High-resolution satellite-gauge merged precipitation climatologies of the Tropical Andes

Bastian Manz[1], Wouter Buytaert[1,2], Zed Zulkafli[1,3], Waldo Lavado[4], Bram Willems[5], Luis Alberto Robles[5], and Juan-Pablo Rodríguez-Sánchez[6]

[1]Department of Civil and Environmental Engineering, Imperial College London, London, U. K., [2]Grantham Institute for Climate Change and the Environment, Imperial College London, London, U. K., [3]Department of Civil Engineering, Universiti Putra Malaysia, Serdang, Malaysia, [4]Servicio Nacional de Meteorología e Hidrología (SENAMHI), Lima, Peru, [5]Laboratorio de Teledetección (LABTEL), Universidad Nacional Mayor de San Marcos, Lima, Peru, [6]Departamento de Ingeniería Civil y Ambiental, Universidad de los Andes, Bogotá, Colombia

Abstract Satellite precipitation products are becoming increasingly useful to complement rain gauge networks in regions where these are too sparse to capture spatial precipitation patterns, such as in the Tropical Andes. The Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar (TPR) was active for 17 years (1998–2014) and has generated one of the longest single-sensor, high-resolution, and high-accuracy rainfall records. In this study, high-resolution (5 km) gridded mean monthly climatological precipitation is derived from the raw orbital TPR data (TRMM 2A25) and merged with 723 rain gauges using multiple satellite-gauge (S-G) merging approaches. The resulting precipitation products are evaluated by cross validation and catchment water balances (runoff ratios) for 50 catchments across the Tropical Andes. Results show that the TPR captures major synoptic and seasonal precipitation patterns and also accurately defines orographic gradients but underestimates absolute monthly rainfall rates. The S-G merged products presented in this study constitute an improved source of climatological rainfall data, outperforming the gridded TPR product as well as a rain gauge-only product based on ordinary Kriging. Among the S-G merging methods, performance of inverse distance interpolation of satellite-gauge residuals was similar to that of geostatistical methods, which were more sensitive to gauge network density. High uncertainty and low performance of the merged precipitation products predominantly affected regions with low and intermittent precipitation regimes (e.g., Peruvian Pacific coast) and is likely linked to the low TPR sampling frequency. All S-G merged products presented in this study are available in the public domain.

1. Introduction

Characterizing the spatiotemporally highly variable nature of precipitation in mountain regions requires high accuracy of measurement at fine spatial and temporal scales, especially in tropical mountain environments such as the Tropical Andes [*Boers et al.*, 2013]. Although point-based rain gauges are considered to be an accurate reflection of local rainfall, gauge networks often do not have the necessary density to represent spatial rainfall patterns adequately [e.g., *Buytaert et al.*, 2010, 2006a]. Deployment of ground-based precipitation radars is restricted by high cost and affected by signal blockage by complex terrain [*Nikolopoulos et al.*, 2013; *Nesbitt and Anders*, 2009]. In this context satellite-based precipitation products (SPPs) have received widespread attention over the last decade for various hydrometeorological applications, such as hydrological modeling [*Zulkafli et al.*, 2014; *Jiang et al.*, 2012; *Li et al.*, 2009], geomorphology and landscape evolution [*Nesbitt and Anders*, 2009; *Bookhagen and Strecker*, 2008], streamflow forecasting [*Nikolopoulos et al.*, 2013; *Li et al.*, 2009], and early warning systems [*Tian et al.*, 2010] as well as investigations into atmospheric processes and storm structures [*Boers et al.*, 2014; *Mohr et al.*, 2014; *Boers et al.*, 2013; *Rasmussen et al.*, 2013; *Demaria et al.*, 2011].

The majority of SPP measurements stem from thermal infrared (IR) or passive microwave (PMW) sensors. IR-based sensors estimate precipitation rates as a function of cold cloud duration (cloud top temperature). This indirect and highly nonstationary relationship tends to result in poor performance under complex meteorological conditions [*Toté et al.*, 2015]. This applies particularly to tropical mountain regions where

infrared sensors have been shown to underestimate precipitation rates from deep convective systems [*Ward et al.*, 2011] as well as warm orographic clouds [*Dinku et al.*, 2010]. PMW sensors rely on a better physical relationship by deriving rainfall estimates from ice scattering in clouds. However, both stratiform and warm rain clouds, which produce little or no ice scattering, are underestimated and cold ground surfaces may be mistaken for rainfall [*Dinku et al.*, 2010]. Furthermore, PMW-based products have been shown to underestimate heavy rainfall, possibly due to nondetection of small-scale convective rainfall [*Thiemig et al.*, 2012].

While the first Global Precipitation Measurement (GPM) products are now being released, the single satellite-based active radar precipitation sensor with long-term historical observations to date is the Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar (TPR). TPR operation recently ended in October 2014 (excluding a brief restart in early 2015 at lower altitude) due to the descent of the TRMM satellite. The high resolution of the sensor (5 km) and length of the TPR record (1998–2014) provide an opportunity to generate high-resolution precipitation climatologies to capture highly spatially variable precipitation patterns in data-sparse tropical regions such as the Tropical Andes. Low-density ground observations of precipitation in these regions preclude a simple interpolation of the rain gauges that would be feasible in high-rain gauge density locations such as in the Alps [see, for example, *Isotta et al.*, 2014]. However, there are multiple limitations associated with TPR measurements, which suggest that a simple reprojection of the orbital TPR scans onto a regular grid may not provide the best estimate of the *true* precipitation climatology. TPR measurements are affected by limitations in rainfall detection by the sensor, simplifying assumptions in the retrieval algorithm and also the low temporal sampling frequency of the TPR.

The TPR detection limit due to instrument sensitivity lies at 18 dBZ, limiting detection of light rainfall to roughly 0.5 mm hr^{-1}, which is estimated to result in an underestimation of long-term rainfall totals by about 1% in humid regions, but by up to 20% in dry regions [*Yang and Nesbitt*, 2014]. Recent validation of the TPR against ground-based radar has affirmed that the largest contribution to rain rate errors stems from the no rain transition with the TPR also experiencing rainfall detection issues when the field of view is filled less than 70% [*Kirstetter et al.*, 2014]. Nonuniform beam filling introduces errors into the reflectivity estimates and hence the rainfall classification. *Kirstetter et al.* [2014] have empirically shown that up to 40% of stratiform rainfall may be misclassified as convective as a result of low field-of-view (FOV) filling.

Aside from failure to detect light rainfall, the TPR strongly underestimates heavy rainfall. *Rasmussen et al.* [2013] have shown that the TPR retrieval algorithm (2A25) underestimates precipitation rates in deep convective systems containing significant mixed phase or frozen hydrometeors. Consequently, the TPR bias for convective rainfall is higher for the Andes, where a larger fraction of convective rainfall stems from deep convective systems, than for the Amazon basin, where horizontally intense storms occur more frequently [*Rasmussen et al.*, 2013]. Recent changes in the TRMM 2A25 algorithm from versions 6 to 7 include an improved elevation map for the Andes and a search algorithm to minimize ground clutter as well as an improved Z-R relationship based on a nonspherical rain drop distribution to increase estimates of heavy rainfall rates [*TRMM Precipitation Radar Team*, 2011; *Iguchi et al.*, 2009]. However, arguably both the stratiform and convective profiling algorithms in 2A25 do not sufficiently represent the dynamics of extreme rainfall [*Kirstetter et al.*, 2014]. In general, the quantification of rainfall rates by the TPR is subject to significant errors due to radar signal attenuation, nonuniform beam filling, and inaccurate Z-R conversion with the combined effects of these error sources often difficult to distinguish based on analysis of the rain rate [*Kirstetter et al.*, 2014, 2012].

Nonetheless, multiple studies have concluded that the TPR delivers accurate estimates of mean rainfall [*Duan et al.*, 2015; *Chen et al.*, 2013; *Nesbitt and Anders*, 2009]. Hence, the TPR can be considered the single most reliable overland precipitation sensor on-board TRMM [*Yang and Nesbitt*, 2014; *Stephens and Kummerow*, 2007] and the high spatial resolution of 5 km [*Seto et al.*, 2013] makes it one of the most suitable data sets for generating high-resolution precipitation climatologies [*Biasutti et al.*, 2012; *Nesbitt and Anders*, 2009]. However, the spatially restricted field of view (FOV) by the TPR of 247 km means a poor temporal sampling frequency with coverage of a particular location only every 1–2 days, which results in considerable sampling error [*Nesbitt and Anders*, 2009]. The impact of this sampling error on the calculation of long-term climatological rainfall can be modeled as a function of the rainfall rate, the number of samples, and the spatial scale [*Nesbitt and Anders*, 2009; *Indu and Nagesh Kumar*, 2014; *Iida et al.*, 2010, 2006]. The rainfall rate to sampling error relationship is known to be noncoherent across spatial scales [*Nesbitt and Anders*, 2009; *Iida et al.*, 2006].

Many studies have shown that local gauge-calibration can significantly improve satellite estimates [e.g., *Cheema and Bastiaanssen*, 2012; *Condom et al.*, 2011; *Vila et al.*, 2009; *Lavado Casimiro et al.*, 2009]. Many end

user (level 3) satellite products (e.g., TRMM 3B42, CHIRPS, TAMSAT among others) already include a first stage of internal gauge correction, mostly based on simple statistical methods such as mean field bias correction [*Huffman et al.*, 2007] or inverse error weighting [*Grimes et al.*, 1999]. Other satellite-gauge merging methods have used a combination of multiplicative and additive bias correction [*Vila et al.*, 2009], linear regression models [*Almazroui*, 2011], residual inverse distance weighting [*Dinku et al.*, 2014], and copula models in combination with satellite-gauge biases [*Moazami et al.*, 2014]. In particular, geostatistical methods have been successfully used for merging of satellite and gauge rainfall estimates [*Grimes and Pardo-Igúzquiza*, 2010]. For example, satellite rainfall data can be incorporated in Kriging with external drift (KED) to interpolate between gauges [*Grimes et al.*, 1999]. *Álvarez-Villa et al.* [2011] explored various Kriging methods to merge TPR and gauge data across Colombia and demonstrated that KED outperformed standardized co-Kriging, colocated co-Kriging, and Markov coregionalization co-Kriging. Recently, *Nerini et al.* [2015] employed a Bayesian Combination method, originally proposed by *Todini* [2001] for merging (urban) rainfall radar and rain gauges, to combine satellite and gauge estimation uncertainties to minimize overall uncertainty of the merged product. For mean climatological TPR estimates, merging with gauge data can allow for correction of the errors induced by the sampling frequency, sensor limitations, and the retrieval algorithm.

This study takes full advantage of the entire TPR data set (1998–2014) to generate high-resolution (5 km) mean monthly climatologies for the highly variable precipitation patterns in the data-sparse Tropical Andes and using multiple satellite-gauge (S-G) merging procedures to identify the optimal combination structure. The objective is that the merged 5 km monthly climatological maps will constitute an improved source of climatological rainfall data for regions with complex precipitation regimes compared to gauge-only or TPR-only climatologies. The derivation of mean monthly precipitation climatologies from the gauges and TPR is presented along with the S-G merging methods (section 2), and the resulting merged precipitation products are evaluated at the point scale (cross validation) against gauge observations and integrated across catchments against discharge observations from 47 tropical catchments (section 3). Spatial and seasonal precipitation patterns are discussed in the context of the merging methods, TPR properties, and climatic controls (section 4). The resulting climatologies are available for research use at doi.org/10.5285/74a588cc-723c-4a35-ac0c-223f5b92ee36.

2. Data and Methods

2.1. Study Area: Precipitation Patterns of the Tropical Andes

The study area extends from 19°S to 12°N and from 67°W to 81.5°W, covering a climatically diverse region from the Guajira region in northern Colombia (average annual rainfall of 300 mm year^{-1}) to the Altiplano in western Bolivia (below 1500 mm year^{-1}), the humid upper Amazon and Orinoco basins (2000–5000 mm year^{-1}), the Peruvian Pacific coast (below 100 mm year^{-1}) and the Colombian Pacific coast (above 10,000 mm year^{-1}; see Figure 1).

Precipitation patterns are controlled by the interaction of synoptic-scale atmospheric currents and the complex Andean topography. Easterly trade winds resulting from the southerly position of the Intertropical Convergence Zone (ITCZ) during the South American monsoon season (austral summer) transport moist air from the tropical Atlantic over the Amazon basin [*Boers et al.*, 2013; *Carvalho et al.*, 2011] and are blocked by the topographic barrier of the Tropical Andes [*Romatschke and Houze*, 2010]. The deflection of air moisture to the southeast gives rise to the South American Low-Level Jet (SALLJ) that transports air moisture along the eastern Andes into the La Plata Basin [*Boers et al.*, 2013]. The amplified state of the SALLJ, characterized by intensified southward low-level wind (also termed Chaco jet), during the South American monsoon coincides with a suppression of the South Atlantic Convergence Zones and increased convection over the La Plata Basin [*Marengo et al.*, 2012; *Vera et al.*, 2006]. Thus, in the region east of the Andes between the equator and 25°S more than 50% of total annual precipitation occur in the austral summer [*Marengo et al.*, 2012].

Along the eastern flanks of the Andes the easterly trade winds and strong topographic gradients result in pronounced orographic effects [*Espinoza et al.*, 2015; *Espinoza-Villar et al.*, 2009a; *Bookhagen and Strecker*, 2008]. This causes deep convection [*Romatschke and Houze*, 2010] and thereby spatiotemporally highly intermittent precipitation patterns with precipitation gradients of up to 190 mm km^{-1} [*Espinoza et al.*, 2015].

Intra-Andean valleys are generally drier (average annual rainfall below 1000 mm year^{-1}) than the Amazon basin and the eastern Andean flanks (2000–5000 mm year^{-1}) but exhibit highly variable rainfall regimes due to the complex Andean topography that can result both in rainfall shielding as well as concentration of
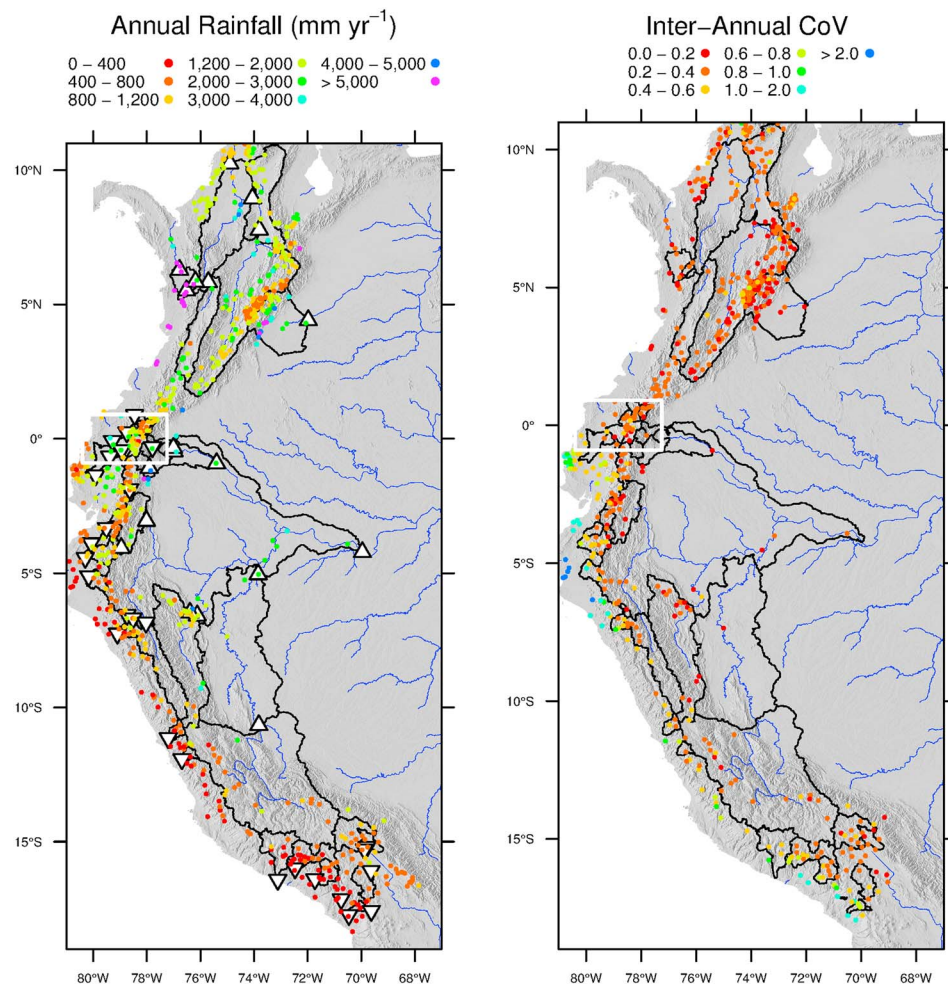
Figure 1. Topographical map of Tropical Andes with rain gauges showing (left) annual precipitation totals (in mm year^{-1}) and (right) the coefficient of variation of the total annual rainfall. Catchments considered in the hydrological analysis are delineated in black and the respective river network in blue. Discharge gauges at catchment outlets are marked by white triangles (upward and downward facing for the Caribbean/Amazonian and Pacific/Andean regions, respectively). The white rectangle delineates the region shown in Figure 4.

precipitation [*Buytaert et al.*, 2006a]. However, annual rainfall totals in intra-Andean valleys are often dominated by convective events [*Mohr et al.*, 2014; *Rasmussen et al.*, 2013].

The northern Tropical Andes exhibit pronounced bimodal precipitation regimes with precipitation maxima in spring and autumn generated by the biannual passage of the ITCZ [*Álvarez-Villa et al.*, 2011]. Furthermore, moisture transport by multiple low-level tropical jets affects the hydrometeorology of the northern Tropical Andes. The Chocó jet results in extremely high annual rainfall along the Colombian Pacific coast exceeding 10,000 mm year^{-1} and transports moisture further inland into the intra-Andean valleys in central Colombia [*Poveda et al.*, 2011]. By contrast, the Peruvian coastline is characterized by very dry conditions as a result of the cold von Humboldt current.

On interannual timescales the El Niño–Southern Oscillation (ENSO) is the major driver of precipitation variability, resulting in regionally contrasting impacts due to the interaction with Andean topography and low-level jets [*Marengo et al.*, 2013]. During the ENSO warm phase rainfall is increased along the dry Peruvian coast, while the Colombian Andes and Caribbean basin experience drier conditions [*Poveda et al.*, 2011]. Easterly trade winds over the Amazon are intensified and consequentially the SALLJ strengthens leading to increased southward moisture transport along the eastern Andes [*Marengo et al.*, 2012]. During ENSO cold phases general regional conditions are reversed.

2.2. Rain Gauge Data

A database of 735 rain gauges was aggregated from quality-checked records by the national meteorological services of Bolivia (SENAMHI), Colombia (IDEAM), Ecuador (INAMHI), and Peru (SENAMHI) as well as the National Climate Data Center (NOAA NCDC) and the Observation Service SO HYBAM. All station records are at least 90% complete at monthly resolution over the 30 year period from 1981 to 2010. To validate the spatial consistency among gauges, the mean relative error (mean bias) between the monthly estimate for each gauge and the monthly estimate of the surrounding five gauges was calculated over the entire 30 year time series. The relative error only acts as a coarse indicator of spatial consistency as it does not account for climatic or orographic differences between gauge locations, which may be significant where gauge density is low. However, 54 (45) gauges with a relative error $>50\%$ ($<-50\%$) were flagged. All 99 gauges were manually inspected with respect to location, elevation, aspect, distance to neighboring gauges and general climate and only those with major relative errors that could not be explained by surrounding precipitation patterns or topography were removed, which led to the elimination of 12 gauges, resulting in a final database of 723 gauges for the S-G merging (see Figure 1). The final set of rain gauges was comprised of 455 gauges located at elevations above 1000 m asl in the Tropical Andes. Of those located below 1000 m asl, 49 are in the upper Amazon basin, 120 in catchments draining into the Caribbean Sea, and 99 in catchments draining into the Pacific Ocean. Finally, mean climatological estimates for each calendar month were derived by averaging the observations of the respective month from all years over the entire 30 year period.

In addition to the rain gauges, discharge records from 47 catchments across the Tropical Andes (see Figure 1), ranging in size from small high mountain catchments (100 km^2) to large continental catchments ($>$100,000 km^2) were assembled for hydrological evaluation of the S-G merging methods (see Figure 2.6 for details). The discharge time series were aggregated and the climatological average annual discharge was determined.

2.3. TPR Data

Satellite rainfall data were obtained from instantaneous near-surface rainfall estimates of the orbital TRMM PR product 2A25 version 7 [*Iguchi et al.*, 2000, 2009] over the entire TRMM observational period (1998–2014). In order to derive 5 km mean monthly climatologies from the instantaneous TPR observations, individual TPR overpasses were reprojected first onto a regular-spaced 1 km grid and then reaggregated to a 5 km grid. For spatial coherence, a metric projection Universal Transverse Mercator (UTM) was used. In this process, for each TPR overpass the 1 km grid cells were assigned the observation of the overpassing 5 km pixel within the TPR swath. Initial reprojection to a 1 km grid has two benefits. First, using a higher resolution target grid allows better delineation of the overpassing TPR pixel onto the regular grid as it minimizes issues of partially filled grid cells. Second, even though for individual time steps the values of neighboring grid cells (within the same 5 km pixel) will be identical, the trajectory of successive TPR overpasses changes over 46 days until a cycle is complete. As a result, the full TPR time series will differ for neighboring 1 km grid cells and vary in length over the total period (1998–2014) from approximately 1900 instantaneous observations at the equator to 2400 at 20°N/S. For each 1 km grid cell these time series were subsequently averaged analogous to the gauge data to derive mean climatological estimates for all 12 calendar months. However, as the original TPR estimates represent average rainfall rates across a 5 km resolution, the climatological means at 1 km should be seen as best estimates of a spatially moving window of 5 km. Consequently, in order for consistency between the spatial scale and the rainfall estimates, the climatological estimates are aggregated back to 5 km. As the TPR reports near-surface rainfall rates in millimeters per hour, the mean monthly climatological estimates were upscaled by the number of hours per calendar month to monthly totals (mm).

2.4. Evaluation of the Climatological Representativeness of the TPR

Climatological precipitation averages are typically defined over a period of 30 years or more, which exceeds the lifetime of the TRMM satellite (17 years). To ascertain whether the gauge observational period (1981–2010) and the satellite record (1998–2014) are suitable for merging, the interannual variability of precipitation totals was investigated and two statistical tests (Mann-Kendall and Kolmogorov-Smirnov) were applied to compare the gauge data over the two periods. Figure 1 shows the coefficient of variation for annual precipitation totals over the period 1981–2010, clearly identifying the Peruvian Pacific coast, in particular the northern Piura region and southwestern Ecuador as the regions showing most variability. It is well known that this region is most directly affected by the interannual ENSO cycles, and this region shows the highest positive correlation between total annual precipitation and the Nino3.4 Index (see Figure 2). Negative correlation is most pronounced in the Colombian Andes, but the resulting variability in precipitation totals is much lower
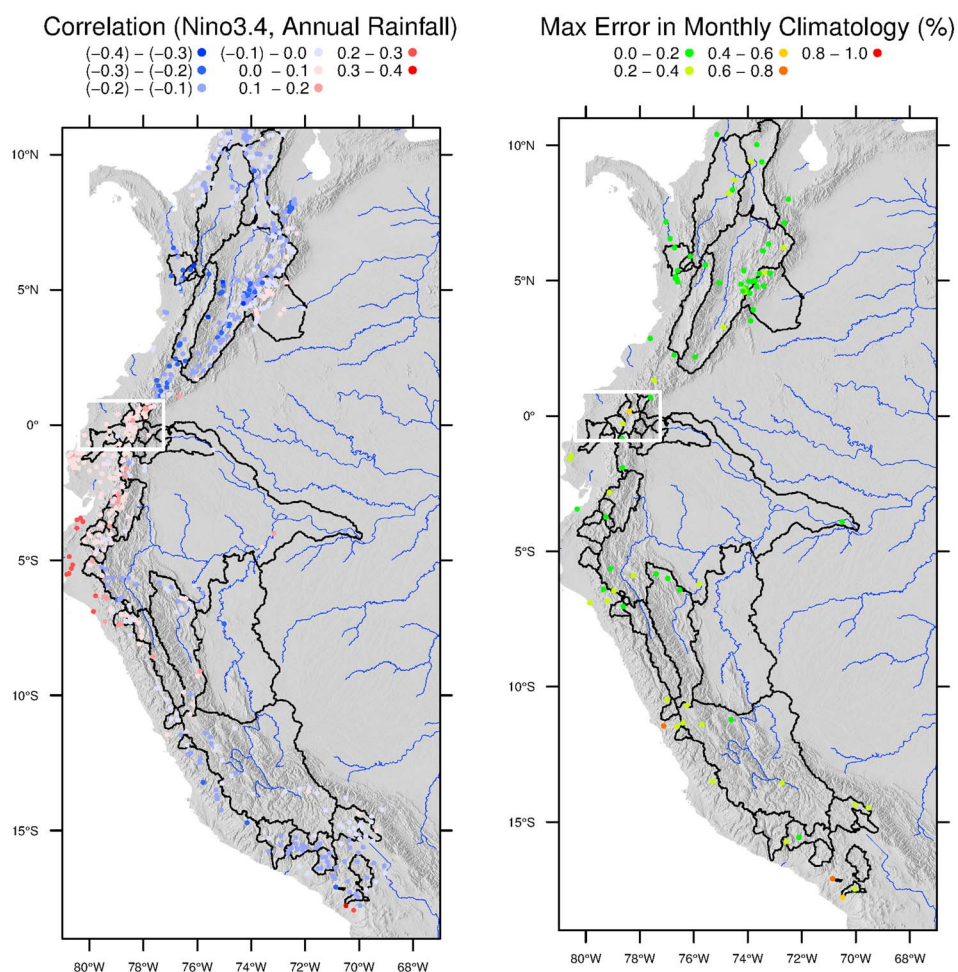
Figure 2. Same as Figure 1 but showing (left) the linear correlation (ρ) between monthly gauge rainfall and the Nino3.4 index as well as (right) the maximum relative difference in climatological mean monthly rainfall between the periods 1981–2010 and 1998–2014 for all gauges flagged by the M-K and K-S tests.

(see Figure 2). In order to identify nonstationarities in the mean of nonnormally distributed rainfall data, the rank-based nonparametric Mann-Kendall (M-K) test is suitable [*Espinoza-Villar et al.*, 2009b]. Where gauge data were available until 2014 (603 gauges) the M-K test was applied to find a trend in the data over the period 1981–2014. At the 95% confidence level 68 (17) gauges showed a weak positive (negative) trend, defined as $\tau > 0.1$ ($\tau < -0.1$). Second, to evaluate if the statistical distribution of these 603 gauges differed between the periods 1981–2010 and 1998–2014 the Kolmogorov-Smirnov (K-S) test was performed. The K-S test is considered adequate in these conditions because it is also nonparametric and makes no assumptions about the statistical distribution of the rainfall data [*Tarnavsky et al.*, 2012]. At the 95% confidence level 27 gauges indicated differences in the empirical distributions between the two periods (D statistic > 0.05) out of which 22 have also been flagged by the M-K test. For all 90 gauges flagged by either the M-K or K-S test, the difference in mean monthly climatologies between 1981–2010 and 1998–2014 was computed. The maximum difference (irrespective of calendar month) ranged 16.3% ± 13.2% (median ± sd). Figure 2 shows that there is no clear spatial pattern or clustering in this gauge set that indicates a direct link to the ENSO correlation or the patterns of the coefficient of variation. It is therefore important to note that the climatological estimates presented in this study only represent mean conditions that are subject to interannual variability, which may be significant locally. Furthermore, the use of different but overlapping periods for the gauges and TPR will be responsible for some limited level of difference between gauge and TPR estimates, but there is no clearly defined spatial region where this is expected to result in major disagreement between the data sources. Given these conditions, the assumption is made that the average TPR estimates over 1998–2014 are suitable descriptors

of the spatial patterns of climatological precipitation across the Tropical Andes and therefore appropriate for merging with the available gauge records.

2.5. S-G Merging Techniques

In this study we compared five S-G merging methods: (1) linear modeling/mean multiplicative bias correction (LM), (2) residual inverse distance weighting (RIDW), (3) ordinary Kriging (OK), (4) residual ordinary Kriging (ROK), and (5) Kriging with external drift (KED). In the context of these methods OK represents a gauge-only interpolation with no satellite data (gauge-only benchmark). The satellite-only benchmark (hereafter *TPR*) is simply the gridded 5 km TPR data, as described previously. All merging techniques are implemented for each climatological month separately.

2.5.1. Linear Modeling (LM)

Satellite estimates are extracted at the gauge locations and a linear regression model is fitted using least squares estimation across all gauges to explain gauge observations in terms of satellite estimates:

$$Z_g = aZ_{s,g} + b, \tag{1}$$

where Z_g is the gauge values and $Z_{s,g}$ the satellite estimates at the gauge locations. The linear model can be seen as a combination of a multiplicative bias correction (via the coefficient; a) and an additive bias correction (via the intercept; b). Exploratory analysis of the coefficient of determination showed that fixing the intercept ($b = 0$) improved the performance of the linear model. This setup was therefore used in the S-G merging. Thereby, the linear model is effectively a multiplicative bias correction.

2.5.2. Residual Inverse Distance Weighting (RIDW)

The difference (residual) between satellite estimates and gauge observations is computed at each gauge location. These residuals are interpolated using inverse distance weighting (IDW) and the interpolated residual surface is added back on to the satellite estimates (for further details on the method see *Dinku et al.* [2014]). In contrast to Kriging methods IDW does not explicitly consider the spatial error structure but simply interpolates as a direct function of distance. Nonetheless, it has been shown that the approach performs comparable to regression Kriging for complex terrain if a high gauge network density is available [*Dinku et al.*, 2014].

2.5.3. Ordinary Kriging (OK)

Kriging interpolates rain gauge locations using a Gaussian process governed by an a priori determined spatial covariance structure. OK provides the best unbiased linear estimate of the ungauged locations. Hereby, the assumptions are made that (a) the interpolated rainfall can be characterized as a stationary random variable with unknown mean and (b) the rainfall can be adequately represented by a normal (Gaussian) distribution. To meet the second assumption gauge rainfall estimates are converted to a normal distribution using a normal score transform and back transformed after interpolation.

In this study, isotropic semivariograms are used to characterize the spatial covariance between different gauge locations. Thereby, variance is explained solely as a function of distance (not location), thus assuming stationarity of the variance of the differences separated by the same distance (see Figures S1–S4 in the supporting information of this article for semivariograms for the different Kriging methods). For further details on the Kriging methods please see section 1 of the supporting information to this article as well as *Grimes and Pardo-Igúzquiza* [2010] and *Goovaerts* [1999].

2.5.4. Residual Ordinary Kriging (ROK)

In ROK, both gauge and satellite data are transformed to a normal state and a linear model is fitted to explain gauge observations as a function of colocated satellite estimates. The residuals of this model are subsequently interpolated onto the whole grid using OK, as described above, before they are added back to the linear model estimate for each grid cell. The combined estimate (linear model + residuals) is then back transformed.

2.5.5. Kriging With External Drift (KED)

In contrast to OK, KED assumes nonstationarity that can be represented by a secondary variable (i.e., TPR). The choice and suitability of the secondary variable can be determined through spatial correlation analysis between the variable and the gauge observations. For the TPR satellite data the Pearson correlation coefficient (ρ) ranged 0.68–0.85 over the 12 climatological months (annual $\rho = 0.82$).

In a separate KED merging the normalized difference vegetation index (NDVI) from the Moderate Resolution Imaging Spectroradiometer (MODIS) was employed as an additional drift term in combination with TPR (hereafter KED_TN for KED-TPR-NDVI). Vegetation response (in terms of NDVI) has been shown to be a suitable

indicator of precipitation on monthly timescales at lags of up to 2–3 months [*Hunink et al.*, 2014; *Duan and Bastiaanssen*, 2013; *Immerzeel et al.*, 2009]. In the present study, correlation was maximized at a lag of 2 months with monthly ρ values of 0.37–0.73 (annual $\rho = 0.71$). More detailed descriptions of implementing KED for S-G merging can be found in *Nerini et al.* [2015], *Álvarez-Villa et al.* [2011], and *Grimes and Pardo-Igúzquiza* [2010].

A limitation to Kriging approaches arises when there are large ungauged regions as the Kriging estimate converges to the mean of the rainfall field at distances beyond the range of the semivariogram. Therefore, OK results beyond the maximum semivariogram range (approximately 185 km) are removed. Estimates over ocean (irrespective of gauge distance) are removed for all methods except TPR.

2.6. Evaluation of Merging Techniques

As no independent distributed data set of precipitation observations exists at such a high resolution to validate independently the S-G merging techniques, performance of the merged rainfall products was assessed using a leave-one-out cross validation (LOOCV) of the gauges. In this method a single gauge is removed prior to the interpolation and the prediction at the gauge location is compared to the gauge observation. This is repeated for all gauges and the goodness of fit between the merged estimates and the gauge observations was evaluated in terms of mean error (ME), root-mean-square error (RMSE), and relative bias (expressed in %). These statistical performance scores, in turn, were analyzed both in terms of their spatial patterns over the study extent as well as the intra-annual variations in the performance of the different S-G methods. While LOOCV can quantify the prediction error for every gauge location, in regions of high gauge density gauges can be highly correlated and so performance is likely to be high irrespective of the interpolation method used. In order to assess the performance of the merging methods at varying gauge densities, a second cross-validation approach as outlined by *Chen et al.* [2008] was implemented. Here 10% (72) of the stations are removed from the data set for validation only. S-G merging is performed repeatedly for all methods using 100% (651), 50% (326), 20% (130), 10% (65), and 5% (33) of the remaining stations and evaluated against observations at the previously removed validation sites.

Lastly, evaluation of individual gauge locations only provides a performance indicator at the point scale but no information on the performance of the merged estimates integrated over larger areas. Therefore, the terrestrial water balance is assessed using simple runoff ratios. The runoff ratio (RR) of a river basin is defined as

$$RR = \frac{Q}{P}, \tag{2}$$

where Q is the climatological average annual discharge observed at the catchment outlet and P is the climatological average annual precipitation over the basin as estimated by the various S-G techniques over the entire study period. For OK, RR estimates are removed for catchments where more than 10% of the surface area is not predicted (as it is outside the semivariogram range distance). The runoff ratio is a common hydrological indicator that summarizes the long-term catchment hydrological property in terms of the propensity to generate runoff, with the assumption that the change to catchment storage is zero over a sufficiently large time period (hence, 1 − RR equals the evapotranspirative flux). In order to have a quantitative reference for each catchment, RR estimates were compared to estimates based on the Budyko curve *Budyko* [1974], which relates runoff ratio and catchment aridity index, i.e., ratio of potential evapotranspiration to precipitation [for details on the derivation of these reference values see *Buytaert and De Bievre* [2012].

3. Results

3.1. Spatial Rainfall Patterns

A visual inspection of the monthly precipitation climatologies Figure 3 shows pronounced differences in the way that climatological features are represented by the different methods. While synoptic processes such as the double passage of the ITCZ and the regional precipitation maximum along the Colombian Pacific coast are discernible from all products, the way precipitation gradients are defined varies strongly between the different methods. For OK the lack of spatial support from the TPR radar results in a poor representation of spatial precipitation patterns in regions of low gauge density as reflected in very smooth precipitation gradients across the region. For instance, the method fails to represent the extended orographic barrier known to exist along the eastern slope of the Andes between 7°S and 15°S [*Espinoza et al.*, 2015; *Bookhagen and Strecker*, 2008].
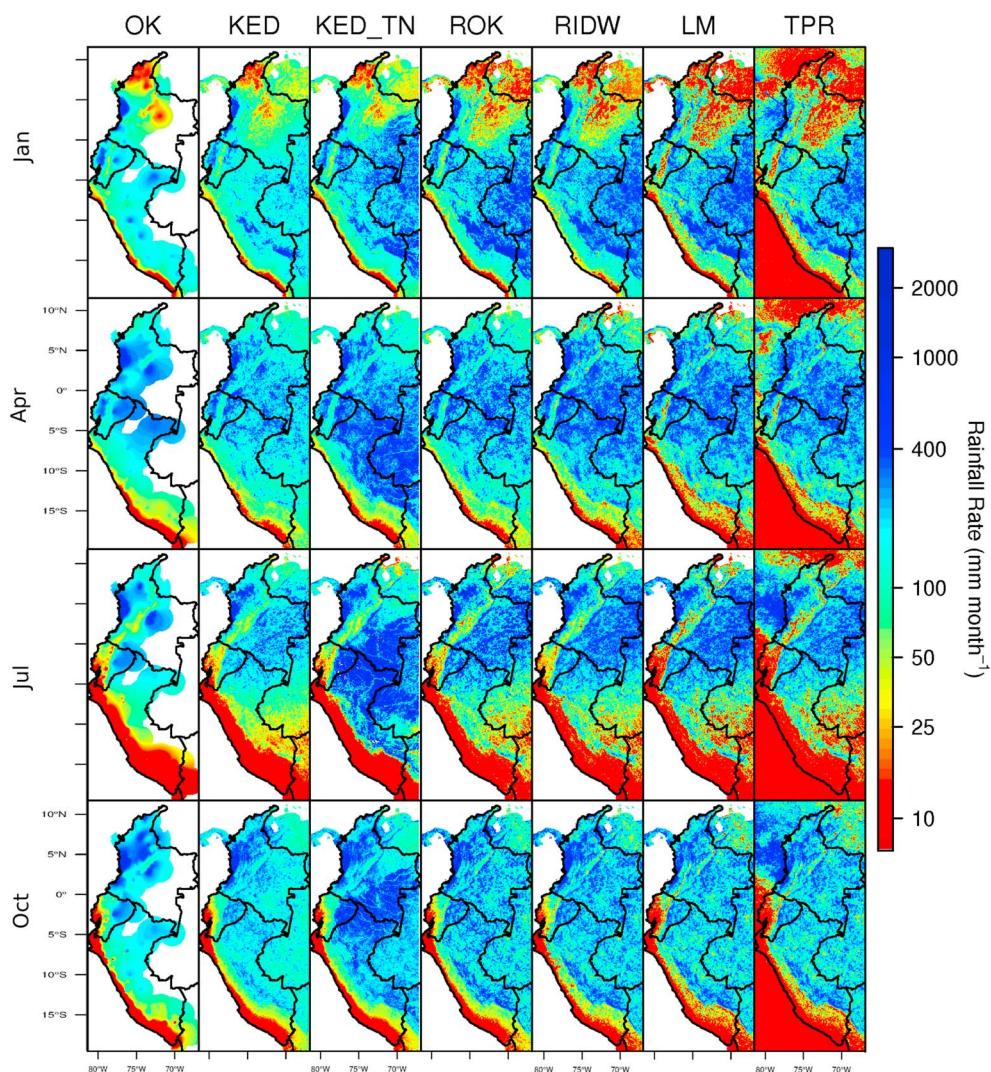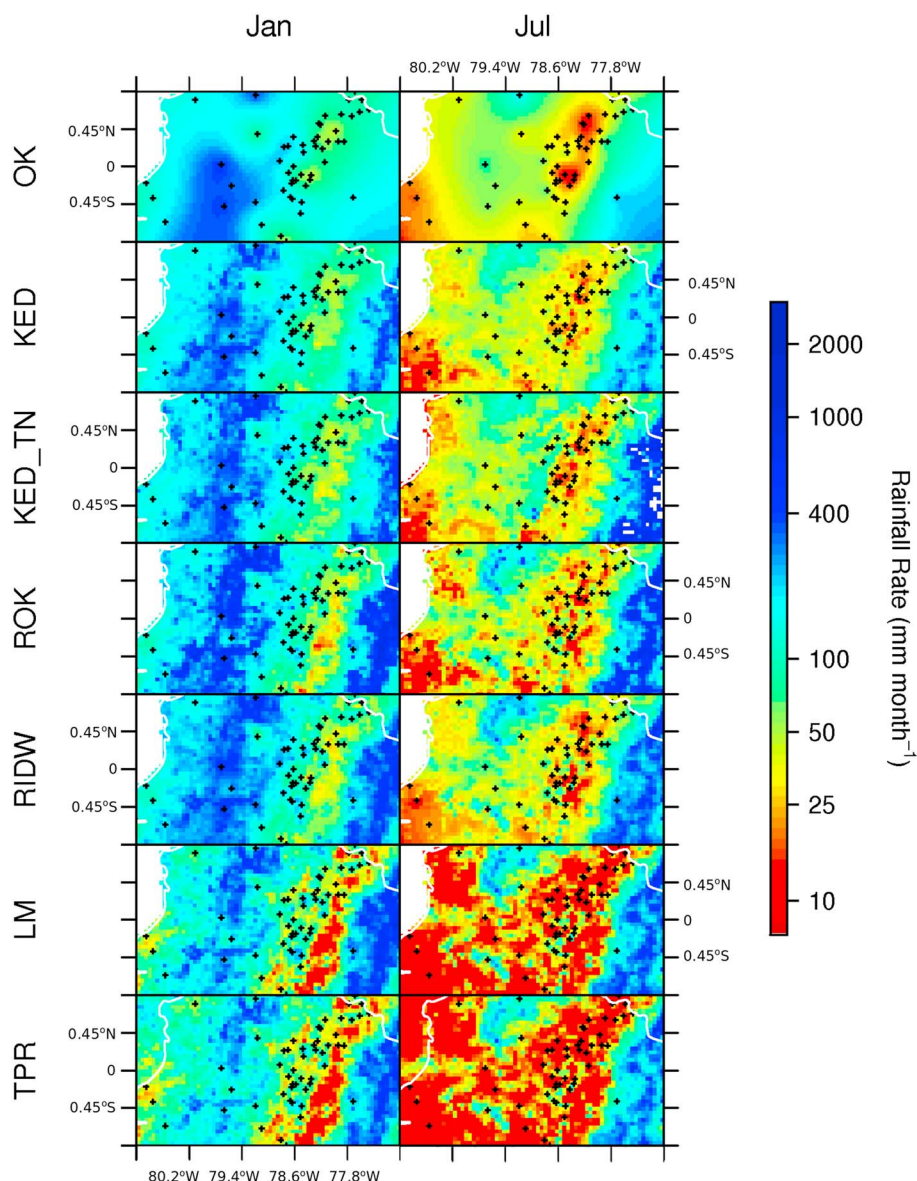
Figure 3. Merged monthly precipitation fields (climatological average) for OK, KED, KED_TN, ROK, RIDW, LM, and TPR in millimeters per month for the months of January, April, July, and October.

In contrast, the TPR spatial patterns are generally maintained by the S-G merging methods. Among these absolute rainfall totals in high rainfall regions (>1000 mm year^{-1}) are lowest for KED, similar in magnitude to OK, and highest for KED_TN with the remaining methods in a range between these. These variations are more evident over the upper Amazon basin compared to the Andes: they are most pronounced during the northerly ITCZ position in July when KED_TN does not show a similar degree of rainfall reduction in the southern Amazon basin and a wider band of high precipitation rates along the eastern Andean slopes compared to the other merging methods. While KED preserves the absolute magnitude of the rain gauge totals similar to OK, KED_TN shows a much wetter upper Amazon basin from April to October (see also Figure S5 in the supporting information of this article). The results suggest that including NDVI as an additional variable in combination with TPR has profound influence on the rainfall estimates by limiting the northward progression of the rainfall field during the northerly position of the ITCZ. In the Peruvian Andean region LM and TPR show very low rainfall totals with widespread underestimation relative to OK throughout the year. This behavior is far less pronounced in the remaining S-G merging methods.

These spatial patterns are reflected at the local scale (Figure 4). In January TPR (and LM) shows a strong precipitation gradient from a high rainfall region (>400 mm) in the east (upper Amazon basin) to a low rainfall region (<25 mm) along the eastern Andes (the high gauge density area roughly corresponds to the intra-Andean region in central Ecuador). The other S-G merging methods maintain the spatial rainfall patterns but rainfall

Figure 4. Merged monthly precipitation fields as in Figure 3 focusing on a section of the Tropical Andes in Ecuador, as a high-resolution example (location shown in Figure 1).

magnitudes are higher with only isolated regions of less than 25 mm. In contrast to all other methods OK does not estimate any rainfall below 50 mm in January and the clustered gauge network does not allow for replication of the strong precipitation gradient along the eastern Andes. In July TPR and LM suggest that large regions along the coast and in the central Andes are very dry (<10 mm). On the other hand, the S-G merging methods show a rainfall field similar to OK with more isolated regions of low rainfall. However, OK again shows the most gradual precipitation gradient in the eastern Andes, followed by KED with the most pronounced gradients evident in KED_TN and ROK.

3.2. Rain Gauge Cross Validation

The spatial distribution of relative bias based on LOOCV (Figure 5) suggests that the gridded TPR data without gauge correction underestimates total monthly gauge rainfall, most noticeably in the intra-Andean region south of the equator (south of the equator median $\pm$ sd of relative bias: -32 ± 74%), gradually improving toward the wetter northern Tropical Andes (north of the equator: -12 ± 28%), although regions with orographic enhancement such as along the Colombian Pacific coast and toward the Amazon show high levels of underestimation. LM showed a similar pattern with major negative bias in the Altiplano and southern
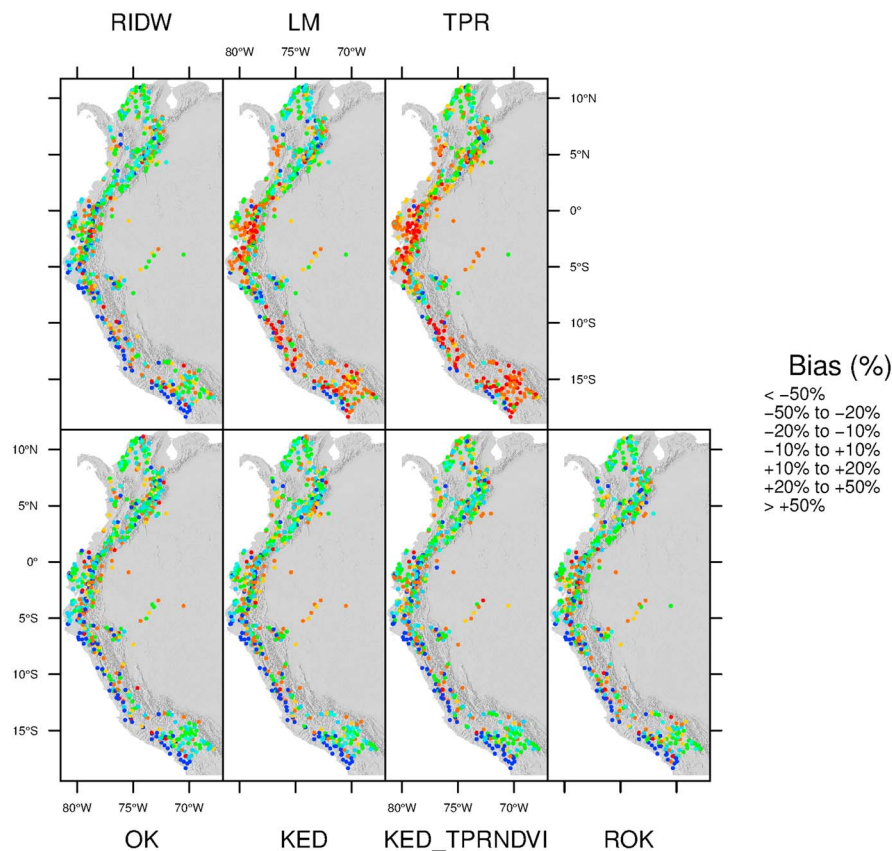
Figure 5. Average cross-validation bias (in %) between merged precipitation products and gauge observations (relative to gauge observations) over the 12 climatological months.

Tropical Andes in general ($-21\pm88\%$) but a more improved and balanced bias in the northern Tropical Andes ($0.8\pm33\%$). The remaining S-G merging methods, RIDW (south of equator/north of equator median $\pm$ sd: $5.9\pm57\%/1.4\pm29\%$), ROK ($2.0\pm53\%/0.55\pm33\%$), KED ($7.6\pm83\%/3.7\pm30\%$), KED_TN ($8.9\pm66\%/3.2\pm28\%$) as well as OK ($8.3\pm81\%/3.1\pm34\%$) showed highly similar spatial performance patterns with very few and isolated stations reporting relative bias below -50% and more comparable performance between the northern and southern Tropical Andes. However, the most notable feature across all these methods, which explains the high standard deviation of the bias in the southern Tropical Andes, is the strong overestimation of rainfall along the Peruvian Pacific coast that is characterized by low rainfall totals (compare Figure 1). By contrast, high-Andean locations close by that are characterized by much higher rainfall totals show a negative bias. This suggests that S-G merged products, despite reducing the bias range, cannot capture entirely the high spatial variability of rainfall patterns in the intra-Andean valleys and along the Peruvian coast that is characterized by a dry, highly intermittent rainfall regime. Over the Altiplano, in the Amazon basin and generally north of the Equator, the bias range for all S-G methods bar LM is much lower than that for the uncorrected TPR.

Temporal analysis of the cross-validation results based on LOOCV (Figure 6) confirms the systematic underestimation of gauge observations by TPR and LM by a large RMSE and a highly negative ME in both cases. The remaining S-G methods are quite consistent in terms of mean error but show strong seasonal variations in terms of RMSE and rRMSE. Absolute random errors (RMSE) are minimized during the northerly (June, July, and August (JJA)) and southerly (December, January, and February (DJF)) position of the ITCZ and maximized during the transitional seasons, whereas relative errors (rRMSE) show the opposite pattern, suggesting that error are not consistently proportional to rainfall magnitude but have a seasonal factor. ROK shows similarly high levels of relative and absolute errors to TPR and LM compared to the remaining S-G methods.

The higher performance of OK compared to the S-G merging methods can be misleading as it is affected by the nonuniform spatial rain gauge distribution, which causes the removal of a single gauge in LOOCV to have little effect on the estimate if there is a cluster of rain gauges surrounding the gauge location. On the other hand,
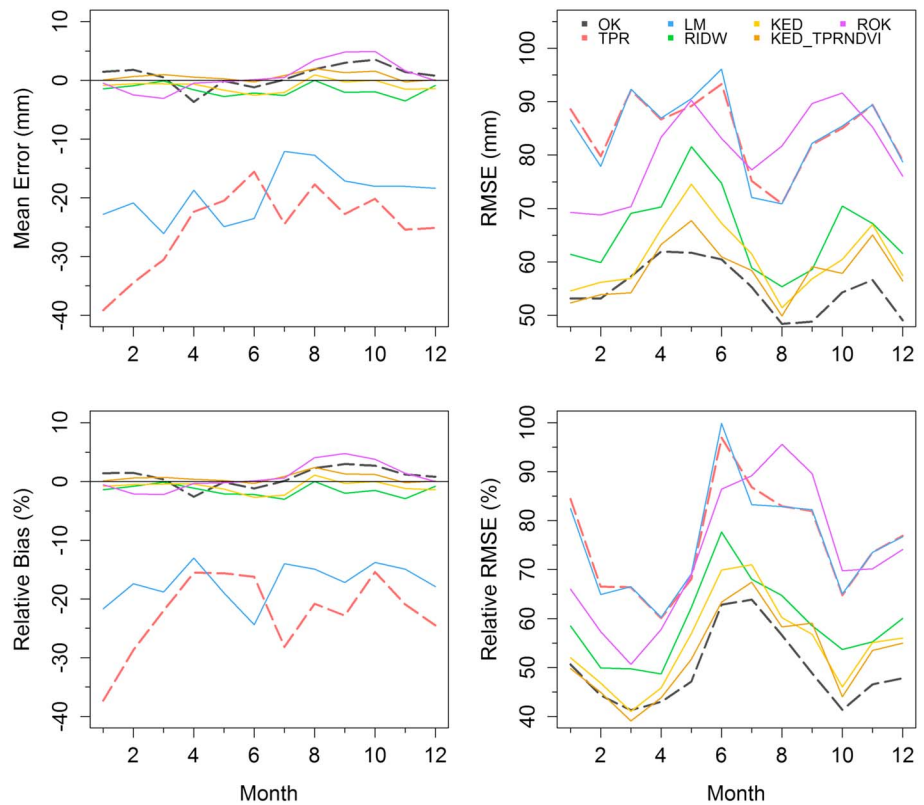
Figure 6. (top row) Seasonal trends of cross-validation absolute mean error (ME) and root-mean-square error (RMSE) of the merged precipitation products. (bottom row) Same as above but relative to mean climatological estimate of gauge observations.

poorly gauged locations where this is more likely to have an effect will by default be less represented in mean performance scores. However, when analyzing explicitly the influence of gauge density on the S-G method performance, a better distinction between OK and the remaining gauge-based merging methods (RIDW, ROK, KED, and KED_TN) can be made (see Figure 7). The correlation between the merged rainfall estimate and the gauge observation shows major deterioration for OK from $\rho > 0.8$ at 100 km to nearest gauge to $\rho < 0.1$
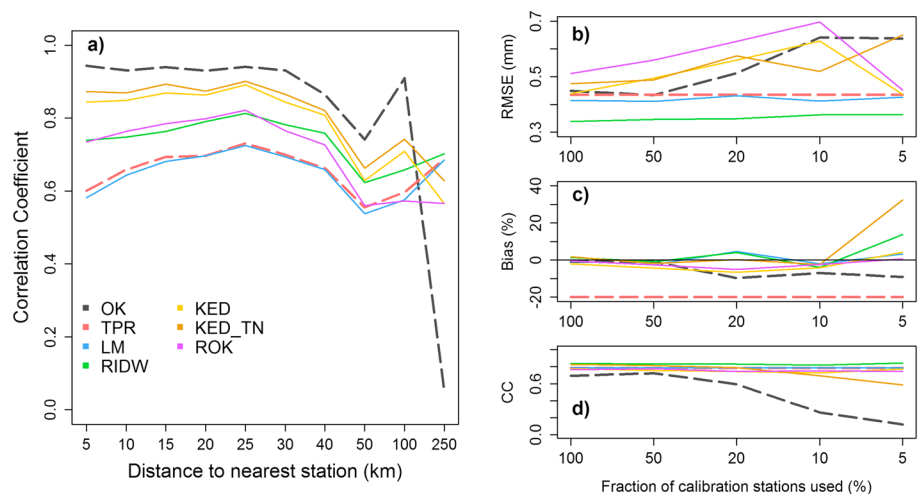
Figure 7. (a) Correlation between cross validation estimates and gauge observations as function of distance to the nearest gauge. (b) Root-mean-square error (RMSE), (c) relative bias, and (d) correlation coefficient (CC) between cross-validation merging estimates and gauge observations as function of network density (percentage of calibration gauges).
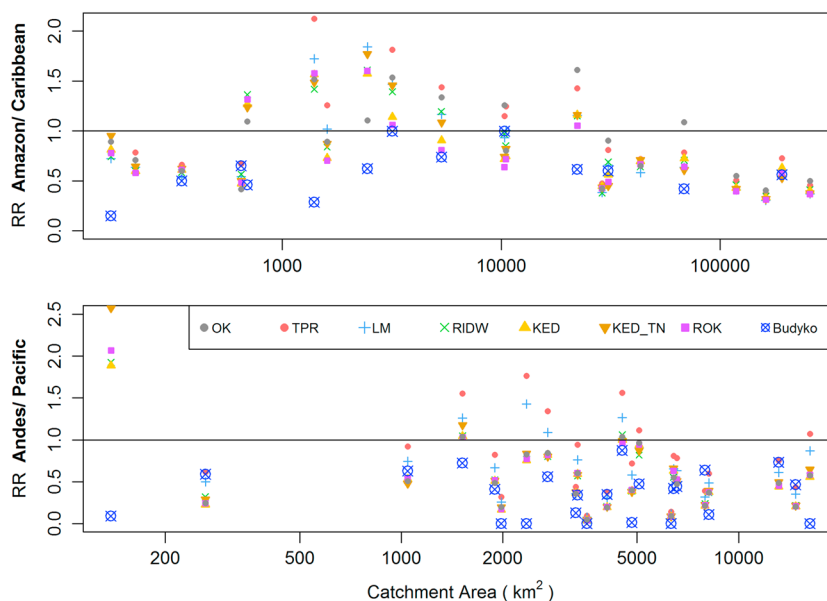
Figure 8. Runoff ratios (RR) of merged precipitation fields to discharge observations (top) for catchments draining to Amazon and Caribbean basins and (bottom) for catchments draining to Pacific and Andean (Lake Titicaca) basins. Catchment area is plotted on a log scale. The black line denotes RR = 1, i.e., total rainfall volume equals total discharge volume. RR < 1 shows the fraction of annual discharge accounted for by the annual rainfall averaged over the catchment, while RR > 1 shows the ratio by which annual discharge exceeds catchment mean precipitation.

at 250 km, whereas the remaining methods show a much weaker deterioration to approximately $\rho = 0.6$ at 250 km. Subsampling of the gauge data set to lower gauge densities prior to merging reflects this result with an increased random error (RMSE), increasingly negative bias, and a deterioration in the correlation coefficient for OK. All other methods show consistently high correlation of $\rho = 0.7$ and little systematic bias. RIDW shows a notably consistent and lower random error than Kriging-based merging methods and TPR irrespective of gauge density. For the geostatistical methods, the ratio between the Kriging estimation standard deviation and the Kriging estimates (RKESD) can also be used to assess the estimation error (results not shown here; see Figure S6 in the supporting information of this article). The RKESD confirms a high dependence of OK on gauge high density, with low error levels clustered around high gauge density areas.

3.3. Water Balance Evaluation (Runoff Ratios)

The hydrological evaluation using runoff ratios (Figure 8) returns some results of RR >1.0, suggesting discharge in excess of average rainfall integrated over the catchment. Discharge may be elevated by groundwater contributions and the accuracy of estimates can be affected by potentially large errors in discharge measurement in this region [*Zulkafli et al.*, 2013]. However, comparison with the cross-validation results suggests that underestimation of total rainfall volume is a more likely explanation for RR >1.0.

Across all S-G methods runoff ratios are highest in the Amazonian and Caribbean basins for catchments in the range of 1000–20000 km^2. Seven of the nine catchments in this size range are located along the eastern Andes. For these catchments underestimation of the orographically enhanced rainfall along the eastern Andes by the TPR (Figure 5) has the most pronounced impact on the total catchment rainfall and therefore on the water balance, resulting in multiple RR < 1.0. OK-based runoff ratios are only slightly lower in this region than TPR-based runoff ratios. However, the S-G methods (with the exception of LM) yield substantially reduced runoff ratios in this region. Where the S-G methods return RR < 1.0 the runoff ratios tend to be very clustered around the same RR value, but for catchments where RR > 1.0, ROK yields the lowest runoff ratios. The RR reduction by the S-G methods compared to either OK or TPR suggests that S-G merging substantially improves the accuracy of the rainfall volume estimates for catchments located in the eastern Andes. Over very large catchments in the Caribbean and Amazonian basin (>50,000 km^2), the zone of orographic enhancement is proportionally smaller and, collectively, the ability of the S-G merging methods (including OK, TPR, and LM) to estimate average precipitation is higher as reflected in RR values ranging 0.4–0.7 and better agreement with the Budyko estimates.

For catchments draining within or to the west of the Andes, most S-G methods return values in the range $0.0 < RR < 1.0$ with the exception of TPR and LM, which have been shown to underestimate rainfall in this drier region (compared to the Amazonian and Caribbean basins). Better agreement among the remaining S-G methods as well as OK suggests that because of the much higher gauge density (compared to the eastern Andes) the methodological differences between the S-G methods have less impact on the total precipitation volume and the introduction of the satellite data does not change total precipitation volume estimates compared to those obtained by the gauges only. In fact, TPR (and to a lesser degree LM) shows consistently higher estimates than those by the S-G methods, often exceeding $RR = 1.0$.

However, comparison with the Budyko estimates shows systematic overestimation for all S-G methods for catchment scales of 2000 km^2 to approximately 75,000 km^2, which reaffirms the results from the cross validation that showed overestimation of the low rainfall rates along the Peruvian coast. In fact, Budyko references values for six catchments are very close to $RR = 0.0$ and should be considered with care. These are likely to be slight underestimations as a result of using mean monthly data but are likely to be below $RR = 0.1$. Estimates are in better agreement with observations in the literature for catchments without major lowland areas along the dry coastline (approximately $0.55 – 0.75$) [*Buytaert et al.*, 2006b].

4. Discussion

4.1. Satellite-Only Product (TPR)

The climatological maps presented in this study support previous findings that the TPR is a suitable tool for identifying spatial precipitation variability, seasonal patterns, and, in particular, delineation of steep precipitation gradients in the Tropical Andes [*Espinoza et al.*, 2015; *Anders and Nesbitt*, 2015; *Nesbitt and Anders*, 2009]. The high-resolution gridded TPR product discernibly defines critical features of tropical Andean precipitation: orographic barriers along the eastern Andes, intra-Andean low precipitation, high precipitation fields in the tropical Pacific exceeding 10,000 mm yr^{-1} as well as the seasonal migration of the ITCZ.

However, cross validation and runoff ratio results also suggest overestimation in regions of low monthly rainfall such as the Peruvian coastline, while a systematic underestimation occurs particularly regions with east-facing windward slopes in the southern Tropical Andes and over the Altiplano as well as in central and western Ecuador. Underestimation of gauge rainfall by the TPR in this region in the order of 35–50% has previously been reported for the Andean highlands [*Condom et al.*, 2011], the Andes-Amazon transition zone and the Altiplano [*Espinoza et al.*, 2015] as well as the lowland Amazon basin [*Franchito et al.*, 2009].

The underestimation by the TPR can be explained by a number of factors. The simple difference between gauge period (1981–2010) and TPR period (1998–2014) being considered has been shown to be lower than these errors, although it can locally be significant (compare Figure 1 to Figure 5). The low TPR sampling frequency is likely to be relevant, especially as across this region the contribution of convective events with extreme rainfall rates to monthly total rainfall is estimated to be over 70 % [*Espinoza et al.*, 2015; *Mohr et al.*, 2014; *Rasmussen et al.*, 2013]. The high relevance of the TPR sampling error on mean climatological estimates has been highlighted in the past [*Nesbitt and Anders*, 2009] and power law models have been proposed to quantify sampling error based on rain rate, spatial grid resolution, and sampling frequency [*Steiner et al.*, 2003; *Iida et al.*, 2006; *Nesbitt and Anders*, 2009]. However, sampling error is likely to be highly nonlinear with respect to spatial scale and at very high spatial resolution will be affected by local rainfall properties. Nonetheless, as we consider the entire historical observational period of the TPR (1998–2014), the generated climatologies represent the maximum information content (i.e., maximum sample size and convergence) that was obtained by TPR.

In addition to sampling error associated with the climatological mean, detection and retrieval errors have been demonstrated to affect the accuracy of estimating different rainfall rates for individual instantaneous TPR measurements. Relevant error sources include signal blockage due to complex terrain, nonuniform beam filling, especially at high inclination angles (off nadir) and the associated misclassification of rainfall [*Kirstetter et al.*, 2014] as well as the reflectivity-to-rainfall rate (Z-R) conversion, which has been shown to be particularly relevant in estimating convective rainfall, especially from deep convective cores [*Rasmussen et al.*, 2013]. While extremely high rainfall rates may be affected by conversion and possible scattering effects, *Yang and Nesbitt* [2014] have shown that the majority of rainfall missed by the TPR sensor is light rain due to the TPR's lower detection limit of 18 dBZ, which translates to approximately 0.5 mm hr^{-1}, depending on the reflectivity to

rainfall rate transformation. This is much higher than the typical rain gauge detection limit of 0.1 mm hr^{-1}. *Yang and Nesbitt* [2014] estimate that the missed light rain contribution would increase the TPR rainfall by 10% and up to 20% in dry regions. Empirical observations suggest that drizzle accounts for 29% of rainfall in Ecuadorian highlands [*Padrón et al.*, 2015], suggesting that the error due to missing light rainfall can locally be amplified by the Andean meteorology. However, it seems unlikely that missing light rainfall alone explains the major underestimations of over 50% in the southern Tropical Andes and the Altiplano, suggesting that a combination of sampling error, underestimation of extreme rainfall rates, and missed light rainfall are contributing factors.

4.2. Gauge-Only Product (OK)

The gauge-only OK product performed comparatively well compared to the TPR and S-G products both in the leave-one-out cross validation and runoff ratio evaluations. While LOOCV results are affected by gauge clustering, the focus of the runoff ratio method on catchment average precipitation, as opposed to regional variations, also favors OK, which converges to the same spatial mean in the absence of local observations. Only catchments with most surface area located in the region of orographic enhancement in the eastern Andes show weaker performance (RR >1.0, hence, rainfall underestimation) by OK compared to the S-G methods. However, explicit analysis of the impact of rain gauge density shows a consistent deterioration of the OK performance with decreasing network density as is reflected in the relative Kriging standard deviation fields. Hereby, the most notable feature is the inability of OK to capture steep precipitation gradients with rather unrealistically gradual spatial variations in rainfall rates, even in highly gauged regions (see Figure 4). *Espinoza-Villar et al.* [2009a] highlight the relevance of gauge density in this respect: The easterly direction of trade winds carrying moisture from the Amazon toward the Andes cause windward (i.e., east facing) stations to experience a unimodal rainfall regime with high annual rainfall totals over 2000 mm, while nearby gauges on the leeward side of intra-Andean catchments experience a bimodal precipitation regime driven by the double passage of the ITCZ with annual rainfall below 1000 mm. A low-density gauge network, and interpolation methods based only on this rain gauge network, may therefore not only fail to accurately estimate rainfall rates but also to capture the general rainfall regime.

4.3. S-G Merged Products

This study has shown that S-G merging can improve precipitation estimation with improved rainfall rates compared to TPR and better representation of spatial variability compared to interpolation of gauges without satellite support (i.e., OK). A simple multiplicative bias correction (LM) was found to be insufficient across the large spatial scale considered in this study, systematically showing the same spatial pattern for overestimation and underestimation as the TPR. Differences are far smaller among the S-G merging methods that perform spatially explicit interpolation with only ROK deviating from the remaining methods in showing higher random errors in the LOOCV. However, when analyzing the impact of reducing gauge density there is no consistent pattern. RIDW shows lower random error and higher correlation than the Kriging-based methods, but this is not consistent with the results for the systematic bias where it exceeds most other methods. KED, which has been shown to perform well for subregions of the study domain [*Álvarez-Villa et al.*, 2011; *Nerini et al.*, 2015], preserves the relative spatial rainfall patterns of the TPR but absolute magnitudes tend to be in the range of the gauge observations, which then places the focus on the quality and representativeness of the gauge network. Addition of NDVI data to satellite observations in the KED has little effect in the high Andes but a profound impact in attenuating the seasonal progression of the high rainfall field that accompanies the ITCZ migration. Higher influence of NDVI as a result of rainforest canopy is to be expected but the resulting seasonal disagreement with S-G methods only based on TPR questions the suitability of using NDVI at this scale. Overall, while the inclusion of TPR is preferable to gauge-only interpolation, especially in poorly gauged regions, the results from this study do not allow for recommending a specific S-G merging method. However, this also reaffirms previous findings that an explicit model of the spatial covariance (as in Kriging) does not necessarily provide an improvement over a simple inverse distance weighted interpolation or gauge-satellite residuals (RIDW) [*Dinku et al.*, 2014]. Choice of S-G merging method will therefore depend on the quality of the gauge network. Geostatistical methods are justified if a good calibration of semivariograms can be achieved. In addition, the use of Kriging methods allows defining the Kriging standard deviation, which can be used as an indicator of relative uncertainty as in RKESD and therefore offers a quantitative tool to characterize and define confidence thresholds.

5. Conclusions

TPR captures the spatial variability of monthly climatological rainfall at high spatial resolutions of 5 km and thereby reflects the intra-annual precipitation patterns across the Tropical Andes, allowing for identification of orographic gradients and the impacts of synoptic-scale controls such as the ITCZ migration. The TPR itself does not provide accurate estimates of point-based rain gauge observations and tends to underestimate rainfall. This effect is amplified in regions where individual convective events represent a large contribution to annual rainfall totals. S-G merging techniques have shown to be beneficial, obtaining both good agreement with gauge observations and preserving spatial patterns observed by the TPR. Spatial variations in performance of the S-G merging techniques can be observed across the Tropical Andes, with particularly poor performance along the dry Peruvian coast. At the regional to synoptic scale, performance differences between the S-G merging methods are small relative to the performance variability of the individual methods across the Tropical Andes. Among the Kriging-based S-G methods, KED performs better than ROK with lower absolute and relative random errors throughout the year, although the inclusion of NDVI as an additional secondary variable should be considered with care and does not perform equal well for all environments. Simple spatial interpolation of gauge residuals (RIDW) yielded very similar results to the Kriging-based S-G merging methods and even performed more consistently during low gauge densities. The quality of the gauge network is therefore critical in the choice of S-G merging method. Furthermore, the availability of estimation uncertainties (Kriging estimation standard deviation) delivers a tool to quantify confidence thresholds when using a Kriging-based S-G merging method.

The merged 5 km monthly climatological maps presented in this study constitute an improved source of climatological rainfall data for the Tropical Andes compared to gauge-only data or TPR-only climatologies. Due to the topographic and hydroclimatic complexity of the Tropical Andes, applications focused on a much smaller spatial extent within the Tropical Andes and a higher local gauge density may consider local S-G merging in order to locally optimize performance of the rainfall products. While the current study analyzed the ability of existing S-G merging methods to capture monthly climatological rainfall, future work should focus on the accurate quantification of different sources of uncertainty (gauge measurement, gauge interpolation, satellite sensor, retrieval algorithm, gridding, and sampling errors) to improve S-G merging, for example, by means of more flexible and unbiased error-weighted merging methods [e.g., *Nerini et al.*, 2015; *Grimes et al.*, 1999].

References

Almazroui, M. (2011), Calibration of TRMM rainfall climatology over Saudi Arabia during 1998–2009, *Atmos. Res.*, *99*(3), 400–414.

Álvarez-Villa, O. D., J. I. Vélez, and G. Poveda (2011), Improved long-term mean annual rainfall fields for Colombia, *Int. J. Climatol.*, *31*(14), 2194–2212.

Anders, A. M., and S. W. Nesbitt (2015), Altitudinal precipitation gradients in the tropics from Tropical Rainfall Measuring Mission (TRMM) precipitation radar, *J. Hydrometeorol.*, *16*(1), 441–448.

Biasutti, M., S. E. Yuter, C. D. Burleyson, and A. H. Sobel (2012), Very high resolution rainfall patterns measured by TRMM precipitation radar: Seasonal and diurnal cycles, *Clim. Dyn.*, *39*(1–2), 239–258.

Boers, N., B. Bookhagen, N. Marwan, J. Kurths, and J. Marengo (2013), Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System, *Geophys. Res. Lett.*, *40*, 4386–4392, doi:10.1002/grl.50681.

Boers, N., R. V. Donner, B. Bookhagen, and J. Kurths (2014), Complex network analysis helps to identify impacts of the El Niño Southern Oscillation on moisture divergence in South America, *Clim. Dyn.*, 1–14.

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

Budyko, M. I. (1974), *Climate and Life*, Intl. Geophys. Ser., 508 pp., Academic, San Diego, Calif.

Buytaert, W., and B. De Bievre (2012), Water for cities: The impact of climate change and demographic growth in the Tropical Andes, *Water Resour. Res.*, *48*, W08503, doi:10.1029/2011WR011755.

Buytaert, W., R. Celleri, P. Willems, B. De Bievre, and G. Wyseure (2006a), Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes, *J. Hydrol.*, *329*(3), 413–421.

Buytaert, W., V. Iñiguez, R. Célleri, B. De Biévre, G. Wyseure, and J. Deckers (2006b), Analysis of the water balance of small páramo catchments in south Ecuador, in *Environmental Role of Wetlands in Headwaters*, edited by J. Krecek and M. Haigh, pp. 143–145, Springer, Dordrecht, Netherlands.

Buytaert, W., M. Vuille, A. Dewulf, R. Urrutia, A. Karmalkar, and R. Célleri (2010), Uncertainties in climate change projections and regional downscaling in the Tropical Andes: Implications for water resources management, *Hydrol. Earth Syst. Sci.*, *14*(7), 1247–1258.

Carvalho, L. M., A. E. Silva, C. Jones, B. Liebmann, P. L. S. Dias, and H. R. Rocha (2011), Moisture transport and intraseasonal variability in the South America monsoon system, *Clim. Dyn.*, *36*(9–10), 1865–1880.

Cheema, M. J. M., and W. G. Bastiaanssen (2012), Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus Basin, *Int. J. Remote Sens.*, *33*(8), 2603–2627.

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

Chen, S., et al. (2013), Evaluation of spatial errors of precipitation rates and types from TRMM spaceborne radar over the southern CONUS, *J. Hydrometeorol.*, *14*(6), 1884–1896.

Condom, T., P. Rau, and J. C. Espinoza (2011), Correction of TRMM 3b43 monthly precipitation data over the mountainous areas of Peru during the period 1998–2007, *Hydrol. Processes*, *25*(12), 1924–1933.

Demaria, E. M. C., D. A. Rodriguez, E. E. Ebert, P. Salio, F. Su, and J. B. Valdes (2011), Evaluation of mesoscale convective systems in South America using multiple satellite products and an object-based approach, *J. Geophys. Res.*, *116*, D08103, doi:10.1029/2010JD015157.

Dinku, T., S. J. Connor, and P. Ceccato (2010), Comparison of CMORPH and TRMM-3b42 over mountainous regions of Africa and South America, in *Satellite Rainfall Applications for Surface Hydrology*, edited by M. Gebremichael and F. Hossain, pp. 3–22, Springer, Dordrecht, Netherlands.

Dinku, T., K. Hailemariam, R. Maidment, E. Tarnavsky, and S. Connor (2014), Combined use of satellite estimates and rain gauge observations to generate high-quality historical rainfall time series over Ethiopia, *Int. J. Climatol.*, *34*, 2489–2504, doi:10.1002/joc.3855.

Duan, Y., A. M. Wilson, and A. P. Barros (2015), Scoping a field experiment: Error diagnostics of TRMM precipitation radar estimates in complex terrain as a basis for IPHEx2014, *Hydrol. Earth Syst. Sci.*, *19*(3), 1501–1520.

Duan, Z., and W. G. M. Bastiaanssen (2013), First results from Version 7 TRMM 3b43 precipitation product in combination with a new downscaling calibration procedure, *Remote Sens. Environ.*, *131*, 1–13.

Espinoza, J. C., S. Chavez, J. Ronchail, C. Junquas, K. Takahashi, and W. Lavado (2015), Rainfall hotspots over the southern tropical Andes: Spatial distribution, rainfall intensity, and relations with large-scale atmospheric circulation, *Water Resour. Res.*, *51*, 3459–3475, doi:10.1002/2014WR016273.

Espinoza-Villar, J. C., J. Ronchail, J. L. Guyot, G. Cochonneau, F. Naziano, W. Lavado, E. De Oliveira, R. Pombosa, and P. Vauchel (2009a), Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador), *Int. J. Climatol.*, *29*(11), 1574–1594.

Espinoza-Villar, J. C., J. L. Guyot, J. Ronchail, G. Cochonneau, N. Filizola, P. Fraizy, D. Labat, E. de Oliveira, J. J. Ordoñez, and P. Vauchel (2009b), Contrasting regional discharge evolutions in the Amazon basin (1974–2004), *J. Hydrol.*, *375*(3), 297–311.

Franchito, S. H., V. B. Rao, A. C. Vasques, C. M. Santo, and J. C. Conforte (2009), Validation of TRMM precipitation radar monthly rainfall estimates over Brazil, *J. Geophys. Res.*, *114*, D02105, doi:10.1029/2007JD009580.

Goovaerts, P. (1999), Geostatistics in soil science: State-of-the-art and perspectives, *Geoderma*, *89*(1), 1–45.

Grimes, D. I., and E. Pardo-Igúzquiza (2010), Geostatistical analysis of rainfall, *Geog. Analysis*, *42*(2), 136–160.

Grimes, D. I. F., E. Pardo-Iguzquiza, and R. Bonifacio (1999), Optimal areal rainfall estimation using raingauges and satellite data, *J. Hydrol.*, *222*(1), 93–108.

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker (2007), The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *J. Hydrometeor.*, *8*(1), 38–55.

Hunink, J. E., W. W. Immerzeel, and P. Droogers (2014), A High-resolution Precipitation 2-step mapping Procedure (HiP2p): Development and application to a tropical mountainous area, *Remote Sens. Environ.*, *140*, 179–188.

Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K. Okamoto (2000), Rain-profiling algorithm for the TRMM precipitation radar, *J. Appl. Meteorol.*, *39*(12), 2038–2052.

Iguchi, T., T. Kozu, J. Kwiatkowski, R. Meneghini, J. Awaka, and K. Okamoto (2009), Uncertainties in the rain profiling algorithm for the TRMM precipitation radar, *J. Meteorol. Soc. Jpn.*, *87*, 1–30.

Iida, Y., K. Okamoto, T. Ushio, and R. Oki (2006), Simulation of sampling error of average rainfall rates in space and time by five satellites using radar-AMeDAS composites, *Geophys. Res. Lett.*, *33*, L01816, doi:10.1029/2005GL024910.

Iida, Y., T. Kubota, T. Iguchi, and R. Oki (2010), Evaluating sampling error in TRMM/PR rainfall products by the bootstrap method: Estimation of the sampling error and its application to a trend analysis, *J. Geophys. Res.*, *115*, D22119, doi:10.1029/2010JD014257.

Immerzeel, W. W., M. M. Rutten, and P. Droogers (2009), Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula, *Remote Sens. Environ.*, *113*(2), 362–370.

Indu, J., and D. Nagesh Kumar (2014), Evaluation of TRMM PR sampling error over a subtropical basin using bootstrap technique, *IEEE Trans. Geosci. Remote Sens.*, *52*, 6870–6881.

Isotta, F. A., et al. (2014), The climate of daily precipitation in the Alps: Development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, *Int. J. Climatol.*, *34*(5), 1657–1675.

Jiang, S., L. Ren, Y. Hong, B. Yong, X. Yang, F. Yuan, and M. Ma (2012), Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method, *J. Hydrol.*, *452*, 213–225.

Kirstetter, P.-E., Y. Hong, J. J. Gourley, S. Chen, Z. Flamig, J. Zhang, M. Schwaller, W. Petersen, and E. Amitai (2012), Toward a framework for systematic error modeling of spaceborne precipitation radar with NOAA/NSSL ground radar-based national mosaic QPE, *J. Hydrometeorol.*, *13*(4), 1285–1300.

Kirstetter, P.-E., Y. Hong, J. J. Gourley, M. Schwaller, W. Petersen, and Q. Cao (2014), Impact of sub-pixel rainfall variability on spaceborne precipitation estimation: Evaluating the TRMM 2a25 product, *Q. J. R. Meteorolog. Soc.*, *141*, 953–966.

Lavado Casimiro, W. S., D. Labat, J. Loup Guyot, J. Ronchail, and J. Julio Ordonez (2009), TRMM rainfall data estimation over the Peruvian Amazon-Andes basin and its assimilation into a monthly water balance model, *IAHS-AISH Publ.*, *333*, 245–252.

Li, L., Y. Hong, J. Wang, R. F. Adler, F. S. Policelli, S. Habib, D. Irwn, T. Korme, and L. Okello (2009), Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa, *Nat. Hazards*, *50*(1), 109–123.

Marengo, J. A., et al. (2012), Recent developments on the South American monsoon system, *Int. J. Climatol.*, *32*(1), 1–21.

Marengo, J. A., L. M. Alves, W. R. Soares, D. A. Rodriguez, H. Camargo, M. P. Riveros, and A. D. Pabló (2013), Two contrasting severe seasonal extremes in tropical South America in 2012: Flood in Amazonia and drought in Northeast Brazil, *J. Clim.*, *26*(22), 9137–9154.

Moazami, S., S. Golian, M. R. Kavianpour, and Y. Hong (2014), Uncertainty analysis of bias from satellite rainfall estimates using copula method, *Atmos. Res.*, *137*, 145–166.

Mohr, K. I., D. Slayback, and K. Yager (2014), Characteristics of precipitation features and annual rainfall during the TRMM era in the Central Andes, *J. Clim.*, *27*, 3982–4001.

Nerini, D., Z. Zulkafli, L.-P. Wang, C. Onof, W. Buytaert, W. Lavado, and J.-L. Guyot (2015), A comparative analysis of TRMM-rain gauge data merging techniques at the daily time scale for distributed rainfall-runoff modelling applications, *J. Hydrometeorol.*, *16*, 2153–2168.

Nesbitt, S. W., and A. M. Anders (2009), Very high resolution precipitation climatologies from the Tropical Rainfall Measuring Mission precipitation radar, *Geophys. Res. Lett.*, *36*, L15815, doi:10.1029/2009GL038026.

Nikolopoulos, E. I., E. N. Anagnostou, and M. Borga (2013), Using high-resolution satellite rainfall products to simulate a major flash flood event in Northern Italy, *J. Hydrometeorol.*, *14*(1), 171–185.

Padrón, R. S., B. P. Wilcox, P. Crespo, and R. Célleri (2015), Rainfall in the Andean Páramo-new insights from high-resolution monitoring in Southern Ecuador, *J. Hydrometeorol.*, *16*, 985–996.

Poveda, G., D. M. Álvarez, and O. A. Rueda (2011), Hydro-climatic variability over the Andes of Colombia associated with ENSO: A review of climatic processes and their impact on one of the Earth's most important biodiversity hotspots, *Clim. Dyn.*, *36*(11–12), 2233–2249.

Rasmussen, K. L., S. L. Choi, M. D. Zuluaga, and R. A. Houze (2013), TRMM precipitation bias in extreme storms in South America, *Geophys. Res. Lett.*, *40*, 3457–3461, doi:10.1002/grl.50651.

Romatschke, U., and R. A. Houze Jr. (2010), Extreme summer convection in South America, *J. Clim.*, *23*(14), 3761–3791.

Seto, S., T. Iguchi, N. Utsumi, M. Kiguchi, and T. Oki (2013), Evaluation of extreme rain estimates in the TRMM/PR standard product version 7 using high-temporal-resolution rain gauge datasets over Japan, *SOLA*, *9*, 98–101.

Steiner, M., T. L. Bell, Y. Zhang, and E. F. Wood (2003), Comparison of two methods for estimating the sampling-related uncertainty of satellite rainfall averages based on a large radar dataset, *J. Clim.*, *16*(22), 3759–3778.

Stephens, G. L., and C. D. Kummerow (2007), The remote sensing of clouds and precipitation from space: A review, *J. Atmos. Sci.*, *64*(11), 3742–3765.

Tarnavsky, E., M. Mulligan, and G. Husak (2012), Spatial disaggregation and intensity correction of TRMM-based rainfall time series for hydrological applications in dryland catchments, *Hydrol. Sci. J.*, *57*(2), 248–264.

Thiemig, V., R. Rojas, M. Zambrano-Bigiarini, V. Levizzani, and A. De Roo (2012), Validation of satellite-based precipitation products over sparsely gauged African river basins, *J. Hydrometeorol.*, *13*(6), 1760–1783.

Tian, Y., C. D. Peters-Lidard, and J. B. Eylander (2010), Real-time bias reduction for satellite-based precipitation estimates, *J. Hydrometeorol.*, *11*(6), 1275–1285.

Todini, E. (2001), A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements, *Hydrol. Earth Syst. Sci. Discuss.*, *5*(2), 187–199.

Toté, C., D. Patricio, H. Boogaard, R. van der Wijngaart, E. Tarnavsky, and C. Funk (2015), Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique, *Remote Sens.*, *7*(2), 1758–1776.

TRMM Precipitation Radar Team (2011), Tropical rainfall measuring mission (TRMM) precipitation radar algorithm: Instruction manual for version 7. [Available at http://www.eorc.jaxa.jp/TRMM/documents/PR_algorithm_product_information/pr_manual/PR_Instruction_Manual_V7_L1.pdf.]

Vera, C., et al. (2006), The South American low-level jet experiment, *Bull. Am. Meteorol. Soc.*, *87*(1), 63–77.

Vila, D. A., L. G. G. De Goncalves, D. L. Toll, and J. R. Rozante (2009), Statistical evaluation of combined daily gauge observations and rainfall satellite estimates over continental South America, *J. Hydrometeorol.*, *10*(2), 533–543.

Ward, E., W. Buytaert, L. Peaver, and H. Wheater (2011), Evaluation of precipitation products over complex mountainous terrain: A water resources perspective, *Adv. Water Res.*, *34*(10), 1222–1231.

Yang, S., and S. W. Nesbitt (2014), Statistical properties of precipitation as observed by the TRMM precipitation radar, *Geophys. Res. Lett.*, *41*, 5636–5643, doi:10.1002/2014GL060683.

Zulkafli, Z., W. Buytaert, C. Onof, W. Lavado, and J.-L. Guyot (2013), A critical assessment of the JULES land surface model hydrology for humid tropical environments, *Hydrol. Earth Syst. Sci.*, *17*(3), 1113–1132.

Zulkafli, Z., W. Buytaert, C. Onof, B. Manz, E. Tarnavsky, W. Lavado, and J.-L. Guyot (2014), A comparative performance analysis of TRMM 3b42 (TMPA) versions 6 and 7 for hydrological applications over Andean-Amazon river basins, *J. Hydrometeorol.*, *15*(2), 581–592.