Global and Regional Increase of Precipitation Extremes under Global Warming

Simon Michael Papalexiou\textsuperscript{1,2} and Alberto Montanari\textsuperscript{3}

\textsuperscript{1}Department of Civil, Geological and Environmental Engineering, University of Saskatchewan, Canada
\textsuperscript{2}Global Institute for Water Security
\textsuperscript{3}University of Bologna, DICAM, Bologna, Italy

Global warming is expected to change the regime of extreme precipitation. Physical laws translate increasing atmospheric heat into increasing atmospheric water content that drives precipitation changes. Within the literature, general agreement is that extreme precipitation is changing, yet different assessment methods, datasets, and study periods, may result in different patterns and rates of change. Here we perform a global analysis of 8730 daily precipitation records focusing on the 1964-2013 period when the global warming accelerates. We introduce a novel analysis of the $N$ largest extremes in records having $N$ complete years within the study period. Based on these extremes, which represent more accurately heavy precipitation than annual maxima, we form time series of their annual frequency and mean annual magnitude. The analysis offers new insights and reveals: (1) global and zonal increasing trends in the frequency of extremes that are highly unlikely under the assumption of stationarity, and (2) magnitude changes that are not as evident. Frequency changes reveal a coherent spatial pattern with increasing trends being detected in large parts of Eurasia, North Australia, and the Midwestern United States. Globally, over the last decade of the studied period we find 7\% more extreme events than the expected number. Finally, we report that changes in magnitude are not in general correlated with changes in frequency.

**Key Points:**
- Global analysis on the 50 largest precipitation extremes over the intensified global warming period (1964-2013)
- Novel method to investigate changes in extremes' frequency shows global and regional changes
- Increasing trends in extremes' frequency are unlikely under the assumption of stationarity
- Trends in frequency and magnitude of precipitation extremes are not correlated

1. **Introduction**

There is a long list of impacts related to extreme precipitation and some of them are societally relevant. Extreme precipitation can stress severely water treatment plants, sewage networks, and play a key role in outbreaks of waterborne disease (Curriero et al., 2001). Heavy rainfall can increase the microbial contaminants on runoff and impact public health (Parker et al., 2010). Intense storms can affect agricultural production by severely damaging crops (Rosenzweig et al., 2002), impose negative consequences in terrestrial ecosystems (Knapp et al., 2008), trigger fatal landslides (e.g., Martelloni et al., 2012), increase the risk of infrastructure failure and damage (e.g., Nissen & Ulbrich, 2017), and worsen the conditions of daily traffic (Cools et al., 2010). Yet the most immense impact of
heavy precipitation regards the prospect to generate heavy flooding—a risk that could be increased in urban areas (impervious surfaces), and also, in coastal communities affected by rising sea levels (e.g., Wdowinski et al., 2016).

For example, the critical role of extreme precipitation is manifested by more than a half million fatalities caused by rain-induced floods from 1980 to 2009 (Doocy et al., 2013) with the number of humans affected by floods reaching almost three billion over the same period (Jonkman, 2005). The cost of future projections of rain-induced flood damages is alarming (e.g., Hallegatte et al., 2013), while the flood damage cost in Unites States has increased consistently throughout the twentieth century (Downton et al., 2005). This does not necessarily imply that these increases are caused by changes in the flood regime, as human and financial losses could be attributed in societal shifts increasing the vulnerability to extremes (Changnon et al., 2000). Interestingly, some studies indicate that increases in heavy precipitation are not reflected in flood magnitudes (Hirsch & Archfield, 2015; Mallakpour & Villarini, 2015; Sharma et al., 2018), yet for the Unites States there is evidence pointing to increased flood frequency (Mallakpour & Villarini, 2015). In any case, there is an undisputed relationship between precipitation and flooding with flooding events following extreme precipitation being reported all over the globe (e.g., Deng et al., 2016; Rebora et al., 2013). For these reasons, understanding and identifying changes in frequency and magnitude of precipitation extremes is vital to develop mitigation strategies ranging from management policies to infrastructure adaptation.

An increase in precipitation extremes is expected (e.g., Allan & Soden, 2008; Fowler & Hennessy, 1995; O’Gorman & Schneider, 2009; Trenberth, 2011), in particular for short duration precipitation (minutes to daily) in convective events, whose dynamics are highly non-linear and therefore more sensitive to perturbations (Lenderink & Van Meijgaard, 2008; Westra et al., 2014). The Clausius-Clapeyron equation dictates a 7% increase in atmospheric capacity to hold water for every 1°C temperature increase (e.g., Pall et al., 2007; G. Wang et al., 2017). While climate models confirm that rainfall extremes may increase under global warming (Wentz et al., 2007), a comparison of modelled and observed precipitation shows that models may underestimate the increases in short-duration rainfall extremes (Formayer & Fritz, 2017; Lenderink & Van Meijgaard, 2008, 2010; Mishra et al., 2012). Interestingly, however, there are studies indicating that a large number of climate models predict precipitation increases at just 2%/°C (see e.g., Richter & Xie, 2008; Wentz et al., 2007). Changes in precipitation extremes have been pointed out regionally, e.g., in the tropics (Allan & Soden, 2008), in India (Goswami et al., 2006), in dry and wet regions (Donat et al., 2016), in North America (Kunkel, 2003) and in China (Y. Wang & Zhou, 2005), while changes in the flooding regime have also been reported, e.g., in Europe (Alfieri et al., 2016; Blöschl et al., 2017) and globally (Hirabayashi et al., 2013; Tanoue et al., 2016). Previous global studies of trends in precipitation extremes, which mainly use gridded data (e.g., Alexander et al., 2006; Donat et al., 2013) and analyzed various extreme precipitation indices detecting changes in some regions of United States and of Eurasia.

Here, we present a global assessment of extreme daily precipitation spanning from local to global scales in order to identify and compare changes in the frequency and magnitude of daily extremes. The focus is on the 1964-2013 period during which global warming was particularly marked (see Fig. A1). We used more than 8730 high quality daily precipitation records (Fig. 1) that were screened from more than 100,000 stations (see
Methods). The key idea lies in decomposing precipitation into frequency and magnitude. Towards this direction we: (1) formed two unique data bases of extremes (one for frequency and one for magnitude) by isolating and identifying for every $N$-year record the $N$ largest precipitation events, and (2) introduced novel approaches for trend assessment (customized for large database analysis) like the exceedance probability profile (EPP). The framework of this analysis expands and augments previous efforts and shows for first time a marked global change in the frequency of daily rainfall extremes.

2. Methods

2.1 Original Data

We use the Global Historical Climatology Network-Daily database (Menne, Durre, Korzeniewski, et al., 2012; Menne, Durre, Vose, et al., 2012) (version 3.22) that comprises approximately 100,000 precipitation stations. First, we screen stations based on their record length and data quality, according to the following criteria: (1) record length of at least 50 years, (2) percentage of missing values less than 20%, (3) percentage of values assigned with quality flags less than 1%. This screening results in a subset of records, which we further process in order to assure the quality of the data by eliminating all values assigned with “G” (failed gap check) and “X” (failed bounds check) flags which indicate unrealistically large precipitation values. After initial screening, we select only records having at least five complete years in each one of the five decades during the 1964-2013 period in order to assure even information coverage over the considered period (similar criteria have been used for previous global analyses of temperature (Alexander et al., 2006; Easterling et al., 1997; Papalexiou et al., 2018)). The 1964-2013 period was selected as there is clear acceleration of global warming during this period (see Supplementary Fig. A1), while the 2014-2018 years were excluded from the analysis as the number of stations in operation drops significantly. Finally, we require no more than 30 missing daily values to accept a year as “complete” (completeness $\geq 91.8\%$) and use it in the analysis. The set of records that has been approved for analysis comprises 8730 stations spread all over the globe (Fig. 1; for number of stations over major geographical zones see Table 1). Note that in some stations the times of observation might have changed over the history of the records. This, at least theoretically, could alter the magnitude of daily extremes due discretization errors (van Montfort, 1990; Papalexiou et al., 2016), as well as, their frequency. Yet there is no reason to assume that this could have a significant effect on the results we present.
Fig. 1. Spatial distribution of suitable stations in 5° × 5° grid cells. We consider 8730 extreme precipitation records over the 1964-2013 period.

2.2 Time Series of Extremes and Basic Framework

To investigate changes in extreme daily precipitation we first form records of extremes by extracting from each daily record of $N$ valid years the $N$ largest values, e.g., from a 50-year complete record we extract the 50 largest daily values (for brevity we term these records as $NyN$ extremes). Particularly, for each daily record of $N$ valid years (Fig. 2a) we identify the occurrence dates of the $N$ extremes (Fig. 2b) and we construct two types of time series: (1) frequency of extremes time series (denoted as EF), i.e., we count the number of extremes per year to obtain a time series of the form $\{(y_1, n_1), ..., (y_N, n_N)\}$, where $n_i$ denotes, the number of observed extremes in the year $y_i$ (Fig. 2c), with $n_1 + \cdots + n_N = N$; and (2) magnitude of extremes time series (denoted as EM), i.e., we average the extreme events that occurred within each year to obtain time series of the form $\{(y_1, \bar{x}_1), ..., (y_k, \bar{x}_k)\}$, where $\bar{x}_i = n_i^{-1} \sum_{j=1}^{n_i} x_j$, with $x_j$ and $n_i$ denoting, respectively, daily extreme precipitation values and number of observed extremes in the year $y_i$ (Fig. 2d). Note that EF series have a regular one-year time step (unless there are missing years) as extreme-free years have the value $n = 0$. In contrast, EM series, which express the average annual magnitude of those extremes, do not necessarily have a regular one-year time step as the $N$ largest daily values are not distributed uniformly throughout the $N$ years, i.e., one extreme per year. Therefore, in some years we do not observe any of the $N$ extremes; yet it is not reasonable to assign a zero magnitude to extreme-free years, as zeros are not representing extremes and this would affect the investigation of magnitude changes.
Fig. 2. Example series from a randomly selected station (database code: AQW00061705) with record length equal to 50 years. Graphs show (a) daily precipitation (prcp) time series, (b) the 50 largest precipitation values, (c) their frequency per year, (d) the mean annual magnitude.

At this point, we stress that there is no unique definition of “precipitation extremes”. For instance, the set of 27 indices recommended by the Expert Team on Climate Change Detection and Indices (ETCCDI; www.climdex.org/indices.html) considers as precipitation extremes annual peak of precipitation intensity, annual totals from days with precipitation larger than the 95th and 99th percentile, or identify as heavy and very heavy daily precipitation amounts those larger than 10 mm and 20 mm, respectively. The empirical exceedance probability (Weibull plotting position) of the smallest of the \( N_yN \) values is 0.00274 (Papalexiou et al., 2013), while the rest of the values of the \( N_yN \) sample having even smaller exceedance probabilities; this result highlights that \( N_yN \) values are indeed extremes. Also, while the maximum of an \( N_yN \) sample and the maximum of an annual maxima sample coincide, the former typically comprises larger values than those found in annual maxima. Thus, \( N_yN \) extremes might describe more accurately the empirical distribution tail, or else, the behavior of the extremes.

The core of this large-scale analysis regards the investigation of temporal changes in the frequency and magnitude of the \( N_yN \) extremes during the recent half century. A key factor is decomposing precipitation into frequency and magnitude and exploiting the definition of stationarity. Stationarity in the EF time series expects on average one \( N_yN \) extreme per year and no changes in average magnitude. We stress that frequency changes in our framework do not reflect changes in extremes above a predefined threshold. For example, a common approach is to define as extreme precipitation, events larger than 10 mm or 20 mm; then one can count their annual frequency and assess if it is changing or not (e.g., Alexander et al., 2006; Donat et al., 2013). Yet this approach may neglect the regional character of extremes as, for instance, events larger than 20 mm might be very frequent in one region and never occurring in another. Here, the framework we introduce allows to
study “relative” changes, i.e., the expected frequency of the NyN extremes under the stationarity assumption is one event per year.

We investigate changes starting at the station level and progress to regional, zonal and global analyses. In all these time series we study the slope of the observed trends aiming to assess their significance and compare results between changes in frequency and magnitude. Throughout the analysis we refer to five major study zones: whole globe (GL), North Hemisphere (NH), Northwest zone (NW; west of the Atlantic Ocean in NH), Northeast zone (NE; east of the Atlantic Ocean in NH) and Southeast zone (SE; east of the Atlantic Ocean in South Hemisphere). The Southwest zone (SW; west of the Atlantic Ocean in the Southern Hemisphere) as well as the Southern Hemisphere (SH) were not considered further because the number of high quality observed records is limited.

2.3 Assessment of Trends

We assess the significance of trends by calculating the exceedance probability \( p_k \) of the observed slopes \( \kappa \) using Monte Carlo (MC) simulation. Therefore, the test statistic is the slope \( \kappa \) of the fitted linear trend, and the null hypothesis is absence of trend. Linear regression trend lines are fitted by the least square error method. The estimated \( p_k \) shows the significance level at which the null hypothesis can be rejected (one-sided test). For example, an observed trend with \( p_k = 4\% \) can be considered significant at the 5% level but not at the 1% level. We deem that this approach, i.e., aiming to calculate the exceedance probability of the observed trends, is more complete and informative than trying to assess significance in a specific and predefined level.

The different statistical properties and nature of the EM, EF and zonal time series demand different MC schemes. We assume the process that generated the observed time series is stationary and we assess its characteristics. Once the process is known, it can be used to evaluate how probable it is to observe given trends with the considered sample size. Particularly, the steps of this scheme are: (1) Assess the stationary stochastic process that could describe the observed time series, i.e., its marginal distribution and dependence structure. (2) Generate 1000 synthetic time series for each observed one preserving the probability distribution and correlation (see Papalexiou, 2018; Papalexiou, Markonis, et al., 2018 for a unified theory for stochastic modelling). (3) Fit a linear trend in each synthetic time series and estimate its slope. (4) Fit a normal distribution \( \mathcal{N}(\mu, \sigma) \) to the 1000 synthetic slopes. (5) Calculate the exceedance probability \( p_k \) of the slope \( \kappa \) that was estimated for the observed time series based on the fitted normal distribution from the previous step, i.e., \( p_k = 1 - F_N(\kappa; \mu, \sigma) \), where \( F_N \) is the cumulative distribution function (cdf) of \( \kappa \sim \mathcal{N}(\mu, \sigma) \). This process can be applied for any time series.

Other methods, such as bootstrapping, could also be used to generate random samples. Yet for small samples (50 years), that may also be autocorrelated, these techniques, based on resampling, can limit the potential of random sample variation. In turn, this may affect the resulting distribution of statistics estimated from these samples, e.g., the distribution of the slope of fitted linear trends. The significance of a trend can be assessed based on methods like the non-parametric Mann-Kendall (MK) test. Yet the MK test does not calculate the exceedance probability of the observed slope, or else, it does not provide a trend magnitude but only significance. Moreover, the MK test is not appropriate for datasets with a large number of ties (Hodgkins et al., 2017). Our approach enables us to
use the exceedance probability profile (EPP) to evaluate how likely the observed trends are under the assumption of stationarity. The EPP does not attempt to assess trends at specific significance level, but rather uses all exceedance probability estimates to judge if a system deviates from the stationarity assumption.

Here, we provide some specific details on the differences between the MC schemes we are using. Specifically, the annual mean of extremes (EM time series) is a continuous random variable above a threshold defined as the minimum of the \( N_yN \) extremes, i.e., its absolute value depends on the record studied. Therefore, we use the three-parameter Weibull \( W(\alpha, \beta, \gamma) \) distribution to generate data, with cdf

\[
F_W(x) = 1 - \exp \left(-\left(\frac{x - \alpha}{\beta}\right)^\gamma\right) \quad (1)
\]

where \( \alpha, \beta, \) and \( \gamma \) denote, respectively, threshold (equal to the minimum of the EM time series), scale, and shape parameters; (evidence, at least for hourly extremes, shows the Weibull tail as a better model than the typically used Pareto (Papalexiou, AghaKouchak, & Foufoula-Georgiou, 2018)). We estimated the parameters \( \beta \) and \( \gamma \) using the method of moments (MoM) in order to preserve the time series standard deviation (maximum likelihood cannot guarantee this), as the standard deviation affects the distribution of the slopes resulting by the MC scheme. We verify that the \( W(\alpha, \beta, \gamma) \) distribution is a proper choice by performing a Chi-squared goodness-of-fit test at the 5% significance level. In 96.2% out of the 8730 EM samples the null hypothesis that annual mean extremes are distributed according to \( W(\alpha, \beta, \gamma) \) is not rejected. Note that at the 5% significance level, a 95% not-rejected rate is expected, which is extremely close to the estimated 96.2%. We note that observed time series do not show evidence of autocorrelation; in fact, the mean value of the sample lag-1 autocorrelation \( \hat{\rho}_1 \) is zero (which under mild assumptions implies independence) and the 90% empirical confidence interval is \([-0.14, 0.13]\) (note due to the irregular time step only pairs of consecutive years are considered in the estimation of \( \hat{\rho}_1 \)). Thus, generating data by sampling from a probability distribution is appropriate. Then, for each EM time series \( \{(y_1, \bar{x}_1), \ldots, (y_k, \bar{x}_k)\} \) 1000 random samples of equal length are generated from the fitted \( W(\alpha, \beta, \gamma) \) distribution replacing the observed values in order to preserve the time structure of the time series, i.e., the synthetic values occur on the years that the real values were observed (note that this is crucial as the way the values are distributed over a period may alter the results).

The number of extremes per year (EF time series) is a discrete random variable, thus we use the Pólya-Aeppli distribution \( \mathcal{PA}(\gamma_1, \gamma_2) \) with probability mass function (pmf)

\[
p_{\mathcal{PA}}(n) = \sum_{k=1}^{n} \gamma_1^k \exp(-\gamma_1) \frac{(1 - \gamma_2)^{n-k}}{k!} \frac{\gamma_2^k}{k} \binom{n-1}{k-1} \quad \text{for } n > 0,
\]

\[
p_{\mathcal{PA}}(0) = \exp(-\gamma_1) \quad \text{for } n = 0
\]

where \( \gamma_1 \) and \( \gamma_2 \) are shape parameters estimated using the MoM. The Chi-squared GoF results (5% significance level) show that in 95.9% of EF samples the null hypothesis of \( n \sim \mathcal{PA}(\gamma_1, \gamma_2) \) cannot be rejected. Time series show no evidence of autocorrelation for EF time series as well; the mean \( \hat{\rho}_1 \) is 0.00 and the 90% confidence interval is \([-0.22, 0.25]\). For
each EF time series, we generate 1000 random samples from the fitted $P\mathcal{A}(\gamma_1, \gamma_2)$
distribution, replacing the observed values in EF time series. Note that when count data are
involved regression methods like the Poisson can also be used; however, this assumes that
the sample is emerging from a Poisson distribution. We had fitted the Poisson distribution
and found that it cannot describe all EF samples, thus, we used the Pólya–Aeppli
distribution which can be considered as a Poisson generalization. Also, we have verified
through MC simulations that using linear regression with count data, works as anticipated,
that is, it reveals significant or non-significant trends.

Global and zonal EM and EF time series result from an averaging process, so we use a
Normal distribution $\mathcal{N}(\mu, \sigma)$ to generate these data; this hypothesis is not rejected (Chi-
squared test at 5% significance level) for all zones and both types of random variables. In
this case we found that time series show weak evidence of autocorrelation. Therefore, for
each 50-year zonal time series we estimate the $\hat{\sigma}_i$, the mean $\hat{\mu}$ and the standard deviation $\hat{\sigma}$
and generate 1000 samples using an autoregressive model AR1 preserving the above
statistics.

Finally, as climatological data are in general spatially correlated it is anticipated for
trends to show spatial clustering (see e.g., Douglas et al., 2000; Lettenmaier et al., 1994).
This would be especially true for trends between nearby sites. The effect of field significant
in estimated statistics, however, is not easy to quantify at the large spatial scales shown
here. Ideally, one would need to perform a multivariate stochastic simulation that respects
the spatial correlation of the stations (8730 stations) and use the simulated series to
investigate the variation of estimated statistics, e.g., the percentage of positive trends.
Clearly, at this scale this is computationally infeasible. Indeed, spatial correlation could
increase the variability of estimated statistics which in turn may affect the accuracy of
significance assessments. Yet analysis of a very large number of stations, spread all over the
world, assures that a large number are independent and thus provide enough information
to make the results robust.

3. Results and Discussion
Here we show results emerging from the trend analysis of the individual 8730 EF and EM
time series. We investigate and compare changes in the frequency and magnitude of NyN
extremes over 1964-2013, at the station level, by quantifying their average rate of change
and its significance based on the slope $\kappa$ of fitted linear trends to EF and EM time series. We
interpret estimated trends as an average rate of change of NyN extremes over the study
period and we acknowledge that climatic variability and global warming may alter these
values in the future (Deser et al., 2013; Trenberth, 2015). Slopes are expressed,
respectively, as number of extreme events per decade and in mm per decade (hereafter, $\kappa_+$
and $\kappa_{s+}$ indicate, respectively, positive and significantly positive trends, while $\kappa_-$ and $\kappa_{s-}$
indicate negative and significantly negative trends). We assess significance based on the
Monte Carlo scheme described in Section 2.3 and we mark as significant (positive or
negative) trends those at the 10% level (one sided).

We find that a high number of stations have $\kappa_+$ and $\kappa_{s+}$ trends in frequency at the zones
studied (Fig. 3a). We also find trends in magnitude although they are less evident than
trends in frequency (Fig. 3b); exception is the SE zone showing similar trends in frequency
and magnitude. The results are better depicted by the positive-to-negative trends ratio,
deefined as $r_{+/-} := N_{k+}/N_{k-}$, with $N_{k+}$ and $N_{k-}$ indicating, respectively, the number of
stations with positive and negative trends. We show that the \( r_{+/-} \) ratio for the frequency of extremes is clearly higher than 1 in all zones except the SE, and reaches a maximum value of 2.8 in the NE zone (Fig. 3c). For magnitude, \( r_{+/-} \) ranges from 0.8 to 1.2 with a global value of 1.1 (Fig. 3d). Under the assumption of stationarity one expects approximately equal numbers of stations having positive and negative trends, i.e., \( N_{K+} \approx N_{K-} \), thus the reported values show an increase at the global level. The corresponding significant (10% level) trends ratio \( r_{s+/-} := N_{K_{s+}}/N_{K_{s-}} \), i.e., number of stations with significant positive trends over stations with significantly negative trends (Fig. 3b) is larger than 2.4 (globally) for frequency and reaches a maximum value of 7.0 for the NE zone. For magnitude, it is larger than 1 in all zones and reaches a maximum value of 1.5 in the NW zone.

**Fig. 3.** Trends in frequency and magnitude of extreme daily precipitation over 1964-2013. Panels (a-b) show the percentage of stations with positive and negative trends in frequency and magnitude, respectively. Panels (c-d) show the ratios of positive to negative and of significant positive to significant negative trends, respectively. Results refer to globe (GL), North hemisphere (NH), Northwest (NW), Northeast (NE) and Southeast (SE) earth’s quadrants.

Additionally, we introduce here a new assessment method that we name the exceedance probability profile (EPP) which is well-suited for the analysis of large databases. The exceedance probabilities \( \overline{p}_K \) of slopes fitted to a set of time series that emerge by a stationary process follow by definition a uniform distribution. This implies
that if we split, e.g., the [0,1] range of \( \widetilde{\mathbf{p}} \) into ten intervals then we expect 10% of the time series to have \( \widetilde{\mathbf{p}} \) lying within each interval. Studying the whole exceedance probability profile instead of focusing just on significant trends at a specific level, offers a more detailed and complete picture. We see that the EPP of the estimated \( \widetilde{\mathbf{p}} \) (Fig. 4) for frequency shows large deviations in all zones (except in SE) indicating that many more stations have trends with smaller \( \widetilde{\mathbf{p}} \) values than those expected. For example, in the NE zone, more than 30% of stations have significant positive trends at the 10% level and more than 45% have at the 20% level. Therefore, MC simulation confirms significant changes for frequency of precipitation. For magnitude, the distribution of \( \widetilde{\mathbf{p}} \) is closer to uniform, therefore indicating that the significance of magnitude trends is less marked.

**Fig. 4.** Profile of exceedance probabilities of the observed slopes in zones. Graphs show the distribution of the estimated exceedance probabilities for: frequency (upper panel), and magnitude (lower panel) of extremes; the solid line indicates the expected profile under stationarity.

We note that there is no significant correlation between magnitude and frequency trends, as the cross-correlation coefficient ranges from 0.02 to 0.07 in the five zones studied showing that positive (negative) changes in frequency do not necessarily imply positive (negative) changes in magnitude. However, among the four possible combinations of trends in magnitude and frequency that can be observed in a station, i.e., (1) positive in magnitude and frequency (F+M+), (2) positive in magnitude and negative in frequency (F−M+), (3) negative in magnitude and positive in frequency (F+M−), and (4) both negative (F−M−), the percentage of stations, in all zones except SE, with F+M+ is higher than the rest, as it varies from 33.0% to 39.1% (see Table 1). The second most probable state in a station corresponds to negative changes in magnitude and positive in frequency (see Table 1). A study related to changes in frequency and magnitude of extremes over the Unites States (Karl & Knight, 1998)—using different methods however—also reports that only a portion of precipitation increases is due to frequency increases. This is an additional evidence that changes in frequency and magnitude do not necessarily coincide. Interestingly, studies focusing on changes on annual maxima also reveal more significant trends than those expected. For example Westra et al. (2012) report 8.5% significant positive trends in annual maxima at the 5% significant level (two-sided test). Of course, these results are not directly comparable with ours as NyN extremes do not coincide to
annual maxima, we used different methods, different periods, and different daily records. Finally, increases in frequency have also been reported for the 2-day precipitation events exceeding station-specific thresholds for a 5-year recurrence interval in the contiguous United States (Wuebbles et al., 2017).

The spatiotemporal variation of frequency and magnitude of precipitation extremes, over 1964-2013, is investigated in 5°×5° cells, by averaging the corresponding EF or EM time series in each cell. It should be clear that these regional EF or EM timeseries do not necessarily coincide with the EF or EM series that would emerge by extracting the $N_yN_x$ extremes from the spatial daily precipitation at the $5°×5°$ resolution. This requires to use either gridded products that assimilate radar, satellite, and observations (e.g., Sun et al., 2018), or observation-based products using interpolation methods (e.g., Schamm et al., 2014). These products provide spatial precipitation time series, yet typically are too short in length, and they may show bias in the variance and consequently in extremes (e.g., Beguería et al., 2016). Here, the regional EF and EM time series we form, and therefore the detected changes or no-changes shown, should be interpreted as a measure of the “average” change of the individual stations within each cell (or zone). Using spatial precipitation or averaging first the daily series and then extracting the $N_yN_x$ might affect the results, yet it is out of the scope of this study to compare the results using different methods.

Note that for frequency we use absolute values since the mean value of EF time series is 1 ($N$-events are selected for an $N$-year record) while for EM we standardized each time series to zero-mean (e.g., Easterling et al., 1997; Jones et al., 2012; Papalexiou et al., 2018; Vose et al., 2005) as anomalies are more representative for large regions than absolute values. We show the results in a series of fifty annual maps starting in 1964 (see Supplementary Movies 1 and 2 for frequency and magnitude, respectively). Spatial variations emerge in every year, but we also found spatial coherence, especially for frequency maps, as many adjacent cells have similar values.

Changes in extreme frequency (Fig. 5a) show strong spatial coherence and relevant changes along time. A large region covering almost the whole of Europe to the western Russia shows strongly positive trends. Marked changes are also observed in eastern Russia, while most of China, excluding a central-north region, shows mild to strong positive trends. In Australia we observe some high-value cells mainly in the north, yet in general there is a balance with 19 and 21 positive and negative trend cells, respectively. The United States, excluding the west coast, shows positive trends in frequency with the most intense changes shown in the north-eastern part. At the global level, 66.4% of the grid cells studied show positive changes. Other recent studies, using gridded data and investigating changes in the frequency of daily precipitation ≥ 10mm also find changes over Europe and Asia (e.g., Donat et al., 2016). These results, however, depend on the dataset analyzed with different datasets revealing different spatial change patterns, and refer also to a different period than the one we analyze.

Some regions with high values in 50-year magnitude trends (Fig. 5b) are detected in Eurasia, e.g, in Vietnam-Cambodia and Thailand and in central Russia (north of Mongolia). Most cells in western Europe, spanning from Portugal to northern Norway, show positive changes while some low-value cells are observed in central and eastern Europe. North Australia has more positive trends in magnitude than southern-centraustralia. Over North America, most cells show positive change, yet a large region with low negative trends
spans from Montana and North Dakota to Texas. At the global level, 56.7% of the 393 analyzed grid cells show positive changes. Analysis of annual daily maxima and of very wet days (defined as days with annual total precipitation >95th percentile) show positive changes in South America, Asia, and Africa (e.g., Donat et al., 2016). Again, these results depend on the gridded product analyzed and a direct comparison with results shown here is not informative; these studies use different datasets, different methods, and refer to different periods.
Fig. 5. Mean trend values in 5° × 5° grid cells in extreme daily precipitation over the period 1964-2013. Maps show trends in (a) frequency as number of extreme events per decade, (b) magnitude as mm per decade.

Global and zonal time series are estimated by area-weighting and averaging the corresponding grid-cell data in each zone (a similar approach has been adopted by other global studies, e.g., Caesar et al., 2006; Easterling et al., 1997; Papalexiou et al., 2018; Vose et al., 2005). The exceedance probabilities (Fig. 6; see Section 2.3 for the assumptions used to estimate these probabilities) of the fitted trends indicate an undisputed difference between changes in frequency and magnitude. The global average change in frequency has an exceedance probability \( \bar{p}_k = 0.3\% \); this provides evidence of a marked increase in the frequency of extremes. This reveals that the distribution of the NyN extremes over the 50-year period deviates markedly from the anticipated behaviour. For example, the fitted trend at the global scale (Fig. 6) shows for 1964 and 2013, respectively, 7.5% less and 7.9% more extremes than those expected. Trends in magnitude are less marked as shown by the exceedance probability of \( \bar{p}_k = 26.5\% \). In summary, in all zones with the exception of the SE zone there is clear evidence of increases in extreme event frequency (Fig. 6) while changes in magnitude are less pronounced, i.e., in all zones magnitude trends have exceedance probabilities larger than 10% with the exception of the NW zone with \( \bar{p}_k = 5.1\% \).
Fig. 6. Mean trend values of frequency (events/decade) and magnitude (mm/decade) of extreme daily precipitation in large geographical zones over the period 1964-2013. Maps show the results for globe, North hemisphere, Northwest quadrant, Northeast quadrant, Southeast quadrant (zones are indicated by insets with global maps). The smooth line shows the 7-year moving average.
4. Conclusions

We used 8730 high quality daily precipitation records from all over the globe in order to investigate changes in the frequency and magnitude of extremes during the 1964-2013 period, when the global warming accelerated. For each record of $N$ complete years we identified as extremes the $N$ largest precipitation values. These extremes represent more accurately the heavy precipitation properties compared to annual maxima series and allow investigation of frequency changes as they are not distributed evenly each year. The initial set of records was used to construct two databases of time series describing: (1) the number of those extremes per year (frequency), and (2) their mean annual magnitude. The analysis reports results at the station level and at regional, zonal and global scales.

Our analysis covers the 1964-2013 period, when the global warming accelerated, and reveals: (1) increasing trends in the frequency of daily precipitation extremes that are highly unlikely under the assumption of stationarity, and (2) magnitude increasing trends that are in general not as evident.

For frequency, most regions of the world have a larger number of stations with positive trends than negative, with a global positive/negative ratio equal to 1.5. In Eurasia (NE zone) this ratio is 2.8 with 74% of records showing positive trends (Fig. 3). The ratio of significant-positive to significant-negative trends, however, is much higher, with a global value of 2.4 and reaching up to 7.0 for the NE zone. We find strong spatial coherence in the regional pattern of frequency changes (Fig. 5a) including a large region of Europe extending up to the western parts of Russia with intense positive trends. Globally, 66.4% of the grid cells studied show positive changes. Global and zonal frequency trends show very low exceedance probabilities (exception is the SE zone) under the stationarity assumption (Fig. 6; left panel); the global value is as low as 0.3%.

For magnitude, analysis of the stations indicates that increasing trends are slightly more frequent than decreasing, e.g., the global positive/negative trends ratio is 1.1. The significant-positive to significant-negative trends ratio is higher (1.3 for the globe), yet it does reveal a striking difference. The spatial pattern of the magnitude of extremes (Fig. 5b) is not as coherent compared to patterns shown for frequency, e.g., some regions in Eurasia show acceleration rates, yet there are also regions with decreasing trends. This fact is also reflected in the exceedance probabilities of the global or zonal magnitude trends (Fig. 6; right panel) which do not indicate highly unlikely trends; expectation is the North America (NW zone) trend having a 5.2% exceedance probability.

We highlight that this analysis and results shown regard the 1964-2013 period and do not claim that the observed trends will continue in the future. Climatic natural variability and global warming might alter markedly the reported rates of change; however, most of the climate models predict increasing future trends. Also, literature shows that results related to trends, such as spatial patterns and rates of change, might be influenced by the data product analysed, the methods used, and the study period, yet a general agreement seems to exist on the changing nature of precipitation extremes. Finally, we note that trends are still not known for many areas where gage records are short and geographically sparse.

Acknowledgements

We thank the reviewers and the AE for their constructive and very detailed reviews. The manuscript has been greatly improved due to their efforts. SMP was funded by the Global
Water Futures program (https://gwf.usask.ca); AM was partially supported by the Italian Government through the grant “Excellent Department” that was awarded to the Department of Civil, Chemical, Environmental and Material Engineering at the University of Bologna.

**Contribution**

SMP and AM conceived, designed the study, and wrote the manuscript. Analysis was performed by SMP.

**Data availability**

The database used in this study is the GHCN-Daily and is freely available by NCEI at: https://www.ncdc.noaa.gov/ghcn-daily-description. The stations’ identification codes are provided in the supplementary file Stations.csv.

**Competing interests**

The authors declare no competing interests.

**References**


Confidential manuscript submitted to Water Resources Research


Tables

Table 1. Percentage of stations with positive and negative trends in frequency (F) and magnitude (M); signs + and – indicate for positive and negative trends, respectively.

<table>
<thead>
<tr>
<th>Station No.</th>
<th>F+M+ (%)</th>
<th>F−M+ (%)</th>
<th>F+M− (%)</th>
<th>F−M− (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GL</td>
<td>33.0</td>
<td>20.2</td>
<td>27.3</td>
<td>19.5</td>
</tr>
<tr>
<td>NH</td>
<td>36.4</td>
<td>18.0</td>
<td>29.2</td>
<td>16.5</td>
</tr>
<tr>
<td>NW</td>
<td>35.1</td>
<td>20.1</td>
<td>27.0</td>
<td>17.9</td>
</tr>
<tr>
<td>NE</td>
<td>39.3</td>
<td>13.0</td>
<td>34.5</td>
<td>13.2</td>
</tr>
<tr>
<td>SE</td>
<td>23.5</td>
<td>26.7</td>
<td>21.7</td>
<td>28.1</td>
</tr>
</tbody>
</table>

Annex A

Fig. A1. Global temperature anomalies. We study the 1964-2013 period when the global warming intensified.
Figure 3.
Intervals of exceedance probability

Percentage of observed slopes with exceedance Pr in various intervals

Frequency of extremes

Magnitude of extremes

GL NH NW NE SE
Figure 5.
a

Trends in frequency of extreme daily precipitation (events/decade) 1964–2013

b

Trends in magnitude of extreme daily precipitation (mm/decade) 1964–2013
Figure A1.