



Consistency of satellite precipitation estimates in space and over time compared with gauge observations and snow-hydrological modelling in the Lake Titicaca region

Frédéric Satgé¹, Denis Ruelland², Marie-Paule Bonnet³, Jorge Molina⁴, Ramiro Pillco⁴

5 ¹CNES, UMR Hydrosiences, University of Montpellier, Place E. Bataillon, 34395 Montpellier Cedex 5, France

²CNRS, UMR Hydrosiences, University of Montpellier, Place E. Bataillon, 34395 Montpellier Cedex 5, France

³IRD, UMR Espace-Dev, Maison de la télédétection, 500 rue JF Breton, 34093 Montpellier Cedex 5, France

⁴Universidad Mayor de San Andres, calle 30 Cota Cota, La Paz, Bolivia

Correspondence to: Frédéric Satgé (frederic.satge@gmail.com), Denis Ruelland (denis.ruelland@um2.fr)

10 **Abstract.** This paper proposes a protocol to assess the space-time consistency of satellite precipitation estimates (SPEs) according to various indicators including: (i) direct comparison of SPEs with 72 precipitation gauges; (ii) sensitivity of streamflow modelling to SPEs at the outlet of four basins; and (iii) the sensitivity of distributed snow models to SPEs using a MODIS snow product as reference in an unmonitored mountainous area. The protocol was applied successively to four different time windows (2000–2004, 2004–2008, 2008–2012 and 2000–2012) to account for the space-time variability of the

15 SPEs and to a large dataset composed of 12 SPEs (CMORPH-RAW, CMORPH-CRT, CMORPH-BLD, CHIRP, CHIRPS, GSMaP, MSWEP, PERSIANN, PERSIANN-CDR, TMPA-RT, TMPA-Adj and SM2Rain), an unprecedented comparison. The aim of using different space-time scales and indicators was to evaluate whether the efficiency of SPEs varies with the method of assessment, time window and location. Results revealed very high discrepancies between SPEs compared to precipitation gauge observations. Some SPEs (CMORPH-RAW, CMORPH-CRT, GSMaP, PERSIANN, TMPA-RT and

20 SM2Rain) are unable to estimate regional precipitation whereas the others (CHIRP, CHIRPS, CMORPH-BLD, MSWEP, PERSIANN-CDR and TMPA-Adj) produce a realistic representation despite recurrent spatial limitation over regions with contrasted emissivity, temperature and orography. In nine out of ten of the cases studied, streamflow was more realistically simulated by the hydrological model tested when SPEs were used as forcing precipitation data rather than precipitation derived from the available precipitation gauge networks. Interestingly, the potential of SPEs to reproduce the observed

25 streamflow varied significantly depending on the basin and period considered and did not systematically corroborate SPE potential compared with gauge precipitation observations. SPE's ability to reproduce the duration of MODIS-based snow cover also showed variable consistency over time with poorer simulations in comparison to those simulated from available precipitation gauges. Using snow cover simulations as indicator led to a different efficiency ranking of the SPEs that the ones obtained when using observed gauge precipitation and streamflow. SPEs thus present space-time errors that may not be

30 detected when short time windows and/or scarce gauge networks and/or single indicators are used, underlining how important it is to carefully consider their space-time consistency before using them for hydro-climatic studies. Moreover SPE



efficiency ranked differently depending on the assessment indicators used, suggesting that SPE efficiency should be assessed using indicators related to their final use. Among all the SPEs assessed, MSWEP showed the highest space-time accuracy and consistency in reproducing gauge precipitation estimates, streamflow and snow cover duration. After some adjustment over Lake Titicaca, MSWEP should thus be preferred for the regional hydro-meteorological survey.

5 1. Introduction

1.1 On the need for and difficulty involved in estimating precipitation fields

Water resources are facing unprecedented pressure due to the combined effects of population growth and climate change. In the 20th century, water extraction underwent a six-fold increase to sustain food needs and economic levels due to the increasing world population (vision, water council, 2000). At the same time, global warming has led to the redistribution of precipitation, which has favoured the occurrence of both drought and extreme flood events (Trenberth, 2011).

As a key component of the hydrologic cycle, it is therefore crucial to have accurate precipitation estimates in many research fields including hydrological and snow modelling (e.g. Hublart et al., 2016), climate studies (e.g. Espinoza Villar et al., 2009), extreme flooding (e.g. Ovando et al., 2016) drought (e.g. Satgé et al., 2017a), and monitoring to understand past and ongoing changes and to optimise water resources management (e.g. Fabre et al., 2016, 2015).

Measurements of precipitation are usually retrieved from point gauge stations. Considered as ground truth at the point level, precipitation estimates are then spatialized to represent the distribution of precipitation in space and over time to be used as inputs for impact modelling. However, in most cases, the gauge network is too sparse and unevenly distributed to correctly capture the spatial variability of precipitation. This is especially true for remote regions such as tropical forests, mountainous areas and deserts where the usual insufficient installation and maintenance operations seriously compromise precipitation monitoring. An alternative approach consists in using precipitation derived from weather radar using the backscattering of electromagnetic waves via hydrometeors (e.g. Mahmoud et al., 2018). Unlike measurements using traditional gauges, this technique monitors large areas in a distributed way thus offering the opportunity to monitor precipitation over remote regions. However, ground radar measurements are rarely available at the global or even regional scale (Tang et al., 2016) and are limited in the case of complex terrain which interferes with the radar signal (Zeng et al., 2018). More recently, some authors (Messer et al., 2006; Overeem et al., 2011; Zinevich et al., 2008) reported on the possibility to estimate precipitation from wireless networks such as commercial cellular phone microwave links. These estimations are based on the attenuation of the electromagnetic signals between telecommunication antennas during precipitation events and their first results are promising as precipitation estimates were found to be very accurate (see e.g. Doumounia et al., 2014). However, this technique faces the problem of private cellular phone company policy about sharing data and telecommunication antenna are mainly located in urban areas, which limits accurate precipitation estimates in space. On the other hand, several satellite precipitation estimates (SPEs) are now available, making possible to monitor



precipitation on regular grids at the near global scale, representing an unprecedented opportunity to complement traditional precipitation measurements.

1.2 Satellite precipitation estimates (SPEs): opportunities and limitations

Several SPEs have become available in recent decades to monitor precipitation at global scale and on regular grids. The first generation of SPEs appeared with the Tropical Rainfall Measuring Mission (TRMM) launched in 1997 by NASA (National Aeronautics and Space Administration) and the Japan Aerospace Exploration Agency (JAXA). Over the last 18 years, the TRMM Multisatellite Precipitation Analysis (TMPA) (Huffman and Bolvin, 2014), the Climate prediction centre MORPHing (CMORPH) (Joyce et al., 2004), the Precipitation Estimation from remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Sorooshian et al., 2000) and the Global Satellite Mapping Precipitation (GSMaP) (GSMaP, 2012) SPE datasets have been developed based on the TRMM mission to deliver precipitation estimates at the 0.25° grid scale. In 2014, the Global Precipitation Measurement (GPM) mission was launched to ensure TRMM continuity. The second generation of SPEs based on GPM missions included the Integrated Multi-SatellitE Retrievals for GPM (IMERG) (Huffman et al., 2017) and a new GSMaP version product which deliver precipitation estimates at a finer grid scale (0.1°) than the first generation of SPEs but estimates are limited to the period from 2014 to the present. At the same time, some SPEs took advantage of previous SPEs and missions to estimate precipitation over larger time window: long-term SPE generation. This is the case of PERSIANN-Climate Data Record (PERSIANN-CDR) (Ashouri et al., 2015), Multi-Source Weighted-Ensemble Precipitation (MSWEP) (Beck et al., 2017) and Climate Hazards Group InfraRed Precipitation (CHIRP) with Station data (CHIRPS) (Funk et al., 2015).

However, SPEs are indirect measurements made from satellite/sensor constellations, including passive microwaves (PMW) and infra-red (IR) sensors on board low earth orbital (LEO) and geosynchronous satellites and are subject to uncertainty due to technical limitations. Indeed, the irregular sampling and limited overpass of LEO PMW measurements impede the correct capture of short term and slight precipitation events (Gebregiorgis and Hossain, 2013; Tian et al., 2009) which can introduce error in precipitation estimates over arid regions and/or during the dry seasons (Prakash et al., 2014; Satgé et al., 2017a; Satgé et al., 2015; Shen et al., 2010). In mountainous regions, the precipitation /no precipitation cloud classification based on cloud top IR temperature may fail in the case of precipitation processes resulting from orographic warm clouds (Dinku et al., 2010; Gebregiorgis and Hossain, 2013; Hirpa et al., 2010). The contrast between temperature and emissivity (i.e. water and snow covered area) of rough land surfaces creates background signals similar to those produce by precipitation, leading to misinterpretation between rainy or not rainy clouds, which can introduce high bias in precipitation estimates (Satgé et al., 2016; Tian and Peters-Lidard, 2007; Ferraro et al. 1998; Hussain et al. 2016; Satgé et al., 2018).

In addition to these spatial inconsistencies, the orbital satellite context implies constantly varying input data for each observation time (snapshot), which likely introduces inhomogeneity in the SPE time records. This could be exacerbated with aging sensors and permanent sensor failures. As an example, the TRMM satellite mission ended on April 8, 2015, making



unavailable the TRMM Microwave Imager (TMI) from input data used for TMPA retrieval (Huffman and Bolvin, 2016). Consequently, the potential of SPEs is expected to present space and time errors whose quantification is crucial before their use for hydro-climatic studies.

1.3 State of the art evaluation of SPEs

5 In the context described above, many studies have reported on the efficiency of SPEs over different regions. The most common way to evaluate SPE potential is to compare their estimates with precipitation gauge measurements, as reviewed by Maggioni et al. (2016) and Sun et al. (2017). The comparison of gauge-based assessment studies confirmed the spatial variability of SPE efficiency in reproducing precipitation, so no single SPE can be said to be the most effective one at global scale. For example, when TMPA, CMORPH, and PERSIANN SPE datasets were compared, TMPA was found to be closer
10 to the observed precipitation in India (Prakash et al., 2014), the Guyana shield (Ringard et al., 2015), Africa (Serrat-Capdevila et al., 2016), Chile (Zambrano-Bigiarini et al., 2017) and South America Andean plateau (Satgé et al., 2016), whereas CMORPH was closer to observed precipitation in Bali, Indonesia (Rahmawati and Lubczynski, 2017), Pakistan (Hussain et al., 2017), and China (Su et al., 2017; Zeng et al., 2018). However, these assessments based on comparison with gauge observations did not assess SPE's potential performance over unmonitored regions. This is especially true for high
15 mountainous regions where available gauge networks (generally located in the valley) cannot correctly represent the local precipitation induced by topographic effects. As a result, evaluating SPE potential over high mountainous regions remains challenging (Hussain et al., 2017; Satgé et al., 2017b).

An alternative method consists in assessing the sensitivity of hydrological models to SPEs. The efficiency of SPEs can be evaluated indirectly via their ability to generate reasonable discharge simulations at the outlet of the basin concerned.
20 Compared to gauge-based assessment studies, fewer authors have reported on hydrological sensitivity to SPEs, as reviewed in (Maggioni and Massari, 2018). For example, TMPA, CMORPH and PERSIANN datasets were compared as forcing data for hydrological modelling in Africa (Casse et al., 2015; Thiemig et al., 2013; Trambly et al., 2016) and South America (Zubieta et al., 2015). These studies provided complementary information to gauge-based assessments, offering an operational overview of SPEs for the management of water resources. However, due to the aggregation process at basin
25 scale, the potential of SPEs over specific ungauged regions remains unclear. Moreover, in these studies, the SPEs were not compared with gauge observations (Thiemig et al., 2013) or provided only a brief comparison at basin (Casse et al., 2015; Trambly et al., 2016) or gauge (Zubieta et al., 2015) scale. Therefore it is difficult to conclude on the respective advantages and limitations of using gauges or streamflow data as indicators to assess the ability of SPEs to reproduce precipitation patterns as SPEs could rank differently depending on the indicator used. For example, considering CMORPH, PERSIANN
30 and TMPA datasets over two African watersheds, TMPA showed the closest estimate in comparison with gauges for both basins (Thiemig et al., 2012) while CMORPH and TMPA provided more accurate streamflow simulations depending on the basin considered (Thiemig et al., 2013).



Whatever the selected approach (based on gauges and/or hydrological modelling), the analysis is performed using a single time window which does not assess the temporal variability of SPEs due to the acquisition process and/or aging sensors. To date, only a few studies have been conducted to observe changes in the efficiency of SPEs over time. For example, CHIRPS precipitation estimates were analysed over two distinct time windows to assess potential changes in precipitation accuracy over Cyprus and Nepal from one period to another (Katsanos et al., 2016; Shrestha et al., 2017). Similarly, Bai et al., 2018 analysed the accuracy of CHIRPS in mainland China separately for each year to assess its inter-annual variability. These studies highlighted temporal inconsistencies in CHIRPS estimates inherent to variations in the input data used for precipitation retrieval. However, only the CHIRPS SPE was considered and similar features are to be expected with other SPEs.

Today more than 20 SPEs are available from the first (TRMM), second (GPM) and long-term SPE generation (Beck et al., 2017). Nevertheless previous studies only considered SPE subsets. SPE assessment has indeed focussed on: (i) a single or a limited sample of SPEs (e.g. Cao et al., 2018; Erazo et al., 2018; Shrestha et al., 2017); (ii) transition from the previous to a new version of an algorithm for a specified SPE (TMPA-v6 to TMPA-v7 for example) (e.g. Chen et al., 2013; Melo et al., 2015; Milewski et al., 2015); and (iii) the effectiveness of the transition from the first (TRMM) to the second (GPM) generation of SPEs (e.g. Satgé et al., 2017; Sharifi et al., 2016; Wang et al., 2017). All these studies provided useful feedback related to their specific objectives but did not really help assess the respective performance of SPEs due to the small sample of SPEs considered.

For these reasons, comprehensive feedback on SPEs including space-time consistency, different indicators, insights over unmonitored regions, and a representative SPEs sample can only be acquired by backcrossing large SPE assessment studies. Even so, as each study is based on different statistical indices, spatial and temporal scales and periods, such an effort is seriously compromised.

1.4 Objectives

This paper proposes a protocol to assess the space-time consistency of a set of 12 SPEs. The efficiency of the SPEs is evaluated at various spatiotemporal scales and indicators in comparison with: (i) gauge point observations; (ii) streamflow observations via sensitivity analysis of a hydrological model in different catchments; and (iii) snow cover observed from satellite imagery via sensitivity analysis of a snow model in an unmonitored mountainous area. The aim of using different indicators is to evaluate whether the efficiency of the SPEs varies with the assessment method, and also to provide feedback on the use of SPEs over poorly monitored regions. The Lake Titicaca region was selected as study area because it includes all the specific features referred as potential limiting factors for SPEs (high mountainous massifs, large water bodies and snow covered areas) to evaluate the potential of SPEs in an extreme context in terms of the limits of sensors with respect to the orographic effect (i.e. mountains) and high temperature/ emissivity contrast (i.e. Lake Titicaca and a snow-covered region).



2. Material: study area and data

2.1. Study area

The Lake Titicaca basin is located between 14°S and 17°S and 71°W and 68°W in the northern part of the South American Andean plateau, known as the Altiplano. It extends over an area of 49,000 km². The Lake Titicaca catchment is bordered to the west and east by the two Cordilleras (Occidental and Real) and includes a few snow-covered areas. With a surface area of 8,560 km², a mean depth of 105 m (284 m max) and a water volume estimated at 903 km³ (Delclaux et al., 2007) Lake Titicaca is the main water body and source of the endorheic Altiplano hydrologic system. The lake is drained by the Desaguadero River to the south (Fig. 1).

Runoff from the lake basin and direct precipitation over the lake contribute approximately 53 and 47% of the lake water supply, respectively (Roche et al., 1992). However, studies on the regional water balance are more than 30 years old (Carmouze et al., 1977; Carmouze and Aquize, 1981; Lozada, 1985) and may not be representative of the current situation. Indeed, a recent study over the Altiplano (López-Moreno et al., 2015) found a temperature increase of between 0.15 and 0.25°C decade⁻¹ between 1965 and 2012, while Heidinger et al. (2018) reported an increase in the intensity of precipitation extremes over the period 1965–2010 in the same region. The temperature is expected to continue to increase until the end of this century while precipitation may decrease by 10 to 30% depending on the climate projections (Bradley, 2006; Minvielle and Garreaud, 2011; Urrutia and Vuille, 2009). There is thus a need to update the Lake Titicaca water balance to support efficient water management to adapt to this changing situation. However, the transboundary, economic and remote context means the sparse hydro-meteorological monitoring network prevents a consistent analysis. Indeed, precipitation presents high spatiotemporal variability due to the geomorphologic and tropical context which is poorly represented by the available monitoring network. With nearly-global scale coverage, SPEs are a promising alternative to monitor regional precipitation in space and over time.

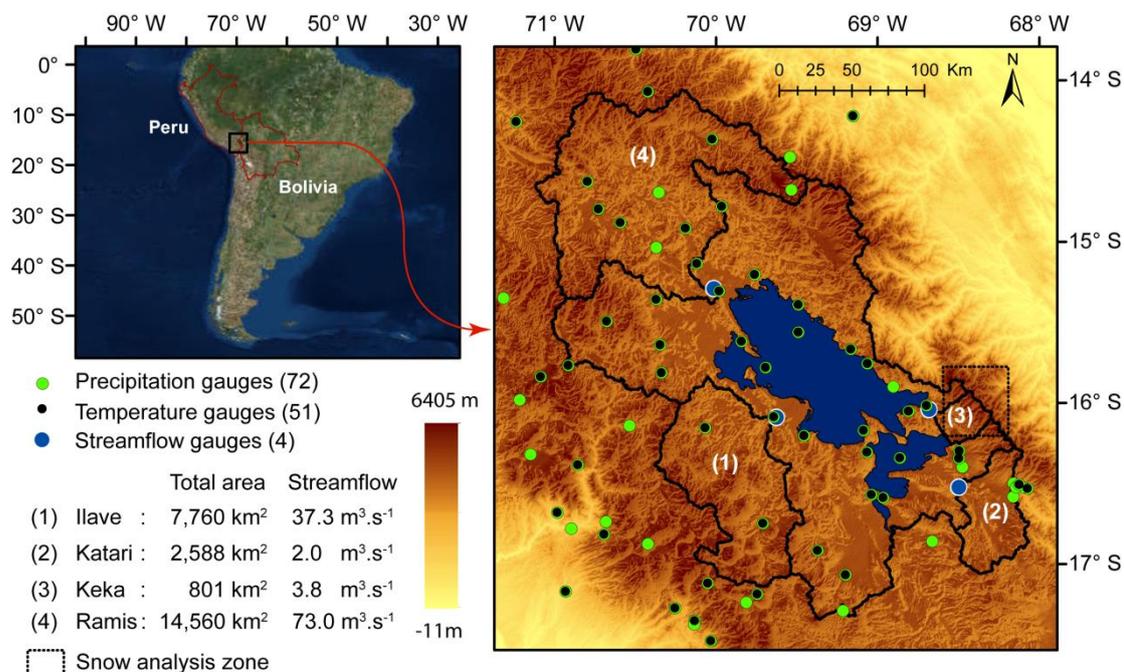


Fig. 1 Study area: location, meteorological stations (precipitation and temperature observations), streamflow gauges, studied catchments and snow analysis zone.

2.2. Hydro-climatic data

5 2.2.1. Hydro-meteorological stations

Precipitation and air temperature data for Bolivia were provided directly by the *Servicio Nacional de Hidrologia e Meteorologia* (SENAMHI) whereas for Peru, data were collected from the Peruvian SENAMHI website. Only weather gauges with less than 20% daily missing data over the 2000–2012 period were selected, giving a total of 72 stations in Bolivia and 51 in Peru (Fig. 1).

10 Water level and discharge data from SENAMHI were managed with the free software HYDRACCESS (Vauchel, 2005) developed by SO-HYBAM to obtain daily discharge records at the outlet of four catchments (Fig. 1): the Ilave (7,766 km²), the Katari (2,588 km²), the Keka (801 km²), and the Ramis (14,560 km²) catchments with respective mean annual discharge estimated at 37.3, 2.0, 3.8 and 73.0 m³.s⁻¹ (Uria & Molina, 2013). Nearly complete discharge observations over the 2000–2012 period were available for two basins (Katari and Keka), whereas discharge observations were only available from 2008
 15 to 2012 for the other two (Ilave and Ramis).

2.2.2. Interpolation of meteorological in-situ observations

To obtain continuous and spatialized meteorological series for the study area, precipitation and temperature gauge data were interpolated using the inverse distance weighted method (IDW) on a 5-km grid at the regional scale (for details on the



purpose of hydrological modelling, see section 3.2) and on a 500-m grid at the local scale (for the purpose of snow modelling, see section 3.3). The choice of the IDW technique as interpolation method was based on the study of Ruelland et al. (2008), which showed low sensitivity of hydrological models to rainfall input datasets derived using different interpolation methods (IDW, Thiessen, spline, ordinary kriging) with IDW yielding the highest hydrological efficiency. Temperature values were interpolated by accounting for a constant lapse rate of $6.5^{\circ}\text{C km}^{-1}$, in a similar way to that described in Ruelland et al. (2014). No orographic effect was accounted for in the interpolation of the point precipitation observations in order to allow for objective comparisons with the satellite products as these estimate precipitation even in the remote, mountainous areas. P_{ref} and T_{ref} , refer to interpolated precipitation and temperature, respectively.

2.3. Remote sensing data

2.3.1. Satellite precipitation estimates (SPEs)

Twelve SPEs with a spatial resolution below or equal to 0.25° (~ 25 km at the Equator) were selected for the 2000–2012 period. SPEs with coarser resolution were not used to ensure accurate representation of spatial precipitation variability. The SPEs include the following datasets: Climate Hazards Group InfraRed Precipitation (CHIRP), Climate Prediction Center MORPHing (CMORPH), Global Satellite Mapping of Precipitation (GSMaP), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network (PERSIANN), Soil Moisture to Rain (SM2Rain) method, Tropical Rainfall Measuring Mission (TRMM), Multisatellite Precipitation Analysis (TMPA) and Multi-Source Weighted-Ensemble Precipitation (MSWEP). All the SPEs used a combination of satellite (S) data gathering information from passive microwave (PMW) radiometers and infra-red (IR) data from Low Earth Orbital (LEO) and geosynchronous satellites, respectively except for the SM2RAIN method, which relies on satellite surface soil moisture derived from passive and active microwave, for precipitation estimates. The selected SPEs differ in terms of the combination of satellite sensors, algorithms and whether the products include reanalysis (R) and/or a calibration step against gauge (G) data in their processing or not. Table 1 provides an overview of these SPEs and relevant references for more information on their respective production. The mean annual precipitation pattern retrieved from all SPEs is presented in Fig. 2.

SPEs were first aggregated to obtain daily time step records using 8h to 8h (local time) time windows to match local daily gauge observations. It should be noted that some SPEs (CHIRP, CHIRPS and GSMaP) are only delivered at daily scale. For those SPEs, the daily aggregation is based on different time windows which could compromise the comparison of SPEs at daily scale. This is why we aimed to analyse SPEs at a 10-day time step for which we assumed that differences in daily temporal windows were negligible. Finally, using the nearest neighbour technique, all the SPEs were spatially resampled to 5 km to facilitate their comparison. The resulting database consists of 10,500 daily virtual stations at 5 km spatial resolution over 12 hydrological years (2000–2012) for each SPE. Additionally, SPEs were resampled to 500 m resolution over the selected subset region to assess SPE potential for snow modelling (see Fig. 1).



Full Name	Acronym	Data source	Temporal coverage	Temporal resolution	Spatial coverage	Spatial resolution	References
Climate Hazard Group InfraRed Precipitation	CHIRP	S,R	1981–present	daily	50°	0.05°	Funk et al., 2015
Climate Hazard Group InfraRed Precipitation with Station	CHIRPS	S,R,G	1981–present	daily	50°	0.05°	Funk et al., 2015
CPC MORPHing technique RAW	CMORPH–RAW	S	1998–present	3 h	60°	0.25°	Joyce et al., 2004
CPC MORPHing technique bias corrected	CMORPH–CRT	S,G	1998–present	3 h	60°	0.25°	Joyce et al., 2004
CPC MORPHing technique blended	CMORPH–BLD	S,G	1998–present	3 h	60°	0.25°	Xie et al., 2011
Global Satellite Mapping of Precipitation Reanalyse v.6	GSMaP-R	S	2000–02/2015	daily	60°	0.1°	Ushio et al., 2009
Global Satellite Mapping of Precipitation Reanalyse Gauges v.6	GSMaP	S,G	2000–02/2014	daily	60°	0.1°	Yamamoto et al., 2015 Ushio et al., 2009 Yamamoto et al., 2014
Multi-Source Weighted- Ensemble Precipitation v.2.1	MSWEP	S,R,G	1979–present	3h	Global	0.1°	Beck et al., 2017
Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks	PERSIANN	S,G	2000–present	6 h	60°	0.25°	Hsu et al., 1997 Sorooshian et al., 2000
PERSIANN-Climate Data Record	PERSIANN– CDR	S	1983–2016	6 h	60°	0.25°	Ashouri et al., 2015
TRMM Multi-Satellite Precipitation Analysis Real Time v.7	TMPA–RT	S	2000–present	3h	60°	0.25°	Huffman et al., 2010 Huffman and Bolvin, 2014
TRMM Multi-Satellite Precipitation Analysis Adjusted v.7	TMPA–Adj	S,G	1998–present	3h	50°	0.25°	Huffman et al., 2010 Huffman and Bolvin, 2014
Soil Moisture to Rain from ESA Climate Change Initiative	SM2RAIN	S	1998–2015	daily	Global	0.25°	Ciabatta et al., 2017

Table 1: Main characteristics and references of the 12 SPEs considered. In the data source column, S stands for satellite, R for reanalysis, and G for gauge information.

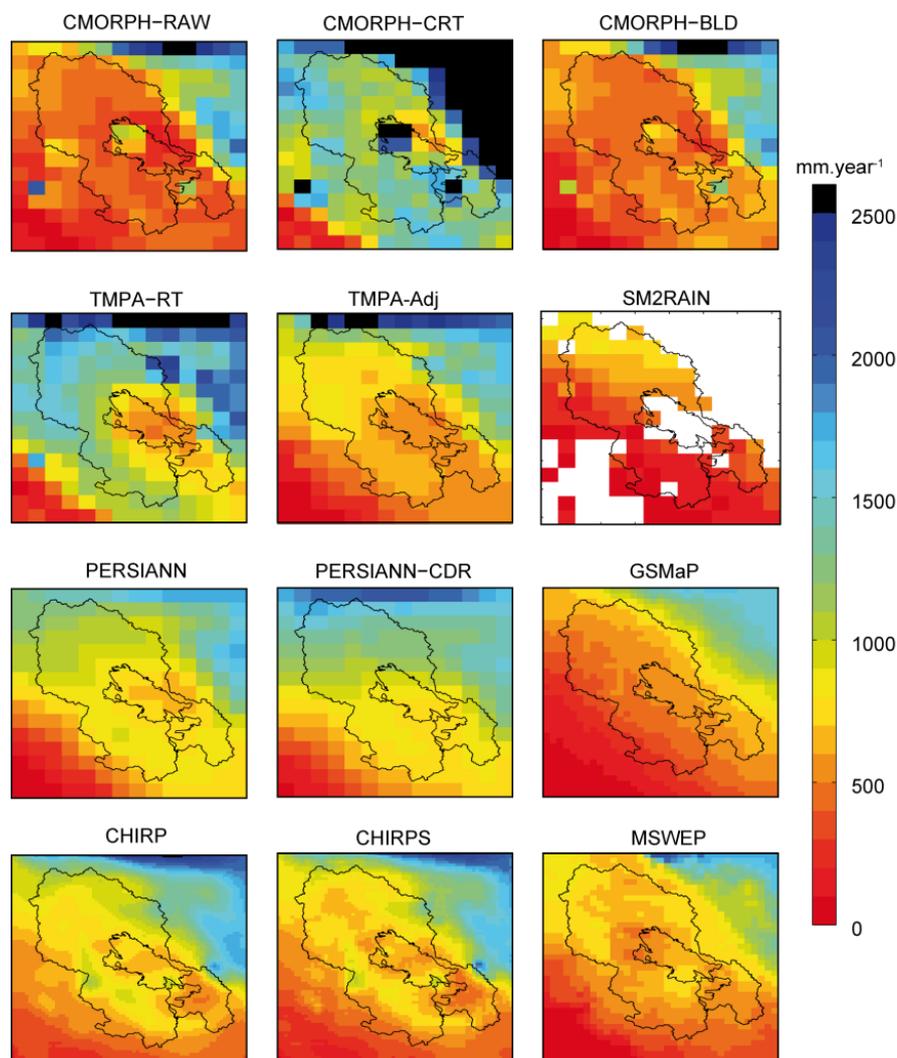


Figure 2: Mean annual precipitation maps for the 2000–2012 period retrieved from all SPEs at their original grid size. For each SPE, only the pixels with more than 80% available daily data were retained. In order to keep the regional precipitation pattern visible, black colour was used to filter pixels whose mean annual precipitation was greater than $2,500 \text{ mm}\cdot\text{year}^{-1}$

5 2.3.2. MODIS snow products

MOD10A1 (Terra) and MYD10A1 (Aqua) snow products version 5 were downloaded from the National Snow and Ice Data Center for the period 24 February 2000–6 January 2016. This corresponds to 5,795 daily values among which 5,697 are available for MOD10A1 (98.3%) and 4,918 for MYD10A1 (84.9%) since Aqua was launched in May 2002 and became operational in July 2002. These snow products are derived from a NDSI (Normalized Differential Snow Index) calculated from the near-infrared and green wavelengths, and for which a threshold has been defined for the detection of snow. Cloud cover represents a significant limit for these products, which are generated from instruments operating in the visible-near-infrared wavelengths.



5 As a result, the pixels were gap-filled so as to produce daily cloud-free snow cover maps of the study area. The different classes in the original products were first merged into three classes: no-snow (no snow or lake), snow (snow or lake ice), no-data (clouds, missing data, no decision, and saturated detector). The missing values were then filled according to a gap-filling algorithm described in Ruelland et al. (submitted). This algorithm works in three sequential steps: (i) Aqua/Terra
combination; (ii) temporal deduction by sliding the time filter up to 6 days; (iii) spatial deduction by elevation and
neighbourhood filter to gap-fill the remaining no-data pixels. The resulting database consists of 8,170 binary (snow/no-
snow) daily stations at 500 m spatial resolution for the period 2000–2012 (hereafter M_{sc}). Finally, Snow Covered
Distribution (SCD) represents the percentage of days with snow and can be retrieved for any pixels and periods.

3. A protocol to evaluate the space-time consistency of satellite precipitation estimates

10 Figure 3 is a flowchart of the main methodological steps. Twelve SPEs were first considered as a representative sample of
currently available SPEs. This is an important consideration to guide potential SPE users towards the most efficient SPE.
However, to avoid overloading the research with unnecessary results, a pre-selection was made at the Titicaca Lake
catchment scale (hereafter denoted regional scale) to discard useless SPEs. The remaining SPEs were then assessed using
15 three successive and complementary methods. The first assessment step consisted of comparing SPEs and gauge
observations at the locations of the 69 pixels which included gauges. The second assessment step consisted of analysing the
sensitivity of streamflow modelling to the SPEs at the four basin outlets using observed streamflow as reference data. The
third step consisted of analysing the sensitivity of snow modelling to the SPEs (precipitation datasets) over a subset
mountainous area in the Andes (see Fig. 1) using snow cover duration as observed from MODIS gap-filled snow products as
reference. The aim of the proposed protocol was to assess the space-time consistency of the SPEs. The spatial consistency
20 was accounted for by the gauge and basin distribution, as well as by the use of a fully distributed snow model for the first,
second and third assessment step, respectively. To evaluate the consistency of the SPEs over time, each assessment step was
analysed according to three 4-year time windows. It is worth mentioning that simulated streamflow and snow cover using
 P_{ref} as forcing data were also used for comparison with SPE simulations. More details of the proposed protocol are presented
in the following sections.

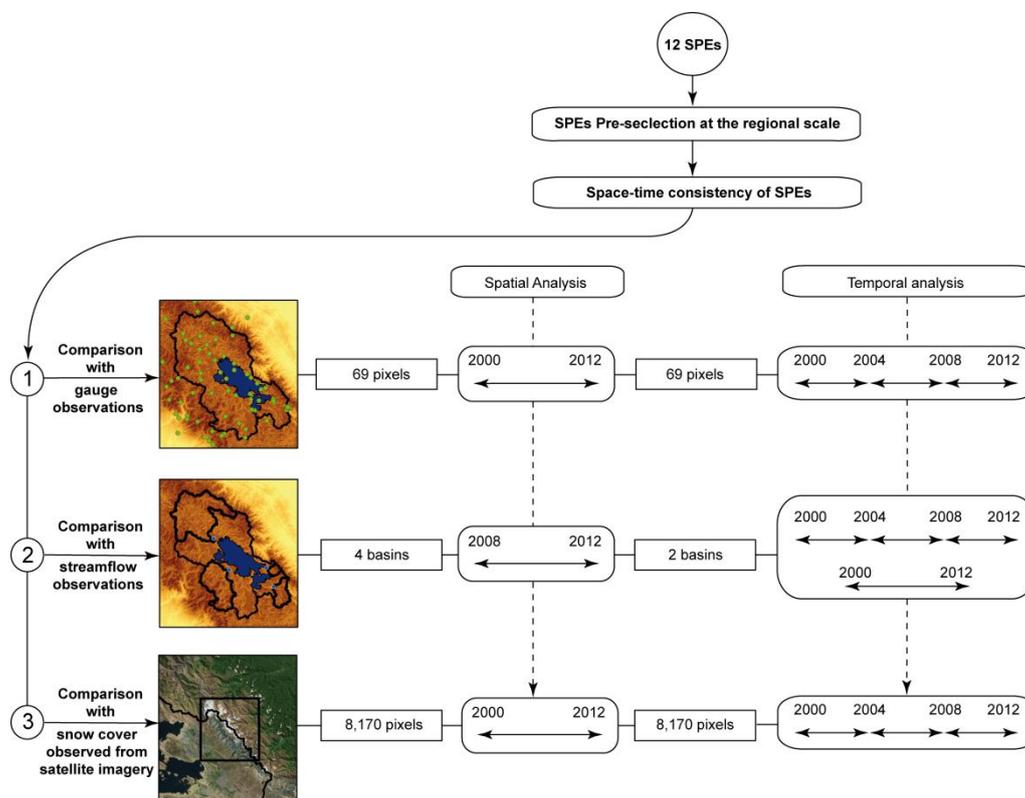


Figure 3: Flowchart of the main steps from SPE preselection to successive assessment approaches including (1) comparison between SPEs and gauge observations, (2) sensitivity analysis of runoff modelling to the various SPEs at the four basin outlets and (3) sensitivity analysis of snow modelling to the various SPEs over a subset mountainous area in the Andes.

3.1 Comparison of SPEs with gauge observations: pre-selection and evaluation

To avoid overloading the analysis with unsuitable SPEs, SPE consistency was first analysed at the regional scale for the 2000–2012 period. For each of the 69 0.05° pixels including at least one precipitation gauge, the daily precipitation series from the gauges and SPEs were first aggregated to the 10-day time step over the 12-year period (2000–2012). The other pixels were not used as P_{ref} is influenced by the interpolation process. For each of the 69 0.05° pixels, only the daily values available from all precipitation datasets (P_{ref} and SPEs) were used and the 10-day records were only computed when more than 80% of daily values records were available. Next, mean regional 10-day precipitation series were computed from P_{ref} and all SPEs by aggregating the values from all 69 pixels. For each pixel, only 10-day values available for all SPEs and P_{ref} were used. It should be noted that SM2Rain estimates rely on soil moisture observations with many missing data over water bodies and mountainous regions (Dorigo et al., 2015), leading to significant spatial gaps over Lake Titicaca and the Cordillera region (Fig. 3). As a result, in comparison to other SPEs, only 44 P_{ref} 0.05° pixels (including gauges) were



available for SM2Rain. Therefore, SM2Rain was analysed separately from other SPEs by computing additional P_{ref} and SM2Rain mean regional 10-day precipitation series based on the 44 available P_{ref} 0.05° pixels.

5 Mean regional 10-day SPEs and P_{ref} series for the 12-year period 2000–2012 were compared according to different statistical criteria, correlation coefficient (CC), standard deviation (STD), percentage bias (%B) and the centred root mean square error (CRMSE) (Eq. 1–4):

$$CC = \frac{Cov(SPE, P_{ref})}{STD_{SPE} \times STD_{ref}} \quad (1)$$

where CC is the correlation coefficient, SPE and P_{ref} are the SPE and P_{ref} precipitation time series, and Cov is the covariance.

10

$$STD = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - \bar{P})^2} \quad (2)$$

where STD is the standard deviation in mm, n is the number of values; and P the precipitation value in mm (SPE or P_{ref}).

$$15 \quad \%B = \frac{\frac{1}{n} \sum_{i=1}^n (SPE_i - P_{ref_i})}{\frac{1}{n} \sum_{i=1}^n P_{ref_i}} \times 100 \quad (3)$$

where %B is the SPE bias value as a percentage, n is the number of values, SPE the precipitation estimate of the considered SPE value in mm; and P_{ref} the reference precipitation value in mm.

$$20 \quad CRMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n ((SPE_i - \overline{SPE}) - (P_{ref_i} - \overline{P_{ref}}))^2} \quad (4)$$

where CRMSE is the centred root mean square error in mm, n is the number of values, SPE the precipitation estimate of the considered SPE value in mm; and P_{ref} the reference precipitation value in mm.

25 To facilitate interpretation of the statistical results, the Taylor diagram (Taylor, 2001) was used to present obtained CC, and normalized values of STD and CRMSE. Normalization was performed by dividing SPEs CRMSE and STD by P_{ref} STD. Therefore, in the Taylor diagram, the reference (black dot) corresponds to CRMSE, STD and CC values of 0, 1 and 1, respectively. Also in the Taylor diagram, the position of the SPEs relative to the reference dot is an integrated indicator (CRMSE, STD, and CC) of SPE efficiency in reproducing gauge precipitation. The shorter the distance between the SPEs and the reference position, the closer the SPEs and P_{ref} estimates. Additionally, %B values were used to observe the potential over/underestimation of each SPE considered. For the following assessment step, only the six SPEs most efficient at the regional scale for the 2000–2012 were considered.

30

To assess the consistency of SPEs over time, mean regional precipitation 10-day series were compared to P_{ref} according to a Taylor Diagram for three 4-year periods corresponding to the 2000–2004, 2004–2008 and 2008–2012 periods,



respectively. For all the periods considered, 98.6% of P_{ref} pixels had more than 90% 10-day records for both the wet and dry season. Consequently, the temporal assessment of the SPEs was not expected to be influenced by any inconsistency of P_{ref} over time in terms of available records.

To assess the spatial consistency of the SPEs, we compared the CC, CRMSE and %B computed between the SPEs and P_{ref} 10-day series for the 2000–2012 period at the location of each pixel which included gauges. For each SPE, CC, CRMSE and %B obtained at the pixel level were plotted to highlight regions potentially concerned by low (high) SPE potential. We repeated the analysis for the three 4-year windows. We only considered CRMSE scores to simplify interpretation of the results, as this score was found to provide more statistical discrimination than CC, STD and %B. It is worth mentioning that resampling SPEs (see section 2.3.1) could affect the assessment of SPE potential at the pixel level. Indeed, at a coarser spatial resolution (0.25°) areal precipitation estimates of SPE pixels are expected to be more representative of the mean precipitation derived from all the gauges in the pixel considered than the one derived from a single gauge. However, for these particular pixels, a preliminary SPE assessment at the original and resample grid size revealed no significant differences (data not shown).

For each SPE, CRMSE values obtained at the pixel level and for each period were plotted to highlight regions possibly concerned by low (high) SPE potential.

Finally, for each pixel, mean and STD CRMSE values were computed from the three CRMSE values obtained from the three 4-year periods. Mean and STD CRMSE values were then plotted to assess the consistency of SPEs in space and over time. The lower the mean and STD of CRMSE, the more stable the SPE considered at the specific pixel location.

3.2. SPEs as input data for hydrological modelling

The GR4j lumped hydrological model (Perrin et al., 2003) was chosen to analyse the sensitivity of streamflow simulations to the various SPEs. The model has demonstrated its ability to perform well under various hydro-climatic conditions (e.g., Coron et al., 2012; Perrin et al., 2003; Grouillet et al., 2016; Dakhlaoui et al., 2017), notably in the Andean region (e.g., Hublart et al., 2016).

This model relies on daily precipitation (P) and potential evapotranspiration (PE), which was computed using the formula proposed by Oudin et al. (2005) (Eq. 5):

$$PE = \frac{R_e T + 5}{\lambda \rho 100} \quad \text{if } (T + 5) > 0 ; \quad \text{else } PE = 0 \quad (5)$$

where PE is daily potential evapotranspiration (mm), R_e is extra-terrestrial solar radiation ($\text{MJ.m}^{-2}.\text{d}^{-1}$), which depends on the latitude of the target point and the Julian day of the year, λ is the net latent heat flux (fixed at 2.45 MJ.kg^{-1}), ρ is water density (fixed at 11.6 kg.m^{-3}) and T is the daily mean air temperature ($^\circ\text{C}$) estimated at the target point by interpolating the gauge observations while correcting for elevation.



Firstly, a production module computes the amount of water available for runoff, i.e., “effective precipitation”. To do so, a soil-moisture accounting (SMA) store is used to separate the incoming precipitation into storage, evapotranspiration and excess precipitation. At each time step, soil drainage is computed as a fraction of the storage and added to excess precipitation to form the effective precipitation. Secondly, a routing function split the effective precipitation into two components: 90% is routed as delayed runoff through an unit hydrograph UH1 in series with a non-linear routing storage while the remaining 10% is routed as direct runoff through an unit hydrograph UH2 (Perrin et al., 2003). UH1 and UH2 consist in a slow and quick routing path, respectively, to account for differences in runoff delays. Finally, the streamflow at the catchment outlet is computed by summing up delayed and direct runoff. This model relies on four calibrated parameters: maximum capacity of the soil moisture accounting store (X1, mm), inter-catchment exchange coefficient (X2, mm), maximum capacity of routing storage (X3, mm), and time base for unit hydrographs (X4, days) for each catchment. Acceptable parameter bounds were defined according to the recommendation of Perrin et al., (2003) and previous experiments in a similar Andean context (Hublart et al., 2016; Ruelland et al., 2014). The following ranges were used for model calibration with all the precipitation datasets tested (P_{ref} and SPEs): $10 \text{ mm} < X1 < 1,800 \text{ mm}$, $-5 \text{ mm} < X2 < 5 \text{ mm}$, $1 \text{ mm} < X3 < 500 \text{ mm}$, $0.5 \text{ days} < X4 < 5 \text{ days}$. These ranges are large enough to compensate for the differences in the different precipitation datasets and basins. Beyond these ranges, we assumed that streamflow simulations could not be realistic. In practice, parameter bounds were rarely reached when calibrating the model with the different datasets tested (data not show here for the sake of brevity).

The area catchment P and PE averaged values were computed from the 5-km pixels P (P_{ref} and SPEs) and PE included in each catchment considered. We used a weighted average based on the 5 km*5 km fraction included in the catchment considered. P_{ref} and SPEs were used sequentially as forcing precipitation datasets for the streamflow simulation. For each run, model parameter calibration was based on the shuffled complex evolution (SCE) algorithm (Duan et al., 1992) by optimising the Nash-Sutcliffe efficiency criterion (NSE, Eq. 6; Nash and Sutcliffe, 1970) at a 10-day time step. The NSE criterion represents the overall agreement of the shape of the hydrograph, while placing more emphasis on high flows. NSE values vary from $-\infty$ to 1 with a maximum score of 1 meaning a perfect agreement between the observed and simulated values. At the contrary, negative values mean that more realistic estimates are obtained using the observed mean values rather than the simulated ones.

$$NSE = 1 - \frac{\sum_{t=1}^N (Q_{obs}^t - Q_{sim}^t)^2}{\sum_{t=1}^N (Q_{obs}^t - \bar{Q}_{sim})^2} \quad (6)$$

where NSE Nash-Sutcliffe efficiency, Q_{obs}^t and Q_{sim}^t are, respectively, the observed and simulated streamflow for time step t ; and N is the number of time steps for which observations are available.

The distribution of the basin in the study region (Fig. 1) provided the opportunity to assess the hydrological consistency of the SPEs in space by running GR4j in the four basins over the common 2008–2012 period of observed discharge



availability. The hydrological consistency of the SPEs over time was then evaluated by running the model over the entire 2000–2012 period for which discharge observations were only available in two catchments (Katari and Keka). For each precipitation input (P_{ref} and SPEs), the model was calibrated against observed streamflow over the whole period (2000–2012) and over three 4-year sub-periods (2000–2004, 2004–2008, and 2008–2012). No validation step was used as the objective was to assess hydrological modelling sensitivity to various precipitation datasets (P_{ref} and SPEs) and not to assess the hydrological model robustness under climate variability. The aim of the analysis of the streamflow simulation accuracy among the basins and periods considered was thus to evaluate the strength of the SPEs in space and over time and discrepancies in reproducing streamflow.

3.3 SPEs as input data for snow modelling in the Andes

A distributed, degree-day model (Ruelland et al., submitted) was chosen to analyse the sensitivity of snow cover simulations to the SPEs. This storage-based model relies on daily, distributed precipitation (P), temperature (T) and potential evaporation (PE) (Eq.7) to represent the main snow accumulation and ablation (sublimation and melt) processes (see Fig. 4). It operates at a daily time step according to a grid of 500 x 500 m corresponding to the spatial resolution of the MODIS data.

Snow accumulation is defined using a temperature threshold T_s fixed at 0°C. Sublimation is accounted for based on daily PE_{sub} (mm) at the target pixel and snowmelt is controlled with a melt factor parameter K_f (°C⁻¹d⁻¹ to be calibrated) according to Eq. (8-9).

$$PE_{sub} = PE \times K_{sub} \quad (7)$$

where PE_{sub} is potential evapo-sublimation, PE is potential evaporation (see Eq. 5); and K_{sub} is a proportional coefficient depending on the mean latitude (lat, decimal degrees) of the study area and varying from 0 to the poles to 1 at the equator (see Ruelland et al., submitted for more details)

$$Mf = K_f \times (T - T_s) \quad (8)$$

$$Melt = \begin{cases} 0 & T_j \leq T_s \\ \text{Min}(SWE, Mf) & T_j > T_s \end{cases}; \quad (9)$$

where Mf is the potential melt (mm), K_f is a melt factor parameter to be calibrated, T is the temperature on day j on the pixel considered, T_s is the threshold temperature parameter (fixed at 0°C) for snow accumulation and melt. $Melt$ cannot exceed the snow water equivalent (SWE) of the snowpack storage.

The snow-covered areas (SCA) are estimated from the SWE . For each pixel, snow is stored in a reservoir which represents the SWE of the pixel snowpack (see Fig. 4). It is fed solely by the solid fraction of precipitation and is emptied according to the simulated sublimation and melt processes. For each model, a pixel is assigned to snow or not depending on a water level threshold SWE_{th} (mm), to be calibrated.



5

As the region contains permanent snow covered areas, for pixels located below and above 5,700 m asl, the *SWE* reservoir was initialized to 0 mm and 300 mm, respectively, at the beginning of the simulations. These values were defined based on MODIS snow observations and on model sensitivity tests to *SWE* initial conditions accounting for different elevation thresholds. The analysis revealed limited sensitivity to initial conditions (data not shown). An initial 3-year warm-up period was used for each simulation to limit the influence of these conditions.

10

The following ranges were used for model calibration with all the tested precipitation datasets (P_{ref} and SPEs): $0.5 \text{ mm} \cdot \text{C}^{-1} \cdot \text{d}^{-1} < K_f < 20 \text{ mm} \cdot \text{C}^{-1} \cdot \text{d}^{-1}$, $1 \text{ mm} < SWE_{th} < 80 \text{ mm}$. K_f ranges were based on ranges adapted from the values reviewed in Hock (2003). Regarding the SWE_{th} value, the assumption is that, using remote sensing, snow cover cannot be detected below a certain threshold (Bergeron et al., 2014). For instance, based on in-situ measurements to detect snow cover from MODIS in the Pyrenees, Gascoin et al., (2015) found a mean threshold of 40 mm. Since this value may be influenced by the local context (vegetation, topography, and climate) and spatial difference between point (in-situ) and areal satellite (MODIS) observations, SWE_{th} was tested according to large ranges. It is worth mentioning that the tested bounds were reached during calibration for all simulations (i.e. $K_f=20 \text{ mm} \cdot \text{C}^{-1} \cdot \text{d}^{-1}$ and $SWE_{th}=1 \text{ mm}$). However we did not consider larger parameter ranges to ensure “realistic” simulations.

15

In association with T_{ref} for temperature forcing data (see section 2.2.), P_{ref} and SPEs were sequentially used as forcing precipitation datasets to simulate snow cover with the model. For each run, model parameters were calibrated based on the shuffled complex evolution (SCE) algorithm by optimising the pixel-to-pixel correlation between the snow cover duration (SCD) simulated by the model and that observed by the gap-filled MODIS snow products (see section 2.3.2.), according to the following Eq. (10):

$$R^2 = 1 - \frac{\sum_{p=1}^n (SCD_{MODIS(p)} - SCD_{MODEL(p)})^2}{\sum_{p=1}^n (SCD_{MODIS(p)})^2}; \quad (10)$$

20

where R^2 is a determination coefficient, n is the total number of pixels in the study area (see Fig. 1); p is a given pixel; SCD_{MODIS} and SCD_{MODEL} are, respectively, the snow cover duration (SCD) observed by MODIS and the SCD simulated by the model as a percentage of days over the analysis period.

25

For each precipitation input (P_{ref} and SPE), the model was calibrated against MODIS observed snow cover over the entire period 2000–2012 and over three 4-year sub-periods (2000–2004, 2004–2008, and 2008–2012). The aim of the analysis of the snow simulation accuracy among the areas and periods considered was to evaluate the strength of SPEs in space and over time and to identify discrepancies in reproducing snow cover in a remote Andean area (see Fig. 1).

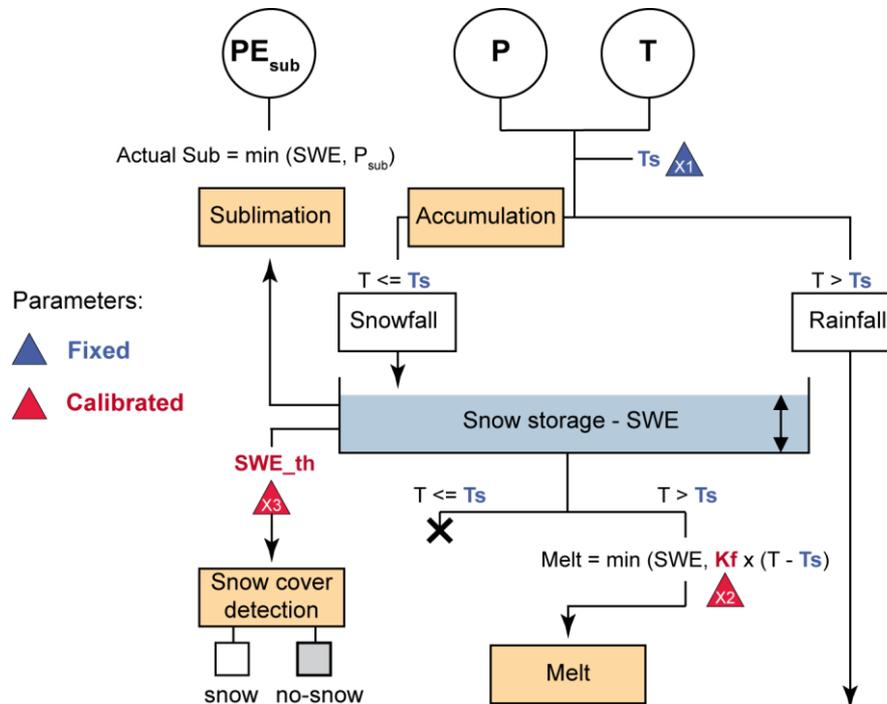


Figure 4: Distributed, degree-day model used in the study (Ruelland et al., submitted).

4. Results

4.1. Space-time consistency of SPEs compared with gauge observations

5 Figure 5a is the Taylor diagram obtained from SPEs at the regional scale for the 2000–2012 period. SPEs are distributed over the Taylor diagram, indicating high discrepancy among them. Some SPEs greatly overestimated precipitation with %B values of 165%, 75%, 40% and 45% for CMORPH–CRT, TMPA–RT, PERSIANN and PERSIANN–CDR, respectively. CMORPH–RAW, GSMaP and SM2Rain, greatly underestimated precipitation with %B values of -24%, -25% and -37%, respectively. However, all SPEs were highly correlated with P_{ref} with CC greater than 0.75. Generally, including gauge data in the SPE processing clearly enhanced precipitation estimates. Indeed, CHIRPS, CMORPH–BLD and PERSIANN–CDR were closer to the reference dot than their respective non-adjusted version, CHIRP, CMORPH–CRT and PERSIANN. With the closest and farthest distance to the reference, MSWEP and CMORPH–CRT were respectively the most and least consistent SPEs to represent the mean regional precipitation over the 2000–2012 period.

15 According to the literature, a quality threshold value can be used to express SPE potential. Some authors (e.g., Hussain et al., 2017; Satgé et al., 2016; Shrestha et al., 2017), considered that a normalized RMSE value lower than 0.5 is associated with a very good SPE performance. Even there were slight differences between CRMSE and RMSE, the use of a normalized CRMSE threshold value of 0.5 to select only the most efficient SPEs remains logical. Therefore, six SPEs were selected for the following assessment steps including CHIRP, CHIRPS, CMORPH–BLD, PERSIANN–CDR, TMPA–Adj, and MSWEP.



Figure 5b presents the performance of the SPEs at the regional scale for the three 4-year time windows considered: 2000–2004, 2004–2008 and 2008–2012, in the form of a Taylor diagram. At the regional scale, SPE rank performance remained stable during the time windows considered and similar to what was observed for the 2000–2012 period (Fig. 5). Therefore, at the regional scale, SPEs were generally consistent over time, MSWEP being the most accurate and PERSIANN-CDR the least accurate SPE. However, for the 2008–2012 period, CHIRPS and CHIRP were closer than for the previously considered period. This might be due to the decrease in the number of available gauges for the adjustment processes applied on CHIRP to produce CHIRPS.

Figure 5c shows the spatial distribution of SPE errors for the 2000–2012 period in terms of %B, CC, and CRMSE. CMORPH-BLD was poorly correlated with P_{ref} with the highest proportion of pixels with CC less than 0.7. This value is generally used as a quality threshold with CC values less than 0.7 indicating poor SPE performance (see e.g., Satgé et al., 2016). MSWEP had the best CC value overall with the highest proportion of pixels with a CC greater than 0.9 and only two pixels with unsatisfactory CC values.

The main differences were in %B and CRMSE. CMORPH-BLD and PERSIANN-CDR precipitation under- and over-estimations at the regional scale (Fig. 5a) were confirmed at the gauge scale (Fig. 5c). CHIRP, CHIRPS, and TMPA-Adj presented similar %B distribution while MSWEP had the most homogeneous %B distribution with the values of almost all pixels ranging between -30% and +30%. This range was previously defined as satisfactory %B for SPEs (Shrestha et al., 2017). The gauge adjustment applied on CHIRP was globally positive with a %B reduction from CHIRP (15.9%) to CHIRPS (0.4%) of almost 100% (Fig. 4a). It considerably increased the numbers of pixels with %B between -15% and +15% from CHIRP to CHIRPS (Fig. 5b) and generally enhanced CRMSE and CC scores.

Interestingly, all the SPEs underestimated precipitation for the two pixels located over the northern Lake Titicaca islands. This is probably linked to SPE's limited ability to detect warm cloud precipitation (see section 5.1).

CMORPH-BLD, PERSIANN-CDR, and TMPA-Adj had the highest proportion of pixels with CRMSE values greater than 0.7 (Fig. 5c). This value can be used as a quality threshold above which SPEs performance is considered as unsatisfactory (see e.g. Shrestha et al., 2017). Therefore, over the 2000–2012 period, precipitation estimates derived from CMORPH-BLD, PERSIANN-CDR and TMPA-Adj are subject to high uncertainties at local scale. The inclusion of gauge observations for CHIRPS estimates reduced the number of pixels with unsatisfactory performance by 50% in comparison with the non-adjusted CHIRP version. With only 14 and 12 pixels with CRMSE greater than 0.7, respectively, CHIRPS and MSWEP showed the highest spatial consistency.

More generally, the spatial analysis highlighted three areas in which all the SPEs considered presented less satisfactory statistical scores (CRMSE, CC, and %B) than over the remaining areas. These regions correspond to the south eastern Lake Titicaca shore, and the south-western and north-eastern borders of the catchment.

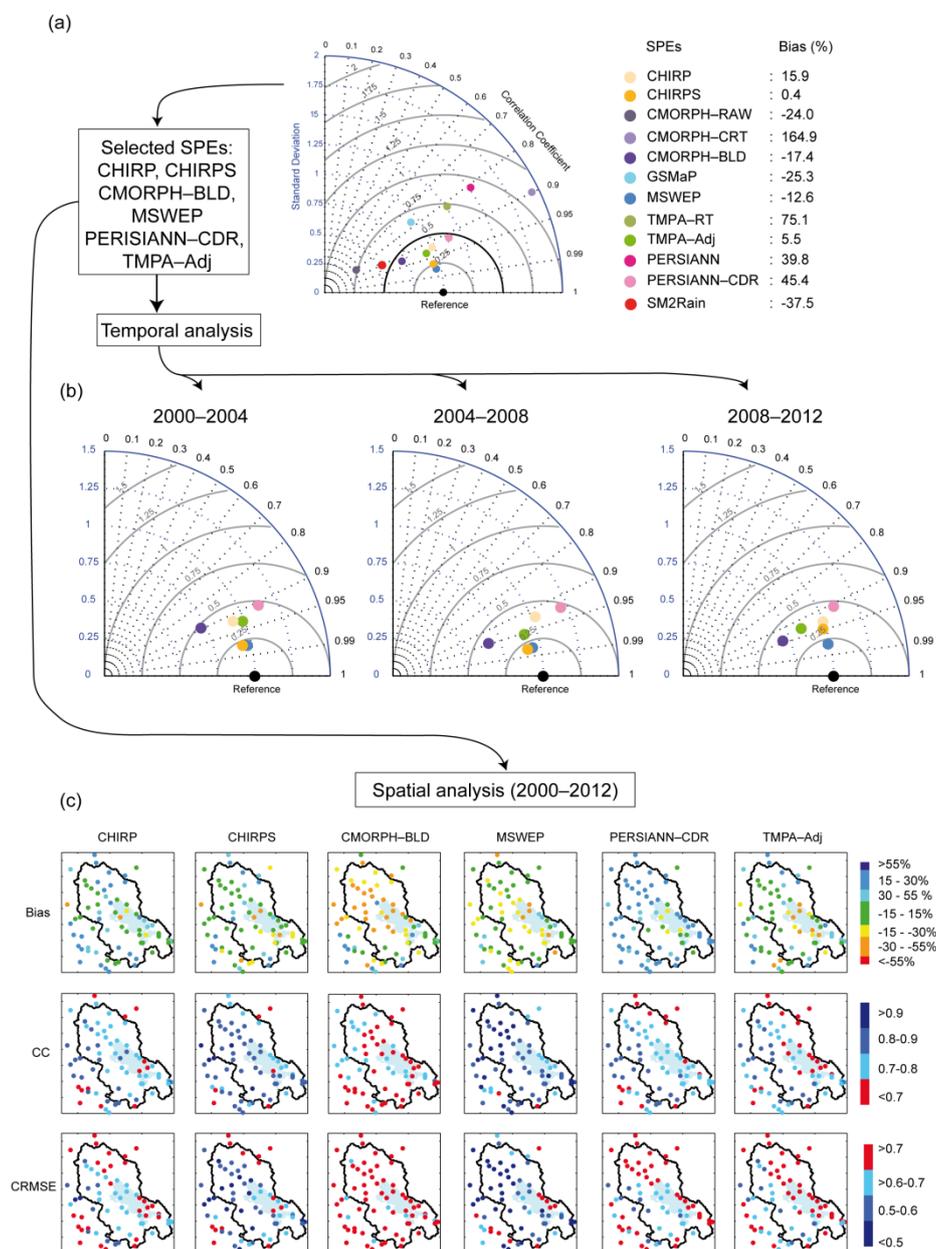


Figure 5: Efficiency of SPEs compared with gauge observations: (a) Pre-selection of SPEs at the regional scale in the form of a Taylor diagram; (b) Consistency of SPEs over time at the regional scale in the form of a Taylor diagram; and (c) Consistency of SPEs in space for the 2000–2012 period from CC, CRMSE and %B values obtained from each P_{ref} pixel including at least one gauge.

Figure 6 shows the performance of the SPEs for three 4-year periods (2000–2004, 2004–2008, and 2008–2012) in terms of CRMSE. The SPEs' lowest potential over the south-eastern shore of Lake Titicaca and at the south-western and north-eastern borders of the catchment highlighted in the 2000–2012 period was confirmed for all four time windows. For each



time window, the CRMSE distribution differed depending on the SPE considered. More interestingly, the distribution of CRMSE of all the SPEs differed depending on the time window considered. As a result, the efficiency of the SPEs varied over time for a specific region (at the pixel scale). As a clear example, for the CMORPH–BLD south-eastern located pixels, the CRMSE values changed drastically over time: these pixels presented unsatisfactory CRMSE scores (above 0.7) for the 2000–2004 period and satisfactory CRMSE scores (below 0.5) for both the 2004–2008 and the 2008–2012 periods.

In this context, Fig. 6d represents the space-time consistency of the SPEs over the 2000–2012 period. The mean and STD CRMSE values obtained at the location of each pixel in the three sub-periods considered are plotted. Overall, all SPEs are stable in space and over time with STD CRMSE values below 0.25 but their consistency differed in accuracy. CMORPH–BLD provided stable but not accurate precipitation estimates with mean CRMSE values systematically above 0.7. In contrast, MSWEP and CHIRPS presented stable and accurate precipitation estimates with many pixels with a mean CRMSE below 0.5. CHIRP space-time consistency was lower than that of CHIRPS, thereby confirming the advantage of the gauge calibration (Fig. 6d). PERSIANN–CDR and TMPA–Adj were more consistent in space and over time over the southern mid and western region, respectively. Interestingly, the close potential precipitation estimates observed for CHIRP and CHIRPS at the regional scale and for the 2008–2012 period were confirmed at the pixel level with similar spatial error distribution for both SPEs. Therefore gauge adjustment appears to be less efficient for the 2008–2012 than for the 2000–2004 and 2004–2008 periods.

MSWEP and CHIRPS were the most stable SPEs in space and over time with the highest proportion of pixels with a mean and a STD CRMSE below 0.5 and 0.25, respectively (Fig. 6d). However, for pixels located on the south-eastern shore of the lake, CHIRPS provided more accurate and stable precipitation estimates than MSWEP and all the other SPEs.

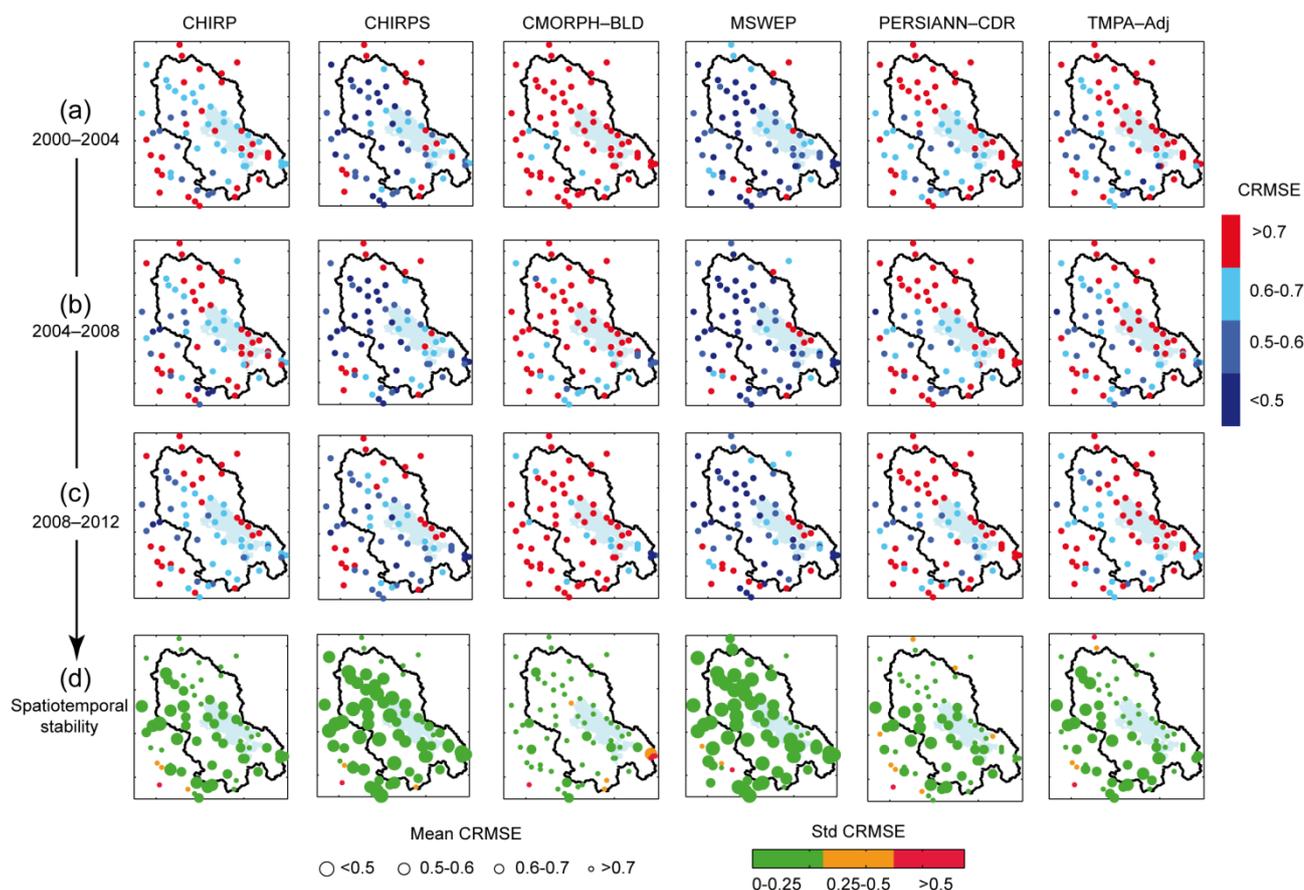


Figure 6: Consistency of SPEs in space and over time compared with gauge observations. CRMSE values obtained from the SPEs at the gauge level for three 4-year periods: (a) 2000–2004; (b) 2004–2008; and (c) 2008–2012. (d) Mean and STD CRMSE values obtained for each pixel using the three respective values obtained for the three 4-year periods.

5 4.2. Space-time consistency of SPEs compared with streamflow simulations

Figure 7 shows the efficiency (in terms of NSE scores) of the hydrological model in reproducing streamflow at the outlet of the four tested basins (Ilave, Katari, Keka, and Ramis) using P_{ref} and SPEs as input data over the period 2008–2012.

Streamflow simulations using P_{ref} varied along the catchments with lowest NSE score of 0.63 for Keka and the highest of 0.89 for Katari. Simulated SPE streamflow efficiency followed the same trend except with CMORPH-BLD and TMPA-Adj. TMPA-Adj provided the best streamflow simulation for the Ilave catchment and the worst for Keka, whereas the opposite was observed using CMORPH-BLD as forcing data. The low efficiency of CMORPH-BLD in the Ilave catchment was related to its erroneous streamflow peaks in the 2009 and 2012 dry seasons. CHIRP and CMORPH-BLD had the lowest scores for the Katari and Ramis catchments with NSE values of 0.59 and 0.45, respectively. For all the catchments, streamflow simulations based on CHIRPS presented systematically higher NSE scores than simulations based on CHIRP, showing that the adjustment provided in CHIRPS with the integration of gauge observations led to better precipitation



estimates. This confirms the enhancement of the gauge-based assessment observed with an overall reduction in bias and an increase in CRMSE and CC.

For all the catchments considered, the best streamflow simulations were obtained with at least one of the SPEs as forcing precipitation data. TMPA–Adj and CMORPH–BLD provided a better streamflow simulation than P_{ref} for Ilave and Keka respectively, and MSWEP outperformed P_{ref} for the Katari and Ramis catchments. These results show that SPEs can efficiently replace the currently available sparse precipitation gauge networks for use in hydrological studies of the region. Overall, MSWEP appears as the most consistent SPE product for streamflow simulations. Indeed, the streamflow simulations forced by MSWEP were more realistic than those forced by P_{ref} over three catchments (Katari, Keka, and Ramis) and were almost the same for the Ilave catchment.

However, as shown in Fig. 7c, the SPE hydrological ranking in the 2008–2012 period changed drastically over time. For example, for the Katari catchment, MSWEP led to the best streamflow simulations for the 2004–2008 and 2000–2012 but not for the 2000–2004 period, for which TMPA–Adj forced streamflow simulations had a higher NSE score of 0.85. Additionally, CMORPH–BLD potential fell drastically over the 2000–2004 period with a negative NSE score, whereas it produced the most realistic streamflow simulation for the period 2008–2012. In the Keka catchment, for each time window, the best streamflow simulation was obtained using different SPEs. CHIRP, PERSIANN–CDR and CMORPH–BLD resulted in the highest NSE scores over the various sub-periods analysed, with respectively 0.73 for the period 2000–2004, 0.83 for 2004–2008 and 0.77 for 2008–2012.

To conclude, the analysis of streamflow model sensitivity to the different precipitation datasets depended to a great extent on the time window and catchment considered. However, MSWEP appeared to be the most ‘stable’ SPE. It provided the most realistic streamflow simulations with higher NSE scores than P_{ref} for seven out of ten simulations and also outperformed other SPEs in almost all streamflow simulations tested. The results obtained for the 2000–2012 period faithfully reflected global SPE performance for all the time windows with MSWEP as the most and CMORPH–BLD as the least suitable SPE for hydrological modelling over the Katari and Keka catchments.

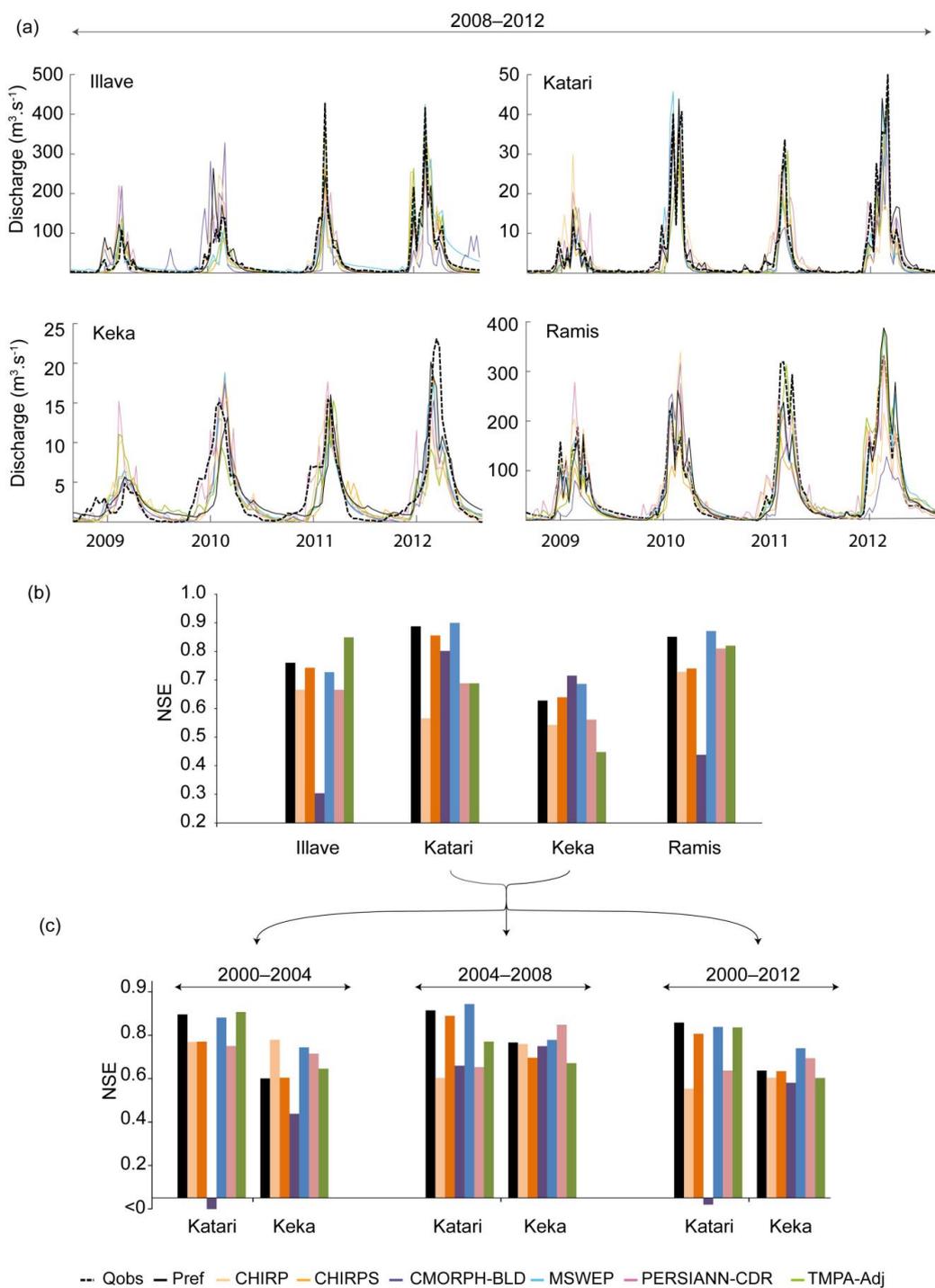


Figure 7: Observed versus simulated streamflow using P_{ref} and SPEs as input data in the hydrological model: (a) 10-day streamflow simulations at the four basin outlets (Ilave, Katari, Keka and Ramis) over the 2008–2012 calibration period; (b) corresponding efficiency (NSE scores) of simulated versus observed streamflow; (c) efficiency (NSE score) of simulated versus observed streamflow over the 2004–2004, 2004–2008 and 2000–2012 calibration periods at two basin outlets (Katari and Keka).

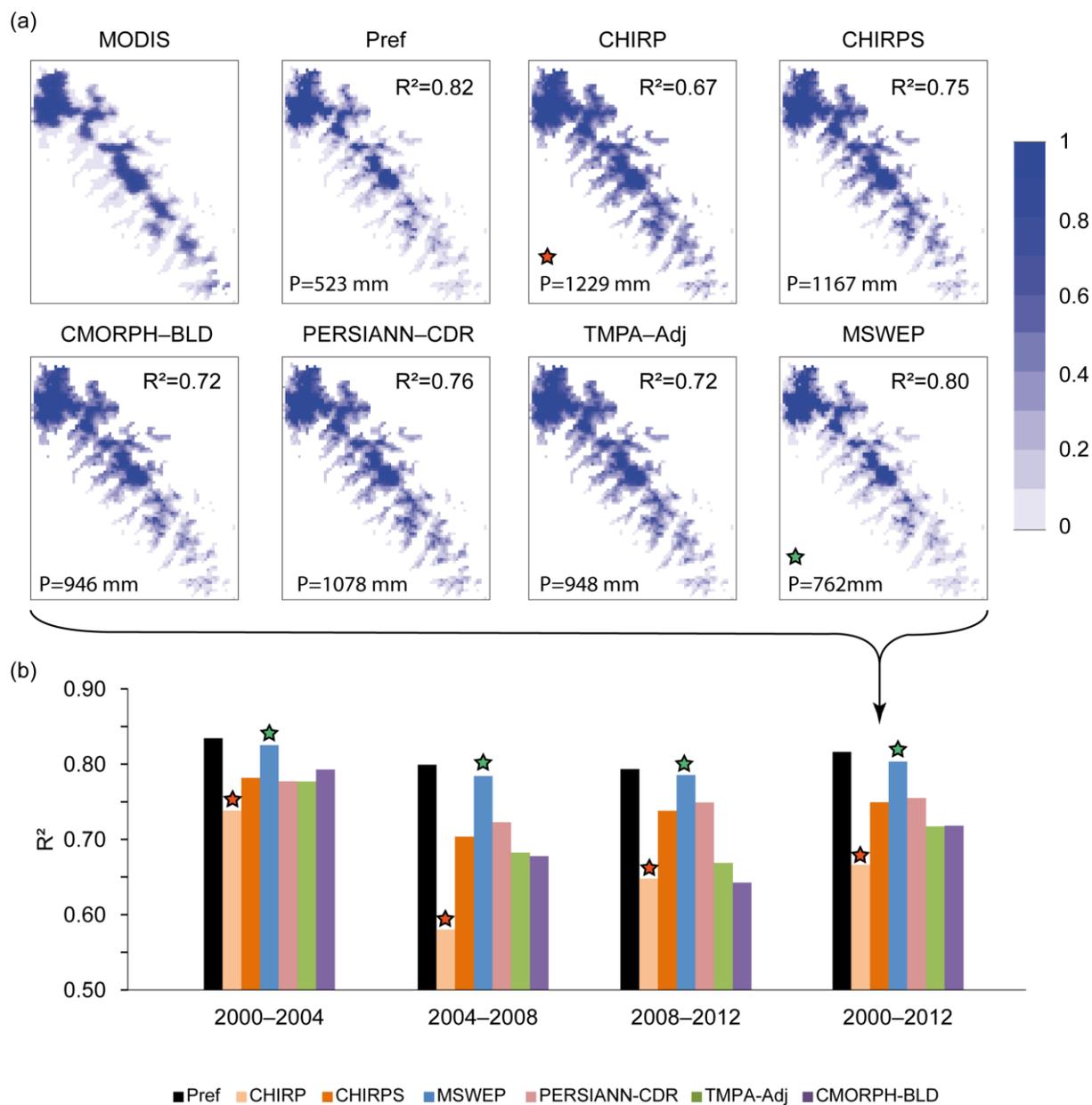
5



4.3. Space-time consistency of SPEs for snow cover duration

Figure 8 compares the snow cover duration (SCD) observed by MODIS and the durations simulated by the snow model using P_{ref} or all the SPEs as input data over the different periods. Whatever the period, the best SCD simulations were obtained using P_{ref} as input data with R^2 values of 0.83, 0.80, 0.79 and 0.82 for the 2000–2004, 2004–2008, 2008–2012 and 2000–2012 periods, respectively. Simulated SCD using SPEs as input data systematically overestimated SCD compared to the SCD computed from the gap-filled MODIS snow products. The MSWEP (CHIRPS) datasets provided the most realistic (unrealistic) SCD simulations over 2000–2012 with R^2 of 0.80 and 0.67 respectively (see Fig. 8a). Interestingly, mean annual precipitation over 2000–2012 (see Fig. 8a) was significantly higher with the SPEs (ranging from 762 mm with MSWEP to 1,229 mm with CHIRP) than with P_{ref} (523 mm). The least realistic SCD simulations with the SPEs may thus be explained by higher precipitation, which increases snowfall input and snow cover duration despite the specific calibration of the snow model. Indeed, MSWEP mean annual precipitation (762 mm) was the closest to the P_{ref} ones and provided the most realistic SCD estimates. Conversely, CHIRP provided the highest mean annual precipitation estimate (1,229 mm) whereas it produced the least realistic SCD estimates with R^2 values of 0.74, 0.58, 0.65, and 0.67, for the 2000–2004, 2004–2008, 2008–2012 and 2000–2012 periods, respectively. This trend was observed for all the periods considered (see Fig. 8b). All the other SPEs presented relatively close mean annual precipitation estimates and therefore relatively close SCD simulation performances (Fig. 8b).

Interestingly, all the SCD simulations based on SPEs as inputs had higher scores for the period 2000–2004 with minimum efficiency for CHIRP ($R^2=0.74$) and maximum for MSWEP ($R^2=0.83$). This is not in line with the space and time analysis based on gauges and hydrological modelling analyses. Indeed, CMORPH-BLD was the least efficient SPE to represent P_{ref} and observed streamflow over 2000–2004, while its SCD estimates was the second most efficient SCD simulation forced by SPEs. Similarly, the CHIRP forced streamflow simulation was the most realistic one in the Keka basin for the period 2000–2004, while the CHIRP SCD simulation was the least efficient in comparison to the other SPEs. The TMPA forced streamflow simulation was the first most realistic one for the Ilave basin and the second most realistic one for the Ramis basin over the period 2008–2012, whereas its SCD simulations were among the least efficient ($R^2=0.67$) in comparison to the other SPEs. This shows that a SPEs can be effective for a specific region or for a given indicator (i.e. gauges and hydrological modelling) but not for another region or indicator (i.e. snow modelling) and inversely. However, the analysis also confirmed the MSWEP overall stability and efficiency in space and over time previously highlighted for the precipitation and streamflow representation. Indeed, MSWEP forced SCD simulations remained relatively stable with similar scores obtained for all the periods considered.



5 **Figure 8: Observed (gap-filled MODIS snow-products) versus simulated snow cover duration (SCD) using P_{ref} and SPEs as input data in the snow model: (a) Maps of snow cover duration for 2000–2012; (b) Efficiency (R^2 scores) of simulated SCD versus MODIS SCD for the different calibration periods. SCD are expressed as the percentage of days over the simulation period. R^2 values stand for the pixel-to-pixel correlation between the observed and simulated SCD. Green and red stars highlight SPEs with closest and further precipitation estimates to P_{ref} over the subset considered.**

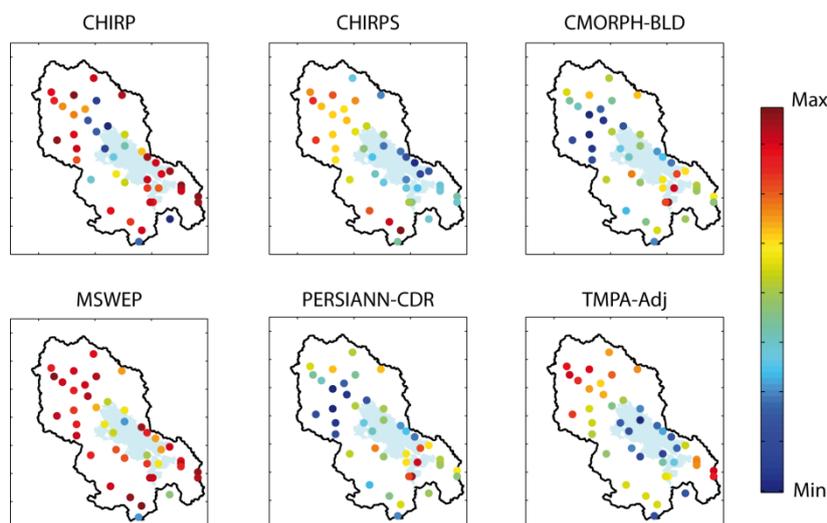


5. Discussion and conclusion

5.1. SPE spatial variability: the Lake Titicaca bias

At the regional scale, the performances of the SPEs were in relatively good agreement with the gauge references. However, their performances differed markedly in space with a general trend to underestimating precipitation over Lake Titicaca.

5 Due to its size, the solar radiation absorption capacity of Lake Titicaca increases the temperature by 4° to 6° over the superficial water layer in comparison to that over the surrounding land (Delclaux et al., 2007; Roche et al., 1992). Additionally, evaporation from Lake Titicaca is very high, estimated a 1,700 mm.year⁻¹ (Pillco Zolá et al., 2018). Therefore, crossing the lake, the air masses pick up lake moisture which increases their temperature and allows their ascension. This convection results in more precipitation over the lake than over the surrounding land (Roche et al., 1992). These
10 precipitation events originate from warm clouds whose detection remains challenging for SPEs. Indeed, SPEs use cloud top IR temperature to discretize rainy and rainless clouds. The IR temperature threshold may be too low to correctly detect warm rain clouds, as suggested over mountainous region (see e.g., Dinku et al., 2010, 2007; Gebregiorgis and Hossain, 2013; Hirpa et al., 2010; Li et al., 2013). Consequently, many precipitation events may be lost leading to underestimation of precipitation over Lake Titicaca. To support our hypothesis, we computed the probability of detection (POD) of daily precipitation events.
15 POD is an indicator of a SPE's ability to correctly forecast precipitation events with values ranging between 0 and 1 and a perfect score of 1. A low POD indicates that precipitation events are not detected by SPEs. Figure 9 shows the POD obtained from all SPEs for each pixel with gauges over the 2000–2012 period. With relatively lower POD values for pixels above the lake than for pixels over the surrounding land, TMPA–Adj and CHIRPS exhibited a clear trend to underestimate daily precipitation occurrence over the lake. MSWEP and CHIRP showed the same trend over the northern part of the lake. These
20 POD “anomalies” could partially explain the negative bias observed over the Lake Titicaca because of the warm cloud precipitation process. CMORPH–BLD and PERSIANN–CDR showed no significant trends, and were the least efficient SPEs overall.



5 **Figure 9: SPE's ability to forecast daily precipitation events represented in the form of POD. $POD = a/a+b$ with "a" being the number of days on which both SPE and P_{ref} detected precipitation and "b" the number of days on which only P_{ref} detected precipitation events. The legend shows the range between min and max to focus on POD anomalies over Lake Titicaca. Pixels located outside the Lake Titicaca basin were removed to facilitate interpretation.**

10 It is noteworthy that the eastern lake border is characterized by significant emissivity and temperature changes (induced by the lake and by the snow cover in the Cordillera), which perturb PMW precipitation retrieval (see e.g., Ferraro et al., 1998; Levizzani et al., 2002; Tian and Peters-Lidard, 2007) and may also contribute to the SPEs relatively lower performance.

15 River streamflow and precipitation over the lake surface account for respectively, 53% and 47% of the total Titicaca water supply (Roche et al., 1992). Even if MSWEP led to realistic simulations of streamflow, its absolute bias greater than 35% over the lake is a major problem when attempting to model the Lake Titicaca water budget. Therefore, MSWEP precipitation estimates need to be adjusted over the lake. Previous studies merging gauges and SPEs successfully applied over the Altiplano (Blacutt et al., 2015; Heidinger et al., 2012; Vila et al., 2009) and elsewhere (Ma et al., 2017; Ringard et al., 2017) should be used as a guideline.

5.3. Gauges versus hydrological modelling based assessment

20 Most of the studies on comparisons of SPE potential rely more on gauge-based assessment (see reviews by Maggioni et al., 2016 and Sun et al., 2017) rather than on hydrological modelling based assessment (see review by Maggioni and Massari, 2018). However, as shown in this study, the SPE's performance compared with gauge observations is not systematically supported by the sensitivity analysis of streamflow modelling. To provide more insight into this discrepancy, Fig. 10 compares SPE performance with precipitation gauges and with observed streamflow at the outlet of the four basins considered (Ilave, Katari, Keka, and Ramis). To simplify the analysis, the comparisons focus on the best and worst SPEs



obtained from both gauges and hydrological analysis. Over the Katari and Ramis basins, MSWEP was the most efficient SPE, according to both gauges and hydrological assessment. However, for the Ilave and Keka basins, the most efficient SPEs varied with the assessment indicator. For the Ilave basin, MSWEP was the most efficient SPE in representing precipitation at the gauge level while TMPA–Adj provided the best streamflow simulation. For the Keka basin, CHIRPS and CMORPH–BLD were the most efficient SPEs according to the gauges and hydrological analysis, respectively. Similar discrepancies were also observed concerning the least efficient SPEs. Indeed, for the Katari, Keka and Ramis basins, the least efficient SPEs varied with the assessment indicator considered (i.e. gauges or hydrological). We identified two main factors to explain those discrepancies.

The first factor is the spatial distribution of precipitation gauges in the catchment. For a given basin, the observed outlet streamflow includes precipitation over the basin. Therefore, unlike the gauge-based assessment, the hydrological assessment allows indirect feedback from the SPE’s ungauged pixels in the basin. The ratio of “suitable” to “unsuitable” SPE pixels at the basin scale may differ from the point gauge assessment of SPE pixels and explain part of the discrepancy between the gauges and hydrological-based assessment. This difference could be even more marked when the gauges are located in specific pixels where SPEs do not provide efficient precipitation estimates due to geomorphological context or to spurious SPE anomalies, as shown by Chen and Li, (2016) and Satgé et al. (2017b). The hydrological-based assessment makes it possible to overcome these problems (distribution and density of the gauges). Besides, errors in the spatial distribution of precipitation can be partially offset by model calibration. Except when a dense network of precipitation gauges adequately captures the precipitation spatial patterns, hydrological-based evaluation appears to be a necessary step to complement SPE gauge-based assessments.

The second factor is the difference in spatial representation between point (gauge) and areal (SPE pixel) measurement of precipitation. An areal measurement (SPE pixel) can be considered as the aggregation of “infinity” point measurements (gauges). “Infinity” is hard to represent with limited spatial gauge measurements. At the SPE pixel scale, precipitation events can occur in the vicinity of the gauges but not at the exact location of the gauges. Consequently, differences in precipitation estimates derived from the gauges and the SPEs are not only related to SPE deficiencies, but also to the difference in their respective measured spatial scale. As a result, as shown by Tang et al. (2018), the use of sparse gauge networks tends to underestimate SPE potential. This would be even more marked for lower spatial resolution SPEs (TMPA–Adj, PERSIANN–CDR, and CMORPH–BLD, in this study) than for higher spatial resolution SPEs (CHIRP, CHIRPS, and MSWEP, in this study) as the difference in spatial representation between a gauge and a SPE is greater. Indeed, the gauge-based assessment ranked the CHIRPS, MSWEP as the best SPEs for this study area. However, aggregation of precipitation at the basin scale eliminated the difference in spatial representation between point (gauge) and areal (SPE) precipitation as both gauge and SPE aimed to represent precipitation at the same spatial scale (basin scale). Therefore, unlike gauge precipitation measurements, streamflow measurements are not expected to be more representative of high than low spatial resolution SPEs. Indeed, as shown in Figure 10, the best performances are not systematically achieved with the highest SPE resolutions compared with streamflow observation. The results obtained over the Ilave and Keka basins illustrate this



feature: over the Keka basin, CHIRPS (5 km) precipitation estimates were better whereas CMORPH–BLD (25 km) provided better streamflow estimates. Similarly, MSWEP (10 km) was more efficient than gauges, but TMPA–Adj (25 km) provided better streamflow simulation over the Ilave basin.

To summarize, SPE assessments based on a gauge network make it possible to detect very local SPE inconsistencies but are influenced by (i) the distribution of the gauges and (ii) the difference in spatial resolution (point vs. pixels). It is also worth mentioning that hydrological-based assessments may be influenced by the model itself. In this study, the SPE assessment relied notably on their ability to provide realistic streamflow simulations at the basin outlet using a lumped model. Distributed physical models could reinforce the potential of each product based on spatial criteria (for instance humidity, water level), but this would be difficult over the Lake Titicaca region due to the scarcity of data. A complete quality assessment of SPEs is therefore difficult to achieve and the choice of the most suitable SPEs should rather be based on assessment steps and include the final use of the SPEs.

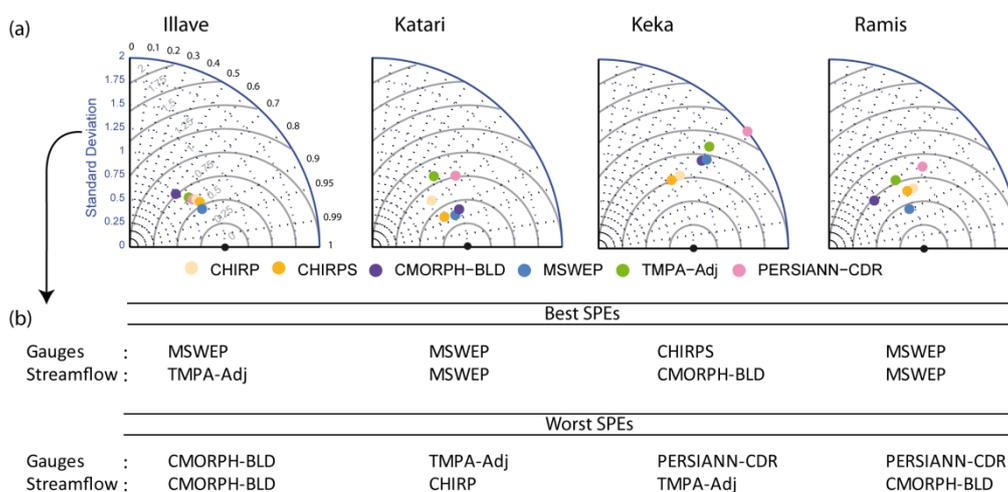


Figure 10. Comparison of the performance of the SPEs with that of precipitation gauges and observed streamflow at the outlet of the four basins (Ilave, Katari, Keka, and Ramis): (a) SPE performance at the gauge level for the four catchments represented according to a Taylor diagram (scores were computed using only the pixels included in the catchments for the 2008–2012 period); (b) Best and worst SPEs according to the gauge-based and hydrological-based assessments for the 2008–2012 period.

5.3. Space-time SPE consistency

For each assessment indicator considered (i.e. gauges, streamflow and snow cover duration), the performances of the SPEs varied depending on the time window considered. As shown in the analysis, a given SPEs can appear to be the best option to represent precipitation for one particular period but not for another. Additionally, for the same period, the SPE may be suitable for a specific regional subset but not for another. This space-time inconsistency is of major concern in the current context of climate variability. Indeed, over remote regions, the scarcity of meteorological stations encourages scientists to use remote sensing data to understand precipitation variability and its contribution to meteorological, agricultural and



hydrological droughts (see e.g. Agutu et al., 2017; Arvor et al., 2017; Tan et al., 2017; Tao et al., 2016; Bayissa et al., 2017; Satgé et al., 2017a). Therefore, for these studies, a consistent analysis of the consistency of SPEs in space and over time is a pioneering step to select the most realistic SPEs. It should make it possible to foresee potential propagation of SPE inconsistencies in the studies to consistently weight the observed results. From our analysis, it will be recalled that some SPEs are relatively stable at regional scale (CHIRPS, MSWEP) and could thus be used to study regional precipitation patterns with a relatively high degree of confidence. For studies on the spatial variability of precipitation, MSWEP is the most suitable SPE for the study region as it is the most stable in space and over time. However, as discussed above, correction methods need to be taken into consideration to enhance precipitation estimates over Lake Titicaca. Moreover, care should be taken in the north-eastern and south-western regions corresponding to the transition from the Amazon to the Altiplano and the Altiplano to the Pacific watershed zones, respectively (see Fig. 6d). For these specific regions, all the SPEs studied here represented the lowest space-time stability. These regions present very significant variations in elevation, which are known to interfere in SPEs (Ochoa et al., 2014; Satgé et al., 2017b) and may partially explain this local discrepancy.

5.4. Snow modelling to assess SPEs over unmonitored regions

SPEs are subject to high uncertainty in mountainous regions. Indeed the precipitation/no precipitation classification based on cloud top IR temperature generally fails because the temperature threshold used for the discretization process is too high. Currently, over high elevation mountainous regions, the top cloud temperature is generally lower than over flat regions resulting in many non-rainy cold clouds being misidentified as rainy clouds (Hussain et al., 2017; Satgé et al., 2017a). In addition, snow and ice cover appear to be similar to ice precipitation aloft in the scattering signal in P channel (Ferraro et al., 1998; Levizzani et al., 2002) leading to the misidentification of snow and ice cover with rainy clouds (Dinku et al., 2010; Hussain et al., 2017; Mourre et al., 2016). As an example, the inconstancy described in both IR and PMW precipitation recovery led to marked overestimation by SPEs over the Himalayan Pakistan region (Hussain et al., 2017, Satgé et al., 2018). However, the accessibility and handling difficulty over high elevation snow covered regions limits the availability of gauges and hence the assessment of SPEs. In this context, the proposed snow modelling-based assessment using snow cover observed by satellite imagery offers an alternative way to obtain initial feedback on the efficiency of SPEs over completely unmonitored regions. The present study shows that gap-filled MODIS snow-products can be used as reference to indirectly assess SPE efficiency and show great promise for the validation of SPEs in regions where distinguishing between rainfall and snowfall is still challenging. However, the distribution of SCD (Fig. 8a) indicates permanent snow-covered areas over which seasonal dynamics from optical imagery (MODIS) are difficult to capture as there is more variation in depth than in spatial extent. Consequently, the regional seasonal snow cycle captured from MODIS is weak and erratic (see Fig. 11). Preliminary modelling tests showed that the dynamics of these snow covered areas were poorly simulated by the snow model when the P_{ref} and SPE precipitation datasets were used.

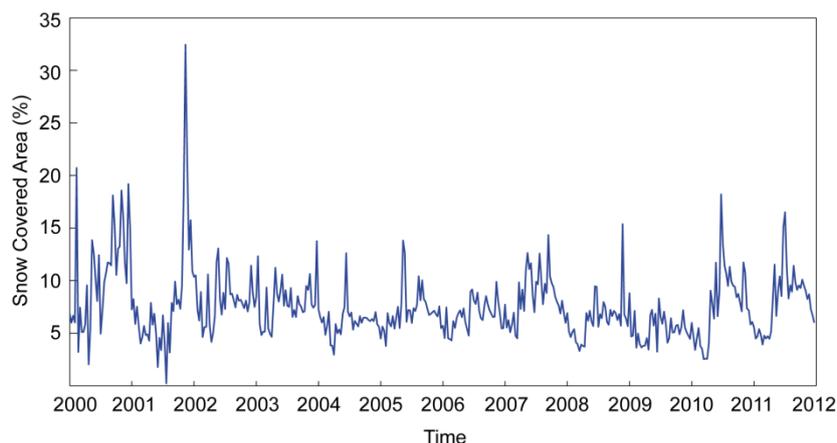


Figure 11: Snow covered area observed from gap-filled MODIS snow products at a 10-day time step over the Andean subset.

Another possible approach could consist of coupling satellite radar altimetry and optical imagery to monitor the seasonal snow cycle in terms of snow depth and volume. The recently launched Sentinel-2 (optical) and Sentinel-3 (altimeter) satellites present an unprecedented opportunity to monitor snow volumetric dynamics at finer spatial resolution than MODIS.

In addition, the SPEs tested in this study were not originally designed to distinguish between liquid and solid precipitation. The newly released SPEs Integrated Multi-satellite Retrievals for the Global Precipitation Mission (IMERG-v3 v4 v5) are the first SPEs which make it possible to discretize liquid to solid precipitation at 10 km spatial resolution and almost global scale. We expect that such products will enhance the monitoring of seasonal snow dynamics. However, as yet, no studies have reported on the potential of IMERG products to estimate snowfall, which should guide the next step in the assessment of SPEs over the region.

6. Conclusions

This paper compared the space-time consistency of 12 SPEs at different spatial and temporal scales with point gauge observations and according to sensitivity analysis of snow-runoff responses. The main results of the study can be summarized as follows:

- Given the currently available precipitation gauge network, SPEs are attractive and efficient tools to monitor local precipitation and to force impact modelling, such as snow-hydrological models. Of the SPEs with a grid scale of 10 km and more than 35 years of observed precipitation, MSWEP provided the best precipitation estimates at the gauge level, globally the most realistic streamflow simulations (with seven out of ten simulations outperforming the ones obtained using the available precipitation gauge network), and the most realistic simulations of snow cover duration compared to those simulated with the other SPEs.



- SPEs present space-time errors that cannot be assessed when the time windows are too short and/or the gauge networks are too scarce. In this context, the space-time consistency of SPEs should be carefully considered before using them for hydro-climatic studies.
- Ranked SPE potential varied depending on the assessment indicators used. It is therefore recommended to assess SPE potential according to indicators which are related to their final use.
- The proposed sensitivity analysis of snow modelling to the SPEs by using MODIS snow-products as control data has great promise for the assessment of SPE potential over completely unmonitored snow-covered regions.
- For the three assessment indicators (gauges, streamflow and snow cover duration) considered here, all SPE versions including gauge data for precipitation estimates (TMPA-Adj, CMORPH-BLD and CHIRPS) outperformed their satellite only based version (TMPA-RT, CMORPH-RAW and CHIRP).
- Soil moisture based precipitation estimates (SM2Rain) were shown to be promising precipitation estimates but are unsuitable for regional contexts with large waterbodies, mountainous and snow covered regions, which include too many gaps in time and space. The much smoothed decreasing north-west south-east precipitation pattern produced by GSMaP was not sensitive to local precipitation variability, reflecting an overall poor performance.

Acknowledgements This work is part of a postdoctoral fellowship funded by the CNES (*Centre National d'Etudes Spatiales*, France). The authors are grateful to SPE and MODIS dataset providers and to the SENAMHI from Bolivia and Peru for providing in situ precipitation, temperature and streamflow observations.

References

- Agutu, N. O., Awange, J. L., Zerihun, A., Ndehedehe, C. E., Kuhn, M. and Fukuda, Y.: Assessing multi-satellite remote sensing, reanalysis, and land surface models' products in characterizing agricultural drought in East Africa, *Remote Sens. Environ.*, 194, 287–302, doi:10.1016/j.rse.2017.03.041, 2017.
- Arvor, D., Funatsu, B., Michot, V. and Dubreuil, V.: Monitoring Rainfall Patterns in the Southern Amazon with PERSIANN-CDR Data: Long-Term Characteristics and Trends, *Remote Sens.*, 9(9), 889, doi:10.3390/rs9090889, 2017.
- Ashouri, H., Hsu, K. L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., Nelson, B. R. and Prat, O. P.: PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies, *Bull. Am. Meteorol. Soc.*, 96(1), 69–83, doi:10.1175/BAMS-D-13-00068.1, 2015.
- Bai, L., Shi, C., Li, L., Yang, Y. and Wu, J.: Accuracy of CHIRPS satellite-rainfall products over mainland China, *Remote Sens.*, 10(3), doi:10.3390/rs10030362, 2018.
- Bayissa, Y., Tadesse, T., Demisse, G. and Shiferaw, A.: Evaluation of satellite-based rainfall estimates and application to monitor meteorological drought for the Upper Blue Nile Basin, Ethiopia, *Remote Sens.*, 9(7), 1–17, doi:10.3390/rs9070669, 2017.



- Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., van Dijk, A. I. J. M., Weedon, G., Brocca, L., Pappenberger, F., Huffman, G. J. and Wood, E. F.: Global-scale evaluation of 23 precipitation datasets using gauge observations and hydrological modeling, *Hydrol. Earth Syst. Sci. Discuss.*, (August), 1–23, doi:10.5194/hess-2017-508, 2017.
- 5 Bergeron, J., Royer, A., Turcotte, R. and Roy, A.: Snow cover estimation using blended MODIS and AMSR-E data for improved watershed-scale spring streamflow simulation in Quebec, Canada, *Hydrol. Process.*, 28(16), 4626–4639, doi:10.1002/hyp.10123, 2014.
- Blacutt, L. A., Herdies, D. L., de Gonçalves, L. G. G., Vila, D. A. and Andrade, M.: Precipitation comparison for the CFSR, MERRA, TRMM3B42 and Combined Scheme datasets in Bolivia, *Atmos. Res.*, doi:10.1016/j.atmosres.2015.02.002, 2015.
- Bradley, R. S.: Threats to Water Supplies in the Tropical Andes, *Science* (80-.), 312(5781), 1755–1756, doi:10.1126/science.1128087, 2006.
- 10 Cao, Y., Zhang, W. and Wang, W.: Evaluation of TRMM 3B43 data over the Yangtze River Delta of China, *Sci. Rep.*, (August 2017), 1–12, doi:10.1038/s41598-018-23603-z, 2018.
- Carmouze, J. and Aquize, E.: La régulation hydrique du lac Titicaca et l’hydrologie de ses tributaires, *Hydrobiol. trop.*, 14(4), 311–328, 1981.
- 15 Carmouze, J., Arce, C. and Quintanilla, J.: La régulation hydrique des lacs Titicaca et Poopo, *Cah. O.R.ST.O.M., Sér. Hydrobiol.*, 11(4), 269–283, 1977.
- Casse, C., Gosset, M., Peugeot, C., Pedinotti, V., Boone, a., Tanimoun, B. a. and Decharme, B.: Potential of satellite rainfall products to predict Niger River flood events in Niamey, *Atmos. Res.*, doi:10.1016/j.atmosres.2015.01.010, 2015.
- Chen, F. and Li, X.: Evaluation of IMERG and TRMM 3B43 Monthly Precipitation Products over Mainland China, , 1–18, doi:10.3390/rs8060472, 2016.
- 20 Chen, S., Hong, Y., Gourley, J. J., Huffman, G. J., Tian, Y., Cao, Q., Yong, B., Kirstetter, P.-E., Hu, J., Hardy, J., Li, Z., Khan, S. I. and Xue, X.: Evaluation of the successive V6 and V7 TRMM multisatellite precipitation analysis over the Continental United States, *Water Resour. Res.*, 49(12), 8174–8186, doi:10.1002/2012WR012795, 2013.
- Condom, T., Rau, P. and Espinoza, J. C.: Correction of TRMM 3B43 monthly precipitation data over the mountainous areas of Peru during the period 1998-2007, *Hydrol. Process.*, 25(12), 1924–1933, doi:10.1002/hyp.7949, 2010.
- 25 Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M. and Hendrickx, F.: Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments, *Water Resour. Res.*, 48(5), 1–17, doi:10.1029/2011WR011721, 2012.
- Dakhlaoui, H., Ruelland, D., Trambly, Y. and Bargaoui, Z.: Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia, *J. Hydrol.*, 550, 201–217, doi:10.1016/j.jhydrol.2017.04.032, 2017.
- 30 Delclaux, F., Coucrain, A. and Condom, T.: Evaporation estimation on Lake Titicaca : a synthesis review and modelling, *Hydrol. Process.*, 1677(January), 1664–1677, doi:10.1002/hyp, 2007.



- Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J. and Ropelewski, C. F.: Validation of satellite rainfall products over East Africa's complex topography, *Int. J. Remote Sens.*, 28(7), 1503–1526, doi:10.1080/01431160600954688, 2007.
- Dinku, T., Connor, S. J. and Ceccato, P.: Satellite Rainfall Applications for Surface Hydrology, in *Satellite Rainfall Applications for Surface Hydrology*, edited by M. Gebremichael and F. Hossain, pp. 193–204, Springer Netherlands, Dordrecht., 2010.
- Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M. and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, *Remote Sens. Environ.*, 162, 380–395, doi:10.1016/j.rse.2014.07.023, 2015.
- 10 Doumounia, A., Gosset, M., Cazenave, F., Kacou, M. and Zougmore, F.: Rainfall monitoring based on microwave links from cellular telecommunication networks: First results from a West African test bed, *Geophys. Res. Lett.*, 41(16), 6015–6021, doi:10.1002/2014GL060724, 2014.
- Duan, Q., Sorooshian, S. and Gupta, V.: Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resour. Res.*, 28(4), 1015–1031, doi:10.1029/91WR02985, 1992.
- 15 Erazo, B., Bourrel, L., Frappart, F., Chimborazo, O., Labat, D., Dominguez-Granda, L., Matamoros, D. and Mejia, R.: Validation of satellite estimates (Tropical Rainfall Measuring Mission, TRMM) for rainfall variability over the Pacific slope and Coast of Ecuador, *Water (Switzerland)*, 10(2), doi:10.3390/w10020213, 2018.
- Espinoza Villar, J. C., Ronchail, J., Guyot, J. L., Cochonneau, G., Naziano, F., Lavado, W., de Oliveira, E., Pombosa, R. and Vauchel, P.: Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador), *Int. J. Climatol.*, 29(11), 1574–1594, doi:10.1002/joc.1791, 2009.
- 20 Fabre, J., Ruelland, D., Dezetter, A. and Grouillet, B.: Simulating past changes in the balance between water demand and availability and assessing their main drivers at the river basin scale, *Hydrol. Earth Syst. Sci.*, 19(3), 1263–1285, doi:10.5194/hess-19-1263-2015, 2015.
- Fabre, J., Ruelland, D., Dezetter, A. and Grouillet, B.: Sustainability of water uses in managed hydrosystems: Human- and climate-induced changes for the mid-21st century, *Hydrol. Earth Syst. Sci.*, 20(8), 3129–3147, doi:10.5194/hess-20-3129-2016, 2016.
- Ferraro, R. R., Smith, E. A., Berg, W. and Huffman, G. J.: A Screening Methodology for Passive Microwave Precipitation Retrieval Algorithms, *J. Atmos. Sci.*, 55(9), 1583–1600, doi:10.1175/1520-0469(1998)055<1583:ASMFPM>2.0.CO;2, 1998.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. and Michaelson, J.: The climate hazards infrared precipitation with stations — a new environmental record for monitoring extremes, , 1–21, doi:10.1038/sdata.2015.66, 2015.
- 30 Gascoin, S., Hagolle, O., Huc, M., Jarlan, L., Dejoux, J. F., Szczypta, C., Marti, R. and Sánchez, R.: A snow cover



- climatology for the Pyrenees from MODIS snow products, *Hydrol. Earth Syst. Sci.*, 19(5), 2337–2351, doi:10.5194/hess-19-2337-2015, 2015.
- Gebregiorgis, A. S. and Hossain, F.: Understanding the Dependence of Satellite Rainfall Uncertainty on Topography and Climate for Hydrologic Model Simulation, *IEEE Trans. Geosci. Remote Sens.*, 51(1), 704–718, doi:10.1109/TGRS.2012.2196282, 2013.
- Grouillet, B., Ruelland, D., Ayar, P. V. and Vrac, M.: Sensitivity analysis of runoff modeling to statistical downscaling models in the western Mediterranean, *Hydrol. Earth Syst. Sci.*, 20(3), 1031–1047, doi:10.5194/hess-20-1031-2016, 2016.
- GSMaP: USER'S GUIDE FOR Global Satellite Mapping of Precipitation Microwave-IR Combined Product (GSMaP_MVK) Version 5., 2012.
- 10 Guo, H., Bao, A., Liu, T., Chen, S. and Ndayisaba, F.: Evaluation of PERSIANN-CDR for Meteorological Drought Monitoring over China, *Remote Sens.*, 8(5), 379, doi:10.3390/rs8050379, 2016.
- Guo, H., Bao, A., Liu, T., Ndayisaba, F., He, D., Kurban, A. and Maeyer, P. De: Meteorological Drought Analysis in the Lower Mekong Basin Using Satellite-Based Long-Term CHIRPS Product, *Sustainability*, (9), doi:10.3390/su9060901, 2017.
- 15 Heidinger, H., Yarlequé, C., Posadas, A. and Quiroz, R.: TRMM rainfall correction over the Andean Plateau using wavelet multi-resolution analysis, *Int. J. Remote Sens.*, 33(14), 4583–4602, doi:10.1080/01431161.2011.652315, 2012.
- Heidinger, H., Carvalho, L., Jones, C., Posadas, A. and Quiroz, R.: A new assessment in total and extreme rainfall trends over central and southern Peruvian Andes during 1965-2010, *Int. J. Climatol.*, doi:10.1002/joc.5427, 2018.
- Hirpa, F. a., Gebremichael, M. and Hopson, T.: Evaluation of High-Resolution Satellite Precipitation Products over Very Complex Terrain in Ethiopia, *J. Appl. Meteorol. Climatol.*, 49(5), 1044–1051, doi:10.1175/2009JAMC2298.1, 2010.
- 20 Hock, R.: Temperature index melt modelling in mountain areas, *J. Hydrol.*, 282(1–4), 104–115, doi:10.1016/S0022-1694(03)00257-9, 2003.
- Hublart, P., Ruelland, D., De Cortazar-Atauri, I. G., Gascoin, S., Lhermitte, S. and Ibacache, A.: Reliability of lumped hydrological modeling in a semi-arid mountainous catchment facing water-use changes, *Hydrol. Earth Syst. Sci.*, 20(9), 3691–3717, doi:10.5194/hess-20-3691-2016, 2016.
- 25 Huffman, G. J. and Bolvin, D. T.: TRMM and Other Data Precipitation Data Set Documentation., 2014.
- Huffman, G. J. and Bolvin, D. T.: TRMM and Other Data Precipitation Data Set Documentation., 2016.
- Huffman, G. J., Bolvin, D. T. and Nelkin, E. J.: Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation., 2017.
- 30 Hussain, Y., Satgé, F., Hussain, M. B., Martinez-Caravajal, H., Bonnet, M.-P., Cardenas-Soto, M., Llacer Roig, H. and Akhter, G.: Performance of CMORPH, TMPA and PERSIANN rainfall datasets over plain, mountainous and glacial regions of Pakistan, *Theor. Appl. Climatol.*, doi:10.1007/s00704-016-2027-z, 2017.
- Joyce, R. J., Janowiak, J. E., Arkin, P. A. and Xie, P.: CMORPH: A Method that Produces Global Precipitation Estimates



- from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution, *J. Hydrometeorol.*, 5(3), 487–803, doi:10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2, 2004.
- Katsanos, D., Retalis, A. and Michaelides, S.: Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period, *Atmos. Res.*, 169, 459–464, doi:10.1016/j.atmosres.2015.05.015, 2016.
- 5 Levizzani, V., Amorati, R. and Meneguzzo, F.: A review of satellite-based rainfall estimation methods., 2002.
- Li, X., Zhang, Q. and Xu, C.-Y.: Assessing the performance of satellite-based precipitation products and its dependence on topography over Poyang Lake basin, *Theor. Appl. Climatol.*, 115(3–4), 713–729, doi:10.1007/s00704-013-0917-x, 2013.
- López-Moreno, J. I., Morán-Tejeda, E., Vicente-Serrano, S. M., Bazo, J., Azorin-Molina, C., Revuelto, J., Sánchez-Lorenzo, A., Navarro-Serrano, F., Aguilar, E. and Chura, O.: Recent temperature variability and change in the Altiplano of Bolivia and
10 Peru, *Int. J. Climatol.*, n/a-n/a, doi:10.1002/joc.4459, 2015.
- Lozada, G.: *BALANCE HIDRICO DE LA CUENCA DEL LAGO TITICACA.*, 1985.
- Ma, Y., Yang, Y., Han, Z., Tang, G., Maguire, L., Chu, Z. and Hong, Y.: Comprehensive evaluation of Ensemble Multi-Satellite Precipitation Dataset using the Dynamic Bayesian Model Averaging scheme over the Tibetan Plateau, *J. Hydrol.*, doi:10.1016/j.jhydrol.2017.11.050, 2017.
- 15 Maggioni, V. and Massari, C.: On the performance of satellite precipitation products in riverine flood modeling: A review, *J. Hydrol.*, 558, 214–224, doi:10.1016/j.jhydrol.2018.01.039, 2018.
- Maggioni, V., Meyers, P. C. and Robinson, M. D.: A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era, *J. Hydrometeorol.*, 17(4), 1101–1117, doi:10.1175/JHM-D-15-0190.1, 2016.
- 20 Mahmoud, M. T., Al-Zahrani, M. A. and Sharif, H. O.: Assessment of Global Precipitation Measurement Satellite Products over Saudi Arabia, *J. Hydrol.*, 559, 1–12, doi:<https://doi.org/10.1016/j.jhydrol.2018.02.015>, 2018.
- Melo, D. D. C. D., Xavier, A. C., Bianchi, T., Oliveira, P. T. S., Scanlon, B. R., Lucas, M. C. and Wendland, E.: Performance evaluation of rainfall estimates by TRMM Multi-satellite Precipitation Analysis 3B42V6 and V7 over Brazil, *J. Geophys. Res. Atmos.*, 9426–9436, doi:10.1002/2015JD023797.Received, 2015.
- 25 Messer, H., Zinevich, A. and Alpert, P.: Environmental monitoring by wireless communication networks, *Science* (80-.), 312(5774), 713, doi:10.1126/science.1120034, 2006.
- Milewski, A., Elkadiri, R. and Durham, M.: Assessment and Comparison of TMPA Satellite Precipitation Products in Varying Climatic and Topographic Regimes in Morocco, *Remote Sens.*, 7(5), 5697–5717, doi:10.3390/rs70505697, 2015.
- Minvielle, M. and Garreaud, R. D.: Projecting rainfall changes over the South American Altiplano, *J. Clim.*, 24(17), 4577–
30 4583, doi:10.1175/JCLI-D-11-00051.1, 2011.
- Mourre, L., Condom, T., Junquas, C., Lebel, T., Sicart, J. E., Figueroa, R. and Cochachin, A.: Spatio-temporal assessment of WRF , TRMM and in situ precipitation data in a tropical mountain environment (Cordillera Blanca , Peru), , 125–141,



- doi:10.5194/hess-20-125-2016, 2016.
- Ochoa, a., Pineda, L., Crespo, P. and Willems, P.: Evaluation of TRMM 3B42 precipitation estimates and WRF retrospective precipitation simulation over the Pacific–Andean region of Ecuador and Peru, *Hydrol. Earth Syst. Sci.*, 18(8), 3179–3193, doi:10.5194/hess-18-3179-2014, 2014.
- 5 Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F. and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2 - Towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling, *J. Hydrol.*, 303(1–4), 290–306, doi:10.1016/j.jhydrol.2004.08.026, 2005.
- Ovando, A., Tomasella, J., Rodriguez, D. A., Martinez, J. M., Siqueira-Junior, J. L., Pinto, G. L. N., Passy, P., Vauchel, P.,
10 Noriega, L. and von Randow, C.: Extreme flood events in the Bolivian Amazon wetlands, *J. Hydrol. Reg. Stud.*, 5, 293–308, doi:10.1016/j.ejrh.2015.11.004, 2016.
- Overeem, A., Leijnse, H. and Uijlenhoet, R.: Measuring urban rainfall using microwave links from commercial cellular communication networks, *Water Resour. Res.*, 47(12), 1–16, doi:10.1029/2010WR010350, 2011.
- Perrin, C., Michel, C. and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, *J. Hydrol.*,
15 279(1–4), 275–289, doi:10.1016/S0022-1694(03)00225-7, 2003.
- Pillco Zolá, R., Bengtsson, L., Berndtsson, R., Martí-Cardona, B., Satgé, F., Timouk, F., Bonnet, M.-P., Mollericon, L., Gamarra, C. and Pasapera, J.: Modeling Lake Titicaca Daily and Monthly Evaporation, *Hydrol. Earth Syst. Sci. Discuss.*, (March), 1–23, doi:10.5194/hess-2018-127, 2018.
- Prakash, S., Sathiyamoorthy, V., Mahesh, C. and Gairola, R. M.: An evaluation of high-resolution multisatellite rainfall
20 products over the Indian monsoon region, *Int. J. Remote Sens.*, 35(9), 3018–3035, doi:10.1080/01431161.2014.894661, 2014.
- Rahmawati, N. and Lubczynski, M. W.: Validation of satellite daily rainfall estimates in complex terrain of Bali Island, Indonesia, *Theor. Appl. Climatol.*, 1–20, doi:10.1007/s00704-017-2290-7, 2017.
- Ringard, J., Becker, M., Seyler, F. and Linguet, L.: Temporal and spatial assessment of four satellite rainfall estimates over
25 French Guiana and north Brazil, *Remote Sens.*, 7(12), 16441–16459, doi:10.3390/rs71215831, 2015.
- Ringard, J., Seyler, F. and Linguet, L.: A quantile mapping bias correction method based on hydroclimatic classification of the Guiana shield, *Sensors (Switzerland)*, 17(6), 1–17, doi:10.3390/s17061413, 2017.
- Roche, M. A., Bourges, J., Cortes, J. and Mattos, R.: *Climatología e hidrología de la cuenca del lago Titicaca*, edited by C. Dejoux and A. Iltis, Springer Netherlands, Boston., 1992.
- 30 Ruelland, D., Ardoin-Bardin, S., Billen, G. and Servat, E.: Sensitivity of a lumped and semi-distributed hydrological model to several methods of rainfall interpolation on a large basin in West Africa, *J. Hydrol.*, 361(1–2), 96–117, doi:10.1016/j.jhydrol.2008.07.049, 2008.



- Ruelland, D., Dezetter, A. and Hublart, P.: Sensitivity analysis of hydrological modelling to climate forcing in a semi-arid mountainous catchment, in *Hydrology in a changing world: Environmental and human dimensions*, vol. 363, pp. 145–150., 2014.
- Ruelland, D., Volff, C., Gascoïn, S., Hublart, P.: Efficiency of distributed models against satellite observations to assess the sensitivity of snow cover to global warming on various mountain ranges, in review in *J. Hydrol*, 2018.
- Satgé, F., Bonnet, M.-P., Gosset, M., Molina, J., Hernan Yuque Lima, W., Pillco Zolá, R., Timouk, F. and Garnier, J.: Assessment of satellite rainfall products over the Andean plateau, *Atmos. Res.*, 167, 1–14, doi:10.1016/j.atmosres.2015.07.012, 2016.
- Satgé, F., Xavier, A., Zolá, R., Hussain, Y., Timouk, F., Garnier, J. and Bonnet, M.-P.: Comparative Assessments of the Latest GPM Mission's Spatially Enhanced Satellite Rainfall Products over the Main Bolivian Watersheds, *Remote Sens.*, 9(4), 369, doi:10.3390/rs9040369, 2017a.
- Satgé, F., Espinoza, R., Zolá, R., Roig, H., Timouk, F., Molina, J., Garnier, J., Calmant, S., Seyler, F. and Bonnet, M.-P.: Role of Climate Variability and Human Activity on Poopó Lake Droughts between 1990 and 2015 Assessed Using Remote Sensing Data, *Remote Sens.*, 9(3), 218, doi:10.3390/rs9030218, 2017b.
- Satgé, F., Hussain, Y., Bonnet, M.-P., Hussain, B., Martinez-Carvajal, H., Akhter, G. and Uagoda, R.: Benefits of the Successive GPM Based Satellite Precipitation Estimates IMERG–V03, –V04, –V05 and GSMaP–V06, –V07 Over Diverse Geomorphic and Meteorological Regions of Pakistan, *Remote Sens.*, 10(9), 1373, doi:10.3390/rs10091373, 2018.
- Serrat-Capdevila, A., Merino, M., Valdes, J. B. and Durcik, M.: Evaluation of the performance of three satellite precipitation products over Africa, *Remote Sens.*, 8(10), doi:10.3390/rs8100836, 2016.
- Sharifi, E., Steinacker, R. and Saghafian, B.: Assessment of GPM-IMERG and Other Precipitation Products against Gauge Data under Different Topographic and Climatic Conditions in Iran: Preliminary Results, *Remote Sens.*, 8(2), 135, doi:10.3390/rs8020135, 2016.
- Shen, Y., Xiong, A., Wang, Y. and Xie, P.: Performance of high-resolution satellite precipitation products over China, *J. Geophys. Res.*, 115(D2), D02114, doi:10.1029/2009JD012097, 2010.
- Shrestha, N. K., Qamer, F. M., Pedreros, D., Murthy, M. S. R., Wahid, S. and Shrestha, M.: Journal of Hydrology : Regional Studies Evaluating the accuracy of Climate Hazard Group (CHG) satellite rainfall estimates for precipitation based drought monitoring in Koshi basin , Nepal, *J. Hydrol. Reg. Stud.*, 13(September), 138–151, doi:10.1016/j.ejrh.2017.08.004, 2017.
- Sorooshian, S., Hsu, K.-L., Gao, X., Gupta, H. V., Imam, B. and Braithwaite, D.: Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall, *Bull. Am. Meteorol. Soc.*, 81(9), 2035–2046, doi:10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2, 2000.
- Su, J., Lü, H., Wang, J., Sadeghi, A. M. and Zhu, Y.: Evaluating the applicability of four latest satellite-gauge combined precipitation estimates for extreme precipitation and streamflow predictions over the upper yellow river basins in China,



- Remote Sens., 9(11), 1–19, doi:10.3390/rs9111176, 2017.
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S. and Hsu, K.-L.: A review of global precipitation datasets: data sources, estimation, and intercomparisons., 2017.
- Tan, M. L., Tan, K. C., Chua, V. P. and Chan, N. W.: Evaluation of TRMM product for monitoring drought in the Kelantan River Basin, Malaysia, *Water*, 9(1), doi:10.3390/w9010057, 2017.
- 5 Tang, G., Ma, Y., Long, D., Zhong, L. and Hong, Y.: Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales, *J. Hydrol.*, 533, 152–167, doi:10.1016/j.jhydrol.2015.12.008, 2016.
- Tang, G., Behrangi, A., Long, D., Li, C. and Hong, Y.: Accounting for spatiotemporal errors of gauges: A critical step to evaluate gridded precipitation products, *J. Hydrol.*, 559, 294–306, doi:10.1016/j.jhydrol.2018.02.057, 2018.
- 10 Tao, H., Fischer, T., Zeng, Y. and Fraedrich, K.: Evaluation of TRMM 3B43 precipitation data for drought monitoring in Jiangsu Province, China, *Water*, 8(6), 1–13, doi:10.3390/w8060221, 2016.
- Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, 106(D7), 7183–7192, doi:10.1029/2000JD900719, 2001.
- 15 Thiemig, V., Rojas, R., Zambrano-Bigiarini, M., Levizzani, V. and De roo, A.: Validation of Satellite-Based Precipitation Products over Sparsely Gauged African River Basins, *J. Hydrometeorol.*, 13(6), 1760–1783, doi:10.1175/JHM-D-12-032.1, 2012.
- Thiemig, V., Rojas, R., Zambrano-Bigiarini, M. and De Roo, A.: Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin, *J. Hydrol.*, 499, 324–338, doi:10.1016/j.jhydrol.2013.07.012, 2013.
- 20 Tian, Y. and Peters-Lidard, C. D.: Systematic anomalies over inland water bodies in satellite-based precipitation estimates, *Geophys. Res. Lett.*, 34(14), L14403, doi:10.1029/2007GL030787, 2007.
- Tian, Y., Peters-Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., Hsu, K.-L., Turk, F. J., Garcia, M. and Zeng, J.: Component analysis of errors in satellite-based precipitation estimates, *J. Geophys. Res.*, 114(D24), D24101, doi:10.1029/2009JD011949, 2009.
- 25 Trambly, Y., Thiemig, V., Dezetter, A. and Hanich, L.: Evaluation of satellite-based rainfall products for hydrological modelling in Morocco, *Hydrol. Sci. J.*, 61(14), 2509–2519, doi:10.1080/02626667.2016.1154149, 2016.
- Trenberth, K. E.: Changes in precipitation with climate change, *Clim. Res.*, 47(1–2), 123–138, doi:10.3354/cr00953, 2011.
- Uria, A. and Molina, J.: Analisis y tratamiento de datos hidrológicos de las cuencas Mauri-Desaguadero y Lago Titicaca, 2013.
- 30 Urrutia, R. and Vuille, M.: Climate change projections for the tropical Andes using a regional climate model: Temperature and precipitation simulations for the end of the 21st century, *J. Geophys. Res. Atmos.*, 114(2), 1–15, doi:10.1029/2008JD011021, 2009.



- Vaze, J., Post, D. A., Chiew, F. H. S., Perraud, J. M., Viney, N. R. and Teng, J.: Climate non-stationarity - Validity of calibrated rainfall-runoff models for use in climate change studies, *J. Hydrol.*, 394(3–4), 447–457, doi:10.1016/j.jhydrol.2010.09.018, 2010.
- 5 Vila, D. a., de Goncalves, L. G. G., Toll, D. L. and Rozante, J. R.: Statistical Evaluation of Combined Daily Gauge Observations and Rainfall Satellite Estimates over Continental South America, *J. Hydrometeorol.*, 10(2), 533–543, doi:10.1175/2008JHM1048.1, 2009.
- Wang, Z., Zhong, R., Lai, C. and Chen, J.: Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility, *Atmos. Res.*, 196(May), 151–163, doi:10.1016/j.atmosres.2017.06.020, 2017.
- 10 Zambrano-Bigiarini, M., Nauditt, A., Birkel, C., Verbist, K. and Ribbe, L.: Temporal and spatial evaluation of satellite-based rainfall estimates across the complex topographical and climatic gradients of Chile, *Hydrol. Earth Syst. Sci.*, 21(2), 1295–1320, doi:10.5194/hess-21-1295-2017, 2017.
- Zeng, Q., Wang, Y., Chen, L., Wang, Z., Zhu, H. and Li, B.: Inter-comparison and evaluation of remote sensing precipitation products over China from 2005 to 2013, *Remote Sens.*, 10(2), doi:10.3390/rs10020168, 2018.
- 15 Zinevich, A., Alpert, P. and Messer, H.: Estimation of rainfall fields using commercial microwave communication networks of variable density, *Adv. Water Resour.*, 31(11), 1470–1480, doi:10.1016/j.advwatres.2008.03.003, 2008.
- Zubieta, R., Getirana, A., Espinoza, J. C. and Lavado, W.: Impacts of satellite-based precipitation datasets on rainfall-runoff modeling of the Western Amazon basin of Peru and Ecuador, *J. Hydrol.*, 528, 599–612, doi:10.1016/j.jhydrol.2015.06.064, 2015.