Australian Journal of Crop Science
AJCS
AJCS 16(04):495-500 (2022) ISSN:1835-2707
doi: 10.21475/ajcs.22.16.04.p3449

Gold Standard in selection of rainfall forecasting models for soybean crops region

Marcio Paulo de Oliveira*[1], Miguel Angel Uribe-Opazo[2], Manuel Galea[3] and Jerry Adriani Johann[2]

[1]Federal University of Technology – Paraná (UTFPR), Toledo, Paraná, Brazil
[2]Western Paraná State University – (UNIOESTE), Cascavel, Paraná, Brazil
[3]Pontificia Universidad Católica de Chile, Santiago, Chile

*Corresponding author: marciooliveira@utfpr.edu.br

Abstract

Rainfall data forecasting is essential in agricultural sciences due to impacts caused by water excess or deficit on crop growth. Our study aimed to develop a method to select rainfall forecast models using references with negligible error denoted as the gold standard. To this end, we used forecasting models from national centers such as Canadian Meteorological Center (CMC), European Center for Medium-Range Weather Forecasts (ECMWF), National Center for Environmental Prediction (NCEP), and Center for Weather Forecasting and Climate Studies (CPTEC). The study area comprised the western mesoregion of Paraná State (Brazil), and data were gathered from October to March between the soybean crop seasons of 2010/2011 and 2015/2016. Ten-day period clusters, corresponding to 240 h forecasts in the centers, were used to assess agreement with the gold standard. Our results showed that forecasting center selection must be based on rainfall value ranges and geographic locations. Selection according to the highest agreement with the gold standard was estimated at 76.9% for range 1 in CPTEC, 38.5% for range 2 and 4 in ECMWF, and 38.5% for range 3 in NCEP. In conclusion, the proposed method was efficient in selecting forecasting centers in areas of interest.

Keywords: agreement, spatial variation, rainfall forecast, model selection.
Abbreviations: CMC_Canadian Meteorological Center; ECMWF_European Center for Medium-Range Weather Forecasts; NCEP_National Center for Environmental Prediction; CPTEC_Center for Weather Forecasting and Climate Studies; EV_Virtual Station; $\hat{\rho}$_Estimated agreement degree; IC_95% confidence interval for population ρ; IC_{Inf}_lower limit of ρ; IC_{Sup}_upper limit of ρ.

Introduction

Water availability is essential during the developmental stages of soybeans (e.g., from germination to emergence and flowering to grain filling). Therefore, rainfall forecast should cover a period within which soybean yield can be affected by water stress (Wang et al., 2020; Seyoum et al., 2020; Islam et al., 2020). Crop monitoring and yield prediction models are influenced by agrometeorological conditions (Araya et al., 2015; Battisti et al., 2018). Accordingly, studies have been focused on agro-meteorological change effects on crop yield (Dastorani and Poormohammadi, 2016; Bao et al., 2015) and rainfall spatial variability (Carbone et al., 2003; Emmanuel et al., 2015; Zhang and Han, 2017).
Rainfall forecasts can be obtained by in-situ (agrometeorological) or virtual (atmospheric forecasting) stations. In-situ stations may be considered the gold standard for agrometeorological data but have limitations, such as lack of data records, generating data loss (Colston et al., 2018). By contrast, virtual stations have as advantage a better spatial and temporal data coverage. As these stations normally show their data in a regular grid, they have an improved spatial representation for agricultural modelling studies. Temporal resolution generally allows daily estimates without data losses, enabling them to be applied in several models (Mahlstein et al., 2015; Cáceres et al., 2018).
World climate centers have forecast models covering agrometeorological parameters such as rainfall. Parameterizations used in each center can produce different forecasts for each geographical location. Thus, reliable forecasts for crop monitoring, rainfall estimation, and yield prediction that are closer to the reality of each site require comparison with a reference, also known as 'the gold standard' ((Papamichail and Metaxa, 1996; Laurent, 1998; Choudhary and Nagaraja, 2005, Galea, 2013, Hermance and Sulieman, 2018).
The state of Paraná (Brazil) is characterized by farming activities and was responsible for 17% of the Brazilian grain production in the 2017/2018 crop season (IPARDES 2018). According to the Paraná Department of Agriculture and Supply/ Department of Rural Economy (SEAB/DERAL 2019), both mid-northern and western mesoregions reached the highest soybean yields, representing 23.4% and 15.9% for the 2010/2011 to 2013/2014 crop seasons, respectively. The study area of this research comprised the western mesoregion of Paraná State. We chose this region for being one of the largest soybean producers and having the largest number of available ANA meteorological stations. These factors contributed to the representativeness of the gold standard.
The spatial and temporal variability of rainfall has been addressed in several studies (Volpi et al., 2012; Mei et al.,

Table 1. Estimated agreements and lower and upper limits of 95% confidence intervals from selected centers in each range at 13 pixels covered by the western mesoregion of Paraná State.

	EV	Center	$\hat{\rho}$	IC_{Inf}	IC_{Sup}		Center	$\hat{\rho}$	IC_{Inf}	IC_{Sup}
Range 1	1	CPTEC	0.308	0.257	0.358	Range 2	CPTEC	0.351	0.288	0.415
	2	CPTEC	0.330	0.277	0.384		ECMWF	0.378	0.312	0.444
	3	CPTEC	0.316	0.266	0.366		ECMWF	0.396	0.326	0.466
	4	CPTEC	0.335	0.285	0.385		NCEP	0.347	0.279	0.416
	5	CPTEC	0.319	0.270	0.368		ECMWF	0.324	0.264	0.384
	6	CPTEC	0.338	0.289	0.387		ECMWF	0.382	0.311	0.454
	7	CPTEC	0.317	0.267	0.367		CMC	0.397	0.324	0.470
	8	CPTEC	0.327	0.276	0.378		CPTEC	0.367	0.293	0.441
	9	CPTEC	0.299	0.251	0.347		CMC	0.316	0.251	0.380
	10	NCEP	0.350	0.297	0.403		NCEP	0.355	0.281	0.428
	11	CPTEC	0.306	0.259	0.354		CPTEC	0.349	0.269	0.428
	12	CMC	0.367	0.311	0.422		CPTEC	0.332	0.263	0.402
	13	NCEP	0.338	0.288	0.387		ECMWF	0.369	0.300	0.437
Range 3	1	ECMWF	0.338	0.229	0.448	Range 4	CPTEC	0.384	0.204	0.564
	2	CPTEC	0.333	0.237	0.430		CMC	0.447	0.236	0.658
	3	CMC	0.343	0.235	0.452		CMC	0.344	0.183	0.504
	4	NCEP	0.372	0.257	0.486		ECMWF	0.331	0.156	0.505
	5	CPTEC	0.366	0.237	0.495		ECMWF	0.416	0.012	0.844
	6	CPTEC	0.293	0.189	0.398		CMC	0.381	0.125	0.638
	7	NCEP	0.338	0.236	0.441		ECMWF	0.441	0.274	0.608
	8	CMC	0.385	0.279	0.491		CPTEC	0.470	0.322	0.619
	9	NCEP	0.348	0.245	0.451		CMC	0.295	0.125	0.466
	10	NCEP	0.354	0.249	0.460		CPTEC	0.303	0.191	0.414
	11	ECMWF	0.407	0.297	0.518		ECMWF	0.325	0.207	0.444
	12	NCEP	0.330	0.241	0.420		CPTEC	0.428	0.285	0.571
	13	CMC	0.306	0.199	0.413		ECMWF	0.452	0.169	0.735

Note: EV: Virtual Station, Center: Selected center of forecast, $\hat{\rho}$: Estimated agreement degree, IC: 95% confidence interval for population ρ, IC_{Inf}: lower limit, IC_{Sup}: upper limit.

2017; Rinat et al., 2018; Oliveira et al., 2020). Some theoretical proposals to study rainfall variability have used geostatistical approaches (Li et al., 2008; Chappell et al., 2013; Jalili et al., 2020). These tools allow data interpolation by kriging the semivariogram model selected, using metrics such as mean error, reduced mean error, reduced error standard deviation, and absolute error (Adhikary, et al., 2017; Jalili et al., 2020). Furthermore, artificial intelligence approaches (e.g., K-nearest neighbor, and artificial and extreme machine learning neural networks) have been widely used to forecast rainfall. Yet, these techniques are focused on available data, with results obtained after algorithm training for selection process (Dash and Panigrahi, 2018; Azimi and Moghaddam, 2020).

The approaches described above have been widely applied in rainfall studies; however, they are not used to select models by relating the variability of reference and estimated data. In this sense, comparisons with the gold standard as proposed in this study contribute to comparison studies using agreement metrics to select models based on the relationship between the variability of reference data with that of data estimated by models. Given the above, our study aimed to propose a method to compare total rainfall forecasts of four climate centers Canadian Meteorological Center (CMC), European Center for Medium-Range Weather Forecasts (ECMWF), National Center for Environmental Prediction (NCEP), and Center for Weather Forecasting and Climate Studies (CPTEC) with database from gauge stations of the National Water Agency (ANA, acronym in Portuguese), between the months of October and March from the crop season of 2010/2011 until 2015/2016.

Results and Discussion

Missing data

From October to March during the crop seasons of 2010/2011 until 2015/2016, there were a predominance of soybean sowing, growing, and harvesting in Paraná State, which can be modelled by several techniques (Tatsumi et al., 2011; Fodor et al., 2017; Silva Fuzzo et al., 2020). Throughout the 977 10-day periods under study, we could observe that there was no loss of data from ANA, unlike what happened for data from the four centers, of which we found 2 10-day periods in CMC, 1 in ECMWF, 10 in NCEP, and 59 in CPTEC. It is worth mentioning that the missing 10-day periods were not considered in comparisons.

Selection of centers in each range

To compare rainfall forecasts and ANA dataset (gold standard), data were divided into four ranges according to gold-standard measurements. According to the highest estimated agreement for range 1, we obtained the following selections: 7.7% in CMC, 15.4% in NCEP, and 76.9% in CPTEC (Figs 1). For range 2, we obtained: 15.4% in CMC, 38.5% in ECMWF, 15.4% in NCEP, and 30.8% in CPTEC (Fig 1). For range 3, agreement selection showed: 23.1% in CMC, 15.4% in ECMWF, 38.5% in NCEP, and 23.1% in CPTEC (Figs 1). For range 4, we obtained: 30.8% in CMC, 38.5% in ECMWF, and 30.8% in CPTEC (Fig 1).

Fig 1. shows the ranges between 1 and 4 that indicate which model should be used at each location as data source for irrigation water management in cultivated areas (Querner et al., 2016; Acheampong et al., 2018). Fig 1. also highlights the agreement differences among ranges, as well as their distinct spatial variation for each center. These outcomes suggest that both forecast selection and calibration should

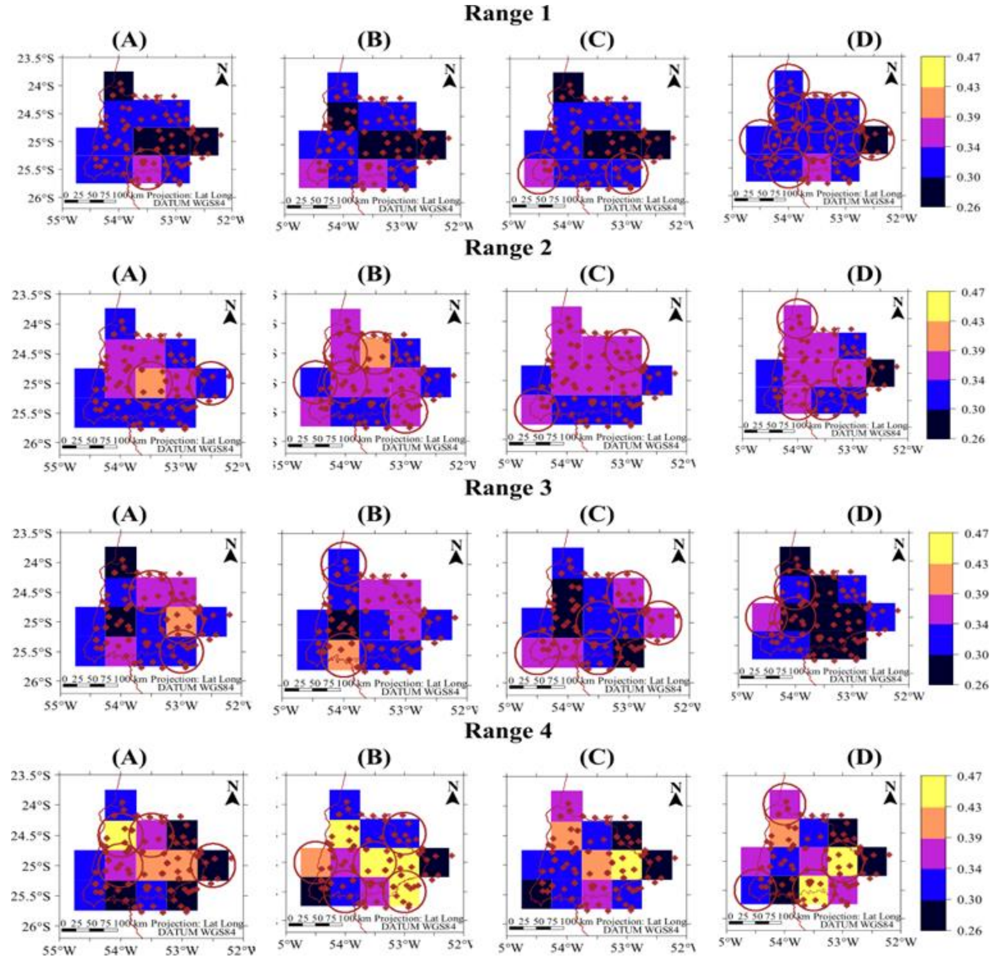

Fig 1. Agreement degrees of 10-day periods for ranges 1 to 4, between 'gold standard' measurements (ANA) and forecasts of the centers CMC (A), ECMWF (B), NCEP (C), and CPTEC (D). Circles indicate the forecast of the center reaching the highest agreement. The pixels in yellow and black are highest and lower agreement, respectively.

consider rainfall ranges and the geographic location of interest, as water availability is one of the main responsible factors for crop yield variability over time and space (Chomsang et al., 2020; Khalifa et al., 2020).

Ten-day forecasts should be incorporated into agricultural projects, considering water demands for each crop. Rainfalls between 650 and 700 mm distributed throughout the development cycle of soybeans can be considered enough to reach good yields and avoid losses (Yang et al., 2021; Dallagnol and Suzana, 2016). On the other side, water shortage from germination to emergence and from flowering to grain filling can lead to soybean yield losses (Fig 1).

Limits of confidence and variability of agreements

In this study, comparisons with the gold standard in each pixel for each of the four ranges aimed to assess the variability of this 'gold standard measure' with respect to the forecast errors of Lin (1989); Feng et al., (2015) and Chabert et al. (2019). Forecasts can be compared in a global context (Voisin et al., 2008; Zhao et al., 2020). Table 1 shows the selected centers, estimated agreement rates ($\hat{\rho}$), in addition to lower (L-inf) and upper (L-sup) limits within 95% confidence intervals, considering each range in each pixel. We observed that the highest rainfall rates in range 4 reached greater agreement (0.47) with CPTEC in pixel 8. The confidence interval allows inferring agreement between 0.322 and 0.619 at 95% reliability.

CMC had the lowest agreement in range 4 (0.295) and a 95% confidence interval from 0.125 to 0.466. These results highlight the spatial variability of agreements in Fig 1 from range 1 to 4. They also suggest that forecasts should be made only after calibration to remove bias and increase accuracy (Li et al., 2008). Therefore, selecting among the four forecasting centers (CMC, ECMWF, NCEP, and CPTEC) for each pixel should follow the indications in Table 1.

Materials and methods

Study Area

The state of Paraná is in southern Brazil and crossed by the Tropic of Capricorn and has a territorial extension of 199,709 km (IBGE 2010). The study area was the western mesoregion of this state since it is one of the producers of soybeans country-wise and has a territorial extension of 22,864.70 km^2. Climates prevailing in the state are humid temperate with hot summer, sub-humid with little water deficiency, mega thermal, and subtropical humid with dry winter.

Agrometeorological Data

Daily rainfall data were gathered from gauge stations belonging to the Brazilian National Water Agency (ANA), of which geographical locations are shown in Fig. 2. These data were grouped into 977 10-day periods. Moreover, daily rainfall forecasts from Canadian (CMC), European (ECMWF), North American (NCEP), and Brazilian (CPTEC) forecasting centers were obtained for the 240-h step in each virtual station (pixel), which have a rectangular coverage area of 0.5 º x 0.5º, with centroids shown in Fig 2 (Cunningham et al., 2015; Abedi et al., 2020).

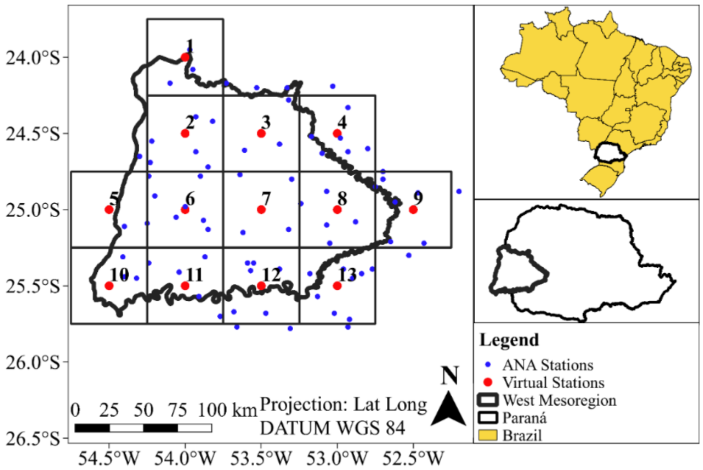

Fig 2. Location map of the western Paraná mesoregion, with ANA gauge stations and virtual stations of the CMC, ECMWF, NCEP, and CPTEC models.

Missing data were disregarded for matchings, and the used data were transformed according to Yeo and Johnson (2000). Thirteen pixels were obtained for the areas within western Paraná (55ºW, 52ºW, 26ºS, 23ºS), and data from a total of 75 ANA gauge stations were considered.

Rainfall means from ANA gauge stations within the virtual station coverage area (stations at a distance less than or equal to 0.36º from the pixel centroid) were considered for suitable spatial correspondence of the ANA stations with virtual ones. For comparisons, data were divided into four equal amplitude ranges.

Our study was carried out from October 1 to March 31 of 2015, for the crop seasons of 2010/2011 until 2015/2016. We selected this time of the year for rainfall forecasts because it is when soybeans are sown and harvested in Paraná State (SEAB/DERAL, 2019), as this crop is one of the economically important one in this Brazilian state.

Gold Standard Modeling

The comparison model used to evaluate agreement degree between two or more measuring instruments under presence of a 'gold-standard' or reference measure was proposed by Laurent (1998) and can be written in matrix notation as:

$$\boldsymbol{Y}_i = x_i \mathbf{1}_p + \boldsymbol{\epsilon}_i, \qquad (1)$$

where $\boldsymbol{Y}_i = \left(y_{i1}, \ldots, y_{ip}\right)^T$ is the measurement vector $p \times 1$ of the approximate methods in unit i, for i = 1,...,n, $\mathbf{1}_p$ is the $p \times 1$ vector of equal elements 1 and $\boldsymbol{\epsilon}_i = \left(\varepsilon_{i1}, \ldots, \varepsilon_{ip}\right)^T$ is the vector of random errors $p \times 1$ of p methods in unit i, for $i = 1, \ldots, n$.

The following formulation allows a solution considering $\boldsymbol{Z}_i = \left(x_i, \mathbf{Y}_i^T\right)^T$, the vector qx1, with q = p + 1 measurements made by the gold standard and the approximate methods in unit i. Thus, the random vectors $\boldsymbol{Z}_i$ for $i = 1, \ldots n$, are independent and identically distributed (iid) with $E(\boldsymbol{Z}_i) = \mu \mathbf{1}_q$ and $Var(\boldsymbol{Z}_i) = \boldsymbol{V}$, where $\boldsymbol{V}$ is a matrix of order $q \times q$ given by:

$$\boldsymbol{V} = \begin{pmatrix} \phi & \phi \mathbf{1}_p^T \\ \phi \mathbf{1}_p & \phi \mathbf{1}_p \mathbf{1}_p^T + \boldsymbol{\Sigma} \end{pmatrix}, \qquad (2)$$

where ϕ is the variance of the gold standard measure in the i-th unit, $\boldsymbol{\Sigma}$ is the $p \times p$ covariance matrix of the approximate p methods.

Assuming that the random vectors $\boldsymbol{Z}_i$for $i = 1, \ldots n$, are iid with $N_q\left(\mu \mathbf{1}_q, \boldsymbol{V}\right)$, a multivariate normal distribution with mean vector $\mu \mathbf{1}_q$ and covariance matrix $\boldsymbol{V}$, follows that the maximum likelihood estimator (MLE) of the parameters μ, ϕ and $\boldsymbol{\sigma}$ are given respectively by $\hat{\mu} = \bar{x}$, $\hat{\phi} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ and $\widehat{\boldsymbol{\sigma}} = \boldsymbol{S}_d = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{D}_i \boldsymbol{D}_i^T$, where $\bar{x} = \sum_{i=1}^{n} x_i / n$. Therefore, the MLE of the coefficient of agreement $\rho_j = \phi / (\phi + \sigma_{jj})$ for the agreement between the methods under test and the gold standard are given by $\hat{\rho}_j = \hat{\phi} / (\hat{\phi} + \hat{\sigma}_{jj})$, with $\hat{\sigma}_{jj} = \frac{1}{n}\sum_{i=1}^{n}\left(y_{ij} - x_i\right)^2$, (Laurent 1998).

According to Harris (1985) being $\boldsymbol{\gamma} = \left(\gamma_1, \ldots, \gamma_q\right)^T = \left(\phi, \sigma_{11}, \ldots, \sigma_{pp}\right)^T$ one has that $\sqrt{n}(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}) \xrightarrow{\mathcal{D}} N_q(\mathbf{0}, \boldsymbol{\Omega})$, in which the symbol $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution when $n \to \infty$ where $\boldsymbol{\Omega}$ is given by:

$$\boldsymbol{\Omega} = \begin{pmatrix} 2\phi^2 & \mathbf{0} \\ \mathbf{0} & 2\boldsymbol{\Sigma} \odot \boldsymbol{\Sigma} \end{pmatrix}, \qquad (3)$$

with $\odot$ denoting the product of Hadamard between matrices. Thus, follows the asymptotic distribution of $\widehat{\boldsymbol{\rho}} = \left(\hat{\rho}_1, \ldots, \hat{\rho}_p\right)^T$ given by:

$$\sqrt{n}(\widehat{\boldsymbol{\rho}} - \boldsymbol{\rho}) \xrightarrow{\mathcal{D}} N_p\left(\mathbf{0}, \boldsymbol{\Omega}_{\boldsymbol{\rho}}\right), \qquad (4)$$

where $\boldsymbol{\Omega}_{\boldsymbol{\rho}} = G \boldsymbol{\Omega} G^T$, $\boldsymbol{G} = \left(\frac{\partial \rho_i(\gamma)}{\partial \gamma_j}\right) = (1/\phi)(\boldsymbol{g}, \boldsymbol{G}_1)$ is a $p \times q$ matrix, $\boldsymbol{g} = \left(g_1, \ldots, g_p\right)^T$, with $g_j = \sigma_{jj}\rho_j^2/\phi$ e $\boldsymbol{G}_1 = -Diag\left(\rho_1^2, \ldots, \rho_p^2\right)$ a diagonal $p \times p$ matrix.

Considering $c_j = z_a$, where $a = 1 - (1 - \alpha_a)^{1/p}$ and z_a the $100(1 - \alpha/2)$ percentile of the standard normal distribution is obtained according to Sidak (1968) and a simultaneous confidence region of $100(1-\alpha)\%$ is given by:

$$\hat{\rho}_j - z_a\sqrt{\frac{\hat{v}_{jj}}{n}} \leq \rho_j \leq \hat{\rho}_j + z_a\sqrt{\frac{\hat{v}_{jj}}{n}}, \qquad (5)$$

where $\hat{v}_{jj}$ are the MLE of v_{jj}, diagonal elements of the $\boldsymbol{\Omega}_{\boldsymbol{\rho}}$ matrix given by Equation (4).

The software R (R DEVELOPMENT CORE TEAM, 2020) was used to devise routine and obtain outputs of equations (1)-(9), which specify the model used to describe agreement between 'gold standard' measurements of rainfall and the forecasts of centers under study. Both agreement degrees and confidence intervals were obtained considering four defined ranges, assuming an amplitude between the highest and the lowest gold standard pixel measurements. This model has not been applied to the selection of rainfall forecasting centers, but it can be very useful for this purpose, a relevant topic for the management of agricultural resources, not only in the Paraná region, Brazil, but also for other regions of the world, that strongly depend on agricultural development.

Conclusion

Choosing a forecasting center for soybean yield estimation or forecasting models should consider rainfall range and geographic location of interest, regardless of the center (CMC, ECMWF, NCEP, and CPTEC). Ranges with higher rainfall values showed greater agreement with the gold standard (ANA), which indicates that the forecast centers CMC, ECMWF, NCEP, and CPTEC, and their models have the most suitable parameterization for detection of extreme events. By comparing gauge station values and center forecasts, additional care should be taken to obtain a gold standard measure of rainfall with representativeness for the events at a given location and within a certain time range, as in the case of 10-day periods. Spatial correspondence of the gold standard measurements with forecasts can be obtained by considering *in-situ* stations within the center circle at the pixel centroid. The radius should cover the entire pixel with forecasts of the centers CMC, ECMWF, NCEP, and CPTEC.

Acknowledgements

The authors are grateful for the partial financial support from UTFPR, UNIOESTE / PGEAGRI, Coordination for the Improvement of Higher Level Personnel – Brazil (CAPES) – Finance Code 001, National Council for Scientific and Technological Development (CNPq), and FONDECYT Chile (Project No. 1150325) and Spatial Statistics Laboratory (LEE/UNIOESTE).

References

Abedi M, Shafizadeh-Moghadam, H, Morid, S, Booij, MJ, Delavar, M (2020) Evaluation of ECMWF mid-range ensemble forecasts of precipitation for the Karun River basin. Theoretical and Applied Climatology. 141(1):61–70.

Acheampong D, Balana BB, Nimoh F, Abaidoo RC (2018) Asssesing the effectiveness and impact of agricultural water management interventions: the case of small reservoirs in northern Ghana. Agricultural Water Management. 209:163–170.

Adhikary SK, Muttil N, Yilmaz AG (2017) Cokriging for enhanced spatial interpolation of rainfall in two Australian catchments. Hydrological Processes. 31(12):2143–2161.

Araya A, Hoogenboom G, Luedeling E, Hadgu KM, Kisekka I, Martorano LG (2015) Assessment of maize growth and yield using crop models under present and future climate in southwestern Ethiopia. Agricultural and Forest Meteorology. 15:252–265.

Azimi S, Azhdary MM (2020) Modeling Short Term Rainfall Forecast Using Neural Networks, and Gaussian Process Classification Based on the SPI Drought Index. Water Resources Management. 34(4):1369–1405.

Bao Y, Hoogenboom G, McClendon RW, Paz JO (2015) Potential adaptation strategies for rainfed soybean production in the south-eastern USA under climate change based on the CSM-CROPGRO-Soybean model. The Journal of Agricultural Science. 153(5):798–824.

Battisti R, Sentelhas PC, Boote KJ (2018) Sensitivity and requirement of improvements of four soybean crop simulation models for climate change studies in Southern Brazil. International Journal of Biometeorology. 62(5):823–832.

Cáceres M, Martin-StPaul N, Turco M, Cabon A, Granda V (2018) Estimating daily meteorological data and downscaling climate models over landscapes. Environmental Modelling & Software. 108:186–196.

Carbone GJ, Kiechle W, Locke C, Mearns LO, McDaniel L, Downton MW (2003) Response of Soybean and Sorghum to Varying Spatial Scales of Climate Change Scenarios in the Southeastern United States. Climatic Change. 60(1):73–98.

Chabert A, Amossé A, Sarthou JP (2019) Assessing landscape composition using visual assessment: accuracy of rapid description compared to digital mapping. Landscape Research, 44(1): 6–18.

Chappell A, Renzullo LJ, Raupach TH, Haylock M (2013) Evaluating geostatistical methods of blending satellite and gauge data to estimate near real-time daily rainfall for Australia. Journal of Hydrology. 493:105–114.

Chomsang K, Morokuma M, Toyota M (2020) Dry matter production and physiological responses to a wide range of irrigation levels in two Japanese soybean cultivars. Plant Production Science. 23(4): 490–503.

Choudhary PK, Nagaraja HN (2005) Selecting the instrument closest to a gold standard. Journal of Statistical Planning and Inference. 129(1): 229–237.

Colston JM, Ahmed T, Mahopo C, Kang G, Kosek M, Sousa Junior F, Shrestha PS, Svensen E, Turab A, Zaitchik B (2018) Evaluating meteorological data from weather stations, and from satellites and global models for a multi-site epidemiological study. Environmental Research. 165:91–109.

Cunningham C, Bonatti JP, Ferreira M (2015) Assessing improved CPTEC probabilistic forecasts on medium-range timescale. Meteorological Applications. 22(3):378–384.

Dallagnol LC, Suzana CS (2016) Soybean yield potential estimated in a central region of RS State, Brazil. Científica. 44(4):584-591.

Dash Y, Mishra SK, Panigrahi BK (2018) Rainfall prediction for the Kerala state of India using artificial intelligence approaches. Computers & Electrical Engineering. 70:66–73.

Dastorani MT, Poormohammadi S (2016) Mapping of climatic parameters under climate change impacts in Iran. Hydrological Sciences Journal. 61(14): 2552–2566.

Donner A (1986) A Review of Inference Procedures for the Intraclass Correlation Coefficient in the One-Way Random

Effects Model. International Statistical Review/Revue Internationale de Statistique. 54(1):67–82.
Emmanuel I, Andrieu H, Leblois E, Janey N, Payrastre O (2015) Influence of rainfall spatial variability on rainfall–runoff modelling: Benefit of a simulation approach? Journal of Hydrology. 531:337–348.
Feng D, Baumgartner R, Svetnik V (2015) A Robust Bayesian Estimate of the Concordance Correlation Coefficient. Journal of Biopharmaceutical Statistics. 25(3): 490–507.
Fodor N, Challinor A, Droutsas I, Ramirez-Villegas J, Zabel F, Koehler AK, Foyer CH (2017) Integrating Plant Science and Crop Modeling: Assessment of the Impact of Climate Change on Soybean and Maize Production. Plant and Cell Physiology. 58(11):1833–1847.
Harris P (1985) Testing for Variance Homogeneity of Correlated Variables. Biometrika. 72(1): 103–107.
Hermance JF, Sulieman HM (2018) Adequacy of the daily TMPA 3B42 high-resolution satellite precipitation product for monitoring hydrometeorological hazards in the Southeast Sahel of Africa. International Journal of Remote Sensing. 39(8):2579–2596.
IBGE (2010) Instituto Brasileiro de Geografia e Estatística. Censo Demográfico. Brasil.
IPARDES (2018). Instituto Paranaense de Desenvolvimento Econômico e Social. BDEweb - Base de Dados do Estado. Curitiba, Paraná.
Islam MS, Roy S, Afrin R, Mia MY (2020) Influence of climate-induced disasters and climatic variability on cropping pattern and crop production in Bangladesh. Environment, Development and Sustainability. 22(7):6709–6726.
Jalili F, Modarres R (2020) Geostatistical and deterministic methods for rainfall interpolation in the Zayandeh Rud basin, Iran. Hydrological Sciences Journal. 65(16): 2678-2692
Khalifa M, Elagib NA, Ahmed BM, Ribbe L, Schneider K (2020) Exploring socio-hydrological determinants of crop yield in under-performing irrigation schemes: pathways for sustainable intensification. Hydrological Sciences Journal. 65(2):153–168.
Laurent RTSt (1998) Evaluating Agreement with a Gold Standard in Method Comparison Studies. Biometrics. 54(2):537–545.
Li B, Eriksson M, Srinivasan R, Sherman M (2008) A geostatistical method for Texas NexRad data calibration. Environmetrics. 19(1):1–19.
Lin LIK (1989) A Concordance Correlation Coefficient to Evaluate Reproducibility. Biometrics. 45(1): 255–268.
Mahlstein I, Spirig C, Liniger MA, Appenzeller C (2015) Estimating daily climatologies for climate indices derived from climate model data and observations. Journal of Geophysical Research: Atmospheres. 120(7):2808–2818.
Mei Y, Shen X, Anagnostou EN (2017) A synthesis of space--time variability in multicomponent flood response. Hydrology and Earth System Sciences. 21(5):2277–2299.
Oliveira MP, Uribe-Opazo MA, Galea M, Johann JA (2020) Concordance Modeling With a Gold Standard for Variables From the Three-Parameter Gamma Distribution. Journal of Agricultural Studies. 8(2): 284–305.
Papamichail DM, Metaxa IG (1996) Geostatistical analysis of spatial variability of rainfall and optimal design of a rain gauge network. Water Resources Management. 10(2):107–127.
Querner EP, Froebrich J, Gallart F, Cazemier MM, Tzoraki O (2016) Simulating streamflow variability and aquatic states in temporary streams using a coupled groundwater-surface water model. Hydrological Sciences Journal. 61(1):146–161.
R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rinat Y, Marra F, Zoccatelli D, Morin E (2018) Controls of flash flood peak discharge in Mediterranean basins and the special role of runoff-contributing areas. Journal of Hydrology. 565:846–860.
SEAB/DERAL - Secretaria da Agricultura e do Abastecimento do Paraná/Departamento de Economia Rural (2019) Banco de Dados da Produção Agropecuária no Paraná. Available in https://www.agricultura.pr.gov.br. Access in 08/2019.
Sidak Z (1968) On Multivariate Normal Probabilities of Rectangles: Their Dependence on Correlations. The Annals of Mathematical Statistics. 39(5):1425–1434.
Silva Fuzzo DF, Carlson TN, Kourgialas NN, Petropoulos GP (2020) Coupling remote sensing with a water balance model for soybean yield predictions over large areas. Earth Science Informatics. 13(2):345–359.
Tatsumi K, Yamashiki Y, Silva VR, Takara K, Matsuoka Y, Takahashi K, Maruyama K, Kawahara N (2011) Estimation of potential changes in cereals production under climate change scenarios. Hydrological Processes. 25(17):2715–2725.
Voisin N, Wood AW, Lettenmaier DP (2008) Evaluation of Precipitation Products for Global Hydrological Prediction. Journal of Hydrometeorology. 9(3):388–407.
Volpi E, Di Lazzaro M, Fiori A (2012) A simplified framework for assessing the impact of rainfall spatial variability on the hydrologic response. Advances in Water Resources. 46:1–10.
Wang J, Liu G, Zhu C (2020) Evaluating precipitation products for hydrologic modeling over a large river basin in the Midwestern USA. Hydrological Sciences Journal. 65(7):1221–1238.
Yang W, Feng G, Adeli A, Tewolde H, Qu Z (2021) Simulated long-term effect of wheat cover crop on soil nitrogen losses from no-till corn-soybean rotation under different rainfall patterns. Journal of Cleaner Production. 280: 124255.
Yeo IK, Johnson RA (2000) A New Family of Power Transformations to Improve Normality or Symmetry. Biometrika. 87(4):954–959.
Zhang J, Han D (2017) Assessment of rainfall spatial variability and its influence on runoff modelling: A case study in the Brue catchment, UK. Hydrological Processes. 31(16):2972–2981.
Zhao T, Chen H, Xu W, Cai H, Yan D, Chen X (2020) Spatial association of anomaly correlation for GCM seasonal forecasts of global precipitation. Climate Dynamics. 55(7):2273–2286.