Evaluating extreme precipitation in gridded datasets with a novel station database in Morocco

Alexandre Tuel[1] and Nabil El Moçayd[2,*]

January 25, 2023

[1] Institute of Geography and Oeschger Centre for Climate Change Research, University of Bern, Switzerland

[2] Institute of Applied Physics, University Mohammed VI Polytechnic, Benguerir, Morocco

[*] Correspondence to: Nabil El Moçayd (nabil.elmocayd@um6p.ma)

Abstract

Morocco is a large country with complex terrain and many sparsely populated regions. With a semi-arid climate, it is highly vulnerable to floods driven by extreme precipitation, whose distribution is highly variable in space and time. Yet, this topic has received little attention. The limited availability of data has so far been the major obstacle to pursue such research in Morocco. Public gridded datasets offer good opportunities to overcome this problem. However, the use of such data should be handled with care, especially when applying extreme value theory. The present work aims at addressing this issue. First, we introduce and analyse a comprehensive set of 120 daily precipitation series which we assembled from different stakeholders in Morocco. Then, we perform quality control of the data and use extreme value statistics to infer trends and large return levels. Finally, we assess the accuracy of nine gridded satellite-based and reanalysis daily precipitation datasets using the station data. These results are intended as a first step towards a comprehensive understanding of extreme precipitation in Morocco, and can help select gridded datasets for future hydrometeorological research.

Keywords: Morocco, station observations, extreme precipitation, reanalysis, remote-sensing of precipitation

1 Introduction

Precipitation is a key variable in climate risk assessment. Its dynamics strongly determine flood risk, especially in arid to semi-arid regions where dry soils and the lack of vegetation prevent precipitation infiltration and often lead to flash floods (Saber and Habib, 2016). Morocco is a case in point. Located in the dry subtropics, at the north-western border of the Sahara desert, the country receives little precipitation overall, but the bulk of it often falls in just a handful of days. This, combined with a high vulnerability, makes floods one of the most frequent, deadly and destructive natural disasters in Morocco (World Bank, 2021; Loudyi et al., 2022). Dozens of destructive and deadly flooding events occurred in the last few decades: in the Ourika Valley in 1995, Al Hoceima in 2003, Driouch in 2008, across the Sebou river in 2010, around Ouarzazate and Sidi Ifni in 2014, Laayoune in 2016, near Tirzit in 2019 or Tetouan in 2021.

Efficient flood risk assessment and preparedness relies on the availability of reliable precipitation data at a fine spatio-temporal resolution. Precipitation extremes are of particular interest, since they are responsible for much of the flood events (Berghuijs et al., 2019). In Morocco, torrential rainfall often leads to floods, especially in the mountainous interior (El Khalki et al., 2018). A careful understanding and monitoring of precipitation is therefore critical to managing flood risk. Precipitation in Morocco is characterised by substantial variability in space and at all time scales. Most of the precipitation occurs north of the Atlas mountains in the winter half-year, when the North Atlantic storm track moves enough to the south to affect Morocco. Precipitation exhibits a large inter-annual variability, following the location of the storm track. Intra-seasonal variability is also considerable. Most of the wet-season precipitation is concentrated in a handful of days when storms hit Morocco, with long dry periods broken by short and intense downpours. While the dynamics of seasonal-mean precipitation in the current and future climates, and their impact on water availability,

have already been addressed by several previous studies (e.g., Knippertz et al., 2003; El Moçayd et al., 2020; Tuel, 2020), extreme precipitation in Morocco remains largely unexplored. Besides, climate projections indicate an overall decrease in extreme precipitation magnitude in the region (Pfahl et al., 2017). Since climate models miss much of Morocco's fine-scale orography and are potentially biased in representing the relationship between large-scale weather patterns and extreme precipitation (Driouech et al., 2009; Tuel, 2020), these projections are still uncertain and need to be better constrained by observations.

The distribution of precipitation is often derived from ground rain gauge observations at networks of meteorological stations. Stations provide direct precipitation measurements, but only at a limited number of locations, usually not dense enough to capture fine-scale variability. Stations are also often lacking in sparsely-populated areas and mountainous regions, where more precipitation falls. There, the accuracy of purely station-based precipitation estimates is thus greatly reduced (Kidd et al., 2017). One can turn instead to satellite-based or reanalysis datasets, many of which have a high spatio-temporal resolution and offer complete spatial coverage. The downside is that their accuracy is not guaranteed: satellite-based estimates always rely on some proxy for precipitation, like radar reflectivity or brightness temperatures, and in reanalyses, precipitation is almost exclusively a prognosis variable, *i.e.* not directly assimilated and therefore subject to parametrisation uncertainties. To assess their accuracy, gridded datasets must consequently be compared with station data. Station data itself is not free of problems – missing data, gauge undercatch, biases in extreme or trace precipitation, wrong readings, changes in instrumentation, unsuitable station location, poor station maintenance, among others – some of which can be addressed with quality control (Hunziker et al., 2017). Still, on average, it represents the best "ground truth" available to determine how accurate gridded products are.

Several studies have used station data to analyse the distribution of precipitation

and the accuracy of gridded precipitation datasets across Morocco. Milewski et al. (2015) and Ouatiki et al. (2017) compared TRMM Multisatellite Precipitation Analysis (TMPA) products to station data over North-Central Morocco and the Oum-Er-Rbia catchment in Central Morocco respectively, and found good agreement at monthly and annual timescales. The Tensift basin around Marrakech has attracted a lot of attention: Habitou et al. (2020) analysed the performance of the CHIRPS dataset in this basin; Saouabe et al. (2020) did the same for the GPM-IMERG dataset and Salih et al. (2022) for 7 satellite-based and reanalysis products. Tramblay et al. (2016) also compared 5 different gridded satellite-based datasets to station data and evaluated their accuracy as inputs to hydrological models in a small catchment of Northern Morocco. Most of these studies did not however look at extreme precipitation, conducting the comparison to station data mostly at the monthly scale. A few recent studies analysed the spatio-temporal variability in extreme precipitation with station data, notably modelling temporal trends and inferring large return levels (e.g., Filahi et al., 2016; El Alaoui El Fels et al., 2021; Driouech et al., 2021; Hadri et al., 2021). Driouech et al. (2009) also assessed the ability of regional climate models to correctly simulate the distribution of extreme precipitation across Morocco, but to our knowledge, no systematic study exists that quantifies the accuracy of the most common available gridded products.

Another important limitation of previous studies is the limited number of stations they use (usually a few to a few dozen, with a temporal coverage of $30-50$ years), in addition often restricted to the major cities or single catchments. The availability of reliable, long-term observation data in Morocco has indeed always been a major obstacle to research. Filahi et al. (2016) and Driouech et al. (2021) limit their analysis to major cities, mostly missing the wetter mountain regions, while El Alaoui El Fels et al. (2021) and Hadri et al. (2021) both only focus on small catchments of a few thousand to 20,000 km^2. Only Milewski et al. (2015) attempted to cover as much of Morocco as possible, by analysing 125 rain gauges spread across Northern Morocco.

This study focuses on extreme precipitation across Morocco. We introduce what is to our knowledge the most complete database of quality-controlled station observations of daily precipitation in Morocco. We briefly describe the spatio-temporal distribution of extreme precipitation. We then provide a comprehensive assessment of the accuracy of 9 commonly used gridded precipitation datasets in representing precipitation extremes.

2 Study area and data

2.1 Study area

Morocco is located in the dry subtropics, between the 21°N and 36°N parallels, at the junction between the Atlantic and Mediterranean coastlines. Its climate is overall warm and dry, though with important regional differences. The Atlas Range, one of Africa's major mountain chains, separates the desert interior, with less than 200 mm of precipitation a year on average, from the relatively wetter coastal plains of the north and northwest and their Mediterranean climate (Figure 1). With 600-800 mm of average annual precipitation, the Atlas and the smaller Rif mountains to the north are the wettest regions in the country. They are Morocco's water tower, from which the country's rivers take their source, storing much of the precipitation as snow during winter and releasing it during warmer months (Tuel, 2020). Precipitation in Morocco is not only scarce, but also highly variable in time. To the north, 80-85% of annual precipitation fall during the wet season, from November to April, when the region stands under the influence of the North Atlantic storm track (Knippertz et al., 2003; Tuel and Eltahir, 2018). A wet season typically consists of 5-10 storm episodes, so that precipitation is often highly concentrated within a handful of days (Born et al., 2008; Tuel, 2020). Precipitation also varies substantially from year to year and across decades, driven by variations in the North Atlantic Oscillation (NAO) that dictate the

mean location of the mid-latitude storm track.

2.2 Station data

We collected data at an ensemble of 120 stations distributed across Morocco (Figure 1). The data come from two main sources: the Moroccan Weather Service ("Direction Générale de la Météorologie", DGM) for 40 stations located in or near cities, and regional watershed agencies ("Agences des Bassins Hydrauliques", ABH) for most of the remaining stations. ABHs are responsible for water monitoring, allocation and management at the regional scale across Morocco.

At each station, daily precipitation accumulation (usually from 12pm to 12pm the next day) is measured in mm, with a precision of 0.1 mm. Snowfall is distinguished only at the DGM stations. Most stations are confined to the northern side of the Atlas, especially in the mountainous regions of the Rif (Sebou and Loukkos river basins) and of the High Atlas (Oum-Er-Rbia and Tensift river basins). Very few stations are available in the arid and sparsely populated regions to the east and south of the Atlas. The earliest station record begins in 1934, and the latest ends in August 2019. We give the station list along with detailed information in Table S1.

2.3 Gridded datasets

We analyse in this study gridded daily precipitation data from several different datasets: eight based on observations (satellite and/or gauge data) and one reanalysis dataset (Table 1). These datasets are not all independent since many rely on the same input data (from the same satellite, for instance) or because one dataset is corrected with another, often at the monthly scale.

- Tropical Rainfall Measuring Mission (TRMM) Version 7 3B42 daily precipitation data, available from 1998 to 2019 at 0.25° resolution. TRMM data incorporates microwave and radar precipitation estimates with infrared brightness observa-

tions, calibrated with station data on a monthly basis (Huffman et al., 2007). TRMM came to an end in 2015; post-2015 data comes from GPM (see below).

- Global Precipitation Measurement (GPM) mission data, available from June 2000 to near-present at 0.1° resolution. GPM began in 2014 as the successor mission to TRMM. Like TRMM, GPM applies precipitation retrieval algorithms to data from multiple satellite sensors (both passive and active) (Hou et al., 2014). We use the GPM Integrated Multi-satellitE Retrievals (IMERG) Version 06 Final Run data, which comes with a 3.5 month latency. Pre-2014 data comes from TRMM and was processed retrospectively with the IMERG algorithm.

- Global Precipitation Climatology Project (GPCP) daily precipitation analysis version 1.3 data at 1° resolution from October 1996 to near-present. GPCP precipitation estimates in the 40°N-S band are obtained from geosynchronous infrared brightness temperatures processed with a threshold-matched precipitation index (Huffman et al., 2001).

- Climate Hazards InfraRed Precipitation with Stations (CHIRPS) daily data, available from 1981 to near-present at 0.05° resolution. CHIRPS estimates precipitation by merging high-resolution remotely-sensed cold cloud duration data and station observations with a precipitation climatology based on station and satellite data, and geographical covariates (elevation, longitude, latitude) (Funk et al., 2015).

- Multi-Source Weighted-Ensemble Precipitation (MSWEP) V2.2 data, available at 0.1° resolution from 1979 to near-present. MSWEP is based on an optimal merging of station data, satellite-based estimates and reanalysis output (Beck et al., 2019).

- Global Satellite Mapping of Precipitation (GSMaP) Version 6 data at 0.1° res-

olution, available from 2000/3 to near-present. A product of the JAXA Global Rainfall Watch, GSMaP precipitation is based on combined microwave-infrared data from a range of satellites (including from the TRMM and GPM constellations) (Kubota et al., 2020). The version 6 data we use was obtained with the Microwave-IR Merged Algorithm (Ushio and Kachi, 2010) applied to remotely-sensed data after 2014 and to Japanese 55-year Reanalysis data from 2000-2014, before operational data became available.

- Climate Prediction Center Morphing Technique (CMORPH) data at 0.25° resolution, available from January 1998 to near-present (Xie et al., 2019). CMORPH is based on retrievals of precipitation rates from many passive microwave satellite measurements that are bias-corrected with CPC daily rain gauge data over land and GPCP data over the ocean. The daily data is based on an hourly version at a resolution of 8×8 km.

- Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) Climate Data Record daily data at 0.25° resolution, available from 1983 to near-present. Precipitation estimates in PERSIANN are obtained from infrared brightness temperature data used as input to a neural network trained on the National Centers for Environmental Prediction (NCEP) stage IV radar-based precipitation data over the continental united States (Ashouri et al., 2015). Precipitation is then corrected at the monthly scale to equal the monthly 2.5° GPCP data.

- ERA5 reanalysis data, available at 0.25° resolution from 1979 to near-present. ERA5 is the ECMWF's latest reanalysis dataset and provides hourly precipitation forecasts based on the Integrated Forecasting System (IFS) cycle 41r2 (Hersbach et al., 2020). Precipitation in ERA5 is therefore not assimilated but only a prognosis variable.

Dataset	Type	Resolution	Period	Reference
CHIRPS	Satellite + stations	0.05°	1981-2019	Funk et al. (2015)
CMORPH	Satellite + stations	0.25°	1998-2019	Xie et al. (2019)
ERA5	Reanalysis	0.25°	1979-2019	Hersbach et al. (2020)
GPM	Satellite + station correction	0.1°	2000/6-2019	Hou et al. (2014)
GPCP	Satellite	0.25°	1996/10-2019	Huffman et al. (2001)
GSMaP	Satellite + reanalysis	0.1°	2000/3-2019	Kubota et al. (2020)
MSWEP	Satellite + stations + reanalysis	0.1°	1979-2019	Beck et al. (2019)
PERSIANN	Satellite	0.25°	1983-2019	Ashouri et al. (2015)
TRMM	Satellite + station correction	0.25°	1998-2019	Huffman et al. (2007)

Table 1: Overview of the gridded datasets used in this study.

3 Methods

3.1 Quality control

We begin by implementing basic quality control checks on the station data following Durre et al. (2010), which we briefly summarise here:

(1) **Basic integrity checks.** They consist in duplication, extreme value and identical value-streak checks. We check for duplication by testing whether all values are the same for two different years, and whether all values in a month are the same as in any other month of the same year, or any other same calendar month in the full record. We exclude months with no precipitation (a frequent occurrence in Morocco). Extreme value checks consist in flagging daily precipitation values above the established world record (1828 mm/day) or below 0 mm. Finally, we flag streaks with identical or frequent (non-zero and non-missing) precipitation values as in Durre et al. (2010).

(2) **Outlier checks.** Outliers are identified by looking for gaps larger than 300mm in the distribution of daily precipitation within each month. We also flag daily values

larger than 9 times the 95$^{\text{th}}$ wet-day percentile.

(3) **Spatial consistency checks.** They consist in identifying observations at a given station which fall significantly outside the range of simultaneous ($\pm$ 1 day) values at neighbouring stations. Neighbouring stations are all available stations within 75 km of the reference station and with non-missing data on the target dates; we require at least 3. Observations at the target station are flagged if they fall outside the range of neighbour values and if they differ from the closest neighbour value by a threshold inversely related to the difference between the climatological percentiles of daily precipitation totals at the target and neighbour stations. This helps detect both dubiously high precipitation values and false zero values (when neighbouring stations all record substantial amounts of precipitation). We run this check twice, removing flagged station values from potential neighbours in the second step so that they may not influence the consistency check results.

3.2 Extreme value analysis

We define extreme daily precipitation events as days when cumulative precipitation exceeds its 99th all-day percentile. Percentiles are calculated empirically for each dataset and location. Results remain qualitatively unchanged with slightly lower percentiles (97.5th or 95th). Choosing an all-day percentile helps avoid the issue of biases in wet-day frequency linked to trace precipitation amounts in satellite and reanalysis datasets. To explore the statistics of extreme daily precipitation, we use concepts of extreme value theory. At each station, we fit Generalized Extreme Value (GEV) distributions to annual maxima of daily precipitation, and Generalized Pareto (GP) distributions to exceedances of daily precipitation above its 99th all-day percentile (Coles, 2001). The

cumulative distribution functions of the GEV and GP distributions are

$$\begin{aligned} F_{\mathrm{GEV}}(x;\mu,\sigma,\xi) &= 1-\left(1+\xi\tfrac{x-\mu}{\sigma}\right)^{-1/\xi} \text{ for } 1+\xi\tfrac{x-\mu}{\sigma}>0 \\ F_{\mathrm{GP}}(x;\sigma,\xi) &= 1-\left(1+\xi\tfrac{x}{\sigma}\right)^{-1/\xi} \text{ for } x>0 \end{aligned} \tag{1}$$

where μ, σ and ξ are the location, shape and scale parameters respectively. In $F_{\mathrm{GP}}(x;\sigma,\xi)$, x represents exceedances above the threshold u: $x = y-u$ conditioned on $y>u$, where y is the series of daily precipitation totals and u its 99th all-day percentile. Although the GEV and GP distributions are theoretically related, the GP fits rely on 3-4 times more data than the GEV fits, and thus tend to yield more robust estimates of high return periods. The GEV nevertheless provides a convenient framework to assess potential temporal trends in extreme precipitation magnitude (see below). This is why we consider both distributions to characterise the extremal behaviour of our precipitation series.

For each station, we systematically discard years with more than 5% missing or flagged data. Additionally, for the GEV/GPD fits, we only keep stations with at least 25 years of data. This is a compromise between data availability, rather constrained by our limited data coverage, and robustness of the distribution fits. The parameters are obtained through maximum likelihood estimation. We assess goodness-of-fit visually with quantile-quantile plots, and statistically by implementing Kolmogorov-Smirnoff tests (Massey, 1951).

From the GPD fits, we estimate extreme return levels (10, 20 and 30 years) by solving

$$\begin{aligned} x^{p}_{\mathrm{GEV}} &= (F_{\mathrm{GEV}}(x))^{-1}(p) \\ x^{p}_{\mathrm{GP}} &= (F_{\mathrm{GP}}(x))^{-1}(p) \end{aligned} \tag{2}$$

where p is the probability of exceedance equal to the reciprocal of the associated return period (e.g., 1/3652 for 10 years).

We also test, at first order, for the presence of temporal trends in extreme daily pre-

cipitation by allowing the location parameter μ to depend linearly on time in the GEV annual maxima fit:

$$\mu(t) = \mu_0 + \beta_\mu t \tag{3}$$

where t is the year index. We assess the significance of the trend by computing the Akaike Information Criterion (AIC) for both the stationary and the time-dependent fits. AIC accounts for both model goodness-of-fit and parameter count, with a smaller AIC indicating a better fit. We restrict the trend analysis to stations with at least 30 years of data between 1980 and 2019 so that trends at different stations are calculated over a similar time period.

3.3 Evaluation metrics

We compare gridded datasets against station observations by assigning each station to the nearest grid cell for each dataset separately. Two neighbouring stations can therefore fall within the same grid cell, especially for the coarser-resolution datasets. However, if a station is located too close to the edge of a grid cell (by less than 1/10th of the cell's spatial resolution), we assign the average value of the closest 2 grid cells. We assess the reliability of gridded products with three simple metrics:

1. Relative bias in the 99th percentile:

$$b = 100 \times \frac{q_{\mathrm{g}}^{99} - q_{\mathrm{s}}^{99}}{q_{\mathrm{s}}^{99}} \tag{4}$$

where q_{g}^{99} and q_{s}^{99} are the 99^{th} all-day percentiles of daily precipitation at a station and its corresponding grid cell respectively.

2. Critical success index (CSI) (Schaefer, 1990): the CSI is an overall measure of the ability of the gridded datasets to correctly capture the timing of extreme

precipitation events. It is defined as

$$\text{CSI} = \frac{\text{Hits}}{\text{Hits} + \text{Misses} + \text{False alarms}} \tag{5}$$

where "Hits" is the number of correctly detected extremes (both station and gridded dataset are extreme), "Misses" the number of undetected extremes (station is extreme but gridded dataset is not extreme), and "False alarm" the number of incorrectly detected extremes (station is not extreme but gridded dataset is extreme). Because the station reporting times are unknown, we allow for an extra day around each station extreme to calculate the number of hits.

3. Ratio of contribution of extreme precipitation to total annual precipitation:

$$r = \frac{\sum_t p_{\text{g}}(t) \mathbb{1}\{p_{\text{g}}(t) > q_{\text{g}}^{99}\} / \sum_t p_{\text{g}}(t)\}}{\sum_t p_{\text{s}}(t) \mathbb{1}\{p_{\text{s}}(t) > q_{\text{s}}^{99}\} / \sum_t p_{\text{s}}(t)} \tag{6}$$

where $p_{\text{g}}(t)$ and $p_{\text{g}}(t)$ are the daily series of precipitation at a station and its corresponding grid cell respectively.

We calculate these metrics on the periods when station and gridded data intersect, only if the gridded dataset shares at least 10 years of data with the station. Again, this is a compromise between the limited station data availability and the need to obtain meaningful and robust bias and CSI estimates.

4 Results

4.1 Quality control and data coverage

We calculate the fraction of missing data for each station between the first day of the month with the earliest non-missing data point and the last day of the month with the latest non-missing data point. On average, 10% of the data is missing across stations,

with a range of 0-58%. The June-August period accounts for almost half (41%) of the missing data, which is not unsurprising since it corresponds to the driest period of the year when most stations will not record a single day with precipitation. Back when stations were manually operated (up to the early 2000s in many parts of Morocco), data were therefore likely frequently not collected during the dry months.

The quality control algorithms flag an average 0.3% of the data, a figure consistent with GHCN-Daily observations (0.24%) (Durre et al., 2010). This percentage ranges from 0 to 2.3% between stations. Most flagged data occurs between May and September (65% of the flagged data). 93% of the flags are due to duplication checks at the monthly scale. Transcription errors from written station records to digital format are likely responsible. It is also possible that missing months (most frequent in summer) were manually replaced with data from non-missing months. The amount of quality-controlled data increases substantially from about 20 stations in the early 1970s to 100-110 stations around 2000 (Figure 1).

4.2 Extreme value analysis in station data

Figure 2 shows the distribution of extreme percentiles and annual maxima of daily precipitation accumulation across stations. This distribution unsurprisingly follows that of average precipitation: the largest values occur in the north – Morocco's wettest region – particularly in the Rif mountains and along the Atlas Range (Knippertz et al., 2003; Born et al., 2008; Tuel, 2020). Extremes are also slightly more intense along the Atlantic coast than on the plains further inland. Regions to the east and south of the Atlas have low to very low extreme percentiles (< 10 mm) and mean annual maxima (< 20 mm).

A visual inspection of the marginal GEV and GPD fits shows a reasonable goodness-of-fit (Figure S1). KS test results confirm this: all stations pass the test for both GEV and GDP fits at a conservative 10% confidence level. Inferred daily precipitation return

levels are shown on Figure 3. The highest values occur along the Mediterranean and Atlantic coastlines, and in the Rif Mountains and Atlas Range. Some northern and mountain stations even have 10-year return levels above 100 mm (Figure 3-a). The plains in Northwestern Morocco have generally small return levels (< 70 mm), even at the 30-year timescale, as do the desert regions to the east and south of the Atlas. The ratio of 30- to 10-year return levels is overall highest (≈ 1.5) for stations in the Sahara and along the Mediterranean coast, and lowest ($\approx$ 1.1-1.2) in the Atlas and Northwestern Morocco.

Out of 47 stations with sufficient data (30 full years between 1980 and 2019), 11 exhibit noticeable temporal trends in their annual daily precipitation maxima, based on the AIC of stationary and non-stationary fits (section 3.2; Figure 4). 10 of these trends are positive. The remaining 36 stations split somewhat evenly into positive and negative trends, with trend magnitudes generally between -1 and +1 mm/decade.

4.3 Comparison of gridded datasets

Figures 5, 6 and 7 show the spatial distribution of the three selected metrics we use to compare gridded products with the station data. The number of locations at which we compare station and gridded data slighlty varies depending on the gridded dataset: a station must have 10 years of common data with a gridded product. However, almost all (105 out of 120) stations have sufficient data to evaluate each dataset (see Figure S2 for the comparison of gridded datasets on these 105 common stations only).

Biases in extreme percentiles generally vary in space with little consistency across the datasets. In ERA5, the bias is negative, though limited (absolute value below 40%) at almost all stations (Figure 5-a). By contrast, CMORPH and GSMAP tend to underestimate extreme percentiles in the Atlas Range and overestimate them in the desert regions to the south (Figure 5-d,f). TRMM, GPM and GPCP have negative biases along the Atlas Range, and mostly positive elsewhere (Figure 5-b,c,g). For CHIRPS,

biases are high (>40%) over much of Northern Morocco, but small elsewhere (Figure 5-h), while PERSIANN mostly underestimates extreme percentiles, with a few positive biases at Saharan stations. MSWEP performs well, with small biases (absolute value below 20%) almost everywhere. By comparing biases at the same stations, MSWEP also stands out, along with GSMaP and ERA5, while other datasets tend to have higher biases with more inter-station variability (Figure S2-a).

The ratio of relative contribution of extreme to annual precipitation does not always reflect biases in extreme percentiles (Figure 6). South of the Atlas, for instance, all datasets underestimate this contribution while they varyingly over- or underestimate extreme percentile magnitudes (Figure 5). In Northwestern Morocco, the ratio is positive in TRMM, GPM, CMORPH, GPCP and CHIRPS, very negative in ERA5 and PERSIANN, and weakly negative in GSMaP and MSWEP. Focusing on the 50 common stations, we again find that GSMaP and MSWEP perform well, along with GPCP and GPM, for which the inter-station variability is however higher (Figure S2-c). PERSIANN and TRMM perform the worst. CSI values also highlight the good performance of GSMaP and MSWEP, with many CSI values above 0.4 (Figures 7 and S2-b). ERA5 also performs well, especially in the Sahara (Figure 7-a), while PERSIANN, GPCP and CHIRPS have poor accuracy across Morocco (Figures 7-e,g,h and S2-b).

5 Discussion

5.1 Challenges with data quality and coverage

While our station dataset is one of the most complete ever put together for Morocco, it still suffers from several limitations. The first is of course related to its limited temporal coverage. Consistent, long-term daily precipitation records are hard to find in Morocco. Out of 120 stations in our database, 24 (20%) have less than 10 full years of quality-controlled data, and only 18 (15%) have more than 40 years, with a median

value of 28 years. The longest records often correspond to urban centers, but a number of stations at relatively high altitudes in the Rif and Atlas, far away from the population centers, also have records of more than 30 years (Figure 1). In recent decades, weather stations have been increasingly automated across Morocco, but the rate of deployment remains uneven, and real-time data is not available for many of the stations. Additionally, station data is not collected in a centralised way in Morocco – each basin agency is responsible for its own network, and the most recent data is not automatically collected by a single entity. This explains why data coverage drops by about 20 stations around 2005 in our dataset (Figure 1).

The second major limitation is the uneven and generally low spatial density of stations. The total number of rain gauges across Morocco in the early 2000s reached 375 (Loudyi et al., 2022), but many of these were discontinued or had only short records, or their data was never digitised. As things stand now, there are few currently operating stations left to include in our inventory. Stations are unsurprisingly few in the largely unpopulated and dry Sahara desert, though even there floods are a major concern whose risk remains badly quantified (Loudyi et al., 2022). The absence of stations is problematic since most of the flood-driving precipitation occurs inland in uninhabited areas. Station density is also low in important agricultural regions in Northern Morocco, like the Doukkala plain between Casablanca and Marrakech. Historically, stations were preferentially installed in the mountains, where most of the precipitation falls, to better assess the potential for irrigation and hydropower dams. Yet given its spatial variability (El Moçayd et al., 2020), precipitation in the plains is also relevant for rainfed agriculture, especially wheat, and for flash floods in many regions of the country. New weather stations have recently been installed by the Moroccan weather service, and could be added to the map as they start collecting more data.

Short records make it difficult to robustly estimate temporal trends in extreme precipitation, all the more so as inter-annual variability is high in Morocco and the signal-to-noise ratio probably low (Tuel, 2020). Additionally, many of the records begin the

1980s, a historically dry decade. One should keep this in mind when looking at Figure 4 – the conclusion being that given the data currently available, it is difficult to conclude whether any meaningful trends in extreme precipitation have occurred in Morocco. One possibility to make the analysis more meaningful would be to include the daily NAO index as a covariate in equation 3, since the NAO is the main driver of precipitation variability in Morocco (Knippertz et al., 2003; Tuel, 2020). Another approach consists in pooling together the values of nearby stations to get more robust parameter values – for instance within a latent variables framework (Cooley et al., 2007).

5.2 Ranking gridded datasets

We find clear differences in accuracy across the nine gridded datasets we analysed. While no dataset is perfect, some are clearly better or worse than others in most respects. MSWEP and GSMaP stand out regarding both the timing and magnitude of daily extremes. GSMaP is slightly more accurate in the timing, but its inter-station variance is higher, and its temporal coverage much lower than that of MSWEP (Table 1). ERA5 and GPM also perform well, especially for the timing of extremes and their contribution to annual precipitation. TRMM, CMORPH and GPCP have a lower, somewhat comparable accuracy, though extreme percentiles tend to be too high in TRMM and too low in CMORPH (Figures 5, S2-a), and GPCP is clearly less reliable for the timing of extremes (Figures 7, S2-b). We also find a clear improvement from TRMM to GPM (Figures 6, 7 and S2). Last in line are CHIRPS and PERSIANN which have bad scores in all three metrics.

Of course, our analysis is limited to precipitation extremes. Other characteristics of the precipitation distribution matter in practice and the datasets could perform better or worse in these respects. Nevertheless, the point of our study was to look at extremes specifically, since the literature was largely missing. Other studies have assessed the performance of gridded precipitation datasets at monthly or annual timescales across

Morocco (see introduction).

All the datasets tend to perform most poorly in the High Atlas. All – excepted CHIRPS – underestimate extreme percentile magnitudes (Figure 5) and most have low CSI values (Figure 7) in this region. ERA5 does a comparatively good job there, but probably overestimates wet-day frequency and small precipitation amounts at high altitudes. Indeed, its contribution of extreme to annual precipitation is too low although it correctly simulates extreme percentiles (Figures 5-a, 6-a). Elevation significantly modulates accuracy, though at varying degrees across the datasets. ERA5 in particular is less affected. The timing of extremes tends to be lower at higher altitudes (Figure S3), although this is not always the case for relative extreme percentile biases (not shown). Still, only 15% of stations are located at an altitude higher than 1000 m, and it remains difficult to infer what biases at high altitudes really are. How accurate the datasets are in the Saharan part of Morocco is also not easy to determine due to stations being really scarce across this vast area (Figure 1). ERA5 and MSWEP and, to a lesser extent, TRMM, GPM and GSMaP, appear to be reliable there, but a more extensive analysis, perhaps articulated around case studies, could help make a better assessment.

6 Conclusion

We introduce in this study what is to our knowledge the most comprehensive quality-controlled daily precipitation dataset for Morocco. The data come from various sources and cover different time periods. About two thirds of stations have reliable data between 1980 and 2015. We estimate extreme percentiles and 10- to 30-year return levels of daily precipitation, and provide a comprehensive assessment of the accuracy of nine commonly used gridded daily precipitation datasets with respect to extreme precipitation in Morocco. MSWEP and GSMaP perform best overall. Since station data is

particularly limited in both space and time in Morocco, and is not available in detail, gridded datasets are important tools for climate and flood risk analyses, and hydrological or agricultural modeling. Our results are therefore important to better guide the choice of gridded datasets in future research and quantify associated uncertainties.

Availability of Data and Materials

Station data are available from the corresponding author upon request. Table 2 shows how the various gridded datasets can be accessed.

Dataset	**URL**
CHIRPS	`https://data.chc.ucsb.edu/products/CHIRPS-2.0/` [last accessed July 1, 2022]
CMORPH	`http://doi.org/10.25921/w9va-q159`
ERA5	`http://doi.org/10.24381/cds.adbb2d47`
GPCP	`http://doi.org/10.7289/V5RX998Z`
GPM	`http://doi.org/10.5067/GPM/IMERGDF/DAY/06`
GSMaP	`https://sharaku.eorc.jaxa.jp/GSMaP/index.htm` [last accessed July 1, 2022]
MSWEP	`http://www.gloh2o.org/mswep/` [last accessed July 1, 2022]
PERSIANN	`http://doi.org/10.7289/V51V5BWQ`
TRMM	`http://doi.org/10.5067/TRMM/TMPA/DAY/7`

Table 2: Links to the gridded datasets used in this study.

Funding

The authors acknowledge partial funding from the Office Chérifien des Phosphates (OCP) through Université Mohammed VI Polytechnic.

Author contributions

A.T.: Conceptualisation; Methodology; Data Curation; Formal analysis; Visualisation; Writing N.E.M.: Conceptualisation; Data collection and curation; Writing; Funding

acquisition.

Acknowledgments

The authors thank M. D. Hasnaoui and D. Ouazar for their help in collecting station data.

References

Ashouri, H., K.-L. Hsu, S. Sorooshian, D. K. Braithwaite, K. R. Knapp, L. D. Cecil, B. R. Nelson, and O. P. Prat, 2015: PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bulletin of the American Meteorological Society*, **96 (1)**, 69–83, doi:10.1175/BAMS-D-13-00068.1, URL https://journals.ametsoc.org/doi/10.1175/BAMS-D-13-00068.1.

Beck, H. E., E. F. Wood, M. Pan, C. K. Fisher, D. G. Miralles, A. I. J. M. van Dijk, T. R. McVicar, and R. F. Adler, 2019: MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, **100 (3)**, 473–500, doi:10.1175/BAMS-D-17-0138.1, URL https://journals.ametsoc.org/view/journals/bams/100/3/bams-d-17-0138.1.xml.

Berghuijs, W. R., S. Harrigan, P. Molnar, L. J. Slater, and J. W. Kirchner, 2019: The Relative Importance of Different Flood-Generating Mechanisms Across Europe. *Water Resources Research*, **55 (6)**, 4582–4593, doi:10.1029/2019WR024841, URL https://onlinelibrary.wiley.com/doi/10.1029/2019WR024841.

Born, K., A. H. Fink, and H. Paeth, 2008: Dry and wet periods in the northwestern Maghreb for present day and future climate conditions. *Meteorologische Zeitschrift*, **17 (5)**, 533–551, doi:10.1127/0941-2948/2008/0313.

Coles, S., 2001: *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, Springer London, London, doi:10.1007/978-1-4471-3675-0, URL http://link.springer.com/10.1007/978-1-4471-3675-0.

Cooley, D., D. Nychka, and P. Naveau, 2007: Bayesian Spatial Modeling of Extreme Precipitation Return Levels. *Journal of the American Statistical Association*, **102 (479)**, 824–840, doi:10.1198/016214506000000780, URL http://www.

tandfonline.com/doi/abs/10.1198/016214506000000780.

Driouech, F., M. Déqué, and A. Mokssit, 2009: Numerical simulation of the probability distribution function of precipitation over Morocco. *Climate Dynamics*, **32 (7-8)**, 1055–1063, doi:10.1007/s00382-008-0430-6, URL http://link.springer.com/10.1007/s00382-008-0430-6.

Driouech, F., H. Stafi, A. Khouakhi, S. Moutia, W. Badi, K. ElRhaz, and A. Chehbouni, 2021: Recent observed country-wide climate trends in Morocco. *International Journal of Climatology*, **41 (S1)**, doi:10.1002/joc.6734, URL https://onlinelibrary.wiley.com/doi/10.1002/joc.6734.

Durre, I., M. J. Menne, B. E. Gleason, T. G. Houston, and R. S. Vose, 2010: Comprehensive Automated Quality Assurance of Daily Surface Observations. *Journal of Applied Meteorology and Climatology*, **49 (8)**, 1615–1633, doi:10.1175/2010JAMC2375.1, URL http://journals.ametsoc.org/doi/10.1175/2010JAMC2375.1.

El Alaoui El Fels, A., M. E. M. Saidi, A. Bouiji, and M. Benrhanem, 2021: Rainfall regionalization and variability of extreme precipitation using artificial neural networks: a case study from western central Morocco. *Journal of Water and Climate Change*, **12 (4)**, 1107–1122, doi:10.2166/wcc.2020.217, URL https://iwaponline.com/jwcc/article/12/4/1107/73613/Rainfall-regionalization-and-variability-of.

El Khalki, E. M., Y. Tramblay, M. El Mehdi Saidi, C. Bouvier, L. Hanich, M. Benrhanem, and M. Alaouri, 2018: Comparison of modeling approaches for flood forecasting in the High Atlas Mountains of Morocco. *Arabian Journal of Geosciences*, **11 (15)**, 410, doi:10.1007/s12517-018-3752-7, URL http://link.springer.com/10.1007/s12517-018-3752-7.

El Moçayd, N., S. Kang, and E. A. B. Eltahir, 2020: Climate change impacts on the water highway project in morocco. *Hydrology and Earth System Sciences*, **24 (3)**, 1467–

1483, doi:10.5194/hess-24-1467-2020, URL https://hess.copernicus.org/articles/24/1467/2020/.

Filahi, S., M. Tanarhte, L. Mouhir, M. El Morhit, and Y. Tramblay, 2016: Trends in indices of daily temperature and precipitations extremes in Morocco. *Theoretical and Applied Climatology*, **124 (3-4)**, 959–972, doi:10.1007/s00704-015-1472-4, URL http://link.springer.com/10.1007/s00704-015-1472-4.

Funk, C., and Coauthors, 2015: The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, **2 (1)**, 150 066, doi:10.1038/sdata.2015.66, URL http://www.nature.com/articles/sdata201566.

Habitou, N., A. Morabbi, D. Ouazar, A. Bouziane, M. D. Hasnaoui, and H. Sabri, 2020: CHIRPS precipitation open data for drought monitoring: application to the Tensift basin, Morocco. *Journal of Applied Remote Sensing*, **14 (3)**, 034 526, doi:10.1117/1.JRS.14.034526.

Hadri, A., M. E. M. Saidi, T. Saouabe, and A. El Alaoui El Fels, 2021: Temporal trends in extreme temperature and precipitation events in an arid area: case of Chichaoua Mejjate region (Morocco). *Journal of Water and Climate Change*, **12 (3)**, 895–915, doi:10.2166/wcc.2020.234, URL https://iwaponline.com/jwcc/article/12/3/895/75303/Temporal-trends-in-extreme-temperature-and.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146 (730)**, 1999–2049, doi:10.1002/qj.3803, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803.

Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bulletin of the American Meteorological Society*, **95 (5)**, 701–722, doi:10.1175/BAMS-D-13-00164.1, URL http://journals.ametsoc.org/doi/10.1175/BAMS-D-13-00164.1.

Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations. *Journal of Hydrometeorology*, **2 (1)**, 36–50, doi:10.1175/1525-7541(2001)002⟨0036:GPAODD⟩2.0.CO;2, URL http://journals.ametsoc.org/doi/10.1175/1525-7541(2001)002{%}3C0036:GPAODD{%}3E2.0.CO;2.

Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *Journal of Hydrometeorology*, **8 (1)**, 38–55, doi:10.1175/JHM560.1, URL http://journals.ametsoc.org/doi/10.1175/JHM560.1.

Hunziker, S., and Coauthors, 2017: Identifying, attributing, and overcoming common data quality issues of manned station observations. *International Journal of Climatology*, **37 (11)**, 4131–4145, doi:10.1002/joc.5037, URL https://onlinelibrary.wiley.com/doi/10.1002/joc.5037.

Kidd, C., A. Becker, G. J. Huffman, C. L. Muller, P. Joe, G. Skofronick-Jackson, and D. B. Kirschbaum, 2017: So, How Much of the Earth's Surface Is Covered by Rain Gauges? *Bulletin of the American Meteorological Society*, **98 (1)**, 69–78, doi:10.1175/BAMS-D-14-00283.1, URL https://journals.ametsoc.org/doi/10.1175/BAMS-D-14-00283.1.

Knippertz, P., M. Christoph, and P. Speth, 2003: Long-term precipitation variability in Morocco and the link to the large-scale circulation in recent and future climates. *Meteorology and Atmospheric Physics*, **83 (1-2)**, 67–88, doi:10.1007/s00703-002-0561-y.

Kubota, T., and Coauthors, 2020: Global Satellite Mapping of Precipitation (GSMaP) Products in the GPM Era. 355–373, doi:10.1007/978-3-030-24568-9_20, URL https://link.springer.com/10.1007/978-3-030-24568-9{_}20.

Loudyi, D., M. D. Hasnaoui, and A. Fekri, 2022: *Flood Risk Management Practices in*

Morocco: Facts and Challenges, 35–94. Springer Singapore, Singapore, doi:10.1007/978-981-16-2904-4_2, URL https://doi.org/10.1007/978-981-16-2904-4{_}2.

Massey, F. J., 1951: The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, **46 (253)**, 68–78, doi:10.1080/01621459.1951.10500769, URL http://www.tandfonline.com/doi/abs/10.1080/01621459.1951.10500769.

Milewski, A., R. Elkadiri, and M. Durham, 2015: Assessment and comparison of TMPA satellite precipitation products in varying climatic and topographic regimes in Morocco. *Remote Sensing*, **7 (5)**, 5697–5717, doi:10.3390/rs70505697.

Ouatiki, H., A. Boudhar, Y. Tramblay, L. Jarlan, T. Benabdelouhab, L. Hanich, M. R. El Meslouhi, and A. Chehbouni, 2017: Evaluation of TRMM 3B42 V7 rainfall product over the Oum Er Rbia watershed in Morocco. *Climate*, **5 (1)**, doi:10.3390/cli5010001.

Pfahl, S., P. A. O'Gorman, and E. M. Fischer, 2017: Understanding the regional pattern of projected future changes in extreme precipitation. *Nature Climate Change*, **7 (6)**, 423–427, doi:10.1038/nclimate3287, URL http://www.nature.com/articles/nclimate3287.

Saber, M., and E. Habib, 2016: *Flash Floods Modelling for Wadi System: Challenges and Trends*, 317–339. Springer International Publishing, Cham, doi:10.1007/978-3-319-18787-7_16, URL https://doi.org/10.1007/978-3-319-18787-7_16.

Salih, W., A. Chehbouni, and T. E. Epule, 2022: Evaluation of the Performance of Multi-Source Satellite Products in Simulating Observed Precipitation over the Tensift Basin in Morocco. *Remote Sensing*, **14 (5)**, 1171, doi:10.3390/rs14051171, URL https://www.mdpi.com/2072-4292/14/5/1171.

Saouabe, T., E. M. El Khalki, M. E. M. Saidi, A. Najmi, A. Hadri, S. Rachidi, M. Jadoud, and Y. Tramblay, 2020: Evaluation of the GPM-IMERG Precipitation

Product for Flood Modeling in a Semi-Arid Mountainous Basin in Morocco. *Water*, **12 (9)**, 2516, doi:10.3390/w12092516, URL https://www.mdpi.com/2073-4441/12/9/2516.

Schaefer, J. T., 1990: The Critical Success Index as an Indicator of Warning Skill. *Weather and Forecasting*, **5 (4)**, 570–575, doi:10.1175/1520-0434(1990)005⟨0570:TCSIAA⟩2.0.CO;2, URL http://journals.ametsoc.org/doi/10.1175/1520-0434(1990)005{%}3C0570:TCSIAA{%}3E2.0.CO;2.

Tramblay, Y., V. Thiemig, A. Dezetter, and L. Hanich, 2016: Evaluation of satellite-based rainfall products for hydrological modelling in Morocco. *Hydrological Sciences Journal*, **61 (14)**, 2509–2519, doi:10.1080/02626667.2016.1154149, URL https://www.tandfonline.com/doi/full/10.1080/02626667.2016.1154149.

Tuel, A., 2020: Precipitation variability and change over Morocco and the Mediterranean. Ph. d. thesis in hydrology, Massachusetts Institute of Technology, 287 pp., URL http://34.201.211.163/handle/1721.1/129036.

Tuel, A., and E. A. Eltahir, 2018: Seasonal Precipitation Forecast Over Morocco. *Water Resources Research*, **54 (11)**, 9118–9130, doi:10.1029/2018WR022984.

Ushio, T., and M. Kachi, 2010: *Kalman Filtering Applications for Global Satellite Mapping of Precipitation (GSMaP)*, 105–123. Springer Netherlands, Dordrecht, doi:10.1007/978-90-481-2915-7_7, URL https://doi.org/10.1007/978-90-481-2915-7_7.

World Bank, 2021: Morocco: Vulnerability. URL https://climateknowledgeportal.worldbank.org/country/morocco/vulnerability.

Xie, P., R. Joyce, S. Wu, S.-H. Yoo, Y. Yarosh, F. Sun, R. Lin, and N. C. Program, 2019: NOAA Climate Data Record (CDR) of CPC Morphing Technique (CMORPH) High Resolution Global Precipitation Estimates, Version 1 [CMORPH 0.25° Daily].

NOAA National Centers for Environmental Information, URL https://doi.org/10.25921/w9va-q159, doi:10.25921/w9va-q159.

Figures

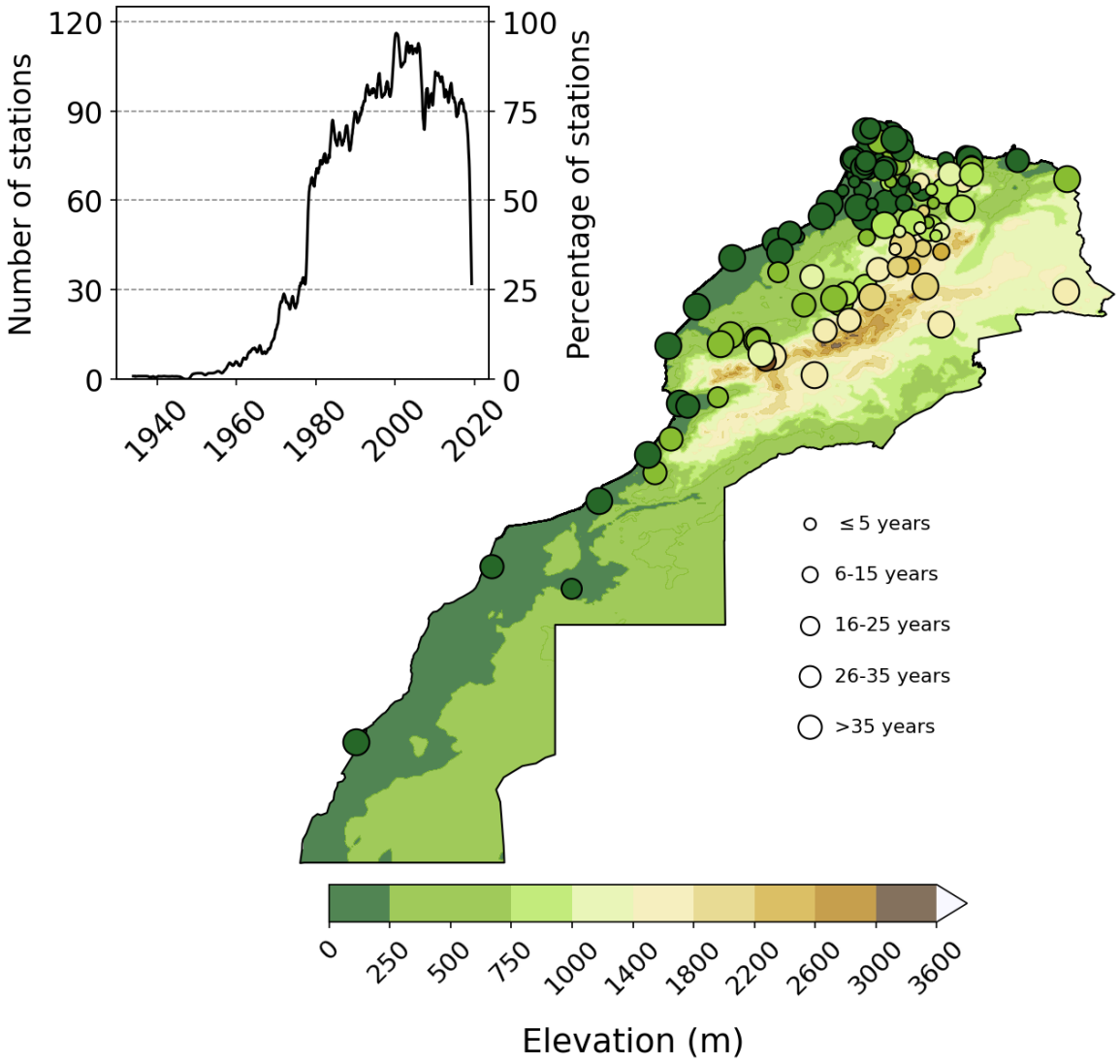

Figure 1: Map of station locations across Morocco (circles) with elevation in shaded contours (data from the USGS GTOPO30 product resampled at 0.05° resolution). The circle diameter indicates the number of years with more than 95% of data both non-missing and passing the quality control checks. The circle color indicates station elevation (taken from station metadata records). The top-left inset shows the time evolution in the number of stations with more than 95% of data both non-missing and passing the quality control checks (evaluated on an annual basis).

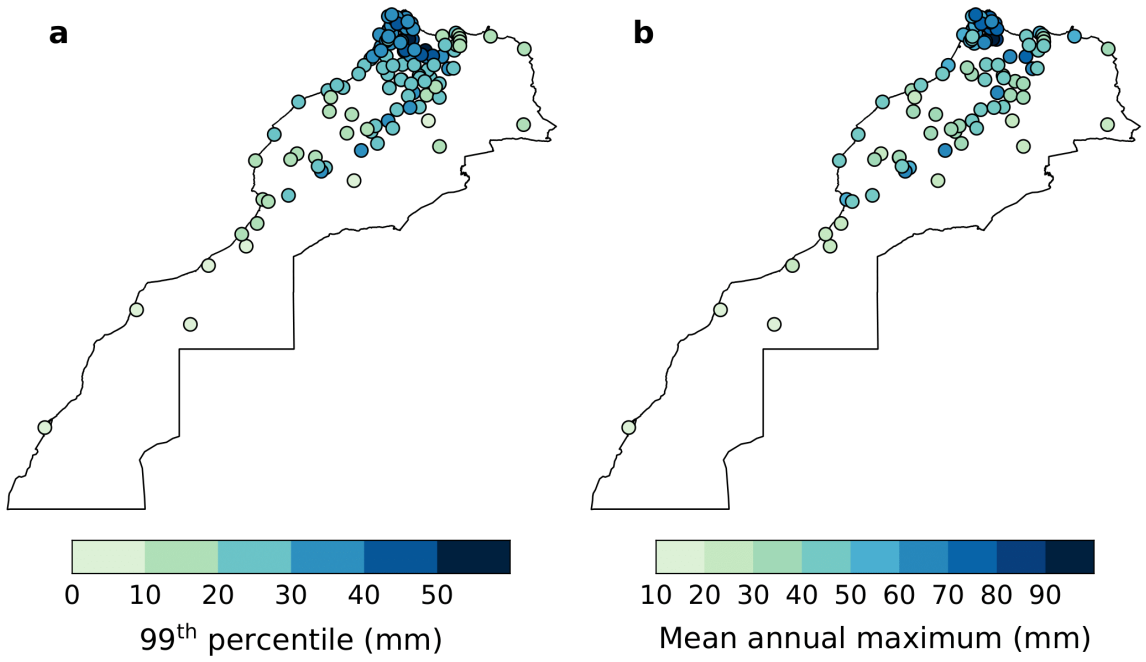

Figure 2: (a) Distribution of (a) 99$^{\text{th}}$ all-day percentiles and (b) mean annual maxima of daily precipitation accumulations at all stations.

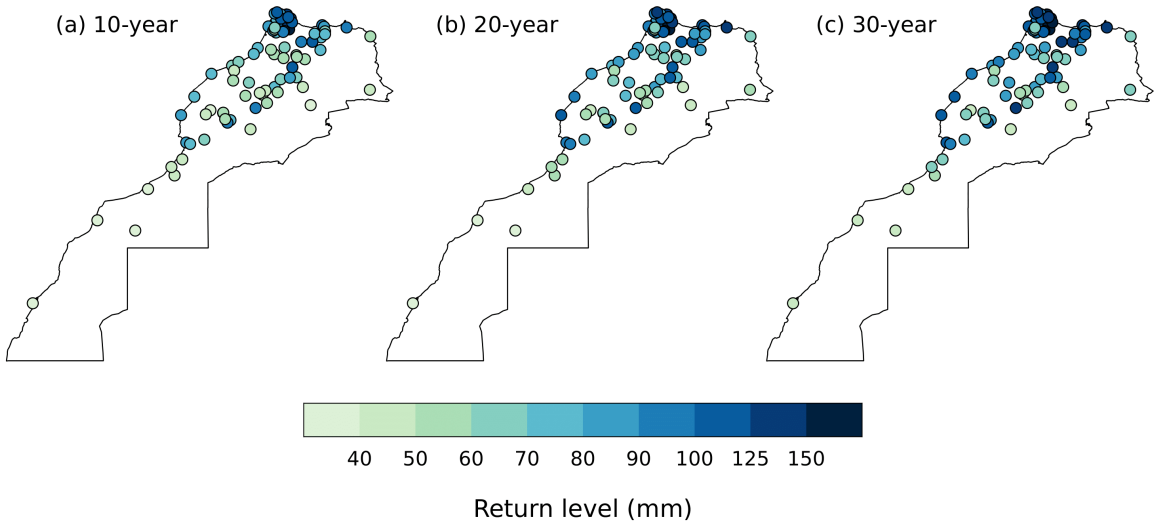

Figure 3: Estimated (a) 10-year, (b) 20-year and (c) 30-year return levels of daily precipitation accumulations at all stations (with at least 25 years with more than 95% of quality-controlled data) inferred from the Generalised Pareto (GP) fits.

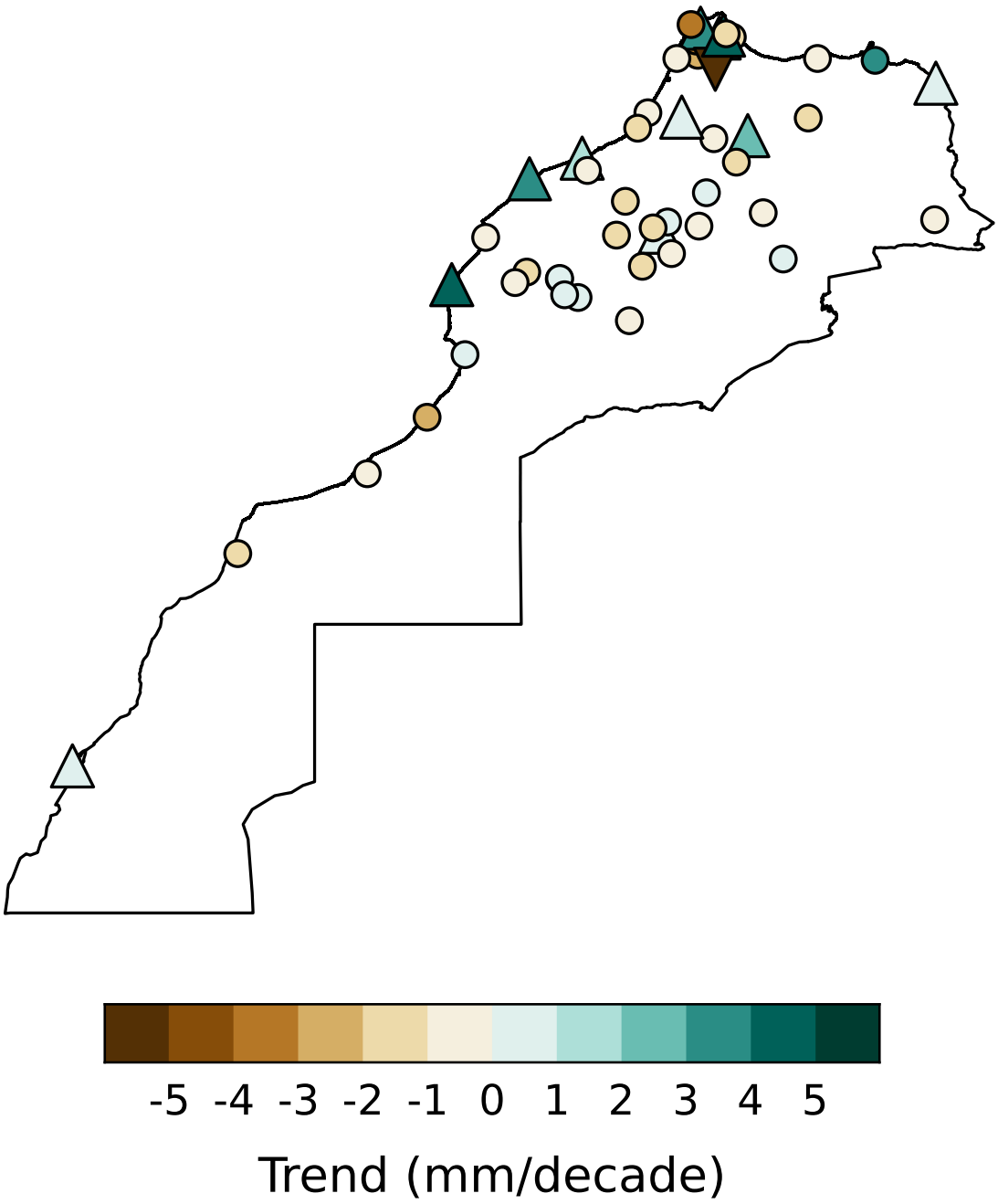

Figure 4: Trends in annual daily precipitation maxima at all stations (with at least 25 years with more than 95% of quality-controlled data) inferred from the Generalised Extreme Value (GEV) fit with the location parameter being a linear function of time (see equation 3). Significant trends are highlighted with triangles (upward: positive trends; downward: negative trends.

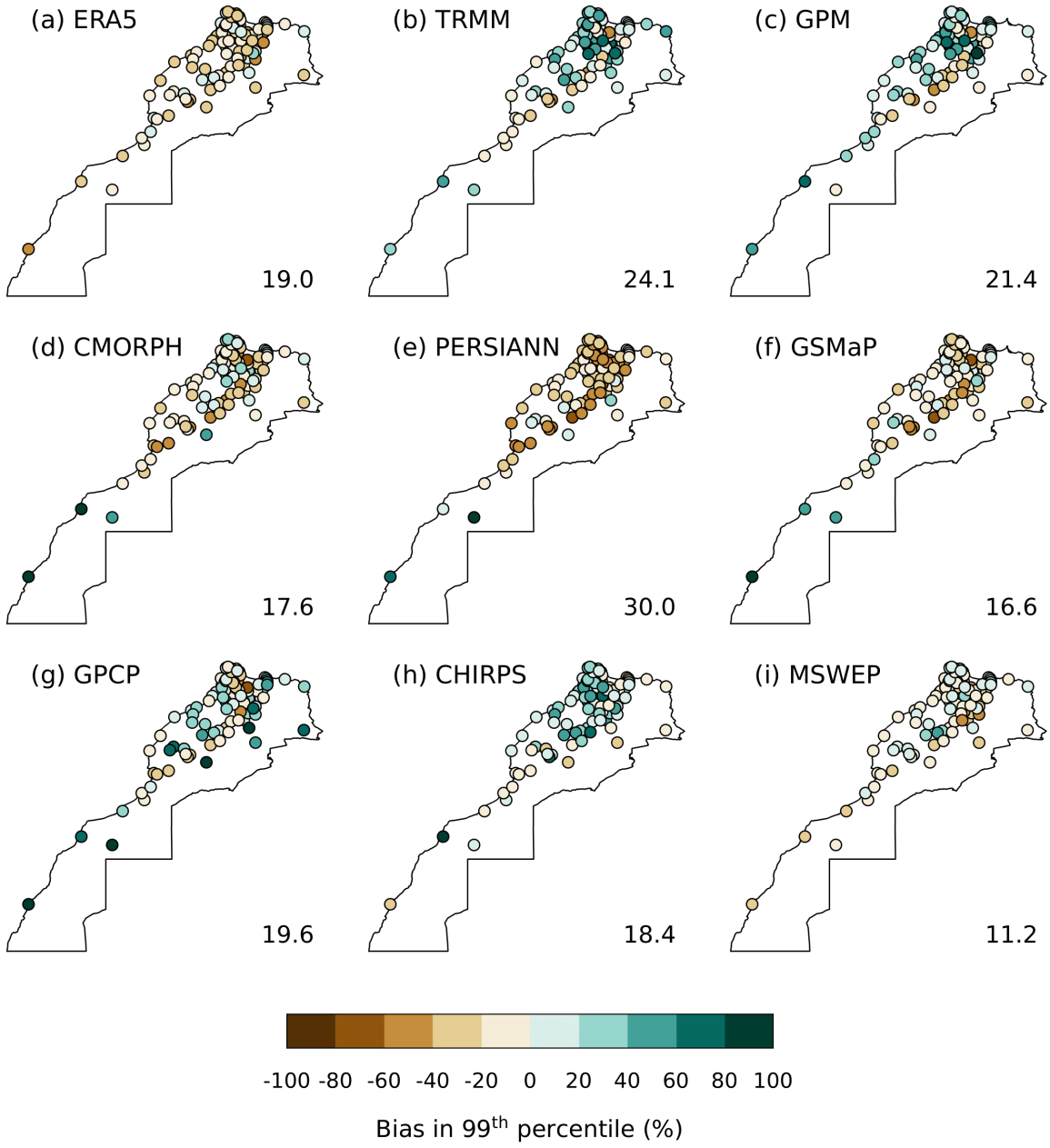

Figure 5: Relative bias in 99th percentiles of daily precipitation accumulations (equation 4) between gridded and station data. Averages across stations are indicated at the bottom right of each panel. The number of plotted locations varies with the gridded dataset as we require at least 10 years of common data with a station to calculate the bias.

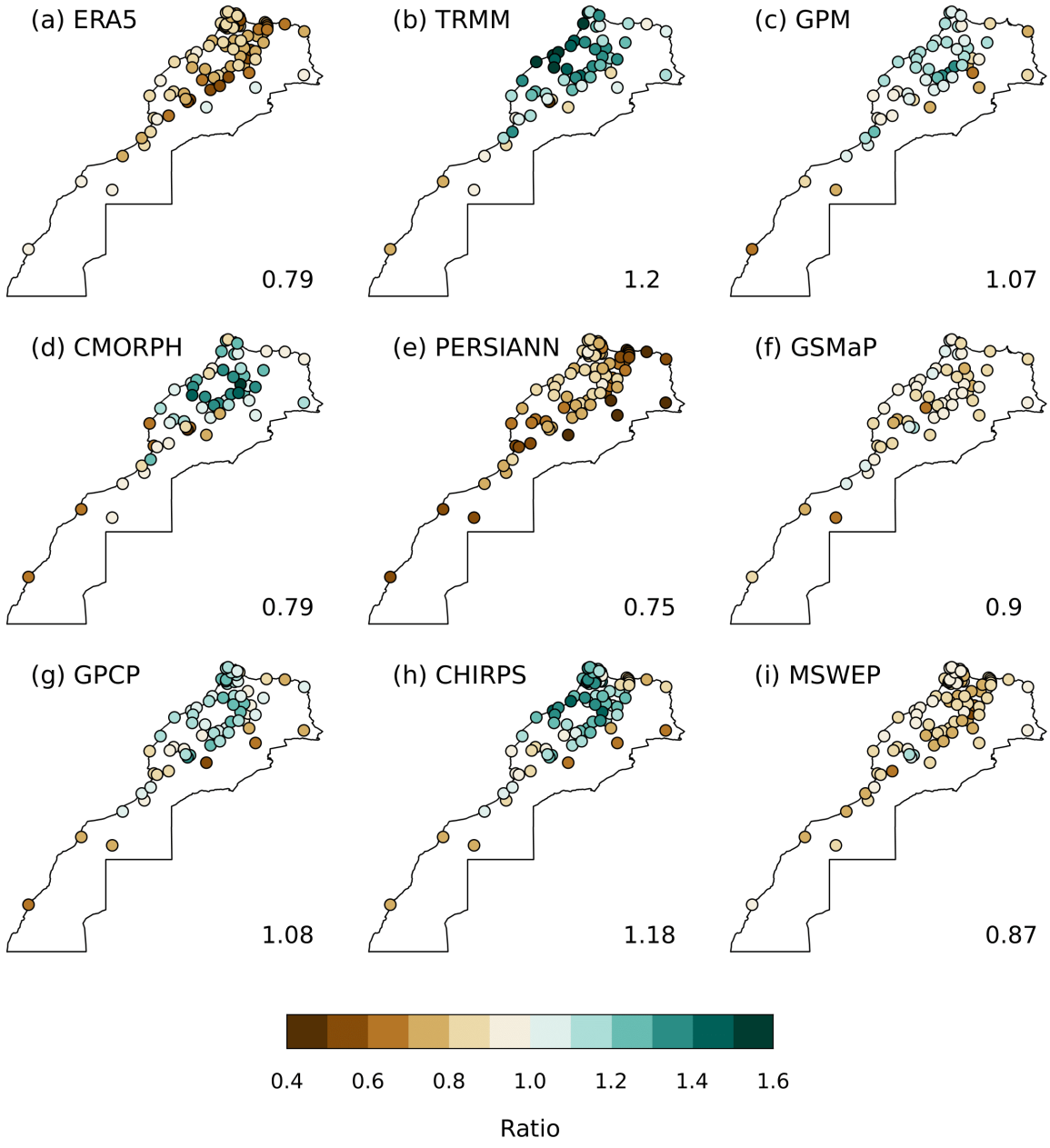

Figure 6: Same as Figure 5, but for the ratio of extreme to annual precipitation (equation 6).

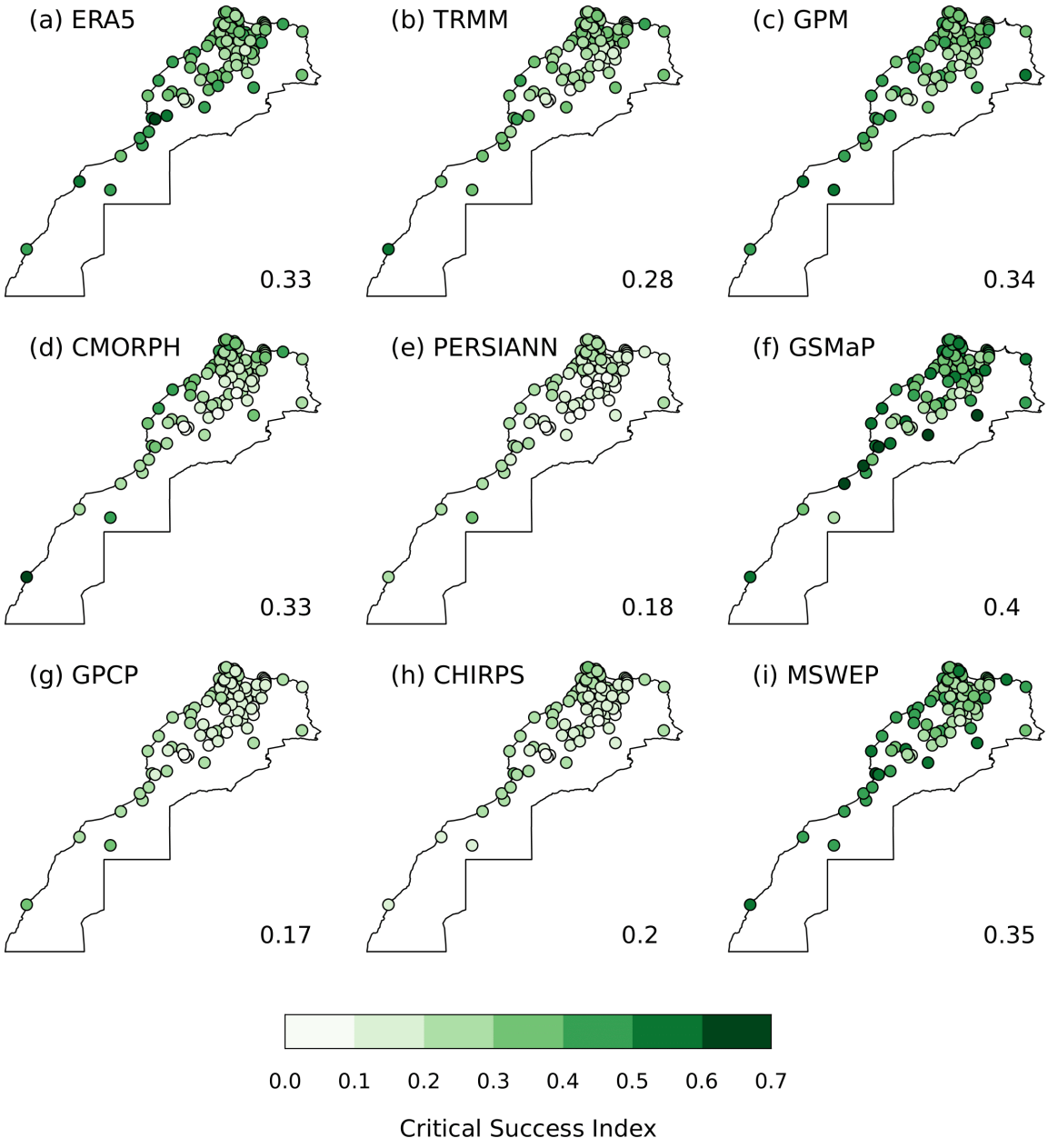

Figure 7: Same as Figure 5, but for the Critical Success Index (equation 5).