Investigating Interannual Variability of Precipitation at the Global Scale: Is There a Connection with Seasonality?

S. FATICHI

Department of Civil and Environmental Engineering, University of Firenze, Florence, Italy, and Institute of Environmental Engineering, ETH Zurich, Zurich, Switzerland

V. YU. IVANOV

Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, Michigan

E. CAPORALI

Department of Civil and Environmental Engineering, University of Firenze, Florence, Italy

ABSTRACT

Interannual variability of precipitation can directly or indirectly affect many hydrological, ecological, and biogeochemical processes that, in turn, influence climate. Despite the significant importance of the phenomenon, few studies have attempted to elucidate spatial patterns of this variability at the global scale. This study uses land gauge precipitation records of the Global Historical Climatology Network, version 2, as well as reanalysis data to provide an assessment of the spatial organization of characteristics of precipitation interannual variability. The coefficient of variation, skewness, and short- and long-range dependence of the precipitation variability are analyzed. Among the major inferences is that the coefficient of variation of annual precipitation shows a significant correlation with intra-annual seasonality. Specifically, subyearly precipitation anomalies occurring in locations with pronounced seasonality affect the total yearly amount, imposing a higher variability in the annual precipitation fluctuations. Furthermore, the study illustrates that a positive skewness of the distribution of annual precipitation is a robust property worldwide and its magnitude is related to the coefficient of variation. Additionally, annual precipitation exhibits very weak small-lag autocorrelation. Conversely, the intensity of long-memory–long-range dependence is significantly larger than zero, hinting that organized long-term variations are an important feature of the interannual variability of precipitation.

1. Introduction

Identifying the nature and patterns of the interannual variability of precipitation can be crucial because these fluctuations exert a long-term control on water resources, affect plant growth and the biogeochemical cycle, and modulate extreme events, such as floods and prolonged dry periods. For instance, several studies suggested that the variability of annual precipitation can be important for the temporal dynamics of aboveground primary production and thus for global vegetation biogeography (Knapp and Smith 2001; Fang et al. 2001; Wiegand et al.

2004; Yang et al. 2008). The interannual variability of hydroclimatic variables (including precipitation) has been often connected to indices of large-scale atmospheric patterns, such as El Niño–Southern Oscillation (ENSO), Arctic Oscillation (AO), North Atlantic Oscillation (NAO), sea level pressure (SLP), or to sea surface temperature (SST) (New et al. 2001; Mason and Goddard 2001; Chen et al. 2002; Whiting et al. 2003; Grimm and Natori 2006; Gu and Adler 2004, 2006, 2009; Gu et al. 2007; Medvigy et al. 2008; Small and Islam 2008; Bartolini et al. 2009). These large-scale atmospheric patterns, especially ENSO, were successful in explaining the time variability of precipitation for certain geographic areas (New et al. 2001; Chen et al. 2002; Gu and Adler 2011). Nonetheless, they cannot explain the spatial interannual dynamics of precipitation at the global scale.

Corresponding author address: Simone Fatichi, Institute of Environmental Engineering, ETH Zürich, Wolfgang-Pauli-Str. 15, HIL D 23.2, 8093 Zurich, Switzerland.
E-mail: simone.fatichi@ifu.baug.ethz.ch

Other analyses have been limited to the investigation of the interannual variability of precipitation during a particular season (e.g., summer or fall) or for a specific region (Cayan et al. 1998; Higgins et al. 1998; Rodriguez-Puebla et al. 1998; Coulibaly 2006; Small and Islam 2008; Mourato et al. 2010). When global datasets of precipitation were analyzed to detect precipitation trends, only time series of anomalies were investigated (Dai et al. 1997; New et al. 2001; Chen et al. 2002), while spatial patterns of precipitation interannual variability were neglected. Practical and theoretical reasons contribute to the complication of analyses of the interannual variability of precipitation. The reasons are (i) the relatively short and spatially sparse observed records and (ii) insufficient understanding of the atmospheric dynamics responsible for this process. For example, a modeling study with general circulation models has highlighted the difficulty of understanding the causes and predictability of the interannual variability of precipitation (Koster et al. 2000).

Despite numerous previous research on global precipitation (e.g., Legates 1991; Xie and Arkin 1997; New et al. 2001; Dai 2006), so far we have not reached a generalized statistical characterization of the structure of the annual variability of land precipitation. Furthermore, no previous study explicitly analyzed the uncertainty of characterizing the interannual variability of precipitation related to short observational records. Yet, interannual variability affects many earth processes and can lead to a range of disturbance phenomena, such as vegetation mortality and establishment and/or biogeochemistry dynamics, feeding back to the global climate. Therefore, a more detailed investigation addressing these issues is still warranted.

A detailed analysis of statistics and spatial patterns of metrics characterizing the interannual variability of land precipitation at the global scale is presented in this study. These include the coefficient of variation, autocorrelation, skewness, and variables characterizing intra-annual seasonality and long-range dependence of the process. Global patterns of dependencies and properties are investigated. Furthermore, the study addresses the issue of uncertainty inherent to inferences based on short observational records. Major inferences of this study are based on data from a precipitation database, the monthly scale of the Global Historical Climatology Network (GHCN-Monthly) version 2, available through the National Climatic Data Center (NCDC) (Peterson and Vose 1997). The retrieved data represent one of the most complete databases and include measurements from precipitation gauges located worldwide with the longest available records.

Although point-scale precipitation exhibits a non-homogeneous coverage of the earth, it avoids problems related to the interpolation of precipitation data on grids (Dai et al. 1997; Xie and Arkin 1997; Chen et al. 2002; Adler et al. 2003; Mitchell and Jones 2005; Ensor and Robeson 2008). Additionally, the presented study summarizes all of the available information and has been carried out at the global scale (for land areas), which overcomes the limitations of local-scale analyses. Similar analyses using the data from reanalysis projects such as the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis (Kalnay et al. 1996; Kistler et al. 2001), the 40-yr European Centre for Medium-Range Weather Forecasts Re-Analysis (ERA-40) (Uppala et al. 2005), and precipitation products from the Global Precipitation Climatology Project (GPCP) (Huffman et al. 1997; Adler et al. 2003; Rudolf et al. 2010) have confirmed the findings of this study. This further assures that the results are not affected by the nonhomogeneity of spatial coverage of the gauge stations.

It must be underlined that no mechanistic explanation of interannual precipitation variability is attempted here. Addressing causes and mechanisms controlling interannual variability would require a different type of study and data. However, statistical inferences characterizing precipitation interannual variability drawn in this study could be important for future research addressing these needs. They can be also important for studies aiming to parameterize the variability of precipitation in hydrological, ecological, or other disciplines.

2. Data and methods

Precipitation data are obtained from GHCN-Monthly (Peterson and Vose 1997; Chen et al. 2002). This archive contains two sets of data. The first type represents the historical precipitation data for thousands of land stations worldwide checked for quality but not adjusted to account for inhomogeneities (Alexandersson 1986; Peterson et al. 1998). The second set contains adjusted data, free of inhomogeneities. The adjusted dataset is significantly smaller and includes a subset of stations from the first dataset, mainly located in the United States and a few additional stations from other countries (Peterson et al. 1998). In this study, the two subsets were merged using homogeneity adjusted data and nonadjusted data. In total 20 319 stations with more than 10 yr of data were identified, corresponding to more than one million station-year data. The period of record varies spatially, with several thousand stations having continuous records that extend back to 1950, that is, about 60 yr of observations and about 150 stations with more than 150 yr of data. The last analyzed year is the 2009. The best spatial coverage is in North America, Europe, Australia, and some
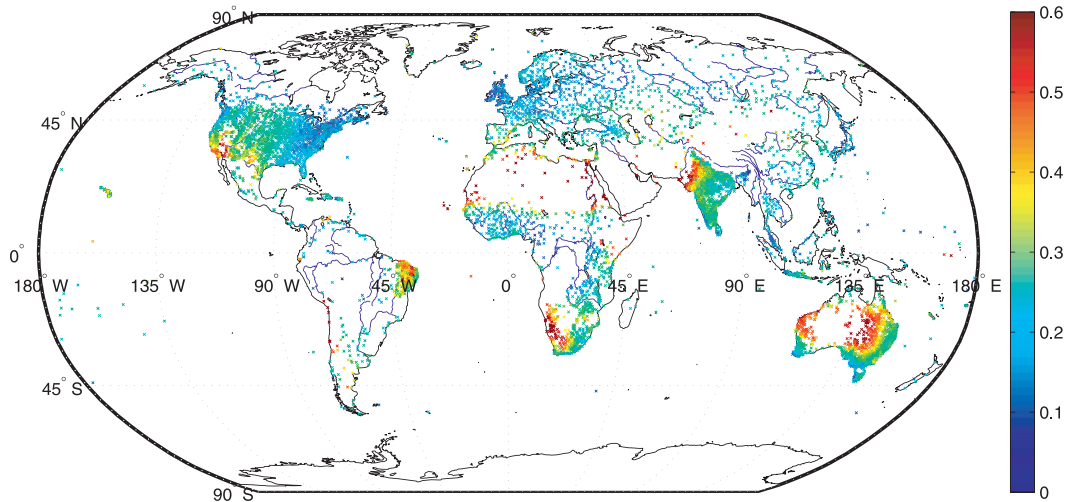
FIG. 1. The global map of C_v. Only stations with >50 yr of observations are included ($n = 8197$).

parts of Asia. Likewise, the coverage in the Northern Hemisphere is better than in the Southern Hemisphere (Fig. 1). The historical data network underwent rigorous quality screening by the NCDC. Additionally, in this study, when one or more months had missing values in a given year, such a year was eliminated from the analysis.

The variability of annual precipitation for each station was characterized by calculating several statistics of the process. In addition to conventional statistics—such as the coefficient of variation C_v, skewness γ, and the first 10 lags of autocorrelation $\rho(1, \ldots, 10)$—the long-range dependence–long-memory parameter of the process d was estimated using the periodogram method (Taqqu et al. 1995; Caballero et al. 2002). Specifically, the long-range dependence is a behavior observable when time series exhibit a slow decay of the autocorrelation function for large lags (Beran 1994; Koutsoyiannis 2003; Fatichi et al. 2009). Long memory can be characterized by a parameter d that can be related to the Hurst coefficient, $H = d + 0.5$ (Beran 1994), and expresses the intensity of the long-memory process. For a stationary time series, d varies within the range $0 \le d \le 0.5$ (Beran 1994).

To explore the linkage of interannual variability with intra-annual seasonality of precipitation, indices characterizing the latter process were calculated. Specifically, the precipitation concentration index (PCI) (De Lùis et al. 2000; Fatichi and Caporali 2009), the seasonality index (SI) (Walsh and Lawler 1981; Pryor and Schoof 2008), the seasonality concentration index (SCI) (Fujita 2008), and the seasonal variability within year, M_r, as defined by Davidowitz (2002) using the Levene's statistic, were calculated for each station-year using monthly data. They were consequently averaged over the observation period for each station. For the PCI, the coefficient of

temporal variation, $C_{v,\text{PCI}}$, was additionally calculated. Details of the definition and estimation of the above-mentioned indices are provided in the appendix.

Since the characterization of metrics of interannual variability is expected to strongly depend on the length of the time series, short records were excluded from the analysis and all of the above-mentioned statistics were calculated only for stations with an observation period longer than 50 yr (the number of stations, $n = 8197$). For the parameter d, characterizing the long-range dependence, this period was increased to 90 yr ($n = 3358$) due to inherent uncertainties in estimating d using short time series.

All of the correlation statistics and their significance (p values) have been computed with classic statistical tests (Wilks 2006). However, the presence of spatial autocorrelation in the variables can invalidate the assumption of independence among samples, artificially increasing the degrees of freedom in the traditional test of significance of the Pearson correlation coefficient (Liebhold and Sharov 1998). Because of this reason, we also tested the significance of the correlations with the modified t test and F test (Clifford et al. 1989; Dutilleul 1993; Dutilleul et al. 2008). These tests of decreasing the degrees of freedom for spatially correlated samples allow one to effectively evaluate the significance of the tests also when variables are spatially autocorrelated.

In addition to station data, C_v, γ, PCI, and $\rho(1, \ldots, 10)$ were calculated using four gridded datasets (Table 1) at the global scale. The NCEP–NCAR precipitation reanalysis product contains 62 yr of data (Kalnay et al. 1996; Kistler et al. 2001). Precipitation product from the reanalysis project ERA-40 contains 45 yr of data (Uppala et al. 2005). Two other products represent updates of the version 2 GPCP reanalysis data (Huffman et al. 1997;

TABLE 1. Characteristics of gridded precipitation datasets used in this study: NCEP–NCAR reanalysis data, ERA-40 data, VASClimO 50 yr, and GPCC Full Data Reanalysis product, version 5. The R^2 of the log-linear relationship between C_v and PCI–MAP; between γ and C_v; and between γ and MAP are also reported. The values in parentheses in the C_v–MAP–PCI column are for grid cells representing sea.

Database	Years	Resolution latitude–longitude grid (°)	R^2, C_v–MAP–PCI	$R^2, \gamma - C_v$	$R^2, \gamma -$ MAP
NCEP–NCAR	Jan 1948–Dec 2009 (62)	≈2	0.72 (0.57)	0.63	0.18
ERA-40	Sep 1957–Aug 2002 (45)	≈2.5	0.50 (0.67)	0.82	0.05
VASClimO	Jan 1951–Dec 2000 (50)	≈0.5	0.60	0.63	0.40
GPCC Full Reanalysis	Jan 1901–Dec 2009 (109)	≈0.5	0.78	0.47	0.05

Adler et al. 2003; Huffman et al. 2009): the Variability Analysis of Surface Climate Observations (VASClimO) dataset covers a period of 50 yr and the Global Precipitation Climatology Centre (GPCC) Full Data Reanalysis product, version 5, covers a period of 109 yr (Adler et al. 2003; Beck et al. 2005; Rudolf 2005; Rudolf and Schneider 2005; Rudolf et al. 2010). These two databases are not independent from the historical GHCN database because gauge data were also used in the compilation of these products, but they also used gauge data from other historical databases (Beck et al. 2005; Rudolf 2005; Rudolf and Schneider 2005). Specifically, the GPCC Full Data Reanalysis product, version 5, uses the complete GPCC station database (about 64 400 stations with at least 10 yr of data) with a monthly data coverage ranging from some 10 000 to more than 47 000 stations (Beck et al. 2005; Rudolf et al. 2010).

We note that precipitation time series derived from the reanalysis project and the GPCP products should be used with caution: the reanalysis data represent model outputs, whereas the GPCP gridded datasets are the result of spatial interpolation techniques (Beck et al. 2005; Rudolf and Schneider 2005; Ensor and Robeson 2008). Nonetheless, these data are still useful for a qualitative evaluation of the influence of using data from precipitation gauges that are irregularly distributed in space (Fig. 1). For example, it has been noted that even though reanalysis estimates of precipitation can be biased, the inferred interannual variability from these products tends to be well correlated with independent observations (Kistler et al. 2001). Statistics from precipitation reanalysis data and gridded datasets are therefore only used to assess the findings of this study; all quantitative inferences are fully based on the GHCN point-scale gauge observations.

3. Results

a. Global properties of annual precipitation variability

The global pattern of the coefficient of variation is illustrated in Fig. 1. About 92% of stations have the coefficient of variation between 0.15 and 0.5; although values as high as 3.5 can be detected. Gradual spatial gradients can be observed in Fig. 1, where smooth transitions from regions of low variability to regions of high variability are more typical than abrupt transitions or scattered patterns. Higher values of C_v can be identified in arid or semiarid areas, such as in the Southwest of the United States and northern Mexico, western India and Pakistan, African Sahara and Kalahari, and inland Australia. Exceptions to the above-mentioned observation are areas of inland Brazil and eastern Australia, where the values of C_v are high despite humid climates. Regions where climate is influenced by the northern Atlantic tend to have similar interannual variabilities: C_v of annual precipitation of the eastern United States and northern Europe exhibit a very narrow range of 0.16–0.21. Similar values can be observed for equatorial Africa, Indochina, Japan, and southern Australia. The observed patterns point to a possible relation between C_v and the mean annual precipitation (MAP), which has been addressed previously (Knapp and Smith 2001). There is indeed a tendency for a large interannual variability in regions characterized by lower precipitation. Nonetheless, the coefficient of determination between C_v and MAP is only $R^2 = 0.16$ (with a value of $p < 0.0001$), which increases to $R^2 = 0.36$ ($p < 0.0001$) when the logarithms of C_v and MAP are used. These results suggest that most of the interannual variability cannot be explained by MAP alone.

In the following, the seasonality of precipitation is analyzed further in attempt to better explain the geographical differences of C_v. Given the subjectivity of seasonality characterization, several indices have been computed, as described previously. While different in mathematical formulation, the PCI, SI, SCI, and M_r indices are strongly correlated among themselves. Specifically, the R^2 between the station long-term averages of PCI and the other three indices are 0.92, 0.96, and 0.77, respectively. For this reason, only the PCI was used in the following to represent precipitation seasonality. In short, PCI characterizes changes in the seasonal distribution of precipitation: values below 12 suggest a fairly uniform
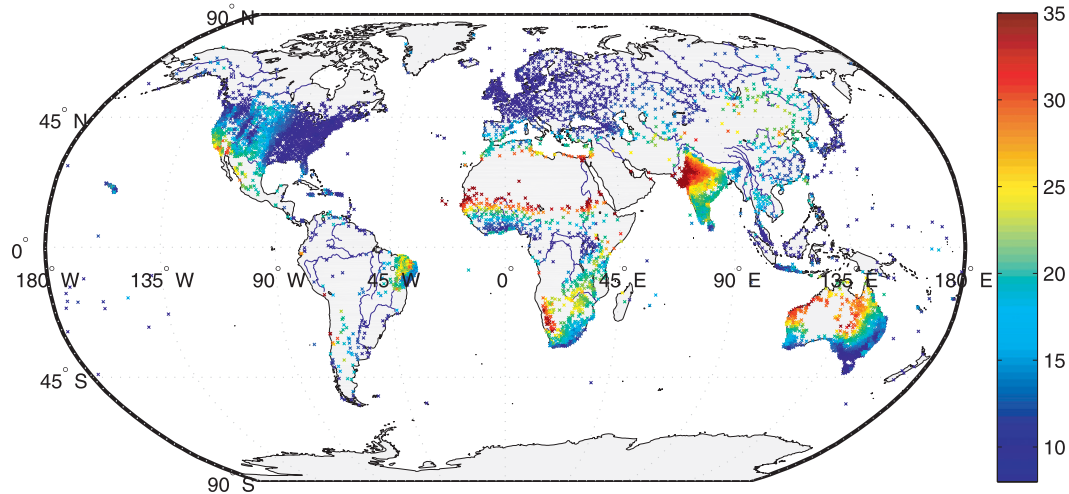
FIG. 2. The global map of the PCI. Stations with >50 yr of observations are included. About 94% of PCI values are contained between 10 and 30, but the range diminishes significantly for the eastern United States and northern Europe.

distribution of rainfall during the year, values between 12 and 20 indicate seasonality, and values above 20 indicate a strong irregularity of precipitation throughout the year. An illustration of the global distribution of PCI over land is available in Fig. 2.

A correlation emerges between the seasonality of precipitation, expressed as PCI, and the interannual variability C_v. A multiple regression analysis between C_v, PCI, and MAP indicates that the interannual variability of precipitation can be largely explained by PCI and MAP. The graphical correlation is shown in Fig. 3. The determination coefficient is $R^2 = 0.64$ ($p < 0.0001$, $n = 8197$) when stations with more than 50 yr of data are considered, and $R^2 = 0.69$ ($p < 0.0001$, $n = 4987$) and $R^2 = 0.74$ ($p < 0.0001$, $n = 3358$) when the data availability threshold is increased to 70 and 90 yr, respectively. This underlines how the length of the observational period can influence the results and that inferences on global patterns of precipitation variability are reinforced when longer time series are available. The increase of R^2 with the series length has been tested to be robust to differences in the composition of the sample, that is, the result is not an artifact of a better correlation for the specific set of stations with longer time series. The correlation illustrated in Fig. 3 is important because it suggests that the interannual variability of precipitation can be evaluated with a certain degree of accuracy, when only the mean annual precipitation and average intra-annual seasonality are known. The latter variables can be also approximated in the absence of long observational records.

A possible statistical interpretation of the positive correlation between irregularity in the precipitation seasonality and C_v can be attributed to a relatively small likelihood of strong interannual variability in areas

where precipitation is uniformly distributed within the year, as opposed to areas where precipitation is concentrated in just few months. When an anomaly of precipitation, such as a drought or a wet period, occurs in locations with pronounced seasonality, it is likely that the total annual amount is also nonnegligibly affected. Conversely, a more uniform within-year distribution of precipitation leads to a less appreciable sensitivity of the total annual precipitation to a seasonal anomaly. The above-mentioned considerations are corroborated by the significant correlation, $R^2 = 0.72$ ($p < 0.0001$), between C_v and the coefficient of variation of the
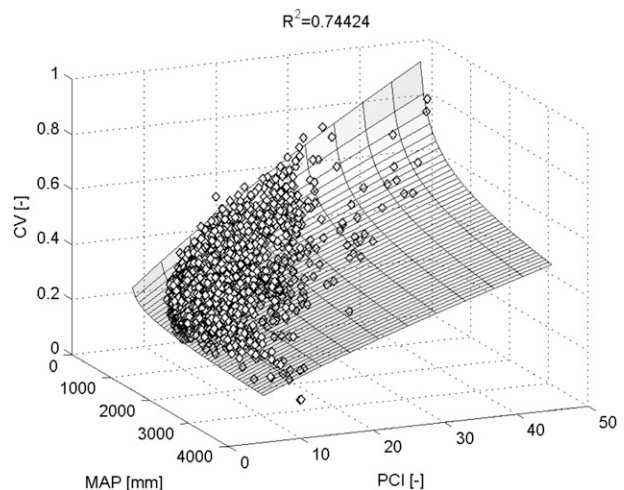

FIG. 3. The C_v as a function of MAP and PCI. Only data for stations with >90 yr of observations are included ($n = 3358$). The surface is a linear least squares fit of the log-transformed C_v, log-transformed MAP, and log-transformed PCI, $\log(C_v) = -2.3581 - 0.179 \log(\text{MAP}) + 0.793 \log(\text{PCI})$, $R^2 = 0.74$, ($p < 0.0001$).

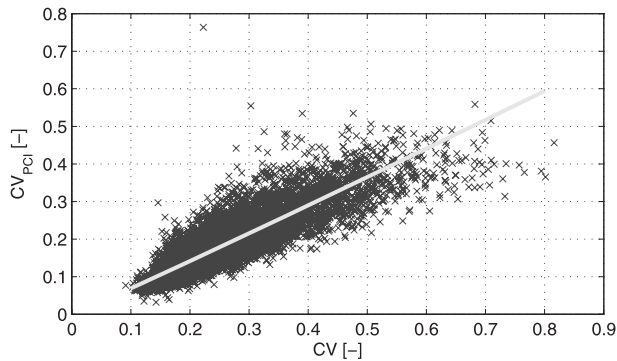FIG. 4. The $C_{v,\mathrm{PCI}}$ as a function of C_v. Only data for stations with >50 yr of observations are included; $C_{v,\mathrm{PCI}} = 0.748 C_v^{1.035}$, $R^2 = 0.72$, ($p < 0.0001$).

FIG. 5. (a). The distribution of skewness of annual precipitation at the global scale (bars) and the distribution of skewness of white-noise time series (line with crosses) equivalent to the observational records in terms of series number and lengths. Stations with >50 yr of data are used. (b) The distribution d as inferred from observations (bars) and from white-noise time series (line with crosses) equivalent to the observational records in terms of series number and lengths. Stations with >90 yr of data are used ($n = 3358$).

seasonality $C_{v,\mathrm{PCI}}$ (Fig. 4). The dynamics of precipitation interannual variability, expressed by C_v, are thus dominated by variations in the precipitation seasonality rather than by uniform increases or decreases of precipitation within the year. This consideration entails that monthly or seasonal anomalies of precipitation are much more likely to occur rather than an entire year or few-years-long anomalies. The statement is reinforced by the generally weak autocorrelation observed in annual precipitation time series, as explained later in this section. The fact that regions subjected to a significant seasonality also experience a strong interannual variability highlights these areas as most critical for water resource management.

Other statistical properties of the interannual precipitation process were analyzed in search of emerging global properties. An empirical distribution of skewness of annual precipitation γ, obtained from the analyzed time series is shown in Fig. 5a. The calculation of skewness is relatively uncertain given the short length of most series. To highlight the degree of this uncertainty, white-noise time series equivalent to the analyzed record in terms of the number of series and their lengths were simulated. Theoretically, all of the simulated white-noise time series have skewness equal to zero, but their lengths are often insufficient for a correct estimation of γ. This is apparent when the distribution of the skewness of the simulated white noise is plotted together with the empirical distribution of γ estimated from the precipitation record (Fig. 5a). This implies that the observed variability in γ is mainly due to the uncertainty of its estimation. However, a clear tendency of γ distribution obtained from the observations toward positive values is appreciable, with the median of γ equal to 0.42. This tendency of annual and seasonal precipitation to have positive skewness was previously noticed (Srikanthan and McMahon 1982; Sardeshmukh et al. 2000) but here it emerges as a global tendency. From the global land map (Fig. 6), no clear
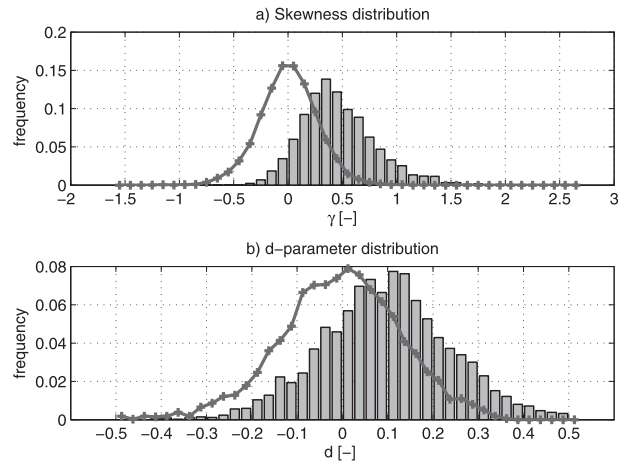
regional or continental pattern of γ can be detected. However, γ is positively correlated with C_v ($R^2 = 0.329$, $p < 0.0001$) and weakly negatively correlated with MAP ($R^2 = 0.063$, $p < 0.0001$), suggesting that underlying patterns may in fact exist. The inability to detect them is almost certainly affected by the short observational records that hinder robust inferences. Note that in this analysis, γ was computed with the method of moments; however, using other estimators, such as the Pearson skewness coefficient, does not affect the results.

Lags from 1 to 10 yr of the autocorrelation process have been also investigated. Similar to the case of skewness estimation, their identification is also complicated by the relatively short duration of time series. Autocorrelations of synthetic time series of white noise calculated following the same approach as for the analysis of skewness were computed for the same number of lags. The variability of autocorrelation is well explained by randomness (Fig. 7) and geographical patterns are hard to distinguish at the global scale. The medians of global autocorrelation for lags 1–10 yr, $\rho(1, \ldots, 10)$, are very low, for example, $\rho(1) = 0.073$ and $\rho(2) = 0.035$, and tend to be almost zero for lags larger than 6 yr, where the medians corresponding to observations are very similar to those calculated using the generated time series of white noise (Fig. 7). Despite this similarity, the hypothesis tests, the Wilcoxon–Mann–Whitney test, and the one-way analysis of variance (ANOVA) (Hollander and Wolfe 1999; Box et al. 2005) accept the null hypothesis of medians–means from data and white-noise series as being not significantly
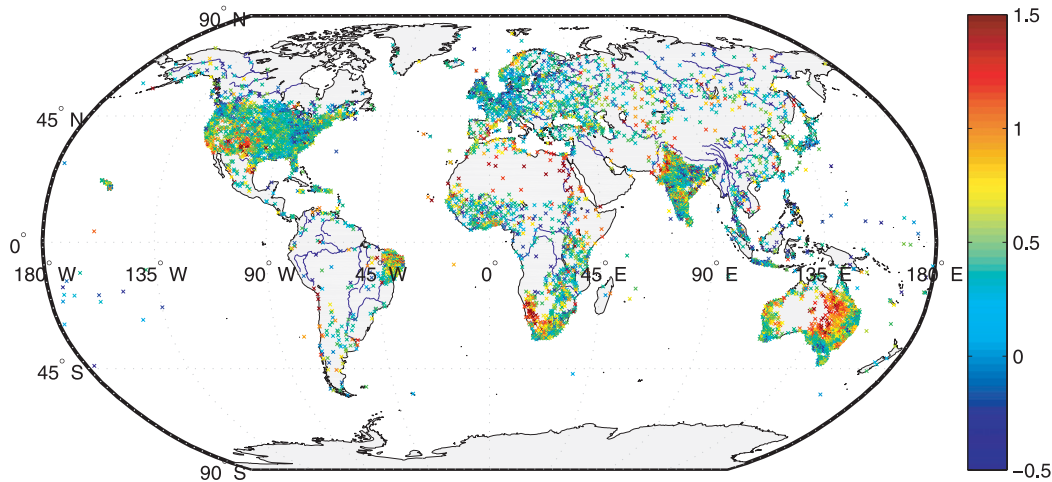
FIG. 6. The global map of γ for stations with >50 yr of observations.

different, only for lags equal to 9 and with a rather low significance level, $\alpha = 0.0001$.

Very low autocorrelation of the annual precipitation represents an interesting observation, given the fact that a rather common assumption in hydrology is that annual precipitation is correlated with the precipitation of preceding years (Thyer and Kuczera 2000; Srikanthan and McMahon 2001). While this cannot be generally excluded and stations with larger values of $\rho(1)$ can be identified in the available records, this analysis argues that this is unlikely to be a robust feature at the global land scale. Areas with a higher "concentration" of $\rho(1)$ larger than 0.25 are located in the east and west coasts of Canada, Atlantic Europe, central Russia, north Brazil, and south Saharian Africa (Fig. 8). Nonetheless, it is difficult to assert whether these outcomes are robust patterns of precipitation climatology or whether they are due to uncertainty in the estimation of $\rho(1)$.

The uncertainty of evaluation of d is even higher than for the previously discussed statistics. The analysis has therefore been limited to stations with more than 90 yr of observational records. The distribution of d is illustrated in Fig. 5b, calculated from both the observational data and the generated white-noise time series. Note that values of d outside of the stationary range, that is, 0–0.5, can be also observed in the results from the synthetic time series. The standard deviation of empirical d is quite large and is equal to 0.145, close to the estimate obtained from the white-noise series. No geographical pattern can be identified (not shown). All of the above-mentioned considerations highlight the large uncertainty in the determination of the long-range dependence and certainly question the values of d at the level of a single station. Nonetheless, calculation at the global scale is considered to be statistically significant, since random errors due to short

observational records tend to compensate in the large sample, as shown in the analysis of the white-noise time series. The median value of $d = 0.097$ is significantly larger than 0 (Wilcoxon–Mann–Whitney test, $p < 0.0001$), corroborating the notion that time series of annual precipitation show a certain degree of long memory. Furthermore, the long-range dependence coefficients d of the longest available time series tend to converge around a common value of $d = 0.1$, which is similar to the global average (Fig. 9). This supports the important role of long-term fluctuations in the occurrence of precipitation (Thyer and Kuczera 2000; New et al. 2001; Whiting et al. 2003). Such an outcome also points to the need to use statistical models that are able to reproduce long-range dependence effects, to better simulate annual precipitation time series.

b. Robustness of inferences

The effect of spatial autocorrelation of the variables in the significance of all of the detected correlations has been tested by decreasing the degrees of freedom of the t test and the F test for the null hypothesis of the absence of correlation (Clifford et al. 1989; Dutilleul 1993; Dutilleul et al. 2008). While some of the p values of the analyzed regression increase in the modified tests, all of the correlations identified as significant with the classical tests (not accounting for spatial autocorrelation) remain significant. The modified tests used a 0.05 level of significance. Therefore, the conclusion of the analysis is robust to spatial autocorrelations of the variables.

The point-scale results concerning the spatial patterns of C_v are fully supported by the gridded precipitation data (Fig. 10), with closer similarities when the higher-resolution GPCC Full Data Reanalysis product, version 5, is considered. The correlation between C_v and PCI–MAP is also confirmed and tends to be stronger when
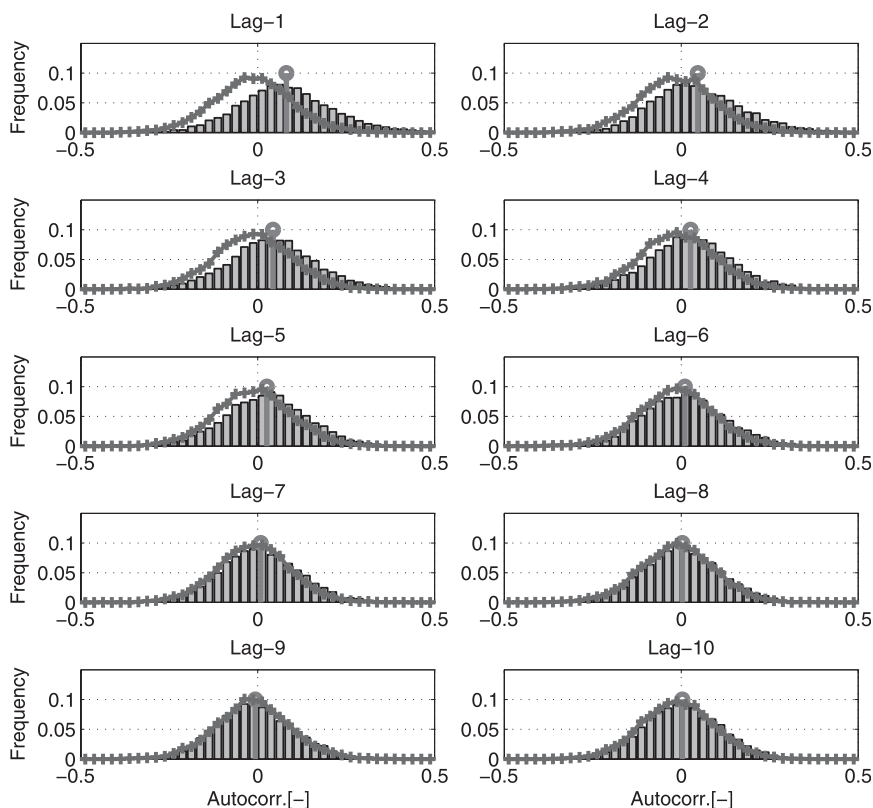
FIG. 7. The distributions of autocorrelation values of annual precipitation (bars) for $\rho(1, \ldots, 10)$, the means of the distributions (bar with circle on top), and the distributions of the auto-correlation values of white-noise time series (line with crosses) equivalent to the observed series in terms of their number and lengths. Only stations with >50 yr of data are included.

the database with longer time series are analyzed (Table 1). Specifically, the coefficients of determination of the multiple log-linear regressions of C_v with PCI and MAP are $R^2 = 0.50$, $R^2 = 0.60$, $R^2 = 0.72$, and $R^2 = 0.78$

when ERA-40, the VASClimO 50-yr product, the NCEP–NCAR reanalysis, and the GPCC Full Data Reanalysis product, respectively, are used. Only grid cells representing land were used in the computation.
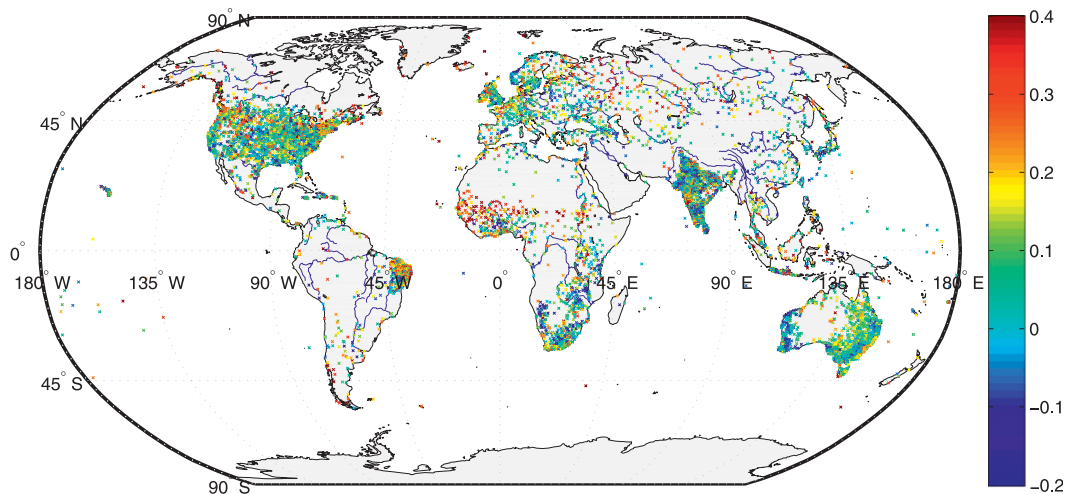

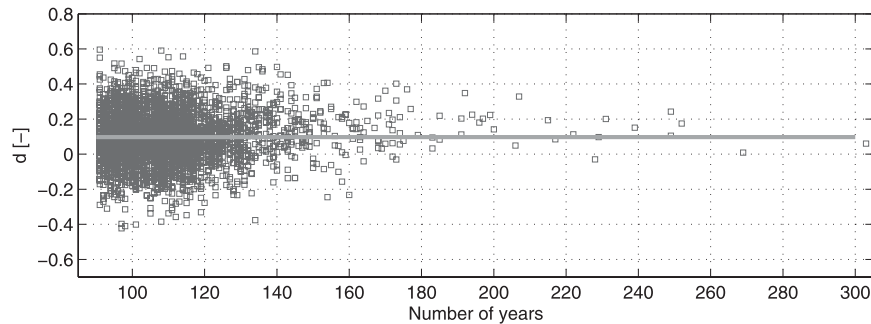FIG. 8. The global map of $\rho(1)$ for stations with >50 yr of observations.

FIG. 9. The parameter d as a function of the number of years of observational records (squares)
and average values of d (solid line).

However, significant correlations with $R^2 = 0.57$ and $R^2 = 0.67$ are also obtained for the grid cells representing sea in the NCEP–NCAR reanalysis and ERA-40 data.

A statistically significant positive skewness of the annual precipitation distribution is also corroborated by the gridded precipitation data (not shown) as well as its positive correlation with C_v ($R^2 = 0.40$–0.82) and a weak negative correlation with MAP ($R^2 = 0.05$–0.18) as reported in Table 1. The relatively larger weight of desert and semiarid areas (where gauge networks are typically sparse) in these datasets tends to increase the significance of these correlations when compared with the GHCN station data. Conversely, the very low values of autocorrelation of annual precipitation obtained using gauge data are not corroborated by the reanalysis data and by the GPCP gridded precipitation that show autocorrelations for lags 1–5 yr significantly larger than zero (especially in the driest and wettest areas of the planet). This difference can be an artifact of the reanalysis data or, most likely, it is due to considering precipitation in grid boxes of 0.5°–2° of latitude–longitude. Precipitation at such spatial scales is likely to show a rather different annual correlation structure as compared to the point-scale process (Ensor and Robeson 2008).

4. Discussion and conclusions

This study illustrates geographic patterns and statistical properties of variability of annual land precipitation at the global scale. Despite the relative simplicity of the analysis, important features of interannual variability have been identified. Specifically, the coefficient of variation of annual precipitation C_v exhibits a well-defined geographical distribution with relatively smooth transitions between the regions of large C_v and regions with low interannual variability.

An important statistical linkage of C_v with MAP and intra-annual precipitation seasonality ($R^2 = 0.74$) has been detected, indicating that the magnitude of interannual variability is related to seasonality and not only to MAP, as previously observed. This can have important consequences for estimating C_v when only short observational records are available.

The estimation of the skewness, the autocorrelation, and the long-memory parameter of annual precipitation is complicated by the relatively short length of available time series. Nonetheless, γ tends to be significantly larger than zero and is positively correlated with C_v. The short-range autocorrelation of the precipitation process is relatively weak worldwide. Conversely, the intensity of the long-range dependence expressed with the long-memory statistic d is significantly different from zero, pointing to an important role of the long-term persistence in the annual precipitation process.

Important variables that can influence interannual variability, such as topographic or local geographic effects (e.g., proximity to sea, orographic precipitation, predominant atmospheric fluxes), cannot be identified with this analysis. Furthermore, no attempt has been made to include indices of atmospheric circulation and no climatological explanations have been sought. The focus of the current analysis is to integrate the current knowledge on interannual variability and provide a comprehensive picture of this phenomenon using "in situ" information, that is, using gauge observations only. A similar statistical analysis using gridded reanalysis data supported all of the major findings of this study, except for those that concern annual autocorrelation. It also supports the fact that the length of the time series influences the estimation of the annual precipitation statistics. Therefore, the principal conclusions of the study can be considered to be essentially unaffected by the spatial irregularity in the distribution of gauge stations.

The results of this study can provide a useful guidance for further studies that aim to characterize the interannual variability of precipitation. For instance, they can be used to validate the results of general circulation models in terms of long-term precipitation variability
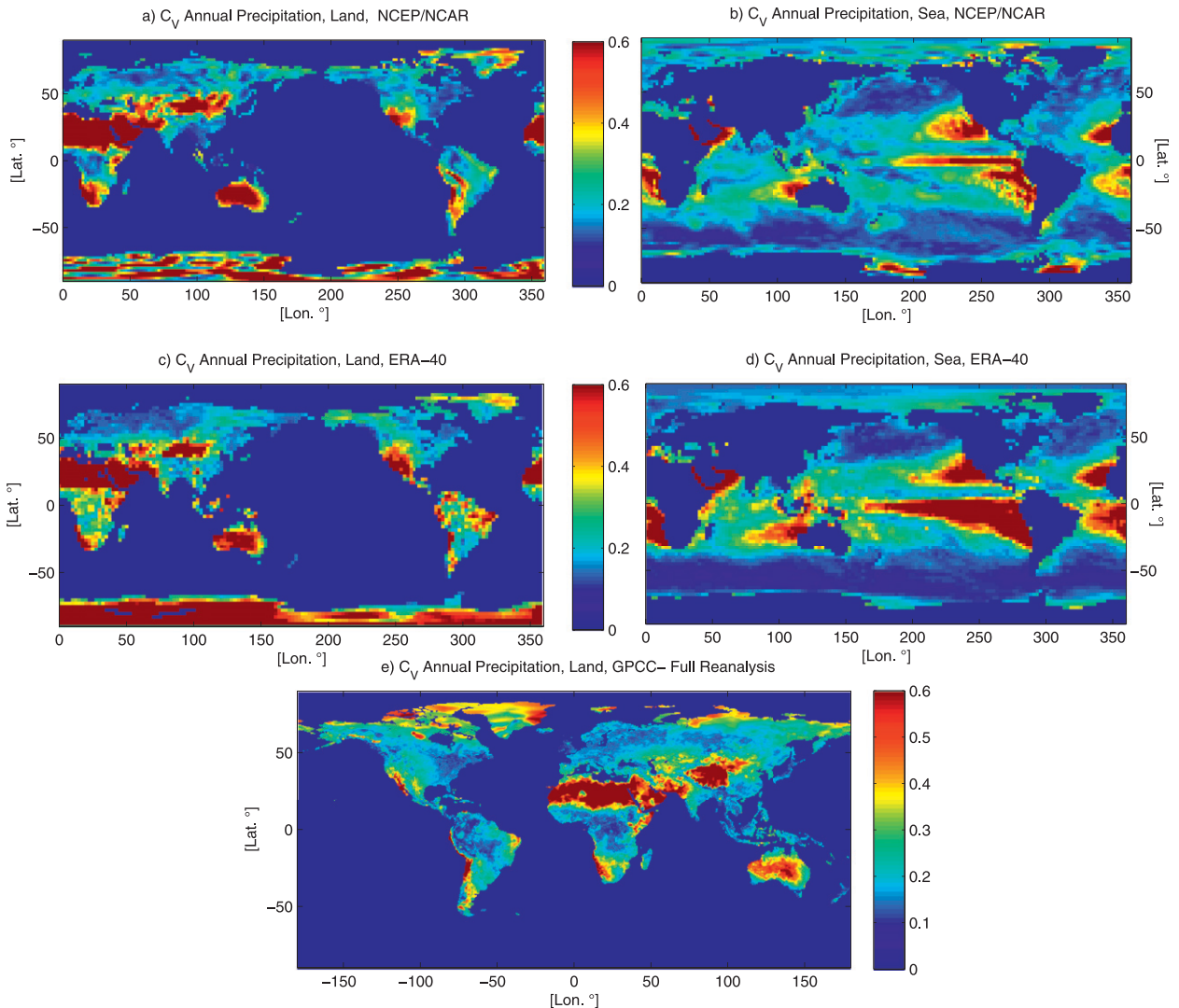
FIG. 10. The global map of C_v derived from (a) NCEP–NCAR reanalysis data land grid cells, (b) NCEP–NCAR reanalysis data sea grid cells, (c) ERA-40 data land grid cells, (b) ERA-40 data sea grid cells, and (e) GPCC Full Data Reanalysis product, version 5.

(Dai 2006) as well as be used as a starting point for analyses aiming to find causal linkages and physical controls driving the observed patterns. Finally, the presented results can have a notable importance for local analyses that need to simulate a realistic annual precipitation variability as input into other earth system components, for example, hydrological, ecological, and geomorphological studies.

APPENDIX

Indices of Seasonality

The PCI (De Lùis et al. 2000; Fatichi and Caporali 2009) for a given year "$_{yr}$" is calculated on the basis of

annual precipitation P_{yr} (mm) and monthly precipitation for a month i, $P_m(i)$ (mm) as follows:

$$\text{PCI} = \frac{100 \sum_{i}^{12} P_m(i)^2}{P_{yr}^2}. \tag{A1}$$

The SI (Walsh and Lawler 1981; Pryor and Schoof 2008) for a given year is calculated as

$$\text{SI} = \frac{\sum_{i}^{12} \left| P_m(i) - \frac{P_{yr}}{12} \right|}{P_{yr}}. \tag{A2}$$

The SCI (Fujita 2008) for a given year is the ratio between the standard deviation σ_{P_m} and the mean of monthly precipitation as follows:

$$\text{SCI} = \frac{\sigma_{P_m}}{\frac{P_{yr}}{12}}. \tag{A3}$$

The M_r for a given year, as defined from Davidowitz (2002) using Levene's statistic, is calculated as

$$M_r = \frac{\sum_{i}^{12} |\log_{10} P_m(i) - \text{Median}[\log_{10} P_m(i)]|}{12}, \tag{A4}$$

where Median is the median operator and months with zero precipitation are replaced by 0.01 (mm), as suggested by Davidowitz (2002), to avoid $-\infty$ values.

The long-term values of the indices are successively calculated for each station averaging the yearly values. The standard deviation of the PCI for each station has also been computed to evaluate the corresponding coefficient of variation $C_{v,\text{PCI}}$.

REFERENCES

Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) monthly precipitation data analysis (1979–present). *J. Hydrometeor.,* **4,** 1147–1167.

Alexandersson, H., 1986: A homogeneity test applied to precipitation data. *J. Climatol.,* **6,** 661–675.

Bartolini, E., P. Claps, and P. D'Odorico, 2009: Interannual variability of winter precipitation in the European Alps: Relations with the North Atlantic Oscillation. *Hydrol. Earth Syst. Sci.,* **13,** 17–25.

Beck, C., J. Grieser, and B. Rudolf, 2005: A new monthly precipitation climatology for the global land areas for the period 1951 to 2000. Climate Status Report 2004, DWD Rep. Klimastatusbericht 2004, 181–190.

Beran, J., 1994: *Statistics for Long-Memory Processes. Monogr. Stat. Appl. Probab.,* Vol. 61, Chapman and Hall/CRC, 315 pp.

Box, G. E. P., J. S. Hunter, and W. G. Hunter, 2005: *Statistics for Experimenters: Design, Innovation, and Discovery.* 2nd ed. Wiley Series in Probability and Statistics, Vol. 559, Wiley, 664 pp.

Caballero, R., S. Jewson, and A. Brix, 2002: Long memory in surface air temperature: Detection, modeling, and application to weather derivative valuation. *Climate Res.,* **21,** 127–140.

Cayan, D., M. D. Dettinger, H. F. Diaz, and N. E. Graham, 1998: Decadal variability of precipitation over western North America. *J. Climate,* **11,** 3148–3165.

Chen, M., P. Xie, J. E. Janowiak, and P. A. Arkin, 2002: Global land precipitation: A 50-yr monthly analysis based on gauge observations. *J. Hydrometeor.,* **3,** 249–266.

Clifford, P., S. Richardson, and D. Hémon, 1989: Assessing the significance of the correlation between two spatial processes. *Biometrics,* **45,** 123–134.

Coulibaly, P., 2006: Spatial and temporal variability of Canadian seasonal precipitation (1900–2000). *Adv. Water Resour.,* **29,** 1846–1865, doi:10.1016/j.advwatres.2005.12.013.

Dai, A., 2006: Precipitation characteristics in eighteen coupled climate models. *J. Climate,* **19,** 4605–4630.

——, I. Y. Fung, and A. D. DelGenio, 1997: Surface observed global land precipitation variations during 1900–88. *J. Climate,* **10,** 2943–2962.

Davidowitz, G., 2002: Does precipitation variability increase from mesic to xeric biomes? *Global Ecol. Biogeogr.,* **11,** 143–154.

De Luís, M., J. Raventós, J. C. González-Hidalgo, J. R. Sánchez, and J. Cortina, 2000: Spatial analysis of rainfall trends in the region of Valencia (east Spain). *Int. J. Climatol.,* **20,** 1451–1469.

Dutilleul, P., 1993: Modifying the *t* test for assessing the correlation between two spatial processes. *Biometrics,* **49,** 305–314.

——, B. Pelletier, and G. Alpargu, 2008: Modified *F* tests for assessing the multiple correlation between one spatial process and several others. *J. Stat. Plann. Inference,* **138,** 1402–1415.

Ensor, L. A., and S. M. Robeson, 2008: Statistical characteristics of daily precipitation: Comparisons of gridded and point datasets. *J. Appl. Meteor. Climatol.,* **47,** 2468–2476.

Fang, J., S. Piao, Z. Tang, C. Peng, and W. Ji, 2001: Interannual variability in net primary production and precipitation. *Science,* **293,** 1723.

Fatichi, S., and E. Caporali, 2009: A comprehensive analysis of changes in precipitation regime in Tuscany. *Int. J. Climatol.,* **29,** 1883–1893, doi:10.1002/joc.1921.

——, S. M. Barbosa, E. Caporali, and M. E. Silva, 2009: Deterministic versus stochastic trends: Detection and challenges. *J. Geophys. Res.,* **114,** D18121, doi:10.1029/2009JD011960.

Fujita, K., 2008: Effect of precipitation seasonality on climatic sensitivity of glacier mass balance. *Earth Planet. Sci. Lett.,* **276,** 14–19, doi:10.1016/j.epsl.2008.08.028.

Grimm, A. M., and A. A. Natori, 2006: Climate change and interannual variability of precipitation in South America. *Geophys. Res. Lett.,* **33,** L19706, doi:10.1029/2006GL026821.

Gu, G., and R. F. Adler, 2004: Seasonal evolution and variability associated with the West African monsoon system. *J. Climate,* **17,** 3364–3377.

——, and ——, 2006: Interannual rainfall variability in the tropical Atlantic region. *J. Geophys. Res.,* **111,** D02106, doi:10.1029/2005JD005944.

——, and ——, 2009: Interannual variability of boreal summer rainfall in the equatorial Atlantic. *Int. J. Climatol.,* **29,** 175–184.

——, and ——, 2011: Precipitation and temperature variations on the interannual time scale: Assessing the impact of ENSO and volcanic eruptions. *J. Climate,* **24,** 2258–2270.

——, ——, G. J. Huffman, and S. Curtis, 2007: Tropical rainfall variability on interannual-to-interdecadal and longer time scales derived from the GPCP monthly product. *J. Climate,* **20,** 4033–4046.

Higgins, R. W., K. C. Mo, and Y. Yao, 1998: Interannual variability of the U.S. summer precipitation regime with emphasis on the southwestern monsoon. *J. Climate,* **11,** 2582–2606.

Hollander, M., and D. A. Wolfe, 1999: *Nonparametric Statistical Methods.* 2nd ed. John Wiley & Sons, Inc., 787 pp.

Huffman, G. J., and Coauthors, 1997: The Global Precipitation Climatology Project (GPCP) combined precipitation dataset. *Bull. Amer. Meteor. Soc.,* **78,** 5–20.

——, R. F. Adler, D. T. Bolvin, and G. Gu, 2009: Improving the global precipitation record: GPCP version 2.1. *Geophys. Res. Lett.,* **36,** L17808, doi:10.1029/2009GL040000.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

Knapp, A. K., and M. D. Smith, 2001: Variation among biomes in temporal dynamics of aboveground primary production. *Science,* **291,** 481–484.

Koster, R. D., M. J. Suarez, and M. Heiser, 2000: Variance and predictability of precipitation at seasonal-to-interannual timescales. *J. Hydrometeor.,* **1,** 26–46.

Koutsoyiannis, D., 2003: Climate change, the Hurst phenomenon, and hydrological statistics. *Hydrol. Sci. J.,* **48,** 3–24.

Legates, D. R., 1991: An evaluation of procedures to estimate monthly precipitation probabilities. *J. Hydrol.,* **122,** 129–140.

Liebhold, A. M., and A. A. Sharov, 1998: Testing for correlation in the presence of spatial autocorrelation in insect count data. *Population and Community Ecology for Insect Management and Conservation,* J. Baumgärtner, P. Brandmayr, and B. F. J. Manly, Eds., A. A. Balkema, 111–117.

Mason, S. J., and L. Goddard, 2001: Probabilistic precipitation anomalies associated with ENSO. *Bull. Amer. Meteor. Soc.,* **82,** 619–638.

Medvigy, D., R. L. Walko, and R. Avissar, 2008: Modeling interannual variability of the Amazon hydroclimate. *Geophys. Res. Lett.,* **35,** L15817, doi:10.1029/2008GL034941.

Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.,* **25,** 693–712, doi:10.1002/joc.1181.

Mourato, S., M. Moreira, and J. Corte-Real, 2010: Interannual variability of precipitation distribution patterns in southern Portugal. *Int. J. Climatol.,* **30,** 1784–1794, doi:10.1002/joc.2021.

New, M., M. Todd, M. Hulme, and P. Jones, 2001: Precipitation measurements and trends in the twentieth century. *Int. J. Climatol.,* **21,** 1899–1922, doi:10.1002/joc.680.

Peterson, T. C., and R. S. Vose, 1997: An overview of the Global Historical Climatology Network temperature database. *Bull. Amer. Meteor. Soc.,* **78,** 2837–2849.

——, and Coauthors, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol.,* **18,** 1493–1517.

Pryor, S. C., and J. T. Schoof, 2008: Changes in the seasonality of precipitation over the contiguous USA. *J. Geophys. Res.,* **113,** D21108, doi:10.1029/2008JD010251.

Rodriguez-Puebla, C., A. H. Encinas, S. Nieto, and J. Garmendia, 1998: Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula. *Int. J. Climatol.,* **18,** 299–316.

Rudolf, B., 2005: Global precipitation analysis products of the GPCC. Climate Status Report 2004, DWD Rep. Klilmastatusbericht 2004, 163–170.

——, and U. Schneider, 2005: Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations. *Proc. Second Workshop of the Int. Precipitation Working Group,* Monterey, CA, Naval Research Laboratory, 231–247.

——, A. Becker, U. Schneider, A. Meyer-Christoffer, and M. Ziese, 2010: The new "GPCC Full Data Reanalysis version 5" providing high-quality gridded monthly precipitation data for the global land-surface is public available since December 2010. GPCC Status Rep., DWD Tech. Rep., 7 pp.

Sardeshmukh, P. D., G. P. Compo, and C. Penland, 2000: Changes of probability associated with El Niño. *J. Climate,* **13,** 4268–4286.

Small, D., and S. Islam, 2008: Low frequency variability in fall precipitation across the United States. *Water Resour. Res.,* **44,** W04426, doi:10.1029/2006WR005623.

Srikanthan, R., and T. A. McMahon, 1982: Simulation of annual and monthly rainfalls—A preliminary study at five Australian stations. *J. Appl. Meteor.,* **21,** 1472–1479.

——, and ——, 2001: Stochastic generation of annual, monthly and daily climate data: A review. *Hydrol. Earth Syst. Sci.,* **5,** 653–670.

Taqqu, M. S., V. Teverovsky, and W. Willinger, 1995: Estimators for long-range dependence: An empirical study. *Fractals,* **3,** 785–798.

Thyer, M. A., and G. A. Kuczera, 2000: Modelling long-term persistence in hydro-climatic time series using a hidden state Markov model. *Water Resour. Res.,* **36,** 3301–3310.

Uppala, S. M., and Coauthors, 2005: The ERA-40 Re-Analysis. *Quart. J. Roy. Meteor. Soc.,* **131,** 2961–3012.

Walsh, P. D., and D. M. Lawler, 1981: Rainfall seasonality: Description, spatial patterns and changes through time. *Weather,* **36,** 201–208.

Whiting, J. P., M. F. Lambert, and A. V. Metcalfe, 2003: Modelling persistence in annual Australia point rainfall. *Hydrol. Earth Syst. Sci.,* **7,** 197–211.

Wiegand, T., H. A. Snyman, K. Kellner, and J. M. Paruelo, 2004: Do grasslands have a memory: Modeling phytomass production of a semiarid South African grassland. *Ecosystems,* **7,** 243–258, doi:10.1007/s10021-003-0235-8.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2nd ed. Elsevier, 627 pp.

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.,* **78,** 2539–2558.

Yang, Y., J. Fang, W. Ma, and W. Wang, 2008: Relationship between variability in aboveground net primary production and precipitation in global grasslands. *Geophys. Res. Lett.,* **35,** L23710, doi:10.1029/2008GL035408.