Bridging the scale gap: enhancing point-scale rainfall estimates by post-processing ERA5

Technical Memo

ECMWF
European Centre for Medium-Range Weather Forecasts

933

Bridging the scale gap: enhancing point-scale rainfall estimates by post-processing ERA5

Fatima M. Pillosu[1,2], Timothy D. Hewson[2], Estibaliz Gascòn[2], Milana Vuckovic[2], Christel Prudhomme[2,3,4], Hannah L. Cloke[1,5]

[1] Department of Geography and Environmental Science, University of Reading, Reading, UK
[2] Forecasts and Services Department, European Centre for Medium-range Weather Forecasts, Reading, UK
[3] Department of Geography and Environment, University of Loughborough, Loughborough, UK
[4] UK Centre for Ecology and Hydrology, Wallingford, UK
[5] Department of Meteorology, University of Reading, Reading, UK

October 2025

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/publications/

Contact: library@ecmwf.int

Abstract

Accurately estimating rainfall distributions, from small to extreme totals, is crucial for addressing various environmental challenges, including flood forecasting, water resource management, and disaster preparedness. Global Numerical Weather Prediction (NWP) models can provide useful rainfall estimates; yet, they often misrepresent point-scale observations from rain gauges, underestimating the frequency of small rainfall totals and underestimating extreme values. This study provides a systematic, global verification of four NWP-modelled rainfall datasets with different resolutions - ERA5's Ensemble Data Assimilation (62 km, probabilistic), ERA5's short-range forecasts (31 km, deterministic), short-range ECMWF reforecasts for cycle 46r1 (18 km, control run), and ERA5-ecPoint (point-scale, probabilistic) - against 20 years of point-rainfall observations from rain gauges around the world. The models' ability to represent the entire rainfall distribution, including extreme rainfall, was assessed. Overall, the higher spatial resolution of NWP models enables a more accurate representation of gauge-based climatologies. Nonetheless, ERA5-ecPoint provides the most accurate representation, capturing the frequency of zeros, the growth rates of rainfall totals, and the wet tails more accurately. Moreover, due to its probabilistic nature, ERA5-ecPoint can estimate long return periods (e.g., 1000 years and more), offering insights into extremely rare or unprecedented events at specific locations. The model significantly improves performance in flat, hilly/mountainous regions. In very mountainous areas (e.g., the Andes), it underestimates zero rainfall totals and overestimates the length of the wet tails. These findings underscore the importance of using post-processing to enhance the local-scale validity of global NWP models. Moreover, as climate change intensifies extreme rainfall events, these findings are crucial for estimating accurate long-period rainfall climatologies, as needed for effective mitigation and resilience building, particularly in areas lacking comprehensive and reliable rain gauge records.

Plain language summary

Historical rainfall datasets, modelled with global weather prediction models, enable us to manage water supplies and prepare for water-related disasters seamlessly around the world. Yet, these models often struggle to match measurements taken by rain gauges at specific locations. They may overestimate the small rainfall totals (including the zeros) whilst underestimating the big rainfall totals that might cause disastrous (flash) floods. This research compared four different weather model systems against twenty years of rain gauge (point-scale) measurements from around the world. Three models had spatial resolutions ranging from very coarse (62 km, ERA5-EDA) to much finer (31 km, from ERA5, and 18 km, from reforecasts). The fourth model post-processed ERA5 rainfall estimates using the post-processing technique ecPoint (ERA5-ecPoint) to obtain point-scale rainfall estimates that closely mirror rain gauge measurements. The study examined how well each model captured the distribution of point-scale rainfall measurements, including extremes. This study demonstrates that while finer resolution generally improves accuracy, ERA5-ecPoint represents the distribution of point-rainfall estimates better, including extremes. This system can estimate extremely rare events (e.g., 1 in 1000 years), providing crucial information for infrastructure planning and risk assessment. The model works particularly well in flat and hilly/mountainous regions. It may encounter difficulties in very mountainous areas such as the Andes, where it underestimates small rainfall totals and overestimates extreme rainfall. These results highlight that even as weather models improve, additional processing remains essential for producing reliable local rainfall estimates, which are particularly vital as climate change intensifies extreme rainfall events worldwide.

1 Introduction

Accurately estimating the full range of past and future rainfall distributions, from light to extreme totals, is one of the biggest challenges in modern meteorology. Yet, it is essential to address a range of critical

issues. In flood forecasting, accurately estimating the spatial distribution of small and extreme rainfall totals influences the catchment response to the rainfall event, impacting runoff generation and streamflow patterns (Cuo et al., 2011; Wang and Karimi, 2022). In water resource management, understanding the full rainfall distribution informs the management of reservoirs for flood control, power generation, and irrigation purposes (Tie et al., 2023). It also helps in agricultural applications such as crop selection and planting schedules (Janmohammadi and Sabaghnia, 2023; Maurya et al., 2024). It also supports urban planning by helping to design effective urban drainage systems and manage storm-water runoff (Hossain et al., 2024; Laouacheria et al., 2019). Analysing changes over time in the entire rainfall distribution provides insights into climate change impacts such as shifts in the frequency and intensity of extreme rainfall events (Tye et al., 2022), droughts' characteristics (Haile et al., 2020), biodiversity and ecosystems stability (Lamprecht et al., 2021), and food security (Balasundram et al., 2023). Extreme rainfall, in particular, has received significant attention in recent literature (Gimeno et al., 2022; Schumacher, 2017) due to its catastrophic impacts for society, infrastructure, and the environment (IPCC, 2023; WMO, 2024). It not only reduces worldwide macroeconomic growth rates and slows global economic rise (Liang, 2022), but also can cause long-term anxiety and post-traumatic stress on affected communities, hindering recovery efforts (Doocy et al., 2013). With climate change expected to intensify both the frequency and severity of extreme rainfall, even in regions where average precipitation is decreasing (Asadieh and Krakauer, 2015; Westra et al., 2014; Zittis et al., 2021), understanding its past and anticipating future trends is crucial to inform disaster preparedness and response efforts.

Precipitation time series can be obtained from various sources. Rain gauges are a primary source of ground truth. They provide highly accurate direct point-scale precipitation measurements when properly maintained and calibrated (Lanza and Stagi, 2008). In regions with dense networks, rain gauges offer a good spatial representation of localised extremes (Haiden and Duffy, 2016). Moreover, stations have been operating for decades in some locations, providing high-quality long-term historical records for trend analysis (Anand and Karunanidhi, 2020; Tadeyo et al., 2020). Rain gauge coverage is notably spatially and temporally uneven, leaving many regions unmonitored (Kidd et al., 2017). In areas with complex topography or low-density networks, gauges may fail to represent the rainfall's spatial variability (Di Curzio et al., 2022). Inadequate rain gauge maintenance can also lead to data gaps or inaccuracies (Lanza and Stagi, 2008). Satellite- and radar-derived gridded datasets provide broader spatial and temporal coverage, particularly in ungauged regions (Herold et al., 2017). Their rainfall estimates may, however, differ from rain gauge measurements, especially extremes which might be severely underestimated and mislocated (Ensor and Robeson, 2008; Gupta et al., 2020; Satgé et al., 2020). Numerical Weather Prediction (NWP) models, such as reanalyses and reforecasts, offer spatially and temporally consistent precipitation datasets with global, multi-decadal coverage. Reanalyses, like ERA5 and its Ensemble Data Assimilation (EDA) component (Hersbach et al., 2020) or NCEP/NCAR Reanalysis (Hamill et al., 2022; Kalnay et al., 1996), integrate historical weather observations with a state-of-the-art NWP model to produce high-resolution precipitation datasets. Reforecasts, such as NCEP's Global Ensemble Forecast System (Hamill et al., 2006) and ECMWF's Integrated Forecast System (Richardson et al., 2014), provide 20-30 years of retrospective forecasts generated with current operational NWP models. Reanalyses capture rainfall's spatial patterns and temporal trends (Lavers et al., 2022) but tend to underestimate extreme precipitation due to their coarse resolution of about 50 or 30 km (Alexandridis et al., 2023; Donat et al., 2016; Espinosa et al., 2024; Gomis-Cebolla et al., 2023)[1]. Reforecasts also capture rainfall's spatial patterns and temporal trends, but still underestimate extreme precipitation even though their resolution is half (i.e. 18 km) (Hewson, 2024)[2]. Statistical post-processing methods can enhance the

[1]Note that the studies comparing both ERA5 and ERA5-Land against rain gauge observations are considered in this study only for their analysis of ERA5. These are somewhat flawed as ERA5-Land simply re-grids, without any statistical or dynamical downscaling, the precipitation in ERA5 onto ERA5-Land's grid (Muñoz-Sabater et al., 2021).

[2]Nowadays, we have reforecasts at 9 km resolution, but Hewson (2024) showed that extremes do not get much bigger in the

local-scale representation of rainfall (Giorgos et al., 2024), but their effectiveness commonly depends on the availability of high-quality observations, leading to a patchy geographical coverage of post-processed reanalysis/reforecasts (Vannitsem et al., 2021). The post-processing method proposed by Hewson and Pillosu (2021), called ecPoint, improves the local-scale representation of NWP model outputs globally, particularly for extremes, without requiring high-density observational networks, using a non-local calibration strategy. The ecPoint approach was applied to ERA5 for rainfall and temperature through the Highlander project (Hewson et al., 2023; Bottazzi et al., 2024).

The primary aim of this study is to assess the fitness-for-purpose of the ERA5-ecPoint dataset by comparing its representation of point rainfall climatologies around the world against the rain gauge-based equivalent. A secondary goal is to evaluate the impact of spatial resolution on the representation of point rainfall climatologies from three additional datasets: ERA5's Ensemble Data Assimilation (EDA, 62 km resolution), ERA5's short-range forecasts (31 km), and ECMWF 46r1 reforecasts (18 km). Two research questions are, therefore, examined. How do NWP models represent the overall distribution of point-rainfall observations (RQ1)? How do NWP models represent, in particular, extreme rainfall (RQ2)? With this goal, one could develop a climatological analysis of extreme rainfall trends over long periods (+80 years) or put extreme rainfall or flooding events into a climatological context. The study is organised as follows. Section 2 describes the rain gauge observations and the NWP models used in this study. Section 3 describes the methods adopted to answer the research questions. Section 4 presents the results from the objective verification and a case study, while Section 5 discusses them. Final remarks are drawn in Section 6.

2 Data

2.1 Point-scale rain gauge precipitation observations

This study considered 24-hourly precipitation from surface synoptic observations (SYNOP) from the Global Telecommunication System (GTS) network and additional gauge data stored internally at ECMWF. SYNOP observations consist of standardised, historical and near-real-time meteorological reports that ensure data quality and format consistency across diverse regions. High-density national rain gauge networks (primarily from European countries and available internally at ECMWF) were also integrated into the analysis (Haiden and Duffy, 2016). The rain gauge rainfall observations underwent manual quality control to remove erroneously high rainfall totals that would have disproportionately affected the upper tails of the point-scale rainfall distributions. The rainfall timeseries were inspected for anomalous spikes, outliers, and odd constant values inconsistent with station and regional climatologies. Flagged values were cross-checked against nearby stations and through the independent CPC Global Unified Gauge-Based Analysis of daily rainfall dataset (gridded, at 50 km spatial resolution)[3], and removed if confirmed to be erroneous. If not corrected, the erroneous high point-rainfall totals could have significantly distorted the upper tails of the observed precipitation distributions, leading to inaccurate results. Rain gauge observations stored at ECMWF have increased considerably since the 2000s. Thus, we consider a 20-year verification period between the 1st of January 2000 to the 31st of December 2019. Only observations ending at 00 UTC were considered, for a total of 7300 sets of daily precipitation realisations within the 20-year verification period (Table 1, row 1). Many rain gauge stations had missing data. To ensure that the timeseries were representative of the considered 20-year period, only sites with at least 75% of valid recordings were considered, which reduced the number of sites in the database from 28834

9 km reforecasts than in the previous lower resolution reforecasts at 18 km resolution.

[3]https://psl.noaa.gov/data/gridded/data.cpc.globalprecip.html

to 4546.

2.2 Gridded NWP-modelled precipitation estimates

2.2.1 ERA5 Reanalysis (ERA5) and ERA5 Ensemble Data Assimilation (ERA5-EDA)

ERA5 is the fifth generation of atmospheric reanalysis produced by the Copernicus Climate Change Service (C3S) run by ECMWF (Hersbach et al., 2020). Compared to its predecessor, ERA-Interim, ERA5 offers high spatial (31 km) and temporal (hourly) resolution and extended temporal coverage from 1940 to near-real time. ERA5 assimilates a diverse range of observational data from satellites, weather balloons, aircraft, and ground stations, employing a 4D-Var assimilation system. This system not only improves the accuracy of the data by adjusting it in four dimensions but also enhances the continuity and stability of the climatological records. No precipitation observations are assimilated into ERA5 (Hersbach et al., 2020).

The ERA5 Ensemble Data Assimilation (EDA) system enhances the robustness of the ERA5 reanalysis by generating multiple simulations with slightly varied initial conditions (Hersbach et al., 2020). Each ensemble member in ERA5 EDA provides an equally probable realisation of the atmospheric state, quantifying the uncertainty associated with observational errors and limitations within the forecasting model itself. ERA5-EDA has 10 ensemble members, running at 62 km spatial resolution and 3-hour temporal resolution.

To match the rain gauge observations, ERA5 and ERA5-EDA data between the 1st of January 2000 and the 31st of December 2019 were extracted, and only 24-hourly precipitation ending at 00 UTC was considered. Hence, ERA5 precipitation distribution was built with 7300 realisations, while ERA5-EDA distribution, considering the 10 ensemble members as equally probable precipitation realisations, was constructed with 73000 realisations (Table 1, rows 2 and 3).

2.2.2 ECMWF Reforecasts

Reforecasts are retrospective weather forecasts generated with a fixed NWP model version. The reforecast uniformity (i.e., with no discrepancies caused by historical changes in model configurations) ensures that differences in climatological patterns are attributable to actual atmospheric variations rather than artefacts of evolving model technologies. To match the temporal span of the precipitation observations as closely as possible, reforecasts from the ECMWF's IFS 46r1 cycle were considered - since 46r1 run operationally from June 2019 to June 2020, the reforecasts span from the 1st of July 1999 to the 30th of June 2019. 46r1 reforecasts are provided at 18 km spatial resolution, and are produced only on Mondays and Thursdays. They consist of an ensemble of one control run and 10 perturbed members, produced at 00 UTC with a 6-hourly resolution up to t+1104 (day 46). The control and the perturbed members' model configurations (e.g., resolution, parametrisations) are the same. However, the control run uses the best estimate of the initial conditions (i.e., the operational analysis), and it has been shown to have a different precipitation climatology than the perturbed members. Hence, in this study, only the control run was used. Since reforecasts have fewer realisations per year (as they are produced only on Mondays and Thursdays), we increased the precipitation realisations by considering lead times up to day 10 as equally probable precipitation realisations. This was possible as there was no drift in the forecasts up to day 10 (not shown). Hence, the precipitation distribution built with ECMWF reforecasts contains 20800 realisations (Table 1, row 4).

No. Column	1	2	3	4	5	6	7	8
No. Row	Type of climatology	Dataset to compute climatology	Dataset description	Spatial coverage & resolution at equator	Temporal Coverage	No. of independent realizations (daily & total in the 20-year period between 2000 and 2019)	Max return period that can be computed in the 20-year period (Max return period actually computed)	Max percentile (%) computed $\frac{1}{100-(\frac{1}{\text{no.total real.}} \times 100)}$
1	Observational, for points	SYNOP + high-resolution national datasets	Rain gauges	Global (patchy), Point-scale	From 01/01/2000 to 31/12/2019	1 daily realization (real.) - 1 real. X 365 days x 20 years = 7300 total real.	7300 real. / 365 days x 0.75* = 1 in a 15-year event (1 in a 10-year event) *required at least 75% of valid obs.	99.97260
2	NWP-modelled, gridded	ERA5-ensemble Data Assimilation (ERA5-EDA)	Reanalysis, probabilistic (10 ensemble members)	Global, 62 km	From 01/01/2000 to 31/12/2019	10 daily real. - 10 real. X 365 days x 20 years = 73000 total real. *(in reality, we have only 66940 real. as some dates were not available)*	73000 real. / 365 days = 1 in a 200-year event (1 in a 100-year event)* *due to the actual reduced number of total real.	99.99726
3	NWP-modelled, gridded	ERA5	Reanalysis, deterministic	Global, 31 km	From 01/01/2000 to 31/12/2019	1 daily real. - 1 real. X 365 days x 20 years = 7300 total real.	7300 real. / 365 days = 1 in a 20-year event (1 in a 20-year event)	99.98630
4	NWP-modelled, gridded	46r1 ECMWF Reforecasts (reforecast_46r1)	Reforecasts, probabilistic (10 ensemble members, up to day 10). Used only control run	Global, 18 km	Past 20 years for period between 01/07/2019 and 30/06/2020. Reforecasts run only on Mondays and Thursdays.	10 daily real. (all lead times for 2 control runs a week were used as daily independent real.) - 10 real. X 2 runs X 52 weeks x 20 years = 20800 total real.	20800 real. / 365 days = 1 in a 56-year event (1 in a 50-year event)* *preferred a round number for return period	99.99452
5	NWP-modelled, gridded	ERA5-ecPoint	Reanalysis, Probabilistic (99 ensemble members)	Global, Point-scale but provided on ERA5's grid at 31 km	From 01/01/2000 to 31/12/2019	99 daily real. (all ensemble members were used as daily independent real.) - 99 real. X 365 days x 20 years = 722700 total real.	722700 real. / 365 days = 1 in a 1980-year event (1 in a 1000-year event)* *preferred a round number for return period	99.99972

Figure 1: Characteristics (columns 1 to 5) of the considered observational datasets and NWP models (rows 1 to 5), and their derived climatologies (columns 6 to 8).

2.2.3 ERA5-ecPoint

Within the Highlander project, co-financed by the EU and coordinated by Italy's Cineca supercomputing centre, ECMWF's ecPoint post-processing technique was applied to the raw ERA5 "deterministic" fields to address ERA5 limitations (Hewson et al., 2023). ecPoint aims to infer sub-grid variability and to correct biases (both according to ongoing weather and geographical scenarios). The ERA5-ecPoint dataset spans from 1950 to the near-present, providing a long-term, continuously updated record of 24-hourly point-scale rainfall estimates with an accumulation period ending at 00 UTC. The dataset is provided in a probabilistic format, i.e. in percentiles 1,2,... 99. Currently, ERA5-ecPoint represents point-scale precipitation, but it is provided on its native (reduced Gaussian) grid at 31 km spatial resolution (TL639). The

99 percentiles can be considered equally probable precipitation outcomes, at a gauge within a grid-box, so that the precipitation distributions are built with 722700 daily realisations (Table 1, row 5).

3 Methods

3.1 RQ1: assessment of the representation by NWP models of the overall distribution of point-rainfall observations

The assessment of how well NWP models represent the overall distribution of point-rainfall observations is conducted by assessing the similarity between observed and the NWP-modelled rainfall distributions over the 20 years. The modelled estimates are extracted at the rainfall observation locations, considering the modelled value at the nearest grid-box. This approach is regarded as standard practice for rainfall, as interpolation may reduce extremes. This study adopts the method proposed by Gudmundsson et al. (2012), which assesses the similarity between the Empirical Cumulative Distribution Functions (ECDFs) of the observed and NWP-modelled rainfall estimates, constructed with the empirical percentiles (Boé et al., 2007). The similarity is assessed by averaging the Mean Absolute Errors (MAE) at corresponding x^{th} percentiles:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|\mathrm{tp_{OBS}}(x_i^{\mathrm{th}}) - \mathrm{tp_{NWP}}(x_i^{\mathrm{th}})\right| \quad (1)$$

$$n = 99 \text{ percentiles}$$

MAE values are expressed in mm and range from 0 (for perfect similarity) to $+\infty$ (for poor similarity). Graphically, the MAE represents the areal difference between the two ECDFs (Figure 2a)[4]. Ninety-nine percentiles (1st to 99th) were used to avoid contamination related to sampling issues in the observational dataset. Moreover, more extreme rainfall events will be considered separately.

Gudmundsson et al. (2012) methodology was, however, adapted to compare ECDFs from different climatologies. MAE values were normalised ($\mathrm{MAE_{NORM}}$) to avoid having consistently bigger MAE values in wetter climates (see Figure 2b and Figure 2c). The normalisation consists of computing dimensionless coefficients by dividing MAE by the corresponding station's average observed rainfall:

$$\mathrm{MAE_{NORM}} = \frac{\mathrm{MAE}}{\mathrm{average}(\mathrm{tp_{OBS}})} \quad (2)$$

To guide the reader on what was considered by the authors a better or worse similarity degree, Figure 2d shows a selection of $\mathrm{MAE_{NORM}}$ values for different similarity degrees between ECDFs. Four categories of similarity degrees were subjectively defined: $\mathrm{MAE_{NORM}}$ values below 0.1, in black, indicate a "good" degree of similarity, $\mathrm{MAE_{NORM}}$ values between 0.1-0.3, in blue, indicate an "acceptable" degree of similarity, $\mathrm{MAE_{NORM}}$ values between 0.3-0.5, in green, indicate an "intermediate" level of similarity degree, $\mathrm{MAE_{NORM}}$ values between 0.5-1, in yellow, indicate a "poor" degree of similarity, and $\mathrm{MAE_{NORM}}$ values greater than 1, in pink, indicate a "very poor" degree of similarity.

[4]The Mean Error (ME) or Bias could have also been used to measure the similarity between the ECDFs. These two measures, although complementary, can, however, provide very different pictures, as we could have big MAEs while the MEs could be very small if they cancel each other. Moreover, the ME has already been computed for ERA5 by Lavers et al. (2022) to assess its performance in climate monitoring. Results from both studies will, however, be compared in the discussion section

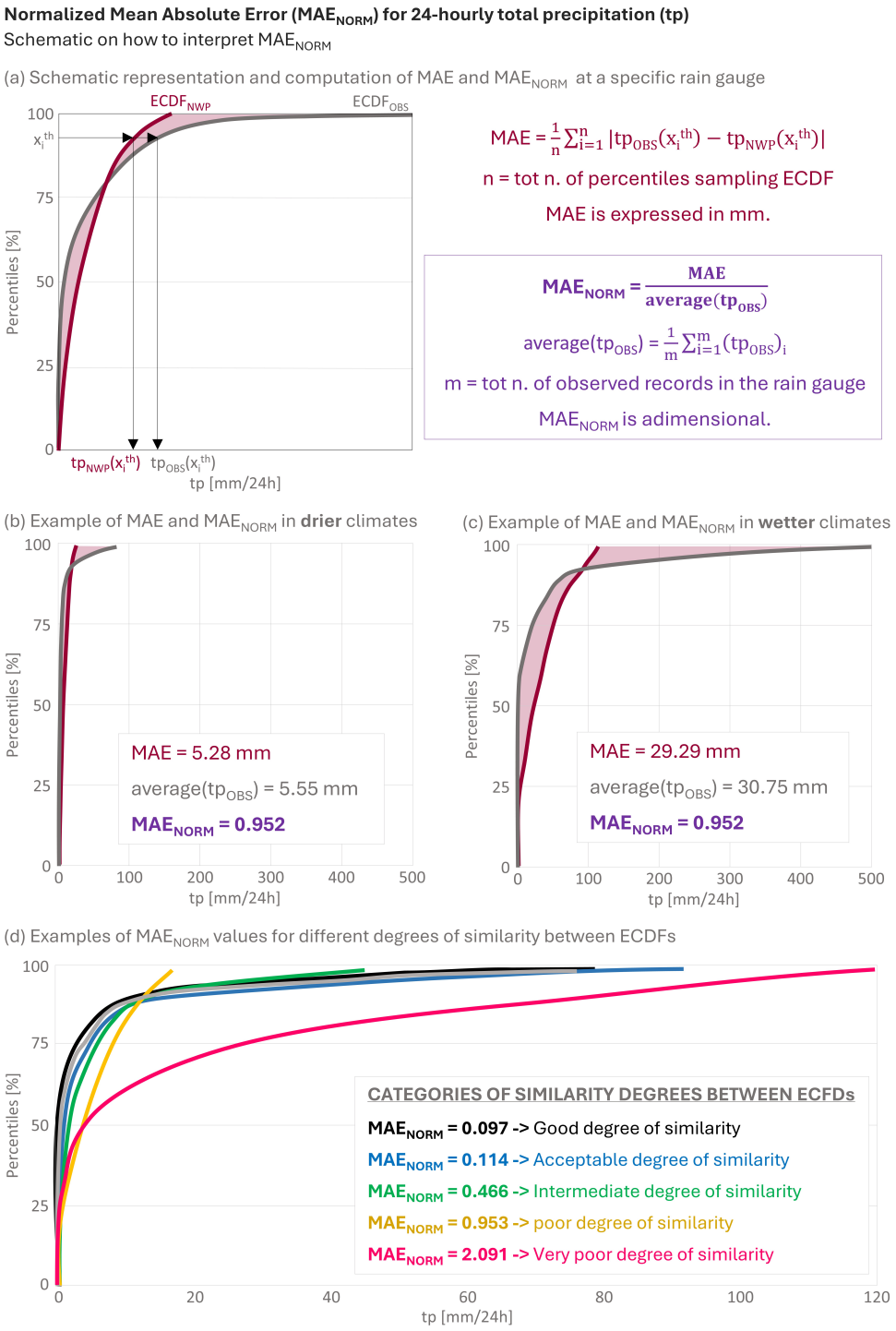

Figure 2: Panel (a) shows a schematic representation and computation of the Mean Absolute Error (MAE) and the Normalised Mean Absolute Error (MAE_{NORM}) for total precipitation (tp). Panels (b) and (c) show how the values of MAE change for drier and wetter climates, respectively, and how dividing them by the average of the observed precipitation helps to normalise the MAE values for different climatologies. Panel (d) shows examples of MAE_{NORM} values for different degrees of similarity between ECDFs: black, blue, green, yellow, and pink represent, respectively, a "good" (less than 0.1), "acceptable" (between 0.1 and 0.3), "intermediate" (between 0.3 and 0.5), "poor" (between 0.5 and 1), and "very poor" degree of similarity (greater than 1). The observational distribution is shown in grey.

Formal statistical tests such as Kolmogorov-Smirnov, Cramer-von-Mises, and Anderson-Darling can also assess the similarity between two distributions (Stephens, 1974). However, for large sample sizes (as in this study, see Table 1, row 6), the tests' statistical significance levels become extremely sensitive to minor differences between distributions that might not be practically significant (Engmann and Cousineau, 2011; Janssen, 2000). This is being referred to as "the problem of practical insignificance" (Kirk, 1996), where the test flags differences that are statistically significant but not meaningful in practice, causing the rejection of the null hypothesis (i.e., the two samples come from the same population) when it is nonetheless practically valid. Gudmundsson's approach was tested to assess whether it was as sensitive to sample size as the formal statistical tests. The ECDFs were sampled with 99, 999, 9999, and 99999 percentiles to assess the sensitivity of MAE_{NORM} to the choice of sampling resolution. The results showed negligible differences across the range of percentiles tested (not shown). Moreover, Gudmundsson's approach assesses similarity between the observed and NWP-modelled rainfall distributions by comparing the whole ECDFs differently to other formal tests that assess similarity only for specific moments of the distribution, such as the mean, standard deviation, skewness, or specific percentiles (Anthanahalli Nanjegowda and Kulamulla Parambath, 2022). Finally, the use of MAE concerning other commonly used scores, such as the Root Mean Squared Error (RMSE), is preferable as it is not unduly sensitive to outliers (e.g., caused by erroneous observations or atypical events), typically observed in the wet tails of the distribution. Hence, MAE should be more representative of the distribution as a whole (Jolliffe and Stephenson, 2011). Moreover, the RMSE is more appropriate when errors follow a normal distribution, which is very atypical for rainfall (Chai and Draxler, 2014). Nonetheless, the RMSE was computed for the examples shown in Figure 2d, and its property of giving more weight to the larger errors (in the wetter part of the distribution) did not change the ranking obtained with the MAE_{NORM} between the different CDFs in 2d (not shown), reassuring the reader that using MAE instead of RMSE should not change the final picture.

Maps plotting MAE_{NORM} values at different rain gauge locations are shown to compare the performance of the four analysed NWP models. The maps are accompanied by pie charts that summarise, for specific regions, the percentage of locations falling in the five MAE_{NORM} categories defined in Figure 2d ($<$ 0.1, between 0.1 and 0.3, 0.3 and 0.5, 0.5 and 1, and $>$ 1). The regions considered are North America, South America, Europe, the Mediterranean, Africa, the Arabian Peninsula, Asia, and Oceania. Finally, a selection of representative ECDFs for all four models against their corresponding observed point-scale precipitation distributions is also shown, to illustrate some differences between the observed and the NWP-modelled precipitation estimates.

3.2 RQ2: assessment of the representation by NWP models of extreme rainfall

The assessment of how well NWP models represent extreme rainfall is conducted by visually comparing the precipitation maps for the 10-year return period to pinpoint geographical differences in estimating extreme precipitation. The 10-year return period was considered because it is the most extreme precipitation event that it was possible to compute with the observational dataset in hand (Table 1, columns 7 and 8, row 1), even though larger events could have been calculated with the NWP-modelled precipitation estimates (Table 1, columns 7 and 8, and rows 2 to 5). The 10-year return period may be exceeded by 10 values in ERA5-EDA, 1 to 2 values in ERA5, 2 to 3 values in the reforecasts, and 100 values in ERA5-ecPoint. A lower number of values exceeding the considered return period may make its estimate noisier. To complement the general global comparison between observed and NWP-modelled extreme precipitation, 24-hourly precipitation estimates from a case of widespread flash floods in Italy are presented. Italy was chosen because, out of all countries in our database, it has the rain gauge network with the highest spatial resolution. This is vital for a case-study-based analysis of extreme rainfall events, as

it increases the chances of capturing extreme localised totals.

4 Results

4.1 RQ1: comparison of rainfall distribution climatologies

Out of all NWP-modelled precipitation estimates, ERA5-ecPoint reproduces observed point-precipitation distributions best. This can be seen by the larger percentage of small MAE_{NORM} values (depicted by the black dots) in Figure 3e compared to Figure 3b-d, where there are bigger percentages of larger MAE_{NORM} values (depicted by the coloured dots). ERA-ecPoint increases the number of MAE_{NORM} values in the black category by a factor of 10, 30, and 27 compared to the reforecasts, ERA5, and ERA5-EDA, respectively (see piecharts and tables in Figure 3a-e). In the baseline NWP models (ERA5-EDA, ERA5, and ECMWF reforecasts), the proportion of grid points with the very high similarity between observed and modelled precipitation (”black” dots) remains consistently low, below 2% in most regions, except in North America, where reforecasts reach about 4%. In South America, the raw NWP models do not yield any such high-similarity points, whereas applying ERA5-ecPoint boosts this proportion to 13%, with representation along Brazil's eastern coast looking particularly good, in relative terms. At the opposite extreme, points with poor similarity (”pink” dots) are substantially reduced when using ERA5-ecPoint. Compared to ERA5-EDA, the number of these poorly performing points declines by about 60%, and relative to ERA5 and reforecasts, by about 50%. These improvements are most pronounced in the Arabian Peninsula, Asia, and North America. Although reforecasts also have a lower count of poorly performing points in these areas, they exhibit slightly worse performance in parts of South America, especially the Bolivian Amazon, increasing the proportion of poor-similarity points by 2% and 5% relative to ERA5-EDA and ERA5, respectively. In contrast, ERA5-ecPoint markedly improves this situation in South America, reducing poor-similarity points by 47%, 23%, and 10% compared to reforecasts, ERA5, and ERA5-EDA, respectively. Much of this improvement occurs in the flatter Amazonian regions east of the Andean highlands. Still, even with ERA5-ecPoint, some challenging areas remain, such as the Andean slopes and the narrow desert-like coastlines of Peru and Chile. For intermediate similarity levels (previously represented by ”blue”, ”green”, and ”yellow” categories), the application of ERA5-ecPoint consistently shifts conditions toward a higher level of agreement across all domains. This results in fewer points showing poor similarity and more points reaching acceptable or good similarity levels. The improvements are especially apparent in South America, Africa, Asia, and Oceania, where ERA5-ecPoint generally transitions more points into categories reflecting moderate to good agreement, thereby offering a notably better representation of precipitation patterns than the baseline NWP models.

It is worth comparing the observed and the NWP-modelled ECDFs to gain insights into how the distributions differ (Figure 4). Each ECDF (in linear scale) has an insert with the ECDF's x-axis in logarithmic scale to compress/expand the small/high rainfall totals, and see more clearly differences in the distributions. In flat areas (Figure 4a), ERA5-ecPoint (in coral) represents the distribution of point-scale precipitation observations better than the baseline raw NWP-models: it captures well the frequency of observed zero precipitation totals (see ECDF in log scale), the growth rate of the precipitation observations[5] (see ECDF in log scale), and the length of the wet tail (going up to the 99th percentile, see ECDFs in linear scale). There are no notable differences between the distributions from ERA5-EDA (in green), ERA5 (in brown), and reforecasts (in blue): they all underestimate, although to different degrees, the frequency of observed zero precipitation totals, and they have similar growth rates, which are greater than that in the observed distribution. They all underestimate the length of the wet tail but to different

[5](Growth rate here is intended as the rate of change of the logarithm of precipitation totals)

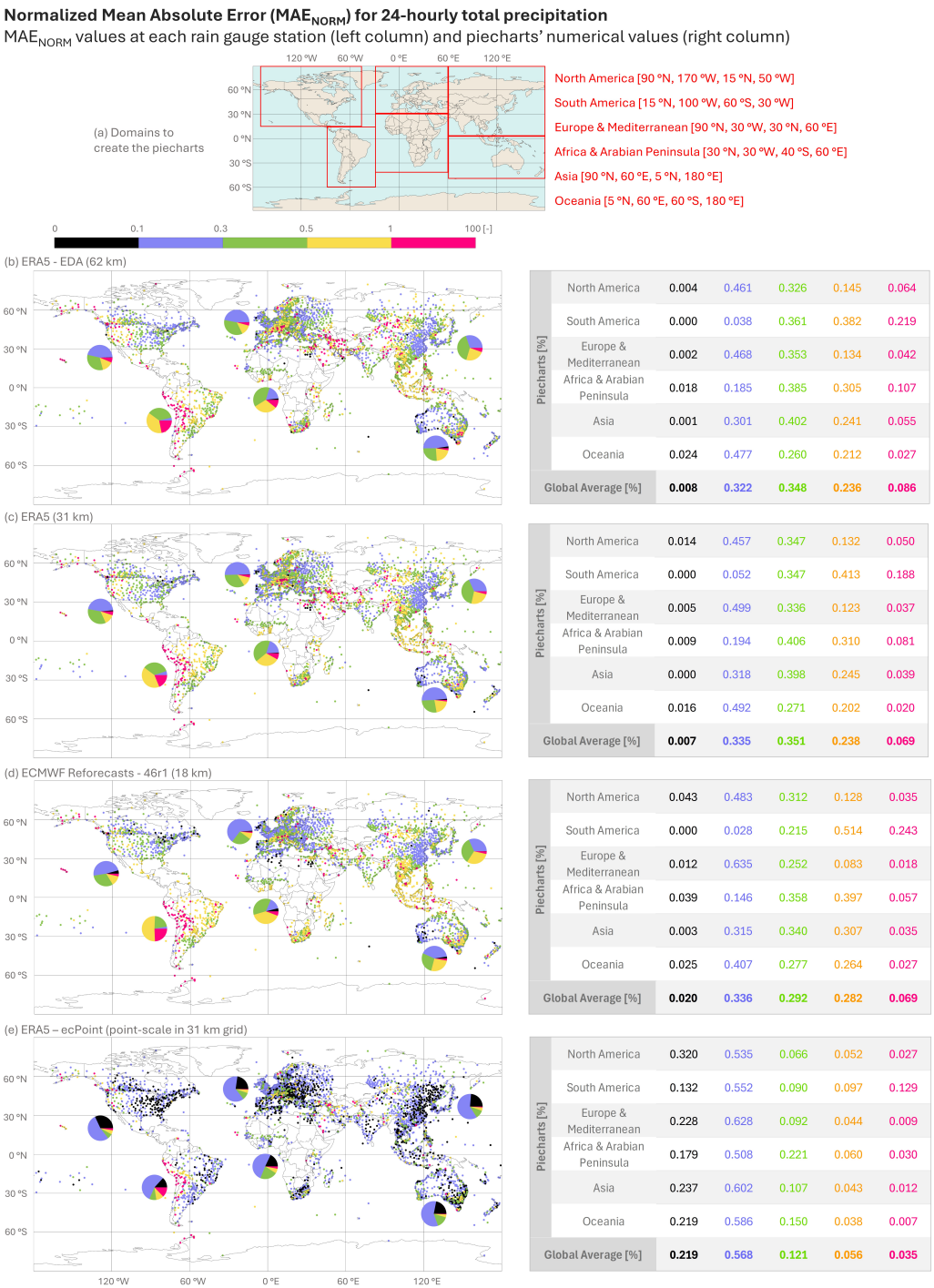

Figure 3: Panels (b) to (e) indicate the Normalised Mean Absolute Error (MAE_{NORM}) for 24-hourly total precipitation at each rain gauge location for ERA5-EDA (62 km), ERA5 (31 km), ECMWF Reforecasts-46r1 (18 km), and ERA5-ecPoint (point-scale, but provided in ERA5's grid), respectively. Dots in black, blue, green, yellow, and pink represent, respectively, a "good", "acceptable", "intermediate", "poor", and "very poor" degree of similarity to the corresponding point observed climatology. The pie charts indicate the frequency of MAE_{NORM} values in the domains defined in panel (a). The tables on the right offer a numerical representation of the pie charts. The numbers in bold in the last rows of each table represent the global average for each representation category.

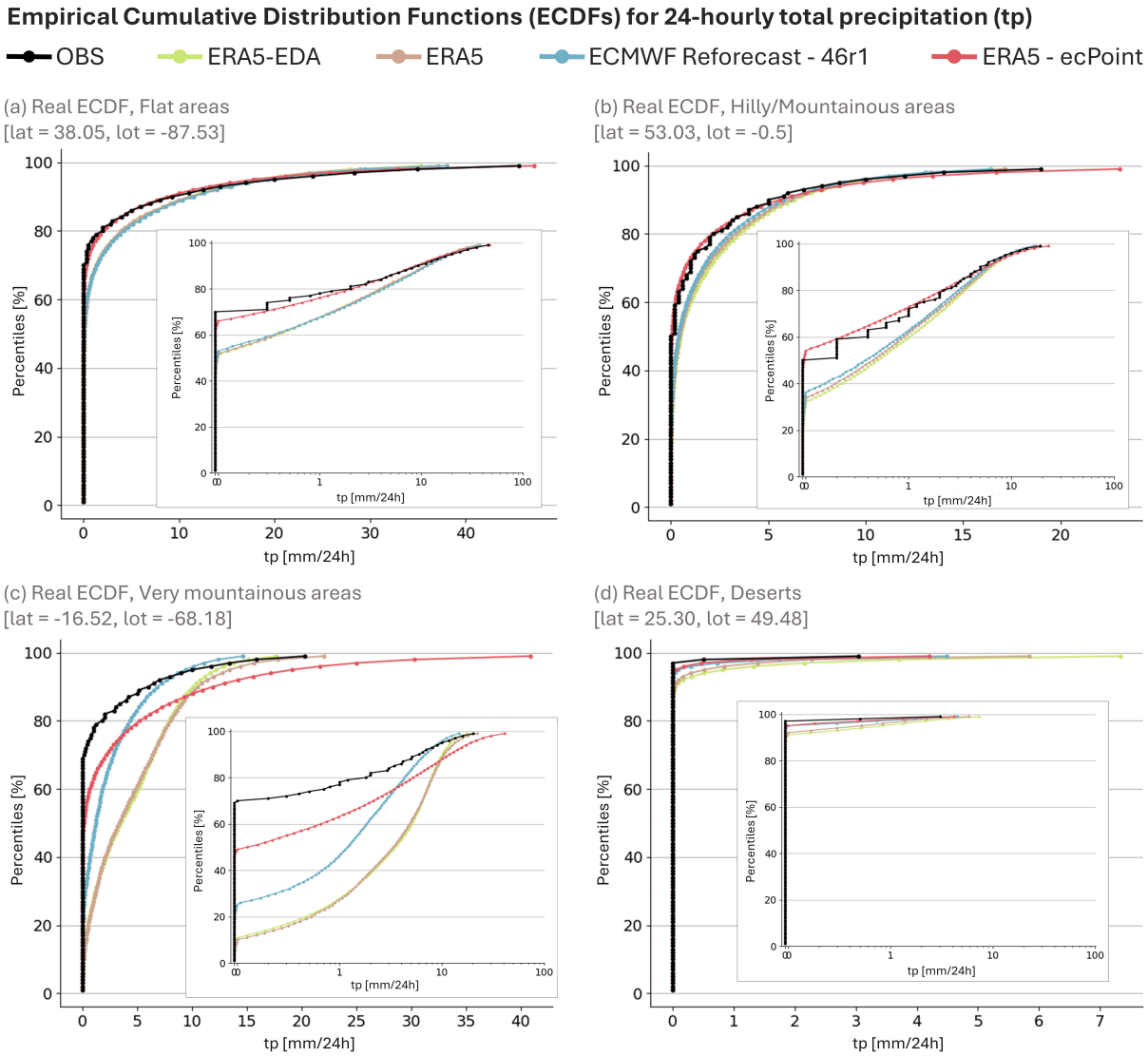

Figure 4: Empirical Cumulative Distribution Functions (ECDFs) for 24-hourly total precipitation (tp) from rain gauge observations (OBS, in black) and the NWP models ERA5-EDA (in green), ERA5 (in brown), ECMWF Reforecasts-46r1 (in light blue), and ERA5-ecPoint (in coral). Panels (a) to (d) show examples of ECDFs, respectively, for flat areas, hilly/mountainous areas, very mountainous areas, and deserts. The inserts represent the same ECDFs but with the x-axis on a logarithmic scale.

degrees: in general, ERA5-EDA shows the biggest underestimation, reforecasts show the smallest, and ERA5 falls in between the two. In hilly/mountainous areas (Figure 4b), ERA5-ecPoint behaves similarly to flat areas. It represents the frequency of zero precipitation totals observed and the growth rate of the precipitation observations well. However, ERA5-ecPoint tends to slightly overestimate the distribution's wet tail of the observed ECDF (in black). Compared to point-rainfall observations, raw NWP models show behaviour similar to that observed in flat areas. The main difference lies in a progression in a better representation of the observed ECDF for NWP models with increasing spatial resolution, i.e., ERA5-EDA at 62 km (in green, in Figure 4b), which shows a worse representation of the observed ECDF compared to ERA5 at 31 km (in brown), and ERA5 shows a worse representation than reforecasts (in blue). This behaviour is seen in other sites too (not shown). In very mountainous areas (Figure 4c), all NWP models fail to represent the observed ECDFs. It is worth noting that this is not surprising as the observations used to train ERA5-ecPoint, and indeed to validate all representations, come primarily from valleys and hilly areas. First, all the NWP model versions underestimate the frequency of observed zero precipitation totals. ERA5-ecPoint tends to double such a frequency, but it does not reach the values in the observed ECDFs. The ECDFs from raw NWP models show a growth rate that is too large compared to the observed ECDFs, while ERA5-ecPoint also improves on that. Finally, while the raw NWP models tend to slightly underestimate the length of the observed ECDFs (with ERA5 providing the best representation out of the three models), ERA5-ecPoint tends to overestimate it. In desert areas (Figure 4d), all NWP models represent the observed ECDFs well, apart from the wet tails that tend to all be overestimated. The overestimation is reduced with the increase in the spatial resolution of the NWP models, with ERA5-ecPoint representing the actual length of the wet tail best.

4.2 RQ2: comparison of the wet tail in the distributions built with NWP-modelled precipitation estimates and rain gauge observations

4.2.1 Comparison of the 10-year period

The precipitation maps for the 10-year return period show that ERA5-ecPoint provides a better representation than the raw NWP models. In North America, the extremes in 24-hourly precipitation over the west coast of Alaska, Canada and North-West USA, which reach peaks up to 500 mm, are better represented in ERA5-ecPoint than in ERA5-EDA, ERA5, and reforecasts that tend not to exceed 125 mm. The peaks around the Gulf of Mexico, the USA's East Coast and the border between Canada and the USA are also better represented in ERA5-ecPoint. However, in the latter case, there seems to also be sampling-related noise in the observations. The raw NWP models better represent the extremes over the Rocky Mountains since ERA5-ecPoint overestimates them. However, the latter shows an overall closer representation of the observed ECDFs apart from the tail. ERA5-ecPoint greatly improves the precipitation peaks over Mexico and South America over the other three NWP models, apart from the Andean region and the desert on the west coast of Peru and Chile, where ERA5-ecPoint overestimates the wet tails (as shown in section 4.1). It is worth noting that the ECMWF reforecasts from 46r1 halved the precipitation extremes over the Amazon compared to ERA5-EDA and ERA5.

The extremes over Europe also verify better on ERA5-ecPoint than the three raw NWP models. The wetter climatology with peaks up to 300-500 mm around the Mediterranean catchment (including the African part), the Alps, the Atlantic coast of Spain and the UK, and the Norwegian Fiords is better captured in ERA5-ecPoint than in the three raw NWP models. The higher spatial resolution in the reforecasts helps to increase the extremes compared to both reanalyses, but they still do not exceed 100 mm in 24-hours.

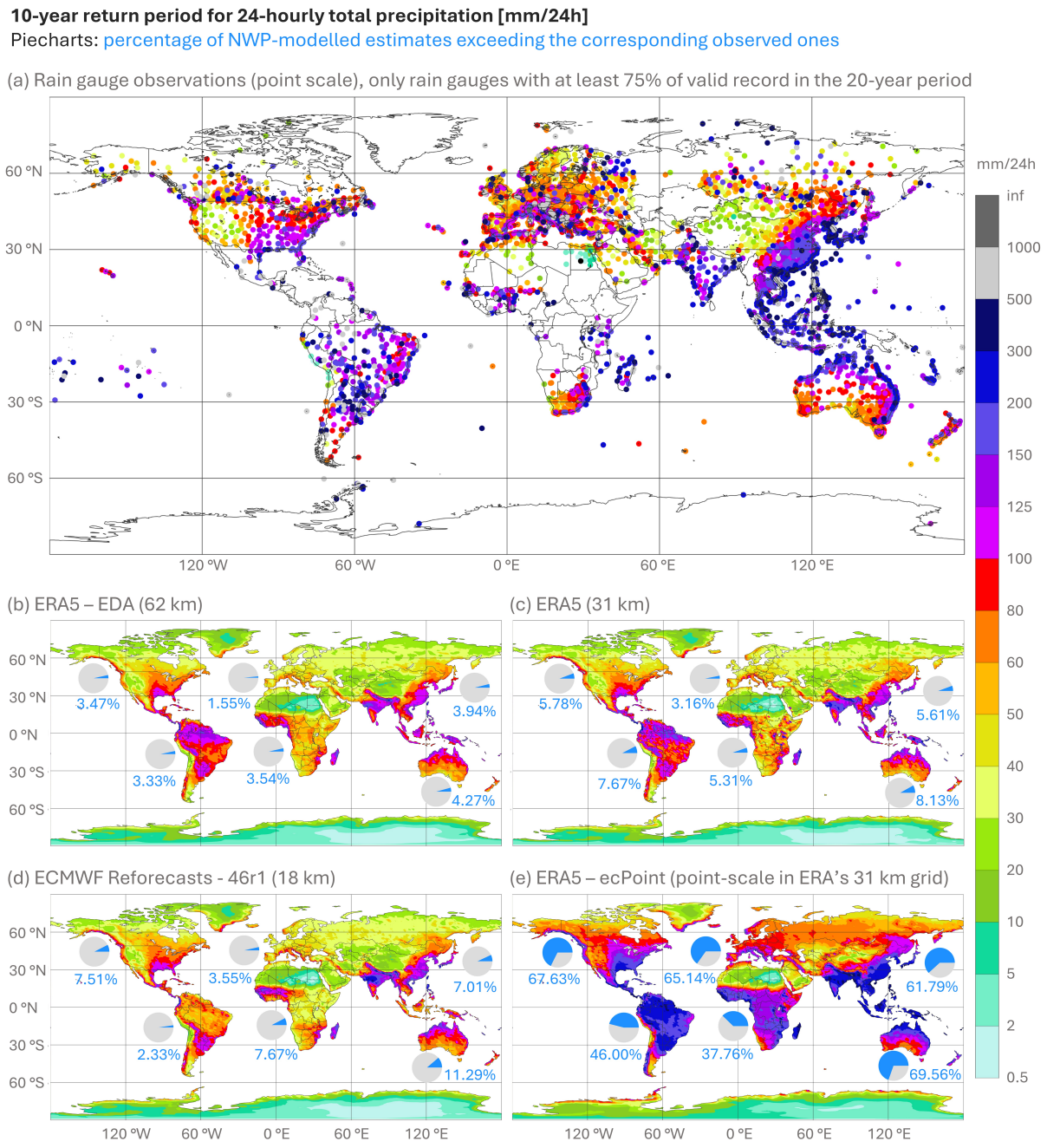

Figure 5: Panel (a) displays the 10-year return period for 24-hourly total precipitation from rain gauge observations, calculated over the 20-year period between 2000 and 2019, and using only rain gauges with at least 75% of valid records. Panels (b) to (e) show the 10-year return period for NWP-modelled 24-hourly total precipitation: ERA5-EDA (62 km), ERA5 (31 km), ECMWF Reforecasts-46r1 (18 km) and ERA5-ecPoint (point-scale, provided on ERA5 grid). The pie charts represent the percentage (in %) of modelled climatologies exceeding the observed climatologies. Reliable modelled estimates should exceed the observed ones, on average, 50% of the time.

In Asia, there is a varied picture. The raw NWP models highlight the wetter climatologies of India (especially the Northeast regions), East China, Japan, Southeast Asia, and the Malay Archipelago. However, they do not reach the peaks of 300-500 mm/24h seen in the observations. ERA5-ecPoint represents such peaks. However, the peaks greater than 500 mm/24h observed in the Malay Archipelago remain underestimated, also in the post-processed ERA5. The overall overestimation in the mountainous regions of Western China has a similar flavour to the ones discussed over the Rocky Mountains in the USA: ERA5-ecPoint shows the best overall representation of the full observed ECDFs, but tends to overestimate the wet tails. In the Arabian Peninsula, all models represent the overall observed ECDF tails quite well. As discussed in section 4.1 for desert areas, such good representation originates from the high frequency of zero precipitation totals well estimated by all NWP models. The only exception is on the peninsula's south coast, where precipitation peaks can reach 200 mm/24h, and raw NWP models estimate a maximum peak of only up to 80 mm/24h. ERA5-ecPoint increases them up to 150 mm/24h. In Oceania, all NWP models show a good overall representation of the observed ECDFs with slight underestimations of the wet tails. The added value of ERA5-ecPoint in this region mainly provides a better representation of the precipitation peaks. There are a few observations in Africa, and nothing can be said about the model representation of precipitation extremes in the numerous ungauged areas of this continent. All NWP models represent the wet climatology of West Africa, including its Atlantic coast. However, ERA5-ecPoint best represents the observed local peaks that vary between 100 and 500 mm/24h. It is worth noting that ECMWF 46r1 reforecasts degrade the representation of the extreme precipitation around the Gulf of Guinea by producing maximum peaks only up to 80-100 mm/24h. Similarly, out of all NWP models, ERA5-ecPoint somewhat better represents the varied precipitation peaks, between 80 and 500 mm/24h, in South Africa, where raw NWP models suggest extreme precipitation might not exceed 80 mm/24h. Also, in East Africa, ERA5-ecPoint provides a more realistic representation of the extreme precipitation peaks (up to 500 mm/24h) than raw NWP models. The reforecasts considerably reduce the precipitation in this area. The wet climatology of Madagascar is well represented in all NWP models, but ecPoint can increase the wet tail of ERA5 and provide extreme precipitation totals that are closer to those observed. Finally, all NWP models seem to represent quite well the observed precipitation distribution in the Sahara, with the caveat that data coverage there is poor. In any case, good performance likely connects to the prevalence of dry weather.

4.2.2 Case study: Storm Vaia in Italy (28th of October 2018)

We now examine the case of widespread (flash) flooding in Italy on the 28th of October 2018 (Figure 6). This event is part of a weather system that persisted over different parts of Italy between the end of October and the beginning of November 2018. It is called Storm Vaia. In the observations (Figure 6a), one can see extreme precipitation amounts between 300-400 mm/24h over Veneto (north-east), up to 200 mm/24h over Lombardi (North) and Liguria (North-East), up to 240 mm/24h in Lazio (centre), up to 130 mm/24h in Puglia (Southwest), and up to 260 mm/24h in Calabria (Southeast). ERA5-EDA (Figure 6b), ERA5 (Figure 6c), and reforecasts (Figure 6d) provide a good signal on which might be the wetter areas in Italy for that day, apart from the south of Italy, that does not stand out as a possible area at risk of extreme precipitation. The precipitation peaks over the Italian Peninsula increase with the increasing spatial resolution of the NWP models, but they do not reach the observed extreme precipitation totals. ERA5-EDA estimated a maximum total of 100 mm/24h, and ERA5 pushed the estimated peaks to 150 mm/24h over Veneto. Reforecasts increased the precipitation peaks in Veneto and Lazio up to 200 mm/24h, but precipitation in Liguria, Puglia, and Calabria remains highly underestimated. ERA5-ecPoint (Figure 6e) represents better the areas where the precipitation peaks were observed. In the north (Figure 6f, Northern Italy), where the Storm created the biggest impacts, roughly 1% of the rainfall observations exceeded the

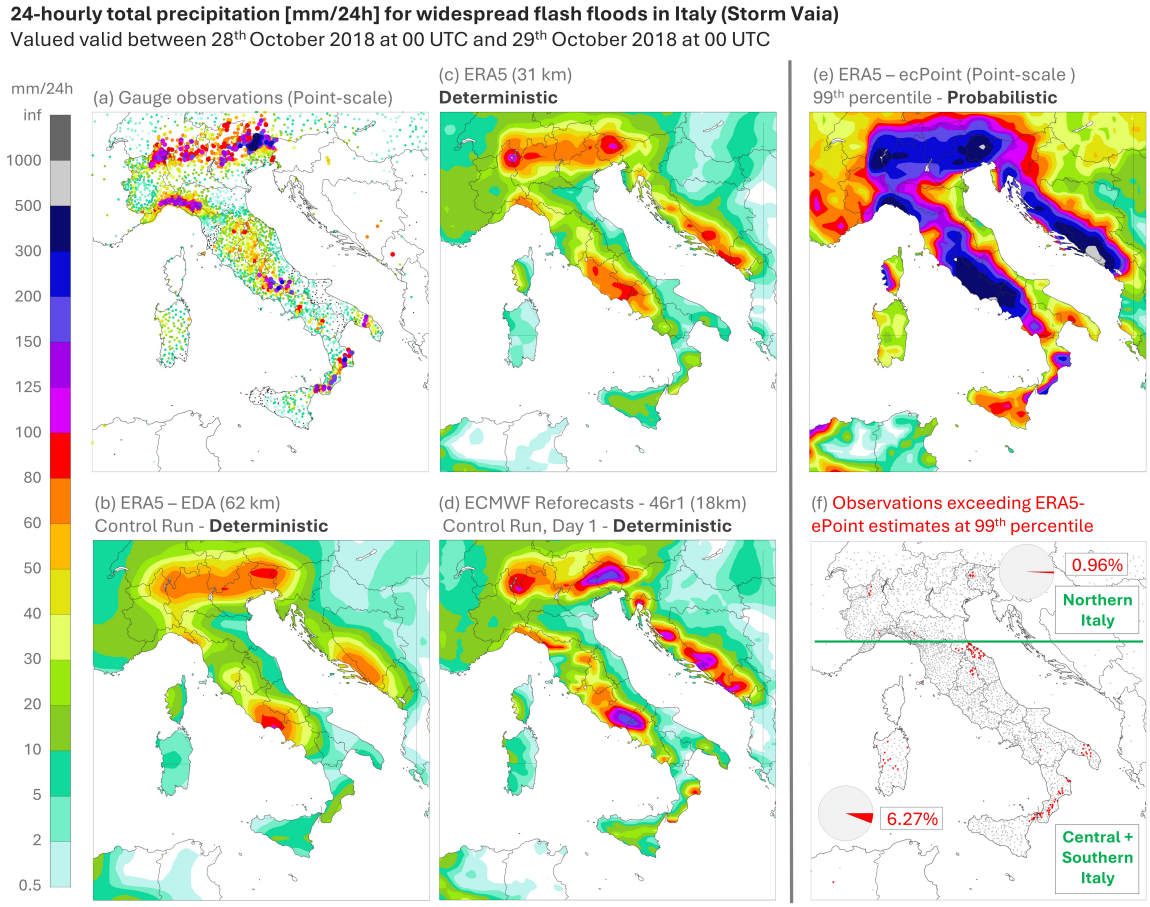

Figure 6: Widespread (flash) flooding in Italy on the 28th of October 2018 due to Storm Vaia. Panel (a) represents the rain gauge observations. Panels (b) to (d) show the deterministic rainfall estimates, respectively, for ERA5-EDA at 62 km (from control run), ERA at 31 km (single realisation), and ECMWF Reforecasts from 46r1 at 18 km (from control run, day 1 lead time). Panel (e) shows the probabilistic rainfall estimates from ERA5-ecPoint (99th percentile). Panel (f) shows the locations of the rain gauge observations exceeding the ERA5-ecPoint estimates at the 99th percentile. The representation is split into two geographical areas: Northern and Southern Italy, with pie charts denoting total counts for these two areas.

99th percentile of ERA5-ecPoint (red dots), indicating reliable point-scale rainfall estimates. In the rest of the peninsula (Figure 6f, Central + Souther Italy), 6% of the observations exceed the ERA5-ecPoint estimates, indicating an under-prediction of point-scale rainfall over Le Marche, Puglia, and Calabria. In this specific case, the location of the red dots along coastlines indicates underestimation primarily due to the known issue of convective cells generated over the sea not moving onto land.

The results from this study show that ERA5-ecPoint provides, overall, the best representation of point-rainfall distributions out of all the NWP models tested. Specifically, ERA5-ecPoint captures better the frequency of the observed zero rainfall totals, the growth rate within the rainfall observation CDFs, and the longer wet tails. The bigger improvements are particularly evident in flat and hilly/mountainous regions. However, in very mountainous areas such as the Andes, ERA5-ecPoint underestimates the frequency of zero rainfall totals and overestimates the length of the wet tails, raising some questions about its effectiveness over very complex orography. This should not surprise, as ERA5-ecPoint is post-

processed with observations primarily coming from valleys and hilly areas, although, on the other hand, verifying data comes from such sites too. Probably, we have a complex interplay, whereby data from non-mountainous regions is sometimes used to train for mountainous areas, despite the inclusion of a sub-grid orography variable in the ERA5-ecPoint decision tree. The growth rate of the ERA5-ecPoint rainfall estimates closely aligns with that of the observations, indicating that the post-processing system is making meaningful adjustments to the rainfall estimates. Additional observational data from regions at high altitudes are necessary to refine the corrections, particularly to increase the accuracy in representing the frequency of zero rainfall totals and to reduce the overestimation observed in the wet tail.

Overall, the raw NWP models (i.e., ERA5-EDA, ERA5, and ECMWF Reforecasts – 46r1) consistently show an underestimation of the zero rainfall totals and the wet tails, and the growth rate of the modelled rainfall estimates is consistently bigger than that observed. This means that the raw NWP models overestimate the frequency of small rainfall totals and underestimate the frequency of extreme rainfall events, as one might expect from representivity considerations, and as has been reported previously by National Meteorological and Hydrological Services around Europe (Hewson and Chevallier, 2024). ERA5 (at 31 km) improves the overall representation of point-rainfall distributions compared to ERA5-EDA (at 62 km), especially in mountainous regions such as the Rocky Mountains, the Alps, and the Norwegian Fjords. However, the improvements in these regions remain modest in proportion, despite the twofold increase in spatial resolution. The ECMWF Reforecasts provide general improvements due to the increased spatial resolution (18 km) and a more up-to-date model version (46r1 rather than 41r2 of ERA5-EDA and ERA5). The observed degradations over Australasia and Africa in 46r1 (see pie charts on Figure 3) are counterintuitive and may be symptomatic of a physics issue that manifests in those areas. Compared to ERA5, the 46r1 improvements are focused again on mountainous areas and extend to most of Europe, the arid regions of Northern Africa, and the Arabian Peninsula.

Focusing on extreme rainfall events, there is a general increase in the values with the increase of the raw NWP models' spatial resolution, which better agrees with the observed wetter tails. The major difference is observed between ERA5-EDA and ERA5, while the differences between the latter and the ECMWF reforecasts are less prominent. Indeed, for the rainfall in the Amazon region, Equatorial Africa, and Indonesia, the reforecasts show rainfall estimates that do not exceed 100 mm/24h. In contrast, both reanalysis, ERA5-EDA and ERA5, show rainfall estimates up to 300 mm/24h, which better represent the observed rainfall totals in the region. These results similarly contradict expectations and may indicate regional limitations in cycle 46r1 employed for the reforecast dataset.

When focusing on extreme precipitation events, ERA5-ecPoint consistently demonstrates a superior ability to replicate observed extremes compared to raw NWP models. For example, the 10-year return period precipitation maps show that ERA5-ecPoint provides a much closer representation of observed extreme rainfall events in regions like North America, Europe, and parts of Asia. The Italian case study on Storm Vaia further underscores this finding. While raw models captured the general distribution of wetter areas, they underestimated the magnitude of precipitation peaks across multiple regions, including Veneto, Lazio, and Liguria. ERA5-ecPoint, on the other hand, was able to capture these extremes better, providing a more realistic forecast of the potential for flash floods. Some underestimation of rainfall along coastlines is highlighted due to non-moving convective cells generated over the sea that fail to generate rain over the land. Convective cell drift is something that has been explored in the ecPoint framework, but not implemented yet. Applying it should bring intrinsic improvements in the areas of triggering, via the bias correction aspect (as shown in the Hewson and Pillosu (2021), Norway example).

Furthermore, ERA5-ecPoint enables one to estimate rainfall events with significantly longer return periods than those presented in this study (e.g., up to a 1000-year return period, as noted in Table 1, row 5, column 8). Hewson et al. (2024) have shown for 2023 Storm Daniel in Libya that applying the ecPoint

post-processing technique to ERA5 can deliver usable estimates of an n-year return period rainfall from m years of data, where $n \gg m$. Consequently, datasets like ERA5-ecPoint offer valuable insights into the potential magnitude of extreme rainfall events, improving our preparedness for unseen events (Heinrich et al., 2024; Ommer et al., 2024) or ones so infrequent that they have faded from collective memory (Ludwig et al., 2023; Merz et al., 2024).

One area where ERA5-ecPoint did not seem to provide significant benefits, and where extremes were overestimated, was in the high-altitude and relatively dry western parts of the USA, where mountain barriers can block external moisture sources. Parts of inland northern China fall into the same class. Dry boundary layers often characterise such areas. We know from experimenting with ecPoint and considering physics that low-level rainfall under-evaporation in such situations can lead to large net positive raw model rainfall biases at the grid scale, particularly in convective situations. Although ERA5-ecPoint includes a low-level humidity parameter within its decision tree, which can combat such biases, it is probably not activated in enough weather-type scenarios to be fully effective. Hence, probably there is some cross-contamination in the calibration from data in areas with much moister boundary layers. This could thus be a focal point for future work. While ecPoint's remote calibration approach has shown significant benefits compared to a purely local approach, as in this paper and in Hewson and Pillosu (2021), there can evidently be some local downsides.

The general improvements provided by ERA5-ecPoint open up significant opportunities across various fields of environmental research that require a more accurate representation of point rainfall estimates. We advocate that such improvements would enhance both long-term strategic planning (e.g., using this dataset for climatological studies) and short-term emergency response (e.g., this dataset to create point-scale rainfall thresholds that are compatible with ecPoint rainfall medium-range forecasts to determine areas at risk of extreme localised rainfall), thereby contributing to developing more resilient societies in the face of climate change. In the realm of flood forecasting, more accurate rainfall estimates at local scales are crucial for predicting runoff and streamflow dynamics, particularly in catchments prone to flash floods. Precise point-scale rainfall data is pivotal in enhancing early warning systems, which are essential for safeguarding communities from the severe impacts of extreme rainfall and flooding. Better rainfall representation could facilitate more efficient management of reservoirs and irrigation planning in water resource management, optimising water storage and distribution for agriculture, power generation, and urban water supply systems. Furthermore, enhanced point-scale precipitation estimates are crucial for designing more resilient stormwater infrastructure and urban drainage systems, which are facing increasing pressure from the intensification of extreme rainfall events due to climate change. In the context of disaster preparedness, ERA5-ecPoint's ability to capture the full spectrum of rainfall values, including zeros and extremes, provides valuable insights into the risks posed by changing precipitation patterns.

5 Conclusions

Modern-day NWP systems and reanalysis products do not provide a good representation of 24h climatological rainfall distributions for gauged sites around the world, whilst ecPoint, in its ERA5 variant form, though not perfect everywhere, does do very much better.

This study provides a systematic, global verification of how well NWP models represent the distribution of point-rainfall observations. It considered point-rainfall observations over 20 years and four different modelled, gridded datasets with distinct spatial resolutions: ERA5-EDA (62 km), ERA5 (31 km), ECMWF Reforecasts for 46r1 (18 km), and ERA5-ecPoint (point-scale but provided over ERA5's grid

at 31 km). Among the tested models, this study shows that ERA5-ecPoint most accurately captures both the frequency of zero rainfall totals and the wet tails of the observed point-rainfall distributions.

Since ERA5-ecPoint provides rainfall totals over a continuous global domain, the post-processed reanalysis could be used to provide seamless point-rainfall estimates, including over regions with sparse or no rain gauge observational data. However, caution is needed when generalising the verification results. While ERA5-ecPoint demonstrates strong performance in estimating point-rainfall totals overall, it is essential to note that the verification dataset contains large regions with sparse or no rain gauge observations. Furthermore, ERA5-ecPoint has shown some limitations in very complex mountainous terrains (e.g. the Andes), where the post-processed reanalysis remains short in representing the frequency of zero rainfall totals and overestimates the wet tails. Hence, this finding highlights the need for further refinement of the post-processed forecasts in these regions by incorporating, when available, more rain gauge observations in the calibration process.

The improved performance of ERA5-ecPoint over raw NWP models in representing point-scale rainfall totals, whether small or large, emphasises post-processing's critical role in addressing the inherent limitations of gridded rainfall estimates in guiding point-scale rainfall. ecPoint effectiveness, however, remains contingent on the quality of the underlying NWP models it post-processes. Without accurate raw NWP estimates at a grid-scale, the skill demonstrated by the ERA5-ecPoint rainfall estimates would be diminished. Moving forward, the authors advocate enhancing the spatial resolution and the skill of raw NWP models alongside ongoing improvements of post-processing techniques such as ecPoint to reduce further errors in estimating the whole distribution of point-rainfall totals. Such improvements will be particularly significant as climate change intensifies the frequency and severity of extreme rainfall, making accurate and reliable point-rainfall estimates indispensable for effective mitigation and response efforts related to droughts, extreme rainfall, flooding, food security, and urban resilience.

Acknowledgments

The authors gratefully acknowledge Hans Hersbach, Paul Berrisford, and Adrian Simmons for their invaluable insights on the characteristics of ERA5 and ERA5-EDA datasets and their implications in the outcomes shown in this paper.

References

Alexandridis, V., Stefanidis, S. and Dafis, S. (2023) Evaluation of era5 and era5-land reanalysis precipitation data with rain gauge observations in greece. *Environmental Sciences Proceedings*, **26**, 104.

Anand, B. and Karunanidhi, D. (2020) Long-term spatial and temporal rainfall trend analysis using gis and statistical methods in lower bhavani basin, tamil nadu, india. *Indian Journal of Geo-Marine Sciences*, **49**, 419–427.

Anthanahalli Nanjegowda, R. and Kulamulla Parambath, S. (2022) A novel bias correction method for extreme rainfall events based on l-moments. *International Journal of Climatology*, **42**, 250–264.

Asadieh, B. and Krakauer, N. Y. (2015) Global trends in extreme precipitation: Climate models versus observations. *Hydrology and Earth System Sciences*, **19**, 877–891.

Balasundram, S. K., Shamshiri, R. R., Sridhara, S. and Rizan, N. (2023) The role of digital agriculture in mitigating climate change and ensuring food security: An overview. *Sustainability*, **15**.

Bottazzi, M., Rodríguez-Muñoz, L., Chiavarini, B., Caroli, C., Trotta, G., Dellacasa, C., Marras, G. F., Montanari, M., Santini, M., Mancini, M. et al. (2024) High performance computing to support land, climate, and user-oriented services: The highlander data portal. *Meteorological Applications*, **31**, e2166.

Boé, J., Terray, L., Habets, F. and Martin, E. (2007) Statistical and dynamical downscaling of the seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, **27**, 1643–1655.

Chai, T. and Draxler, R. R. (2014) Root mean square error (rmse) or mean absolute error (mae)? - arguments against avoiding rmse in the literature. *Geoscientific Model Development*, **7**, 1247–1250.

Cuo, L., Pagano, T. C. and Wang, Q. J. (2011) A review of quantitative precipitation forecasts and their use in short- to medium-range streamflow forecasting. *Journal of Hydrometeorology*, **12**, 713–728.

Di Curzio, D., Di Giovanni, A., Lidori, R., Montopoli, M. and Rusi, S. (2022) Comparing rain gauge and weather radar data in the estimation of the pluviometric inflow from the apennine ridge to the adriatic coast (abruzzo region, central italy). *Hydrology*, **9**, 225.

Donat, M. G., Alexander, L. V., Herold, N. and Dittus, A. J. (2016) Temperature and precipitation extremes in century-long gridded observations, reanalyses, and atmospheric model simulations. *Journal of Geophysical Research*, **121**, 11174–11189.

Doocy, S., Daniels, A., Murray, S. and Kirsch, T. D. (2013) The human impact of floods: A historical review of events 1980-2009 and systematic literature review. *PLoS Currents*.

Engmann, S. and Cousineau, D. (2011) Comparing distributions: The two-sample anderson-darling test as an alternative to the kolmogorov-smirnoff test. *Journal of Applied Quantitative Methods*, **6**. URL: https://jaqm.ro/issues/volume-6,issue-3/1_engmann_cousineau.php.

Ensor, L. A. and Robeson, S. M. (2008) Statistical characteristics of daily precipitation: Comparisons of gridded and point datasets. *Journal of Applied Meteorology and Climatology*, **47**, 2468–2476.

Espinosa, L. A., Portela, M. M. and Gharbia, S. (2024) Assessing changes in exceptional rainfall in portugal using era5-land reanalysis data (1981/1982–2022/2023). *Water*, **16**, 628.

Gimeno, L., Sorí, R., Vázquez, M., Stojanovic, M., Algarra, I., Eiras-Barca, J., Gimeno-Sotelo, L. and Nieto, R. (2022) Extreme precipitation events. *Wiley Interdisciplinary Reviews: Water*, **9**.

Giorgos, N., Nastos, P. T. and Kapsomenakis, Y. (2024) Creating high-resolution precipitation and extreme precipitation indices datasets by downscaling and improving on the era5 reanalysis data over greece. *Eng*, **5**, 1885–1904.

Gomis-Cebolla, J., Rattayova, V., Salazar-Galán, S. and Francés, F. (2023) Evaluation of era5 and era5-land reanalysis precipitation datasets over spain (1951–2020). *Atmospheric Research*, **284**.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E. and Engen-Skaugen, T. (2012) Technical note: Downscaling rcm precipitation to the station scale using statistical transformations - a comparison of methods. *Hydrology and Earth System Sciences*, **16**, 3383–3390.

Gupta, V., Jain, M. K., Singh, P. K. and Singh, V. (2020) An assessment of global satellite-based precipitation datasets in capturing precipitation extremes: A comparison with observed precipitation dataset in india. *International Journal of Climatology*, **40**, 3667–3688.

Haiden, T. and Duffy, S. (2016) Use of high-density observations in precipitation verification. *ECMWF Newsletter*, 20–25.

Haile, G. G., Tang, Q., Hosseini-Moghari, S. M., Liu, X., Gebremicael, T. G., Leng, G., Kebede, A., Xu, X. and Yun, X. (2020) Projected impacts of climate change on drought patterns over east africa. *Earth's Future*, **8**.

Hamill, T. M., Whitaker, J. S. and Mullen, S. L. (2006) Reforecasts: An important dataset for improving weather predictions. *Bulletin of the American Meteorological Society*, **87**, 33–46.

Hamill, T. M., Whitaker, J. S., Shlyaeva, A., Bates, G., Fredrick, S., Pegion, P., Sinsky, E., Zhu, Y., Tallapragada, V., Guan, H., Zhou, X. and Woollen, J. (2022) The reanalysis for the global ensemble forecast system, version 12. *Monthly Weather Review*, **150**, 59–79.

Heinrich, D., Stephens, E. and Coughlan de Perez, E. (2024) More than magnitude: Towards a multi-dimensional understanding of unprecedented weather to better support disaster management. *Water Security*, **23**, 100181.

Herold, N., Behrangi, A. and Alexander, L. V. (2017) Large uncertainties in observed daily precipitation extremes over land. *Journal of Geophysical Research: Atmospheres*, **122**, 668–681.

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. and ... Thépaut, J. N. (2020) The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1999–2049.

Hewson, T. (2024) Capturing extreme rainfall events. *ECMWF Newsletter*.

Hewson, T., Ashoor, A., Bousetta, S., Emanual, K., Lagouvardos, K., Lavers, D., Magnusson, L., Pillosu, F. and Zsoter, E. (2024) Medicane daniel: an extraordinary cyclone with devastating impacts. *ECMWF Newsletter*.

Hewson, T. and Chevallier, M. (2024) Use and verification of ecmwf products. *ECMWF Technical Memoranda*.

Hewson, T. and Pillosu, F. (2021) A low-cost post-processing technique improves weather forecasts around the world. *Communications Earth and Environment*, **2**.

Hewson, T., Pillosu, F., Gascòn, E. and Vučković, M. (2023) Post-processing era5 output with ecpoint. *ECMWF Newsletter*.

Hossain, I., Gato-Trinidad, S., Imteaz, M. and Rayburg, S. (2024) Future scenarios of design rainfall due to upcoming climate changes in nsw, australia. *Atmosphere*, **15**, 1101.

IPCC (2023) *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland: IPCC.

Janmohammadi, M. and Sabaghnia, N. (2023) Strategies to alleviate the unusual effects of climate change on crop production: a thirsty and warm future, low crop quality. a review. *Biologija*, **69**, 121–133.

Janssen, A. (2000) Global power functions of goodness of fit tests. *The Annals of Statistics*, **28**, 239–253.

Jolliffe, I. T. and Stephenson, D. B. (2011) *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley & Sons, Ltd., 2nd edn.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A. and ... Joseph, D. (1996) The ncep/ncar 40-year reanalysis project. *Bulletin of the American Meteorological Society*, **77**, 437–472.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G. and Kirschbaum, D. B. (2017) So, how much of the earth's surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, **98**, 69–78.

Kirk, R. E. (1996) Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, **56**, 746–759.

Lamprecht, A., Rutzinger, M., Pauli, H., Bardy-Durchhalter, M., Euller, K., Niederheiser, R., Steinbauer, K., Wilfing, K., Erschbamer, B., Calzado, R. F., Cecco, V. D., Gattringer, A., Kazakis, G., Mallaun, M., Mesa, J. M., Moser, D., Remoundou, H., Stanisci, A., Theurillat, J.-P. and ... Winkler, M. (2021) Disentangling anthropogenic drivers of climate change impacts on alpine plant species: Alps vs. mediterranean mountains. final report. *Earth System Sciences (ESS)*, 1–72.

Lanza, L. G. and Stagi, L. (2008) Certified accuracy of rainfall data as a standard requirement in scientific investigations. *Advances in Geosciences*, **16**, 43–48.

Laouacheria, F., Kechida, S. and Chabi, M. (2019) Modelling the impact of design rainfall on the urban drainage system by storm water management model. *Journal of Water and Land Development*, **40**, 119–125.

Lavers, D. A., Simmons, A., Vamborg, F. and Rodwell, M. J. (2022) An evaluation of era5 precipitation for climate monitoring. *Quarterly Journal of the Royal Meteorological Society*, **148**, 3152–3165.

Liang, X. Z. (2022) Extreme rainfall slows the global economy. *Nature*, **601**, 193–194.

Ludwig, P., Ehmele, F., Franca, M. J., Mohr, S., Caldas-Alvarez, A., Daniell, J. E., Ehret, U., Feldmann, H., Hundhausen, M., Knippertz, P., Küpfer, K., Kunz, M., Mühr, B., Pinto, J. G., Quinting, J., Schäfer, A. M., Seidel, F. and Wisotzky, C. (2023) A multi-disciplinary analysis of the exceptional flood event of july 2021 in central europe - part 2: Historical context and relation to climate change. *Natural Hazards and Earth System Sciences*, **23**, 1287–1311.

Maurya, S. K., Kalhapure, A., Verma, V. K., Tiwari, A., Chaubey, C., Maurya, D. K. and Kumar, M. (2024) Irrigation scheduling and cultivar management for increasing water productivity under dryland condition: A review. *International Journal of Environment and Climate Change*, **14**, 461–470.

Merz, B., Nguyen, V. D., Guse, B., Han, L., Guan, X., Rakovec, O., Samaniego, L., Ahrens, B. and Vorogushyn, S. (2024) Spatial counterfactuals to explore disastrous flooding. *Environmental Research Letters*, **19**.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C. and Thépaut, J. N. (2021) Era5-land: A state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, **13**, 4349–4383.

Ommer, J., Neumann, J., Kalas, M., Blackburn, S. and Cloke, H. L. (2024) Surprise floods: The role of our imagination in preparing for disasters. *Natural Hazards and Earth System Sciences*, **24**, 2633–2646.

Richardson, D., Hemri, S., Bogner, K., Gneiting, T., Haiden, T., Pappenberger, F. and Scheuerer, M. (2014) Calibration of ecmwf forecasts. *ECMWF Newsletter*. URL: https://doi.org/10.21957/45t3o8fj.

Satgé, F., Defrance, D., Sultan, B., Bonnet, M. P., Seyler, F., Rouché, N., Pierron, F. and Paturel, J. E. (2020) Evaluation of 23 gridded precipitation datasets across west africa. *Journal of Hydrology*, **581**, 124412.

Schumacher, R. S. (2017) Heavy rainfall and flash flooding. *Oxford Research Encyclopedia of Natural Hazard Science*, **1**.

Stephens, M. A. (1974) Edf statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, **69**, 730.

Tadeyo, E., Chen, D., Ayugi, B. and Yao, C. (2020) Characterization of spatio-temporal trends and periodicity of precipitation over malawi during 1979-2015. *Atmosphere*, **11**.

Tie, J., Kim, S. and Sharma, A. (2023) How does increasing temperature affect the sub-annual distribution of monthly rainfall? *Environmental Research: Climate*, **2**.

Tye, M. R., Dagon, K., Molina, M. J., Richter, J. H., Visioni, D., Kravitz, B. and Tilmes, S. (2022) Indices of extremes: geographic patterns of change in extremes and associated vegetation impacts under climate intervention. *Earth System Dynamics*, **13**, 1233–1257.

Vannitsem, S., Bremnes, J. B., Demaeyer, J., Evans, G. R., Flowerdew, J., Hemri, S., Lerch, S., Roberts, N., Theis, S., Atencia, A., Bouallègue, Z. B., Bhend, J., Dabernig, M., de Cruz, L., Hieta, L., Mestre, O., Moret, L., Plenković, I. O., Schmeits, M. and ... Ylhaisi, J. (2021) Statistical postprocessing for weather forecasts review, challenges, and avenues in a big data world. *Bulletin of the American Meteorological Society*, **102**, E681–E699.

Wang, Y. and Karimi, H. A. (2022) Impact of spatial distribution information of rainfall in runoff simulation using deep learning method. *Hydrology and Earth System Sciences*, **26**, 2387–2403.

Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., Kendon, E. J., Lenderink, G. and Roberts, N. M. (2014) Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, **52**, 522–555.

WMO (2024) State of the global climate 2023. *WMO-No.1347*.

Zittis, G., Bruggeman, A. and Lelieveld, J. (2021) Revisiting future extreme precipitation trends in the mediterranean. *Weather and Climate Extremes*, **34**.