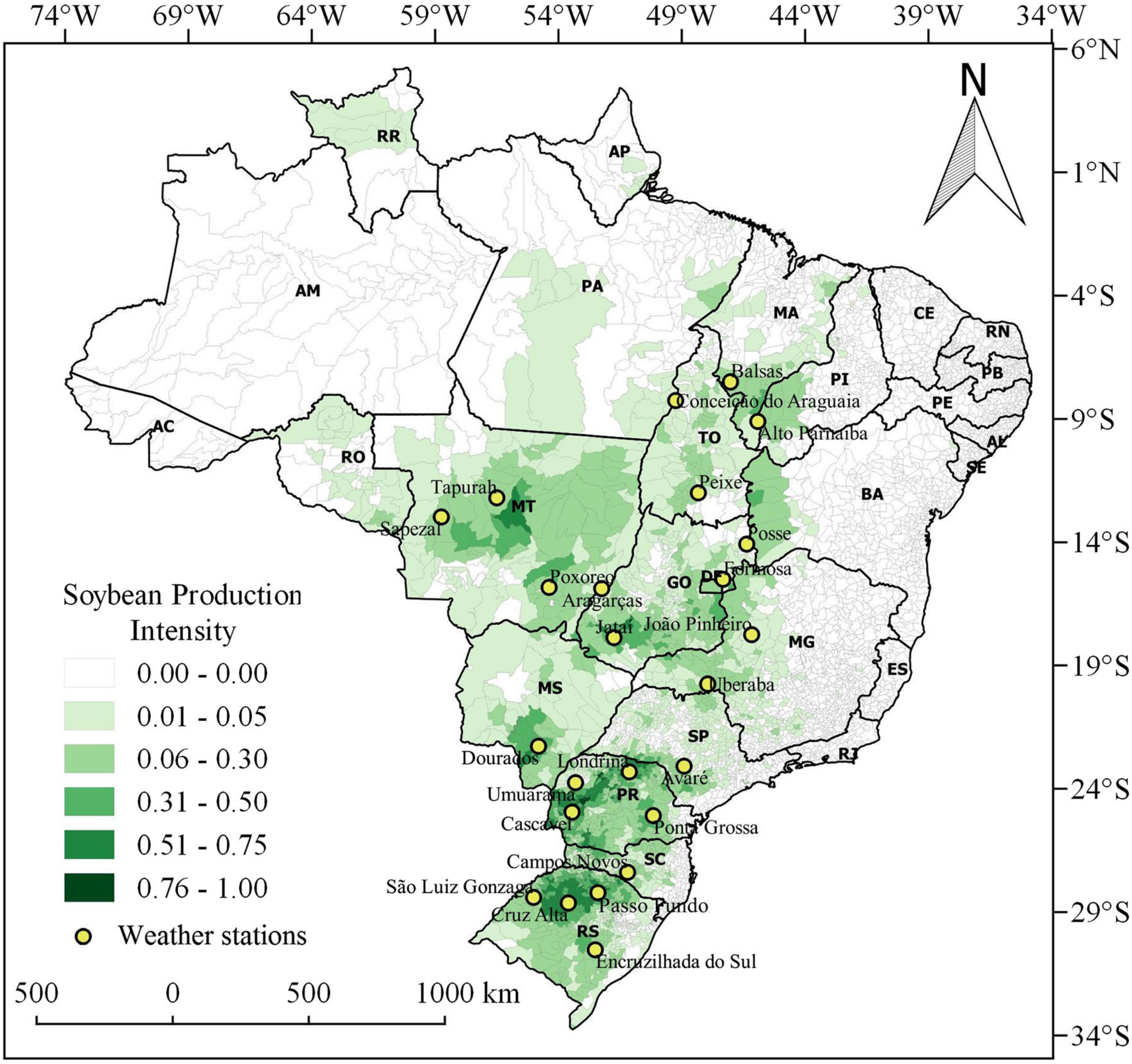

Fig. 1 Distribution of weather stations used for soybean yield simulations and the distribution of this crop during the 2014/2015 growing season in Brazil. The values in the legend represent the decimal part of a pixel cultivated with soybean. Adapted from Battisti and Sentelhas (2017) and based on IBGE (2016) data

System for Agrotechnology Transfer (DSSAT) (Jones et al. 2003). This model was calibrated by Battisti et al. (2017b) for Brazilian cultivars, with the model presenting a root mean square error below of 550 kg ha^{-1} for calibration and validation phases. CSM-CROPGRO-Soybean model considers the following approaches: reference evapotranspiration estimated by Penman-Monteith FAO 56 method (Allen et al. 1998); infiltration of water into the soil defined by soil curve number (Soil Conservation Service 1972); soil water balance determined by Ritchie tipping-bucket method (Ritchie 1998); soil evaporation estimated by Suleiman-Ritchie (Suleiman and Ritchie 2003); and leaf-level photosynthesis response from Boote and Pickering (1994).

The common soybean management practices were considered in the simulations, such as sowing date on 15 Nov with a plant density of 30 plants m^{-2} for all locations, considering a rainfed crop with nitrogen in the soil coming from biological fixation. Across locations, three cultivar maturity groups were used based on the latitude. Maturity group 5.8 was used for latitudes higher than 23° S; maturity group 6.8 between 15° and 22.9° S; and maturity group 7.8 for latitudes lower than 14.9° S. The cultivar coefficients for each of the three maturity groups were obtained from Battisti (2016). The use of different maturity groups resulted in a crop cycle between 110 and 130 days from sowing to maturity, representing the main cultivars used in Brazil. In these regions, the soil type was classified as clay, having a permanent wilting point of 0.296 cm^3 cm^{-3}, field capacity of 0.458 cm^3 cm^{-3}, and saturation point of 0.578 cm^3 cm^{-3}, which resulted in a soil water holding capacity of 1.6 mm cm^{-1} of soil depth. For simulating the root system, an intermediate value for root growth factor was used, which resulted in a maximum root depth of 120 cm (Battisti and Sentelhas 2017).

2.3 Statistical analysis

The following variables were used to compare the performance of XAVIER and AgMERRA GWD against MWD: potential and attainable grain yield as defined by Sentelhas et al. (2015); maximum and actual crop evapotranspiration accumulated between sowing and maturity; water deficit represented by the difference between potential and actual crop evapotranspiration; days for vegetative (from sowing to beginning of flowering) and reproductive (from beginning of flowering to maturity) periods; accumulated rainfall from sowing to maturity; and mean solar radiation and minimum and maximum air temperatures from sowing to maturity. The analysis was done considering 24 locations and 30 growing seasons as replications. For attainable grain yield, an individual analysis by location was also considered. These variables were all outputs from the crop model used to simulate the soybean growth, considering MWD and GWD as the input in the simulation.

The statistical indexes used to assess the performance of each database were as follows: Pearson's correlation coefficient (r), mean error (ME), and root mean square error (RMSE) (Wallach et al. 2006):

$$r = \frac{\sum_{i=1}^{N}[(Mi - M)(Si - S)]}{\sqrt{\sum_{i=1}^{N}\left[(Mi - M)^2\right]\sum_{i=1}^{N}\left[(Si - S)^2\right]}}$$

$$ME = \frac{1}{N}\sum_{i=1}^{N}(Si - Mi)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Si - Mi)2}$$

where Mi and M are respectively the variable of each i replication and the mean value from MWD or estimated based on it, Si and S are respectively the variable of each i replication and the mean value from GWD or estimated based on it, and N is the number of replications.

3 Results and discussions

3.1 Air temperature, solar radiation, and rainfall

GWD showed a good agreement with MWD for minimum and maximum air temperature with r above 0.79 for AgMERRA (Fig. 2a) and above 0.87 for XAVIER (Fig. 2b), considering the period when soybean is cultivated. The AgMERRA showed a mean error (ME) of − 0.32 and − 0.30 °C, respectively, for minimum and maximum air temperatures (Fig. 2a), while XAVIER had a ME of 0.18 and − 0.06 °C, respectively, for minimum and maximum air temperatures (Fig. 2b). XAVIER presented lower RMSE than AgMERRA for minimum and maximum air temperatures. These errors were around 1.2 °C for AgMERRA and between 0.7 and 1.1 °C for XAVIER. These errors in air temperature can affect several crop development and growth processes in the crop model, such as crop cycle duration, photosynthesis, and evapotranspiration (Martre et al. 2015). Bai et al. (2010) observed that GWD systematic errors of − 2.8 and − 1.4 °C, respectively, for minimum and maximum air temperatures, limited the potential yield estimation for maize in China.

The agreement between GWD and MWD solar radiation data, based on their average during the soybean cycle, was poor for both databases (Fig. 2c and d), with r of 0.15 and 0.19, respectively, for AgMERRA and XAVIER. AgMERRA underestimated solar radiation (ME = − 0.28 MJ m^{-2} day^{-1}), while XAVIER overestimated it (ME = 0.30 MJ m^{-2} day^{-1}). The poor correlations resulted in high RMSE, around of 2.50 MJ m^{-2} day^{-1} for both databases. The worst performance of the GWD for solar radiation was in part expected once as MWD is derived from conventional weather stations, which do not report solar radiation, but effective sunshine hours (Xavier et al. 2015). Based on sunshine hours from MWD, the solar radiation was estimated by Angstrom-Prescott method (Angstrom 1924; Prescott 1940), to the locations where calibrated coefficients (a and b) were available, or by Glover and McCulloch (1958) coefficients, where a and b were not calibrated. Xavier et al. (2015) used Glover and McCulloch (1958) coefficients to estimate solar radiation through Angstrom-Prescott model for conventional weather stations and daily solar radiation measured in automatic stations that probably caused the differences.

The accumulated rainfall during soybean cycle showed good agreement for both databases (Fig. 2e and f). AgMERRA underestimated rainfall by 9.14 mm cycle^{-1}, while XAVIER underestimated it by 21.17 mm cycle^{-1}. For rainfall above 1200 mm cycle^{-1}, both sources underestimated the observed values. Such performance mainly occurs as a function of the spatial rainfall distribution, once intense rainfall that occurs at a specific point is normally underestimated by the GWA that reports a mean value for the grid (Ruane et al., 2015; Xavier et al. 2015).

3.2 Evapotranspiration and water deficit

The maximum and actual crop evapotranspiration estimated with both databases were compared (Fig. 3). For crop evapotranspiration, XAVIER data showed better performance than AgMERRA, with similar performance for maximum (r = 0.82, ME = −11 mm cycle^{-1}, and RMSE = 63 mm cycle^{-1}) and actual (r = 0.80, ME = 4 mm cycle^{-1}, and RMSE =

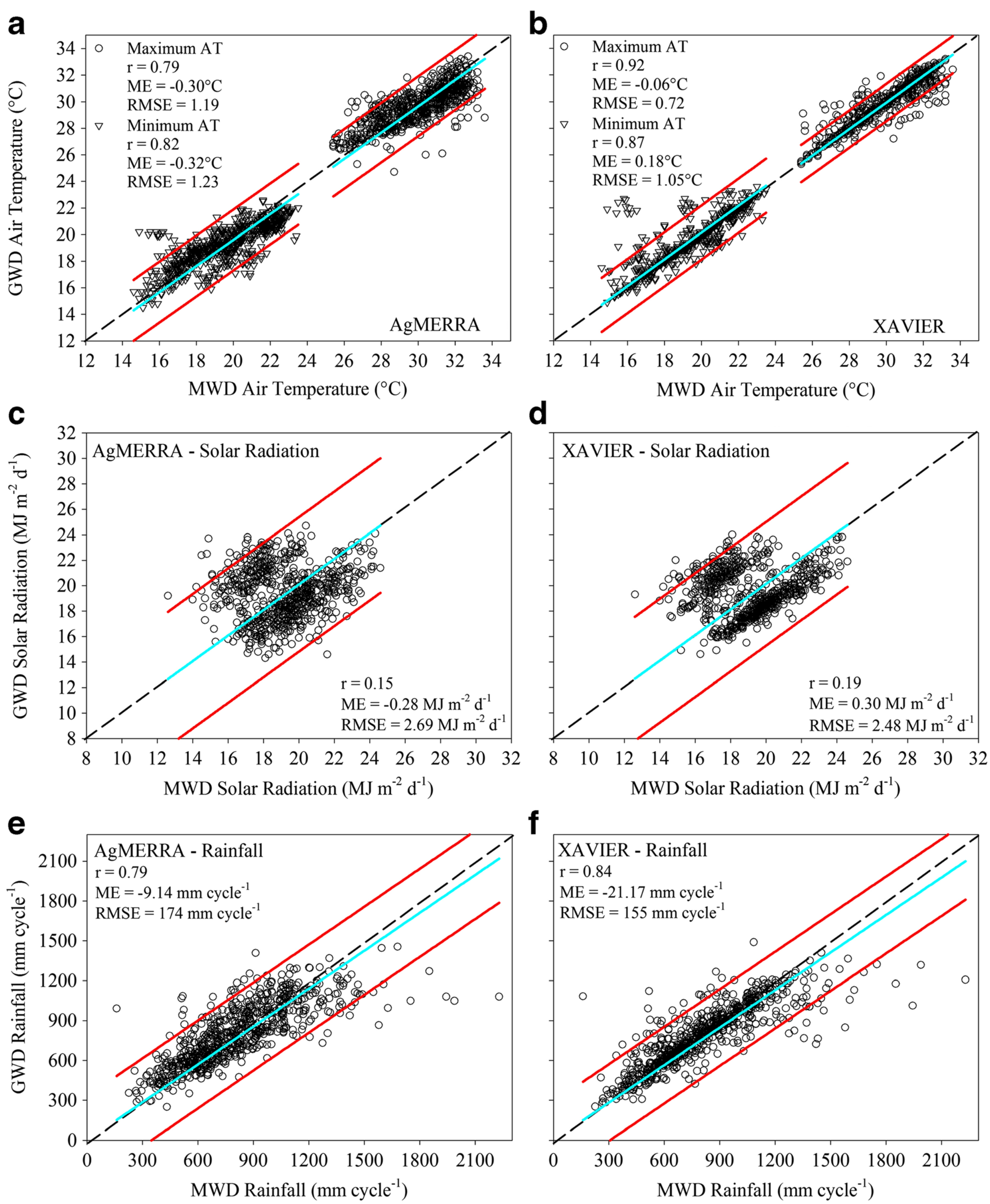

Fig. 2 Relationship between measured (MWD) and gridded (GWD) weather data for the following variables: minimum and maximum air temperature (**a** and **b**), solar radiation (**c** and **d**), and accumulated rainfall (**e** and **f**), during soybean cycle, considering two sources of GWD, AgMERRA (**a**, **c**, and **e**) and XAVIER (**b**, **d**, and **f**). Person correlation coefficient (r), mean error (ME), and root mean square error (RMSE) are presented in the graphs. The red line is the 95% confidence interval for predicted individual values, the cyan line is the trend line, and the black dashed line is 1:1 line

49 mm cycle^{-1}) crop evapotranspiration (Fig. 3). Maximum crop evapotranspiration showed higher dispersion than actual crop evapotranspiration for both databases, which is basically associated with the errors in solar radiation, once Penman-Monteith ETo method (Allen et al. 1998) considers net radiation as an input data.

The total water deficit during soybean cycle, obtained by the difference between the maximum and actual crop evapotranspiration, is presented in Fig. 4. The water deficit penalizes the final yield and is the main yield gap cause for soybean in Brazil, which has different levels across the country (Sentelhas et al. 2015). When determined with

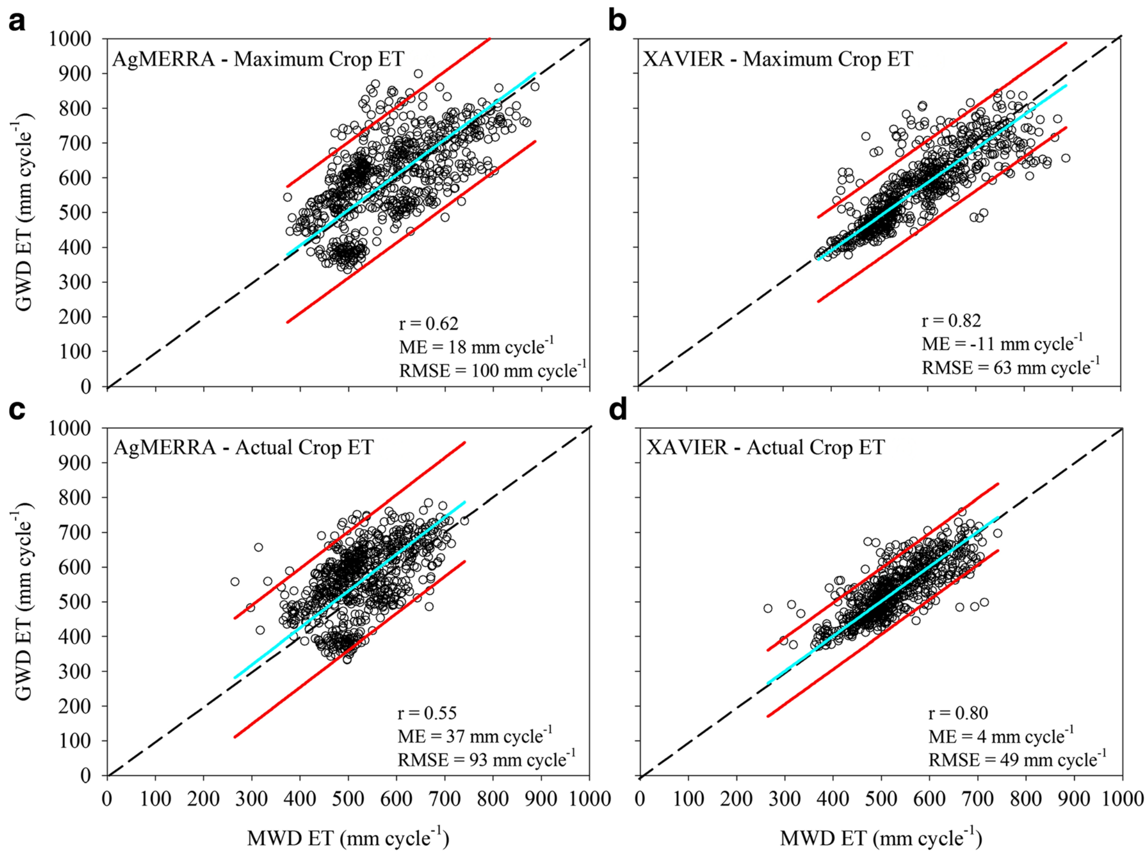

Fig. 3 Relationship between evapotranspiration estimated with measured and gridded data, using AgMERRA (**a** and **c**) and XAVIER (**b** and **d**) database, for maximum (**a** and **b**) and actual (ET) (**c** and **d**) crop evapotranspiration accumulated during the soybean cycle. Person correlation coefficient (r), mean error (ME), and root mean square error (RMSE) are presented in the graphs. The red line is the 95% confidence interval for predicted individual values, the cyan line is the trend line, and the black dashed line is 1:1 line

MWD, the water deficit showed more values lower than 100 mm cycle^{-1}, but in some cases, it reached near 400 mm cycle^{-1}. Both databases showed an underestimation of water deficit in comparison to the values estimated by MWD, with ME of −22 and −15 mm cycle^{-1}, respectively, for AgMERRA and XAVIER (Fig. 4a and b). The water deficit estimated with XAVIER database presented higher correlation ($r = 0.81$) and lower RMSE (46 mm cycle^{-1})

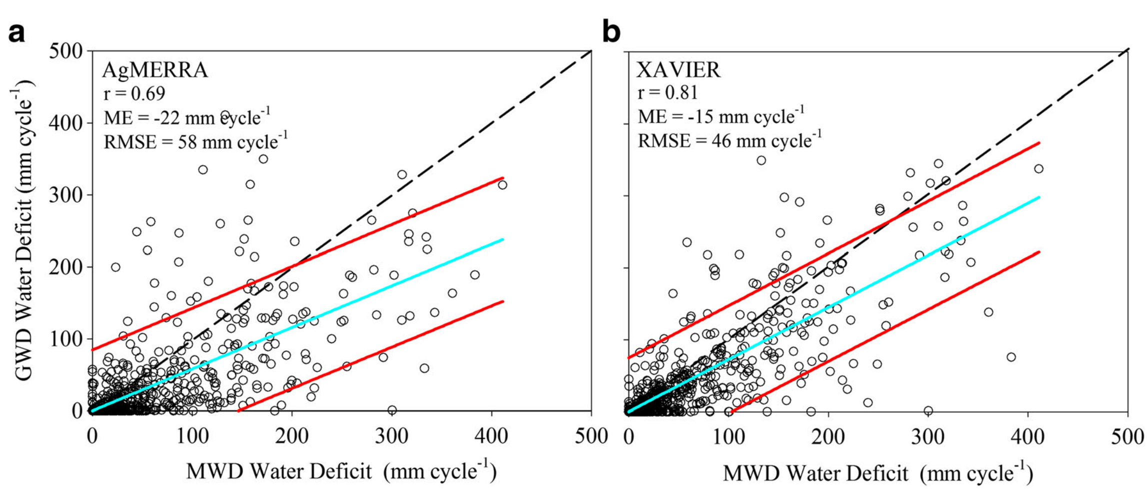

Fig. 4 Relationship between water deficits estimated with measured and gridded data, using AgMERRA (**a**) and XAVIER (**b**) database, accumulated during the soybean cycle. Person correlation coefficient (r), mean error (ME), and root mean square error (RMSE) are presented in the graphs. The red line is the 95% confidence interval for predicted individual values, the cyan line is the trend line, and the black dashed line is 1:1 line

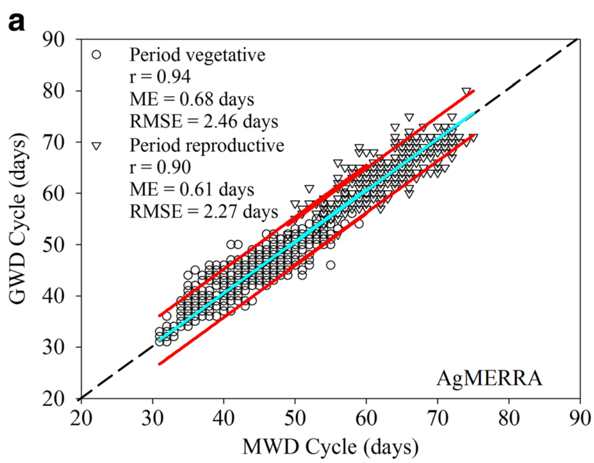

Fig. 5 Relationship between vegetative and reproductive periods estimated with measured and gridded data, using AgMERRA (**a**) and XAVIER (**b**) database. Person correlation coefficient (*r*), mean error (ME), and root mean square error (RMSE) are presented in the graphs. The red line is the 95% confidence interval for predicted individual values, the cyan line is the trend line, and the black dashed line is 1:1 line

with the estimates with MWD than AgMERRA ($r = 0.69$; RMSE = 58 mm cycle^{-1}). The higher water deficit tended to have higher errors, which was a consequence of the errors in rainfall amount and distribution (Xavier et al. 2015; Ruane et al. 2015).

3.3 Crop development

The vegetative and reproductive periods were well estimated by both GWD, with similar performance for AgMERRA and XAVIER when compared with MWD (Fig. 5). The correlation

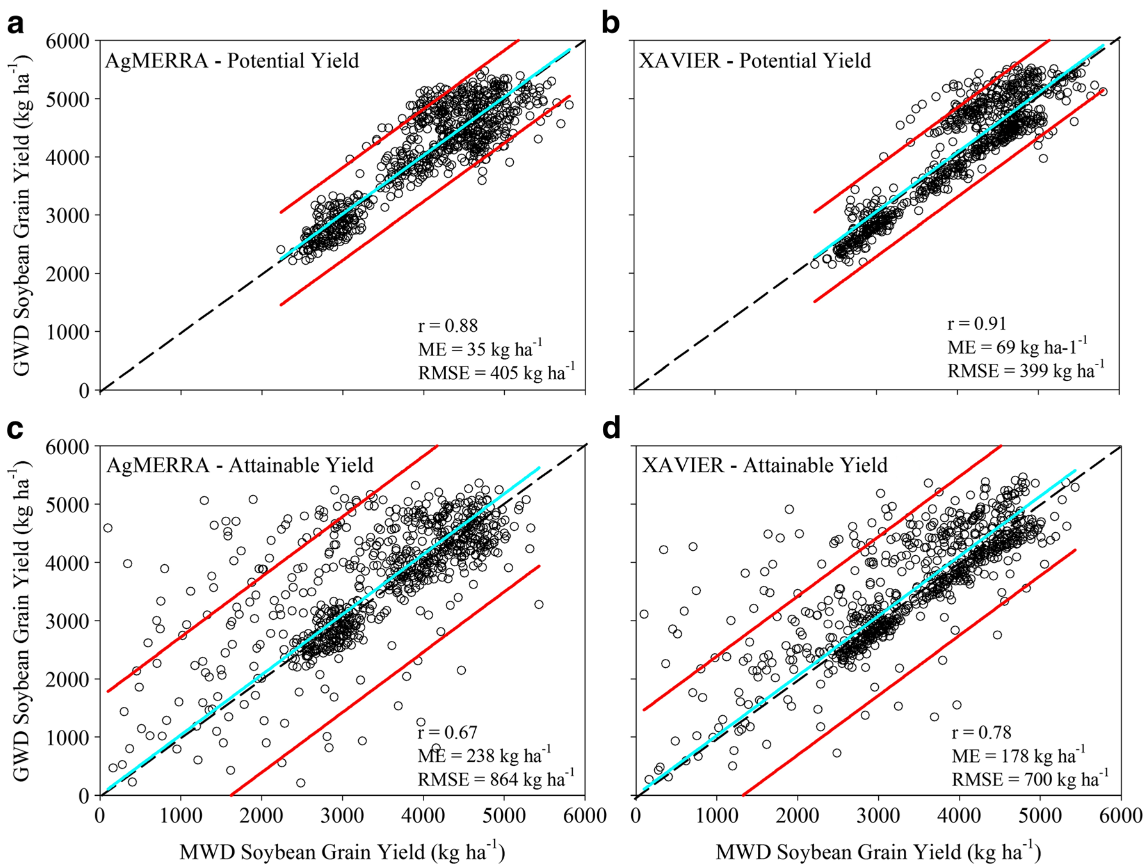

Fig. 6 Relationship between soybean yield estimated with measured and gridded data, using AgMERRA (**a** and **c**) and XAVIER (**b** and **d**) database, for potential (**a** and **b**) and attainable (**c** and **d**) yield. Person correlation coefficient (*r*), mean error (ME), and root mean square error (RMSE) are presented in the graphs. The red line is the 95% confidence interval for predicted individual values, the cyan line is the trend line, and the black dashed line is 1:1 line

Table 1 Mean attainable soybean yield estimated by the CSM-CROPGRO-Soybean, using measured (MWD) and gridded weather data (GWD) from AgMERRA and XAVIER databases, totalizing 30 crop seasons, for 24 locations in Brazil

Site	MWD	GWD							
		AgMERRA	XAVIER	AgMERRA			XAVIER		
		Soybean grain yield kg ha^{-1}		r	ME kg ha^{-1}	RMSE	r	ME kg ha^{-1}	RMSE
Alto Parnaíba	2784	2693	2638	0.44	−92	417	0.97	−146	204
Aragarças	3738	3878	3724	−0.18	139	737	−0.35	−14	750
Avaré	4598	4414	4435	−0.14	−184	719	0.89	−162	366
Balsas	2480	2485	2385	0.30	5	299	0.30	−95	312
Campos Novos	4149	4753	4531	0.64	605	933	0.90	382	568
Cascavel	3808	4837	4277	0.53	1029	1240	0.61	470	813
Con. do Araguaia	2669	2542	2499	0.49	−127	199	0.57	−170	214
Cruz Alta	3363	3253	3330	0.86	−110	774	0.92	−34	577
Dourados	2451	2802	3611	0.80	352	901	0.62	1160	1586
Enc. do Sul	3117	2381	2827	0.57	−736	1418	0.64	−289	1111
Formosa	4220	4098	4067	0.64	−123	416	0.95	−153	236
Jataí	4263	4244	4162	0.23	−20	245	0.49	−102	207
João Pinheiro	3553	3955	3790	0.68	401	982	0.85	237	695
Londrina	3602	3762	4101	0.66	160	820	0.62	499	735
Passo Fundo	2845	3801	3932	0.85	956	1187	0.84	1087	1326
Peixe	3036	2956	2885	0.42	−80	243	0.89	−152	187
Ponta Grossa	4015	4728	4567	−0.14	713	1003	0.22	553	870
Posse	3485	3356	3349	0.67	−128	448	0.96	−136	242
Poxoréo	3838	3908	3784	0.26	71	283	0.42	−53	255
S. L. Gonzaga	2316	3955	2980	0.61	1639	1986	0.89	664	959
Sapezal	2909	3003	3032	0.25	94	249	0.27	123	275
Tapurah	3003	2983	2922	0.15	−20	218	0.41	−81	188
Uberaba	4610	4523	4459	0.42	−87	298	0.37	−151	285
Umuarama	3041	4289	3883	0.72	1247	1404	0.87	842	960
General	3412	3650	3590	0.67	238	864	0.78	178	700

r is the Person correlation coefficient, ME is the mean error, and RMSE is the root mean square error

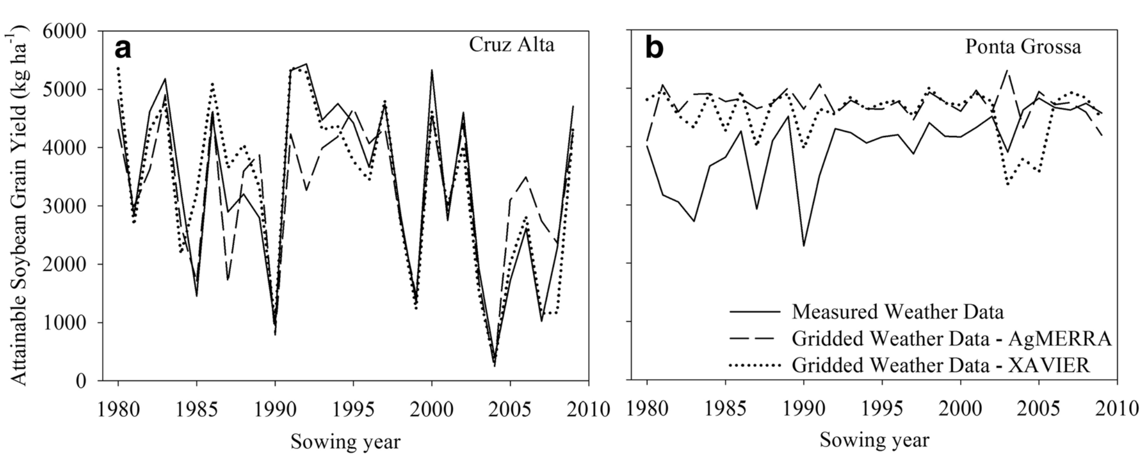

Fig. 7 Attainable soybean yield simulated using measured (MWD) and gridded weather data (GWD) from AgMERRA and XAVIER databases along 30 crop seasons for the locations with one of the best (Cruz Alta, **a**) and the worst (Ponta Grossa, **b**) performances, as presented in Table 1

coefficient was higher than 0.90, with ME ranging from − 0.36 to 0.68 days and RMSE below 2.46 days. These results are similar to those observed by Mourtzinis et al. (2017) for maize in the U.S. Corn Belt for different GWD (RMSE from 3 to 7 days). As the duration of vegetative and reproductive periods are defined by air temperature, photoperiod, and water deficit (Boote et al. 2003), the errors observed of GWD for these variables were not enough to cause errors of higher magnitude for crop phenology estimation. The correct estimation of crop phases and cycle is the first step to have good crop growth and yield estimates by crop simulation models since they also affect photosynthesis rate and crop sensitivity to water deficit (Battisti and Sentelhas 2015).

3.4 Potential and attainable soybean grain yield

Potential yield is determined by solar radiation, air temperature, and crop cycle duration (Sentelhas et al. 2015). The AgMERRA and XAVIER showed a good agreement with MWD for potential yield, respectively, with r of 0.88 and 0.81, ME of 35 and 69 kg ha^{-1}, and RMSE of 405 and 399 kg ha^{-1} (Fig. 6a and b). As presented in Fig. 6a and b, most of the data are close to 1:1 line, showing the high agreement between potential yield estimated with MWD and GWD. As the temperature was well estimated by both gridded databases, the majority of the dispersion observed for potential yield is related to the errors in solar radiation (Fig. 2c and d). The results indicate that both GWD were able to estimate efficiently soybean potential yield across Brazil, which disagrees from the study of Bai et al. (2010), which observed that potential yield can be estimated efficiently using solar radiation from satellite GWD, but air temperature needs to be from ground-stations for improving maize yield simulations in China.

For attainable yield, estimated by considering the effect of water deficit on potential yield (Sentelhas et al. 2015), the estimates with XAVIER GWD presented better performance than with AgMERRA, when compared with MWD. The correlation coefficients were of 0.78 and 0.67, ME of 179 and 238 kg ha^{-1}, and RMSE of 700 and 864 kg ha^{-1}, respectively, for the estimates with XAVIER and AgMERRA (Fig. 6c and d). In both cases, the comparison between the attainable yield estimated by GWD and MWD presented higher dispersion than the same analysis for potential yield, which is mainly caused by the errors observed for water deficit (Fig. 6c and d). According to Banduru et al. (2017), it is important to account for the uncertainties in biomass estimation by crop simulation models when using GWD.

The attainable yield is normally the main result of a crop simulation model, once it represents the maximum yield that can be achieved by the crop in a rainfed system, allowing to elaborate agricultural plan, regarding the main options for crop management (Battisti and Sentelhas 2014; Do Rio et al. 2015). Taking into account the attainable yield estimated for the 24 locations assessed in this study (Table 1), an expressive variability was observed in the efficiency of the estimates done with GWD. The variability was associated with the local climate conditions (Alvares et al. 2013) as well as to the density of the weather stations in the area (Xavier et al. 2015; Ruane et al. 2015; Mourtzinis et al. 2017). One of the best performances was obtained in Cruz Alta, with $r > 0.86$, ME > − 110 kg ha^{-1}, and RMSE < 774 kg ha^{-1} (Table 1). On the other hand, one of the worst performances was observed in Ponta Grossa, with r of − 0.14 and 0.22, ME of 713 and 553 kg ha^{-1}, and RMSE of 1003 and 870 kg ha^{-1}, respectively, for AgMERRA and XAVIER GWD (Table 1).

Besides the performance of the estimates done by GWD, these databases were able to represent the attainable yield variability along the crop seasons, as showed for two sites in Fig. 7. It is possible to observe that attainable yields estimated by both GWD resulted in similar yield tendency along the crop seasons when the locations with one of the best and one of the worst performances were assessed (Fig. 7). In Cruz Alta, RS (Fig. 7a), the AgMERRA and XAVIER followed the same yield tendency of the estimates obtained with MWD. In Ponta Grossa, PR (Fig. 7b), where one of the worst performance was obtained (Table 1), a systematic error was observed, with the GWD simulating higher yields than MWD (Fig. 7b), mainly with AgMERRA. When the errors are systematic, as in Fig. 7b, it is even possible to capture the effects of climate on yield variability (Pirttioja et al. 2015).

The systematic error observed in Ponta Grossa (Fig. 7b) was associated with uncertainties in the solar radiation between MWD and GWD. The solar radiation was estimated from sunshine hours for MWD, while in the GWD was estimated based on sunshine hours, and measured daily solar radiation, especially after 2000s with the increase of automatic weather stations in Brazil (Xavier et al. 2015). This type of trend was not clear for other sites analyzed, which is associated with the number of regional weather station that were variable for both GWD along the period of simulation.

4 Conclusions

The AgMERRA and XAVIER gridded weather databases showed a good agreement for minimum and maximum air temperature, accumulated rainfall, potential and actual evapotranspiration, and water deficit when considered the soybean-growing seasons. Solar radiation was identified being a limiting variable for both GWD in comparison to MWD, but these errors did not affect potential yield simulations by the crop model. AgMERRA and XAVIER GWD were efficient to simulate soybean vegetative and reproductive periods and potential and attainable yields, representing well their variability along the crop seasons. XAVIER GWD presented a better performance in all variables than AgMERRA, which makes its use preferential for simulating soybean growth and development in Brazil.

Funding information 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 (process n° 152868/2016-0) and third author (process n° 300582/2013-7).

References

Allen GR, Pereira LS, Raes D, Smith M (1998) Crop evapotranspiration—guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, Rome

Alvares CA, Stape JL, Sentelhas PC, Gonçalves JLM, Sparovek G (2013) Köppen's climate classification map for Brazil. Meteorol Z 22(6): 711–728. https://doi.org/10.1127/0941-2948/2013/0507

Angstrom A (1924) Solar and terrestrial radiation. Q J R Meteorol Soc 50: 121–125. https://doi.org/10.1002/qj.49705021008

Bai J, Chen X, Dobermann A, Yang H, Cassman KG, Zhang F (2010) Evaluation of NASA satellite- and model-derived weather data for simulation of maize yield potential in China. Agron J 102(1):9–16. https://doi.org/10.2134/agronj2009.0085

Bandaru V, Pei Y, Hart Q, Jenkins BM (2017) Impact of biases in gridded weather datasets on biomass estimates of short rotation woody cropping systems. Agric Forest Meteorol 233:71–79. https://doi.org/10.1016/j.agrformet.2016.11.008

Battisti R (2016) Calibration, uncertainties and use of soybean crop simulation models for evaluation strategies to mitigate the effects of climate change in Southern Brazil. Thesis, Luiz de Queiroz College of Agriculture, University of São Paulo

Battisti R, Sentelhas PC (2014) New agroclimatic approach for soybean sowing dates recommendation: a case study. Rev Bras Engenharia Agrícola Ambient 18(11):1149–1156. https://doi.org/10.1590/1807-1929/agriambi.v18n11p1149-1156

Battisti R, Sentelhas PC (2015) Drought tolerance of Brazilian soybean cultivars simulated by a simple agrometeorological yield model. Exp Agric 51(02):285–298. https://doi.org/10.1017/S0014479714000283

Battisti R, Sentelhas PC (2017) Improvement of soybean resilience to drought through deep root system in Brazil. Agron J 109(4):1612–1622. https://doi.org/10.2134/agronj2017.01.0023

Battisti R, Sentelhas PC, Boote KJ (2017b) Inter-comparison of performance of soybean crop simulation models and their ensemble in southern Brazil. Field Crop Res 200:28–37. https://doi.org/10.1016/j.fcr.2016.10.004

Battisti R, Sentelhas PC, Boote KJ, Câmara GMS, Farias JRF, Basso CJ (2017a) Assessment of soybean yield with altered water-related genetic improvement traits under climate change in Southern Brazil. Eur J Agron 83:1–14. https://doi.org/10.1016/j.eja.2016.11.004

Boote KJ, Jones JW, Batchelor WD, Nafziger ED, Myers O (2003) Genetic coefficients in the CROPGRO-Soybean model: link to field performance and genomics. Agron J 95(1):32–51. https://doi.org/10.2134/agronj2003.3200

Boote KJ, Pickering NB (1994) Modeling photosynthesis of row crop canopies. Hort Sci 29:1423–1434

Davis KF, Gephart JA, Emery KA, Leach AM, Galloway JN, D'Odorico P (2016) Metting future food demand with current agricultural resources. Glob Environ Chang 36:125–132. https://doi.org/10.1016/j.gloenvcha.2016.05.004

Do Rio A, Sentelhas PC, Farias JRB, Sibaldelli RNR, Ferreira RC (2015) Alternative sowing dates as a mitigation measure to reduce climate change impacts on soybean yield in southern Brazil. Int J Climatol 36(11):3664–3672. https://doi.org/10.1002/joc.4583

EMBRAPA (2015) Soybean in number http://www.cnpso.embrapa.br/sojaemnumeros/app/graf2.html. Accessed 10 July 2015

Grassini P, Van Bussel LGJ, Wart J, Wolf J, Claessens L, Yang H, Boogaard H, De Groot H, Van Ittersum MK, Cassman KG (2015) How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. Field Crop Res 177:49–63. https://doi.org/10.1016/j.fcr.2015.03.004

Glover J, McCulloch JSG (1958) The empirical relation between solar radiation and hours of sunshine. Q J R Meteorol Soc 84(360):172–175. https://doi.org/10.1002/qj.49708436011

Hargreaves GH, Samani ZA (1982) Estimation of potential evapotranspiration. J Irrig Drain Div 108:223–230

Heinemann AB, Ramirez-Villegas J, Souza TLPO, Didonet AD, Di Stefano JG, Boote KJ, Jarvis A (2016) Drought impact on rainfed common bean production areas in Brazil. Agric For Meteorol 225: 57–74. https://doi.org/10.1016/j.agrformet.2016.05.010

IBGE (2016) Agricultural production (In Portuguese). http://www.sidra.ibge.gov.br/bda/pesquisas/pam/default.asp?o=18&i=P. Accessed 18 February 2016

Jones JW, Hoogenboom G, Porter CH, Boote KJ, Batchelor WD, Hunt LA, Wilkens W, Singh U, Gijsman AJ, Ritchie JT (2003) The DSSAT cropping system model. Eur J Agron 18(3-4):234–265. https://doi.org/10.1016/S1161-0301(02)00107-7

Martre P, Wallach D, Asseng S, Ewert F, Jones JW, Rötter RP, Boote KJ, Ruane AC, Thorburn PJ, Cammarano D, Hatfield JL, Rosenzweig C, Aggarwal PK, Angulo C, Basso B, Bertuzzi P, Biernath C, Brisson N, Challinor AJ, Doltra J, Gayler S, Goldberg R, Grant RF, Heng L, Hooker J, Hunt LA, Ingwersen J, Izaurralde RC, Kersebaum KC, Müller C, Kumar SN, Nendel C, O'leary G, Olesen JE, Osborne TM, Palosuo T, Priesack E, Ripoche D, Semenov MA, Shcherbak I, Steduto P, Stöckle CO, Stratonovitch P, Streck T, Supit I, Tao F, Travasso M, Waha K, White JW, Wolf J (2015) Multimodel ensembles of wheat growth: many models are better than one. Glob Chang Biol 21(2):911–925. https://doi.org/10.1111/gcb.12768

Monteiro LA, Sentelhas PC (2017) Sugarcane yield gap: can it be determined at national level with a simple agrometeorological model? Crop & Pasture Science 68:272–284. https://doi.org/10.1071/CP16334

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

Nendel C, Kersebaum KC, Mirschel W, Wenkel KO (2014) Testing farm management options as climate change adaptation strategies using the MONICA model. Eur J Agron 52:47–56. https://doi.org/10.1016/j.eja.2012.09.005

Pereira AR, Angelocci LR, Sentelhas PC (2002) 'Agrometeorologia: Fundamentos e aplicações práticas.' (Agropecuária: Guaíba, RS, Brazil)

Pirttioja N, Carter TR, Fronzek S, Bindi M, Hoffmann H, Palosuo T, Ruiz-Ramos M, Tao F, Trnka M, Acutis M, Asseng S, Baranowski P, Basso B, Bodin P, Buis S, Cammarano D, Deligios P, Destain MF, Dumont B, Ewert F, Ferrise R, François L, Gaiser T, Hlavinka P, Jacquemin I, Kersebaum KC, Kollas C, Krzyszczak J, Lorite IJ, Minet J, Minguez MI, Montesino M, Moriondo M, Müller C, Nendel C, Öztürk I, Perego A, Rodríguez A, Ruane AC, Ruget F, Sanna M, Semenov MA, Slawinski C, Stratonovitch P, Supit I, Waha K, Wang E, Wu L, Zhao Z, Rötter RP (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

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

Ritchie JT (1998) Soil water balance and plant water stress. In: Tsuji GY, Hoogenboom G, Thornton K (eds) Understanding options of agricultural production. Kluwer Academic Publishers and International Consortium for Agricultural Systems Applications, Dordrecht, The Netherlands, pp 41–53. https://doi.org/10.1007/978-94-017-3624-4_3

Rosenzweig C, Jones JW, Hatfield JL, Ruane AC, Boote KJ, Thorburn P, Antle JM, Nelson GC, Porter C, Janssen S, Asseng S, Basso B, Ewert F, Wallach D, Baigorria G, Winter JM (2013) The agricultural model intercomparison and improvement project (AgMIP): protocols and pilot studies. Agric Forest Meteorol 15:166–182. https://doi.org/10.1016/j.agrformet.2012.09.011

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.

Sentelhas PC, Battisti R, Câmara GMS, Farias JRB, Hampf A, Nendel C (2015) The soybean yield gap in Brazil—magnitude, causes and possible solutions for a sustainable production. J Agric Sci 153(08):1394–1411. https://doi.org/10.1017/S0021859615000313

Soil Conservation Service (SCS) (1972) National engineering handbook. Hydrology Section 4, Chapters 4/10.

Suleiman AA, Ritchie JT (2003) Modeling soil water redistribution during second-stage evaporation. Soil Sci Soc Am J 67(2):377–386. https://doi.org/10.2136/sssaj2003.3770

van Ittersum MK, Cassman KG, Grassini P, Wolf J, Tittonell P, Hochman Z (2013) Yield gap analysis with local to global relevance—a review. Field Crops Res 143:4–17. https://doi.org/10.1016/j.fcr.2012.09.009

Wallach D, Makowski D, Jones JW (2006) Working with dynamic crop models: evaluation, analysis, parameterization, and application. ELSEVIER, Amsterdam

WMO (1989) Calculation of monthly and annual 30-year standard normals (WCDP928 No. 10, WMO-TD/No.341). World Meteorological Organization, Geneva.

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

Zanon AJ, Streck NA, Grassini P (2016) Climate and management factors influence soybean yield potential in a subtropical environment. Agron J 104(4):1447–1454. https://doi.org/10.2134/agronj2015.0535