Correction of TRMM 3B43 monthly precipitation data over the mountainous areas of Peru during the period 1998–2007

Thomas Condom,1,2* Pedro Rau3 and Jhan Carlo Espinoza3,4
1 IRD-France, UMR-HydroSciences, Montpellier, France
2 EGID Institute, GHIYMAC, University of Bordeaux, France
3 Universidad Nacional Agraria La Molina, Lima, Peru
4 CNES-France, UMR-LMTG, Toulouse, France

Abstract:
In an attempt to estimate the spatial and temporal behaviour of rainfall over the mountainous areas of the Peruvian Andes, a new in situ monthly rainfall dataset has been collected (1998–2007) and compared with Tropical Rainfall Measuring Mission (TRMM) 3B43 monthly precipitation data for regions located above 3000 m. The reliability of the TRMM 3B43 data varies depending on the root mean squared error ratio (%RMSE) and correlation coefficient. Because of the discrepancy between the two datasets, the use of additive and multiplicative correction models is proposed for the TRMM 3B43 data. In the Peruvian mountain ranges, these correction models better approximate TRMM rainfall monthly values, as already verified for annual values. Copyright © 2010 John Wiley & Sons, Ltd.

KEY WORDS rainfall; Andes; TRMM 3B43; Peru

Received 15 June 2010; Accepted 29 October 2010

INTRODUCTION
In terms of climate, mountain systems develop considerably complex local and regional systems. Beniston et al. (1997) described an uncertainty reflecting the varying temperature and rainfall records, associated with the complexity of the weather and terrain. Garreaud et al. (2003) published an article on the climate in the Altiplano, showing the current conditions and mechanisms of past changes. According to these authors, the climatic conditions on all timescales are closely related to the upper-air circulation with an easterly zonal flow aloft favouring wet conditions and a westerly flow causing dry conditions. Therefore, the annual cycle of dry winter and wet summer conditions, which is reflected in the precipitation regime, is caused by the seasonal expansion of the equatorial easterlies in the upper troposphere. Interannual variability is primarily controlled by changes in the mean zonal flow over the Altiplano, reflecting changes in the meridional baroclinicity between tropical and subtropical latitudes, which in turn is a response to sea-surface temperature (SST) changes in the tropical Pacific. More recently, Espinoza et al. (2009a) described the strong spatial variability of rainfall in the Amazon-Andes region.

Currently, due to a lack of observational data at an adequate spatial and temporal resolution, there is a limited understanding of the multiple interconnections between climate and hydrology in mountain systems around the world such as the Himalayas, the Rockies, the Andes and the Alps (Beniston et al., 1997). In situ rainfall data are scarce in high mountainous regions and it is necessary to develop a method to spatialize rainfall in these zones. Satellite data can be used for this purpose. In this article, two types of monthly rainfall data are compared: (i) in situ data from the ‘Servicio Nacional de Meteorología e Hidrología de Perú’ (SENAMHI) and (ii) data issued from the TRMM (Tropical Rainfall Measuring Mission) 3B43 algorithm, which is the single best estimate of monthly precipitation obtained as it combines 3-hourly integrated high-quality data, infrared (IR) estimates (3B42) with the monthly accumulated Climate Assessment Monitoring System or Global Precipitation Climatology Center rain gauge analyses (3S45) (Huffman et al., 1995, 1997, 2007). These second type of data are partly obtained from a satellite project operating under the TRMM and will be referred to as the ‘TRMM 3B43 data’ in the remaining text. Several studies are dedicated to the use of the TRMM products for hydrological purposes. These studies focused on the lowlands in Africa with the Kafue River Basin in Zambia, the Okavango River Basin in Angola, both the Thukela and Kat River Basins in South Africa (Hughes, 2006; Wilk et al., 2006), in Ghana (Endreny and Imbeah, 2009) and the Tapajos River Basin in Brazil, which is a tributary of the Amazon River in its lower part (Colischonn et al., 2008). The TRMM data collected at daily to monthly intervals are often underestimated for mountainous regions in the tropical and middle latitude mountain systems (Berg et al., 2006;
Huffman *et al.*, 2007). This is particularly true for the Himalayan range (Barros *et al.*, 2004; Anders *et al.*, 2006). In this article, we analyse the correlation between two monthly precipitation time series in high-altitude Peruvian regions; the TRMM 3B43 data and in situ rain gauge data. With respect to the monthly 3B43 TRMM data, no adequate algorithms are available to correct the bias in the mountainous region. In this study, we present a robust bias correction algorithm to solve this problem for the Peruvian Andes (>3000 m asl) and provide a correction of the TRMM 3B43 data that allow a better estimation of the rainfall data in the mountains covered with glaciers. In Peru, the glaciers occur between 6700 and 4800 m asl. Furthermore, in order to compare the in situ data with the TRMM grid cells (see section on Data and Methodology), we need to include at least two stations per zone, therefore we chose to limit the stations included in this study to those at or above an altitude of 3000 m asl. This article is organized as follows: the main characteristics of the rainfall variability in the Andes mountains are given in Section on Main Characteristics of the Rainfall Variability in the Tropical Andes; the data and methodology used in this study are described in Section on Data and Methodology and finally, we provide the results in Section on Results.

**MAIN CHARACTERISTICS OF THE RAINFALL VARIABILITY IN THE TROPICAL ANDES**

The elevation of the Peruvian Andes Mountains is above 5 km and represents a climatic barrier for the tropospheric flow, allowing drier conditions to the west compared to the moist Amazon region to the east. Water vapour in this region is supplied from the Atlantic Ocean and Amazon Rainforest by means of evapotranspiration (e.g. Salati and Vose, 1984; Nogues-peagle and Mo, 1997; Zhou and Lau, 1998; Vera *et al.*, 2006). At low elevations, abundant rainfall is related to moist warm air over the eastern slope of the Andes (Espinoza *et al.*, 2009a). The rainfall stations registering more than 3000 mm/year are located lower than 1500 m asl. These authors also showed strong spatial variability at altitudes lower than 2000 m asl, where rainfall varies from 500 to 3000 mm/year. The highest rainfall values were recorded at stations that were located in positions characterized by strong air uplift, such as the southeastern Peruvian Amazon, when the San Gabán and Quincemil stations (around 800 m asl) have an average of 6000 mm rain per year. They are located in a concavity in the Carabaya Mountain Range, close to steep slopes. On the west side of the Andes Cordillera, there is no clear relationship between annual in situ rainfall and altitude; values that are close to zero or very low are often observed near sea level. Above the inversion layer, rainfall is generally more abundant (Johnson, 1976; Lavado, 2010); however, very low amounts of rainfall can also be recorded, probably due to the lee wind exposure of the stations.

The southern tropical regime characterizes the seasonal rainfall variability in the southern part of the Peruvian Andes, where rainfall is concentrated during the austral summer (around 75% of annual rainfall is recorded from December to March). In the high mountains and on the west side of the Andes, dry conditions are noticed during the austral winter, whereas a less marked dry season characterizes the lowland on the east side of the Andes. Towards the equatorial Andes, two rainfall regimes are seen: an equatorial regime, which has a better rainfall repartition throughout the year, and a north equatorial regime, where rainfall is largely concentrated during the austral winter (June, July and August). The spatial variability of the rainfall regimes may be even greater; only a few stations are exposed to the different regimes or easterlies. For a detailed description of the seasonal rainfall regimes in the Andes region, see Espinoza *et al.*, 2009a), Laraque *et al.*, 2007), Garraud et al. (2003) and Ronchail and Gallaire (2006).

In the southern part of the Andes, including the Altiplano region, many studies have shown that the interannual rainfall variability is greatly controlled by the El Niño Southern Oscillation (ENSO) (e.g. Aceituno, 1988; Ronchail, 1995; Lenters and Cook, 1999; Vuille, 1999; Garreaud and Aceituno, 2001). All these studies concluded that El Niño years tend to be dry, whereas La Niña years (ENSO cold phase) are often associated with wet conditions. In addition, an inverse and weak signal has been observed in the Amazon plain of Bolivia (Ronchail, 1998; Ronchail *et al.*, 2002, 2005; Ronchail and Gallaire, 2006). However, in the Peruvian Andes, no clear signal is found (Tapley and Waylen, 1990; Rome and Ronchail, 1998; Espinoza *et al.*, 2009a). The rainfall anomaly is not as pronounced in Ecuador, which is characterized by a slight rainfall increase during El Niño events (i.e. Rossel *et al.*, 1999; Ronchail *et al.*, 2002; Bendix *et al.*, 2003) In conclusion, these studies indicate that the relationship between SST in the tropical Pacific and precipitation anomalies in the central Andes is not a simple one.

More recently, Espinoza *et al.* (2009a) and Lavado (2010) documented that rainfall in the central and southern Andes is also related to the tropical North Atlantic SST. Rainfall is less than normal when positive anomalies are noticed in the tropical North Atlantic SST. Moreover, the higher anomalies noticed at the beginning of the 1990s in this oceanic region can be explained by the negative trend noticed in the rainfall patterns in the Peruvian Andes (Espinoza *et al.*, 2009a; Lavado, 2010), which was also detected in the discharge flow of the southern Andean rivers of the Amazon Basin (Espinoza *et al.*, 2006, 2009b).

**DATA AND METHODOLOGY**

**Digital elevation model**

Using SRTM data (Shuttle Radar Topography Mission, NASA-NGA, USA), the resolution of the model was a 90-m grid (http://srtm.csi.cgiar.org/SELECTION/inputCoord.asp). The study zone lay between latitudes...
CORRECTION OF TRMM 3B43 MONTHLY PRECIPITATION DATA

5–20°S and longitudes 80–70°W. To focus on the highlands, a topographical limit of 3000 m was used. The SRTM data were used to delineate the major river basins or watersheds representing the different regions. In other words, the SRTM data also served as a reference for verifying the location of the meteorological stations as well as for the regionalization (see Section on Rainfall Analysis Methods).

In situ rainfall data

The in situ rainfall data for the 29 rainfall stations in the different Peruvian basins were classified into seven regions: Santa (R1), North Pacific (R2), Colca (R3), Quilca (R4), Apurimac (R5), Urubamba (R6) and Ocoña (R7). The distribution of the stations, regions and SRTM digital elevation model are shown in Figure 1.

All the stations were located above 3000 m, except for two: Abancay (2750 m asl) and Yungay (2537 m asl). The in situ measurements fit into the sampling error category, ranging from errors in the handling of equipment by field staff to weather impacts on the measurements (Franchito et al., 2009). The data have passed rigorous quality assessments and quality control processes by the SENAMHI. The SENAMHI verified the level of consistency between the daily and monthly values given by the operators. Furthermore, all extreme values have been checked and validated by comparing data from the stations situated in the same region. It is well known that snow and/or wind causes errors in rain gauge data (Legates and Willmott, 1990). As far as snowfall is concerned, the occurrence of snow is very rare in this tropical country at altitudes situated below 4600 m asl because the rainy season occurs during the

Figure 1. Altitude and location map of the meteorological stations (in situ data, SENAMHI). Location of the seven regions: Santa (R1), North Pacific (R2), Colca (R3), Quilca (R4), Apurimac (R5), Urubamba (R6) and Ocoña (R7). The heights are taken from the digital elevation model using SRTM terrain data with 90-m resolution.
summertime when the temperature is relatively high. In fact, the monthly temperatures are always higher than 2 °C/month for all the \textit{in situ} meteorological stations used in this study. There were no missing data for the monthly time series rainfall data for any of the 29 stations. Finally, considering the quasi-absence of snowfall at the meteorological stations and the regionalization process (see Section on Rainfall Analysis Methods), we believe that the errors for the \textit{in situ} values are low.

**TRMM 3B43 data**

The TRMM 3B43 data used are freely available from the NASA database (http://trmm.gsfc.nasa.gov/data_dir/data.html). The 3B43 algorithm provides the best estimate of the total monthly rainfall in a complete record from January 1998 to December 2007. This dataset is the result of the combination of precipitation datasets (Microwave Imager TMI, Precipitation Radar PR, Visible and Infrared Scanner VIRS with the Special Sensor Microwave Imager SSM/I and rain gauge data). Spatially, each record is a $0.25 \times 0.25^\circ$ grid ($\sim$770 km$^2$). These data are compressed 3B43 HDF formats, and a free source (Collischonn et al., 2008) was used to process these files in matrix form to present each TRMM rainfall grid in a time series. The common types of errors for TRMM measurements are due to the discrete sampling characteristics. Indeed, the TRMM satellite is in a low Earth orbit $\sim$350 km high; the rain sensors sample the regional atmosphere only at discrete time intervals, sometimes missing short duration storms. Many theoretical studies have shown that the temporal sampling error for the average rainfall ranges from $\pm 8$ to $\pm 12\%$ per month (Shin and North, 1988; North and Nakamoto, 1989; Bell et al., 1990).

**Rainfall analysis methods**

The use of TRMM satellite data and its comparison with the \textit{in situ} data was initially based on the methodology developed in Franchito et al. (2009). In this section, we present the methodology used for the regionalization and the different quality criteria.

**Regionalization.** The digital elevation model SRTM was used to delineate the watersheds, which were classified into main drainage areas, and to obtain an acceptable correlation between the total average annual rainfall and altitude, longitude and latitude. The stations were grouped into seven representative watersheds: R1 Santa, R2 North Pacific, R3 Colca, R4 Quilca, R5 Apurimac, R6 Urubamba and R7 Ocoña. Table I and Figure 1 show these seven regions (or watersheds) and the availability of \textit{in situ} and TRMM 3B43 data for the 29 meteorological stations. Then for the seven regions, we calculated the mean monthly values for the \textit{in situ} (punctual points) and TRMM data (cell values). For example, for region R1 Santa, the mean monthly values were based on the time

### Table I. \textit{In situ} rain gauge data (SENAMHI) and their correlation with the TRMM 3B43 grid cells

<table>
<thead>
<tr>
<th>Region</th>
<th>Conventional meteorological station</th>
<th>Grid TRMM</th>
<th>Common period</th>
<th>\textit{In situ} mean annual rainfall (mm/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1: Santa</td>
<td>YUNGAY (2537 m asl)</td>
<td>7787912/7762912</td>
<td>1998–2007</td>
<td>731</td>
</tr>
<tr>
<td></td>
<td>RECUAY (3394 m asl)</td>
<td>7737962</td>
<td>1998–2007</td>
<td>909</td>
</tr>
<tr>
<td></td>
<td>MILPO (4400 m asl)</td>
<td>7712987</td>
<td>1998–2006</td>
<td>1185</td>
</tr>
<tr>
<td>R2: North Pacific</td>
<td>PIRA (3570 m asl)</td>
<td>7762962</td>
<td>1998–2006</td>
<td>701</td>
</tr>
<tr>
<td></td>
<td>MALVAS (3500 m asl)</td>
<td>7762987</td>
<td>1998–2007</td>
<td>513</td>
</tr>
<tr>
<td>R3: Colca</td>
<td>CHIQUIAN (3350 m asl)</td>
<td>77121012</td>
<td>1998–2007</td>
<td>745</td>
</tr>
<tr>
<td></td>
<td>ORCOPAMPA (3779 m asl)</td>
<td>72371537</td>
<td>1998–2007</td>
<td>457</td>
</tr>
<tr>
<td></td>
<td>ANDAHUA (3587 m asl)</td>
<td>72371537</td>
<td>1998–2007</td>
<td>373</td>
</tr>
<tr>
<td></td>
<td>PAMPACOLCA (2950 m asl)</td>
<td>77621562</td>
<td>1998–2007</td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>CHIVAY (3633 m asl)</td>
<td>71621562</td>
<td>1998–2007</td>
<td>475</td>
</tr>
<tr>
<td></td>
<td>SIBAYO (3810 m asl)</td>
<td>71371537</td>
<td>1998–2007</td>
<td>674</td>
</tr>
<tr>
<td>R4: Quilca</td>
<td>TISCO (4175 m asl)</td>
<td>71371537</td>
<td>1998–2007</td>
<td>797</td>
</tr>
<tr>
<td></td>
<td>PORPERA (4195 m asl)</td>
<td>71371537</td>
<td>1998–2007</td>
<td>881</td>
</tr>
<tr>
<td>R5: Apurimac</td>
<td>LA ANGOSTURA (4150 m asl)</td>
<td>71621612</td>
<td>1998–2007</td>
<td>376</td>
</tr>
<tr>
<td></td>
<td>YAURI (3927 m asl)</td>
<td>71371487</td>
<td>1998–2007</td>
<td>856</td>
</tr>
<tr>
<td>R6: Urubamba</td>
<td>ABANCAY (2750 m asl)</td>
<td>72871362</td>
<td>1998–2007</td>
<td>840</td>
</tr>
<tr>
<td></td>
<td>SICUANI (3574 m asl)</td>
<td>71214377121412</td>
<td>1998–2007</td>
<td>905</td>
</tr>
<tr>
<td></td>
<td>POMACANCHI (3700 m asl)</td>
<td>71621412</td>
<td>1998–2007</td>
<td>754</td>
</tr>
<tr>
<td></td>
<td>POMACANCHI (3700 m asl)</td>
<td>71621362</td>
<td>1998–2007</td>
<td>673</td>
</tr>
<tr>
<td></td>
<td>COLQUEPATA (3729 m asl)</td>
<td>71621237</td>
<td>1998–2007</td>
<td>621</td>
</tr>
<tr>
<td>R7: Ocoña</td>
<td>NUCAYLLAPA (4100 m asl)</td>
<td>72621487</td>
<td>1998–2007</td>
<td>783</td>
</tr>
<tr>
<td></td>
<td>PULHUAY (4600 m asl)</td>
<td>72371512</td>
<td>1998–2007</td>
<td>682</td>
</tr>
</tbody>
</table>

Copyright © 2010 John Wiley & Sons, Ltd.
CORRECTION OF TRMM 3B43 MONTHLY PRECIPITATION DATA

series comprising 108 values for the 1998–2007 period for the three meteorological stations (Yungay, Recuay and Milpo). The mean monthly values for the TRMM data were calculated based on the 108 values for the four cell numbers: 778912, 7762912, 7737962 and 7712987. As there are only a few stations for each region, we did not use a spatial interpolation. We compared the two data sets, in situ and TRMM, and subsequently propose two correction models (the additive and multiplicative models). For each region, Equation (1) was used to compare the two sources of data.

\[ \Delta P_{i,j} = \mu_{\text{TRMM},i,j} - \mu_{\text{SENAMHI},i,j} \]  

where \( i \) is the month number (1, \ldots, 12), \( j \) the year number (1998, \ldots, 2007), \( \mu_{\text{TRMM}} \), the TRMM 3B43 record for the month \( i \) and the year \( j \) and \( \mu_{\text{SENAMHI}} \), is the SENAMHI record (in situ) for the month \( i \) and the year \( j \).

Quality criteria to compare the in situ and TRMM 3B43 data. A frequency histogram analysis of the full set of total monthly in situ and TRMM 3B43 data for each month and region was performed (not shown here). The probabilistic behaviour was compared to a normal distribution to determine the potential similarities between the points and matching cells and to facilitate the use of statistical tests. The reliability of the TRMM 3B43 estimates for each in situ station in each of the seven regions was validated using the average error ratio (MBE, Equation (2)) mean absolute relative error (AE) and RMSE (Equation (3))

\[ \text{MBE} = \frac{\sum (P_{\text{TRMM},i,j} - P_{\text{SENAMHI},i,j})}{N} \]  

where \( P_{\text{TRMM},i} \) is the TRMM rainfall estimated in a month \( i \) and a year \( j \) (mm/month), \( P_{\text{SENAMHI},i} \), the rainfall measured by SENAMHI in a month \( i \) and a year \( j \) (mm/month), \( i \) the month number (1,\ldots,12), \( j \) the year number (1998,\ldots,2007) and \( N \) the number of pairs compared and

\[ \text{RMSE} = \sqrt{\frac{\sum (P_{\text{TRMM},i,j} - P_{\text{SENAMHI},i,j})^2}{N}} \]  

The AE is the average of the absolute differences between the estimated and measured rainfall. The above errors represent the best estimation parameters. They are more accurate compared with other classic estimation parameters, i.e. the mean square error, which is too sensitive to estimate large errors for rainfall (Franchito et al., 2009). These errors are absolute and maintain the unit of measurement in millimetres. Absolute errors do not clearly indicate which value or range of values can be considered reliable when considering regions with different ranges of rainfall values.

Then, the absolute errors were divided by the average in situ rainfall record in each region and for each month in order to obtain the following relative errors, which were expressed in percentages: mean relative error (\( \%\text{MBE} \)), relative absolute error (\( \%\text{AE} \)) and root mean squared relative error (\( \%\text{RMSE} \)). In this analysis, \( \%\text{RMSE} \) and \( \%\text{MBE} \) were used to determine the systematic and random components of TRMM data error (Adeyewa and Nakamura, 2003). The \( \%\text{RMSE} \) evaluated the reliability of each data source in different seasons and was reliable when the RMSE of the estimated rainfall was <50% of the measured rainfall amount. Conversely, when the RMSE was equal to or greater than 50% of the magnitude of the reference rainfall, the estimate was considered unreliable for the region considered (Franchito et al., 2009).

A quantitative relationship was established between these two sources of information based on the correlation coefficient (CC) estimated from a linear regression with the intercept on the axis of the coordinates. The correlations were significant (\( p < 0.01 \)) when the CC was greater than or equal to 0.7.

Another way to validate these corrections was to use an annual-level analysis. This was based on a comparison of the total average annual rainfall in each region for the 1998–2007 period. The aim was to determine a range showing by how much the original TRMM 3B43 data, corrected with the additive and multiplicative models, underestimate and overestimate the total annual in situ rainfall.

The in situ monthly data for the 29 rainfall stations data were visually checked and showed that the probability distribution has a strong negative asymmetry due to a wide range of rainfall levels and a minimum value of zero in the dry periods. This asymmetry decreased when the data are log transformed, plus one unit, using Equation (4):

\[ \text{Rainfall}_{\text{transformed}} = \log (\text{Rainfall} + 1) \]  

Jones and Hulme (1996) transformed these data as a percentage of the mean and standard deviation. Diaz et al. (1989) and Hutchinson (1995) validated the transformation using other distributions including logarithms. The TRMM 3B43 records mostly showed the same negative asymmetrical behaviour, thus requiring transformation. Finally, Equation 4 was applied to the additive and multiplicative models.

Correction models for the TRMM 3B43 data

Additive model. It is a standard procedure to consider that the in situ rainfall data are correct (Adeyewa and Nakamura, 2000; Vila et al., 2009). It was possible to calculate the mean monthly difference between the TRMM 3B43 and in situ data for the 1997–2008 period using Equation (5). This vector \( \mathbf{F}_1 \) was composed of 12 monthly values

\[ \sum_{i=1}^{j} (\mu_{\log(\text{TRMM}_{i,j}+1}) - \mu_{\log(\text{SENAMHI}_{i,j}+1)}) = F_{1i} \]  

where \( j \) is the year number, \( i \) the month number (1, \ldots, 12), \( \mu_{\log(\text{SENAHMI}_{i,j}+1)} \) the SENAMHI record transformed in a month \( i \) and a year \( j \) and \( \mu_{\log(\text{TRMM}_{i,j}+1)} \) the TRMM record transformed in a month \( i \) and a year \( j \).

Vector \( F_1 \) was applied to each TRMM 3B43 value. Finally, we obtained the total monthly TRMM corrected rainfall values for each month over the 10-year period of records using the following equation (Equation (6))

\[
\text{TRMM}_{i,j} = \sqrt[12]{\text{TRMM}_{i,j} + 1} - 1
\]

Equations 6 and 8 are the proposed models for the correction of TRMM 3B43 data for the Peruvian Andes.

**RESULTS**

**In situ rainfall**

The mean annual values for the 1998–2007 period were between 285 mm year\(^{-1}\) at the Pampacolca station and 1185 mm year\(^{-1}\) at the Milpo station (Table I). Considering the behaviour of in situ rainfall on a monthly scale, a seasonal pattern showed the average monthly rainfall as ranging from 0 to \( \sim 200 \) mm/month (Figure 2). The zero values occurred in the dry months of June, July and August at the stations in the North Pacific area in region 2 (i.e. the Pira, Malvas and Chiquian stations) and part of the Quilca and Colca river basins, which are closer to the Pacific Ocean. The maximum values occurred in January (i.e. the middle of the wet season) in the Apurimac and Urubamba basin areas near the Amazon Basin (i.e. a forest area). All the regions were characterized by rainfall that was largely concentrated during the austral winter (June, July and August). These results are in agreement with previous studies described in Section on Main Characteristics of the Rainfall Variability in the Tropical Andes.

**In situ and TRMM 3B43 data**

Table II shows the reliability of the original TRMM 3B43 data (without correction) in each region (see lines S/C). This reliability is defined by a \( \% \text{RMSE} < 50\% \), a CC >0.7, the \( \% \text{MBE} \) and the \( \% \text{AE} \). Figure 3 shows the monthly variations in the differences between the TRMM 3B43 and in situ data. This figure shows that during the wet period, the in situ data were underestimated for the majority of the regions, and conversely, a positive value for the difference between the TRMM 3B43 and in situ data was seen in the dry months. The \( \% \text{AE} \) presented values below 50% in the wet months (November to April). Below is a summary detailing for which months the original TRMM 3B43 data were reliable per region: R1 region (Santa), November–April; R2 region (North Pacific), October–April, with the best results in November; R3 (Colca) and R4 (Quilca) regions, none; R5 region (Apurimac), August–April; R6 region (Urubamba), January–February and R7 region (Ocoña), November—April. The TRMM 3B43 data were considered accurate in the wet months in contrast to
### Table II. Validity summary of the monthly TRMM 3B43 data

<table>
<thead>
<tr>
<th>Region</th>
<th>Model</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1: Santa</td>
<td>S/C</td>
<td>0.55</td>
<td>0.57</td>
<td>0.4</td>
<td>0.41</td>
<td>X</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>1.02</td>
<td>1.1</td>
<td>0.8</td>
<td>0.65</td>
<td>X</td>
<td>0.87</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.92</td>
<td>0.99</td>
<td>0.73</td>
<td>0.63</td>
<td>X</td>
<td>0.51</td>
</tr>
<tr>
<td>R2: North Pacific</td>
<td>S/C</td>
<td>0.51</td>
<td>0.42</td>
<td>0.47</td>
<td>0.47</td>
<td>X</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.98</td>
<td>0.82</td>
<td>0.75</td>
<td>1.24</td>
<td>X</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.83</td>
<td>0.7</td>
<td>0.85</td>
<td>0.7</td>
<td>X</td>
<td>0.81</td>
</tr>
<tr>
<td>R3: Colca</td>
<td>S/C</td>
<td>0.49</td>
<td>0.47</td>
<td>0.34</td>
<td>0.29</td>
<td>X</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.89</td>
<td>0.88</td>
<td>0.54</td>
<td>0.23</td>
<td>X</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.8</td>
<td>0.78</td>
<td>0.85</td>
<td>0.47</td>
<td>X</td>
<td>0.19</td>
</tr>
<tr>
<td>R4: Quilca</td>
<td>S/C</td>
<td>0.4</td>
<td>0.36</td>
<td>0.38</td>
<td>0.24</td>
<td>X</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.76</td>
<td>0.69</td>
<td>0.79</td>
<td>0.42</td>
<td>X</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.64</td>
<td>0.58</td>
<td>0.63</td>
<td>0.33</td>
<td>X</td>
<td>0.3</td>
</tr>
<tr>
<td>R5: Apurimac</td>
<td>S/C</td>
<td>0.59</td>
<td>0.55</td>
<td>0.59</td>
<td>0.47</td>
<td>X</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>1.07</td>
<td>0.98</td>
<td>1.09</td>
<td>0.62</td>
<td>X</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>1.01</td>
<td>0.91</td>
<td>0.98</td>
<td>0.56</td>
<td>X</td>
<td>0.29</td>
</tr>
<tr>
<td>R6: Urubamba</td>
<td>S/C</td>
<td>1.1</td>
<td>1.09</td>
<td>0.97</td>
<td>1.31</td>
<td>X</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>1.98</td>
<td>1.77</td>
<td>2.23</td>
<td>1.54</td>
<td>X</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>1.98</td>
<td>1.71</td>
<td>2.2</td>
<td>1.55</td>
<td>X</td>
<td>0.6</td>
</tr>
<tr>
<td>R7: Ocoña</td>
<td>S/C</td>
<td>0.48</td>
<td>0.45</td>
<td>0.38</td>
<td>0.37</td>
<td>X</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>0.86</td>
<td>0.83</td>
<td>0.58</td>
<td>0.47</td>
<td>X</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>0.8</td>
<td>0.76</td>
<td>0.68</td>
<td>0.52</td>
<td>X</td>
<td>0.37</td>
</tr>
</tbody>
</table>

**S/C =** original 3B43-TRMM data without correction; **A =** TRMM data corrected by the additive model; **M =** TRMM data corrected by the multiplicative model; **p =** reliable TRMM data; **X =** unreliable TRMM data. The numeric value indicates the slope or ratio of the average in situ rainfall, obtained using a linear regression without constant term with the original TRMM 3B43 and TRMM corrected data using the additive and multiplicative models. Values in bold indicate that their reliability has been verified by all the parameters used (%RMSE, %CC and slope). The slope must be greater than 0.7 and must not exceed 1.3; a slope close to unity is considered as the best approach.

the dry period, where the TRMM 3B43 estimates were not guaranteed for any of the parameters evaluated. Two statistical models, additive and multiplicative, were scored for the correction of the TRMM 3B43 data.

**Correction of the TRMM 3B43 data**

**Additive and multiplicative models.** These models were validated using the %RMSE, %CC and slope from a linear regression without constant term. The correction model is considered more accurate when the slope value is close to 1.

Table II summarizes the applicability of the correction models to the TRMM 3B43 data from all the regions. A positive sign (+) indicates an increase in the %RMSE using either the additive or multiplicative model, with respect to the %RMSE of the original TRMM 3B43 data. A decrease in the %RMSE using either model, with respect to the %RMSE of the original TRMM 3B43 data, is represented with a negative sign (−).

TRMM 3B43 data corrected with the additive model (TRMM A) were very close to the in situ data in

Figure 3. Box plot of the monthly variations of the differences between the TRMM 3B43 and in situ data for each region (R1 to R7) according to Equation (1). The x-axis shows the months from 1 (January) to 12 (December). The box plot summarizes five numbers: the sample minimum, the sample maximum, the lower quartile, the median, the upper quartile, the sample maximum and the outliers mark with asterisks. In general, the TRMM 3B43 data underestimate the in situ data during the wet period.

Table III. Variation expressed in the percentage (%) of TRMM 3B43 data in terms of the total annual average in situ rainfall (SENAMHI) in each region. In bold we indicate the better agreement between the two data set.

<table>
<thead>
<tr>
<th></th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
<th>R4</th>
<th>R5</th>
<th>R6</th>
<th>R7</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRMM 3B43</td>
<td>76</td>
<td>65</td>
<td>60</td>
<td>49</td>
<td>58</td>
<td>122</td>
<td>60</td>
</tr>
<tr>
<td>TRMM A</td>
<td>106</td>
<td>99</td>
<td>98</td>
<td>82</td>
<td>93</td>
<td>188</td>
<td>97</td>
</tr>
<tr>
<td>TRMM M</td>
<td>91</td>
<td>85</td>
<td>86</td>
<td>69</td>
<td>84</td>
<td>177</td>
<td>87</td>
</tr>
</tbody>
</table>

At an annual time scale, corrections from either model provided an accurate approximation with regards to the SENAMHI data for all regions except R6 (Figure 4). This behaviour was due to wet weather and a very strong influence from the lowland Amazon Rainforest.

CONCLUSIONS

The monthly total rainfall at the 29 SENAMHI stations located above 3000 m was characterized by strong seasonality with elevated values recorded during the wet period (January to March) and the lowest values occurring during the dry season (June to August). The main problem was that the poor distribution of meteorological stations did not allow a spatialization of the precipitation in mountainous regions. The TRMM 3B43 product based partly on satellite data would help to do this. Nevertheless, this study demonstrated that before these data can be used, they must be corrected. Therefore, we propose two

Figure 4. Total annual average rainfall (1997–2008 period) for the in situ measurements (SENAMHI) estimated by uncorrected TRMM 3B43 data, corrected TRMM data using both the additive model (TRMM A) and multiplicative model (TRMM M). The comparison is made for each region based on the two time series: the mean monthly values for the in situ (punctual points) and TRMM data (cell values). In a second step, the TRMM data time series were corrected with the additive and multiplicative models and then the mean annual values were calculated.

CORRECTION OF TRMM 3B43 MONTHLY PRECIPITATION DATA


Ronchail J, Gallaire R. 2006. ENSO and rainfall along the Zongo valley (Bolivia) from the Altiplano to the Amazon basin. International Journal of Climatology 26: 1223–1236.


ACKNOWLEDGEMENTS

The authors thank SENAMHI for their help gathering data. Furthermore, the authors greatly appreciate the constructive comments made by Waldo Lavado. We are very grateful to Sara Mullin who corrected the English. We are thankful to the two anonymous reviewers who have greatly improved the quality of this article.

REFERENCES


