

## Assessment of Seasonal Soil Moisture Forecasts over Southern South America with Emphasis on Dry and Wet Events

PABLO C. SPENNEMANN

*Centro de Investigaciones del Mar y la Atmósfera, Consejo Nacional de Investigaciones Científicas y Técnicas–Universidad de Buenos Aires, UMI–Instituto Franco-Argentino sobre Estudios de Clima y sus Impactos/CNRS, Buenos Aires, Argentina*

JUAN A. RIVERA

*Instituto Argentino de Nivología, Glaciología y Ciencias Ambientales, Consejo Nacional de Investigaciones Científicas y Técnicas, Centro Científico Tecnológico Mendoza, and Universidad Juan Agustín Maza, Mendoza, Argentina*

MARISOL OSMAN

*Centro de Investigaciones del Mar y la Atmósfera, Consejo Nacional de Investigaciones Científicas y Técnicas–Universidad de Buenos Aires, UMI–Instituto Franco-Argentino sobre Estudios de Clima y sus Impactos/CNRS, Buenos Aires, Argentina*

A. CELESTE SAULO

*Centro de Investigaciones del Mar y la Atmósfera, Consejo Nacional de Investigaciones Científicas y Técnicas–Universidad de Buenos Aires, UMI–Instituto Franco-Argentino sobre Estudios de Clima y sus Impactos/CNRS, and Servicio Meteorológico Nacional, Buenos Aires, Argentina*

OLGA C. PENALBA

*Departamento de Ciencias de la Atmósfera y los Océanos, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, and Consejo Nacional de Investigaciones Científicas y Técnicas, Buenos Aires, Argentina*

(Manuscript received 17 January 2017, in final form 21 June 2017)

### ABSTRACT

The importance of forecasting extreme wet and dry conditions from weeks to months in advance relies on the need to prevent considerable socioeconomic losses, mainly in regions of large populations and where agriculture is a key value for the economies, such as southern South America (SSA). To improve the understanding of the performance and uncertainties of seasonal soil moisture and precipitation forecasts over SSA, this study aims to 1) perform a general assessment of the Climate Forecast System, version 2 (CFSv2), soil moisture and precipitation forecasts against observations and soil moisture simulations based on GLDAS, version 2.0; 2) evaluate the ability of CFSv2 to represent wet and dry events through the forecasted standardized precipitation index (SPI) and standardized soil moisture anomalies (SSMA); and 3) analyze the capability of a statistical methodology (merging observations and forecasts) in representing a severe drought event. Results show that both SPI and SSMA forecast skill are regionally and seasonally dependent. In general, a fast degradation of the forecasts skill is observed as the lead time increases, resulting in almost no added value with regard to climatology at lead times longer than 3 months. Additionally, a better performance of the SSMA forecasts is observed compared to SPI calculated using three months of precipitation (SPI3), with a higher skill for dry events against wet events. The CFSv2 forecasts are able to represent the spatial patterns of the 2008/09 severe drought event, although it shows crucial limitations regarding the identification of drought onset, duration, severity, and demise, considering both meteorological (SPI) and agricultural (SSMA) drought conditions.

---

*Corresponding author:* Juan A. Rivera, [jrivera@mendoza-conicet.gob.ar](mailto:jrivera@mendoza-conicet.gob.ar)

DOI: 10.1175/JHM-D-17-0015.1

© 2017 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy \(www.ametsoc.org/PUBSReuseLicenses\)](https://www.ametsoc.org/PUBSReuseLicenses).

## 1. Introduction

Extreme climate events are one of the most costly natural disasters in southern South America (SSA), a region that is prone to experience prolonged drought events and flooding episodes that can even cause human casualties. Within the regions of SSA affected by these extremes lies the La Plata basin (LPB), a reservoir of enormous biological wealth in which agriculture is the main economic activity. The region is vulnerable to several types of precipitation extremes spanning a wide range of temporal and spatial scales (Carril et al. 2016) that can produce flooding in certain areas (Cavalcanti 2012). Conversely, periods of extremely dry conditions affect the agriculture and undermine the water resources needed for hydroelectric power (Cavalcanti et al. 2015).

Several works have addressed the relationship between local and remote forcings on droughts and flood events in LPB. Previous studies have reported the prevalence of humid conditions (i.e., excess of precipitation) over LPB during the warm phase of El Niño–Southern Oscillation (ENSO) while dry conditions are observed during the cold phase (Grimm et al. 2000; Fraisse et al. 2008; Penalba and Rivera 2016). On the other hand, Mo and Berbery (2011) mention that local influences, like soil moisture and atmospheric moisture transport, provide the frame for extreme events to persist and evolve with more intensity. In particular, soil moisture is a key variable of the earth–atmosphere system that not only reflects the soil conditions of a given region (e.g., as an indicator of agricultural droughts), but also has the potential to influence the atmospheric variability by controlling the water and energy balances at the surface, from synoptic to seasonal time scales (e.g., Kanamitsu et al. 2003; Betts 2009; Seneviratne et al. 2010). For instance, the coupling strength between soil moisture, evapotranspiration, and temperature shows a hot spot over LPB (Ruscica et al. 2015; Spennemann and Saulo 2015; Sörensson and Menéndez 2011), which might explain some degree of local control over precipitation variability.

The importance that wet and dry events have on socioeconomic activities has motivated the development of drought monitoring and prediction tools (e.g., Dutra et al. 2014a,b; Yoon et al. 2012, hereafter Y12). In this sense, the standardized precipitation index (SPI; McKee et al. 1993) has been recommended by the World Meteorological Organization (WMO) as a reference drought index. The SPI is widely used for drought definition and monitoring due to its simplicity and flexibility, given that it is based only on precipitation data and can be calculated on any time scale. In this context, Y12

developed a method that combines 3 and 6 months of forecasted precipitation with observations to generate predictions of SPI over the United States using the National Centers for Environmental Prediction (NCEP) Climate Forecast System, version 1 (CFSv1; Saha et al. 2006). They found that the forecast performance strongly depends on the region and the season. Also, for the first 3-month forecast the skill comes mainly from the observations and then it drops, which limits its practical usefulness once that period is exceeded. On the other hand, recent studies use simulated soil moisture fields to monitor and predict drought conditions (Zhang et al. 2017). Soil moisture simulations represent the physical mechanisms of water and energy balances in the soil and at the surface, including the interactions with the vegetation. In this sense, Dirmeyer (2013) analyzes the performance of the soil moisture NCEP Climate Forecast System, version 2 (CFSv2; Saha et al. 2014), ensemble mean forecasts against Climate Forecast System Reanalysis (Saha et al. 2010), which is used as the initial condition of the CFSv2. The author documents that precipitation forecasts are in general more accurate in areas where there are larger soil moisture anomalies at the initial stage, concluding that the CFSv2 forecast skill for soil moisture is higher and for a longer time than for precipitation. Despite the fact that both approaches—drought monitoring through SPI or through simulations—have been extensively used worldwide, their potential use in regions like SSA has not yet been addressed.

The abovementioned assessments are limited to the number of in situ soil moisture observations, a factor that stands out in South America (see Dorigo et al. 2013, their Fig. 1). Instead of direct observations, several studies used the SPI as a proxy for soil moisture variations (Mueller and Seneviratne 2012), or soil moisture estimations derived from a combination of satellite microwave sensors (Wagner et al. 2012). In this context, the study by Spennemann et al. (2015) assesses the variability and physical consistency between the Global Land Data Assimilation System (GLDAS; Rodell et al. 2004) standardized soil moisture anomalies (SSMA) against the SPI and multisatellite soil moisture estimations over SSA. They conclude that the simulation of GLDAS, version 2.0 (GLDAS-2.0), is a useful tool to describe soil moisture anomalies over SSA and can be employed to develop soil moisture indices for agricultural production management. These findings provide the opportunity to evaluate soil moisture forecasts using GLDAS dataset, a strategy on which the present work is based. However, it is important to mention that GLDAS should not be considered as the ground truth soil

moisture, but as a reference soil moisture condition based on our previous studies over this region of South America.

The latter studies open a line of research regarding the potential for soil moisture predictions in a region where agriculture is a fundamental part of the economy. The main goal in this study is to evaluate the possible advantages of using SSMA against SPI forecasts, with emphasis on wet and dry events. This evaluation will be performed against simulated soil moisture and observed precipitation using different verification metrics. Then the performance of CFSv2 will be evaluated in detail for the 2008/09 severe drought. This assessment will incorporate the methodology developed by Y12, that is, merging in this case simulations with forecasted SSMA. This study will help to determine if the quality of SSMA forecasts is higher than that of precipitation forecasts, a result that could have potential benefits for the agricultural practices of the region. The structure of the paper is the following: section 2 describes the different datasets and the region of study, section 3 describes the metrics used for the forecast assessment and the application of Y12 methodology, section 4 exposes the results, and section 5 presents the discussion and conclusions.

## 2. Study area and data

The study focuses on the southern portion of LPB, also called southeastern South America (SESA; 35°–25°S, 63°–50°W). The region encompasses the southern portion of Brazil and Paraguay, Uruguay, and the northeastern portion of Argentina, including the Pampas region, one of the most productive soils of the world (Penalba and Bettolli 2011).

NCEP developed the CFSv2, a fully coupled land–ocean–atmosphere dynamical seasonal prediction system, which became operational in March 2011

(Saha et al. 2014). The CFSv2 improved the model parameterizations and assimilation systems and increased the spatial resolution with respect to the previous version of the Climate Forecast System (Mo et al. 2012; Saha et al. 2014). Both versions of the Climate Forecast System were widely used for seasonal forecasting applications considering variables such as 2-m temperature and precipitation (Wang et al. 2010), sea surface temperature (Kumar et al. 2012), and soil moisture (Mo et al. 2012). In this study, the retrospective 9-month CFSv2 forecasts from January 1982 to December 2010 were used. The reforecasts consist of 24 ensemble members (28 for November) initiated every fifth day, for all four cycles of each day (0000, 0600, 1200, and 1800 UTC). The CFSv2 spatial resolution was interpolated to GLDAS-2.0 (1° × 1°) regular latitude–longitude grid. The evaluation focuses on the complete set of forecast lead times (from 0 to 9 months), although special attention will be given to the ensemble mean lead times 0, 1, 2, and 5 (F0, F1, F2, and F5, respectively). Lead times 0–2 are chosen for the evaluation assuming that the model will have the highest performance during the early leads. The notation for the lead times 0–9 or F0–F9 correspond to the 0–1/9–10 month mean value of the forecast.

The reference soil moisture is based on GLDAS-2.0 (1° × 1°) monthly simulations, which use the Noah land surface model (LSM), version 3.3 (Ek et al. 2003). The CFSv2 also uses the Noah LSM, making the analysis compatible (i.e., same soil depths and dynamics). The root zone layer (RZ; 0–100 cm) was considered to calculate the SSMA for the evaluation of the soil moisture forecasts and the identification of agricultural drought conditions, based on the agreement between RZ SSMA and SPI3 and SPI6 (where 3 and 6 refer to the number of months of precipitation used to calculate SPI) shown in Spennemann et al. (2015). The SSMA is calculated for each grid point ( $i, j$ ) as

$$\text{SSMA}(i, j, t) = \{\text{SM}(i, j, t) - \text{mean}[\text{SM}(i, j, t_{\text{month}})]\} / \text{std dev}[\text{SM}(i, j, t_{\text{month}})], \quad (1)$$

where  $t_{\text{month}}$  corresponds to each month (i.e., January–December), std dev is the standard deviation, and SM is the soil moisture. Thus, for each month the mean value of that month is subtracted and then divided by its corresponding interannual variability.

The observational precipitation dataset was obtained from the Global Precipitation Climatology Centre (GPCC) full data reanalysis, version 7, at 1.0° grid resolution (Schneider et al. 2015). Figure 1 shows the spatial distribution of the mean monthly RZ soil moisture

values over the region under study. Higher values are located in central and eastern portions of Argentina and Paraguay, southern Brazil, and Uruguay, ranging from 280 to 350 mm, favoring rain-fed agriculture over that region. The spatial distribution of mean monthly precipitation amounts is similar to that for soil moisture, with regions with over 140 mm located in the northeastern part of the study area.

The SPI was calculated using the GPCC dataset considering precipitation accumulated over 3 and 6 months.

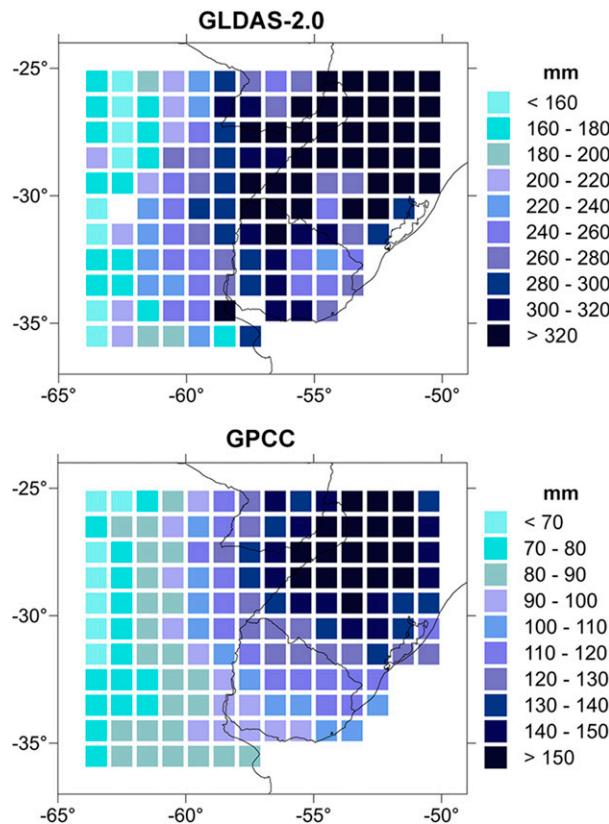


FIG. 1. (top) Climatological mean RZ soil moisture monthly values (mm) and (bottom) average accumulated monthly precipitation (mm) over 1982–2010.

The SPI3 and SPI6 were chosen because they are currently used to monitor dry and wet conditions by different institutions [e.g., Climate Prediction Center (CPC) and Servicio Meteorológico Nacional—Argentina (SMN)]. Also, temporal variability depicted by SPI3 and SPI6 is in agreement with that for SSMA based on GLDAS-2.0 over the same region (Spennemann et al. 2015). Following Penalba and Rivera (2015) for the calculation of the SPI, the accumulated precipitation time series were divided in 12 monthly series of 29 years, and each of them were fitted to a gamma probability density function. The 12 probability density functions for each time scale were translated to 12 cumulative density functions. Finally, an equiprobability transformation from the cumulative density functions to the standard normal distribution with mean of 0 and variance of 1 were performed to obtain the SPI.

When considering the CFSv2 data, forecasted SPI3 was calculated based on the accumulation of precipitation using F0 + F1 + F2 for each month. For example, for April 1982 the forecast of February 1982 (F0) + March 1982 (F1) + April 1982 (F2) was accumulated to construct the SPI3 for April 1982. Then it was

compared to the observed SPI3 (based on GPCP data). In this way, the accumulation over different lead times of a single forecast is used. The same criterion is applied to build up the SPI6, considering the accumulation of precipitation from F0 to F5.

### 3. Methodology

Model ensemble mean performance is evaluated through the computation of the root-mean-square error (RMSE), the anomaly correlation coefficient (ACC), the mean square skill score (MSSS), and the relative operating characteristic (ROC). The RMSE has been used as a standard statistical metric to measure model performance in meteorology, air quality, and climate research studies (Chai and Draxler 2014). The RMSE gives an indication of the mean deviation of the estimated values  $F$  compared to the magnitude of the estimated parameter  $O$ , defined as

$$\text{RMSE}(i, j, \tau) = \sqrt{\frac{1}{N} \sum_{t=1}^N [F(i, j, t, \tau) - O(i, j, t)]^2}, \quad (2)$$

where  $i$  and  $j$  represent latitude and longitude and the summation is over time  $t$  (1:  $N$ ) for a specific forecast lead time  $\tau$  (0–9 months). The RMSE gives information about the average magnitude of the forecast errors, assigning greater weight to the larger errors.

The ACC is one of the most widely used measures in the verification of spatial fields (Jolliffe and Stephenson 2011), regarded as a skill score relative to the climate:

$$\text{ACC}(i, j, \tau) = \frac{\sum_t F'(i, j, t, \tau) O'(i, j, t)}{\left\{ \sum_t [F'(i, j, t, \tau)]^2 \sum_t [O'(i, j, t)]^2 \right\}^{1/2}}, \quad (3)$$

where  $i$  and  $j$  correspond to the grid position, the summation is done over time  $t$  for a specific forecast lead time  $\tau$ , and the prime denotes the temporal anomaly after removing the mean annual cycle. The ACC can also be calculated over space and for a specific month of the year, that is,  $\text{ACC}(\text{month}, \tau)$ . As detailed in Becker et al. (2013, hereafter B13), the AAC returns a number between  $-1.0$  and  $1.0$ , where  $1.0$  refers to a perfect forecast and  $0$  to a random forecast. The ACC significance was assessed through a bootstrap resampling procedure (Efron and Tibshirani 1993), based on 1000 resamples for a 95% confidence interval.

The MSSS is essentially the mean square error (MSE) of the forecasts compared to the standard deviation of the variable for a grid point  $(i, j)$ , given by

$$\text{MSSS}(i, j, \tau) = 1 - \frac{\text{MSE}_{i,j,\tau}}{\text{MSEc}_{ij}}, \quad (4)$$

where

$$\text{MSE} = \frac{1}{N} \sum_{t=1}^N [(F_{i,j,t,\tau} - O_{i,j,t}) - (\overline{F_{i,j,t,\tau}} - \overline{O_{i,j,t}})]^2 \quad (5)$$

and

$$\text{MSEc} = \frac{1}{N} \sum_{t=1}^N (\overline{O_{ij}} - O_{ij})^2, \quad (6)$$

where the letter c in MSEc represents the climatology. Values of MSSS higher (lower) than 0 mean that the forecast error is lower (higher) than the standard deviation of the variable.

The ROC curves are useful for assessing the accuracy of predictions, measuring the ability of the forecasting system to discriminate between events (wet and dry) and nonevents. For the definition of SPI and SSMA dry and wet events, we considered a threshold of plus/minus one standard deviation for the observed and forecasted time series. This threshold, in the case of the SPI, defines droughts and excess conditions (Penalba and Rivera 2015). The ROC curves are constructed plotting the hit rate (HR) against the false alarm rate (FAR). The HR and FAR indicate the proportion of events forecasted correctly and the proportion of nonevents incorrectly forecasted, respectively. Different thresholds were fixed in order to provide the dichotomous wet/dry or nonwet/dry predictions, resulting in 100 pairs of (FAR, HR). The area under the ROC curve is commonly used as a metric representing the skill of the forecast system. The area is standardized against the total area of the figure, such that a perfect forecast has an area of 1 and a curve lying along the diagonal (no information, HR = FAR) has an area of 0.5 (Dutra et al. 2014b).

Before applying any metrics to the SSMA and SPI forecasts, it was necessary to correct the systematic errors (i.e., bias) of the model, to highlight the skillful part of the forecast (B13). Thus, the forecasted soil moisture and precipitation anomalies were constructed subtracting, for each different lead time and month, the corresponding forecasted climatology as described in B13.

The 2008/09 drought event is further analyzed following Y12. First, a bias correction based on the 1982–2006 forecast climatology is applied to the CFSv2 forecasts initialized in November 2007. Then, the CFSv2 SSMA from November 2007 to August 2008 is added to each of the corresponding months of the GLDAS-2.0 1982–2006 mean annual cycle. These new 10-month soil moistures (CFSv2 + GLDAS-2.0) are appended to

GLDAS-2.0 from January 1980 to October 2007, and finally the SSMA is calculated based on this new combined dataset. This calculation is applied to the ensemble mean and to the 24 members of the ensemble.

## 4. Results

### a. General evaluation of CFSv2 SSMA and SPI forecasts

The ACC of the CFSv2 ensemble mean SSMA and precipitation forecast as a function of the lead time and the target month averaged over the domain is shown in Fig. 2. ACC values higher than 0.4 are significant at the 95% confidence level. ACC is higher for SSMA compared to precipitation for each month and lead time, which highlights the reason while variables/indices with stronger memory (i.e., persistence), like SSMA and SPI (3–6) are used for agricultural drought monitoring instead of using just precipitation. In addition, the ACC for SSMA is significant for all forecast months at lead 0 month, with values from 0.4 up to 0.7, and is significant for 5 months (target month: April–August) at 1-month lead time. For the winter season, SSMA presents the highest skill with significant ACC values up to 2-month lead time. The higher skill achieved by CFSv2 during winter over South America was also observed in other studies using precipitation (B13). Finally, the lack of skill of CFSv2 for monthly precipitation over SESA should be pointed out, as ACC is below 0.2 for most of the forecast months and lead times.

An evaluation of forecasted SSMA, SPI3, and SPI6 spatial representation based on the CFSv2 ensemble mean was carried out. Figure 3 compares for the same verifying date/months the spatial pattern of ACC for each grid point for the SSMA, SPI3, and SPI6 forecasted at different lead times: 0, 1, 2, and 5 months. As the SPI is calculated using the accumulation over different forecast lead times, depending on the SPI time window, the comparison of SSMA against SPI3 is only possible for F2 onward and against SPI6 for F5 onward. For instance, SPI3 needs three months of precipitation to perform the accumulation (e.g., F0 + F1 + F2). For SSMA, ACC at lead 0 shows maximum and statistically significant values over Uruguay (>0.6), southern Brazil, and central and eastern Argentina and nonsignificant values elsewhere. For F1 (lead 1) the SSMA ACC is lower than for F0, being significant only over east-central Argentina and over parts of Uruguay. In general, as the lead time increases, a fast degradation of the SSMA and SPI3 skill is observed. The SSMA ACC values, for example, decrease from 0.4 for F1 to 0.3 in the case of F2 and 0.1 for F5 (see Fig. 3). In particular, for longer lead times (i.e., F5), ACC is nonsignificant

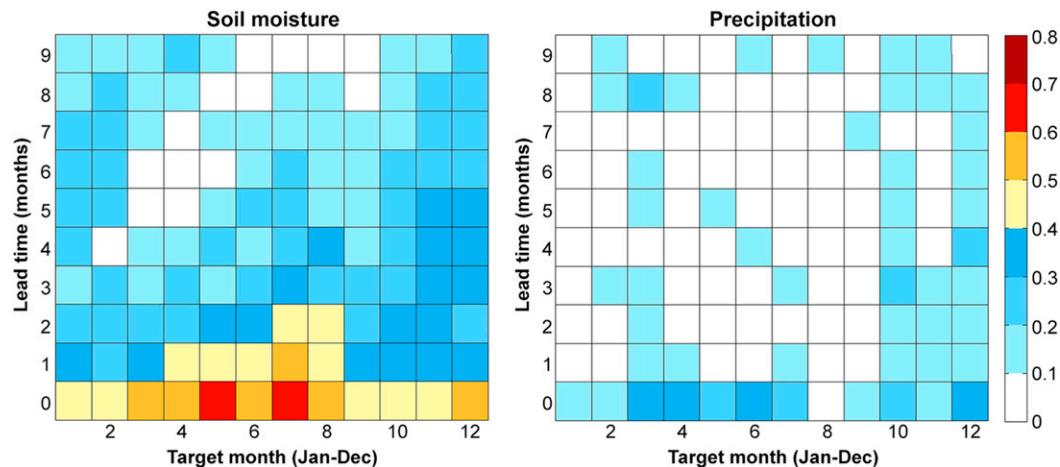


FIG. 2. Area-averaged ACC for CFSv2 ensemble mean forecast of (left) SSMA (against GLDAS-2.0) and (right) precipitation (against GPCC). The  $x$  axis shows the target month, and the  $y$  axis shows the forecast lead times (months; where number 0 corresponds to F0 and so on).

everywhere for SSMA and SPI3, and even negative for SSMA over some grid points.

In contrast to the region where SSMA ACC maximizes, the ACC values for SPI3 (F2) and SPI6 (F5) forecasts maximize over southern Brazil. The SPI6 shows higher ACC values compared to SPI3 (F5) and SSMA (F5). This can be attributed to the higher auto-correlation of SPI6 due to the accumulation of precipitation over a larger period of time.

The lack of skill for longer lead times was also documented by other studies analyzing the performance of the CFSv2 forecasting different hydrometeorological variables (e.g., Dirmeyer 2013; Dutra et al. 2014b; Siegmund et al. 2015). In this case, the ACC results highlight some potential advantages of SSMA over SPI3 for particular regions. A possible explanation for these regional differences in the performance of CFSv2 for SSMA and SPI3 might be in part related to differences between GLDAS-2.0 and GPCC precipitation datasets [e.g., see Fig. 3 of Spennemann et al. (2015)]. But, the comparison of the CFSv2 precipitation forecast against both GLDAS-2.0 and GPCC precipitation shows a similar spatial pattern and range of ACC values (Fig. 4), suggesting that the differences in skill observed are not related to the CFSv2 performance for precipitation. In a similar way to the ACC analysis, the RMSE for the different forecast indices shows that, in general, for all lead times, the SSMA forecast errors are smaller than SPI3 over central Argentina (not shown). The areas with larger RMSE values (typically higher than 1.4 mm) are located over the areas with nonsignificant ACC values in Fig. 3.

In summary, the forecast skill analysis reveals that the CFSv2 SSMA forecast exhibits a better performance over central Argentina compared to SPI3 based on the

ACC and RMSE metrics. On the other hand, the performance of SPI6 shows higher skill than SSMA (F5) over southern Brazil and a similar behavior over a few grid points over central Argentina and Uruguay. The better skill can also be attributed to the higher auto-correlation of SPI6 compared to SSMA. The regional SSMA and SPI differences are not related to differences in the precipitation datasets.

#### b. CFSv2 performance: Seasonal, wet, and dry events

To further document and understand seasonal differences in the SSMA spatial skill, the mean ACC was computed for the austral summer [i.e., target months of December–February (DJF)], autumn [March–May (MAM)], winter [June–August (JJA)], and spring [September–November (SON)] (Fig. 5, left) with initial conditions from November, February, May, and August, respectively (that is, lead 1–3 months). To complement this analysis, the spatial distribution of the mean MSSS (Fig. 5, right) was also calculated. The winter season presents the highest percentage of grid points at  $ACC > 0.4$  with 36.6%, located in the center of the domain, followed by spring, with 33.7%. Meanwhile summer and autumn show the lowest percentages, both with 18.2%. However, the range of values across the domain changes among the seasons. For example, summer shows a higher ACC mean value ( $ACC = 0.28$ ) than autumn ( $ACC = 0.25$ ), while the maximum mean ACC values are in winter ( $ACC = 0.37$ ) and spring ( $ACC = 0.33$ ). The spatial distribution of the MSSS shows positive values over the areas where ACC maximize in each of the seasons. On the other hand, negative MSSS values are distributed among the regions with nonsignificant ACC values, indicating higher forecasted MSE

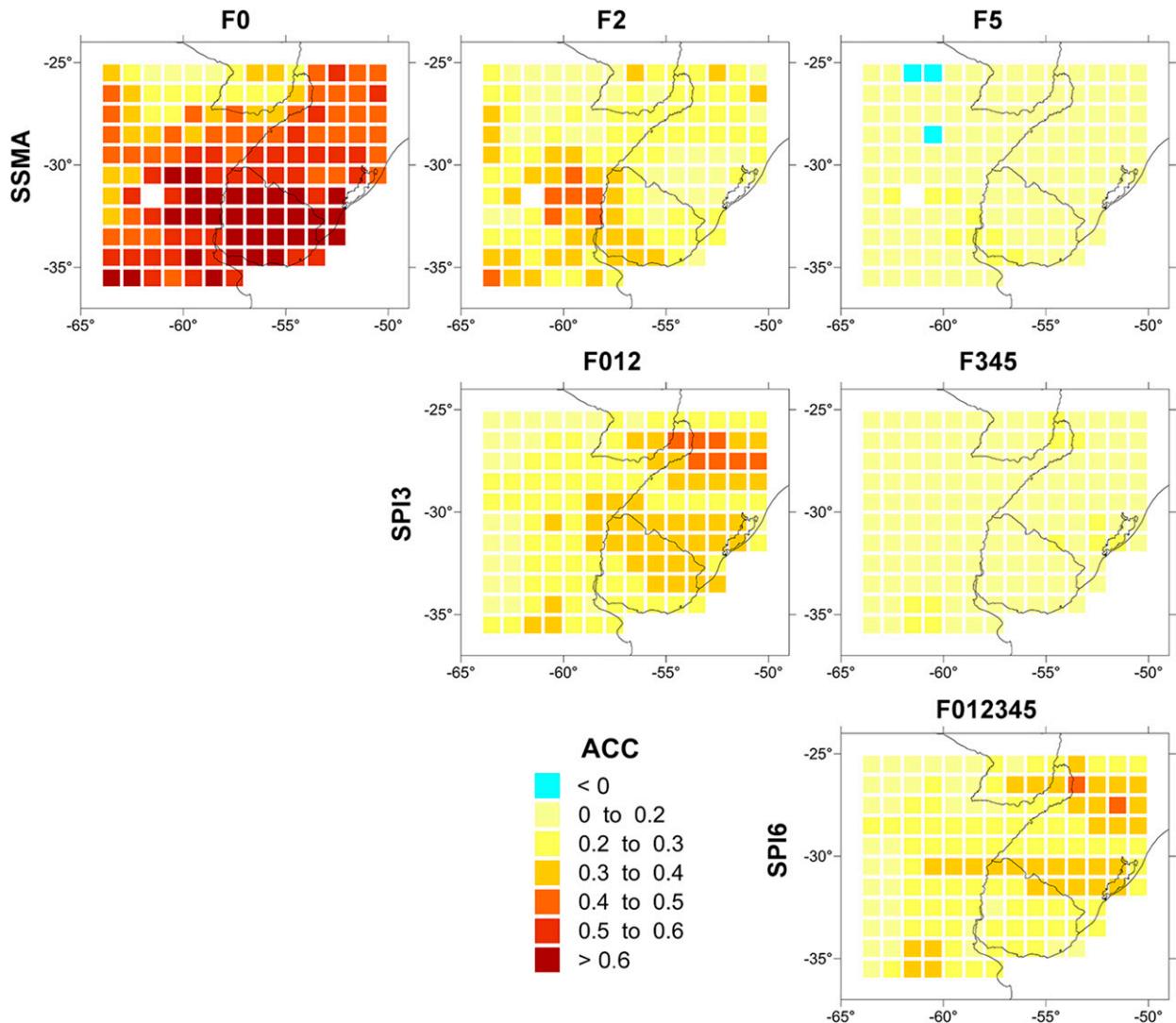


FIG. 3. ACC of CFSv2 ensemble mean forecast for (top) SSMA, (middle) SPI3, and (bottom) SPI6. For SSMA, forecast lead times are F0, F2, and F5. For SPI3, the lead times span from F0 to F2 and from F3 to F5. For SPI6, the lead times span from F0 to F5.

compared to the observed standard deviation. In turn, the ensemble MSSS shows lower values for summer–autumn than in winter–spring. This result denotes that seasonal forecasts do not provide valuable information, except over limited areas.

The ability of CFSv2 to predict the observed wet or dry events is further evaluated through ROC curves. These events were defined as the periods of time exceeding plus/minus one standard deviation, considering the regional time series of SSMA, SPI3, and SPI6. For instance, 22 wet events and 24 dry events were obtained for the SPI3, with average durations between 2 and 3 months. In the case of SPI6, the effect of increasing the time scale—that is, accumulating a larger number of monthly precipitation totals—led to a decrease in the

analyzed wet and dry events (15 and 14 events, respectively) and an increase in its mean duration (between 3 and 5 months). For each variable—SSMA (F0, F2, and F5), SPI3 (F0–F2 and F3–F5), and SPI6 (F0–F5)—we computed the ROC curve for all the grid points in the domain and the entire period (Fig. 6). In all cases the area under the ROC curve decreases with lead time, which means that the forecasts are less able to discriminate between occurrences and nonoccurrences of the wet or dry event as lead time increases. For F0, F2, and F5 lead times, SSMA ROC area is higher for dry events than for wet events. Even when the performance of the F2 forecast showed in general a low ACC and a high RMSE, the ROC curves for SSMA dry and wet events show some skill, with the curves bending above the

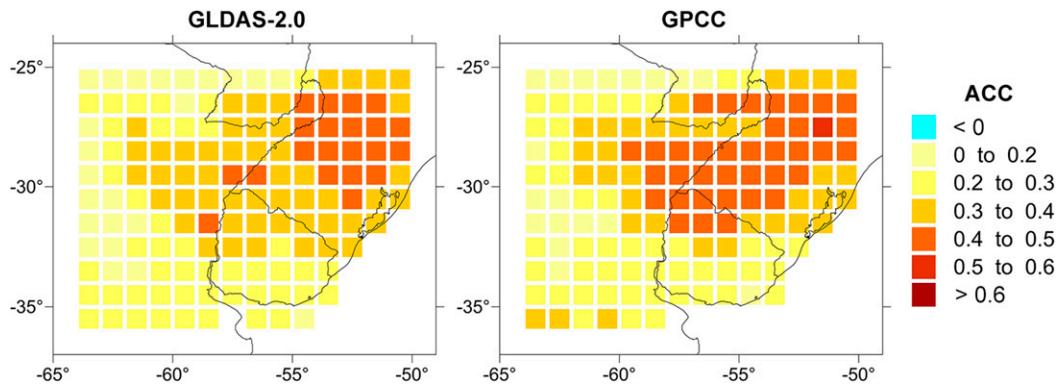


FIG. 4. ACC for F0 CFSv2 precipitation forecast against (left) GLDAS-2.0 and (right) GPCC precipitation databases over the 1982–2010 period.

diagonal—that is, the HR exceeds the FAR. For SPI3 and SPI6, the ROC curves also show higher predictability for the dry events, except when using lead times F3–F5 for SPI3.

The comparison between SSMA and SPI3 (F0–F2) forecast shows that, in general, the SSMA skill for dry events is higher than the skill of SPI3. But, for longer lead times, SSMA (F5) and SPI3 (F3–F5) both show the same skill for dry events. Moreover, the SPI6 shows higher predictability for dry events compared to SSMA (F5) and SPI3 (F3–F5) forecasts. The latter result is in agreement with the ACC analysis, which is related to the accumulation of precipitation for longer time scales. This accumulation acts as a low-pass filter, increasing the persistence of dry and wet events, which leads to forecasts that provide more valuable information. Based on this assessment, SSMA forecasts show an added value for the first few months compared to the SPI forecasts, particularly for dry events.

### c. Case study: The 2008/09 drought

In this section, the performance of CFSv2 in forecasting the 2008/09 drought event based on SSMA is analyzed. This drought event was severe in terms of intensity and spatial extension, related to a combination of La Niña conditions with large tropical North Atlantic warm sea surface temperature anomalies (Müller et al. 2014). It was the fourth most severe meteorological drought event in the 1980–2010 period after the 1988/89, 1996, and 2004 events. However, considering that the SSMA is indicative of agricultural drought occurrences, this drought event was characterized by the lowest SSMA value on record. The temporal evolution of the SSMA and SPI3 averaged over SESA (Fig. 7) shows good agreement between both indices ( $r = 0.75$ ,  $p < 0.001$ ), as previously shown by Spennemann et al. (2015), a fact that is also observed considering the SPI6

( $r = 0.81$ ,  $p < 0.001$ , not shown). The figure also shows that from both meteorological (SPI3) and agricultural (SSMA) points of view, the 2008/09 drought event reached severe conditions (SPI3 and SSMA less than or equal to  $-1.5$ ) during 2009 winter. The temporal evolution also shows three relative minima in both indices between 2008 and 2009. Moreover, the difference between SPI3 and SSMA during the end of the drought event shows that the recovery in terms of precipitation is faster than the response in terms of soil moisture anomalies, with a lag of approximately three months between the two variables.

Three characteristics of the CFSv2 SSMA forecasts were analyzed: 1) the determination of the onset of the drought event, 2) the representation of the persistence, and 3) the ability in terms of capturing the severity during the peak in winter 2009. Figure 8 shows the evolution of CFSv2 ensemble mean and individual members' SSMA forecasts for three different initial dates (October 2007, May 2008, and February 2009), after applying Y12 methodology. The representation of the onset of the drought event, analyzed from November 2007 to August 2008 (10 forecasted months, Fig. 8a), reveals that the dispersion increases rapidly with forecast length, being approximately constant after February 2008 (i.e., drifting to its own climatology). This can be linked to the increase in the dispersion—that is, negative MSSS over larger areas—observed in the spatial analysis during summer and autumn seasons (Fig. 5). During the first forecasted month, most of the members show an increase in SSMA, opposite to what was observed with GLDAS-2.0 data. The following six forecasted months are in line with the occurrence of a dry pulse but neither the intensity nor the timing of the drought are accurately forecasted by the CFSv2 ensemble. Regarding the persistence of this drought event, forecasts initialized during the development of the event

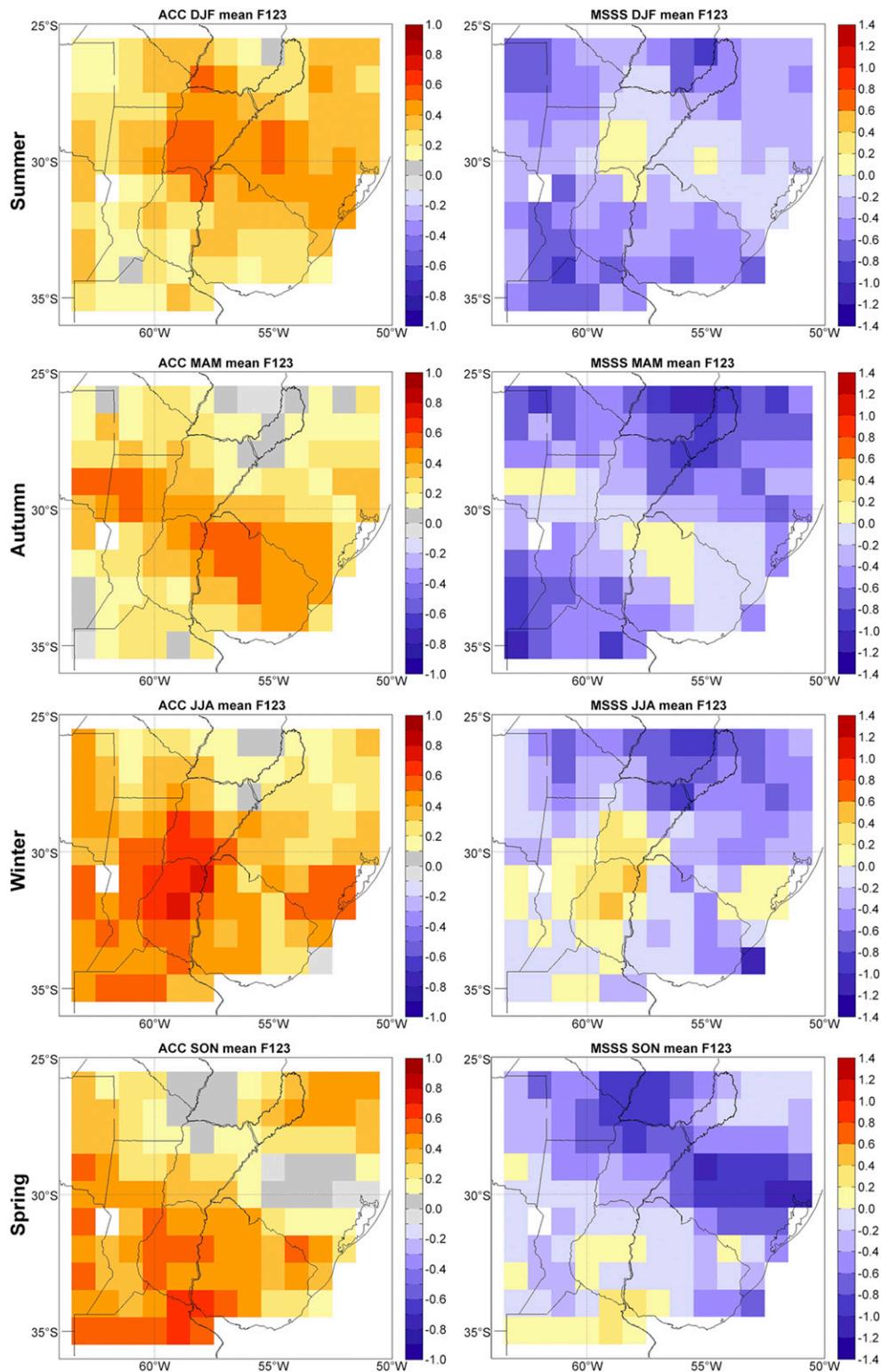


FIG. 5. CFSv2 ensemble mean SSMA forecasts of (left) mean seasonal ACC and (right) seasonal MSSS. Seasonal forecasts are constructed with CFSv2 F1, F2, and F3 lead times initialized the month before each corresponding season.

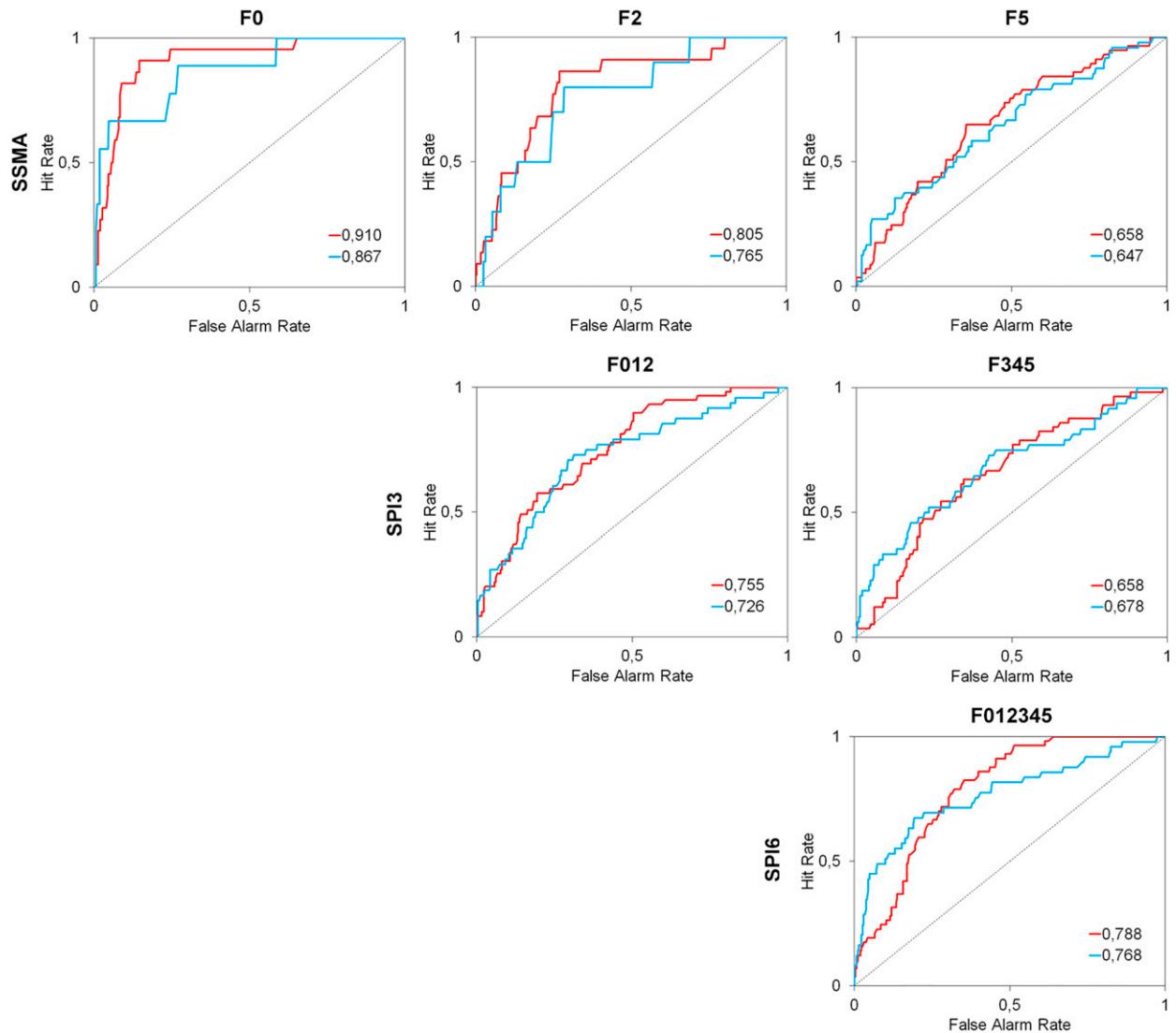


FIG. 6. ROC curves for wet (blue line) and dry (red line) conditions for (top) SSMA, (middle) SPI3, and (bottom) SPI6. For SSMA, forecast lead times are F0, F1, and F2. For SPI3, the lead times span from F0 to F2 and from F3 to F5. For SPI6, the lead times span from F0 to F5. Wet and dry events are based on the 1.0 and  $-1.0$  thresholds, respectively.

(May 2008) tend to progress rapidly toward nearly normal conditions, while the observations show a slight recovery until October 2008 (Fig. 8b), in line with an increase in precipitation as shown in the SPI3 time series (Fig. 7). Nevertheless, observations show that the dry event persisted for several months, reaching values of SSMA lower than  $-1.0$  during the summer of 2008. A final and relevant aspect to be assessed is the skill in the representation of the maximum severity peak between autumn and winter of 2009 (Fig. 8c). It can be observed that both the ensemble and the individual members initialized in January 2009 show a sharp increase in the SSMA values from February to March 2009, from  $-0.68$  to  $0.23$  considering the ensemble

(Fig. 8c). The ensemble for the forecasted months from March to November 2009 shows values around  $0.4$ , while the individual members exhibit larger dispersion between May and August, mostly with positive SSMA values. Conversely, observations show that the peak of the drought severity occurs during May 2009, reaching a value of  $-1.96$  and six consecutive months with SSMA below  $-1.0$ , which is not reproduced by CFSv2. Only two members show a peak with values lower than  $-1.0$ , but neither the timing nor the severity are accurately estimated. As previously shown in Fig. 5, autumn was characterized by low ACC values and a large proportion of negative MSSS values, which might be reflected in the behavior observed in Fig. 8c.

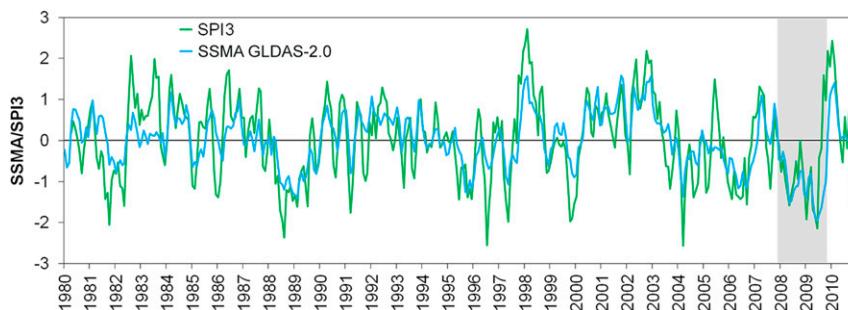


FIG. 7. Temporal evolution of SPI3 and GLDAS-2.0 SSMA averaged over SESA (1980–2010). The gray shading indicates the 2008/09 drought event.

In addition to the evaluation of the performance in forecasting the temporal evolution, the ability of CFSv2 to capture the SSMA spatial features/patterns of the event was analyzed at each stage (onset, persistence, and maximum intensity). Figure 9 shows the spatial distribution of the 3-month average SSMA values from GLDAS-2.0 (Fig. 9, left) for DJF 2007/08, July–September (JAS) 2008, and April–June (AMJ) 2009, compared with the CFSv2 forecast, after applying Y12 methodology, for those periods initialized in October 2007, May 2008, and February 2009, respectively. The SSMA forecast is able to represent the spatial dry condition of the onset (Fig. 9, top); forecasting the relative minimum over Uruguay and southern Brazil, but the severity of the drought is underestimated. The spatial features of drought persistence are captured accurately (Fig. 9, middle), especially over central Argentina, although the severity is still underestimated. The underestimation of the drought intensity observed at both stages of the analysis is in agreement with Fig. 8. Regarding the maximum drought severity, CFSv2 SSMA forecast is far from reproducing both the location and intensity of drought during June 2009, even forecasting wet conditions over the northeast portion of the domain (Fig. 9, bottom). This result is in line with the temporal evolution of both the ensemble and most of the individual members (Fig. 8c), progressing toward wet conditions while observations show an intensification of drought conditions, both in temporal and spatial scales.

## 5. Discussion and conclusions

This study assessed the performance of the CFSv2 in forecasting three commonly used drought indices relevant to agricultural activities (SSMA, SPI3, and SPI6) over SSA. The CFSv2 forecast performance was evaluated for the 1982–2010 period considering the general and seasonal behavior of the indices. Then, the analysis focused on overall dry and wet events, and

particularly on the 2008/09 extreme drought event that affected SSA.

For each month and lead time considered, SSMA are better forecasted than precipitation. This result confirms the value of variables/indices with stronger memory (i.e., persistence), like SSMA and SPI rather than precipitation for drought monitoring. The predictive skill of SSMA and SPI3 declines with increasing lead time, being nonsignificant for lead times longer than three months. The longer time scale used to calculate the SPI adds autocorrelation to the time series, which results in a better performance of the forecasts for the SPI6 in comparison with SPI3 and SSMA. The highest signals for SSMA and SPIs are located over different regions within SSA, being highest over Uruguay and central Argentina (southern Brazil) for SSMA (SPI).

These results open the question of whether this dissimilar behavior between the SSMA and SPI forecasts is related to the precipitation datasets used in GLDAS-2.0, different from those used for SPI calculations. Results showed that the regional differences are not related to differences in the precipitation dataset. Then, the regional differences in skill between SSMA and SPI might be linked to 1) a stronger persistence of SSMA compared to the SPI3 over particular regions and 2) the usage of Noah LSM by both CFSv2 and GLDAS-2.0, thus reflecting the impact of the soil model on these results. Further studies using different datasets and numerical experiments to analyze the physical mechanism that has taken place (e.g., the role of evapotranspiration) are needed to clarify this issue. As this type of analysis was beyond the scope of this study, the results found here should motivate future studies focusing on these regional differences.

A strong seasonality of the SSMA forecasts skill was observed, showing a better performance for austral winter than for summer. Thus, the better performance for the winter season can provide, for example, more valuable information for regional (i.e., in relation with

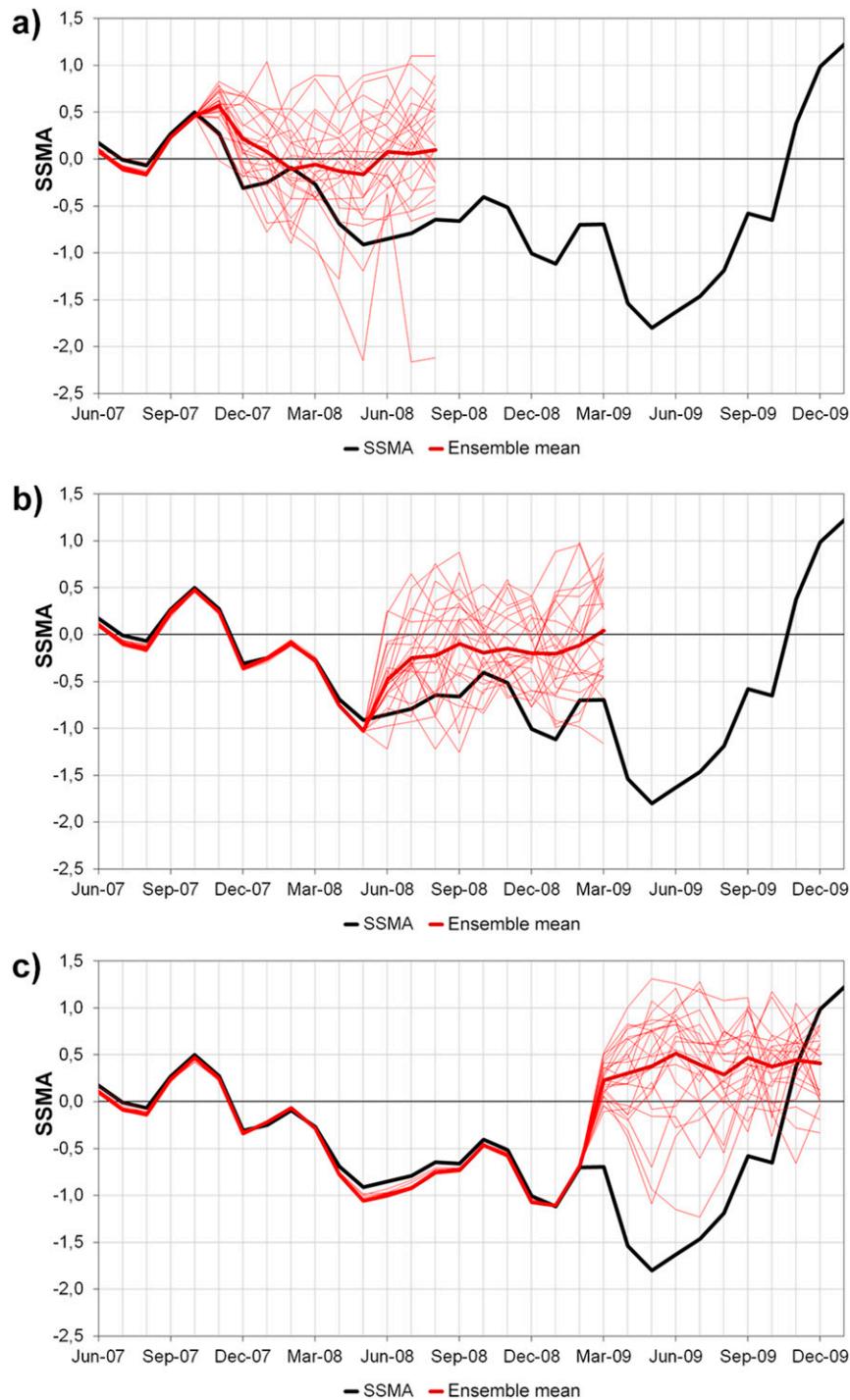


FIG. 8. Evolution of CFSv2 SSMA forecasts for three different initial dates: (a) October 2007, (b) May 2008, and (c) February 2009. The black line represents GLDAS-2.0 SSMA based on the 1982–2010 period. The thick red line represents the CFSv2 ensemble mean, while the thin red lines correspond to each of the 24 CFSv2 members. Forecasts were corrected following Y12 methodology (see text).

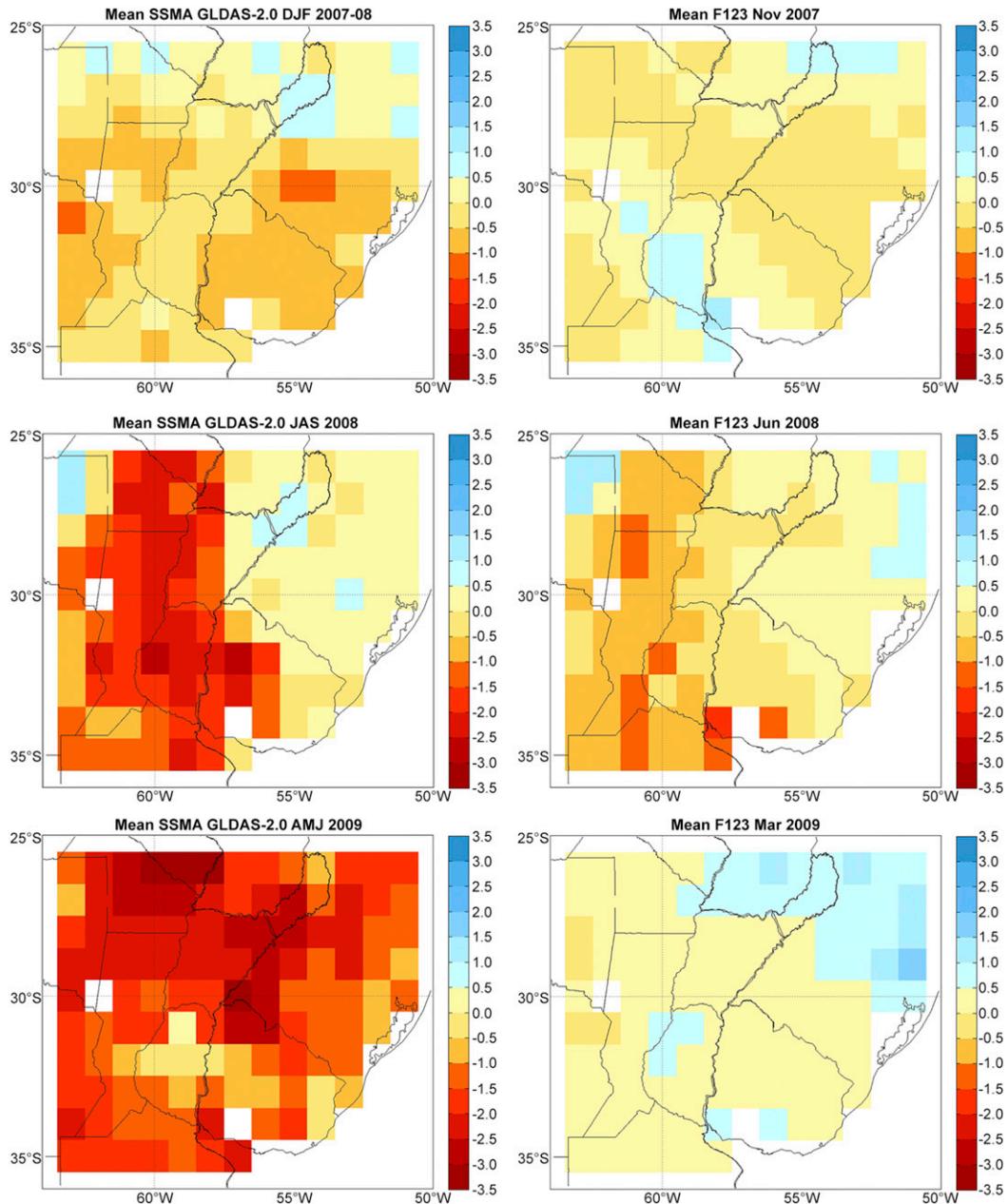


FIG. 9. (left) Spatial features of GLDAS-2.0 SSMA mean value for (top) DJF 2007/08, (middle) JAS 2008, and (bottom) AMJ 2009. (right) Y12 methodology applied to the mean values of F1, F2, and F3 initialized in (top) November 2007, (middle) June 2008, and (bottom) March 2009. The GLDAS-2.0 SSMA is based on 1982–2010 climatology.

$1^\circ \times 1^\circ$  spatial resolution) planning of wheat crops over SSA than for summer crops (e.g., soybean). In addition, given that this region is characterized by strong interactions between soil moisture and surface variables (e.g., 2-m temperature), which maximize during summer, the lower SSMA forecast signal during this season might be translated into the forecast of surface variables, thus degrading CFSv2 forecast performance.

In general, a higher skill of SSMA and SPI for dry events was observed compared to wet episodes. This result can be attributed, in the case of SPI, to a large number of correct zero precipitation forecasts matching in time the precipitation observations. Also, as in the case of the ACC, the effect of autocorrelation of SPI6 led to better metrics when compared with the SPI3 results. In the case of SSMA, it showed a higher skill of

SSMA forecasts for dry and wet events compared to SPI3 and lower skill compared to SPI6 only for the dry events.

Despite the fact that CFSv2 is able to discriminate between the occurrence or not of dry events, the assessment of the 2008/09 extreme drought event showed that the CFSv2 ensemble forecasts have strong limitations regarding the identification of the onset, duration, severity, and demise of the event, based on the forecasted SSMA derived following the methodology proposed by Y12. While the temporal evolution of the negative anomalies of SSMA was underestimated in all stages, the spatial patterns of SSMA derived from this methodology, except for the most severe months, were represented in an accurate manner. As stated in Y12, the CFSv2 model is able to capture the impact of ENSO on precipitation only when the initial conditions already contain the ENSO signal. In this case, the drought event was related to La Niña conditions that extended from July–September 2007 to May–July 2008, followed by cold SST anomalies but neutral conditions until March–May 2009. Therefore, this might partially explain the poor performance of CFSv2 forecasts on reproducing the maximum drought intensity, which was observed between autumn and winter of 2009, after a La Niña event. However, neither the onset nor the persistence were reproduced accurately, even under a clear signal of La Niña. A potential limitation of the Y12 methodology is that it requires at least 30 years of data and a representative spatial coverage, which in some regions, like the one under analysis, is difficult to achieve. This limitation can prevent the assessment of past severe wet or drought events like those observed during 1986 or 1988/89 and can degrade the skill achieved by CFSv2. Moreover, considering that lead times longer than three months do not have forecast skill, their inclusion in the calculation of the Y12 methodology can degrade forecast performance. Nevertheless, the adaptation of Y12 methodology to SSMA ensemble forecasts showed some added value, in particular when the regional average series are used in combination with the spatial analysis.

The results documented in this study show the added value for considering SSMA for agricultural drought monitoring and forecasting in combination with the commonly recommended SPI. Both indices complement each other in a regional/spatial perspective, as each one maximizes its skill over different regions. However, before agricultural planning studies can be carried out based on this forecast system, further evaluations are needed, as it is not straightforward yet if the use of these seasonal indices can be determining in, for example, a sowing planning strategy. Last, the potential use of the

information derived from the seasonal forecast must be complemented with short-term forecasts and real-time monitoring to give an integrated picture of the soil and atmosphere conditions to provide the necessary tools for decision-makers.

*Acknowledgments.* This work was carried out with the aid of the following projects: the Inter-American Institute for Global Change Research (IAI) CRN3035, which is supported by the U.S. National Science Foundation (Grant GEO-1128040); the CLIMAX Project, which is funded by Belmont Forum; the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) PICT2010-2110; and the Secretaría de Ciencia y Técnica de la Universidad de Buenos Aires (UBACyT) 20020100100434BA and 20020130100263BA. The GLDAS data used in this study were acquired as part of the mission of NASA's Earth Science Division and archived and distributed by the Goddard Earth Sciences Data and Information Services Center (GES DISC). Marisol Osman is supported by a Ph.D. grant from the National Scientific and Technical Research Council (CONICET), Argentina.

#### REFERENCES

- Becker, E. J., H. Van Del Dool, and M. Peña, 2013: Short-term climate extremes: Prediction skill and predictability. *J. Climate*, **26**, 512–531, doi:10.1175/JCLI-D-12-00177.1.
- Betts, A. K., 2009: Land-surface-atmosphere coupling in observations and models. *J. Adv. Model. Earth Syst.*, **1**, doi:10.3894/JAMES.2009.1.4.
- Carril, A. F., and Coauthors, 2016: Extreme events in La Plata basin: A retrospective analysis of what we have learned during CLARIS-LPB project. *Climate Res.*, **68**, 95–116, doi:10.3354/cr01374.
- Cavalcanti, I. F. A., 2012: Large scale and synoptic features associated with extreme precipitation over South America: A review and case studies for the first decade of the 21st century. *Atmos. Res.*, **118**, 27–40, doi:10.1016/j.atmosres.2012.06.012.
- , and Coauthors, 2015: Precipitation extremes over La Plata basin—Review and new results from observations and climate simulations. *J. Hydrol.*, **523**, 211–230, doi:10.1016/j.jhydrol.2015.01.028.
- Chai, T., and R. R. Draxler, 2014: Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.*, **7**, 1247–1250, doi:10.5194/gmd-7-1247-2014.
- Dirmeyer, P. A., 2013: Characteristics of the water cycle and land-atmosphere interactions from a comprehensive reforecast and reanalysis data set: CFSv2. *Climate Dyn.*, **41**, 1083–1097, doi:10.1007/s00382-013-1866-x.
- Dorigo, W. A., and Coauthors, 2013: Global automated quality control of in situ soil moisture data from the International Soil Moisture Network. *Vadose Zone J.*, **12**, doi:10.2136/vzj2012.0097.
- Dutra, E., F. Wetterhall, F. Di Giuseppe, G. Naumann, P. Barbosa, J. Vogt, W. Pozzi, and F. Pappenberger, 2014a: Global

- meteorological drought—Part 1: Probabilistic monitoring. *Hydrol. Earth Syst. Sci.*, **18**, 2657–2667, doi:10.5194/hess-18-2657-2014.
- , and Coauthors, 2014b: Global meteorological drought—Part 2: Seasonal forecasts. *Hydrol. Earth Syst. Sci.*, **18**, 2669–2678, doi:10.5194/hess-18-2669-2014.
- Efron, B., and R. J. Tibshirani, 1993: *An Introduction to the Bootstrap*. Chapman & Hall/CRC, 456 pp.
- Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, doi:10.1029/2002JD003296.
- Fraisse, C. W., V. E. Cabrera, N. E. Breuer, J. Baez, J. Quispe, and E. Matos, 2008: El Niño–Southern Oscillation influences on soybean yields in eastern Paraguay. *Int. J. Climatol.*, **28**, 1399–1407, doi:10.1002/joc.1641.
- Grimm, A. M., V. R. Barros, and M. E. Doyle, 2000: Climate variability in southern South America associated with El Niño and La Niña events. *J. Climate*, **13**, 35–58, doi:10.1175/1520-0442(2000)013<0035:CVISSA>2.0.CO;2.
- Jolliffe, I. T., and D. B. Stephenson, 2011: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley and Sons, 292 pp.
- Kanamitsu, M., C.-H. Cheng-Hsuan, J. Schemm, and W. Ebisuzaki, 2003: The predictability of soil moisture and near-surface temperature in hindcasts of the NCEP seasonal forecast model. *J. Climate*, **16**, 510–521, doi:10.1175/1520-0442(2003)016<0510:TPOSMA>2.0.CO;2.
- Kumar, A., M. Chen, L. Zhang, W. Wang, Y. Xue, C. Wen, L. Marx, and B. Huang, 2012: An analysis of the non-stationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2. *Mon. Wea. Rev.*, **140**, 3003–3016, doi:10.1175/MWR-D-11-00335.1.
- McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. Preprints, *Eighth Conf. Applied Climatology*, Anaheim, CA, Amer. Meteor. Soc., 179–184.
- Mo, K. C., and E. H. Berbery, 2011: Drought and persistent wet spells over South America based on observations and the U. S. CLIVAR drought experiments. *J. Climate*, **24**, 1801–1820, doi:10.1175/2010JCLI3874.1.
- , S. Shukla, D. P. Lettenmaier, and L.-C. Chen, 2012: Do Climate Forecast System (CFSv2) forecasts improve seasonal soil moisture prediction? *Geophys. Res. Lett.*, **39**, L23703, doi:10.1029/2012GL053598.
- Mueller, B., and S. I. Seneviratne, 2012: Hot days induced by precipitation deficits at the global scale. *Proc. Natl. Acad. Sci. USA*, **109**, 12 398–12 403, doi:10.1073/pnas.1204330109.
- Müller, O. V., E. H. Berbery, D. Alcaraz-Segura, and M. B. Ek, 2014: Regional model simulations of the 2008 drought in southern South America using a consistent set of land surface properties. *J. Climate*, **27**, 6754–6778, doi:10.1175/JCLI-D-13-00463.1.
- Penalba, O. C., and M. L. Bettolli, 2011: Climate Change Impacts on Atmospheric Circulation and Daily Precipitation in the Argentine Pampas Region. *Climate Change—Geophysical Foundations and Ecological Effects*, J. Blanco and H. Kheradmand, Eds., Intech, 137–156, doi:10.5772/24163.
- , and J. A. Rivera, 2015: Comparación de seis índices para el monitoreo de sequías meteorológicas en el sur de Sudamérica. *Meteorologica*, **40** (2), 33–57.
- , and —, 2016: Precipitation response to El Niño/La Niña events in Southern South America—Emphasis in regional drought occurrences. *Adv. Geosci.*, **42**, 1–14, doi:10.5194/ageo-42-1-2016.
- Rodell, M., and Coauthors, 2004: The Global Land Data Assimilation System. *Bull. Amer. Meteor. Soc.*, **85**, 381–394, doi:10.1175/BAMS-85-3-381.
- Ruscica, R. C., A. A. Sörensson, and C. G. Menéndez, 2015: Pathways between soil moisture and precipitation in southeastern South America. *Atmos. Sci. Lett.*, **16**, 267–272, doi:10.1002/asl2.552.
- Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19**, 3483–3517, doi:10.1175/JCLI3812.1.
- , and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.*, **91**, 1015–1057, doi:10.1175/2010BAMS3001.1.
- , and Coauthors, 2014: The NCEP Climate Forecast System version 2. *J. Climate*, **27**, 2185–2208, doi:10.1175/JCLI-D-12-00823.1.
- Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, B. Rudolf, and M. Ziese, 2015: GPCC full data reanalysis version 7.0 at 1.0°: Monthly land-surface precipitation from rain-gauges built on GTS-based and historic data. DWD, accessed 10 June 2015, doi:10.5676/DWD\_GPCC/FD\_M\_V7\_100.
- Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Sci. Rev.*, **99**, 125–161, doi:10.1016/j.earscirev.2010.02.004.
- Siegmund, J., J. Bliefernicht, P. Laux, and H. Kuntzmann, 2015: Toward a seasonal precipitation prediction system for West Africa: Performance of CFSv2 and high-resolution dynamical downscaling. *J. Geophys. Res. Atmos.*, **120**, 7316–7339, doi:10.1002/2014JD022692.
- Sörensson, A. A., and C. G. Menéndez, 2011: Summer soil–precipitation coupling in South America. *Tellus*, **63A**, 56–68, doi:10.1111/j.1600-0870.2010.00468.x.
- Spennemann, P. C., and A. C. Saulo, 2015: An estimation of the land–atmosphere coupling strength in South America using the Global Land Data Assimilation System. *Int. J. Climatol.*, **35**, 4151–4166, doi:10.1002/joc.4274.
- , J. A. Rivera, A. C. Saulo, and O. C. Penalba, 2015: A comparison of GLDAS soil moisture anomalies against the standardized precipitation index and multisatellite estimations over South America. *J. Hydrometeorol.*, **16**, 158–171, doi:10.1175/JHM-D-13-0190.1.
- Wagner, W., W. Dorigo, R. de Jeu, D. Fernandez, J. Benveniste, E. Haas, and M. Ertl, 2012: Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture. *XXII ISPRS Congress*, Melbourne, Australia, ISPRS, 315–321. [Available online at <http://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/I-7/315/2012/isprsanals-I-7-315-2012.pdf>.]
- Wang, W., M. Chen, and A. Kumar, 2010: An assessment of the CFS real-time seasonal forecasts. *Wea. Forecasting*, **25**, 950–969, doi:10.1175/2010WAF2222345.1.
- Yoon, J.-H., K. Mo, and E. F. Wood, 2012: Dynamic-model-based seasonal prediction of meteorological drought over the contiguous United States. *J. Hydrometeorol.*, **13**, 463–482, doi:10.1175/JHM-D-11-038.1.
- Zhang, X., Q. Tang, X. Liu, G. Leng, and Z. Li, 2017: Soil moisture drought monitoring and forecasting using satellite and climate model data over southwestern China. *J. Hydrometeorol.*, **18**, 5–23, doi:10.1175/JHM-D-16-0045.1.