Construction of a high-resolution gridded rainfall dataset for Peru from 1981 to the present day

Journal:	*Hydrological Sciences Journal*
Manuscript ID	HSJ-2019-0097
Manuscript Type:	Special Issue Paper
Date Submitted by the Author:	23-Feb-2019
Complete List of Authors:	Aybar Camacho, Cesar Luis; Servicio Nacional de Meteorología e Hidrología del Perú Fernández, Carlos; Servicio Nacional de Meteorología e Hidrología del Perú; Potsdam Institute for Climate Impact Research Huerta, Adrian; Servicio Nacional de Meteorología e Hidrología del Perú Lavado, Waldo; Servicio Nacional de Meteorología e Hidrología del Perú Vega, Fiorella; Servicio Nacional de Meteorología e Hidrología del Perú Felipe-Obando, Oscar; Servicio Nacional de Meteorología e Hidrología del Perú
Keywords:	quality control, gap-infilling, CHIRP, TRMM 2A25, quantile mapping, PISCOp V2.1

Construction of a high-resolution gridded rainfall dataset for Peru from 1981 to the present day

Cesar Aybar[a]; Carlos Fernández[a,b]; Adrian Huerta[a]; Waldo Lavado[a]; Fiorella Vega[a]; Oscar Felipe-Obando[a]

[a] *Servicio de Meteorología e Hidrología del Perú (SENAMHI), Lima, Peru*

[b] *Potsdam Institute for Climate Impact Research, Potsdam, Germany*

Corresponding author: Cesar Aybar (csaybar@gmail.com)

Abstract

This paper introduces a new gridded rainfall dataset available for Peru called PISCOp V2.1 (Peruvian Interpolated data of SENAMHI's Climatological and Hydrological Observations) that have been developed for the period 1981 to the present with an average latency of eight weeks at 0.1° spatial resolution. The merging algorithm is based on geostatistical and deterministic interpolation methods including three different rainfall sources: (i) the national quality-controlled and infilled rain gauge dataset, ii) radar-gauge merged precipitation climatologies and (iii) the Climate Hazards Group Infrared Precipitation (CHIRP) estimates. The validation results suggest that precipitation estimates are acceptable showing the highest performance for the Pacific coast and the western flank of the Andes. Furthermore, a meticulous quality-control and gap-infilling procedure allowed to reduce the formation of inhomogeneities (non-climatic breaks). The dataset is publicly available at https://piscoprec.github.io/ and is intended to support hydrological studies and water management practices.

Keywords: quality control; gap-infilling; CHIRP; TRMM 2A25, quantile mapping, PISCOp V2.1

1. Introduction

Accurate spatial-temporal rainfall estimations are essential for the development of scientific and operational applications which allow to understand the water cycle and its impact on natural and human systems. Conventional observations from rain gauge stations are an ideal input for the aforementioned applications. Unfortunately, strong spatial variability (Garreaud et al., 2009) and the heterogeneous and sparse distribution of rain gauges combined with systematic data quality deficiencies (Hunziker et al., 2017) precludes their widespread use within Peru.

In the last decades, new algorithms based on the indirect estimations from advanced infrared and microwave satellites, have led to the construction of different gridded rainfall datasets (Grd) that are used as auxiliary data to overcome the lack of rain gauge stations, increase the spatio-temporal resolution and reduce uncertainties in rainfall predictions (Baik et al., 2015; Bi et al., 2017; Sun et al., 2017; Verdin et al., 2015). The most recent Grd with global and near real-time coverage are the Multi-Source Weighted-Ensemble Precipitation (MSWEP; Beck et al., 2018) and Climate Hazards Group Infrared Precipitation with Stations (CHIRPS; Funk et al., 2015) products. These blended datasets are based on the merge of climate, satellite, reanalysis and gauge (the latter optional for the last versions) rainfall sources. MSWEP provides three-hour precipitation at a spatial resolution of ~10 km for the period 1979-2017, while CHIRPS covers daily precipitation at ~5 km for 1981-present. Only a few studies have been done to analyze the performance of these new blended Grd in adjacent regions of Peru. For instance, Zambrano-Bigiarini et al. (2017) show that both CHIRPS and MSWEP perform well at high temporal scales, presenting problems of overestimation (underestimation) in events of light rain (violent rain) in Chile, whereas Perdigón-Morales et al. (2017) and Javier et al. (2016) mention that CHIRPS is acceptable in reproducing climatological values of monthly accumulated precipitation in Mexico and Venezuela, respectively. Nonetheless, these studies re-used rain gauges that had been incorporated previously in the merging algorithm which alters the reliability of blended Grd performance results. In line with this, Beck et al. (2017b) reported that MSWEP performs better than CHIRPS only in regions with extensive rain gauge networks (e.g., in temperate regions). However, these findings cannot be directly applied to regions with sparse and irregular monitoring networks, such as Peru.

Recently, several studies have indicated essential considerations when generating daily and monthly blended Grd in data-scarce regions: (1) efficiency in blended Grd largely depends on the predictor and the interpolation method used must be able to adapt to scenarios with high spatial heterogeneity (Dinku et al., 2014); (2) geostatistical interpolation methods outperformed deterministic methods at annual and monthly time step, whereas for daily time step, geostatistic and deterministic methods have been proven to be comparable (Ly et al., 2013); (3) difficulties exist for assuming space-stationary hypothesis and establish an adequate theoretical semivariogram at daily time step (Nerini et al., 2015); (4) the optimized interpolation parameters in deterministic methods significantly improve the final results (Chen & Liu, 2012); (5) the gap-infilling in precipitation time series is highly important to minimize inhomogeneities in the gridded datasets for periods of missing data, especially in heterogeneous regions (Beguería et al., 2015; Peterson et al., 1998; Yanto et al., 2017a); and (6) the use of simple ratios based on very high rainfall climatologies can significantly decrease the systematic bias (Strauch et al., 2016).

On this basis, this study presents the development of PISCOp V2.1, a new local blended Grd headed by the National Service of Meteorology and Hydrology of Peru (SENAMHI). PISCOp V.2.1 contains daily and monthly rainfall grids at 0.1° computed for the 1981-

present period covering entire Peru with an average latency of eight weeks. It is built using serially complete rain gauge datasets, CHIRP V2.0 (without rain gauges), radar-gauge merged precipitation climatologies, geostatistics and deterministic interpolation methods and a simple monthly correction factor applied to daily estimates. The objective of this paper is to provide detailed and transparent information about the construction of PISCOp V2.1 as well as to evaluate its performance and stress its limitations.

2. Material and methods

2.1. Study Area

Peru is located in South America's central-western region from 0°S-18°S and 68°-82°W (Figure 1), covering climatically extremely variable regions with diverse precipitation regimes that result from the interaction between synoptic-scale atmospheric currents, the complex orography of the Andes, the cold Humboldt Current System (HCS) and the El Niño Southern Oscillation (ENSO, Garreaud et al., 2009; Lavado Casimiro et al., 2012b).

Figure 1

In austral summer, easterly trade winds from the southerly position of the Intertropical Convergence Zone (ITCZ) transport humid air masses from the tropical Atlantic towards the Amazon basin (Carvalho et al., 2011; Manz et al., 2017; Marengo et al., 2012) and to the south along the Andes through the South American Low-Level Jet (SALLJ; Boers et al., 2013; Vera et al., 2006). This period determines a marked wet season in most of Peru

(Marengo et al., 2012). Conversely, when the ITCZ is located further north (austral winter), convection and, consequently, precipitation levels are significantly reduced.

In the Peruvian Andes, the climate is complex and primarily controlled by orography that acts as a topographic barrier to moisture flow, causing the formation of strong precipitation gradients at the eastern flanks of the Andes (Bookhagen & Strecker, 2008). The inter-Andean valleys (> ~500 mm/year) are principally dominated by convective processes (Campozano et al., 2016; Garreaud, 1999) channeling moisture intrusions of the Amazon (Chavez & Takahashi, 2017). At the same time, the influence of cold and dry air masses originating from the HCS cause the driest conditions of our study area at the Pacific coast and the western flanks of the Andes (< ~500 mm/year). However, during the occurrence of ENSO the HCS weakens and the formation of severe convective storms can occur especially over the northern Pacific coast (Antico, 2009).

Based on the hydro-climatic heterogeneity described above and according to the classification of Manz et al. (2017), the study area was divided into five sub-regions (Figure 1a and Table 1): (i) The Pacific coast (PC, average annual precipitation of ~150 mm/year). (ii) The western Andean slopes (AW; ~400 mm/year), the eastern Andean slopes (AE; ~1100 mm/year), the Andes-Amazon transition (AAT; ~3200 mm/year) and the Amazon lowlands (AL; ~2250 mm/year).

Table 1

2.2. Rain gauge dataset

The raw gauge dataset comprised 945 daily observation data provided by SENAMHI. The rain gauge data spans the period 1981-present and is characterized by numerous sources of systematic errors and gaps (Hunziker et al., 2017b). Even though metadata information would help to perform the data quality control, this was not used because of its limited availability. Therefore, our analysis in rain gauge data mainly focused on gross error detection and gap-infilling methods.

2.2.1. Quality control (QC)

Most methods for quality control (QC) on rain gauge observations are designed for dense station networks (Isotta et al., 2014; Notivoli et al., 2017; Vicente-Serrano et al., 2010) which are difficult to assume in this study. As expected, assuring the quality of a dataset is more problematic for data-scarce regions due to a reduced number of neighboring stations (Hunziker et al., 2017b). Considering this fact and the lack of an established quality management system in Peru, we propose a three-step QC approach which can be considered as conservative because only gross errors are deleted if there is strong evidence for implausibility. Hence it is still expected to find some remaining errors after QC, especially in areas with a lower density of gauges. The methods that were applied are:

1. Check for general problems (automatic): The first step was to delete obvious non-consistent values, such as negative and non-physical precipitation, decimal-point related errors, repetitive dates, repetitive consecutive values and unexpected changes in latitude and longitude coordinates.
2. Spatial extreme values check (automatic): Secondly, for the detection of extreme events of precipitation a threshold of 200 years of return period was used. If the atypical values occur in at least two neighboring (< 50 km) gauges for the same date,

they were preserved, otherwise they were deleted. This procedure was done automatically and is similar to Keller et al. (2015).

3. Break and bad segments check (manual): A visual control was performed in order to recognize segments with asymmetric rounding patterns and obvious inhomogeneities.

2.2.2. Gap-infilling

Another source of uncertainty provides the temporal non-consistent gauge network which is susceptible to produce systematic bias during the merging phase (New et al., 2000). This gains primary importance in data-scarce regions where several rain gauges come in and out of use (Hunziker et al., 2017b). Similar to the QC approach, there is no established methodology for gap-infilling. Therefore, we propose a two-step approach in order to generate a serially complete rain gauge dataset.

First, relying on neighboring gauges, the relatively newer and effective spatio-temporal imputation method CUTOFF (Feng et al., 2014) was applied to infill the previously quality controlled gauge datasets at daily and monthly time step. For this purpose, each rain gauge was grouped with other rain gauges if the following conditions were met: (1) A distance < 100 km, (2) Sharing as minimum 10 years of data, and (3) A daily (monthly) linear relationship > 0.5 (> 0.8). Secondly, if the previous condition was neither fulfilled nor enough to create serially complete time series, the quantile mapping bias correction (Qm), produced by matching the empirical cumulative distribution of the collocated grid cell (CHIRPM, see section below) to the available gauge data (Gudmundsson et al., 2012), was used to infill the remaining gaps of each rain gauge.

2.3. Modification of CHIRP (CHIRPM)

CHIRP products at monthly ($CHIRP_m$) and daily ($CHIRP_d$) time step are initially calculated from preliminary pentad time step product ($CHIRP_{pentad}$) by using the following equations (Funk et al., 2015a):

$$IRP_{pentad} = b_o + b_1 * (TIR\ CCD_{pentad}\ \%) \ldots (1)$$

$$CHIRP_{pentad} = CHP_{clim} * \frac{IRP_{pentad}}{IRP_{clim}} \ \ldots (2)$$

Where IRP_{pentad} is the precipitation calculated from the linear model between Thermal Infrared Cold Cloud Duration Percentage (TIR CCD%) and TRMM 3B42 V7 product; CHP_{clim} are the monthly precipitation climatologies generated by Funk et al. (2015b) and IRP_{clim} represents the climatology of IRP.

Even though the use of climatic or monthly correction factors can reduce the systematic bias (Beck et al., 2017c; Funk et al., 2015a; Keller et al., 2015; Parmentier et al., 2015; Strauch et al., 2016; van Osnabrugge et al., 2017), it is crucial that the predictor of larger temporary aggregations (e.g. CHP_{clim}) is well represented, otherwise a reverse process could occur. For Peru, we found that CHP_{clim} extremely overestimates precipitation (> 500%) at the Peruvian coast between 8°-18°S (Figure S1). Furthermore, it does not adequately represent the orographic rainfall hotspots over the Andes-Amazon transition and considers rain gauges with poor reliability. Based on this, CHP_{clim} was replaced by our own climatology $PISCOp_{clim}$ (see Section 2.3.1), resulting in CHIRP modified (CHIRPM):

$$CHIRPM_m = CHIRP_m * \frac{\mathrm{PISCOp}_{clim} + \mathcal{E}}{CHP_{clim} + \mathcal{E}} \quad \ldots (3)$$

$$CHIRPM_d = CHIRP_d * \frac{\mathrm{PISCOp}_{clim} + \mathcal{E}}{CHP_{clim} + \mathcal{E}} \quad \ldots (4)$$

Where Ɛ is a threshold defined as 0.5 in the denominator and the numerator in order to deal with values of zero or near zero. This equation is applied to monthly and daily CHIRP estimates resulting in CHIRPMm and CHIRPMd, respectively. CHIRP was before resampled to a spatial resolution of 0.1 ° through cubic spline interpolation.

2.3.1. PISCOp climatology (PISCOp$_{\mathbf{clim}}$)

For Peru it has been found that the TRMM precipitation radar product 2A25 (TPR, Iguchi et al., 2000) is the most suitable rainfall data source for identifying spatial precipitation variability and seasonal patterns, even for the complex orographic rainfall hotspots located in the eastern Andes (Bookhagen & Strecker, 2008; Manz et al., 2016; Nesbitt & Anders, 2009). Based on this dataset, we constructed monthly climatologies at 0.1° spatial resolution. The TPR data used corresponds to the 1998-2013 period excluding the year 2014 because during that year the satellite was carrying out anomalous maneuvers related to its dismantling (Houze et al., 2015). The construction procedure of PISCOp$_{\mathrm{clim}}$ (Figure 2) is summarized as follows:

1. Extraction of suspicious pixels with a rain rate > 300 mm/h (Hamada & Takayabu, 2014).
2. Aggregation of the entire TPR dataset to mean climatologies estimates for each calendar month considering the delineation of the overpassing TPR pixel proposed by Manz et al. (2016).
3. Applying a spatial bias thresholding filter to replace the pixels with large ratios (> 5 median) by the average of the surrounding 3×3 kernel.

4. Smoothing of rain rate through cubic spline interpolation.
5. Merging with the long-term (1981-2010) monthly climatologies rain gauge dataset (Figure 1a) using residual ordinary kriging (ROK, Section 2.4.2).

2.4. PISCOp V.2.1 Merging Phase

The merging phase (Figure 2) can be divided into four steps. Firstly, the provisional product P-PISCOpd was created by merging CHIRPMd and serially complete daily gauge datasets applying Residual Inverse Distance Weighting (RIDW). Secondly, PISCOpm was estimated by merging CHIRPMm and completed monthly gauge datasets using Residual Ordinary Kriging (ROK). Thirdly, a monthly correction factor (Mcf) was derived from the comparison of PISCOpm and P-PISCOpd aggregated at monthly time step. Finally, PISCOpd was estimated by multiplying Mcf with P-PISCOpd.

Figure 2

2.4.1. Residual Inverse Distance Weighting (RIDW)

RIDW is used to generate P-PISCOpd. In this deterministic prediction method, the residuals are defined in each gauged location s_i, as follows:

$$r_o(s_i) = X_B(s_i) - X_O(s_i) \quad \ldots (5)$$

$$S_i \in S, \mathrm{i} = 1, \ldots, \mathrm{N} \quad \ldots (6)$$

Where N is the number of gauge observations, r_o the residuals, X_O the daily rain gauge values and X_B the CHIRPMd value computed at each gauge location. The collocation of a rain gauge to each CHIRPMd grid cell was performed using a Smoothed Merging (SM) approach described by Li & Shao (2010). To estimate the residual field (μ_j), r_o was interpolated by IDW at each grid point j = 1, …, M, given by:

$$\mu_j = \begin{cases} \dfrac{\sum_{i=1}^{N} w_i(||S_j - S_i||) r_o(s_i)}{\sum_{i=1}^{N} w_i(||S_j - S_i||)}; & if\ ||S_j - S_i|| \neq 0 \ldots (7) \\ u_i & ; \quad if\ ||S_j - S_i|| = 0 \ \ldots (8) \end{cases}$$

$$w_i\left(||S_j - S_i||\right) = \frac{1}{||Sj - S_i||^{\alpha}} \ldots (9)$$

Where $||.||$ is the euclidean distance, w_i is the weight assigned to the gauge observation s_i and α is the power parameter. The α parameter controls the desired smoothness and the local behavior in the spatial prediction. High (low) values of α increase (decrease) the influence of the furthermost observations, generating low (high) variance in the residual field. For additional details on IDW see Babak & Deutsch (2009).

Different studies have examined the effects of varying α for the spatial prediction of rainfall (Adhikary, 2017; Chen & Liu, 2012). Accordingly, the optimal α was estimated minimizing the root mean square error (RMSE) obtained from the 10-fold cross-validation. Finally, P-PISCOpd was defined as:

$$\text{P-PISCOpd} = \text{CHIRPMd} - \mu \ldots \ (10)$$

2.4.2. Residual Ordinary Kriging (ROK)

ROK is used for the generation of PISCOpm. Similar to RIDW, with ROK the residuals are estimated by the equations (5) and (6) with the main difference that X_O corresponds to monthly gauge estimates and X_B represents the CHIRPMm values computed at each gauge location. Unlike RIDW, with ROK the residual field r_o is interpolated by ordinary kriging (Grimes et al., 1999) at each grid point and added back to CHIRPMm.

To ensure the non-stationarity assumption, the residuals are converted to logarithmic scale and back-transformed after the merging phase. In this study, the variogram adjustment was automatically performed based on Hiemstra et al. (2009). For more details on the implementation of ROK refer to Goovaerts (2000).

2.4.3. Monthly correction factor (Mcf)

Given that a higher spatial relationship is achieved at monthly rather than at daily time step (Ly et al., 2013), PISCOpm is expected to present a higher performance compared with the monthly aggregation of P-PISCOpd. Therefore, based on Keller et al. (2015), an Mcf was added after the creation of P-PISCOpd and PISCOpm under two purposes: provide higher spatial consistency to daily predictions and ensure that the monthly aggregation of the daily product matches with the monthly product at each grid point. Mcf is, thus, calculated by using the following equation:

$$Mcf = \begin{cases} \frac{PISCOpm}{\sum_{i=1}^{N} P-PISCOpd_{(i)}} & if\ Mcf > 0 \\ 1 & if\ \sum_{i=1}^{N} P-PISCOpd_{(i)} = 0 \end{cases} \quad \ldots (11)$$

Where N represents the number of days of the corresponding month.

Finally, PISCOpd is defined as follows:

$$PISCOpd = P-PISCOpd * Mcf \quad \ldots (12)$$

2.5. Evaluation of PISCOp V.2.1

The process for evaluating the performance of PISCOp V2.1 was performed within the period 1981-2016 through two steps: Firstly, a pixel-to-point evaluation is carried out using an independent rainfall network (ID) which consists of 100 rain gauges (Figure 1a) not previously used for the development of PISCOp V2.1.

We selected all rain gauges with > 12 months of data between January 1981 and December 2016 that were located within a minimum distance of 20 km from PISCOp rain gauge network. Three continuous statistics were computed comparing the time series of PISCOp V2.1 and ID (Table 2). The Pearson Correlation Coefficient (CC) was used to evaluate the capability of PISCOp V2.1 to capture rainfall variability. The RMSE measuring the average magnitude of the error and the Percent bias (PBIAS) indicates the degree to which each PISCOp V2.1 value is over- or underestimated (Teng et al., 2014).

Table 2

Additionality, following the criteria defined by Zambrano-Bigiarini et al. (2017), the three categorical statistics (Table 2) probability of detection (POD), false alarm rate (FAR) and threat score (TS) were used to determine PISCOp V2.1 rainfall detection capabilities within five precipitation intensity classes (Table 3). The POD and FAR indicate which fraction of the observed events were correctly detected and which fraction of the events reported by the Grd's did not occur. TS is a general categorical statistic sensitive to hits and penalizes both missing and false alarms affected by the climatological frequency of the event. More information on the aforementioned indices can be found in Wilks (2006).

Table 3

Secondly, a water balance evaluation using two simple runoff ratios (RR and RR_f) is carried out in 19 Peruvian catchments following the equations:

$$RR = \frac{Q}{P} \ \ldots (13)$$

$$RR_f = \frac{Q}{(P-ET)} \ldots (14)$$

Where Q, P, and ET are the annual long-term average of the discharge, precipitation, and real evapotranspiration, respectively. Unlike the pixel-to-point approach, the runoff ratios (RR and RR_f) allowed assessing the long-term capacity of PISCOp V2.1 in more extensive areas. Due to the fact that the annual time step is a sufficiently large time period, we assume that the catchment storage is zero. Hence, Q and (P – ET) are expected to adopt similar values. The discharge gauges were obtained from the Environmental Research Observatory SO HYBAM (www.ore-hybam.org) and SENAMHI (Table 4 and Figure 1b). ET was calculated from Fu's equation (Yao et al., 2016) using the gridded maximum and minimum temperature dataset generated for Peru by Vicente-Serrano et al. (2017). Four Grd were used to estimate P: (1) PISCOp V2.1, (2) CHIRPM, (3) ORE-HYBAM (HOP; Guimberteau et al., 2012), which is only available for the Amazon and generated at 1° spatial resolution using ordinary kriging and (4) the previous PISCOp version (V1.0) which is based on the merging between CHIRPS and rain gauges applying kriging with external drift (Lavado et al., 2015). We used CHIRPM as a reference to explore possible improvements after the merging phase. We compare PISCO V2.1 and PISCOp V1.0 to examine the repercussions of changing CHIRPS by CHIRPM. The comparison to HOP serves to understand the role of the introduction of spatial predictors (CHIRPM). It is important to note that HOP, PISCOp V1.0, and PISCOp V2.1 present almost the same availability of rain gauges within selected upstream catchments. Therefore, the influence of rain gauge density can be handled as a common systematic variation for all Grd used.

Table 4

3. Results and Discussion

3.1. State and gap-infilling of the Peruvian rainfall dataset

According to the three-step QC approach, 3.51% out of total data (Table 5) was data with gross error which was deleted for the next steps. The most (least) affected sub-region with data exclusion was the AW (PC). Hunziker et al. (2017b) indicated that a large fraction of these gross errors are caused by the observers during data recording, while Hunziker et al. (2017a) mentioned that due to these errors and a large amount of missing data, 40% of the available rain gauges are inappropriate for climate analyses. Following our approach and considering that within the 1981-2016 period at least 10 years of continuous information must be available after applying QC, the number of stations had reduced from initially 945 to 441 rain gauges (Figure 1a). These rain gauges (henceforth called PISCOp rainfall network) form the basis and constitute the most valuable source of information for the construction of the gridded dataset.

Table 5

The density of PISCOp rainfall network lies at around 282/ (10^6 km²) for Peru (Table 5), with maximum (minimum) density in AW (AL). These results suggest a remarkably heterogeneous distribution and very sparse conditions across whole Peru. Despite data-scarcity conditions, the Peruvian rainfall network (Table 5) lies within the minimum requirements for hydrological analyses defined by the WMO (World Meteorological Organization, 1994) and it is similar to the rain gauge density presented in previous works related to the regional runoff simulation in adjacent regions of Peru (Getirana et al., 2014;

Guimberteau et al., 2012; Zulkafli et al., 2013). Also, most of the rain gauges are located in mountain recharge zones which is an important condition for rainfall-runoff simulation in Pacific and Andean drainage catchments (Figure 1). Nonetheless, the density of PISCOp rainfall network is far below conventional climatological datasets worldwide (Hofstra et al., 2009; Lussana et al., 2018; Newman et al., 2015; Yatagai et al., 2012). As the rain gauge density strongly influences the merging phase (Hofstra et al., 2010) is expected that regions with lesser density introduce biases in the mean and variance of the gridded dataset.

The gaps in the rain gauge series were identified as the main problem for the construction of a serially complete gauge dataset in Peru. In general, most rain gauges belonging to the PISCOp rainfall network were installed after the year 2000, which explains the high percentage of missing data (34.74%) before this date (Figure 3). For this reason, the amount of data available for the period 2001-2016 exceeds by 117% the period 1981-2000, with the most significant difference (204%) in the AL. The monthly (daily) gap-infilling approach based on neighboring stations allowed infilling the 44% (58%) of total missing data (Table 5).

Considering the fact that the neighboring information was not enough to create serially complete time series for all rain gauges, bias-corrected CHIRPM was added. Similar to other studies (Chaney et al., 2014; Teegavarapu & Nayak, 2017), we used the statistical tests Mann-Kendall (MK) and Kolmogorov-Smirnov (KS) to determine whether the performance of the gap-infilling procedure is appropriate.

Figure 3

Firstly, the KS nonparametric test that does not use any distributional assumptions, was employed as a metric to estimate if the distribution before and after the gap-infilling procedure was identical. Considering a 5% significance level, 54 (41) rain gauges at daily (monthly) time step were flagged for evidencing a poor matching between the cumulative distribution functions (Dn > 0.1, Figure S2). In addition, Table 5 indicated a spatial average of Dn bellows to 0.07 for all the sub-regions. Next, to identify the possible formation of spurious trends, the MK test was performed in annual time series from the period 1981-2016. At a 5% significance level, 18 rain gauges (Figure S2) were identified with positive trends (τ > 0.1). Both results suggest that the gap-infilling procedure adopted in this study is acceptable showing an unclear spatial pattern on results (Figure S2). Nonetheless, note that ~14 % of the rain gauge was flagged by at least one of the tests. Therefore, is expected that inhomogeneities are present in the serially complete gauge dataset. For an individual inspection of the gap-infilling procedure see http://piscoprec.github.io/gauge.

3.2. Spatial description of PISCOp

The rainfall pattern for January 1998 (Figure 4) and day twenty-five of this month (Figure 5) were used to perform a visual check of PISCOp products (PISCOpm and PISCOpd, Figure 2). This date was selected since an atypical and violent rain (Table 3) caused by the ENSO phenomenon was experienced in the north of Peru particularly.

The outputs analysis for January 1998 was performed considering as a reference the rainfall measured in this month by PISCOp rainfall network (Figure 4a). Thus, CHIRPm (Figure 4b) only showed an acceptable representation of the spatial structure of the rainfall field within the AW. For the AE, AAT, and AL a remarkable underestimation was evident with increasing rainfall rates. Underestimation of TIR-based Grd's, such as CHIRPm, in front of high rainfall values is a well-known condition (Kidd & Huffman, 2011). In contrast, the PC indicates unrealistically high rain values for the entire study period (1981-2016). This pattern is a consequence of the used CHP_{clim} as a spatial predictor. CHP_{clim} reasonably depicts the spatial structure in the PC. Nonetheless, the lack of rain gauges for calibration produces severe overestimation rainfall amounts (Figure 1S). Regarding the CHIRPMm product (Figure 4c), in general it improved the rainfall characterization in most of the sub-regions in comparison to CHIRPm. This improvement is explained by the use of $PISCO_{clim}$ instead of CHP_{clim} in the CHIRPMm construction. However, it is still unable to detect the convective storms caused by the ENSO in the northern PC.

Figure 4

In contrast to CHIRPm and CHIRPMm, the PISCOpm product (Figure 4d) underlies different rainfall intensities and spatial structure of the entire study area. These changes are the result of the interaction of the following three factors: (1) spatial autocorrelation among residuals (measured through the semivariogram), (2) the distribution of PISCOp rainfall network and (3) the magnitude and sign of the residuals. For the analyzed month, negative residuals and a considerable number of rain gauges lead to the rainfall increase in the AE and AW as well as to a better representation of the spatial pattern caused by ENSO in the northern PC. A similar behavior can be observed in the Peruvian Amazon (AAT and AL), where negative residuals and a reduced number of rain gauges lead to the negative local average, systematically increasing rainfall amounts when the de-correlation distance was overcome.

Even though the use of geostatistical interpolation methods allows the spatial coherence to be maximized, it must be taken into account that the predicted values in PISCOpm may differ from the rain gauge values particularly related to the formation of large residuals. This behavior has already been extensively described in Tozer et al. (2012) and further studies (Ensor & Robeson, 2008; Erdin et al., 2012; Hofstra et al., 2009). Therefore, special care must be taken when using PISCOp V2.1 for analyzing extreme events, such as those related to ENSO.

For the blend process at daily time step, several products (CHIRPMd, P-PISCOpd, and Mcf), were created before generating the final dataset PISCOpd. In previous studies it has been demonstrated that in data-scarce regions the areal rainfall estimates are better represented in blended than in only gauge-based Grd (Buytaert et al., 2006; Nerini et al., 2015; Schuurmans et al., 2007). Hence, P-PISCOp (Figure 5c) was produced merging the serially complete PISCOpd rainfall network (Figure 5a) with CHIRPMd (Figure 5b). Unlike CHIRPMd, P-PISCOp allowed an admissible characterization of convective storms that occurred at the northern PC and improved the rainfall scenario in AW. Nonetheless, a marked underestimation concerning CHIRPMd, was found for the rainy pixels located at the center of the AAT and AL. The explanation of this change is similar to the previous month analyzed, with the difference that at daily time step the predictor (CHIRPMd) represents the spatial variance worse causing higher instability in the residuals. Due to the absence of rain gauges, the residuals of the western sub-regions (PC, AW, and AE) are continuously and omnidirectionally transferred to the eastern sub-regions (AAT and AL) by IDW. The efficiency of this process directly depends on the residuals variance and intermittent rainfall regime (Chappell et al., 2012). A scenario with low variance and intermittency should improve the systematic bias, otherwise, this transfer would result in inaccurate precipitation values (Figure 5c). Finally, PISCOpd (Figure 5d), unlike P-PISCOpd, decreases the bull's-eye effect formation around the gauge observations relying on the spatial structure of PISCOpm. In addition, a clear precipitation increase can be observed for the AL and AAT due to the formation of Mcf values > 1 in these sub-regions.

Figure 5

3.3. Performance of PISCOp V.2.1

As Figure 6 shows, the sub-regions PC and AW indicate the most significant improvements for PISCOp V2.1 compared to CHIRPM, although the spread of their scores slightly increments. In these sub-regions, P-PISCOpd and PISCOpd (PISCOpm) increase the accuracy of the CC to 213% and 210% (14%) compared to CHIRPMd (CHIRPMm). The RMSE show consistent reduction in random error and the systematic PBIAS is close to 0. For the AE and AAT, the performance of PISCOp V2.1 continues indicating a substantial increase (reduction) of the CC (RMSE) score with respect to CHIRPM, although these values are worse in comparison to the PC and AW. This can be explained by a lesser density of rain gauges and that TIR-based retrieval algorithms imply a lower performance under high influence of topographic complexity (Derin et al., 2016; Mantas et al., 2015; Thouret et al., 2013). At monthly time step, PISCOpm provides the higher accuracy to capture the influence of the ITCZ migration through the tropical Andes despite a remarkable underestimation on the precipitation gradients for the eastern Andean slope. On the contrary, at daily time step, P-PISCOpd and PISCOpd showed a low performance that did not lead to any improvement compared to CHIRPMd. Finally, the AL presented the largest (lowest) RMSE (CC) score of our study area. Even at monthly time step, PISCOpm presented a lower performance than CHIRPMm. Similar results have recently been reported for the Ecuadorian Amazon by Ulloa et al. (2017) who stated that the reduced number of rain gauges used for the merging phase generates spatial inconsistencies for the entire analysis period.

Figure 6

According to the categorical indices applied in this study, both P-PISCOpd and PISCOpd presented similar scores that were higher than CHIRPMd in the entire study area (Figure 7). In general, the detection capacity of the three products weakens as far as the precipitation intensity category (Table 3) increases, regardless of the geographical position. Comparing the categories ‘no rain’ and ‘violent rain’, the POD (FAR) score decreased (increased) drastically by 452% (245%). The TS evidenced that the daily CHIRPMd, PISCOpd and P-PISCOpd products were not able to correctly capture the fraction of rainfall events for all sub-regions. These results evidence that daily PISCOp products (CHIRPMd, P-PISCOpd, and PISCOpd) are most likely not sufficiently accurate for capturing heavy rainfall events. Hence, the use of PISCOp products to describe the intensity of extreme precipitation events is not recommended if no high-density rainfall network exist nearby.

Figure 7

Figure 8 illustrates the water balance evaluation of PISCOp V2.1, CHIRPM and other two Grd (PISCOp V1.0 and HOP) using runoff ratios (RR_f and RR). The widest spread in RR and RR_f scores were observed within the Amazon basin indicating that PISCOp V.2.1 reaches a similar score compared to HOP and an underestimation (overestimation) of 15% (28%) with regard to CHIRPM (PISCOp V1.0). For the Amazon, RR estimates are typically below 0.8 (Costa & Foley, 1997; Gusev et al., 2017; Rudorff et al., 2014) while the average values of PISCOp V.2.1, CHIRPM, PISCOp V1.0, and HOP are situated around 0.89, 0.75, 1.05 and 0.88, respectively. Similar results have been found in RR_f, where PISCOp V2.1, CHIRPM, and HOP provide values close to 1 ($-0.2 < RR_f < 0.2$). Although these high runoff ratios could be explained by excessive groundwater contribution (Zubieta *et al.*, 2015), a rainfall underestimation scenario is a more likely explanation, especially when considering the independent validation results which indicate the PBIAS trend to be negative (Figure 6). The RR_f within the Amazon evidences that CHIRPM achieves the best agreement with discharge values. Additionally, the climate correction based on $PISCOp_{clim}$ (used in CHIRPM and PISCOp V2.1) leads to a better performance than CHP_{clim} (used in PISCOp V1.0) in order to eliminate the underestimation of rainfall, particularly at the eastern slopes of the Andes. For catchments covering the Andes or Pacific, CHIRPM, PISCOp V1.0, and PISCOp V2.1 presented very slight differences in their RR and RR_f scores with values mostly below 1. Unlike the Amazon, in the Andes-Pacific the change of CHP_{clim} to $PISCOp_{clim}$ does not affect the areal rainfall estimations, possibly due to a better distribution and higher density of the rainfall network. The RR_f values of CHIRPM, PISCOp V2.1, and PISCOp V1.0 exhibited a systematic overestimation mainly for catchments >10,000 km² (Figure 8). This uncertainty could be related to streamflow alteration caused by anthropogenic factors or low ET estimates. However, it is difficult to predict and beyond the scope of this investigation.

Figure 8

3.4. Impacts and detection of inhomogeneities

All inputs used for the development of PISCOp V2.1 are not thoroughly homogenized. The Peruvian long-term gauge dataset is affected by a plethora of non-climatic factors, such as changes of instruments or bad observer practices (Brönnimann, 2015; New et al., 2000; Peterson et al., 1998). In addition, inconsistencies are also present in the CHIRP algorithm. For instance, they arise when infilling missing CHIRP values with the Coupled Forecast System version 2 (Saha et al., 2014) or in the overlap between the TIR archives: Globally Gridded Satellite (GriSat) and NOAA Climate Prediction Center (CPC, Funk et al., 2015). In order to detect the spatial pattern of these inhomogeneities, we apply the Pettitt test with a significance level of 5% for annual time series of each PISCOp V2.1 grid cell, and as for the water balance estimation, we use CHIRPM, PISCOp V1.0, and HOP as a reference.

Figure 9a shows the years when a breakpoint was detected. In general, the three Grd indicate a wide variability in their break years and spatial extent. CHIRPM obtained the smaller inhomogeneity area within the Amazon basin (8.1%), followed by PISCOp V.2.1 (8.2%), HOP (34%) and PISCOp V1.0 (89%). The breakpoint year observed for CHIRPM is associated with the transition from GriSat to CPC, whereas the breakpoint for PISCOp V1.0 notably coincides with the changes of the density of PISCOp rainfall network (Figure 3). Regarding PISCOp V2.1 and HOP, the inhomogeneity area drastically decreased due to the data gap infilling performed at each station and a further balancing-out during the geostatistical interpolation. Although breakpoints in rainfall time series can naturally occur, no evidence was found for any of the 56 rain gauges with more than 95% of complete time series (Figure 1a).

Figure 9

Based on the breakpoints detected for each cell, a sensitive area (S1, Figure 9a) was defined to analyze the plausibility of the time series in more detail. As shown in Figure 8b, the intensities, break year at 5% significance level (red dotted line) and the seasonality of the three Grd at monthly time step vary considerably despite using similar inputs. The inhomogeneities in Grd imply severe impacts for the analysis of the seasonal (not shown here) and annual trend (Figure 9c). For the assessment of these impacts, the Sen's slope estimator at 95% confidence level was used. For the 1981-2016 period, PISCOp V2.1, PISCOp V1.0, and HOP revealed a significant positive trend that exceeded 55%, 81%, and 72%, respectively, the slope of CHIRPM. Artificial trends (Hofstra et al., 2010; Kingdom, 2014; Nicolas & Bromwich, 2011; Tozer et al., 2012) are principally spread across the entire Peruvian Amazon, especially where data-scarcity prevails and a high amount of gaps exists.

4. Summary and conclusions

In this paper, we presented the development of PISCOp V2.1, a new daily and monthly long-term Grd for 1981 until the present. This gridded product was generated based on the integration of serially complete gauge datasets, CHIRP data, radar-based climatologies, and geostatistical and deterministic interpolation methods. The quality of PISCOp V2.1 was assessed within six hydro-meteorological sub-regions with an independent rain gauge network. Additionality, the runoff ratio estimation within 19 catchments allowed to evaluate the performance of PISCOp V2.1 in more extensive areas.

For the first time, PISCOp V2.1 comprehensively presents the state of the Peruvian rain gauge dataset for the period 1981-2016. It has been identified that the gaps in rainfall time series represent the most determining problem for the construction of a temporally consistent Grd. This study shows that the combinations between both CUTOFF and Qm are a conservative and efficient method for successful data gap-infilling, especially for large data gaps that prevail at the beginning of the PISCOp V2.1 period. However, this method strongly depends on the predictor and the proximity of rain gauges. Hence, it is expected that our approach applied to areas with low station density and poor performance of CHIRPM might lead to unsatisfactory results.

The independent and water balance evaluation confirm that PISCOp V2.1 is the most suitable product for representing area rainfall estimates, except for the Amazon lowland where CHIRPM obtained better results. Additionally, we note the climatological correction based on $\text{PISCOp}_{\text{clim}}$ significantly improved the results compared to CHP_{clim}. At daily time step, PISCOpd did not capture the convective storm intensity regardless of the geographical position. Although it seems highly attractive to use gridded data with full spatial and temporal coverage, such as PISCOp V2.1, all inhomogeneities inherent to this merged product and presented in this study must be entirely taken into account. Therefore, like for other blended Grd products (Beck et al., 2017a; Yanto et al., 2017b), we recommend to the users special care when using PISCOp V.2.1 for the analysis of trends, extreme events or other applications related to e.g. climate change.

New versions of PISCOp have been planned and it is expected to make use of new existing information sources such as IMERG (Huffman et al., 2015), cross-border rain gauges as well as to improve the data gap-infilling of rainfall time series.

5. Data Access

The PISCOp V2.1 product, source code, and additional information are freely available to users in NetCDF (1981-2016) and GeoTIFF (1981-present) format at the following website: https://piscoprec.github.io/.

Abbreviations

The full name for the abbreviations used in this paper are:

CHIRP	Climate Hazards Group Infrared Precipitation estimates (without station data).
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station data
CHIRPM	CHIRP modified
CHPclim	Climate Hazards Group's Precipitation Climatology
IRP	Product calculated from the linear model between TIR CCD% and TRMM 3B42 V7.
QC	Quality control
Qm	Quantile mapping
PISCOp	Peruvian Interpolated data of SENAMHI's Climatological and Hydrological Observations - precipitation product
PISCOpclim	PISCOp climatological product
PISCOpd	Daily PISCOp product
P-PISCOp	Provisional daily PISCOp product
PISCOpm	Provisional monthly PISCOp product
TRMM 3B42 V7	3B42 product of TRMM version 7

TRMM 2A25	2A25 product of TRMM
TIR CCD%	Thermal Infrared Cold Cloud Duration Percentage

Acknowledgments

We thank SENAMHI and CLIMANDES for supporting this research and acknowledge hundreds of hydrometeorological observers around Peru for their work which has made possible this study. The authors would like to thank Stephanie Gleixner, Fred F. Hattermann and Fabian Drenkhan for his valuable comments.

References

Adhikary, S.K., 2017. Cokriging for enhanced spatial interpolation of rainfall in two Australian catchments 2143–2161. https://doi.org/10.1002/hyp.11163

Antico, P.L., 2009. Relationships between autumn precipitation anomalies in southeastern South America and El Niño event classification 727, 719–727. https://doi.org/10.1002/joc

Babak, O., Deutsch, Æ.C. V, 2009. Statistical approach to inverse distance interpolation 543–553. https://doi.org/10.1007/s00477-008-0226-6

Baik, J., Park, J., Ryu, D., Choi, M., 2015. Geospatial blending to improve spatial mapping of precipitation with high spatial resolution by merging satellite- and ground based data. https://doi.org/10.1002/hpy.10786

Beck, H.E., Van Dijk, A.I.J.M., Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B., De Roo, A., 2017a. MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data. Hydrol. Earth Syst. Sci. 21, 589–615. https://doi.org/10.5194/hess-21-589-2017

Beck, H.E., Vergopolan, N., Pan, M., Levizzani, V., Dijk, A.I.J.M. Van, Brocca, L., Pappenberger, F., Huffman, G.J., Wood, E.F., 2017b. Global-scale evaluation of 23 precipitation datasets using gauge observations and hydrological modeling.

Beck, H.E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A.I.J.M., Weedon, G., Brocca, L., Pappenberger, F., Huffman, G.J., Wood, E.F., 2017c. Global-scale evaluation of 23 precipitation datasets using gauge observations and hydrological modeling. Hydrol. Earth Syst. Sci. Discuss. 1–23. https://doi.org/10.5194/hess-2017-508

Beguería, S., Vicente-serrano, S.M., Tomás-Burgueraa, M., Maneta, M., 2015. Bias in the variance of gridded data sets leads to misleading conclusions about changes in climate variability Bias in the variance of gridded data sets leads to misleading. https://doi.org/10.1002/joc.4561

Bi, S., Bi, S., Chen, D., Pan, J., Wang, J., 2017. A Double-Smoothing Algorithm for Integrating Satellite Precipitation Products in Areas with Sparsely Distributed In Situ Networks. Isprs Int. J. Geo-Information 6. https://doi.org/10.3390/ijgi6010028

Boers, N., Bookhagen, B., Marwan, N., Kurths, J., 2013. Complex networks identify spatial patterns of extreme rainfall events of the South American Monsoon System 40, 1–7. https://doi.org/10.1002/grl.50681

Bookhagen, B., Strecker, M.R., 2008. Orographic barriers, high-resolution TRMM rainfall, and relief variations along the eastern Andes. Geophys. Res. Lett. 35, 1–6. https://doi.org/10.1029/2007GL032011

Brönnimann, S., 2015. Climatic Changes Since 1700. Springer International Publishing: Cham, Switzerland. https://doi.org/10.1007/978-3-319- 19042-6_4

Buytaert, W., Celleri, R., Willems, P., Biévre, B. De, Wyseure, G., 2006. Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. J. Hydrol. 329, 413–421. https://doi.org/10.1016/j.jhydrol.2006.02.031

Campozano, L., Célleri, R., Trachte, K., Bendix, J., Samaniego, E., 2016. Rainfall and Cloud Dynamics in the Andes: A Southern Ecuador Case Study. Adv. Meteorol. 2016. https://doi.org/10.1155/2016/3192765

Carvalho, L.M. V, Silva, A.E., Jones, C., Liebmann, B., Silva, P.L., Humberto, D., 2011. Moisture transport and intraseasonal variability in the South America monsoon system 1865–1880. https://doi.org/10.1007/s00382-010-0806-2

Chaney, N.W., Sheffield, J., Villarini, G., Wood, E.F., 2014. Development of a high-resolution gridded daily meteorological dataset over sub-Saharan Africa: Spatial analysis of trends in climate extremes. J. Clim. 27, 5815–5835. https://doi.org/10.1175/JCLI-D-13-00423.1

Chappell, A., Renzullo, L., Haylock, M., 2012. Spatial uncertainty to determine reliable daily precipitation maps 117, 1–14. https://doi.org/10.1029/2012JD017718

Chavez, S.P., Takahashi, K., 2017. Orographic rainfall hot spots in the Andes-Amazon transition according to the TRMM precipitation radar and in situ data. J. Geophys. Res. Atmos. https://doi.org/10.1002/2016JD026282

Chen, F.W., Liu, C.W., 2012. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. Paddy Water Environ. 10, 209–222. https://doi.org/10.1007/s10333-012-0319-1

Costa, M.H., Foley, J. a., 1997. Water balance of the Amazon Basin: Dependence on vegetation cover and canopy conductance. J. Geophys. Res. 102, 23973. https://doi.org/10.1029/97JD01865

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., Chang, C.-H., Delrieu, G., Hong, Y., Hsu, Y.C., Lavado-Casimiro, W., Manz, B., Moges, S., Nikolopoulos, E.I., Sahlu, D., Salerno, F., Rodríguez-Sánchez, J.-P., Vergara, H.J., Yilmaz, K.K., 2016. Multiregional Satellite Precipitation Products Evaluation over Complex Terrain. J. Hydrometeorol. 17, 1817–1836. https://doi.org/10.1175/JHM-D-15-0197.1

Dinku, T., Hailemariam, K., Maidment, R., Tarnavsky, E., Connor, S., 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. https://doi.org/10.1002/joc.3855

Ensor, L.A., Robeson, S.M., 2008. Statistical characteristics of daily precipitation: Comparisons of gridded and point datasets. J. Appl. Meteorol. Climatol. 47, 2468–2476. https://doi.org/10.1175/2008JAMC1757.1

Erdin, R., Frei, C., Künsch, H.R., 2012. Data Transformation and Uncertainty in Geostatistical Combination of Radar and Rain Gauges. J. Hydrometeorol. 13, 1332–1346. https://doi.org/10.1175/JHM-D-11-096.1

Feng, L., Nowak, G., Neill, T.J.O., Welsh, A.H., 2014. CUTOFF : A spatio-temporal imputation method. J. Hydrol. 519, 3591–3605. https://doi.org/10.1016/j.jhydrol.2014.11.012

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015a. The climate hazards infrared precipitation with stations — a new environmental record for monitoring extremes 1–21. https://doi.org/10.1038/sdata.2015.66

Funk, C., Verdin, A., Michaelsen, J., Peterson, P., Pedreros, D., Husak, G., 2015b. A global satellite-assisted precipitation climatology. Earth Syst. Sci. Data 7, 275–287. https://doi.org/10.5194/essd-7-275-2015

Garreaud, R., 1999. Multiscale Analysis of the Summertime Precipitation over the Central Andes. Mon. Weather Rev. 127, 901–921. https://doi.org/10.1175/1520-0493(1999)127<0901:MAOTSP>2.0.CO;2

Garreaud, R.D., Vuille, M., Compagnucci, R., Marengo, J., 2009. Present-day South American climate. Palaeogeogr. Palaeoclimatol. Palaeoecol. 281, 180–195. https://doi.org/10.1016/j.palaeo.2007.10.032

Getirana, A.C. V., Dutra, E., Guimberteau, M., Kam, J., Li, H.-Y., Decharme, B., Zhang, Z., Ducharne, A., Boone, A., Balsamo, G., Rodell, M., Toure, A.M., Xue, Y., Peters-Lidard, C.D., Kumar, S. V., Arsenault, K., Drapeau, G., Ruby Leung, L., Ronchail, J., Sheffield, J., 2014. Water Balance in the Amazon Basin from a Land Surface Model Ensemble. J. Hydrometeorol. 15, 2586–2614. https://doi.org/10.1175/JHM-D-14-0068.1

Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J. Hydrol. 228, 113–129. https://doi.org/10.1016/S0022-1694(00)00144-X

Grimes, D.I.F., Pardo-Igúzquiza, E., Bonifacio, R., 1999. Optimal areal rainfall estimation using raingauges and satellite data. J. Hydrol. 222, 93–108. https://doi.org/10.1016/S0022-1694(99)00092-X

Gudmundsson, L., Bremnes, J.B., Haugen, J.E., Engen-Skaugen, T., 2012. Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – A comparison of methods. Hydrol. Earth Syst. Sci. 16, 3383–3390. https://doi.org/10.5194/hess-16-3383-2012

Guimberteau, M., Drapeau, G., Ronchail, J., Sultan, B., Polcher, J., Martinez, J.M., Prigent, C., Guyot, J.L., Cochonneau, G., Espinoza, J.C., Filizola, N., Fraizy, P., Lavado, W., De Oliveira, E., Pombosa, R., Noriega, L., Vauchel, P., 2012. Discharge simulation in the sub-basins of the Amazon using ORCHIDEE forced by new datasets. Hydrol. Earth Syst. Sci. 16, 911–935. https://doi.org/10.5194/hess-16-911-2012

Gusev, Y.M., Nasonova, O.N., Kovalev, E.E., Aizel, G. V., 2017. Modelling river runoff and estimating its weather-related uncertainty for 11 large-scale rivers located in different regions of the globe. Hydrol. Res. 2013, nh2017015. https://doi.org/10.2166/nh.2017.015

Hiemstra, P.H., Pebesma, E.J., Twenho, C.J.W., 2009. Computers & Geosciences Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network 35, 1711–1721. https://doi.org/10.1016/j.cageo.2008.10.011

Hofstra, N., Haylock, M., New, M., Jones, P.D., 2009. Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature 114. https://doi.org/10.1029/2009JD011799

Hofstra, N., New, M., McSweeney, C., 2010. The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. Clim. Dyn. 35, 841–858. https://doi.org/10.1007/s00382-009-0698-1

Houze, R.A., Rasmussen, K.L., Zuluaga, M.D., Brodzik, S.R., 2015. The variable nature of convection in the tropics and subtropics: A legacy of 16 years of the Tropical Rainfall Measuring Mission satellite. Rev. Geophys. 994–1021. https://doi.org/10.1002/2015RG000488.Received

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., 2015. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. NASA/GSFC Code 612, 47. https://doi.org/10.1136/openhrt-2016-000469

Hunziker, S., Brönnimann, S., Calle, J.M., Moreno, I., Andrade, M., Ticona, L., Huerta, A., Lavado-Casimiro, W., 2017a. Effects of undetected data quality issues on climatological analyses. Clim. Past Discuss. 1–31. https://doi.org/10.5194/cp-2017-64

Hunziker, S., Gubler, S., Calle, J., Moreno, I., Andrade, M., Velarde, F., Ticona, L., Carrasco, G., Castellón, Y., Oria, C., Croci-Maspoli, M., Konzelmann, T., Rohrer, M., Bronnimann, S., 2017b. Identifying, attributing, and overcoming common data quality issues of manned station observations. Int. J. Climatol. https://doi.org/10.1002/joc.5037

Iguchi, T., Kozu, T., Meneghini, R., Awaka, J., Okamoto, K., 2000. Rain-Profiling Algorithm for the TRMM Precipitation Radar. J. Appl. Meteorol. 39, 2038–2052. https://doi.org/10.1175/1520-0450(2001)040<2038:RPAFTT>2.0.CO;2

Isotta, F.A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S.M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., Vertačnik, G., 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, 1657–1675. https://doi.org/10.1002/joc.3794

Javier, F., Trejo, P., Civil, D.D.I., Llanos, U.D.L., Zamora, E., Carlos, S., Peñaloza-murillo, M.A., Moreno, M.A., Farías, A., 2016. Intercomparison of improved satellite rainfall estimation with CHIRPS gridded product and rain gauge data over Venezuela 29, 323–342. https://doi.org/10.20937/ATM.2016.29.04.04

Kang, S., Ahn, J.B., 2015. Global energy and water balances in the latest reanalyses. Asia-Pacific J. Atmos. Sci. 51, 293–302. https://doi.org/10.1007/s13143-015-0079-0

Keller, V.D.J., Tanguy, M., Prosdocimi, I., Terry, J.A., Hitt, O., Cole, S.J., Fry, M., Morris, D.G., Dixon, H., 2015. CEH-GEAR: 1 Km resolution daily and monthly areal rainfall estimates for the UK for hydrological and other applications. Earth Syst. Sci. Data 7, 143–155. https://doi.org/10.5194/essd-7-143-2015

Kidd, C., Huffman, G., 2011. Review Global precipitation measurement 353, 334–353. https://doi.org/10.1002/met.284

Kingdom, U., 2014. Advances in the Stochastic Modeling of Satellite-Derived Rainfall Estimates Using a Sparse Calibration Dataset 1810–1831. https://doi.org/10.1175/JHM-D-13-0145.1

Lavado Casimiro, W.S., Ronchail, J., Labat, D., Espinoza, J.C., Guyot, J.L., 2012a. Basin-scale analysis of rainfall and runoff in Peru (1969–2004): Pacific, Titicaca and Amazonas drainages. Hydrol. Sci. J. 57, 625–642. https://doi.org/10.1080/02626667.2012.672985

Lavado Casimiro, W.S., Ronchail, J., Labat, D., Espinoza, J.C., Guyot, J.L., 2012b. Basin-scale analysis of rainfall and runoff in Peru (1969–2004): Pacific, Titicaca and Amazonas drainages. Hydrol. Sci. J. 57, 625–642. https://doi.org/10.1080/02626667.2012.672985

Lavado, W., Fernandez, C., Vega, F., Caycho, T., Endara, S., Huerta, A., Obando, O.F., 2015. PISCO: Peruvian Interpolated data of the SENAMHI's Climatological and hydrological Observations. Precipítación v1.0.

Li, M., Shao, Q., 2010. An improved statistical approach to merge satellite rainfall estimates and raingauge data. J. Hydrol. 385, 51–64. https://doi.org/10.1016/j.jhydrol.2010.01.023

Lussana, C., Saloranta, T., Skaugen, T., Magnusson, J., Einar Tveito, O., Andersen, J., 2018. SeNorge2 daily precipitation, an observational gridded dataset over Norway from 1957 to the present day. Earth Syst. Sci. Data 10, 235–249. https://doi.org/10.5194/essd-10-235-2018

Ly, S., Charles, C., Degré, A., 2013. Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: a review. Base 17, 392–406.

Mantas, V.M., Liu, Z., Caro, C., Pereira, A.J.S.C., 2015. Validation of TRMM multi-satellite precipitation analysis (TMPA) products in the Peruvian Andes. Atmos. Res. 163, 132–145. https://doi.org/10.1016/j.atmosres.2014.11.012

Manz, B., Buytaert, W., Zulkafli, Z., Lavado, W., Willems, B., Robles, L.A., 2016. High-resolution satellite-gauge merged precipitation climatologies of the Tropical Andes. https://doi.org/10.1002/2015JD023788.Received

Manz, B., Páez-Bimos, S., Horna, N., Buytaert, W., Ochoa-Tocachi, B., Lavado-Casimiro, W., Willems, B., 2017. Comparative ground validation of IMERG and TMPA 2 at variable spatio-temporal scales in the Tropical Andes. J. Hydrometeorol. https://doi.org/10.1175/JHM-D-16-0277.1

Marengo, J.A., Liebmann, B., Grimm, A.M., Misra, V., Silva Dias, P.L., Cavalcanti, I.F.A., Carvalho, L.M.V., Berbery, E.H., Ambrizzi, T., Vera, C.S., Saulo, A.C., Nogues-Paegle, J., Zipser, E., Seth, A., Alves, L.M., 2012. Recent developments on the South American monsoon system. Int. J. Climatol. 32, 1–21. https://doi.org/10.1002/joc.2254

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

Hydrometeorol. 150904104740009. https://doi.org/10.1175/JHM-D-14-0197.1

Nesbitt, S.W., Anders, A.M., 2009. Very high resolution precipitation climatologies from the Tropical Rainfall Measuring Mission precipitation radar. Geophys. Res. Lett. 36, 1–5. https://doi.org/10.1029/2009GL038026

New, M., Hulme, M., Jones, P., 2000. Representing Twentieth-Century Space – Time Climate Variability . Part II : Development of 1901 – 96 Monthly Grids of Terrestrial Surface Climate. J. Clim. 13, 2217–2238.

Newman, A.J., Clark, M.P., Craig, J., Nijssen, B., Wood, A., Gutmann, E., Mizukami, N., Brekke, L., Arnold, J.R., 2015. Gridded Ensemble Precipitation and Temperature Estimates for the Contiguous United States. J. Hydrometeorol. 16, 2481–2500. https://doi.org/10.1175/JHM-D-15-0026.1

Nicolas, J.P., Bromwich, D.H., 2011. Precipitation Changes in High Southern Latitudes from Global Reanalyses: A Cautionary Tale. Surv. Geophys. 32, 475–494. https://doi.org/10.1007/s10712-011-9114-6

Notivoli, R.S., Beguería, S., Saz, M.Á., Luis, M.D.E., 2017. Reconstrucción e Incertidumbre en Series de Precipitación Diaria Instrumental. X Congr. Int. AEC Clima, Soc. riesgos y Ord. del Territ.

Parmentier, B., McGill, B.J., Wilson, A.M., Regetz, J., Jetz, W., Guralnick, R., Tuanmu, M.N., Schildhauer, M., 2015. Using multi-timescale methods and satellite-derived land surface temperature for the interpolation of daily maximum air temperature in Oregon. Int. J. Climatol. 35, 3862–3878. https://doi.org/10.1002/joc.4251

Perdigón-Morales, J., Romero-Centeno, R., Pérez, P.O., Barrett, B.S., 2017. The midsummer drought in Mexico: perspectives on duration and intensity from the CHIRPS precipitation database. Int. J. Climatol. https://doi.org/10.1002/joc.5322

Peterson, T.C., Easterling, D.R., Karl, T.R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Salinger, J., Hanssen-bauer, I., Alexandersson, H., Jones, P., Parker, D., 1998. Homogeneity adjustments of in of in situ atmospheric climate data: a review 1517, 1493–1517.

Rudorff, C.M., Melack, J.M., Bates, P.D., 2014. Flooding dynamics on the lower Amazon floodplain: 2. Seasonal and interannual hydrological variability. Water Resour. Res. 50, 635–649. https://doi.org/10.1002/2013WR014714

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.T., Chuang, H.Y., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M.P., Van Den Dool, H., Zhang, Q., Wang, W., Chen, M., Becker, E., 2014. The NCEP climate forecast system version 2. J. Clim. 27, 2185–2208. https://doi.org/10.1175/JCLI-D-12-00823.1

Schuurmans, J.M., Bierkens, M.F.P., Pebesma, E.J., Uijlenhoet, R., 2007. Automatic Prediction of High-Resolution Daily Rainfall Fields for Multiple Extents: The Potential of Operational Radar. J. Hydrometeorol. 8, 1204–1224. https://doi.org/10.1175/2007JHM792.1

Stisen, S., Tumbo, M., 2015. Interpolation of daily raingauge data for hydrological modelling in data sparse regions using pattern information from satellite data. Hydrol. Sci. J. 1–16. https://doi.org/10.1080/02626667.2014.992789

Strauch, M., Kumar, R., Eisner, S., Mulligan, M., Reinhardt, J., Santini, W., Vetter, T., Friesen, J., 2016. Adjustment of global precipitation data for enhanced hydrologic modeling of tropical Andean watersheds. Clim. Change 1–14. https://doi.org/10.1007/s10584-016-1706-1

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., Hsu, K.-L., 2017. A review of global precipitation datasets: data sources, estimation, and intercomparisons. Rev. Geophys. 79–107. https://doi.org/10.1002/2017RG000574

Teegavarapu, R.S.V., Nayak, A., 2017. Evaluation of long-term trends in extreme precipitation: Implications of in-filled historical data use for analysis. J. Hydrol. 550, 616–634. https://doi.org/10.1016/j.jhydrol.2017.05.030

Teng, H., Shi, Z., Ma, Z., Li, Y., 2014. Estimating spatially downscaled rainfall by regression kriging using TRMM precipitation and elevation in Zhejiang Province, southeast China. Int. J. Remote Sens. 35, 7775–7794. https://doi.org/10.1080/01431161.2014.976888

Thouret, J.C., Enjolras, G., Martelli, K., Santoni, O., Luque, J.A., Nagata, M., Arguedas, A., MacEdo, L., 2013. Combining criteria for delineating lahar-and flash-flood-prone hazard and risk zones for the city of Arequipa, Peru. Nat. Hazards Earth Syst. Sci. 13, 339–360. https://doi.org/10.5194/nhess-13-339-2013

Tozer, C.R., Kiem, A.S., Verdon-Kidd, D.C., 2012. On the uncertainties associated with using gridded rainfall data as a proxy for observed. Hydrol. Earth Syst. Sci. 16, 1481–1499. https://doi.org/10.5194/hess-16-1481-2012

Ulloa, J., Ballari, D., Campozano, L., Samaniego, E., 2017. Two-Step Downscaling of Trmm 3b43 V7 Precipitation in Contrasting Climatic Regions With Sparse Monitoring: The Case of Ecuador in Tropical South America. Remote Sens. 9, 758. https://doi.org/10.3390/rs9070758

van Osnabrugge, B., Weerts, A.H., Uijlenhoet, R., 2017. genRE: A Method to Extend Gridded Precipitation Climatology Data Sets in Near Real-Time for Hydrological Forecasting Purposes. Water Resour. Res. 53, 9284–9303. https://doi.org/10.1002/2017WR021201

Vera, C., Baez, J., Douglas, M., Emmanuel, C.B., Marengo, J., Meitin, J., Nicolini, M., Nogues-Paegle, J., Paegle, J., Penalba, O., Salio, P., Saulo, C., Dias, M.A.S., Dias, P.S., Zipser, E., 2006. The South American Low-Level Jet Experiment 63–77. https://doi.org/10.1175/BAMS-87-1-63

Verdin, A., Rajagopalan, B., Kleiber, W., Funk, C., 2015. A Bayesian kriging approach for blending satellite and ground precipitation observations. Water Resour. Res. 1359–1371. https://doi.org/10.1002/2014WR015432.Received

Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., García-Vera, M.A., Stepanek, P., 2010. A complete daily precipitation database for northeast Spain: Reconstruction, quality control, and homogeneity. Int. J. Climatol. 30, 1146–1163. https://doi.org/10.1002/joc.1850

Vicente-Serrano, S.M., López-Moreno, J.I., Correa, K., Avalos, G., Bazo, J., Azorin-Molina, C., Domínguez-Castro, F., Kenawy, A. El, Gimeno, L., Nieto, R., 2017. Recent changes in monthly surface air temperature over Peru, 1964-2014. Int. J. Climatol. https://doi.org/10.1002/joc.5176

Wilks, D.S., 2006. Statistical methods in the atmospheric sciences, Statistical Methods in the Atmospheric Sciences.2nd ed.

WMO, 1994. Guide to hydrological practices,World Meteorological Organization.

Yanto, Livneh, B., Rajagopalan, B., Gu, G., Wang, L., 2017a. Development of a gridded meteorological dataset over Java island, Indonesia 1985–2014. Sci. Data 4, 170072. https://doi.org/10.1038/sdata.2017.72

Yanto, Livneh, B., Rajagopalan, B., Gu, G., Wang, L., 2017b. Development of a gridded meteorological dataset over Java island, Indonesia 1985–2014. Sci. Data 4, 170072. https://doi.org/10.1038/sdata.2017.72

Yao, J., Mao, W., Yang, Q., Xu, X., Liu, Z., 2016. Annual actual evapotranspiration in inland river catchments of China based on the Budyko framework. Stoch. Environ. Res. Risk Assess. https://doi.org/10.1007/s00477-016-1271-1

Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., Kitoh, A., 2012. Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. Bull. Am. Meteorol. Soc. 93, 1401–1415. https://doi.org/10.1175/BAMS-D-11-00122.1

Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K., Ribbe, L., 2017. Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile. Hydrol. Earth Syst. Sci. 21, 1295–1320. https://doi.org/10.5194/hess-21-1295-2017

Zubieta, R., Getirana, A., Espinoza, J.C., Lavado, W., 2015. Impacts of satellite-based precipitation datasets on rainfall-runoff modeling of the Western Amazon basin of

Peru and Ecuador. J. Hydrol. 528, 599–612. https://doi.org/10.1016/j.jhydrol.2015.06.064

Zulkafli, Z., Buytaert, W., Onof, C., Lavado, W., Guyot, J.L., 2013. A critical assessment of the JULES land surface model hydrology for humid tropical environments. Hydrol. Earth Syst. Sci. 17, 1113–1132. https://doi.org/10.5194/hess-17-1113-2013

Zulkafli, Z., Buytaert, W., Onof, C., Manz, B., Tarnavsky, E., Lavado, W., Guyot, J.-L., 2014. A Comparative Performance Analysis of TRMM 3B42 (TMPA) Versions 6 and 7 for Hydrological Applications over Andean–Amazon River Basins. J. Hydrometeorol. 15, 581–592. https://doi.org/10.1175/JHM-D-13-094.1

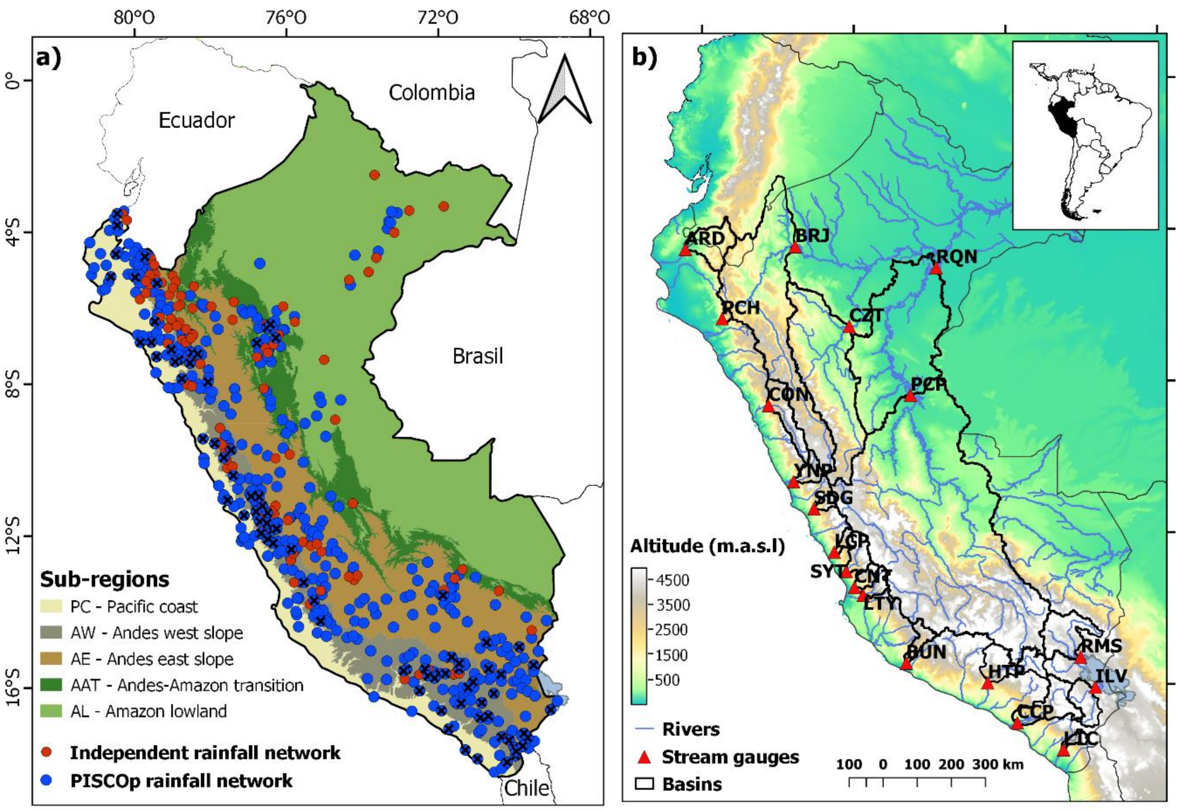

Figure 1. a) Spatial extension of PISCOp V2.1.The blue and red points represent PISCOp V2.1 and the independent rainfall network, respectively. The points with a cross indicate stations with more than 95% of data within the 1981-2016 period; b) Location and upstream catchments of the selected stream gauges.

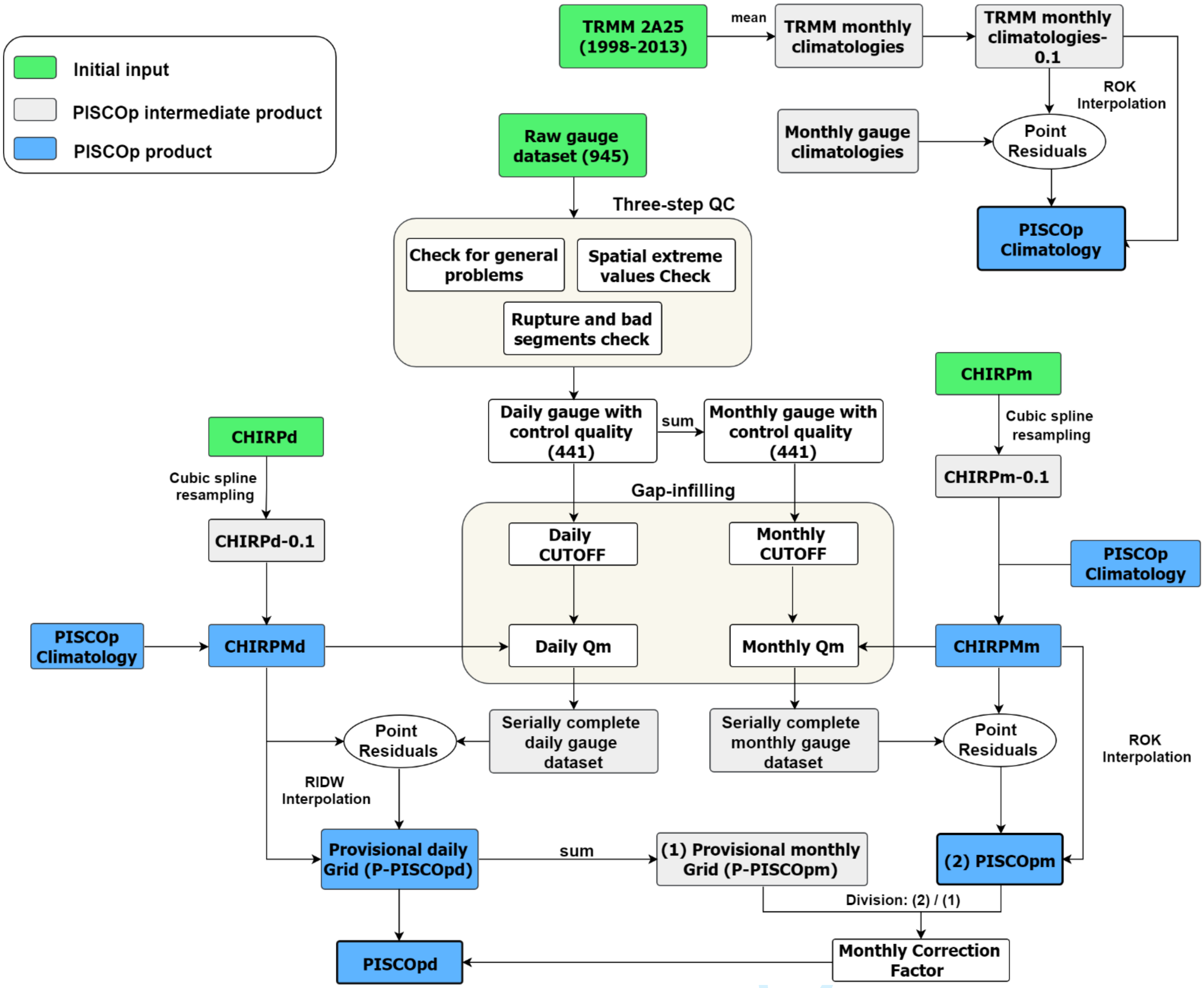

Figure 2. Schematic overview of the development of PISCOp V2.1.

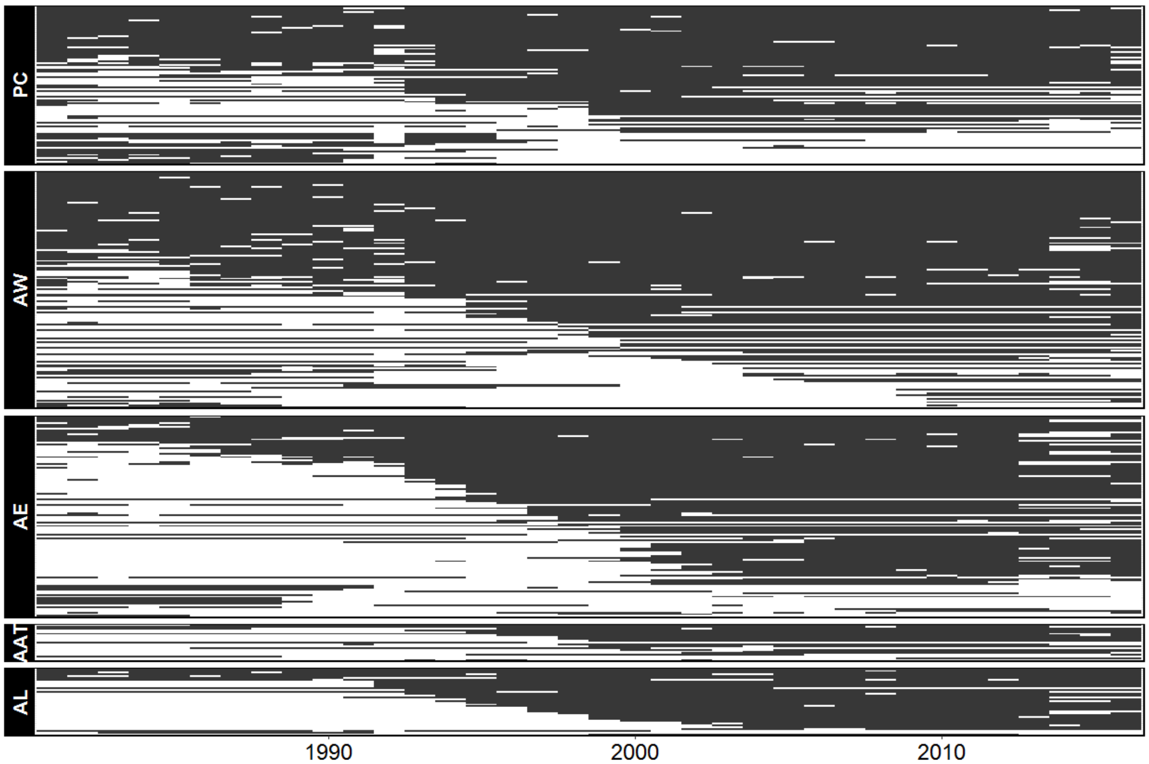

Figure 3. Missing data for each sub-region. Note: the rain gauges are ordered from lowest (upper line) to highest (lower line) missing values.

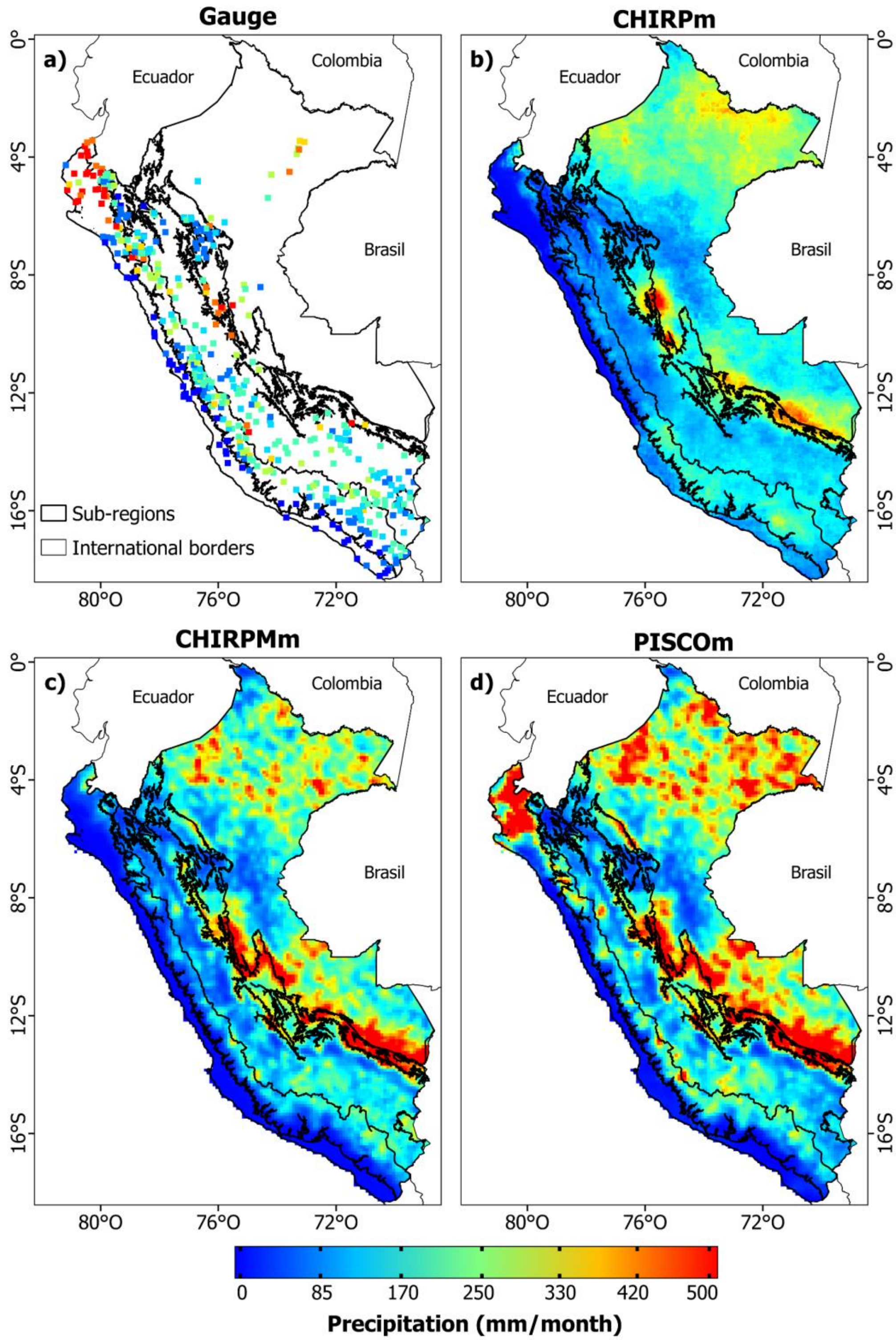

Figure 4. Spatial distribution of rainfall for January 1998. The four maps represent: (a) rainfall network, (b) CHIRPm, (c) CHIRPMm and (d) PISCOpm.

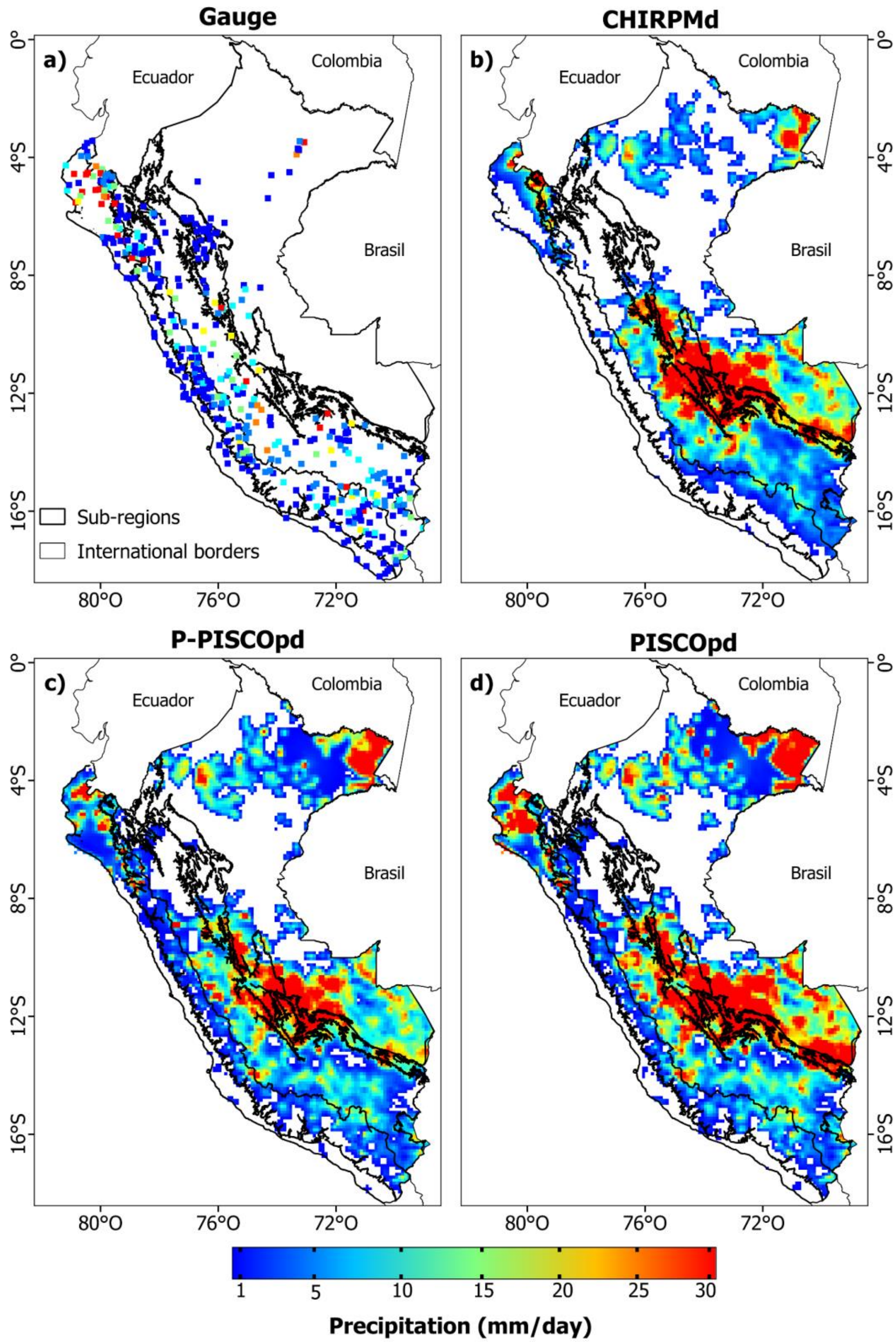

Figure 5. Spatial distribution of rainfall for January 25, 1998. Only values > 1 mm are plotted. The four maps represent: (a) rainfall network, (b) CHIRPMd, (c) P-PISCOpd and (d) PISCOpd.

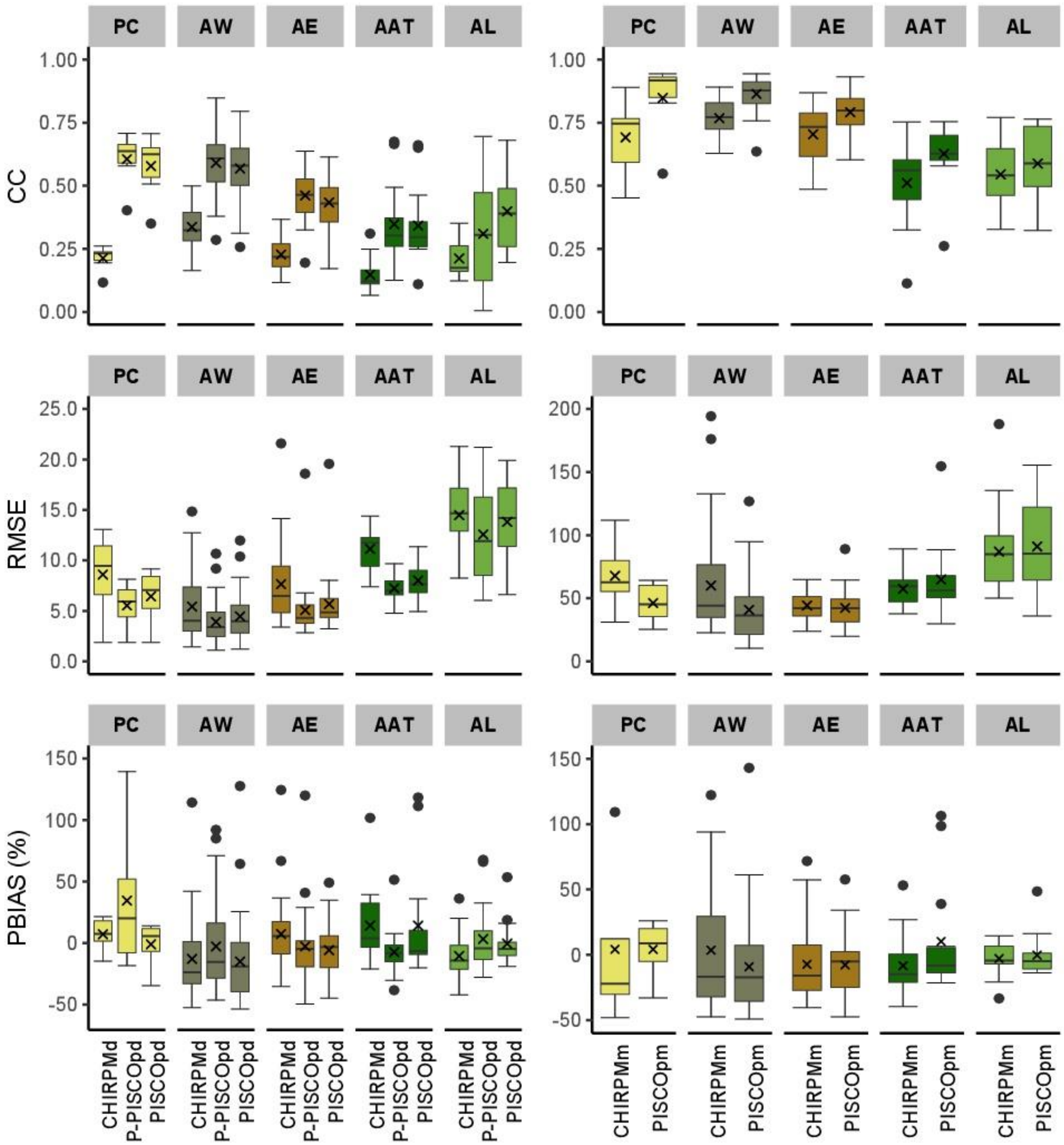

Figure 6. Box plots of continuous statistics (CC, RMSE, and PBIAS) between the daily and monthly products of PISCOp V.2.1 and ID. The cross represents the spatial average.

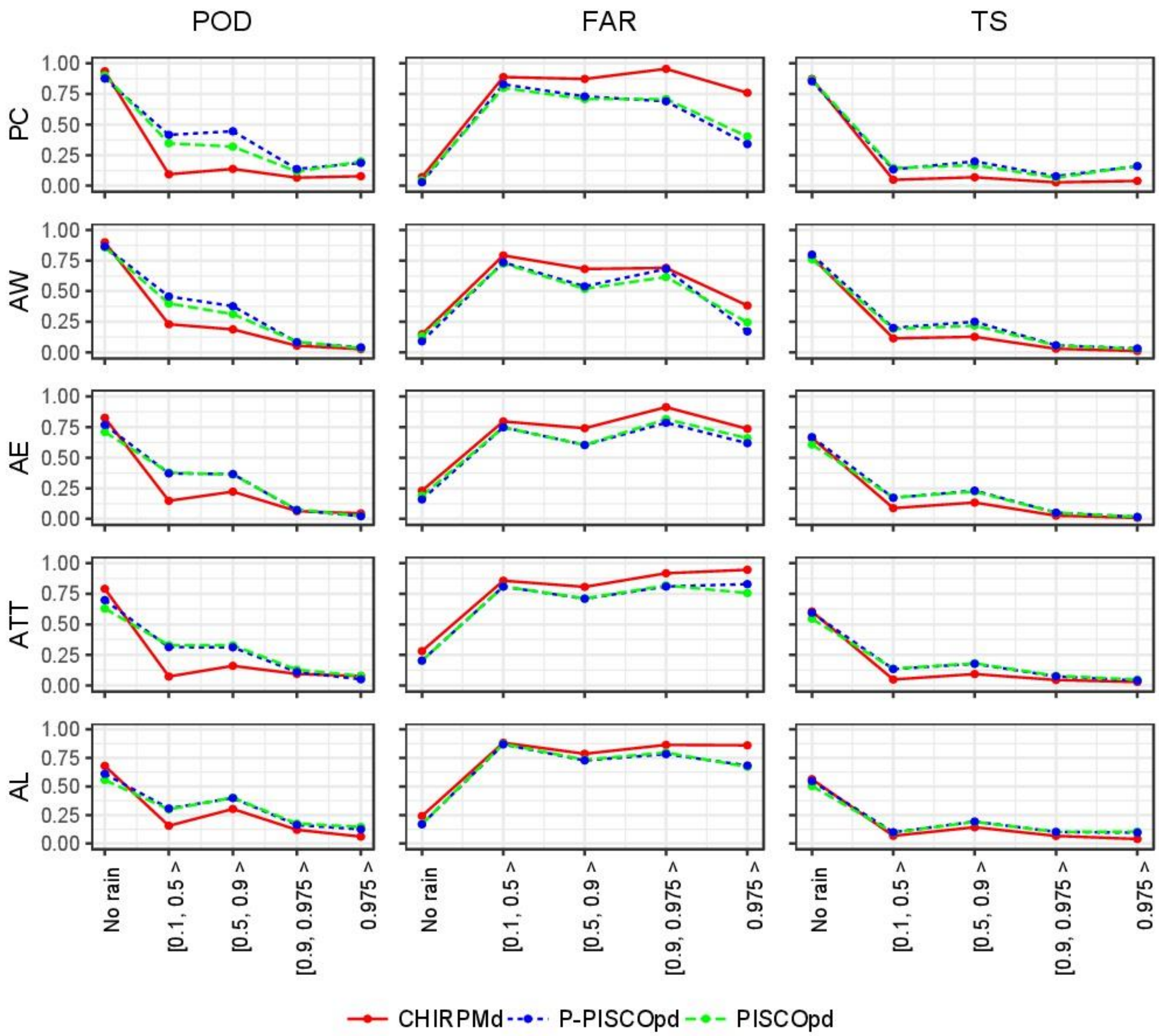

Figure 7. Categorical validation statistics of PISCOpd products in five quantiles classes of rainfall intensity for the five sub-regions (see Table 3).

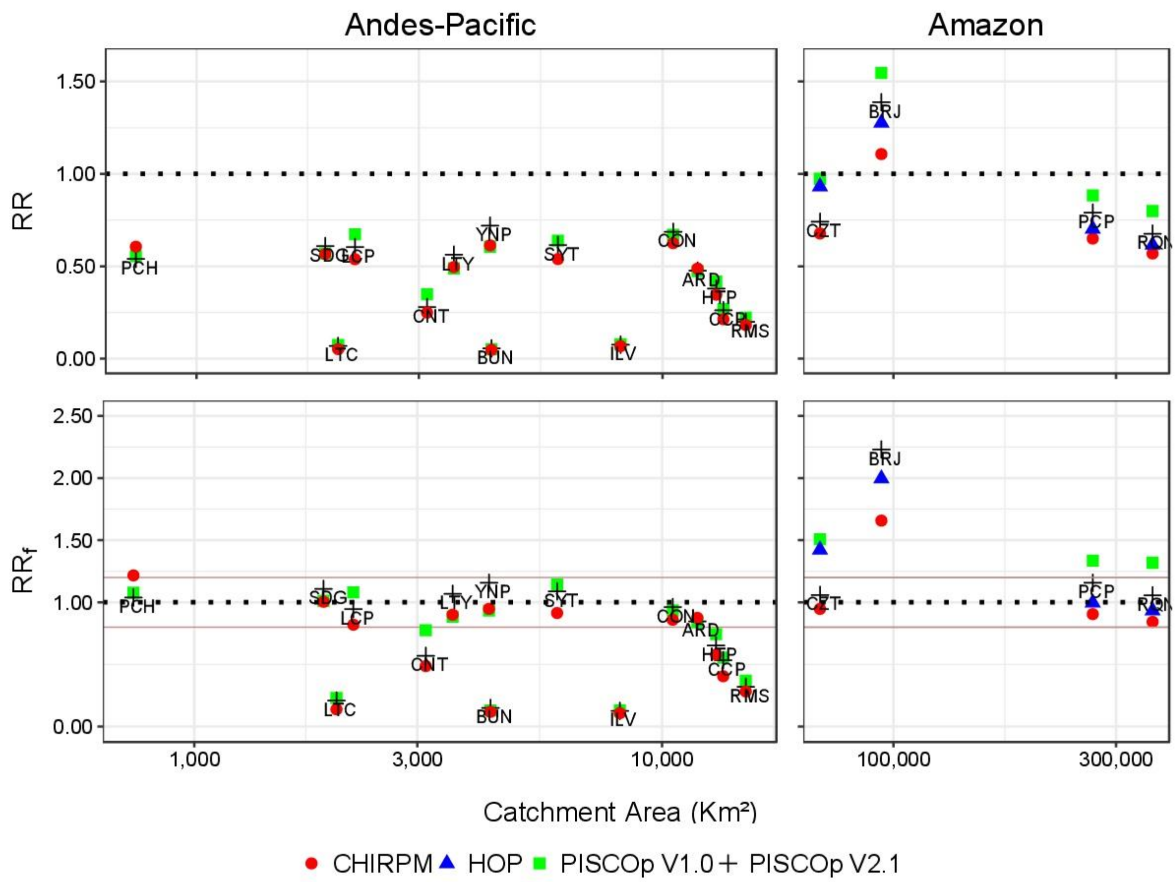

Figure 8. Runoff ratios (RR and RR_f) between different Grd's (PISCOp V2.1, CHIRPM, PISCOp V1.0 and HOP) and discharge observations for catchments draining within the Andes-Pacific (first column) and Amazon (second column).

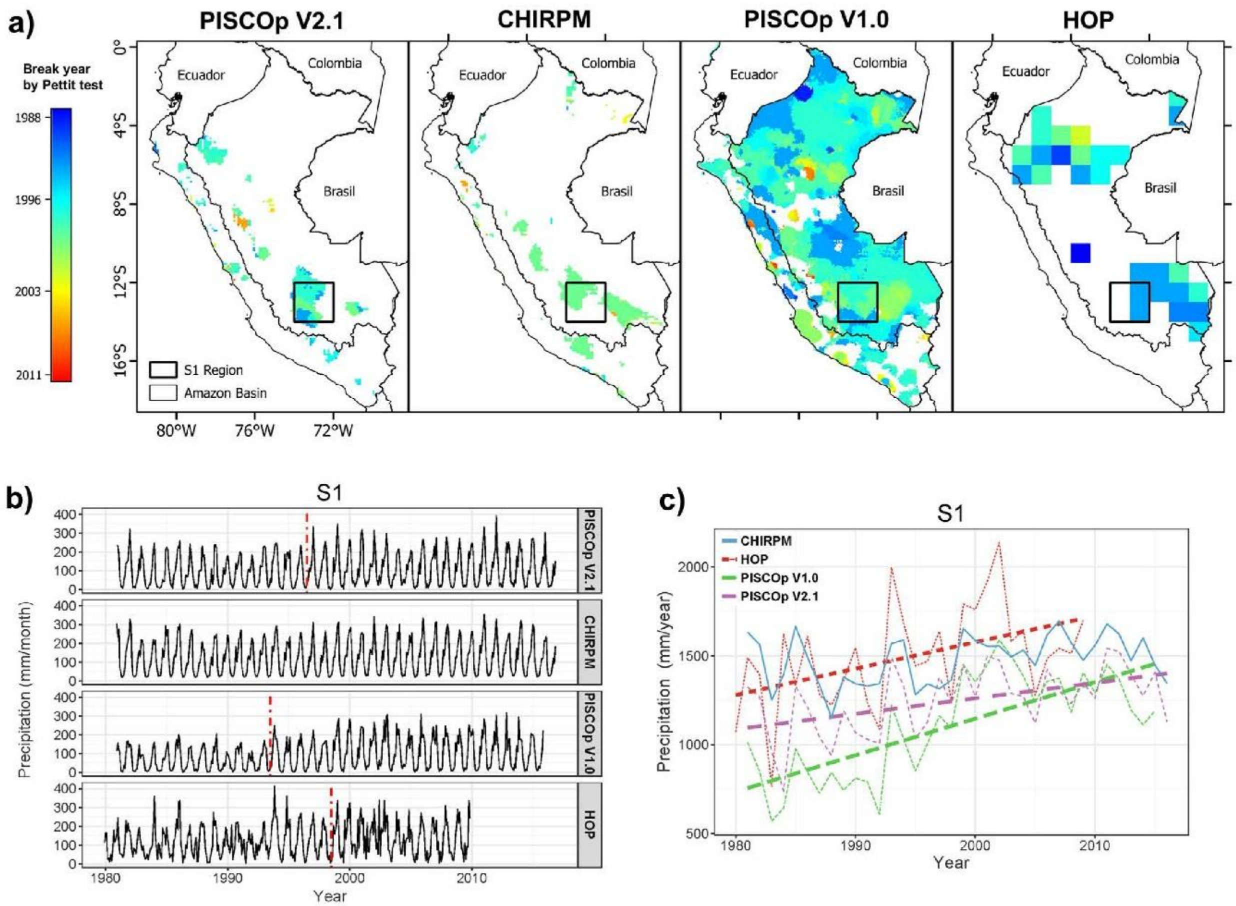

Figure 9. a) Spatial distribution of break year (calculated by Pettit test) in annual time series of PISCOp V2.1, CHIRPM, PISCOp V1.0 and HOP. Only values at 95% significance level ($p < 0.05$) are plotted. b) Areal monthly precipitation for the S1 region, the breaks at 95% significance level are plotted using a vertical red dotted line. c) Evolution of the average annual rainfall in the S1 region, only trend lines with a significant level of 95% were plotted.

Table 1. Sub-regions defined for the analysis of PISCOp V.2.1. Adapted from (Manz *et al.*, 2016), n represents the number of rain gauges within each sub-region.

Sub-region	Elevation (m.a.s.l)	Climate Driver	Rainfall Regime	N
Peruvian Pacific Coast (PC)	0 – 1500	ITCZ, HCS, ENOS	Wet (Dec - May) Dry (Jun - Nov)	97
Andes western slope (AW)	> 1500	Elevation, ITCZ	Wet (Dec - May) Dry (Jun - Nov)	151
Andes eastern slope (AE)	> 1500	Elevation, Orography, ITCZ	Weak seasonality, drier JJA	128
Andes-Amazon transition (AAT)	500 – 1500	Orography, ITCZ, SALLJ	Weak seasonality, drier JJA	26
Amazon lowland (AL)	0 – 500	ITCZ, trade winds	Weak seasonality, drier JJA	39

Table 2. Continuous and categorical statistics. X = Grd estimate, Y = ID measurement, $\bar{X}$ = Grd average, $\bar{Y}$ = ID average, N= number of data pairs, A = number of hits, B = number of false alarms, C = number of misses, and D = number of correct negatives.

	Name	Formula	Perfect Score
Continuous statistics	Correlation Coefficient (CC)	$CC = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2(Y-\bar{Y})^2}}$	1
	Root Mean Squared Error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum(X-Y)^2}$	0
	Percentage Bias (PBIAS)	$\text{PBIAS} = 100 \times (\sum(X-Y)/\sum X)$	0
Categorical statistics	Probability of detection (POD)	$POD = A/(A+C)$	1
	False alarm ratio (FAR)	$FAR = B/(A+B)$	0
	Threat score (TS)	$TS = A/(A+B+C)$	1

Table 3. Classification of rainfall events based on quantiles

Quantile	Daily rain (mm/day)	Daily Rainfall event
[0 - 0.1> *	[0 - 1.5>	No rain
[0.1 - 0.5>	[1.5 - 5.3>	Light rain
[0.5 - 0.9>	[5.3 – 19.5>	Moderate rain
[0.9 - 0.975 >	[19.5 – 38.4>	Heavy rain
0.975 >	38.4>	Violent rain

*** This rainfall class is considered as no rain.**

Table 4. Summary of in-situ discharge gauge characteristics: ETP is the potential evapotranspiration and ET is the real evapotranspiration.

N°	Name	Code	Catchment Area (km²/1000)	Rain gauge density (/ 10^6 km²)	Discharge mean (m³/s)	Precipitation mm/year (PISCOp V2.1)	ETP	ET
1	Borja	BRJ	94.21	371.5	49.43	1730	973	799
2	Chazuta	CZT	69.57	445.6	30.76	2333	1043	936
3	Pucallpa	PCP	267.38	254.3	101.78	2047	996	881
4	Requena	RQN	359.26	203.2	123.34	2236	1351	1116
5	Ardilla	ARD	11.89	1093.2	1.49	916	1051	554
6	Puchaca	PCH	0.74	0	0.06	717	834	448
7	Condorcerro	CON	10.54	664.3	1.39	743	653	380
8	Yanapampa	YNP	4.27	703.3	0.41	470	678	199
9	Santo Domingo	SDG	1.89	1587.2	0.16	463	676	206
10	La Capilla	LCP	2.19	1370.8	0.19	541	695	242
11	S&T Imperial	SYT	5.96	1341.3	0.55	477	658	210
12	Conta	CNT	3.12	960.0	0.11	374	727	268
13	Letrayoc	LTY	3.57	1121.4s	0.26	502	669	304
14	Huatiapa	HTP	13.04	997.2	0.76	513	551	328
15	Chucarapi	CCP	13.51	592.2	0.32	300	604	235
16	La Tranca	LTC	2.01	993.1	0.02	123	665	93
17	Bella Union	BUN	4.30	465.5	0.12	204	698	122
18	Puente Ilave	ILV	8.12	369.2	0.32	304	523	173
19	Puente Ramis	RMS	15.09	596.2	0.72	597	549	422

Table 5. Overview of the state and gap-infilling of PISCOp rainfall network: BCC (%) is the percentage of gaps completed by bias-corrected CHIRPM, D_n is the spatial average of KS statistic and MK_{bef-af} is the number of spurious trends after the gap-infilling procedure.

		Sub-regions (Total of rain gauges: 441)					
		PC	**AW**	**AE**	**AAT**	**AL**	**Total**
	Density (/10^6 km^2)	480	754	381	170	59	282
	Gross error (%)	0.62	7.5	2.19	3.60	1.2	3.51
Monthly	**No data (%)**	34.8	30.94	41.44	45.14	41.72	37.41
	CUTOFF (%)	14.16	11.20	12.49	35.28	40.76	16.40
	BCC (%)	20.64	19.74	28.95	10.05	0.96	21.01
	D_n	0.07	0.05	0.04	0.03	0.02	0.05
Daily	**No data (%)**	32.98	29.39	39.82	43.09	39.40	35.65
	CUTOFF (%)	12.22	13.88	27.04	34.07	33.72	20.56
	BCC (%)	20.76	15.51	12.78	9.02	5.68	15.09
	D_n	0.03	0.03	0.05	0.07	0.03	0.03
	MK_{bef-af}	4	1	7	3	3	18