A comparison of moderate and extreme ERA-5 daily precipitation with two observational data sets

Pauline Rivoire[1], Olivia Martius[1], and Philippe Naveau[2]

[1]University of Bern
[2]IPSL-CNRS

November 22, 2022

Abstract

A comparison of moderate to extreme daily precipitation from the ERA-5 reanalysis by the European Centre for Medium-Range Weather Forecasts (ECMWF) against two observational gridded data sets, EOBS and CMORPH, is presented. We assess the co-occurrence of precipitation days and compare the full precipitation distributions. The co-occurrence is quantified by the hit rate. An extended generalized Pareto distribution is fitted to the positive precipitation distribution at every grid point and confidence intervals of quantiles compared. The Kullback-Leibler divergence is used to quantify the distance between the entire extended generalized Pareto distributions obtained from ERA-5 and the observations. For days exceeding the local 90th percentile, the mean hit rate is 65% between ERA-5 and EOBS (over Europe) and 60% between ERA-5 and CMORPH (globally). Generally, we find a decrease of the co-occurrence with increasing precipitation intensity. The agreement between ERA-5 and EOBS is weaker over the southern Mediterranean region and Iceland compared to the rest of Europe. Differences between ERA-5 and CMORPH are smallest over the oceans. Differences are largest over North-West America, Central Asia and land areas between 15°S and 15°N. The confidence intervals on quantiles are overlapping between ERA-5 and the observational data sets for more than 80% of the grid points on average. The intensity comparisons indicate an excellent agreement between ERA-5 and EOBS over Germany, Ireland, Sweden and Finland, and a disagreement over areas where EOBS uses sparse input stations. ERA-5 and CMORPH precipitation intensity agree well over the mid-latitudes and disagree over the tropics.

A comparison of moderate and extreme ERA-5 daily precipitation with two observational data sets

Pauline Rivoire[1], Olivia Martius[1], Philippe Naveau[2]

[1]Oeschger Centre for Climate Change Research and Institute of Geography, University of Bern, Bern, Switzerland
[2]Laboratoire des Sciences du Climat et de l'Environnement, CNRS-CEA-UVSQ, Gif-sur-Yvette, France

Key Points:

- The timing and the intensity of daily precipitation are assessed over Europe and globally
- Extended generalized Pareto distributions are fitted to precipitation from ERA-5 and station and satellite data
- Agreement between data sets is highest in the midlatitudes

Corresponding author: Pauline Rivoire, `pauline.rivoire@giub.unibe.ch`

Abstract

A comparison of moderate to extreme daily precipitation from the ERA-5 reanalysis by the European Centre for Medium-Range Weather Forecasts (ECMWF) against two observational gridded data sets, EOBS and CMORPH, is presented. We assess the co-occurrence of precipitation days and compare the full precipitation distributions. The co-occurrence is quantified by the hit rate. An extended generalized Pareto distribution is fitted to the positive precipitation distribution at every grid point and confidence intervals of quantiles compared. The Kullback-Leibler divergence is used to quantify the distance between the entire extended generalized Pareto distributions obtained from ERA-5 and the observations. For days exceeding the local 90^{th} percentile, the mean hit rate is 65% between ERA-5 and EOBS (over Europe) and 60% between ERA-5 and CMORPH (globally). Generally, we find a decrease of the co-occurrence with increasing precipitation intensity. The agreement between ERA-5 and EOBS is weaker over the southern Mediterranean region and Iceland compared to the rest of Europe. Differences between ERA-5 and CMORPH are smallest over the oceans. Differences are largest over North-West America, Central Asia and land areas between 15°S and 15°N. The confidence intervals on quantiles are overlapping between ERA-5 and the observational data sets for more than 80% of the grid points on average. The intensity comparisons indicate an excellent agreement between ERA-5 and EOBS over Germany, Ireland, Sweden and Finland, and a disagreement over areas where EOBS uses sparse input stations. ERA-5 and CMORPH precipitation intensity agree well over the mid-latitudes and disagree over the tropics.

1 Introduction

Natural hazards related to extreme precipitation (river floods, flash floods, landslides, debris flows and avalanches) cause casualties, damages to infrastructures and buildings and have direct and indirect economic impacts (MunichRE, 2018). For infrastructure planning and prevention measures, information about rare events, e.g., events that occur on average only once in a hundred years, is important. Such information can be obtained from precipitation data with statistical tools. Assessing the accuracy in high quantiles depends on spatial domain sizes and temporal availability. Different types of global precipitation data sets are available (Sun et al., 2018): global precipitation data sets are based on ground observations, satellite observations, combinations of ground observations and satellite observations and on short term weather model forecasts in reanalyses data sets. Reanalyses combine past observations with weather forecast models to reconstruct past weather. The main advantage of this type of precipitation data set is its regular spatial and temporal coverage. Reanalyses ensure consistency of the precipitation data with the atmospheric conditions, which is important for weather and climate process studies. Here, we focus on ERA-5 precipitation (C3S, 2017). ERA-5 is the latest reanalysis product from the European Centre for Medium-Range Weather Forecasts (ECMWF). ERA-5 precipitation is computed in short-term forecast started from reanalysis initial conditions (Hennermann, 2020). The ERA-5 precipitation production process does not include precipitation observation inputs. Hence comparison with observational data makes sense, keeping in mind that observation data have (partly substantial) uncertainties as well (Sun et al., 2018; Kulie et al., 2010; Prein & Gobiet, 2017). ERA-5 precipitation has already been widely used since its release in 2018, but very few assessments of this data set have been conducted over large regions. Only precipitation over restricted areas and precipitation associated with specific type of events have been assessed (Wang et al., 2018; Hénin et al., 2018; Mahto & Mishra, 2019; Tarek et al., 2020; Amjad et al., 2020). The goal of this article is to assess daily precipitation in ERA-5 against observational data sets over large regions: Europe, comparing with the station-based data set EOBS (Haylock et al., 2008), and the entire globe, comparing with the satellite-based data set CMORPH (Joyce et al., 2004). The verification of daily precipitation will focus on the intensity distribution and on the temporal consistence, i.e., the co-occurrence of events. We use the extended generalized Pareto distribution (Tencaliec et al., 2019) to evaluate the intensity distribution and we calculate co-occurrence hit rates to assess the joint occurrence of precipitation events.

This paper is structured as follows. Section 2 describes the data used for this study. We introduce methods used for the comparison of co-occurrence and intensity in section 3. The results of our analysis are presented in section 4. Finally, the results are summarized and discussed in section 5.

2 Data

2.1 ERA-5 Precipitation

Reanalysis precipitation in this study are extracted from ERA-5 reanalysis data set. ERA-5 is the latest global reanalysis data set provided by the European Center for Medium-Range Weather Forecasts (C3S, 2017; Hersbach et al., 2020). In this data set, precipitation stem from short-term forecasts and are available at an hourly resolution that we aggregate to daily precipitation. The precipitation data calculation does not rely on observed precipitation. The data is interpolated to a regular grid with 0.25° resolution.

2.2 Observation-based Data Sets

The two gridded observation-based precipitation data sets used in this study are EOBS (Haylock et al., 2008) that is based on European station observations and CMORPH that is based on satellite observations (Joyce et al., 2004).

The EOBS data set is provided by the European Climate Assessment & Dataset and is a daily gridded data set based on spatially interpolated station data. The version used is 19.0e, with a 0.25° by 0.25° grid. The interpolation to a 0.25° by 0.25° grid is a combination of monthly precipitation totals and daily anomalies products. Figure 1a in Cornes et al. (2018) displays the station coverage for version 16.0. This coverage is heterogeneous, with a very dense network in Ireland, the Netherlands, Germany, Switzerland, and northern Italy, for example, and very few stations in northern Africa, in the Middle East, in Iceland, in Norway, and in Sweden. EOBS covers land precipitation only, and the comparison with ERA-5 is conducted for the time period between January 1979 and December 2018.

The second observational data, CMORPH, is provided by the National Center for Atmospheric Research (NCAR) (Climate Prediction Center, National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2011). This gridded precipitation product combines passive microwave satellite scans and geostationary satellite infrared data and provides 3 hour accumulations that we aggregate in daily accumulation. CMORPH stands for climate prediction center morphing method, the name of this combination technique. The precipitation estimation algorithm of this data set is not able to capture snow (Joyce et al., 2004). The spatial resolution of the gridbox is also 0.25°. The comparison with ERA-5 is conducted for the period 2003-2016, for latitudes between 60° S and 60° N.

The two observation-based data sets have the same grid resolution as ERA-5 but a shift of 0.125° in latitude and longitude is present for the coordinates of the grid points compared to ERA-5.

2.3 Data Processing

The study evaluates seasonal precipitation for September, October, November (SON); December, January, February (DJF); March, April, May (MAM); and June, July, August (JJA). Separation between seasons ensures stationarity of the time series. The intensity distribution analyses are based on wet days, defined as days with precipitation accumulations exceeding 1 mm. The 1 mm threshold corresponds to standard recommendations for station data (Hofstra et al., 2009) and eliminates potential drizzle effect in reanalysis data (Maraun, 2013). The co-occurrence analysis is conducted on the entire seasonal time series, including days with precipitation lower than 1 mm.

The precipitation time series are not de-trended in the present study as the response of precipitation to increasing atmospheric CO_2 varies with the precipitation intensity (Pendergrass & Hartmann, 2014) and trends depend on the length of time series (Scherrer et al., 2016). Moreover, Donat et al. (2014) identified mostly small or insignificant trends for the past thirty years.

3 Methods

For the sake of simplicity, in the method section OBSER denotes the observation-based data sets. OBSER can be either EOBS or CMORPH, as the comparison procedures between ERA-5 and EOBS and between ERA-5 and CMORPH are identical.

3.1 Co-occurrence of Precipitation Events

Binary events are defined here as occurrences of daily precipitation above the P^{th} seasonal percentile with $P \in \{75, 90, 95, 99\}$. Percentile values can be different between ERA-5 and OBSER. In Figure 1, the 95^{th} precipitation percentiles in SON is displayed for: (a) ERA-5 over the EOBS domain (1979-2018), (b) EOBS over its entire domain (1979-2018), (c) ERA-5 over the CMORPH domain (2003-2016) and (d) CMORPH over the entire domain (2003-2016).

We define a co-occurrence between two data sets when two exceedances occur either at the same grid point on the same day, at the same grid point with one day lag, or at one of the eight surrounding grid points on the same day. During the spatial and temporal shift, a single event is never used more than once when looking for co-occurrences. We allow for one day shift to bypass uncertainties that arise having a fixed 24h time window (Haylock et al., 2008). The extension to the eight grid points around the centre point addresses potential issues arising from the precipitation interpolation to the different grids.

The hit rate is the ratio between the number of joint events and the total number of events (Rhodes et al., 2015). The total number of events is the same if computed from ERA-5 or from OBSER.

3.2 Intensity Assessment

Extreme value theory is often used in hydrology and climate sciences (e.g., Lamb & Kay, 2004; Cooley et al., 2007; Tramblay et al., 2013; Kang & Song, 2017). This approach states that peaks over high thresholds, i.e., amounts of rain exceeding a given threshold u, may be approximated by a generalized Pareto distribution, provided the threshold and the number of observations are large enough and some additional mild conditions are satisfied (see section 3.2 for generalized Pareto distribution definition). However, the generalized Pareto distribution fitting has drawbacks. First, it only captures the upper tail behavior. A distribution combining gamma behavior for small and moderate precipitation amounts with generalized Pareto distribution behavior for high amounts can be a solution. Second, a threshold has to be determined for every station or grid point to separate the upper tail from the rest of the distribution (Dupuis, 1999). To overcome these challenges, here we use the extended generalized Pareto distribution (EGPD) (Naveau et al., 2016; Tencaliec et al., 2019).

To study wet day precipitation intensity distributions, this section presents a comparison of quantiles and a homogeneity test based on the Kullback-Leibler divergence. Both parts rely on our EGPD fit.

For the intensity comparison, we discard grid points where the number of wet days is smaller than 500 days for the comparison with EOBS and smaller than 200 days for the comparison with CMORPH. Moreover, auto-correlation can be present in daily time series, for example when two consecutive wet days are fostered by the same weather system (e.g., Lenggenhager et al., 2019a). To address the auto-correlation in time, we consider that two precipitation events separated by two days are independent (Barton et al., 2016; Lenggenhager & Martius, 2019b; Fukutome et al., 2015). To ensure independence of the time series, the intensity assessment is conducted on one-third of the data that is randomly drawn. This approach is a trade-off between keeping enough data in the sub-samples to ensure robust fitting and best removing the auto-correlation in the data.

3.2.1 Extended General Pareto Distribution

In extreme value theory, one way to model the extremal tail behavior is the so-called peak-over-threshold approach (see, e.g. Coles, 2001; Katz et al., 2002). Under this framework, rainfall exceedances above a large threshold u are assumed to follow a Generalized Pareto distribution defined as

$$H_\xi(z) = \begin{cases} 1 - (1 + \xi z)_+^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - e^{-z} & \text{otherwise,} \end{cases} \tag{1}$$

where the positive scalar σ represents a scale parameter and the real ξ drives the upper tail behavior. A negative, null and positive ξ corresponds respectively to the "bounded", "light" and "heavy" tail case, i.e. an upper tail that is bounded for $\xi < 0$, with exponential decay for $\xi = 0$ or polynomial decay for $\xi > 0$. The selection of the threshold u is not trivial for large data sets, as each grid point may need a different optimal threshold (e.g., Deidda, 2010). A large threshold implies a small sample size of extremes and consequently, large uncertainties in the estimation of σ and ξ. Conversely, a moderate threshold leads to a possible incorrect approximation by a Generalized Pareto distribution, i.e. a large model error. To bypass this complex threshold selection step, Naveau et al. (2016) proposed a simple scheme to smoothly transition between the main body of the distribution and its upper tail, while keeping the constraint of modeling extremes with a Generalized Pareto distribution. The proposed model can be written as

$$F(x) \quad = \quad G\{H_\xi(x/\sigma)\}, \qquad \text{for all } x > 0, \tag{2}$$

where G, the transition function, is a continuous cumulative distribution function (cdf) in the unit interval. By imposing the two constraints, $\lim_{u \downarrow 0} \frac{1-G(1-u)}{u}$ is finite and positive and $\lim_{u \downarrow 0} \frac{G(u)}{u^s}$ is finite and positive for some $s > 0$, the new cdf $F(.)$ is bound to be in compliance with extreme value theory for its lower and upper tails. This class of cdf is called extended generalized Pareto distribution (EGPD) family. In this study, the cdf $G(.)$ is estimated using a specific Bernstein polynomial approximation, more information can be found in Tencaliec et al. (2019). The R code is available upon request.

3.2.2 Quantile Confidence Intervals

Confidence intervals for the quantiles of ERA-5 and OBSER precipitation are computed using a semi-parametric bootstrap on EGPD fitting. For each grid point, the following bootstrap procedure is conducted. Two subsamples containing one-third of the time series each are randomly drawn from the initial wet day time series. Each of these two subsamples is bootstrapped 100 times each, and each bootstrapped sample is fitted by a EGPD. Having our disposal 200 bootstrapped estimates of $G(.)$, σ and ξ, quantiles for any given non-exceedance probability can be computed from Eq. (2). In particular, 95% confidence intervals of the quantiles are obtained by calculating the empirical 2.5% and 97.5% quantiles of the 200 bootstrapped quantiles values. An important feature to assess the proximity of our different data sets is to check if the confidence intervals from ERA-5 overlap (or not) with the observational data sets.

3.2.3 Kullback-Leibler Divergence Test

The well-known Kullback-Leibler divergence used in various fields "measures" the distance between two probability density functions, say f_1 and f_2. It is given by equation

$$\mathbb{E}_{f_1}\left[\log\left\{\frac{f_1(\mathbf{X})}{f_2(\mathbf{X})}\right\}\right]+\mathbb{E}_{f_2}\left[\log\left\{\frac{f_2(\mathbf{Y})}{f_1(\mathbf{Y})}\right\}\right] \tag{3}$$

with $\mathbf{X}$ and $\mathbf{Y}$ being random variables following respectively the probability density functions f_1 and f_2.

Let X_{ERA-5}=$(X_i)_{i=1,...,n}$ and Y_{OBSER}=$(Y_j)_{j=1,...,m}$ be the time series (after removing the auto-correlation) of wet day precipitation in ERA-5 and OBSER. With $\hat{f}_1$ and $\hat{f}_2$ estimated by the EGPD fitting, we obtain the empirical value of the Kullback-Leibler divergence with equation

$$\frac{1}{n}\sum_{i=1}^{n}\log\frac{\hat{f}_1(X_i)}{\hat{f}_2(X_i)}+\frac{1}{m}\sum_{j=1}^{m}\log\frac{\hat{f}_2(Y_j)}{\hat{f}_1(Y_j)}. \tag{4}$$

The null hypothesis of our test is "X_{ERA-5} and Y_{OBSER} have the same distribution", i.e. $\hat{f}_1=\hat{f}_2$. The alternative hypothesis is $\hat{f}_1 \neq \hat{f}_2$.

The distribution of the Kullback-Leibler divergence under the null hypothesis is estimated using 300 values of the divergence between two vectors randomly drawn from a concatenation of X_{ERA-5} and Y_{OBSER}. The probability of "The Kullback-Leibler divergence between X_{ERA-5} and Y_{OBSER} is greater than the Kullback-Leibler divergence under the null hypothesis" is the p-value of the test. This p-value is empirically determined from the 300 values of the Kullback-Leibler divergence under the null hypothesis. The null hypothesis is rejected with a confidence level of 5% if the p-value is greater than 0.05.

3.3 Difference in Number of Wet Days

The intensity comparison is based on the EGPD, which is fitted to wet days only. Discrepancies in the number of wet days between ERA-5 and the observational data sets could have an impact when comparing the EGPD fitted to ERA-5 and the observational data sets. To quantify these discrepancies at a fixed grid point, we use two simple measures. The first measure is the ratio of the seasonal number of wet days, defined by :

$$\frac{N^{wd}_{ERA-5}}{N^{wd}_{OBSER}} \tag{5}$$

where, for a fixed season, N^{wd}_{ERA-5} and N^{wd}_{OBSER} are the number of wet days in ERA-5 and OBSER, respectively.

The second measure is the absolute value of the difference in the number of wet days between ERA-5 and OBSER, normalized by the time series length of OBSER, given by:

$$100 \times \left| \frac{N^{wd}_{ERA-5} - N^{wd}_{OBSER}}{N^{wd}_{OBSER}} \right|. \tag{6}$$

Note that the absolute difference quantifies the distance between the ratio of the number of wet days and 1.

4 Results

4.1 Number of Wet Days

Table 1 presents the mean absolute value of the difference in the number of wet days. The differences are computed only over grid points retained for the intensity comparison, i.e. with time series longer than 200 days for CMORPH and 500 days for EOBS.

Over Europe, the mean absolute difference is between 11% (SON) and 21% (MAM) of EOBS number of wet days. In SON and DJF, the number of wet days is lower in ERA-5 than in EOBS in northern Europe and higher in southern Europe. In MAM and JJA the number of wet days is almost always higher in ERA-5 than in EOBS (see Figure A1a in the appendix for a map of the ratio of the number of wet days in SON). The difference in number of wet days between ERA-5 and EOBS can be considered as low and will not have impact on the EGPD fitting.

The global comparison with CMORPH reveals larger discrepancies in the seasonal number of wet days than with EOBS. The mean difference in the number of wet days corresponds to between 66% (DJF) and 76% (SON) of the CMORPH number of wet days. This difference is mainly due to the ERA-5 wet days being more numerous than in CMORPH. The number of wet days in ERA-5 is twice as large as in CMORPH for 19% (DJF) to 25% (JJA) of grid points. Figure A1b in appendix displays the map of the ratio of number of wet days in SON. This ratio is greater than 2 over bands at fixed latitudes e.g. over bands close 60° S, 20° S, 20° N and 60° N.

4.2 Co-occurrence of Precipitation Events

In the comparison between ERA-5 and both observation-based data sets, the hit rate decreases with increasing intensity of the events and is similar across the seasons (Table 2). Grid points with a given percentile of less than 1 mm are not considered.

Over Europe the average hit rate between ERA-5 and EOBS for the 75^{th} percentile is between 73 % (in JJA) and 77 % (in SON), i.e. about three quarters of the events exceeding the 75^{th} percentiles coincide. For the 95^{th} percentile, the mean hit rate is between 53% (JJA) and 60% (SON). For the 99^{th} percentile, the hit rate varies between 39 % and 45 % depending on the season. The global mean hit rate is of the same order of magnitude as for Europe. The mean hit rate between ERA-5 and CMORPH for the 95^{th} percentiles is above 50%. The mean hit rate associated with the 99^{th} percentile is between 35% and 37% depending on the season.

Maps of the hit rate for the 95^{th} percentile can be found in the appendix (Figure B1 and Figure B2). The spatial pattern does not strongly depend on the season or the percentile. For Europe, the hit rate has a large variability near arid regions (Maghreb and Turkey). The rest of Europe is quite homogeneous. A lower hit rate is observed in Iceland and southern Italy. For the global comparison, the best hit rate is reached over the oceans in the mid-latitudes. The hit rate is substantially lower in Eastern China, along the equator, in South America and in tropical Africa.

4.3 Intensity Verification

4.3.1 Confidence Intervals on Quantiles

The confidence interval overlap between ERA-5 and EOBS is independent of the probability of non-exceedance, i.e., the intensity of the events. Figure 2a shows the relative position of the 95% confidence intervals for quantiles associated with probability of non-exceedance 0.9 in SON, between ERA-5 and EOBS. A grid point is displayed in yellow if the confidence intervals are overlapping, in orange if the upper boundary of ERA-5 confidence interval is lower than the lower boundary of the EOBS confidence interval, and in blue if the lower boundary of ERA-5 confidence interval is larger than the upper boundary of the EOBS confidence interval. Figure 3a shows the number of seasons with a confidence intervals overlap for quantiles with non-exceedance probability 0.9 between ERA-5 and EOBS. The confidence intervals overlap during all the seasons for a major part of Europe. The exceptions are Iceland, Norway and Western Russia, Romania, the Adriatic sea coast and some grid points in the Alps. Non-overlapping confidence intervals correspond primarily to an underestimation of the quantiles by ERA-5 for low precipitation intensities (not shown), and overestimation for large intensities (Figure 2a.). Quantiles with probability of 0.3 have a larger number of grid points with disagreement in JJA. ERA-5 quantiles for probability 0.3 during JJA are underestimated compared to EOBS quantiles for a major part of Europe (not shown).

The global comparison of ERA-5 with CMORPH shows less overlap of the confidence intervals with increasing precipitation intensity, as we will see in section 5. Figure 2b displays the relative position of the 95% confidence intervals for quantiles associated with probability of non-exceedance 0.9 in SON, between ERA-5 and CMORPH and Figure 3b shows the number of seasons with an overlap of the confidence intervals for quantiles with non-exceedance probability 0.9 between ERA-5 and CMORPH. The confidence intervals for non-exceedance probabilities of 0.3 and 0.5 overlap for more than 85% of the grid points. For probabilities 0.75 to 0.95, in a band along the equator the confidence intervals do not overlap for all the seasons for many grid points, see e.g. Figure 2b. For all seasons and between 15° S and 15° N, ERA-5 quantiles are smaller compared to CMORPH. Another disagreement that deserves to be highlighted is the higher ERA-5 quantiles compared to CMORPH over the mountainous regions of the west coast of North America, the south of Chile, in Papua New Guinea, the Himalayas and the Alps, regardless of the non-exceedance probabilities and seasons.

4.3.2 Comparison of the Full Distributions Using the Kullback-Leibler Test

The Kullback-Leibler divergence of the full precipitation distribution points to regions of agreement and disagreement independent of the seasons, for both the comparison between ERA-5 and EOBS, and ERA-5 and CMORPH, see Figure 4. Figures 5a and 5b display the number of seasons for which the p-value of the Kullback-Leibler test is greater than 0.05, i.e. where the EGPD distributions fitted to ERA5 and to EOBS and CMORPH do not differ significantly.

ERA-5 and EOBS wet day precipitation intensities agree best over Germany, Ireland, Sweden and Finland. Wet day precipitation intensity follows the same distribution in ERA-5 and EOBS for most grid points in these countries. Regions with the least agreement, i.e. where the null hypothesis is rejected for all the seasons, are Iceland, Norway, Hungary and the Balkans. The area with at least one season where the null hypothesis is rejected is rather large. The Kullback-Leibler test gives weight to differences in the entire distribution, thus the low intensity precipitation disagreement in JJA mentioned previously in section 4.3.1 has an impact on the Kullback-Leibler divergence (see Figure 4a and Figure 4c for the plots of the p-value in JJA and SON). Note the pattern following the border between Norway and Sweden, and Finland and Russia: a very good agreement is observed in Sweden and Finland, whereas the null hypothesis is rejected for almost all seasons in Norway and Karelia.

The Kullback-Leibler test between ERA-5 and CMORPH has a clear signal of agreement in the mid-latitudes and disagreement in the tropics, for all seasons. The summary over all seasons mainly informs about intensity agreement over the oceans, because of the time series length constrain removes most land grid points. Figures 4b and 4d present the p-value for JJA and SON, and show over land the same general pattern of disagreement in the tropics and agreement in the mid-latitudes. One exception to this pattern is the disagreement over mountainous regions of the mid-latitudes (western North America, Himalayas, South Chile), in agreement with the results in section 4.3.1.

5 Summary and Discussion

The analysis of precipitation event co-occurrence between ERA-5 and EOBS and ERA-5 and CMORPH reveals a decreasing agreement with increasing intensity of events, independently of the season.

Key results of the intensity comparison of ERA-5 with EOBS over Europe depend on the season (Table 3). Quantiles in MAM and SON show a good agreement for all non-exceedance probabilities p. Indeed, between 81% (for $p = 0.3$ in MAM and for $p = 0.5$ in DJF) and 90% (for $p = 0.5$ in MAM and SON, and for $p = 0.75$ in SON) of the grid points have overlapping confidence intervals. The percentages of grid points where the distributions agree (Kullback-Leibler test) are highest for MAM (34%) and SON (39%). In DJF, the agreement between quantiles is between 69% ($p = 0.9$) and 82% ($p = 0.3$). In JJA agreement is much better for high quantiles (up to 94% for $p = 0.95$) than low ones, the confidence intervals for quantiles with probability $p = 0.3$ overlap for only 39% of grid points. This discrepancy for low precipitation intensity has an impact on the Kullback-Leibler test: the null-hypothesis could not be rejected for only 10% of grid points in JJA. The other seasons show a higher fraction of grid points for which the null-hypothesis was not rejected (between 29% in DJF and 39% in SON).

The study of the wet day precipitation distribution (Kullback-Leibler test) between ERA-5 and EOBS over Europe reveals a robust agreement over regions where the station coverage of EOBS is dense. Areas with the largest differences in the distribution are areas with thin station coverage, e.g. in southern Europe and Russia. Cornes et al. (2018) highlighted that “station coverage is the most important factor in determining the success of the gridded data”. In areas with sparse station data, the precipitation is interpolated from distant stations (Hofstra et al., 2009). Additionally, extreme precipitation is smoothed by the spatial interpolation (Hofstra et al., 2010), which justifies our quantiles larger in ERA-5 than in EOBS for extreme precipitation.

The intensity comparison between ERA5 and CMORPH indicates a decreasing agreement between the two data sets with increasing precipitation intensity (Table 3). The percentage of grid points with confidence intervals overlapping is between 92% (JJA) and 94% for $p = 0.3$, and between 70% (SON) and 75% (DJF) for $p = 0.9$. One exception is the slightly better agreement of quantiles for extreme precipitation (non-exceedance probability $p = 0.95$) than for moderately extreme precipitation ($p = 0.9$), with an overlap rate 1% higher, for all seasons. This can be explained by confidence intervals becoming larger for larger quantiles, and there is thus a higher chance of overlap between ERA-5 and the observational data set. This remark holds also for the EOBS results. The Kullback-Leibler test presents little variation with season, like for the quantile study. Between 52% (SON) and 57% (DJF) of the grid points studied did not reject the null-hypothesis of the test.

The analysis of the entire precipitation distribution reveals proportionally more grid points agreeing between ERA-5 and CMORPH than between ERA-5 and EOBS (see last row of Table 3). This can be due to the longer time series in EOBS leading to a stricter test. Another explanation can be that there are proportionally more challenging regions for a model over Europe, with the Alps for example, whereas globally the largest regions compared are oceans, where the agreement is good in general.

The global comparison of the wet day precipitation distributions between ERA-5 and CMORPH over the period 2003-2016 shows a rather good agreement in the mid-latitudes, and a strong disagreement over the tropics. This result is robust over the seasons and does not depend on the method used. The band along the equator where precipitation intensities are lower in ERA-5 compared to CMORPH corresponds to a region with a ratio of the number of wet days rather close to 1 (see Figure A1). The number of wet days does not differ substantially in this region and therefore does not play a role in the robust disagreement between ERA-5 and CMORPH over the tropics. In this region, ERA-5 quantiles are lower than CMORPH ones, especially for non-exceedance probabilities larger than 0.75. This feature has already been observed by Pfahl and Wernli (2012) with ERA-interim, another reanalysis product from ECMWF. They computed the empirical 99^{th} percentiles of 6-hourly precipitation in ERA-interim and CMORPH for the period 2003-2016 (Figure 2 in their article). The authors showed a strong underestimation of ERA-interim precipitation over the tropics compared to CMORPH. They concluded that the deep convection, a central process in tropical extreme precipitation, was not properly captured by the reanalysis data set. Even though Nogueira (2020) found an improvement of the precipitation simulation over the tropics in ERA-5 compared to ERA-interim the previous reanalysis data set from ECMWF, we assume ERA-5 to still contain an underestimation of the tropical extreme rainfall.

One limit of CMORPH that it is important to emphasize is the limitations of this data set to capture snow (Joyce et al., 2004) and low-intensity precipitation during winter in the mid-latitudes (Sun et al., 2018). This property leads to smaller number of wet days in winter (DJF or JJA depending on the hemisphere) as seen in section 4.1, hence the large areas where our intensity analysis can not be conducted. Some regions at high latitudes (close to 60° S and 60° N) and some mountainous regions have a number of wet days large enough for the intensity comparison to be conducted even if snow can be expected. However a substantial difference in the number of wet days is observed. The higher ERA-5 precipitation compared to CMORPH over mountainous regions might be related to snow. Timmermans et al. (2019) revealed a disagreement between CMORPH and gauge-based product for extreme precipitation in mountainous regions of western USA in DJF and interpreted it as a consequence of the post-processing performed in CMORPH leading to lots of missing data in winter (Xie et al., 2017).

Hénin et al. (2018) assessed ERA-5 daily accumulated precipitation during extreme precipitation events over the Iberian Peninsula for the period 2000-2008 against precipitation from a ground based gridded data set. They found an overestimation of daily sums for moderate extreme events and and underestimation for the most extreme events. Our study of quantiles in ERA-5 and EOBS for non-exceedance probabilities greater than 0.75 reveals a moderate signal of overestimation of ERA-5 precipitation in the same region. One exception is the southern Basque Country where an underestimation of ERA-5 quantiles for these probabilities in DJF. Our comparison of moderate and large extremes over the period 1979-2018 is therefore only partially in agreement with their study over the period 2000-2008.

For the period 2014-2018, Amjad et al. (2020) showed that ERA-5 overestimates the precipitation observed over Turkey compared to ground based stations, independently of the wetness and slope classes. Our comparison with EOBS indicates the same signal for high quantiles, but also a dry bias for lower quantiles. This can be due to the fact that our study period is longer or that EOBS has a poorer station coverage in this region.

In their study, Tarek et al. (2020) compared the mean seasonal precipitation in ERA-5 with station observations over North America between 1979 and 2018. In JJA, they found an underestimation of precipitation in ERA-5 over Florida, and an overestimation along the west coast of Canada, which is in agreement with our comparison of quantiles between ERA-5 and CMORPH for this season and for all probabilities of non-exceedance. In DJF, they showed an underestimation precipitation in ERA-5 over the west coast of USA and Florida, and an overestimation over the west coast of Canada. Our analysis highlighted that ERA-5 presents larger quantiles than CMORPH over the west coast of North America for all non-exceedance probabilities. Our results are thus in agreement for Canada but not for USA. This can be due to the fact that the time periods studied are different and that CMORPH underestimates precipitation during the cold months (Sun et al., 2018).

Mahto and Mishra (2019) assessed ERA-5 precipitation in India against observation comparing precipitation sums during the monsoon season (June-September) between 1980 and 2018. They found a wet bias over Indo-Gangetic Plain and foothills of Himalaya and a dry bias in semi arid regions of western India. These results are in agreement with our quantile analysis in ERA-5 and CMORPH in JJA for the period 2003-2016.

6 Conclusion

We compare daily precipitation from the ERA-5 reanalysis data set with daily precipitation from two observation-based data sets, EOBS and CMORPH. The comparison addresses three aspects i) the temporal co-occurrence of moderate to high extreme events in two data sets, ii) the agreement of return values for moderate to extreme non-exceedance probabilities derived from the extended generalized Pareto distribution (EGPD), and iii) a comparison of the full precipitation distribution captured by the EGPD using the Kullback-Leibler divergence. We quantify the co-occurrence of precipitation events with the hit rate. We compare the EGPD distributions between ERA-5 and the observational data sets with confidence intervals for several non-exceedance probabilities and with a test based on the Kullback-Leibler divergence.

Between ERA-5 and EOBS over Europe the hit rate is above 65% for moderate precipitation and approximately 50% for extreme precipitation. Between ERA-5 and CMORPH globally the hit rate is above 60% for moderate precipitation and around 40% for extreme precipitation. Over Europe areas with the least agreement are the southern Mediterranean region and Iceland and for the global comparison areas with the least agreement are land areas between 15°S and 15°N, North-West America and Central Asia.

For a majority of grid points confidence intervals for non-exceedance probabilities of 0.3 to 0.95 overlap between ERA-5 and EOBS. We find a disagreement between ERA-5 and EOBS in areas where EOBS uses fewer input stations. We therefore hypothesize that the reanalysis data set might better capture moderate to extreme precipitation in regions where the station coverage is sparse. The analysis also showed that ERA-5 underestimates extreme precipitation compared to CMORPH in the tropics. In general, the magnitudes of the non-exceedance probabilities agree between ERA-5 and the observation-based data sets in the mid-latitudes.

The Kullback-Leibler test on the entire precipitation distributions over Europe shows an agreement of the EGPD distributions in ERA-5 and EOBS over Germany, Ireland, Sweden and Finland. The precipitation distributions differ significantly between in ERA-5 and EOBS in all four seasons in Iceland, Norway, Karelia, Hungary and the Balkan. The Kullback-Leibler test between ERA-5 and CMORPH shows that precipitation distributions are generally in agreement over the mid-latitudes and differ significantly over the tropics for all seasons, confirming the results of the quantile comparison. ERA-5 should only be used with great care to study extreme precipitation over the tropics.

The strengths of ERA-5 daily precipitation data are the regular spatial and temporal resolution and the consistency with the large-scale circulation and there is generally a good agreement with observation-based data sets in the extra-tropics. The reanalysis data set provides valuable complementary information to observational data in regions where observational data sets are sparse, e.g. in areas where the EOBS station coverage is poor or for CMORPH in regions and seasons where snow is prevalent. In the tropics, an observational data set should be preferred over ERA-5.

Acknowledgments

P.R. acknowledges funding from the the Swiss National Science Foundation (grant number 178751). Part of P.N. work was supported by the French national program FRAISE-LEFE/INSU, ANR-Melody and ANR-TRex. The support of DAMOCLES-COST-ACTION on compound events is also acknowledged.

The authors thank the assistance of Andrey Martynov (Institute of Geography, Bern), who prepared the precipitation data sets.

We acknowledge the E-OBS data set from the EU-FP6 project UERRA (http://www.uerra.eu) and the Copernicus Climate Change Service, and the data providers in the ECA&D project (https://www.ecad.eu).

The article processing charges for this open-access publication were covered by the university of Bern.

The authors declare that they have no conflict of interest.

All data sets used in the current study are publicly available from the indicated references or sources. The codes for the intensity assessment are available from GitHub (`https://github.com/PauRiv/characterization_ERA-5_daily_precipitation`). The codes for the co-occurrence verification are available upon request.

References

Amjad, M., Yilmaz, M. T., Yucel, I., & Yilmaz, K. K. (2020). Performance evaluation of satellite- and model-based precipitation products over varying climate and complex topography. *Journal of Hydrology*, *584* (February), 124707. doi: 10.1016/j.jhydrol.2020.124707

Barton, Y., Giannakaki, P., von Waldow, H., Chevalier, C., Pfahl, S., & Martius, O. (2016, jan). Clustering of Regional-Scale Extreme Precipitation Events in Southern Switzerland. *Monthly Weather Review*, *144* (1), 347–369. doi: 10.1175/MWR-D-15-0205.1

C3S, C. C. C. S. (2017). *Era5: Fifth generation of ecmwf atmospheric reanalyses of the global climate.* `https://cds.climate.copernicus.eu/cdsapp#!/home`. (Accessed: 2020-03-02)

Climate Prediction Center, National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce. (2011). *Noaa cpc morphing technique (cmorph) global precipitation analyses.* Boulder CO: Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. Retrieved from `https://doi.org/10.5065/D6CZ356W` (Accessed: 2017-01-02)

Coles, S. (2001). *An introduction to statistical modeling of extreme values.* Springer London. Retrieved from `https://books.google.fr/books?id=SonbBwAAQBAJ`

Cooley, D., Nychka, D., & Naveau, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, *102* (479), 824-840. doi: 10.1198/016214506000000780

Cornes, R. C., van der Schrier, G., van den Besselaar, E. J., & Jones, P. D. (2018). An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets. *Journal of Geophysical Research: Atmospheres*, *123*(17), 9391–9409. doi: 10.1029/2017JD028200

Deidda, R. (2010). A multiple threshold method for fitting the generalized Pareto distribution to rainfall time series. *Hydrology and Earth System Sciences*, *14*(12), 2559–2575. doi: 10.5194/hess-14-2559-2010

Donat, M. G., Sillmann, J., Wild, S., Alexander, L. V., Lippmann, T., & Zwiers, F. W. (2014). Consistency of temperature and precipitation extremes across various global gridded in situ and reanalysis datasets. *Journal of Climate*, *27*(13), 5019–5035. doi: 10.1175/JCLI-D-13-00405.1

Dupuis, D. (1999, Jan 01). Exceedances over high thresholds: A guide to threshold selection. *Extremes*, *1*(3), 251–261. doi: 10.1023/A:1009914915709

Fukutome, S., Liniger, M. A., & Süveges, M. (2015). Automatic threshold and run parameter selection: a climatology for extreme hourly precipitation in Switzerland. *Theoretical and Applied Climatology*, *120*(3-4), 403–416. doi: 10.1007/s00704-014-1180-5

Haylock, M. R., Hofstra, N., Klein Tank, A. M., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research Atmospheres*, *113*(20). doi: 10.1029/2008JD010201

Hénin, R., Liberato, M. L., Ramos, A. M., & Gouveia, C. M. (2018). Assessing the use of satellite-based estimates and high-resolution precipitation datasets for the study of extreme precipitation events over the Iberian Peninsula. *Water (Switzerland)*, *10*(11). doi: 10.3390/w10111688

Hennermann, K. (2020). *Era5: data description.* `https://confluence.ecmwf.int/pages/viewpage.action?pageId=85402030`. (Accessed: 2020-03-06)

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., … Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*(September 2019), 1–51. doi: 10.1002/qj.3803

Hofstra, N., Haylock, M., New, M., & Jones, P. D. (2009). Testing E-OBS European high-resolution gridded data set of daily precipitation and surface temperature. *Journal of Geophysical Research Atmospheres*, *114*(21). doi: 10.1029/2009JD011799

Hofstra, N., New, M., & McSweeney, C. (2010). The influence of interpolation and station network density on the distributions and trends of climate variables in gridded daily data. *Climate Dynamics*, *35*(5), 841–858. doi: 10.1007/s00382-009-0698-1

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *Journal of Hydrometeorology*. doi: 10.1175/1525-7541(2004)005⟨0487:CAMTPG⟩2.0.CO;2

Kang, S., & Song, J. (2017). Parameter and quantile estimation for the generalized Pareto distribution in peaks over threshold framework. *Journal of the Korean Statistical Society*, *46*(4), 487–501. doi: 10.1016/j.jkss.2017.02.003

Katz, R. W., Parlange, M. B., & Naveau, P. (2002). Statistics of Extremes in Climatology and Hydrology. *Advances in Water Resources*, *25*, 1287–1304. doi: 10.1111/j.1467-9574.1989.tb01244.x

Kulie, M. S., Bennartz, R., Greenwald, T. J., Chen, Y., & Weng, F. (2010). Uncertainties in microwave properties of frozen precipitation: Implications for remote sensing and data assimilation. *Journal of the Atmospheric Sciences*, *67*(11), 3471–3487. doi: 10.1175/2010JAS3520.1

Lamb, R., & Kay, A. L. (2004). Confidence intervals for a spatially generalized, continuous simulation flood frequency model for Great Britain. *Water Resources Research*, *40*(7), 1–13. doi: 10.1029/2003WR002428

Lenggenhager, S., Croci-Maspoli, M., Brönnimann, S., & Martius, O. (2019a). On the dynamical coupling between atmospheric blocks and heavy precipitation events: A discussion of the southern Alpine flood in October 2000. *Quarterly Journal of the Royal Meteorological Society*, *145*(719), 530–545. doi: 10.1002/qj.3449

Lenggenhager, S., & Martius, O. (2019b). Atmospheric blocks modulate the odds of heavy precipitation events in Europe. *Climate Dynamics*, *53*(7-8), 4155–4171. doi: 10.1007/s00382-019-04779-0

Mahto, S. S., & Mishra, V. (2019). Does ERA-5 Outperform Other Reanalysis Products for Hydrologic Applications in India? *Journal of Geophysical Research: Atmospheres*, *124*(16), 9423–9441. doi: 10.1029/2019JD031155

Maraun, D. (2013). Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *Journal of Climate*, *26*(6), 2137–2143. doi: 10.1175/JCLI-D-12-00821.1

MunichRE. (2018). A stormy year: Natural Catastrophes 2017. *Topics Geo*, 65. Retrieved from `https://www.munichre.com/topics-online/en/climate-change-and-natural-disasters/natural-disasters/topics-geo-2017.html`

Naveau, P., Huser, R., Ribereau, P., & Hannart, A. (2016). Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. *Water Resources Research*, *52*(4), 2753-2769. doi: 10.1002/2015WR018552

Nogueira, M. (2020). Inter-comparison of ERA-5, ERA-interim and GPCP rainfall over the last 40 years: Process-based analysis of systematic and random differences. *Journal of Hydrology*, *583*(August 2019), 124632. doi: 10.1016/j.jhydrol.2020.124632

Pendergrass, A. G., & Hartmann, D. L. (2014). Changes in the distribution of rain frequency and intensity in response to global warming. *Journal of Climate*, *27*(22), 8372–8383. doi: 10.1175/JCLI-D-14-00183.1

Pfahl, S., & Wernli, H. (2012, oct). Quantifying the Relevance of Cyclones for Precipitation Extremes. *Journal of Climate*, *25*(19), 6770–6780. doi: 10.1175/JCLI-D-11-00705.1

Prein, A. F., & Gobiet, A. (2017). Impacts of uncertainties in European gridded precipitation observations on regional climate analysis. *International Journal of Climatology*, *37*(1), 305–327. doi: 10.1002/joc.4706

Rhodes, R. I., Shaffrey, L. C., & Gray, S. L. (2015). Can reanalyses represent extreme precipitation over England and Wales? *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 1114–1120. doi: 10.1002/qj.2418

Scherrer, S. C., Fischer, E. M., Posselt, R., Liniger, M. A., Croci-Maspoli, M., & Knutti, R. (2016). Emerging trends in heavy precipitation and hot temperature extremes in Switzerland. *Journal of Geophysical Research*, *121*(6), 2626–2637. doi: 10.1002/2015JD024634

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018, mar). A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons. *Reviews of Geophysics*, *56*(1), 79–107. doi: 10.1002/2017RG000574

Tarek, M., Brissette, F. P., & Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrology and Earth System Sciences*, *24*(5), 2527–2544. doi: 10.5194/hess-24-2527-2020

Tencaliec, P., Favre, A. C., Naveau, P., Prieur, C., & Nicolet, G. (2019). Flexible semiparametric generalized Pareto modeling of the entire range of rainfall amount. *Environmetrics*(May), 1–20. doi: 10.1002/env.2582

Timmermans, B., Wehner, M., Cooley, D., O'Brien, T., & Krishnan, H. (2019). An evaluation of the consistency of extremes in gridded precipitation data sets. *Climate Dynamics*, *52*(11), 6651–6670. doi: 10.1007/s00382-018-4537-0

Tramblay, Y., Neppel, L., Carreau, J., & Najib, K. (2013). Non-stationary frequency analysis of heavy rainfall events in southern France. *Hydrological Sciences Journal*, *58*(2), 280–294. doi: 10.1080/02626667.2012.754988

Wang, C., Graham, R. M., Wang, K., Gerland, S., & Granskog, M. A. (2018). Comparison of ERA5 and ERA-Interim near surface air temperature and precipitation over Arctic sea ice: Effects on sea ice thermodynamics and evolution. *The Cryosphere Discussions*, 1–28. doi: 10.5194/tc-2018-245

Xie, P., Joyce, R., Wu, S., Yoo, S. H., Yarosh, Y., Sun, F., & Lin, R. (2017). Reprocessed, bias-corrected CMORPH global high-resolution precipitation estimates from 1998. *Journal of Hydrometeorology*, *18*(6), 1617–1641. doi: 10.1175/JHM-D-16-0168.1

Table 1. Mean Absolute value of the Difference in the Number of Wet Days

EOBS				CMORPH			
DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
14.5%	20.7%	18.5%	10.9%	66.1%	73.7%	75.2%	76.0%

Note: the mean absolute value of the difference in the number of wet days is defined in equation 6, and is computed here for grid points with more than 500 wet days for ERA-5 vs EOBS and for grid points with more than 200 wet days for ERA-5 vs CMORPH.

Table 2. Mean Hit Rate ERA-5 vs EOBS and ERA-5 vs CMORPH

Percentile	EOBS				CMORPH			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
75^{th}	76%	75%	73%	77%	74%	73%	72%	73%
90^{th}	66%	65%	61%	67%	62%	61%	60%	60%
95^{th}	59%	58%	53%	60%	53%	52%	51%	52%
99^{th}	44%	45%	39%	45%	37%	35%	35%	36%

Note: for a given percentile, the mean is computed over all grid points where the precipitation percentile is larger than 1 mm. See section 3.1 for the definition of the hit rate.

Appendix A Number of Wet Days

Appendix B Hit Rate

Table 3. Summary of the Wet Day Precipitation Distribution Comparison of ERA-5 With the Observational Data Sets

Precipitation intensity	EOBS				CMORPH			
	DJF	MAM	JJA	SON	DJF	MAM	JJA	SON
Low $p = 0.3$	82%	81%	39%	82%	94%	93%	92%	93%
Median $p = 0.5$	81%	90%	72%	90%	90%	86%	87%	88%
Moderate $p = 0.75$	72%	89%	93%	90%	79%	76%	76%	75%
High $p = 0.9$	69%	85%	93%	87%	75%	72%	73%	70%
Extreme $p = 0.95$	73%	87%	94%	87%	76%	73%	74%	71%
whole distrib.	29%	34%	10%	39%	57%	55%	53%	52%

Note: for a given non-exceedance probability p, the percentage denotes the proportion of grid points for which the confidence intervals are overlapping. For the whole distribution, the percentage denotes the proportion of grid points where the null-hypothesis can not be rejected, i.e. where the distributions are similar.

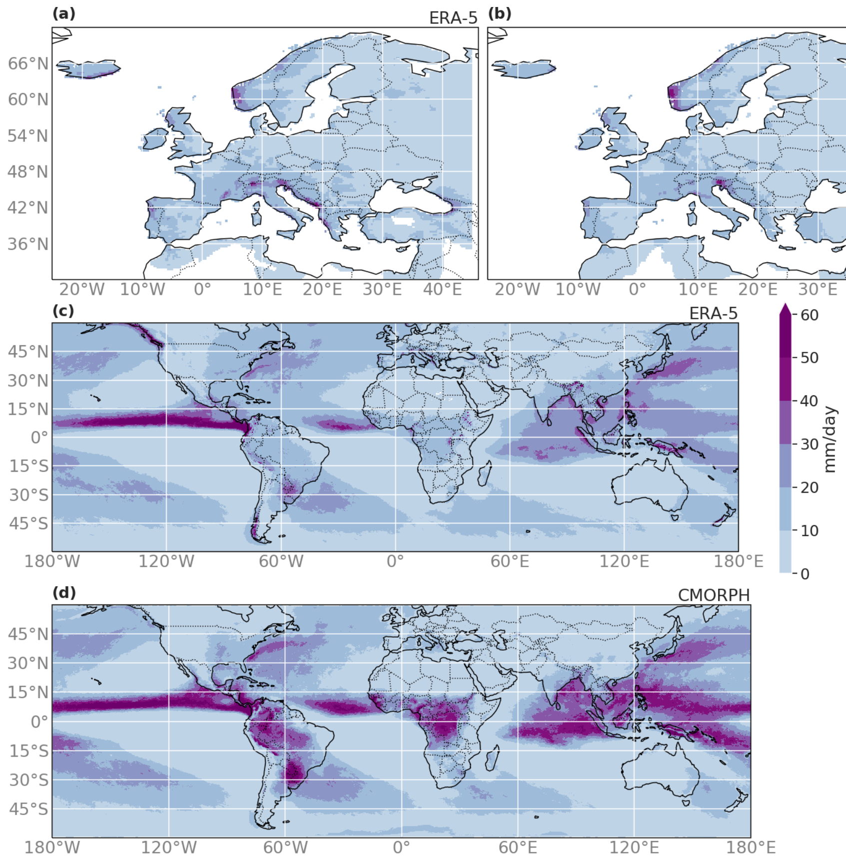

Figure 1. 95^{th} precipitation percentile (mm) for all days in SON for (a) ERA-5 over Europe 1979-2018 (b) EOBS 1979-2018 (c) ERA-5 globally 2003-2016 (d) CMORPH 2003-2016.

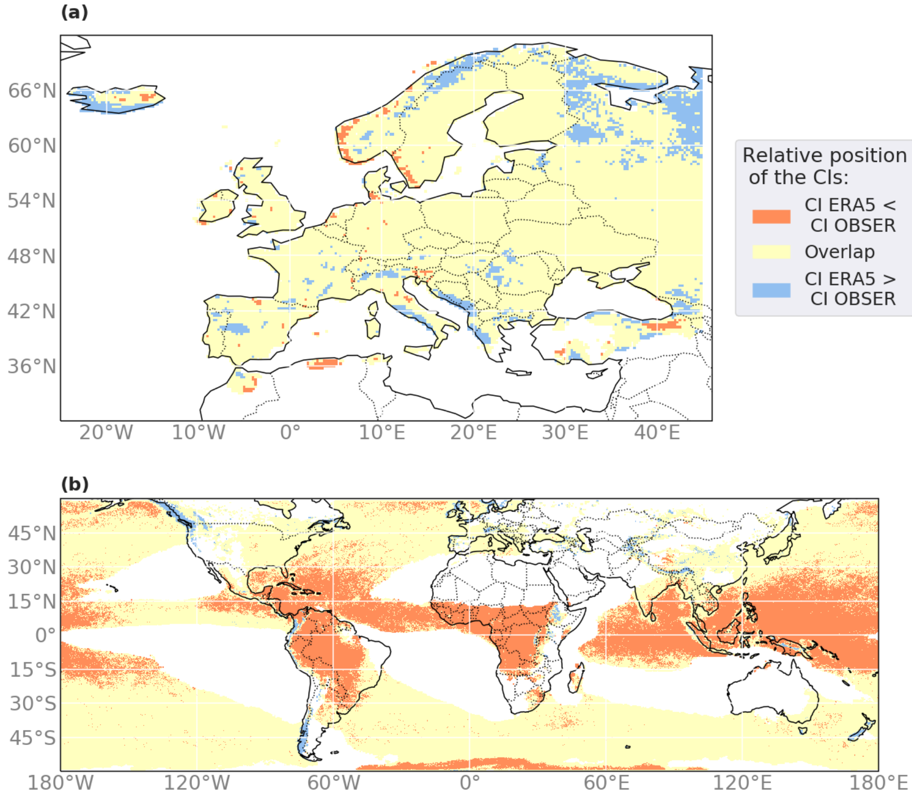

Figure 2. Relative position of the confidence intervals (CIs) for SON quantiles associated with non-exceedance probability 0.9 between (a) ERA-5 and EOBS (1979-2018) and between (b) ERA-5 and CMORPH (2003-2016). See section 3.2.2 for computational details. Grid points with an insufficient number of wet days (see section 3.2) are discarded and displayed in white.

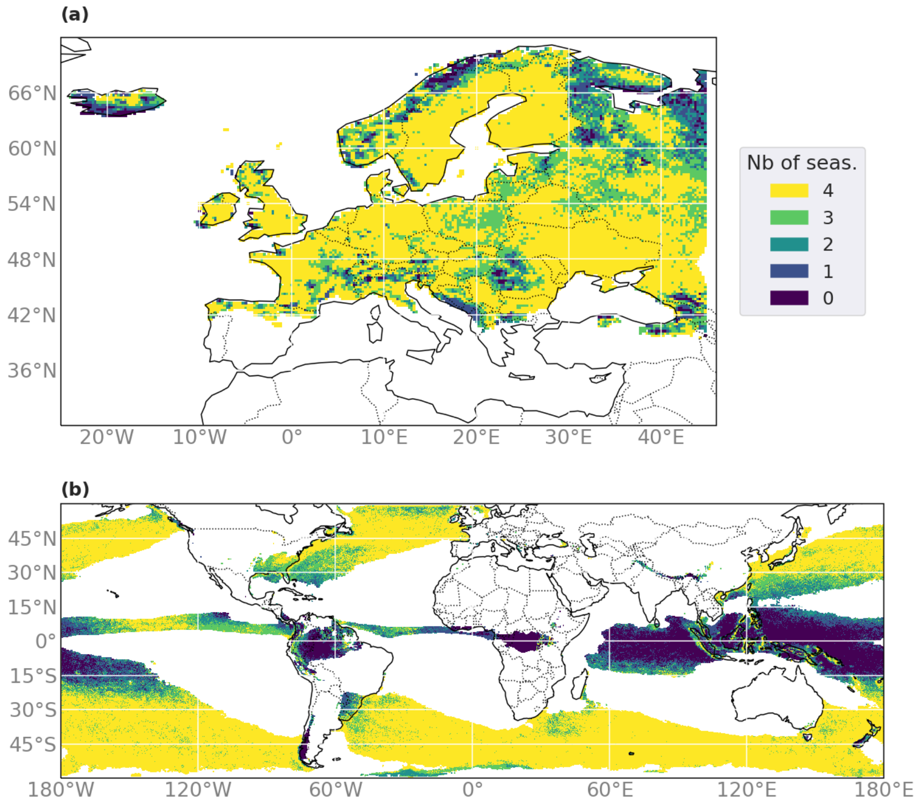

Figure 3. Number of seasons with overlapping confidence intervals for quantiles associated with non-exceedance probability 0.9 between (a) ERA-5 and EOBS (1979-2018) and between (b) ERA-5 and CMORPH (2003-2016). See section 3.2.2 for computational details. Grid points with an insufficient number of wet days (see section 3.2) are discarded and displayed in white.

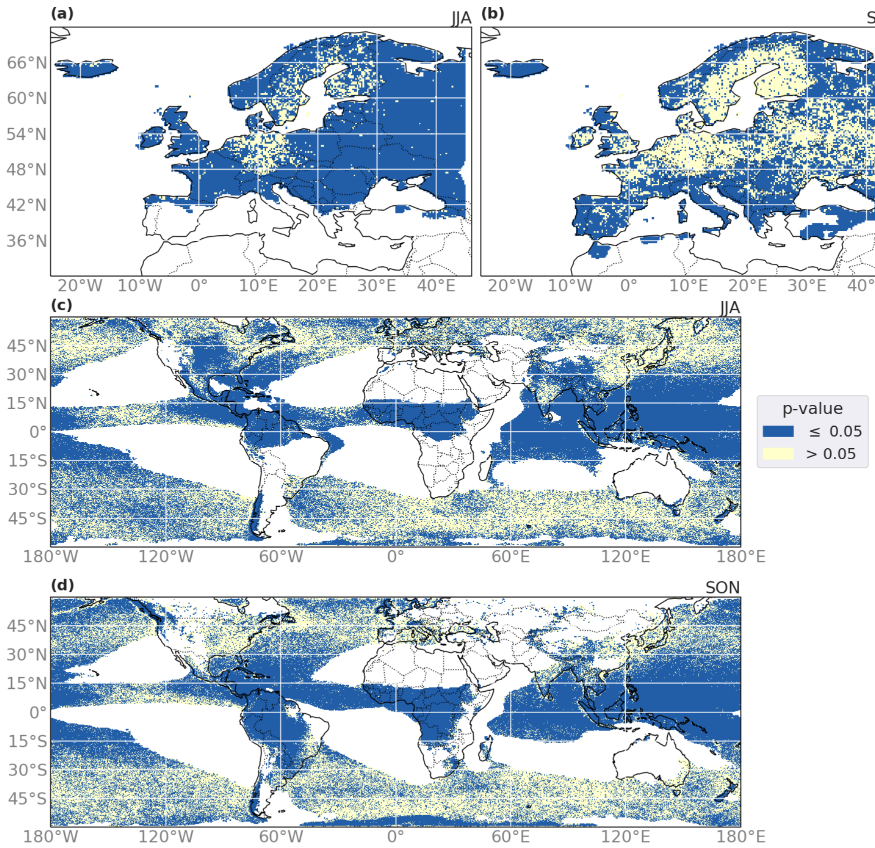

Figure 4. p-value of the Kullback-Leibler divergence test (as defined in section 3.2.3) between ERA-5 and EOBS (1979-2018) in JJA (a) and SON (b) and between ERA-5 and CMORPH (2003-2016) in JJA (c) and SON (d). p-values ¡0.05 indicate that the distributions differ significantly. Grid points with an insufficient number of wet days (see section 3.2) are discarded and displayed in white.

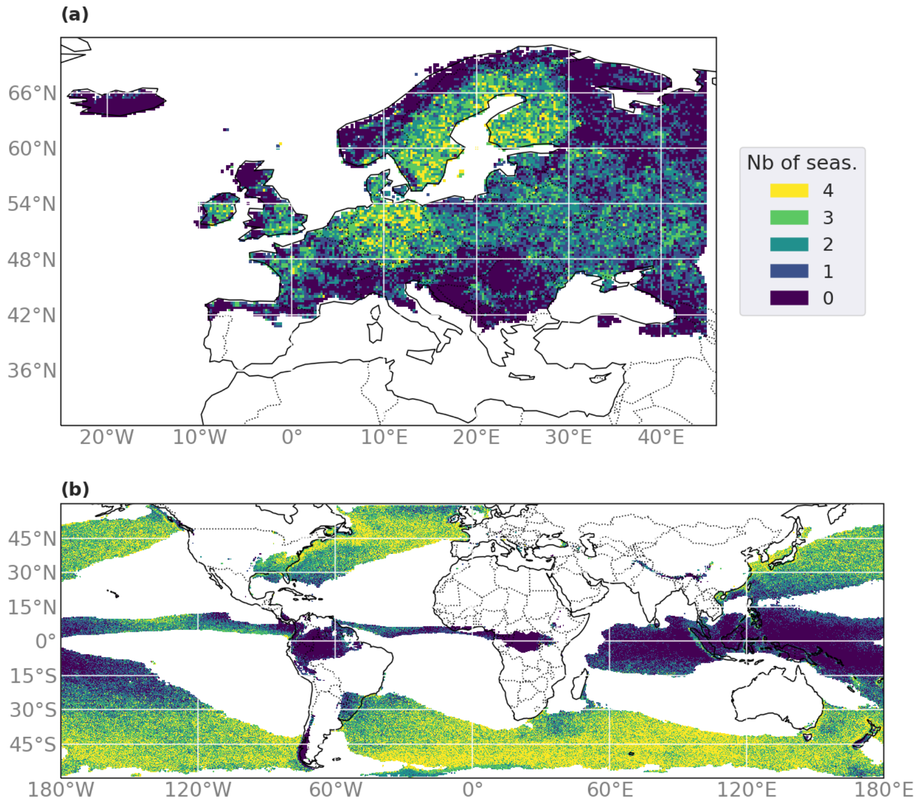

Figure 5. Number of seasons where the distributions are similar, i.e., without rejection of the null hypothesis of the Kullback-Leibler test (as defined in section 3.2.3) (a) between ERA-5 and EOBS (1979-2018) and (b) between ERA-5 and CMORPH (2003-2016). Grid points with an insufficient number of wet days (see section 3.2) are discarded and displayed in white.

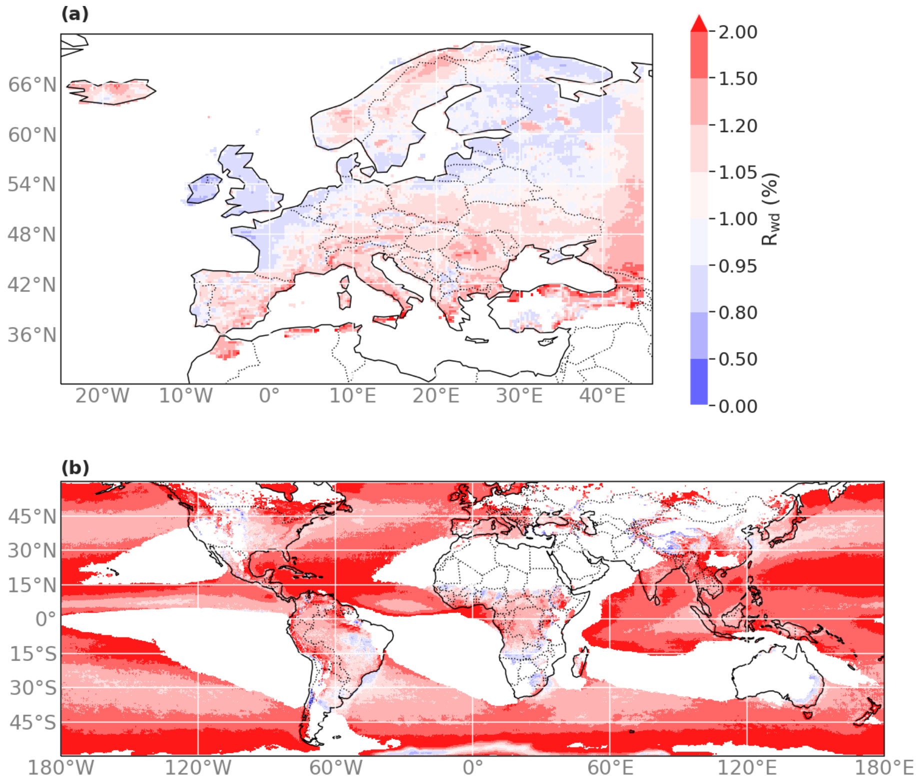

Figure A1. Ratio of the number of wet days as defined with Eq. (5) in SON between (a) ERA-5 and EOBS (1979-2018) and between (b) ERA-5 and CMORPH (2003-2016). Grid points with an insufficient number of wet days (see section 3.2) are discarded and displayed in white.

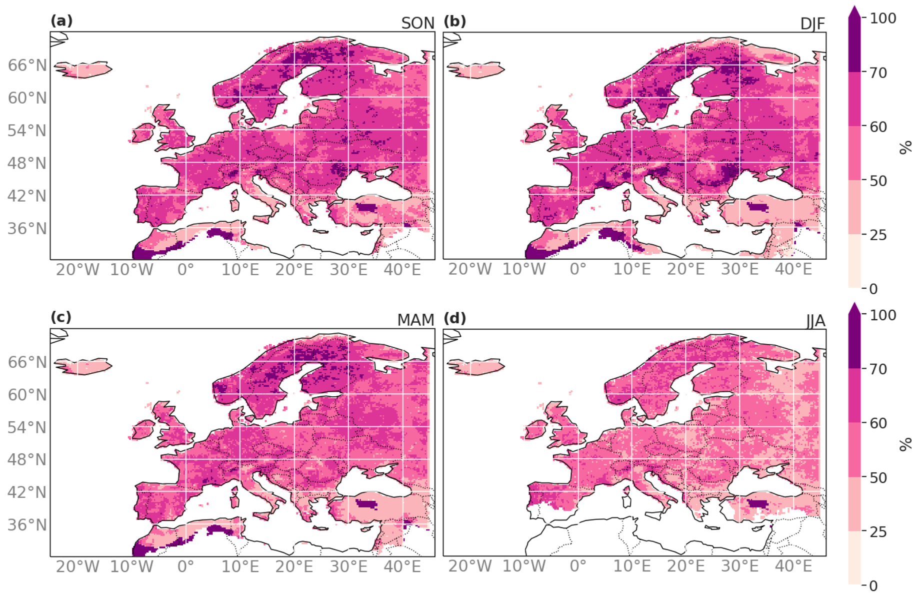

Figure B1. Hit rate for events greater than the 95^{th} percentile between ERA-5 and EOBS in (a) SON, (b) DJF, (c) MAM and (d) JJA.

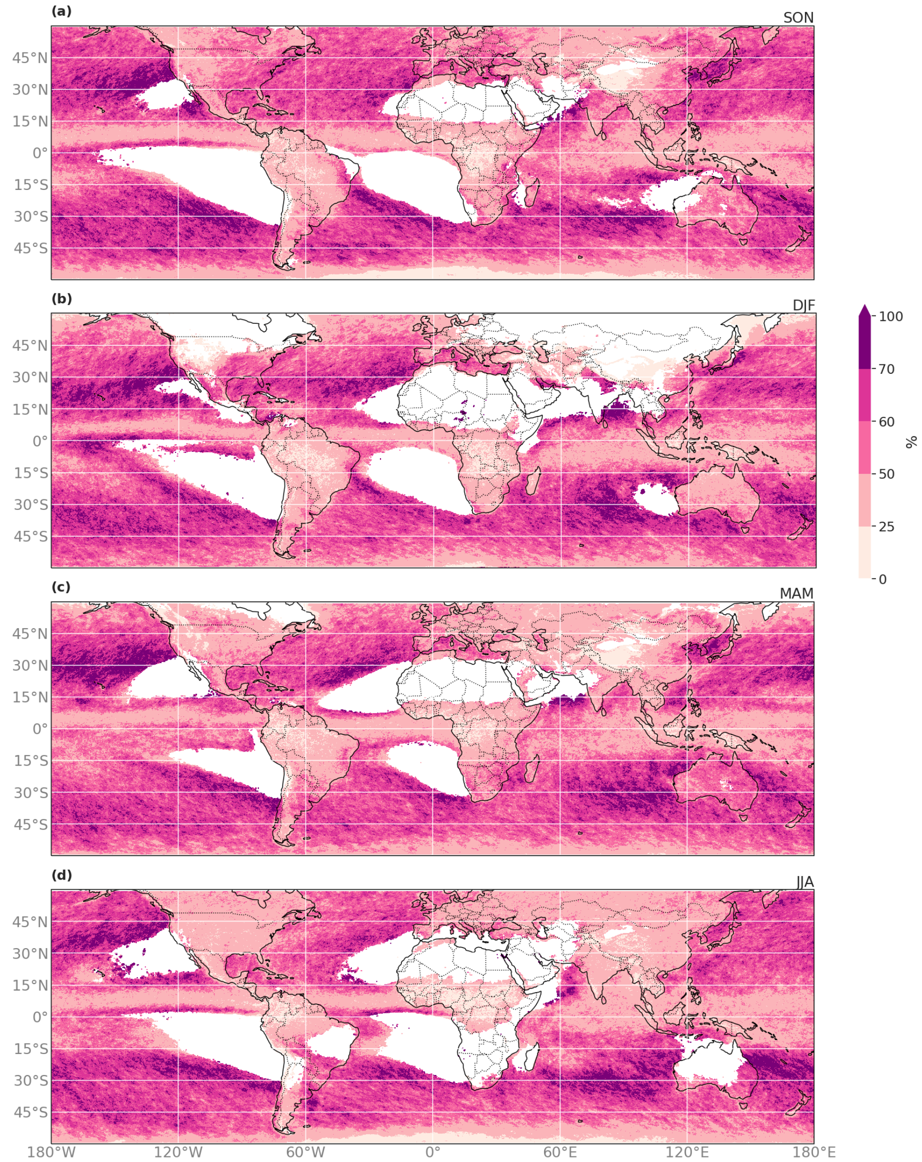

Figure B2. Hit rate for events greater than the 95^{th} percentile between ERA-5 and CMORPH in (a) SON, (b) DJF, (c) MAM and (d) JJA. Grid points with an insufficient number of wet days (see section 3.2) are discarded and displayed in white.