Temporal and spatial evaluation of long-term satellite-based precipitation products across the complex topographical and climatic gradients of Chile

Mauricio Zambrano-Bigiarini[a,b]

[a]Department of Civil Engineering, Universidad de La Frontera, Temuco, Chile
[b]Center for Climate and Resilience Research, Universidad de Chile, Santiago, Chile

ABSTRACT

Satellite-based rainfall estimates (SRE) have become a promising data source to overcome some limitations of ground-based rainfall measurements, in particular for hydrological and other environmental applications. This study evaluates the spatial and temporal performance of four long-term SRE products (TMPA 3B42v7, CHIRPSv2, MSWEPv1.1 and MSWEPv2.2) over the complex topography and climatic gradients of Chile. Time series of precipitation measured at 371 stations are compared against the corresponding grid cell of each SRE (in their original spatial resolution) at different temporal scales (daily, monthly, seasonal, annual). The modified Kling-Gupta efficiency along with its three individual components were used to assess the performance of each SRE, while two categorical indices (POD, and fBIAS) were used to evaluate the skill of each SRE to correctly capture different precipitation intensities.

Results revealed that all SREs performed best in Central-Southern Chile (32.18-36.4°S), in particular at low- and mid-elevation zones (0-1000 m a.s.l.). Seasonally, all products performed best in terms of KGE' during the wet autumn and winter seasons (MAM-JJA) compared to summer (DJF). In addition, all SREs were able to correctly identify no rain events, but during rainy days all SREs that did not use a local dataset of precipitation to recalibrate their estimates presented a low skill in providing an accurate classification of different precipitation intensities.

Overall, MSWPEPv22 showed the best performance at all time scales and country-wide, due to the use of a Chilean dataset of daily data for calibrating its precipitation estimates, making it a good candidate for hydrological applications in Chile. Finally, we conclude that when the in situ precipitation dataset used in the evaluation of different SREs does not cover the headwaters of the catchments, the obtained performances should only be considered as first guess about how well a given SRE represent the real amount of water in an area.

Keywords: precipitation, rainfall, Chile, 3B42, CHIRPS, MSWEP, POD, fBIAS

1. INTRODUCTION

In the last decades, satellite-based estimates of precipitation have become a promising data source to overcome some limitations of ground-based rainfall measurements.[1–5] Moreover, the temporal and spatial resolution of these precipitation products have continuously been improved due to developments in sensor technology and algorithms for merging different data sources.[6] However, several researchers have found that SREs still have important biases, false detection of precipitation, and fail in capturing some precipitation intensities.[7–10]

When SREs are used as input for hydrological applications, the previously mentioned errors are non-linearly propagated into simulated streamflows,[2,11,12] and they might lead to wrong conclusions and poor management decisions. Furthermore, several time scales of precipitation data are required for different hydrological applications (e.g., flood forecasting, water resources management), ranging from (sub)hourly to seasonal.[1] Finally, to allow SREs to achieve their full potential, we require a better understanding on how different climatological regimes and other environmental factors impact on the quality of precipitation fields produced by different SREs.[1,8,12]

Further author information: (Send correspondence to Av. Francisco Salazar 01145, Temuco, Chile)
E-mail: mauricio.zambrano@ufrontera.cl, Telephone: 56 45 259 2812

Recently, Ref. 13 carried out an exhaustive comparison of seven state-of-the-art SRE products in the data-scarce and complex mountainous region of Chile, finding that there was no single SRE product that performed best always and everywhere, and that a site-specific assessment is recommended before using any SRE in hydrological studies. The present study attempts to exhaustively compare the best long-term SREs found by Ref. 13 against the latest version of the Multi-Source Weighted-Ensemble Precipitation (MSWEP v2.2), which has been continuously updated since 2017, and -in comparison with other SREs- it should not underestimate observed precipitation amounts in mountain regions.[14]

2. STUDY AREA

Continental Chile has more than 4000 km of latitudinal extension (17.50°-66.42°S), bounded by the Pacific Ocean to the west and by the Andes mountain range in the east. Its climate is mainly controlled by the topography, latitude, and the oceanic influence of the Pacific Ocean.[15] An hyperarid climate characterise the northern part of Chile, while important precipitation amounts are observed in the humid south, reaching values of above 5000 mm/year.[16] In most of the central-southern territory the wet season occurs mainly during winter (JJA), while in the northernmost regions it is observed during summer (DJF), and no dry season is clearly defined in the southern part of Chile.[15] Five macroclimatic areas are identified to evaluate the performance of the selected SREs: (i) Far North (17.50-26.00°S), (ii) Near North (26.00-32.18°S), (iii) Central Chile (32.18-36.40°S), (iv) South (36.40-43.70°S) and (v) Austral/Far South (43.70-56.00°S), the same as Ref. 13.

3. DATASETS

3.1 Rain gauges

In situ precipitation data were downloaded from a dataset of 874 rain gauges provided by the Center of Climate and Resilience Research (CR2), with daily data from 01-Jan-1900 to 31-Dec-2017 (available on http://www.cr2.cl/datos-de-precipitacion/). Only rain gauges with less than 5% of missing data during the period Jan/1998-Dec/2014 were selected for this study, resulting in 371 stations within our study area (see Section 4).

3.2 Satellite-based data

In contrast to the exhaustive comparison carried out by Ref. 13, in this study we focused only on long-term and high-quality SREs, as in the future we are interested in carrying out hydrological simulations in selected basins. A short description of each one of the selected SREs is given in the next paragraphs.

3.2.1 3B42v7

The Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) combines IR data from geosynchronous earth orbit (GEO) satellites with four passive microwave (PMW) sensors, to provide rainfall estimates at 0.25° of spatial resolution and 3-hourly time steps, in both real and post-real time.[17,18] For this study, only the 3B42_daily product was used, which was then accumulated to monthly, seasonal or annual values depending on the analysis.

3.2.2 CHIRPSv2

The Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS version 2) is a quasi-global SRE explicitly developed for monitoring global environmental change over land and agricultural drought.[19] It combines information from polar-orbiting and geo-synchronous satellites with more than 2000 station records to calibrate global Cold Cloud Duration (CCD) rainfall estimates.[19] The product is available from 1981 to present, at daily, pentadal and monthly time scales, with a spatial resolution of 0.05°, from 50°S to 50°N (although since November 2012 data are only available up to 46°S).

Table 1. Summary of satellite-based products providing (sub)daily data at a quasi-global scale

Dataset	Full Name	Latitudinal Coverage	Spatial Resolution	Temporal Coverage	Temporal Resolutions	References
3B42v7	**T**RMM **M**ulti-satellite **P**recipitation **A**nalysis research product 3B42 Version 7	50°N-50°S	0.25°	Jan-1998 - present	3-hourly, daily	Ref. 17, 18
CHIRPSv2	**C**limate **H**azards group **I**nf**r**ared **P**recipitation with **S**tations Version 2.0	50°N-50°S	0.05°	Jan-1981 - present	daily, pentadal, monthly	Ref. 19
MSWEPv1.1	**M**ulti-**S**ource **W**eighted-**E**nsemble **P**recipitation Version 1.1	90°N-90°S	0.25°	Jan-1979 Dec-2014	daily	Ref. 20
MSWEPv2.2	**M**ulti-**S**ource **W**eighted-**E**nsemble **P**recipitation Version 2.2	90°N-90°S	0.10°	Jan-1979 Dec-2017	3-hourly, daily	Ref. 14

3.2.3 MSWEPv1.1

The Multi-Source Weighted-Ensemble Precipitation (MSWEP[20]) dataset is a new global precipitation dataset specifically developed to be used in hydrological modelling studies, aiming at improving the representation of precipitation in mountainous, tropical and snowmelt-driven regions. Data are provided for the 1979-2014 period, with 3-hourly and 0.25° resolution. It is based on data from rain gauges, satellite observations, and reanalysis. In this study we used the MSWEPv1.1, released on August 2nd 2016.

3.2.4 MSWEPv2.2

In comparison to v1.1., the most important features of the latest version[14] of MSWEP are: (i) correction of the cumulative distribution function and precipitation frequency, to account for spurious drizzle and attenuated peaks; (ii) higher spatial resolution, from 0.25° to 0.1°; (iii) the inclusion of ocean areas; (iv) the addition of precipitation estimates derived from Gridded Satellite (GridSat) thermal infrared imagery for the pre-TRMM era; (v) the use of a daily (rather than monthly) gauge correction scheme, which considers regional differences in reporting times; (vi) the use of a large database of daily gauge observations compiled from several sources to replace the 0.5° CPC Unified dataset; and (vii) extension of the data record to 2017. In this study we used the MSWEPv2.2,[14] released on March 2nd, 2018.

Table 1 provides a summary of the satellite-based datasets, including their full names, spatial and temporal resolutions for the versions used in this study.

4. METHODOLOGY

4.1 Point-to-pixel comparison

A point-to-pixel analysis[9] was used to compare time series of observed precipitation at selected rain gauges to the corresponding SRE pixel. All SREs were used in their original spatial resolution, because upscaling them to a unified grid might affect the comparison.[7] Daily observations at the 371 rain gauges (see Section 3.1) and their corresponding satellite estimates were accumulated into monthly, seasonal (DJF, MAM, JJA, SON) and annual values, to assess the accuracy of each SRE at different time scales. Because a given amount of rainfall falling down with two different durations will lead to different hydrological processes, daily precipitation events were classified and analysed using the classes shown in Table 2. Based on the availability of satellite data (see Table 1), the evaluation period for this study extends from Jan 1998 to Dec 2014.

Table 2. Classification of rainfall events based on its daily intensity i. Modified for daily values from Ref. 21.

Rainfall event	Intensity (i), [mm d^{-1}]
No rain	[0 , 1)
Light rain	[1 , 5)
Moderate rain	[5 , 20)
Heavy rain	[20 , 40)
Violent rain	$\geq$ 40

4.2 Performance indices

The four SREs described in Section 3.2 were exhaustively evaluated, using both continuous and categorical indices of performance at different temporal scales. Continuous indices are the modified Kling-Gupta efficiency[22,23] along with its three individual components, which is used here to decompose the total performance of the SREs into a linear correlation (r, Equation 2), bias (β, Equation 3) and a variability (γ, Equation 4) term. The optimum value of *KGE'*, r, β and γ is 1.0.

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{1}$$

$$r = \frac{\sum_{i=1}^{n}(O_i - \bar{O})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2}\sqrt{\sum_{i=1}^{n}(S_i - \bar{S})^2}} \tag{2}$$

$$\beta = \frac{\mu_s}{\mu_o} \tag{3}$$

$$\gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o} \tag{4}$$

In addition to separate rainfall events into *no rain* and *rain*,[24,25] in this study we further classify daily rainfall in five types of intensities, ranging from *no rain* (*dry day*) ($<$ 1 mm d^{-1}) to *violent rain* ($>$ 40 mm d^{-1}), as shown in Table 2. Following Ref. 13, only two categorical indices were used to assess the ability of each SRE to correctly identify different precipitation intensities. The *probability of detection* (*POD*, Equation 5 measure the fraction of events that are correctly captured by the satellite product, ranging from 0 to 1. The *frequency bias* (*fBias*, Equation 6) compares the number of events identified by the satellite product to the number of events that actually occurred at the corresponding rain gauge. Its optimal value is 1.0 (unbiased), with $fBias > 1$ and $fBias < 1$ indicating an overestimation and underestimation of the occurrences, respectively.

All the aforementioned indices of SRE performance were computed based on,[26] using the R environment 3.5.1[27] and the `raster`,[28] `hydroGOF`[29] and `hydroTSM`[30] R packages.

$$POD = \frac{H}{H+M} \tag{5}$$

$$fBias = \frac{H+F}{H+M} \tag{6}$$

5. RESULTS

5.1 Spatial variability of SRE performance

Spatial maps showing the performance of each SREs in each one of the 371 selected stations and for different time scales (daily, monthly, annual, four seasons) were produced, including the performance at different elevation zones and macro-climatic areas. As an example, Figure 1 shows the modified Kling-Gupta efficiency (KGE') with MSWEPv22 estimates compared against observed precipitation at a daily time scale. It shows a general good performance ($KGE' > 0.75$) in low and mid-elevation areas (0-1000 m a.s.l.). However, it shows a poor performance ($KGE' < 0$) for some stations located in the Far North. Analysing in detail the individual components of KGE' (figures not shown here) revealed that the poor performance is due to a large overestimation, with values as high as 24 times the observed precipitation (note that average annual values for those stations is, in general, lower than 10 mm).

Figure 2 shows boxplots with the modified Kling-Gupta efficiency between different daily SREs and their corresponding observations for different macro-climatic regions. It illustrates that the best performance for all SREs was obtained in the South-Central (36.4-32.18°S), with MSWEPv22 as the best product followed closely by MSWEPv1.1, CHIRPSv2, and 3B42v7. On the other hand, all SREs presented an acceptable performance in the Near North (26.0-32.18°S) and a poor performance in the Far North (17.5-26.0°S). Figure 3 shows that the best performance for all SREs was obtained in low- and mid-elevation zones (0-1000 m a.s.l.), only with MSWEPv22 showing good performance for higher elevations.

5.2 Temporal variability of SRE performance

Figure 4 shows boxplots of KGE' for all the satellite products and time scales. Most of the SREs, except MSWEPv22, presented a limited performance at daily time scales (median values of $KGE' \leq 0.5$), which improved when aggregated into monthly and annual values. MSWEPv22 was the product with the best and worst performance for all scales, followed by CHIRPSv2 and 3B42v7. At the seasonal scale all the products performed best during the humid autumn (MAM) and winter (JJA), and showed the worst performance and larger variability during the dry summer (DJF).

5.3 SRE performance for different precipitation intensities

Figure 5 shows the median values of the two categorical indices of performance computed between different SREs and the observations at the corresponding grid cell, for the five classes of daily rainfall intensity defined in Table 2. Most of the products showed a low probability of detection ($POD \leq 0.4$) for all rainfall events except by *no rain* ([0, 1) mm d^{-1}), which was well captured by all SREs ($POD \geq 0.8$). MSWEPv22 was the product that best captured all classes of rainfall events. On the other hand, the frequency bias ($fBIAS$) panel showed that the amount of *no rain* events identified by most of SREs was in excellent agreement with the corresponding observed amount (median $fBIAS \sim 1$). All SREs overestimated the amount of light and moderate rain events ([1, 20) mm d^{-1}) and underestimated the number of heavy and violent events (≥ 20 mm d^{-1}).

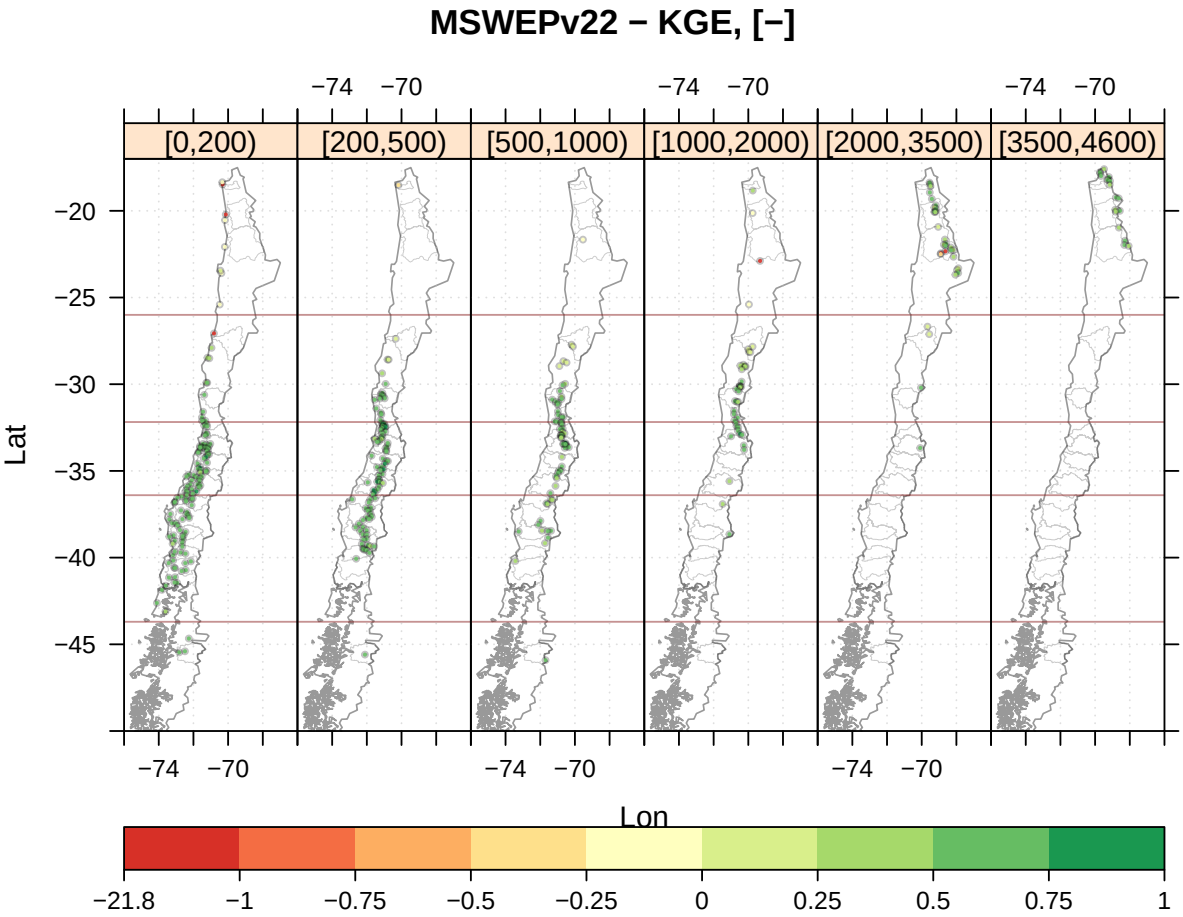

Figure 1. Modified Kling-Gupta efficiency (*KGE'*) between daily satellite estimations (MSWEPv22) and observations at the corresponding grid cell, for six different elevation zones: [0,200), [200,500), [500,1000), [1000,2000), [2000,3500), [3500,4560) m a.s.l. Colours for KGE' range from intense red to dark green, representing very poor and optimum performance, respectively.

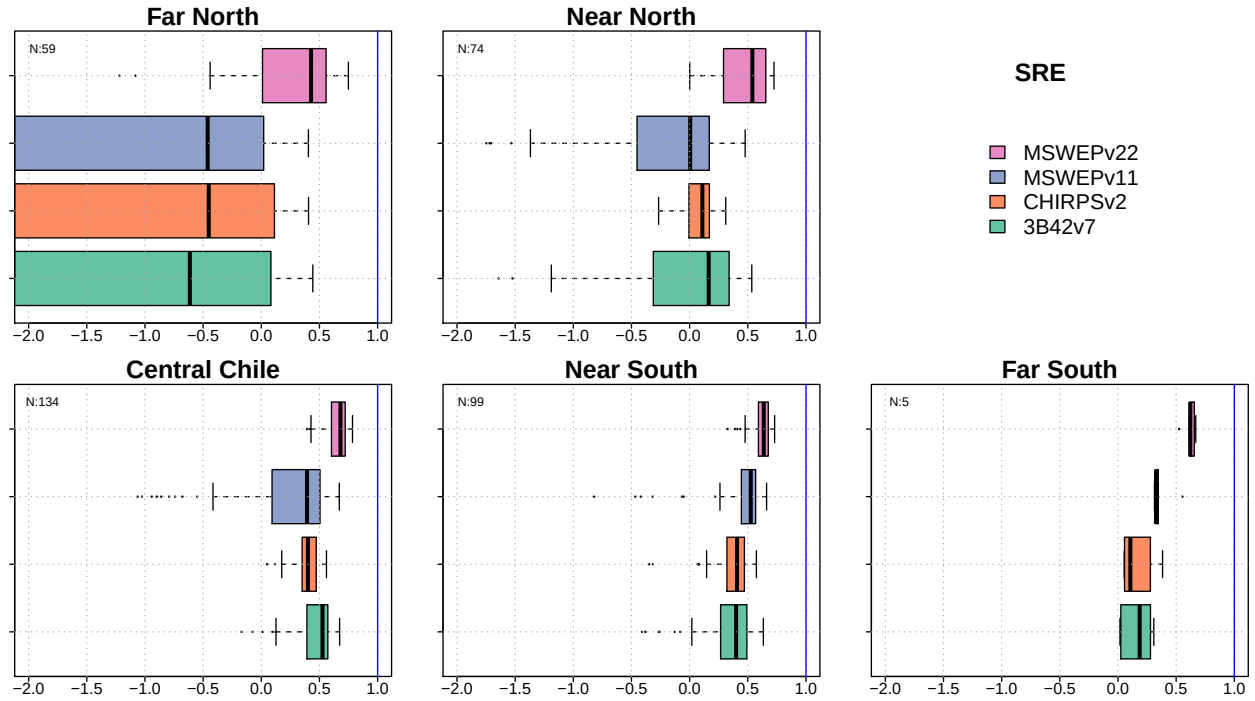

Figure 2. Modified Kling-Gupta efficiency (KGE') between different daily satellite estimations and observations at the corresponding grid cell, for five different macroclimate zones: Far North, Near North, Centre, South and Far South. The vertical blue line indicates the optimum value for KGE', while N indicates the number of stations in each macroclimate zone.

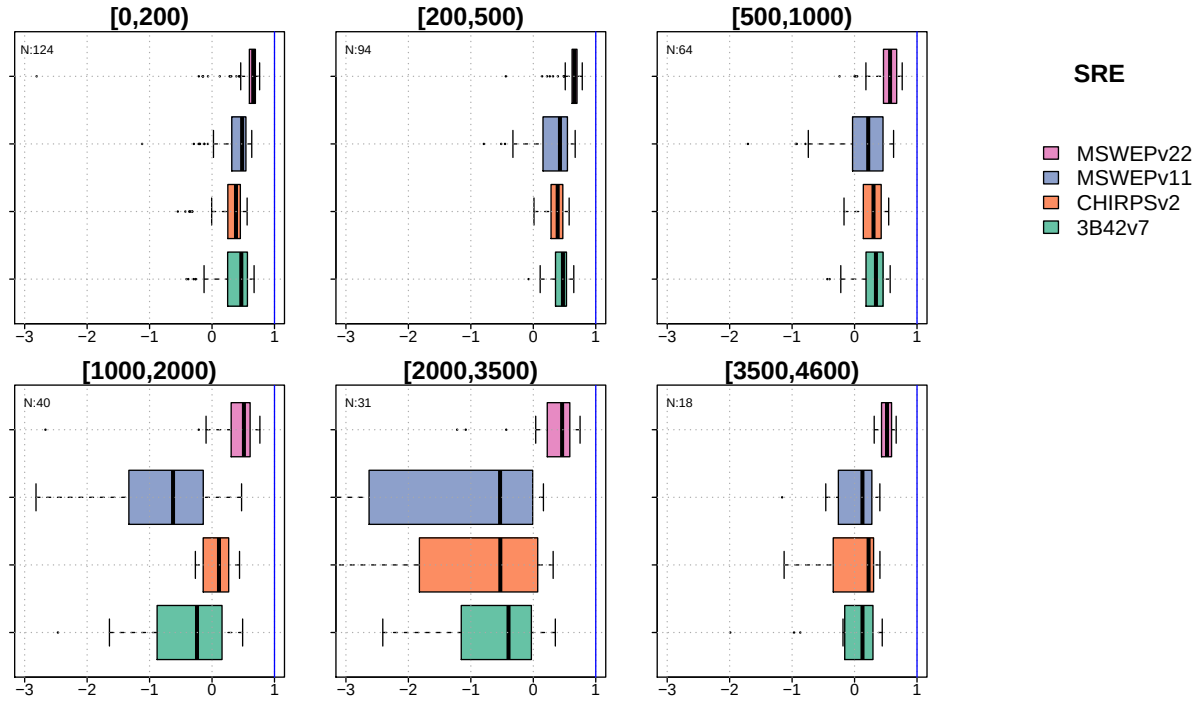

Figure 3. Modified Kling-Gupta efficiency (KGE') between different daily satellite estimations and observations at the corresponding grid cell, for six different elevation zones: [0,200), [200,500), [500,1000), [1000,2000), [2000,3500), [3500,4560) m a.s.l. The vertical blue line indicates the optimum value for KGE', while N indicates the number of stations in each elevation zone.

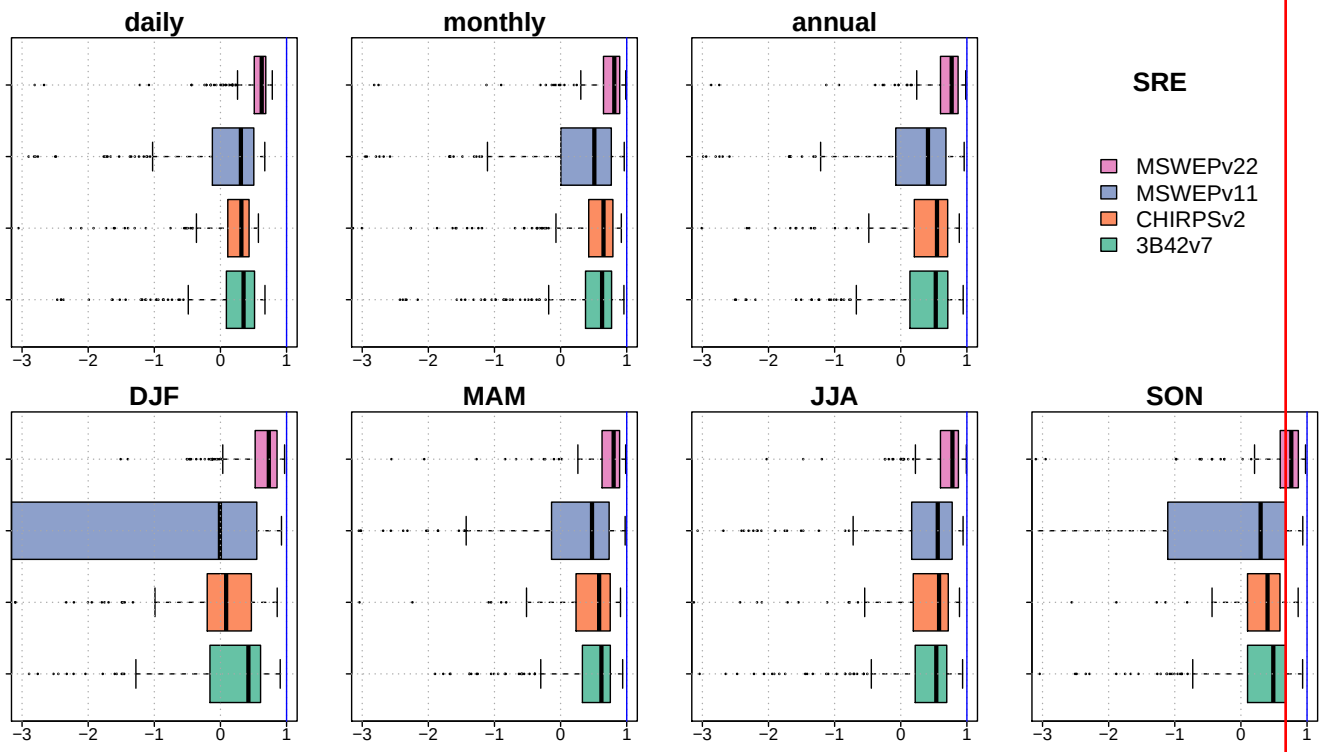

Figure 4. Modified Kling-Gupta efficiency (KGE') between different satellite estimations and the observations at the corresponding grid cell, for six different temporal scales. From left to right and up to bottom: daily, monthly, annual, summer (DJF), autumn (MAM), winter (JJA), and spring (SON). The vertical blue line indicates the optimum value for KGE'.

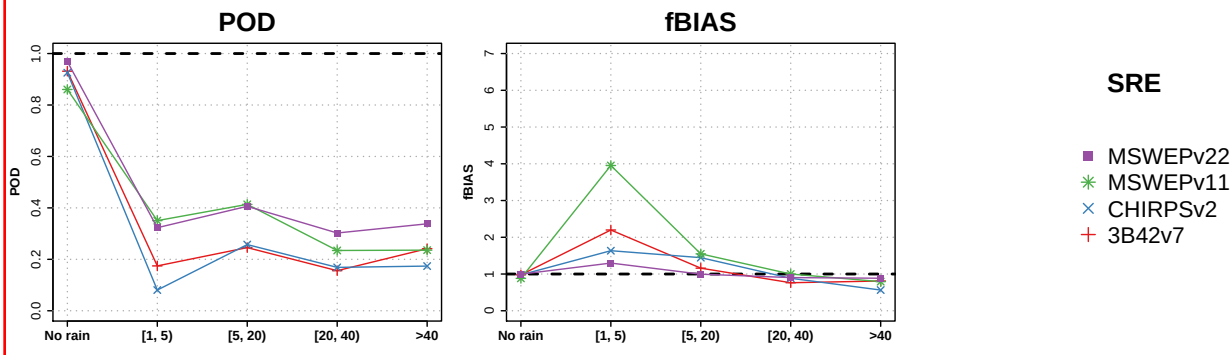

Figure 5. Median values of categorical indices of performance for different SREs and the five classes of rainfall intensity defined in Table 2. From left to right: probability of detection (POD) and frequency bias ($fBias$). The horizontal dashed black line indicates the optimum value for each index.

6. CONCLUSIONS

This work aimed at providing insights about the performance of four long-term and state-of-the-art SREs (3B42v7, CHIRPSv2, MSWEPv1.1 and MSWEPv2.2) at different temporal scales (daily, monthly, seasonal, annual) over the Chilean territory, using 371 stations located from sea level to 4600 m a.s.l. in the Andean altiplano. As expected, the latest version of MSWEP showed the best performance at all time scales and country-wide (Figure 4), due to the use of a Chilean dataset of daily data for calibrating its precipitation estimates. Moreover, considering its high spatial and temporal resolution and continuous updates, MSWEPv22 is a good candidate for being used in hydrological applications in Chile. The following paragraphs summarise the main findings of this work:

- Most SRE products performed the best in Central-Southern Chile (32.18-36.4°S), in particular at low- and mid-elevation zones (0-1000 m a.s.l.), compared to the arid northern regions and the Far South.
- All SREs but MSWEPv22 performed worst in the high elevation areas ($\geq$ 2000 m a.s.l.) of the hyper-arid northern region.
- All SREs performed best in terms of KGE' during the wet autumn and winter seasons (MAM-JJA) compared to summer (DJF).
- All SREs were able to correctly identify the occurrence of no rain events (i.e., dry days). However, during rainy days all SREs that did not use a local dataset of precipitation to recalibrate their estimates presented a low skill in providing an accurate classification of different precipitation intensities.
- All SREs overestimated the amount of light and moderate rain events ([1, 20) mm d^{-1}) and underestimated the number of heavy and violent events ($\geq$ 20 mm d^{-1}).
- Lack of rain gauges at higher elevation zones (over 2000 m a.s.l.) over most of the Chilean territory (south of 26.0°S), prevented an exhaustive assessment of SRE products in such areas, which are the ones that get most of the total annual precipitation. In this circumstance, the obtained SRE performances should only be considered as a first guess about how well a given SRE represents the real amount of water in an area. Recently, Ref 31 showed that an SRE that presented the worst evaluation when compared to in situ precipitation measurements resulted the best when used to force hydrological simulations in a humid basin of southern Chile.

ACKNOWLEDGMENTS

The author thanks Conicyt-Fondecyt 11150861 for financial support and Conicyt-Fondap 15110009 for providing the precipitation dataset used in this study.

REFERENCES

[1] Tobin, K. J. and Bennett, M. E., "Satellite precipitation products and hydrologic applications," *Water International* **39**(3), 360–380 (2014).
[2] 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," *Journal of Hydrology* **499**, 324–338 (2013).
[3] Werren, G., Reynard, E., Lane, S. N., and Balin, D., "Flood hazard assessment and mapping in semi-arid piedmont areas: a case study in Beni Mellal, Morocco," *Natural Hazards* **81**(1), 481–511 (2016).
[4] AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., and Hain, C. R., "Remote sensing of drought: Progress, challenges and opportunities," *Reviews of Geophysics* **53**, 452–480 (jun 2015).
[5] Naumann, G., Barbosa, P., Carrao, H., Singleton, A., and Vogt, J., "Monitoring drought conditions and their uncertainties in Africa using TRMM data," *Journal of Applied Meteorology and Climatology* **51**, 1867–1874 (oct 2012).

[6] Kidd, C., Levizzani, V., Turk, J., and Ferraro, R., "Satellite precipitation measurements for water resource monitoring," *JAWRA Journal of the American Water Resources Association* **45**(3), 567–579 (2009).

[7] Baez-Villanueva, O. M., Zambrano-Bigiarini, M., Ribbe, L., Nauditt, A., Giraldo-Osorio, J. D., and Thinh, N. X., "Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America," *Atmospheric Research* (2018).

[8] Gebregiorgis, A. S. and Hossain, F., "Understanding the dependence of satellite rainfall uncertainty on topography and climate for hydrologic model simulation," *IEEE Transactions on Geoscience and Remote Sensing, vol. 51, issue 1, pp. 704-718* **51**, 704–718 (2013).

[9] 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," *Journal of Hydrometeorology* **13**(6), 1760–1783 (2012).

[10] Ebert, E. E., Janowiak, J. E., and Kidd, C., "Comparison of near-real-time precipitation estimates from satellite observations and numerical models," *Bulletin of the American Meteorological Society* **88**(1), 47–64 (2007).

[11] Bisselink, B., Zambrano-Bigiarini, M., Burek, P., and de Roo, A., "Assessing the role of uncertain precipitation estimates on the robustness of hydrological model parameters under highly variable climate conditions," *Journal of Hydrology: Regional Studies* **8**, 112–129 (2016).

[12] Nikolopoulos, E. I., Anagnostou, E. N., Hossain, F., Gebremichael, M., and Borga, M., "Understanding the scale relationships of uncertainty propagation of satellite rainfall through a distributed hydrologic model," *Journal of Hydrometeorology* **11**(2), 520–532 (2010).

[13] 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," *Hydrology and Earth System Sciences* **21**, 1295–1320 (mar 2017).

[14] Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. M., van Dijk, A. I. J. M., McVicar, T. R., and Adler, R. F., "MSWEP V2 global 3-hourly 0.1° precipitation: methodology and quantitative assessment," *Bulletin of the American Meteorological Society* (2019). *in press.*

[15] INE, *Medio Ambiente - Informe Anual 2015.* Chile (2015). ISBN: 978-956-323-173-1. Available on line at http://www.ine.cl/canales/chile_estadistico/estadisticas_medio_ambiente/2015/informe-medio-ambiente2015.pdf. [Accessed on 29-Aug-2016].

[16] Valdés-Pineda, R., Pizarro, R., García-Chevesich, P., Valdés, J. B., Olivares, C., Vera, M., Balocchi, F., Pérez, F., Vallejos, C., Fuentes, R., Abarza, A., and Helwig, B., "Water governance in Chile: Availability, management and climate change," *Journal of Hydrology* (2014).

[17] Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B., "The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales," *Journal of Hydrometeorology* **8**(1), 38–55 (2007).

[18] Huffman, G. J., Adler, R. F., Bolvin, D. T., and Nelkin, E. J., "The TRMM multi-satellite precipitation analysis (TMPA)," in [*Satellite Rainfall Applications for Surface Hydrology*], Gebremichael, M. and Hossain, F., eds., 3–22, Springer Dordrecht Heidelberg, London New York (2010).

[19] Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., and Michaelsen, J., "The climate hazards infrared precipitation with stations-a new environmental record for monitoring extremes," *Sci Data* **2**, 150066 (2015).

[20] Beck, H. E., van Dijk, A. I. J. M., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., and de Roo, A., "MSWEP: 3-hourly 0.25° global gridded precipitation (1979-2015) by merging gauge, satellite, and reanalysis data," *Hydrology and Earth System Sciences* **21**, 589–615 (jan 2017).

[21] World Meteorological Organization, [*Guide to meteorological instruments and methods of observation. WMO-No. 8*], ch. Present and past weather; state of the ground, I.14–1–I.14–9, World Meteorological Organization, Geneva, Switzerland, Seventh ed. (2008).

[22] Kling, H., Fuchs, M., and Paulin, M., "Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios," *Journal of Hydrology* **424-425**, 264–277 (2012).

[23] Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F., "Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling," *Journal of Hydrology* **377**(1-2), 80–91 (2009).

[24] Guo, H., Chen, S., Bao, A., Hu, J., Gebregiorgis, A., Xue, X., and Zhang, X., "Inter-comparison of high-resolution satellite precipitation products over Central Asia," *Remote Sensing* **7**, 7181–7212 (jun 2015).

[25] 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," *Atmospheric Research* **163**, 117–131 (2015).

[26] Jolliffe, I. T. and Stephenson, D. B., eds., [*Forecast verification: A practitioners guide in atmospheric science*], John Wiley & Sons Ltd, England (2003).

[27] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria (2018). https://www.R-project.org/.

[28] Hijmans, R. J., *raster: Geographic Data Analysis and Modeling* (2016). R package version 2.5-8. https://CRAN.R-project.org/package=raster.

[29] Zambrano-Bigiarini, M., *hydroGOF: Goodness-of-fit functions for comparison of simulated and observed hydrological time series* (2017). R package version 0.3-10.

[30] Zambrano-Bigiarini, M., *hydroTSM: Time Series Management, Analysis and Interpolation for Hydrological Modelling*. R package version 0.5-1 . https://doi.org/10.5281/zenodo.83964.

[31] Hann, H., Nauditt, A., Zambrano-Bigiarini, M., Thurner, J., McNamara, I., and Ribbe, L., "Hydrological evaluation of satellite-based rainfall estimates in a data-scarce Andean catchment to support drought management," *Hydrological Processes* (2018). (*submitted*).