

## Comparing GEFS, ECMWF, and Postprocessing Methods for Ensemble Precipitation Forecasts over Brazil

HANOI MEDINA AND DI TIAN

*Department of Crop, Soil and Environmental Sciences, Auburn University, Auburn, Alabama*

FABIO R. MARIN

*Department of Biosystems Engineering, Escola Superior de Agricultura Luiz de Queiroz, University of São Paulo, São Paulo, Brazil*

GIOVANNI B. CHIRICO

*Department of Agricultural Sciences, Water Resources Management and Biosystems Engineering Division, University of Naples Federico II, Portici, Naples, Italy*

(Manuscript received 10 June 2018, in final form 20 February 2019)

### ABSTRACT

This study compares the performance of Global Ensemble Forecast System (GEFS) and European Centre for Medium-Range Weather Forecasts (ECMWF) precipitation ensemble forecasts in Brazil and evaluates different analog-based methods and a logistic regression method for postprocessing the GEFS forecasts. The numerical weather prediction (NWP) forecasts were evaluated against the Physical Science Division South America Daily Gridded Precipitation dataset using both deterministic and probabilistic forecasting evaluation metrics. The results show that the ensemble precipitation forecasts performed commonly well in the east and poorly in the northwest of Brazil, independent of the models and the postprocessing methods. While the raw ECMWF forecasts performed better than the raw GEFS forecasts, analog-based GEFS forecasts were more skillful and reliable than both raw ECMWF and GEFS forecasts. The choice of a specific postprocessing strategy had less impact on the performance than the postprocessing itself. Nonetheless, forecasts produced with different analog-based postprocessing strategies were significantly different and were more skillful and as reliable and sharp as forecasts produced with the logistic regression method. The approach considering the logarithm of current and past reforecasts as the measure of closeness between analogs was identified as the best strategy. The results also indicate that the postprocessing using analog methods with long-term reforecast archive improved raw GEFS precipitation forecasting skill more than using logistic regression with short-term reforecast archive. In particular, the postprocessing dramatically improves the GEFS precipitation forecasts when the forecasting skill is low or below zero.

### 1. Introduction

Precipitation is an important source of water resources and a major driving factor in the functioning of agriculture, forest, and freshwater ecosystems. Accurate precipitation forecasting is one of the most sensible aspects of weather prediction to the society. It strongly affects daily decisions in different sectors, such as public health, water resources, energy production, agriculture, and environmental protection. Numerical weather prediction (NWP) models are the state-of-art technology for forecasting

medium-range precipitation at daily or subdaily time step over the globe. Practically every aspect of the NWP has dramatically improved (Hamill et al. 2013) over the last decades, which has led to significant increments in the skill of the model forecasts (Bauer et al. 2015), and has encouraged their use in a wide range of applications.

NWP has global applicability (Bauer et al. 2015) and potential for improving regional precipitation, runoff, and water storage forecasting over the globe (e.g., Hamill 2012; Su et al. 2014; He et al. 2010; Cloke and Pappenberger 2009). However, few studies have focused on assessing the NWP precipitation predictability associated with large and intense mesoscale convective systems

Corresponding author: Di Tian, tiandi@auburn.edu

(Bechtold et al. 2012), such as tropical rainfall. Atmospheric convection is an essential process for understanding and modeling the weather dynamics over the tropics (Bony et al. 2015), but one very difficult to analytically represent in global NWP models (Bauer et al. 2015). The representation of these processes is especially challenging over continental areas from the Southern Hemisphere where the abundant vegetation and the sparse observations for evaluation and data assimilation have limited the models' accuracy. Recent progress in forecasting tropical convection (Bechtold et al. 2014; Subramanian et al. 2017) and the increasing quantity and quality of global information encourage the use of NWP for tropical precipitation forecasting. It is therefore necessary to conduct comprehensive assessments of the NWP's ability to forecast heavy and highly variable rainfall regimes in tropical and near-tropical regions dominated by large mesoscale convective systems (Mohr and Zipser 1996).

The National Centers for Environmental Prediction (NCEP) Global Ensemble Forecast System (GEFS) and the European Centre for Medium-Range Weather Forecasts (ECMWF) are two leading NWPs for medium-range weather forecasting at the global scale. In particular, the ECMWF global ensemble precipitation forecasts have consistently been the most skillful in Northern Hemisphere regions (Atger 2001; Hamill et al. 2008; Su et al. 2014) compared with those produced by other global NWP models. An advantage of the GEFS model is that it archives retrospective forecast (reforecast) datasets for long past periods at no cost, which are useful for statistically postprocessing to correct weather forecasts against observed data, thus reducing the uncertainty and improving forecast performance (Hamill et al. 2006; Hagedorn et al. 2008). Statistical postprocessing methods are commonly effective to amend systematic inconsistencies in forecasts, while taking into account actual modes of the spatial weather variability that are impossible to represent in NWPs (e.g., Glahn and Lowry 1972; Gneiting 2014; Pelosi et al. 2017). However, little is yet known about the relative performance of GEFS and ECMWF precipitation forecasts, and the effectiveness of the statistical postprocessing over the tropical and near-tropical regions dominated by large and intense mesoscale convective systems.

Analog-based postprocessing methods are an efficient approach to improve probabilistic precipitation forecasts (Voisin et al. 2010; Ben Daoud et al. 2016) and in general several other hydrometeorological forecasts (Tian and Martinez 2012, 2014). In one analog-based implementation, the current forecast from a fixed NWP is compared against the past forecasts of the same NWP at a similar time of the year within a limited region, and an ensemble is formed considering the observations on

the dates of the closest matches (Hamill et al. 2006). Studies have explored different strategies for implementing analog methods with GEFS reforecast, such as testing different similarity criteria (Hamill and Whitaker 2006), and multivariate (Hamill and Whitaker 2006; Delle Monache et al. 2011, 2013) versus univariate similarity metrics, and evaluating different sizes of the search region (Hamill and Whitaker 2006; Hamill et al. 2015; Tian and Martinez 2012, 2014) and number of ensemble members (Hamill et al. 2015). Nevertheless, guidelines regarding the optimal implementing strategies to efficiently postprocessing tropical convective precipitations are still lacking.

A disadvantage of the analog approaches is that it needs long-term reforecasts for finding the closest matching analogs. When the forecasted precipitation is a large, rare event, it becomes a challenge to find sufficient number of analogs if the reforecast archive is not sufficiently long enough (Hamill et al. 2015). There are alternative approaches that are less reliant on the size of the training data. The logistic regression method is one of these methods and has been found suitable for dealing with medium-range precipitation forecasts in several regions (Wilks 2006; Wilks and Hamill 2007). Few previous studies have compared the relative performance of analog techniques and logistic regression techniques for postprocessing GEFS precipitation forecasts. For selecting optimal postprocessing methods, it would be informative to compare the performance of analog methods, which requires long-term reforecast archives, with logistic regression, which only needs a small set of training data.

Given the research gaps we have identified, this study aims to 1) document the performance of the GEFS and ECMWF daily precipitation ensemble forecasts using Brazil as case study, 2) evaluate the GEFS-based precipitation forecasts postprocessed using analog methods with different strategies, and 3) compare the performance of analog-based methods with the logistic regression method.

Brazil covers a large area and is considerably affected by large and intense mesoscale convective systems within which severe weather events develop (Mohr and Zipser 1996). The complexity of the spatial and temporal variability of rainfall patterns over Brazil may provide a unique setting for assessing progresses of global scale NWPs and postprocessing techniques for rainfall prediction.

## 2. Data and methods

### a. Study region

Brazil is one of the mega-diverse countries and the world's fifth-most populous. It has the second-largest forest area in the world (FAO 2015), is a country with high risks of

transmission of water-borne diseases (e.g., Guerrant et al. 1983), is one of the top hydropower-potential countries (Zhou et al. 2015), and is one of the world's main producers of food and biofuels (Ferreira et al. 2012). It ranks first in production of sugarcane, coffee, and oranges and sixth in the world's cereal production (FAO 2015). Given the significant impact of precipitation in those sectors, forecasting medium-range daily precipitation for Brazil has great implications for its agriculture, natural resources, hydropower generation, and public health management. The study focused on the six major natural biomes of Brazil: Amazon, Caatinga, Cerrado, Atlantic Forest, Pampa, and Pantanal, representing climatologically consistent regions (Fig. 1). A brief description of each biome is provided as follows:

- 1) The Brazilian Amazon covers around 4 million  $\text{km}^2$  (almost half the national territory), representing 69% of the Amazon basin. Annual rainfall is generally above 2000 mm and decreases from the equatorial regions toward the tropics and the northeast of Brazil (under 1500 mm).
- 2) Caatinga is among the semiarid regions with larger population and biodiversity in the world (MMA 2011). Annual rainfall is commonly less than 750 mm (Leal et al. 2005), and it is highly variable (Moura and Shukla 1981). The region experiments a peculiar intra-annual rainfall regime, with a maximum in March–April over the north and the center part and in November–March over the southern part.
- 3) Cerrado is a tropical savanna covering 22% of Brazil's territory. The overall amount of rain is usually between 800 and 2000  $\text{mm yr}^{-1}$  (Ratter et al. 1997), mostly distributed between October and April.
- 4) Atlantic Forest is the second-largest rain forest of the American continent and one of the world's regions hosting the biggest biodiversity. Annual rainfall is between 1000 and 3000 mm.
- 5) The Brazilian Pampa represents 2.07% of the national territory and lies within the South Temperate Zone (Roesch et al. 2009). The annual precipitation in the region is around 1200–1600 mm.
- 6) The Pantanal wetland is a complex of seasonally inundated floodplains along the upper Paraguay River, located mostly in Brazil (Hamilton 2002). Annual rainfall is 1000–1500 mm across the basin, with most rainfall occurring between November and March.

#### b. Verification dataset

The choice of the verification dataset is important in the context of medium-range forecasting, especially in data-sparse regions affected by complex patterns of variability. Using gridded data based on rain gauge

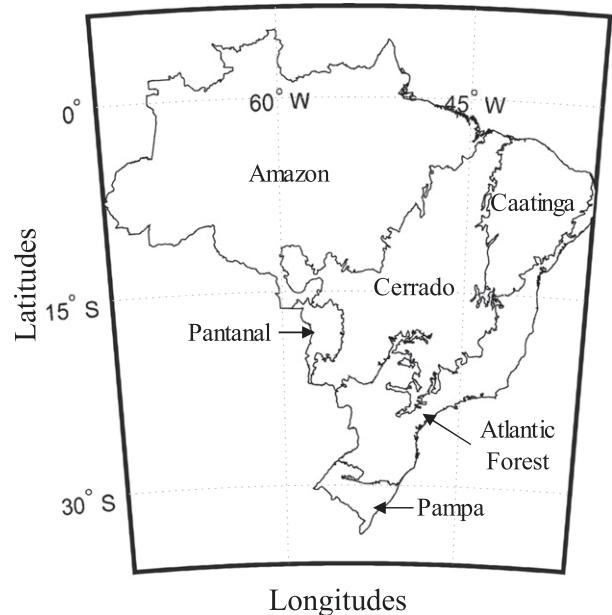


FIG. 1. Regions of Brazil involved in this study corresponding to the six major natural biomes as defined in IBGE (2004).

observations has the advantage of being independent of all models (Hagedorn et al. 2012). Carvalho et al. (2012) found that the Physical Science Division South America Daily Gridded Precipitation dataset (Liebmann and Allured 2005; Liebmann and Allured 2006) consistently represents the variability of the South American monsoon system, which is the most important climatic feature in South America, and provides a similar spatial pattern of mean precipitation compared with other gridded precipitation products such as the Global Precipitation Climatology Project (Huffman et al. 2001) and Climate Prediction Center unified gauge (Silva et al. 2007). In this study, we use this dataset for evaluating rainfall forecasts over each biome in Brazil. It consists of  $1^\circ \times 1^\circ$  grid of daily precipitation values over Brazil over 1985–2010, interpolated using the average of rain gauge records within a geographic ellipse. Measurements have been taken at 1200 UTC, while precipitation is recorded as having occurred on the day on which the rain gauge reading is taken. This dataset is available at [http://www.esrl.noaa.gov/psd/data/gridded/data.south\\_america\\_precip.html](http://www.esrl.noaa.gov/psd/data/gridded/data.south_america_precip.html). Figure 2 shows the cumulative probabilistic distribution of the daily precipitations over each biome generated from the verification dataset.

It is worth noting that at least two other datasets based on rain gauge observations are available for Brazil. Silva et al. (2007) produced the Climate Prediction Center unified gauge. This is a  $1^\circ \times 1^\circ$  dataset using a Cressman (1959) scheme of interpolation (Glahn et al. 1985) that

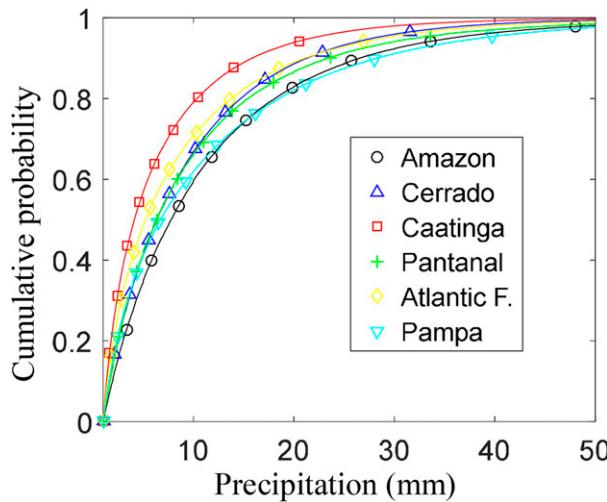


FIG. 2. Cumulative distribution of the precipitation higher than 1 mm over 1985–2010.

corrects the background gridpoint value by a linear combination of residuals between calculated and observed values. However, this dataset has fewer rain gauges over the Brazilian Amazonian domain compared to the adopted dataset (see Fu et al. 2013). Recently, Xavier et al. (2016) produced a high-resolution dataset over a  $0.25^\circ \times 0.25^\circ$  grid based upon the inverse distance weighting interpolation method; this method had been identified as the most skillful when compared against several other methods. However, the grid coordinates in this dataset do not coincide with the grid coordinates of the forecast datasets in our study, meaning that a further interpolation would be needed to use it as our verification dataset.

### c. Forecast datasets

#### 1) GEFS REFORECAST DATA

We used  $1^\circ \times 1^\circ$  gridded reforecasts over 1985–2010 (26 years) produced from the second-generation global medium-range ensemble reforecast dataset (Hamill et al. 2013). This a retrospective weather forecast dataset generated with the currently operational NCEP GEFS, available at <http://esrl.noaa.gov/psd/forecasts/reforecast2/download.html>. The daily precipitation ensemble reforecasts considered both the control forecast and the 10 perturbed forecasts issued at 0000 UTC at 1.5-, 3.5-, and 5.5-day leads. A lead time of 1.5 days matches up the observation of day  $n$  with the sum of the 6-h total precipitation at 18, 24, 30, and 36 h of the forecast issued at day  $n - 1$ .

#### 2) ECMWF FORECASTS DATA

ECMWF reforecasts archived in the TIGGE database at ECMWF (see <http://apps.ecmwf.int/datasets/data/tigge>)

were also considered. We used the 50-member ensembles of perturbed ECMWF forecasts issued at 0000 UTC over October 2006–10 at lead times of 1.5, 3.5, and 5.5 days. Forecasts, which originally have a horizontal resolution of about 32 km, were converted into a  $1^\circ \times 1^\circ$  grid before downloading, using the software available at the ECMWF website. About 2.0% of the records accounted for negative, mostly negligible values, which were set to zero. Probabilities were also calculated directly from the ensemble relative frequency

### d. Postprocessing methods

#### 1) THE ANALOG FORECAST METHOD

In the analog forecast method, the real-time forecast is adjusted using a long time series of past forecasts and associated observations (Hamill et al. 2015). Suppose that we want to produce an ensemble of  $n$  analog forecasts for today's forecast at a specific point and a given lead. The first step is to compare today's forecasts within a region surrounding that point with the forecasts from the historical reforecast archive in that same region and at the same forecast lead, and then find the  $n$  dates with the best matching. In a second step, the analog ensemble is formed from the verification dataset on those dates. This process is repeated for each lead day and location across the study region, and the forecast over the grid point is produced by grouping together the analog forecasts (Hamill et al. 2006; Tian and Martinez 2014). Notice that in some analog-based applications the resolution of the verification dataset is considerably larger than that of the forecast dataset, such that the search region is linked to a tile of verification data points rather than just one point. In applications like these, the procedure of tiling the analog forecasts may lead to spatial inconsistencies at the boundaries between tiles (Hamill and Whitaker 2006; Hamill et al. 2006). However, this is not an issue present in this study. Leave-one-out cross validation are carried out by excluding the current year from the list of potential analogs. For a detailed description and theoretical basis of the analog method, the readers can refer to Hamill and Whitaker (2006).

#### 2) LOGISTIC REGRESSION METHOD

In the logistic regression (LR) method a nonlinear function is fitted to past pairs of the predictor(s), and the predictand, which as an observed value takes on a probability of either 1.0 (event occurred) or 0.0 (event did not occur), according to the adopted threshold  $T$  (Wilks 2006). The fitted function is then used to estimate the probability  $P$  that the current unknown observed amount  $O$  be higher than the threshold  $T$  given the current predictor values, associated to the forecast.

TABLE 1. Configurations of the six analog approaches. Parameters  $F_{\text{pr}}^{i,t}$  and  $F_{\text{pw}}^{i,t}$  are the 24-h cumulative precipitation (pr) and the total-column precipitable water (pw) forecasts, respectively, at time  $t$  and over grid point  $i$ , while  $F_{\text{pr}}^{i,\text{tc}}$  and  $F_{\text{pw}}^{i,\text{tc}}$  are the corresponding forecasts at current time (tc) (involving the current grid point and the set of  $N_s$  supplemental points surrounding the current grid point) and time  $t$ .

| ID_Method    | Ensemble size | Grid points | Closeness metric  |
|--------------|---------------|-------------|---|
| Control      | 50            | 9           | $\sum_{i=1}^{N_s+1} (F_{\text{pr}}^{i,t} - F_{\text{pr}}^{i,\text{tc}})^2$  |
| Short_reg    | 50            | 5           | $\sum_{i=1}^{N_s+1} (F_{\text{pr}}^{i,t} - F_{\text{pr}}^{i,\text{tc}})^2$  |
| 100_Ens      | 100           | 9           | $\sum_{i=1}^{N_s+1} (F_{\text{pr}}^{i,t} - F_{\text{pr}}^{i,\text{tc}})^2$  |
| LogF         | 50            | 9           | $\sum_{i=1}^{N_s+1} [\log(F_{\text{pr}}^{i,t} + 1) - \log(F_{\text{pr}}^{i,\text{tc}} + 1)]^2$  |
| Prec_water01 | 50            | 9           | $0.9 \times \sqrt{\sum_{i=1}^{N_s+1} (F_{\text{pr}}^{i,t} - F_{\text{pr}}^{i,\text{tc}})^2} + 0.1 \times \sqrt{\sum_{i=1}^{N_s+1} (F_{\text{pw}}^{i,t} - F_{\text{pw}}^{i,\text{tc}})^2}$ |
| Prec_water05 | 50            | 9           | $0.5 \times \sqrt{\sum_{i=1}^{N_s+1} (F_{\text{pr}}^{i,t} - F_{\text{pr}}^{i,\text{tc}})^2} + 0.5 \times \sqrt{\sum_{i=1}^{N_s+1} (F_{\text{pw}}^{i,t} - F_{\text{pw}}^{i,\text{tc}})^2}$ |

In this study we adopted the same nonlinear function as [Hamill et al. \(2008\)](#):

$$P(O > T) = 1 - 1/[1 + \exp(a + b\bar{F}_{\text{pr}}^{0.25} + c\sigma_{F_{\text{pr}}}^{0.25})], \quad (1)$$

where  $\bar{F}_{\text{pr}}$  and  $\sigma_{F_{\text{pr}}}$  represent the mean and the standard deviation of the ensemble of precipitation forecast, respectively, while  $a$ ,  $b$ , and  $c$  are the fitting parameters. Following [Hamill and Whitaker \(2006\)](#), we also pondered using a one-half power transformation of the predictors, instead of the one-quarter adopted here, but our results were practically the same.

As for the analog method, the logistic regression technique is performed separately for each location and each forecast lead time, within the verification period using all historical data available.

#### e. Experimental design

The first experiment is to compare the performance of GEFS and ECMWF raw forecasts, as well as GEFS analog forecasts over January, April, July, and October, from October 2006 to December 2010. These four months are representative of the summer, fall, winter, and spring season, respectively. In this and the subsequent experiments, training of the GEFS forecasts considered the 26-yr dataset of retrospective forecasts. The analog forecasts for a current date and time were formed by finding the lowest sum of the square differences (ssd) between the current forecasts and the similar historical forecasts in the other years (25 in total) from the reforecast archive, considering a limited region of 9 grid points. Preliminary analyses showed that by considering tiles of 21 grid points (i.e., those located at a distance lower than  $2\sqrt{2}^{\circ}$  with respect to the current grid point), the results tended to be slightly worse compared

with 9 grid points, especially at 1.5 and 3.5 days and during rainy days. [Hamill et al. \(2015\)](#) suggested using relatively small grids for short lead times and under scenarios of heavy rainfalls. The forecasts were selected within a  $\pm 45$ -day window around the date of the forecast, and the best 50 analogs were chosen to construct the forecast ensemble. The size of the window takes into consideration the rainfall seasonality over the region (e.g., [Espinoza Villar et al. 2009](#)), and the fact that the forecast bias might considerably change across seasons. A size of the ensemble of 50 members seems reasonable following the results of [Hamill et al. \(2015\)](#); in our study it represents just the 2% of the search period (of 91 days  $\times$  25 years). This analog procedure is adopted as the control variant of the method and referred henceforth as “Control” forecast.

The second experiment is to conduct an intercomparison among six GEFS-based analog approaches and one logistic regression method. In this case, the forecasts were verified over January, April, July, and October from 1985 to 2010. The six analog-based methods included the Control method plus five modified versions of this procedure (see Table 1), with each version considering only one modification with respect to the Control procedure. Each method is described as follows:

- 1) Short\_reg considered a search region with five grid points, that is, the current grid point and the four adjacent grid points at a distance of  $1^{\circ}$ .
- 2) The 100\_Ens was produced with 100 analog members, instead of only 50.
- 3) LogF considered the differences between the logarithm of the current and past precipitation forecasts plus one, as the measure of the closeness among forecasts.
- 4) Prec\_water01 included the mean ensemble of the column precipitable water as a predictor variable.

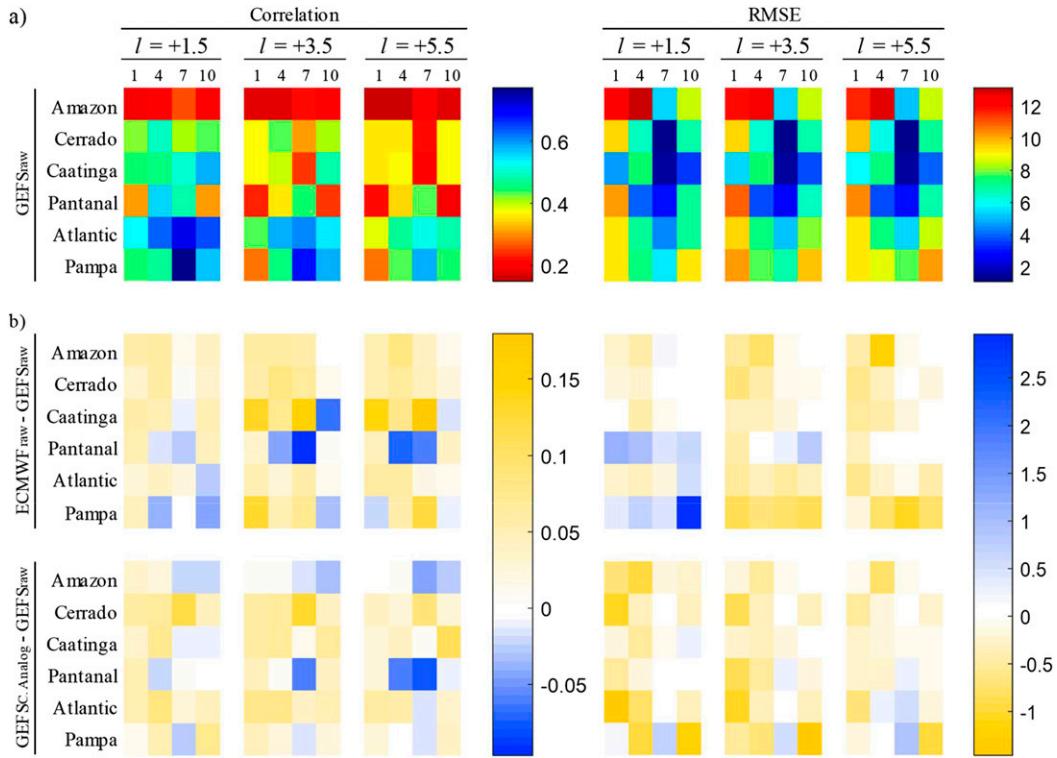


FIG. 3. (a) Correlation and RMSE (mm) of the GEFS precipitation raw forecasts at each biome in January (1), April (4), July (7), and October (10), for lead times of 1.5, 3.5, and 5.5 days, and (b) differences between correlation and RMSE of the ECMWF raw forecasts as well as the Control calibrated forecasts and the GEFS raw forecasts.

The analogs were produced by pondering the 90% of the ssd of total precipitation plus the 10% of the ssd of precipitable water.

5) Prec\_water05 is similar to Prec\_water01, but considering the 50% of ssd of both total precipitation and precipitable water.

Besides these five methods, we had pondered the rank analog technique (Hamill and Whitaker 2006), which used the differences between the rank of the precipitation forecasts within the search region as the similarity measure. However, this method was found unsuitable for the conditions of Brazil and was therefore excluded. By matching ranks instead of the actual values, many members of the analog ensemble corresponded to dates whose precipitations over the search region follows the same order (rank) compared to the current day, but whose total amounts are dramatically different. For example, the method often matched a heavy rainfall at the current day with drizzle in the past.

#### f. Verification analysis

In this study we compare point and regionally aggregated values (see Medina et al. 2018) of several deterministic and probabilistic metrics. For the deterministic

metrics, we used the mean error (ME), the root-mean-square error (RMSE), and correlation coefficient  $\rho$ , which are among the most commonly reported measures of agreement between forecasts and observations. For the probabilistic metrics we used the Brier skill score (BSS) and the reliability diagram (e.g., Toth et al. 2003; Wilks 2011) associated with the precipitation events above 2.5 mm. In the study the forecast probability is calculated from the ensemble forecast, while the climatological probability is computed as an average probability taken over  $\pm 30$  days of the forecast date. A bootstrapping procedure involving 1000 samples was used to quantify the uncertainty of the probabilistic statistics (see Medina et al. 2018).

### 3. Results and discussion

#### a. Intercomparisons between GEFS, ECMWF, and Control analog postprocessed forecasts

Figure 3 shows the average correlation and RMSE of the raw GEFS forecasts in each region, and their differences with the ECMWF forecasts and the Control analog forecasts. The average correlation varied especially among regions: from high values over Atlantic Forest and Pampa

to very weak values over Amazon. The RMSE was proportional to the total rainfall and therefore more seasonally driven, with maximums during warm seasons, and minimums during cold seasons. [Janowiak et al. \(2010\)](#) noted very weak correlation between the ECMWF and GEFS forecasts and the Global Precipitation Climatology analyses over the northwest of South America in warm seasons. The ECMWF and GEFS raw forecasts performed comparably at 1.5 days, but the former performed better, even compared with the Control analog forecasts, at 3.5 and 5.5 days. The performance of the GEFS forecasts mostly improved through postprocessing; the analog control forecast provided the best correlation and RMSE at 1.5 days. Moreover, as indicated in [Fig. 4](#), the Control analog GEFS forecasts showed greater ME than both GEFS and ECMWF raw forecasts. They tended to underestimate the precipitations in most regions, seasons, and lead times.

[Figure 5](#) presents the distribution of the bootstrapped BSS values over each region and month at lead times of 1.5 and 5.5 days. The Control analog forecasts in most cases improved the BSS compared to the raw ensemble forecasts. The improvements tend to be higher over regions and seasons, such as the spring month in Amazon and the fall month in Pantanal, where raw forecasts are less skillful. Practically all the Control analog forecasts provided a positive BSS, although it was still close to zero in Amazon. The raw ECMWF forecasts showed higher probabilistic forecast skill than the raw GEFS forecasts at 3.5 and 5.5 lead days, but lower at 1.5 days. Both of the ECMWF and GEFS raw forecasts showed no skill over Amazon and Pantanal at any lead time, indicating that the climatological predictions are better here compared to the raw forecasts. This result is consistent with the study based on regional ensemble forecasts over South America ([Ruiz et al. 2009](#)). The reason for that is due to the convection in the Amazon exhibiting more pronounced diurnal and seasonal variability than in the east region ([Jones and Schemm 2000](#)).

To provide a better insight in space, [Fig. 6](#) shows the differences between the Brier scores of the raw ECMWF and GEFS ensemble forecasts at each grid point. Positive differences indicate the GEFS forecasts are better, since the lower the Brier score the better the forecasts. The ECMWF forecast seems relatively weak over the northwest, mainly at 1.5 days, probably due to issues with the model representation of the daily precipitation cycle over Amazon ([Betts and Jakob 2002a,b](#)). Similarly, the GEFS forecast are unskillful over this region as well. The differences in October, a period associated with the onset of the precipitations in most Brazil ([Marengo et al. 2001; Grimm and Zilli 2009](#)), are the most favorable for the GEFS forecasts, while those in January (a period of higher convection) are the most unfavorable, at both lead times.

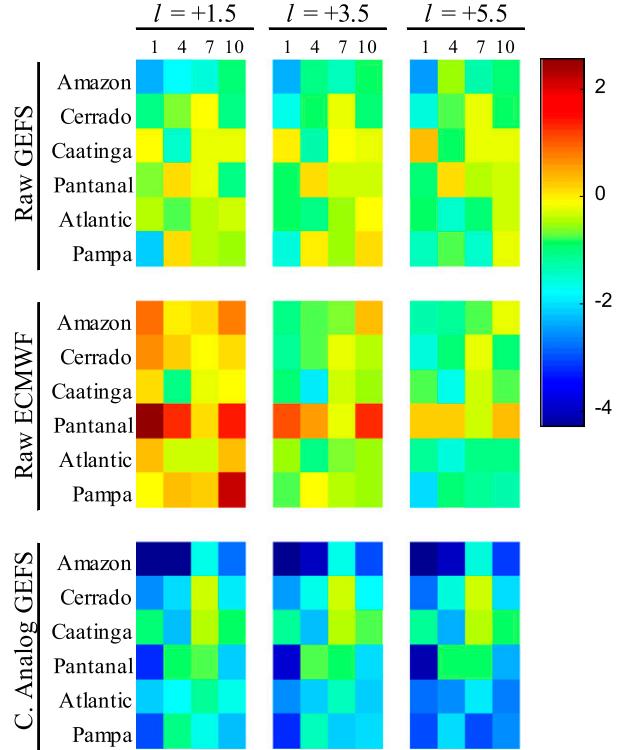


FIG. 4. ME (mm) of the raw GEFS and ECMWF and the Control analog precipitation forecasts at each biome in January (1), April (4), July (7), and October (10), for lead times of 1.5, 3.5, and 5.5 days.

The ECMWF forecasts at 3.5 and 5.5 days in several cases provided better RMSE and BSS than the forecasts at 1.5 days, while the bias of the ECMWF at 5.5 days tended to be negative while the bias at 1.5 days is positive ([Fig. 4](#)). These trends were not observed for the raw and postprocessed GEFS forecasts. This finding is consistent with [Janowiak et al. \(2010\)](#), who found that the 9-day ECMWF raw forecasts had lower bias than the day-2 forecasts over much of central South America. To investigate what caused the better performance at longer lead times, we analyzed the spread–skill relationships of different forecasts at 1.5 and 5.5 leads, by comparing the average standard deviation of the ensembles to the RMSE of the ensemble means for different intervals of the deviations ([Fig. 7](#)). The result showed that, while the spread of the GEFS ensembles (both, raw and postprocessed) was similar at different lead times, the ECMWF ensembles at 1.5 lead days were more underdispersed than 5.5 lead days. The wider spread of the ECMWF ensemble forecasts at longer lead times compared to shorter lead times may cause the better performance for longer lead times. The results also suggest that the forecast postprocessing with the analog technique considerably improves the spread–skill relationship of the ensembles.

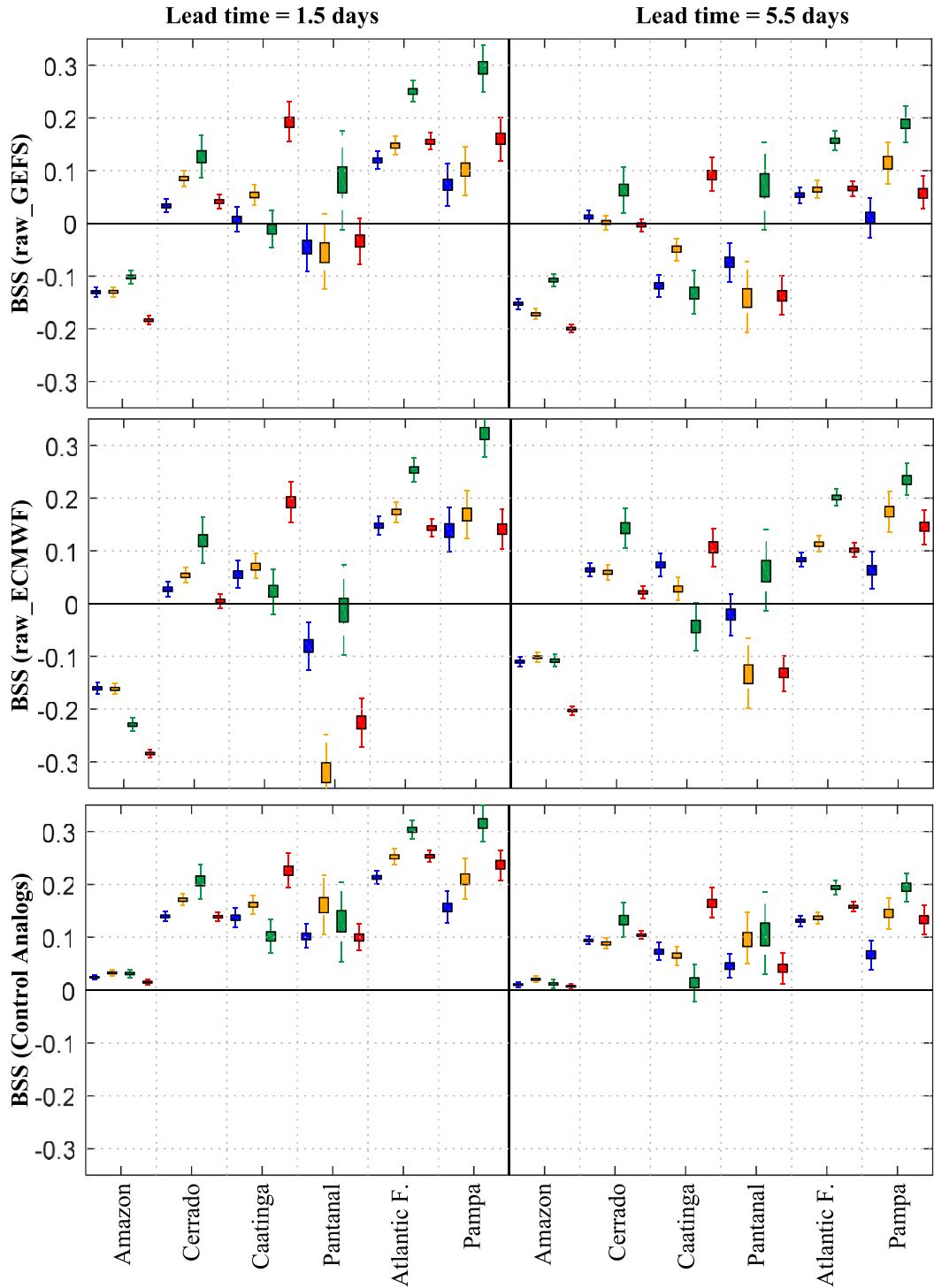


FIG. 5. BSS of the raw GEFS and ECMWF and the Control analog precipitation in (from left to right) January (blue), April (yellow), July (green), and October (red) for lead times of 1.5 and 5.5 days.

Figure 8 shows the reliability diagrams over January at lead days of 1.5 and 5.5. In general, the forecasts were slightly less reliable in drier months when high-probability precipitation forecasts are issued less frequently. The

postprocessed forecasts were considerably more reliable but less sharp than raw forecasts. The frequency of medium-probability forecasts grows after postprocessing mainly at the expense of the high-probability forecasts, as

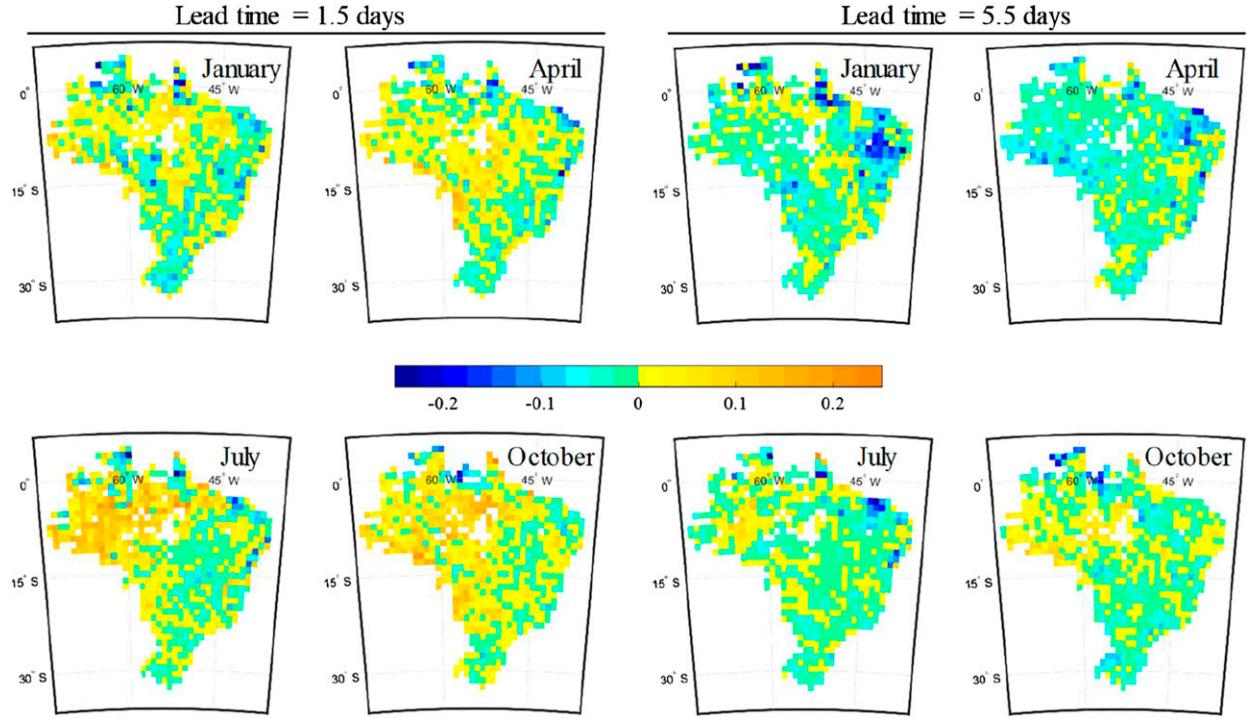
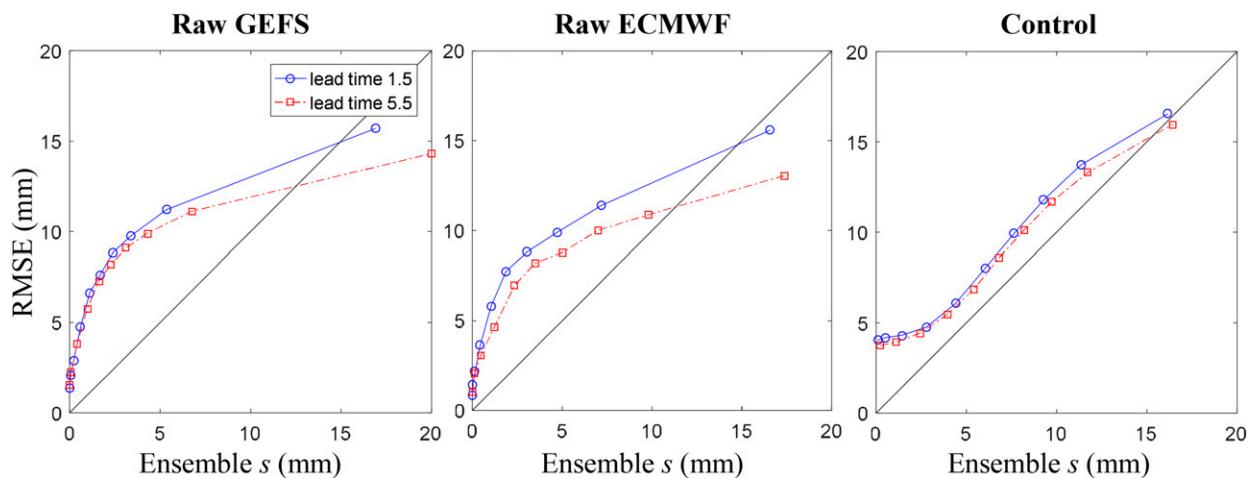


FIG. 6. Differences between the ECMWF and the GEFS BSS for 1.5 and 5.5 lead days.

found by Hamill et al. (2008). The GEFS and ECMWF raw forecasts provided similar reliability at 1.5 lead days, while the former one seemed slightly less reliable at 5.5 days. While the reliabilities are not considerably changed with lead times, the sharpness at shorter lead times is slightly higher than longer lead times, especially for the ECMWF forecasts. This may be caused by the narrower ensemble spread at shorter lead times. It is also worth noting that through postprocessing, the

reliability of the precipitation forecasts improved more than the skill score, which is in agreement with previous studies based on, either analog postprocessing techniques (e.g., Voisin et al. 2010), or other methods (Hamill et al. 2008).

In summary, the Control analog forecasts considerably improved the probabilistic forecasting performance but more systematically biased compared to the GEFS and ECMWF raw forecasts. They were also slightly less

FIG. 7. RMSE of the ensemble forecasts vs the mean standard deviation  $s$  of the ensemble members over all grid points and at 1.5 and 5.5 lead days for (from left to right) raw GEFS, raw ECMWF, and Control analog forecasts.

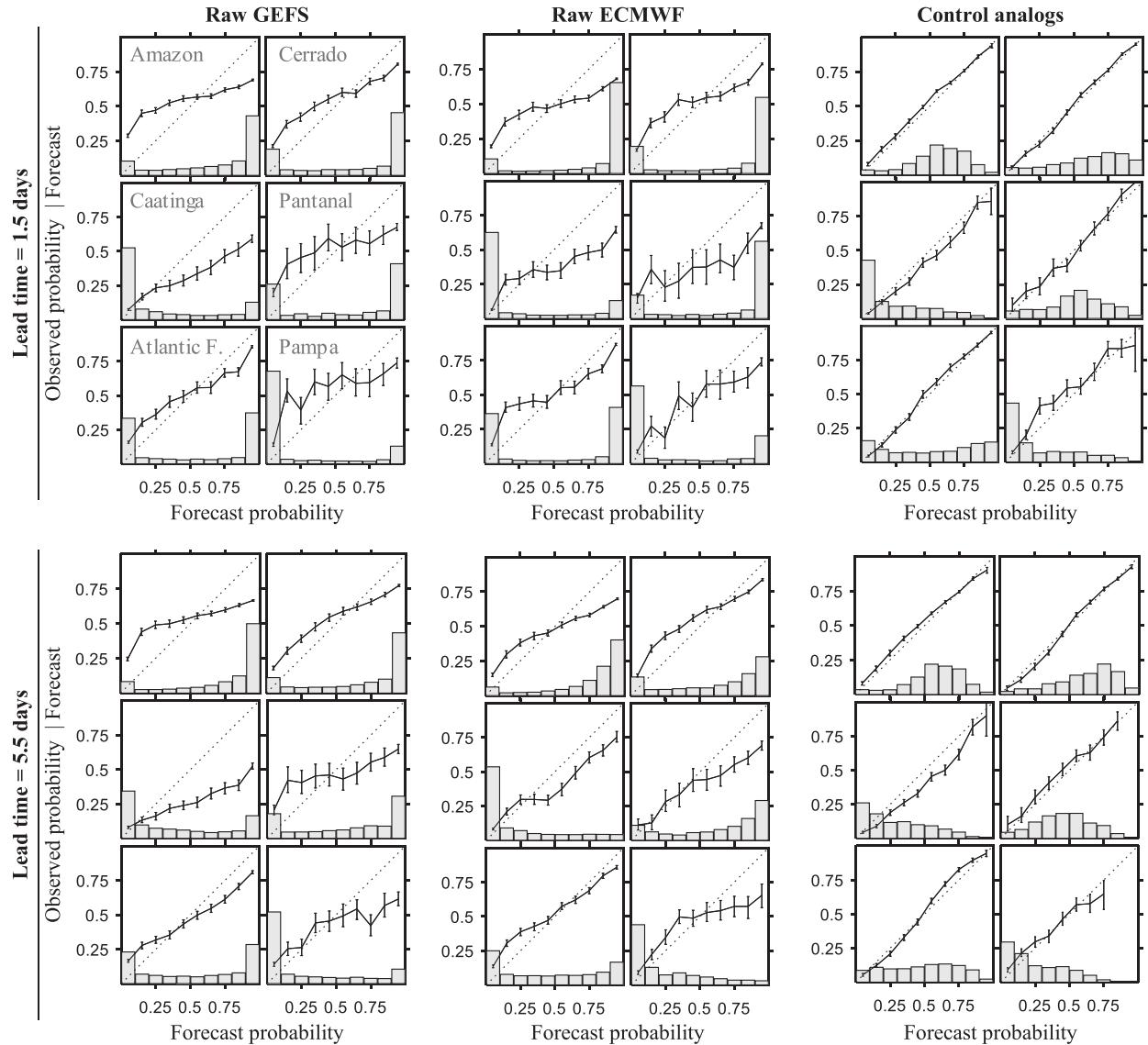


FIG. 8. Reliability diagrams of the GEFS and ECMWF raw forecasts, and the Control analog forecasts for January for days +1.5 and +5.5.

correlated with observations and less accurate than the ECMWF forecasts at 3.5 and 5.5 days. The performance of the raw ECMWF and GEFS forecasts was comparable at a 1.5-day lead, but the ECMWF forecasts performed better at the longer lead times.

*b. Comparing multiple analog approaches and the logistic regression approach*

Figure 9 shows cumulative probability distributions of the correlations, the ME, and the RMSE using raw GEFS forecasts and the six analog approaches over January and at 1.5-day lead. The trends were similar across different regions and lead times. The analog forecasts improve the correlation and the RMSE compared to the raw forecasts, while they are more biased

than the raw forecasts. This last is a deficiency principally caused by the skewness of the density distribution of the forecast precipitation (e.g., Scheuerer and Hamill 2015), meaning that the probability density of lighter precipitation events in a neighborhood around the mean is higher than the probability of heavier events. The used loss function, that is, the squared differences between current and past forecasts, might have also contributed to that issue, since it tends to provide large penalties to the large discrepancies (Casella and Berger 2002). Hence, reforecasts associated to lighter events are more likely to be (indirectly) represented in the analog ensembles (e.g., Hamill and Whitaker 2006).

A paired sample *t* test was conducted to compare the performance of the six analog forecasts. The result

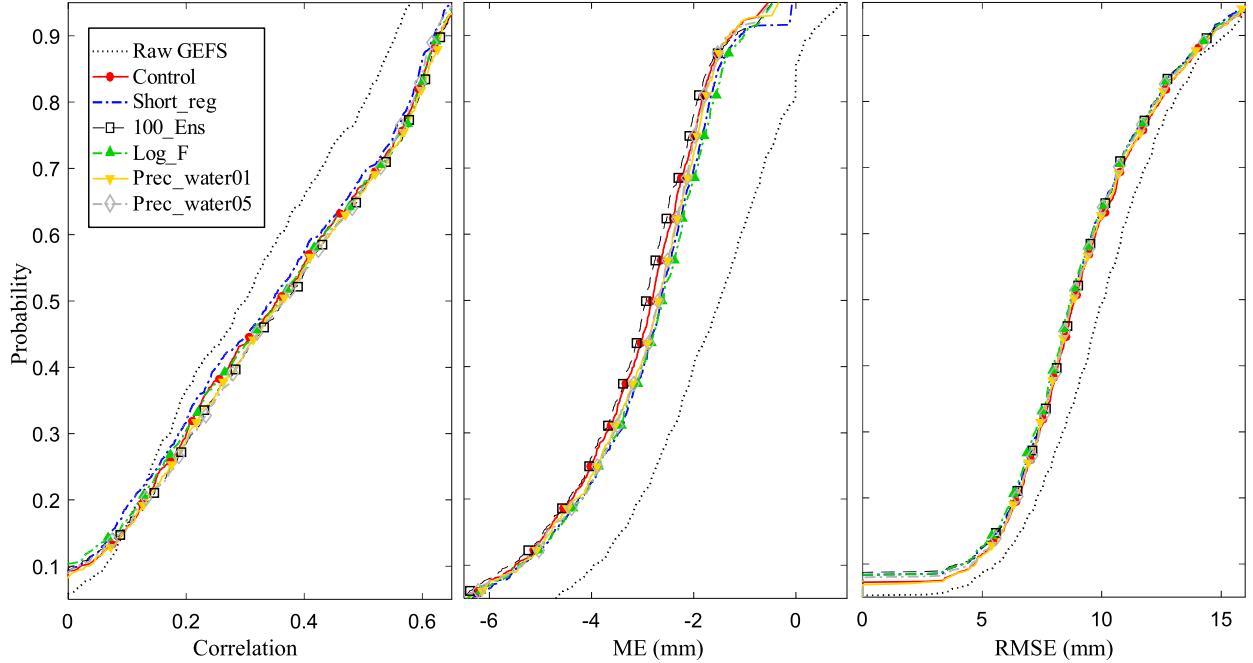


FIG. 9. Cumulative distribution of the correlations, ME, and RMSE for the raw and analog GEFS forecasts at 1.5 lead days in January.

shows that the differences of the six analog methods are small in correlation and ME but mostly significant at the 1% significance level; the differences in RMSE values are less significant, especially for comparisons among the Control, Sort\_reg, and Prec\_water01 approaches. The changes in ME and RMSE after postprocessing seemed roughly constant among grid points, while the correlation improved more over grid points with higher correlations, that is, regions with better correlation were more benefited through postprocessing. Among all the six methods, the 100\_Ens and LogF forecasts commonly provided the best correlations and ME and RMSE, respectively. In most cases, the Prec\_water05 forecasts perform the worst among all the analog forecasts.

For probabilistic forecasts, all the analog forecasts, as well as the LR forecasts considerably improved the BSS compared to the raw forecasts (Fig. 10). The improvements were similar but mostly significantly different at 1% significance level. In agreement with previous studies (e.g., Hamill and Whitaker 2006; Delle Monache et al. 2013), the forecasts produced with analog methods provided better skill compared to the logistic regression. Only the Prec\_water05 forecasts performed similarly or slightly worse than the LR forecasts. While the 100\_Ens forecasts commonly provided better BSS over West regions, where the skill is consistently low, the Log\_F forecasts provided better skill over the East regions. The average BSSs were mostly below 0.3, and affected by the considerable spatial and temporal variability of the BSS

(Fig. 11). As suggested by the maps of Brier score of the climatology in Fig. 12, this variability seemed associated with the climate predictability. Climate predictability changes from high in the winter all over the center of Brazil to very low in summer practically all over the country. It is influenced by the interannual migration of deep tropical convection from the central and southern portion of the Amazon basin in summer to the northwestern sector of South America in winter (Rao and Hada 1990).

Figure 13 shows the reliability diagrams of the raw and postprocessed forecasts over January at 1.5 days lead. The postprocessed forecasts performed very similarly: they were considerably more reliable but less sharp than the raw forecasts. Only the 100\_Ens forecasts showed a slightly dry bias over a few regions. Our postprocessed forecasts are slightly more reliable than the analog forecasts from Hamill and Whitaker (2006) and the logistic regression forecasts from Hamill et al. (2008), probably due to the new improvements of the GEFS model compared to its first version (Hamill et al. 2015).

#### c. Discussing analog postprocessing methods for precipitation forecasts

While the comparison between raw and postprocessed forecasts in section 3a only involved one analog approach, the results from this comparison also seemed valid for the other postprocessing methods. This is because the differences among postprocessing methods were lower than the

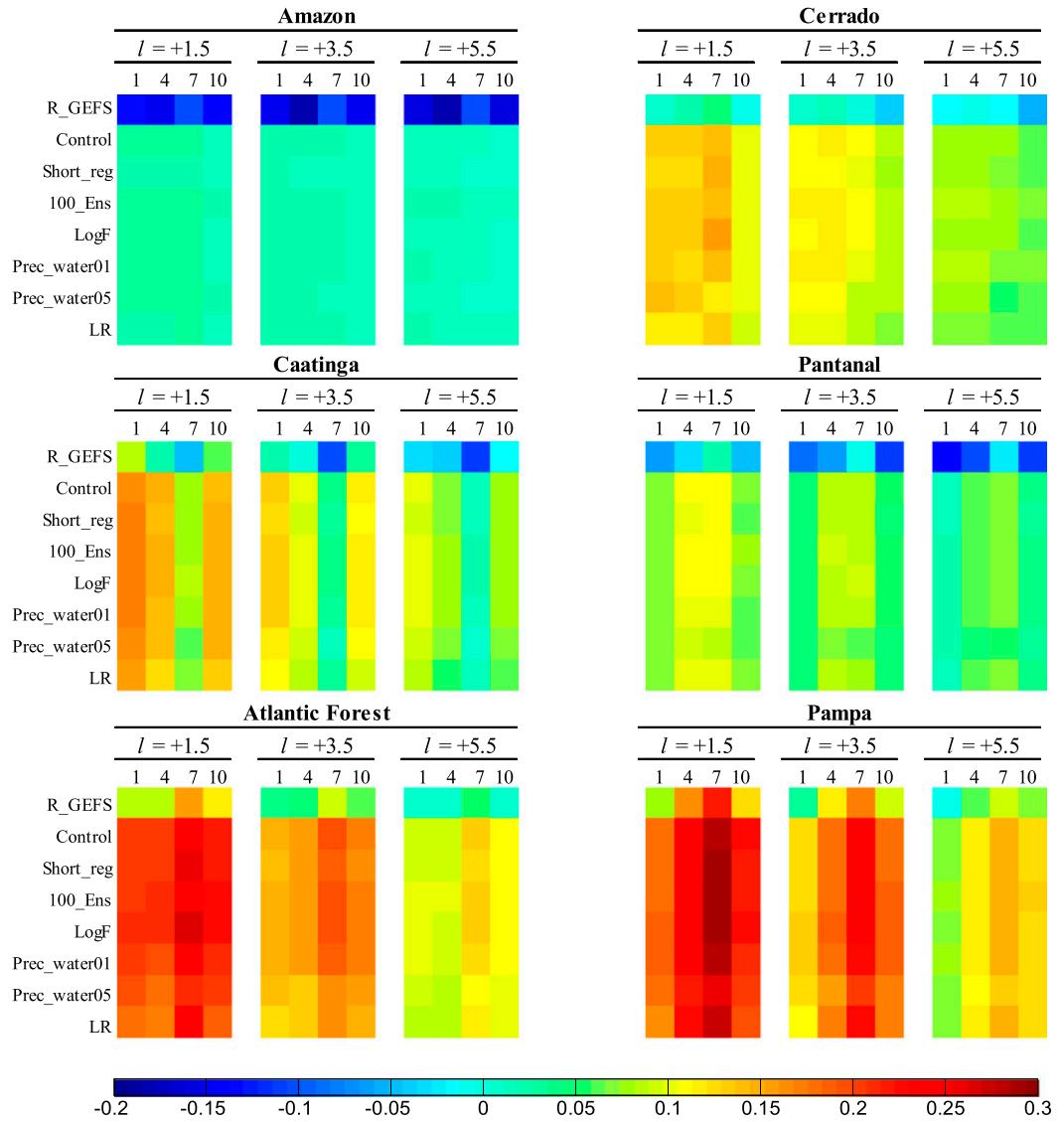


FIG. 10. Mean BSS of the raw GEFS forecasts, the analog calibration methods, and the logistic regression method in the six regions.

differences among raw and postprocessed forecasts. Nonetheless, different methods in most cases provided significantly different statistics, allowing us to identify a best and a worst strategy. In Table 2 we show the number of experiments (of a total of 72, i.e., 6 regions  $\times$  4 months  $\times$  3 lead times) where each analog approach performed best and worst in terms of correlation, ME, RMSE, and BSS. We also provide a ranking of the methods by sorting the differences between totals, from better to worse. The LR-based performance was considered when comparing the BSS, but not ranked. LogF was identified as the best approach, ranking as the best or the second best in terms of all the performance statistics. Taking the logarithm of the mean precipitation

is a way of partially addressing the issues associated to the skewness of the forecast density distribution, by making the approach less prone to penalize the heavier events with respect to the lighter events. This confirms the results of Foresti et al. (2015) for who the logarithmic transformation is necessary to get a suitable representation of the forecast uncertainty.

The 100\_Ens approach was the second best since it provides the highest BSS and correlations but also has relatively large ME and RMSE. The larger the sample of ensemble members, the better the probabilistic description of the events, but also the larger the issues for finding sound analogs on rainy days, especially for regions with lower pluviometry. Therefore the choice between the

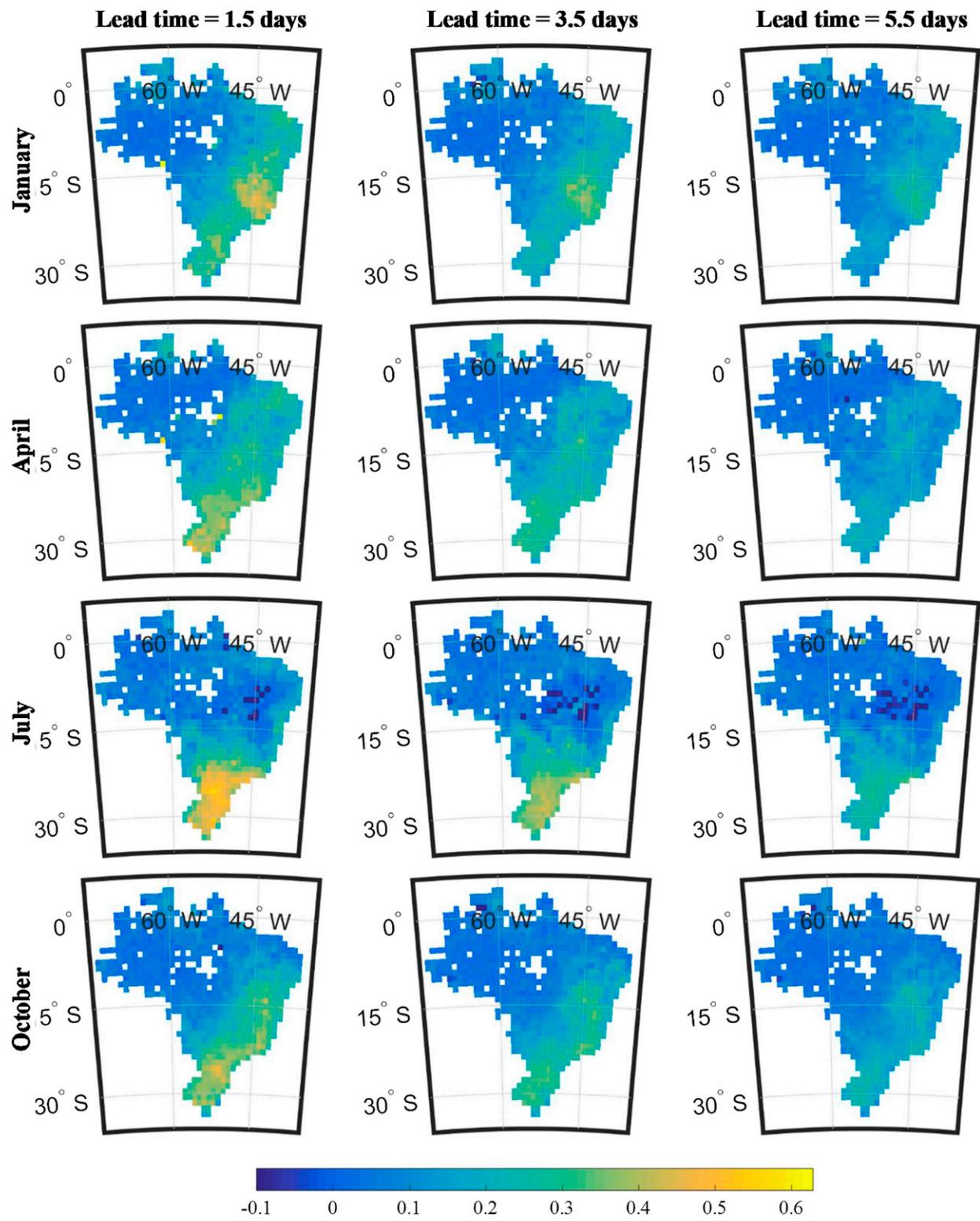


FIG. 11. BSS values of the basic analog technique in space from 1985 to 2010.

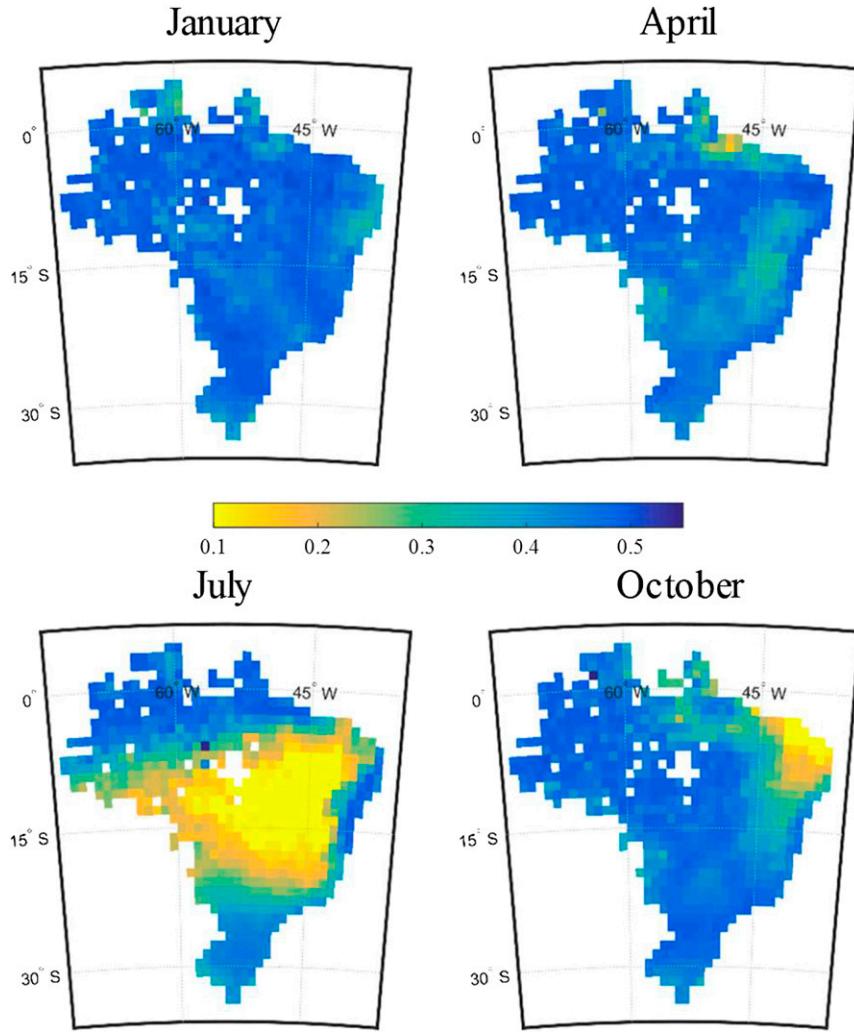


FIG. 12. Brier score of the climatology in space; 2.5 mm is used as a threshold.

LogF and the 100\_Ens scheme might be subject to the specific purposes of the application. These results led to the question whether an analog method combining the modifications adopted in Log\_F and 100\_Ens (100\_Ens-LogF) improve the performance compared to each of these two approaches. We found that it slightly improved the performance of the original approaches by taking advantage of their best features, since it is less vulnerable than the 100\_Ens method to large deterministic errors because of the log transformation, and at the same time performs probabilistically slightly better because of the larger sample size. It provided consistently better BSS than the Log\_F and 100\_Ens strategies, although the improvements only affected the third significant figure. It also improved the ME compared with the 100\_Ens strategy, but not compared with the Log\_F strategy.

Finally, the Prec\_water05 scheme performed worst for most statistics, especially during fall and winter times

(e.g., Fig. 10), distantly followed by the Short\_reg scheme. Studies suggest that there is a weak relationship between precipitable water and precipitation in most regions (e.g., Rao and Da Silva Marques 1984; Teixeira and Satyamurty 2007), presumably because of the complex regimes of water vapor and rainfall, which are generally regulated by the ocean on the east, a steep orography on the west, and the Amazon forest on the north (Berberry and Collini 2000; Rao et al. 1996).

#### 4. Concluding remarks

This study conducted intercomparisons between raw GEFS forecasts, raw ECMWF forecasts, and postprocessed GEFS forecasts with six analog methods and the logistic regression method over six biome regions in Brazil. To the authors' knowledge, this study is the first to comprehensively examine the performance of these global-scale

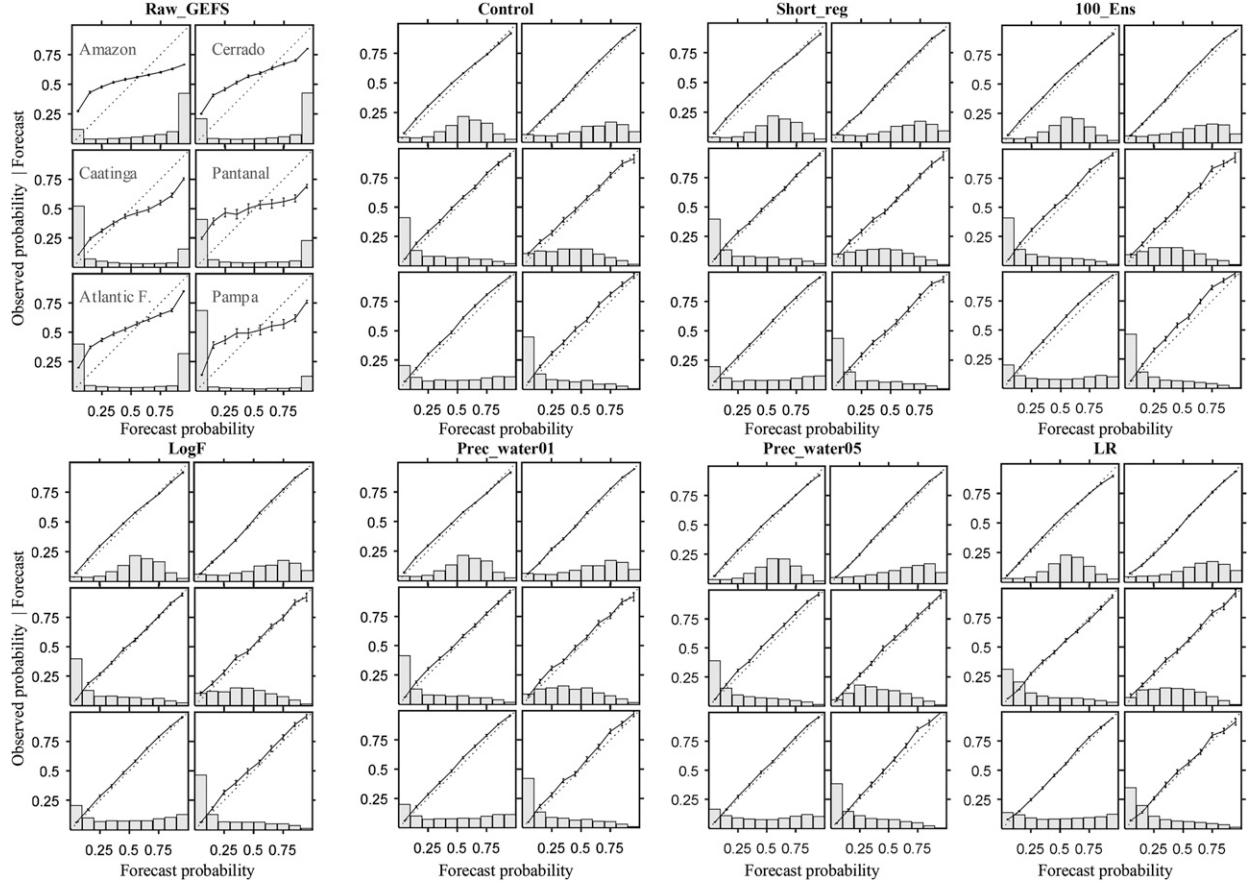


FIG. 13. Reliability diagrams for the GEFS-based precipitation forecasts in January. Each panel indicates each reliability diagram of the methods including raw GEFS, six analog methods, and logistic regression.

NWP models and statistical postprocessing methods over South America, specifically over a region severely affected by large mesoscale convective systems.

The article showed that the global-scale NWP's raw forecasts are helpful for precipitation forecasting over the east, and particularly the southeast, of Brazil, but unskillful over the northwest. The ECMWF raw forecasts are better than the GEFS raw forecasts since they perform similarly or better over the east. However, the

postprocessed GEFS forecasts, particularly the analog forecasts, are strongly recommended over the raw ECMWF forecasts as they performed probabilistically much better; unlike the raw forecasts they improved the skill of climatological forecasts in all the evaluated regions, seasons and lead times. Our results also confirmed previous findings showing that the analog forecasts tend to be negatively biased: this study suggests that the larger the size of the analog ensemble, the higher the bias.

TABLE 2. Number of experiments (considering 6 regions, 4 months, and 3 lead times) where the alternative analog approaches performed the best and worst in terms of different metrics.

| Method       | Correlation |       | ME   |       | RMSE |       | BSS  |       | Ranking |
|--------------|-------------|-------|------|-------|------|-------|------|-------|---------|
|              | Best        | Worst | Best | Worst | Best | Worst | Best | Worst |         |
| Control      | 2           | 3     | 0    | 3     | 2    | 1     | 1    | 1     | 4       |
| Short_reg    | 1           | 22    | 3    | 2     | 2    | 4     | 0    | 9     | 5       |
| 100_Ens      | 57          | 1     | 0    | 33    | 8    | 16    | 37   | 1     | 2       |
| LogF         | 7           | 1     | 60   | 0     | 48   | 2     | 24   | 0     | 1       |
| Prec_water01 | 3           | 1     | 4    | 1     | 11   | 4     | 5    | 0     | 3       |
| Prec_water05 | 2           | 44    | 5    | 33    | 1    | 45    | 3    | 28    | 6       |
| LR           |             |       |      |       |      |       | 2    | 33    |         |

The forecast performance showed less sensitivity to the postprocessing strategy than to the postprocessing itself. Nevertheless, different postprocessing strategies are significantly different statistically, with the analog forecasts being as reliable as the logistic regression forecast but slightly more skillful. The strategy considering the log of current and past reforecasts as the measure of closeness performed slightly better among all the analog forecasts, followed by that considering 100 analog members (instead of the regular 50). The analog method combining modifications adopted in these two approaches performed slightly better than the individual approaches, whereas the strategies that included precipitable water as a predictor variable were among the worst.

This study provides useful information for precipitation forecasting over tropical and subtropical regions affected by large mesoscale convective systems. While we have addressed the impact of the forecast uncertainty on the performance by using bootstrap analysis, we have not addressed the impact of the verification dataset uncertainty. The quality of interpolated datasets in data-sparse regions is always a source of concern. While we tried several analog-based schemes, new analog strategies are emerging in literature, which may further improve the precipitation forecasting in Brazil. For example, the hybrid NWP–analog method (Eckel and Delle Monache 2016), seizes both the strengths of the numerical weather prediction ensemble and the analog ensemble, and could be used for efficiently combining the ECMWF and GEFS raw forecasts. Recent approaches also explore methods based on different synoptic-scale predictors (Horton and Brönnimann 2019; Horton et al. 2018; Ben Daoud et al. 2016), such as the geopotential height, and temperature, while excluding precipitation as a predictor, as a way to avoid common issues with precipitation reanalysis (e.g., Caillouet et al. 2016).

More research is needed for further decreasing the forecasting uncertainty, especially over the Amazon. We foresee future studies will evaluate the efficacy of multimodel forecasts and other postprocessing methods with the consideration of the uncertainty from the verification dataset, in particular focusing on the methods that can perform well with much shorter training datasets.

**Acknowledgments.** This work was supported in part by the Auburn University Intramural Grants Program, by the USGS Alabama Water Resources Institute 104(b) Grant Program, and by the Alabama Agricultural Experimental Station (AAES) grant under the USDA National Institute of Food and Agriculture Hatch project 421 Accession 1012578.

## REFERENCES

Atger, F., 2001: Verification of intense precipitation forecasts from single models and ensemble prediction systems. *Non-linear Processes Geophys.*, **8**, 401–417, <https://doi.org/10.5194/npg-8-401-2001>.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525**, 47–55, <https://doi.org/10.1038/nature14956>.

Bechtold, P., and Coauthors, 2012: Progress in predicting tropical systems: The role of convection. ECMWF Tech. Memo. 686, 63 pp., <https://www.ecmwf.int/sites/default/files/elibrary/2012/8019-progress-predicting-tropical-systems-role-convection.pdf>.

—, N. Semane, P. Lopez, J. P. Chaboureau, A. Beljaars, and N. Bormann, 2014: Representing equilibrium and non-equilibrium convection in large-scale models. *J. Atmos. Sci.*, **71**, 734–753, <https://doi.org/10.1175/JAS-D-13-0163.1>.

Ben Daoud, A., E. Sauquet, G. Bontron, C. Obled, and M. Lang, 2016: Daily quantitative precipitation forecasts based on the analogue method: Improvements and application to a French large river basin. *Atmos. Res.*, **169**, 147–159, <https://doi.org/10.1016/j.atmosres.2015.09.015>.

Berbery, E. H., and E. A. Collini, 2000: Springtime precipitation and water vapor flux over southeastern South America. *Mon. Wea. Rev.*, **128**, 1328–1346, [https://doi.org/10.1175/1520-0493\(2000\)128<1328:SPA WVF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<1328:SPA WVF>2.0.CO;2).

Betts, A. K., and C. Jakob, 2002a: Evaluation of the diurnal cycle of precipitation, surface thermo-dynamics and surface fluxes in the ECMWF model using LBA data. *J. Geophys. Res.*, **107**, 8045, <https://doi.org/10.1029/2001JD000427>.

—, and —, 2002b: Study of diurnal cycle of convective precipitation over Amazonia using a single column model. *J. Geophys. Res.*, **107**, 4732, <https://doi.org/10.1029/2002JD002264>.

Bony, S., and Coauthors, 2015: Clouds, circulation and climate sensitivity. *Nat. Geosci.*, **8**, 261–268, <https://doi.org/10.1038/ngeo2398>.

Caillouet, L., J. P. Vidal, E. Sauquet, and B. Graff, 2016: Probabilistic precipitation and temperature downscaling of the Twentieth Century Reanalysis over France. *Climate Past*, **12**, 635–662, <https://doi.org/10.5194/cp-12-635-2016>.

Carvalho, L. M., C. Jones, A. N. Posadas, R. Quiroz, B. Bookhagen, and B. Liebmann, 2012: Precipitation characteristics of the South American monsoon system derived from multiple datasets. *J. Climate*, **25**, 4600–4620, <https://doi.org/10.1175/JCLI-D-11-00335.1>.

Casella, G., and R. L. Berger, 2002: *Statistical Inference*. 2nd ed. Duxbury, 660 pp.

Cloke, H. L., and F. Pappenberger, 2009: Ensemble flood forecasting: A review. *J. Hydrol.*, **375**, 613–626, <https://doi.org/10.1016/j.jhydrol.2009.06.005>.

Cressman, G. P., 1959: An operational objective analysis system. *Mon. Wea. Rev.*, **87**, 367–374, [https://doi.org/10.1175/1520-0493\(1959\)087<0367:AOOAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2).

Delle Monache, L., T. Nipen, Y. Liu, G. Roux, and R. Stull, 2011: Kalman filter and analog schemes to post-process numerical weather predictions. *Mon. Wea. Rev.*, **139**, 3554–3570, <https://doi.org/10.1175/2011MWR3653.1>.

—, F. A. Eckel, D. L. Rife, B. Nagarajan, and K. Searight, 2013: Probabilistic weather prediction with an analog ensemble. *Mon. Wea. Rev.*, **141**, 3498–3516, <https://doi.org/10.1175/MWR-D-12-00281.1>.

Eckel, F. A., and L. Delle Monache, 2016: A hybrid NWP–analog ensemble. *Mon. Wea. Rev.*, **144**, 897–911, <https://doi.org/10.1175/MWR-D-15-0096.1>.

Espinoza Villar, J. C., and Coauthors, 2009: Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador). *Int. J. Climatol.*, **29**, 1574–1594, <https://doi.org/10.1002/joc.1791>.

FAO, 2015: FAOSTAT database. Food and Agriculture Organization of the United Nations, Statistics Division, <http://faostat3.fao.org/home/E>.

Ferreira, J., R. Pardini, J. P. Metzger, C. R. Fonseca, P. S. Pompeu, G. Sparovek, and J. Louzada, 2012: Towards environmentally sustainable agriculture in Brazil: Challenges and opportunities for applied ecological research. *J. Appl. Ecol.*, **49**, 535–541, <https://doi.org/10.1111/j.1365-2664.2012.02145.x>.

Foresti, L., L. Panziera, P. V. Mandapaka, U. Germann, and A. Seed, 2015: Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall. *Meteor. Appl.*, **22**, 141–155, <https://doi.org/10.1002/met.1416>.

Fu, R., and Coauthors, 2013: Increased dry-season length over southern Amazonia in recent decades and its implication for future climate projection. *Proc. Natl. Acad. Sci. USA*, **110**, 18 110–18 115, <https://doi.org/10.1073/pnas.1302584110>.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11**, 1203–1211, [https://doi.org/10.1175/1520-0450\(1972\)011<1203:TUOMOS>2.0.CO;2](https://doi.org/10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2).

—, T. L. Chambers, W. S. Richardson, and H. P. Perrotti, 1985: Objective map analysis for the local AFOS MOS Program. NOAA Tech. Memo. NWS TDL 75, 42 pp., <https://www.nws.noaa.gov/mdl/pubs/Documents/TechMemo/TechMemo75.pdf>.

Gneiting, T., 2014: Calibration of medium-range weather forecasts. ECMWF Tech. Memo. 719, 30 pp., <https://www.ecmwf.int/sites/default/files/elibrary/2014/9607-calibration-medium-range-weather-forecasts.pdf>.

Grimm, A. M., and M. T. Zilli, 2009: Interannual variability and seasonal evolution of summer monsoon rainfall in South America. *J. Climate*, **22**, 2257–2275, <https://doi.org/10.1175/2008JCLI2345.1>.

Guerrant, R. L., and Coauthors, 1983: Prospective study of diarrheal illnesses in northeastern Brazil: Patterns of disease, nutritional impact, etiologies, and risk factors. *J. Infect. Dis.*, **148**, 986–997, <https://doi.org/10.1093/infdis/148.6.986>.

Hagedorn, R., T. M. Hamill, and J. S. Whitaker, 2008: Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part I: Two-meter temperatures. *Mon. Wea. Rev.*, **136**, 2608–2619, <https://doi.org/10.1175/2007MWR2410.1>.

—, R. Buizza, T. M. Hamill, M. Leutbecher, and T. N. Palmer, 2012: Comparing TIGGE multimodel forecasts with reforecast-calibrated ECMWF ensemble forecasts. *Quart. J. Roy. Meteor. Soc.*, **138**, 1814–1827, <https://doi.org/10.1002/qj.1895>.

Hamill, T. M., 2012: Verification of TIGGE multimodel and ECMWF reforecast-calibrated probabilistic precipitation forecasts over the contiguous United States. *Mon. Wea. Rev.*, **140**, 2232–2252, <https://doi.org/10.1175/MWR-D-11-00220.1>.

—, and J. S. Whitaker, 2006: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. *Mon. Wea. Rev.*, **134**, 3209–3229, <https://doi.org/10.1175/MWR3237.1>.

—, —, and S. L. Mullen, 2006: Reforecasts: An important dataset for improving weather predictions. *Bull. Amer. Meteor. Soc.*, **87**, 33–46, <https://doi.org/10.1175/BAMS-87-1-33>.

—, R. Hagedorn, and J. S. Whitaker, 2008: Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part II: Precipitation. *Mon. Wea. Rev.*, **136**, 2620–2632, <https://doi.org/10.1175/2007MWR2411.1>.

—, G. T. Bates, J. S. Whitaker, D. R. Murray, M. Fiorino, T. J. Galarneau Jr., Y. Zhu, and W. Lapenta, 2013: NOAA's second-generation global medium-range ensemble reforecast dataset. *Bull. Amer. Meteor. Soc.*, **94**, 1553–1565, <https://doi.org/10.1175/BAMS-D-12-00014.1>.

—, M. Scheuerer, and G. T. Bates, 2015: Analog probabilistic precipitation forecasts using GEFS reforecasts and climatology-calibrated precipitation analyses. *Mon. Wea. Rev.*, **143**, 3300–3309, <https://doi.org/10.1175/MWR-D-15-0004.1>.

Hamilton, S. K., 2002: Hydrological controls of ecological structure and function in the Pantanal wetland (Brazil). *The Ecohydrology of South American Rivers and Wetlands*, M. E. McClain, Ed., IAHS Special Publ. 6, International Association of Hydrological Sciences, 133–158.

He, Y., and Coauthors, 2010: Ensemble forecasting using TIGGE for the July–September 2008 floods in the Upper Huai catchment: A case study. *Atmos. Sci. Lett.*, **11**, 132–138, <https://doi.org/10.1002/asl.270>.

Horton, P., and S. Brönnimann, 2019: Impact of global atmospheric reanalyses on statistical precipitation downscaling. *Climate Dyn.*, <https://doi.org/10.1007/s00382-018-4442-6>, in press.

—, M. Jaboyedoff, and C. Obled, 2018: Using genetic algorithms to optimize the analogue method for precipitation prediction in the Swiss Alps. *J. Hydrol.*, **556**, 1220–1231, <https://doi.org/10.1016/j.jhydrol.2017.04.017>.

IBGE, 2004: Mapa de Biomas do Brasil: Primeira aproximação. Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.

Jones, C., and J. K. E. Schemm, 2000: The influence of intra-seasonal variations on medium-to-extended-range weather forecasts over South America. *Mon. Wea. Rev.*, **128**, 486–494, [https://doi.org/10.1175/1520-0493\(2000\)128<0486:TIOIVO>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<0486:TIOIVO>2.0.CO;2).

Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multi-satellite observations. *J. Hydrometeorol.*, **2**, 36–50, [https://doi.org/10.1175/1525-7541\(2001\)002<0036:GPAODD>2.0.CO;2](https://doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2).

Janowiak, J. E., P. Bauer, W. Wang, P. A. Arkin, and J. Gottschalck, J., 2010: An evaluation of precipitation forecasts from operational models and reanalyses including precipitation variations associated with MJO activity. *Mon. Wea. Rev.*, **138**, 4542–4560, <https://doi.org/10.1175/2010MWR3436.1>.

Leal, I. R., J. M. C. da Silva, M. Tabarelli, and T. E. Lacher, 2005: Changing the course of biodiversity conservation in the Caatinga of northeastern Brazil. *Conserv. Biol.*, **19**, 701–706, <https://doi.org/10.1111/j.1523-1739.2005.00703.x>.

Liebmann, B., and D. Allured, 2005: Daily precipitation grids for South America. *Bull. Amer. Meteor. Soc.*, **86**, 1567–1570, <https://doi.org/10.1175/BAMS-86-11-1567>.

—, and —, 2006: Reply. *Bull. Amer. Meteor. Soc.*, **87**, 1096–1096, <https://doi.org/10.1175/BAMS-87-8-1096>.

Marengo, J. A., B. Liebmann, V. E. Kousky, N. P. Filizola, and I. C. Wainer, 2001: Onset and end of the rainy season in the Brazilian Amazon Basin. *J. Climate*, **14**, 833–852, [https://doi.org/10.1175/1520-0442\(2001\)014<0833:OAEOTR>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0833:OAEOTR>2.0.CO;2).

Medina, H., D. Tian, P. Srivastava, A. Pelosi, and G. B. Chirico, 2018: Medium-range reference evapotranspiration forecasts for the contiguous United States based on multi-model numerical weather predictions. *J. Hydrol.*, **562**, 502–517, <https://doi.org/10.1016/j.jhydrol.2018.05.029>.

MMA, 2011: Quarto relatório nacional para a convenção sobre diversidade biológica. Ministério do Meio Ambiente, 248 pp.

Mohr, K. L., and E. J. Zipser, 1996: Mesoscale convective systems defined by their 85-GHz ice scattering signature: Size and intensity comparison over tropical oceans and continents. *Mon. Wea. Rev.*, **124**, 2417–2437, [https://doi.org/10.1175/1520-0493\(1996\)124<2417:MCSDBT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<2417:MCSDBT>2.0.CO;2).

Moura, A. D., and J. Shukla, 1981: On the dynamics of droughts in northeast Brazil: Observations, theory and numerical experiments with a general circulation model. *J. Atmos. Sci.*, **38**, 2653–2675, [https://doi.org/10.1175/1520-0469\(1981\)038<2653:OTDODI>2.0.CO;2](https://doi.org/10.1175/1520-0469(1981)038<2653:OTDODI>2.0.CO;2).

Murphy, A. H., 1973: Hedging and skill scores for probability forecasts. *J. Appl. Meteor.*, **12**, 215–223, [https://doi.org/10.1175/1520-0450\(1973\)012<0215:HASSFP>2.0.CO;2](https://doi.org/10.1175/1520-0450(1973)012<0215:HASSFP>2.0.CO;2).

Pelosi, A., H. Medina, J. Van den Bergh, S. Vannitsem, and G. B. Chirico, 2017: Adaptive Kalman filtering for post-processing ensemble numerical weather predictions. *Mon. Wea. Rev.*, **145**, 4837–4854, <https://doi.org/10.1175/MWR-D-17-0084.1>.

Rao, V. B., and V. Da Silva Marques, 1984: Water vapor characteristics over Northeast Brazil during two contrasting years. *J. Climate Appl. Meteor.*, **23**, 440–444, [https://doi.org/10.1175/1520-0450\(1984\)023<0440:WVCONB>2.0.CO;2](https://doi.org/10.1175/1520-0450(1984)023<0440:WVCONB>2.0.CO;2).

—, and K. Hada, 1990: Characteristics of rainfall over Brazil: Annual variations and connections with the Southern Oscillation. *Theor. Appl. Climatol.*, **42**, 81–91, <https://doi.org/10.1007/BF00868215>.

—, I. F. Cavalcanti, and K. Hada, 1996: Annual variation of rainfall over Brazil and water vapor characteristics over South America. *J. Geophys. Res.*, **101**, 26 539–26 551, <https://doi.org/10.1029/96JD01936>.

Ratter, J. A., J. F. Ribeiro, and S. Bridgewater, 1997: The Brazilian cerrado vegetation and threats to its biodiversity. *Ann. Bot.*, **80**, 223–230, <https://doi.org/10.1006/anbo.1997.0469>.

Roesch, L. F. W., F. C. B. Vieira, V. A. Pereira, A. L. Schünemann, I. F. Teixeira, A. J. T. Senna, and V. M. Stefenon, 2009: The Brazilian Pampa: A fragile biome. *Diversity*, **1**, 182–198, <https://doi.org/10.3390/d1020182>.

Ruiz, J., C. Saulo, and E. Kalnay, 2009: Comparison of methods used to generate probabilistic quantitative precipitation forecasts over South America. *Wea. Forecasting*, **24**, 319–336, <https://doi.org/10.1175/2008WAF2007098.1>.

Scheuerer, M., and T. M. Hamill, 2015: Statistical postprocessing of ensemble precipitation forecasts by fitting censored, shifted gamma distributions. *Mon. Wea. Rev.*, **143**, 4578–4596, <https://doi.org/10.1175/MWR-D-15-0061.1>.

Silva, V. B., V. E. Kousky, W. Shi, and R. W. Higgins, 2007: An improved gridded historical daily precipitation analysis for Brazil. *J. Hydrometeor.*, **8**, 847–861, <https://doi.org/10.1175/JHM598.1>.

Su, X., H. Yuan, Y. Zhu, Y. Luo, and Y. Wang, 2014: Evaluation of TIGGE ensemble predictions of Northern Hemisphere summer precipitation during 2008–2012. *J. Geophys. Res. Atmos.*, **119**, 7292–7310, <https://doi.org/10.1002/2014JD021733>.

Subramanian, A., A. Weisheimer, T. Palmer, F. Vitart, and P. Bechtold, 2017: Impact of stochastic physics on tropical precipitation in the coupled ECMWF model. *Quart. J. Roy. Meteor. Soc.*, **143**, 852–865, <https://doi.org/10.1002/qj.2970>.

Teixeira, M. S. and P. Satyamurty, 2007: Dynamical and synoptic characteristics of heavy rainfall episodes in southern Brazil. *Mon. Wea. Rev.*, **135**, 598–617, <https://doi.org/10.1175/MWR3302.1>.

Tian, D., and C. J. Martinez, 2012: Comparison of two analog-based downscaling methods for regional reference evapotranspiration forecasts. *J. Hydrol.*, **475**, 350–364, <https://doi.org/10.1016/j.jhydrol.2012.10.009>.

—, and —, 2014: The GEFS-based daily reference evapotranspiration (ET<sub>0</sub>) forecast and its implication for water management in the southeastern United States. *J. Hydrometeor.*, **15**, 1152–1165, <https://doi.org/10.1175/JHM-D-13-0119.1>.

Toth, Z., O. Talagrand, G. Candille, and Y. Zhu, 2003: Probability and ensemble forecasts. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, I. T. Jolliffe and D. B. Stephenson, Eds., Wiley, 137–163.

Voisin, N., J. C. Schaake, and D. P. Lettenmaier, 2010: Calibration and downscaling methods for quantitative ensemble precipitation forecasts. *Wea. Forecasting*, **25**, 1603–1627, <https://doi.org/10.1175/2010WAF2222367.1>.

Wilks, D. S., 2006: Comparison of ensemble-MOS methods in the Lorenz'96 setting. *Meteor. Appl.*, **13**, 243–25, <https://doi.org/10.1017/S1350482706002192>.

—, 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. International Geophysics Series, Vol. 100, Academic Press, 704 pp.

—, and T. M. Hamill, 2007: Comparison of ensemble-MOS methods using GFS reforecasts. *Mon. Wea. Rev.*, **135**, 2379–2390, <https://doi.org/10.1175/MWR3402.1>.

Xavier, A. C., C. V. King, and B. R. Scanlon, 2016: Daily gridded meteorological variables in Brazil (1980–2013). *Int. J. Climatol.*, **36**, 2644–2659, <https://doi.org/10.1002/joc.4518>.

Zhou, Y., M. Hejazi, S. Smith, J. Edmonds, H. Li, L. Clarke, K. Calvin, and A. Thomson, 2015: A comprehensive view of global potential for hydro-generated electricity. *Energy Environ. Sci.*, **8**, 2622–2633, <https://doi.org/10.1039/C5EE00888C>.

Copyright of Journal of Hydrometeorology is the property of American Meteorological Society and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.