Journal of Hydrology: Regional Studies 63 (2026) 103083

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

journal homepage: www.elsevier.com/locate/ejrh

Spatiotemporal Non-Linear Dynamics Assessment (SNLDA) of ERA5-Land precipitation in the Magdalena River Basin

Santiago Duarte [a,b,*], Gerald Corzo [a], Dimitri Solomatine [a,b,c], Remko Uijlenhoet [b]

[a] IHE Delft Institute for Water Education, Delft, the Netherlands
[b] Delft University of Technology, Delft, the Netherlands
[c] Water Problems Institute of the Russian Academy of Sciences, Moscow, Russia

ARTICLE INFO

Keywords:
Reanalysis
ERA5
Spatiotemporal
Attractors
Error
Tropical regions

ABSTRACT

Study Region: The study region is the Magdalena River basin in Colombia. The basin was divided into three distinct regions (Andean, Caribbean, and Pacific) and analyzed across different elevations.
Study Focus: The study proposes a Spatiotemporal Non-Linear Dynamics Assessment (SNLDA) framework to compare ERA5-Land reanalysis data with in-situ rain gauge observations. It specifically examines the constraints imposed by nonlinear dynamical processes and their associated space-time complexities on the representation of precipitation, particularly in a tropical region. The SNLDA framework incorporates three main components: (i) standard performance metrics (e.g., correlations, RMSE, and dry spell duration), (ii) rainfall spatiotemporal objects (characterizing precipitation events through attributes such as volumes and start-end centroids), and (iii) nonlinear dynamics complexity (reconstructing dynamical behavior from time series and evaluating attractors properties, including the Hurst and Lyapunov exponents). These elements were analyzed both individually and in combination. Daily ERA5-Land information (0.1°x0.1°) and in-situ rain gauge data comprising 558 stations from 1980 to 2020 were used, enriched by an Inverse Distance Weighting (IDW) interpolation (0.1°x0.1°) to facilitate comparison across spatial scales.
New Hydrological Insights for the Region: Overall, ERA5-Land overestimates precipitation, producing shorter, more frequent events while poorly representing extreme wet and dry spells.
Andean region: ERA5-Land overestimates rainfall, with largest errors at low elevations, driven by unresolved spatiotemporal object volumes displacements and nonlinear processes.
Caribbean region: ERA5-Land shows the highest errors in nonlinear dynamics and extremes, despite lower annual bias and RMSE.
Pacific region: ERA5-Land strongly overestimates precipitation volumes and RMSE, while nonlinear errors remain low; these biases are mainly driven by spatiotemporal objects displacement.

1. Introduction

Reanalysis datasets have become a valuable alternative for obtaining climatic information in basins with sparse rain gauge networks (Wang et al., 2019). These products integrate satellite data, such as GPM, with model outputs to generate comprehensive

* Corresponding author.
E-mail address: santiagoduarte09@gmail.com (S. Duarte).

https://doi.org/10.1016/j.ejrh.2025.103083
Received 2 September 2025; Received in revised form 15 December 2025; Accepted 23 December 2025
Available online 6 January 2026

gridded datasets. Traditional methods for evaluating precipitation estimates, however, often show discrepancies in their spatiotemporal accuracy (Laverde-Barajas et al., 2020). While efforts to improve local rain gauge networks through spatiotemporal interpolation and incorporation of additional variables have reduced uncertainty (Liu et al., 2017), limitations remain. Many regions, particularly in developing or remote areas, suffer from sparse observations, complex terrain, and limited spatial coverage, hindering accurate precipitation measurement (Derin et al., 2019). As a result, alternative sources, including satellite products and merged estimates, are increasingly used in hydro-climatological analyses (Beck et al., 2019). In this context, reanalysis datasets have gained global prominence due to their consistency, comprehensive coverage, and growing availability (Wang et al., 2019; Di Napoli et al., 2023).

Regarding reanalysis products, the ERA5 reanalysis family (ERA5 and ERA-Land), developed by the European Centre for Medium-Range Weather Forecasts (ECMWF), serves as a global benchmark for atmospheric studies (Albergel et al., 2018). ERA5 provides worldwide coverage, while ERA-Land offers higher-resolution data over land at shorter time scales (Balsamo et al., 2015). Rather than directly assimilating satellite precipitation products such as IMERG or CHIRPS, ERA5 assimilates various observations, including satellite radiances, to estimate atmospheric states and precipitation (Fessehaye et al., 2022). These datasets deliver consistent, long-term atmospheric information at high temporal resolution over multiple pressure levels, making ERA5 a comprehensive reference for hydrometeorological research.

In tropical regions, global reanalysis products often overestimate precipitation relative to rain-gauge observations (Sun et al., 2018). This bias is driven in part by monsoon circulation, the El Niño–Southern Oscillation (ENSO) (Salas et al., 2020), and the Intertropical Convergence Zone (ITCZ) (Durán-Quesada, 2012), all of which pose challenges for precipitation modeling (Sorí et al., 2021). Although newer products such as ERA5-Land show improvements, notable limitations remain in representing precipitation in areas with complex tropical hydroclimatology (May et al., 2022; Wright et al., 2020; Muñoz-Sabater et al., 2021). The Magdalena River basin in Colombia was selected as a case study, as it exemplifies these precipitation representation challenges through its marked spatiotemporal variability, strong oceanic influence and complex orographic conditions.

A wide range of metrics has been used to evaluate gridded precipitation datasets on a global scale (including reanalyses), spanning from standard statistical metrics to extreme event indicators. For example, Gomis-Cebolla et al. (2023) used KGE, RMSE, relative bias, and spatial correlations to assess datasets in Spain. Steinkopf and Engelbrecht (2022) evaluated CHRIPS, ERA5, and ERA-Interim over Africa at multiple time scales using statistical metrics (e.g., RMSE and correlation) as well as spatial correlation metrics (Spearman's rank correlation). Bodjrènou et al. (2025) assessed ERA5 and ERA5-Land over West Africa using Pearson correlation and relative mean absolute error. Gbode et al. (2023) analyzed extremes in CHRIPS, ERA5, and ERA-Interim using six extreme indices and examined the spatiotemporal variability across different scales.

In the Magdalena River basin, diverse metrics have also been employed: Valencia et al. (2023) used KGE, Pearson correlation, RMSE, and detection-skill metrics in multiple precipitation products in Colombia; Morales-Acuña et al. (2021) applied Spearman correlation, MAE, and bias over CHIRPS v2.0 in the Magdalena Department, Colombia; Romero-Hernández et al. (2024) combined EOF analysis with KGE and RMSE in Upper Cauca River Basin; López-Bermeo et al. (2022) incorporated PCA and wavelet analysis for CHRIPS validation; and Bojacá et al. (2025) evaluated the RMSE and a model-agreement index for the ERA5-Land and NASA POWER climate reanalysis products

Although multiple metrics are widely applied to evaluate limitations in reanalysis datasets such as ERA5-Land, these assessments remain largely statistical and often fail to capture the complex spatiotemporal and non-linear characteristics of precipitation, particularly in dynamically complex tropical regions (Salazar et al., 2025). This limitation complicates the comprehensive characterization of dataset deficiencies and underscores the persistent challenge of integrating multiple error sources within a unified framework, as conflicting metrics and unclear error interrelations limit robust decision-making (Zhao and Zhong, 2024). To address this, several methodological frameworks have been proposed, including multi-criteria evaluation techniques that synthesize diverse performance metrics (Lawal et al., 2021), multi-scale analytical approaches that explicitly represent scale-dependent variability and uncertainty (Singh and Mohanty, 2025), and statistical intercomparison methods aimed at reconciling inconsistencies among observational and gridded datasets (Centella-Artola et al., 2020).

Regarding advances in spatiotemporal dynamics, recent studies have proposed novel representations of precipitation behavior across space and time. (Li et al., 2021). One approach represents precipitation as spatiotemporal objects (3D entities defined by latitude, longitude, and time) from which volumetric characteristics can be derived. For instance, Laverde-Barajas et al. (2020) constructed precipitation objects for extreme events in Brazil using satellite data to decompose associated errors, while Díaz et al. (2023) analyzed spatiotemporal objects of meteorological droughts under varying thresholds. Another approach reconstructs attractors from precipitation time series to explore system dynamics, including deterministic chaos and anomaly detection. Duarte et al. (2018) identified chaotic behavior in precipitation records from local stations and GCMs in the upper Magdalena River Basin, and Agarwal et al. (2017) examined multiscale event relationships through chaos theory.

This study addresses the research question: To what extent does ERA5-Land reproduce the statistical, nonlinear dynamical, and spatiotemporal characteristics of precipitation across the Magdalena River basin, and how are errors from these dimensions interconnected?. For this purpose, we developed a novel Spatiotemporal Non-Linear Dynamics Assessment (SNLDA) framework, We hypothesize that by integrating standard, non-linear, and spatiotemporal metrics, the SNLDA framework will reveal error patterns in ERA5-Land precipitation that are not captured by standard metrics alone, particularly highlighting spatiotemporal dynamics in the Magdalena River basin. SNLDA integrates three complementary dimensions of evaluation: (1) Precipitation standard metrics (i.e., RMSE, correlation, annual maxima and longest dry spells), (2) properties of reconstructed attractors derived from time series (i.e., embedding dimension, Lyapunov spectrum and Hurst exponent), and (3) characteristics and distribution of spatiotemporal objects of precipitation events (i.e. volume, duration, and centroid location). No prior framework integrates statistical, attractor, and spatiotemporal object metrics in a single methodology. By conducting individual analyses and integrating error metrics from each dimension

within a multi-metric error plot, the framework enables a thorough and multidimensional assessment of ERA5-Land precipitation. This methodology reveals its main strengths and limitations and provides a clearer understanding of error attribution. For this analysis, the basin was divided into three regions (Pacific, Caribbean, and Andean).

This paper is organized into six sections. Section 2, describe the Case Study (Magdalena River basin), Section 3 Methodology, details the Data Sources (Local stations and ERA5-Land) and techniques for Spatiotemporal Non-Linear Dynamics Assessment (SNLDA). Section 4 presents the Results, followed by a Discussion in Section 5. Finally, Section 6 offers the Conclusion.

2. Case study

2.1. Magdalena River basin overview

The Magdalena River basin, situated in Colombia, Northern South America (see Fig. 1), has an area of 273,459 km^2, representing 23.9 % of Colombia's total land area, with an average precipitation of 2050 mm/year (Vega-Viviescas & Rodríguez-S, 2019). Water resources play a crucial role in this area, supporting hydroelectric power generation, agriculture, biodiversity, and recreation (Alvaréz-Villa et al., 2011). However, the basin is also susceptible to the impact of extreme hydrological events, with a long history of disasters related to floods and droughts. Notably, significant events occurred in 2010 and 2014 (Avila et al., 2019).

Regarding the climatology in the region, multiple drivers influence climate dynamics across the basin. Given the latitudes of the basin in the range of 1°–12°N, it is located in a tropical region characterized by Intertropical Convergence Zone (ITCZ) conditions. In this context, Urrea et al. (2019) highlighted the migration of the ITCZ as a main driver of Colombia's hydroclimatology. Additionally, other factors such as the Choco low-level jet in the Pacific and the Caribbean jet in the Atlantic have a significant impact on the basin. Moreover, the interaction between the two oceans is relevant for precipitation studies (Taylor et al., 2002). Furthermore, the effects of the El Niño Southern Oscillation (ENSO) are considerable in the country (Bedoya-Soto et al., 2019). These elements influence the magnitude, volume, duration, and frequency of extreme hydrological events (Wang et al., 2021). Additionally, physical parameters, such as elevation and wind, can also impact the generation and magnitude of extreme event conditions in the region (López et al., 2018).

The Magdalena River Basin exhibits substantial spatial precipitation variability driven by its complex orography and climatological conditions, which generate heterogeneous rainfall patterns across elevations and seasons, as noted by (Elgamal et al., 2017) in their evaluation of satellite rainfall products for flow simulations. This orographic complexity (Gomez-Rios et al., 2023) arises from the division of the Andes mountains into three separated mountain chains in the south, as well as the presence of the Sierra Nevada de Santa Marta massif (5730 m) in the northeast. The spatial distribution of the orography is illustrated in Fig. 2a. Accordingly, distinct precipitation regimes are evident along the basin, highlighting contrasts among the Andean, Caribbean, and Pacific sub-regions, as

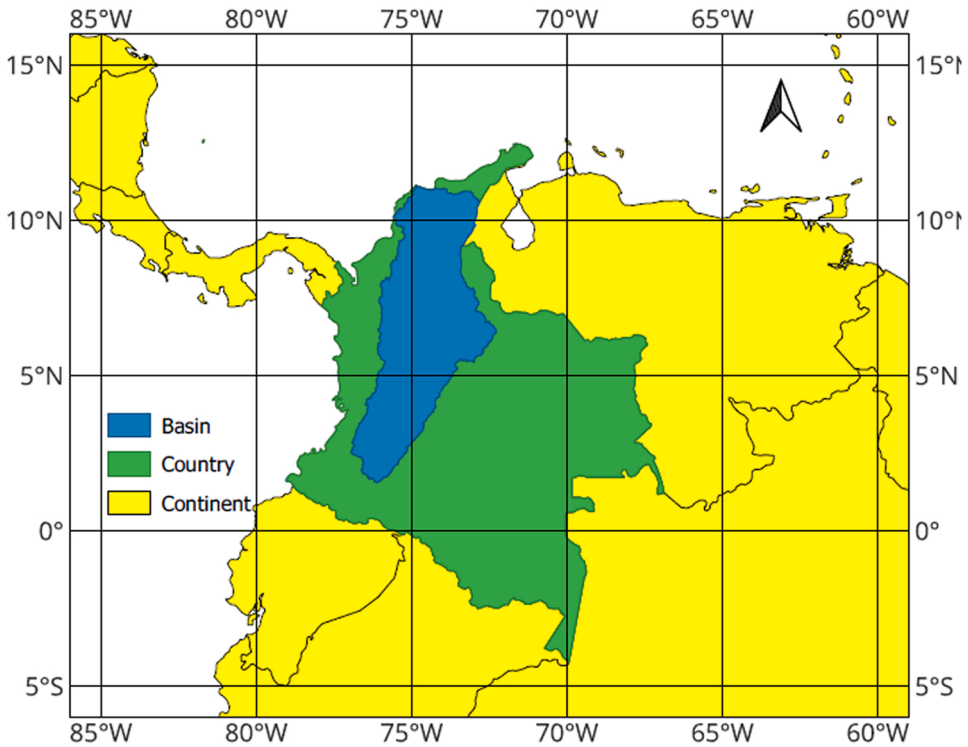

Fig. 1. Magdalena River Basin location in country and continent. Blue-Basin, Green-Country (Colombia), Yellow-Continent.

shown in Fig. 2b, which presents the regional division used in this study along with key reference rain gauge stations. This division was based on the sub-basins of the Magdalena River basin. The Pacific region comprises all sub-basins between the central mountain range and the western mountain range, while the Andean region includes the extension between the central mountain range and the eastern mountain range. On the other hand, the Caribbean region involves sub-basins in the most northern region.

Regarding the temporal distribution of precipitation, strong interannual variations in precipitation along the basin are observed. Two wet periods and two dry periods are identified: wet seasons in March-May (MAM) and September–November (SON), and dry seasons between December–February (DJF) and June–August (JJA) (Bolaños Chavarría et al., 2022). Notably, no seasonal temperature-related changes are observed across the basin; instead, the region is characterized by alternating wet and dry seasons, as reported in previous studies (Vega-Viviescas & Rodríguez-S, 2019). Furthermore, this basin has been shown to exhibit chaotic (non-linear) behavior in precipitation data at the daily time scale (Duarte et al., 2018). Add variations among stations.

On the other hand, interannual variations are highlighted in Fig. 3, which presents multiannual annual precipitation data for three rain gauge stations across different regions. For instance, annual precipitation at station pr_21150030 (Andean region) ranges from 640 mm to 2189 mm, while in the Caribbean region (station pr_25021000), it varies from 200 mm to 2557 mm. Furthermore, the influence of ENSO events on wet and dry periods is evident, particularly during the years 2000 and 2015 (Giraldo-Osorio et al., 2022).

2.2. Data sources

This study employed two primary datasets covering the period from 1980 to 2020. The first consists of daily precipitation records from local gauge stations provided by the Institute of Hydrology, Meteorology, and Environmental Studies (IDEAM), which served as the reference in-situ observations. The second is the ERA5-Land reanalysis dataset from the European Centre for Medium-Range Weather Forecasts (ECMWF); available at hourly intervals with a spatial resolution of $0.1^\circ \times 0.1$.

Our assessment primarily focused on daily precipitation, with potential applications in medium-term forecasting and the analysis of extreme events. Accordingly, hourly ERA5-Land data was processed to daily totals. To complement the study of precipitation spatiotemporal dynamics, the analysis was extended to monthly and annual time scales.

2.2.1. Local data

We obtained local information through the IDEAM open access portal for hydrometeorological data download (DHIME: http://dhime.ideam.gov.co/atencionciudadano/). Stations within the region of interest were pre-selected, and data were subsequently filtered based on the ratio of missing data [15 %] (Norazizi and Deni, 2019), the longest continuous missing information (366 days), full-time availability, and homogeneity with the closest stations. As a result, 558 stations were utilized out of the 1264 available in the region. Of these, 509 stations (91 %) collected data manually, while 49 (9 %) used automatic measurements, with a similar distribution across regions. In addition, 437 stations (78 %) measured only precipitation, whereas 121 (22 %) also recorded other climatic variables. In the Caribbean region, 88 % of stations measured precipitation exclusively, compared with 76 % in the Andean region and 72 % in the Pacific region. Regarding temporal resolution, 470 stations (84 %) recorded data at a daily scale, while 88 stations (16 %)

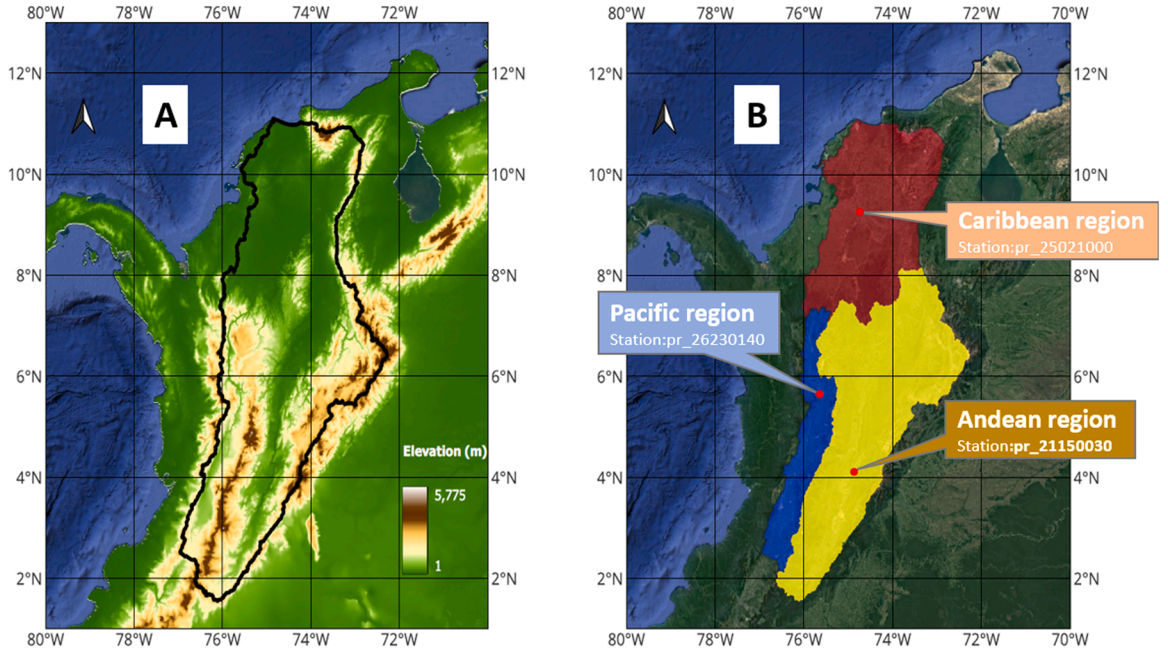

Fig. 2. A. Digital Elevation Model (DEM) obtained from the USGS map service. B. Basin's sub-regions regarding its location.

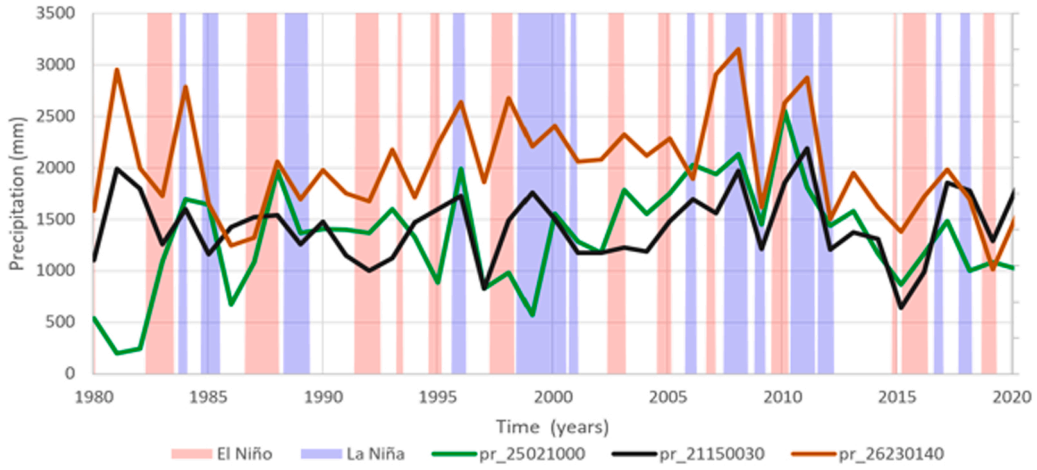

Fig. 3. Multiyear annual precipitation for rain gauge stations Andean (black), Caribbean (green) and Pacific (orange). ENSO periods are indicated by red ("El Niño") and blue ("La Niña") bands.

measured precipitation at finer temporal resolutions. The spatial distribution of all 558 stations is presented in Fig. 4A, and a comparison of network statistics by region is shown in Table 1, where is shown that the Pacific region has the highest station density, while the Caribbean has the lowest.

Gap filling for a given date was performed by comparing the station's historic average values (daily, monthly, and annual) with the corresponding daily, monthly, and annual conditions of the given date. Although any data imputation may affect the statistics of the data and, subsequently, the baseline for error comparison, we anticipate a minimal impact. This is attributed to both the small proportion of missing data and the spatial distribution of the stations.

To enable a consistent spatiotemporal comparison across the study domain, the station observation data was spatially interpolated

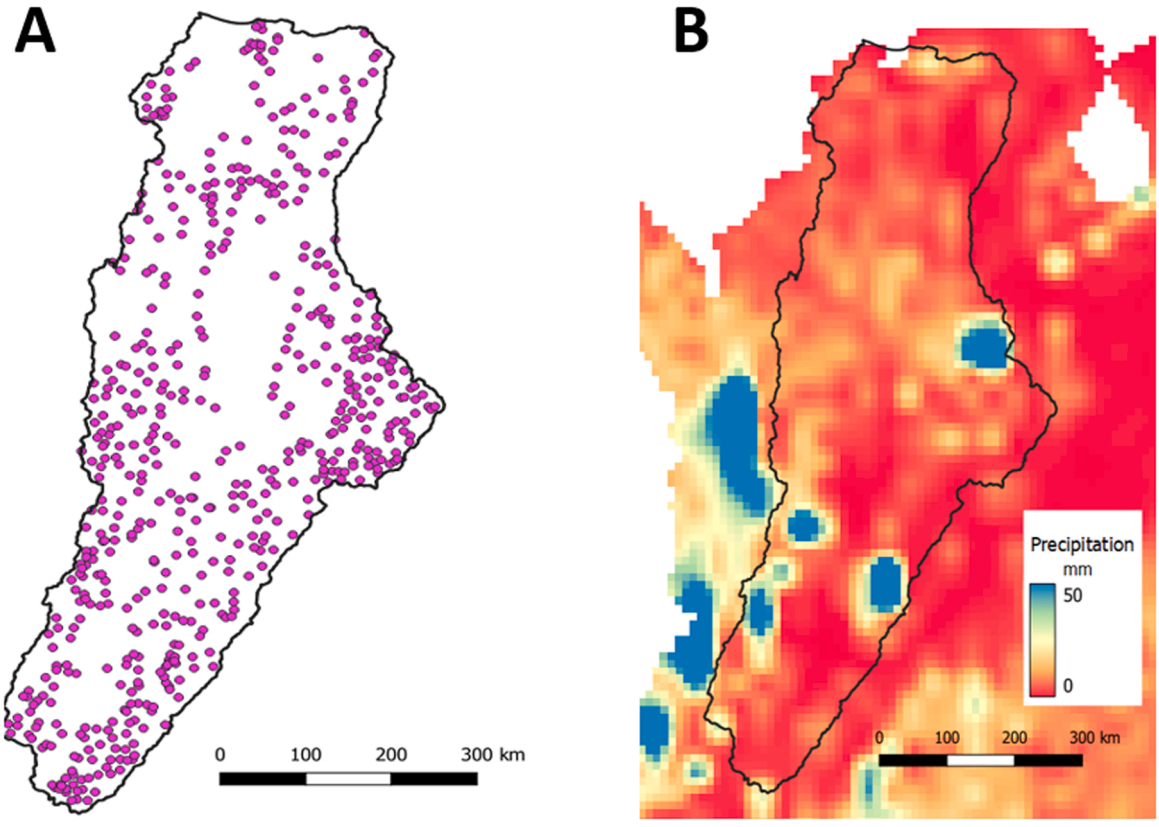

Fig. 4. A. In-situ rain gauge distribution and B. ERA5-Land Data. 1st January-1980. Blue: High precipitation – Red: low precipitation.

Table 1
Stations network statistics by region.

Region	Number of stations	Average station density (km^2/station)	Average nearest neighbor station distance (km)
Andean	313	432.8	11.1
Caribbean	136	721.9	13.4
Pacific	109	342.5	10.7

to match the resolution of ERA5-Land (0.1° × 0.1°). Inverse Distance Weighting (IDW) served as the spatial interpolation method for station observations, due to its approach of interpolating based on constant distances for all time steps. IDW was tested previously by López et al. (2018) in the Magdalena River Basin showing to outperform other methodologies (nearest neighbor algorithm and cubic spline). However, this study also suggested further study the effects of complex orography and non-homogeneous spatial distribution. Consequently, the interpolation was validated with the local station data through Root Mean Squared Error (RMSE) and Pearson's Correlation (R).

2.2.2. *ERA5-Land data*

ERA5-Land belongs to the fifth generation and is the latest set of reanalysis products from the ERA5 family dataset. ERA5 reanalysis data is obtained by merging climate model outputs with observations. Regarding the involved models, ERA5 uses the IFS model and the Cy41r2 implementation, which considers multiple physical processes, including but not limited to radiative transfer, convection, clouds, and orographic drag. In terms of observations, preprocessing and assimilation operations are applied to link them to their equivalents in the physical model. This process is performed for conventional local stations, as well as satellite-derived variables. Furthermore, ERA5 uses a 4D-Var data assimilation, based on a multi-resolution incremental method (Tompkins and Janiskov, 2004). These methods have been shown to provide consistent data. For example, Gomis-Cebolla et al. (2023) tested the data in Spain for accuracy (RMSE between 2 and 8 mm/d) over a long time series (70 years) against gridded precipitation data from the Spanish Meteorological Agency. Regarding the ERA5-Land dataset, it is a derivation of ERA5 with increased accuracy and spatial resolution in the land surface. This is obtained through a recalculation of the ERA5 land component with a spatial resolution grid of ∼ 9 km and has a focus on land surface hydrology (Muñoz-Sabater et al., 2021). Fig. 4B shows the precipitation spatial variability under this dataset for the 1st of January 1980. For the case of this study, daily precipitation was obtained directly from the dataset, based on the hourly

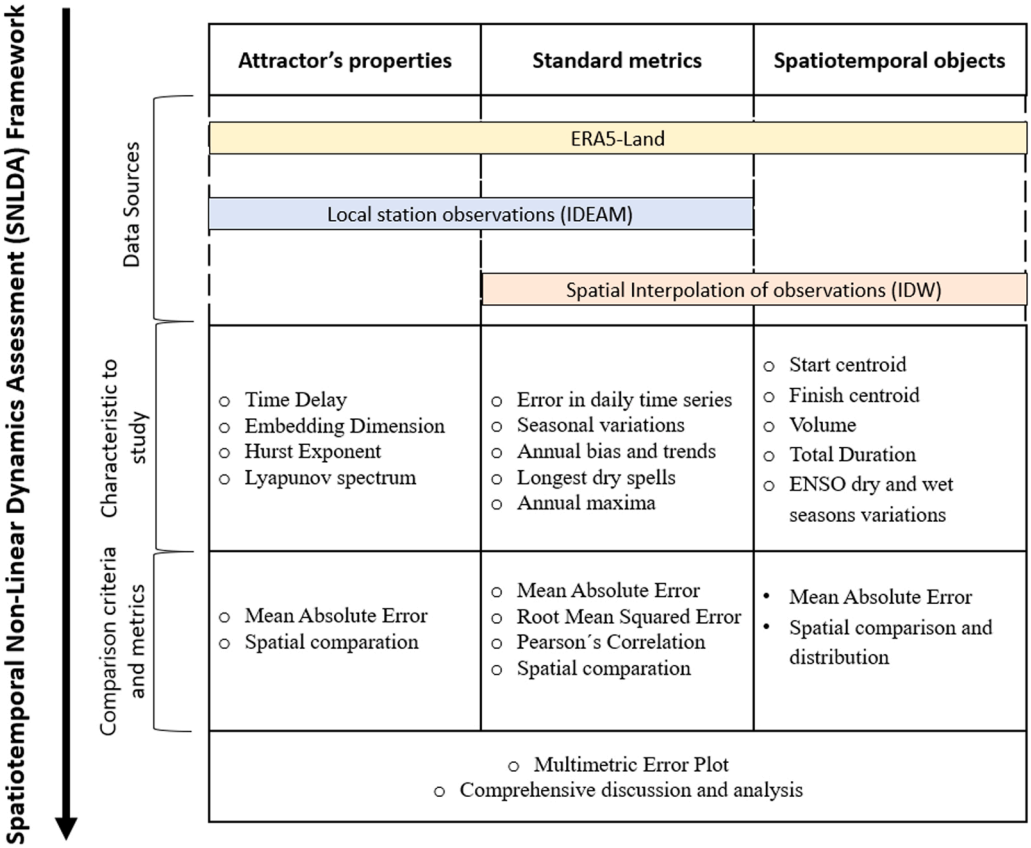

Fig. 5. General methodology: SNLDA methods, data sources and comparison metrics.

accumulations included in the NetCDF file. ERA5-Land hourly data was accessed through the Copernicus portal (https://cds.climate.copernicus.eu/datasets/reanalysis-era5-land).

3. Methodology

The ERA5-Land dataset was used to validate the limitations in representing precipitation within tropical landscapes using multiple methodologies: standard metrics, attractor properties, and spatiotemporal objects. The employed methodology is summarized in Fig. 5.

These representations were evaluated and compared with those obtained from the reference (local observations from IDEAM). To represent local point observations as spatiotemporal objects, the station data were interpolated using the IDW method ($0.1^\circ \times 0.1^\circ$ grid) to create the baseline dataset.

Additionally to the subregions division (Andean, Caribbean and Pacific), three elevation categories (low, medium and high) were established. Given the elevation variability among regions (Fig. 2A), thresholds of 150 m and 1850 m were defined to classify the elevation ranges. These thresholds were derived by calculating the 30th and 70th elevation percentiles of elevation for each region, using the elevation data from the 558 rain gauge stations. The resulting percentile values were then averaged to obtain comparable elevation classes, ensuring consistency across regions.

3.1. SNLDA (Spatiotemporal Non-Linear Dynamics Assessment)

Standard metrics have limitations in assessing errors in reanalysis data, particularly due to the spatiotemporal complexity and nonlinear behavior of climate dynamics (Schutgens et al., 2017). To address this, the Spatiotemporal Non-Linear Dynamics Assessment (SNLDA) offers a comprehensive framework that incorporates these components. Thus, time series statistics, attractor properties, and spatiotemporal characteristics were calculated and compared across different regions using various metrics (Fig. 5). General basin analyses are conducted, while regional variations are explored in depth. Here, the performance of multiple metrics (i.e., RMSE, maximum Lyapunov exponent differences, and MAE of object volumes) is compared to providing complementary information about precipitation dynamics. Additionally, as part of our basin characterization, we analyzed data from local stations to estimate the annual average and standard deviation of precipitation across the sub-regions.

3.1.1. Standard metrics

Regarding the evaluation of ERA5-Land precipitation using standard metrics ERA5-Land was compared against Local Observations. Here, we performed a direct point-to-pixel validation, in which each rain gauge was paired with the ERA5-Land grid cell whose center lay closest to the gauge location. This nearest-grid-cell selection, commonly referred to as the Simple Nearest Point (SNP) method, is widely used in satellite and reanalysis validation studies, as observed in Valencia et al. (2023) who applied this methodology in this region; Muñoz-Sabater et al. (2021) when describing ERA5-Land state of the art applications and Bodjrènou et al. (2025) in West-Africa ERA5-Land validation. Additionally, the performance of the interpolated gauge observations (IDW interpolation), was assessed against the raw local observations using the same SNP approach. For both cases, performance was evaluated using Root Mean Squared Error (RMSE) and Pearson's correlation coefficient (R), which are defined in Eqs. 1 and 2 respectively. To further analyze performance regarding associated extremes, differences in annual maxima and longest dry spells were computed. For the longest dry spells, a threshold of 1 mm was used to identify dry days. Additionally, time series of daily precipitation were aggregated to derive monthly and annual time series. Biases at monthly and annual time scales were computed. These results were then compared across the four seasons comprising the dataset.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \tag{1}$$

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\sum_{i=1}^{n}(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{2}$$

Where $x_i$and $y_i$denote the precipitation values, of the reference and the evaluated precipitation datasets, and n is the length of the time series.

3.1.2. Attractor properties

Phase space is a vectorial space of $\boldsymbol{m}$ dimensions in which time series (i.e. precipitation) are transformed, such that every possible state of the system's dynamics (i.e. climatic system) can be represented completely (Koutsoyiannis, 2006). Here, the trajectory describes the evolution of the system for a given initial state within a region of attraction; this is the attractor (Sivakumar and Berndtsson, 2010). Two parameters were used for the phase space reconstruction, time delay and the embedding dimension. The time delay (τ) was selected as the value between where the autocorrelation first decayed to 1/e of its initial value, and the lag where the mutual information reaches its first minimum (Casdagli, 1991). The embedding dimension was computed using the False Nearest Neighbors

(FNN) method (Abarbanel and Kennel, 1993). In this method, first, a time series is reshaped in the phase space following the selected lag. Second, for a range of given dimensions, the nearest neighbor for each point is identified. Third, the percentage of false neighbors is calculated. Finally, given a threshold of acceptance of false neighbors (20 %), the lowest dimension for which this criterion is achieved was selected as the embedding dimension (***m***). Once the attractor was reconstructed, the properties of its dynamic stability were estimated. The Lyapunov spectrum ($\lambda_1, \lambda_2, \ldots, \lambda_i,$) and the Hurst exponent (H) were then computed. The Lyapunov spectrum describes the exponential instability of close trajectories in the phase space under different initial conditions (Eckmann and Ruelle, 1992). Positive finite maximum Lyapunov exponents ($\lambda_{max} > 0$) are related to deterministic chaotic systems, while negative exponents are associated with stable motions. The algorithm implementation of Eckmann et al. (1992) was used in Python under the package "nolds" to obtain the whole Lyapunov spectrum and the Maximum Lyapunov exponent (λ_{max}). The Hurst exponent provides insights related to the persistence (H < 0.5) or anti-persistence (H > 0.5) of the time series. This indicator is highly related to the fractal dimension that characterizes deterministic chaotic systems (Planinić et al., 2004). For the comparison between the attractor characteristics of ERA5 and local stations, the absolute difference metric was used.

3.1.3. Spatiotemporal object characteristics

Spatiotemporal objects were generated through spatial aggregation. A connected component label approach in three dimensions was the method selected for this task. The application of this method followed the approaches of Laverde-Barajas et al. (2020) and Diaz et al. (2023) for the analysis of hydrometeorological-related objects. Moreover, given the climatic characteristics of the region, special attention was given to periods when ENSO influence was larger. In this sense, the Oceanic Niño Index (ONI) was used to identify these periods, using as thresholds ONI < -0.5 ('La Niña') and ONI > 0.5 ('El Niño') (Peña-Q et al., 2015). Fig. 6 illustrates the process involved in object creation.

The spatiotemporal object creation and characterization steps are summarized below:

1. A binary mask representing the presence or absence of precipitation was generated by filtering the spatiotemporal dataset (latitude, longitude, time, value) using a threshold of > 1 mm/day to identify valid precipitation days.
2. The connected component label method was then used to identify 3D objects with the binary mask. Here, contiguous, and adjacent elements connected by N positions in space (lat, lon) and time are assumed to belong to the same object and are then labelled. N was fixed at 1 step for this study given the spatial and temporal resolution of information, as well as the spatial and temporal correlation of precipitation at the daily time scale.
3. Objects were counted and filtered based on whether they occurred during "El Niño" or "La Niña" periods. Only the 1000 largest objects (by precipitation volume) were retained for further analysis.
4. Object characteristics were computed using the Python package *skimage*. These properties included centroid (start, average, finish), volume, and duration. Duration refers to the total object's lifespan, while centroids are represented by 3D coordinates (latitude, longitude, time).

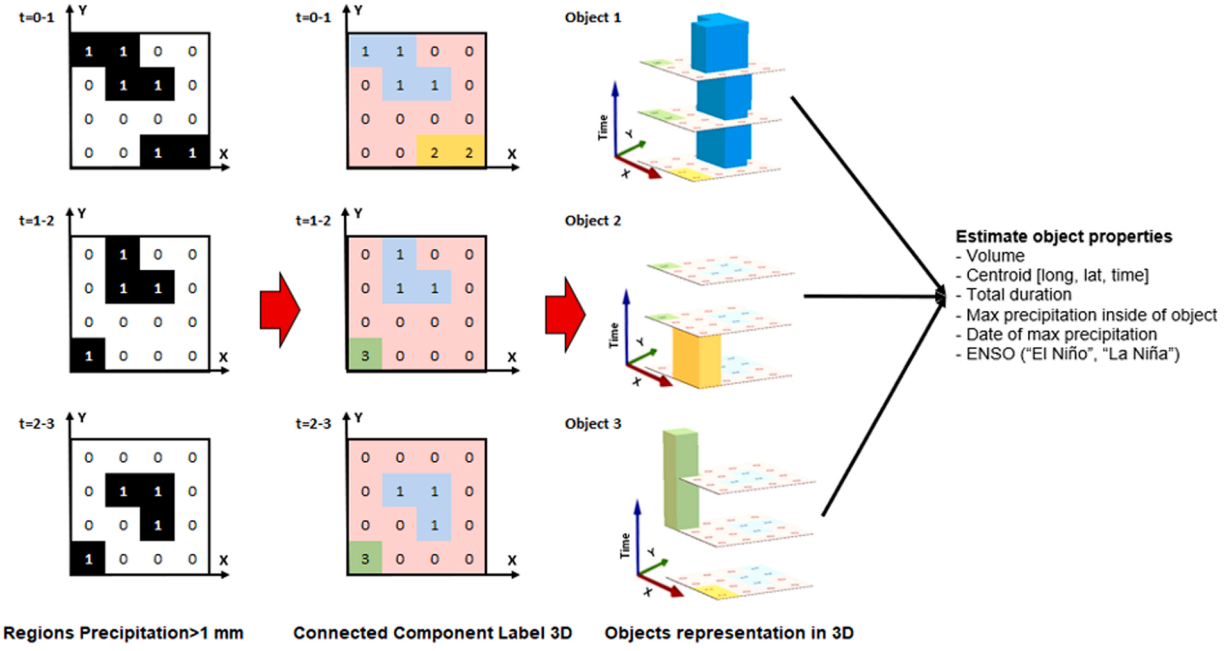

Fig. 6. Precipitation spatiotemporal object-events [long, lat, time] construction and analysis.

5. Objects were classified according to the position of their start and finish centroids within the basins: regions (Andean, Caribbean and Pacific) and elevations.
6. Finally, the spatiotemporal characteristics were compared between IDW local stations and ERA5-Land using MAE, for each combination of ENSO season, region, and elevation. The distribution of the identified objects was also analyzed.

In addition, the analysis accounted for the boundary effect by considering only spatiotemporal objects that both start and end within the river basin (i.e., those with at least some portion of their area inside the basin) were included in the study.

3.1.4. Spatiotemporal non-linear dynamics assessment

Once the error metrics are computed for each region and elevation, the discrepancies between local station data and ERA5-Land are analyzed using a multi-metric error plot, as shown in Fig. 7. This analysis includes an assessment of spatiotemporal non-linear dynamics, which examines how non-linear and spatiotemporal characteristics contribute to the observed differences. To achieve this, standard error metrics are plotted against the average differences in attractor's properties across regions and elevations. Additionally, characteristics of spatiotemporal objects are used to identify inter-regional relationships that may provide insights into the underlying causes of error. In this study, the most significant relation for each region (higher error in volume or duration), was plotted. This process was repeated separately for "El Niño" and "La Niña" periods.

4. Results

4.1. Standard metrics

Overall, Our analysis indicates averages of 1809 ± 209 mm/year in the Andean, 1931 ± 209 mm/year in the Caribbean, and 1989 ± 298 mm/year in the Pacific region. As a baseline reference for the standard metrics, the average daily rainfall is presented for each region and elevation. Andean region: low (6.76 mm/d), medium (5.33 mm/d), high (3.69 mm/d); Caribbean region: low (5.34 mm/d), medium (5.13 mm/d) and Pacific region: low (4.41 mm/d), medium (5.38 mm/d), high (5.68 mm/d). For the first experiment (ERA5-Land vs local observations), higher error was found in the Pacific (21 mm/d), and lower in the Andean region (14 mm/d). The Caribbean region presented intermediate values (16 mm/d). Fig. 8 illustrates the variation of RMSE along the basin. Regarding the variation of metrics with elevation, the error decreases with altitude in the Caribbean and Andean region, while in the Pacific this pattern is the opposite. In the Caribbean, the average RMSE was 16 mm/d for lower elevations and 17 mm/d for medium elevations; no stations with high elevations are present in this region. For the Andean region, values of 21, 16 and 10 mm/d were found for lower, medium, and high elevations, respectively.

On the other hand, in the Pacific region, RMSE increased from 12 to 21 and then 19 mm/d for the three elevation categories. Regarding the temporal distribution of error, the DJF season exhibits the lowest RMSE values for all regions. Conversely, the highest RMSE values were found in MAM for the Pacific and Andean region, and in SON for the Caribbean as shown in Fig. 8A. Moreover, comparable results were found for other metrics. The value of Pearson's correlation between local stations and ERA5-Land ranged between 0.04 and 0.20. On average correlations were less than 0.132, distribution pattern is presented in Fig. 8B.

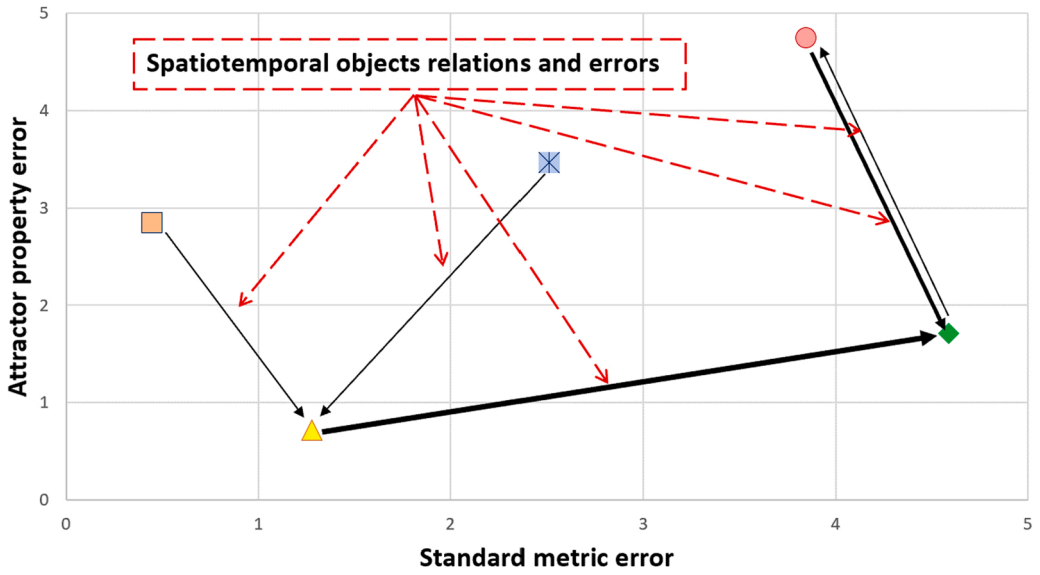

Fig. 7. Scheme of a multi-metric error plot for a spatiotemporal non-linear dynamics assessment. Points represent regions, and lines indicate spatiotemporal relationships.

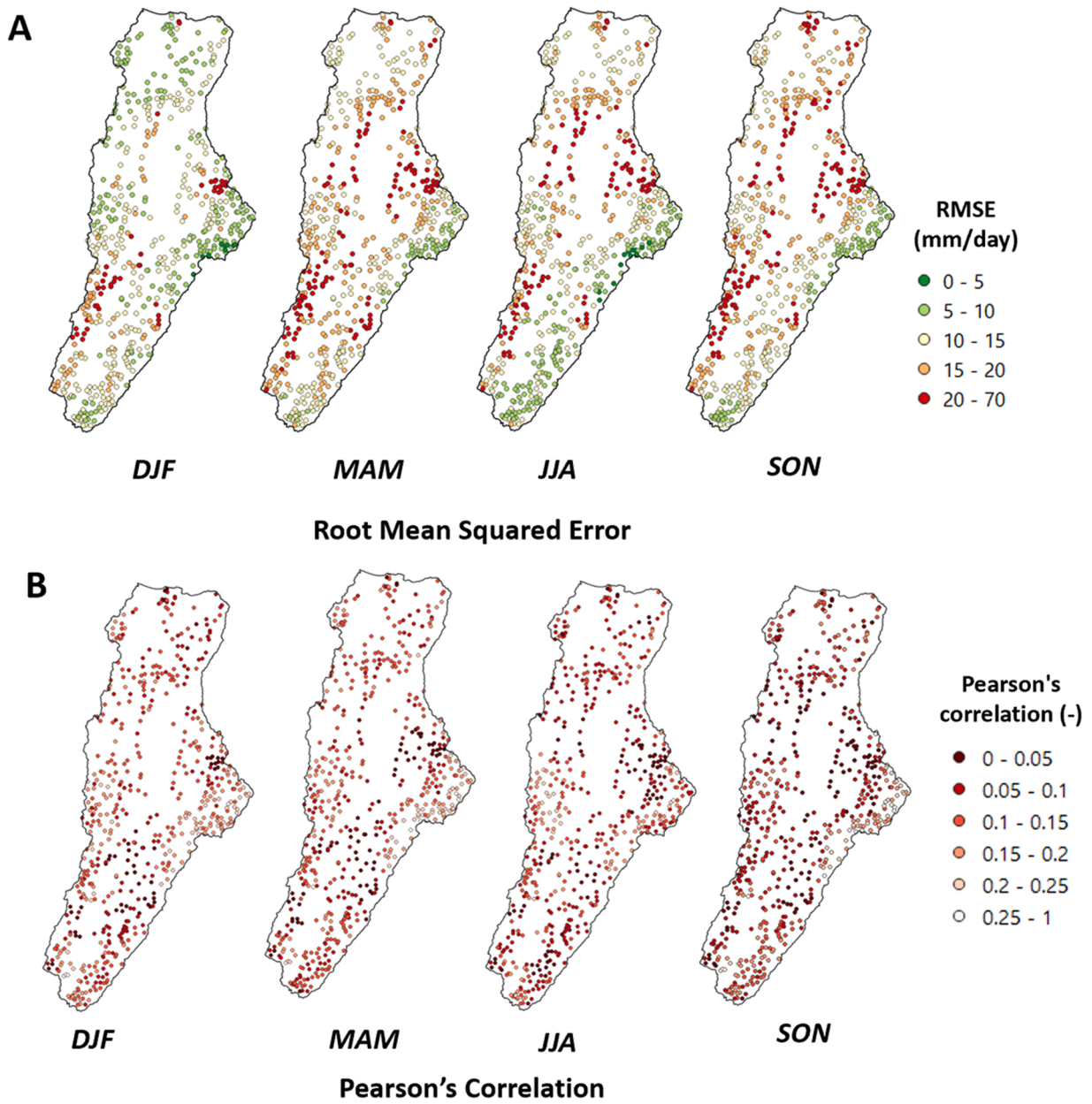

Fig. 8. Spatial comparison of Local observations vs ERA5-Land for each season with A. RMSE, larger error (red colour). B. Pearson's correlation, larger error (darker red shading).

In the second experiment (interpolation of local stations vs rain gauge local observations), average difference between mean daily values across regions was less than 2 %, while RMSE value was 4.3 mm/d for the data time range. Lower values were found in dry (3.9 mm/d) and higher in wet (4.8 mm/d) periods. Andean and Caribbean regions showed the highest errors for low elevations (5 mm/d). Pearson correlation oscillates between 0.89 and 0.94 in the regions.

Regarding the extremes in the datasets, from the dry perspective, on average the longest dry spells are longer at local stations than in ERA5-Land (for which this value is 41 days). This pattern is more intense in the Caribbean (77 days) than in the Andean (26 days) and Pacific (41 days), as observed in Fig. 9A. Moreover, for all the regions, the last decade (2010–2020) presents the largest differences in the length of dry spells with ERA5-Land, showing an increase of 77.8 % from the previous decade's average. On the other hand, from the wet viewpoint, on average the maximum precipitation values have a larger magnitude for local stations than in ERA5-Land. However, the pattern varies in the basin, as observed in Fig. 9B. In the Pacific, ERA5 values exceed gauge stations by 15.7 mm on average, while in the Caribbean the stations show values greater than 35.6 mm. No significant decadal change was observed for this characteristic.

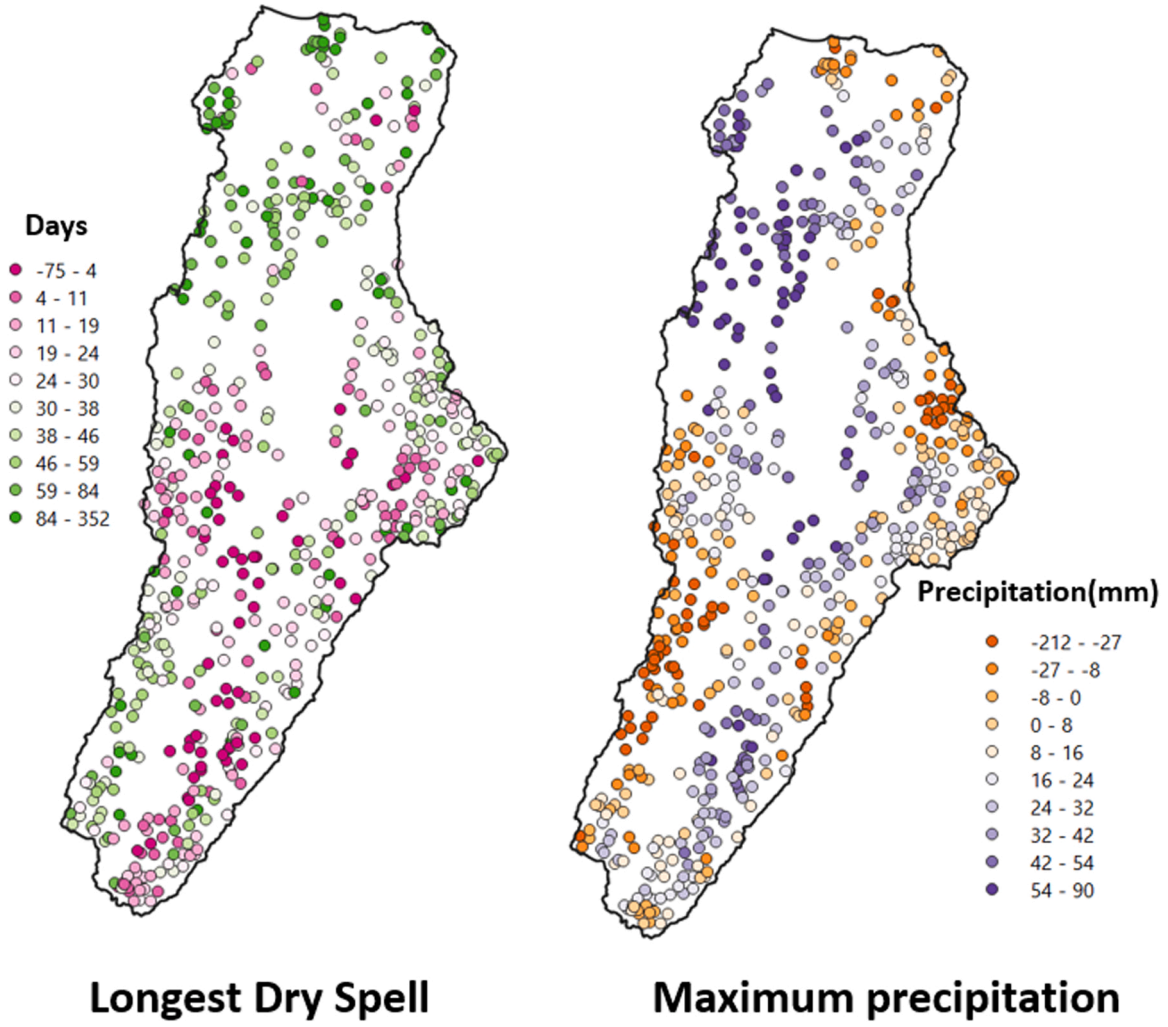

Fig. 9. Spatial comparison of differences between Local stations and ERA5-Land for: A. Longest dry spell and B. Maximum precipitation.

Additionally, about the annual characteristics, ERA5-Land precipitation volumes were 2.2 times larger on average than gauge stations. This value fluctuates around the basin, with values of 1.5, 2.0 and 3.4 for the Caribbean, Andean and Pacific regions, respectively. At the monthly time scale, average monthly precipitations of 259, 223 and 496 mm/month were found at local stations for the Andean, Caribbean and Pacific regions, while for ERA5-Land precipitations these values were 151, 161 and 166 mm/month for the same regions.

4.2. attractor properties

Fig. 10 illustrates the spatial variation of the attractor's properties along the basin for local observations. Time delay values ranged between 2 and 12 days, with an average of 3.8 days. The embedding dimension ($\boldsymbol{m}$) was on average 9.5, with minimum and maximum values of 4 and 16. However, for 51 stations in the Caribbean region the threshold acceptance was not achieved, for the number of dimensions explored. Regarding λ_{max}, they ranged from 0.038 to 0.987. All stations showed persistent behaviour (> 0.5) for the Hurst exponent, oscillating between 0.5 and 0.85.

As for ERA5-Land, Fig. 11 shows the main results obtained for the attractor's properties. Time delay oscillated in a wider range between 2 and 23 days. On the other hand, the embedding dimension varied from 3 to 6. Similarly to the results for local stations, Hurst exponents were found to vary between 0.56 and 1. Finally, λ_{max} ranged from 0.093 to 0.598.

Regarding the differences between ERA5 and the baseline reference, these results are summarised in Table 2 for the regions and elevation categories. The time delay increased slightly with elevation for the Andean and Caribbean regions, while no pattern was found for the Pacific. The embedding dimension was observed to decrease with increased elevation for all cases. For example, in the Andean region differences were 7.7, 5.5 and 5.0 for low, medium and high elevations. Moreover, larger values were noted in the Caribbean (9.9), and lower in the Pacific (4.5). As for λ_{max}, the larger differences were found at low elevations for the Andean (0.241)

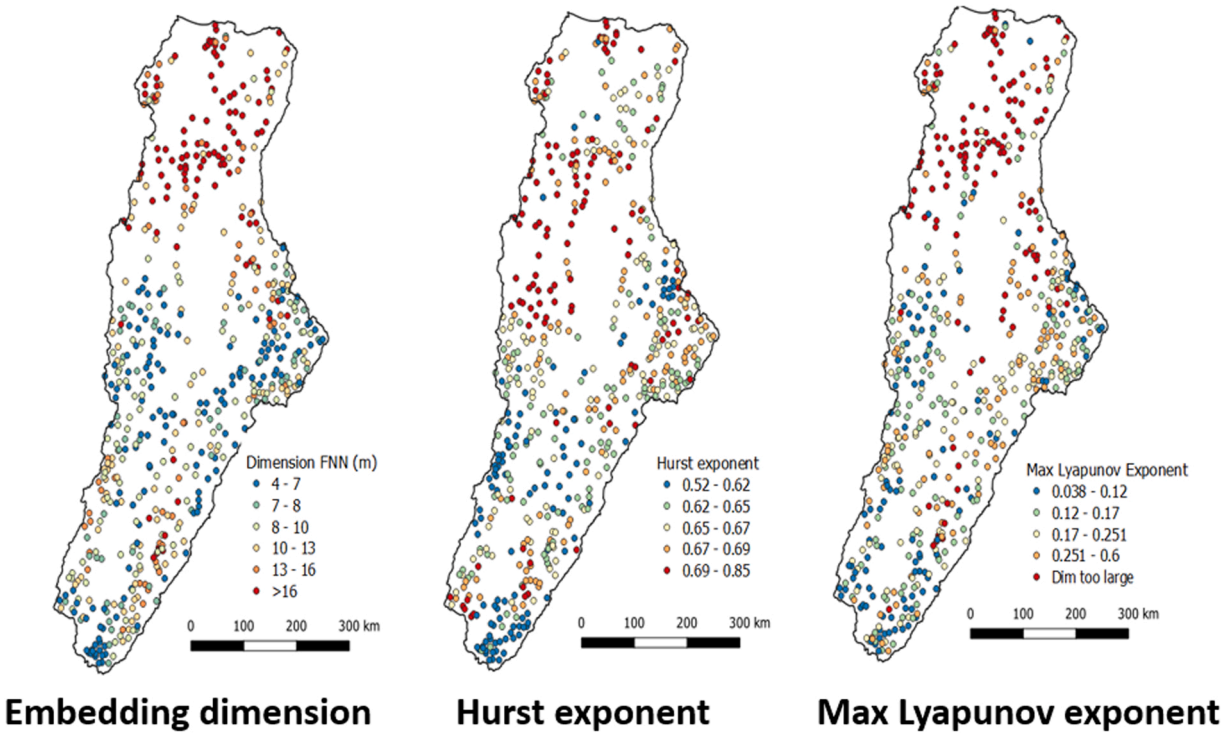

Fig. 10. Spatial comparison of attractor properties for Local stations: A. Embedding dimension, B. Hurst exponent and C. Max Lyapunov exponent.

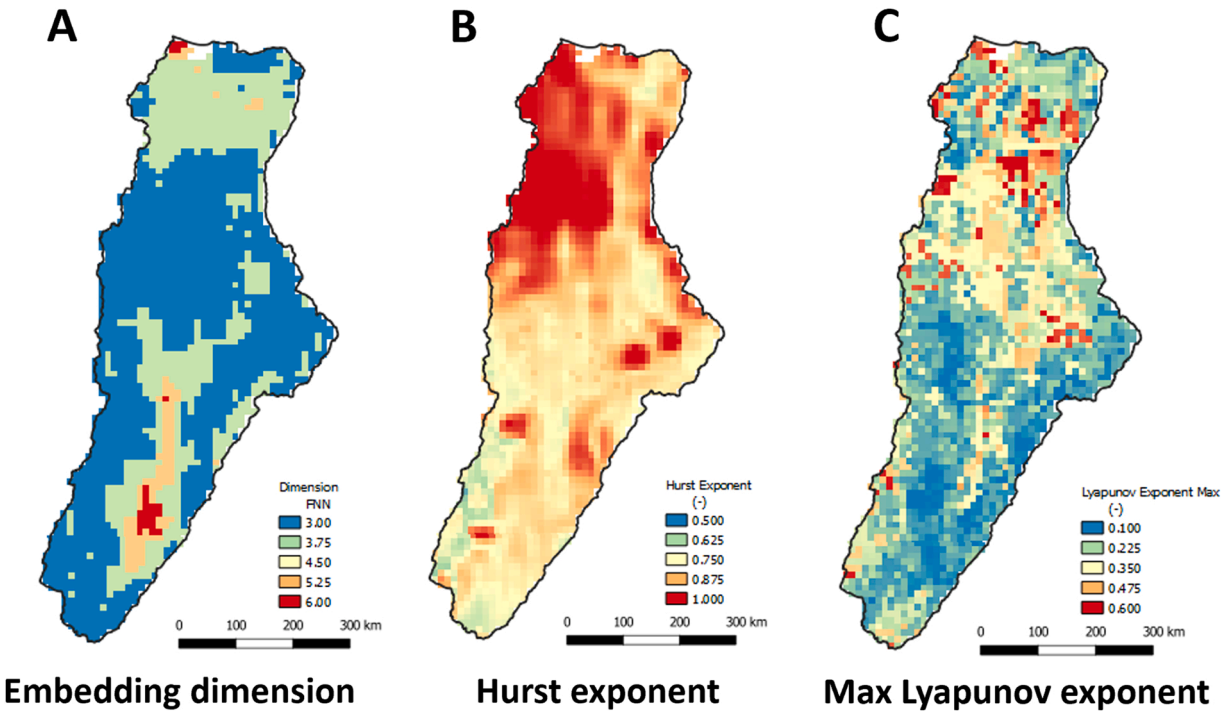

Fig. 11. Spatial comparison of attractor properties for ERA5-Land (red-high, blue-low): A. Embedding dimension B. Hurst exponent and C. Max Lyapunov exponents.

Table 2
Attractor properties average differences by region and elevation.

Region	Elevation	Delay (days)	Dimension (-)	Hurst (-)	Lyapunov (-)
Andean	Low	4.25	7.7	0.15	0.24
	Medium	4.77	5.5	0.17	0.12
	High	5.84	5.0	0.17	0.09
Caribbean	Low	3.2	9.6	0.26	0.23
	Medium	4.5	7.7	0.22	0.18
Pacific	Low	6.0	7.0	0.15	0.05
	Medium	3.6	6.9	0.12	0.11
	High	4.2	4.5	0.17	0.10

and Caribbean (0.229) regions, while the smaller differences were found for the higher elevations of the Andean region. Hurst exponent values are consistent with those obtained from the local stations, exhibiting persistent behaviour in both cases, with lower differences in the Pacific and higher in the Caribbean.

4.3. Spatiotemporal object characteristics

Once the spatiotemporal objects were constructed, the number of objects was counted before any filtering was applied. A greater number of objects were identified in ERA5-Land ("El Niño": 444,294 - "La Niña": 490,755) compared to the local stations IDW interpolation ("El Niño": 147,260 - "La Niña": 187,980). From these totals, the 1000 largest objects from each dataset (Local stations and ERA5-Land) and ENSO period (("El Niño" and "La Niña") were then selected for further analysis.

A comparison revealed that spatiotemporal objects derived from local stations had higher precipitation volumes than those from the reanalysis dataset. For the 1000 largest events, the average absolute difference was 429 Hm^3 during "El Niño" and 460 Hm^3 during "La Niña". This gap was even larger when comparing the single most intense events, with maximum differences of 4630 Hm^3 and 3146 Hm^3, respectively. Additionally, the average object duration was longer for the "El Niño" period (4.0 days) than for "La Niña" (2.2 days).

The classification of the object's initial and final centroid locations within the regions and elevations is summarized in Table 3 for the local station's interpolation, and in Table 4 for ERA5-Land.

Among the objects from the local interpolation, 59 % of objects originating in the Andean region remain there, while 31 % end up in the Caribbean and the remaining 10 % in the Pacific. A similar distribution pattern was found in the Caribbean, where most of the objects stay in the region (78.2 %). However, in the Pacific, most of the objects end up in the Andean region (65 %). Compared to local interpolation, ERA5-Land shows more objects staying within their region of origin: 79 % in the Andean and 78 % in the Caribbean.

Additionally, the most significant elevation-related differences were observed in the Andean and the Caribbean. In the Andean, objects starting at any elevation and ending at mid-elevations showed discrepancies of 8–20 %. In the Caribbean, objects starting at any elevation and ending at low elevations exhibited larger errors, ranging from 13 % to 38 %. Furthermore, in ERA5-Land, no objects originated in the Pacific Low region. For the local station interpolation, objects originated in the Pacific Low region only during "El

Table 3
Distribution of objects start-finish distribution within the regions for Local IDW interpolations. Green indicates a higher number of objects finishing in a given region.

			Centroid Finish								
	Region/Elevation		Andean			Caribbean			Pacific		
			Low	Medium	High	Low	Medium	High	Low	Medium	High
Centroid Start	Andean	Low	26%	22%	8%	20%	13%	2%	3%	5%	2%
		Medium	13%	31%	16%	14%	13%	3%	1%	6%	3%
		High	9%	32%	19%	13%	12%	3%	1%	8%	3%
	Caribbean	Low	10%	15%	7%	44%	21%	2%	0%	1%	1%
		Medium	7%	19%	9%	26%	30%	7%	0%	2%	1%
		High	1%	21%	6%	15%	32%	19%	1%	2%	1%
	Pacific	Low	44%	15%	7%	0%	0%	7%	22%	4%	0%
		Medium	18%	35%	15%	10%	8%	0%	5%	7%	2%
		High	2%	37%	22%	5%	10%	0%	0%	10%	16%

Table 4
Distribution of objects start-finish distribution within the regions for ERA5-Land. Green indicates a higher number of objects finishing in a given region.

			Centroid Finish								
	Region/Elevation		Andean			Caribbean			Pacific		
			Low	Medium	High	Low	Medium	High	Low	Medium	High
Centroid Start	Andean	Low	12%	42%	13%	23%	7%	0%	0%	1%	1%
		Medium	4%	50%	24%	9%	7%	0%	0%	3%	3%
		High	4%	41%	38%	4%	3%	1%	0%	5%	4%
	Caribbean	Low	2%	11%	2%	56%	24%	2%	0%	1%	0%
		Medium	4%	17%	5%	46%	22%	1%	0%	2%	3%
		High	0%	8%	0%	54%	23%	8%	0%	0%	8%
	Pacific	Low	---	---	---	---	---	---	---	---	---
		Medium	0%	36%	25%	13%	7%	0%	0%	12%	6%
		High	3%	40%	33%	7%	4%	0%	0%	7%	7%

Niño" period.

Furthermore, average differences in spatiotemporal object properties (volume and total duration) were discretized by region and elevation. These results are summarized in Fig. 12, which presents the absolute volume differences between local station interpolation and ERA5-Land data for objects moving within the basin.

During periods of "La Niña", volume differences were concentrated in objects that started or ended at medium elevations, with the largest errors occurring in those originating from the Andean region. The largest volume error observed was −2619 Hm^3 for objects that began in the Andean region at medium elevations and ended in the Pacific region at low elevations. In contrast, during "El Niño" periods, aggregated volume differences were more uniform across elevations. However, the largest regional error occurred in objects terminating in the Andean region. The largest volume error in this case was 3919 Hm^3 , corresponding to objects that originated in the Pacific region at high elevations and ended in the Andean region at similarly high elevations.

Regarding the total duration of events, during "El Niño" periods, the largest average errors were observed in objects that originated at high elevations and ended at low elevations (16.5 days). Regionally, the largest errors occurred in objects that started in the Caribbean and ended in the Pacific (17.2 days). In "La Niña" periods, the largest average errors were found in objects originating from medium elevations and terminating at low elevations (14.2 days). In terms of regions, the greatest errors were seen in objects that started in the Andean region and ended in the Pacific (12.2 days).

4.4. Spatiotemporal non-linear dynamics assessment

After computing the differences between local stations and ERA5-Land using multiple approaches, multi-metric error plots were generated for various metrics. Based on the results obtained and the regional variations they revealed, special focus was given to the following metrics combination: RMSE (from the standard metrics), the Maximum Lyapunov exponent (from attractor properties), and volume differences (from spatiotemporal object characteristics), as presented in Fig. 13. Here, regions are shown and connected based on their spatiotemporal relationships, with links indicating the direction and magnitude of the largest volumetric differences in objects originating from each region.

The Andean region at low elevations exhibited the largest error among all regions in terms of λ_{max} (0.24), with a RMSE (20.8 mm/d). Although the Pacific region at medium elevations showed a similar RMSE (21.1), while it had a lower λ_{max} error (0.11), highlighting the difference in precipitation dynamic across the basin. Furthermore, the objects with the largest volumetric errors (720 Hm^3) originating from Andean low elevations were directed toward the Pacific medium elevation zone during both ENSO phases. Objects moving in the opposite direction showed larger errors of volume (863 Hm^3) only in the Pacific region during the "La Niña" period.

Conversely, high elevation areas in both the Andean and Pacific regions exhibited similar λ_{max} errors (0.092 and 0.098). However, the Andean region demonstrated a lower RMSE (10.4 mm/d). Moreover, the high elevation Andean region proved to be a key finishing area for objects with larger volume errors originating from other regions during "El Niño". It was also during these periods that objects with the largest volumetric errors across regions were observed, starting at high elevations in the Pacific (3919 Hm^3).

This plot also highlights that the Caribbean region, across all elevations, shows similar values for both RMSE (15.7–16.7 mm/d) and λ_{max} differences (0.18–0.23), while the largest volumetric errors for objects originating in this region occur when moving toward higher elevations.

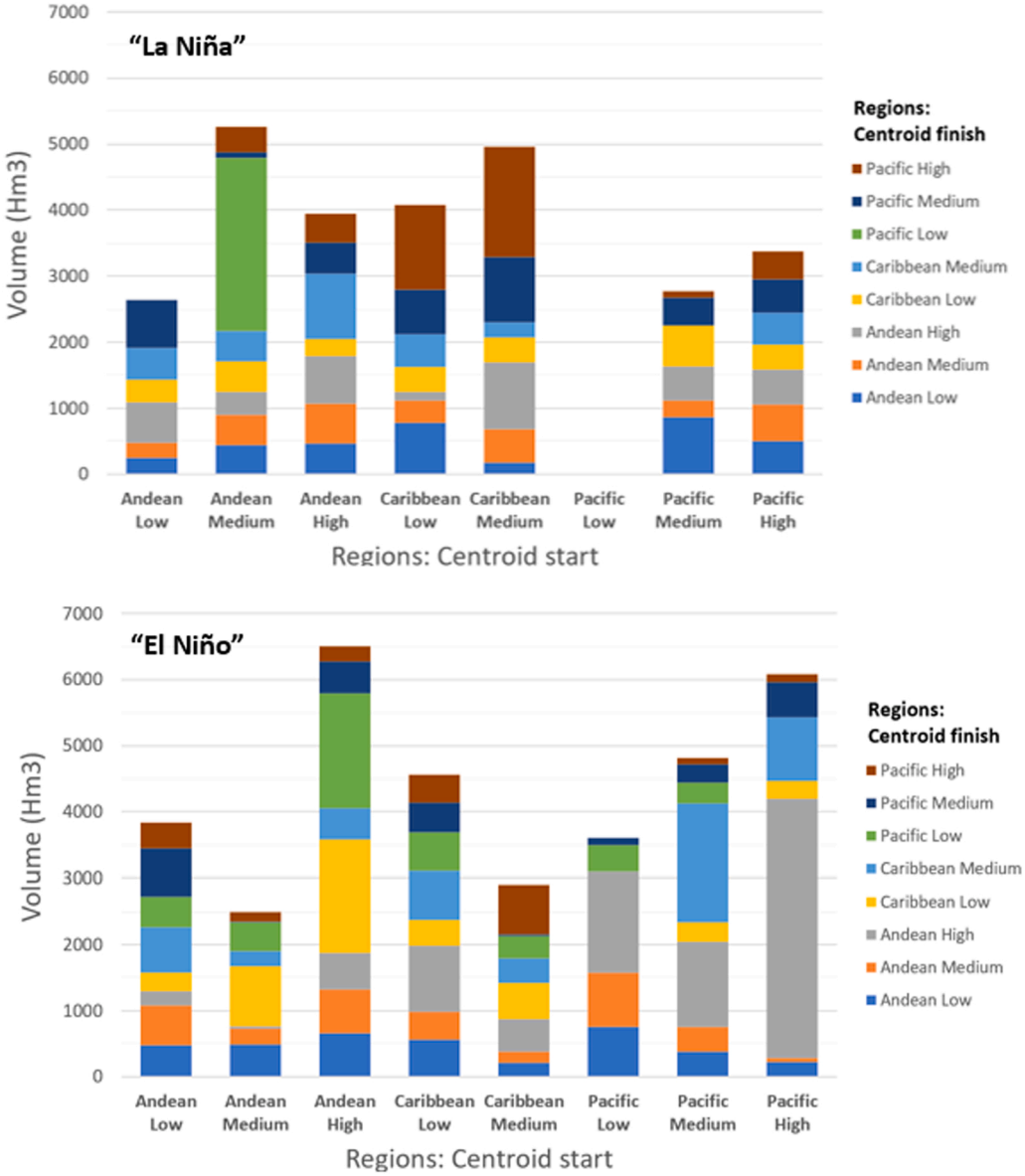

Fig. 12. Absolute volume differences of spatiotemporal objects between Local stations interpolation and ERA5-Land for "El Niño" (down) and "La Niña" (up) periods, categorized by their start and finish centroid regions.

5. Discussion

5.1. Basin overview

SNLDA general results across the basin showed that ERA5-Land's ability to represent precipitation is quite limited. In most cases, it overestimates overall precipitation volumes, a pattern consistent with other studies (e.g., Vega-Durán et al., 2021, in the Sinu River basin). While lower RMSE values were observed in the dry seasons (DJF, JJA), errors were higher in the wet seasons (MAM, SON), highlighting challenges for ERA5-Land in representing ITCZ-driven precipitation, as also noted by Lavers et al. (2022).

From an attractor perspective, error differences in embedding dimensions and Maximum Lyapunov exponents decreased with elevation, suggesting that nonlinear dynamics are relatively similar between ERA5-Land and local observations at higher altitudes. The Hurst exponent indicated anti-persistent behavior ($H > 0.5$), while the Maximum Lyapunov exponent showed deterministic chaos ($\lambda_{max} > 0$), reflecting the complexity of the local climate system and raising questions about ERA5-Land's assumptions during data assimilation. These findings align with Salas et al. (2020), who observed general synchronization between ENSO and rainfall anomalies in Colombia's Pacific, southern Caribbean, and western Andean regions.

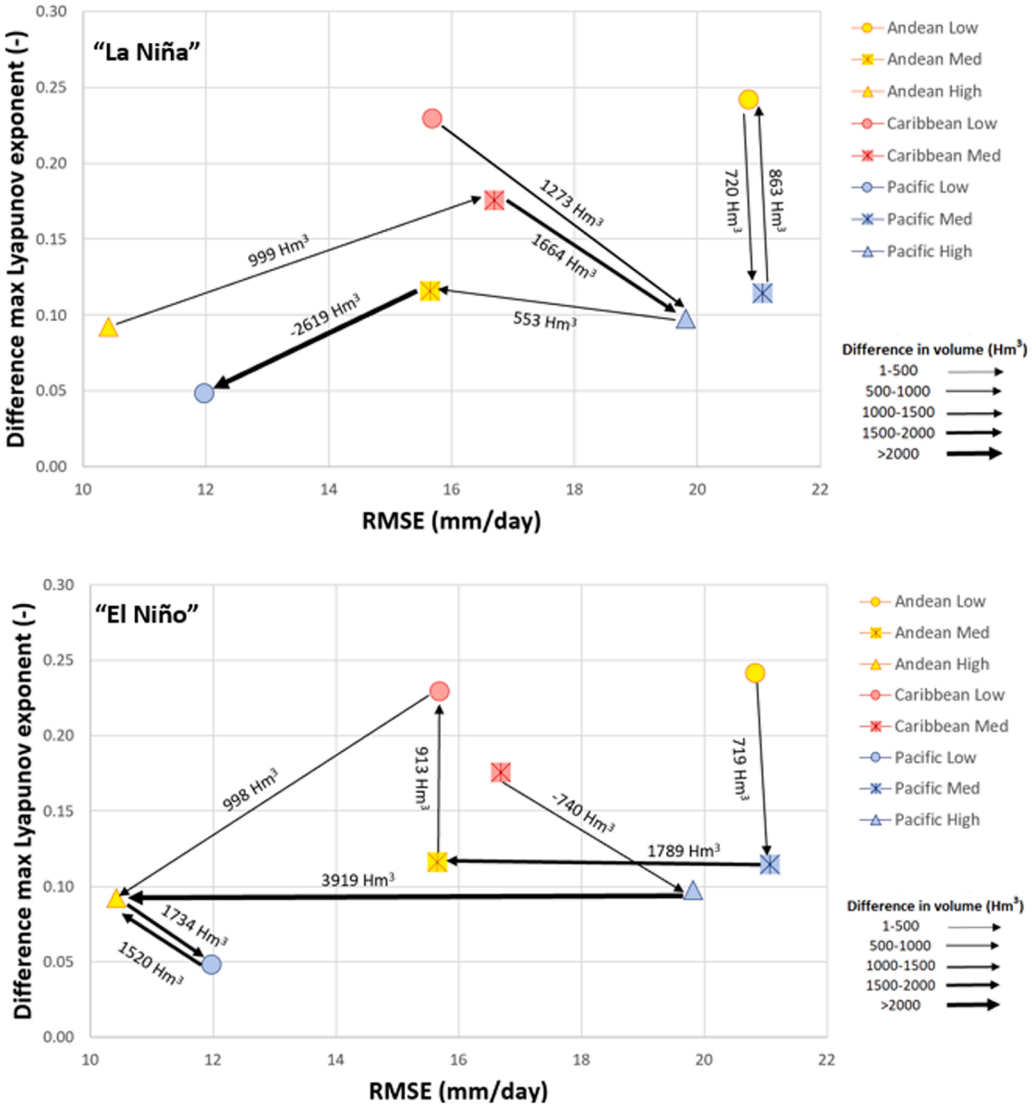

Fig. 13. Comparison of RMSE and Lyapunov exponent errors for local station interpolation vs. ERA5-Land data. Points represent regions. Arrows indicate the direction and magnitude of the largest volume differences in spatiotemporal objects during "El Niño" (down) and "La Niña" (up) periods.

Similarly, spatiotemporal object analysis revealed differences in the distribution and displacement of precipitation events, with ERA5-Land producing more objects of lower precipitation magnitude, consistent with its overall volume overestimation (see Tables 3 and 4, Figs. 12 and 13; Gomis-Cebolla et al., 2023; Xu et al., 2022).

These basin-level findings demonstrate that ERA5-Land errors manifest across all three evaluated dimensions (statistical, nonlinear, and spatiotemporal) highlighting both the limitations of the dataset and the specific patterns of misrepresentation. The errors are not isolated but interconnected: Overestimation of precipitation volumes (statistical) corresponds with an excess of spatiotemporal objects (spatiotemporal) and, in some regions, misrepresentation of nonlinear dynamics (nonlinear). By revealing these multidimensional interconnections, the SNLDA framework directly addresses the study's research question regarding the extent to which ERA5-Land reproduces precipitation characteristics across the Magdalena basin and how errors in these dimensions are related.

This connection frames the subsequent Regional Analysis, which details how these errors vary across the Andean, Caribbean, and Pacific regions, providing a foundation for understanding the spatial patterns and error sources of ERA5-Land misrepresentations.

5.2. Regional analysis

5.2.1. Andean region

On average, the volumes presented in ERA5-Land for this region duplicate those of local stations at an annual scale; this difference is also reflected in the magnitude of RMSE results at the daily level. Furthermore, RMSE values reduce with elevation, with the highest

error for elevations below 150 m. The same pattern was also observed for maximum precipitation. Similarly, this characteristic is also observed with λ_{max}, where the highest error (0.24) is also located in this range of elevations, which may indicate that non-linear processes in this region and elevation may affect ERA5-Land accuracy. Additionally, regarding the displacement of spatiotemporal objects, it was found that objects starting or ending in this region exhibited the largest differences in their distribution. Furthermore, during both ENSO phases, this region was associated with the largest volume errors across all regions, as shown in Fig. 12 and summarized in Fig. 13.

In this sense, other research indicates that while ERA5-Land's resolution is an improvement over ERA5 (which includes a thermodynamic orographic adjustment of temperature), it is still not sufficient to reproduce local observations in complex orographies (Khadka et al., 2022; Xu et al., 2022; Gomis-Cebolla et al., 2023;), which may explain the errors in the Andean region. While these results differ from those obtained by Valencia et al. (2023) using standard metrics and categorical indices, where ERA5-Land overestimation was found at high elevations, the error is also associated with the coarse spatial resolution.

Taking all of this into consideration, the integrated SNLDA results suggest that ERA5-Land struggles to accurately represent precipitation processes in the Andean region due to its location between two mountain chains. This leads to the generation of more rainfall objects than there should be, increasing the overall precipitation volume in the region. Although ERA5-Land improves upon ERA5 through higher spatial resolution and thermodynamic orographic adjustments, previous studies indicate that this resolution remains insufficient to reproduce local precipitation in complex topography (Khadka et al., 2022; Xu et al., 2022; Gomis-Cebolla et al., 2023). This limitation provides a plausible explanation for the errors identified here. While these results differ from those of Valencia et al. (2023), who reported overestimation at high elevations using standard metrics, both studies associate the errors with coarse spatial resolution, underscoring the added value of the SNLDA framework in refining error attribution.

5.2.2. Caribbean region

On average, the error in the embedding dimension (8.6), Hurst exponent (0.24), and Maximum Lyapunov exponent (0.20) were the highest in the entire basin. However, the annual bias factor and the RMSE were lower than in other regions. It was also found that the longest dry spell (97 days) and maximum precipitation (40 mm/d) errors in the region were the highest in the basin. This could imply that, while systematic errors exist in this region, errors in extreme events may be affected by non-linear processes in the climate system that are not fully represented in ERA5-Land. This missing process was also observed when analyzing the spatiotemporal objects. For example, the largest error in object durations is found in this region, particularly during "El Niño" periods.

These results support the findings of Vega-Durán et al. (2021) and Estrada and Arregocés (2024). Both studies, which focused on Colombian Caribbean basins using ERA5-Land, identified a systematic underestimation of precipitation in the region using standard metrics (NRMSE, BIAS, NSE). Estrada and Arregocés (2024) further suggested that adjustments to ERA5-Land associated models' parameterization are needed, especially for drought analysis. This aligns with our findings, which also show ERA5-Land's limitations in representing non-linear processes within the climate system. Furthermore, prior research indicates that gridded precipitation datasets frequently misrepresent precipitation in this region, as shown by evaluations using correlation and bias metrics (Valencia et al., 2023; Morales-Acuña et al., 2021). These studies also highlight additional error sources, including extended dry periods and the scarcity of IDEAM stations. These issues are further supported by our results for this region, as presented in Fig. 9.

5.2.3. Pacific region

In this region, individual metrics suggest conflicting performance. However, their integration within the SNLDA framework reveals a consistent and interpretable error structure in ERA5-Land precipitation. Opposite to the Caribbean region, attractor properties' errors showed the best results in the basin ($\lambda_{max} < 0.11$ and $H < 0.12$), while standard metrics presented the poorest results within the basin (see Fig. 13), with a precipitation volume 3.4 times higher in ERA5-Land, and RMSE values between 16 mm/d and 23 mm/d for different seasons. This error is also observed in maximum precipitation, which is higher for ERA5-Land than for local stations, at 17 and 13 mm/d for medium and high elevations, respectively. From the spatiotemporal object perspective, beyond the significant interaction with the Andean, where the highest error in volume was found, it was identified that 65 % of the objects that started in the Pacific had as a destination the Andean region or beyond in the east. In contrast, few objects have as their destination the Caribbean (21 %) or stay in the Pacific (14 %).

These findings suggest the influence of the orography and ocean-land local climate dynamics in the region over the errors. In this sense, Sagero et al., (2024) observed ERA5 well performance in a tropical region over the South Pacific, while also highlighting the bias under high topography variations. Moreover, ENSO findings in this region are consistent with those presented by Bedoya-Soto et al. (2019) when describing the characteristics of the 2010–2011 La Niña and its rainfall contribution in the basin. Furthermore, given the identified volume mismatch, Romero-Hernández et al., (2024) underscore the need for bias correction in reanalysis products over this region. They also highlight the variation of statistical metrics (e.g., KGE) during "El Niño" and "La Niña" periods, which aligns with our study of spatiotemporal patterns for both periods.

The SNLDA framework effectively integrated three perspectives into a single, comprehensive methodology. This approach enabled the quantification of how well ERA5-Land reproduces statistical, non-linear dynamics, and spatiotemporal characteristics, as well as their interconnections, across the Magdalena River basin and its subregions. The novelty and strengths of this methodology include: (a) A multi-metric error plot to facilitate analysis. (b) The comparison of standard errors and attractor properties to identify representations of complex non-linear dynamics. (c) Representation of spatiotemporal objects displacements for physical processes understanding. (d) High suitability for complex tropical regions with varied elevations. However, the following limitations were identified: (a) Reliance on gridded input data (b) High sensitivity to noise in time-series data (c) An indirect approach to nonlinearity; while SNLDA does not explicitly model nonlinear patterns, the resulting attractor properties capture the system's underlying complexity.

5.3. Uncertainties in rain-gauge comparisons

Instrumental, procedural, and observational heterogeneities in the rain-gauge network inevitably influence the comparison with ERA5-Land and must be considered when interpreting the results. Although rain gauges are commonly treated as "ground truth", they are affected by both systematic and human-induced errors related to exposure, maintenance, reading practices, and sub-daily turbulence (Vega-Durán et al., 2021). The predominance of manual stations (91 %) and daily measurements (84 %) implies limited representation of short-duration and high-intensity events, which are better resolved by ERA5-Land at sub-daily scale (Dai et al., 2024). In addition, The heavy reliance on precipitation-only measurements (78 %) limits the diagnosis of ERA5-Land's rainfall physical drivers (Xie et al., 2022). This is most acute in the Caribbean (88 %), where the lack of auxiliary variables hides the causes of the model's failure to capture non-linear dynamics and local complexity.

Procedural aspects, especially data pre-processing, also play a role. The infilling of missing values and the use of IDW interpolation unavoidably modify rainfall statistics, potentially smoothing extremes and altering spatial gradients (Ensor and Robeson, 2008). However, the high correlation (0.92) and low RMSE (<5 mm/d) between observed and interpolated values indicate that the dataset preserves the dominant precipitation signals and is suitable for spatiotemporal analyses. Nonetheless, as noted by Xie et al. (2022), gauge bias can affect evaluations of reanalysis products by influencing estimates of magnitude, frequency, and extremes. While no bias correction was applied here, this uncertainty should be considered when interpreting both the statistical and non-linear discrepancies between ERA5-Land and observations.

Observational heterogeneity across regions further modulates the comparison. The Caribbean region exhibits the lowest station density (721.9 km^2/station) and the largest average inter-station distance (13.4 km), increasing representativeness errors relative to the Andean and Pacific regions. This is due to the sparse gauge network, which may not fully capture precipitation patterns, as also noted by Zhang et al. (2025) in their study of ERA5-Land in Southwest China. In contrast, the Pacific region, despite having the highest station density, shows the largest ERA5-Land overestimation, suggesting that model limitations related to regional conditions (e.g. orography and moisture transport) dominate over observational uncertainty (Sagero et al., 2024).

Overall, the combined evidence suggests that instrumental and procedural uncertainties in the gauge network can amplify or reduce differences with ERA5-Land, especially at fine temporal scales and for extreme events. However, the consistency of the results across multiple metrics (including standard, non-linear, and spatiotemporal analyses) indicates that many observed discrepancies arise from structural limitations of the reanalysis in representing complex precipitation dynamics. In this context, the SNLDA framework provides a robust analytical approach not only for evaluation but also for future bias-correction and uncertainty-aware comparisons between observational networks and reanalysis products.

6. Conclusions

This study demonstrates the strengths and limitations of the ERA5-Land dataset in representing daily precipitation across the complex climatic and topographic regions of the Magdalena River Basin. To evaluate this, we utilized the SNLDA comprehensive framework, incorporating traditional metrics, spatiotemporal precipitation objects, and nonlinear dynamics. This approach reveals that while ERA5-Land captures some broad statistical patterns, it largely struggles to reproduce the critical nonlinear dynamical and spatiotemporal characteristics of precipitation. The assessment found that traditional metrics alone overlook significant errors in the dataset's representation of complex local processes, highlighting the dataset's limitations in this tropical region.

The sources of error and their interconnections vary significantly by region. In the Andean region, ERA5-Land systematically overestimates precipitation, an error linked to its failure to resolve nonlinear and orographic processes, resulting in discrepancies during ENSO phases. In the Caribbean, the dataset showed high errors in nonlinear dynamics and extreme characteristics (like dry spells), even where traditional statistical errors were low. In the Pacific, strong biases in precipitation volume were linked to the misrepresentation of spatiotemporal objects displacement, orography, and ENSO influences, rather than fundamental errors in the local dynamics. Across the basin, these interconnected errors translate into a misrepresentation of precipitation events: ERA5-Land often generates more frequent, shorter-duration rain events and poorly captures extreme wet and dry spells compared to local observations

The SNLDA framework used in this study provides critical insights beyond traditional validation, revealing that ERA5-Land has difficulty synchronizing with local climate dynamics. This highlights the importance of integrating nonlinear and spatiotemporal analysis into validation, especially for hydrometeorological modeling in complex tropical areas. While the specific results are regional, the framework itself is transferable to any basin. Our findings underscore the need to refine reanalysis data assimilation techniques and apply bias correction to better account for local climatic complexities, ultimately improving water resource management and disaster preparedness in these sensitive environments.

7. FAIR data statement

This research follows the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles. All analyses are based on publicly available open datasets (DHIME and ERA5-LAND), which are fully documented to support transparency, reproducibility, and reuse. Methodological details are provided to ensure discoverability and interoperability in line with community standards. The scripts used for data processing and analysis are openly available on GitHub at https://github.com/sduarte09/SNLDA.git

CRediT authorship contribution statement

Remko Uijlenhoet: Writing – review & editing, Writing – original draft, Supervision. **Duarte Freddy Santiago:** Writing – original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Gerald Corzo:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology. **Dimitri Solomatine:** Supervision, Methodology.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Santiago Duarte reports financial support was provided by Colombian Foundation for the Future (2019-05.932). Santiago Duarte reports financial support was provided by Colombian Institute of Educational Credit and Technical Studies Abroad (5322338). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared the data links and code respoitory within the manuscript

References

Abarbanel, H.D.I., Kennel, M.B., 1993. Local false nearest neighbors and dynamical dimensions from observed chaotic data. Phys. Rev. E 47 (5), 3057–3068. https://doi.org/10.1103/PhysRevE.47.3057.

Agarwal, A., Marwan, N., Rathinasamy, M., Merz, B., Kurths, J., 2017. Multi-scale event synchronization analysis for unravelling climate processes: a wavelet-based approach. Nonlinear Process. Geophys. 24 (4), 599–611. https://doi.org/10.5194/npg-24-599-2017.

Albergel, C., Dutra, E., Munier, S., Calvet, J.-C., Munoz-Sabater, J., de Rosnay, P., Balsamo, G., 2018. ERA-5 and ERA-Interim driven ISBA land surface model simulations: which one performs better? Hydrol. Earth Syst. Sci. 22 (6), 3515–3532. https://doi.org/10.5194/hess-22-3515-2018.

Álvarez-Villa, O.D., Vélez, J.I., Poveda, G., 2011. Improved long-term mean annual rainfall fields for Colombia. Int. J. Climatol. 31 (14), 2194–2212. https://doi.org/10.1002/joc.2232.

Avila, H., Gutierrez, R., Otero, L., Amaris, G., 2019. Navigability of the Magdalena River: opportunities and challenges. BASED Sci. Evid. 2645. https://doi.org/10.3850/38WC092019-1113.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., Vitart, F., 2015. ERA-interim/land: a global land surface reanalysis data set. Hydrol. Earth Syst. Sci. 19 (1), 389–407. https://doi.org/10.5194/hess-19-389-2015.

Beck, H.E., Pan, M., Roy, T., Weedon, G.P., Pappenberger, F., van Dijk, Huffman, G.J., Adler, R.F., Wood, E.F., 2019. Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. Hydrology and Earth System Sciences 23, 207–224. https://doi.org/10.5194/hess-23-207-2019.

Bedoya-Soto, J.M., Poveda, G., Trenberth, K.E., Vélez-Upegui, J.J., 2019. Interannual hydroclimatic variability and the 2009–2011 extreme ENSO phases in Colombia: from Andean glaciers to Caribbean lowlands. Theor. Appl. Climatol. 135 (3), 1531–1544. https://doi.org/10.1007/s00704-018-2452-2.

Bodjrènou, R., Sintondji, L.O., Comandan, F., 2025. Revealing the spatiotemporal precipitation patterns of ECMWF fifth-generation reanalyses since the mid-20th century in West-Africa. ISSN 2666-7657 Environ. Adv. 20, 100636. https://doi.org/10.1016/j.envadv.2025.100636.

Bojacá, A.J., Patiño, D.M., Pérez, G.M., 2025. Assessing climate reanalysis products in tropical lowlands: a validation study from Colombian oil palm regions. Rev. Bras. De. Meteorol. https://doi.org/10.1590/0102-778640230053.

Bolaños Chavarría, S., Werner, M., Salazar, J.F., Betancur Vargas, T., 2022. Benchmarking global hydrological and land surface models against GRACE in a medium-sized tropical basin. Hydrol. Earth Syst. Sci. 26, 4323–4344. https://doi.org/10.5194/hess-26-4323-2022.

Casdagli, M., Eubank, S., Farmer, J.D., Gibson, J., 1991. State space reconstruction in the presence of noise. Physica D Nonlinear Phenomena 51 (1), 52–98. https://doi.org/10.1016/0167-2789(91)90222-U.

Centella-Artola, A., Bezanilla-Morlot, A., Taylor, M.A., Herrera, D.A., Martinez-Castro, D., Gouirand, I., Sierra-Lorenzo, M., Vichot-Llano, A., Stephenson, T., Fonseca, C., Campbell, J., Alpizar, M., 2020. Evaluation of sixteen gridded precipitation datasets over the caribbean region using gauge observations. Atmosphere 11 (12), 1334. https://doi.org/10.3390/atmos11121334.

Dai, Y., Abhishek, Li, L., Gong, Y., Wu, X., Sheng, B., Zhao, W., 2024. Variations in present and future hourly extreme rainfall: insights from high-resolution data and novel temporal disaggregation model. Water 16 (23), 3463. https://doi.org/10.3390/w16233463.

Derin, Y., Anagnostou, E., Berne, A., Borga, M., Boudevillain, B., Buytaert, W., Chang, C.-H., Chen, H., Delrieu, G., Hsu, Y.C., Lavado-Casimiro, W., Manz, B., Moges, S., Nikolopoulos, E.I., Sahlu, D., Salerno, F., Rodríguez-Sánchez, J.-P., Vergara, H.J., Yilmaz, K.K., 2019. Evaluation of GPM-era global satellite precipitation products over multiple complex Terrain regions. Article 24. Remote Sens. 11 (24). https://doi.org/10.3390/rs11242936.

Di Napoli, C., Romanello, M., Minor, K., Chambers, J., Dasgupta, S., Escobar, L.E., Hang, Y., Hänninen, R., Liu, Y., Lotto Batista, M., Lowe, R., Murray, K.A., Owfi, F., Rabbaniha, M., Shi, L., Sofiev, M., Tabatabaei, M., Robinson, E.J.Z., 2023. The role of global reanalyses in climate services for health: insights from the lancet countdown. Meteorol. Appl. 30 (2), e2122. https://doi.org/10.1002/met.2122.

Diaz, V., Corzo, G., Van Lanen, H., Solomatine, D., 2023. Three-dimensional clustering in the characterization of spatiotemporal drought dynamics. Clust. SIZE FILTER DROUGHT Indic. THRESHOLD Optim. 319–342. https://doi.org/10.1002/9781119639268.ch11.

Duarte, F., Corzo, G., Santos, G., Hernández, O., 2018. Chaotic statistical downscaling (CSD): application and comparison in the Bogotá River Basin. EPiC Ser. Eng. 3, 626–634. https://doi.org/10.29007/wkcx.

Durán-Quesada, A.M., Reboita, M., Gimeno, L., 2012. Precipitation in tropical America and the associated sources of moisture: a short review. Hydrol. Sci. J. 57 (4), 612–624. https://doi.org/10.1080/02626667.2012.673723.

Eckmann, J.-P., Ruelle, D., 1992. Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems. Physica D Nonlinear Phenomena 56 (2), 185–187. https://doi.org/10.1016/0167-2789(92)90023-G.

Elgamal, A., Reggiani, P., Jonoski, A., 2017. Impact analysis of satellite rainfall products on flow simulations in the Magdalena River Basin, Colombia. Journal of Hydrology: Regional Studies 9, 85–103. https://doi.org/10.1016/j.ejrh.2016.09.001.

Ensor, L.A., & Robeson, S.M. (2008). Statistical Characteristics of Daily Precipitation: Comparisons of Gridded and Point Datasets. https://doi.org/10.1175/2008JAMC1757.1.

Estrada, E., Arregocés, H.A., 2024. Drought trends in a coastal region with complex topography in northern colombia. Wit Trans. Ecol. Environ. 262, 609–618. https://doi.org/10.2495/sdp240501.

Fessehaye, M., Franke, J., Brönnimann, S., 2022. Evaluation of satellite-based (CHIRPS and GPM) and reanalysis (ERA5-Land) precipitation estimates over Eritrea. Meteorol. Z. 401–413. https://doi.org/10.1127/metz/2022/1111.

Gbode, I.E., Babalola, T.E., Diro, G.T., et al., 2023. Assessment of ERA5 and ERA-interim in reproducing mean and extreme climates over West Africa. Adv. Atmos. Sci. 40, 570–586. https://doi.org/10.1007/s00376-022-2161-8.
Giraldo-Osorio, J.D., Trujillo-Osorio, D.E., Baez-Villanueva, O.M., 2022. Analysis of ENSO-driven variability, and long-term changes, of extreme precipitation indices in Colombia, using the satellite rainfall estimates CHIRPS (Article). Water 14 (11), 11. https://doi.org/10.3390/w14111733.
Gomez-Rios, S., Zuluaga, M.D., Hoyos, C.D., 2023. Orographic controls over convection in an inter-andean valley in Northern South America. Mon. Weather Rev. 151 (1), 145–162. https://doi.org/10.1175/MWR-D-21-0231.1.
Gomis-Cebolla, J., Rattayova, V., Salazar-Galán, S., Francés, F., 2023. Evaluation of ERA5 and ERA5-Land reanalysis precipitation datasets over Spain (1951–2020). Atmos. Res. 284, 106606. https://doi.org/10.1016/j.atmosres.2023.106606.
Khadka, A., Wagnon, P., Brun, F., Shrestha, D., Lejeune, Y., & Arnaud, Y. (2022). Evaluation of ERA5-Land and HARv2 Reanalysis Data at High Elevation in the Upper Dudh Koshi Basin (Everest Region, Nepal). https://doi.org/10.1175/JAMC-D-21-0091.1.
KOUTSOYIANNIS, D., 2006. On the quest for chaotic attractors in hydrological processes. Hydrol. Sci. J. 51 (6), 1065–1091. https://doi.org/10.1623/hysj.51.6.1065.
Laverde-Barajas, M., Corzo Perez, G.A., Chishtie, F., Poortinga, A., Uijlenhoet, R., Solomatine, D.P., 2020. Decomposing satellite-based rainfall errors in flood estimation: hydrological responses using a spatiotemporal object-based verification method. J. Hydrol. 591, 125554. https://doi.org/10.1016/j.jhydrol.2020.125554.
Lavers, D.A., Simmons, A., Vamborg, F., Rodwell, M.J., 2022. An evaluation of ERA5 precipitation for climate monitoring. Q. J. R. Meteorol. Soc. 148 (748), 3152–3165. https://doi.org/10.1002/qj.4351.
Lawal, I.M., Bertram, D., White, C.J., Jagaba, A.H., Hassan, I., Shuaibu, A., 2021. Multi-criteria performance evaluation of gridded precipitation and temperature products in data-sparse regions. Atmosphere 12 (12), 1597. https://doi.org/10.3390/atmos12121597.
Li, G., Gao, C., Lu, B., Chen, H., 2021. Inter-annual variability of spring precipitation over the Indo-China Peninsula and its asymmetric relationship with El Niño-Southern Oscillation. Clim. Dyn. 56 (7), 2651–2665. https://doi.org/10.1007/s00382-020-05609-4.
Liu, M., Xu, X., Xu, C., Sun, A.Y., Wang, K., Scanlon, B.R., Zhang, L., 2017. A new drought index that considers the joint effects of climate and land surface change. Water Resour. Res. 53 (4), 3262–3278. https://doi.org/10.1002/2016WR020178.
López López, P., Immerzeel, W.W., Rodríguez Sandoval, E.A., Sterk, G., Schellekens, J., 2018. Spatial downscaling of satellite-based precipitation and its impact on discharge simulations in the Magdalena River Basin in Colombia. Front. Earth Sci. 6. https://doi.org/10.3389/feart.2018.00068.
López-Bermeo, C., Montoya, R.D., Caro-Lopera, F.J., Díaz-García, J.A., 2022. Validation of the accuracy of the CHIRPS precipitation dataset at representing climate variability in a tropical mountainous region of South America. Phys. Chem. Earth Parts A/B/C. 127, 103184. https://doi.org/10.1016/j.pce.2022.103184.
May, P.T., Trewin, B., Nairn, J.R., Ostendorf, B., Su, C.-H., Moise, A., 2022. Diurnal and seasonal variability of near-surface temperature and humidity in the maritime continent. J. Appl. Meteorol. Climatol. 61 (11), 1819–1834. https://doi.org/10.1175/JAMC-D-22-0032.1.
Morales-Acuña, E., Linero-Cueto, J.R., Canales, F.A., 2021. Assessment of precipitation variability and trends based on satellite estimations for a heterogeneous Colombian region. Hydrology 8 (3), 128. https://doi.org/10.3390/hydrology8030128.
Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. Earth Syst. Sci. Data 13 (9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021.
Norazizi, N.A.A., Deni, S.M., 2019. Comparison of artificial neural network (ANN) and other imputation methods in estimating missing rainfall data at Kuantan station. In: Berry, M.W., Yap, B.W., Mohamed, A., Köppen, M. (Eds.), Soft Computing in Data Science. Springer, pp. 298–306. https://doi.org/10.1007/978-981-15-0399-3_24.
Peña-Q, A.J., Bermudez-F, L.N., Ramírez-C, C., Riaño-H, N.M., 2015. Oceanic Niño index as a tool to determine the effect of weather on coffee plantation in Colombia. J. Exp. Agric. Int. 7 (6), 395–404. https://doi.org/10.9734/AJEA/2015/15876.
Planinić, J., Vuković, B., Radolić, V., 2004. Radon time variations and deterministic chaos. J. Environ. Radioact. 75 (1), 35–45. https://doi.org/10.1016/j.jenvrad.2003.10.007.
Romero-Hernández, C.M., Ávila Díaz, Quesada, B.R., Medeiros, F., Cerón, W.L., Guzman-Escalante, J., Ocampo-Marulanda, C., Rodrigues Torres, Zuluaga, C.F., 2024. Bias-corrected high-resolution precipitation datasets assessment over a tropical mountainous region in Colombia: A case study in Upper Cauca River Basin. Journal of South American Earth Sciences 140, 104898. https://doi.org/10.1016/j.jsames.2024.104898.
Sagero, P.O., Pratap, A., Magiri, R., et al., 2024. Validation of ERA5 rainfall data over the South Pacific Region: case study of Fiji Islands. Meteor. Atmos. Phys. 136, 28. https://doi.org/10.1007/s00703-024-01025-z.
Salas, H.D., Poveda, G., Mesa, Ó.J., Marwan, N., 2020. Generalized synchronization between ENSO and hydrological variables in Colombia: a recurrence quantification approach. Front. Appl. Math. Stat. 6, 3. https://doi.org/10.3389/fams.2020.00003.
Salazar, Á., Larrea-Alcázar, D.M., Bertin, A., Gouin, N., Pareja, A., Morales, L., Maillard, O., Ocampo-Melgar, D., Squeo, F.A., 2025. Spatio-temporal evaluation of MSWEP, CHIRPS and ERA5-Land reveals regional-specific responses across complex topography in Bolivia. Atmosphere 16 (11), 1281. https://doi.org/10.3390/atmos16111281.
Schutgens, N., Schutgens, N., Tsyro, S., Gryspeerdt, E., Gryspeerdt, E., Goto, D., Weigum, N., Schulz, M., Stier, P., 2017. On the spatio-temporal representativeness of observations. Atmos. Chem. Phys. 17 (16), 9761–9780. https://doi.org/10.5194/ACP-17-9761-2017.
Singh, H., Mohanty, M.P., 2025. Multi-scale assessment and entropy-MCDM framework for evaluating reanalysis precipitation datasets over Indian basins. Int. J. Appl. Earth Obs. Geoinf. 144, 104919. https://doi.org/10.1016/j.jag.2025.104919.
Sivakumar, B., Berndtsson, R., 2010. Nonlinear dynamics and chaos in hydrology. In: Advances in Data-Based Approaches for Hydrologic Modeling and Forecasting. World Scientific Publishing, pp. 411–461. https://doi.org/10.1142/9789814307987_0009.
Sorí, R., Nieto, R., Liberato, M.L.R., Gimeno, L., 2021. Oceanic versus terrestrial origin of El Niño Southern Oscillation–associated continental precipitation anomalies. Ann. N. Y. Acad. Sci. 1504 (1), 202–214. https://doi.org/10.1111/nyas.14665.
Steinkopf, J., Engelbrecht, F., 2022. Verification of ERA5 and ERA-Interim precipitation over Africa at intra-annual and interannual timescales. Atmos. Res 280, 106427. https://doi.org/10.1016/j.atmosres.2022.106427.
Sun, A.Y., Xia, Y., Caldwell, T.G., Hao, Z., 2018. Patterns of precipitation and soil moisture extremes in Texas, US: a complex network analysis. Adv. Water Resour. 112, 203–213. https://doi.org/10.1016/j.advwatres.2017.12.019.
Taylor, M.A., Enfield, D.B., Chen, A.A., 2002. Influence of the tropical Atlantic versus the tropical Pacific on Caribbean rainfall. J. Geophys. Res. Oceans 107 (C9), 10-1–10–14. https://doi.org/10.1029/2001JC001097.
Tompkins, A.M., Janisková, M., 2004. A cloud scheme for data assimilation: description and initial tests. Q. J. R. Meteorol. Soc. 130 (602), 2495–2517. https://doi.org/10.1256/qj.03.162.
Urrea, V., Ochoa, A., Mesa, O., 2019. Seasonality of Rainfall in Colombia. Water Resour. Res. 55 (5), 4149–4162. https://doi.org/10.1029/2018WR023316.
Valencia, S., Marín, D.E., Gómez, D., Hoyos, N., Salazar, J.F., Villegas, J.C., 2023. Spatio-temporal assessment of gridded precipitation products across topographic and climatic gradients in Colombia. Atmos. Res. 285, 106643. https://doi.org/10.1016/j.atmosres.2023.106643.
Vega-Durán, J., Escalante-Castro, B., Canales, F.A., Acuña, G.J., Kaźmierczak, B., 2021. Evaluation of areal monthly average precipitation estimates from MERRA2 and ERA5 reanalysis in a Colombian Caribbean Basin (Article). Atmosphere 12 (11), 11. https://doi.org/10.3390/atmos12111430.
Vega-Viviescas, C., Rodríguez-S, E.A., 2019. Evaluation of reanalysis data in the study of meteorological and hydrological droughts in the Magdalena-Cauca river basin, Colombia. Dyna 86 (211), 268–277. https://doi.org/10.15446/dyna.v86n211.80530.
Wang, G., Hope, P., Lim, E.-P., Hendon, H.H., Arblaster, J.M., 2021. An initialized attribution method for extreme events on subseasonal to seasonal time scales. J. Clim. 34 (4), 1453–1465. https://doi.org/10.1175/JCLI-D-19-1021.1.
Wang, N., Sun, F., Yao, Z., Wang, H., Liu, W., 2019. Evaluating satellite-based and reanalysis precipitation datasets with gauge-observed data and hydrological modeling in the Xihe River Basin, China. Atmos. Res., 104746 https://doi.org/10.1016/j.atmosres.2019.104746.
Wright, J.S., Sun, X., Konopka, P., Krüger, K., Legras, B., Molod, A.M., Tegtmeier, S., Zhang, G.J., Zhao, X., 2020. Differences in tropical high clouds among reanalyses: origins and radiative impacts. Atmos. Chem. Phys. 20 (14), 8989–9030. https://doi.org/10.5194/acp-20-8989-2020.

Xie, W., Yi, S., Leng, C., Xia, D., Li, M., Zhong, Z., Ye, J., 2022. The evaluation of IMERG and ERA5-Land daily precipitation over China with considering the influence of gauge data bias. Sci. Rep. 12, 8085. https://doi.org/10.1038/s41598-022-12307-0.

Xu, J., Ma, Z., Yan, S., Peng, J., 2022. Do ERA5 and ERA5-land precipitation estimates outperform satellite-based precipitation products? A comprehensive comparison between state-of-the-art model-based and satellite-based precipitation products over mainland China. J. Hydrol. 605, 127353. https://doi.org/10.1016/j.jhydrol.2021.127353.

Zhang, L., Yan, Z., Huang, K., et al., 2025. Evaluation and statistical bias correction of ERA5-Land meteorological variables for a humid river basin in Southwest China. Sci. Rep. 15, 41101. https://doi.org/10.1038/s41598-025-24942-4.

Zhao, K., Zhong, S., 2024. Evaluation and error analysis of multi-source precipitation datasets during summer over the Tibetan Plateau. Atmosphere 15 (2), 165. https://doi.org/10.3390/atmos15020165.