

A High-resolution Precipitation 2-step mapping Procedure (HiP2P): Development and application to a tropical mountainous area



J.E. Hunink ^{a,*}, W.W. Immerzeel ^{b,c}, P. Droogers ^b

^a FutureWater, Paseo Alfonso XIII, 48, 30203, Cartagena, Spain

^b FutureWater, Costerweg 1V, 6702 AA, Wageningen, The Netherlands

^c Universiteit Utrecht, Department of Physical Geography, The Netherlands

ARTICLE INFO

Article history:

Received 23 May 2013

Received in revised form 24 July 2013

Accepted 24 August 2013

Available online xxxx

Keywords:

Precipitation

TRMM

Vegetation

NDVI

Elevation

Proxy

ABSTRACT

Understanding the spatial and temporal variability of precipitation in tropical high mountain areas remains a key challenge. Point measurements are often not sufficient to capture the strong spatial variability particularly in mountain regions. Satellite remote sensing allows capturing the spatial heterogeneity of precipitation, yet it is generally characterized by significant bias. Rainfall satellite products such as those coming from the Tropical Rainfall Measuring Mission (TRMM) are being continuously improved and an increasing amount of high- and medium-resolution remote sensing data on biophysical surface properties is becoming available. Here we present a methodology that blends two TRMM products with remote sensing data on vegetation and topography to quantify the spatial distribution of precipitation in areas where direct observations are lacking. The approach assumes that vegetation cover, the topography and satellite-derived estimates of rainfall are reasonable indirect measures of ground-based precipitation. The methodology is evaluated for an area in the Andes of Ecuador. The results show that around 40% of the variance in weekly precipitation is explained by these proxies. During the drier periods of the year, vegetation is the strongest proxy. In the very wet areas and during the wet periods vegetation is usually in a climax development phase with no development trends to correlate with rain, and the other proxies dominate precipitation estimation. A cross-validation procedure in which each one of the weather stations is sequentially excluded from the analysis, was applied to test the performance of the methodology. The performance was satisfactory, and as expected it is related to the density of the weather station network and temporal rainfall variability. Overall we conclude that the methodology is useful for areas with very high variable conditions, where sufficient ground-data is available to establish the relationships with the different remote sensing datasets.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Understanding and quantifying the spatial and temporal variability of rainfall are of key importance in hydrological studies as precipitation drives most hydrological, environmental and agricultural processes. Especially in mountainous areas the spatial and temporal variability of rainfall can be very high. Climatic conditions may vary strongly depending on topography, aspect, and slope. This causes additional challenges for its correct estimation (Immerzeel, Pellicciotti, & Shrestha, 2012). One of the areas where understanding rainfall variability is of special relevance is the Andes mountain range (Buytaert, Celleri, Willems, Bièvre, & Wyseure, 2006; Vuille, Bradley, & Keimig, 2000). Several areas in the Andes currently experience economic growth, ongoing land use changes, and increasing agricultural and environmental pressures.

Strong precipitation gradients over short distances are difficult to capture with point measurements from weather stations. Weather

stations are generally located in areas which are readily accessible, so in mountainous areas the density of the monitoring network is usually low and insufficient for the use of conventional spatial interpolation techniques (Celleri, Willems, Buytaert, & Feyen, 2007; Ward, Buytaert, Peaver, & Wheeler, 2011).

Satellite rainfall products offer the unique opportunity to improve rainfall estimates in such areas. Advantages of these products are their consistency and uniformity in estimating temporal and spatial variability. A disadvantage is the need to use ground data in order to remove bias to obtain accurate enough estimates for regional assessments (Cheema & Bastiaanssen, 2012; Dinku, Chidzambwa, Ceccato, Connor, & Ropelewski, 2008). Uncertainty in TRMM estimates depends principally on the topography in the area (Bookhagen & Burbank, 2006; Gebregiorgis & Hossain, 2013). TRMM estimates have been reported to compare relatively well with ground-based measurements in low-altitude environments (Stampoulis & Anagnostou, 2012; Tian & Peters-Lidard, 2010). There are more difficulties in areas where orographic effects are important (Chen, Ebert, Walsh, & Davidson, 2013; Dinku et al., 2008), but others have found even reasonable performance in areas

* Corresponding author.

E-mail address: j.hunink@futurewater.es (J.E. Hunink).

of complex topography (Montero-Martínez, Zarraluqui-Such, & García-García, 2012).

The limited spatial resolution of satellite rainfall products makes them less adequate for certain applications as for example watershed planning (Su, Hong, & Lettenmaier, 2008). Several authors therefore have developed interpolation and aggregation algorithms that combine remote sensing data with ground data (e.g. Immerzeel, Rutten, & Droogers, 2009; Scheel et al., 2011; Yatagai et al., 2012). Data resulting from these algorithms have temporal resolutions that range from daily to monthly, and typically spatial resolutions from 10 km to 250 km. Performance of these algorithms is directly related to the temporal aggregation level and spatial scale, and aggregation of data on a weekly or monthly basis is needed for adequate performance (Ouma, Owiti, Kipkorir, Kibiyi, & Tateishi, 2012; Scheel et al., 2011). Spatial aggregation may increase performance (Brunsell & Young, 2008; Cheema & Bastiaanssen, 2012) but others have not found a significant effect (e.g. Scheel et al., 2011).

There is a large amount of remote sensing data available on biophysical variables that are related to precipitation. Examples are the Normalized Difference Vegetation Index (NDVI), Leaf Area Index (LAI), vegetation fraction cover, Fraction of Absorbed Photosynthetically Active Radiation (FAPAR) and elevation. Recently, several authors have made use of these proxies together with satellite rainfall data to estimate precipitation. Immerzeel et al. (2009) successfully applied remote sensing-based NDVI as proxy to downscale precipitation of the Iberian Peninsula. The authors investigated the effect of the spatial resolution on the relationships between NDVI and precipitation. Quiroz, Yarlequé, Posadas, Mares, and Immerzeel (2010) and Heidinger, Yarlequé, Posadas, and Quiroz (2012) applied wavelet multi-resolution analysis using NDVI data to improve TRMM daily rainfall estimates at meteorological stations located in the Andean Plateau. Jia, Zhu, Lü, and Yan (2011) used linear relationships between annual NDVI and elevation to downscale TRMM data. They explored regression relationships at different scales to downscale precipitation to 1 km resolution.

So far most studies focused on the use of one remote sensing-derived dataset of land surface attributes to enhance estimates of the spatial distribution of rainfall. Also, most studies have used a single TRMM product for analyses. In this study we explore the use of three proxies in a multiple regression approach, using two different TRMM products to optimize both the spatial and temporal precipitation resolutions. The objective of this study is to develop and validate a procedure to quantify the spatial distribution of precipitation at high spatial resolution with a weekly time step in a tropical mountainous area based on observed data, vegetation response, elevation and TRMM derived

estimates. The methodology is evaluated in the Andean mountains with a renowned strong spatial and temporal variability of rainfall.

2. Study area

The study area comprised the Tungurahua province (3389 km²), located in the central part of Ecuador and which is part of the Patate river basin, draining into the Amazon. Altitudes in the province range from 1200 to 5000 m above sea level (Fig. 1). The morphology of the area is characterized by its volcanic origin and activity, and the inter-Andean depression flanked by the Cordillera Real to the east and the Cordillera Occidental to the west.

The region has strong climatic gradients and is well-known for its micro-climates: small areas with very specific climatic conditions. Semi-arid conditions prevail in the inter-Andean valley, with annual rainfall around 500 mm and an average annual temperature of 14 °C. The high mountain areas have a temperate climate with volcanic peaks covered with snow for most of the year. The eastern slopes of the Cordillera Real receive annual amounts of rainfall around 3000 mm, originating from influx of moist air from the Amazon basin (Garreaud, 2009).

In terms of land use, the high mountains are characterized by the moorlands (*páramos*) with a renowned environmental and hydrological value (Céleri & Feyen, 2009). The inter-Andean depression is filled with alluvial fan, fluvial, lacustrine, and contemporaneous volcanic deposits (Winkler et al., 2005), providing fertile grounds for cultivation. Around 50% of the total surface area is used for agriculture, while the remaining is natural and urban area. Due to the high agro-climatic variability a wide range of fruits, vegetables and grains is cultivated in the region. Most of the national fruit production originates from this region.

3. Methodology

3.1. Approach

The rainfall satellite products which have been studied and evaluated most are those based on data from the TRMM satellite (Simpson, Kummerow, Tao, & Adler, 1996). The objective of the TRMM is to provide accurate global tropical rainfall estimates by using a combination of instruments designed entirely for rainfall observation. Moreover, some of the derived products are calibrated with selected ground-based observation sites. Since its launch, the TRMM algorithms and products have been continuously improved, and since May 2012 version 7 is available.

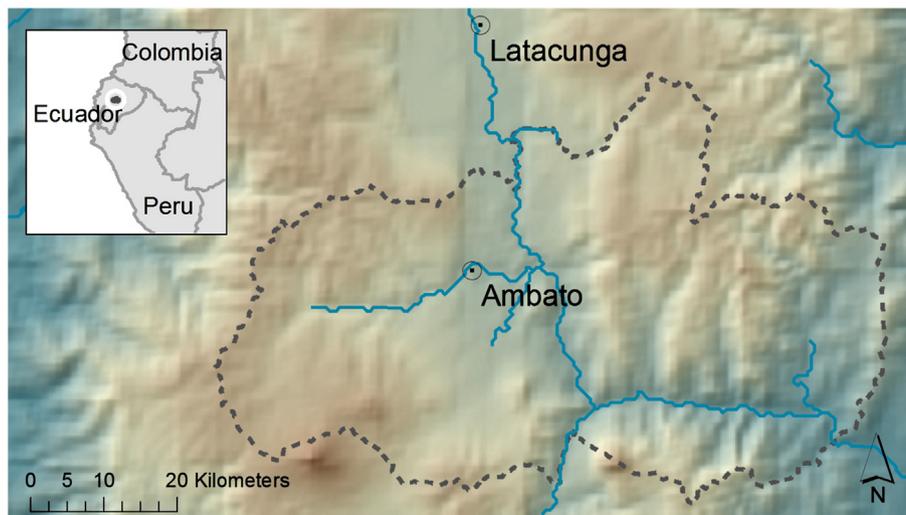


Fig. 1. Location of the Tungurahua province in Ecuador.

The methodology presented here, referred to as High-resolution Precipitation 2-step Procedure (HiP2P), combines ground-based rainfall measurements with two TRMM products and two remote sensing-based datasets with different temporal and spatial resolutions. The aim is to generate reliable spatial and temporal estimates of rainfall over large areas at a very high spatial resolution.

The guiding principles for the procedure presented here are: (i) combining coarse spatial, but high temporal resolution TRMM (daily, 25×25 km) with high spatial, but low temporal TRMM (monthly, 4×4 km) resolution; (ii) bias correct TRMM with observations; (iii) use NDVI (vegetation index) to enhance spatial resolution to 1×1 km; and (iv) use DEM to correct for missing high altitude observations.

The precipitation mapping procedure is based on the following datasets:

- Precipitation records of the weather stations.
- Daily spatial rainfall estimates from TRMM.
- Average monthly rainfall patterns derived from TRMM.
- Digital Elevation Model obtained from SRTM.
- NDVI derived from SPOT-VEGETATION.

The final outcome of the procedure is a time series of spatial distributions over a representative period that should allow better insight in the temporal and spatial patterns of rainfall in the area. However, for detecting climatic trends it is less useful because human factors and land use change can influence NDVI significantly on the long-term. It is challenging to separate this effect from a pure climate signal.

3.2. Datasets

3.2.1. Precipitation records

Data from 28 weather stations within the study area (15) and up- and downstream in the basin outside the study area were used (Fig. 2). Daily rainfall data were available covering the period 1998–2011.

3.2.2. TRMM rainfall

The Tropical Rainfall Measuring Mission (TRMM) is a satellite launched and operated by the US Space Agency (NASA) and the Japanese Aerospace Exploration Agency (JAXA). The satellite mission is focused on providing data on tropical and subtropical rainfalls and to estimate its associated latent heating. TRMM is operational since November 1997 and is releasing products since 1998.

The rainfall measuring instruments on the TRMM satellite include the Precipitation Radar (PR), an electronically scanning radar operating at 13.8 GHz, the TRMM Microwave Image (TMI), a nine-channel passive microwave radiometer, and a Visible and Infrared Scanner (VIRS), a five-channel visible/infrared radiometer.

The development of TRMM products has gone through various phases. Continuous improvements are implemented to obtain more reliable products with higher accuracy. Significant changes were made to the current version 7 compared to the previous version, including additional satellite data and a standardized data reprocessing and calibration facility, using global data from rain gauges. Initial assessments of version 7 showed significant improvements in accuracy (Duan & Bastriaanssen, 2013; Junzhi, A-Xing, & Zheng, 2012).

In spite of these improvements, bias removal and ground-truthing remain necessary when using TRMM products for regional assessments (Hong, Hsu, Moradkhani, & Sorooshian, 2006). Temporal patterns may be captured correctly by TRMM but still significant bias has been found when comparing rainfall amounts with measured quantities at weather stations (Scheel et al., 2011; Su et al., 2008).

For this study, two TRMM products were used.

- 3B42_V07: The purpose of the 3B42 algorithm is to produce TRMM-adjusted merged-infrared (IR) precipitation and root-mean-square (RMS) precipitation-error estimates. The algorithm consists of two

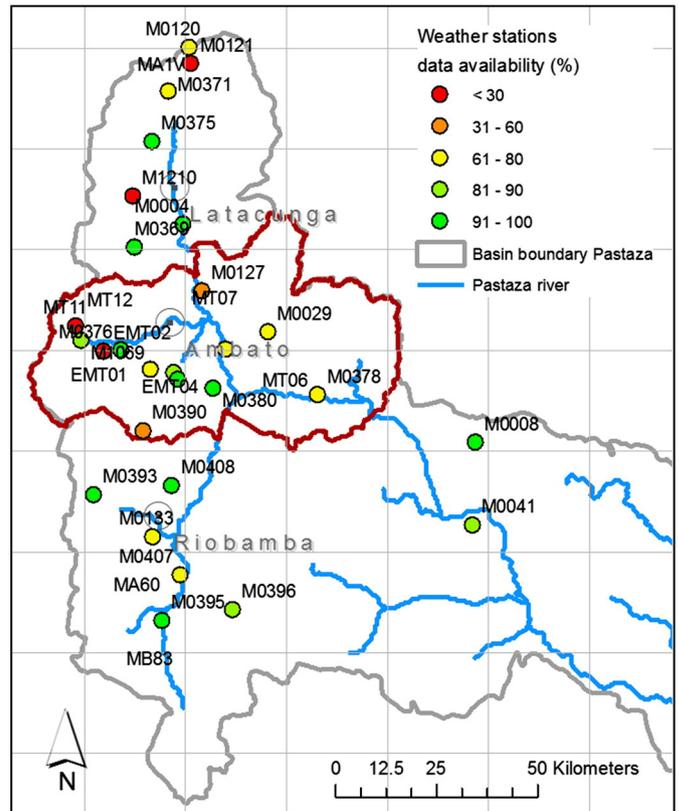


Fig. 2. Map showing for each weather station the amount (%) of valid data available for the study (dots). The grid shows the resolution of the TRMM 3B41 product.

separate steps. The first step uses the TRMM VIRS and TMI orbit data and the monthly TMI/TRMM Combined Instrument (TCI) calibration parameters to produce monthly IR calibration parameters. The second step uses these derived monthly IR calibration parameters to adjust the merged-IR precipitation data, which consists of GMS, GOES-E, GOES-W, Meteosat-7, Meteosat-5, and NOAA-12 data. Then, gauge correction is performed using the uniformly processed database by the Global Precipitation Climatology Centre (GPCC). The final gridded, adjusted merged-IR precipitation (mm/h) and RMS precipitation-error estimates have a 3-hourly temporal resolution and a 0.25-degree by 0.25-degree spatial resolution. Spatial coverage extends from 50 degrees south to 50 degrees north latitude. The final product has a 1-day temporal resolution and a 25×25 km spatial resolution (Simpson et al., 1996).

- 2B31_V07: The 2B31 product is a combined PR/TMI rain rate and path-integrated attenuation at 4 km horizontal, and 250 m vertical, resolution, over a 220 km swath. For this product, data was processed from 1998 to 2006 and each of the orbits (approx. 16/day) was fitted to an equally-spaced grid with a bilinear interpolation to account for projection and resolution differences (Bookhagen & Strecker, 2008). Prior to rainfall calibration, the TRMM 2B31 data is scaled with the number of measurements in each grid cell, because higher latitudes are more frequently measured due to the TRMM orbital paths. In order to convert the 2B31 rainfall rate into absolute rainfall, the scaled data is calibrated with gauged rainfall data from 1970 stations. The final product has an average monthly temporal resolution and a 4×4 km spatial resolution.

By using these two products we aimed to make optimal use of the high spatial resolution of the 2B31_V07 products and high temporal resolution of the 3B42_V07 product.

3.2.3. Digital Elevation Model

A Digital Elevation Model (DEM) was used to include topographical effects on precipitation patterns. Data from the NASA Shuttle Radar Topographic Mission (SRTM), currently distributed by the USGS, were used. The SRTM data are available as 3 arc second (approx. 90 m) resolution. The vertical error of the DEMs is less than 16 m (Sun, Ranson, Kharuk, & Kovacs, 2003). SRTM is currently the de-facto standard for elevation models in hydrological assessments. For this study the SRTM data were resampled to 1×1 km to match with the NDVI data.

3.2.4. The Normalized Difference Vegetation Index (NDVI)

The NDVI was used as a proxy for the temporal and spatial dynamics of precipitation. The vegetative response to precipitation can be relatively fast under arid, semi-arid and sub-humid conditions. The relationship of NDVI and precipitation has been used previously to convert the spatial distribution of TRMM precipitation from daily to annual temporal resolutions (Immerzeel et al., 2009; Quiroz et al., 2010). NDVI data from the French SPOT satellite have been used and evaluated in many studies previously. SPOT consists of two observation instruments in orbit, VEGETATION 1 and VEGETATION 2, as well as ground infrastructures. The first of the two instruments in orbit is aboard the SPOT 4 satellite, launched on 24 March 1998. The second is aboard SPOT 5, which was placed into orbit on 4 May 2002.

The VGT-S10 (ten day synthesis) products are composite (maximum-value) products. All the segments of this period are compared to pick out the 'best' ground reflectance values. These products provide data from all spectral bands, the NDVI, and auxiliary image acquisition parameter data. The continental S10-composite data products (spectral band data, data quality layer, and NDVI) are provided by Vito, Belgium. The individual composite NDVI data for each period were extracted from the S10-HDF file and post processed. The post-processing steps include a re-projection from the native global Mercator projection, regional sub-setting, and incorporation of flags for bad data, clouds, and a land mask. The data flag, cloud and land mask information are obtained from the Status Map (SM) layer (data quality information) included as part of the S10 synthesis product catalogue. The original NDVI dataset is distributed in a format using digital numbers (DN). The true value of NDVI has to be calculated from the 8 bit decadal images.

3.3. Rainfall mapping procedure

The spatial rainfall mapping consisted of two main steps: (i) a prediction of spatial distribution based on regression models obtained from the different spatial datasets, and (ii) an interpolation of observations of spatial distribution to include the variability not explained by the models. Several pre-processing steps had to be done as explained in the following section. The complete procedure is detailed in the flow chart in Fig. 3.

3.3.1. Pre-processing of datasets

All datasets were downloaded, re-projected, re-sampled and clipped to the area of interest. The analysis was done for the period covered by all datasets from 1-Apr-1998 to 31-Dec-2011 (almost 14 years).

The decadal distributions of NDVI show gaps in some areas and during some periods due to persistent cloud cover. These were filled by using the spatial distribution of mean monthly NDVI. The result is a cloud-filled decadal time series of NDVI grids for the period 1-Apr-1998 to 31-Dec-2011.

The observed precipitation records (OBS) were subjected to a quality check for missing values, and were aggregated to monthly values. The resulting data gaps were filled by calculating the bias factor K between OBS and the monthly TRMM observations at each weather station and were summarized as follows:

$$OBS_i = K * TRMM_i \quad (1)$$

with OBS_i as the monthly corrected observed rainfall at raingauge i , and $TRMM_i$ as the monthly TRMM rainfall at raingauge i based on TRMM 3B42_V07, and K as the monthly specific bias factor.

3.3.2. Step 1: regression modeling and prediction

Data representing conditions at the weather station locations were extracted from the five spatial datasets (TRMM 3B42_V07, TRMM 2B31_V07, NDVI, OBS, DEM). Weekly time series were created to form a dataset that contained for each of the 28 weather stations and for each week in the 14-year period, the weekly rainfall amount, the NDVI, the monthly average rainfall from the 2B31 dataset (further on in text: CLIM) and the altitude (DEM). This dataset was split into separate datasets by the week of the year, resulting in 52 separate datasets. For each of these datasets, a linear regression model was fitted, as follows:

$$PCP_j = a + b * DEM + c * CLIM_j + d * NDVI_{j+1} \quad (2)$$

In which a , b , c and d are the fitting coefficients of the regression, and j is the week number in the year. This resulted in a total of 52 fitted regression equations, each based on 392 data points (28 stations \times 14 years). For NDVI we assume a lag time of 1 week in the regression models to mimic the response of vegetation to precipitation.

The 52 regression models were subsequently applied to the spatial datasets at the resolution of 1×1 km to obtain a model estimate for each pixel and for each week in the 14-year analyzed period. This resulted in a precipitation estimate based on elevation, NDVI and the climatic conditions at a much higher resolution than the TRMM and observed precipitation.

3.3.3. Step 2: spatial correction with observations

Only part of the variability is explained by the regression models. This is because the input variables represent only some of the factors that determine the complex and erratic behavior of rainfall. The obtained models describe average patterns but do not accurately capture extremes.

Therefore, in the second step precipitation variability, which was not explained by the regression models, was included by deriving the residuals (modeled precipitation minus observed) for each station. These residuals were then spatially interpolated (spline) and added to the predicted maps as described by Immerzeel et al. (2009).

3.4. Predictive uncertainty analysis

The two-step procedure guarantees a perfect match between the predicted and observed values at the weather station locations. For this reason, the procedure does not provide a good estimate of prediction performance and there is no effective measure to indicate how well the observed dataset is reproduced. Cross-validation is a useful method to obtain a more reliable estimate of the predictive quality at each point.

Cross-validation consists of removing sequentially one of the stations and evaluating model performance in the absence of the particular station. This is also called leave-one-out cross-validation. This process was repeated for each of the 28 stations such that each observation in the sample is used once for validation.

This cross validation procedure gives an estimate of the predictive quality of the algorithm used for this purpose. The estimate can be spatially represented and interpolated to obtain an indication of the accuracy of the final spatial output. The accuracy of the model was assessed by the coefficient of determination (R^2) and the Normalized Root Mean Square Error (NRMSE: root mean square error divided by the difference between the maximum and minimum observed precipitations).

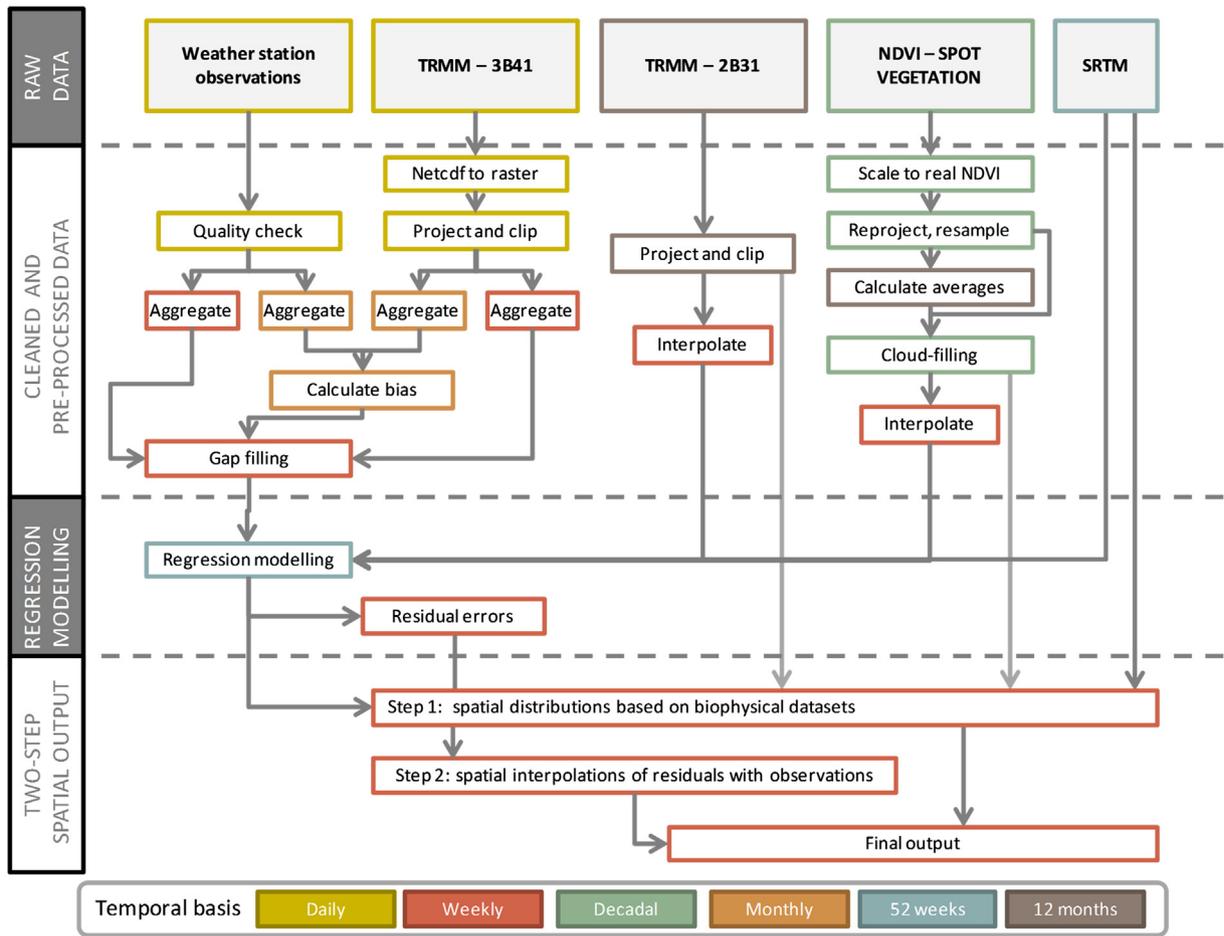


Fig. 3. Flow chart of the algorithm used.

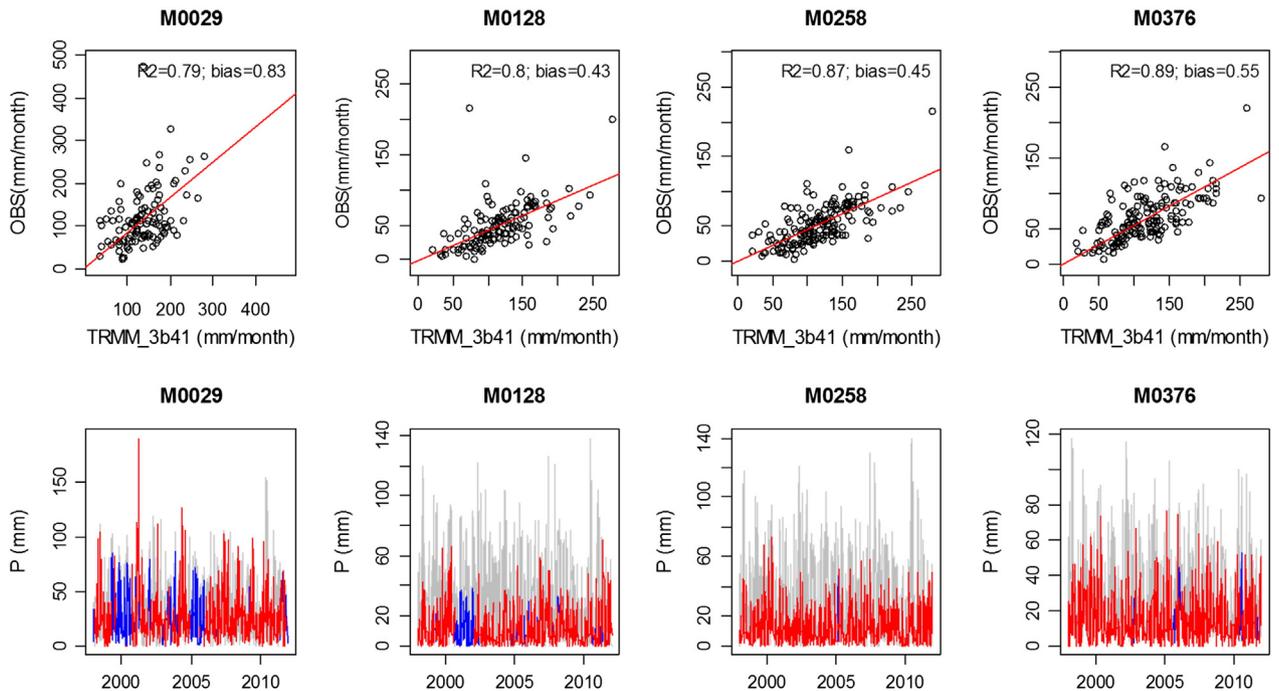


Fig. 4. The linear relationship established for each station between ground observations and TRMM to derive the bias factor for four stations (top plots) and the corrected time series for the same stations (bottom plots) (blue = corrected, red = original observations, gray = TRMM 3B43v07). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

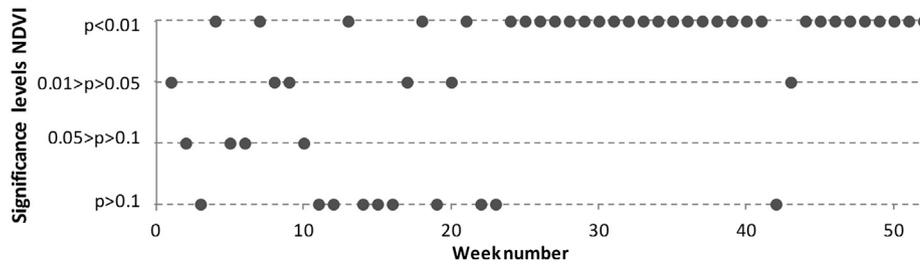


Fig. 5. Significance levels for each of the factors in the regression, for each week of the year (x-axis).

4. Results

4.1. Rainfall time series

The observed precipitation records were quality checked to detect outliers in the dataset. Most stations contained missing values. Fig. 2 shows the integrity (days with valid values divided by total days) of the dataset visually.

The missing values were filled by deriving the bias factor with the TRMM_3B41 dataset at each location, as explained before. The bias factor ranges from 0.33 to 1.35 with an average of 0.63. The R² between OBS and TRMM ranges from 0.63 to 0.99 with an average of 0.81. We conclude that there is a reasonable match between TRMM and OBS, but there are significant biases in some cases and TRMM generally overestimates precipitation in this area. Fig. 4 shows an example for four of the stations in which in the upper part the red line corresponds with the slope through the origin, i.e. the bias factor, based on monthly rainfall amounts.

4.2. Regression models

The linear regression carried out for each week of the year resulted in 52 regression models specific to the area of interest that relate altitude (DEM), vegetation index (NDVI) and the average monthly rainfall based on TRMM 2B31 (CLIM), with gap-filled weekly precipitation (OBS). The R² of the obtained models varies between 0.23 and 0.56, with an average of 0.40. This indicates that an average of 40% of the variance in precipitation is explained by the regression model. By normalizing the RMSE values (dividing them by the difference between the maximum and minimum of the observed values), a relative error of 11% is obtained on average.

The regression function fits four coefficients, as in Eq. (2). Fig. 5 shows the significance levels for each of the factors. These levels were established according to the p-values of the resulting regression coefficients. The lowest level refers to p-values > 10%, meaning that the factor

is not significant, the following level indicates reasonable significance (p-value between 5% and 10%) and the upper levels indicate good to very good significance.

These results show that the significance level of NDVI is higher during the second half of the year, which are drier months in most of the basin than the first half of the year. This seems logical as the vegetation responds more directly to rainfall during drier periods than during wet periods. For the other coefficient no yearly pattern is observed in the significance levels.

The relevance of each of the explanatory variables in the regression function can also be interpreted from the value or weight of each of the fitted coefficients. Fig. 6 shows the value of the coefficients *b*, *c* and *d* (respectively for DEM, CLIM and NDVI), for all the models obtained for each week of the year. Some clear trends can be observed: the DEM coefficient is lower between weeks 15 and 25, the CLIM coefficient is higher around week 3; around week 12 (start of April) and around week 42. Also, the NDVI coefficient is generally higher in the second half of the year than in the first half. The relation between NDVI and CLIM is positive, e.g. a high NDVI value corresponds to high precipitation. The DEM coefficient is generally negative, i.e. generally speaking precipitation decreases with altitude in this area.

4.3. Spatial analysis

As a first step, the regression models were applied to the spatial datasets (DEM, CLIM and NDVI) to obtain a prediction of the spatial distribution of precipitation for each week in the 14-year period. Then, in the second step of the procedure the interpolated residuals between the observed and estimated values at the station locations were calculated for each week and added to the predicted maps from the regression models. This resulted in a weekly time series of precipitation covering the entire region on a 1 × 1 km resolution. The dataset allows insight in the spatial distributions on a weekly basis. Fig. 7 shows the mean rainfall distribution of week 22, which is the week when the study area receives on average most rainfall (left) and of week 40 which is on

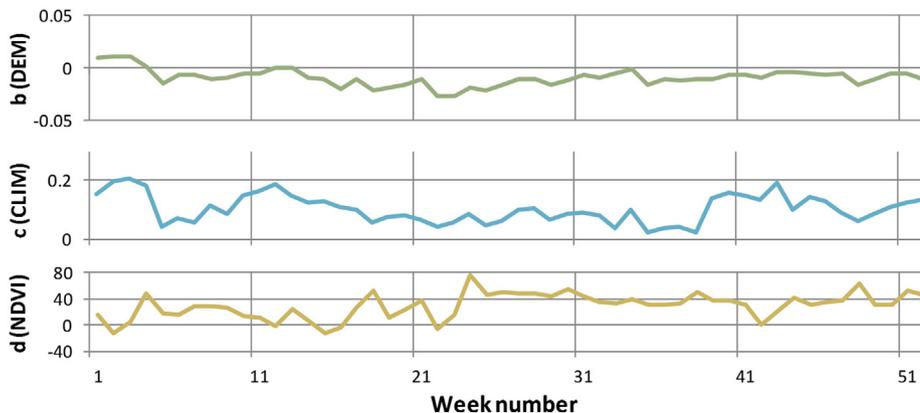


Fig. 6. Values of the fitted coefficients in the regression models for each week of the year.

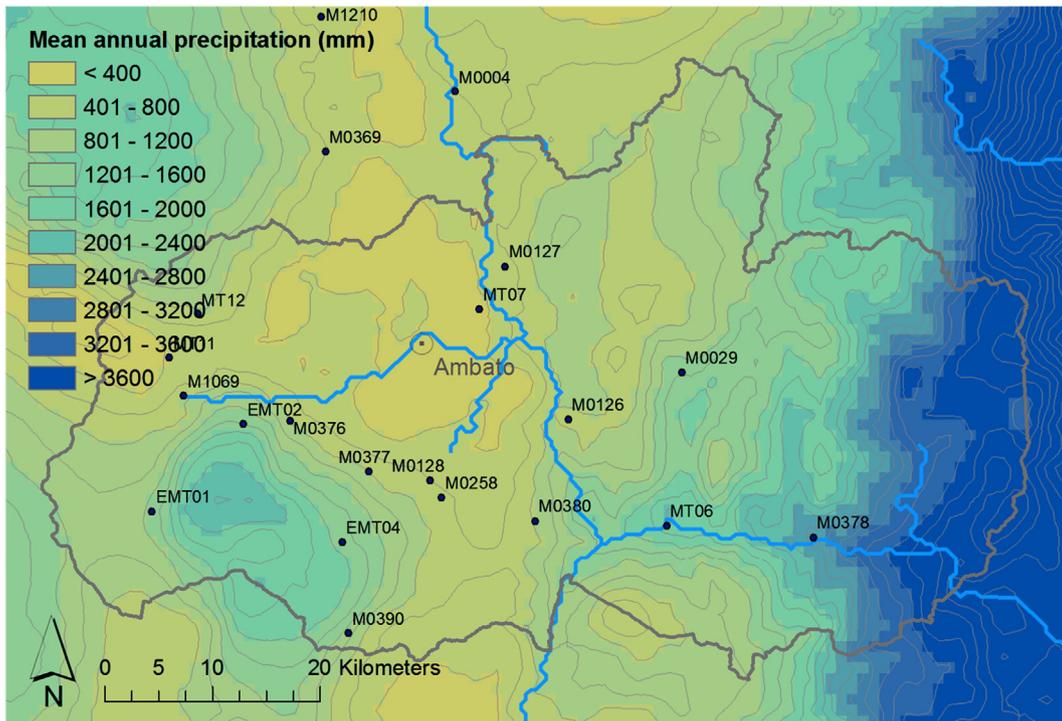


Fig. 7. Mean rainfall distribution of week 22 – wet (left) and of week 40 – dry (right) based on 14-year weekly spatial rainfall maps.

average the driest week. Fig. 8 shows the mean annual rainfall based on the weekly time series of the entire study period.

The results show that spatial rainfall variability in the region is very high. Three major regions can be distinguished with different pluvial regimes. Firstly, the moorlands of the eastern Ecuadorian Andes received more than 1000 mm during a wet period between April and August. The second region corresponds with the eastern flanks of the same mountain range which received an average of more than 3000 mm per year in the lowest parts of the basin. The third zone is the inter-Andean valley with annual precipitation of around 500 mm. This area is highly variable in terms of rainfall with some small areas receiving significantly more than this amount (locally referred to as ‘micro-climates’).

It is not possible to understand the high variability in the area from common spatial interpolation methods. Fig. 9 (left) shows the annual mean rainfall based on a frequently used interpolation method (spline) which was applied on the annual means at the weather stations. Clearly such method, based purely on weather station data, does not reflect the high spatial variability which exists in the area. On the other hand,

remote sensing has the disadvantage of low resolution and low accuracy, as demonstrated in Fig. 9 (right).

4.4. Intra-annual variability

The intra-annual variability can be expressed by calculating the coefficient of variation (the ratio of the standard deviation to the mean rainfall at each point). This gives an indication of the relative differences between wet and dry periods. Fig. 10 shows the spatial representation of the coefficient of variation, based on the entire 14-year period of weekly rainfall.

The map shows that in the region around the city of Ambato, the temporal variation is highest. This relatively high variability of rainfall coincides with the areas of principally annual crop cultivation. The area south of Ambato shows lower variability and coincides with areas where perennials are cultivated: principally fruit trees. The coefficient of variation is also low in the downstream areas and the high precipitation amounts allow fruit production.

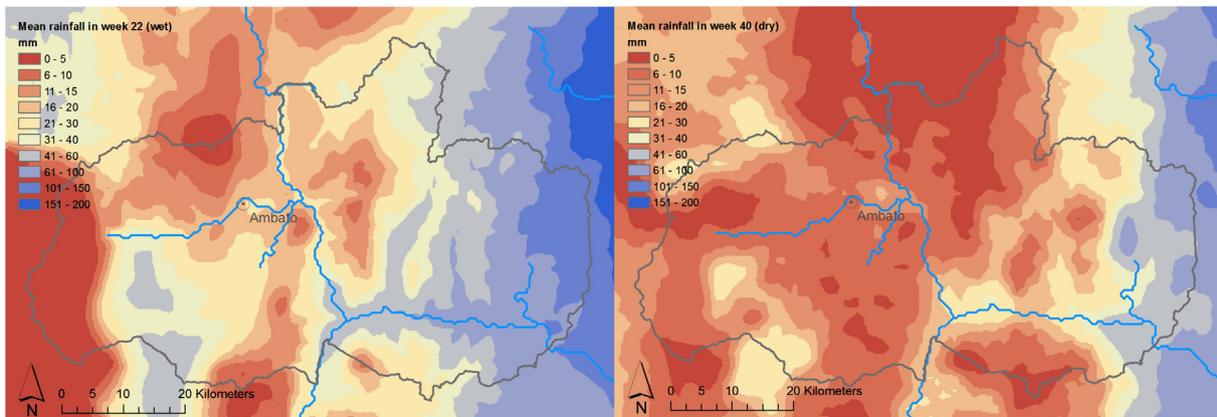


Fig. 8. Annual downscaled rainfall based on 14-year weekly spatial rainfall maps.

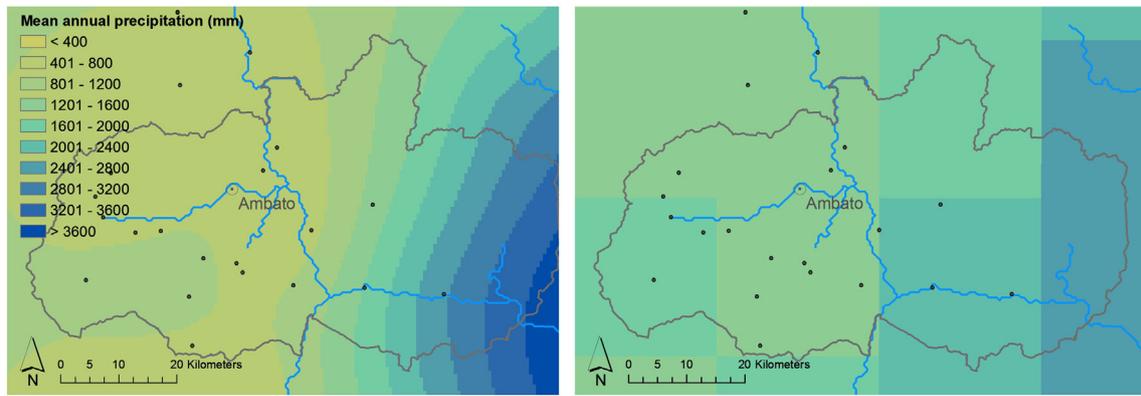


Fig. 9. Left: annual rainfall distribution by interpolating (spline) the 14-year means at the weather stations. Right: annual rainfall based on the daily grids of TRMM 3B41V7 product.

4.5. Accuracy assessment

The predictive quality of the model was assessed by using cross-validation, as explained in the methodological section. The R^2 obtained by sequentially excluding one station from the modeling, was plotted and spatially interpolated (using spline) (see Fig. 11). This gives an indication of the accuracy of the final output over the entire area and the predictive quality of the procedure. The average R^2 based on monthly rainfall amounts over all the stations is 0.75.

As can be expected, the prediction quality is highest in the areas with a relatively high density of weather stations, as in the south-western part of the province. In the south-eastern part where the temporal variability is low and rainfall amounts are high, the correlation is also relatively high. The performance in the northern part of the province is lower as a result of the high rainfall variability and low density of weather stations.

The lower areas around Ambato, receiving smaller amounts of rainfall, clearly show a lower predictive quality, probably due to the higher temporal variability in these areas. The wetter areas in the province are predicted with the highest accuracy.

5. Conclusions

In this study we present a methodology to estimate spatial distributions of precipitation at high spatial resolution with a weekly time step in a tropical mountainous region in Ecuador. Data from meteorological stations and four remote sensing datasets were combined to improve the quantification of the spatial distribution of precipitation in areas where direct observations are lacking.

The TRMM 3B42 product provides a relatively coarse resolution of rainfall ($0.25 \times 0.25^\circ$). An algorithm has been developed to combine this rainfall dataset with datasets on climate, vegetation index (as a proxy for rainfall) and altitude, in the Tungurahua province in order to: (i) correct TRMM rainfall data with rainfall field observations and (ii) obtain weekly rainfall maps of the region with a spatial resolution of 1×1 km. The product of such an approach is weekly spatial rainfall estimates based on NDVI, TRMM-based monthly average rainfall, DEM, and field observations.

The approach assumes that the state of the vegetation, the elevation and satellite derived estimates of precipitation explain part of the variability in actual precipitation. The results show that this is indeed the

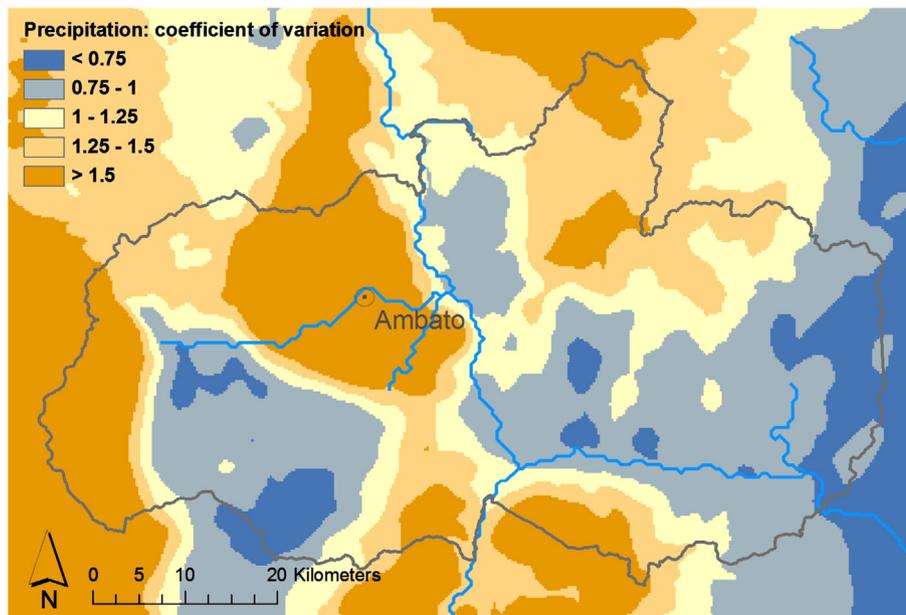


Fig. 10. Coefficient of variation based on weekly variation in rainfall over the whole period 1998–2011.

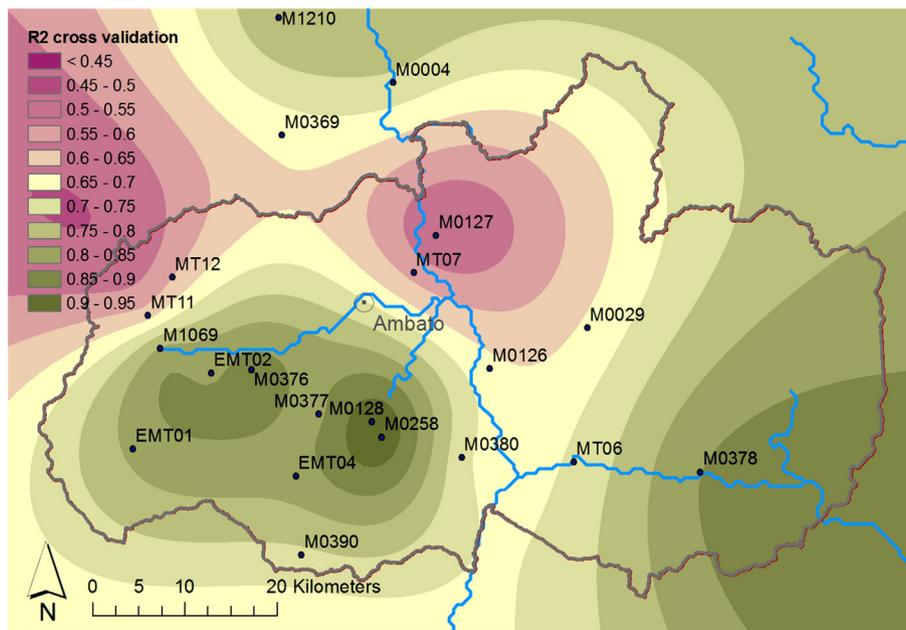


Fig. 11. Interpolated R^2 as indicator for the predictive performance calculated using cross-validation.

case and that for this highly variable area 40% of the variance in weekly precipitation is explained by these proxies, with vegetation being the strongest proxy. The unexplained part of the variance, determined as a residual at the weather station locations, is spatially interpolated to obtain a perfect fit with observations at the weather stations.

A leave-one-out cross-validation procedure was applied to test the performance of the methodology. The performance was satisfactory, and as expected related to the density of the weather station network and temporal rainfall variability. In areas with high temporal variability, the prediction quality decreases with the distance from the weather stations. However, in the eastern and wet part of the study area this effect is less clear and the performance is relatively high due to the lower temporal variability.

Our results may be more uncertain in the higher mountain areas where the variation in elevation is larger. This is difficult to quantify because most meteorological stations are located in the valleys. The reason for this possible higher uncertainty is that on windward slopes of mountains there is often a strong positive relation between elevation and precipitation, whereas on the lee side a rain shadow is formed. This relation is usually not straightforward as the saturation vapor pressure is not linear through an atmospheric column. A second reason is that at higher altitudes the temperature is lower and the NDVI may be limited by temperature rather than precipitation. We therefore strongly recommend increasing in-situ observations at high altitude in combination with wind fields to improve the quantification of precipitation patterns in mountain regions.

The results show that the procedure performs best during the drier period of the year, when vegetation development responds more directly to precipitation. Also, the high precipitation of around 3000 mm in the eastern humid areas, causes vegetation to be fully developed and thus do not have development trends which can be related to trends in precipitation. This makes the methodology more reliant on the other proxies in these tropical areas. The use of other vegetation indices that is less dependent on trends in vegetation structural development can be investigated in future studies, as for example EVI (Enhanced Vegetation Index), which is reported to exhibit less saturated signals for high biomass conditions (Huete et al., 2002).

Apart from the TRMM-based monthly average rainfall, possibly also other gridded climate variables could be included in the procedure that take into account the seasonality of the rainfall regime and could

explain an additional part of the variability, such as wind direction, temperature or potential evapotranspiration. It is also recommended to examine the lag time for NDVI in the regression models as this may be specific to vegetation in each area of study and may improve the spatial predictions.

Overall we conclude that the methodology is useful for areas with very high variable conditions, where sufficient ground-data is available to establish the relationships with the different remote sensing datasets.

Acknowledgments

The authors are grateful to the PACT program of the Tungurahua province in Ecuador for providing financial support to conduct this study, and to the National Institute of Meteorology and Hydrology (INAMHI) for providing the meteorological station data.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at <http://dx.doi.org/10.1016/j.rse.2013.08.036>. These data include Google maps of the most important areas described in this article.

References

- Bookhagen, B., & Burbank, D. W. (2006). Topography, relief, and TRMM-derived rainfall variations along the Himalaya. *Geophysical Research Letters*, 33(8), 1–5. <http://dx.doi.org/10.1029/2006GL026037>.
- Bookhagen, B., & Strecker, M. R. (2008). Orographic barriers, high-resolution TRMM rainfall, and relief variations along the eastern Andes. *Geophysical Research Letters*, 35(6), L06403. <http://dx.doi.org/10.1029/2007GL032011>.
- Brunsell, N. A., & Young, C. B. (2008). Land surface response to precipitation events using MODIS and NEXRAD data. *International Journal of Remote Sensing*, 29(7), 1965–1982. <http://dx.doi.org/10.1080/01431160701373747>.
- Buytaert, W., Celleri, R., Willems, P., Bièvre, B. De, & Wyseure, G. (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of Hydrology*, 329(3–4), 413–421. <http://dx.doi.org/10.1016/j.jhydrol.2006.02.031>.
- Céleri, R., & Feyen, J. (2009). The hydrology of tropical andean ecosystems: Importance, knowledge status, and perspectives. *Mountain Research and Development*, 29(4), 350–355. <http://dx.doi.org/10.1659/mrd.00007>.
- Celleri, R., Willems, P., Buytaert, W., & Feyen, J. (2007). Space-time rainfall variability in the Paute basin, Ecuadorian Andes. *Hydrological Processes*, 21(24), 3316–3327. <http://dx.doi.org/10.1002/hyp.6575>.

- Cheema, M. J. M., & Bastiaanssen, W. G. M. (2012). Local calibration of remotely sensed rainfall from the TRMM satellite for different periods and spatial scales in the Indus basin. *International Journal of Remote Sensing*, 33(8), 2603–2627. <http://dx.doi.org/10.1080/01431161.2011.617397>.
- Chen, Y., Ebert, E. E., Walsh, K. J. E., & Davidson, N. E. (2013). Evaluation of TRMM 3B42 precipitation estimates of tropical cyclone rainfall using PACRAIN data. *Journal of Geophysical Research-Atmospheres*, 118(5), 2184–2196. <http://dx.doi.org/10.1002/jgrd.50250>.
- Dinku, T., Chidzambwa, S., Ceccato, P., Connor, S. J., & Ropelewski, C. F. (2008). Validation of high-resolution satellite rainfall products over complex terrain. *International Journal of Remote Sensing*, 29(14), 4097–4110. <http://dx.doi.org/10.1080/01431160701772526>.
- Duan, Z., & Bastiaanssen, W. G. M. (2013). First results from Version 7 TRMM 3B43 precipitation product in combination with a new downscaling–calibration procedure. *Remote Sensing of Environment*, 131, 1–13. <http://dx.doi.org/10.1016/j.rse.2012.12.002> (null).
- Garreaud, R. D. (2009). The Andes climate and weather. *Advances in Geosciences*, 22, 3–11.
- Gebregiorgis, A. S., & Hossain, F. (2013). Understanding the dependence of satellite rainfall uncertainty on topography and climate for hydrologic model simulation. *IEEE Transactions on Geoscience and Remote Sensing*, 51(1), 704–718. <http://dx.doi.org/10.1109/TGRS.2012.2196282>.
- Heidinger, H., Yarlequé, C., Posadas, A., & Quiroz, R. (2012). TRMM rainfall correction over the Andean Plateau using wavelet multi-resolution analysis. *International Journal of Remote Sensing*, 33(14), 4583–4602.
- Hong, Y., Hsu, K., Moradkhani, H., & Sorooshian, S. (2006). Uncertainty quantification of satellite precipitation estimation and Monte Carlo assessment of the error propagation into hydrologic response. *Water Resources Research*, 42(8). <http://dx.doi.org/10.1029/2005WR004398>.
- Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., & Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2), 195–213. [http://dx.doi.org/10.1016/S0034-4257\(02\)00096-2](http://dx.doi.org/10.1016/S0034-4257(02)00096-2).
- Immerzeel, W. W., Pellicciotti, F., & Shrestha, A.B. (2012). Glaciers as a proxy to quantify the spatial distribution of precipitation in the Hunza basin. *Mountain Research and Development*, 32(1), 30–38. <http://dx.doi.org/10.1659/MRD-JOURNAL-D-11-00097.1>.
- Immerzeel, W. W., Rutten, M., & Droogers, P. (2009). Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. *Remote Sensing of Environment*, 113(2), 362–370. <http://dx.doi.org/10.1016/j.rse.2008.10.004>.
- Jia, S., Zhu, W., Lü, A., & Yan, T. (2011). A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam basin of China. *Remote Sensing of Environment*, 115(12), 3069–3079. <http://dx.doi.org/10.1016/j.rse.2011.06.009>.
- Junzhi, L., A-Xing, Z., & Zheng, D. (2012). Evaluation of TRMM 3B42 precipitation product using rain gauge data in Meichuan Watershed, Poyang Lake basin, China. *Journal of Resources and Ecology*, 3(4), 359–366. <http://dx.doi.org/10.5814/j.issn.1674-764x.2012.04.009>.
- Montero-Martínez, G., Zarraluqui-Such, V., & García-García, F. (2012). Evaluation of 2B31 TRMM-product rain estimates for single precipitation events over a region with complex topographic features. *Journal of Geophysical Research*, 117(D2), D02101. <http://dx.doi.org/10.1029/2011JD016280>.
- Ouma, Y. O., Owiti, T., Kipkorir, E., Kibiyi, J., & Tateishi, R. (2012). Multitemporal comparative analysis of TRMM-3B42 satellite-estimated rainfall with surface gauge data at basin scales: Daily, decadal and monthly evaluations. *International Journal of Remote Sensing*, 33(24), 7662–7684. <http://dx.doi.org/10.1080/01431161.2012.701347>.
- Quiroz, R., Yarlequé, C., Posadas, A., Mares, V., & Immerzeel, W. W. (2010). Improving daily rainfall estimation from NDVI using a wavelet transform. *Environmental Modelling and Software*, 26, 201–209. <http://dx.doi.org/10.1016/j.envsoft.2010.07.006>.
- Scheel, M. L. M., Rohrer, M., Huggel, C., Santos Villar, D., Silvestre, E., & Huffman, G. J. (2011). Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrology and Earth System Sciences*, 15(8), 2649–2663. <http://dx.doi.org/10.5194/hess-15-2649-2011>.
- Simpson, J., Kummerow, C., Tao, W. -K., & Adler, R. F. (1996). On the Tropical Rainfall Measuring Mission (TRMM). *Meteorology and Atmospheric Physics*, 60(1–3), 19–36. <http://dx.doi.org/10.1007/BF01029783>.
- Stampoulis, D., & Anagnostou, E. N. (2012). Evaluation of global satellite rainfall products over Continental Europe. *Journal of Hydrometeorology*, 13(2), 588–603. <http://dx.doi.org/10.1175/JHM-D-11-086.1>.
- Su, F., Hong, Y., & Lettenmaier, D. P. (2008). Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata basin. *Journal of Hydrometeorology*, 9(4), 622–640. <http://dx.doi.org/10.1175/2007JHM944.1>.
- Sun, G., Ranson, K., Kharuk, V., & Kovacs, K. (2003). Validation of surface height from shuttle radar topography mission using shuttle laser altimeter. *Remote Sensing of Environment*, 88(4), 401–411. <http://dx.doi.org/10.1016/j.rse.2003.09.001>.
- Tian, Y., & Peters-Lidard, C. D. (2010). A global map of uncertainties in satellite-based precipitation measurements. *Geophysical Research Letters*, 37(24). <http://dx.doi.org/10.1029/2010GL046008>.
- Vuille, M., Bradley, R. S., & Keimig, F. (2000). Climate variability in the Andes of Ecuador and its relation to tropical pacific and atlantic sea surface temperature anomalies. *Journal of Climate*, 13(14), 2520–2535. [http://dx.doi.org/10.1175/1520-0442\(2000\)013<2520:CVITAO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2000)013<2520:CVITAO>2.0.CO;2).
- Ward, E., Buytaert, W., Peaver, L., & Wheeler, H. (2011). Evaluation of precipitation products over complex mountainous terrain: A water resources perspective. *Advances in Water Resources*, 34(10), 1222–1231. <http://dx.doi.org/10.1016/j.advwatres.2011.05.007>.
- Winkler, W., Villagómez, D., Spikings, R., Abegglen, P., Tobler, S., & Egüez, A. (2005). The Chota basin and its significance for the inception and tectonic setting of the inter-Andean depression in Ecuador. *Journal of South American Earth Sciences*, 19(1), 5–19. <http://dx.doi.org/10.1016/j.jsames.2004.06.006>.
- Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., & Kitoh, A. (2012). APH-RODITE: Constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, 93(9), 1401–1415. <http://dx.doi.org/10.1175/BAMS-D-11-00122.1>.