

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

Fatima M. Pillosu^{a,b} Timothy D. Hewson^b Estibaliz Gascòn^b Milana Vuckovic^b Christel
Prudhomme^{b,c,d} Hannah Cloke^{a,e}

^a *Department of Geography and Environmental Science, University of Reading, Reading, UK*

^b *Forecast and Services Department, European Centre for Medium-range Weather Forecasts,
Reading, UK*

^c *Department of Geography and Environment, University of Loughborough, Loughborough, UK*

^d *UK Centre for Ecology and Hydrology, Wallingford, United Kingdom*

^e *Department of Meteorology, University of Reading, Reading, UK*

¹¹ Corresponding author: Fatima Pillosu, tj835682@student.reading.ac.uk

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

35 SIGNIFICANCE STATEMENT: The purpose of this study is to understand better how modelled
36 rainfall datasets represent the distribution of point-scale rainfall as measured by rain gauges,
37 including extremes. We considered four modelled rainfall datasets, including ERA5 at 62 and
38 31 km, ECMWF reforecasts at 18 km, and the ecPoint post-processed version of ERA5 (ERA5-
39 ecPoint), which produces probabilistic point-scale rainfall estimates. All models were evaluated
40 against twenty years of rain gauge measurements worldwide. Our findings demonstrate that ERA5-
41 ecPoint largely improves the estimates of point-rainfall, correctly capturing both small and extreme
42 rainfall totals. This advancement is crucial since many regions lack (appropriate) rain gauge
43 coverage. Moreover, as climate change intensifies extreme rainfall globally, ERA5-ecPoint enables
44 planners to quantify rare, unseen events, providing crucial information for infrastructure design
45 and disaster preparedness.

46 1. Introduction

47 Accurately estimating the full range of past and future rainfall distributions, from light to extreme
48 totals, is one of the biggest challenges in modern meteorology. Yet, it is essential to address a
49 range of critical issues. In flood forecasting, accurately estimating the spatial distribution of small
50 and extreme rainfall totals influences the catchment response to the rainfall event, impacting runoff
51 generation and streamflow patterns (Cuo et al. 2011; Wang and Karimi 2022). In water resource
52 management, understanding the full rainfall distribution informs the management of reservoirs for
53 flood control, power generation, and irrigation purposes (Tie et al. 2023). It also helps in agricultural
54 applications such as crop selection and planting schedules (Janmohammadi and Sabaghnia 2023;
55 Maurya et al. 2024). It also supports urban planning by helping to design effective urban drainage
56 systems and manage storm-water runoff (Hossain et al. 2024; Laouacheria et al. 2019). Analysing
57 changes over time in the entire rainfall distribution provides insights into climate change impacts
58 such as shifts in the frequency and intensity of extreme rainfall events (Tye et al. 2022), droughts'
59 characteristics (Haile et al. 2020), biodiversity and ecosystems stability (Lamprecht et al. 2021),
60 and food security (Balasundram et al. 2023). Extreme rainfall, in particular, has received significant
61 attention in recent literature (Gimeno et al. 2022; Schumacher 2017) due to its catastrophic impacts
62 for society, infrastructure, and the environment (IPCC 2023; WMO 2024). It not only reduces
63 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

¹Note 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).

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 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 Pillou (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

a. 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

²Nowadays, we have reforecasts at 9 km resolution, but Hewson (2024) showed that extremes do not get much bigger in the 9 km reforecasts than in the previous lower resolution reforecasts at 18 km resolution.

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 to 4546.

b. Gridded NWP-modelled precipitation estimates

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

³<https://psl.noaa.gov/data/gridded/data.cpc.globalprecip.html>

148 enhances the continuity and stability of the climatological records. No precipitation observations
149 are assimilated into ERA5 (Hersbach et al. 2020).

150 The ERA5 Ensemble Data Assimilation (EDA) system enhances the robustness of the ERA5
151 reanalysis by generating multiple simulations with slightly varied initial conditions (Hersbach et al.
152 2020). Each ensemble member in ERA5 EDA provides an equally probable realisation of the
153 atmospheric state, quantifying the uncertainty associated with observational errors and limitations
154 within the forecasting model itself. ERA5-EDA has 10 ensemble members, running at 62 km
155 spatial resolution and 3-hour temporal resolution.

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

163 2) ECMWF REFORECASTS

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

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 $100 - \left(\frac{1}{\text{no. total real.}} \times 100 \right)$
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) <i>* required at least 75% of valid obs.</i>	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. <i>(in reality, we have only 66940 real. as some dates were not available)</i>	73000 real. / 365 days = 1 in a 200-year event (1 in a 100-year event)* <i>* due to the actual reduced number of total real.</i>	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)* <i>* preferred a round number for return period</i>	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)* <i>* preferred a round number for return period</i>	99.99972

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

179 equally probable precipitation realisations. This was possible as there was no drift in the forecasts
 180 up to day 10 (not shown). Hence, the precipitation distribution built with ECMWF reforecasts
 181 contains 20800 realisations (Table 1, row 4).

182 3) ERA5-ecPOINT

183 Within the Highlander project, co-financed by the EU and coordinated by Italy's Cineca su-
184 percomputing centre, ECMWF's ecPoint post-processing technique was applied to the raw ERA5
185 "deterministic" fields to address ERA5 limitations (Hewson et al. 2023). ecPoint aims to infer
186 sub-grid variability and to correct biases (both according to ongoing weather and geographical
187 scenarios). The ERA5-ecPoint dataset spans from 1950 to the near-present, providing a long-term,
188 continuously updated record of 24-hourly point-scale rainfall estimates with an accumulation pe-
189 riod ending at 00 UTC. The dataset is provided in a probabilistic format, i.e. in percentiles 1,2,...
190 99. Currently, ERA5-ecPoint represents point-scale precipitation, but it is provided on its native
191 (reduced Gaussian) grid at 31 km spatial resolution (TL639). The 99 percentiles can be considered
192 equally probable precipitation outcomes, at a gauge within a grid-box, so that the precipitation
193 distributions are built with 722700 daily realisations (Table 1, row 5).

194 **3. Methods**

195 *a. RQ1: assessment of the representation by NWP models of the overall distribution of point-*
196 *rainfall observations*

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

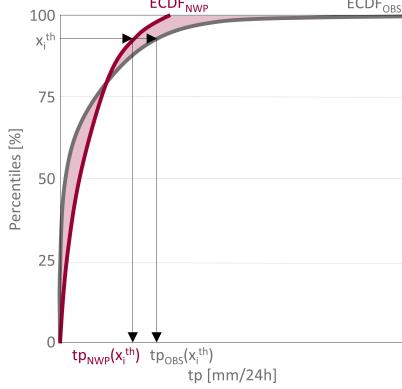
$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\text{tp}_{\text{OBS}}(x_i^{\text{th}}) - \text{tp}_{\text{NWP}}(x_i^{\text{th}})| \quad (1)$$

214 $n = 99$ percentiles

Normalized Mean Absolute Error (MAE_{NORM}) for 24-hourly total precipitation (tp)

Schematic on how to interpret MAE_{NORM}

(a) Schematic representation and computation of MAE and MAE_{NORM} at a specific rain gauge



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

n = tot n. of percentiles sampling ECDF

MAE is expressed in mm.

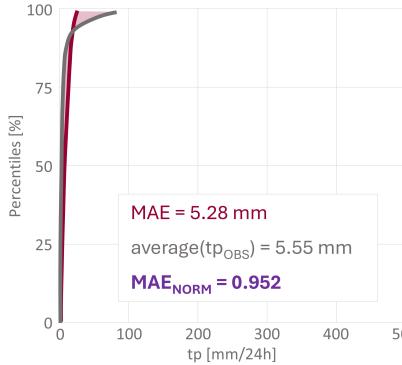
$$MAE_{\text{NORM}} = \frac{MAE}{\text{average}(tp_{OBS})}$$

$$\text{average}(tp_{OBS}) = \frac{1}{m} \sum_{i=1}^m (tp_{OBS})_i$$

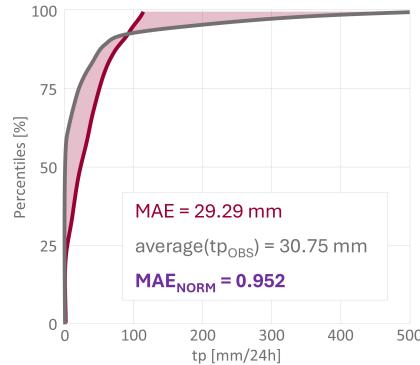
m = tot n. of observed records in the rain gauge

MAE_{NORM} is adimensional.

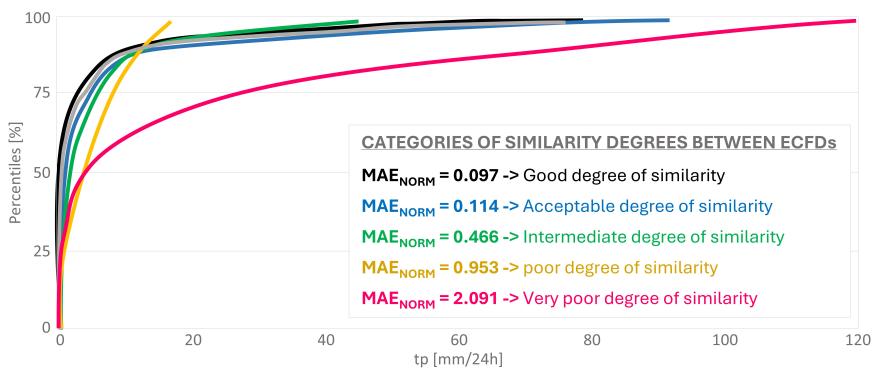
(b) Example of MAE and MAE_{NORM} in drier climates



(c) Example of MAE and MAE_{NORM} in wetter climates



(d) Examples of MAE_{NORM} values for different degrees of similarity between ECDFs



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

215 MAE values are expressed in mm and range from 0 (for perfect similarity) to $+\infty$ (for poor
216 similarity). Graphically, the MAE represents the areal difference between the two ECDFs (Figure
217 2a)⁴. Ninety-nine percentiles (1st to 99th) were used to avoid contamination related to sampling
218 issues in the observational dataset. Moreover, more extreme rainfall events will be considered
219 separately.

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

$$\text{MAE}_{\text{NORM}} = \frac{\text{MAE}}{\text{average}(\text{tp}_{\text{OBS}})} \quad (2)$$

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

233 Formal statistical tests such as Kolmogorov-Smirnov, Cramer-von-Mises, and Anderson-Darling
234 can also assess the similarity between two distributions (Stephens 1974). However, for large sample
235 sizes (as in this study, see Table 1, row 6), the tests' statistical significance levels become extremely
236 sensitive to minor differences between distributions that might not be practically significant (Eng-
237 mann and Cousineau 2011; Janssen 2000). This is being referred to as "the problem of practical
238 insignificance" (Kirk 1996), where the test flags differences that are statistically significant but not
239 meaningful in practice, causing the rejection of the null hypothesis (i.e., the two samples come
240 from the same population) when it is nonetheless practically valid. Gudmundsson's approach was

⁴The 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

241 tested to assess whether it was as sensitive to sample size as the formal statistical tests. The ECDFs
242 were sampled with 99, 999, 9999, and 99999 percentiles to assess the sensitivity of MAE_{NORM} to
243 the choice of sampling resolution. The results showed negligible differences across the range of
244 percentiles tested (not shown). Moreover, Gudmundsson's approach assesses similarity between
245 the observed and NWP-modelled rainfall distributions by comparing the whole ECDFs differently
246 to other formal tests that assess similarity only for specific moments of the distribution, such as
247 the mean, standard deviation, skewness, or specific percentiles (Anthanahalli Nanjegowda and
248 Kulamulla Parambath 2022). Finally, the use of MAE concerning other commonly used scores,
249 such as the Root Mean Squared Error (RMSE), is preferable as it is not unduly sensitive to outliers
250 (e.g., caused by erroneous observations or atypical events), typically observed in the wet tails of
251 the distribution. Hence, MAE should be more representative of the distribution as a whole (Jolliffe
252 and Stephenson 2011). Moreover, the RMSE is more appropriate when errors follow a normal
253 distribution, which is very atypical for rainfall (Chai and Draxler 2014). Nonetheless, the RMSE
254 was computed for the examples shown in Figure 2d, and its property of giving more weight to the
255 larger errors (in the wetter part of the distribution) did not change the ranking obtained with the
256 MAE_{NORM} between the different CDFs in 2d (not shown), reassuring the reader that using MAE
257 instead of RMSE should not change the final picture.

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

266 *b. RQ2: assessment of the representation by NWP models of extreme rainfall*

267 The assessment of how well NWP models represent extreme rainfall is conducted by visually
268 comparing the precipitation maps for the 10-year return period to pinpoint geographical differences
269 in estimating extreme precipitation. The 10-year return period was considered because it is the

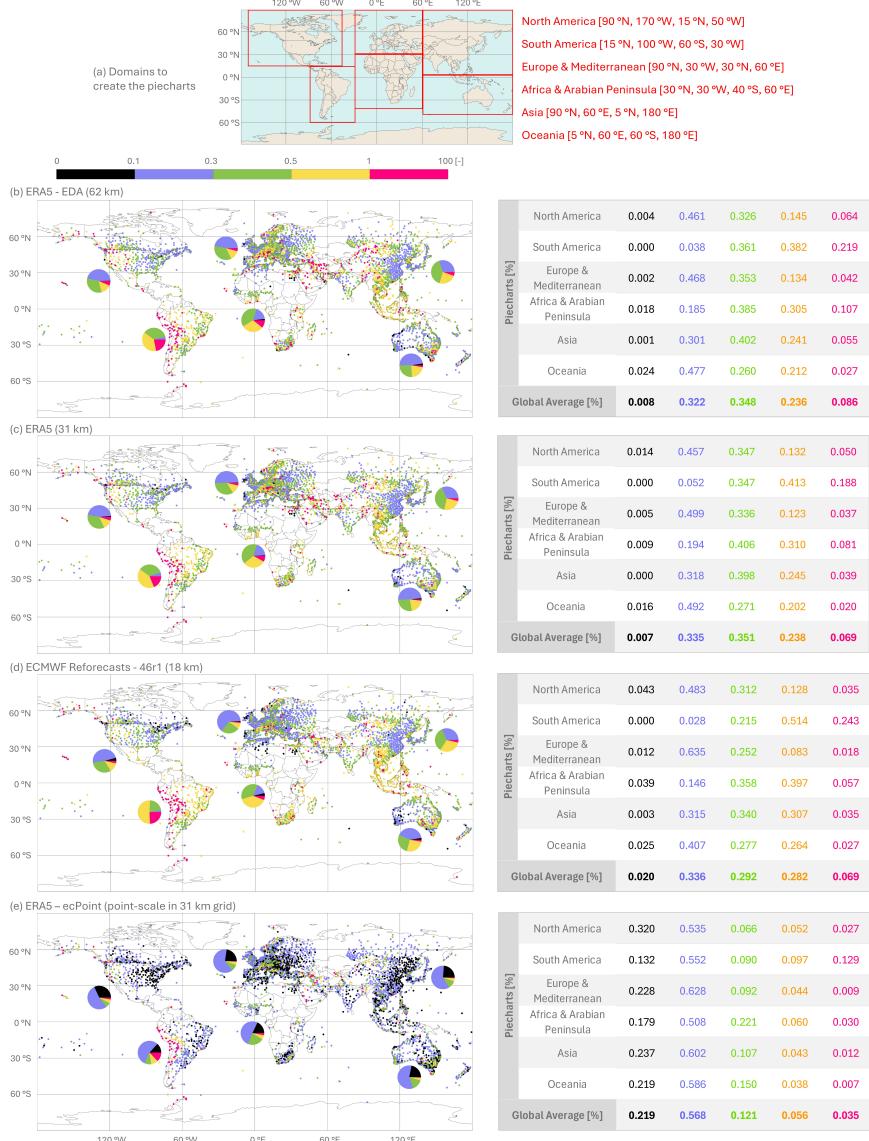
most extreme precipitation event that it was possible to compute with the observational dataset in hand (Table 1, columns 7 and 8, row 1), even though larger events could have been calculated with the NWP-modelled precipitation estimates (Table 1, columns 7 and 8, and rows 2 to 5). The 10-year return 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

a. 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

Normalized Mean Absolute Error (MAE_{NORM}) for 24-hourly total precipitation
 MAE_{NORM} values at each rain gauge station (left column) and piecharts' numerical values (right column)



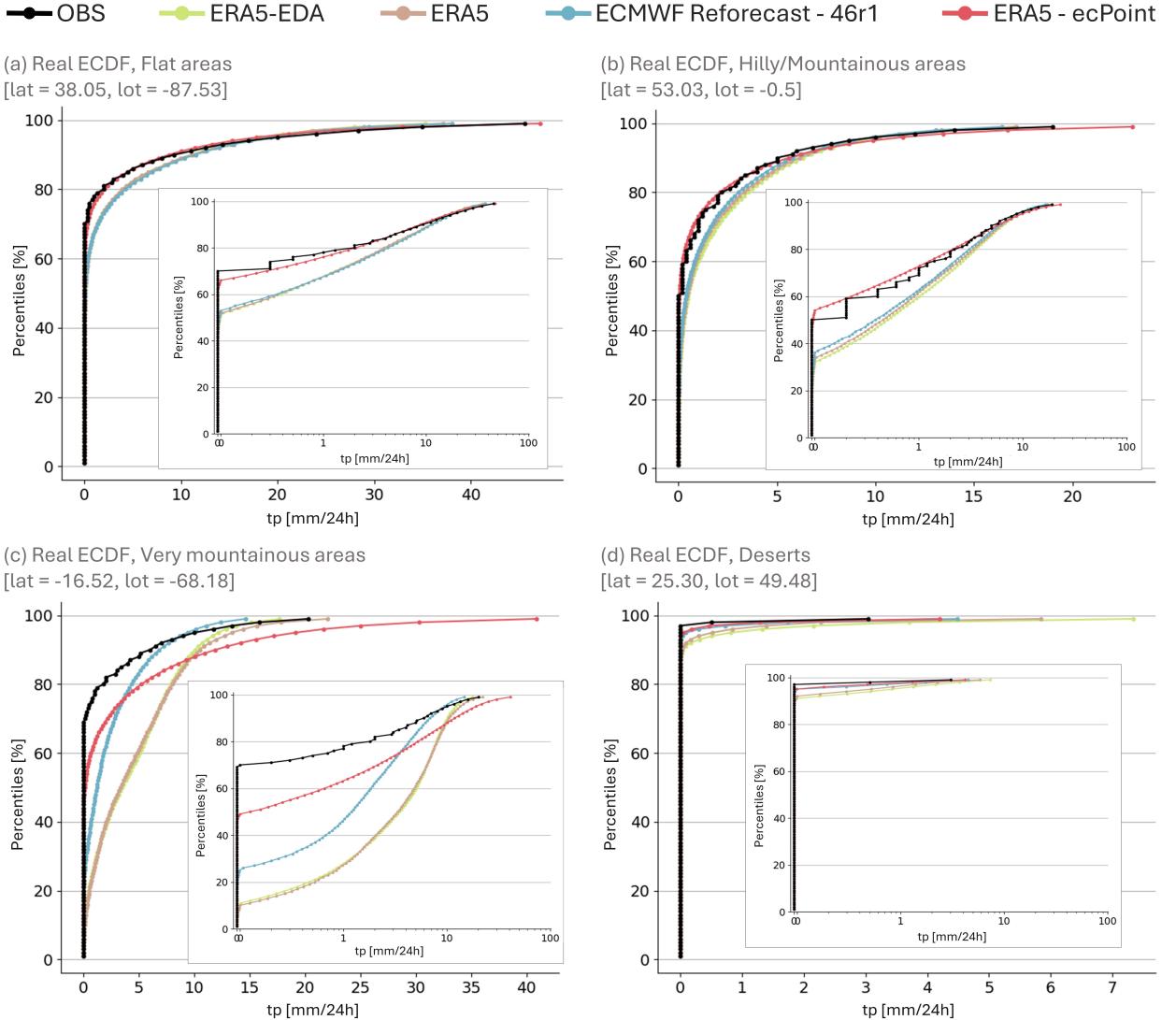
283 FIG. 3. Panels (b) to (e) indicate the Normalised Mean Absolute Error (MAE_{NORM}) for 24-hourly total
 284 precipitation at each rain gauge location for ERA5-EDA (62 km), ERA5 (31 km), ECMWF Rerofcasts-46r1
 285 (18 km), and ERA5-ecPoint (point-scale, but provided in ERA5's grid), respectively. Dots in black, blue, green,
 286 yellow, and pink represent, respectively, a "good", "acceptable", "intermediate", "poor", and "very poor" degree
 287 of similarity to the corresponding point observed climatology. The pie charts indicate the frequency of MAE_{NORM}
 288 values in the domains defined in panel (a). The tables on the right offer a numerical representation of the pie
 289 charts. The numbers in bold in the last rows of each table represent the global average for each representation
 290 category.

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 observations⁵ (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 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

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

Empirical Cumulative Distribution Functions (ECDFs) for 24-hourly total precipitation (tp)



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

342 tail of the observed ECDF (in black). Compared to point-rainfall observations, raw NWP models
 343 show behaviour similar to that observed in flat areas. The main difference lies in a progression in a
 344 better representation of the observed ECDF for NWP models with increasing spatial resolution, i.e.,

345 ERA5-EDA at 62 km (in green, in Figure 4b), which shows a worse representation of the observed
346 ECDF compared to ERA5 at 31 km (in brown), and ERA5 shows a worse representation than
347 reforecasts (in blue). This behaviour is seen in other sites too (not shown). In very mountainous
348 areas (Figure 4c), all NWP models fail to represent the observed ECDFs. It is worth noting that
349 this is not surprising as the observations used to train ERA5-ecPoint, and indeed to validate all
350 representations, come primarily from valleys and hilly areas. First, all the NWP model versions
351 underestimate the frequency of observed zero precipitation totals. ERA5-ecPoint tends to double
352 such a frequency, but it does not reach the values in the observed ECDFs. The ECDFs from raw
353 NWP models show a growth rate that is too large compared to the observed ECDFs, while ERA5-
354 ecPoint also improves on that. Finally, while the raw NWP models tend to slightly underestimate
355 the length of the observed ECDFs (with ERA5 providing the best representation out of the three
356 models), ERA5-ecPoint tends to overestimate it. In desert areas (Figure 4d), all NWP models
357 represent the observed ECDFs well, apart from the wet tails that tend to all be overestimated. The
358 overestimation is reduced with the increase in the spatial resolution of the NWP models, with
359 ERA5-ecPoint representing the actual length of the wet tail best.

360 *b. RQ2: comparison of the wet tail in the distributions built with NWP-modelled precipitation
361 estimates and rain gauge observations*

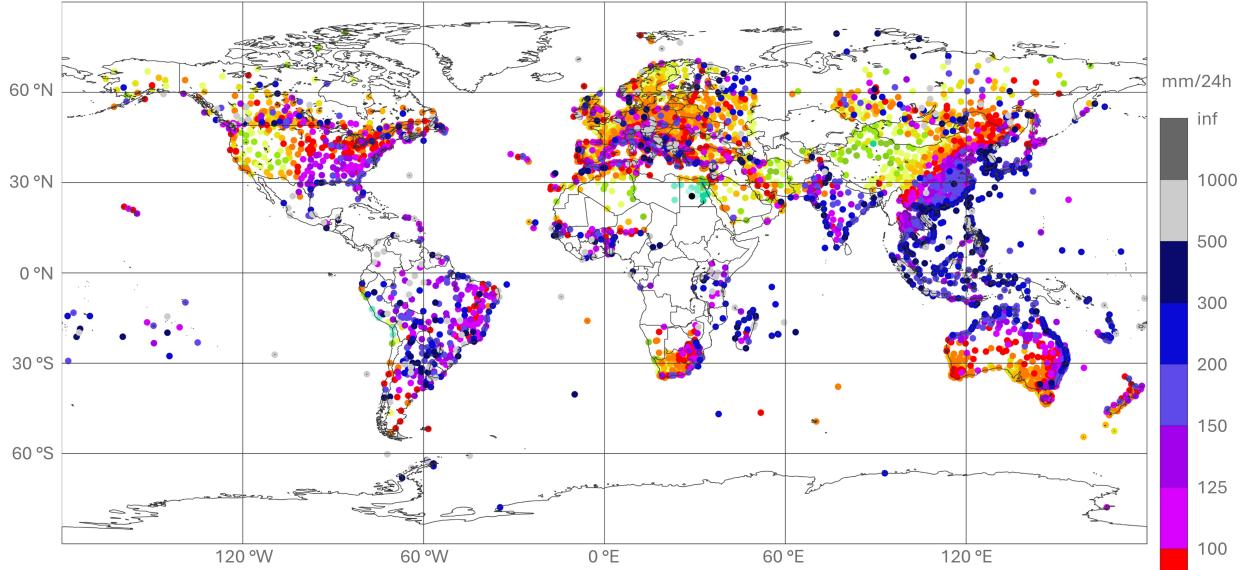
362 1) COMPARISON OF THE 10-YEAR PERIOD

363 The precipitation maps for the 10-year return period show that ERA5-ecPoint provides a better
364 representation than the raw NWP models. In North America, the extremes in 24-hourly precipitation
365 over the west coast of Alaska, Canada and North-West USA, which reach peaks up to 500 mm,
366 are better represented in ERA5-ecPoint than in ERA5-EDA, ERA5, and reforecasts that tend not
367 to exceed 125 mm. The peaks around the Gulf of Mexico, the USA's East Coast and the border
368 between Canada and the USA are also better represented in ERA5-ecPoint. However, in the latter
369 case, there seems to also be sampling-related noise in the observations. The raw NWP models
370 better represent the extremes over the Rocky Mountains since ERA5-ecPoint overestimates them.
371 However, the latter shows an overall closer representation of the observed ECDFs apart from the
372 tail. ERA5-ecPoint greatly improves the precipitation peaks over Mexico and South America over
373 the other three NWP models, apart from the Andean region and the desert on the west coast of Peru
374 over the other three NWP models, apart from the Andean region and the desert on the west coast of Peru
375 over the other three NWP models, apart from the Andean region and the desert on the west coast of Peru
376 over the other three NWP models, apart from the Andean region and the desert on the west coast of Peru
377 over the other three NWP models, apart from the Andean region and the desert on the west coast of Peru
378 over the other three NWP models, apart from the Andean region and the desert on the west coast of Peru
379 over the other three NWP models, apart from the Andean region and the desert on the west coast of Peru
380 over the other three NWP models, apart from the Andean region and the desert on the west coast of Peru

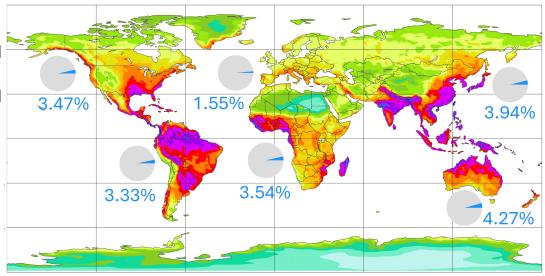
10-year return period for 24-hourly total precipitation [mm/24h]

Piecharts: percentage of NWP-modelled estimates exceeding the corresponding observed ones

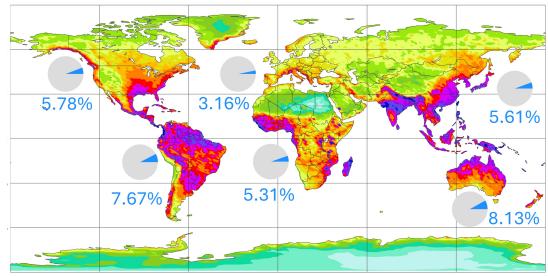
(a) Rain gauge observations (point scale), only rain gauges with at least 75% of valid record in the 20-year period



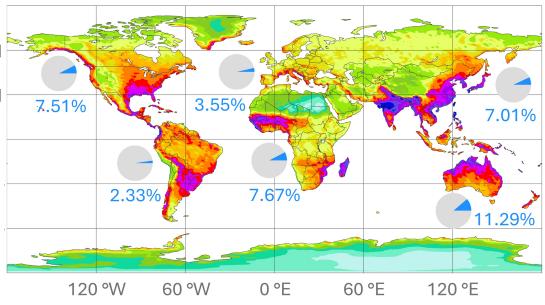
(b) ERA5 – EDA (62 km)



(c) ERA5 (31 km)



(d) ECMWF Reforecasts - 46r1 (18 km)



(e) ERA5 – ecPoint (point-scale in ERA's 31 km grid)

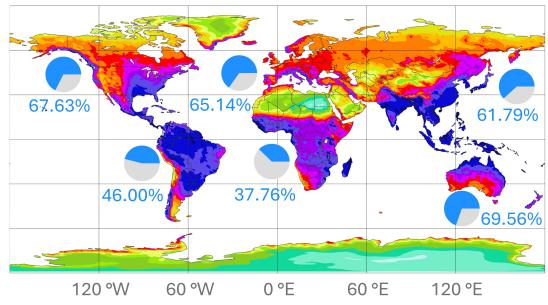


FIG. 5. Panel (a) displays the 10-year return period for 24-hourly total precipitation from rain gauge observations, calculated over the 20-year period between 2000 and 2019, and using only rain gauges with at least 75% of valid records. Panels (b) to (e) show the 10-year return period for NWP-modelled 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.

381 and Chile, where ERA5-ecPoint overestimates the wet tails (as shown in section 4.1). It is worth
382 noting that the ECMWF reforecasts from 46r1 halved the precipitation extremes over the Amazon
383 compared to ERA5-EDA and ERA5.

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

390 In Asia, there is a varied picture. The raw NWP models highlight the wetter climatologies of India
391 (especially the Northeast regions), East China, Japan, Southeast Asia, and the Malay Archipelago.
392 However, they do not reach the peaks of 300-500 mm/24h seen in the observations. ERA5-ecPoint
393 represents such peaks. However, the peaks greater than 500 mm/24h observed in the Malay
394 Archipelago remain underestimated, also in the post-processed ERA5. The overall overestimation
395 in the mountainous regions of Western China has a similar flavour to the ones discussed over
396 the Rocky Mountains in the USA: ERA5-ecPoint shows the best overall representation of the full
397 observed ECDFs, but tends to overestimate the wet tails. In the Arabian Peninsula, all models
398 represent the overall observed ECDF tails quite well. As discussed in section 4.1 for desert
399 areas, such good representation originates from the high frequency of zero precipitation totals
400 well estimated by all NWP models. The only exception is on the peninsula's south coast, where
401 precipitation peaks can reach 200 mm/24h, and raw NWP models estimate a maximum peak of only
402 up to 80 mm/24h. ERA5-ecPoint increases them up to 150 mm/24h. In Oceania, all NWP models
403 show a good overall representation of the observed ECDFs with slight underestimations of the
404 wet tails. The added value of ERA5-ecPoint in this region mainly provides a better representation
405 of the precipitation peaks. There are a few observations in Africa, and nothing can be said
406 about the model representation of precipitation extremes in the numerous ungauged areas of this
407 continent. All NWP models represent the wet climatology of West Africa, including its Atlantic
408 coast. However, ERA5-ecPoint best represents the observed local peaks that vary between 100
409 and 500 mm/24h. It is worth noting that ECMWF 46r1 reforecasts degrade the representation
410 of the extreme precipitation around the Gulf of Guinea by producing maximum peaks only up to

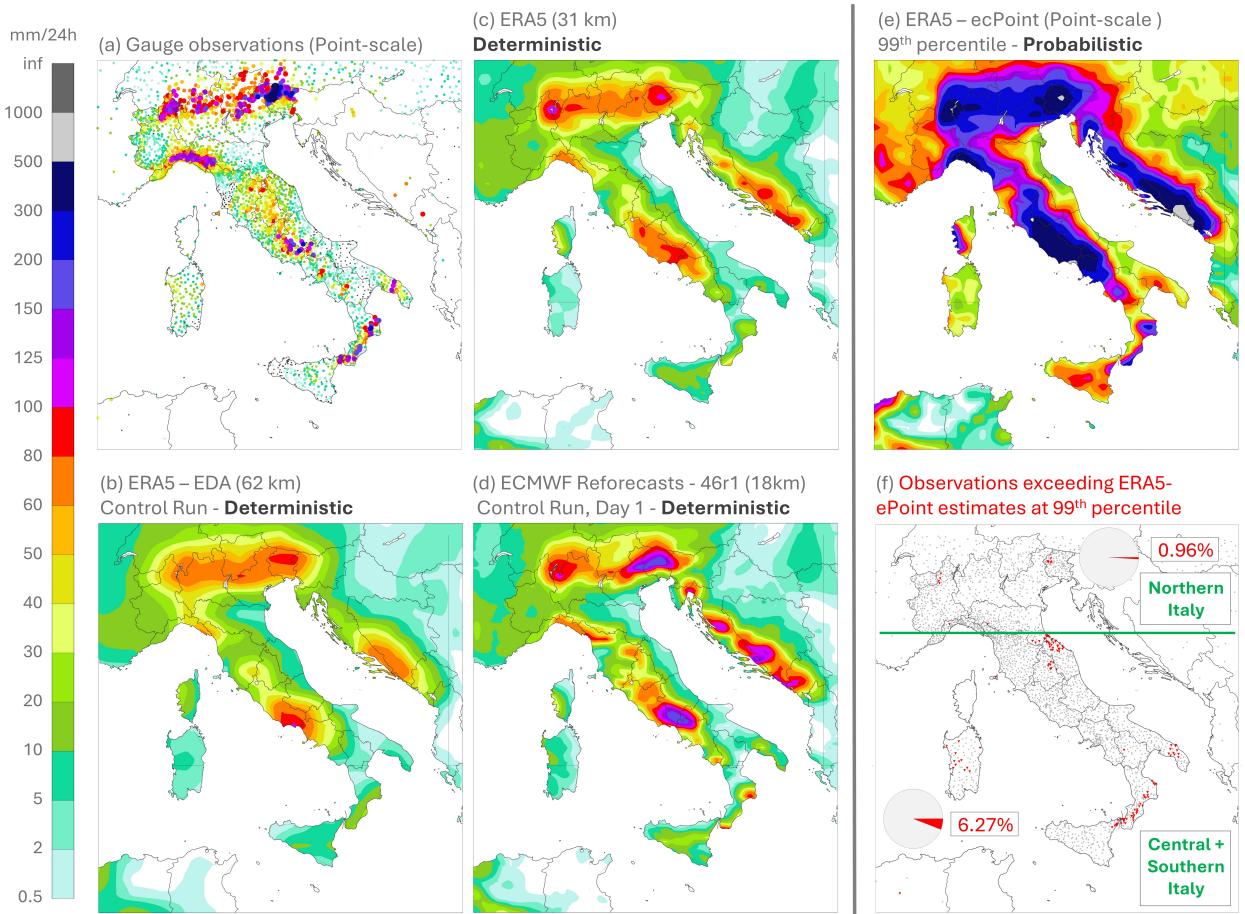
80-100 mm/24h. Similarly, out of all NWP models, ERA5-ecPoint somewhat better represents the varied precipitation peaks, between 80 and 500 mm/24h, in South Africa, where raw NWP models suggest extreme precipitation might not exceed 80 mm/24h. Also, in East Africa, ERA5-ecPoint provides a more realistic 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.

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 Lombardia (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 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

24-hourly total precipitation [mm/24h] for widespread flash floods in Italy (Storm Vaia)

Valued valid between 28th October 2018 at 00 UTC and 29th October 2018 at 00 UTC



444 FIG. 6. Widespread (flash) flooding in Italy on the 28th of October 2018 due to Storm Vaia. Panel (a)
445 represents the rain gauge observations. Panels (b) to (d) show the deterministic rainfall estimates, respectively,
446 for ERA5-EDA at 62 km (from control run), ERA at 31 km (single realisation), and ECMWF Reforecasts from
447 46r1 at 18 km (from control run, day 1 lead time). Panel (e) shows the probabilistic rainfall estimates from
448 ERA5-ecPoint (99th percentile). Panel (f) shows the locations of the rain gauge observations exceeding the
449 ERA5-ecPoint estimates at the 99th percentile. The representation is split into two geographical areas: Northern
450 and Southern Italy, with pie charts denoting total counts for these two areas.

440 the observations exceed the ERA5-ecPoint estimates, indicating an under-prediction of point-scale
441 rainfall over Le Marche, Puglia, and Calabria. In this specific case, the location of the red dots
442 along coastlines indicates underestimation primarily due to the known issue of convective cells
443 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 representativity 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

481 to ERA5, the 46r1 improvements are focused again on mountainous areas and extend to most of
482 Europe, the arid regions of Northern Africa, and the Arabian Peninsula.

483 Focusing on extreme rainfall events, there is a general increase in the values with the increase
484 of the raw NWP models' spatial resolution, which better agrees with the observed wetter tails.
485 The major difference is observed between ERA5-EDA and ERA5, while the differences between
486 the latter and the ECMWF reforecasts are less prominent. Indeed, for the rainfall in the Amazon
487 region, Equatorial Africa, and Indonesia, the reforecasts show rainfall estimates that do not exceed
488 100 mm/24h. In contrast, both reanalysis, ERA5-EDA and ERA5, show rainfall estimates up
489 to 300 mm/24h, which better represent the observed rainfall totals in the region. These results
490 similarly contradict expectations and may indicate regional limitations in cycle 46r1 employed for
491 the reforecast dataset.

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

505 Furthermore, ERA5-ecPoint enables one to estimate rainfall events with significantly longer
506 return periods than those presented in this study (e.g., up to a 1000-year return period, as noted in
507 Table 1, row 5, column 8). Hewson et al. (2024) have shown for 2023 Storm Daniel in Libya that
508 applying the ecPoint post-processing technique to ERA5 can deliver usable estimates of an n-year
509 return period rainfall from m years of data, where $n \gg m$. Consequently, datasets like ERA5-
510 ecPoint offer valuable insights into the potential magnitude of extreme rainfall events, improving

511 our preparedness for unseen events (Heinrich et al. 2024; Ommer et al. 2024) or ones so infrequent
512 that they have faded from collective memory (Ludwig et al. 2023; Merz et al. 2024).

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

526 The general improvements provided by ERA5-ecPoint open up significant opportunities across
527 various fields of environmental research that require a more accurate representation of point rainfall
528 estimates. We advocate that such improvements would enhance both long-term strategic planning
529 (e.g., using this dataset for climatological studies) and short-term emergency response (e.g., this
530 dataset to create point-scale rainfall thresholds that are compatible with ecPoint rainfall medium-
531 range forecasts to determine areas at risk of extreme localised rainfall), thereby contributing to
532 developing more resilient societies in the face of climate change. In the realm of flood forecasting,
533 more accurate rainfall estimates at local scales are crucial for predicting runoff and streamflow
534 dynamics, particularly in catchments prone to flash floods. Precise point-scale rainfall data is
535 pivotal in enhancing early warning systems, which are essential for safeguarding communities from
536 the severe impacts of extreme rainfall and flooding. Better rainfall representation could facilitate
537 more efficient management of reservoirs and irrigation planning in water resource management,
538 optimising water storage and distribution for agriculture, power generation, and urban water supply
539 systems. Furthermore, enhanced point-scale precipitation estimates are crucial for designing more
540 resilient stormwater infrastructure and urban drainage systems, which are facing increasing pressure

541 from the intensification of extreme rainfall events due to climate change. In the context of disaster
542 preparedness, ERA5-ecPoint's ability to capture the full spectrum of rainfall values, including zeros
543 and extremes, provides valuable insights into the risks posed by changing precipitation patterns.

544 5. Conclusions

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

548 This study provides a systematic, global verification of how well NWP models represent the
549 distribution of point-rainfall observations. It considered point-rainfall observations over 20 years
550 and four different modelled, gridded datasets with distinct spatial resolutions: ERA5-EDA (62
551 km), ERA5 (31 km), ECMWF Reforecasts for 46r1 (18 km), and ERA5-ecPoint (point-scale but
552 provided over ERA5's grid at 31 km). Among the tested models, this study shows that ERA5-
553 ecPoint most accurately captures both the frequency of zero rainfall totals and the wet tails of the
554 observed point-rainfall distributions.

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

565 The improved performance of ERA5-ecPoint over raw NWP models in representing point-scale
566 rainfall totals, whether small or large, emphasises post-processing's critical role in addressing
567 the inherent limitations of gridded rainfall estimates in guiding point-scale rainfall. ecPoint
568 effectiveness, however, remains contingent on the quality of the underlying NWP models it post-
569 processes. Without accurate raw NWP estimates at a grid-scale, the skill demonstrated by the

570 ERA5-ecPoint rainfall estimates would be diminished. Moving forward, the authors advocate
571 enhancing the spatial resolution and the skill of raw NWP models alongside ongoing improvements
572 of post-processing techniques such as ecPoint to reduce further errors in estimating the whole
573 distribution of point-rainfall totals. Such improvements will be particularly significant as climate
574 change intensifies the frequency and severity of extreme rainfall, making accurate and reliable point-
575 rainfall estimates indispensable for effective mitigation and response efforts related to droughts,
576 extreme rainfall, flooding, food security, and urban resilience.

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

580 *Data availability statement.* Data and/or the codes used to generate the figures incorporated into
581 this manuscript will be made available upon reasonable request. Unauthenticated — Downloaded
582 10/14/25 09:11 PM UTC

583 References

- 584 Alexandridis, V., S. Stefanidis, and S. Dafis, 2023: Evaluation of era5 and era5-land reanalysis
585 precipitation data with rain gauge observations in greece. *Environmental Sciences Proceedings*,
586 **26** (1), 104, <https://doi.org/10.3390/ENVIRONSCIPROC2023026104>.
- 587 Anand, B., and D. Karunanidhi, 2020: Long-term spatial and temporal rainfall trend analysis
588 using gis and statistical methods in lower bhavani basin, tamil nadu, india. *Indian Journal of*
589 *Geo-Marine Sciences*, **49** (3), 419–427.
- 590 Anthanahalli Nanjegowda, R., and S. Kulamulla Parambath, 2022: A novel bias correction method
591 for extreme rainfall events based on l-moments. *International Journal of Climatology*, **42** (1),
592 250–264, <https://doi.org/10.1002/JOC.7242>.
- 593 Asadieh, B., and N. Y. Krakauer, 2015: Global trends in extreme precipitation: Climate models
594 versus observations. *Hydrology and Earth System Sciences*, **19** (2), 877–891, <https://doi.org/10.5194/HESS-19-877-2015>.

- 596 Balasundram, S. K., R. R. Shamshiri, S. Sridhara, and N. Rizan, 2023: The role of digital agriculture
597 in mitigating climate change and ensuring food security: An overview. *Sustainability*, **15** (6),
598 <https://doi.org/10.3390/SU15065325>.
- 599 Bottazzi, M., and Coauthors, 2024: High performance computing to support land, climate, and
600 user-oriented services: The highlander data portal. *Meteorological Applications*, **31** (2), e2166,
601 <https://doi.org/10.1002/met.2166>.
- 602 Boé, J., L. Terray, F. Habets, and E. Martin, 2007: Statistical and dynamical downscaling of
603 the seine basin climate for hydro-meteorological studies. *International Journal of Climatology*,
604 **27** (12), 1643–1655, <https://doi.org/10.1002/joc.1602>.
- 605 Chai, T., and R. R. Draxler, 2014: Root mean square error (rmse) or mean absolute error (mae)?
606 -arguments against avoiding rmse in the literature. *Geoscientific Model Development*, **7** (3),
607 1247–1250, <https://doi.org/10.5194/GMD-7-1247-2014>.
- 608 Cuo, L., T. C. Pagano, and Q. J. Wang, 2011: A review of quantitative precipitation forecasts
609 and their use in short- to medium-range streamflow forecasting. *Journal of Hydrometeorology*,
610 **12** (5), 713–728, <https://doi.org/10.1175/2011JHM1347.1>.
- 611 Di Curzio, D., A. Di Giovanni, R. Lidori, M. Montopoli, and S. Rusi, 2022: Comparing rain
612 gauge and weather radar data in the estimation of the pluviometric inflow from the apennine
613 ridge to the adriatic coast (abruzzo region, central italy). *Hydrology*, **9** (12), 225, <https://doi.org/>
614 [10.3390/HYDROLOGY9120225](https://doi.org/10.3390/HYDROLOGY9120225).
- 615 Donat, M. G., L. V. Alexander, N. Herold, and A. J. Dittus, 2016: Temperature and precip-
616 itation extremes in century-long gridded observations, reanalyses, and atmospheric model
617 simulations. *Journal of Geophysical Research*, **121** (19), 11 174–11 189, <https://doi.org/>
618 [10.1002/2016JD025480](https://doi.org/10.1002/2016JD025480).
- 619 Doocy, S., A. Daniels, S. Murray, and T. D. Kirsch, 2013: The human impact of floods: A historical
620 review of events 1980-2009 and systematic literature review. *PLoS Currents*, <https://doi.org/>
621 [10.1371/CURRENTS.DIS.F4DEB457904936B07C09DAA98EE8171A](https://doi.org/10.1371/CURRENTS.DIS.F4DEB457904936B07C09DAA98EE8171A).

- 622 Engmann, S., and D. Cousineau, 2011: Comparing distributions: The two-sample anderson-
623 darling test as an alternative to the kolmogorov-smirnoff test. *Journal of Applied Quantitative*
624 *Methods*, **6** (3), URL https://jaqm.ro/issues/volume-6,issue-3/1_engmann_cousineau.php.
- 625 Ensor, L. A., and S. M. Robeson, 2008: Statistical characteristics of daily precipitation: Compar-
626 isons of gridded and point datasets. *Journal of Applied Meteorology and Climatology*, **47** (9),
627 2468–2476, <https://doi.org/10.1175/2008JAMC1757.1>.
- 628 Espinosa, L. A., M. M. Portela, and S. Gharbia, 2024: Assessing changes in exceptional rain-
629 fall in portugal using era5-land reanalysis data (1981/1982–2022/2023). *Water*, **16** (5), 628,
630 <https://doi.org/10.3390/W16050628>.
- 631 Gimeno, L., R. Sorí, M. Vázquez, M. Stojanovic, I. Algarra, J. Eiras-Barca, L. Gimeno-Sotelo, and
632 R. Nieto, 2022: Extreme precipitation events. *Wiley Interdisciplinary Reviews: Water*, **9** (6),
633 <https://doi.org/10.1002/WAT2.1611>.
- 634 Giorgos, N., P. T. Nastos, and Y. Kapsomenakis, 2024: Creating high-resolution precipitation and
635 extreme precipitation indices datasets by downscaling and improving on the era5 reanalysis data
636 over greece. *Eng*, **5** (3), 1885–1904, <https://doi.org/10.3390/ENG5030101>.
- 637 Gomis-Cebolla, J., V. Rattayova, S. Salazar-Galán, and F. Francés, 2023: Evaluation of era5 and
638 era5-land reanalysis precipitation datasets over spain (1951–2020). *Atmospheric Research*, **284**,
639 <https://doi.org/10.1016/J.ATMOSRES.2023.106606>.
- 640 Gudmundsson, L., J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, 2012: Technical note:
641 Downscaling rcm precipitation to the station scale using statistical transformations - a comparison
642 of methods. *Hydrology and Earth System Sciences*, **16** (9), 3383–3390, <https://doi.org/10.5194/hess-16-3383-2012>.
- 644 Gupta, V., M. K. Jain, P. K. Singh, and V. Singh, 2020: An assessment of global satellite-based
645 precipitation datasets in capturing precipitation extremes: A comparison with observed precipi-
646 tation dataset in india. *International Journal of Climatology*, **40** (8), 3667–3688, <https://doi.org/10.1002/JOC.6419>.
- 648 Haiden, T., and S. Duffy, 2016: Use of high-density observations in precipitation verification.
649 *ECMWF Newsletter*, (147), 20–25, <https://doi.org/10.21957/hsacrde>.

- 650 Haile, G. G., and Coauthors, 2020: Projected impacts of climate change on drought patterns over
651 east africa. *Earth's Future*, **8** (7), <https://doi.org/10.1029/2020EF001502>.
- 652 Hamill, T. M., J. S. Whitaker, and S. L. Mullen, 2006: Reforecasts: An important dataset for
653 improving weather predictions. *Bulletin of the American Meteorological Society*, **87** (1), 33–46,
654 <https://doi.org/10.1175/BAMS-87-1-33>.
- 655 Hamill, T. M., and Coauthors, 2022: The reanalysis for the global ensemble forecast system, version
656 12. *Monthly Weather Review*, **150** (1), 59–79, <https://doi.org/10.1175/MWR-D-21-0023.1>.
- 657 Heinrich, D., E. Stephens, and E. Coughlan de Perez, 2024: More than magnitude: Towards a mul-
658 tidimensional understanding of unprecedented weather to better support disaster management.
659 *Water Security*, **23**, 100 181, <https://doi.org/10.1016/J.WASEC.2024.100181>.
- 660 Herold, N., A. Behrangi, and L. V. Alexander, 2017: Large uncertainties in observed daily
661 precipitation extremes over land. *Journal of Geophysical Research: Atmospheres*, **122** (2),
662 668–681, <https://doi.org/10.1002/2016JD025842>.
- 663 Hersbach, H., and Coauthors, 2020: The era5 global reanalysis. *Quarterly Journal of the Royal
664 Meteorological Society*, **146** (730), 1999–2049, <https://doi.org/10.1002/qj.3803>.
- 665 Hewson, T., 2024: Capturing extreme rainfall events. *ECMWF Newsletter*, **(178)**.
- 666 Hewson, T., and M. Chevallier, 2024: Use and verification of ecmwf products. *ECMWF Technical
667 Memoranda*, **(919)**, <https://doi.org/10.21957/734d5d4d39>.
- 668 Hewson, T., and F. Pillosu, 2021: A low-cost post-processing technique improves weather
669 forecasts around the world. *Communications Earth and Environment*, **2** (1), <https://doi.org/10.1038/s43247-021-00185-9>.
- 671 Hewson, T., F. Pillosu, E. Gascòn, and M. Vučković, 2023: Post-processing era5 output with
672 ecpoint. *ECMWF Newsletter*, **(176)**.
- 673 Hewson, T., and Coauthors, 2024: Medicane daniel: an extraordinary cyclone with devastating
674 impacts. *ECMWF Newsletter*, **(179)**, <https://doi.org/10.21957/th3wxk861d>.

- 675 Hossain, I., S. Gato-Trinidad, M. Imteaz, and S. Rayburg, 2024: Future scenarios of design rainfall
676 due to upcoming climate changes in nsw, australia. *Atmosphere*, **15** (9), 1101, <https://doi.org/10.3390/ATMOS15091101>.
- 678 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*. IPCC,
679 Geneva, Switzerland, 35–115 pp., <https://doi.org/10.59327/IPCC/AR6-9789291691647>.
- 681 Janmohammadi, M., and N. Sabaghnia, 2023: Strategies to alleviate the unusual effects of climate
682 change on crop production: a thirsty and warm future, low crop quality. a review. *Biologija*,
683 **69** (2), 121–133, <https://doi.org/10.6001/BIOLOGIJA.2023.69.2.1>.
- 684 Janssen, A., 2000: Global power functions of goodness of fit tests. *The Annals of Statistics*, **28** (1),
685 239–253.
- 686 Jolliffe, I. T., and D. B. Stephenson, 2011: *Forecast Verification: A Practitioner’s Guide in Atmo-*
687 *spheric Science*. 2nd ed., John Wiley & Sons, Ltd., <https://doi.org/10.1002/9781119960003>.
- 688 Kalnay, E., and Coauthors, 1996: The ncep/ncar 40-year reanalysis project. *Bulletin of the Amer-*
689 *ican Meteorological Society*, **77**, 437–472, [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2).
- 690 Kidd, C., A. Becker, G. J. Huffman, C. L. Muller, P. Joe, G. Skofronick-Jackson, and D. B.
691 Kirschbaum, 2017: So, how much of the earth’s surface is covered by rain gauges? *Bulletin of the*
692 *American Meteorological Society*, **98** (1), 69–78, <https://doi.org/10.1175/BAMS-D-14-00283.1>.
- 694 Kirk, R. E., 1996: Practical significance: A concept whose time has come. *Educational and*
695 *Psychological Measurement*, **56** (5), 746–759, <https://doi.org/10.1177/0013164496056005002>.
- 696 Lamprecht, A., and Coauthors, 2021: Disentangling anthropogenic drivers of climate change
697 impacts on alpine plant species: Alps vs. mediterranean mountains. final report. *Earth System*
698 *Sciences (ESS)*, 1–72, <https://doi.org/10.1553/ESS-MEDIALPSS1>.
- 699 Lanza, L. G., and L. Stagi, 2008: Certified accuracy of rainfall data as a standard re-
700 quirement in scientific investigations. *Advances in Geosciences*, **16**, 43–48, <https://doi.org/10.5194/ADGEO-16-43-2008>.

- 702 Laouacheria, F., S. Kechida, and M. Chabi, 2019: Modelling the impact of design rainfall on
703 the urban drainage system by storm water management model. *Journal of Water and Land*
704 **Development**, **40** (1), 119–125, <https://doi.org/10.2478/JWLD-2019-0013>.
- 705 Lavers, D. A., A. Simmons, F. Vamborg, and M. J. Rodwell, 2022: An evaluation of era5
706 precipitation for climate monitoring. *Quarterly Journal of the Royal Meteorological Society*,
707 **148** (748), 3152–3165, <https://doi.org/10.1002/QJ.4351>.
- 708 Liang, X. Z., 2022: Extreme rainfall slows the global economy. *Nature*, **601** (7892), 193–194,
709 <https://doi.org/10.1038/d41586-021-03783-x>.
- 710 Ludwig, P., and Coauthors, 2023: A multi-disciplinary analysis of the exceptional flood event of july
711 2021 in central europe - part 2: Historical context and relation to climate change. *Natural Hazards*
712 and *Earth System Sciences*, **23** (4), 1287–1311, <https://doi.org/10.5194/NHESS-23-1287-2023>.
- 713 Maurya, S. K., A. Kalhapure, V. K. Verma, A. Tiwari, C. Chaubey, D. K. Maurya, and M. Kumar,
714 2024: Irrigation scheduling and cultivar management for increasing water productivity under
715 dryland condition: A review. *International Journal of Environment and Climate Change*, **14** (1),
716 461–470, <https://doi.org/10.9734/IJECC/2024/V14I13856>.
- 717 Merz, B., and Coauthors, 2024: Spatial counterfactuals to explore disastrous flooding. *Environmental*
718 *Research Letters*, **19** (4), <https://doi.org/10.1088/1748-9326/AD22B9>.
- 719 Muñoz-Sabater, J., and Coauthors, 2021: Era5-land: A state-of-the-art global reanalysis dataset
720 for land applications. *Earth System Science Data*, **13** (9), 4349–4383, <https://doi.org/10.5194/ESSD-13-4349-2021>.
- 722 Ommer, J., J. Neumann, M. Kalas, S. Blackburn, and H. L. Cloke, 2024: Surprise floods: The
723 role of our imagination in preparing for disasters. *Natural Hazards and Earth System Sciences*,
724 **24** (8), 2633–2646, <https://doi.org/10.5194/NHESS-24-2633-2024>.
- 725 Richardson, D., S. Hemri, K. Bogner, T. Gneiting, T. Haiden, F. Pappenberger, and M. Scheuerer,
726 2014: Calibration of ecmwf forecasts. *ECMWF Newsletter*, **(142)**, URL <https://doi.org/10.21957/45t3o8fj>.

- 728 Satgé, F., D. Defrance, B. Sultan, M. P. Bonnet, F. Seyler, N. Rouché, F. Pierron, and J. E. Paturel,
729 2020: Evaluation of 23 gridded precipitation datasets across west africa. *Journal of Hydrology*,
730 **581**, 124 412, <https://doi.org/10.1016/J.JHYDROL.2019.124412>.
- 731 Schumacher, R. S., 2017: Heavy rainfall and flash flooding. *Oxford Research Encyclopedia of*
732 *Natural Hazard Science*, **1**, <https://doi.org/10.1093/acrefore/9780199389407.013.132>.
- 733 Stephens, M. A., 1974: Edf statistics for goodness of fit and some comparisons. *Journal of the*
734 *American Statistical Association*, **69 (347)**, 730, <https://doi.org/10.2307/2286009>.
- 735 Tadeyo, E., D. Chen, B. Ayugi, and C. Yao, 2020: Characterization of spatio-temporal trends and
736 periodicity of precipitation over malawi during 1979-2015. *Atmosphere*, **11 (9)**, <https://doi.org/10.3390/ATMOS11090891>.
- 738 Tie, J., S. Kim, and A. Sharma, 2023: How does increasing temperature affect the sub-annual
739 distribution of monthly rainfall? *Environmental Research: Climate*, **2 (1)**, <https://doi.org/10.1088/2752-5295/ACB5B9>.
- 741 Tye, M. R., K. Dagon, M. J. Molina, J. H. Richter, D. Visioni, B. Kravitz, and S. Tilmes, 2022:
742 Indices of extremes: geographic patterns of change in extremes and associated vegetation
743 impacts under climate intervention. *Earth System Dynamics*, **13 (3)**, 1233–1257, <https://doi.org/10.5194/ESD-13-1233-2022>.
- 745 Vannitsem, S., and Coauthors, 2021: Statistical postprocessing for weather forecasts review,
746 challenges, and avenues in a big data world. *Bulletin of the American Meteorological Society*,
747 **102 (3)**, E681–E699, <https://doi.org/10.1175/BAMS-D-19-0308.1>.
- 748 Wang, Y., and H. A. Karimi, 2022: Impact of spatial distribution information of rainfall in runoff
749 simulation using deep learning method. *Hydrology and Earth System Sciences*, **26 (9)**, 2387–
750 2403, <https://doi.org/10.5194/HESS-26-2387-2022>.
- 751 Westra, S., and Coauthors, 2014: Future changes to the intensity and frequency of short-duration ex-
752 treme rainfall. *Reviews of Geophysics*, **52 (3)**, 522–555, <https://doi.org/10.1002/2014RG000464>.
- 753 WMO, 2024: State of the global climate 2023. *WMO-No.1347*.

⁷⁵⁴ Zittis, G., A. Bruggeman, and J. Lelieveld, 2021: Revisiting future extreme precipitation trends in
⁷⁵⁵ the mediterranean. *Weather and Climate Extremes*, **34**, <https://doi.org/10.1016/J.WACE.2021.100380>.