https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

PRECISi: A High-Resolution Daily Gridded Precipitation Dataset for Sicily (1951–2022) Derived from a Doubly Conditional Geostatistical Framework

Niloufar Beikahmadi[1], Antonio Francipane[1], Dario Treppiedi[1], Leonardo Valerio Noto[1]

[1] Department of Engineering, University of Palermo, Palermo, Italy

Correspondence to: Niloufar Beikahmadi (Niloufar.Beikahmadi@gmail.com)

Abstract Long-term, high-resolution precipitation datasets are indispensable for understanding hydroclimatic variability and supporting water-resource and climate-impact studies. Yet, the reconstruction of spatially consistent rainfall fields over extended historical periods remains a major challenge, particularly in data-sparse Mediterranean regions characterized by complex terrain and pronounced precipitation heterogeneity. This study introduces the first long-term, high-resolution gridded precipitation dataset for Sicily, delivering a continuous historical reconstruction since 1951 from sparse gauge observations through a novel two-phase, doubly conditional spatial interpolation framework. Calibration was performed on a dense modern network of automated rain gauges; a Transfer-Informed Modelling strategy, whereby all geostatistical parameters calibrated on this high-density contemporary period are transferred without modification to epochs of substantially lower network density, ensures temporal coherence across the full historical record. A merged archive of historical stations was assembled from two independent networks following rigorous homogeneity assessment of inter-network spatial coherence prior to reconstruction. Three principal methodological innovations distinguish the framework: (i) a double classification of daily events by intermittency and magnitude into distinct hydrometeorological regimes having unique spatial correlation structures; (ii) a comparison of regime-specific intensity-modelling paradigms based on geostatistics or regression; and (iii) a contrast of binary masking against probabilistic hurdle-weighting for occurrence-conditional intensity. Leave-one-out cross-validation demonstrated high rainfall detection capability across all intermittency classes and established the consistent superiority of the geostatistical framework over its regression-kriging counterpart across all hydrometeorological regimes. The superior optimal framework was consequently selected for operational historical reconstruction. Validated against an independent monthly climatological benchmark over the full study period, the bias-corrected primary dataset reproduces the reference climatology with high fidelity, affirming its fitness for long-term hydroclimatic analysis across topographically complex Mediterranean terrain (data is available at https://doi.org/10.5281/zenodo.19228150; BEIKAHMADI et al., 2026).

Keywords: Spatial Modelling; High-Resolution Dataset; Precipitation Reconstruction; Mediterranean; Hydrometeorological Regime Stratification; Transfer-Informed Modelling

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Open Access Earth System Science Data Discussions

1. Introduction

High-resolution observational gridded precipitation datasets constitute a foundational resource for distributed hydrological modelling, flood and drought early-warning systems, sustainable water resources management, calibration and validation of satellite-derived precipitation products, and the attribution of observed trends in hydroclimatic extremes (Beikahmadi and Rahimzadegan, 2023; Haylock et al., 2008; Hegerl et al., 2007; Huffman et al., 1997; Nosratpour et al., 2022; Refsgaard, 1997). Despite the increasing availability of spatially distributed precipitation products, including weather radar estimates, satellite-based datasets, and reanalysis products, the rain gauge remains the primary high-fidelity instrument for direct precipitation measurement; however, point observations from sparse and irregularly distributed networks cannot directly describe the continuous spatial fields required by environmental models or operational services (Jones and Hulme, 1996; Singh, 1992), especially in areas with complex topography, where precipitation exhibits strong spatial variability.

Transforming these discrete records into reliable, spatially coherent gridded products is not merely a technical exercise but a critical enabling step for a broad spectrum of scientific inquiry and societal decision-making. Compared to coarse temporal resolution (e.g., annual or monthly), the difficulty of this transformation increases substantially as the time scale decreases (e.g., daily or hourly). Indeed, precipitation can be highly variable in both space and time and exhibits an intermittent behavior characterized by frequent zero values and highly variable positive amounts, posing significant challenges for any interpolation approach (Goovaerts, 1997; Wilks, 2011). This is particularly true in regions characterized by strong climatic gradients, complex topography, and temporally evolving observation networks. In this case, a robust strategy is to condition well-established geostatistical models through event stratification, explicit occurrence–intensity separation, and variance-stabilizing transformations, allowing even simple estimators to be applied to more homogeneous subsets of data whose spatial structure can be reliably captured (Haberlandt, 2007; Li and Heap, 2011).

In recent years, the growing scientific demand for spatially continuous precipitation records has driven a proliferation of national and sub-national high-resolution gridded observational datasets. At global and continental scales, the CPC Global Unified Precipitation dataset (Chen et al., 2002), E-OBS (~10 km; Haylock et al., 2008; Cornes et al., 2018), the Greater Alpine Region High-Resolution Temperature dataset (GAR-HRT; Chimani et al., 2013), and the Alpine precipitation grids APGD and LAPrec (~5 km; Isotta et al., 2014, 2019) provide valuable multi-decadal climatological references. More recently, a rich portfolio of sub-continental products has emerged, including Iberia01 for the Iberian Peninsula (Herrera et al., 2019), SPREAD for Spain (Serrano-Notivoli et al., 2017), SiCLIMA at 500 m × 500 m for Aragon (Serrano-Notivoli et al., 2024), HYRAS for Germany (Krähenmann et al., 2018), HadUK-Grid for the United Kingdom (Hollis et al., 2019), seNorge2 for Norway (Lussana et al., 2018), SLOCLIM for Slovenia (Škrk et al., 2021), MeteoSerbia1km for Serbia (Sekulić et al., 2021), CLIMADAT-GRid at 1 km resolution for Greece (Varotsos et al., 2025), and a 250 m resolution product for the Trentino-South Tyrol Alpine region (Crespi et al., 2021).

Despite this remarkable expansion, Sicily, which is the largest island in the Mediterranean Sea, a recognized hotspot of anthropogenic climate change (Pecorino et al., 2024; Treppiedi et al., 2023, 2025, 2026), and a region acutely exposed to devastating flash floods and persistent droughts (Francipane et al., 2021, 2023), remains conspicuously absent. Moreover, the spatial resolution of broader European products is fundamentally too coarse to resolve the sharp precipitation gradients generated by Sicily's complex orography, where mean annual rainfall spans from below 360 mm to over 1,900 mm across distances of a few hundred kilometers (Beguería et al., 2016; Caracciolo et al., 2018; Di Piazza et al., 2011).

This gap is particularly critical given that the methodological evolution from early deterministic techniques, such as Thiessen polygons (Thiessen, 1911), Inverse Distance Weighting (Shepard, 1968), and thin-plate splines (Hutchinson and Gessler, 1994), to geostatistical and hybrid frameworks has progressively improved the accuracy of spatial precipitation estimation.

Within this context, two landmark meta-analyses provide critical guidance. These studies, respectively surveying over 70 comparative analyses across the environmental sciences (Li and Heap, 2011), and 48 operational gridded precipitation

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Open Access Earth System Science Data Discussions

products worldwide (Serrano-Notivoli and Tejedor, 2021), independently conclude that geostatistical methods, particularly Ordinary Kriging and its variants, are the most frequently employed and most consistently high-performing techniques in hydrometeorological mapping. These methods are valued for their theoretical optimality as Best Linear Unbiased Estimators (BLUE), their explicit modelling of spatial autocorrelation through the variogram, and their capacity to provide spatially explicit prediction uncertainty (Goovaerts, 1997; Matheron, 1971; Stein, 1999). Importantly, both studies highlight that reconstruction skill is driven primarily by the quality and appropriate stratification of input data, rather than by the specific choice of estimator.

In parallel, an increasingly influential paradigm extends geostatistical interpolation by incorporating physically meaningful covariates through Regression-Kriging (RK; Hengl, 2007; Odeh et al., 1995), where a deterministic trend, often estimated by Generalized Additive Models (GAM; Hastie and Tibshirani, 1986), captures non-linear relationships between precipitation and controlling factors such as topography, coastal proximity, and atmospheric dynamics, while spatially autocorrelated residuals are modelled via kriging (Aalto et al., 2013). Whether this added covariate-driven complexity provides systematic accuracy improvements over carefully conditioned geostatistical approaches, particularly when reanalysis-based atmospheric predictors are available only at coarser spatial resolution than the target grid, remains an open and practically important question. This question extends well beyond Sicily and has direct implications for the design of operational precipitation mapping systems in data-sparse and topographically complex regions worldwide.

Despite these methodological advances and the broad consensus on the effectiveness of geostatistical approaches, several important challenges remain in the reconstruction of daily precipitation fields. First, the spatial structure of precipitation is strongly regime-dependent: widespread frontal events tend to exhibit smooth spatial patterns and long correlation ranges, whereas convective storms are characterized by localized and highly variable rainfall fields (Bárdossy and Pegram, 2014; Pegram and Clothier, 2001). Consequently, treating all rainfall events as a single statistical population fitted by a global variogram model is a well-documented oversimplification (Li and Heap, 2011).

Additional uncertainties concern how rainfall occurrence and intensity should be combined within a unified interpolation framework (Allcroft and Glasbey, 2003; Haberlandt, 2007), how to maintain consistent reconstruction skill when rain gauge network density changes over time (Alexander et al., 2006), and which variance-stabilizing transformation is most suitable for daily rainfall data. These issues have received comparatively limited attention and remain open questions for operational precipitation mapping.

The present study addresses these challenges by developing a high-resolution daily precipitation reconstruction framework for Sicily, which represents an exceptional natural laboratory for precipitation reconstruction, combining pronounced climatic gradients, complex terrain, and a rain-gauge archive extending back more than seven decades. The central hypothesis is that a carefully conditioned geostatistical approach, based on a novel double stratification of daily rainfall events into physically meaningful hydrometeorological regimes and supported by appropriate variance stabilization and parameter-transfer strategies, can match or outperform a more complex covariate-based Regression-Kriging framework.

To test this hypothesis, four explicit objectives are pursued: (i) to develop and evaluate a two-phase occurrence-intensity interpolation framework; (ii) to compare the performance of Ordinary Kriging and Regression-Kriging under identical experimental conditions using leave-one-out cross-validation (LOOCV); (iii) to assess the feasibility of Transfer-Informed Modelling for the reconstruction of a continuous 72-year precipitation record (1951–2022), thereby addressing the common trade-off between record length and station density that affects long-term climatological reconstructions; and (iv) to produce the first long-term, high-resolution (2 km × 2 km) daily gridded precipitation dataset for Sicily (PRECISi; BEIKAHMADI et al., 2026) available at https://doi.org/10.5281/zenodo.19228150.

The remainder of the paper is organized as follows: Sect. 2 describes the study area and datasets, Sect. 3 presents the methodological framework, Sect. 4 reports the results and discusses the findings; and Sect. 5 states the principal conclusions. The overall workflow is illustrated in Fig. 1.

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

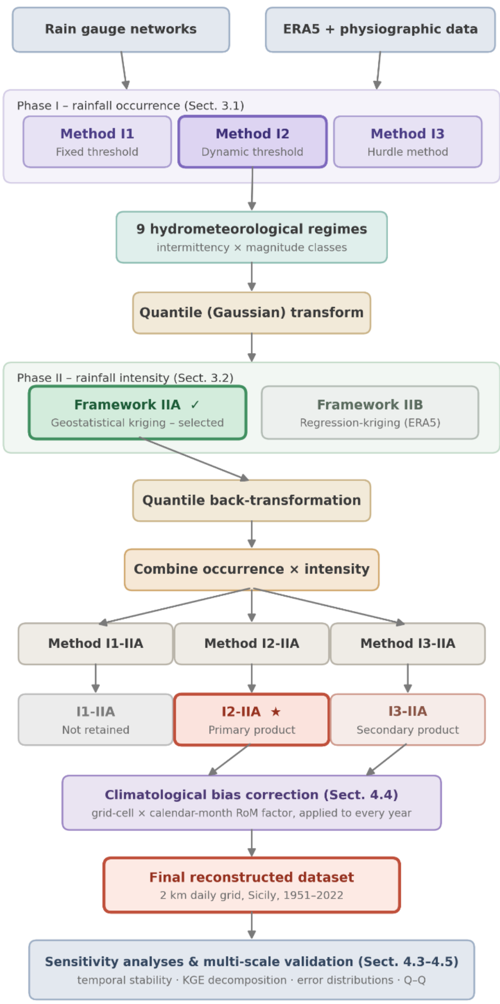

Figure 1: The methodology adopted for generating the PRECISi dataset.

2. Domain and Data

2.1 Study Area

Sicily is the largest island in the Mediterranean Sea (~25,711 km²), located between approximately 36°–38° N and 12°–15° E at the crossroads of Atlantic westerly influences from the north and arid subtropical air masses from North Africa (Fig. 2a), making it among the most climatically sensitive regions of the central Mediterranean (Beikahmadi et al., 2023; Giorgi, 2006). The island is predominantly hilly and mountainous, including the Northern Mountain Chain (reaching 1,979 m a.s.l. at Pizzo Carbonara), which intercepts northerly moist flows, and the isolated volcanic massif of Mount Etna (~3,357 m a.s.l.), which simultaneously enhances precipitation on its windward flanks while generating pronounced rain shadows on the leeward side, leading to sharp spatial gradients in precipitation over distances of only a few kilometers (Arnone et al., 2013; Sottile et al., 2022).

The island's climate broadly conforms to the Köppen Csa Mediterranean classification, characterized by hot, dry summers and mild, wet winters. However, this classification masks a pronounced spatial heterogeneity: mean annual precipitation spans from less than 360 mm in the semi-arid southeastern coastal areas to over 1,900 mm in the orographically exposed northeastern

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

highlands. This nearly fivefold gradient reflects the combined influence of topographic forcing, distance from maritime moisture sources, and the diversity of synoptic and convective precipitation-generating mechanisms acting across the region (Sottile et al., 2022).

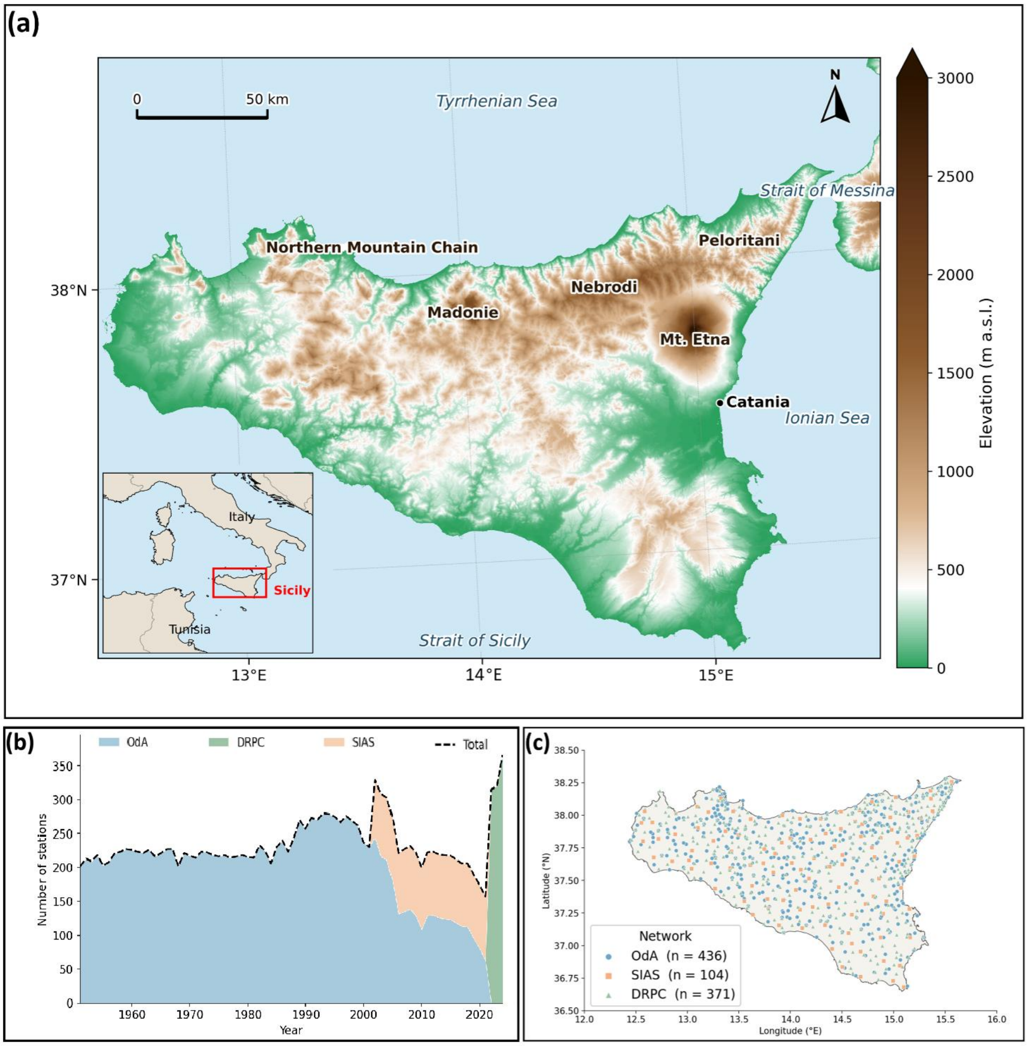

Figure 2: Area of study overlaid on the digital elevation model (a); number of active stations for the three precipitation networks over time (b); and the spatial distribution of the in situ network over the study region (c).

This meteorological diversity is not merely a descriptive feature of Sicily's climatology; it is the primary scientific motivation for the doubly conditional stratification framework at the core of this study. Precipitation over Sicily is generated by at least three physically distinct meteorological mechanisms, each producing rainfall fields with fundamentally different spatial structures (Sottile et al., 2022; Treppiedi et al., 2023; Trigo et al., 2002).

Large-scale frontal systems and Mediterranean cyclones produce widespread, spatially coherent precipitation fields of light-to-moderate intensity, characterized by long spatial correlation lengths and low intermittency. These events dominate wet-season totals and account for the bulk of the island's annual water budget (Trigo et al., 2002).

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Open Access Earth System Science Data Discussions

In some cases, embedded within or following synoptic disturbances, organized Mesoscale Convective Systems generate spatially extensive but high-intensity rainfall events, where elevated mean field intensities are still associated with relatively coherent spatial structures (Caccamo et al., 2017; Sottile et al., 2022).

In stark contrast, isolated convective cells produce short-duration, high-intensity downpours whose frequency has increased since the 1990s (Arnone et al., 2013; Treppiedi et al., 2021) and whose seasonal occurrence is shifting toward the summer (Treppiedi et al., 2026), in what could be described as a "tropicalization" of the regional rainfall regime. These events generate highly spatially fragmented and strongly intermittent rainfall fields, with correlation lengths of only a few kilometers and extreme local intensities occurring adjacent to dry areas.

Consequently, the proposed stratification is not an arbitrary data-partitioning strategy but a physically based response to a precipitation regime in which synoptic, mesoscale, and convective processes coexist, making single-model approaches statistically untenable.

2.2 Data Collection

Four datasets were employed in this study, including ground-based precipitation observations, dynamic atmospheric predictors, static physiographic variables, and a benchmark dataset. Their characteristics and roles within the modelling framework are described in the following subsections.

2.2.1 The Observation database

Three ground-based precipitation datasets were employed (Fig. 2 panels b-c), each serving a distinct functional role within the work (Table 1).

The primary dataset for model development and calibration was provided by the *Dipartimento Regionale della Protezione Civile* (DRPC–Sicilia), the regional Civil Protection Department. This modern, telemetered network consists of 372 automated tipping-bucket rain gauges (0.1 mm resolution) operating at sub-hourly time steps. For this study, records were aggregated to daily totals. Following quality control, 369 active stations were retained, corresponding to an average inter-station distance of approximately 8.3 km.

Two complementary historical datasets were used to reconstruct long-term daily precipitation over the period 1951–2022.

The first is the network of the *Servizio Informativo Agrometeorologico Siciliano* (SIAS), the regional agro-meteorological information service, comprising 107 tipping-bucket rain gauge stations distributed relatively homogeneously across the island (approximately one station per 250 km²). These gauges record precipitation at 10-minute resolution with a 0.2 mm tipping increment and were aggregated to daily totals for consistency. Data span January 2002 to December 2022.

The second dataset is the long-term daily rainfall archive of the *Autorità di Bacino* (AdB), successor to the former Italian National Hydrographic Service (*Servizio Idrografico Nazionale*). This archive extends back to 1916 and represents one of the most extensive daily historical rainfall records in southern Italy. For consistency and network stability, the analysis was restricted to the 1951–2022 period, to maximize the number of stations with continuous records in the post-World-War-II period, providing 72 years of data for historical reconstruction, yielding up to 441 stations with variable temporal coverage due to network evolution.

The three networks employ different native daily conventions (OdA: 09:00–09:00 UTC+1; SIAS: 10-min records; DRPC: 00:00–24:00). To place all observations on a single, physically consistent daily definition, the sub-daily SIAS and DRPC series were aggregated to the OdA 09:00–09:00 window prior to any joint analysis, and the delivered PRECISi product inherits this 09:00–09:00 daily accumulation. The 09:00–09:00 'rainfall day' is a widely used gauge convention (e.g. the UK/Ireland climatological day) that assigns the accumulation to the day on which most rain fell. Because the historical OdA archive is available only as fixed daily totals, its records could not be re-windowed; the homogeneity assessment below therefore verifies

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

that the aggregated SIAS series and the native OdA series are mutually consistent on this common definition and can be merged without a systematic offset.

Internal validation within the AdB network showed Pearson correlations ranging from 0.60 to 0.90 for station pairs within 0–15 km, decreasing smoothly with separation as expected for a spatially coherent precipitation field (Fig. 3, left). Fitting an exponential decay, $\rho(d) = n_0 + (1-n_0)\, e^{(-d/L)}$, yielded a decorrelation length of $L \approx 11$ km for the direct (same-day) series, confirming internal spatial homogeneity across the historical network. Cross-network agreement between AdB and SIAS over their common period (2002–2022) followed the same pattern (Fig. 3 right): short-range correlations of 0.6–0.9, with a decorrelation length $L \approx$ 12-14 km comparable to the internal AdB–AdB values, indicating that the two archives sample a consistent underlying field. In both cases, applying a 3-day moving average raised the correlations and lengthened the decay, as the smoothing absorbs sub-daily timing offsets between stations. This is a relevant check given the differing reporting conventions of the two networks. Only pairs within 40 km and with at least 70% joint data availability were retained; pair counts increased with separation, providing dense sampling across the full distance range.

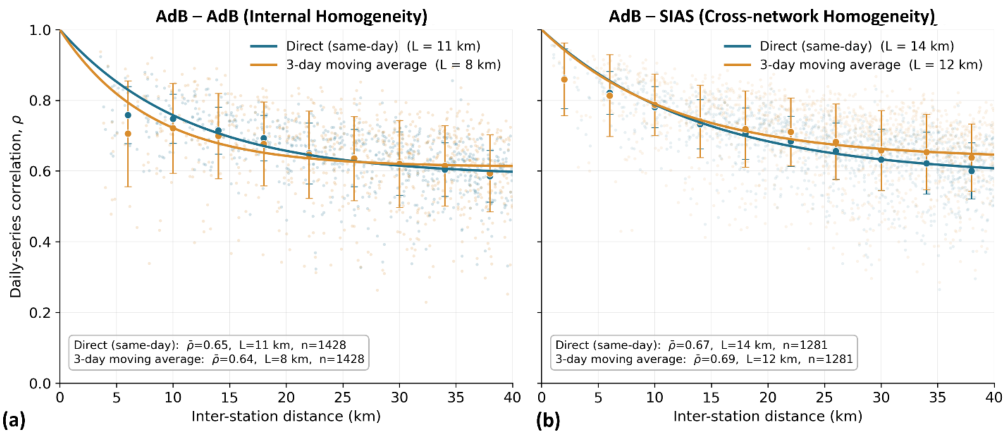

Figure 3: Spatial homogeneity of the merged record, assessed as the decay of daily series correlation with inter-station distance for station pairs within 40 km over the common period 2002–2022 (pairs with ≥70% joint availability). Left: AdB–AdB (internal homogeneity); right: AdB–SIAS (cross-network homogeneity).

Overall, these results indicate that the two networks are sufficiently homogeneous to be merged. The resulting combined AdB–SIAS archive (1951–2022), comprising up to 547 stations with time-varying availability, forms the historical input for the gridded precipitation reconstruction. Table 1 provides a consolidated overview of three in-situ precipitation datasets, including their provenance, temporal coverage, spatial characteristics, and their respective roles within the modelling framework.

2.2.2 Dynamic covariates

To inform the regression model (see Sect. 3), two dynamic atmospheric covariates were sourced from the ERA5 global reanalysis, namely the Convective Available Potential Energy (CAPE, $J\,kg^{-1}$) and Vertically Integrated Moisture Flux Divergence (VIMFD, $kg\,m^{-2}\,s^{-1}$) (Hersbach et al., 2020). These variables were selected because they represent the primary thermodynamic (atmospheric instability) and dynamic (moisture convergence) drivers of Mediterranean precipitation across both frontal and convective systems, respectively (Cipolla et al., 2020).

Both variables were obtained at hourly temporal resolution and 0.25° spatial resolution. Daily values were derived by retaining the maximum hourly CAPE and the minimum hourly VIMFD (i.e., the most negative value, corresponding to the strongest moisture convergence) within each day, under the assumption that daily rainfall totals are more strongly governed by peak instability and moisture supply than by mean diurnal conditions. The resulting daily covariate fields were then bilinearly resampled to the 2-km target grid to ensure spatial consistency with the modelling framework.

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

2.2.3 Static covariates

Static covariates represent time-invariant physiographic properties of the landscape that exert a persistent modulating influence on local precipitation patterns. Two static predictors were incorporated into the regression component of the modelling framework. Surface elevation was derived from the Shuttle Radar Topography Mission Version 3.0 (SRTMGL1) global dataset at approximately 30 m spatial resolution (Farr and Kobrick, 2000) and subsequently bilinearly resampled to the 2 km modelling grid. Elevation was included to represent the large-scale topographic controls on orographic precipitation enhancement and regional atmospheric circulation patterns.

Distance to the coastline was computed from a high-resolution OpenStreetMap vector polygon of the Sicilian coastline (Haklay and Weber, 2008). After projection to UTM Zone 33N, the Euclidean distance from each 2 km grid-cell centroid to the nearest coastline was calculated. This variable captures the influence of proximity to marine moisture sources and interior continental areas where local convection and orographic interactions dominate (Marra et al., 2022; Mazon and Pino, 2017).

2.2.4 Benchmark dataset

An independent monthly climatological dataset was used to assess the long-term performance of the reconstructed precipitation product. This dataset consists of 1,320 monthly precipitation maps at 1 km spatial resolution covering the full island and was produced from the AdB archive using a two-step hybrid methodology that combines multiple regression models and geostatistical techniques.

Although based on the same underlying observational network, the benchmark dataset was generated through an independent modelling workflow and therefore provides a valuable reference for evaluating the climatological consistency and spatial realism of the reconstructed daily fields after aggregation to monthly time scales.

Table 1: Summary of in situ precipitation datasets used in this study.

Category / variable	Source	Temporal coverage	Spatial / station detail	Native resolution (spatial / temporal)	Role or physical rationale
In situ precipitation networks					
DRPC	Dipartimento della Protezione Civile – Sicilia	2022–2025	369 stations (post-QC)	1 hour (daily aggregated)	Model development and calibration
AdB	Autorità di Bacino	1951–2022	Variable (up to 441 stations)	Daily (climatological daily aggregated)	Historical reconstruction input
SIAS	Servizio Informativo Agrometeorologico Siciliano	2002–2022	107 stations (post-QC)	10-minute (climatological daily aggregated)	Historical reconstruction input
AdB–SIAS (merged)	Combined historical archive	1951–2022	Up to 547 stations (variable)	Daily (homogenized)	Full historical reconstruction (1951–2022)
Auxiliary covariates					

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

CAPE (dynamic)	ERA5 (ECMWF)	1951-2025	0.25° grid; resampled to 2 km	1-hourly; Daily (climatological daily aggregated)	Convective instability; predictor for intensity of convective rainfall
VIMFD (dynamic)	ERA5 (ECMWF)	1951-2025	0.25° grid; resampled to 2 km	1-hourly; Daily (climatological daily aggregated)	Moisture convergence; predictor for moisture supply sustaining stratiform rainfall
Elevation (DEM)	SRTM v3.0 (NASA/DLR)	Static	~30 m native; bilinearly resampled to 2 km	–	Orographic forcing: lift, rain-shadow, adiabatic cooling on windward slopes
Distance to coastline	Derived (OpenStreetMap)	Static	Computed at 2-km grid-cell centroids	–	Maritime influence proxy: land–sea breezes, moisture advection, SST effects

2.3. Preliminary analysis of the dataset

All station coordinates were reprojected from geographic to the planar coordinate system Universal Transverse Mercator Zone 33N (EPSG:32633), using the *pyproj* library (Snow et al., 2025), ensuring that the Euclidean distance calculations used in the variogram and kriging framework are metrically consistent and free from the angular distortions inherent in geographic coordinate representation. The reconstruction domain spans 12°–16° E and 36.5°–38.5°N, discretized onto a regular 2 km × 2 km grid of 100 × 203 cells (20,300 in total). A land mask was constructed by querying an R-tree spatial index against the Sicily boundary polygon (Guttman, 1984), retaining the 6,786 cells lying on the mainland (≈27,100 km², consistent with the island's area including coastal cells intersected by the 2 km buffer of shoreline). Offshore cells were assigned missing values and excluded from all subsequent analyses.

3. Methodology

3.1 Phase I: Geostatistical Modelling of Rainfall Occurrence

Daily precipitation is a compound, intermittent process whose spatial footprint varies fundamentally with the governing meteorological mechanism. The coexistence of wet and dry observations with the presence of zero-rainfall within a single spatial field inflates the empirical variogram nugget, distorts spatial autocorrelation estimates, and degrades kriging performance when the field is treated as a single statistical population (Goovaerts, 1997; Webster and Oliver, 2007).
To address this issue, each calendar day was stratified into three mutually exclusive spatial intermittency classes based on the fraction of dry gauges (Barancourt et al., 1992), F_{dry} , defined as the ratio of stations recording precipitation below the 0.2 mm threshold to the total number of active stations on that day:

- **Widespread** (low intermittency; $0 \leq F_{dry} < 0.25$): synoptically driven events with spatially coherent precipitation across most of the domain (large-scale) and long correlation ranges;

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

- **Intermediate** (moderate intermittency; $0.25 \leq F_{\text{dry}} \leq 0.75$): mixed-regime events with partial spatial coverage and combined stratiform-convective structures;
- **Localized** (high intermittency; $0.75 < F_{\text{dry}} \leq 1.00$): isolated convective events characterized by high intermittency (most of the domain is dry) and short correlation lengths.

This stratification represents the primary conditioning axis of the framework: all subsequent modelling steps (Phase I and Phase II) are performed independently within each class, ensuring that variogram estimation, kriging weights, and intensity modelling are derived from physically homogeneous event populations.

Rainfall occurrence probability at each grid cell was estimated using Indicator Kriging (IK; (Journel, 1983)). Each observation $Z(s_i, t)$ was first transformed into a binary indicator using a 0.2 mm threshold consistent with the gauge resolution:

$$I(s_i, t) = \begin{cases} 1 & if \quad Z(s_i, t) > trace\_threshold \qquad Rain \\ 0 & otherwise \end{cases} \tag{1}$$

The IK predictor at an unsampled location s_o is given by:

$$\begin{cases} \sum_{j=1}^{n} \lambda_j \, \gamma(\mathbf{s}_i, \mathbf{s}_j) + \mu = \gamma(\mathbf{s}_i, \mathbf{s}_0), \quad i = 1, \ldots, n \\ \sum_{j=1}^{n} \lambda_j = 1 \end{cases} \tag{2}$$

where λ_i are kriging weights obtained under the unbiasedness constraint $\Sigma \lambda_i = 1$ (Cressie, 2015; Matheron, 1963). The resulting estimate $\hat{I}(s_0) \in [0, 1]$ represents the probability of rainfall occurrence.

Class-specific empirical semivariograms were computed by pooling all station pairs belonging to each intermittency class, rather than averaging daily variograms, to obtain a more stable representation of the characteristic spatial structure of each regime (Haberlandt, 2007). Three semivariance estimators were evaluated: (i) the classical Matheron estimator (Matheron, 1963); (ii) the robust Cressie–Hawkins estimator (Cressie & Hawkins, 1980), and (iii) the median-based Dowd estimator (Dowd, 1984).

For the Matheron and Cressie–Hawkins estimators, within-bin aggregation was performed using a distance-weighted mean to reduce bin-edge sensitivity. Two binning strategies were tested: uniform, which partitions the lag axis into equal-width intervals, and log-spaced, which provides higher resolution at short distances where spatial autocorrelation is strongest (Bowman and Crujeiras, 2013; Goovaerts, 1997). Each estimator-binning strategy was evaluated via leave-one-out cross-validation (LOOCV; Sect. 3.3).

Four theoretical models (Spherical, Exponential, Gaussian, and Matérn) were fitted to the empirical semivariograms, covering a wide range of spatial continuity behaviors from rough to smooth fields (Cressie, 2015; Goovaerts, 1997; Stein, 1999). Parameters (nugget (C_0), sill ($C_0 + C_1$), range (a), and, for the Matérn model, the smoothness parameter ν) were estimated by weighted least squares minimization. The optimal configuration for each class was selected via LOOCV (Sect. 3.3).

The resulting probability-of-occurrence field $\hat{I}(x)$ was then used to condition intensity estimation through three alternative strategies:

Method I1: Fixed class-specific threshold: A single probability cutoff τ_{fixed} is defined per intermittency class based on observed rain frequency matching:

$$\tau_{fixed} = \widehat{F_p}^{-1}(1 - \bar{o}) \tag{3}$$

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Method I2: Dynamic daily threshold: A daily adaptive cutoff $F_{\mathrm{allocation}}(p)$ is derived from the daily Empirical Cumulative Distribution Function (ECDF) of predicted probabilities across mainland grid cells over the calibration period:

$$F_{allocation}(p) = median_{days\ in\ class}\left(ECDF_{day}(p)\right) \tag{4}$$

This ensures that the predicted wet area matches the observed daily wet fraction on a day-by-day basis, rather than only on average.

Method I3: Probabilistic weighting (hurdle approach): No binarization is applied, and the full probability-of-occurrence field $\hat{I}_{occ}(x)$ is retained and multiplied cell-wise by the interpolated intensity field $M(x)$:

$$\widehat{P}(day) = I(day) \times M(day) \tag{5}$$

consistent with the conditional expectation identity:

$$E[R \mid data] = E[R \mid rain\ occurs, data] \times Pr(rain\ occurs \mid data) \tag{6}$$

This formulation preserves probabilistic information and avoids the information loss induced by thresholding (Cragg, 1971; Long et al., 2016).

All three conditioning strategies are carried forward independently into Phase II, enabling a controlled assessment of their impact on reconstruction skill.

3.2 Phase II: Conditional Modelling of Rainfall Intensity

Phase II estimates daily rainfall depth at each grid cell conditional on the occurrence information generated in Phase I. Because the statistical properties and spatial correlation structure of rainfall intensity depend strongly on the governing precipitating mechanism, ranging from widespread stratiform systems with long correlation ranges and moderate skewness to isolated convective events characterized by short ranges, strong nugget effects, and heavy-tailed distributions, intensity modelling is performed within a doubly conditional stratification framework that accounts for both event spatial footprint and magnitude regime.

The three intermittency classes (i.e., Widespread, Intermediate, Localized) define the first stratification axis. The second axis characterizes event magnitude using two complementary metrics computed from wet stations: the spatial mean intensity (μ, mm day^{-1}) and the maximum observed point intensity ($d_{\max}$, mm day^{-1}). Both metrics were initially classified according to the Mediterranean daily rainfall-rate taxonomy from Alpert et al. (2002), which distinguishes Light (< 4 mm day^{-1}), Light-Moderate (4–16), Moderate-Heavy (16–32), Heavy (32–64), Heavy-Torrential (64–128), and Torrential (> 128 mm day^{-1}) categories, yielding for each day a pair of intensity categories (μ-class, $d_{\max}$-class). To ensure sufficient sample sizes for robust variogram estimation within each stratum, the six Alpert categories were consolidated into three magnitude super-classes (Light, Moderate, Heavy) through a frequency-based alignment procedure described below. The super-class assignment of each day is determined by a look-up rule that maps the two-dimensional (μ-class, $d_{\max}$-class) pair to a super-class, with the mapping conditioned on the intermittency class of that day (Table 1). The rule is constructed so that the marginal frequency of each magnitude super-class matches, as closely as possible, the marginal frequency of the corresponding intermittency class, yielding balanced regime sample sizes across the nine strata. In detail: a day is classified as Light if both μ and $d_{\max}$ fall in the A (Light) or B (Light-Moderate) categories (pair A–A or A–B); as Heavy if $d_{\max}$ reaches C2 (Heavy) or above while μ simultaneously reaches B or above (pairs B–C2 through C2–D2); and as Moderate for all remaining (μ, $d_{\max}$) combinations. This rule is applied identically across all three intermittency classes, including Localized days: because μ is computed only over the small number of reporting wet stations on such days, an isolated but intense convective cell can drive $d_{\max}$ into the

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Heavy-Torrential or Torrential range while μ remains Light-Moderate, so Localized-Heavy is retained as a genuine, if sparsely populated, regime (Table 3). Because the mapping table is fully specified, each day belongs to exactly one of the nine regimes; no day is left unassigned or doubly assigned. The complete frequency-based partition of the 1096 calibration days across the nine regimes, consistent with Table 3, is provided in the Supplementary Materials (Table S.2).

3.2.1 Framework IIA: Geostatistical Intensity Modelling

Non-zero rainfall intensities exhibit strong right-skewness and heavy tails, violating the near-Gaussian assumption required by OK (Cressie, 2015; Goovaerts, 1997; Wilks, 2011), a regime-specific quantile transformation was applied to wet-station observations:

$$Z' = \emptyset^{-1}\left(\hat{F}(Z)\right) \quad (7)$$

where Φ^{-1} is the probit function and $\hat{F}$ is the ECDF within each regime over the calibration period. This rank-based transformation was preferred over parametric alternatives (square-root, logarithmic, Box-Cox, and Yeo-Johnson) based on the Shapiro–Wilk normality test and quantile–quantile (Q-Q) plot evaluation across all nine regimes a representative comparison against the four parametric alternatives, illustrated for three regimes, is provided in the Supplementary Materials (Fig. S.1). OK was then applied in the Gaussian space, using regime-specific semivariogram models. Empirical semivariograms were computed using pooled observations within each regime, as described in Phase I (Sect. 3.1). For each regime, an exhaustive model selection was performed across combinations of, namely three semivariogram estimators (Matheron, Cressie–Hawkins, Dowd), two lag-binning strategies (uniform and log-spaced), and a suite of 13 theoretical semivariogram models (Spherical, Exponential, Gaussian, Matérn, Power Law, Linear, Rational Quadratic, J-Bessel, Cubic, hole-effect, Circular, Stable, Super-Spherical, and nested or hybrid combinations) (Chilès and Delfiner, 2012; Cressie, 2015).
The optimal configuration for each regime was selected via LOOCV (Sect. 3.3, see Fig. S.2-4); the semivariance estimators, binning strategies, and theoretical variogram models considered, together with the criteria guiding their selection, are summarized in the Supplementary Materials (Tables S.3–S.5). Uncertainty in the original space was propagated via Monte Carlo simulations ($N = 1000$ samples per grid cell).
For Methods I1 and I2, OK was restricted to grid cells classified as rainy by Phase I mask, whereas dry cells were set to zero. For Method I3, OK was performed over the entire mainland domain and modulated by the Phase I probability field. All computations were performed using the *PyKrige* library.

3.2.2 Framework IIB: Regression-Kriging Intensity Modelling

Framework IIB decomposes rainfall intensity into a physically interpretable large-scale deterministic trend and a spatially correlated residual:

$$Z(s) = m(s) + e(s) \quad (8)$$

where *m(s)* is the trend component and *e(s)* is the residual field. The trend was estimated using a Generalized Additive Model (GAM; Hastie and Tibshirani, 1986), assuming a Gamma distribution with a log-link function to enforce positivity and account for heteroscedasticity (Wilks, 2011; Wood, 2017).
The covariate set included static physiographic predictors (elevation and distance to coastline), dynamic atmospheric predictors (CAPE and VIMFD), and the Phase I probability-of-occurrence field (Sect. 2.2). Candidate GAM structures, including smooth and tensor-product iterations, were evaluated and optimized using Generalized Cross-Validation (GCV) in *pyGAM* (Servén et al., 2025).

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

The best-performing GAM structure for each hydrometeorological regime was selected by daily LOOCV using the continuous performance metrics described in Sect. 3.3. The final selected model for each regime took the general form:

$$log(\mu) = \beta_0 + s(\text{Prob}_i) + s(\text{CAPE}_i) + te(\text{DEM}_i, \text{dist-coast}_i) + s\left(\text{doy}_i, \text{basis} = ' cp'\right) \quad (9)$$

where $\mu = E[R]$ is the expected rainfall intensity and β_0 the intercept; $s(\cdot)$ denotes a univariate penalized spline, a smooth, data-driven function that captures the (possibly non-linear) individual effect of each covariate without prescribing its functional form; and $te(\cdot,\cdot)$ denotes a tensor-product smooth, the two-dimensional extension of $s(\cdot)$ that models joint, interacting effects of covariate pairs, which here are the orographic enhancement modulated by moisture availability (elevation × VIMFD) and the convective instability conditioned on synoptic-scale favorability (CAPE × occurrence probability). The GAM residuals, representing the component of spatial variability unexplained by the deterministic covariate-driven trend, were interpolated using OK. Model selection was again performed via LOOCV (Sect. 3.3).

3.3 Evaluation, Validation, and Historical Reconstruction

LOOCV served as the primary tool for model selection and performance evaluation throughout Phases I and II. In each iteration, one station is withheld from the fitting dataset, the model is calibrated on the remaining $N - 1$ daily observations, a prediction is generated at the withheld location, and the predicted value is compared with the withheld observation.

This procedure mimics the intended application of the framework, namely the estimation of precipitation at ungauged grid cells, while maximizing the use of available observations. Because the mean inter-station distance (8.3 km) is smaller than the fitted variogram ranges (15–39 km), non-buffered LOOCV may yield an optimistic estimate of predictive skill (Cressie, 2015; Roberts et al., 2017). LOOCV was applied independently for each calendar day within each intermittency class (Phase I) or hydrometeorological regime (Phase II); day-level metrics were then aggregated to class- or regime-level summaries.

Phase I performance was evaluated using several verification metrics, including contingency table-based metrics such as the Probability of Detection (POD), False Alarm Ratio (FAR), Critical Success Index (CSI), Overall Accuracy (ACC), together with the Brier Score (BS):

$$POD = {}^{H}\!/_{(H + M)} \quad (10)$$

$$FAR = {}^{FA}\!/_{(H + FA)} \quad (11)$$

$$CSI = {}^{H}\!/_{(H + M + FA)} \quad (12)$$

$$ACC = (H + CN)/(H + M + FA + CN) \quad (13)$$

$$BS = {}^{1}\!/_{N}\ \Sigma\,(p_i - o_i)^2 \quad (14)$$

where H, M, FA, and CN indicate Hits, Misses, False Alarms, and Correct Negatives, respectively, while p_i is the predicted occurrence probability and o_i is the binary observed outcome at the i-th withheld station.

Phase II performance was instead evaluated in the original rainfall units (mm day^{-1}) following back-transformation of kriging predictions. Both Framework IIA and Framework IIB were evaluated using Relative Bias (RB), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), and Kling–Gupta Efficiency coefficient (KGE):

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Open Access Earth System Science Data Discussions

$$\mathrm{RB} = 100 \times \frac{\frac{1}{n}\sum_{i=1}^{n}(\hat{z}_i - z_i)}{\frac{1}{n}\sum_{i=1}^{n} z_i} \tag{15}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{z}_i - z_i)^2} \tag{16}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(z_i - \hat{z}_i)^2}{\sum_{i=1}^{n}(z_i - \bar{z})^2} \tag{17}$$

$$KGE = 1 - \sqrt{[(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2]} \tag{18}$$

where r is the Pearson correlation between observed and predicted values, $\alpha = \sigma_{\hat{z}}/\sigma_z$ is the ratio of standard deviations, $\beta = \mu_{\hat{z}}/\mu_z$ is the ratio of means and μ and σ denote the mean and standard deviation, respectively. (Gupta and Kling, 2011). For Framework IIA exclusively, the calibration reliability of the kriging variance was additionally assessed through standardized residual diagnostics computed in the transformed space:

$$\begin{gathered} \epsilon_i = \frac{z_i - \hat{z}_i}{\sigma_i} \\ s_1 = \frac{1}{n}\sum_{i=1}^{n}\epsilon_i \\ s_2 = \frac{1}{n}\sum_{i=1}^{n}\epsilon_i^2 \end{gathered} \tag{19}$$

where s_1 is the mean standardized prediction error (ideal value 0, with values within ±0.1 indicating negligible bias in the transformed-space) and s_2 is the mean squared standardized error (ideal value 1, with acceptable values in the range 0.8–1.2) (Kitanidis, 1997). Here, z_i, $\hat{z}_i$, and σ denote the observed value, the kriging prediction, and the kriging standard error in transformed space, respectively.

Framework IIB does not provide a single kriging variance for the full prediction, as total predictive uncertainty results from the combined contributions of the GAM and the residual kriging component; therefore, standardized residual diagnostics were not computed for this framework.

LOOCV performance metrics for Frameworks IIA and IIB were computed separately for each of the nine hydrometeorological regimes, enabling a direct and consistent comparison of the two intensity-modelling paradigms under identical stratification and evaluation conditions.

The comparative performance of the two frameworks is presented in Sect. 4 and used to guide the selection of the operational reconstruction framework.

3.3.1 Historical reconstruction

The calibrated two-phase framework was subsequently applied to the merged AdB–SIAS dataset spanning 1951–2022 (Table 1). All parameters estimated during calibration were transferred unchanged to the historical reconstruction period, implementing the Transfer-Informed Modelling strategy introduced in Sect. 1.

For days on which fewer than four active stations were available, and the kriging system was therefore considered spatially under-constrained, each grid cell was assigned the value of its geographically nearest active station via nearest-neighbor

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

assignment. Over the full 72-year period, only 43 days met this criterion (~6 days per decade on average), and no single year contained more than two such occurrences. All affected days are explicitly flagged in the metadata of released dataset. This procedure yielded three parallel historical reconstructions (one per conditioning strategy), each consisting of a complete daily 2 km gridded precipitation field with an accompanying daily uncertainty map over the 1951–2022 period.

3.3.2 Climatological adjustment

Comparison of the reconstructed fields with the independent monthly benchmark revealed a systematic negative bias arising from the smooth-support limitation of gauge-based interpolation, whereby the 2 km kriging estimate recovers an areal mean and cannot reproduce the higher spatial moments captured by point gauges. To correct long-term magnitude while leaving interannual variability free, and to remain applicable to years for which no benchmark exists, a grid-cell-wise, month-specific climatological multiplicative correction was adopted. For each cell x and calendar month m, a single corrector $C_{(x,m)}$ was formed as the ratio of the long-term mean reconstructed total to the long-term mean benchmark total, pooled over all years 1951–2022 (ratio-of-means):

$$C_{(x,m)} = {\Sigma_y P_{benchmark(x,m)}} \Big/ {\Sigma_y P_{model(x,m)}} \qquad P_{adj(x,d,m)} = C_{(x,m)} \times P_{raw(x,d,m)} \tag{20}$$

where $P_{\mathrm{benchmark}\,(x,m)}$ is the benchmark monthly precipitation total and $P_{\mathrm{model}\,(x,m)}$ is the corresponding reconstructed monthly precipitation total derived from the daily fields, both averaged over all years (1951–2022). The resulting ratio, $C_{(x,m)}$, was then applied uniformly to all raw daily values within the same *(x,m)* block.

Sea cells retained their NaN designation; cells with no reconstructed rainfall in any year received unit correction. The ratio-of-means form is volume-weighted and mass-conserving over the calibration period and was preferred over a mean-of-ratios estimator, which is inflated by low-rainfall months. Because one fixed corrector scales every year identically, the correction cannot reproduce year-to-year variability by construction; the residual interannual scatter against the benchmark therefore constitutes a genuine consistency test rather than an identity. A year-specific corrector was deliberately avoided, as it forces the corrected fields to integrate exactly into the benchmark and is inapplicable in near-real time. The effectiveness of the bias correction procedure is assessed in Sect. 4.

4. Results and Discussion

4.1 Phase I: Geostatistical Modelling of Rainfall Occurrence

The optimal variogram configuration identified through LOOCV for all three classes combined a Matérn variogram model with the Cressie–Hawkins semivariance estimator and a log-spaced binning. Table 2 shows the results of LOOCV performance metrics for Phase I.

Table 2: LOOCV performance metrics for Phase I rainfall occurrence model, stratified by the three intermittency classes (Widespread, Intermediate, Localized)

	CLASS WIDESPREAD	CLASS INTERMEDIATE	CLASS LOCALIZED
Best fit variogram model	Matérn	Matérn	Matérn
Variogram parameters			
[ϑ,	0.20	0.24	0.20
nugget,	0.01	0.03	0.02
range,	38887.74	39054.93	15256.83
sill]	0.09	0.25	0.04

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

POD	0. 91	0.89	0.84
FAR	0.18	0.19	0.23
ACC	0.87	0.82	0.77
CSI	0.85	0.79	0.75
Brier Score (0< BS<1)	0.05	0.07	0.04

The fitted variogram parameters are physically consistent with the expected characteristics of the different precipitation regimes. The sill hierarchy, Intermediate (0.25) >> Widespread (0.09) ≈ Localized (0.04), directly mirrors the variance structure of the underlying binary occurrence field and the degree of spatial heterogeneity expected from each meteorological regime: widespread frontal days drive a near-homogeneous, spatially coherent indicator field of intrinsically low variance; intermediate days juxtapose contiguous rainy belts with dry subregions, maximizing binary contrast and hence sill; and localized days are dominated by zeros, constraining variance not through coherence but through base-rate scarcity. Similarly, the range hierarchy aligns with known meso-meteorological spatial scales, with Localized events exhibiting a shorter correlation range (~15 km), typical of isolated Mediterranean convective cells (Cristiano et al., 2017), and Widespread and Intermediate events displaying larger ranges (~39 km), indicative of synoptic-scale organization (Bárdossy and Pegram, 2014). LOOCV metrics confirm the expected decrease in predictive skill from Widespread to Localized events, with ACC declining from 0.87 to 0.77 and CSI from 0.85 to 0.75 (Table 2). CSI, which directly measures the overlap between predicted and observed rain areas and is insensitive to the abundance of correctly predicted dry cells, is adopted as the summary metric for the rarer classes (Wilks, 2011).

4.2 Phase II: Comparative Framework Assessment and Intensity Model Selection

Table 3 reports the LOOCV performance metrics (back-transformed to mm day^{-1}) for Frameworks IIA (geostatistical Ordinary Kriging) and IIB (Regression-Kriging) across the nine hydrometeorological regimes.

Table 3: LOOCV performance metrics for Framework IIA (on the left) and for Framework IIB (on the right) across nine hydrometeorological regimes.

Regime	n. of wet days/ total active stations	Median/ Min of active stations	RMSE IIA	KGE IIA	R^2 IIA	RB% IIA	s_1	s_2	RMSE IIB	KGE IIB	R^2 IIB	RB% IIB	ΔKGE
W-Light	22-7134	318-313	1.33	0.70	0.79	6.0	−0.001	1.16	5.85	0.58	0.65	8.4	+0.12
W-Moderate	86- 27751	319-309	4.10	0.72	0.83	~8	−0.002	1.09	8.2	0.61	0.70	12.2	+0.11
W-Heavy	49- 1604	320-308	11.81	0.70	0.84	~19	−0.002	1.19	16.3	0.55	0.68	21.6	+0.15
I-Light	105- 34150	320-307	1.31	0.68	0.76	~9	−0.008	1.27	6.1	0.51	0.58	6.8	+0.17
I-Moderate	102- 33277	319-307	3.97	0.61	0.72	~18	−0.007	1.25	7.45	0.48	0.54	19.2	+0.13
I-Heavy	44-14830	318-306	8.12	0.59	0.67	~21	−0.005	0.97	14.9	0.42	0.51	29.4	+0.17
L-Light	612- 200352	318-305	1.27	0.42	0.46	~11	−0.003	1.36	4.45	0.27	0.31	15.2	+0.15
L-Moderate	45- 14739	317-311	5.53	0.49	0.61	~22	−0.016	1.27	7.9	0.32	0.42	26.8	+0.17
L-Heavy	31- 10345	319-308	9.89	0.44	0.44	26.2	−0.011	1.17	16.6	0.21	0.29	39.5	+0.23

Framework IIA shows a clear regime-dependent performance gradient. KGE values range from 0.70–0.72 for Widespread regimes, decrease to 0.59–0.68 for Intermediate regimes, and reach their minimum for Localized regimes (0.42–0.49), mirroring the decline in spatial coherence from frontal to convective precipitation and confirming that predictive skill is controlled primarily by the spatial structure of the underlying regime.

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Bias diagnostics indicate negligible systematic error in the transformed space ($|s_i| \leq 0.016$). Relative Bias after back-transformation is positive across regimes (6–26%), an expected consequence of the asymmetric inverse quantile transformation; importantly, it reflects the spatial mean over wet stations only, while the day-to-day spatial pattern and the extent of dry areas are correctly reproduced.

The two right-most columns of Table 3 report the calibration support behind each regime: while the frequent regimes rest on hundreds of days and $>10^4$ station pairs, the rarest convective class (Localized-Heavy) is constrained by only 31 days and 10345 wet station-days within the 2022–2025 window, so its fitted variogram and the upper tail of its quantile transform carry correspondingly wider uncertainty and should be interpreted accordingly. Framework IIB performs worse than IIA across all metrics and all regimes, with lower KGE (0.21–0.61 vs 0.42–0.72), higher RMSE, and larger Relative Bias; the largest discrepancies occur for Localized events (ΔKGE up to 0.23 for Localized-Heavy). Bootstrapped 95% confidence intervals for the per-regime ΔKGE (resampled over calibration days; Table 3) exclude zero for all regimes, confirming that the superiority of IIA is not an artefact of sampling variability; the interval for Localized-Heavy is correspondingly wide, consistent with its limited 31-day support (Sect. 4.2).

This inferior performance stems from the scale mismatch between the ERA5 covariates (~28 km native resolution) and the 2 km target grid: at this native resolution ERA5 resolves only a few cells across Sicily, so the resampled covariate fields vary little from one part of the island to another and add little information beyond the spatial structure already captured by the gauge-derived variogram, even in the Widespread regimes where physical predictors should be most informative (ΔKGE 0.11–0.15). For Localized events the day-by-day correlations between CAPE/VIMFD and local precipitation approach zero, so the GAM trend contributes noise rather than signal, and any trend misspecification propagates directly into the residual field entering the kriging step; the resulting Relative Bias (39.5% vs 26.2%) is a structural signature of this compounding, not random noise. The limitation of IIB therefore lies in the informativeness of the available covariates, not in the kriging component, which is identical in both frameworks. Overall, Framework IIA provides more accurate and robust performance across all regimes and is therefore selected for the historical reconstruction. Subsequent results refer to IIA unless otherwise stated.

4.3 Sensitivity Analyses

To verify the stationarity of the spatial covariance structure assumed by the fixed-parameter transfer and to address the possibility of a trend, the per-regime variograms were re-estimated independently on two sub-periods of comparable length (early: 1951–1986 and late: 1987–2022) of the historical network. Holding the network fixed, the early and late variograms coincide within their bootstrapped 95% confidence intervals across all well-resolved regimes, and none of the diagnostics reveals the range contraction or relative-nugget increase that would signal a systematic change in spatial structure. For the least-populated Heavy regimes, however, the bootstrapped confidence intervals are wide enough that a moderate shift in spatial structure cannot be excluded; the stationarity conclusion is therefore best regarded as robust for the frequent, well-sampled regimes and provisional for the rarest ones, consistent with the wider calibration uncertainty already noted for these regimes (Sect. 4.2). Within the resolution of this test, and acknowledging that the regime definitions and temporal blocking inherited from the PRECISi construction constrain what such a comparison can detect, the frequent, well-sampled regimes show no evidence of a material change in covariance structure between sub-periods. This supports the practical assumption underlying the reconstruction, namely that a single set of calibration parameters can be transferred across the record; it concerns the second-order spatial structure of daily fields and is not in tension with the long-term changes in rainfall totals and intensity documented for the region, which act on different statistical properties. Representative variogram comparisons are shown in Fig. 4.

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Open Access Earth System Science Data Discussions

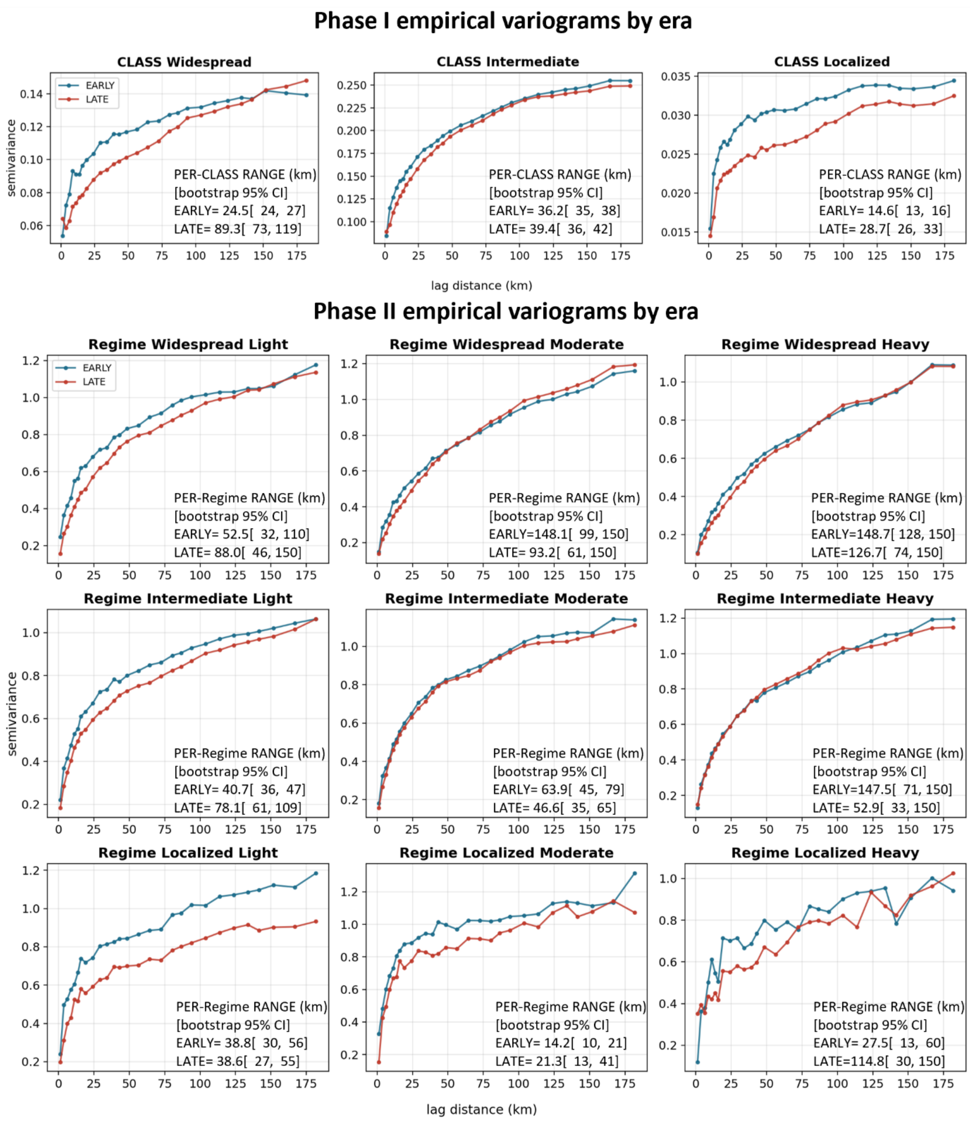

Figure 4: Empirical variograms re-estimated by sub-period and network for the (a) occurrence (indicator) and (b) intensity (normal-score) regimes.

520 A second concern is that the sparser historical network biases the regime classification: with fewer gauges the per-day maximum depth d_{max}, a sample extreme whose expectation grows with the number of gauges, is under-estimated, drifting days toward lighter magnitude classes. This is tested by subsampling the modern network down to a range of densities spanning the realized record and recomputing the classification, with network density read directly from the observed daily gauge availability. The d_{max} bias and resulting class lightening decay smoothly as density increases and are negligible across all realized densities (≥150 gauges day^{-1}: mean d_{max} bias < 1 mm and < 2% of days reclassified) (see Fig. 5), becoming material only below ~60 gauges day^{-1}, a density the Sicilian network never reached. The regime classification is therefore robust over the entire reconstruction period; the historical under-estimation is consequently not a density artefact, confirming the magnitude classification underlying both delivered products.

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Open Access Earth System Science Data Discussions

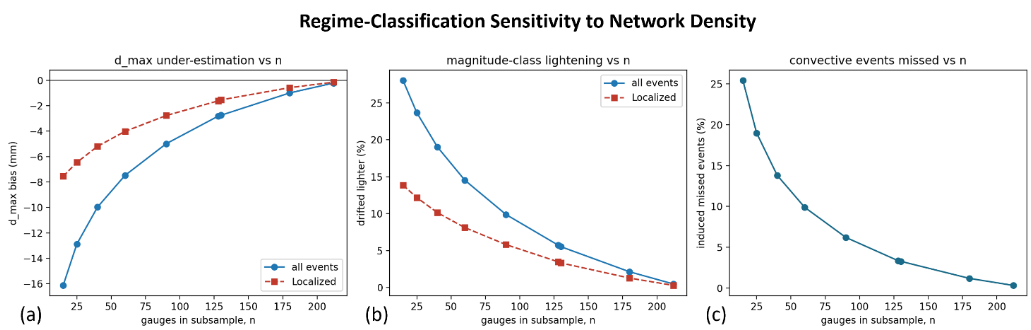

Figure 5: Regime-classification sensitivity to network density: panel (a) d_{max} bias, panel (b) magnitude-class lightening, and panel (c) induced missed events as a function of subsample size n, for all events and the convective (Localized) subset.

4.4 Model Selection, Historical Reconstruction and Bias Assessment

The calibrated Phase I–IIA pipeline was applied to the merged AdB–SIAS historical archive spanning 1951–2022, producing three daily 2 km gridded precipitation products, which were first evaluated against the independent monthly climatological benchmark in their raw, uncorrected form.

Following the framework selection (Sect. 4.2), the geostatistical variant IIA was carried forward to reconstruct the full 1951–2022 record for the two delivered occurrence schemes, I2-IIA and I3-IIA; the intermediate scheme I1-IIA is omitted hereafter, as its fixed-threshold occurrence field is effectively indistinguishable from I2-IIA and yields near-identical totals; for this reason, I1-IIA is not carried forward into the multi-scale validation of Sect. 4.5 or the corresponding Supplementary Materials. Evaluated in raw form against the independent monthly benchmark (Fig. 6, grey points), both products exhibit the systematic underestimation anticipated from the smooth-support limitation of gauge interpolation, and is present at both temporal scales (I2-IIA: $R^2 = 0.94$ monthly, 0.47 annual; I3-IIA: $R^2 = 0.73$ monthly, −2.12 annual). The mean bias grows from −8.6 to −103 mm for I2-IIA and from −20.6 to −247 mm for I3-IIA between the monthly and annual scales, reflecting the accumulation of a persistent per-month deficit into a larger annual shortfall. The negative annual score for I3-IIA indicates squared errors exceeding the observed interannual variance, a consequence of accumulated bias, not of poor correlation. The offset was removed with the grid-cell-wise, month-specific climatological factor $C_{(x,m)}$ of Sect. 3.2.4, computed as the ratio of the long-term mean benchmark to the long-term mean reconstruction (Ratio-of-Means, RoM) and applied identically to every year, so that interannual variability remains unconstrained and the residual scatter constitutes a genuine consistency test rather than an identity. The volume-weighted, mass-conserving RoM estimator was preferred over a mean-of-ratios alternative, which is inflated by low-rainfall months and degraded I3-IIA. After correction (Fig. 6, blue), the mean bias vanishes at both scales by construction, while the freely varying component improves markedly. I2-IIA attains monthly $R^2 = 0.99$ and annual $R^2 = 0.98$, whereas I3-IIA recovers only to 0.91 and 0.46, a residual gap rooted in the variance structure of the hurdle product, which mean scaling cannot repair. Its spatial and mechanistic origin is examined next.

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Open Access Earth System Science Data Discussions

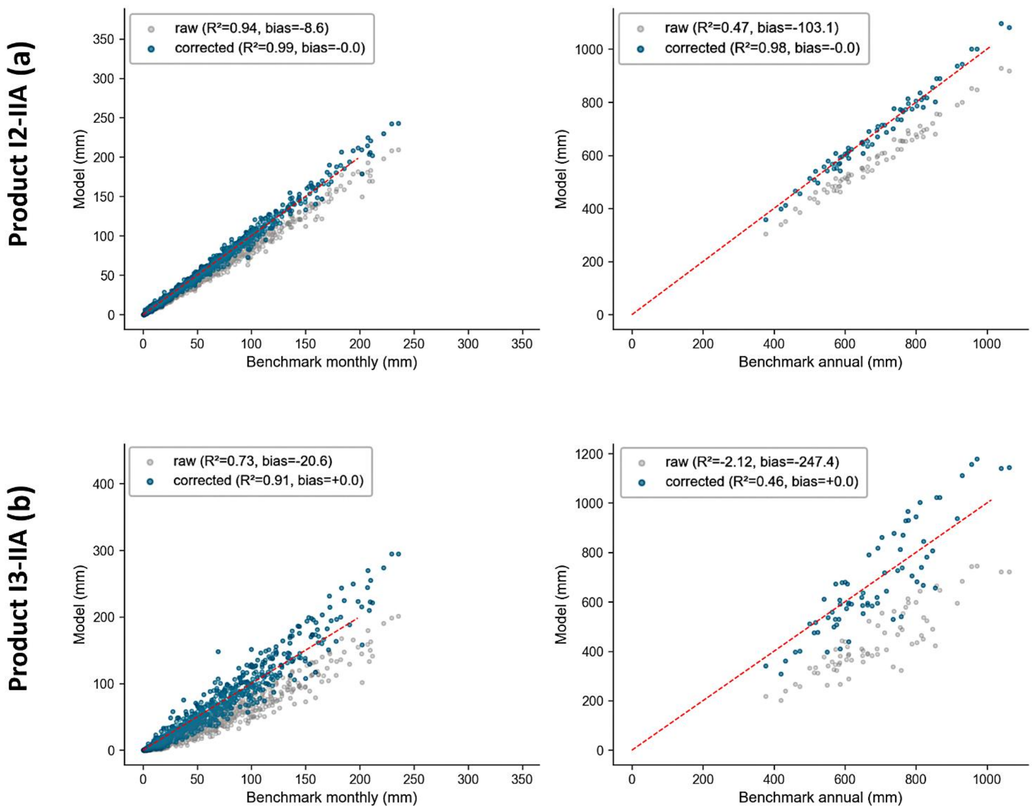

Figure 6: Monthly (left) and annual (right) spatial-mean totals of the raw (grey) and climatologically corrected (blue) reconstructions against the independent benchmark for I2-IIA (a) and I3-IIA (b).

4.5 Multi-Scale Validation

4.5.1 Temporal stability of the reconstruction

The full 1951–2022 spatial-mean series (Fig. 7) demonstrates that the correction restores long-term magnitude without distorting the temporal signal. For I2-IIA the corrected series tracks the benchmark across the entire record, reproducing the seasonal cycle and individual extreme months, including the >300 mm peaks of the early 1950s, while the raw series runs persistently below it. The residual panel makes the improvement explicit. The raw deficit, typically −5 to −25 mm and exceeding −50 mm in some early wet months, collapses after correction onto the zero line, with symmetrically distributed departures (≈ ±10 mm) and no trend, seasonality, or drift across the 72-year span. Because a single climatological factor is applied identically to every year, this stability is not imposed, because the corrector removes the systematic offset while leaving interannual and intra-annual variability entirely free, so the close month-to-month agreement is a falsifiable result rather than an artefact. The bias removal is temporally homogeneous (equally effective in the sparsely observed pre-1975 period and the dense modern era) consistent with the density-sensitivity analysis of Sect. 4.3, which excluded network sparsity as the source of the historical deficit. The I3-IIA panel exposes the limits of mean-only scaling. Its corrected residuals are markedly wider and, from the 1990s onward, become increasingly positive. This occurs because the raw deficit of the hurdle product diminishes through time, whereas the climatological factor is by design constant, remains constant; the corrector applied to a time-varying deficit therefore over-corrects the later decades. This progressive, time-dependent drift is far weaker in I2-IIA, where any residual trend stays within the ≈ ±10 mm envelope, indicating that the deficiencies of I3-IIA are structural rather than climatological, and anticipating the variance diagnostics that follow.

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Earth System Science Data Discussions
Open Access

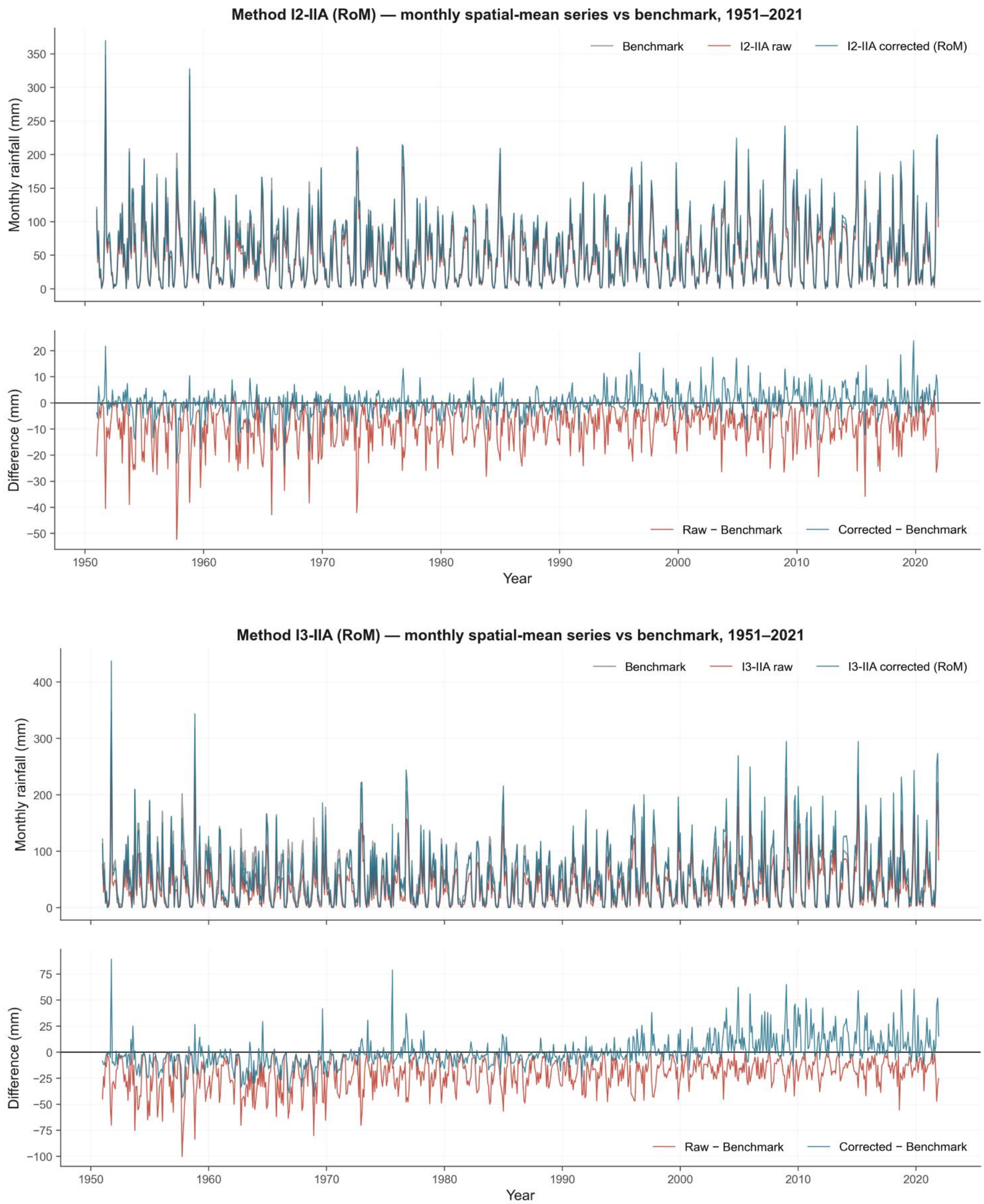

Figure 7: Monthly spatial-mean precipitation series, raw and bias-corrected versus benchmark, 1951–2022, for (a) Method I2-IIA and (b) Method I3-IIA. (upper panel) Raw series (benchmark gray, I2-IIA red, I2-IIA corrected blue). (bottom panel) Raw differences (red) and the corrected residual errors (blue).

4.5.2 Spatial structure of interannual skill

To assess skill beyond the climatological mean, which the RoM corrector matches by construction, the interannual agreement between the corrected products and the benchmark was decomposed per grid cell into the three independent KGE components (Gupta et al., 2009; Kling et al., 2012): temporal correlation r, variability ratio $\alpha = \sigma_{model}/\sigma_{benchmark}$ and bias ratio β, complemented by the interannual RMSE (Fig. 8). Three representative months are shown: January (wet frontal month), July (dry convective stress month) and October (autumn transitional month); the full monthly panels are provided in the

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Supplementary Materials (Fig. S.5-6). The β maps are uniformly ≈1 for both products, confirming that the correction constrains the long-term mean and nothing else; the remaining components therefore constitute a genuine, non-circular validation. Both products capture interannual timing well ($r > 0.8$ almost everywhere). The discriminating component is α. For I2-IIA it remains close to unity in all three months, with only mild under-dispersion in July, indicating that the dynamic-threshold product transmits the benchmark's year-to-year variance essentially unaltered. For I3-IIA the hurdle product is variance-distorted, and the mean-only correction cannot restore the variance structure, leaving α systematically above unity, most severely in July, where sparse convective rainfall renders the back-transformed hurdle field over-disperse. The RMSE column translates these ratios into millimeters. July inflation costs little in absolute terms, whereas the wet and transitional months concentrate the real error, with I3-IIA exceeding 50 mm over the north-eastern relief in January and October, roughly double the co-located I2-IIA values. This spatially resolved, mechanistic contrast explains the divergent annual skill of the two products ($R^2 = 0.98$ vs 0.46) and consolidates the selection of I2-IIA as the reference product.

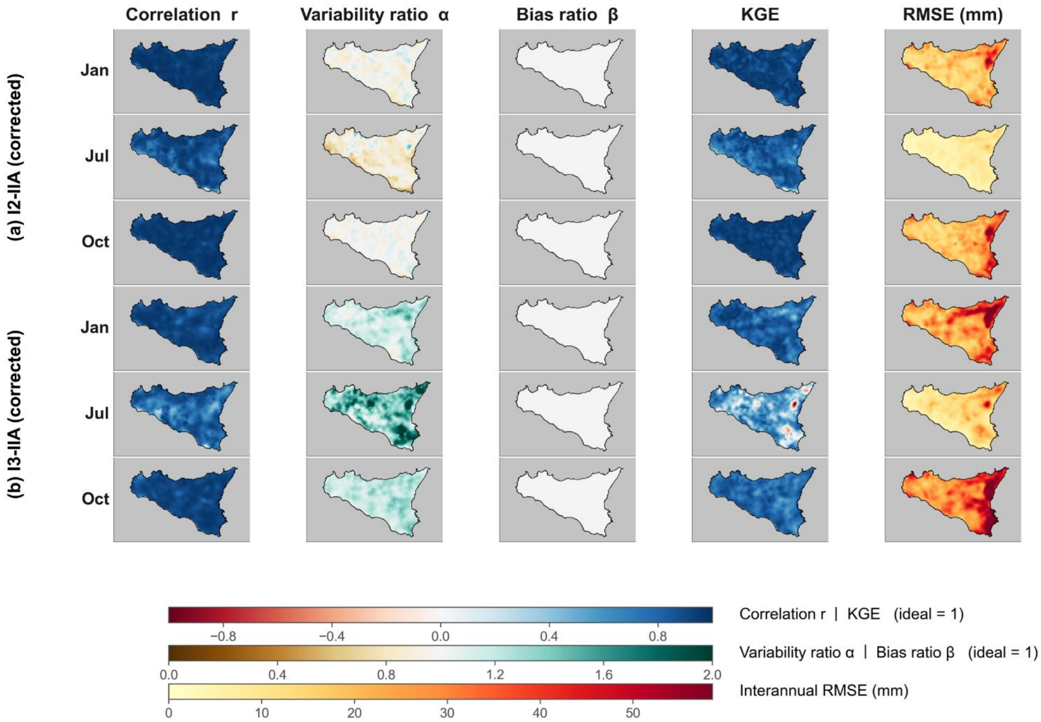

Figure 8: Per-grid-cell interannual validation of the corrected (a) I2-IIA and (b) I3-IIA products against the benchmark for three representative months: January (wet frontal), July (dry convective stress) and October (autumn transitional). Columns: the KGE decomposition into temporal correlation r, variability ratio α and bias ratio β, the composite KGE, and the interannual RMSE (mm).

4.5.3 Error distributional analysis

The per-grid-cell error distributions (Fig. 9; Tables 4-5) resolve the bias correction beyond domain means. In the raw state both products underestimate systematically, with the magnitude of the median bias increasing toward the wet season: I2-IIA reaches −8.4 mm in October and −8.2 mm in December, whereas the hurdle-based I3-IIA is two-to-three times more biased over the same months (−23.6 mm in December, −23.3 mm in October), confirming its stronger wet-season deficit. After correction, the medians collapse to near zero in every month and season ($|\text{median}| \leq 1.3$ mm for I2-IIA; ≤ 2.0 mm for I3-IIA), while the distributions retain a finite spread, the expected outcome of a climatological corrector that removes the systematic offset but deliberately leaves year-to-year departures free. This residual spread behaves physically, scaling with the seasonal rainfall magnitude, contracting from ~3 mm (*IQR* for July) to ~18 mm (October–December) for I2-IIA, and is systematically narrower for I2-IIA than for I3-IIA by roughly 40% in the wet seasons (seasonal *IQR* 16.0 vs 26.8 mm in DJF; 16.7 vs 28.0

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

mm in SON). A second, diagnostic feature emerges from the divergence between the two dispersion measures: for I3-IIA the standard deviation exceeds the *IQR* most severely in summer (August: 21.9 vs 6.5 mm), revealing heavy tails produced by localized convective cells that the hurdle formulation reproduces poorly; indeed, its JJA correction tightens the distribution body (*IQR* 10.6 → 4.5 mm) while the tails grow heavier (*SD* 13.8 → 17.6 mm). I2-IIA shows the same qualitative pattern but far less severely (JJA corrected *SD*/*IQR* ≈ 2.0 vs ≈ 3.9 for I3-IIA), indicating that its error distribution, while not strictly Gaussian, is markedly less prone to extreme-tail contamination than the hurdle-based product.

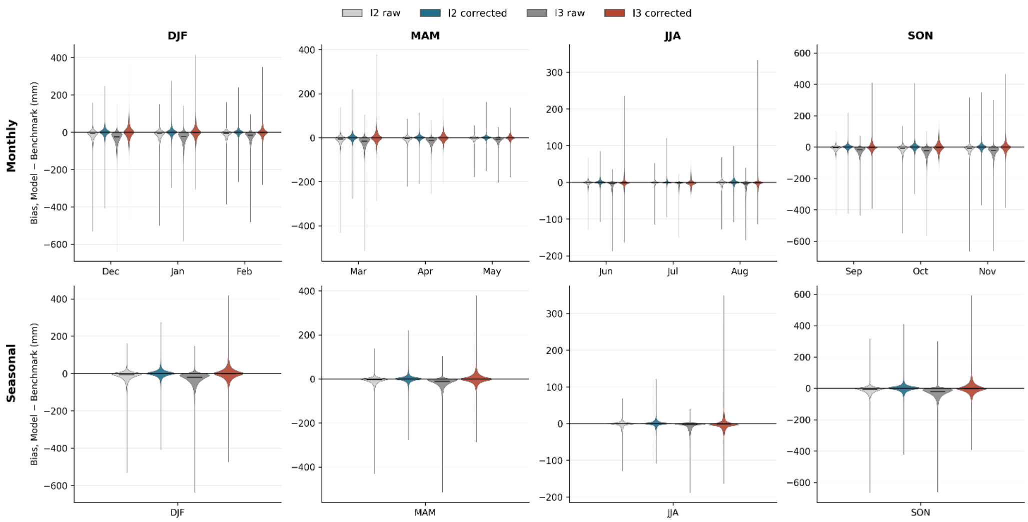

Figure 9: Per-grid-cell bias (model - benchmark) distributions. The I2 (I3) raw distribution is in light (dark) gray, while the I2 (I3) corrected one is in blue (red) and they are grouped per month (upper row) and per season (lower row); note the independent y-scales.

Table 4: Error spread (median, *IQR*, *SD*; mm), Method I2-IIA, raw vs. bias-corrected

	Raw median	Raw IQR	Raw SD	Corrected median	Corrected IQR	Corrected SD
Dec	−23.6	35.0	37.0	−0.1	30.7	35.8
Jan	−21.5	33.2	33.7	−0.6	28.6	32.4
Feb	−16.3	27.1	27.2	−0.8	21.9	27.4
Mar	−15.9	25.1	27.4	−0.7	21.5	26.5
Apr	−12.6	18.8	18.6	−0.7	18.5	20.3
May	−7.9	14.0	15.3	−0.2	12.0	15.6
Jun	−3.2	9.7	12.6	−0.4	4.9	14.8
Jul	−1.2	6.9	11.2	−0.2	2.8	14.9
Aug	−4.2	15.3	16.6	−0.7	6.5	21.9
Sep	−17.5	26.3	24.0	−2.0	25.3	28.3
Oct	−23.3	33.9	35.0	−1.8	31.5	38.2
Nov	−20.8	31.0	33.9	−0.5	27.3	33.6
DJF	−20.2	32.2	33.1	−0.5	26.8	32.1
MAM	−11.7	19.4	21.5	−0.5	16.9	21.3
JJA	−2.6	10.6	13.8	−0.4	4.5	17.6

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

Open Access Earth System Science Data Discussions

SON	−20.4	30.4	31.6	−1.5	28.0	33.6

Table 5: Same as Table 4, for Method I3-IIA

	Raw median	Raw IQR	Raw SD	Corrected median	Corrected IQR	Corrected SD
Dec	+23.2	35.1	36.1	+0.1	31.0	35.9
Jan	+21.9	33.4	32.2	+0.8	27.4	31.3
Feb	+16.4	26.1	26.9	+0.8	21.6	27.5
Mar	+16.2	24.7	28.6	+0.9	22.5	26.3
Apr	+11.8	18.6	17.9	+0.7	17.6	20.3
May	+8.2	13.8	15.9	+0.2	11.7	15.4
Jun	+3.4	9.6	12.3	+0.3	4.7	14.9
Jul	+1.3	7.2	10.6	+0.3	3.1	12.7
Aug	+4.0	14.9	16.4	+0.9	7.1	21.4
Sep	+17.8	26.9	24.2	+1.7	24.9	27.0
Oct	+23.0	33.6	34.1	+1.4	31.4	41.0
Nov	+21.2	31.0	33.7	+0.5	28.1	34.5
DJF	+20.3	33.1	32.2	+0.6	26.4	31.3
MAM	+11.6	19.3	21.1	+0.3	16.4	20.9
JJA	+2.4	10.3	13.7	+0.4	4.2	16.0
SON	+20.2	30.3	31.6	+1.4	27.7	35.2

4.5.4 Quantile-based distributional fidelity

While the KGE decomposition evaluates co-located temporal skill, the quantile–quantile analysis (Fig. 10) asks a complementary, pooled-distribution question: do the corrected products reproduce the benchmark's full climatological distribution of monthly totals, including the wet tail on which hydrological and extreme-event applications depend? For I2-IIA the answer is affirmative to within measurement precision: in DJF, MAM and SON the model/benchmark quantile ratios are 0.98–0.99 simultaneously at the median, the 90th and the 99th percentile, i.e. the entire distribution, body and extremes alike is transmitted essentially unaltered. Its only departures occur in summer, where the near-zero median (ratio 1.36 on totals of a few millimeters) is numerically inflated by the dry denominator and the wet tail is mildly damped (P99 ratio 0.93), consistent with the smoothing of isolated convective cells. I3-IIA, by contrast, exhibits a characteristic S-shaped deviation in all seasons: the distribution body is deficient (P50 ratios 0.89–0.92) while the upper tail is inflated (P99 ratios 1.13–1.14), a redistribution of probability mass toward the extremes that becomes drastic in summer (P50 = 0.24; P99 = 1.47). Because the ratio-of-means corrector conserves the climatological mean by construction, this pattern is precisely the signature expected when a mean-only rescaling is applied to a variance-distorted field: the hurdle product's compressed body and its correction-amplified tail must compensate to preserve the mean. The Q–Q analysis thus provides the distribution-level counterpart of the variance-ratio (α) anomalies mapped in the previous section, and carries a practical implication: quantile-dependent applications, return-level estimation, drought and flood indices, threshold exceedances, can rely on I2-IIA across the full range of totals, whereas I3-IIA would require an additional distributional (e.g. quantile-mapping) adjustment before such use.

https://doi.org/10.5194/essd-2026-679
Preprint. Discussion started: 21 September 2026

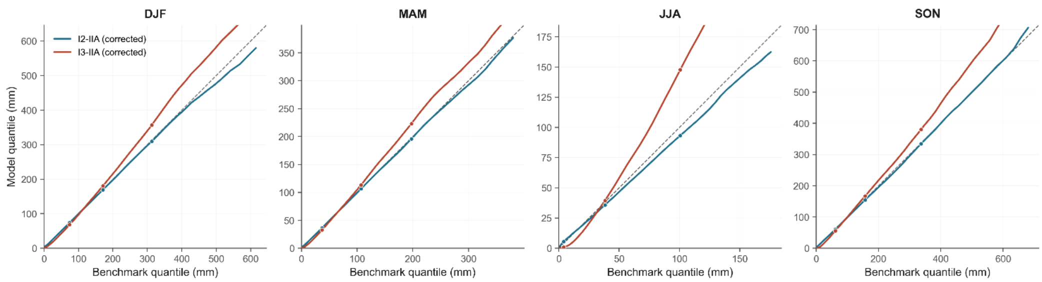

Figure 10: Seasonal quantile–quantile comparison of corrected per-cell monthly totals against the benchmark, pooled over all land cells and years (1951–2022). Dots mark the P50, P90 and P99 quantiles.

5. Conclusions

This study delivered PRECISi, a 72-year (1951–2022), 2 km daily gridded precipitation dataset for Sicily, produced by a doubly conditional geostatistical framework that models occurrence and intensity as separate, regime-stratified processes and merges three heterogeneous gauge networks into a single homogenized archive. Of the candidate formulations, these results confirm the central hypothesis: the regime-conditioned geostatistical framework matches or outperforms the covariate-based Regression-Kriging alternative, and binary occurrence masking proves superior to probabilistic hurdle weighting. Validation against an independent monthly climatological benchmark showed that the raw reconstruction carries the systematic underestimation inherent to smooth-support gauge interpolation (Beguería et al., 2016), negligible at the monthly scale but compounding to −103 mm yr^{-1}. A grid-cell-wise, calendar-month climatological correction, applied identically to every year so that interannual variability remains a falsifiable test rather than a fitted identity, removed this deficit: the corrected I2-IIA attains $R^2 = 0.99$ (monthly) and 0.98 (annual), temporally homogeneous residuals of ≈ ±10 mm over the full 72-year record, near-unit KGE bias and variability ratios per grid cell, and quantile ratios of 0.98–0.99 from the median to the 99th percentile in all wet seasons, distributional fidelity sufficient for return-level and threshold-based applications. The probabilistic hurdle alternative I3-IIA, while equally unbiased in the mean after correction, retains a structural variance distortion that mean scaling cannot repair, and is therefore recommended only with an additional distributional adjustment. Sensitivity analyses confirmed the temporal stationarity of the fitted spatial covariance structure and the robustness of the regime classification across all historically realized network densities, excluding tropicalization trends and gauge sparsity as confounders of the reconstruction. The main limitation remains the representation of localized summer convection, where sub-grid variability bounds the achievable skill of any gauge-based product. Its preservation of spatial gradients and wet/dry boundary structure makes the reference product directly suitable for distributed hydrological modelling, flash-flood risk assessment, and extreme-event analysis. PRECISi provides the first high-resolution, multi-decadal daily precipitation reference for Sicily, and its doubly conditional, benchmark-anchored design is directly transferable to other Mediterranean regions with fragmented observational archives.

Data and code availability statement

The DRPC data are freely accessible through AEGIS platform (https://www.protezionecivilesicilia.it:8443/aegis/map/map2d). The AdB and SIAS historical rainfall data that support the findings of this study are archived by the Autorità di Bacino (https://www.regione.sicilia.it/istituzioni/regione/strutture-regionali/presidenza-regione/autorita-bacino-distretto-idrografico-sicilia/) and the Servizio Informativo Agrometeorologico Siciliano (http://www.sias.regione.sicilia.it/); These data can be requested from the respective agencies or obtained from the corresponding author upon formal authorization. ERA5 reanalysis data are publicly available from the Copernicus Climate Data Store at https://cds.climate.copernicus.eu. The SRTM Version 3.0 DEM is publicly available from NASA. The reconstructed 1951–2022 daily gridded precipitation dataset generated in this study is available on Zenodo at https://doi.org/10.5281/zenodo.19228150 (BEIKAHMADI et al., 2026)(embargoed until publication) and is distributed under CC BY 4.0 licence. For review purposes, a private link can be provided by the corresponding author. All processing and analysis codes are publicly accessible in a modular form at the GitHub Repository: https://github.com/Niloufarbeikahmadi/High-Resolution-ATLAS-of-Observational-Precipitation-Reconstruction.

Author contributions

NB and LVN jointly conceptualized the structure of the study. NB and DT collected the data, while NB processed them, implemented the methodology and analysis code, performed the validation analyses, and prepared the original manuscript draft. AF and DT contributed to the discussion of results. AF collaborated on the editing of the manuscript and supervised and supported all phases of the work. All authors reviewed and approved the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

The authors are grateful to the Dipartimento della Protezione Civile–Sicilia (DRPC), the Autorità di Bacino (AdB), and the Servizio Informativo Agrometeorologico Siciliano (SIAS) for providing the rain gauge data used in this study. This work was carried out within the PhD program in Chemical, Environmental, Biomedical, Hydraulic and Materials Engineering at the University of Palermo.

References

Aalto, J., Pirinen, P., Heikkinen, J., and Venäläinen, A.: Spatial interpolation of monthly climate data for Finland: comparing the performance of kriging and generalized additive models, Theor. Appl. Climatol., 112, 99–111, https://doi.org/10.1007/s00704-012-0716-9, 2013.

Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Tank, A. M. G. K., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Kumar, K. R., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D. B., Burn, J., Aguilar, E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., and Vazquez-Aguirre, J. L.: Global observed changes in daily climate extremes of temperature and precipitation, J. Geophys. Res. Atmos., 111, https://doi.org/10.1029/2005JD006290, 2006.

Allcroft, D. J. and Glasbey, C. A.: A Latent Gaussian Markov Random-Field Model for Spatiotemporal Rainfall Disaggregation, J. R. Stat. Soc. Ser. C Appl. Stat., 52, 487–498, https://doi.org/10.1111/1467-9876.00419, 2003.

Alpert, P., Ben-Gai, T., Baharad, A., Benjamini, Y., Yekutieli, D., Colacino, M., Diodato, L., Ramis, C., Homar, V., Romero, R., Michaelides, S., and Manes, A.: The paradoxical increase of Mediterranean extreme daily rainfall in spite of decrease in total values, Geophys. Res. Lett., 29, 31-1-31–4, https://doi.org/10.1029/2001GL013554, 2002.

Arnone, E., Pumo, D., Viola, F., Noto, L. V., and La Loggia, G.: Rainfall statistics changes in Sicily, Hydrol. Earth Syst. Sci., 17, 2449–2458, https://doi.org/10.5194/hess-17-2449-2013, 2013.

Barancourt, C., Creutin, J. D., and Rivoirard, J.: A method for delineating and estimating rainfall fields, Water Resour. Res., 28, 1133–1144, https://doi.org/10.1029/91WR02896, 1992.

Bárdossy, A. and Pegram, G.: Infilling missing precipitation records – A comparison of a new copula-based method with other techniques, J. Hydrol., 519, 1162–1170, https://doi.org/10.1016/j.jhydrol.2014.08.025, 2014.

Beguería, S., Vicente-Serrano, S. M., Tomás-Burguera, M., and Maneta, M.: Bias in the variance of gridded data sets leads to misleading conclusions about changes in climate variability, Int. J. Climatol., 36, 3413–3422, https://doi.org/10.1002/joc.4561, 2016.

Beikahmadi, N. and Rahimzadegan, M.: Evaluating IMERG Precision: Comparative Study of Daily and Six-Hourly Scales in Diverse Climate Zones of Iran, Atmos. Ocean. Opt., 36, 256–271, https://doi.org/10.1134/S1024856023030223, 2023.

Beikahmadi, N., Francipane, A., and Noto, L. V.: Smart Data Blending Framework to Enhance Precipitation Estimation through Interconnected Atmospheric, Satellite, and Surface Variables, Hydrology, 10, 128, https://doi.org/10.3390/hydrology10060128, 2023.

BEIKAHMADI, N., NOTO, L. V., Francipane, A., and Treppiedi, D.: PRECISi: A High-Resolution Daily Gridded Precipitation Dataset for Sicily (1951–2022) Derived from a Doubly Conditional Geostatistical Framework, https://doi.org/10.5281/zenodo.19228150, 2026.

Bowman, A. W. and Crujeiras, R. M.: Inference for variograms, Comput. Stat. Data Anal., 66, 19–31, https://doi.org/10.1016/j.csda.2013.02.027, 2013.

Caccamo, M. T., Castorina, G., Colombo, F., Insinga, V., Maiorana, E., and Magazù, S.: Weather forecast performances for complex orographic areas: Impact of different grid resolutions and of geographic data on heavy rainfall event simulations in Sicily, Atmos. Res., 198, 22–33, https://doi.org/10.1016/j.atmosres.2017.07.028, 2017.

Caracciolo, D., Francipane, A., Viola, F., Noto, L. V., and Deidda, R.: Performances of GPM satellite precipitation over the two major Mediterranean islands, Atmos. Res., 213, 309–322, https://doi.org/10.1016/j.atmosres.2018.06.010, 2018.

Chen, M., Xie, P., Janowiak, J. E., and Arkin, P. A.: Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations, J. Hydrometeorol., 3, 249–266, https://doi.org/10.1175/1525-7541(2002)003%3C0249:GLPAYM%3E2.0.CO;2, 2002.

Chilès, J.-P. and Delfiner, P.: Geostatistics: Modeling Spatial Uncertainty, John Wiley & Sons, 750 pp., 2012.

Chimani, B., Matulla, C., Böhm, R., and Hofstätter, M.: A new high resolution absolute temperature grid for the Greater Alpine Region back to 1780, Int. J. Climatol., 33, 2129–2141, https://doi.org/10.1002/joc.3574, 2013.

Cipolla, G., Francipane, A., and Noto, L. V.: Classification of Extreme Rainfall for a Mediterranean Region by Means of Atmospheric Circulation Patterns and Reanalysis Data, Water Resour. Manag., 34, 3219–3235, https://doi.org/10.1007/s11269-020-02609-1, 2020.

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J. M., and Jones, P. D.: An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets, J. Geophys. Res. Atmos., 123, 9391–9409, https://doi.org/10.1029/2017JD028200, 2018.

Cragg, J. G.: Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods, Econometrica, 39, 829–844, https://doi.org/10.2307/1909582, 1971.

Crespi, A., Matiu, M., Bertoldi, G., Petitta, M., and Zebisch, M.: A high-resolution gridded dataset of daily temperature and precipitation records (1980–2018) for Trentino-South Tyrol (north-eastern Italian Alps), Earth Syst. Sci. Data, 13, 2801–2818, https://doi.org/10.5194/essd-13-2801-2021, 2021.

Cressie, N.: Statistics for Spatial Data, John Wiley & Sons, 931 pp., 2015.

Cristiano, E., ten Veldhuis, M.-C., and van de Giesen, N.: Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review, Hydrol. Earth Syst. Sci., 21, 3859–3878, https://doi.org/10.5194/hess-21-3859-2017, 2017.

Di Piazza, A., Conti, F. L., Noto, L. V., Viola, F., and La Loggia, G.: Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy, Int. J. Appl. Earth Obs. Geoinformation, 13, 396–408, https://doi.org/10.1016/j.jag.2011.01.005, 2011.

Dowd, P. A.: The Variogram and Kriging: Robust and Resistant Estimators, in: Geostatistics for Natural Resources Characterization, edited by: Verly, G., David, M., Journel, A. G., and Marechal, A., Springer Netherlands, Dordrecht, 91–106, https://doi.org/10.1007/978-94-009-3699-7_6, 1984.

Farr, T. G. and Kobrick, M.: Shuttle radar topography mission produces a wealth of data, Eos Trans. Am. Geophys. Union, 81, 583–585, https://doi.org/10.1029/EO081i048p00583, 2000.

Francipane, A., Pumo, D., Sinagra, M., La Loggia, G., and Noto, L. V.: A paradigm of extreme rainfall pluvial floods in complex urban areas: the flood event of 15 July 2020 in Palermo (Italy), Nat. Hazards Earth Syst. Sci., 21, 2563–2580, https://doi.org/10.5194/nhess-21-2563-2021, 2021.

Francipane, A., Arnone, E., and Noto, L. V.: Combining a data-driven approach with seasonal forecast data to predict reservoir water volume in the Mediterranean area, Hydrol. Sci. J., 68, 764–781, https://doi.org/10.1080/02626667.2023.2189521, 2023.

Giorgi, F.: Climate change hot-spots, Geophys. Res. Lett., 33, https://doi.org/10.1029/2006GL025734, 2006.

Goovaerts, P.: Geostatistics for Natural Resources Evaluation, Oxford University Press, 502 pp., 1997.

Gupta, H. V. and Kling, H.: On typical range, sensitivity, and normalization of Mean Squared Error and Nash-Sutcliffe Efficiency type metrics, Water Resour. Res., 47, https://doi.org/10.1029/2011WR010962, 2011.

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, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Guttman, A.: R-trees: a dynamic index structure for spatial searching, in: Proceedings of the 1984 ACM SIGMOD international conference on Management of data - SIGMOD '84, the 1984 ACM SIGMOD international conference, 47, https://doi.org/10.1145/602259.602266, 1984.

Haberlandt, U.: Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event, J. Hydrol., 332, 144–157, https://doi.org/10.1016/j.jhydrol.2006.06.028, 2007.

Haklay, M. and Weber, P.: OpenStreetMap: User-Generated Street Maps, IEEE Pervasive Comput., 7, 12–18, https://doi.org/10.1109/MPRV.2008.80, 2008.

Hastie, T. and Tibshirani, R.: Generalized Additive Models, Chapman & Hall/CRC, New York, NY, USA, 335 pp., https://doi.org/10.1214/ss/1177013604, 1986.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006, J. Geophys. Res. Atmos., 113, https://doi.org/10.1029/2008JD010201, 2008.

Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Orsini, J. A. M., Nicholls, N., Penner, J. E., Stott, P. A., Allen, M., Ammann, C., Andronova, N., Betts, R. A., Clement, A., Collins, W. D., Crooks, S., Delworth, T. L., Forest, C., Forster, P., Goosse, H., Gregory, J. M., Harvey, D., Jones, G. S., Joos, F., Kenyon, J., Kettleborough, J., Kharin, V., Knutti, R., Lambert, F. H., Lavine, M., Lee, T. C. K., Levinson, D., Masson-Delmotte, V., Nozawa, T., Otto-Bliesner, B., Pierce, D., Power, S., Rind, D., Rotstayn, L., Santer, B. D., Senior, C., Sexton, D., Stark, S., Stone, D. A., Tett, S., Thorne, P., Dorland, R. V., Wang, M., Wielicki, B., Xu, L., Zhang, X., Zorita, E., Hegerl, G. C., Zwiers, F. W., Karoly, D. J., Ogallo, L., and Planton, S.: Understanding and Attributing Climate Change, in: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp., 2007.

Hengl, T.: A Practical Guide to Geostatistical Mapping of Environmental Variables, Office for Official Publications of the European Communities, Luxembourg, 143 pp., 2007.

Herrera, S., Cardoso, R. M., Soares, P. M., Espírito-Santo, F., Viterbo, P., and Gutiérrez, J. M.: Iberia01: a new gridded dataset of daily precipitation and temperatures over Iberia, Earth Syst. Sci. Data, 11, 1947–1956, https://doi.org/10.5194/essd-11-1947-2019, 2019.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Q. J. R. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hollis, D., McCarthy, M., Kendon, M., Legg, T., and Simpson, I.: HadUK-Grid—A new UK dataset of gridded climate observations, Geosci. Data J., 6, 151–159, https://doi.org/10.1002/gdj3.78, 2019.

Huffman, G. J., Adler, R. F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., Janowiak, J., McNab, A., Rudolf, B., and Schneider, U.: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset, Bull. Am. Meteorol. Soc., 78, 5–20, https://doi.org/10.1175/1520-0477(1997)078%3C0005:TGPCPG%3E2.0.CO;2, 1997.

Hutchinson, M. F. and Gessler, P. E.: Splines — more than just a smooth interpolator, Geoderma, 62, 45–67, https://doi.org/10.1016/0016-7061(94)90027-2, 1994.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, Int. J. Climatol., 34, 1657–1675, https://doi.org/10.1002/joc.3794, 2014.

Isotta, F. A., Begert, M., and Frei, C.: Long-Term Consistent Monthly Temperature and Precipitation Grid Data Sets for Switzerland Over the Past 150 Years, J. Geophys. Res. Atmos., 124, 3783–3799, https://doi.org/10.1029/2018JD029910, 2019.

Jones, P. D. and Hulme, M.: Calculating Regional Climatic Time Series for Temperature and Precipitation: Methods and Illustrations, Int. J. Climatol., 16, 361–377, https://doi.org/10.1002/(SICI)1097-0088(199604)16:4%3C361::AID-JOC53%3E3.0.CO;2-F, 1996.

Journel, A. G.: Nonparametric estimation of spatial distributions, J. Int. Assoc. Math. Geol., 15, 445–468, https://doi.org/10.1007/BF01031292, 1983.

Kitanidis, P. K.: Introduction to Geostatistics: Applications in Hydrogeology, Cambridge University Press, 276 pp., 1997.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, J. Hydrol., 424–425, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

Krähenmann, S., Walter, A., Brienen, S., Imbery, F., and Matzarakis, A.: High-resolution grids of hourly meteorological variables for Germany, Theor. Appl. Climatol., 131, 899–926, https://doi.org/10.1007/s00704-016-2003-7, 2018.

Li, J. and Heap, A. D.: A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors, Ecol. Inform., 6, 228–241, https://doi.org/10.1016/j.ecoinf.2010.12.003, 2011.

Long, Y., Zhang, Y., and Ma, Q.: A merging framework for rainfall estimation at high spatiotemporal resolution for distributed hydrological modeling in a data-scarce area, Remote Sens., 8, 599, https://doi.org/10.3390/rs8070599, 2016.

Lussana, C., Saloranta, T., Skaugen, T., Magnusson, J., Tveito, O. E., and Andersen, J.: seNorge2 daily precipitation, an observational gridded dataset over Norway from 1957 to the present day, Earth Syst. Sci. Data, 10, 235–249, https://doi.org/10.5194/essd-10-235-2018, 2018.

Marra, F., Armon, M., and Morin, E.: Coastal and orographic effects on extreme precipitation revealed by weather radar observations, Hydrol. Earth Syst. Sci., 26, 1439–1458, https://doi.org/10.5194/hess-26-1439-2022, 2022.

Matheron, G.: Principles of geostatistics, Econ. Geol., 58, 1246–1266, https://doi.org/10.2113/gsecongeo.58.8.1246, 1963.

Matheron, G.: The Theory of Regionalized Variables and Its Applications, École national supérieure des mines, 228 pp., 1971.

Mazon, J. and Pino, D.: The Influence of an Increase of the Mediterranean Sea Surface Temperature on Two Nocturnal Offshore Rainbands: A Numerical Experiment, Atmosphere, 8, 58, https://doi.org/10.3390/atmos8030058, 2017.

Nosratpour, R., Rahimzadegan, M., and Beikahmadi, N.: Introducing a merged precipitation satellite model using satellite precipitation products, land surface temperature, and precipitable water vapor, Geocarto Int., 37, 11782–11812, https://doi.org/10.1080/10106049.2022.2060324, 2022.

Odeh, I. O. A., McBratney, A. B., and Chittleborough, D. J.: Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging, Geoderma, 67, 215–226, https://doi.org/10.1016/0016-7061(95)00007-B, 1995.

Pecorino, V., Matteo, T. D., Milazzo, M., Pasotti, L., Pluchino, A., and Rapisarda, A.: Evidence of climate change in Sicily by means of rainfall data analysis, https://doi.org/10.21203/rs.3.rs-3920420/v1, 2024.

Pegram, G. G. S. and Clothier, A. N.: High resolution space–time modelling of rainfall: the "String of Beads" model, J. Hydrol., 241, 26–41, https://doi.org/10.1016/S0022-1694(00)00373-5, 2001.

Refsgaard, J. C.: Parameterisation, calibration and validation of distributed hydrological models, J. Hydrol., 198, 69–97, https://doi.org/10.1016/S0022-1694(96)03329-X, 1997.

Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., and Dormann, C. F.: Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure, Ecography, 40, 913–929, https://doi.org/10.1111/ecog.02881, 2017.

Sekulić, A., Kilibarda, M., Protić, D., and Bajat, B.: A high-resolution daily gridded meteorological dataset for Serbia made by Random Forest Spatial Interpolation, Sci. Data, 8, 123, https://doi.org/10.1038/s41597-021-00901-2, 2021.

Serrano-Notivoli, R. and Tejedor, E.: From rain to data: A review of the creation of monthly and daily station-based gridded precipitation datasets, WIREs Water, 8, e1555, https://doi.org/10.1002/wat2.1555, 2021.

Serrano-Notivoli, R., Beguería, S., Saz, M. Á., Longares, L. A., and de Luis, M.: SPREAD: a high-resolution daily gridded precipitation dataset for Spain – an extreme events frequency and intensity overview, Earth Syst. Sci. Data, 9, 721–738, https://doi.org/10.5194/essd-9-721-2017, 2017.

Serrano-Notivoli, R., Saz, M. Á., Longares, L. A., and de Luis, M.: SiCLIMA: High-resolution hydroclimate and temperature dataset for Aragón (northeast Spain), Data Brief, 56, 110876, https://doi.org/10.1016/j.dib.2024.110876, 2024.

Servén, D., Brummitt, C., Király, F., JodesL, Schmitz, D., Carmichael, Z., Ouslan, A., Miccoli, F., IlyesD, Taylor, J., Egan, J., Lapenta, L., Marín, L. S., Mo, S., Ranjan, S., Rose, S. A., Kraxner, T., Fasci, U., Szépe, V., Haddar, Y. E., hlink, Abedi, H., and nicoleta-kyo: pyGAM: Generalized Additive Models in Python, , https://doi.org/10.5281/zenodo.17981758, 2025.

Shepard, D.: A two-dimensional interpolation function for irregularly-spaced data, in: Proceedings of the 1968 23rd ACM national conference on -, the 1968 23rd ACM national conference, 517–524, https://doi.org/10.1145/800186.810616, 1968.

Singh, V. P.: Elementary Hydrology, Prentice Hall, 1000 pp., 1992.

Škrk, N., Serrano-Notivoli, R., Čufar, K., Merela, M., Črepinšek, Z., Kajfež Bogataj, L., and de Luis, M.: SLOCLIM: a high-resolution daily gridded precipitation and temperature dataset for Slovenia, Earth Syst. Sci. Data, 13, 3577–3592, https://doi.org/10.5194/essd-13-3577-2021, 2021.

Snow, A. D., Cochran, M., Hoese, D., Miara, I., Bossche, J. V. den, Mayo, C., Lucas, G., Cochrane, P., Kloe, J. de, Karney, C., Shaw, J. J., Anh, T. Q., Filipe, Ouzounoudis, G., Couwenberg, B., Lostis, G., Dearing, J., Jurd, B., Gohlke, C., Schneck, C., McDonald, D., Taves, M., Itkin, M., May, R., Stewart, A. J., Bittencourt, H. de, Little, B., Hugonnet, R., and Rahul, P. S.: pyproj: Python interface to PROJ, , https://doi.org/10.5281/zenodo.16874869, 2025.

Sottile, G., Francipane, A., Adelfio, G., and Noto, L. V.: A PCA-based clustering algorithm for the identification of stratiform and convective precipitation at the event scale: an application to the sub-hourly precipitation of Sicily, Italy, Stoch. Environ. Res. Risk Assess., 36, 2303–2317, https://doi.org/10.1007/s00477-021-02028-7, 2022.

Stein, M. L.: Interpolation of Spatial Data: Some Theory for Kriging, Springer Science & Business Media, 280 pp., 1999.

Thiessen, A. H.: PRECIPITATION AVERAGES FOR LARGE AREAS, Mon. Weather Rev., 39, 1082–1089, https://doi.org/10.1175/1520-0493(1911)39%3C1082b:PAFLA%3E2.0.CO;2, 1911.

Treppiedi, D., Cipolla, G., Francipane, A., and Noto, L. v.: Detecting precipitation trend using a multiscale approach based on quantile regression over a Mediterranean area, Int. J. Climatol., 41, 5938–5955, https://doi.org/10.1002/joc.7161, 2021.

Treppiedi, D., Cipolla, G., Francipane, A., Cannarozzo, M., and Noto, L. V.: Investigating the Reliability of Stationary Design Rainfall in a Mediterranean Region under a Changing Climate, Water, 15, 2245, https://doi.org/10.3390/w15122245, 2023.

Treppiedi, D., Francipane, A., and Noto, L. V.: Projecting Depth-Duration-Frequency Curves for Future Climate: a Case Study in the Mediterranean Area, Water Resour. Manag., 39, 4409–4427, https://doi.org/10.1007/s11269-025-04162-1, 2025.

Treppiedi, D., Villarini, G., Mattina, C., Francipane, A., and Noto, L. V.: Unveiling Changes in the Seasonality of Extreme Precipitation Shows an Anticipation of Short-Duration Extremes, Earth Syst. Environ., https://doi.org/10.1007/s41748-026-01178-4, 2026.

Trigo, I. F., Bigg, G. R., and Davies, T. D.: Climatology of Cyclogenesis Mechanisms in the Mediterranean, Mon. Weather Rev., 130, 549–569, https://doi.org/10.1175/1520-0493(2002)130%3C0549:COCMIT%3E2.0.CO;2, 2002.

Varotsos, K. V., Katavoutas, G., Kitsara, G., Karali, A., Lemesios, I., Patlakas, P., Hatzaki, M., Tenentes, V., Sarantopoulos, A., Psiloglou, B., Koutroulis, A. G., Grillakis, M. G., and Giannakopoulos, C.: CLIMADAT-GRid: a high-resolution daily gridded precipitation and temperature dataset for Greece, Earth Syst. Sci. Data, 17, 4455–4477, https://doi.org/10.5194/essd-17-4455-2025, 2025.

Webster, R. and Oliver, M. A.: Geostatistics for Environmental Scientists, John Wiley & Sons, 333 pp., 2007.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Academic Press, 697 pp., 2011.

Wood, S. N.: Generalized Additive Models: An Introduction with R, Second Edition, 2nd ed., Chapman and Hall/CRC, New York, 496 pp., https://doi.org/10.1201/9781315370279, 2017.