



Spatial random downscaling of rainfall signals in Andean heterogeneous terrain

A. Posadas^{1,2}, L. A. Duffaut Espinosa^{1,3}, C. Yarlequé⁴, M. Carbajal¹, H. Heidinger⁵, L. Carvalho⁵, C. Jones⁵, and R. Quiroz¹

¹Production Systems and the Environment Division, International Potato Center (CIP), Lima, Peru

²World Agroforestry Centre (ICRAF), Nairobi, Kenya

³Electrical and Computer Engineering Department, George Mason University, Virginia, USA

⁴Department of Atmospheric and Environmental Sciences, University at Albany, State University of New York, Albany, USA

⁵Earth Research Institute (ERI), University of California Santa Barbara, Santa Barbara, USA

Correspondence to: L. A. Duffaut Espinosa (l.duffautespinosa@gmail.com)

Received: 01 December 2013 – Revised: 09 May 2015 – Accepted: 18 June 2015 – Published: 16 July 2015

Abstract. Remotely sensed data are often used as proxies for indirect precipitation measures over data-scarce and complex-terrain areas such as the Peruvian Andes. Although this information might be appropriate for some research requirements, the extent at which local sites could be related to such information is very limited because of the resolution of the available satellite data. Downscaling techniques are used to bridge the gap between what climate modelers (global and regional) are able to provide and what decision-makers require (local). Precipitation downscaling improves the poor local representation of satellite data and helps end-users acquire more accurate estimates of water availability. Thus, a multifractal downscaling technique complemented by a heterogeneity filter was applied to TRMM (Tropical Rainfall Measuring Mission) 3B42 gridded data (spatial resolution ~ 28 km) from the Peruvian Andean high plateau or *Altiplano* to generate downscaled rainfall fields that are relevant at an agricultural scale (spatial resolution ~ 1 km).

precipitation in the Andes, mainly due to little spatial coherence (Vuille et al., 2003). This uncertainty in precipitation patterns is recurrent in projected climate change scenarios (Thibeault et al., 2011) and therefore studying the behavior of precipitation in this region is crucial.

In the last 20 years, satellite missions, such as the Tropical Rainfall Measuring Mission (TRMM), were developed with the objective of measuring agroclimatic events around of the planet. TRMM is able to measure rainfall intensities in places of scarce local information. Due to their global nature, these missions generate data at very low resolution. The coarse resolution represents poorly rainfall variation in areas with complex terrain. In places such as the Peruvian Andes, better resolution information is either scarce, too expensive or nonexistent. Hence, the resolution of the available information is not high enough to assess climate change variability on smallholder agriculture and thus new approaches to further *downscaling* the data to model crop yields are required. Moreover, in the case of using global circulation models (GCM), the computational power together with parameterizations necessary to simulate processes at small scales may introduce nonlinear errors and biases in the simulated variables (Baron et al., 2005). Therefore, an alternative to increasing the scales is by performing downscaling of the variables involved such that relevant information can be obtained at the smallest usable scale.

This paper addresses a statistical downscaling model involving a cascade disaggregation process of satellite precipitation data. The downscaling procedure is based on the con-

1 Introduction

The Andes is the longest and highest mountain range in the tropics and a suitable site to study precipitation patterns in relation to complex topography. Assessing the variability in precipitation plays a key role in the sustainability of agriculture, which is one of the most important human activities in the Andes (Sanabria et al., 2009; Vuille, 2011; Garreaud, 2000). There is no clear pattern of decreasing or increasing

cept of multiplicative cascades. A general cascade process assumes that data on a given area is available at a coarse resolution of l_0 km over some interval of time (Fig. 1). However, a more realistic assessment of the information requires a resolution of l_r km (local scale), where $l_0 \gg l_r$. Initially, an arbitrary value ρ_0 , corresponding to the height of the box at the top of Fig. 1, is distributed uniformly on an area of l_0^2 km². Then, to increase the spatial resolution, a statistical model is used to generate a set of weights whose values are used to disaggregate ρ_0 into subdivisions with area $l_1^2 = l_0^2/b$ km², where b is the number of subdivisions (*branching parameter*). The random variable producing the weights at each level is usually called the *generator* of the cascade. The procedure is repeated until the desired resolution is reached at level r , where the area of each subdivision is $l_r^2 = l_0^2/b^r$ km². An important requirement of the cascade process is that the values at level $k - 1$ must be preserved by the corresponding means at level k . Thus, only the spatial distribution of the measured precipitation changes and therefore provides a realistic rainfall distribution at a higher spatial resolution.

One way to infer the parameters of the generator is by characterizing its behavior at different scales based in a discrete form of multiplicative random cascades having the ability to generate rainy and non-rainy outcomes (Schertzer and Lovejoy, 1987; Over and Gupta, 1994). An advantage of this approach lies in that it describes the complex process of rainfall precipitation in a wide range of scales using very few parameters. The evidence concerning the multiple scaling or multifractal behavior of the rainfall fields has been well documented in the scientific literature (Schertzer and Lovejoy, 1989; Over and Gupta, 1996; Perica and Foufoula-Georgiou, 1996; Deidda et al., 1999; Stolle et al., 2009; Lovejoy et al., 2012). Multifractal applications to rainfall includes multifractal objective analysis, statistics of extreme values, multifractal modeling, space–time transformations, the multifractal radar observer problem, stratification, and texture of rain. These have been discussed at large in Lovejoy and Schertzer (1995, 2013). The rainfall modeling based on discrete multiplicative random cascades has been tested for spatial and temporal disaggregation under different climate conditions with the conclusion that it is possible to capture rainfall variability of subgrid scale using this multifractal approach (Olsson, 1998; Güntner et al., 2001; Molnar and Burlando, 2005). Moreover, there has been some research in which spatial heterogeneity in the random cascade process was incorporated in different manners (Jothityangkoon et al., 2000; Pathirana and Herath, 2002; Tachicawa et al., 2004; Sharma et al., 2007). For instance, the application of a pointwise filtering approach, which accounts for the local details at the desired final resolution.

The main goal of the paper is to obtain a reliable representation of the spatial variability of rainfall intensities by applying the multifractal downscaling technique over the Altiplano, a region that is characterized by its complex terrain and a relatively high spatial (and temporal) coverage of

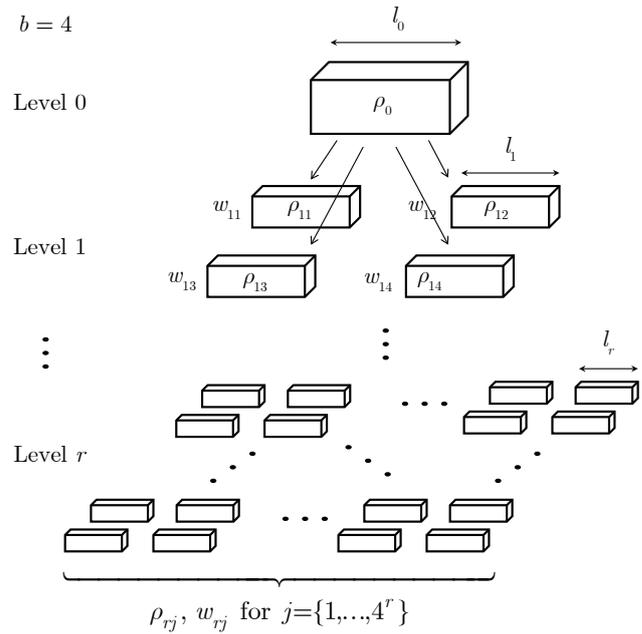


Figure 1. Cascade procedure.

weather stations (compared to other regions in Peru). Specifically, this study characterized the consistency of physical parameters in conditions where topography and spatial distribution of precipitation are moderately heterogeneous, and applies a lognormal disaggregation of rainfall measurements via the generation of a multifractal random cascade model by using the Mandelbrot–Kahane–Peyriere (MKP) function to characterize the sample moments and scalability of the process.

The paper is organized as follows. In Sect. 2, a description of the studied region, the TRMM data and gauge rainfall data are provided. A brief description of the correction of TRMM data and a simple assessment of its lognormality are also given in this section. The multifractal downscaling procedure is described in Sect. 3. This includes the description of the disaggregation model and the introduction of the information heterogeneity in the method. In particular, TRMM heterogeneity and local heterogeneity produced using standard techniques in hydrology is presented and discussed. Section 4 shows the results of the application of the multifractal downscaling method on the Andean high plateau and its vicinity, using TRMM rainfall data. Also, the quality of the generator is assessed on a spatial and temporal basis. Finally, Sect. 5 gives some conclusions and future research comments.

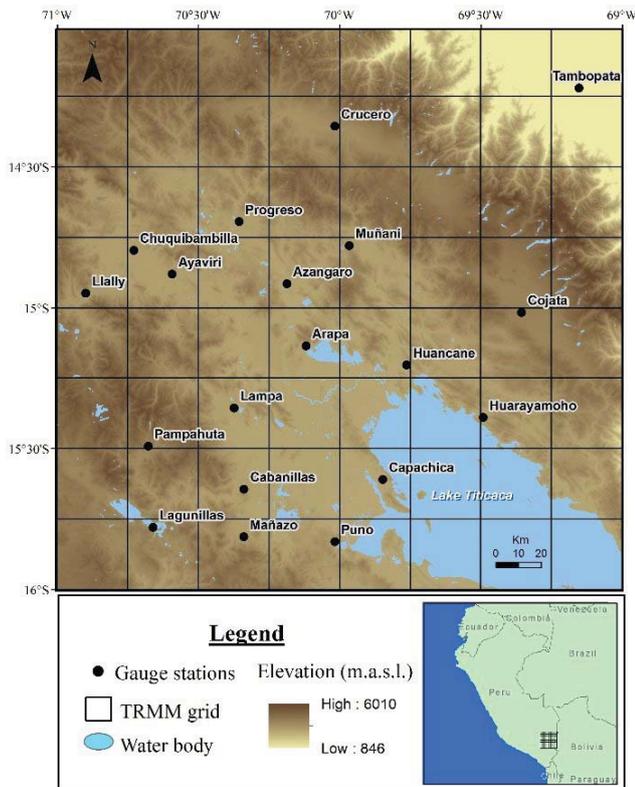


Figure 2. Study region: stations (black dots) and TRMM 8×8 cell grid (black line).

2 Materials

2.1 Study area

The studied region is defined by 8×8 TRMM cells from the southern Peruvian Andes (Fig. 2), with a spatial resolution of approximately 28 km. The majority of these TRMM cells are placed in the Peruvian Andean high plateau or Altiplano and a few cells over the east side of the Andes. Geographical coordinates of the study area are between latitudes 14 and 16° S and longitudes 69 and 71° W constituting an area of approximately $225 \times 225 \text{ km}^2$. The altitudes range between 800 and 6500 m a.s.l., approximately. The annual rainfall varies, on average, from $\sim 250 \text{ mm}$ in the arid southwest to $\sim 5000 \text{ mm}$ in the Amazon basin at the northeast corner of the study site (Garreaud et al., 2003). The year to year variability with respect to extreme phenomena, such as *El Niño*, is mitigated by the Andes and therefore the region's climate can be considered as yearly invariant (Garreaud et al., 2003). The rainfall over the Altiplano is largely concentrated in the austral summer months, when more than 70 % of the precipitation occurs from December to February. On all timescales, the climatic conditions on the Altiplano are closely related to the upper-air circulation, with an easterly zonal flow aloft favoring wet conditions and westerly flow causing dry conditions (Garreaud et al., 2003). At a diurnal timescale, convective clouds

are particularly common during the afternoon and evening (Vera et al., 2006), and are produced by the insolation-driven surface heating and the consequent destabilization of the local lower troposphere (Garreaud et al., 2003). At intraseasonal timescales, within the rainy season, the Altiplano experiences rainy and dry episodes lasting between 5 and 10 days (Aceituno and Montecinos, 1993), as a reflection of the position and intensity of the Bolivian High (Lenters and Cook, 1999), which is modulated by Rossby waves emanating from the midlatitude South Pacific. Interannual variability is primarily related to changes in the mean zonal flow over the Altiplano, reflecting changes in meridional baroclinicity between tropical and subtropical latitudes, which in turn is a response to sea-surface temperature changes in the tropical Pacific (Garreaud et al., 2003).

In this area of study, there are 19 stations (black dots in Fig. 2) from which observed rainfall measurements are obtained during a period of 8 years (from 1999 to 2006). A tipping bucket rain gauge was used to obtain these on-site rainfall measurements (Garreaud et al., 2003). The role of the data from the stations in this paper is twofold. The first one is to validate the generation of downscaled rainfall data, and the second is to aid in the generation of the deterministic local heterogeneity; see Sect. 3.2.

2.2 TRMM rainfall data

The Tropical Rainfall Measuring Mission is a joint space mission between the National Aeronautics and Space Administration (NASA) and Japan's National Space Development Agency (NASDA) designed to monitor and study tropical and subtropical precipitation and the associated release of energy. The primary rainfall instruments on TRMM are the TRMM Microwave Imager (TMI), the PR (polarization radiometers) and the Visible and Infrared Radiometer System (VIRS). In addition, TRMM satellites carry two related Earth Observing System (EOS) instruments: the Clouds and the Earth's Radiant Energy System (CERES) and the Lightning Imaging Sensor (LIS) (Kummerow et al., 1998). The TMI is the main instrument used for precipitation; see <http://climatedataguide.ucar.edu/guidance/trmm-tropical-rainfall-measuring-mission> for additional information about the TRMM mission. Our interest is in daily TRMM 3B42 v7 grid data. The information was downloaded directly from http://iridl.ldeo.columbia.edu/SOURCES/.NASA/.GES-DAAC/.TRMM_L3/.TRMM_3B42/.v7/.daily/ for the period of 1 January 1999–31 December 2006. TRMM 3B42 v7 is a 3 h average, extrapolated to a daily temporal scale, 0.25° spatial resolution ($\sim 27.8 \text{ km}$) TRMM product derived from calibrated geosynchronous IR imagery merged with TRMM and other satellite data. Hereafter, TRMM 3B42 v7 will be called simply TRMM. TRMM rainfall (mm) is distributed uniformly in each grid cell. A total of 8×8 TRMM cells cover the study area with rainfall

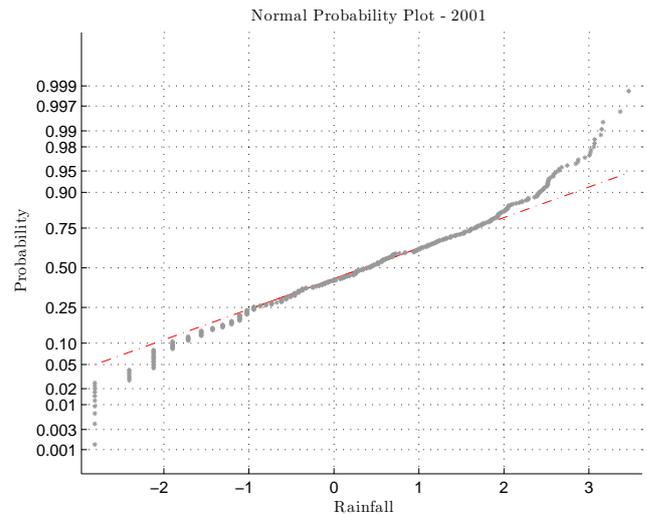
Table 1. Weather station locations and altitudes.

Weather station	Longitude (degrees)	Latitude (degrees)	Altitude (m a.s.l.)
Arapa	-70.12	-15.14	3920
Ayaviri	-70.59	-14.88	3920
Azángaro	-70.19	-14.91	3863
Cabanillas	-70.35	-15.64	3890
Capachica	-69.84	-15.62	3819
Chuquibambilla	-70.73	-14.80	3910
Cojata	-69.36	-15.02	4344
Crucero Alto	-70.02	-14.36	4130
Huancané	-69.76	-15.20	3860
Huaraya Moho	-69.49	-15.39	3890
Lagunillas	-70.66	-15.77	4250
Lampa	-70.37	-15.36	3900
Llally	-70.90	-14.95	4111
Mañazo	-70.34	-15.81	3942
Muñani	-69.97	-14.78	4119
Pampahuta	-70.68	-15.49	4320
Progreso	-70.36	-14.69	3965
Puno	-70.02	-15.82	3840
Tambopata	-69.15	-14.22	1340

between 450 and 2000 mm yr⁻¹. Although TRMM tends to underestimate rainfall in terrain with complex topography, we used the correction protocol described in Heidinger et al. (2012). To correct TRMM daily rainfall over the Andean plateau, the procedure in Heidinger et al. (2012) incorporates the high-frequency component (detail or noise) from rain gauge signals into the low-frequency component (tendency or base) derived from TRMM using wavelet multiresolution analysis (MRA). For each TRMM cell, the high-frequency component of the closest meteorological station was added to the low-frequency component of TRMM. The MRA reconstruction was started at the third level of the Haar wavelet decomposition of TRMM and rain gauge signals. It is important to emphasize that the main reason for selecting the study site was the rainfall heterogeneity found in that small area including arid, semi-arid, and humid tropics and the availability of a rain gauges (Table 1)¹. As mentioned in the introduction, precipitation data is not widely available in this region and therefore missions such as TRMM become essential as the starting point of downscaling techniques.

Another fact about the corrected TRMM information is that its spatial distribution fits approximately a lognormal distribution, which can be confirmed by a simple lognormality test. Such test is applied to the corrected TRMM by eliminating the zero intensity rainfall and computing its logarithm. As a result, the logarithm of the information mostly

¹The Tambopata station was not used to correct TRMM rainfall information since there were not enough neighboring stations to accurately represent the details in the region surrounding the station (Heidinger et al., 2012).

**Figure 3.** Lognormal test of TRMM data for 15 April 2003.

agrees with a normal distribution and thus any test for normality should give the desired answer. This procedure must be applied on a daily basis since the parameters of the distribution vary with time. Figure 3 exemplifies the method for 15 April 2003 and provides evidence of lognormality. The lognormality assumption is also supported, for example, by Lin (1978) and Kedem and Long (1987). The interested reader can find another perspective on the nature of satellite information and its relationship to the cascade structure of precipitation in Lovejoy et al. (2012).

As the reader can observe in Fig. 3, there are some discrepancies with respect to the high values of rainfall intensity. There are several factors that may contribute to such discrepancies. First, the northeast corner of the studied region lacks enough meteorological stations to correctly perform the MRA correction of TRMM. That region is in the rain forest and its rainfall variability differs greatly with respect to the rest of the studied area (there is an abrupt change of topography, also the Andes is a high elevation area whereas the rain forest is at a low elevation). Also, each snapshot is comprised of 64 pixels from which the zero values must be removed to perform the test due to the fact that zero is not included in the domain of lognormal distributions. Even though this is a small sample of points for applying an statistical test, this very simple test shows a good relationship between the rainfall field and a lognormal distribution, but we argue that there are not enough extreme values (corresponding to the tail of the distribution), which contributes to the discrepancies shown in Fig. 3.

3 Methods

3.1 Discrete disaggregation model

To describe a general disaggregation model, it is first assumed that the initial information, at level 0, is comprised by only one value of TRMM rainfall rate (Fig. 1). That is, the rainfall total rate ρ_0 is uniformly spread over the initial area l_0^2 . At level 1, the initial area is partitioned into four identical boxes of length $l_0/2^1$. For subdivision i at level n , a part of the initial rainfall rate ρ_0 , denoted as ρ_{ni} , is uniformly distributed over an area of $l_0^2/4^n$, and is expressed in terms of ρ_0 as

$$\rho_n^i = \rho_0 \prod_{j=1}^n w_{i,j},$$

where w_{ij} are the weights at which the rainfall rate is disaggregated throughout the cascade process. In particular, these weights are outcomes of a random variable W that follows some distribution law and satisfies $E[W] = 1$, where $E[\cdot]$ denotes the expected value of a random variable. This condition also means that if n tends to infinity then ρ_∞ will exist and not be degenerated; i.e., $\rho_\infty \neq 0$. The procedure is repeated until the n th level is reached. After n cascade steps, the original information is disaggregated into 4^n pieces.

A good disaggregation model is necessary for a realistic and accurate distribution of the information. This is done by describing the weights w in a probabilistic manner. The class of log-stable random variables constitutes a broad class of generators used for this purpose and therefore characterizes the class of multifractal fields that can be generated; see Royer et al. (2008) for an accessible introduction to general multifractal analysis. In particular, such a log-stable generator has the form

$$Y = b^{\gamma + \sigma X},$$

where X is a stable (Lévy) random variable (Over, 1995; Schertzer and Lovejoy, 1987). The goal of this paper is to utilize a particular type of such generators. That is, those whose generator is lognormal. In the language of log-stable generators, this corresponds to those characterized by the stability index $\alpha = 2$ (Samorodnitsky and Taqqu, 1994). In order to preserve $E[Y] = 1$, it follows that

$$\begin{aligned} E[Y] &= E \left[b^{\gamma + \sigma X} \right] \\ &= b^\gamma E \left[b^{\sigma X} \right] \\ &= b^\gamma E \left[e^{\sigma X \ln(b)} \right], \end{aligned}$$

where the expectation in the right hand side is simply the moment generating function of an unitary normal distribution. Thus, $e^{\gamma \ln(b) + \frac{1}{2} \sigma^2 \ln(b)^2} = 1$, which implies $\gamma = -\frac{\sigma^2 \ln(b)}{2}$ and

$$Y = b^{-\frac{\sigma^2 \ln(b)}{2} + \sigma X}.$$

It is obvious that the generator Y is not able to generate zeros as expected from a precipitation simulation. In this regard, a more realistic generator, W , is given by including an *atom at zero* in a composition manner (representing the presence and absence of data). That is,

$$W = BY, \tag{1}$$

where

$$P[B = 0] = 1 - b^{-\beta} \text{ and } P[B = b^\beta] = b^{-\beta} \tag{2}$$

and is known as the β model. This composition results in the β -lognormal model specified by

$$P[W = 0] = 1 - b^{-\beta} \text{ and } P \left[W = b^{\beta - \sigma^2 \frac{\log(b)}{2} + \sigma X} \right] = b^{-\beta}, \tag{3}$$

where X is a standard normal variable, and β and σ fully describe the generator W probability distribution (Over and Gupta, 1994, 1996; Sharma et al., 2007; Kang and Ramírez, 2010; Pathirana and Herath, 2002). Other models, in one dimension, that introduce a measure of zero rainfall can be found in Gires et al. (2013) and Schmitt et al. (1998). Since β and σ completely characterize the downscaling process, the key for the generation of a downscaling rainfall model is to estimate these parameters via a scale invariant method. In Over (1995), Pathirana and Herath (2002) and Sharma et al. (2007), it was shown that the parameters β and σ can be easily computed using multifractal sample moments obtained directly from rainfall information. Specifically, the *multifractal sample moments* are defined as

$$M_n(q) = \sum_{i=1}^{b^n} \mu_{ni}^q, \tag{4}$$

where $\mu_{ni} = b^{-n} \rho_{ni}$ is the value of the field at the i th box at the resolution scale $\lambda_n = l_n/l_0$ with l_n being the scale of interest, l_0 is the current scale, $n \gg 1$ and $M_n(q) > 0$. To the multifractal moments in Eq. (4), one can associate λ_n to a parameter τ as follows:

$$M_n(q) \propto \lambda_n^{-\tau(q)}, \tag{5}$$

which provides

$$\tau(q) = \lim_{n \rightarrow \infty} \frac{\log M_n(q)}{-\log \lambda_n}.$$

An underlying assumption is that the measures of rainfall are independent and identically distributed (iid), and that there exists ergodicity in the rainfall process (at least approximately), which allows us to approximate $E[M_n(q)]$ by $M_n(q)$ (Over and Gupta, 1994).

The multifractal parameter τ characterizes the downscaling process and has the property of being scale invariant (Chhabra and Jensen, 1989; Chhabra et al., 1989). The disaggregation process is then characterized by computing τ for all desired levels for each multifractal moment q . This characterization is analogous to the one provided by the parameters β and σ in Eq. (3). Therefore, there exists a relationship between τ and the parameters β and σ . To see such relationship, one has to rely on the MKP function (Mandelbrot, 1974; Kahane et al., 1976). The MKP function is literally the slope of the linear regression of the log–log plot of $M_n(q)$ versus λ_n for different values of q , and it is specifically given by

$$\chi_b(q) = \log_b E[W^q] - (q - 1).$$

Theorem 3.2 in Over (1995) adapted a theorem of Holley and Waymire (1992) to account for the presence of zero rainfall giving the relationship

$$\tau(q) = d\chi_b(q), \tag{6}$$

where again d constitutes the domain of the cascade (here $d = 2$). Thus, one can explicitly compute the MKP function for the generator W and equate it with the multifractal sample q moments. That is,

$$\begin{aligned} \chi_b(q) &= \log_b E[W^q] - (q - 1) \\ &= 1 - q + \log_b (E[W = 0] + E[W^q = B^q Y^q]) \\ &= 1 - q + \log_b E[B^q] + \log_b E[Y^q], \end{aligned}$$

where

$$E[B^q] = b^{q\beta} P[B = b^\beta] = b^{\beta(q-1)}$$

and

$$\begin{aligned} E[Y^q] &= b^{\frac{-q\sigma^2 \log b}{2}} E[b^{q\sigma X}] = b^{\frac{-q\sigma^2 \log b}{2}} b^{\frac{q^2 \sigma^2 \log b}{2}} \\ &= b^{\frac{\sigma^2 \log b}{2}(q^2 - q)}. \end{aligned}$$

Thus,

$$\chi_b(q) = (\beta - 1)(q - 1) + \frac{\sigma^2 \log b}{2} (q^2 - q).$$

By considering the first and second derivatives of $\chi_b(q)$ in Eq. (6) with respect to the moment q , the parameters β and σ can be obtained as follows:

$$\beta = 1 + \frac{\tau'(q)}{d} - \sigma^2 \frac{\log(b)}{2} (2q - 1) \text{ and } \sigma^2 = \frac{\tau''(q)}{d \log(b)}. \tag{7}$$

The validity of Eq. (6) is given by the conditions in Theorem 3.2 in Over (1995); however, one can empirically test the linearity of $\log M_n(q)$ vs. $\log \lambda_n$ directly from measurements coming from a multifractal field. Thus, the formalism above

is valid for the q moments in which linearity is observed. The first and second derivatives of τ with respect to q can be computed, for example, by using finite differences. It is normal to compute the parameters of the β -lognormal model using $q = 1$. Note that $\tau(1)$ is always zero, and this obviously holds for all multifractal fields analyzed in this paper's application. This provides a common point of reference when computing the parameters β and σ . One way to argue the use of $q = 1$ is that it is related to the mean of the generator, and this is what is preserved at every step of the disaggregation process.

3.2 Heterogeneity information

The random generator model is only good for an isotropic distribution of the information at every step of the disaggregation process. From the spatial point of view, the presence of highly variable mountainous terrain in the Andean high plateau adds spatial heterogeneity to any rainfall measurement. From the temporal point of view, averaging rainfall over a given timescale, say hourly or daily, decreases randomness in the sense that averaging acts as a smoothing filter for the rainfall time series. The rainfall field after averaging shows the tendency of the field (usually dependent on the topography of the area), which reveals the heterogeneity in the data due to rainfall intensity. This was originally considered in Pathirana and Herath (2002). In meteorological wording, convective activity from cumulus clouds is much more heterogeneous than the total amount of precipitation observed over a month or larger timescale. Even temporal aggregation on a minute scale to obtain an hourly scale can cause spatial heterogeneity, however, this is usually masked by the information variability. For instance, the accumulated rainfall over all seasons in the Andean high plateau may cause the winter and summer to compensate each other to a large degree and produce a more uniform field than the winter or summer rainfall being measured separately. Since direct downscaling produces homogeneous multifractal fields (each TRMM cell is replaced by a set of random numbers distributed according to the β -lognormal model), the process of the multifractal downscaling model (as any other type of statistical downscaling model) is not enough to keep the heterogeneity information through the downscaling process. Pathirana and Herath (2002), Over and Gupta (1996) and Kang and Ramírez (2010) have applied suitable (pointwise) filters that highlight the effect of spatial heterogeneity. In this respect, rainfall is considered a combined effect of two processes: (1) a multifractal process which is highly variable in space, at least at regional and smaller scales, but statistically uniform; and (2) a deterministic process that represents the heterogeneity of rainfall in space and modifies the multifractal process. This is why prior to applying the β -lognormal model to the TRMM data described in Sect. 2, it is first necessary to remove the heterogeneity from the TRMM rainfall field. The rainfall field (spatial) under study is denoted as R . R is assumed to have the following decomposition:

$R_{i,j} = M_{i,j} G_{i,j}$, with M being a homogeneous multifractal field and G the deterministic component of weights corresponding to long-term monthly averages. For the application of multifractal downscaling, one requires the multifractal field M , which is obtained simply as

$$M_{i,j} = \frac{R_{i,j}}{G_{i,j}},$$

where $M_{i,j} = 0$ if $G_{i,j} = 0$. Specifically, G at position (i, j) is computed as

$$G_{i,j} = N \frac{\bar{R}_{i,j}}{\sum_{i,j} \bar{R}_{i,j}}, \quad (8)$$

where $\bar{R}_{i,j}$ is the monthly mean for the pixel location (i, j) , and N is the total number of pixels in a daily snapshot. In Sect. 4, the effect of extracting the long-term monthly averages during a month of the rainy season is shown in the exceedance probability plots at the stations' locations (see Fig. 15). Also, in Fig. 19 the limitation of this type of heterogeneity is observed during a dry month (July).

After downscaling, the heterogeneity is introduced back into the downscaled rainfall field. In this work, however, downscaling is performed beyond the largest resolution of TRMM (to reach a subkilometer rainfall resolution). Therefore, details about the local heterogeneity at the final downscaling resolution are required in order to account for spatial tendencies. Information at this scale is scarce or inexistent in the Andean high plateau; however, we rely on a rainfall field based on the stations' measurements (stations' locations are shown in Fig. 2) and elevation maps widely used in hydrological, agricultural and other models in the area under study (Hijmans et al., 2005). These rainfall fields are generated using the thin-plate smoothing spline algorithm implemented in the ANUSPLIN 4.4 package (Hutchinson, 2006; ANUSPLIN, 2007) for interpolation using scattered points of on-field measured rainfall together with latitude, longitude, and elevation as independent variables (Hutchinson, 1995). It is worth mentioning that several references indicate its higher accuracy compared to other methods in several parts of the world (Hijmans et al., 2005; Hartkamp et al., 1999; Jarvis and Stuart, 2001; Price et al., 2000). Also, the method has been widely applied on well-known and used climate products such as WorldClim ((Hijmans et al., 2005), <http://www.worldclim.org>) and IWMI (International Water Management Institute) Climate Atlas/CRU gridded data ((New et al., 2002), <http://www.iwmi.org>, <http://www.cru.uea.ac.uk>). The scalability properties of ANUSPLIN outputs are presented in Fig. 6. It is observed that the linearity of the plot of $\log M_n(q)$ vs. $\log \lambda_n$ is very good (determination coefficient is above 0.98 for all ANUSPLIN snapshots in the studied period from 1999 to 2006). Although these fields are "smooth" because of the usage of splines, these fields serve as a source for finding the monthly tendencies of rainfall at the desired

subkilometer resolution. More specifically, the local heterogeneity G' is the normalized field of spatial heterogeneities obtained as the monthly average of daily rainfall synthesized from ANUSPLIN output rainfall fields, which is equivalent to Eq. (8). The inclusion of the local heterogeneity provides the corrected downscaled rainfall field R' as follows:

$$R'_{i,j} = K \frac{M_{i,j} G'_{i,j}}{\sum_{i,j} M_{i,j} G'_{i,j}}, \quad (9)$$

where K is the large-scale forcing scale factor preserving the rainfall mean magnitude (Over and Gupta, 1994; Pathirana and Herath, 2002). Finally, the heterogeneity for the months of February and August is shown in Fig. 4 for both TRMM and ANUSPLIN outputs.

To end this section, the lognormality of the corrected ANUSPLIN outcomes is tested (see Fig. 5). That is, the monthly averages were pointwise filtered and then run over a lognormality test as it was performed on the TRMM data. Some disagreement is expected from the high rainfall intensities because of the northeast region being located in the rain forest rather than the Andean high plateau, and the fact that the ANUSPLIN outcomes did not use any station on that area for its construction. Note that since a lognormal distribution is stable, assembling all days in a period of time should amount to a set of numbers also lognormally distributed. Thus, Fig. 5 shows that the ANUSPLIN outcomes resemble a lognormal behavior for not too high rainfall intensities. Although, as already mentioned, it appears that the distribution may have a heavy tail for the reasons argued before.

4 Results: a regional downscaling application using TRMM rainfall data on the southern Andes

This section is divided in four parts. First, we describe the range of scales in which a multifractal downscaling is allowed. Then, the cascade generator is parameterized using the β -lognormal parameters β and σ . Next, the TRMM, the on-site stations and the downscaled corrected data are compared. This is followed by the spatial and temporal assessment of the generated downscaled rainfall intensity information.

4.1 Multifractal-scale range

Prior to applying a multifractal technique, an assessment of the available information was performed in order to detect multifractality in the data (TRMM precipitation in this case). This is summarized by the multifractal linear analysis (Over and Gupta, 1994; Deidda, 2000; Deidda et al., 2006). In summary, a multifractal characterization can only be performed if the log–log plot of the sample moments versus the resolution parameter (Eqs. 4, 5) shows a linear relationship for each moment order q (see Fig. 6). The multifractal parameter τ is then obtained as the slope of the linear tendency

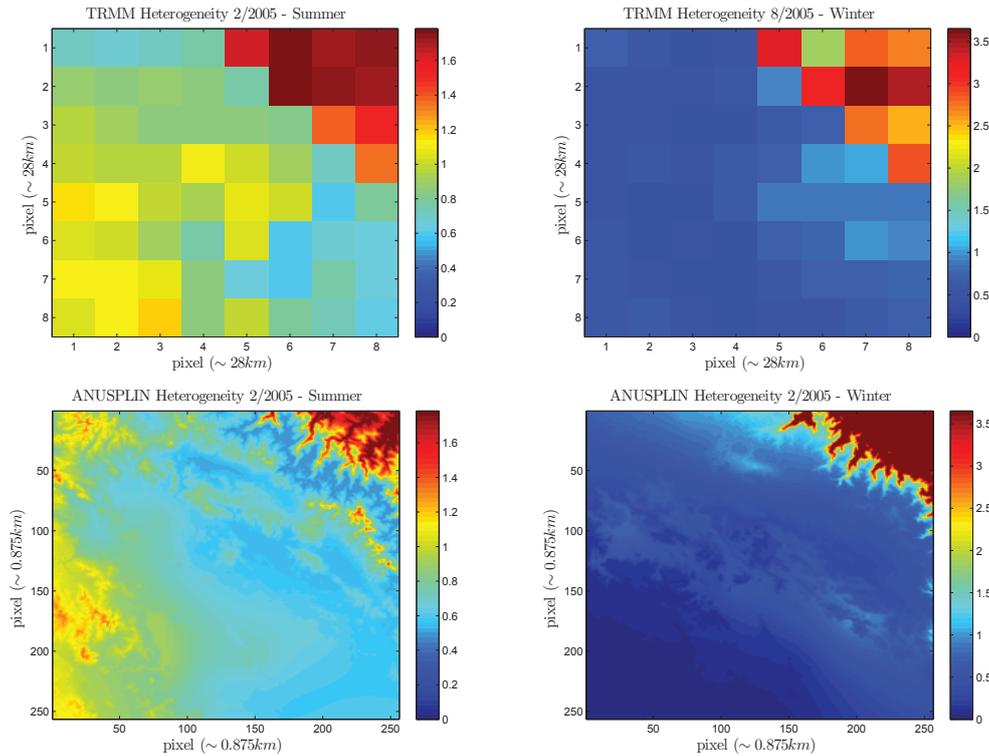


Figure 4. Illustration of summer and winter heterogeneity information from TRMM and ANUSPLIN outcomes.

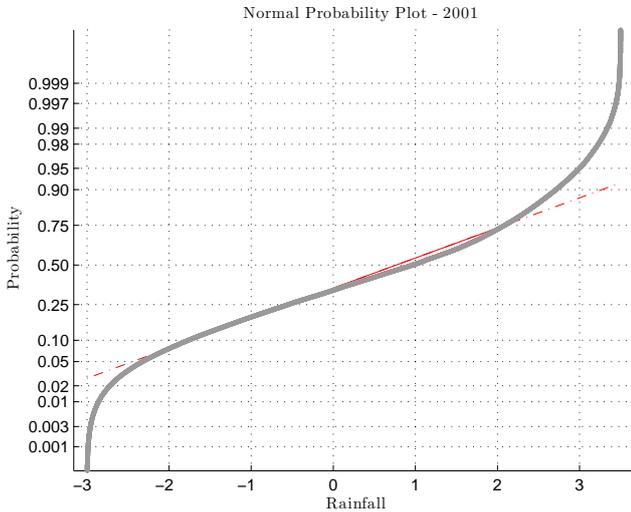


Figure 5. Lognormal test of ANUSPLIN data outcomes for the entire 2001 year.

for each q . It has been reported in Gires et al. (2012) and Tchiguirinskaia et al. (2011) the possibility of two different multifractal behaviors at approximately the 16 km resolution (scale breaking). However, there is a fundamental difference between the data used in those studies and that used in this study. First, TRMM is obtained on a daily temporal framework and it is the result of a 3 h accumulation and extrapo-

lated to reflect daily rainfall, whereas in Gires et al. (2012) and Tchiguirinskaia et al. (2011) the temporal framework is on a minute scale. Therefore, if such scale breaking were to occur, the described temporal averaging or dressing of the TRMM data accounts for the hindering of such scale breaking. Figure 7 shows the difference of τ as a function of the moments q for the summer and winter seasons for the years 2001 and 2004. Each curve $\tau(q)$ corresponds to a nonzero rain day in the corresponding month.

The multifractal linear analysis was performed on TRMM data and also on the ANUSPLIN output snapshots. The latter was done in order to assess the validity of the equations presented in Sect. 3 when downscaling beyond the highest resolution of TRMM data. For TRMM data, three cascade levels were assessed in which the resolutions run from ~ 225 to ~ 28 km. For ANUSPLIN data, eight cascade levels were considered sweeping from ~ 225 to ~ 0.875 km. The cascade levels correspond to the scale parameter $\lambda_n = 2^n$. The range of q 's considered in the analysis was $[0, 5]$; see Fig. 6 for the analysis on 26 February 1999 for both TRMM and ANUSPLIN data. Specifically, the coefficients of determination for the linear regression for q 's in the interval of $[0, 5]$ are above 0.98 throughout the period of time studied (every day from 1999 to 2006). This result allows for the application of a multifractal procedure; i.e., the parameter τ is well defined in this range and therefore the MKP function can be used to infer the parameters β and σ . Here we have used finite difference

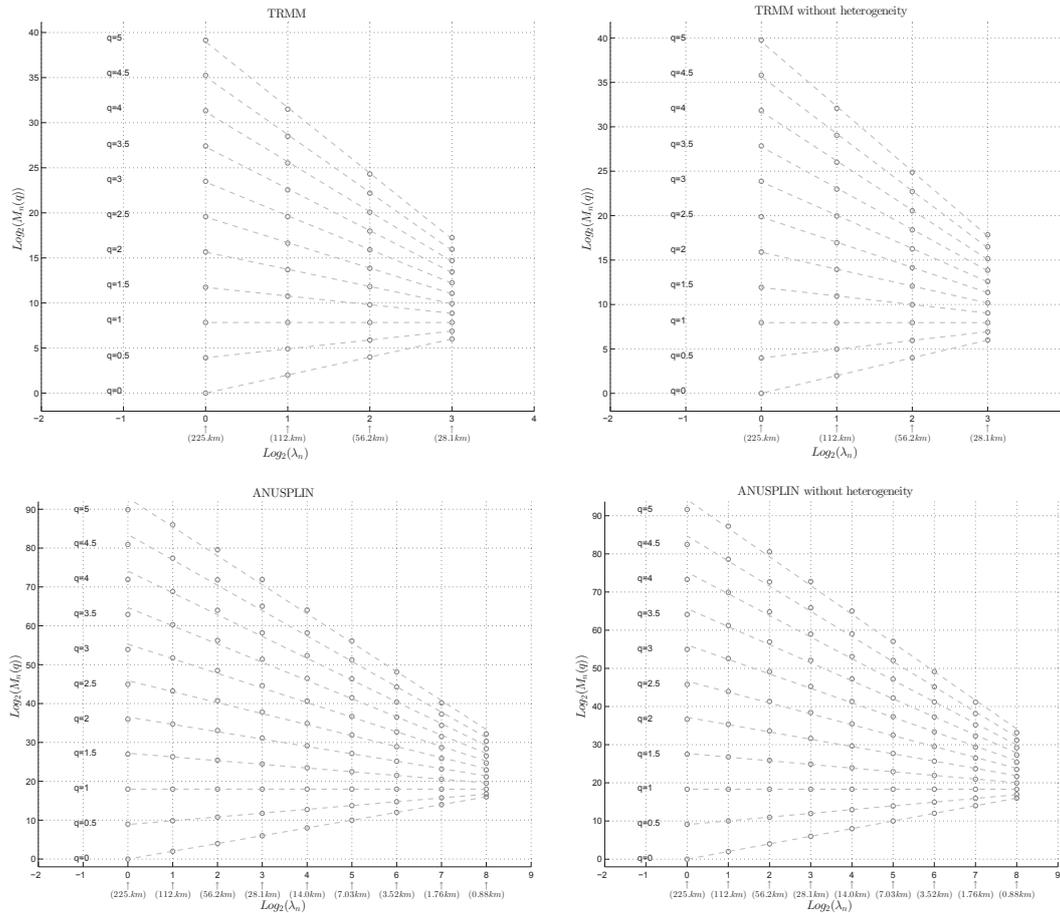


Figure 6. Multifractal linear analysis of TRMM and ANUSPLIN data for 26 February 1999 and q values in the range $[0, 5]$.

methods to evaluate the first and second derivatives of $\tau(q)$, with respect to q as implemented by Over (1995), Over and Gupta (1994, 1996), and Sharma et al. (2007). In particular, these derivatives were estimated using a partition of step 0.1 on the q values, so that $\tau'(1)$ and $\tau''(1)$ were estimated using $q = 0.9$ and $q = 1.1$.

4.2 Multifractal characterization and lognormal parameters

Lognormal parameters β and σ were computed for each snapshot (daily timescale). Their features, in this paper’s application, correspond to specific aspects of precipitation dynamics. For example, β is associated with the degree of intermittency in the precipitation field, and σ determines the variance of the cascade generator. The relationship between the given information (rainfall intensity, variance, energy, entropy indicators, etc.) and the lognormal parameters is key to understand the role of the cascade parameters and the measurements. Here only the rainfall intensity (mean over all TRMM grids at time t) is compared with respect to the β and σ lognormal parameters. This relationship between data and the lognormal parameters is inferred from Eqs. (4), (5)

and (6). In Sect. 3.1, τ was computed from the sample moments, which are direct consequence of the rainfall intensity data in Eq. (5). Given that β and σ are obtained from Eq. (7), they are implicitly related to τ . The monthly values of β and σ vs. the logarithm of mean rainfall intensity (per snapshot), for the period running from 1999 to 2006, are fitted with respect to the empirical models used in Pathirana and Herath (2002) and Jothityangkoon et al. (2000). The nonlinear regressions for the months of February and July are provided in Fig. 8, and only snapshots with nonzero rainfall intensity were considered. These plots visually express the relationship of the parameters β and σ with respect to the large-scale forcing, which here is the mean. One can observe that while β shows high variability the parameter σ does not. In Pathirana and Herath (2002), σ was set to a fixed number due to its lack of sensitivity with respect to rainfall intensity. A similar behavior is observed in the Andean high plateau, but in this paper σ is allowed to vary. Given that the probability of nonzero rain is $b^{-\beta}$, one can observe that in February the majority of β values are close to zero (probability close to 1), which is expected from a rainy month, whereas for July the probability of rain for snapshots with rain on them decreases

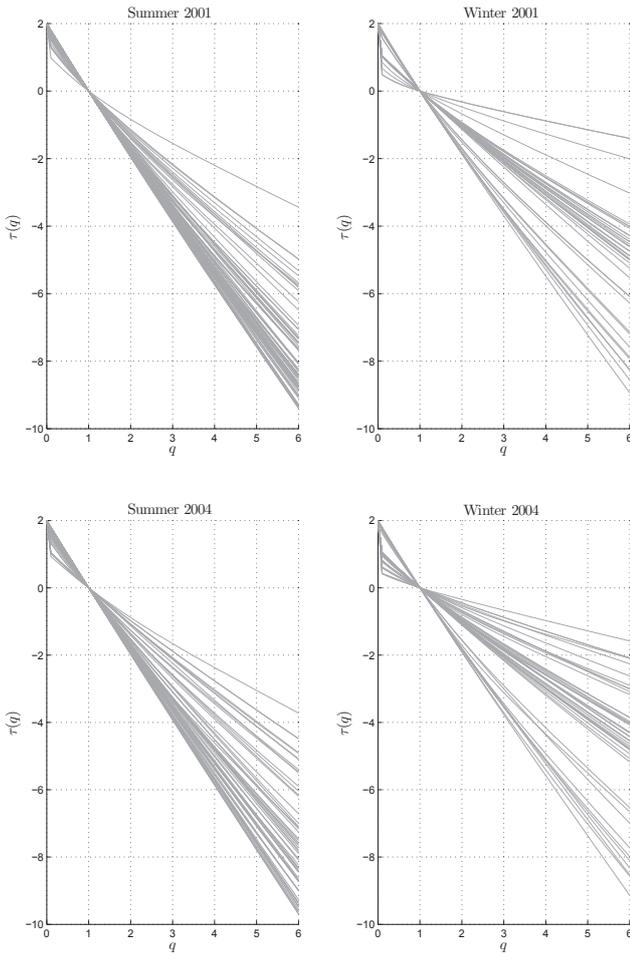


Figure 7. $\tau(q)$ for summer and winter of 2001 and 2004.

to around 50 % indicated by the value of $\beta \approx 0.4$. Note that the scatter plot for July is much less populated than the one for February since there are less rainy days during July. The stability analyses of β and σ are given in Tables 2 and 3. It is observed that the minimum value of the average β occurs in February ($\beta = 0.1568$ gives probability $\approx 80\%$) and the maximum occurs in June ($\beta = 0.7284$ gives probability $\approx 35\%$), as expected from the usual seasonal cycle of precipitation in the Andes. On the contrary, the exact opposite happens with the values of σ in which the wet season has more variance during the wet season compared to the localized rainfall during the dry season given by low values of σ . This last fact simply tells us that the rainfall is much more spread out in the region during the rainy season while it is more localized in the dry season. This is of course observed when comparing February and July in Fig. 8.

4.3 Generation of expected scenario

The main tool in the generation of rainfall data at smaller scales, as described in previous sections, is a disaggregation

Table 2. Monthly statistics for β (1999–2006).

Month	Mean	Median	SD	Skew
January	0.1693	0.0434	0.2586	1.9906
February	0.1568	0.0419	0.2572	2.1882
March	0.1843	0.0490	0.2709	1.7880
April	0.4713	0.4226	0.3872	0.2663
May	0.6560	0.6337	0.3492	-0.3614
June	0.7284	0.9491	0.3030	-0.5844
July	0.7142	1.0000	0.3262	-0.5900
August	0.6927	1.0000	0.3674	-0.6224
September	0.5840	0.4847	0.4063	-0.2053
October	0.3795	0.4192	0.3603	0.6665
November	0.4244	0.4224	0.3683	0.4507
December	0.2656	0.1196	0.3065	1.1686

Table 3. Monthly statistics for σ^2 (1999–2006).

Month	Mean	Median	SD	Skew
January	0.0382	0.0273	0.0345	1.4670
February	0.0387	0.0299	0.0326	1.0053
March	0.0403	0.0323	0.0344	1.2436
April	0.0268	0.0136	0.0334	1.6286
May	0.0168	0.0018	0.0269	1.6558
June	0.0163	0.0004	0.0297	2.1931
July	0.0130	0.0000	0.0236	2.4033
August	0.0128	0.0000	0.0234	2.0872
September	0.0211	0.0036	0.0337	2.0339
October	0.0284	0.0135	0.0368	2.1377
November	0.0313	0.0164	0.0375	1.5117
December	0.0339	0.0245	0.0319	1.0955

procedure having a random generator associated with a β -lognormal model with parameters β and σ . Recall that this generator includes the possibility of no rain in the generator (see Eq. 3). Figure 9 shows TRMM data with no heterogeneity along with its downscaling to a subkilometer resolution averaging 1000 realizations and its heterogeneity correction as described in Sect. 3. Specifically, the parameters of the β -lognormal distribution for this day are $\beta = 0.0266$ and $\sigma = 0.1109$.

Based on this consideration, the objective of this section is to compare the quality of the disaggregation procedure with respect to TRMM precipitation. The information to be generated is the mean rainfall intensity of all days in a month over the period of 8 years. Figure 10 shows two scatter plots comparing the means of observed and generated rainfall snapshots. The plots for February and July were chosen since they correspond to distinct rainfall patterns. The former is representative of a summer month (rainy season) and the latter is representative of a winter month (dry season). The scatter plots show all information over the period of 8 years to assure a good statistical assessment. Similar behavior is observed for all the other months (not shown). One can observe

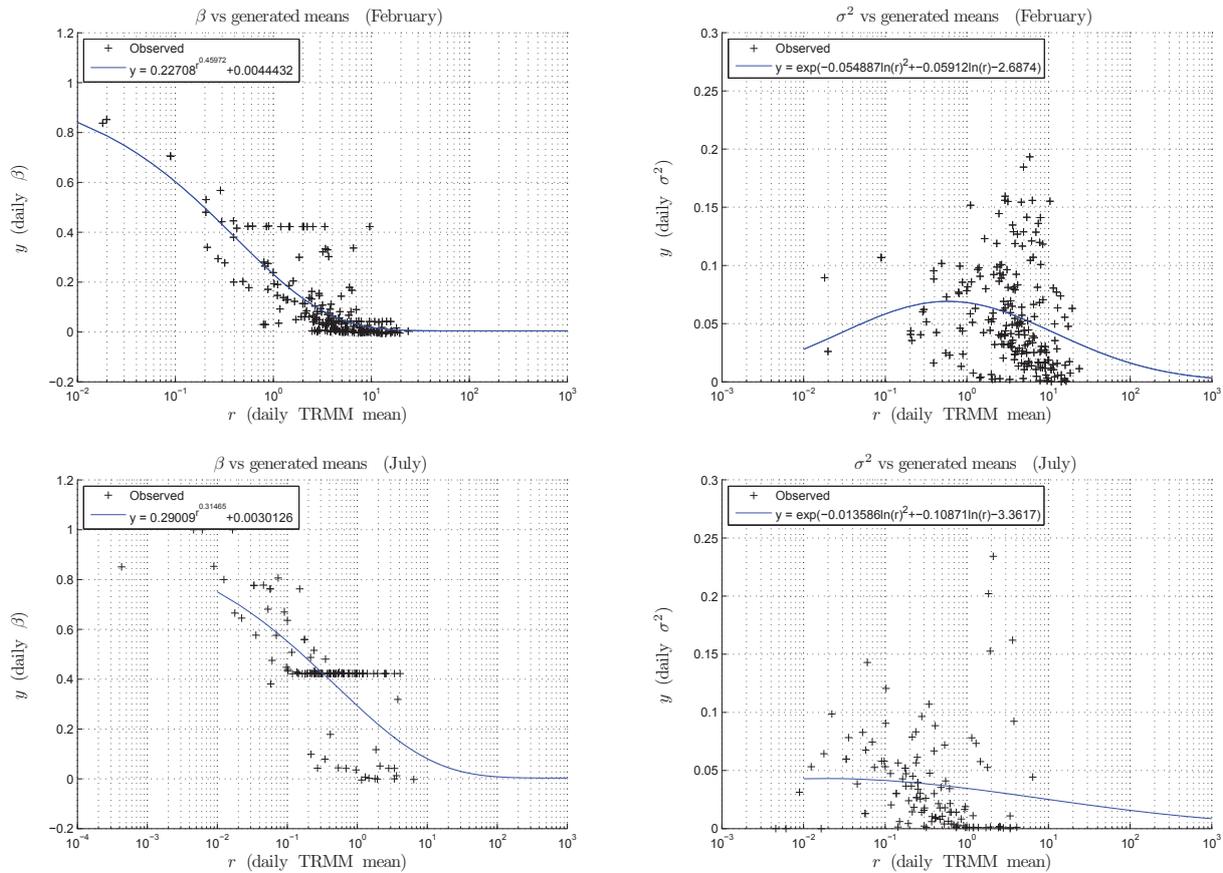


Figure 8. Relationship between β and the daily TRMM mean values for the months of February, May, August and December for the period of 1999–2006.

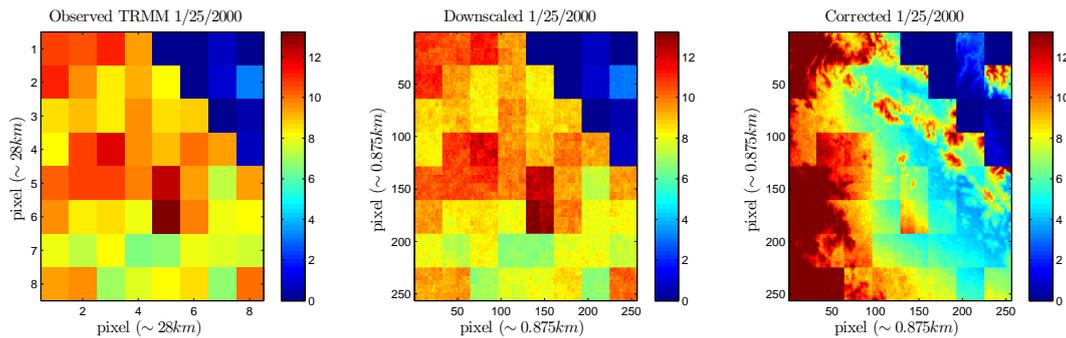


Figure 9. Comparison of observed, downsampled and corrected rainfall measurements for day 25 January 2000 and 1000 realizations ($\beta = 0.0266$ and $\sigma = 0.1109$).

that the mean of each snapshot is preserved accurately as expected. This is caused by the large-scale forcing factor, K , used in the correction procedure described in Sect. 3.2. On the other hand, increasing the resolution via the downscaling process increases the variability (or randomness) of rainfall as expected from the fact that the sample $q = 2$ moment is proportional to $\lambda_n^{\tau(2)}$. This is observed when comparing the variances of observed and generated snapshots over a period

of time. In addition, the heterogeneity correction procedure was devised for preserving the means and not other statistical moments. Also, note that the increment in variability changes depending on the season, i.e., February (wet season) shows on average about 4 times more variability after downscaling and correction (values outside the regression line). Observe that the variance can only increase, and that the very few cases in Fig. 10 where the variance decreases are mainly

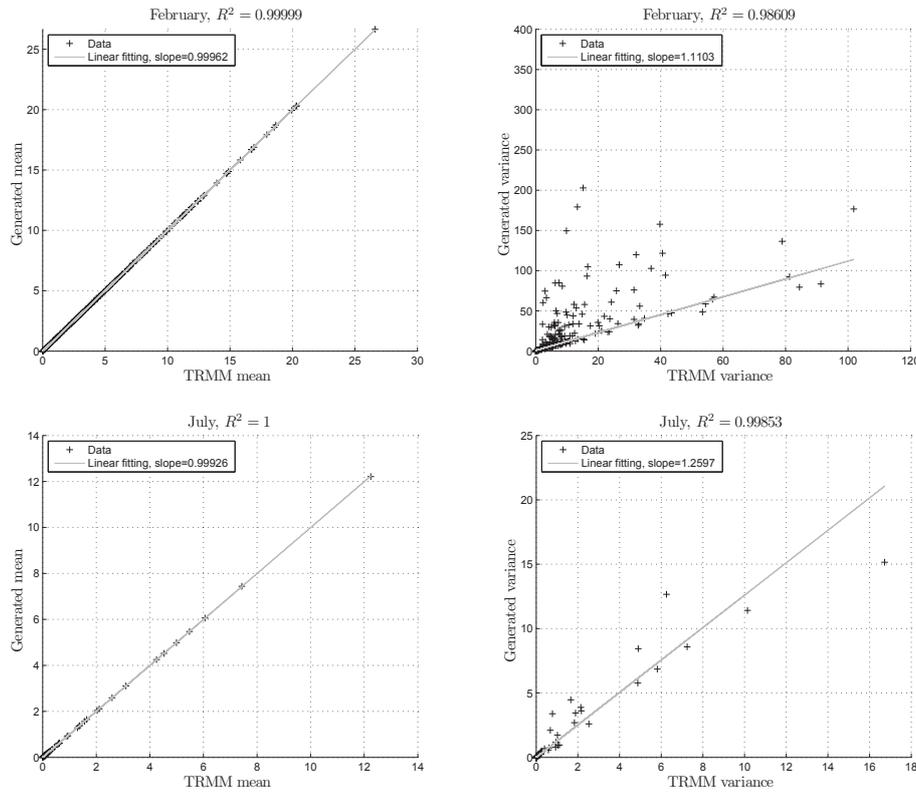


Figure 10. Comparison between TRMM and simulated rainfall means from 1999 to 2006 in the months of February and July.

due to the correction process. On the other hand, July (dry season) exhibits about 2.5 times more variability after downscaling. In general, an argument for such changes in variance is that, from a temporal point of view, the second moments are proportional to $\lambda_n^{\tau(2)}$ in contrast with the case for $q = 1$, where all $\tau(q)$ curves coincide ($\tau(1) = 0$ always) (see Fig. 7).

With the objective of illustrating the disaggregation procedure used in the paper, Fig. 11 presents the level by level downscaling for 25 January 2000.

Also, the ensemble mean realization is given in Fig. 12 for 10, 100 and 1000 realizations of the downscaling and correction procedure described in Sect. 3.

4.4 Validation of downscaled and corrected rainfall

This section begins with a pointwise (spatial) comparison of some statistics between observed and generated (downscaled and corrected) time series for the Capachica and Cojata stations. Such time series are shown in Fig. 13. In Table 4 the statistics of observed and generated time series are provided. In particular, it is observed that the Hurst exponents for the whole 8 years (all seasons) between on-site observed and generated (downscaled and corrected) rainfall mostly agree. For the Capachica station, it was found that $H_{\text{obs}} = 0.7453$ and $H_{\text{gen}} = 0.6750$; for the Muñani station $H_{\text{obs}} = 0.6766$ and $H_{\text{gen}} = 0.6838$. A Hurst exponent in

the range $0.5 < H < 1$ indicates a long-term positive auto-correlation, which implies the tendency of a high value to be followed by another high value. This behavior, $H > 0.5$, is shared by the other stations as well. However, the fact that $H \neq 0.5$ may indicate that the multifractal field is not conservative, which is usually handled by studying the field fluctuations (Lovejoy and Schertzer, 2013). However, H here is in general close to 0.5 and the differences can be attributed to, for example, uncertainty/error in the precipitation measurements of both TRMM and the stations. A fluctuation analysis would be the concern of future research. Also, it is important to highlight that the only source of correlation among snapshots of downscaled rainfall is carried over from the correlation of TRMM data since the downscaling process is exclusively spatial.

4.4.1 Wet season (November–March)

Figure 14 shows the quantile–quantile (Q–Q) plots of four stations during the wet season: Capachica, Chuquibambilla, Cojata and Mañazo. The plots show that the distributions of the observed and generated rainfall information largely agree. The very last point in the Q–Q plot represents the difference between 99 and 100 % quantile information. It exhibits peaks originated from the downscaling procedure. We also point out that for Capachica and Chuquibambilla stations the agreement for high rainfall values may improve by

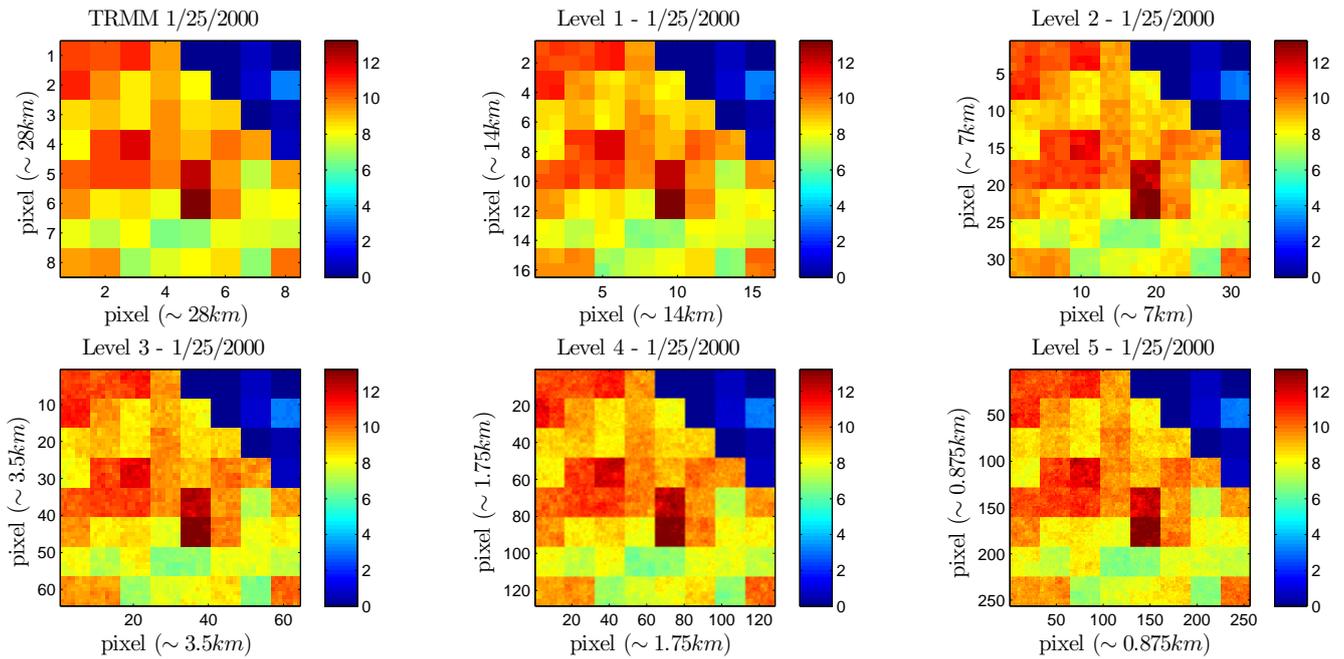


Figure 11. Downscaling for TRMM snapshot on 25 January 2000 for $\lambda_n = 14, 7, 3.5, 1.75,$ and 0.875 (km).

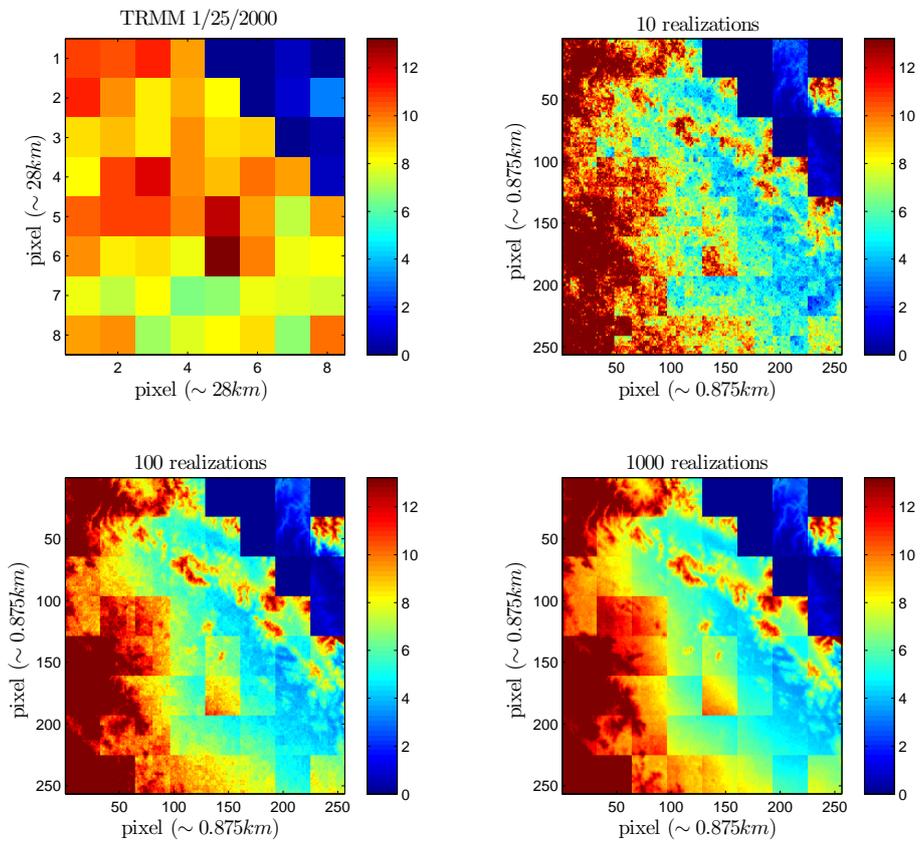


Figure 12. Ensemble test for the downscaling and correction of TRMM information for 25 January 2000.

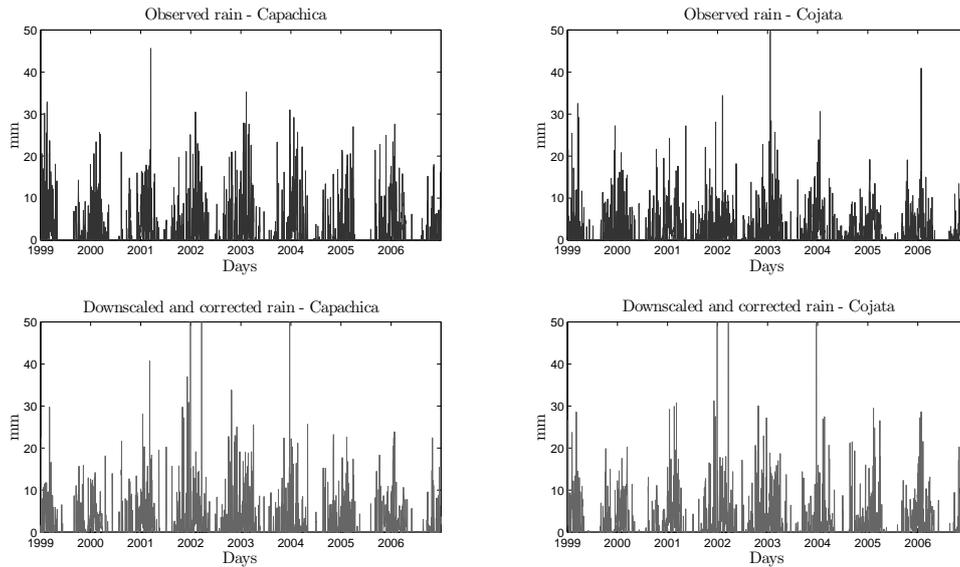


Figure 13. Observed and generated time series for the Capachica and Cojata stations.

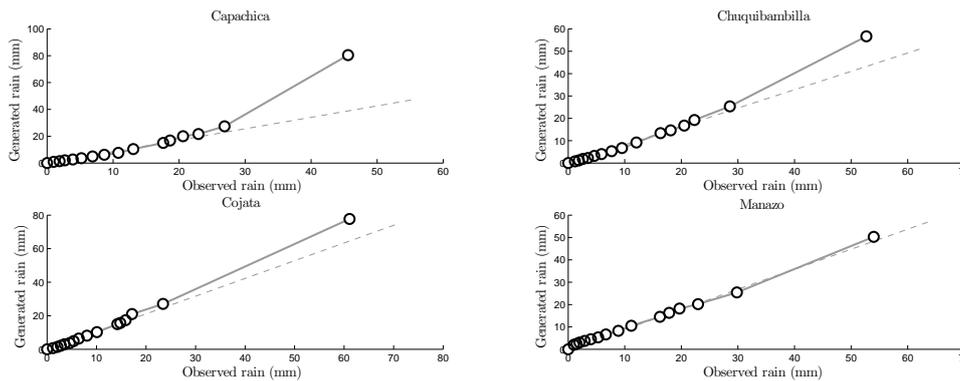


Figure 14. Quantile–quantile plots.

increasing the number of realizations in the rainfall generation (downscaling) procedure since the probability of these rainfall intensities is low. Specifically, the Capachica station is largely affected by Lake Titicaca, which explains some differences in the Q–Q plots for higher values of rainfall. Other reasons for the discrepancy are (1) temporal heterogeneity only being introduced indirectly via the monthly correction using the fields G' , (2) the uncertainty about extremes that the local heterogeneity (from ANUSPLIN) cannot mitigate due to the intrinsic smoothness of the splines used by the ANUSPLIN package, and (3) that perhaps a period of 8 years is not long enough to statistically produce correctly the extreme values causing the differences in the Q–Q plots presented here.

Another form of temporal assessment consists in comparing the temporal variation of rainfall intensity at a particular location by using cumulative plots. First, this analysis is used to illustrate the effect of the heterogeneity correction on

the TRMM data. Obviously, in Fig. 15 (left panel) one can group the on-site stations in at least two classes, which is due to the heterogeneity of the studied area. During the wet seasons this heterogeneity mainly correlates to rainfall intensity long-term averages. Thus, by weighting TRMM with the heterogeneity G , the rainfall TRMM fields get homogenized. This results in the overlapping of all the exceedance curves (see Fig. 15, right panel).

In Fig. 16, the exceedance probabilities over the period of 8 years for station-measured precipitation and the generated (downscaled and corrected) from TRMM data are overlaid over four places in which a meteorological station is located (indicated by black dots in Fig. 2). They were chosen to represent the terrain and precipitation heterogeneities in the Andean plateau. Specifically, the Capachica station is located close to Lake Titicaca, which is considered a humid zone due to the lake's influence. The Chuquibambilla station is located in a semi-arid zone. The Cojata station corresponds to

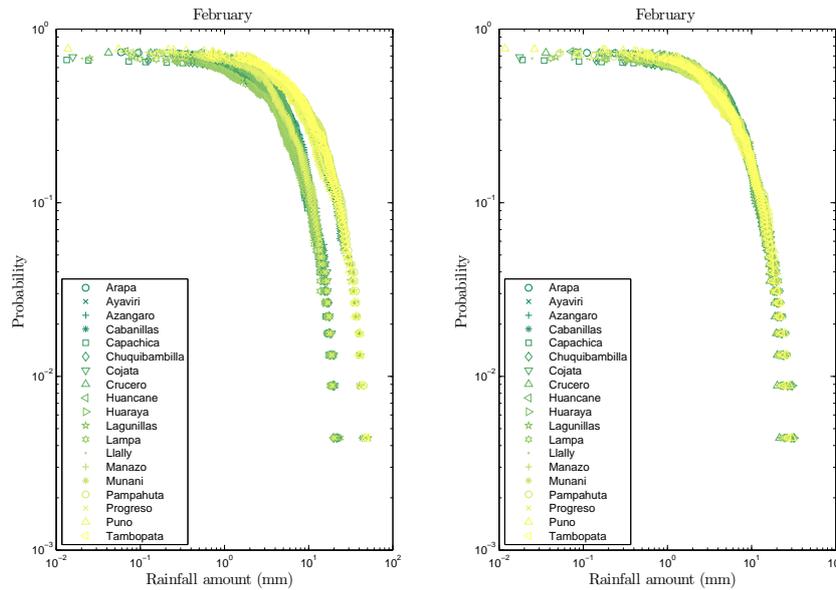


Figure 15. Exceedance probability plots for February over an 8-year period of (right) TRMM data (*R* rainfall field) and (left) TRMM with long-term averages (monthly) removed (*M* rainfall field).

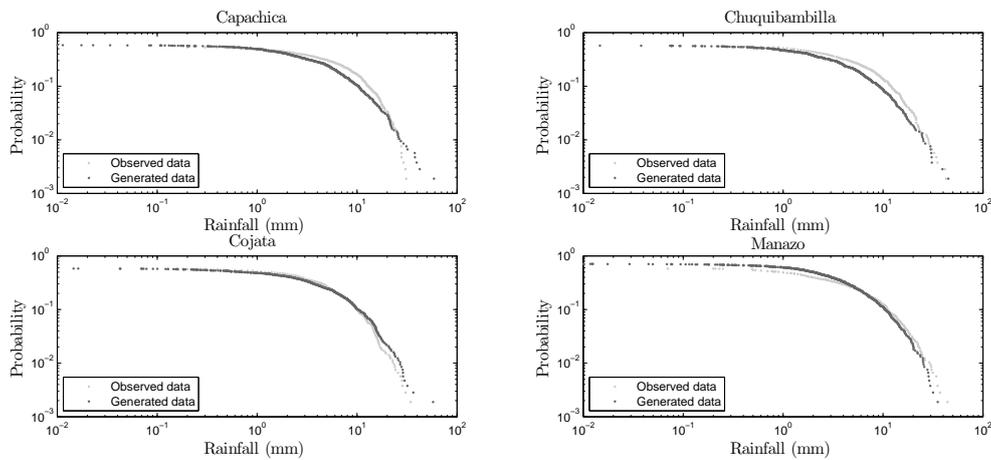


Figure 16. Exceedance probability plots for the rainy season (November–March) over an 8-year period of observed and simulated TRMM measurements for stations Capachica, Chuquibambilla, Cojata and Mañazo.

an abrupt mountain terrain. The Mañazo station is located in the arid section of the study area, with a slight influence of the mountains, and the Tambopata station is located in the rain forest. In Table 4, the statistics of the on-site observed and generated via downscaling and correction rainfall for the mentioned stations are compared. It shows a good agreement taking into consideration that the information provided by the stations has been assumed comparable with respect to the generated rainfall at a spatial resolution of ~ 0.875 km. That is, while the downscaled and corrected information has a resolution of ~ 0.875 km, the measurements at the stations are punctual. This is an intrinsic source of error and it justifies

some of the differences in the statistics between observed and generated rainfall information.

Table 5 shows the following indicators of the goodness of fit of the exceedance probability curves: MAE (mean average error), RMSE (root mean square error), CORR (correlation coefficient), PBIAS (percent Bias), NSE (Nash–Sutcliffe efficiency) and RSR (ratio of RMSE to the standard deviation of the observations); see Moriasi et al. (2007) for more information about goodness of fit indicators. In general, the model simulation can be considered satisfactory if the indicator NSE is greater than 0.50 and the indicator RSR is around 0.80 or less.

Table 4. Statistics of temporal information during the wet season (November–March) from 1999 to 2006.

Station		<i>H</i>	Mean	Max	<i>Q</i> ₅₀	<i>Q</i> ₇₅	Var
Capachica	Obs	0.65	4.26	45.60	1.00	6.90	39.55
	Gen	0.57	3.28	79.27	0.89	4.42	38.13
Chuquibambilla	Obs	0.55	4.05	52.70	1.20	5.90	40.01
	Gen	0.38	3.03	54.14	0.71	4.00	29.65
Cojata	Obs	0.61	3.55	61.10	1.20	5.38	28.60
	Gen	0.58	3.52	70.66	0.69	4.97	36.95
Mañazo	Obs	0.71	3.76	54.00	1.00	5.30	38.69
	Gen	0.61	3.12	40.03	1.52	4.50	20.23
Tambopata	Obs	0.50	6.42	105.1	1.90	8.18	109.8
	Gen	0.57	4.16	80.66	1.40	5.98	48.23

Table 5. Goodness of fit for exceedance probability curves on the wet season (November–March) from 1999 to 2006.

Station	MAE	RMSE	CORR	PBIAS	NSE	RSR
Capachica	0.05	0.06	0.98	25.58	0.85	0.39
Chuquibambilla	0.05	0.05	0.99	24.72	0.88	0.34
Cojata	0.01	0.02	1.00	7.36	0.98	0.12
Mañazo	0.04	0.04	0.99	17.96	0.93	0.26
Tambopata	0.05	0.06	0.99	26.86	0.88	0.35

Table 6. Goodness of fit for exceedance probability curve and statistics on the dry season for the Santa Rosa station (November–March) from 1999 to 2006.

	Obs	Gen
<i>H</i>	0.57	0.45
Mean	4.03	3.56
Min	0	0
Max	43.50	62.81
Skew	2.19	3.74
<i>Q</i> ₂₅	0	0
<i>Q</i> ₅₀	1.80	1.24
<i>Q</i> ₇₅	5.70	4.68
VAR	33.46	35.09
MAE	0.025	0.025
RMSE	0.029	0.029
CORR	0.997	0.997
PBIAS	12.65	12.65
NSE	0.98	0.98
RSR	0.17	0.17

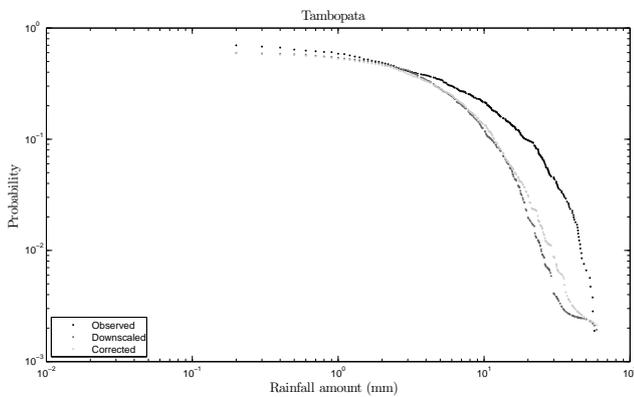


Figure 17. Exceedance probability of the Tambopata station during the wet season over an 8-year period. Comparison of observed, generated and corrected rainfall for the Tambopata station.

It is clear that the multifractal downscaling and correction procedures give appropriate results and, in particular, they are very close to the ones measured on-site by the weather stations in Capachica, Chuquibambilla, Cojata and Mañazo (see Fig. 16). However, the correction for the Tambopata station clearly fails to provide an acceptable result, even though the NSE and RSR are within the acceptable ranges, because this station was not employed in the generation of the local heterogeneity matrix (ANUSPLIN outcomes). This is because there are no other stations at a reasonable distance, which introduces undesired errors in the interpolation procedure utilized by the ANUSPLIN package. Reasons for the lack of stations in the area are the inaccessibility, cost, and human factor, all of which are needed to maintain more stations in the tropical forest. This fact can be clearly observed in Fig. 17, where the correction improves the exceedance curve but not enough for its usage in agriculture models for the rain forest region. The Tambopata station illustrates how sensitive the correction procedure is with respect to the heterogeneity used.

Finally, a validation during the wet season is performed on a station that has not been used for either correcting

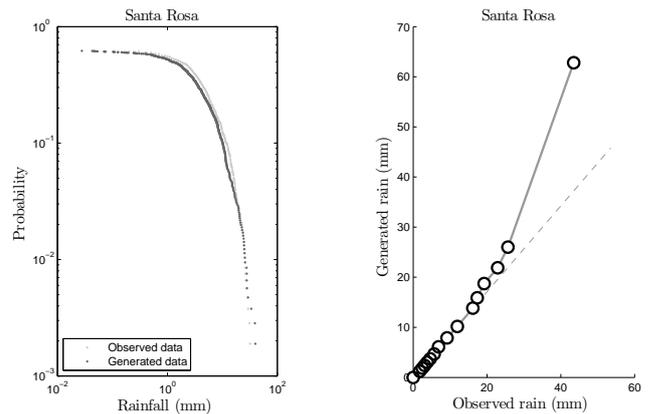


Figure 18. Exceedance probability and quantile–quantile plot of the Santa Rosa station during the dry season over an 8-year period. Comparison of observed, generated and corrected rainfall for the Santa Rosa station.

TRMM information or to generate the heterogeneity matrices using the ANUSPLIN package. This is the Santa Rosa station. The Santa Rosa station is geographically located at longitude 70.79° W, latitude 14.62° S and has an altitude of 3940 m a.s.l. This location corresponds to the pixel at col-

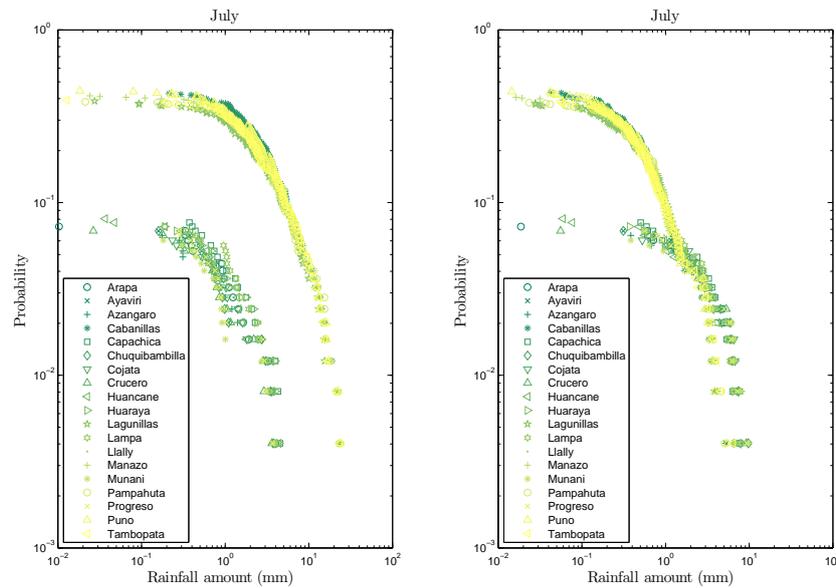


Figure 19. Exceedance probability plots for July over an 8-year period of (right panel) TRMM data (*R* rainfall field) and (left panel) TRMM with long-term averages (monthly) removed (*M* rainfall field).

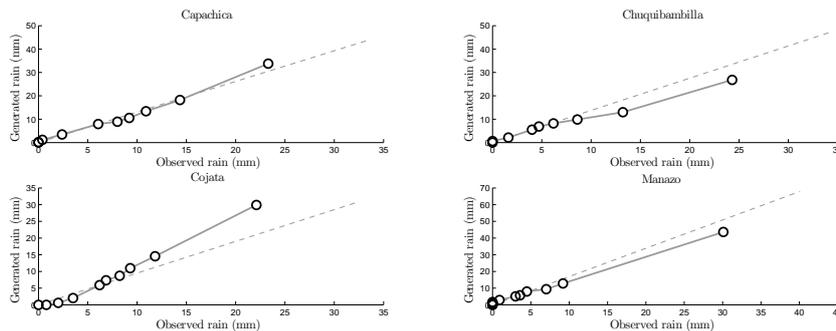


Figure 20. Quantile–quantile plots dry season.

umn 27 and row 81 of the 256×256 grid constituting the scale after downscaling. The exceedance probability plot and quantile–quantile plots are shown in Fig. 18 as well as its quantified time series statistics and exceedance probability plot of goodness of fit in Table 6. A good agreement, similar to the one shown in Tables 4 and 5, is observed. This is due to the fact that the location where the Santa Rosa station is located possesses enough stations to characterize the area’s heterogeneity. As mentioned in the previous paragraph, an isolated station like Tambopata fails the validation due to the fact that the area does not have other stations that characterize the heterogeneity of the rain forest region in the area under study.

4.4.2 Dry season (June–October)

The exceedance probability assessment is repeated for the dry season. This is to confirm the limitations of the heterogeneity correction as indicated in Pathirana and Herath

(2002). Note first that the heterogeneity is not completely removed by weighting TRMM data using long-term averages (see Fig. 19). The heterogeneity now is due to rainfall intermittency rather than rainfall intensity (Pathirana and Herath, 2002). Observe that the stations’ exceedance probability plots do not overlap after removing the long-term rainfall average heterogeneity. In spite of that, the Q–Q plots in Fig. 20 show very good agreement, but this result could be misleading since during the dry season percentiles are mostly zero except for the high ones (above 90 %).

The results of the downscaling and correction procedures show mixed results. Table 7 shows a comparison of the observed and generated rainfall statistics of these stations. All stations except Tambopata show significant agreement. However, the analysis of the exceedance probability shows the limitations of the technique during the dry season. More precisely, the Cojata and Mañazo stations show some discrepancies reflected in their NSE and RSR indexes (Table 8).

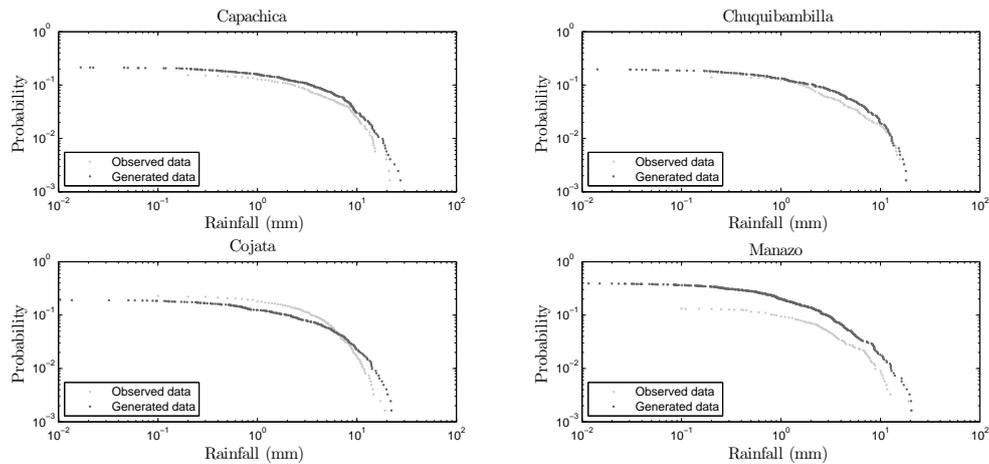


Figure 21. Exceedance probability plots for the non-rainy season (June–October) over an 8-year period of observed and simulated TRMM measurements for stations Capachica, Chuquibambilla, Cojata and Mañazo.

Mañazo results, for example, seem to depend heavily on the intermittency heterogeneity, which has not been considered and is reflected in the bad $NSE = -2.51$ and $RSR = 1.86$. Also, the effect of the surrounding water sources on the Capachica and Chuquibambilla stations seems to be minimal in comparison to the wet season. This is shown in their strong goodness of fit indicators ($NSE \approx 0.88$ and $RSR = 0.34$ for both stations).

Finally, the Tambopata station is analyzed for validation during the dry season. In general, the rain forest behavior is quite homogeneous during both wet and dry seasons. Removing the long-term averages of TRMM basically removes a constant value to the whole northeast corner of the studied area as observed in the downscaled curve in Fig. 22. The ANUSPLIN outcomes provides a constant factor, as observed in the corrected curve in Fig. 22, that is not large enough to agree with the on-site measurements at the station. Therefore, more information is needed around Tambopata in order to produce better local heterogeneity, which is the same conclusion reached in the analysis during the wet season for the rain forest region.

5 Conclusions

The downscaling multifractal technique was used for obtaining rainfall estimations at small spatial scales in the Andean high plateau region. It was pointed out that the mountains in the Andean high plateau alter the cloud dynamics of the TRMM rainfall measurements. Some of these effects were attenuated by removing the monthly averages of TRMM data. This procedure worked consistently for the wet season whereas for the dry season the results were mixed. Overall, the downscaled and corrected TRMM rainfall data showed that the procedure employed in this paper is a reliable downscaling technique for the Andean high plateau with potential

Table 7. Statistics of temporal information during the dry season (April–October) from 1999 to 2006.

Station		<i>H</i>	Mean	Max	<i>Q</i> ₅₀	<i>Q</i> ₇₅	Var
Capachica	Obs	0.50	0.82	23.30	0	0	7.60
	Gen	0.48	0.96	31.39	0	0	9.02
Chuquibambilla	Obs	0.61	0.62	24.30	0	0	5.19
	Gen	0.43	0.76	28.59	0	0	6.48
Cojata	Obs	0.55	0.95	22.10	0	0	6.22
	Gen	0.53	0.79	29.43	0	0	7.55
Mañazo	Obs	0.44	0.47	30.10	0	0	3.82
	Gen	0.56	0.77	28.78	0	0.51	4.40
Tambopata	Obs	0.56	2.41	82.30	0	1.65	37.62
	Gen	0.53	0.38	19.00	0	0	2.67

Table 8. Goodness of fit for exceedance probability curves on the dry season (April–October) from 1999 to 2006.

Station	MAE	RMSE	CORR	PBIAS	NSE	RSR
Capachica	0.01	0.02	0.99	17.72	0.88	0.35
Chuquibambilla	0.01	0.01	0.98	21.87	0.87	0.35
Cojata	0.02	0.03	0.99	27.90	0.78	0.47
Mañazo	0.05	0.08	0.98	88.61	-2.51	1.86
Tambopata	0.10	0.13	0.46	88.24	-0.97	1.40

application as an input for agricultural models requiring sub-kilometer precipitation information, but some extra research is required in order to overcome the heterogeneity related to intermittency, which was shown in the analysis of the dry season.

It was also observed that the multifractal technique is sensitive in relation to the heterogeneity of local corrections. In particular, the Tambopata station had no surrounding stations helping in the generation of local heterogeneity in the northeast corner of the area under study. This showed in the

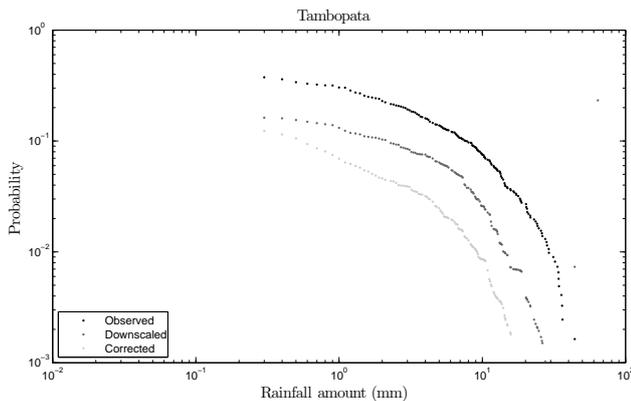


Figure 22. Exceedance probability of the Tambopata station during the dry season over an 8-year period. Comparison of observed, generated and corrected rainfall for the Tambopata station.

large differences between on-site measurements and generated (downscaled and corrected) precipitation values. The disagreement occurred in both wet and dry seasons.

Acknowledgements. This work is part of the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). Also, the authors want to thank the useful comments and observations provided by the anonymous reviewers.

Edited by: I. Tchiguirinskaia

Reviewed by: two anonymous referees

References

- Aceituno, P. and Montecinos, A.: Circulation anomalies associated with dry and wet periods in the South American Altiplano, Proc. Fourth Int. Conf. on Southern Hemisphere Meteorology, Hobart, Australia, Am. Meteor. Soc., 330–331, 1993.
- ANUSPLIN: The ANUSPLIN package, Version 4.4, <http://fennerschool.anu.edu.au/research/products/anusplin-vrsn-44> (last access: 6 July 2015), 2007.
- Baron, C., Sultan, B., Balme, M., Sarr, B., Traore, S., Lebel, T., Janicot, S., and Dingkuhn, M.: From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact, Philos. T. Roy. Soc. B, 360, 2095–2108, 2005.
- Chhabra, A. and Jensen, R. V.: Direct determination of the $f(\alpha)$ Singularity Spectrum, Phys. Rev. Lett., 62, 1327, 1989.
- Chhabra, A., Meneveau, C., Jensen, R. V., and Sreenivasan, K. R.: Direct determination of the $f(\alpha)$ singularity spectrum and its application to fully developed turbulence, Phys. Rev. A, 40, 5284–5294, 1989.
- Deidda, R.: Rainfall downscaling in a space-time multifractal framework, Water Resour. Res., 36, 1779–1794, 2000.
- Deidda, R., Benzi, R., and Siccaldi, F.: Multifractal modeling of anomalous scaling laws in rainfall, Water Resour. Res., 35, 1853–1867, 1999.
- Deidda, R., Badas, M. G., and Piga, E.: Space-time multifractality of remotely sensed rainfall fields, J. Hydrol., 322, 2–13, 2006.
- Garraud, R.: Notes and correspondence Intraseasonal Variability of Moisture and Rainfall over the South American Altiplano, Am. Meteorol. Soc., 128, 3337–3346, 2000.
- Garraud, R., Vuille, M., and Clement, A.: The climate of the Altiplano: Observed current conditions and mechanisms of past changes, Palaeogeogr. Palaeoclimatol., 194, 5–22, 2003.
- Gires, A., Onof, C., Maksimovic, C., Schertzer, D., Tchiguirinskaia, I., and Simoes, N.: Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: A case study, J. Hydrol., 442–443, 117–128, 2012.
- Gires, A., Tchiguirinskaia, I., Schertzer, D., and Lovejoy, S.: Development and analysis of a simple model to represent the zero rainfall in a universal multifractal framework, Nonlin. Processes Geophys., 20, 343–356, doi:10.5194/npg-20-343-2013, 2013.
- Güntner, A., Olsson, J., Calver, A., and Gannon, B.: Cascade-based disaggregation of continuous rainfall time series: the influence of climate, Hydrol. Earth Syst. Sci., 5, 145–164, doi:10.5194/hess-5-145-2001, 2001.
- Hartkamp, A. D., De Beurs, K., Stein, A., and White, J. W.: Interpolation techniques for climate variables, CIMMYT Mexico, DF, 1999.
- Heidinger, H., Yarlequé, C., Posadas, A., and Quiroz, R.: TRMM rainfall correction over the Andean Plateau using wavelet multi-resolution analysis, Int. J. Remote Sens., 33, 4583–4602, 2012.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol., 25, 1965–1978, 2005.
- Holley, R. and Waymire, E.: Multifractal dimensions and scaling exponents for strongly bounded random cascades, Ann. Appl. Probab., 2, 819–845, 1992.
- Hutchinson, M. F.: Interpolating mean rainfall using thin plate smoothing splines, Int. J. Geogr. Inf. Syst., 9, 305–403, 1995.
- Hutchinson, M. F.: ANUDEM Version 5.2, Centre for Resource and Environmental Studies, Australian National University, Canberra, <http://fennerschool.anu.edu.au/research/products/anudem-vrsn-53> (last access: 6 July 2015), 2006.
- Jarvis, C. H. and Stuart, N.: A Comparison among Strategies for Interpolating Maximum and Minimum Daily Air Temperatures, Part II: The Interaction between Number of Guiding Variables and the Type of Interpolation Method, J. Appl. Meteorol., 40, 1075–1084, 2001.
- Jothityangkoon, C., Sivapalan, M., and Viney, N. R.: Test of a space-time model of daily rainfall in southwestern Australia based on nonhomogeneous random cascades, Water Resour. Res., 36, 267–284, 2000.
- Kahane, J. P. and Peyriere, J.: Sur certaines martingales de Benoit Mandelbrot, Adv. Math., 22, 131–145, 1976.
- Kang, B. and Ramirez, J. A.: A coupled stochastic space time intermittent random cascade model for rainfall downscaling, Water Resour. Res., 46, W10534, doi:10.1029/2008WR007692, 2010.
- Kedem, B. and Long, C.: On the lognormality of rain rate, P Natl. Acad. Sci. USA, 84, 901–905, 1987.
- Kummerow, C., Barnes, W., Kozu, T., Shiue, J., and Simpson, J.: The Tropical Rainfall Measuring Mission (TRMM) sensor package, J. Atmos. Ocean. Tech., 15, 809–817, 1998.
- Lenters, J. D. and Cook, K. H.: Summertime precipitation variability over South America: role of the large-scale circulation, Mon. Weather Rev., 127, 409–431, 1999.

- Lin, S. H.: More on rain rate distributions and extreme value statistics, *Bell Syst. Tech. J.*, 57, 1545–1568, 1978.
- Lovejoy, S. and Schertzer, D.: Multifractals and rain, New uncertainty concepts in hydrology and hydrological modeling, edited by: Kundzewicz, A. W., Cambridge Press, 62–103, 1995.
- Lovejoy, S. and Schertzer, D.: *The Weather and Climate: Emergent Laws and Multifractal Cascades*, The Cambridge University Press, Cambridge, 2013.
- Lovejoy, S., Pineda, J., and Schertzer, D.: The Global space-time Cascade structure of precipitation: satellites, gridded gauges and re-analyses, *Adv. Water Resour.*, 45, 37–50, 2012.
- Mandelbrot, B. B.: Intermittent turbulence in self-similar cascades: Divergence of high moments and dimension of the carrier, *J. Fluid Mech.*, 62, 331–358, 1974.
- Molnar, P. and Burlando, P.: Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model, *Atmos. Res.*, 77, 137–151, 2005.
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., and Veith, T. L.: *Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations*, T. ASABE, 50, 885–900, 2007.
- New, M., Lister, D., Hulme, M., and Makin, I.: A high-resolution data set of surface climate over global land areas, *Clim. Res.*, 21, 1–25, 2002.
- Olsson, J.: Evaluation of a scaling cascade model for temporal rain-fall disaggregation, *Hydrol. Earth Syst. Sci.*, 2, 19–30, doi:10.5194/hess-2-19-1998, 1998.
- Over, T. M.: Modeling space-time rainfall at the mesoscale using random cascades, PhD Dissertation, University of Colorado, Boulder, 1995.
- Over, T. M. and Gupta, V.: Statistical analysis of mesoscale rainfall: Dependence of random cascade generator on large-scale forcing, *J. Appl. Meteorol.*, 33, 1526–1542, 1994.
- Over, T. M. and Gupta, V. K.: A space-time theory of mesoscale rainfall using random cascade, *J. Geophys. Res.*, 101, 26319–26331, 1995.
- Pathirana, A. and Herath, S.: Multifractal modelling and simulation of rain fields exhibiting spatial heterogeneity, *Hydrol. Earth Syst. Sci.*, 6, 695–708, doi:10.5194/hess-6-695-2002, 2002.
- Perica, S. and Fofoula-Georgiou, E.: Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions, *J. Geophys. Res.*, 101, 26347–26361, 1996.
- Price, D. T., McKenney, D. W., Nalder, I. A., Hutchinson, M. F., and Kesteven, J. L.: A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data, *Agr. Forest Meteorol.*, 101, 81–94, 2000.
- Royer, J., Biaou, A., Chauvin, F., Schertzer, D., and Lovejoy, S.: Multifractal analysis of the evolution of simulated precipitation over France in a climate scenario, *C. R. Geosci.*, 340, 431–440, 2008.
- Samorodnitsky, G. and Taq, M.: *Stable Non-Gaussian Random Processes*, Chapman & Hall, New York, 1994.
- Sanabria, J., Marengo, J., and Valverde, M.: Climate change scenarios using regional models for the Peruvian Altiplano, *Revista Peruana Geo-Atmosférica Rpga*, 1, 134–149, 2009.
- Schertzer, D. and Lovejoy, S.: Physical modeling and analysis of rain and clouds by anisotropic scaling of multiplicative processes, *J. Geophys. Res.*, 92, 9693–9714, 1987.
- Schertzer, D. and Lovejoy, S.: Nonlinear variability in geophysics: multifractal analysis and simulations, in: *Fractals: Their physical origins and properties*, edited by: Pietronero, L., Plenum Press, New York, 49–79, 1989.
- Schmitt, F., Vannitsem, S., and Barbosa, A.: Modeling of rainfall time series using two-state renewal processes and multifractals, *J. Geophys. Res.-Atmos.*, 103, 23181–23193, 1998.
- Sharma, D., Das Gupta, A., and Babel, M. S.: Spatial disaggregation of bias-corrected GCM precipitation for improved hydrologic simulation: Ping River Basin, Thailand, *Hydrol. Earth Syst. Sci.*, 11, 1373–1390, doi:10.5194/hess-11-1373-2007, 2007.
- Stolle, J., Lovejoy, S., and Schertzer, D.: The stochastic multiplicative cascade structure of deterministic numerical models of the atmosphere, *Nonlin. Processes Geophys.*, 16, 607–621, doi:10.5194/npg-16-607-2009, 2009.
- Tachikawa, Y., Hiwasa, M., and Takara, K.: Spatial rainfall field simulation with random cascade introducing orographic effects on rainfall, in: *Proc. 2nd APHW International Conference*, Singapore, 2004.
- Tchiguirinskaia, I., Schertzer, D., Hoang, C. T., and Lovejoy, S.: Multifractal study of three storms with different dynamics over the Paris region, in: *12th International Conference on Urban Drainage*, Porto Alegre, Brazil., 2011.
- Thibeault, J., Seth, A., and Wang, G.: Mechanisms of summertime precipitation variability in the Bolivian Altiplano: present and future, *Int. J. Climatol.*, 32, 2033–2041, 2011.
- Vera, C., Baez, J., Douglas, M., Emmanuel, B. C., Marengo, J., Meitin, J., Nicolini, M., Nogues-Paegle, J., Penalba, O., Salio, P., Saulo, C., Silvia Dias, P., and Zipser, E.: The south American low-level jet experiment, *B. Am. Meteorol. Soc.*, 87, 63–77, 2006.
- Vuille, M.: Climate variability and high altitude temperature and precipitation, in: *Encyclopedia of snow, ice and glaciers*, Encyclopedia of Earth Science Series, Springer, the Netherlands, 153–156, 2011.
- Vuille, M., Bradley, R. S., Werner, M., and Keimig, F.: 20th century climate change in the tropical Andes: Observations and model results, *Climatic Change*, 59, 75–99, 2003.