Precipitation Characteristics of the South American Monsoon System Derived from Multiple Datasets

LEILA M. V. CARVALHO

Department of Geography, and Earth Research Institute, University of California, Santa Barbara, Santa Barbara, California

CHARLES JONES

Earth Research Institute, University of California, Santa Barbara, Santa Barbara, California

ADOLFO N. D. POSADAS AND ROBERTO QUIROZ

International Potato Center (CIP), Lima, Peru

BODO BOOKHAGEN

Department of Geography, and Earth Research Institute, University of California, Santa Barbara, Santa Barbara, California

BRANT LIEBMANN

CIRES Climate Diagnostics Center, Boulder, Colorado

(Manuscript received 14 June 2011, in final form 27 January 2012)

ABSTRACT

The South American monsoon system (SAMS) is the most important climatic feature in South America and is characterized by pronounced seasonality in precipitation during the austral summer. This study compares several statistical properties of daily gridded precipitation from different data (1998–2008): 1) Physical Sciences Division (PSD), Earth System Research Laboratory [1.0° and 2.5° latitude (lat)/longitude (lon)]; 2) Global Precipitation Climatology Project (GPCP; 1° lat/lon); 3) Climate Prediction Center (CPC) unified gauge (CPC-uni) (0.5° lat/lon); 4) NCEP Climate Forecast System Reanalysis (CFSR) (0.5° lat/lon); 5) NASA Modern-Era Retrospective Analysis for Research and Applications (MERRA) reanalysis (0.5° lat/0.3° lon); and 6) Tropical Rainfall Measuring Mission (TRMM) 3B42 V6 data (0.25° lat/lon). The same statistical analyses are applied to data in 1) a common 2.5° lat/lon grid and 2) in the original resolutions of the datasets.

All datasets consistently represent the large-scale patterns of the SAMS. The onset, demise, and duration of SAMS are consistent among PSD, GPCP, CPC-uni, and TRMM datasets, whereas CFSR and MERRA seem to have problems in capturing the correct timing of SAMS. Spectral analyses show that intraseasonal variance is somewhat similar in the six datasets. Moreover, differences in spatial patterns of mean precipitation are small among PSD, GPCP, CPC-uni, and TRMM data, while some discrepancies are found in CFSR and MERRA relative to the other datasets. Fitting of gamma frequency distributions to daily precipitation shows differences in the parameters that characterize the shape, scale, and tails of the frequency distributions. This suggests that significant uncertainties exist in the characterization of extreme precipitation, an issue that is highly important in the context of climate variability and change in South America.

1. Introduction

The monsoon [hereafter the South American monsoon system (SAMS)] is the most important climatic feature in South America (Zhou and Lau 1998; Vera et al. 2006; Marengo et al. 2012). The main feature of the SAMS is the enhanced convective activity and heavy precipitation in tropical South America, which typically starts in October–November, is fully developed during December–February, and retreats in late April or early May (Kousky 1988; Horel et al. 1989; Marengo et al. 2001; Grimm et al. 2005; Gan et al. 2006; Liebmann et al. 2007; Silva and Carvalho 2007). Associated with intense

Corresponding author address: Dr. Leila M. V. Carvalho, Dept. of Geography, University of California, Santa Barbara, Santa Barbara, CA 93106.
E-mail: leila@eri.ucsb.edu

DOI: 10.1175/JCLI-D-11-00335.1

© 2012 American Meteorological Society

latent heat release in the region of heavy precipitation, the large-scale atmospheric circulation is characterized by the upper-level "Bolivian high" and "Nordeste" trough, the "Chaco" surface low pressure, low-level jet east of the Andes (Silva Dias et al. 1983; Gandu and Silva Dias 1998; Lenters and Cook 1999; Marengo et al. 2002), and the South Atlantic convergence zone (SACZ) (Kodama 1992, 1993; Carvalho et al. 2004).

Several studies have shown that the SAMS varies on broad ranges of time scales including diurnal, synoptic, intraseasonal, seasonal, interannual, and decadal (Hartmann and Recker 1986; Robertson and Mechoso 1998; Liebmann et al. 1999; Robertson and Mechoso 2000; Liebmann et al. 2001; Carvalho et al. 2002b; Jones and Carvalho 2002; Grimm 2003; Carvalho et al. 2004; Grimm 2004; Liebmann et al. 2004; Marengo 2004; Grimm and Zilli 2009; Marengo 2009; Carvalho et al. 2011a,b). In addition, precipitation is not uniformly distributed over tropical South America. Complex terrain such as the Andes and the coastal mountain ranges in eastern South America and variations in land use and cover are among the most important causes of spatial variability of precipitation in the SAMS domain (Berbery and Collini 2000; Carvalho et al. 2002a; Durieux et al. 2003; Bookhagen and Strecker 2008).

Although the variability of precipitation in the SAMS has been extensively investigated over the years, one of the main challenges has been the availability of datasets with suitable spatial and temporal resolutions able to resolve the large range of meteorological systems observed within the monsoon. While some stations in South America have precipitation records going back several decades, the density of stations is not sufficient to characterize mesoscale precipitation systems. To overcome this difficulty, considerable effort has been devoted to collecting precipitation records from stations over South America and producing quality-controlled gridded precipitation datasets (Legates and Willmott 1990; Liebmann and Allured 2005; Silva et al. 2007). Although the statistical properties of gridded precipitation may differ from observations at individual stations (Silva et al. 2007), an advantage of the gridded, complete data is that multivariate statistical analyses are more easily performed and teleconnection patterns can be studied in detail. In addition to station data, satellite-derived precipitation estimates (Kummerow et al. 1998, 2000; Huffman et al. 2001; Xie et al. 2003) have been developed over the years and provide important information to further investigate the variability of the SAMS.

Recently, a new generation of reanalysis products has been completed (Saha et al. 2010; Dee et al. 2011; Rienecker et al. 2011). The new reanalyses, which are derived from state-of-the-art data assimilation systems and high-resolution climate models, provide substantial improvements in the spatiotemporal variability of precipitation relative to the first generation of reanalyses (Higgins et al. 2010; Saha et al. 2010; Rienecker et al. 2011; Silva et al. 2011). It is worth noting, however, that precipitation from reanalysis is not an observed variable but is derived from data assimilation and a background forecast model and, therefore, uncertainties resulting from model physics are present (e.g., Bosilovich et al. 2008).

Although the variability of precipitation in the SAMS has been investigated in many previous studies, comparisons among datasets have been only partially addressed (e.g., Silva et al. 2011). The objective of this paper is to evaluate and compare several statistical properties of daily precipitation in three types of datasets: gridded station data, satellite-derived precipitation, and reanalyses. Specifically, this study employs several analyses to determine consistencies and disagreements in the representation of precipitation over the SAMS region. The period 1998–2008 was selected because all datasets used cover that period. In addition, since the datasets are available with different horizontal resolutions, the comparison is performed in two ways: 1) all datasets regridded to a common resolution and 2) datasets with their original resolutions. The paper is organized as follows. Section 2 describes the datasets, and section 3 discusses the methodology. Section 4 compares two precipitation datasets both derived from surface stations but different gridding methods. Section 5 compares the variability of precipitation in the datasets regridded to a common grid resolution, whereas section 6 compares the datasets with their original grid resolution. Section 7 summarizes the main conclusions.

2. Data

The statistical properties of precipitation in the SAMS region are investigated with daily gridded data from multiple sources. Each dataset has a different spatial resolution and the period of available data varies. While some datasets are available for the entire 1979–present period, other datasets have a large number of missing data over several regions in South America (e.g., Amazon). Likewise, some datasets cover only land areas (i.e., those derived from rain gauges), whereas others cover land and ocean regions. To develop a consistent comparison, we chose the period from 1 January 1998 to 31 December 2008 for analysis. Moreover, the domain of analysis is limited to 40°S–15°N, 85°–30°W and grid points over the ocean are masked out in all statistical calculations. The following datasets are used.

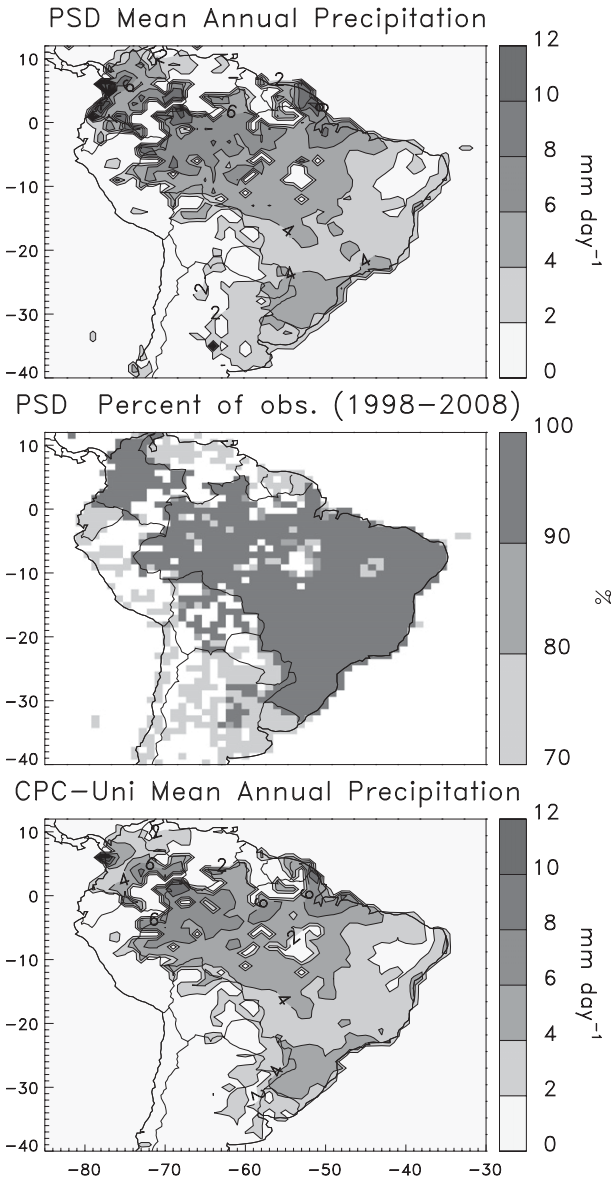

FIG. 1. (top) Mean annual precipitation (mm day^{-1}) from PSD data and (middle) percentage of available daily observations used in PSD during 1 Jan–31 Dec 1998–2008. Grid points with less than 70% of observations are masked out. (bottom) Mean annual precipitation (mm day^{-1}) from CPC-Uni data during 1 Jan–31 Dec 1998–2008. Data grid spacing is 1° lat/lon.

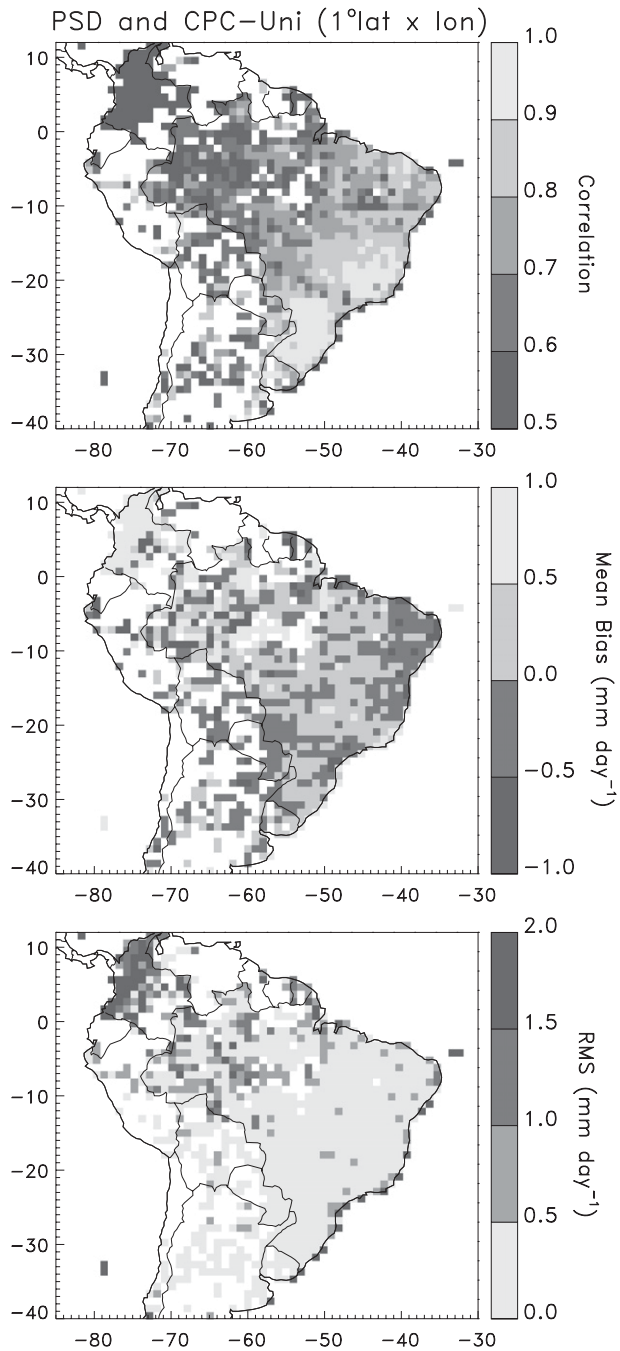

FIG. 2. (top) Correlation between daily precipitation from PSD and CPC-Uni, (middle) mean daily precipitation bias (PSD minus CPC-Uni), and (bottom) root-mean-square difference in daily precipitation. Grid points with fewer than 70% of observations are masked out. The period is 1 Jan–31 Dec 1998–2008. Data grid spacing is 1° lat/lon.

1) Physical Sciences Division (PSD), Earth System Research Laboratory: this dataset is computed from observed precipitation collected at stations throughout South America. A detailed discussion is found in Liebmann and Allured (2005, 2006). The daily gridded precipitation is constructed by averaging all observations available within a specified radius of each grid point. It is important to note that the density of stations varies significantly in space and time as discussed next. Two grid resolutions [1° and 2.5° latitude (lat)/longitude (lon)] are used in this study.
2) Global Precipitation Climatology Project (GPCP): the daily GPCP combines Special Sensor Microwave

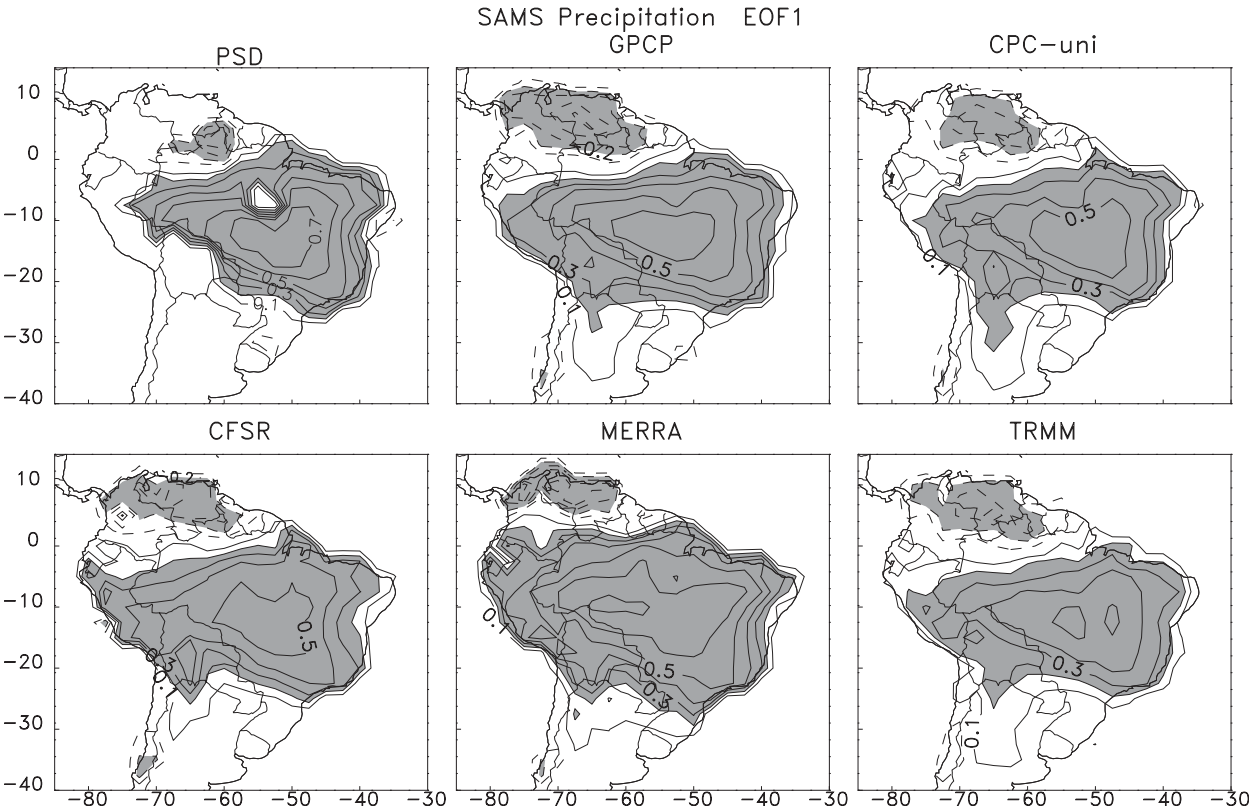

FIG. 3. First EOF patterns described as correlations between the first temporal coefficient (PC1) and precipitation anomalies. Solid (dashed) contours indicate positive (negative) correlations at 0.1 intervals (zero contours omitted). Shading indicates correlations ≥ 0.2 (≤ -0.2) and is significant at 5%. Data grid spacing is 2.5° lat/lon.

Imager (SSM/I), GPCP version 2.1 satellite gauge, geosynchronous-orbit infrared (IR), (geo-IR) brightness temperature T_b histograms (1° × 1° grid in the band 40°N–40°S, 3-hourly), low-orbit IR Geostationary Operational Environmental Satellite (GOES) precipitation index (GPI), Television and Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS), and Atmospheric Infrared Sounder (AIRS) data (Huffman et al. 2001). The GPCP data used in this study have 1° lat/lon grid spacing.

3) Climate Prediction Center (CPC) unified gauge (CPC-uni): the National Oceanic and Atmospheric Administration (NOAA) CPC unified gauge uses an optimal interpolation technique to reproject precipitation reports to a grid (Higgins et al. 2000; Silva et al. 2007; Chen et al. 2008; Silva et al. 2011). This study uses data with 0.5° lat/lon grid spacing. Although the PSD and CPC-uni datasets share some of the same station observations, it is worth noting that the quality control and gridding methods are different. In addition, it is likely that the number and origin of station data in both datasets are different.
4) Climate Forecast System Reanalysis (CFSR): the National Centers for Environmental Prediction (NCEP) have recently concluded the latest reanalysis based on the Climate Forecast System (CFS) model (Saha et al. 2010). The advantages of CFSR relative to previous reanalyses include higher horizontal and vertical resolutions, improvements in data assimilation, and first-guess fields originated from a coupled atmosphere–land–ocean–ice system (Higgins et al. 2010; Saha et al. 2010). Although the CFSR reanalysis is produced at about 35 km in the horizontal, this study uses daily precipitation at 0.5° lat/lon grid spacing [available from the National Center for Atmospheric Research (NCAR)]. It is also important to note that precipitation is not assimilated in the CFSR production but rather is a forecast (first-guess) product.
5) Modern-Era Retrospective Analysis for Research and Applications (MERRA): the new reanalysis produced by the National Aeronautics and Space Administration (NASA) was used in this study. MERRA utilizes an advanced data assimilation system and was generated with the Goddard Earth Observing System

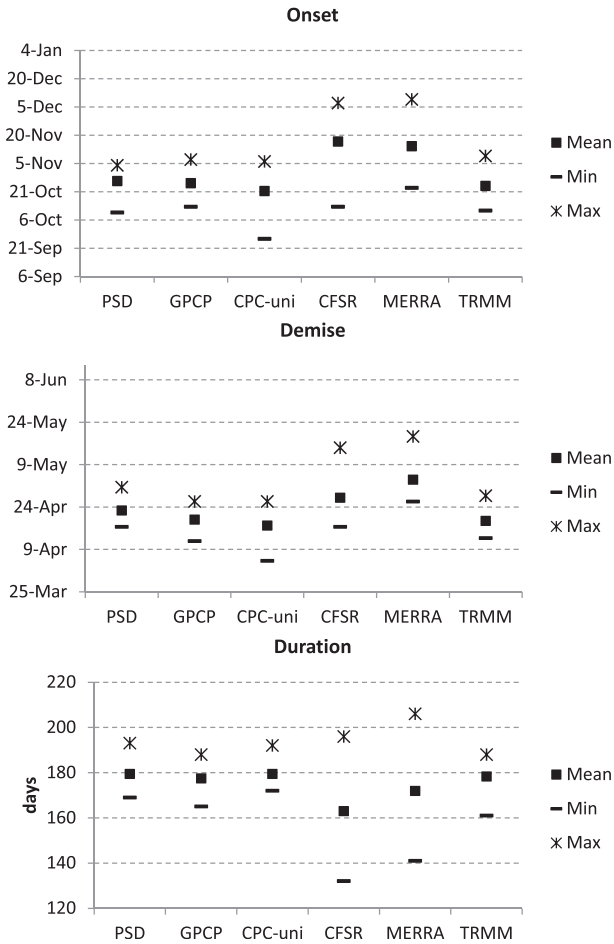

FIG. 4. (top) Mean (squares), minimum (dashes), and maximum (asterisks) dates of SAMS onset. (middle) Mean (squares), minimum (dashes), and maximum (asterisks) dates of SAMS demise. (bottom) Mean (squares), minimum (dashes), and maximum (asterisks) durations of SAMS. Datasets are indicated in the horizontal axis. Data grid spacing is 2.5° lat/lon.

(GEOS) atmospheric model (Rienecker et al. 2011). Daily precipitation at 0.5° latitude/0.3° longitude is used. As in the CFSR, precipitation is a forecast product.

6) Tropical Rainfall Measurement Mission (TRMM 3B42 V6): the TRMM Multisatellite Precipitation Analysis (TMPA) (Huffman et al. 2007) provides rainfall estimates at 0.25° × 0.25° spatial resolution and 3-h intervals. The gridded rainfall algorithm uses an optimal combination of TRMM 2B31 and TRMM 2A12 data products, SSM/I, Advanced Microwave Scanning Radiometer (AMSR), and Advanced Microwave Sounding Units (AMSI) (Kummerow et al. 1998; Kummerow et al. 2000). These data were processed to generate daily precipitation. Previous studies have used TRMM 3B42 V6 data to identify rainfall-extreme events and their impact on river discharge in the southwestern part of the Amazonian catchment in Bolivia and Brazil (Bookhagen and Strecker 2010). To achieve high spatial resolution and identify orographic rainfall processing, TRMM 2B31 data with a spatial resolution of about 5 km × 5 km and approximately daily snapshots have been used to relate topographic characteristics and orographic rainfall along the eastern Andes (Bookhagen and Strecker 2008). Additional details of data processing are described in Bookhagen and Burbank (2011).

TABLE 1. Cross correlations among seasonal amplitudes of SAMS derived from daily PC1. Correlations ≥ 0.63 (≤−0.63) are significant at 5%. Data grid spacing is 2.5° lat/lon.

	PSD	GPCP	CPC-uni	CFSR	MERRA	TRMM
PSD	1.00	—	—	—	—	—
GPCP	0.85	1.00	—	—	—	—
CPC-uni	0.46	0.64	1.00	—	—	—
CFSR	0.44	0.41	−0.09	1.00	—	—
MERRA	0.54	0.50	−0.06	0.82	—	—
TRMM	0.79	0.95	0.76	0.24	0.41	1.00

3. Methodology

Comparisons of precipitation variability are performed with several statistical methods applied to the six datasets regridded to a common horizontal resolution as well as to their original grid spacings. A grid of 2.5° lat/lon spacing is selected for the same resolution comparison because the number of missing observations in the 1° lat/lon PSD dataset is very high in some locations over South America (see section 4). The 2.5° lat/lon common regrid is obtained by regridding the GPCP, CPC-uni, CFSR, MERRA, and TRMM data to the PSD grid according to the following: $P_R(i,j) = [P(i,j) + P(i-1,j) + P(i+1,j) + P(i,j-1) + P(i,j+1)]/5$, where $P_R(i,j)$ is the regridded value, and the five terms on the right-hand side are precipitation values from the data being transformed. $P_R(i,j)$ is centered on the same coordinates of the PSD dataset. Note that this regridding method is

TABLE 2. Anomaly cross correlations among daily PC1 derived from precipitation datasets. Correlations ≥ 0.20 (≤−0.20) are significant at 5%. Data grid spacing is 2.5° lat/lon.

	PSD	GPCP	CPC-uni	CFSR	MERRA	TRMM
PSD	1.00	—	—	—	—	—
GPCP	0.67	1.00	—	—	—	—
CPC-uni	0.88	0.66	1.00	—	—	—
CFSR	0.68	0.6	0.62	1.00	—	—
MERRA	0.66	0.62	0.61	0.78	1.00	—
TRMM	0.75	0.9	0.74	0.63	0.64	1.00

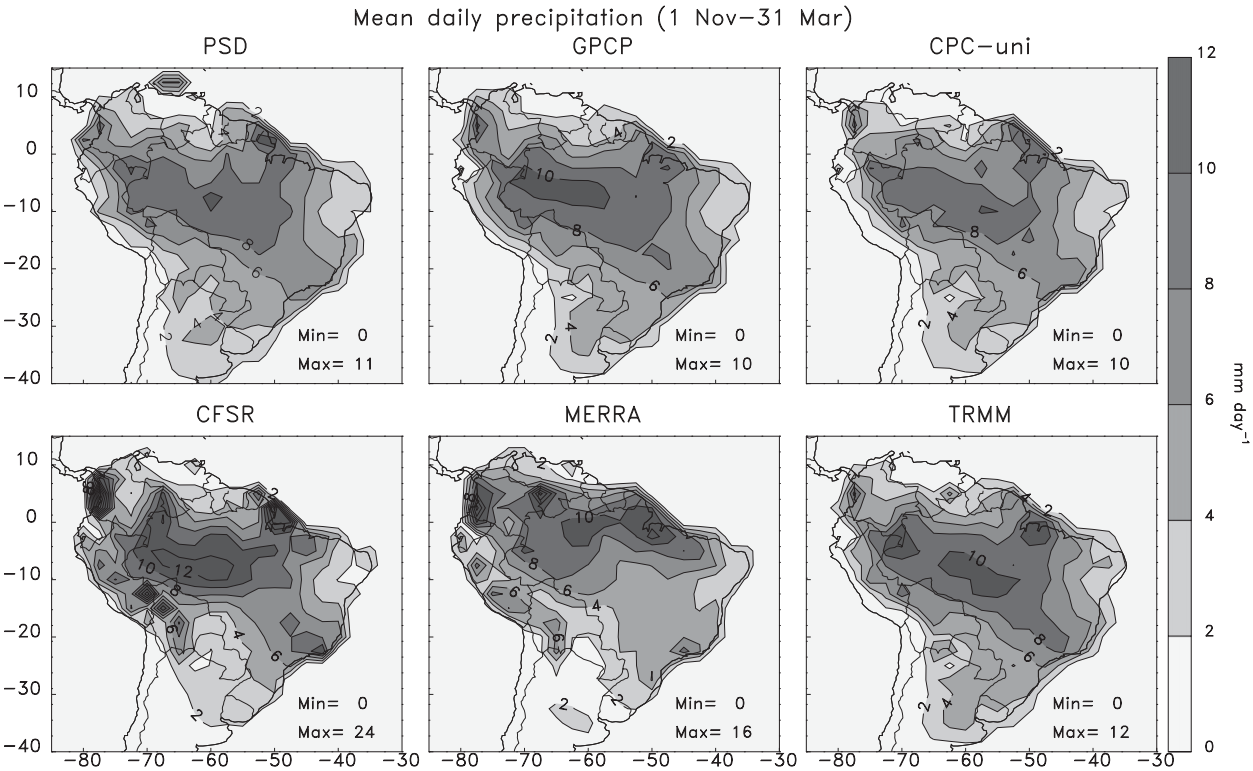

FIG. 5. Mean daily precipitation during 1 Nov–31 Mar 1979–2010. Contour interval and shading is 2 mm day^{-1}. Datasets and minimum/maximum values are indicated in each panel. Data grid spacing is 2.5° lat/lon.

different than the procedure used in the PSD data. The PSD data represents gridded precipitation as averages of precipitation over ''circles'' with radius of about 1.8°; see Liebmann and Allured (2005, 2006) for details. Comparisons between both regridding methods indicated insignificant differences in the statistics described in section 4.

The annual evolution of SAMS is examined to determine consistencies and disagreements among the datasets. The large-scale features of interest are as follows: the dominant spatial precipitation pattern, dates of onset and demise, duration, and amplitude of the monsoon. These characteristics are determined with empirical orthogonal functional (EOF) analysis (Wilks 2006) applied to the daily precipitation (only land grid points) from each dataset separately. Before computation of the EOFs, the time series of precipitation in each grid point are scaled by the square root of the cosine of the latitude and the long-term mean removed (1 January–31 December 1998–2008). The EOFs are calculated from correlation matrices. The first mode (EOF1) and associated temporal coefficient (PC1) explain the largest fraction of the total variance of precipitation over land and are used to describe the annual evolution of SAMS.

To determine dates of onset, demise, and duration of SAMS, the daily PC1 is smoothed with 10 passes of a 15-day moving average. This smoothing procedure is obtained empirically and is used to decrease the influence of high-frequency variations during the transitional phases of SAMS. The large-scale onset of SAMS is defined as the date when the smoothed PC1 changes from negative to positive values. This implies that positive precipitation anomalies during that time become dominant over the SAMS domain. Likewise, the demise of SAMS is defined as the date when the smoothed PC1 changes from positive to negative values. The duration of the monsoon is defined as the period between onset and demise dates. The seasonal amplitude of the monsoon is defined as the integral of positive unsmoothed PC1 values from onset to demise. Therefore, the seasonal amplitude index represents the sum of positive precipitation anomalies and minimizes the effect of ''break'' periods in the monsoon especially near the onset and demise. Active/break periods in SAMS are particularly frequent on intraseasonal time scales (Jones and Carvalho 2002).

The distribution of precipitation variance is examined with a power spectrum of the daily PC1 during 1 October–30 April 1998–2008. The following methodology is used: 1) the mean, linear trend, and annual cycle are removed from the PC1 time series during each season; 2) the resulting time series is tapered with a split cosine bell function (5% at each end); 3) fast Fourier transform

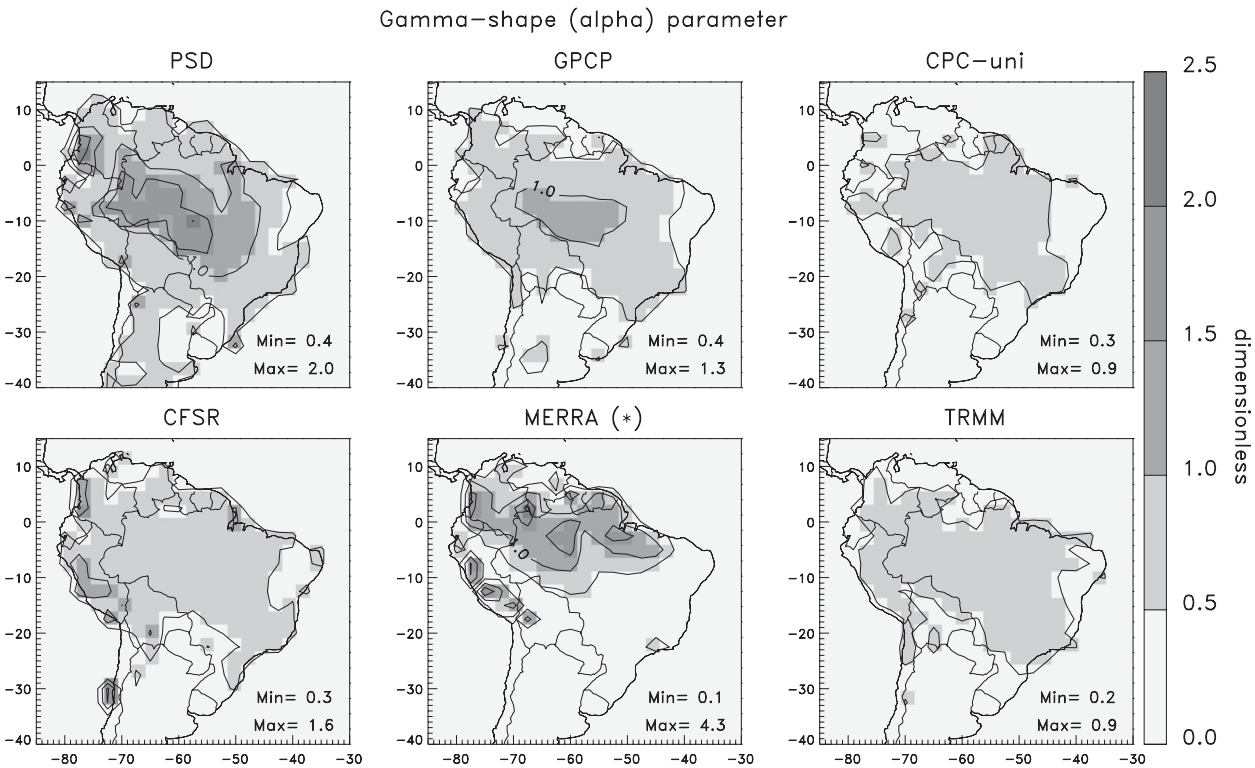

FIG. 6. Shape parameter of gamma probability distribution function (dimensionless). Datasets and minimum/maximum values are indicated in each panel. Note that to fit α from MERRA in the same scale the shape parameter has been divided by 2. Data grid spacing is 2.5° lat/lon. Time period is 1 Nov–31 Mar 1998–2008.

(FFT) is used to obtain raw spectral estimates for each season; 4) the raw spectral estimates are smoothed with a running average of length $L = 3$ raw spectral estimates; 5) the spectra computed for each season are normalized by the seasonal variance and averaged to obtain a 10-yr ensemble mean; and 6) the degrees of freedom are estimated initially as 60 [(2 for every raw spectral estimate) × (3 for smoothing the raw spectrum) × (10 for ensemble average)]. The actual degrees of freedom are reduced to 52.38 because of the tapering of the time series [see Madden and Julian (1971) for further details]. The red-noise background spectrum and 95% significance level are computed following the methodology of Mitchell (1966).

Statistical properties of precipitation are further studied in the following way. Because of the skewness in precipitation, parameters of frequency distributions are estimated and compared among the six datasets. Gamma frequency distributions are fitted to daily precipitation following the maximum likelihood approach (Wilks 2006). The fitting is done on the entire sample (1 November–31 March 1998–2008) of daily precipitation values P_i (zero values excluded). Next, the sample statistic D is computed as follows:

$$D = \ln(\overline{P}) - \frac{1}{N}\sum_{i=1}^{N}\ln(P_i). \tag{1}$$

Then the shape α and scale β parameters are estimated by the polynomial approximations:

$$\alpha = \frac{0.500\,087\,6 + 0.164\,885\,2D - 0.054\,427\,4D^2}{D}, \quad 0 \le D \le 0.5772, \tag{2}$$

$$\alpha = \frac{8.898\,919 + 9.059\,950D - 0.977\,537\,3D^2}{17.797\,28D + 11.968\,477D^2 + D^3}, \quad 0.5772 < D \le 17.0, \quad \text{and} \tag{3}$$

$$\beta = \frac{\overline{P_i}}{\alpha}. \tag{4}$$

The gamma distribution is thus expressed as

$$f(P) = \frac{(P/\beta)^{\alpha-1}\exp(-P/\beta)}{\beta\Gamma(\alpha)}, \tag{5}$$

where P, α, and $\beta > 0$, and $\Gamma(\alpha)$ is the gamma function. To compute the cumulative distribution function, we

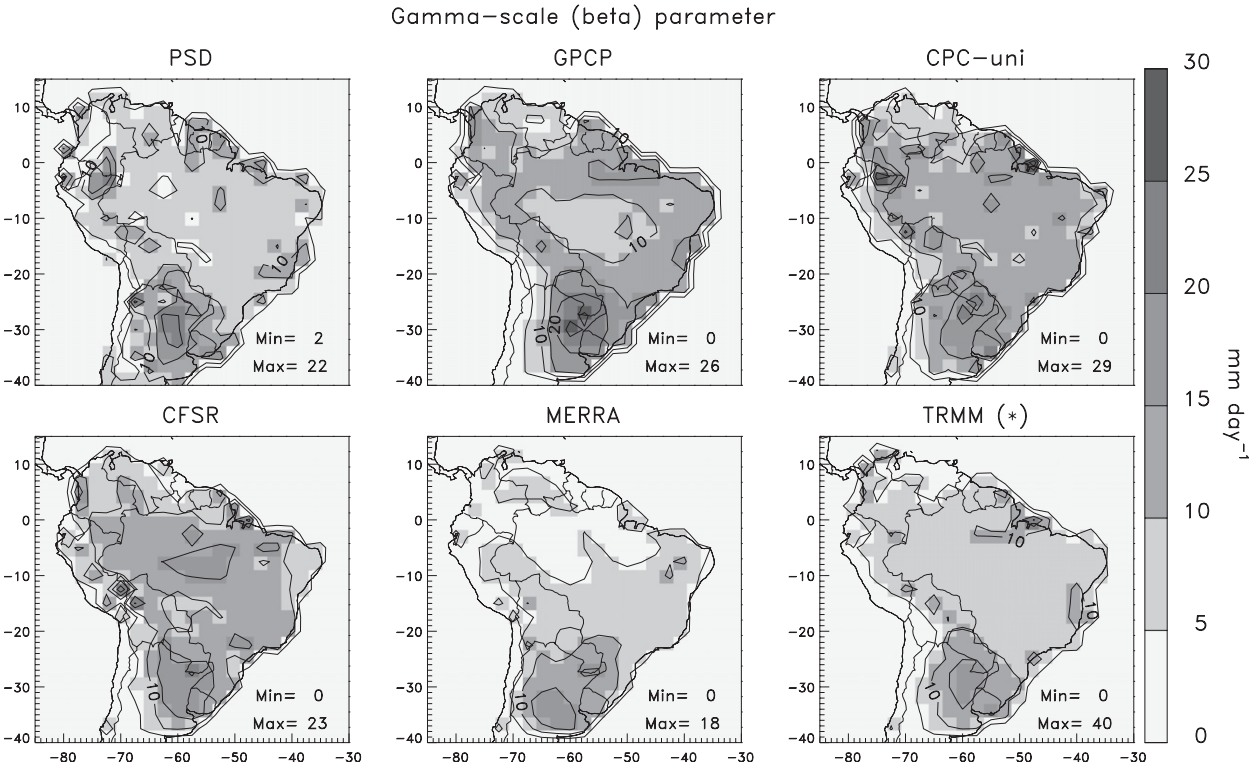

FIG. 7. Scale parameter of gamma probability distribution function (mm day^{-1}). Datasets and minimum/maximum values are indicated in each panel. Note that to fit β from TRMM in the same scale the scale parameter has been divided by 2. Data grid spacing is 2.5° lat/lon. Time period is 1 Nov–31 Mar 1998–2008.

first scale the precipitation by $\xi = P/\beta$ and numerically integrate the incomplete gamma function $F(\alpha, \xi)$ [see Wilks (2006) for details].

Small values of the shape parameter α indicate that the distribution is strongly skewed to small precipitation values (i.e., skewed to the left), whereas large values of α indicate that the distribution tends to approximate the form of Gaussian distributions. The scale parameter β represents the "stretch" or "squeeze" in the gamma density function to the right or left (Wilks 2006). Examples of α and β estimated over several precipitation regimes are discussed in Jones et al. (2004) (see their Fig. 3).

4. Comparison between PSD and CPC-uni with 1° lat/lon grid spacing

It is instructive to begin by first considering the issue of precipitation sampling from surface stations in South America. While the total number of stations shows a positive trend over the past three decades, this aspect is in fact more complicated. Stations can be frequently deactivated after a few years of operation, while new stations are brought online. The variable temporal record of precipitation is clearly reflected in the spatial density of surface stations in South America (Liebmann and Allured 2005).

A comparison between PSD and CPC-uni is made since gridded precipitation in these datasets is derived exclusively from surface stations. The comparison is performed at 1° lat/lon grid spacing, such that the CPC-uni data are transformed to the PSD grid. It is important to note that PSD and CPC-uni have distinct gridding methods. The CPC-uni is based on optimal interpolation method and, therefore, the quality of interpolated values depends on the spatial density of stations (Chen et al. 2008). The PSD is based on averaging precipitation from stations within specified distances from grid points; if no stations are present, missing values are assigned.

Figures 1 (top, middle) respectively show the mean annual precipitation and percentage of available observations in the PSD data. Grid points with fewer than 70% of observations are masked. The sampling clearly shows some geographical boundaries. For instance, adequate number of observations (≥90%) is seen over Brazil, although it is still quite deficient over the Amazon. In addition, surface stations are very sparse over the Andes and northern parts of South America. Sampling becomes even more critical in previous decades, when

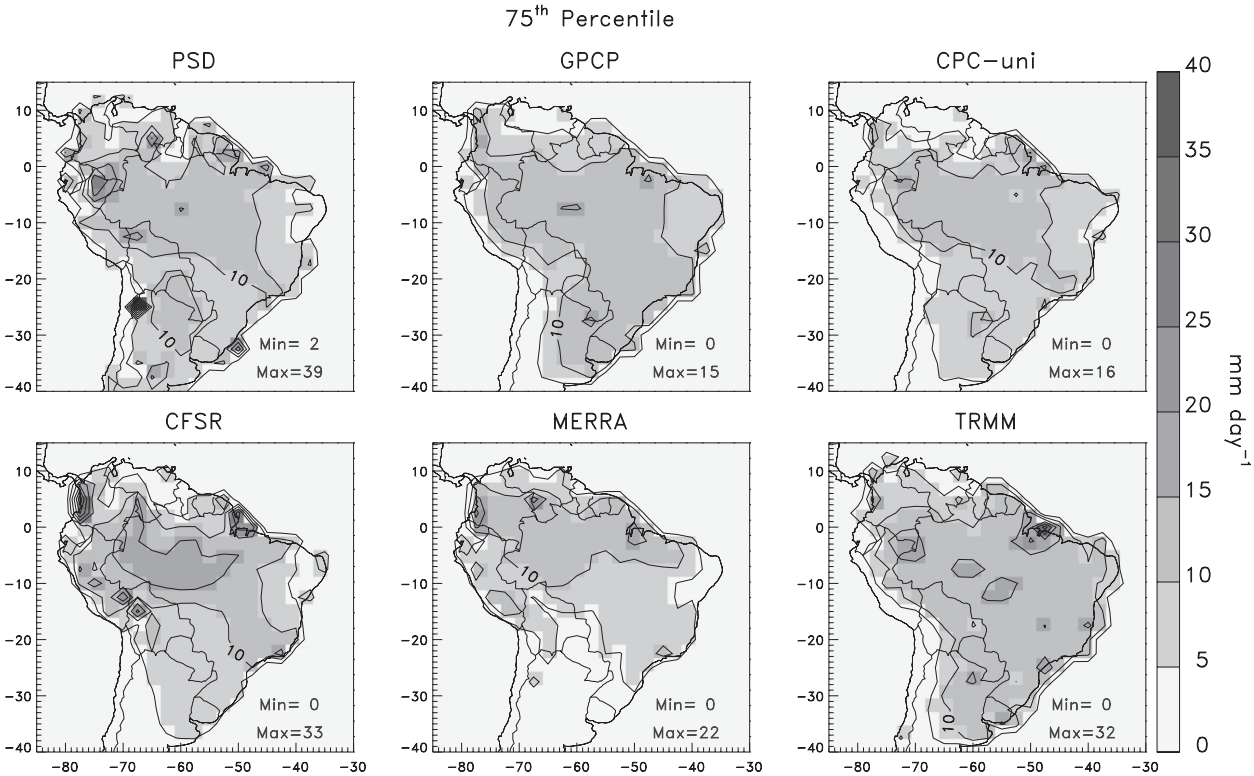

FIG. 8. 75th percentile of daily precipitation (mm day^{-1}). Datasets and minimum/maximum values are indicated in each panel. Data grid spacing is 2.5° lat/lon. Time period is 1 Nov–31 Mar 1998–2008.

the density of stations over the SAMS region was low. For comparison, Fig. 1 (bottom) shows the mean annual precipitation from CPC-uni data regridded to the same 1° lat/lon of the PSD data. While the spatial patterns of mean precipitation from both datasets are comparable, differences in intensity are noticeable especially over the northwestern parts of South America.

A direct comparison is made by computing correlations between precipitation from PSD and CPC-uni in each 1° lat/lon (Fig. 2, top). These correlations are performed on the raw time series (1 January–31 December 1998–2008) and, therefore, include subseasonal, seasonal, and interannual variations. Because of that, one would expect a high degree of agreement between the two datasets over the SAMS since the amplitude of the annual cycle is large. Thus, correlations are high and spatially coherent (above 0.8) mostly over eastern Brazil, where the density of stations is high (Liebmann and Allured 2005). In other locations over South America, correlations are less spatially coherent and significantly low (correlations $\geq$ 0.2 are significant at 5%). For instance, correlations over the Amazon are on the order of 0.5–0.6.

The mean daily precipitation bias (Fig. 2, middle) is in the range of $\pm$1.0 mm day^{-1}, which is a reasonable amount. However, the mean bias is not spatially coherent and even changes sign among neighboring grid points. In contrast, the root-mean-square (rms) difference (Fig. 2, bottom) is quite uniform (0.5 mm day^{-1}) over most of Brazil and other countries. The relatively large rms difference in coastal grid points in Brazil results from the regridding process since the land masks between the two datasets do not match exactly. More importantly, rms differences are consistently high ($\geq$1.0 mm day^{-1}) over a large portion of Colombia, which is surprising given the high number of available observations (Fig. 1, bottom). The results above highlight important differences in the gridding methods in the PSD and CPC-uni datasets. Additionally, since the number of missing data is substantially high in the PSD at 1° lat/lon, further comparisons among the datasets are performed with a common grid of 2.5° lat/lon. This is important in the EOF analysis because missing data vary widely among grid points and the calculation becomes extremely difficult with small sample sizes.

5. Comparisons among datasets with 2.5° lat/lon grid spacing

In this section, the six datasets are compared at the same 2.5° lat/lon grid (gridded according to the description

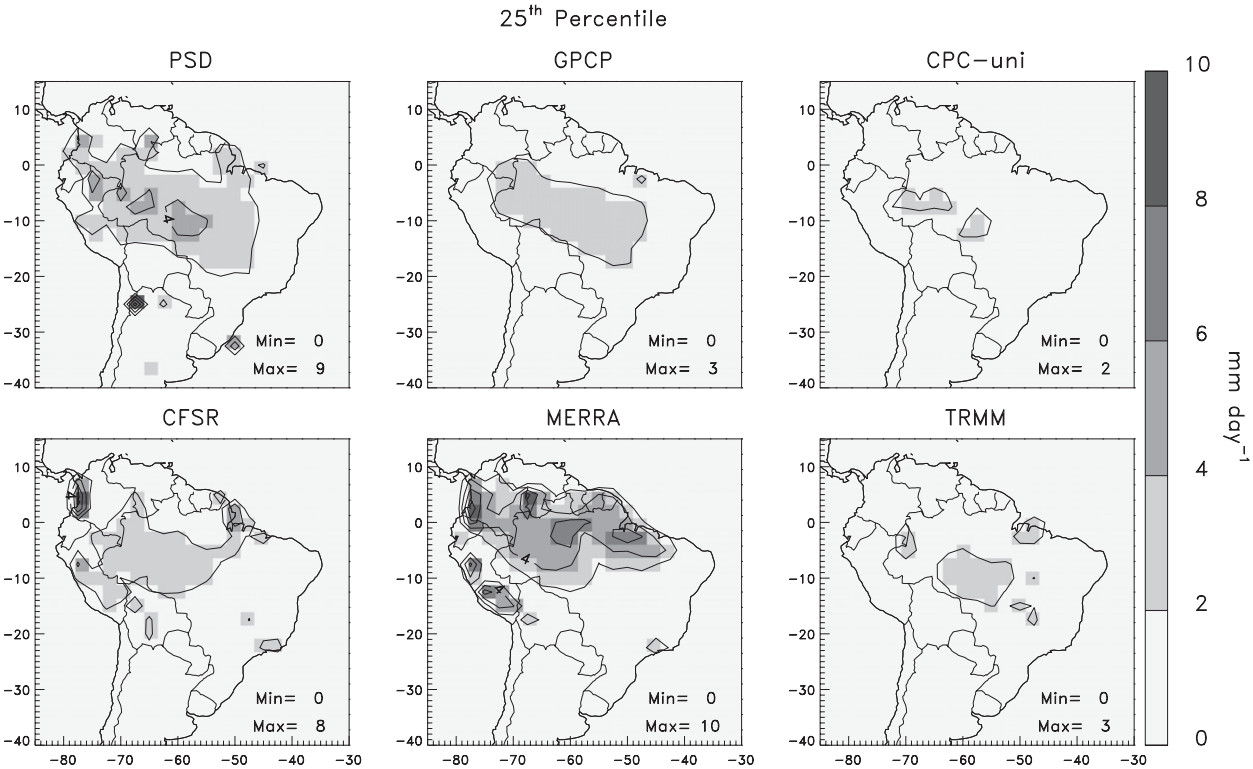

FIG. 9. 25th percentile of daily precipitation (mm day^{-1}). Datasets and minimum/maximum values are indicated in each panel. Data grid spacing is 2.5° lat/lon. Time period is 1 Nov–31 Mar 1998–2008.

in section 3). These results can be evaluated with the comparison at the original grid spacing resolution in section 6.

EOF analysis (section 3) is used to characterize the large-scale features of SAMS as represented in the different datasets. Figure 3 shows the spatial patterns of EOF1 derived from each dataset and expressed as correlations between PC1 and precipitation anomalies. Positive correlations are interpreted as positive precipitation anomalies and indicative of active SAMS. In general, all datasets show similar features such as positive precipitation anomalies over central South America and negative anomalies over the northern parts of the continent. The region of negative anomalies over northern South America is substantially smaller in the PSD because of missing data (the ''bull's eye'' near 10°S, 60°W is a grid point with missing data). The magnitude of positive correlations varies slightly and is highest for PSD. Also, the largest positive correlation in MERRA is slightest to the west relative to the other datasets. As discussed in section 6, however, significant spatial differences are seen in EOF1 patterns from the datasets at their original resolutions.

The percentages of explained variance by EOF1 are the following: 20.5% (PSD), 11.6% (GPCP), 8.4% (CPC-uni), 10% (CFSR), 17.9% (MERRA), and 6.9% (TRMM). EOF1 captures the largest fraction of the total variance, which includes subseasonal, seasonal, and interannual variations, since the EOF analysis is performed removing only the long-term mean. Main differences in explained variance are associated with how much each PC1 represents the distribution of subseasonal, seasonal, and interannual variations. These percentages are comparable to the percentages obtained with the datasets at their original resolutions (section 6), which suggests that spatial resolution of the datasets is not the main issue but rather how each dataset represents temporal variations.

The seasonal variation of SAMS is represented by the dates of onset, demise, and duration. The mean onset date (Fig. 4, top) is highly coherent among PSD, GPCP, CPC-uni, and TRMM (~21 October) including the ranges of minimum and maximum onset dates. In contrast, the mean onset dates in the CFSR and MERRA reanalyses occur in the second week of November and the ranges of minimum and maximum onset dates are larger than in the other datasets. Although there is more variability in the dates of mean demise (Fig. 4, middle), the agreement among PSD, GPCP, CPC-uni, and TRMM is relatively good. The mean demise dates in the

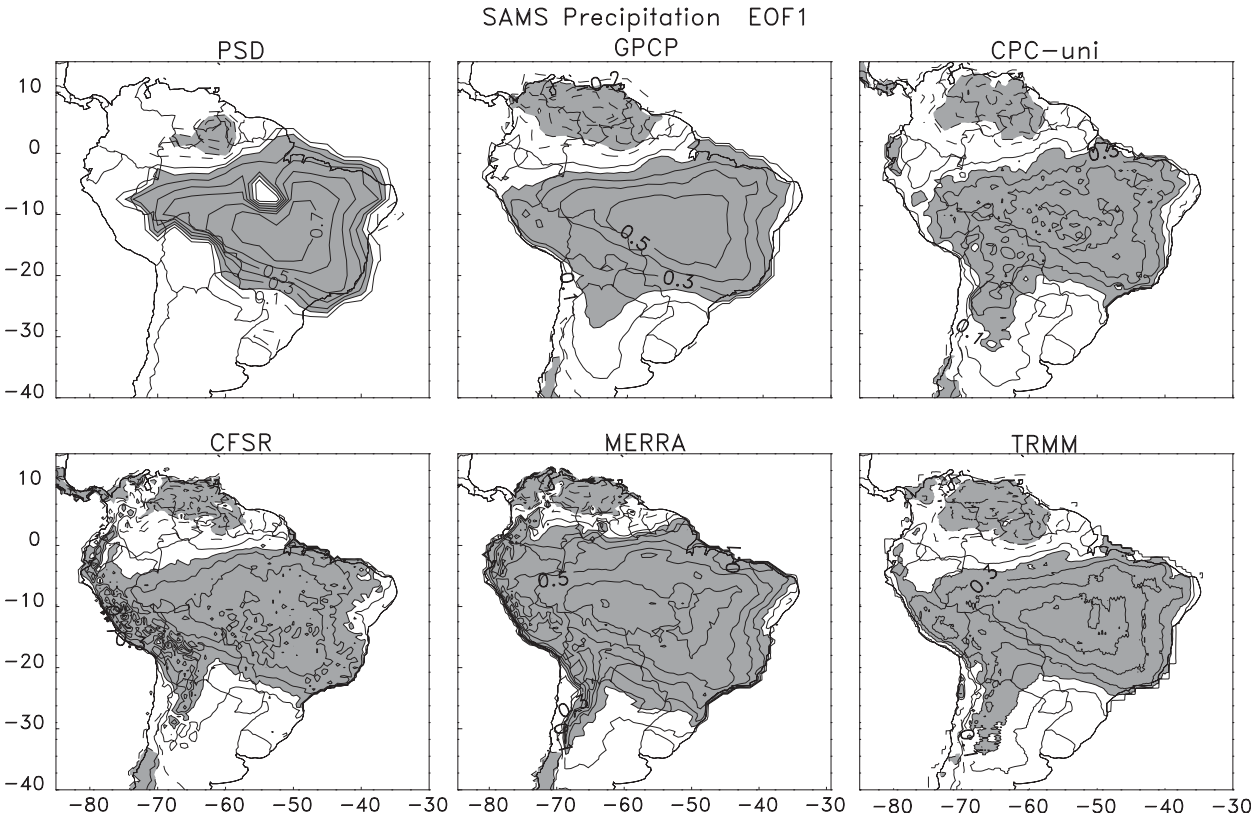

FIG. 10. First EOF patterns described as correlations between the first temporal coefficient (PC1) and precipitation anomalies. Solid (dashed) contours indicate positive (negative) correlations at 0.1 intervals (zero contours omitted). Shading indicates correlations ≥ 0.2 (≤ -0.2) and is significant at 5%. Datasets have different grid spacings.

reanalyses occur later than in the other datasets and CFSR is more consistent with the other data than MERRA. As a consequence, the mean duration of SAMS (~180 days) agrees reasonably well among PSD, GPCP, CPC-uni, and TRMM data and is shorter and more variable in the CFSR and MERRA (Fig. 4, bottom). The characteristics shown in Fig. 4 are practically identical to similar results obtained from the datasets at their original resolutions (section 6).

The seasonal amplitude (section 3) is another important characteristic of SAMS and is compared by computing cross correlations among the seasonal amplitudes obtained from each dataset (Table 1). Significant correlations are seen among PSD–GPCP (0.85), PSD–TRMM (0.79), GPCP–TRMM (0.95), CPC-uni–TRMM (0.76), and CFSR–MERRA (0.82). Surprisingly, the seasonal amplitude correlation between PSD–CPC-uni is not statistically significant, even though both datasets are derived from station data. Additionally, the seasonal amplitudes derived from CFSR and MERRA are correlated (although not statistically significant) with PSD and GPCP but not CPC-uni. The correlations in Table 1 do not differ substantially from similar correlations calculated from seasonal amplitudes derived from each dataset at their original resolutions (section 6). Therefore, because of space limitations, the temporal variability of seasonal amplitudes (see Fig. 12) obtained from each dataset with 2.5° lat/lon is not shown.

A more detailed comparison of PC1 is carried out by removing the mean seasonal cycle in PC1 from each dataset. Next, cross correlations are calculated among the daily PC1 anomalies during 1 November–31 March 1998–2008 (Table 2). The best agreements are found among GPCP–TRMM (0.90), PSD–CPC-uni (0.88), CFSR–MERRA (0.78), PSD–TRMM (0.75), and CPC-uni–TRMM (0.74), whereas the smallest correlations are among MERRA–CPC-uni (0.61) and CFSR–GPCP (0.6). This result is important because it shows that anomalies in PC1, which are largely related to intraseasonal variations, can have different degrees of representation in the six datasets. Spectra of daily PC1 anomalies from datasets with 2.5° lat/lon grid are virtually identical to spectra of daily PC1 anomalies from datasets with their original resolution (section 6) and are not shown (see Fig. 13 instead).

The results above indicate that, although several aspects of the large-scale characteristics of the SAMS are

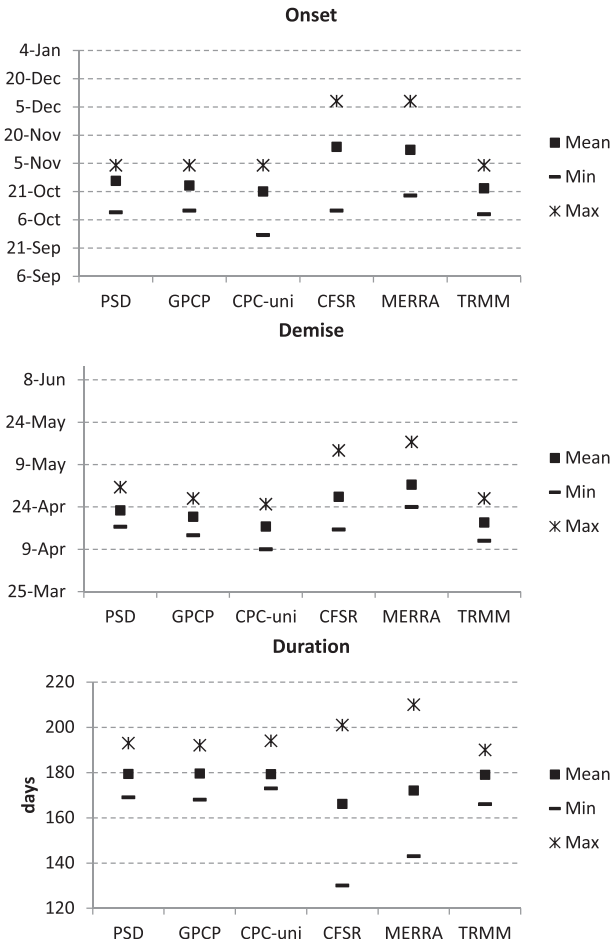

FIG. 11. (top) Mean (squares), minimum (dashes), and maximum (asterisks) dates of SAMS onset. (middle) Mean (squares), minimum (dashes), and maximum (asterisks) dates of SAMS demise. (bottom) Mean (squares), minimum (dashes), and maximum (asterisks) durations of SAMS. Datasets are indicated in the horizontal axis. Datasets have different grid spacings.

well represented in most of the datasets, some important inconsistencies are also found. To better characterize differences among the datasets, the following analyses are performed on the raw precipitation (i.e., no EOF filtering). The mean precipitation during 1 November–31 March 1998–2008 is shown in Fig. 5. The maximum is over central Amazon and the pattern extends to southeastern Brazil in all datasets, except in MERRA where the maximum is clearly displaced to the north relative to the other datasets; a pattern that is more typical of early spring rather than austral summer (Kousky 1988; Horel et al. 1989). The range of mean precipitation is nearly the same in PSD, GPCP, CPC-uni, and TRMM and considerably high in CFSR and MERRA. In addition, the mean precipitation with 2.5° lat/lon shows substantial spatial differences from the same field obtained from datasets with their original resolutions especially over the Andes (section 6).

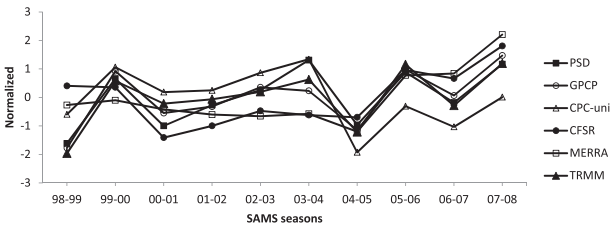

FIG. 12. Normalized seasonal amplitudes of SAMS as derived from different datasets. Datasets are indicated on the side and have different grid spacings. Normalization of time series is done by subtracting the mean and dividing by the standard deviation during 1998–2008.

TABLE 3. Cross correlations among seasonal amplitudes of SAMS derived from daily PC1. Correlations ≥ 0.63 (≤−0.63) are significant at 5%. Datasets have different grid spacings.

	PSD	GPCP	CPC-uni	CFSR	MERRA	TRMM
PSD	1.00	—	—	—	—	—
GPCP	0.91	1.00	—	—	—	—
CPC-uni	0.58	0.48	1.00	—	—	—
CFSR	0.43	0.54	−0.14	1.00	—	—
MERRA	0.48	0.65	−0.03	0.87	—	—
TRMM	0.92	0.96	0.56	0.36	0.54	1.00

The distribution of precipitation can be evaluated with the parameters of the Gamma frequency distribution. The shape parameter (Fig. 6) is in the range 0.5–1.0 over most of South America (except the Andes) and indicates the skewness of precipitation to small values. This feature is consistent in the PSD, GPCP, CPC-uni, CFSR, and TRMM data, whereas PSD and GPCP data show values larger than 1.0 over the core of the monsoon. It is also worth noting that, while PSD and CPC-uni are based on station data, the precipitation distribution over the core of the monsoon can be very different in these datasets. Another obvious feature is that α obtained from MERRA has a spatial pattern and magnitudes substantially different than any other dataset.

The scale parameter β (Fig. 7) from all datasets consistently shows large values (≥10) over southern South America with maximum centered over northern

TABLE 4. Anomaly cross correlations among daily PC1 derived from precipitation datasets. Correlations ≥ 0.20 (≤−0.20) are significant at 5%. Datasets have different grid spacings.

	PSD	GPCP	CPC-uni	CFSR	MERRA	TRMM
PSD	1.00	—	—	—	—	—
GPCP	0.67	1.00	—	—	—	—
CPC-uni	0.87	0.67	1.00	—	—	—
CFSR	0.66	0.61	0.6	1.00	—	—
MERRA	0.62	0.62	0.61	0.8	1.00	—
TRMM	0.77	0.93	0.76	0.65	0.66	1.00

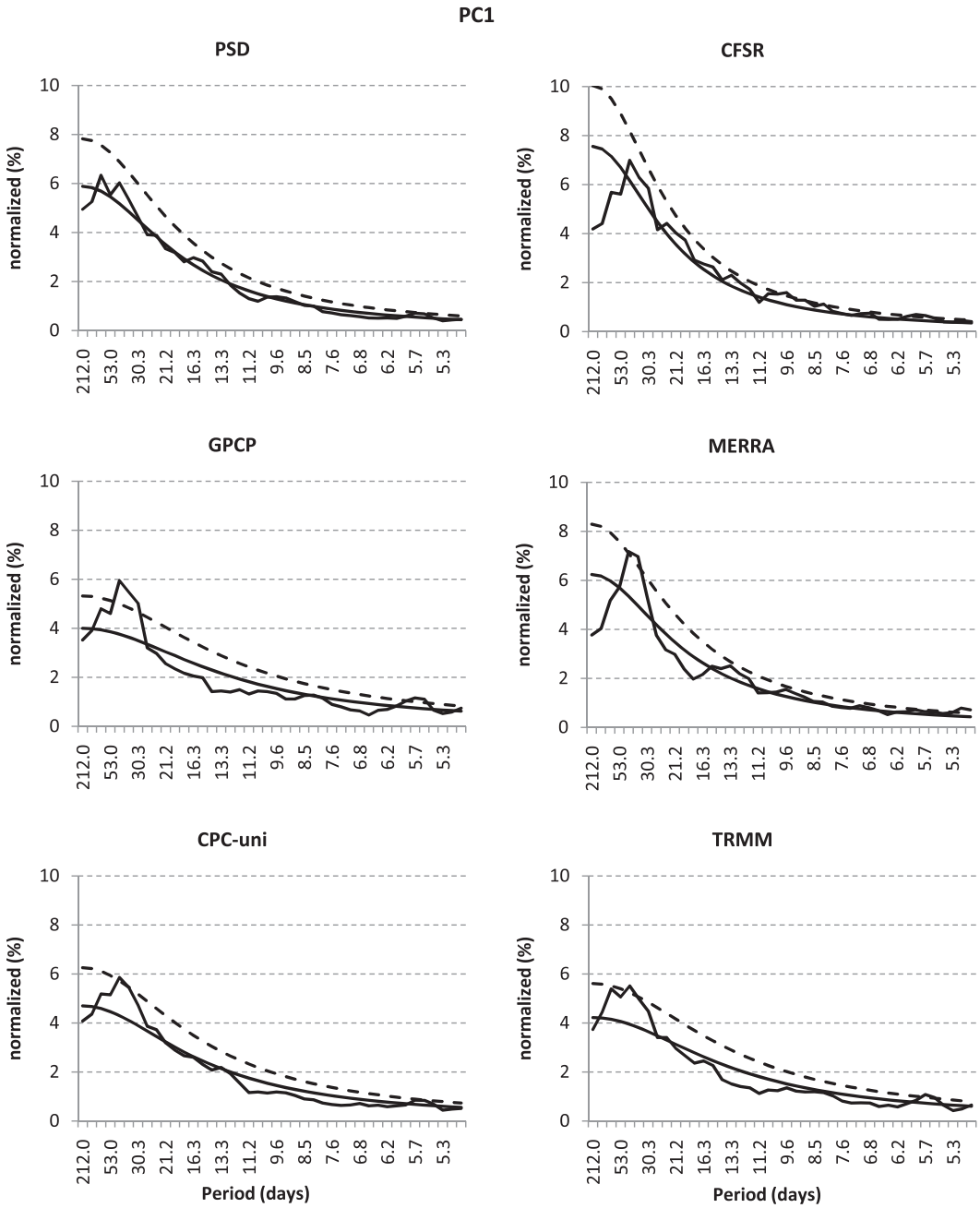

FIG. 13. Power spectrum of PC1. Smoothed (dashed) lines indicate the background red-noise spectrum (95% confidence level). Datasets have different grid spacings. Spectra were computed from daily PC1 during 1 Oct–30 Apr 1998–2008.

Argentina. The spatial pattern of β shows some similarities between PSD and CPC-uni over tropical South America, although the magnitudes are different. Different spatial patterns and magnitudes are seen in GPCP and CFSR over the core of the monsoon. While the β spatial pattern from TRMM is consistent with PSD and CPC-uni over the SAMS, the magnitudes are excessively large. It is interesting to note that β from MERRA is significantly different than any other dataset over the northern Amazon.

Further insight about the distribution of precipitation is noted on the 75th percentile (P75) (Fig. 8). It is encouraging that the northwest–southeast orientation and the range of magnitudes of P75 from PSD and CPC-uni

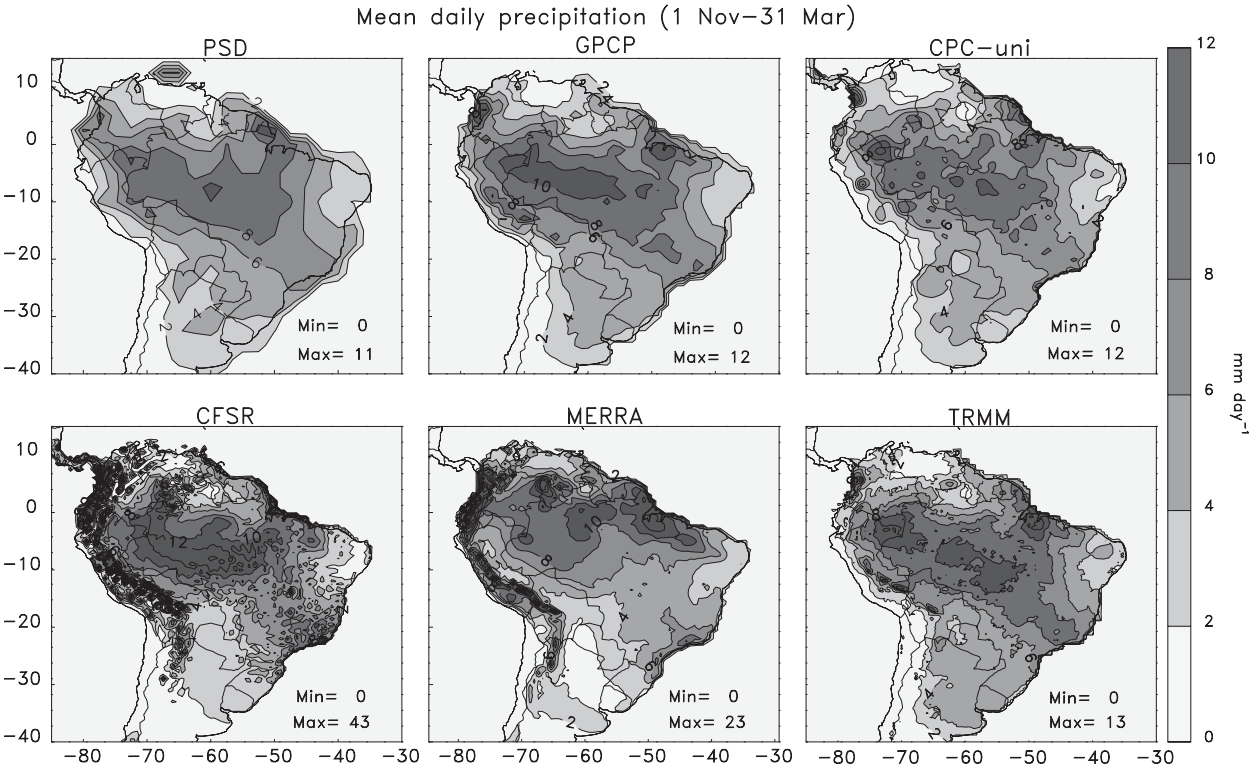

FIG. 14. Mean daily precipitation during 1 Nov–31 Mar 1979–2010. Contour interval and shading are 2 mm day^{-1}. Datasets and minimum/maximum values are indicated in each panel. Datasets have different grid spacings.

agree over the SAMS. TRMM and GPCP P75 show similar range of magnitudes over SAMS, although TRMM indicates large values where the Amazon River meets the Atlantic Ocean. P75 from CFSR also shows a northwest–southeast orientation as in PSD and CPC-uni, while large values are found over the Amazon. In contrast, MERRA shows P75 $\geq$ 10 mm day^{-1} displaced over northern Amazon clearly indicative of difficulties in properly representing the precipitation pattern over the core of the monsoon.

The 25th percentile (P25) (Fig. 9) is less than 2 mm day^{-1} over large portions of South America in all datasets. Over the core of the monsoon, PSD, GPCP, CPC-uni, CFSR, and TRMM indicate P25 in the range of 2.0–4.0 mm day^{-1}, although the spatial extent varies among these datasets. As before, P25 derived from MERRA shows values larger than 2.0 mm day^{-1} displaced over the northern Amazon.

6. Comparisons among datasets with original resolutions

The same type of analysis is performed for the datasets with their original resolution: PSD (2.5° lat/lon), GPCP (1.0° lat/lon), CPC-uni (0.5° lat/lon), CFSR (0.5° lat/lon), MERRA (0.5° lat/0.3° lon), and TRMM (0.25° lat/lon). Figure 10 shows the spatial patterns of EOF1. As expected, the higher the spatial resolution, the more details are represented in the horizontal, although some features are obviously suspicious, for example, over the central Andes. Some of these can be explained by a lack of precipitation gauges or abnormal land cover conditions that influence satellite-derived precipitation values. All datasets consistently indicate a region of positive precipitation anomalies over the core of the monsoon region and negative anomalies over northern South America. Moreover, although positive precipitation anomalies over the central Andes are consistently represented in the GPCP, CPC-uni, CFSR, MERRA, and TRMM, the large gradients in correlations shown in CFSR and MERRA suggest that the reanalysis products overestimate precipitation in that region.

The percentages of total variance explained by EOF1 are the following: 20.5% (PSD), 9.7% (GPCP), 7.6% (CPC-uni), 8.9% (CFSR), 17.5% (MERRA), and 7.3% (TRMM), respectively. The distribution of percentages are comparable to the percentages obtained with all datasets with the same resolution (section 5), which suggests that the spatial resolution is not the main factor

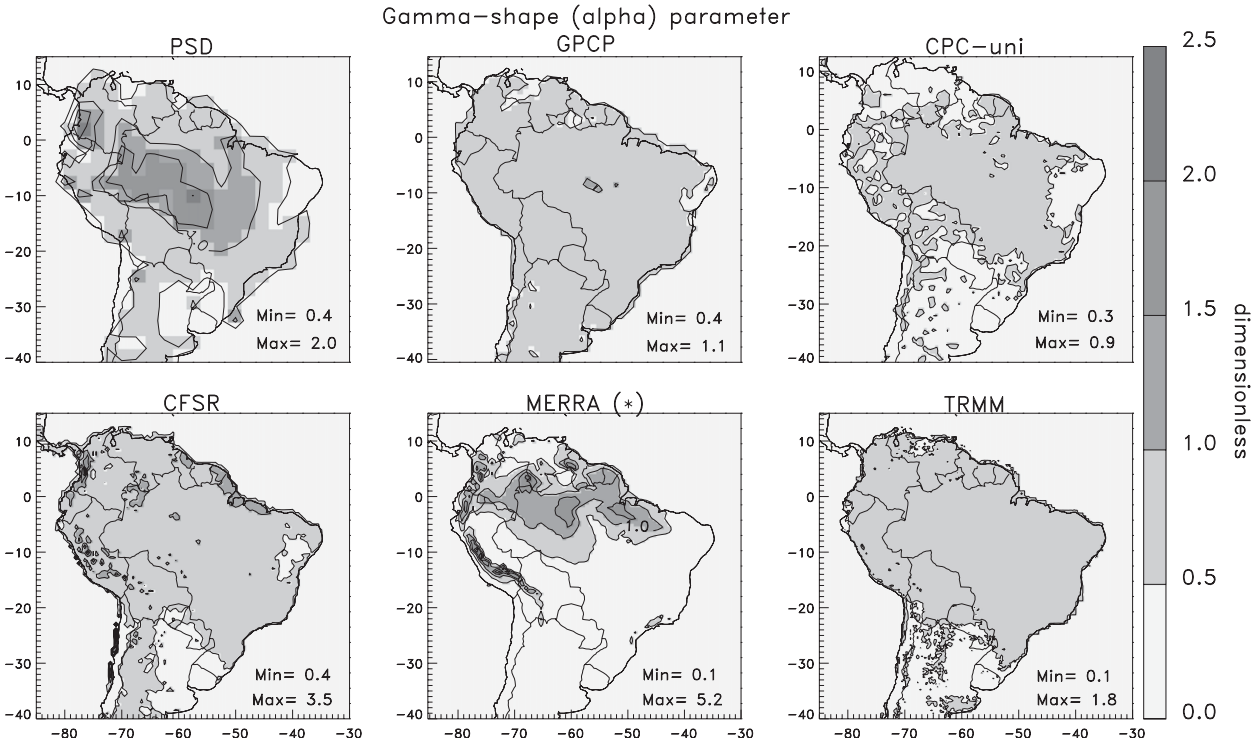

FIG. 15. Shape parameter of gamma probability distribution function (dimensionless). Datasets and minimum/maximum values are indicated in each panel. Note that to fit α from MERRA in the same scale the shape parameter has been divided by 2. Datasets have different grid spacings. Time period is 1 Nov–31 Mar 1998–2008.

in explaining these differences. The temporal variability of PC1, especially the distribution of intraseasonal variance (see power spectra next), may explain some of these differences.

Dates of onset, demise, and duration of SAMS are shown in Fig. 11 and can be compared with the same data resolution results (Fig. 4). Identical patterns are noted such that the mean onset date (Fig. 11, top) is highly coherent among PSD, GPCP, CPC-uni, and TRMM (~21 October) including the ranges of minimum and maximum onset dates. In contrast, the mean onset dates in CFSR and MERRA are off by several weeks. The variability in dates of mean demise (Fig. 11, middle) indicates agreements among PSD, GPCP, CPC-uni, and TRMM and some differences in CFSR and large disagreement in MERRA. Consequently, the mean durations of SAMS (~180 days) agree reasonably well among PSD, GPCP, CPC-uni, and TRMM data and is shorter and more variable in the CFSR and MERRA reanalyses (Fig. 11, bottom). These results indicate that differences in data resolution do not explain disagreements in the annual evolution of SAMS especially between CFSR and MERRA and the other datasets.

Fig. 12 shows the temporal variability of the seasonal amplitude of SAMS in each dataset. Since the magnitudes of PC1 vary significantly among the datasets, the amplitudes are normalized by removing the mean seasonal amplitude and dividing by the standard deviation of the 10 seasons (1998–2008). This allows visualizing the seasonal amplitudes on the same scale. Table 3 shows the cross correlations in seasonal amplitudes and can be compared with Table 1. The largest differences relative to the same resolution correlations are between PSD–TRMM (0.92), CPC-uni–GPCP (0.48), GPCP–CFSR (0.54), GPCP–MERRA (0.65), CPC-uni–TRMM (0.56), and MERRA–TRMM (0.54). The correlations between reanalyses and CPC-uni are surprisingly negative (−0.03, −0.14). It is interesting to note that all datasets consistently indicate a short-term trend in seasonal amplitudes of SAMS after 2004, which agrees with L. M. V. Carvalho et al. (2011, unpublished manuscript), who investigated long-term changes in SAMS large-scale circulation and warming in South America and Atlantic Ocean.

As before, the mean annual cycle is removed from the PC1 series and cross correlations among the datasets are computed (Table 4). A comparison between these results and Table 2 shows only very minor differences. This indicates that differences in the representation of subseasonal and interannual variations among the datasets

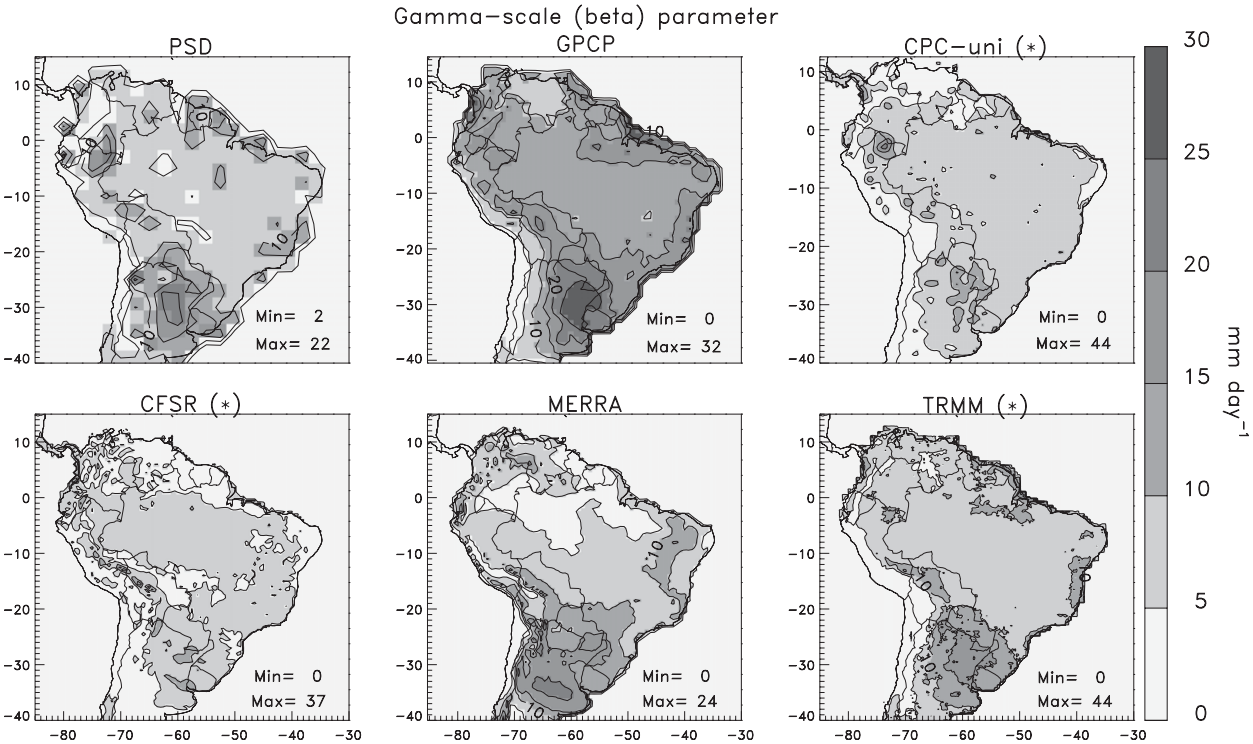

FIG. 16. Scale parameter of gamma probability distribution function (mm day^{-1}). Datasets and minimum/maximum values are indicated in each panel. Note that to fit β from CPC-Uni, CFSR, and TRMM in the same scale the scale parameters have been divided by 2. Datasets have different grid spacings. Time period is 1 Nov–31 Mar 1998–2008.

do not vary significantly with spatial resolution. Evidently, these conclusions hold for the large-scale aspects of SAMS inferred from one EOF mode. Locally, differences in subseasonal, seasonal, and interannual variations can vary significant among the datasets and spatial resolutions can be important.

Several studies have shown that intraseasonal variations play an important role in modulating the variability of SAMS (Nogues-Paegle and Mo 1997; Liebmann et al. 1999; Nogues-Paegle et al. 2000; Jones and Carvalho 2002; Cunningham and Cavalcanti 2006; Gonzalez et al. 2008; Muza et al. 2009; Carvalho et al. 2011b). When calculated from outgoing longwave radiation (OLR) anomalies, the first EOF shows regions of positive and negative anomalies associated with the eastward propagation of the MJO as well as midlatitude wave trains propagating over eastern South America.

The spectra from daily PC1 (Fig. 13) indicate that intraseasonal variations (~20–90 days) are captured by all datasets, although with different degrees of amplitude. Most noticeably, the intraseasonal spectral peak from GPCP exceeds the 95% confidence level above the background red noise. Similar peaks are evident in CPC-uni, MERRA, and TRMM, although they barely exceed the 95% significance level. It should be noted that the spectra derived from 2.5° lat/lon data are identical to the results in Fig. 13 indicating that the large-scale representation of subseasonal variations vary among the datasets but spatial resolution is not the dominant factor.

Fig. 14 shows the mean precipitation during 1 November–31 March 1998–2008 and can be compared with Fig. 5. In general, differences in spatial patterns of mean precipitation are small among PSD, GPCP, CPC-uni, and TRMM data. These datasets consistently show maximum precipitation over the central Amazon including the extensions to southeastern Brazil and toward Amapá state, where the Amazon River meets the Atlantic Ocean. The mean precipitation in the PSD data is slightly less than in GPCP, CPC-uni, and TRMM data (minimum and maximum values are indicated in each panel). It is also interesting to note that GPCP, CPC-uni, and TRMM data agree in the high precipitation over Colombia, whereas the agreement is less obvious over Peru. Although precipitation from the new reanalyses certainly improves from previous reanalyses (Bosilovich et al. 2008; Rienecker et al. 2011; Silva et al. 2011), substantial discrepancies are observed in the CFSR and MERRA reanalyses relative to the other datasets. While the spatial pattern of mean CFSR precipitation over the central Amazon is somewhat consistent with

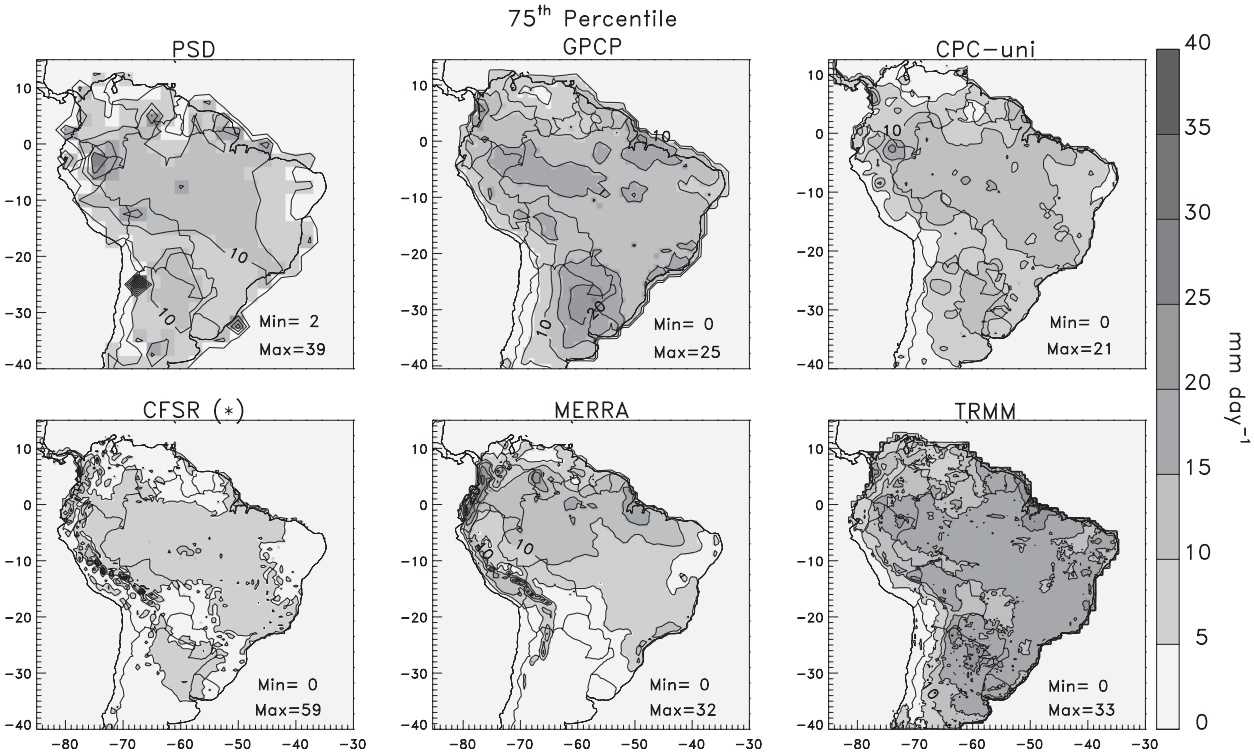

FIG. 17. 75th percentile of daily precipitation (mm day^{-1}). Datasets and minimum/maximum values are indicated in each panel. Note that to fit the 75th percentile from CFSR in the same scale the percentile has been divided by 2. Datasets have different grid spacings. Time period is 1 Nov–31 Mar 1998–2008.

the other datasets (except MERRA), CFSR precipitation is unrealistically large and spotty over the Andes and southeastern Brazil and the maximum precipitation (43 mm day^{-1}) is significantly larger than in the PSD, GPCP, CPC-uni, and TRMM data. These features are much less evident in the 2.5° lat/lon comparison (Fig. 5). The spatial pattern of mean precipitation in MERRA appears shifted to the north relative to the other datasets. Although the MERRA precipitation is not as spotty as the CFSR data in some locations, precipitation is excessive over the Andes Mountains and the maximum precipitation (23 mm day^{-1}) is higher than in PSD, GPCP, CPC-uni, and TRMM data.

The shape parameter (Fig. 15) is quite comparable with 2.5° lat/lon data (Fig. 6), although some differences are worth noting. For instance, the region of $\alpha \geq 1.0$ over the Amazon obtained with 2.5° lat/lon GPCP data (Fig. 6) is practically absent in the 1° lat/lon comparison (Fig. 15). This suggests that spatially averaging precipitation increases α and, therefore, might explain $\alpha \geq 1.0$ over the Amazon obtained from PSD data. Except for small spatial variations, $0.5 \leq \alpha \leq 1.0$ over a large area over South America as derived from GPCP, CPC-uni, CFSR, and TRMM. The shape parameter estimated from MERRA reanalysis is markedly different than any other dataset both in spatial pattern and magnitudes (i.e., the field was divided by 2 to fit the same scale).

The patterns of scale parameter (Fig. 16) obtained from datasets at their original resolution are somewhat similar to the β derived with 2.5° lat/lon grids (Fig. 7). Note, however, that the range of values increase in some datasets (CPC-uni, CFSR, and TRMM), and β has been divided by 2 to fit in the same scale as the other datasets. Values $\beta \geq 10$ are seen over southern South America with maximum centered over northern Argentina. The spatial patterns of β between PSD and CPC-uni are similar over tropical South America, although the magnitudes are twice as large in CPC-uni. The β spatial pattern from TRMM is consistent with PSD, although the magnitudes are much larger in the former. MERRA is significantly different than any other dataset over the northern Amazon.

To further characterize the frequency distributions of precipitation, Fig. 17 shows the 75th percentile (P75) and can be compared with Fig. 8. The spatial pattern of P75 is quite similar between PSD and CPC-uni and, to some extent, between GPCP and TRMM, although the magnitudes vary among these datasets. While the spatial pattern of P75 from CFSR is similar to the GPCP–TRMM patterns, the magnitudes are very different. Interestingly,

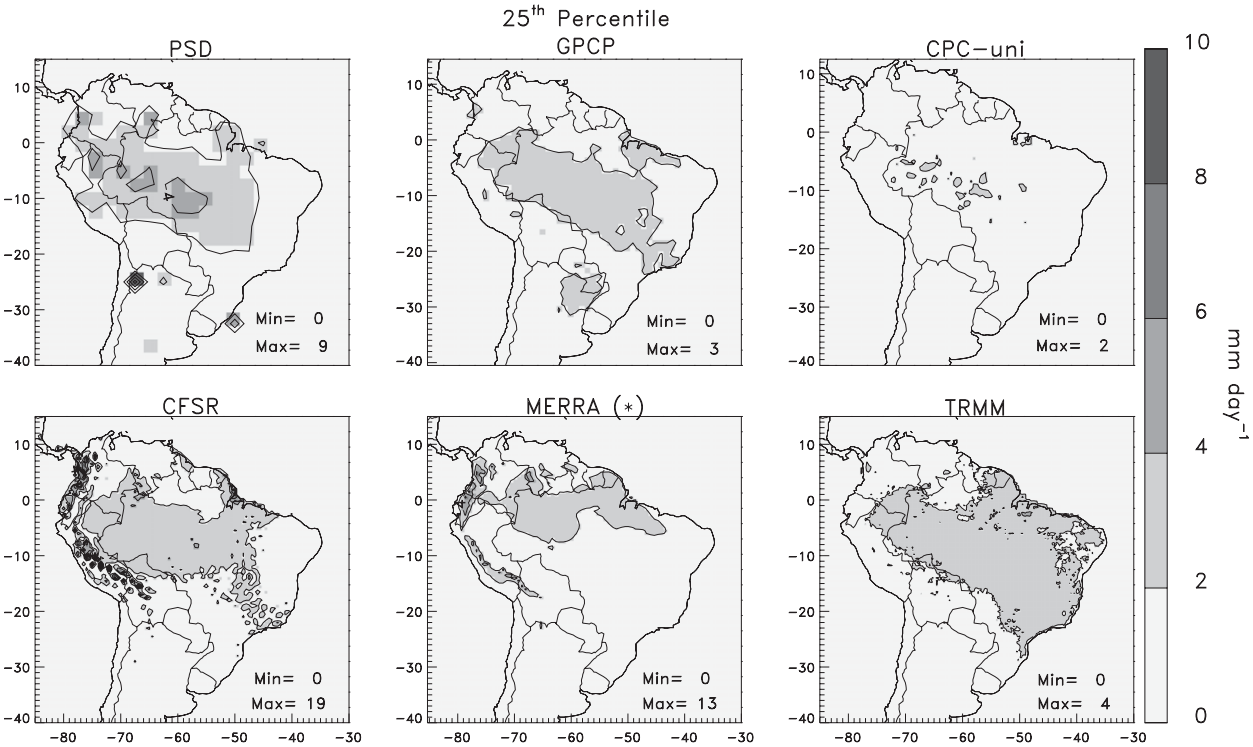

FIG. 18. 25th percentile of daily precipitation (mm day^{-1}). Datasets and minimum/maximum values are indicated in each panel. Note that to fit the 25th percentile from MERRA in the same scale the percentile has been divided by 2. Datasets have different grid spacings. Time period is 1 Nov–31 Mar 1998–2008.

MERRA shows P75 comparable to PSD, GPCP, and CPC-uni over central–southeastern Brazil, but the patter is different over the northern Amazon.

Fig. 18 shows comparisons in the 25th percentile (P25) derived from original grids and can be further evaluated with Fig. 9. P25 agrees among PSD, GPCP, CFSR, and TRMM over a large portion of the SAMS region (except over the Andes), whereas CPC-uni shows smaller P25 values. CFSR also shows much large P25 values over the central and northern Andes. P25 derived from MERRA is evidently shifted over the northern Amazon.

7. Summary and conclusions

The variability of gridded precipitation over SAMS in six different datasets is investigated during a common period (1998–2008). The range of grid spacings is 0.25°–2.5° lat/lon, and precipitation is derived from surface stations (PSD, CPC-uni), satellite data (GPCP, TRMM), and first-guess forecasts from reanalysis products (CFSR, MERRA). The same statistical analyses are applied to data in 1) a common 2.5° lat/lon grid and 2) in the original resolutions of the datasets.

Empirical orthogonal function analysis of daily precipitation allows the identification of the main mode of large-scale variability in the SAMS (EOF1/PC1). All datasets consistently indicate a region of positive precipitation anomalies over the core of the monsoon and negative anomalies over northern South America. Although positive precipitation anomalies over the central Andes are consistently represented in the GPCP, CPC-uni, CFSR, MERRA, and TRMM, the large spatial gradients in correlations derived from CFSR and MERRA suggest that the new reanalysis products overestimate precipitation in that region. This important aspect requires further investigation. Because of the sparseness of station data over the Andes, the details of precipitation in that region are still unknown. The early study by Hoffman (1975) suggests that gridded precipitation may underestimate the heavy precipitation over the eastern flanks of the Andes.

The onset, demise, and duration of SAMS are consistently represented among the PSD, GPCP, CPC-uni, and TRMM datasets in both 2.5° lat/lon and original resolutions. In contrast, CFSR and MERRA reanalyses show surprising shifts in the dates of onset and demise relative to the other datasets and seem to have problems in capturing

the correct timing of SAMS. The spectral variance in precipitation is examined by computing power spectra of daily PC1 time series. It is encouraging that the distribution of intraseasonal variance is somewhat similar in both comparisons (2.5° lat/lon and original grids), indicating that the large-scale representation of subseasonal variations agree among the datasets. Because of the limited data sampling (10 yr), differences in interannual variability among the datasets are not explored here.

Comparisons of mean precipitation during the monsoon indicate that differences in spatial patterns are small among PSD, GPCP, CPC-uni, and TRMM data and some discrepancies are found in the CFSR and MERRA reanalyses. While the spatial pattern of mean CFSR precipitation over the central Amazon is somewhat consistent with the other datasets (except MERRA), CFSR precipitation is unrealistically large and spotty over the Andes and southeastern Brazil and the maximum precipitation (43 mm day^{-1} with 0.5° lat/lon) is significantly larger than in the PSD, GPCP, CPC-uni, and TRMM data. These results appear consistent with Silva et al. (2011), who compared CFSR against the NCEP–NCAR (R1) and NCEP–Department of Energy (DOE) (R2) reanalyses and the CPC-uni precipitation. They noted that, although precipitation from CFSR is more consistent with observations than R1 and R2, large biases are still present, particularly during the SAMS when CFSR appears to overestimate precipitation. Moreover, Silva et al. (2011) showed that CFSR has substantial biases in intensity and frequency of precipitation events. One of the most worrisome findings is that the spatial pattern of mean precipitation from MERRA appears shifted to the north relative to the other datasets, which is not typical of the summer monsoon.

Last, the results of this study highlight an important issue that has yet to be resolved before reliable assessments of climate changes in South America can be achieved. The fitting of gamma frequency distributions to daily precipitation shows significant differences among all datasets in the parameters that characterize the shape, scale, and tails of the distributions. This implies that significant uncertainties exist in the characterization of extreme precipitation in these datasets. These discrepancies are not only relevant over the Amazon, where the maximum climatological precipitation amount is observed, but are also remarkable over the eastern Andes and east-central Brazil. These regions are highly susceptible to long-term changes in SAMS characteristics (L. M. V. Carvalho et al. 2011, unpublished manuscript; Carvalho et al. 2011a).

Acknowledgments. L. M. V. Carvalho, C. Jones, and B. Liebmann thank the support of NOAA's Climate Program Office (NA07OAR4310211 and NA10OAR4310170). L. M. V Carvalho, C. Jones, A. Posadas, and R. Quiroz thank USAID-CIP (Subcontract SB100085). L. M. V Carvalho and C. Jones thank the NSF Rapid Program (AGS-1126804). NCEP–NCAR reanalysis and OLR data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado (www.esrl.noaa.gov). TRMM data were acquired by an international joint project sponsored by the Japan National Space Development Agency (NASDA) and the U.S. National Aeronautics Space Administration (NASA) Office of Earth Science. The help from Bob Dattore, NCAR-CISL, in providing the NCEP CFSR data is greatly appreciated.

REFERENCES

Berbery, E. H., and E. A. Collini, 2000: Springtime precipitation and water vapor flux over southeastern South America. *Mon. Wea. Rev.,* **128,** 1328–1346.

Bookhagen, B., and M. R. Strecker, 2008: Orographic barriers, high-resolution TRMM rainfall, and relief variations along the eastern Andes. *Geophys. Res. Lett.,* **35,** L06403, doi:10.1029/2007GL032011.

——, and ——, 2010: Modern Andean rainfall variation during ENSO cycles and its impact on the Amazon basin. *Neogene History of Western Amazonia and Its Significance for Modern Diversity,* C. Hoorn, H. Vonhof, and F. Wesselingh, Eds., Blackwell Publishing, 223–41.

——, and D. W. Burbank, 2011: Towards a complete Himalayan hydrologic budget: The spatiotemporal distribution of snow melt and rainfall and their impact on river discharge. *J. Geophys. Res.,* **115,** F03019, doi:10.1029/2009JF001426.

Bosilovich, M. G., J. Chen, F. R. Robertson, and R. F. Adler, 2008: Evaluation of global precipitation in reanalyses. *J. Appl. Meteor. Climatol.,* **47,** 2279–2299.

Carvalho, L. M. V., C. Jones, and B. Liebmann, 2002a: Extreme precipitation events in southeastern South America and large-scale convective patterns in the South Atlantic convergence zone. *J. Climate,* **15,** 2377–2394.

——, ——, and M. A. F. Silva Dias, 2002b: Intraseasonal large-scale circulations and mesoscale convective activity in tropical South America during the TRMM-LBA campaign. *J. Geophys. Res.,* **107,** 8042, doi:10.1029/2001JD000745.

——, ——, and B. Liebmann, 2004: The South Atlantic convergence zone: Intensity, form, persistence, and relationships with intraseasonal to interannual activity and extreme rainfall. *J. Climate,* **17,** 88–108.

——, ——, A. E. Silva, B. Liebmann, and P. L. S. Dias, 2011a: The South American monsoon system and the 1970s climate transition. *Int. J. Climatol.,* **31,** 1248–1256, doi:10.1002/joc.2147.

——, A. E. Silva, C. Jones, B. Liebmann, and H. Rocha, 2011b: Moisture transport and intraseasonal variability in the South America monsoon system. *Climate Dyn.,* **36,** 1865–1880.

Chen, M. Y., W. Shi, P. P. Xie, V. B. S. Silva, V. E. Kousky, R. W. Higgins, and J. E. Janowiak, 2008: Assessing objective techniques for gauge-based analyses of global daily precipitation. *J. Geophys. Res.,* **113,** D04110, doi:10.1029/2007JD009132.

Cunningham, C. A. C., and I. F. D. Cavalcanti, 2006: Intraseasonal modes of variability affecting the South Atlantic convergence zone. *Int. J. Climatol.,* **26,** 1165–1180.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quart. J. Roy. Meteor. Soc.,* **137,** 553–597, doi:10.1002/qj.828.

Durieux, L., L. A. T. Machado, and H. Laurent, 2003: The impact of deforestation on cloud cover over the Amazon arc of deforestation. *Remote Sens. Environ.,* **86,** 132–140.

Gan, M. A., V. B. Rao, and M. C. L. Moscati, 2006: South American monsoon indices. *Atmos. Sci. Lett.,* **6,** 219–223.

Gandu, A. W., and P. L. Silva Dias, 1998: Impact of tropical heat sources on the South American tropospheric upper circulation and subsidence. *J. Geophys. Res.,* **103** (D6), 6001–6015.

Gonzalez, P. L. M., C. S. Vera, B. Liebmann, and G. Kiladis, 2008: Intraseasonal variability in subtropical South America as depicted by precipitation data. *Climate Dyn.,* **30,** 727–744.

Grimm, A. M., 2003: The El Niño impact on the summer monsoon in Brazil: Regional processes versus remote influences. *J. Climate,* **16,** 263–280.

——, 2004: How do La Niña events disturb the summer monsoon system in Brazil? *Climate Dyn.,* **22,** 123–138.

——, and M. T. Zilli, 2009: Interannual variability and seasonal evolution of summer monsoon rainfall in South America. *J. Climate,* **22,** 2257–2275.

——, C. S. Vera, and C. R. Mechoso, 2005: The South American monsoon system. *The American Monsoon Systems: An Introduction,* C.-P. Chang, B. Wang, and N.-C. G. Lau, Eds., World Meteorological Organization, 197–206.

Hartmann, D. L., and E. E. Recker, 1986: Diurnal variation of outgoing longwave radiation in the tropics. *J. Climate Appl. Meteor.,* **25,** 800–812.

Higgins, R. W., W. Shi, E. Yarosh, and R. Joyce, 2000: Improved United States precipitation quality control system and analysis. NCEP/Climate Prediction Center Atlas 7, 40 pp.

——, V. E. Kousky, V. B. S. Silva, E. Becker, and P. Xie, 2010: Intercomparison of daily precipitation statistics over the United States in observations and in NCEP reanalysis products. *J. Climate,* **23,** 4637–4650.

Hoffman, J., 1975: Maps of mean temperature and precipitation. *Climatic Atlas of South America,* Vol. 1, World Meteorological Organization, 1–28.

Horel, J. D., A. N. Hahmann, and J. E. Geisler, 1989: An investigation of the annual cycle of convective activity over the tropical Americas. *J. Climate,* **2,** 1388–1403.

Huffman, G. J., and Coauthors, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeor.,* **2,** 36–50.

——, and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeor.,* **8,** 38–55.

Jones, C., and L. M. V. Carvalho, 2002: Active and break phases in the South American monsoon system. *J. Climate,* **15,** 905–914.

——, D. E. Waliser, K. M. Lau, and W. Stern, 2004: Global occurrences of extreme precipitation and the Madden–Julian oscillation: Observations and predictability. *J. Climate,* **17,** 4575–4589.

Kodama, Y. M., 1992: Large-scale common features of subtropical precipitation zones (the baiu frontal zone, the SPCZ, and the SACZ). 1. Characteristics of subtropical frontal zones. *J. Meteor. Soc. Japan,* **70,** 813–836.

——, 1993: Large-scale common features of subtropical convergence zones (the baiu frontal zone, the SPCZ, and the SACZ). 2. Conditions of the circulations for generating the STCZs. *J. Meteor. Soc. Japan,* **71,** 581–610.

Kousky, V. E., 1988: Pentad outgoing longwave radiation climatology for the South American sector. *Rev. Bras. Meteor.,* **3,** 217–231.

Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.,* **15,** 809–817.

——, and Coauthors, 2000: The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteor.,* **39,** 1965–1982.

Legates, D. R., and C. J. Willmott, 1990: Mean seasonal and spatial variability in gauge-corrected, global precipitation. *Int. J. Climatol.,* **10,** 111–127.

Lenters, J. D., and K. H. Cook, 1999: Summertime precipitation variability over South America: Role of the large-scale circulation. *Mon. Wea. Rev.,* **127,** 409–431.

Liebmann, B., and D. Allured, 2005: Daily precipitation grids for South America. *Bull. Amer. Meteor. Soc.,* **86,** 1567–1570.

——, and ——, 2006: Reply. *Bull. Amer. Meteor. Soc.,* **87,** 1096–1096.

——, G. N. Kiladis, J. A. Marengo, T. Ambrizzi, and J. D. Glick, 1999: Submonthly convective variability over South America and the South Atlantic convergence zone. *J. Climate,* **12,** 1877–1891.

——, C. Jones, and L. M. V. Carvalho, 2001: Interannual variability of daily extreme precipitation events in the state of Sao Paulo, Brazil. *J. Climate,* **14,** 208–218.

——, G. N. Kiladis, C. S. Vera, A. C. Saulo, and L. M. V. Carvalho, 2004: Subseasonal variations of rainfall in South America in the vicinity of the low-level jet east of the Andes and comparison to those in the South Atlantic convergence zone. *J. Climate,* **17,** 3829–3842.

——, and Coauthors, 2007: Onset and end of the rainy season in South America in observations and the ECHAM 4.5 atmospheric general circulation model. *J. Climate,* **20,** 2037–2050.

Madden, R. A., and P. R. Julian, 1971: Detection of a 40–50 day oscillation in the zonal wind in the tropical Pacific. *J. Atmos. Sci.,* **28,** 702–708.

Marengo, J. A., 2004: Interdecadal variability and trends of rainfall across the Amazon basin. *Theor. Appl. Climatol.,* **78,** 79–96.

——, 2009: Long-term trends and cycles in the hydrometeorology of the Amazon basin since the late 1920s. *Hydrol. Proc.,* **23,** 3236–3244.

——, 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.

——, M. W. Douglas, and P. L. S. Dias, 2002: The South American low-level jet east of the Andes during the 1999 LBA-TRMM and LBA-WET AMC campaign. *J. Geophys. Res.,* **107,** 8079, doi:10.1029/2001JD001188.

——, and Coauthors, 2012: Recent developments on the South American monsoon system. *Int. J. Climatol.,* **32,** 1–21, doi:10.1002/joc.2254.

Mitchell, J. M., Jr, 1966: Climate change. World Meteorological Organization Tech. Note 79, 79 pp. [Available from WMO, 41 Avenue Giuseppe-Motta-1211, Geneva 2, Switzerland.]

Muza, M. N., L. M. V. Carvalho, C. Jones, and B. Liebmann, 2009: Intraseasonal and interannual variability of extreme dry and wet events over southeastern South America and subtropical Atlantic during the austral summer. *J. Climate,* **22,** 1682–1699.

Nogues-Paegle, J., and K. C. Mo, 1997: Alternating wet and dry conditions over South America during summer. *Mon. Wea. Rev.,* **125,** 279–291.

——, L. A. Byerle, and K. C. Mo, 2000: Intraseasonal modulation of South American summer precipitation. *Mon. Wea. Rev.,* **128,** 837–850.

Rienecker, M. M., and Coauthors, 2011: MERRA—NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate,* **24,** 3624–3648.

Robertson, A. W., and C. R. Mechoso, 1998: Interannual and decadal cycles in river flows of southeastern South America. *J. Climate,* **11,** 2570–2581.

——, and ——, 2000: Interannual and interdecadal variability of the South Atlantic convergence zone. *Mon. Wea. Rev.,* **128,** 2947–2957.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.,* **91,** 1015–1057.

Silva, A. E., and L. M. V. Carvalho, 2007: Large-scale index for South America monsoon (LISAM). *Atmos. Sci. Lett.,* **8,** 51–57.

Silva, V. B. S., V. E. Kousky, W. Shi, and R. W. Higgins, 2007: An improved gridded historical daily precipitation analysis for Brazil. *J. Hydrometeor.,* **8,** 847–861.

——, ——, and R. W. Higgins, 2011: Daily precipitation statistics for South America: An intercomparison between NCEP reanalyses and observations. *J. Hydrometeor.,* **12,** 101–117.

Silva Dias, P. L., W. H. Schubert, and M. DeMaria, 1983: Large-scale response of the tropical atmosphere to transient convection. *J. Atmos. Sci.,* **40,** 2689–2707.

Vera, C., and Coauthors, 2006: Toward a unified view of the American monsoon systems. *J. Climate,* **19,** 4977–5000.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Vol. 91, Academic Press, 648 pp.

Xie, P. P., and Coauthors, 2003: GPCP pentad precipitation analyses: An experimental dataset based on gauge observations and satellite estimates. *J. Climate,* **16,** 2197–2214.

Zhou, J. Y., and K. M. Lau, 1998: Does a monsoon climate exist over South America? *J. Climate,* **11,** 1020–1040.