



Evaluating satellite-derived long-term historical precipitation datasets for drought monitoring in Chile



Francisco Zambrano^{a,*}, Brian Wardlow^b, Tsegaye Tadesse^c, Mario Lillo-Saavedra^a, Octavio Lagos^a

^aDepartment of Water Resources, University of Concepción, Chile

^bCenter for Advanced Land Management Information Technologies, University of Nebraska, Lincoln, NE68588, USA

^cNational Drought Mitigation Center, University of Nebraska, Lincoln, NE68583-0988, USA

ARTICLE INFO

Article history:

Received 5 August 2016

Received in revised form 3 November 2016

Accepted 7 November 2016

Available online 11 November 2016

ABSTRACT

Precipitation is a key parameter for the study of climate change and variability and the detection and monitoring of natural disaster such as drought. Precipitation datasets that accurately capture the amount and spatial variability of rainfall is critical for drought monitoring and a wide range of other climate applications. This is challenging in many parts of the world, which often have a limited number of weather stations and/or historical data records. Satellite-derived precipitation products offer a viable alternative with several remotely sensed precipitation datasets now available with long historical data records (+30years), which include the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) and Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks-Climate Data Record (PERSIANN-CDR) datasets. This study presents a comparative analysis of three historical satellite-based precipitation datasets that include Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) 3B43 version 7 (1998–2015), PERSIANN-CDR (1983–2015) and CHIRPS 2.0 (1981–2015) over Chile to assess their performance across the country and for the case of the two long-term products the applicability for agricultural drought were evaluated when used in the calculation of commonly used drought indicator as the Standardized Precipitation Index (SPI). In this analysis, 278 weather stations of in situ rainfall measurements across Chile were initially compared to the satellite data. The study area (Chile) was divided into five latitudinal zones: North, North-Central, Central, South-Central and South to determine if there were a regional difference among these satellite products, and nine statistics were used to evaluate their performance to estimate the amount and spatial distribution of historical rainfall across Chile. Hierarchical cluster analysis, k-means and singular value decomposition were used to analyze these datasets to better understand their similarities and differences in characterizing rainfall patterns across Chile. Monthly analysis showed that all satellite products highly overestimated rainfall in the arid North zone. However, there were no major difference between all three products from North to South-Central zones. Though, in the South zone, PERSIANN-CDR shows the lowest fit with high underestimation, while CHIRPS 2.0 and TMPA 3B43 v7 had better agreement with in situ measurements. The accuracy of satellite products were highly dependent on the amount of monthly rainfall with the best results found during winter seasons and in zones (Central to South) with higher amounts of precipitation. PERSIANN-CDR and CHIRPS 2.0 were used to derive SPI at time-scale of 1, 3 and 6 months, both satellite products presented similar results when it was compared in situ against satellite SPI's. Because of its higher spatial resolution that allows better characterizing of spatial variation in precipitation pattern, the CHIRPS 2.0 was used to mapping the SPI-3 over Chile. The

* Corresponding author at: Av. Vicente Méndez 595, Casilla 527, Chillán, Chile.

E-mail addresses: frzambr@gmail.com (F. Zambrano), bwardlow2@unl.edu (B. Wardlow), ttadesse2@unl.edu (T. Tadesse), malillo@udec.cl (M. Lillo-Saavedra).

URL: <https://frzambr.github.io> (F. Zambrano).

Keywords:
 Precipitation
 Chile
 Drought
 Agriculture
 CHIRPS 2.0
 PERSIANN-CDR
 TMPA 3B43 v7

results of this study show that in order to use the CHIRPS 2.0 and PERSIANN-CDR datasets in Chile to monitor spatial patterns in the rainfall and drought intensity conditions, these products should be calibrated to adjust for the overestimation/underestimation of rainfall geographically specially in the North zone and seasonally during the summer and spring months in the other zones.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Precipitation is one of the key parameters for climate monitoring, particularly to detect climatically-extreme events such as drought, which impacts most regions of the world. A simple definition of drought is an extended period of abnormal dryness that has negative impacts on agricultural and water resources (WMO, 1986). Conceptually, there are several different sectoral definitions of drought that are defined by the duration of the precipitation deficit including short-term meteorological drought (spanning days to weeks), agricultural drought (month to several months), and hydrological drought (months to years) (Wilhite and Glantz, 1985). According to IPCC (2013), changes in the patterns of precipitation is expected globally over the next few decades. The change in precipitation patterns coupled with the sustained increase in global temperature since 1880 and the anthropogenic factors due to human activities (e.g. increased emission of the Green House Gasses and land use change such as cutting down forests to create farmland) is likely to increase the frequency and intensity of natural disasters like drought throughout the world (IPCC, 2013; Loon et al., 2016). Knowledge of the amount and spatial variability of precipitation historically is important to map and monitor drought condition globally. In situ-based rainfall measurements at weather station locations have traditionally been used for this application, but the number, geographic distribution, and length of record of these measurements are often lacking in many countries including Chile. The creation of such datasets is challenging because it is costly to maintain a dense network of weather stations over a long period of time as a result, there are often spatial gaps and a lack of local resolution in the rainfall data and the drought patterns mapped from these point-based weather station data using spatial interpolation techniques. Many weather stations have a relatively short or incomplete historical record of observations, which is problematic for determining the magnitude of specific precipitation deficit period and the severity of the corresponding drought.

Accurate historical precipitation data and effective drought monitoring tools are of considerable interest for Chile. The IPCC (2013) indicate that precipitation is expected to decrease in the near-future in the central part of Chile, which is a primarily agricultural area within the country. Studies made during the last years from North to Central Chile found important results about drought frequency. A trend in the increase of drought frequency in the Coquimbo region of northern Chile, particularly in Limarí Valley was identified (Meza, 2013). Also, was found that a rainfall deficit of 40% had a return interval of approximately once every 4 years in the northern, semi-arid Coquimbo region of Chile and a 22-year return interval in the more humid O'Higgins regions of central Chile (Núñez, et al., 2011). Lately, Zambrano et al. (2016) evaluated agricultural drought using satellite-based vegetation index data and found that in the Bío-Bío region (South-Central Chile) over the last sixteen years had experienced three drought event during the 2007–2008, 2008–2009 and 2014–2015 growing seasons. All these studies were carried out using precipitation data obtained from a limited number of weather stations across Chile. These results could be extended both on the spatial and temporal scale to improve the findings using accurate satellite long-term precipitation datasets.

Satellite datasets are becoming increasingly important to fill in the spatial and temporal data gaps for climate-based applications such as drought monitoring. Several global remotely sensed datasets now have historical records spanning 18 years or more, and lately the long-term products having more than 30 years which are appropriate for climate studies and represent a viable information source in many parts of the world. One widely used remotely sensed precipitation dataset has been acquired by the Tropical Rainfall Measuring Mission (TRMM), which is jointly supported by National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA). The TRMM precipitation datasets (Huffman et al., 2007) spans since November 1997 until present, although the mission comes to its end on April 2015, but thanks to its successor the Global Precipitation Measurement (GPM) which is the continuity of TRMM the dataset have continued. The Global Precipitation Measurement (GPM) is an international satellite mission to provide next-generation observations of rain and snow worldwide starting from temporal resolution of three hours and at spatial resolution as high as 0.1–30 min. Other precipitation datasets have been produced using a combination of infrared (IR) and passive microwave (PMW) observations from multiple satellite sensors using different precipitation estimation methods. These include Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) (Hsu et al., 1997) and Climate Prediction Center Morphing (CMORPH) technique (Joyce et al., 2004). Most products have short-term data and have been evaluated in different part of the world such as: South America (Salio et al., 2015), Colombia (Dinku et al., 2009), Saudi Arabia (Almazroui, 2011), Greece (Nastos et al., 2016), Ethiopia (Duan and Bastiaanssen, 2013), China (Guo et al., 2016b), India (Shah and Mishra, 2015), Iran (Moazami et al., 2016) and Himalayas (Bharti and Singh, 2015) and many other more (Dinku et al., 2007; Kenawy et al., 2015; Pipunic et al., 2015; Tan et al., 2015).

The study of climate change and climate variability requires a long-record data to permit the evaluation of climate and associated natural disasters like drought. The National Research Council (NRC) defined the Climate Data Records (CDR) as time-series measurements of sufficient length, consistency, and continuity to determine variability and climate change (National Research Council, 2004). Thus, two new satellite products for long-record precipitation studies were considered in this study. They are the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Climate Data Record (PERSIANN-CDR) (Ashouri et al., 2015) and CHIRPS 2.0 (Funk et al., 2014) datasets which both have more than 30-year records of data. Both products represent potentially valuable data sources for monitoring drought in data-limited countries such as Chile because they have an adequate length of record to detect and quantify drought conditions within a longer historical context. These products are relatively new and still there are only a limited number of studies evaluating their performance of estimating the amount and spatial distribution of precipitation. Miao et al. (2015) evaluated PERSIANN-CDR over China and found that the agreement between the dataset with in situ measurements in dry regions is not strong. Ashouri et al. (2015) tested PERSIANN-CDR during the Hurricane Katrina (2005) and the flooding on Sydney, Australia (1986); and found in both that PERSIANN-CDR is performing reasonably well when compared to radar and ground-based

observations. However, Ashouri et al. (2015) also examined the frequency distribution of precipitation from PERSIANN-CDR as compared to those of CPC gauge observations and TMPA v7 concluding that generally PERSIANN-CDR tends to underestimate the frequency distribution. Lately, Guo et al. (2016a) analyzed PERSIANN-CDR for the assessment of meteorological drought over China using ground-based gridded China monthly Precipitation Analysis Product (CPAP) from 1983 to 2014, their results shows that 6-month SPI has the best agreement with CPAP in identifying drought months; however, large differences between PERSIANN-CDR and CPAP in depicting drought patterns and identifying specific drought events were found over northwestern China. Katsanos et al. (2016) evaluated CHIRPS over Cyprus for a 30-year period and their results showed good correlation with in situ measurement with an overestimation noted during the decade 2001–2010 possibly due to the incorporation in the latter of TRMM estimates, which tend to overestimate rainfall (Katsanos et al., 2016).

The aim of this study is to evaluate the performance and fit of monthly, long-term satellite-based precipitation products over Chile for mapping and quantifying historical rainfall and drought patterns. In this study, the CHIRPS 2.0 and PERSIANN-CDR datasets are compared with a well-studied satellite-based product, the TMPA 3B43 v7 and with in situ measurements obtained from weather stations across Chile. The Chilean territory was divided in five zones primarily based on climate characteristics to evaluate the performance of these datasets regionally across the country. The goal was to evaluate the accuracy and applicability of these products to characterize rainfall patterns across Chile and transform the data into a precipitation-based drought index, the Standardized Precipitation Index (SPI, McKee et al. (1993)) technique, for agricultural drought monitoring in Chile.

2. Study area

According to the Köppen climate classification system (Kottek et al., 2006), Chile has five primary climate regimes that include the cold desert climate (Bwk) in the North, temperate Mediterranean climate (Csb) in the Central part, temperate oceanic climate (Cfb) in the South-Central, cool oceanic climate (Cfc) in the South and tundra climate in the austral part of the extreme South. To regionally evaluate the satellite-derived precipitation data with in situ measurements, the Chilean territory (for areas north of 50° South latitude) were divided into five zones that capture geographic variations in climatic conditions. The five zones were geographically defined as: (1) North (17.6° S to 28° S latitude), (2) North-Central (28° S to 32° S latitude), (3) Central (32° S to 36° S latitude), (4) South-Central (36° S to 40° S latitude) and (5) South (40° S to 50° S latitude). Fig. 1a shows the five zones with the 278 weather stations locations used in this study. Annual rainfall is mostly below 80 mm in the North zone, below 100 mm in the North-Central zone, around 800 mm in the Central zone, from 1000 mm to 1500 mm in the South-Central zone and reaching up to 2000 mm in the South zone. To describe vegetative features of Chile, the land cover MCD12Q1 product with the scheme IGBP was used (Friedl et al., 2010). The land cover of the study area is shown in Fig. 1b. The North zone is mostly barren and in the North-Central region is dominated by shrubland with some isolated cropland along the river valleys. Agriculture in Chile is primarily concentrated from the 32° S to 40° S in the Central and Central-South zones where forest, cropland and grassland are the dominant land cover types. Between 40° S and 45° S there are mainly forest and some areas covered with cropland which are close to the 40° S. Finally, to the South of 45° S there are a mixed vegetation of forest and grassland as the principal land cover type. Fig. 1c shows the altitude difference through Chile, highest altitudes are in the North and North-Central zones,

range mainly from 700 m to more than 3500 m. In the Central and South-Central zones is possible to note the valleys in the middle part with altitudes between 100 m and 700 m. South zone close to 40° S latitude has topography around 10 m to 200 m. Also, toward eastern from 40° S to the North the highest altitudes corresponds to the Andes Mountains.

3. Methods

3.1. Data

Monthly time-series of precipitation (mm month^{-1}) from the three satellite products and the weather station-based in situ rainfall measurements were compared and analyzed. The first satellite product evaluated was the 0.25-degree, spatial resolution TMPA products with near-global (between 50° S and 50° N) coverage. The standard monthly TMPA product 3B43 version 7 was already temporally aggregated to a monthly time-step and the data required no additional temporal modifications prior to analysis. The TMPA data record spanned a period from January 1998 to present (Huffman et al., 2007, 2010). The second satellite product was the PERSIANN Precipitation Climate Data Record (PERSIANN-CDR), which is a daily quasi-global precipitation product covering the period starting from January 1983. The PERSIANN-CDR data covers from 60° S to 60° N latitude and 0° to 360° longitude at 0.25° spatial resolution (Ashouri et al., 2015). The third satellite products evaluated was the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) dataset, which has a 30+ year, quasi-global rainfall dataset, available at the monthly, dekad, pentad and daily time steps. CHIRPS spatial data coverage spans from 50° S to 50° N (over all longitudes) with the data record extending from 1981 to present. CHIRPS data have a 0.05° spatial resolution with the satellite data calibrated with in situ station data to create gridded rainfall time-series appropriate for trend analysis and seasonal drought monitoring (Funk et al., 2014).

The General Water Directorate of Chile (DGA) has the densest weather stations network through Chile and with the longest historical records too. The network consists of stations with observers, which, recorded on a daily basis each data. Also, it has datalogger equipment to store data and transmit to a central database. About the quality control, the first verification is performed regionally for experienced staff which checks coarse errors and data gaps. Then, the quality control is made by spatial consistency comparing rainfall measurement with weather stations of neighboring sub-basin of the same time interval. Distance and altitude considerations are included in the selection of neighboring stations. The effect of an air mass over the area and the impact on the weather stations values are analyzed. For measurements validation, the analyst experience allows discovering the errors of direct readings from an observer or datalogger malfunction.

For the evaluation of satellite precipitation products, ground observation of rainfall data were obtained from the database of the Center for Climate and Resilience Research (CR2) at the University of Chile (<http://www.cr2.cl>). The data-base consisted of 780 stations with monthly rainfall data collected from 1940 to 2015. They are comprised of data records consolidated from the DGA and the Meteorological Directorate of Chile (DMC) for use in research. From the total of 780 available meteorological stations with in situ data, only 278 stations were selected for analysis in this study based on the criteria that a station must have at least a 25 year rainfall data record (1981 to 2015), less than 10% missing data, and be geographically located between 17° S and 50° S in continental Chile. The dramatic reduction in the number of stations that could be used in this analysis illustrates the relative lack of long-term, temporally-complete rainfall records from in situ gauges that are available in Chile.

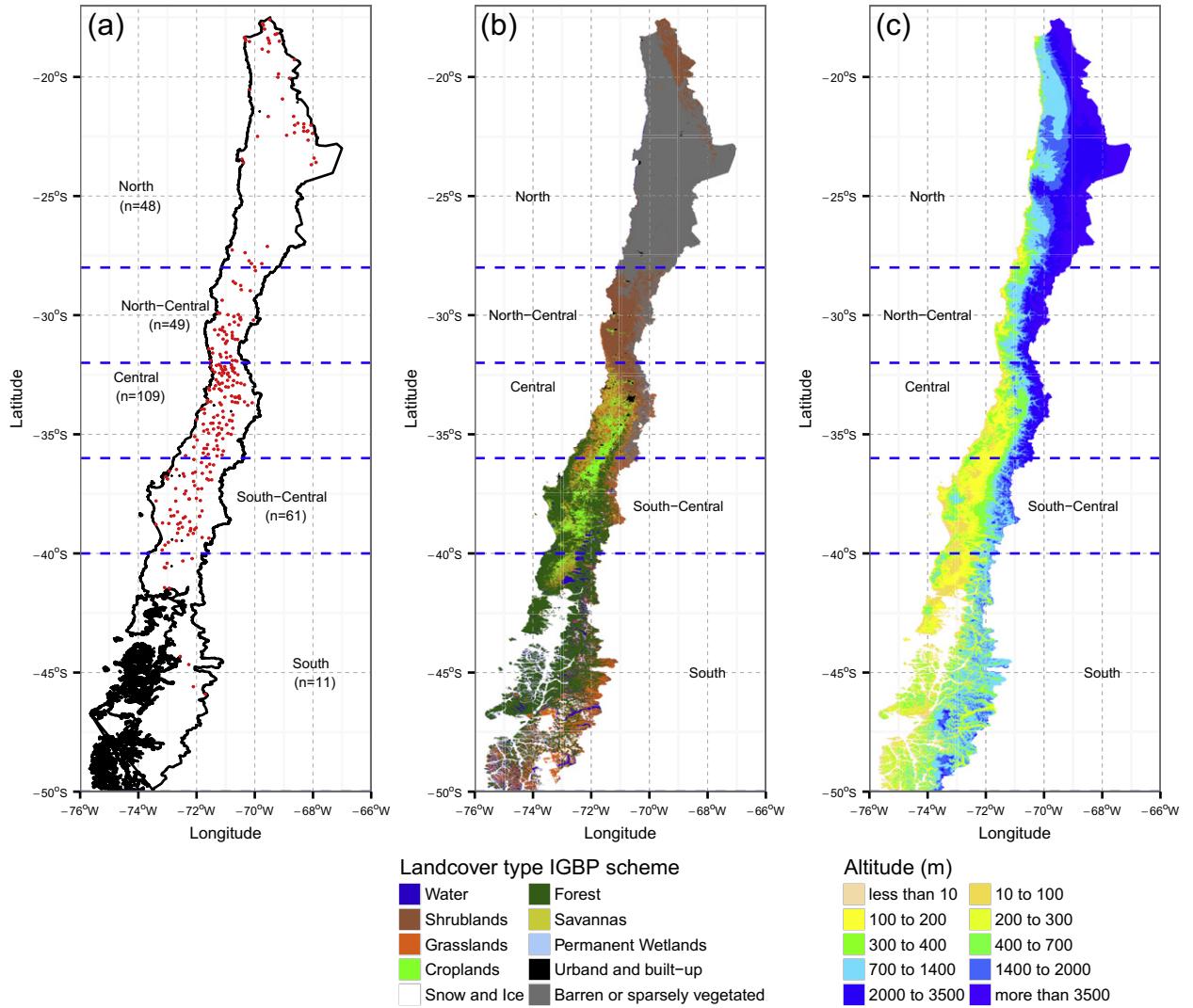


Fig. 1. Study area map with (a) 278 weather stations, (b) landcover type and (c) terrain elevations.

3.2. Preparation and data analysis

With the daily precipitation data from the product PERSIANN-CDR from January 1983 to June 2015, monthly cumulative precipitation was calculated obtaining a total of 402 precipitation grids. In the case of CHIRPS 2.0 and TMPA 3B43 v7, these datasets are in monthly accumulated precipitation. For TMPA 3B43 v7, the data spanned from 1998 to 2015 and 210 data grid were used between January 1998 and June 2015. For CHIRPS 2.0, a total of 414 data grids were used from January 1981 to June 2015. Monthly precipitation values across the historical record were extracted from the grid cell locations that corresponded to the 278 weather stations locations from the TMPA 3B43 v7, CHIRPS 2.0 and the PERSIANN-CDR precipitation datasets, respectively. The extracted precipitation data from each remote sensing and in situ dataset were then spatially averaged with the five regional zones defined for the regional analysis part of this study. The aggregation was done by averaging the values of precipitation (in situ and satellite) for each monthly period for all weather station locations within each zone. The aggregated data were then statistically analyzed using box-plot and heatmap graphs, as well as summary statistics that included mean, median, first quartile ($Q_{25\%}$), third quartile ($Q_{75\%}$), maximum value and number of observations (n).

The monthly precipitation time-series data were compared for each zone and between the satellite products and in situ measurements. Using the spatially-aggregated data, the monthly difference between satellite products and in situ data for each year was calculated to measure the level of underestimation and overestimation among datasets. To complement this analysis, the data were also temporally aggregated calculating the mean cumulative season for winter, summer, spring and autumn time periods to determine if there were differences in the inter-relationships among the datasets at different times of the year.

To further evaluate and compare the time-series precipitation datasets nine statistics were calculated for each of the 278 weather station locations that included: lineal coefficient of correlation (CC), the magnitude of underestimation (ME) (Eq. (B.1)), mean absolute error (MAE) (Eq. (B.2)), multiplicative bias ($bias$) (Eq. (B.3)), efficiency (E_{ff}) (Eq. (B.4)), the Root Mean Square Error ($RMSE$) (Eq. (B.5)), frequency bias (FBS) (Eq. (B.6)), the probability of detection (POD) (Eq. (B.7)) and false-alarm ratio (FAR) (Eq. (B.8)). These station-based statistics were then aggregated for each zone generating 135 statistics in total among the five zones. These statistics were analyzed using hierarchical cluster analysis, singular value decomposition (SVD) and k-mean to better understand their similarities and differences in characterizing spatial and temporal rainfall patterns. To

evaluate the behavior during the year, the statistics were monthly calculated, obtaining 1620 statistics (i.e., 12 months \times 9 statistics \times 3 products \times 5 zones), which were also analyzed using hierarchical cluster, k-means and singular value decomposition. Also, we used the Root Mean Squared Error (RMSE) statistic to have a measure of the error between the in situ time-series of rainfall and SPI values against those derived from satellite products.

Finally, to analyze the application of these satellite products and evaluate the spatial variability of agricultural drought patterns, the CHIRPS 2.0 precipitation data were used to calculate the SPI at time-scales (i.e., one, three and six months SPI) for the period of 1981 to 2015. For SPI calculation, a Gamma distribution was adjusted using a method for parameter fitting based on unbiased Probability Weighted Moments (Vicente-Serrano et al., 2010). To compute the SPI the 'SPEI' (Beguería and Vicente-Serrano, 2013) package within the R environment (R Core Team, 2016) was used.

4. Results

4.1. Satellite and rain gauge precipitation

The summary of monthly rainfall and distribution comparison for 278 selected weather stations with extracted data from the grids of satellite precipitation products estimates are presented in Table 1, Fig. 2a and b. In the North, the monthly precipitation is mainly under 200 mm and to the South is possible to observe the winter pattern where the amount of rainfall increase to more than 200 mm in most locations and as high as more than 400 mm in the South. There are some extreme precipitation values (i.e., more than 1200 mm) from the Central zone and southward. This extreme was 1232 mm on May 1981, 1241 mm on July 1987, 1236 mm on July 1987, 1373 mm on June 2000, 1394 mm on June 2000 and 1258 on July 2001 for the weather stations of Rio Malleco y Vergara, Rio Maule Medio, Rio Loncomilla, Rio Maule Medio, Rio Loncomilla and Rio Bío-Bío Alto, respectively. The seasonal variations are also evident for the weather station located in the central and northern zones as shown by the alternating seasonally higher (red color) and lower (green color) for each year in the multi-year historical record presented.

Annual drier periods also were more pronounced across the stations moving northward from central to northern parts of Chile.

The boxplots of Fig. 2b compare the statistical distribution of rainfall data between in situ and satellite products for the five zones. In the North zone, all three products overestimate the distribution of precipitation as shown by Fig. 2b where the box of three satellite products is higher than the box for in situ measurements. In situ rainfall totals for 75% of the stations in the North were below 0.5 mm, whereas for the three satellite rainfall products the totals ranged between 0.84 mm and 12.44 mm as shown in Table 1 and Fig. 2b. In the North-Central zone, the satellite products overestimate the lower precipitation values with $X_{25\%}$ equal to 0.52 mm in the case of TMPA 3B43 v7 compared to the in situ data that equaled 0 mm, but the three satellite products capture most of the remaining rainfall distribution across the intermediate and higher values. In the Central zone, the three satellite products had a $X_{75\%}$ between 53 mm (CHIRPS 2.0) and 64 mm (PERSIANN-CDR) close to in situ values which has 54 mm. However, like in the North-Central zone, the lower rainfall values ($X_{25\%}$) for the PERSIANN-CDR and CHIRPS 2.0 datasets were approximately 5 mm and TMPA 3B43 v7 was 1.8 mm compared to 0 mm for in situ data (see Fig. 2a and Table 1). The South and South-Central zones show rainfall distributions very similar between satellite products and in situ measures, in the range for interquartile range (IQR) of 29 mm to 178 mm for South-Central and between 50 mm to 200 mm in the South zone (see Fig. 2b and Table 1).

4.2. Time series, annual difference and seasonal variation of rainfall

Fig. 3 shows the time series of the spatially-averaged precipitation datasets for the five zones. Like the previous results, the values of in situ with satellite products had the close agreement in the Central zone as compared to the North zone, where the largest differences were found, particularly when there is low precipitation near 0 mm. The South-Central and South zones exhibited good correspondence between in situ and satellite data across the range of precipitation totals. Values of RMSE for the North zone were around 7 mm for all three products, in the North-Central TMPA 3B43v7 shows the lowest values with 12.2 mm and CHIRPS 2.0 the highest with 20.8 mm. Similar values had PERSIANN-CDR and CHIRPS 2.0 in the Central zones

Table 1

Summary statistics for precipitation data from weather stations (in situ) and PERSIANN-CDR, TMPA 3B43 v7, CHIRPS 2.0 satellite products, in five zones of Chile. Total number of observations (n), weather stations by zone (Stations), missing observations, mean (\bar{X}), first quartile ($X_{25\%}$), median ($X_{50\%}$), third quartile ($X_{75\%}$) and maximum (X_{max}). The time-period used was 1981–2015, 1983–2015 and 1998–2015 for in situ, CHIRPS 2.0, PERSIANN-CDR and TMPA 3B43 v7, respectively.

Summary statistics									
Zone	Product	n	Stations	Missing	\bar{X} [mm months $^{-1}$]	$X_{25\%}$ [mm months $^{-1}$]	$X_{50\%}$ [mm months $^{-1}$]	$X_{75\%}$ [mm months $^{-1}$]	X_{max} [mm months $^{-1}$]
North	in situ	19,872	48	473	6.95	0.00	0.00	0.50	413.40
	PERSIANN-CDR	18,720	48	0	11.90	1.24	4.50	12.44	213.64
	TMPA 3B43 v7	10,080	48	0	10.19	0.84	3.31	10.23	284.99
	CHIRPS 2.0	19,872	48	0	10.47	1.43	3.90	10.93	221.72
North-Central	in situ	20,286	49	226	12.40	0.00	0.00	7.50	597.50
	PERSIANN-CDR	19,110	49	0	18.85	2.02	8.25	24.10	390.18
	TMPA 3B43 v7	10,290	49	0	16.26	0.52	3.23	16.03	568.67
	CHIRPS 2.0	20,286	49	0	10.72	1.23	3.41	11.55	220.66
Central	in situ	45,126	109	853	46.06	0.00	9.20	54.20	1394.10
	PERSIANN-CDR	42,510	109	0	45.99	5.62	22.12	64.29	504.89
	TMPA 3B43 v7	22,890	109	0	45.33	1.80	14.91	61.71	517.03
	CHIRPS 2.0	45,126	109	0	44.34	5.22	14.74	52.53	787.23
South-Central	in situ	25,254	61	500	124.93	29.00	81.00	177.60	1258.60
	PERSIANN-CDR	23,790	61	0	85.49	19.13	55.87	127.58	554.02
	TMPA 3B43 v7	12,810	61	0	105.30	28.26	77.21	159.97	688.65
	CHIRPS 2.0	25,254	61	0	121.81	34.80	82.31	170.37	1238.03
South	in situ	4554	11	130	145.56	57.18	114.90	201.12	1139.00
	PERSIANN-CDR	4290	11	0	107.73	50.44	86.80	147.28	441.86
	TMPA 3B43 v7	2310	11	0	142.40	69.64	125.33	197.01	594.76
	CHIRPS 2.0	4554	11	0	128.34	56.29	102.70	170.85	827.94

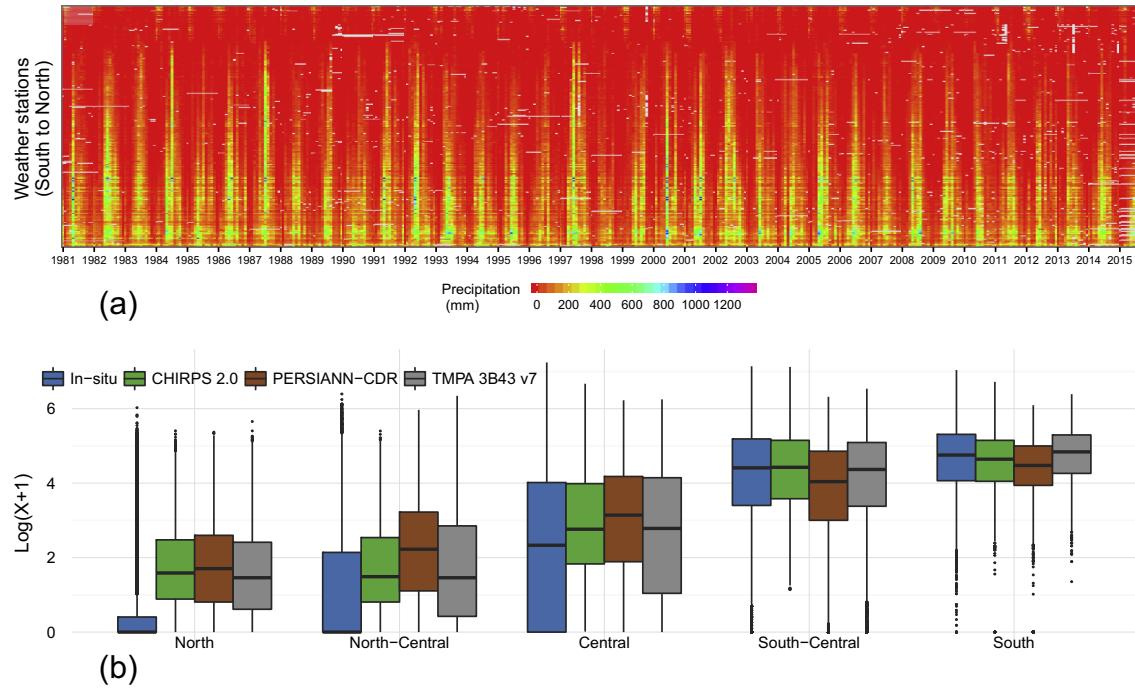


Fig. 2. (a) Monthly rainfall data from 1981 to 2015 for 278 weather stations, and (b) comparison of distribution for precipitation data with logarithmic transformation, from in situ and satellite products TMPA 3B43 v7, PERSIANN-CDR and CHIRPS 2.0; for North, North-Central, Central, South-Central and South zones in Chile. The y-axis presents the 278 weather stations from North (top) to South (bottom), and the x-axis presents the period from 1981 to 2015. The white spaces represent missing data.

reaching around 20 mm. In the South-Central and South zones, TMPA 3B43 v7 has the lowest RMSE value with 37.6 mm and 24.2 mm, respectively; and PERSIANN-CDR reaches the highest values with 54.8 mm and 48 mm, respectively (Fig. 3).

Fig. 4a shows the yearly averaged values of monthly precipitation difference between the three satellite products (S) and the in situ rain gauge (G) data for the five regional zones. In the North zone, the rainfall totals were consistently overestimated by as much as 10 mm in all years except 2000 and 2001. This pattern was also observed for all satellite products over the North zone. The North-Central and Central zones had a similar rainfall pattern, in most of the years with satellite products overestimating rainfall. The CHIRPS 2.0 data had the lowest difference values and the best fit among the remotely sensed products, but there were still some anomaly years (1987 and 1997) in which CHIRPS 2.0 data had differences by as much as 20 mm in North-Central and Central zones where the mean annual precipitation are around 80 mm (~25%) and 500 mm (~4%), respectively. In the South-Central and South zones, the satellite products consistently underestimated monthly precipitation with differences as high as 69 mm and 83 mm, respectively; in the case of PERSIANN-CDR which correspond as much as 5% for both. In contrast, CHIRPS 2.0 has the best correspondence in these zones with in situ measurements difference in South-Central and South zones of 20 mm (1%) and 47 mm (2%), respectively (Fig. 4a).

In addition, a comparison of the seasonal variation of precipitation between products for each zone is presented in Fig. 4b. The greatest difference with in situ data was during the Spring season in the North zone, where each product had accumulated seasonal precipitation values higher than the in situ measurements. Also, PERSIANN-CDR and CHIRPS 2.0 overestimates, while TMPA 3B43 v7 underestimates seasonal rainfall. Over the North zone, the precipitation amount is higher during the Summer season, which corresponds to the commonly known as 'Bolivian winter' (Romero et al., 2013). In the North-Central zone, there was a significant difference with

in situ measurements, CHIRPS 2.0 underestimates by 30 mm and PERSIANN-CDR overestimates by 10 mm during the Winter season and in the other zones, the behavior is similar with those found in the North zone. However, in the other three seasons (i.e., Spring to Autumn), CHIRPS 2.0 had the best agreement with in situ data while PERSIANN-CDR highly overestimates seasonal rainfall. For the Central, South-Central and South zones, CHIRPS 2.0 had the best agreement for all seasons with the lowest precipitation in Summer and the highest in Winter (see Fig. 4b), followed by PERSIANN-CDR and ending TMPA 3B43 v7 with the lowest seasonal agreement.

4.3. Statistics of comparison between satellite-derived and in situ measurements

The statistics used to compare the different precipitation datasets were calculated for each 278 meteorological stations and the average values of the satellite products over the five zones are presented in Table 2. Hierarchical cluster analysis, k-means and singular value decomposition were used to analyze these datasets to better understand their similarities and differences in characterizing rainfall patterns across Chile. As an important source of variation monthly precipitation in rain gauges (G) was included to the 9 statistics for the analysis. The results revealed no significant difference between the products from North to South-Central zones and differences exists between them in the South zone. The singular value decomposition indicated that the CC, E_{ff} , POD, G and RMSE are the major contributors to the variation among these datasets. The statistics showed that differences were greater in areas that receive lower precipitation than areas with high rates of precipitation. For example, the South and South-Central zones, which normally receive between 1500 mm and 1700 mm of precipitation annually instead of 80 mm and 150 mm in the North and North-Central zones. The South and South-Central zones had the best results reflected by the FAR and high POD

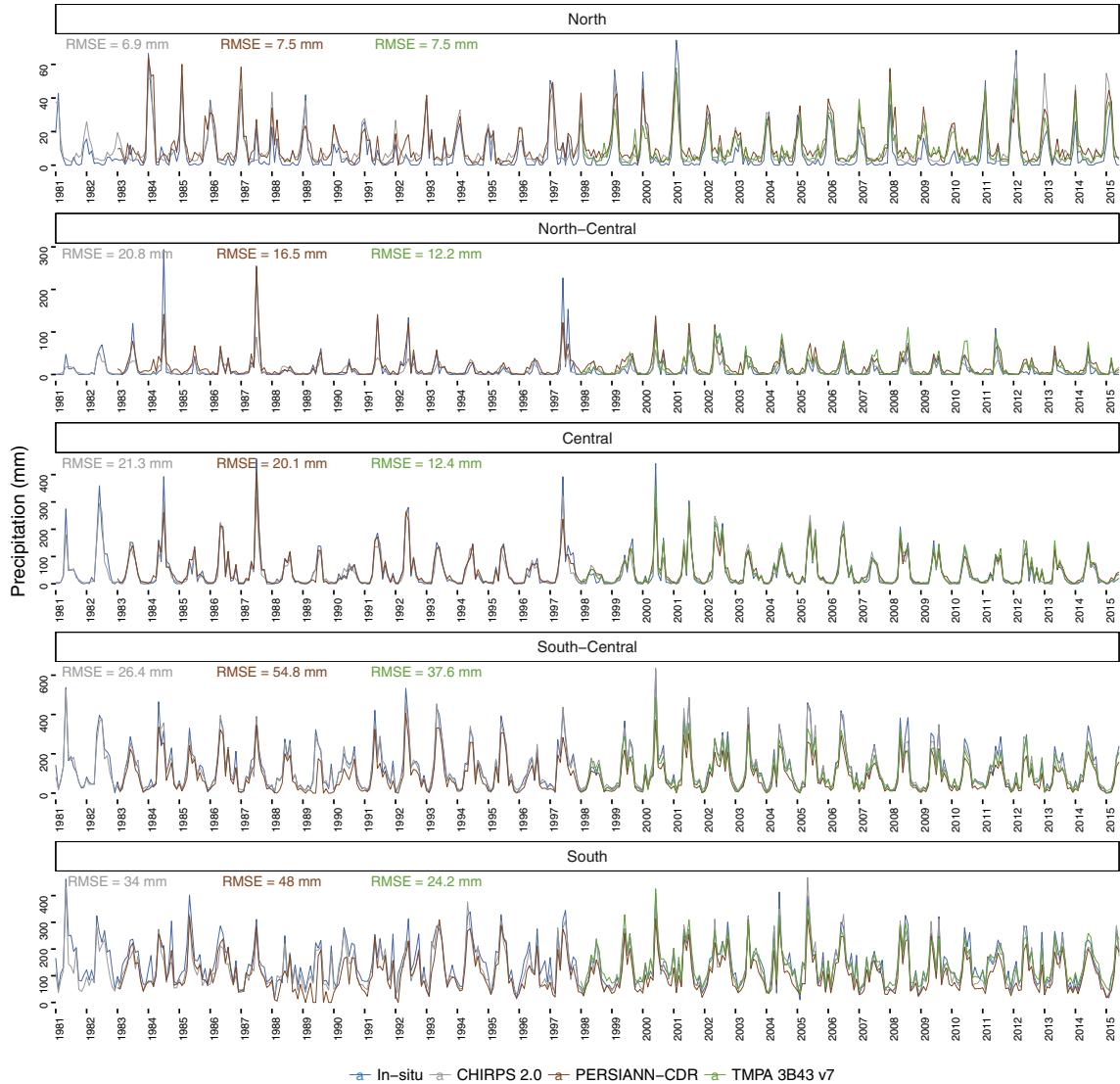


Fig. 3. Spatial-averaged monthly time-series of in situ rainfall data and extracted from satellite products TMPA 3B43 v7, PERSIANN-CDR, CHIRPS 2.0 in five zones of Chile: North, North-Central, Central, South-Central and South.

values among the zones, as well as the highest CC and good E_{ff} values. PERSIANN-CDR in the South and South-Central zones presents higher underestimations reaches -36 mm of ME in both zones. Cluster analysis identifies two big groups, one corresponding to North zone, and the second from North-Central to South zones. This reflects that North zone (group 1) presents the lower agreement with in situ measurements and this was similar for the three satellite products. The second group has two main sub-groups, one for North-Central and Central, showing better results than in North zone, and the second sub-group with the higher performance for the three satellite products showed in South-Central and South zones.

In order to further analyze the goodness of fit of the satellite products during the year, several monthly statistics were also studied. Fig. 5 presents the heatmap diagrams for the hierarchical cluster analysis performed by month for five zones and 9 statistics along with in situ precipitation (G) that was added as an additional measure of accuracy variation. For Fig. 5, the top horizontal axis represents the dendrogram cluster by statistics and the left vertical axis corresponds to dendrogram cluster by specific month. Also the vertical color palette between the left vertical dendrogram and

the first column of the heatmap shows the specific remote sensing precipitation product that corresponds to each row. The colors are gray, brown and green, corresponding to CHIRPS 2.0, PERSIANN-CDR, and TMPA 3B43 v7, respectively. As example, in Fig. 5a, in first row, the left dendrogram indicates that was included in one of the main clusters, next the green color in the vertical palette indicate that correspond to TMPA 3B43 v7 product, then the nine cells shows the nine scaled statistics values, beside that on the right the month is showed, in the case of the first row was 'November'.

The major contributors statistics in the North zone were ME, MAE, bias, FBS and FAR as showed in the horizontal dendrogram in Fig. 5a. TMPA 3B43 v7 has the highest values among these statistics from September to November, and CHIRPS 2.0 in September; which collectively formed the first group with the lowest fit. In vertical dendrogram in Fig. 5a, the second and third group are also defined. The second group had the best seasonal results from December to March, which corresponds to the 'Bolivian winter', when higher precipitation are received during this Summer period over the North zone. The third group that corresponded to the April to November time period had the poorest fit among all products.

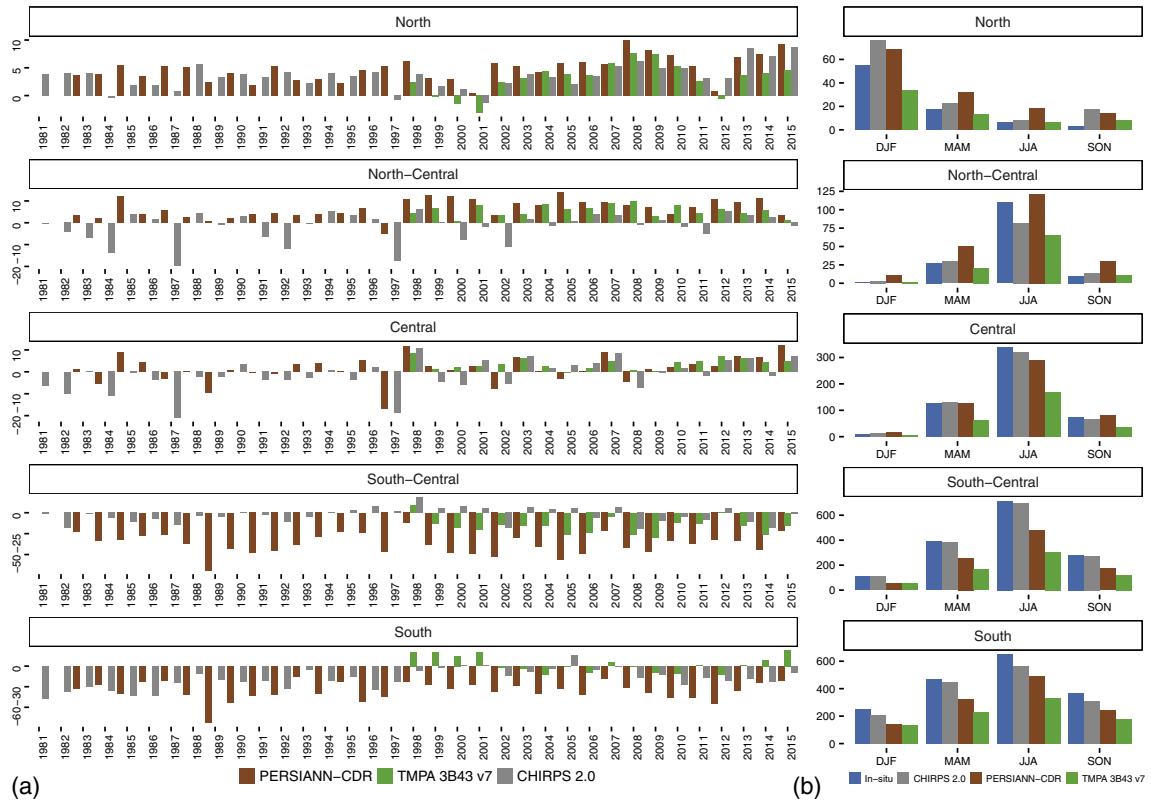


Fig. 4. (a) Annual averaged difference of monthly precipitation for in situ data and from satellite products TMPA 3B43 v7, PERSIANN-CDR and CHIRPS 2.0. (b) Seasonal comparison of precipitation for five zones of Chile and between in situ, CHIRPS 2.0, PERSIANN-CDR and TMPA 3B43 v7. The seasonal periods were: March, April and May (MAM) for Autumn; June, July and August (JJA) for Winter; September, October and November (SON) for Spring; and December, January and February (DJF) for Summer.

In the ‘North-Central’ and ‘Central’ zones, the statistics that made major contribution to the variation were CC , E_{ff} , $RMSE$, G and POD as showed in Fig. 5b and c, respectively. For the North-Central zone (Fig. 5b), the TMPA 3B43 v7 and PERSIANN-CDR had the poorest fit and with the PERSIANN-CDR having the lowest fit in February. Collectively for these months, the precipitation datasets poor fit is shown by the high values of MAE , $bias$, FBS and FAR . The best fit was found during the period from May to August (mainly Winter season) for the second group across all the precipitation products. The third group had lower results particularly from September to April (Spring

and Summer seasons). The Central zone (Fig. 5c) exhibited similar results than North-Central zone, with a lower fit in December and February for CHIRPS 2.0 and PERSIANN-CDR, and during January for all three products. Also the Central zone, shows that from June to September had the best fit mainly for PERSIANN-CDR and CHIRPS 2.0; and the highest values of ME , $bias$, FBS , MAE and FAR showing lower agreement from October to May, similar for the three products.

The indicators that were most relevant for the South-Central zone were $RMSE$, G and POD ; and for the South zone were MAE , $RMSE$, G and FAR as presented in Fig. 5d and e. PERSIANN-CDR was grouped

Table 2

Summary of statistics aggregate for zones North, North-Central, Central, South-Central and South; for products TMPA 3B43 v7, PERSIANN-CDR and CHIRPS 2.0. CC, ME, MAE, bias, E_{ff} , FBS, POD, FAR and HSS.

Zone	Product	Statistics of comparison								
		CC	ME [mm months $^{-1}$]	MAE	bias	E_{ff}	RMSE [mm months $^{-1}$]	FBS	POD [%]	FAR [%]
North	CHIRPS 2.0	0.48	3.34	10.69	10.89	-44.44	15.32	7.31	0.87	0.73
	PERSIANN-CDR	0.59	4.64	8.86	9.33	-30.90	13.95	6.66	0.95	0.72
	TMPA 3B43 v7	0.51	2.97	9.20	9.43	-36.57	14.78	7.86	0.91	0.72
North-Central	CHIRPS 2.0	0.72	-1.71	0.86	0.82	0.43	24.20	2.14	0.95	0.54
	PERSIANN-CDR	0.80	6.45	1.07	1.63	0.57	20.99	2.38	0.99	0.56
	TMPA 3B43 v7	0.75	5.59	1.04	1.55	0.05	21.29	1.94	0.97	0.48
Central	CHIRPS 2.0	0.88	-1.50	0.51	1.01	0.72	38.25	1.57	0.99	0.35
	PERSIANN-CDR	0.91	1.16	0.55	1.19	0.73	37.53	1.48	0.98	0.32
	TMPA 3B43 v7	0.93	3.47	0.46	1.19	0.76	31.23	1.32	0.99	0.24
South-Central	CHIRPS 2.0	0.90	-3.00	0.31	1.01	0.76	58.49	1.05	1.00	0.05
	PERSIANN-CDR	0.91	-36.94	0.39	0.78	0.60	77.47	1.02	0.97	0.04
	TMPA 3B43 v7	0.91	-15.08	0.32	0.93	0.74	63.20	1.02	0.99	0.03
South	CHIRPS 2.0	0.81	-17.52	0.30	0.93	0.56	64.76	1.00	0.99	0.00
	PERSIANN-CDR	0.79	-35.45	0.42	0.89	0.23	82.60	0.98	0.97	0.00
	TMPA 3B43 v7	0.88	2.56	0.28	1.08	0.69	50.86	1.00	1.00	0.00

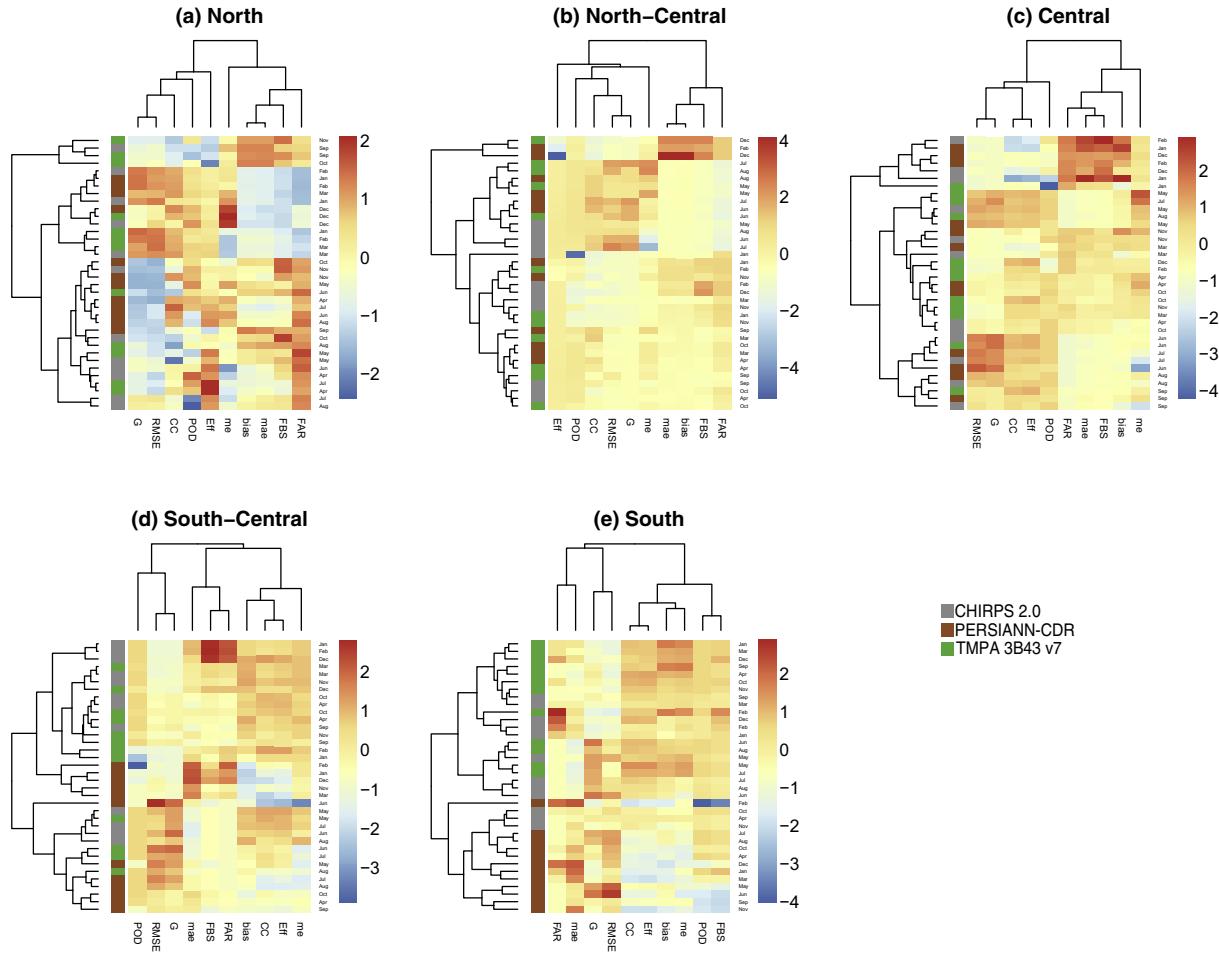


Fig. 5. Heatmap for hierarchical cluster analysis of monthly statistics of RMSE, G, CC, E_{ff} , POD, FAR, MAE, FBS, bias, and ME; for North, North-Central, Central, South-Central and South zones. In the left vertical axis, the dendrogram shows groups made by month. The top axis present the dendrogram by statistics. On the left of the column showing POD, there is a palette legend showing the satellite product at which each row correspond.

from November to January with high values of MAE, FBS and FAR; and low values of bias (indicating underestimation), CC, and E_{ff} showing the lowest agreement with in situ measurements in South-Central zone (Fig. 5d). CHIRPS 2.0 and TMPA 3B43 v7 had better results from September to March with high values of CC, E_{ff} , bias and ME. The best results for all three precipitation data products occurred from May to September. Finally, in the South zone (Fig. 5e) the major contributors were MAE, RMSE, G and FAR, and the cluster shows that PERSIANN-CDR had the lowest fit during all the year, similar was CHIRPS 2.0 in April, October and November. In this zone better result were found for CHIRPS 2.0 and TMPA 3B43 v7 with high values of CC, E_{ff} , bias (showing overestimation), POD and FBS, during the year.

Fig. 6 shows the analysis of the variation of the main statistics with respect to the amount of monthly rainfall. In the 'North' and 'North-Central' zones, low monthly precipitation below 5 mm had a clear impact on the linear correlation coefficient (CC), bias, FAR and FBS, compared to the other zones. In the zones where the days with rain are very limited, the ability of the satellite products to detect rain is reduced and often detected precipitation when there was no rain, which is reflected in high values of FAR that produce high values of bias and FBS too. Precipitation estimates were improved as shown in the statistical results as the rainfall increased, which was consistent across all zones. The 'South-Central' and 'South' statistics shows good agreement when compared to observed precipitation primarily

because of the higher monthly rainfall amount received over these areas. Further, there are considerably more days with rain than in the other zones, that has direct impact on the detection of rain, showing very high POD value of close to 1 and a FAR value close 0.

4.4. Spatial variation and comparison of products with long data-record

Spatial variations of the CC and E_{ff} for the 278 stations and the three satellite precipitation products across Chile are shown in Fig. 7 and compared with the spatial variation of in situ rainfall (G). The values of CC are very high in central Chile, mainly over 0.78 (Fig. 7a). The TMPA 3B43 v7 data had the highest correlation with values around 0.92 over this area with correlations decreasing in a northward direction for all products as seen in Fig. 7a. The E_{ff} was below 0.45 from 30° latitude and northward with values increasing from this latitude northward reaching peak value of 0.7 in central Chile (Fig. 7b) with similar values for all three satellite products. When these statistical results (CC and E_{ff}) are compared with the monthly average of precipitation (G) in weather stations, the spatial variation patterns in monthly rainfall are very similar, indicating that the fit of the precipitation products was strongly related to the amount of rainfall mainly for CHIRPS 2.0, followed by PERSIANN-CDR and TMPA 3B43 v7 with the greatest spatial pattern differences.

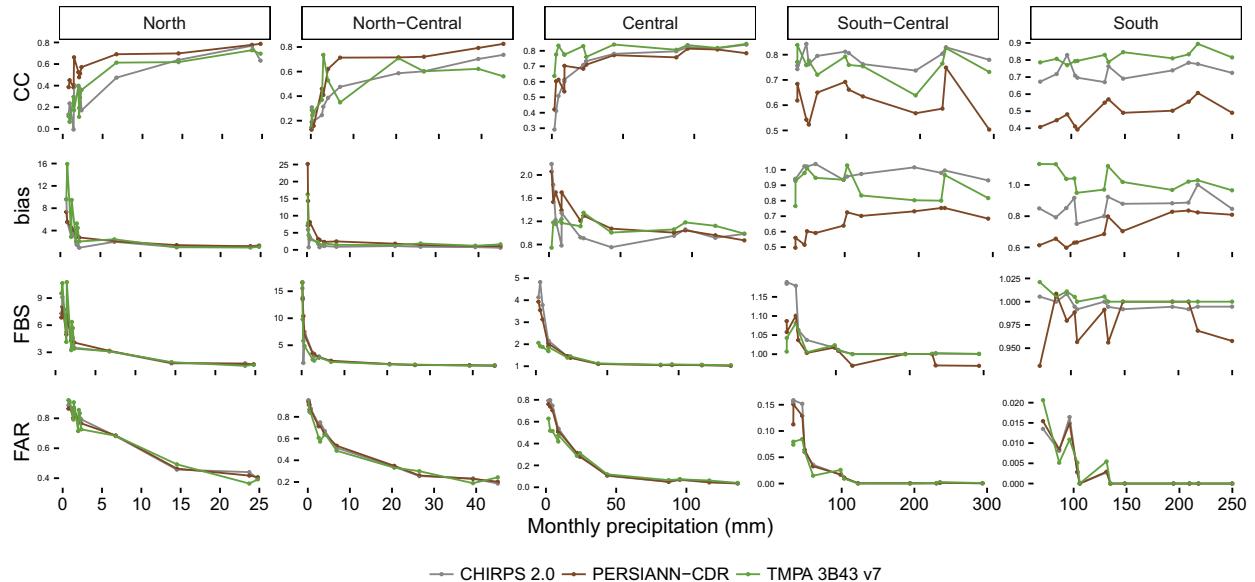


Fig. 6. Variation of statistics linear coefficient of correlation (CC), bias, false-alarm ratio (FAR) and frequency bias (FBS), with the amount of monthly rainfall, for North, North-Central, Central, South-Central and South zones of Chile and products TMPA 3B43 v7, PERSIANN-CDR and CHIRPS 2.0.

The climatology of monthly averaged precipitation maps from 1983 to 2015 were compared between the high spatial resolution, CHIRPS 2.0 and the coarse spatial resolution, PERSIANN-CDR in Fig. 8. Both products were compared to evaluate the difference on how well these products measure the spatial patterns variation of precipitation. TMPA 3B43 v7 was not included because its shorter

historical data, which has limited utility for climatological analysis. As would be expected, the high resolution CHIRPS 2.0 data captured more spatial variability in precipitation patterns than the PERSIANN-CDR. Months with the greatest spatial pattern differences between these datasets occurred in February, March, September, October, November and December; particularly in the southern part of Chile.

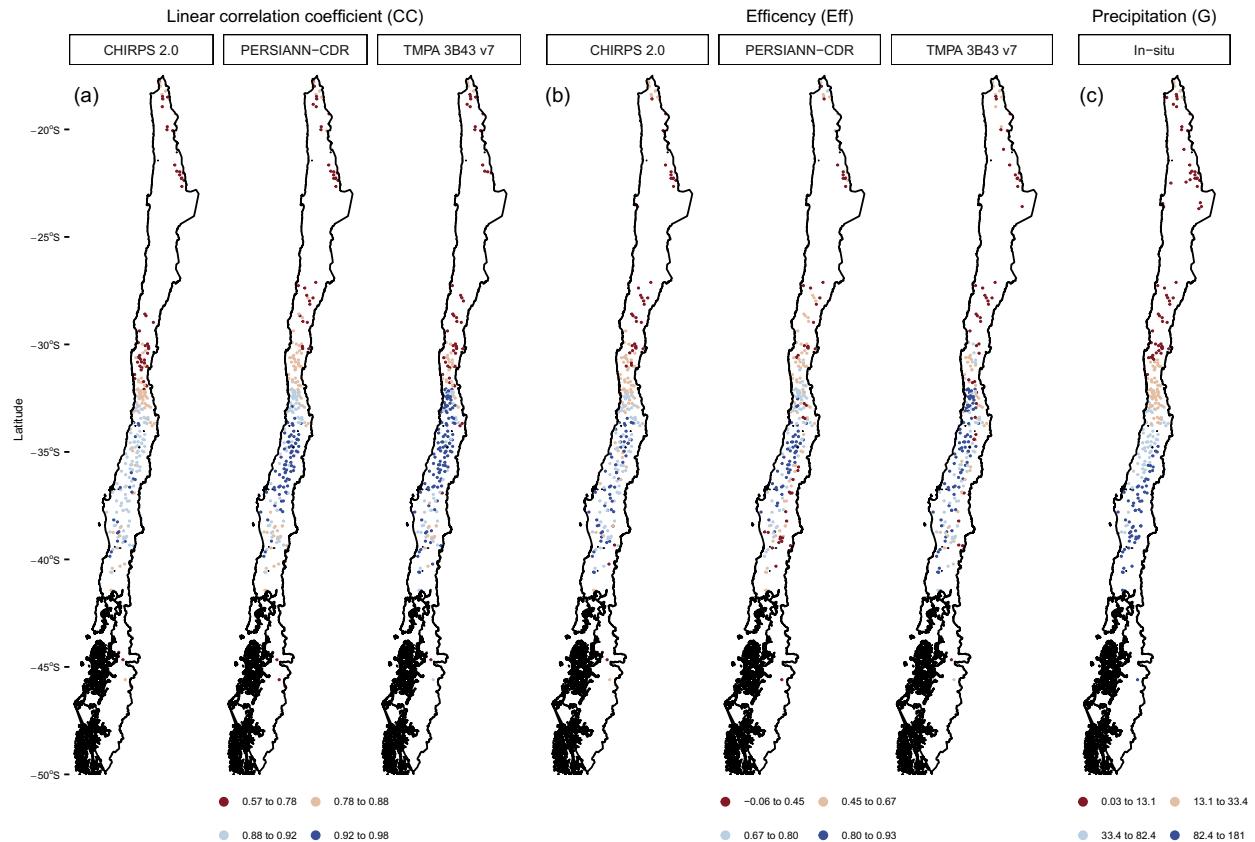


Fig. 7. Spatial variation of statistics of (a) linear correlation coefficient (CC), (b) efficiency (E_{ff}) for products CHIRPS 2.0, PERSIANN-CDR and TMPA 3B43 v7 and (c) In situ precipitation in 278 rain gauge (G) over Chile.

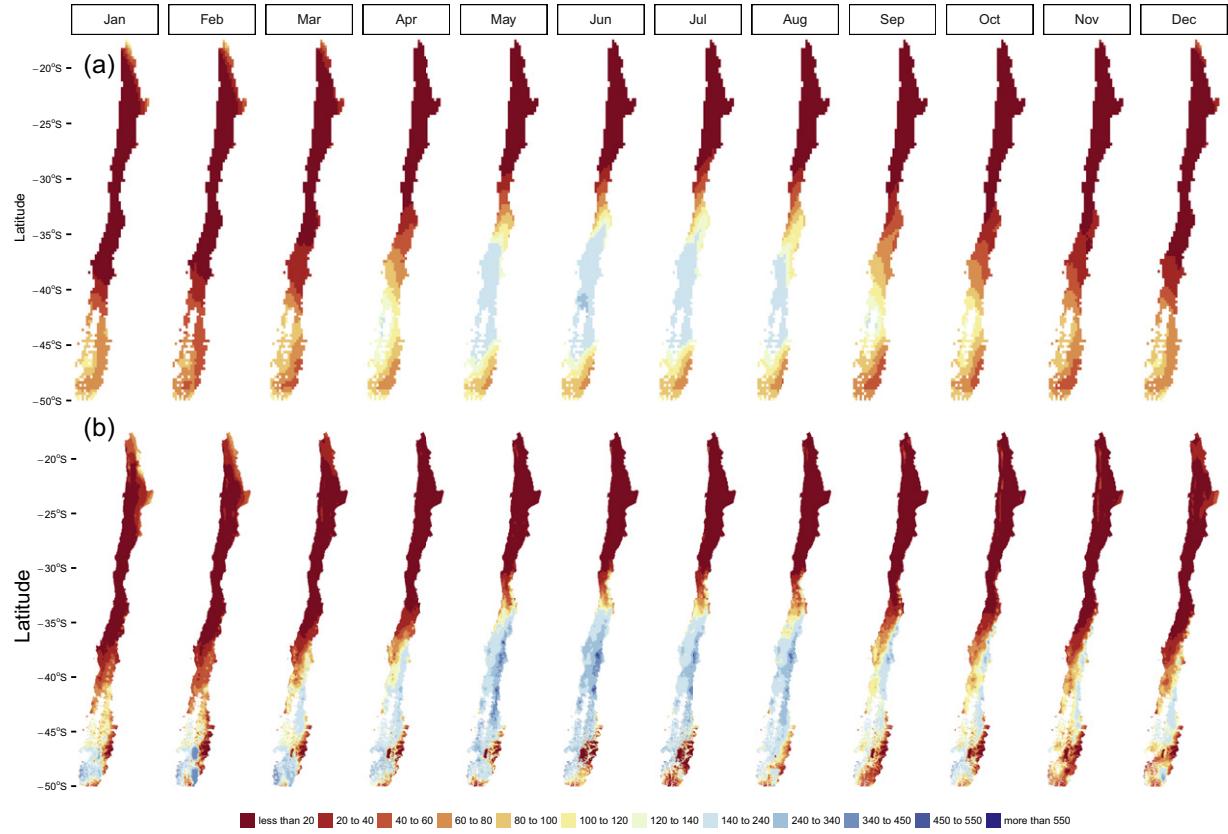


Fig. 8. Monthly satellite-derived precipitation in Chile during 1983 to 2015 for products (a) PERSIANN-CDR and (b) CHIRPS 2.0.

Most of the spatial differences in the precipitation patterns captured in these two products occurred mainly in the Central to South part of Chile. One of the notable discrepancies occurred in eastern Chile, where the Andes Mountains is located and few, meteorological stations are located at the higher altitudes. During May to August in eastern Chile near the 40° S latitude, precipitation can exceed 400 mm as estimated for CHIRPS 2.0, compared to 200 mm estimated by PERSIANN-CDR. In the South between 45° S and 50° S in far eastern Chile, CHIRPS 2.0 had lower precipitation values of less than 50 mm and PERSIANN-CDR estimated approximately 100 mm. It is well known that precipitation raises as altitude increases, further, [Garreaud \(2009\)](#) indicates that orographic air uplift produces 2–3 times more annual precipitation up in the Andes Mountains relative to the coastal values at the same latitudes. CHIRPS 2.0, captures the increasing of monthly precipitation climatology as showed in Fig. 8b during Winter months from May to July and toward the Andes Mountains between 35° and 45° South latitude. Fig. 8b shows that monthly rainfall raises up to 340 mm toward Andes Mountains and picking more than 550 mm during Winter. Those patterns were not captured by PERSIANN-CDR product as showed in Fig. 8a.

When the spatial variation of annual precipitation is considered for the maps presented in Fig. 9, together with the wettest month and driest month during the 30+ year historical period estimated from the satellite-based CHIRPS 2.0 and PERSIANN-CDR products, they were similar to the monthly averaged results in Fig. 8. The majority of differences occurred mainly south of 35° S and over the Andes Mountain range in far eastern Chile. The annual precipitation pattern showed in Fig. 9a is similar with those of Fig. 9b that shows the maximum monthly rainfall. This pattern reflects mainly the monthly

rainfall from May to August, also shows that from 35° S and northern there is the driest pattern and in the other direction from 35° S to the South, there is a wettest pattern. Moreover, CHIRPS 2.0 capture precipitation increasing with elevation toward the east, which is not reflected by PERSIANN-CDR. Fig. 9c, minimum monthly rainfall reflects the driest pattern of Spring and Summer seasons as shown in Fig. 8 (September to December and January to March, respectively), and in this case, the driest pattern moves it to the 40° S latitude. The areas southern to 35° S had higher precipitation variability and the amount of monthly rain is also high as depict maps of annual precipitation (Fig. 9a), wettest month (Fig. 9b) and driest month (Fig. 9c). CHIRPS 2.0 was found to capture the spatial variation of rainfall better than PERSIANN-CDR, as reflected in Figs. 9a, 8b and 9c, where the spatial pattern of rainfall are highly variable specially in the East part of Chile as showed by CHIRPS 2.0. In Central to South zones in Chile, the PERSIANN-CDR showed more homogeneous spatial variability compared to CHIRPS 2.0, where the gradient of precipitation are more coarse due to the low spatial resolution of PERSIANN-CDR, this is readily apparent in Fig. 9a and b.

4.5. Application for agricultural drought analysis

In this section, historical time-series of SPI data ([McKee et al., 1993](#)) derived from both the PERSIANN-CDR and CHIRPS 2.0 precipitation datasets were evaluated. This index was selected because is recommended for the World Meteorological Organization (WMO) as an index to characterize droughts ([Hayes et al., 2011](#)). TMPA 3B43 v7 was not included because of its short historical record. The SPI at 3

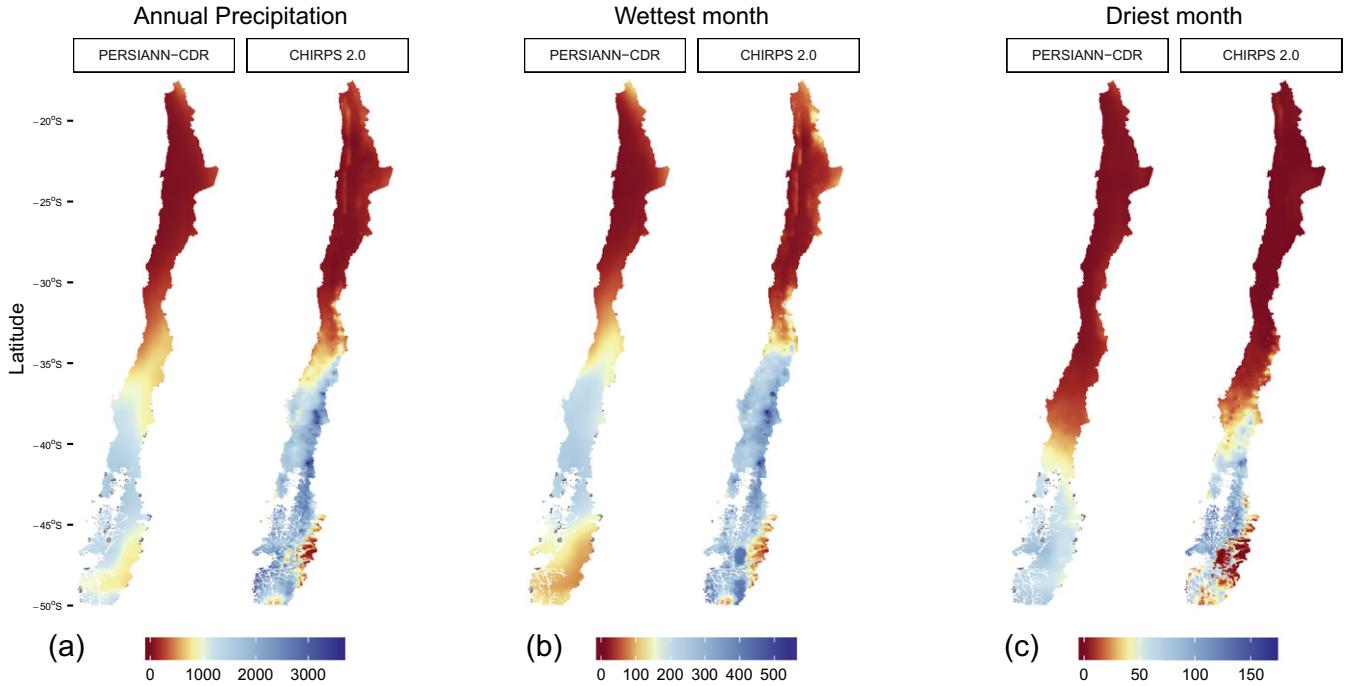


Fig. 9. Rainfall climatology for period 1983 to 2015 of (a) annual precipitation, (b) wettest month and (c) driest month, for satellite products CHIRPS 2.0 and PERSIANN-CDR.

time-scales (i.e., 1-, 3-, 6-month SPI) commonly associated to agricultural drought were produced and the correlation with SPI derived from weather station observations was evaluated. However, the analysis was mainly focused on the time-scales of three months or less because agricultural drought has generally short-term response. Further, in the South-Central zone of Chile, Zambrano et al. (2016) found the 3-month SPI (SPI-3) had the best correlation over cropland areas.

The SPI-3 times-series data is presented for the in situ, PERSIANN-CDR and CHIRPS 2.0 precipitation data in Fig. 10. The time-series line graphs represent the geographically aggregated SPI values over each of the five zones. Also, the RMSE indicator was calculated to evaluate the error for derived SPI between satellite and in situ values as showed in Fig. 11. From the results in the previous section, found the precipitation products had a better fit in the Central to South zones. Results in Fig. 11 supports these findings, showing that the lowest error was found in South-Central and Central zones for SPI-1, SPI-3, and SPI-6; and they were very similar for both satellite products. However, the lowest value of RMSE is 0.6 which is also high and could induce to find a place under drought when it is not.

The variation of SPI-3 during the growing season (September–April) from 1983 to 2015 were compared in Fig. 12, for the index data derived from the in situ, PERSIANN-CDR and CHIRPS 2.0 data. In the North and North-Central zones, major difference are apparent between the satellite and in situ derived SPI-3 data. Between 1988–1989 and 1996–1997 seasons, PERSIANN-CDR present 8 seasons with greatest differences, showing lower values than those from in situ SPI-3. On the other hand, from 2001–2002 to 2010–2011 CHIRPS 2.0 has highest differences with in situ SPI-3 values in 11 seasons in the North and North-Central zones. The SPI-3 bar plot calculated from the PERSIANN-CDR and CHIRPS 2.0 data for Central and southern zone had better fits with the in situ SPI-3, which is consistent with the previous results presented. However, large discrepancies with in situ SPI-3 were found for the PERSIANN-CDR-based SPI-3 data during 1989–1990 growing season. In 1989 the PERSIANN-CDR was found to largely underestimate precipitation as

shown earlier (see Fig. 4), which resulted in the larger discrepancies of SPI-3 that year. In the Central, South-Central and South zones, the 1998–1999, 2007–2008, 2008–2009 and 2014–2015 growing seasons exhibited the lowest SPI-3 values that were calculated in each of the three SPI-derived from the in situ, PERSIANN-CDR and CHIRPS 2.0, respectively. The SPI-3 results for these years are consistent with Zambrano et al. (2016), who found in the South-Central zone of Chile, severe drought occurrence during the 2007–2008, 2008–2009 and 2014–2015, note that the Zambrano et al. (2016) study did not evaluate the period 1981–1999 that low SPI values were calculated from the various precipitation datasets; however, the three later events were well represented in the SPI-3 time series over this area.

Because agricultural drought is associated with abnormal dryness over shorter-term time scales (<6 months), and considering the previous results that found high monthly variability of adjustment between satellite products and in situ data. Fig. 13 presents the monthly CC that compared the two satellite-based sets of SPI datasets with the in situ SPI data at time-scales of one, three and six months. For SPI-1, the results were very similar between the in situ and satellite-based index results. Higher correlations among these datasets might be expected for the SPI-1 because calculation is a standardization of monthly precipitation, which was the time set of all three input datasets. In the North-Central and Central zones, the correlation improved as the time-scales of SPI lengthened from one month (SPI-1) to six months (SPI-6), with the greatest improvement in the last months in the year. That could be explained because as seen in Subsection 4.3, the higher correlation of rainfall with CHIRPS 2.0 and PERSIANN-CDR were in winter months on the middle of the year (May to August), and this has an accumulated effect in the following months, which allow to improve the correlation on SPI as the time-scales increase until six months. This did not occur in the South and South-Central zones where the correlation does not have as high variability between the years, this may have been because in these zones the seasonality difference of rainfall between Winter and Summer is lower than in Central and North-Central zones.

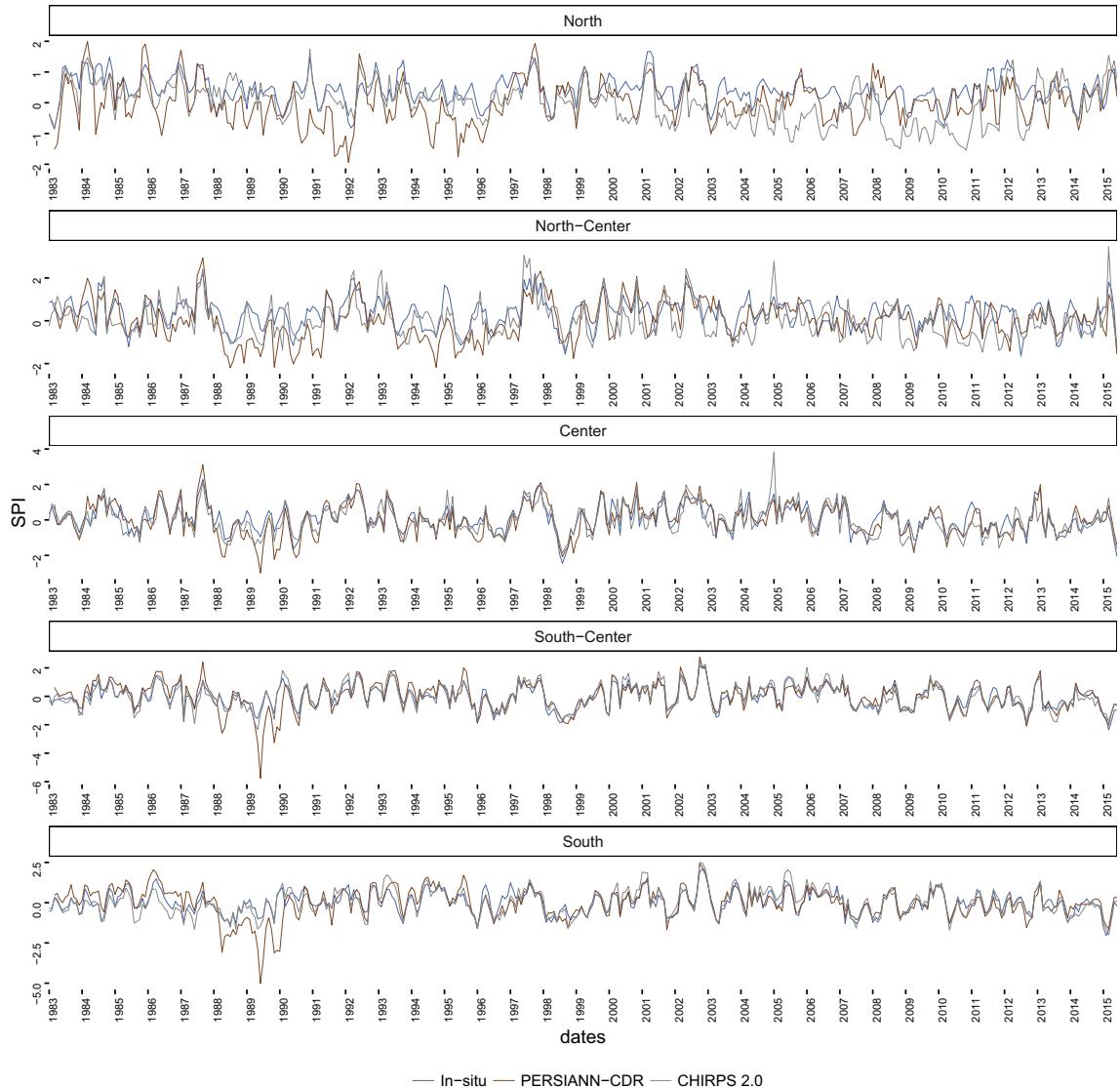


Fig. 10. Time-series of SPI-3 for in situ precipitation data and satellite products PERSIANN-CDR and CHIRPS 2.0 data with spatial aggregation for five zones of Chile: North, North-Central, Central, South-Central and South.

In Fig. 14, the mapping of SPI-3 derived from CHIRPS 2.0 is presented. CHIRPS 2.0 map results are presented and discussed in more detail here because this remote sensing dataset has a higher spatial resolution applicable for regional applications and had a better fit in the precipitation estimates with in situ observations than PERSIANN-CDR. From the results presented in Fig. 12, the 1998–1999, 2007–2008, 2008–2009 and 2014–2015 growing season were selected as the four most severe drought events because of significant rainfall deficit over a three-month period (or longer) during the last thirty

years in Chile. In Fig. 14, the large spatial extent of drought conditions is clearly identified during the 2007–2008 growing season compared to the other drought years. However, Fig. 12 shows that the CHIRPS 2.0 derived SPI products overestimated the severity of drought during the 2007–2008 and 2008–2009 growing seasons in North-Central and Central zones which is supported in Fig. 4 with the monthly precipitation difference with in situ measurements. North-Central zone in 2007–2008 season has a in situ SPI-3 of 0.2 against CHIRPS 2.0-derived SPI-3 of −0.68 and in 2008–2009 the in situ SPI-3

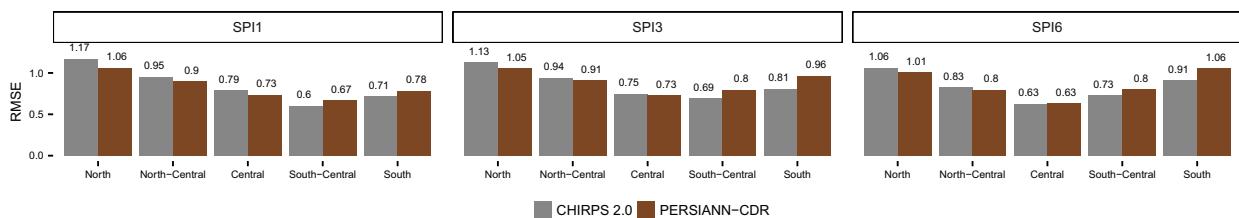


Fig. 11. Comparison of RMSE of in situ SPI with satellite derived SPI, for time-scales of 1, 3 and 6 months and zones North, North-Central, Central, South-Central and South.

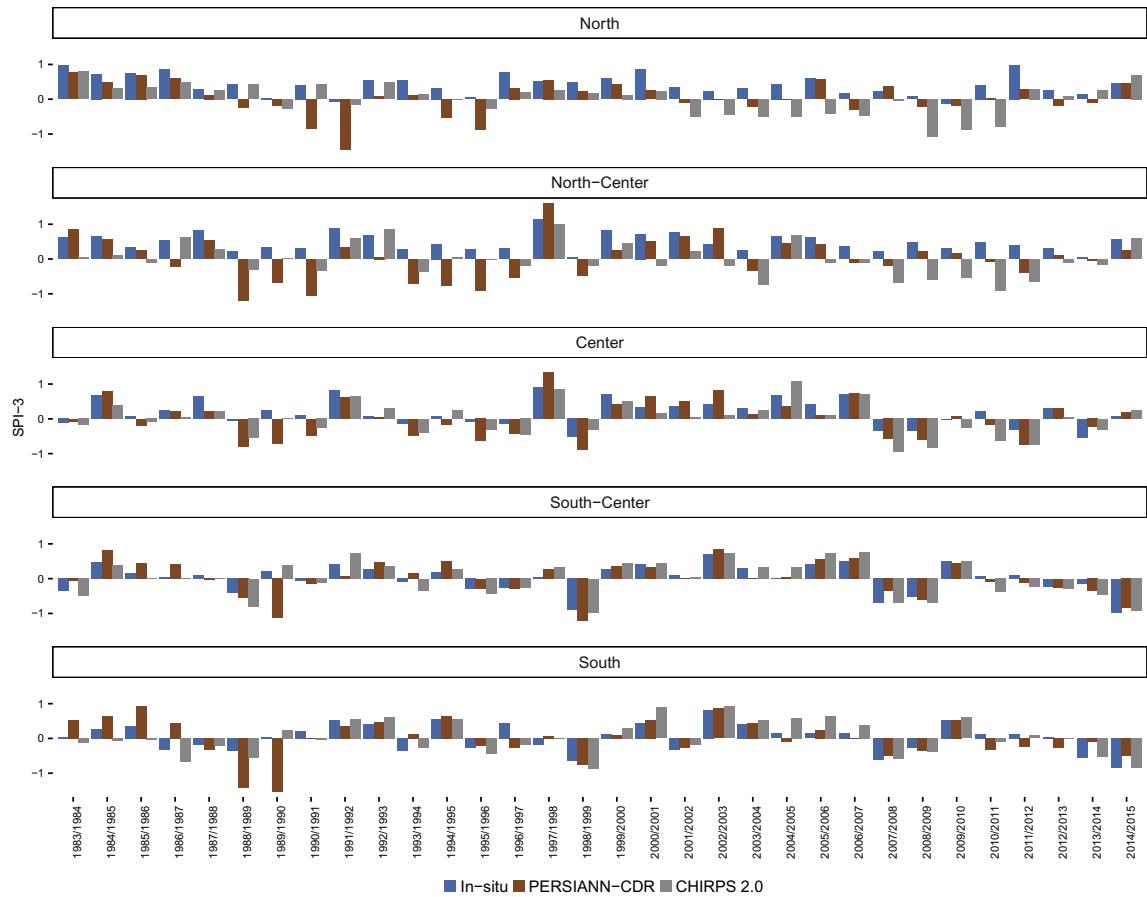


Fig. 12. Averaged SPI-3 during growing season (September to April) spatially aggregated for zones North, North-Central, Central, Central-South and South; compared for in situ, PERSIANN-CDR and CHIRPS 2.0 data for period 1983 to 2015.

has 0.49 against CHIRPS 2.0-derived SPI-3 of −0.6. Then, Central zone in 2007–2008 season has −0.33 and −0.95 for in situ and CHIRPS 2.0 derived, SPI's, respectively; and 2008–2009 season has −0.35 and −0.84 for in situ and CHIRPS 2.0-derived SPI's, respectively. The drought condition during the 2014–2015 growing season over the

South-Central and South zones of Chile, are evident with SPI-3 values corresponding to 'mild dry' to 'moderate dry' conditions. The result for North and North-Central zone are unreliable due the higher discrepancies found in the precipitation estimated from CHIRPS 2.0 as reported earlier. For example, in Fig. 12, the North zone in 2007–2008

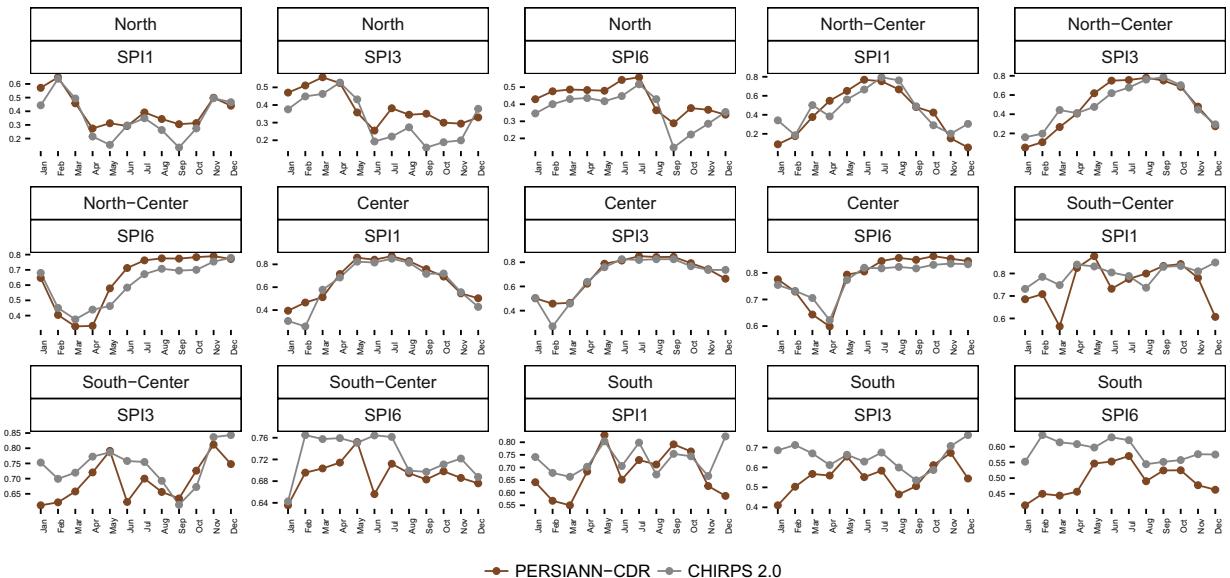


Fig. 13. Monthly correlation of SPI-1, SPI-3 and SPI-6 between satellite products PERSIANN-CDR and CHIRPS 2.0 with SPI from in situ data.

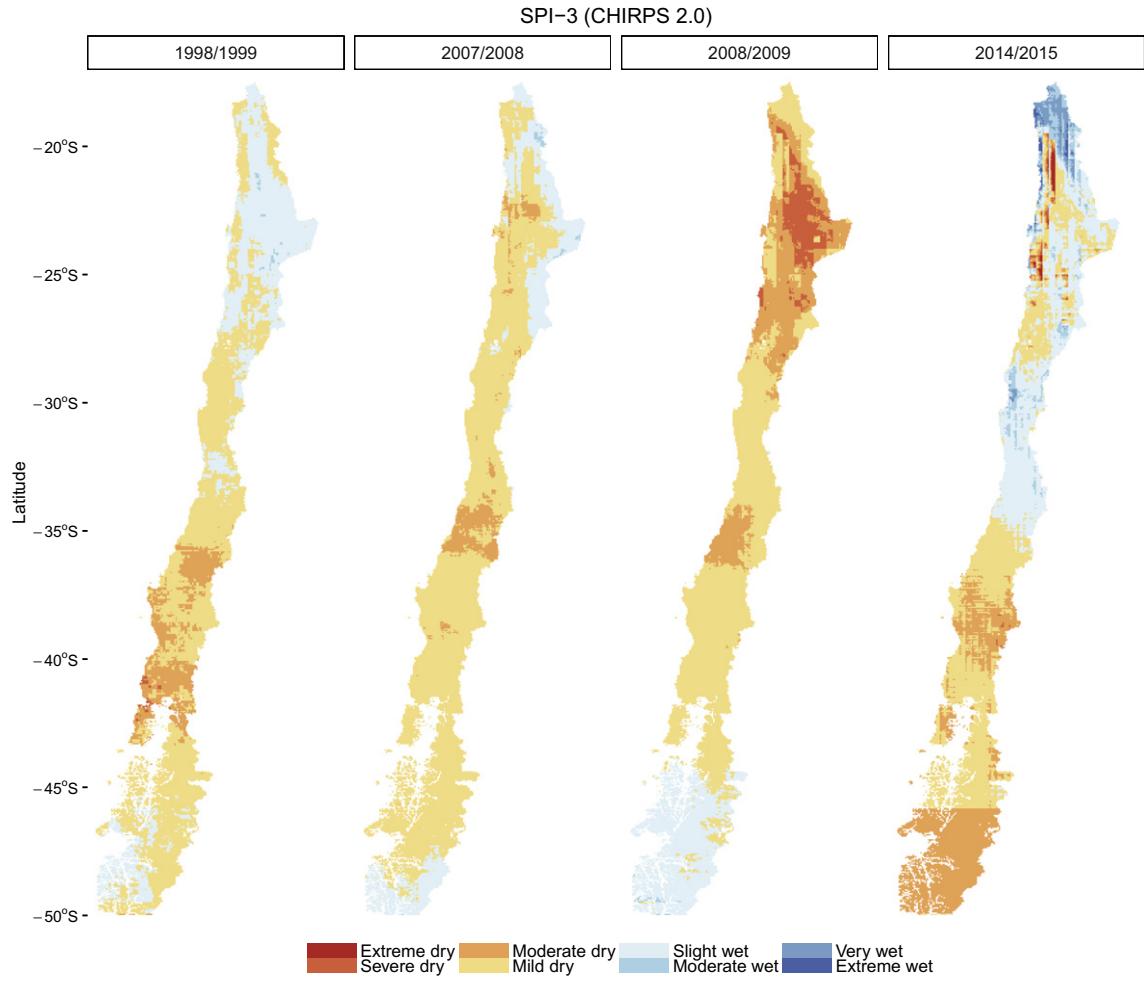


Fig. 14. Maps of SPI-3 for the growing seasons (September–April) with the most severe values from 1983 to 2015, using CHIRPS 2.0 dataset.

and 2008–2009 growing season had a high underestimation of SPI-3, similar results has the North-Central zone.

5. Conclusions

In this study, three satellite-based precipitation products with varying spatial resolutions (i.e., 0.05° and 0.25°), and extended historical data records (ranging from 18 to 30+ years) were evaluated for their accuracy of estimating the amount and spatial patterns of precipitation across Chile and their applicability for calculating the SPI to monitor agricultural drought. The precipitation and SPI estimates were compared to in situ precipitation measurements and nine statistics indices were calculated for 278 selected weather station locations across Chile. In order to assess potential regional variations of precipitation values across Chile, the country was divided into five latitudinal regional zones based on climate.

Satellite precipitation products with long, 30+ year historical records such as PERSIANN-CDR and CHIRPS 2.0 when compared to the TMPA 3B43 v7 data with a shorter 17-year historical data record and station-based precipitation, for the five latitudinal zones with nine statistics, were found to have similar results from North to South-Central zones. However, differences exist in the South zone of Chile in the case of PERSIANN-CDR which highly underestimate rainfall.

The monthly statistics analysis, showed that for the North zone, the precipitation estimates were more accurate from December to

March during the so called period of the ‘Bolivian winter’, when higher rainfall amounts occur. For the North-Central the precipitation results from satellite were more reliable specially during the Winter months (May to August) than in either the Summer (December to March) or Spring (September to December). The Central zone exhibited similar results than North zone, with the best agreement with in situ measurement from June to September in the case of PERSIANN-CDR and CHIRPS 2.0; and in May, July, August for TMPA 3B43 v7. The South-Central zone had the best results from May to September for the three products. In the South zone PERSIANN-CDR had the lowest fit and the highest was achieved by CHIRPS 2.0 and TMPA 3B43 v7, during the year. This was reflected in the results of hierarchical clustering, singular value decomposition and k-means, carry over the 9 elements of statistical comparison (i.e., CC, MAE, ME, bias, E_{ff} , FBD, POD, FAR and RMSE) for the three satellite products over the five zones and for each month of the year. Also, results showed that the clusters were made according with seasonality, where the months with higher monthly precipitation (Winter/Autumn) were grouped together and the satellite products had better fit with in situ measurement.

The spatial variation pattern for CC and E_{ff} showed that the fit of the precipitation satellite products with ground measures was strongly related to the amount of monthly rainfall for the three satellite products and CHIRPS 2.0 showing the highest agreement. Also, results showed that in the North zone where the average monthly rainfall is less than 1 mm, the results were highly inaccurate. Further, during the so called ‘Bolivian Winter’ which is when the amount

of rainfall is higher (wet season) the accuracy of satellite products improves. For the North-Central to southern zones, the accuracy of satellite products were more accurate from May to August (Autumn–Winter) when higher amounts of rainfall is received and the estimates were less accurate from September to April (Spring–Summer), which is a seasonally drier period. More detailed spatial rainfall patterns were captured by the higher spatial resolution CHIRPS 2.0 data than the PERSIANN-CDR data for most areas of Chile. Results in the higher altitudes locations of the Andes Mountains were difficult to evaluate because of the lack of in situ measurements to validate the satellite data.

The averaged time-series analysis by zone of the satellite-derived SPI-1, SPI-3 and SPI-6, showed that PERSIANN-CDR and CHIRPS 2.0 have similar results for each zone, and the lowest error was for the South-Central zone for SPI-1 and both products; South-Central for CHIRPS 2.0 and Central for PERSIANN-CDR, with SPI-3; and Central zone for SPI-6. Then, the averaged SPI-3 for zones and growing season, shows high discrepancies in the North and North-Central zones and better results in South-Central zone. Also, identified three most severe drought events between 1983–1984 and 2014–2015 growing seasons, these were 1998–1999, 2007–2008, 2008–2009 and 2014–2015.

Overall, the CHIRPS 2.0 precipitation dataset with its high spatial resolution ($0.05^\circ \times 0.05^\circ$) and long (+30 years) historical record was found to be a very useful dataset for characterizing precipitation patterns across Chile. It also provided a valuable data source to calculate a precipitation-based drought index like the SPI, which is commonly used to monitor drought. In this study, two satellite-derived datasets were tested for monitoring agricultural drought by transforming the precipitation data into the SPI over multiple time intervals (1, 3, and 6 months). The results of SPI analysis, particularly derived from the CHIRPS 2.0 data, were promising as the SPI-3 results identified drought events during the growing season that have occurred in Chile over the past thirty years. Most of the lowest SPI-3 values, which represent severe to extreme drought conditions, were in close agreement with the SPI-3 values calculated from ground-based rainfall measurements at most weather station locations across the country. The best performance of the SPI-3 calculations from CHIRPS 2.0 geographically occurred in the Central and South zones of Chile. SPI-3 results for North-Central and North zones were highly inaccurate, showing years in which the condition measure in rain gauges were wet and satellite-derived SPI-3 shows dry condition. The SPI-3 results derived from the CHIRPS 2.0 identified the key drought years of 1998–1999, 2007–2008, 2008–2009 and 2014–2015, particularly in the Central and South zones.

The long-record (more than 30 years) precipitation satellite datasets evaluated in this study, the PERSIANN-CDR and the CHIRPS 2.0, were found to be viable options for precipitation information for countries such as Chile, which have limited in situ precipitation measurements both in number and historical length. These datasets provide alternative data options for the scientific community to improve and extend the hydro-meteorological models and analysis in data poor countries and regions of the world. For drought object, future research must be done to evaluate the use of these datasets to derive drought indices spatially distributed with a higher resolution than traditional based drought maps that are produced from the spatial interpolation of in situ measurements. Like the SPI used here, these satellite-based precipitation datasets could also be used to derive the Precipitation Condition Index (Du et al., 2013) and the Standardized Precipitation Evapotranspiration Index (Vicente-Serrano et al., 2010). Collectively, the increased use of these remotely sensed precipitation datasets has the potential to improve the drought monitoring and early warning tools in Chile and other parts of the world.

However, in order to use the PERSIANN-CDR and CHIRPS 2.0 data for monitoring drought in Chile, these product should be calibrated

specifically for all months in the North zone and for the summer and spring months in the North-Central, Central, South-Central and South zones of Chile. To calibrate, and considering the higher correlations found between in situ and satellite products different techniques could be used such as, regression models, bias correction or spatial interpolations like regression-kriging Hengl (2009). In order to use long-term precipitation satellite products CHIRPS 2.0 and PERSIANN-CDR in other countries and regions of the world similar studies need to be done to firstly assess its accuracy.

6. Acknowledgment

We would like to express our gratitude to the 'Center for Advanced Land Management Information Technologies (CALMIT)', 'National Drought Mitigation Center (NDMC)', and 'The Robert B. Daugherty Water for Food Institute' from the University of Nebraska for the collaboration and support during the six months time in which this manuscript was written. Thanks for the support and additional funding provided for project UCO-1407 from the University of Concepcion. Finally, thanks to the 'Center for Climate and Resilience Research (CR²) (CONICYT/FONDAP/15110009)' from the University of Chile for leaving the precipitation data available which were used in this study. This study was funded by the CONICYT Scholarship/National Ph.D. 21141028.

Appendix A. Data processing and data analysis

Appendix B. Statistics

To compare between in situ data from measured in situ using the rain gauge (G) and the estimates from the satellite products (S), nine statistics were used, following Eq. (B.1) for magnitude of underestimation (ME), Eq. (B.2) for mean absolute error (MAE), Eq. (B.3) for multiplicative bias (bias), Eq. (B.4) for efficiency (E_{ff}), and Eq. (B.5) for the Root Mean Square Error (RMSE). These statistics evaluate the performance of the satellite products in estimating the amount of the rainfall (Dinku et al., 2009).

$$ME = \frac{1}{N} \sum (S - G) \quad (B.1)$$

$$MAE = \frac{\frac{1}{N} \sum |S - G|}{\bar{G}} \quad (B.2)$$

$$bias = \frac{\sum S}{\sum G} \quad (B.3)$$

$$E_{ff} = 1 - \frac{\sum (S - G)^2}{\sum (G - \bar{G})^2} \quad (B.4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum (S - G)^2} \quad (B.5)$$

To evaluate the rainfall detection capabilities of the satellite products, several statistics were calculated including the frequency

Table A.1

Packages from R environment used for the processing of the remote sensing data and the data analysis.

Procedure	Package
Remote sensing processing	<i>raster</i> (Hijmans, 2015)
Data analysis	<i>data.table</i> , <i>dplyr</i> (Dowle et al., 2015; Wickham and Francois, 2015)
Data visualization	<i>ggplot2</i> (Wickham, 2007)
Data transformation	<i>reshape2</i> (Wickham, 2007)

Table B.2

Contingency table for comparing rain gauge measurements and satellite rainfall estimates. The threshold correspond to the value above which rainfall is considered detected. In this case a value of 1mm was used.

	Gauge \geq threshold	Gauge $<$ threshold
Satellite \geq threshold	A	B
Satellite $<$ threshold	C	D

bias (*FBS*; Eq. (B.6)), probability of detection (*POD*; Eq. (B.7)) and false-alarm ratio (*FAR*; Eq. (B.8)). **Table B.2**

$$FBS = \frac{A + B}{A + C} \quad (\text{B.6})$$

$$POD = \frac{A}{A + C} \quad (\text{B.7})$$

$$FAR = \frac{B}{A + B} \quad (\text{B.8})$$

shows a contingency table, where A, B, C and D represent hits, false alarms, misses, and correct negatives, respectively (Dinku et al., 2009).

References

- Almazroui, M., 2011. Calibration of TRMM rainfall climatology over Saudi Arabia during 1998–2009. *Atmos. Res.* 99, 400–414.
- Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D.K., Knapp, K.R., Cecil, L.D., Nelson, B.R., Prat, O.P., 2015. PERSIANN-CDR: daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bull. Amer. Meteor. Soc.* 96, 69–83.
- Beguería, S., Vicente-Serrano, S.M., 2013. SPEI: calculation of the standardised precipitation-evapotranspiration index. R Package Version 1.6 Edition
- Bharti, V., Singh, C., 2015. Evaluation of error in TRMM 3B42V7 precipitation estimates over Himalayan region. *J. Geophys. Res. Atmos.* 120 (12), 458–473.
- Dinku, T., Cecato, P., Grover-Kopec, E., Lemma, M., Connor, S., Ropelewski, C., 2007. Validation of satellite rainfall products over East Africa's complex topography. *Int. J. Remote Sens.* 28 (7), 1503–1526.
- Dinku, T., Ruiz, F., Connor, S.J., Ceccato, P., 2009. Validation and intercomparison of satellite rainfall estimates over Colombia. *J. Appl. Meteorol. Climatol.* 49, 114–1004.
- Dowle, M., Srinivasan, A., Short, T., 2015. with contributions from R. In: Saporta, S.L., Antonyan, E. (Eds.), *Data.Table: Extension of Data.Frame*. R package version 1.9.6.
- Du, L., Tian, Q., Yu, T., Meng, Q., Jancso, T., Udvardy, P., Huang, Y., 2013. A comprehensive drought monitoring method integrating MODIS and TRMM data. *Int. J. Appl. Earth Obs. Geoinf.* 23, 245–253.
- Duan, Z., Bastiaanssen, W., 2013. First results from version 7 TRMM 3B43 precipitation product in combination with a new downscaling-calibration procedure. *Remote. Sens. Environ.* 131, 1–13.
- Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., Huang, X., 2010. MODIS collection 5 global land cover: algorithm refinements and characterization of new datasets. *Remote. Sens. Environ.* 114 (1), 168–182.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Rowland, J., Romero, B., Husak, G., Michaelsen, J., Verdin, A., 2014. A quasi-global precipitation time series for drought monitoring. Technical Report, U.S Geological Survey Data Series 832, Garreaud, R., 2009. The Andes climate and weather. *Adv. Geosci.* 7 (1–9).
- Guo, H., Bao, A., Liu, T., Chen, S., Ndayisaba, F., 2016a. Evaluation of PERSIANN-CDR for meteorological drought monitoring over China. *Remote Sens.* 8 (379), 17.
- Guo, H., Chen, S., Bao, A., Behrang, A., Hong, Y., Ndayisaba, F., Hu, J., Stepanian, P., 2016b. Early assessment of integrated multi-satellite retrievals for global precipitation measurement over China. *Atmos. Res.* 176–177, 121–133.
- Hayes, M., Svoboda, M., Wall, N., Widhalm, M., 2011. The lincoln declaration on drought indices: universal meteorological drought index recommended. *Bull. Am. Metor. Soc.* 92 (4), 485–488.
- Hengl, T., 2009. *A Practical Guide to Geostatistical Mapping*. 2nd edition, University of Amsterdam, Amsterdam.
- Hijmans, R.J., 2015. *Raster: geographic data analysis and modeling*. R Package Version 2.5–2 Edition
- Hsu, K.L., Gao, X., Sorooshian, S., Gupta, H.V., 1997. Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteor.* 36, 1176–1190.
- Huffman, G.J., Adler, R.F., Bolvin, D.T., Gu, G., Nelkin, E.J., Bowman, K.P., Hong, Y., Stocker, E.F., Wolff, D.B., 2007. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.* 8 (1), 38–55.
- Huffman, G.J., Adler, R.F., Bolvin, D.T., Nelkin, E.J., 2010. The TRMM Multi-satellite Precipitation Analysis (TMPA). *Satellite Applications for Surface Hydrolog.*
- IPCC, 2013. *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge UK; New York, USA.
- Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorol.* 5, 487–503.
- Katsanos, D., Retailis, A., Michaelides, S., 2016. Validation of a high-resolution precipitation database (chirps) over Cyprus for a 30-year period. *Atmos. Res.* 169, 459–464.
- Kenawy, A.M.E., Lopez-Moreno, J.I., McCabe, M.F., Vicente-Serrano, S.M., 2015. Evaluation of the TMPA-3b42 precipitation product using a high-density rain gauge network over complex terrain in northeastern Iberia. *Glob. Planet. Chang.* 133, 188–200.
- Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Koppen-Geiger climate classification updated. *Meteor. Z.* 15 (3), 259–263.
- Loon, A.F.V., Gleeson, T., Clark, J., Dijk, A.J.M.V., Stahl, K., Hannaford, J., Baldassarre, G.D., Teuling, A.J., Tallaksen, L.M., Uijlenhoet, R., Hannah, D.M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N., Lanen, H.A.J.V., 2016. Drought in the Anthropocene. *Nat. Geosci.* 9 (2), 89–91.
- McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. In: Proceedings of the International 8th Conference on Applied Climatology. American Meteorological Society, Anaheim, CA, USA, 17–22 January, pp. 179–184.
- Meza, F.J., 2013. Recent trends and ENSO influence on droughts in northern Chile: an application of the standarized precipitation evapotranspiration index. *Weather Clim. Extremes* 1, 51–58.
- Miao, C., Ashouri, H., Hsu, K.-L., Sorooshian, S., Duan, Q., 2015. Evaluation of the PERSIANN-CDR daily rainfall estimates in capturing the behavior of extreme precipitation events over China. *Am. Meteorol. Soc.* 16, 1387–1396.
- Moazami, S., Golian, S., Hong, Y., Sheng, C., Kavianpour, M.R., 2016. Comprehensive evaluation of four high-resolution satellite precipitation products under diverse climate conditions in Iran. *Hydrol. Sci. J.* 61 (2), 420–440.
- Nastos, P., Kapsomenakis, J., Philandras, K., 2016. Evaluation of the TRMM 3b43 gridded precipitation estimates over Greece. *Atmos. Res.* 169, 497–514.
- National Research Council, 2004. *Climate Data Records From Environmental Satellites: Interim Report*. The National Academies Press, Washington DC.
- Núñez, J., Verbist, K., Wallis, J.R., Schaefer, M.G., Morales, L., Cornelis, W.M., 2011. Regional frequency analysis for mapping drought events in north-central Chile. *J. Hydrol.* 405, 352–366.
- Pipunic, R.C., Ryu, D., Costelloe, J.F., Su, C.-H., 2015. An evaluation and regional error modelling methodology for near-real-time satellite rainfall data over Australia. *J. Geophys. Res. Atmos.* 120 (10), 767–783.
- Core Team, R., 2016. *R: A Language and Environment for Statistical Computing* R Foundation for Statistical Computing, Vienna, Austria.
- Romero, H., Smith, P., Mendonca, M., Méndez, M., 2013. Macro y mesoclimas del altiplano andino y desierto de Atacama: desafíos y estrategias de adaptación social ante su vulnerabilidad. *Rev. Geogr. Norte Gd.* 55, 19–41.
- Salio, P., Hobouchiand, M.P., Skabar, Y.G., Vila, D., 2015. Evaluation of high-resolution satellite precipitation estimates over southern South America using a dense rain gauge network. *Atmos. Res.* 163, 146–161.
- Shah, H.L., Mishra, V., 2015. Uncertainty and bias in satellite-based precipitation estimates over Indian subcontinental basins: implications for real-time streamflow simulation and flood prediction. *J. Hydrometeorol.* 17, 615–636.
- Tan, M.L., Ibrahim, A.L., Duan, Z., Cracknell, A.P., Chaplot, V., 2015. Evaluation of six high-resolution satellite and ground-based precipitation products over Malaysia. *Remote Sens.* 7 (2), 1504–1528.
- Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *J. Clim.* 23 (7), 1696–1718.
- Wickham, H., 2007. Reshaping data with the reshape package. *J. Stat. Softw.* 21 (12), 1–20.
- Wickham, H., Francois, R., 2015. *dplyr: a grammar of data manipulation*. R Package Version 0.4.2
- Wilhite, D.A., Glantz, M.H., 1985. Understanding: the drought phenomenon: the role of definitions. *Water Int.* 10 (3), 111–120.
- WMO, 1986. Report on drought and countries affected by drought during 1974–1985. WMO, Geneva, , pp. 118.
- Zambrano, F., Lillo-Saavedra, M., Verbist, K., Lagos, O., 2016. Sixteen years of agricultural drought assessment of the BíoBío Region in Chile using a 250 m resolution vegetation condition index (VCI). *Remote Sens.* 8 (6), 530.