Observed Scaling of Precipitation Extremes with Surface Temperature and Convective Available Potential Energy (Invited Chapter for the AGU Geophysical Monograph Series “Clouds and Climate”)

Wenhao Dong\(^1\) and Yanluan Lin\(^1\)

\(^1\)Tsinghua University

November 26, 2022

Abstract

Changes in precipitation extremes remain a key uncertainty as the climate warms. Improved understanding of their evolution is crucial for effective water management. A number of studies have demonstrated various scaling relationships between precipitation extremes and several different environmental variables. In this chapter, we review recent important advances in two of these relationships primarily based on observations: The scaling of precipitation extremes with surface temperature (both air temperature and dew point temperature) and convective available potential energy (CAPE). Two up-to-date global daily datasets are also used to provide a further check on the generality of earlier findings. Known scaling relationships are used to quantify the impacts of these two factors on precipitation extremes. Results show that both of them play important roles, but their impacts vary over different regions on various time scales, highlighting the challenges of constructing global relationships to explain the changing nature of precipitation extremes.
Observed Scaling of Precipitation Extremes with Surface Temperature and Convective Available Potential Energy

Wenhao Dong¹,² and Yanluan Lin¹

¹Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China
²Now at NOAA/Geophysical Fluid Dynamics Laboratory and Cooperative Programs for the Advancement of Earth System Science, University Corporation for Atmospheric Research

Keywords
Precipitation extremes, Surface temperature, Clausius-Clapeyron relationship, Convective available potential energy

Corresponding author: Yanluan Lin, yanluan@tsinghua.edu.cn
Abstract
Changes in precipitation extremes remain a key uncertainty as the climate warms. Improved understanding of their evolution is crucial for effective water management. A number of studies have demonstrated various scaling relationships between precipitation extremes and several different environmental variables. In this chapter, we review recent important advances in two of these relationships primarily based on observations: The scaling of precipitation extremes with surface temperature (both air temperature and dew point temperature) and convective available potential energy (CAPE). Two up-to-date global daily datasets are also used to provide a further check on the generality of earlier findings. Known scaling relationships are used to quantify the impacts of these two factors on precipitation extremes. Results show that both of them play important roles, but their impacts vary over different regions on various time scales, highlighting the challenges of constructing global relationships to explain the changing nature of precipitation extremes.

1. Introduction
Precipitation extremes, referred to in particular as large precipitation events in terms of intensity or accumulation in this study, have substantial impact on human activity, agriculture, and water resources. These extreme events also play a major role in many hydrological catastrophes (Ahern et al., 2005; Easterling et al., 2000; Hallegatte et al., 2013; Knapp et al., 2008; Wasko and Sharma, 2017; Westra et al., 2014; Yin et al., 2018). For instance, flooding induced by precipitation extremes is among the top-ranking destructive natural hazards in terms of economic losses worldwide, resulting in substantial damages to infrastructure, ecosystems, and loss of life. Given the practical significance of these extreme events, a great deal of effort has been made by multiple scientific communities to assess whether and to what extent they are changing in frequency and intensity (O’Gorman and Schneider, 2009a; Trenberth, 2011; Westra et al., 2013). Many regions around the world have experienced an increase in precipitation extremes on various time scales in the past several decades (Donat et al. 2016; Min et al. 2011). The increasing tendency, although of varying magnitudes in different geographical locations, is expected to continue as temperature increases (Alexander et al., 2006; Allan and Soden, 2008; Bao et al., 2017; Berg et al., 2013; Fischer and Knutti, 2014; Groisman
et al., 2005; Kao and Ganguly, 2011; Kharin et al. 2007; Meehl et al., 2000; O’Gorman, 2014; O’Gorman and Schneider, 2009a; Pall et al. 2007).

Benefitting from extensive researches using theoretical studies, station observations, and model simulations [including idealized cloud-resolving models (CRMs) and global/regional climate models (GCMs/RCMs)], significant advances in our understanding of the response of precipitation extremes to a warming climate have been achieved (Muller and Takayabu, 2020, and references therein). Several scaling relationships have been established between precipitation extremes and various measures of temperature, including surface air temperature, surface dew point temperature, tropospheric air temperature, and atmospheric dew point temperature (Ali & Mishra, 2017; Allen and Ingram, 2002; Bui et al., 2019; Lenderink and Van Meijgaard, 2010; Martinkova and Kysely, 2020; Mishra et al., 2012; Roderick et al., 2020), and other environmental properties, such as precipitation efficiency, convective available potential energy (CAPE), vertical wind velocity, wind shear, etc. (Muller et al., 2011; O’Gorman, 2015; Pfahl et al., 2017). CAPE is an energy-based measure of atmospheric potential instability, which estimates the theoretical maximum velocity that a positive buoyant air parcel could acquire through adiabatic ascent (DeMott and Randall, 2004; North and Erukhimova, 2009).

In this chapter, we focus on the scaling relationship of precipitation extremes with surface temperature, including air temperature ($P_e-Ta$) and dew point temperature ($P_e-Td$), as well as convective available potential energy ($P_e-CAPE$) based on observational records. We do not attempt to summarize all the existing literature on these topics since similar works have been carried out over different regions. Instead, we focus more on works which have come up with different results and reasonable interpretations. A large number of studies have demonstrated a close relationship between precipitation extremes and surface temperature linked to moisture availability (Allen and Ingram, 2002; Held and Soden, 2006; Kharin et al., 2007; Lenderink et al., 2011; Lenderink and Van Meijgaard, 2008; 2010; O’Gorman and Schneider, 2009b; Pall et al., 2007). While CAPE is often used to estimate the vertical velocity of convective updrafts. An increase in CAPE indicates enhanced atmospheric instability or the positive buoyancy that would be experienced by a lifted air parcel, which could complicate the moisture-driven intensification of precipitation extremes (DeMott and Randall, 2004;
Lepore et al., 2015; Muller and Takayabu, 2020; O’Gorman, 2015; Seeley and Romps, 2015; Singh and O’Gorman, 2013). Exploring these relationships would be an important step to understand the evolution of precipitation extremes and to constrain its future changes. For each relationship, results from pioneering studies are compared and summarized. Considering that most of the previous observational studies are carried out on a regional or local scale, we complement it by analyzing two up-to-date global sets of daily station records (see section 3) to provide a further check on the universality of earlier findings. The paper is structured as follows: we commence by the definition of precipitation extremes and scaling methodology in section 2, followed by a brief description of the data used in this study. We then review the scaling relationship of $P_e$–$T_a$ and $P_e$–$T_d$ in section 4 and $P_e$–CAPE in section 5, respectively. Discussion and summary are presented in Section 6.

2. Definition of precipitation extremes and scaling methodology

How to define precipitation extremes is not a trivial thing. It can introduce considerable complexity and influence the results significantly (Pendergrass et al., 2017). Usually, a definition of precipitation extremes need to take three aspects into consideration: a metric (a threshold based on an absolute value, a certain percentile, a return period, etc.), a timescale (the length of precipitation accumulation), and a spatial scale (station records, domain-average, contiguous rain area, etc.). Different combinations of these aspects correspond to different types of weather systems across a wide range of spatial and temporal scales, characterized by different impacts. For instance, flash floods are typically associated with short but high-intensity rainstorms at small spatial scales while riverine floods are generally caused by a long-lasting rainfall system over a larger area. As a result, various extreme precipitation definitions are in use, which complicates the direct comparison and interpretation of previous results. Most commonly, a higher percentile, such as 99th or 99.9th percentile (i.e., the top 1% or 0.1% events), is chosen at a specific location to emphasize the extreme characteristics of the precipitation events (see examples in Table 1). In addition, as suggested by previous studies, including non-precipitating events or snowfall records could affect the results significantly (Schleiss, 2018). Large uncertainties were found in snowfall observation, and the response of snowfall to warming is different from that of rainfall (Lute et al., 2015; O’Gorman, 2014;
On account of these impacts, for the dataset used in this study, a precipitation event is defined conventionally with daily accumulated precipitation larger than 0.1 mm. Precipitation extremes are identified using the 99th percentile (hereafter called \( P_{99} \)) of all the precipitation events after removing snowfall records.

A binning method based on a single predictor variable is often used to calculate the \( P_e \)-related scaling. Take \( P_e-Ta \) scaling as an example, the precipitation events are binned into either equal width or equal sample size according to surface air temperature (Guo et al., 2020; Hardwick Jones et al., 2010; Lenderink and Van Meijgaard, 2008; Lepore et al., 2015). Extreme precipitation percentiles are calculated for each bin and are then used to determine the scaling rate (\( \alpha \)). Other methods, like the quantile regression approach, are also used in the literature (Ali and Mishra, 2017; Pumo and Noto, 2021; Wasko and Sharma, 2014). Unlike regular linear regression which uses the least squares to calculate the conditional mean of a response variable, quantile regression can calculate any quantile of the response variable. It is an extension of linear regression and is found to be unbiased with sample size. A comparison of different scaling methodology based on the Global Summary of the Day (GSOD) dataset can be found in Ali et al. (2018). Scaling relationships of precipitation extremes with multiple variables have also been explored before (e.g., Lepore et al., 2015; Lepore et al., 2016), but they are not considered here. In this chapter, the \( P_e \)-related scaling is determined using a least-squares linear relationship based on the binning approach with 20 equal sample sizes according to the predictor variables. A minimum of 100 samples is required in each bin (i.e., at least 2,000 effective pairs of data in total) to optimize statistical robustness by balancing the competing needs to reduce noise and to retain enough samples. The main conclusions presented here are qualitatively insensitive to the number of bins (i.e., 10 or 30 bins).

To conclude, for each station analyzed in this study, the precipitation extreme (\( P_e \)) is identified using the 99th percentile of all the precipitating days (daily precipitation larger than 0.1 mm) after removing snowfall records. While the \( P_e-Ta \) scaling is determined using a least-squares linear relationship based on the binning approach with 20 equal sample sizes of surface air temperature. Only stations with more than 2,000 pairs of observations are analyzed. Similar analyses are performed for the dew point temperature (\( P_e-Td \) scaling) and the CAPE (\( P_e-CAPE \) scaling).
3. Observational dataset

Many efforts have been made to compile precipitation products on various time scales from global weather stations, which are the principal source providing direct measures of precipitation. Products like the Global Sub-Daily Rainfall (GSDR) dataset (Lewis et al. 2019), the GSOD dataset, and the Global Historical Climatology Network (GHCN) datasets (Menne et al., 2012) have been widely used. In this chapter, we utilize daily precipitation and surface air temperature from the GHCN-Daily dataset, which is collected at over 100,000 stations in 180 countries and territories (Menne et al., 2012). The record length and period of record vary largely but, on average, have a span of decades or more. The daily surface temperature in the GHCN dataset is represented by the average of daily maximum and minimum values. As a supplementary, we also utilize daily precipitation, surface air temperature, and dew point temperature from the GSOD dataset, which is available for 13110 stations global wide. The daily GSOD dataset is derived from the hourly observations archived in the Integrated Surface Hourly (ISH) dataset. The availability of dew point temperature in the GSOD dataset enables us to explore the $P_e–Td$ scaling relationship. CAPE values are taken from version 2 of the Integrated Global Radiosonde Archive (IGRA), which contains over 2,700 globally distributed radiosonde stations (Durre et al., 2006; Durre et al., 2009). CAPE is calculated from radiosonde profiles as the vertical integral of the positive portion of the parcel buoyancy between the level of free convection and the level of neutral buoyancy. Bi-daily CAPE values (00 and 12 UTC), when available, are averaged to daily means for the subsequent analysis to reduce diurnal sampling biases.

Data from the GHCN, GSOD, and IGRA during 1950–2019 are first used to identify stations reporting data for more than 10 years. Since our aim is to establish the statistical relationship between precipitation extremes and surface temperature ($Ta$ or $Td$) or CAPE, the continuity of these variables is not required. For the $P_e–Ta$ and $P_e–Td$ scaling relationships, we make use of the GSOD dataset considering the consistency in source data among the three variables and the fact that daily temperature variables calculated from hourly records could be more representative compared to the GHCN dataset, which is merely the average of two records, i.e., the maximum and minimum daily temperature. Following the quality check procedure in Ali et al. (2018), for each GSOD station, we remove days
with the daily precipitation record covered less than 24 hours. There are 7491 stations available for the subsequent analysis. A similar $P_e$–$Ta$ scaling relationship is obtained based on the GHCN dataset (figures not shown). While for the $P_e$–CAPE scaling relationship, we pair each radiosonde station with its nearest neighbor among GHCN stations. Though similar results are achieved if we pair the radiosonde station with the GSOD dataset (figures not shown), the GHCN dataset is chosen here because it shows more effective pairs of stations compared to the GSOD dataset. Any station pair with a distance larger than 1-km is excluded from the analysis. The 1-km criterion is justified as one can associate the recorded precipitation amount and temperature with the reported CAPE from IGRA sounding. The pairing procedure translates to 712 pairs of stations global wide. As shown in Fig. 1, these stations, either from the GSOD dataset or the GHCN dataset, are reasonably well distributed over different continents, enabling a global analysis. On average, larger precipitation amounts and CAPE values are distributed over the deep tropical and coastal regions while surface air temperature distribution shows a strong dependence on latitude regulated by the solar insolation (Fig. 1). The dew point temperature has a similar distribution as surface air temperature, but it shows relatively larger values over the coastal region, which is indicative of the impacts of moisture from the surrounding oceans. Note the available number of stations is slightly different among these variables.

Figure 1. Climatological distributions of precipitation (upper-left), surface air temperature (upper-right), dew point temperature (lower-left), and CAPE (lower-right). The first three variables are based on the GSOD dataset while the last one is based on the IGRA dataset after matching with the GHCN dataset.
4. Scaling of precipitation extremes with surface temperature

We focus on the scaling relationship between precipitation extremes and surface temperature ($P_e-T_a$ and $P_e-T_d$) in this section. According to the Clausius-Clapeyron (CC) equation, the response of precipitation to temperature is linked to changes in the atmosphere’s moisture-holding capacity (Allen and Ingram, 2002; Lenderink et al., 2011; Lenderink and Van Meijgaard, 2008; 2010; O’Gorman and Schneider, 2009a; Moustakis et al., 2020). Under the assumption of constant relative humidity, CC relationship implies an increase in specific humidity at the rate of ~7% per degree of warming, a prediction roughly borne out by plenty of observational and modeling studies (Allen and Ingram, 2002; Held and Soden, 2006; Kharin et al., 2007; Muller et al., 2011; Pall et al., 2007; Romps, 2011; Sherwood et al., 2010). One may expect precipitation to increase at a similar rate as the water vapor. However, global mean precipitation is found to increase with warming at a rate of 2–3% °C$^{-1}$ across multiple observations and climate model simulations, much slower than the CC scaling (Allen and Ingram, 2002; Attema et al., 2014; Emori and Brown, 2005; Held and Soden, 2006; Norris et al., 2019; Panthou et al., 2014; Pfahl et al., 2017; Utsumi et al., 2011). Unlike the mean precipitation, precipitation extremes show a much larger increasing rate with warming, albeit with significant variations across different geographic locations (Barbero et al., 2017; Chan et al., 2016; Drobinski et al., 2018; Guerreiro et al., 2018; Lenderink and Attema, 2015; Lenderink et al., 2011; Prein et al., 2017). Table 1 lists part of the previous studies investigating the $P_e-T_a$ scaling rate ($\alpha$) over different regions. Most of these studies are evaluated in Europe and Australia, with a few recent studies in North America, East Asia, and South Asia. Generally, the $P_e-T_a$ scaling can be classified into three categories: a monotonic increasing scaling, a monotonic decreasing scaling, and a peak-like structure where $P_{99}$ first increases with $T_a$ until reaching a maximum at a threshold temperature and then declines with a further increase in $T_a$ (Berg et al., 2009; Hardwick Jones et al., 2010). The third category is also known as a hook shape or a parabolic shape of the scaling, in which two regimes emerge: one at low temperatures with the scaling rate approach CC or super-CC (i.e., scaling rate larger than 7% °C$^{-1}$) rates while the other one at higher temperatures with sub-CC rates (i.e., scaling rate smaller than 7% °C$^{-1}$) or even negative rates. The temperature threshold, usually called peak-point temperature ($T_{Peak}$), is defined as the temperature at which the fitted curve maximizes. We
calculate $T_{\text{Peak}}$ by applying a locally weighted regression (LOWESS) smoothing method (Chan et al., 2016; Utsumi et al., 2011).

Table 1. Summary of selected previous studies of the $P_e-Ta$ scaling and the $P_e-Td$ scaling in a chronological order. The references are included in the first column. $P$, $T$, and $T_d$ stands for precipitation, surface air temperature, and surface dewpoint temperature. $T_{\text{peak}}$ is defined as the temperature at which the fitted scaling curve maximizes.

<table>
<thead>
<tr>
<th>Region[ref]</th>
<th>Temporal resolution</th>
<th>$P_e$ definition (percentile)</th>
<th>Super-CC</th>
<th>Peak-like structure</th>
<th>$T_{\text{Peak}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netherlands [61]</td>
<td>Hourly/daily $P$ &amp; Daily $T$</td>
<td>99&amp;99.9th</td>
<td>YES</td>
<td>NO</td>
<td>NA</td>
</tr>
<tr>
<td>Netherlands, Belgium, and Switzerland [62]</td>
<td>Hourly/daily $P$ &amp; Daily $T$</td>
<td>99&amp;99.9th</td>
<td>YES</td>
<td>NO</td>
<td>NA</td>
</tr>
<tr>
<td>Australia [49]</td>
<td>Hourly $P$ &amp; Daily $T$</td>
<td>99th</td>
<td>YES</td>
<td>YES</td>
<td>20~26°C</td>
</tr>
<tr>
<td>Hong Kong and the Netherlands [60]</td>
<td>Hourly $P$ &amp; $T$</td>
<td>99&amp;99.9th</td>
<td>YES</td>
<td>YES</td>
<td>24°C</td>
</tr>
<tr>
<td>United States [105]</td>
<td>Hourly $T$ &amp; $P$</td>
<td>99&amp;99.9th</td>
<td>YES</td>
<td>NO</td>
<td>NA</td>
</tr>
<tr>
<td>Japan [117]</td>
<td>10-min $P$ &amp; $T$</td>
<td>99th</td>
<td>NO</td>
<td>YES</td>
<td>24°C</td>
</tr>
<tr>
<td>United States [76]</td>
<td>Hourly $P$ &amp; Daily $T$</td>
<td>99th</td>
<td>YES</td>
<td>YES</td>
<td>28°C</td>
</tr>
<tr>
<td>Germany [16]</td>
<td>5-min $P$ &amp; Daily $T$</td>
<td>99th</td>
<td>YES</td>
<td>YES</td>
<td>22°C</td>
</tr>
<tr>
<td>Canada [91]</td>
<td>Hourly $P$ &amp; Daily $T$</td>
<td>95th</td>
<td>YES</td>
<td>YES</td>
<td>20°C</td>
</tr>
<tr>
<td>United Kingdom [17]</td>
<td>Hourly $P$ &amp; Daily $T$</td>
<td>99th</td>
<td>YES</td>
<td>NO</td>
<td>NA</td>
</tr>
<tr>
<td>India [3]</td>
<td>Daily $P$ &amp; $T/T_d$</td>
<td>95th</td>
<td>YES</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Australia [10]</td>
<td>Daily $P$ &amp; $T$</td>
<td>99th</td>
<td>YES</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>South Korea [92]</td>
<td>Hourly $P$ &amp; Daily $T$</td>
<td>99th</td>
<td>YES</td>
<td>YES</td>
<td>24°C</td>
</tr>
<tr>
<td>Mediterranean [33]</td>
<td>3-hour $P$ &amp; $T$</td>
<td>99th</td>
<td>NO</td>
<td>YES</td>
<td>15~18°C</td>
</tr>
<tr>
<td>United States [103]</td>
<td>Sub-hourly $P$ &amp; $T/T_d$</td>
<td>99th</td>
<td>YES</td>
<td>YES</td>
<td>20~25°C</td>
</tr>
<tr>
<td>Mediterranean [98]</td>
<td>Hourly $P$ &amp; Daily $T$</td>
<td>99th</td>
<td>YES</td>
<td>YES</td>
<td>22°C</td>
</tr>
<tr>
<td>Global [130]</td>
<td>Daily $P$ &amp; $T/T_d$</td>
<td>99th</td>
<td>YES</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>China [46]</td>
<td>Hourly $P$ &amp; $T$</td>
<td>99th</td>
<td>YES</td>
<td>YES</td>
<td>24°C</td>
</tr>
<tr>
<td>Six macro-regions [77]</td>
<td>Hourly $P$ &amp; Daily $T$</td>
<td>99th</td>
<td>YES</td>
<td>YES</td>
<td>Wide range</td>
</tr>
</tbody>
</table>
In the following analysis, we re-examine the $P_e$–$Ta$ scaling for each station using the up-to-date GSOD dataset. The requirement of an effective sample size of 2,000 reduces the available number of stations to 5227. As shown in Fig. 2, more than 96% of the available stations report a value of $\alpha$ less than 7 $^\circ$C$^{-1}$. Only a few stations (<5%) have $\alpha$ greater than the CC rate. Positive $\alpha$ is observed within mid-to-high latitudes over 81% of the total stations while negative values (19%) are mainly observed in the tropics. This finding is consistent with previous studies on regional or local scales over tropical regions (Ali et al., 2018; Hardwick Jones et al., 2010; Maeda et al., 2012; Utsumi et al., 2011) as well as mid-latitudes (Berg and Haerter, 2013; Berg et al., 2013; Lenderink et al., 2011; Lenderink and Van Meijgaard, 2008; 2010; Shaw et al., 2011). But obviously, latitude cannot solely explain the nature of this dependence, small scale differences are noted over many regions. The negative scaling relationship over the tropics has been attributed to different processes. Most precipitation extremes in the tropics result from synoptic systems, which are usually accompanied by drops in temperature caused by colder air masses intrusion or evaporative cooling. While for higher temperature bins, moisture availability becomes the dominant limiting factor, the drier condition usually leads to fewer and less intense precipitation (Bao et al., 2017; Barbero et al., 2018; Lenderink et al., 2011, 2017; Lenderink and Meijgaard, 2011; Park and Min, 2017). When taken together, these two factors result in a negative scaling. Ali et al. (2018) pointed out the relatively smaller seasonality of surface air temperature in tropics also contributes to the observed negative scaling. After removing the temperature seasonality, they found that most locations show a positive scaling with surface air temperature.

In addition, there are 1713 out of 5227 stations (~33%) showing a peak-like structure. Among these stations, high-latitude stations or stations over mountainous regions generally have lower $T_{Peak}$, supporting the idea that $T_{Peak}$ might be related to the local temperatures. On average, the observed $T_{Peak}$ is within the range of 18–26 $^\circ$C for North America, East Asia, and Australia while a relatively smaller
Peak of about 10–18 °C is noted over Europe, which is in agreement with previous studies (Drobinski et al., 2016; Guo et al., 2020; Lenderink et al., 2011; Lepore et al., 2015; Mishra et al., 2012; Panthou et al., 2014; Park and Min, 2017; Wasko and Sharma, 2014). A number of arguments have been proposed to explain this peak-like structure. The decrease in precipitation intensity (Berg et al., 2009; Hardwick Jones et al., 2010) and precipitation duration (Haerter and Berg, 2009; O’Gorman, 2012; 2015; Utsumi et al., 2011; Wasko et al., 2015; Wei et al., 2013) at higher temperatures are argued to be possible reasons, which is realized through the impact of humidity as discussed above. Several studies have found that relative humidity flattens or decreases for surface temperature above ~25°C (Hardwick Jones et al., 2010; Lenderink et al., 2011; Singh et al., 2019; Utsumi et al., 2011; Wasko et al., 2018).

To overcome moisture limitations at higher temperatures, the dew point temperature has been advocated to be used to calculate rainfall-temperature sensitivities (Ali et al., 2018; Barbero et al., 2017; Lenderink and Van Meijgaard, 2010; Wasko et al., 2018; Zhang et al., 2019). Using dew point temperature instead of surface temperature in the assessment has been found to increase the consistency in scaling relationship across different temperature regimes (Ali et al., 2018; Barbero et al., 2018; Lenderink et al., 2018; Panthou et al., 2014; Wasko et al., 2015). This is evident in Fig. 3 as one can find that there is a significant difference in scaling relationship with $T_a$ and $Td$, especially in the tropics where for the majority of stations, positive rather than negative scaling rates closer to CC relationship are obtained when surface dew point temperature is used. More specifically, the number of stations with negative scaling rates has been halved from 19% to 9%. These contrasting results of scaling rate for $T_a$ and $Td$ are consistent with the previous studies analyzed at the regional and global scale (Ali and Mishra, 2017; Ali et al., 2018; Lenderink and Van Meijgaard, 2010; Lenderink et al., 2011; Wasko et al., 2018; Barbero et al., 2018). But similar to Ali et al. (2018), negative scaling rates of $P_e–Td$ can still be observed in the deep tropical stations (Fig. 3). And the peak-like structure is also evident in the $P_e–Td$ scaling but with fewer stations (~17%) when compared to the $P_e–Ta$ scaling. The value of $T_{peak}$ based on the dew point temperature is close to that of surface air temperature in some of the tropical stations. This may indicate a limited role of moisture in the formation of this peak-like structure in these tropical stations. Instead, it could be partly related to the argument that the
temperature measured during the occurrence of a rainfall event may not be the most relevant to calculate the scaling rate. The temperature may be affected by the rainfall event via evaporative cooling or cold air advection associated with the synoptic systems (Bao et al., 2017). Some studies found negative scaling rates could turn to positive ones in some tropical stations when the temperatures three days prior to rainfall events are used (Ali and Mishra, 2017; Visser et al., 2020; Zhang et al., 2017). Other factors, like larger-scale dynamics and orography could also affect the $P_e$–$T_a$ scaling (Mounstakis et al., 2020).

**Figure 2.** Global distribution of mean $P_e$–$T_a$ scaling rate (upper-left), peak-point temperatures (upper-right), $P_e$–$T_a$ scaling rate prior to the peak-point temperature (lower-left), and $P_e$–$T_a$ scaling rate after the peak-point temperature (lower-right). For each station, the $P_e$–$T_a$ scaling rate is determined using a least-squares linear relationship based on the binning approach with 20 equal sample sizes of air temperature. The peak-like temperature is calculated by applying a locally weighted regression (LOWESS) smoothing method. The scaling rate prior to and after the peak-like temperature are calculated similarly as the $P_e$–$T_a$ scaling rate but over different temperature ranges. Crosses in the $P_e$–$T_a$ scaling rate distribution maps denote the least-squares linear relationships that are not statistically significant ($P>0.05$).

Another distinct feature of $P_e$–$T_a/P_e$–$T_d$ relationship is the reported super-CC scaling. Several studies from Europe, North America, Australia, and Hong Kong have found a super-CC scaling at low to moderate temperature, followed by negative rates at moderate to high temperatures (Berg and Haerter, 2013; Berg et al., 2013; Blenkinsop et al., 2015; Drobinski et al., 2016; Lenderink
et al., 2017; Lenderink and Van Meijgaard, 2008; 2010; Panthou et al., 2014; Wasko et al., 2015; Westra et al., 2014). Similar results can be seen over many stations with the temperature prior to the \( T_{\text{Peak}} \) based on the daily precipitation extremes as shown in Figs. 2 and 3. This type of super-CC scaling is part of the peak-like structure as discussed above, which is commonly valid over certain temperature ranges. In addition, the existence of super-CC scaling was also found in higher-frequency precipitation dataset (see Table 1). An example can be found in Fowler et al. (2021; their Fig. 2), based on a same dataset over the Netherlands, the scaling relationship can change from CC rates for daily precipitation extremes to a 2CC rate (~14% \(^{\circ}\text{C}^{-1}\)) for 10-min precipitation extremes. There has been debate about the reason for the reported super-CC relationship. A positive feedback could exist in which excessive latent heat release increases upward motions in clouds and enhances moisture convergences (Berg et al., 2013; Catto et al., 2013; Fowler et al., 2021; Lenderink et al., 2017; Lenderink and Van Meijgaard, 2008; 2010; Trenberth et al., 2003). But their exact role is uncertain because of the very small cloud-scale dynamics involved. A shift in the circulation may lead to either enhancement or suppression of rainfall over different regions, which could make the local assessment of changes in extremes difficult.

In addition, the super-CC relationship was also found to be linked to changes in weather regimes with temperature, such as a change in rainfall type from lower intensity stratiform (large scale frontal precipitation) to higher intensity convective precipitation (Fowler et al., 2021; Haerter and Berg, 2009). This argument suggests that super-CC scaling may be most prevalent in regions where both convective and large-scale rainfall events coexist. Other viewpoints, like the mixing of large-scale flow conditions varying substantially between seasons (Berg et al., 2009; Zhang et al., 2017) may also influence the scaling. Readers are recommended to Fowler et al. (2021), O’Gorman (2015), and Westra et al. (2014) for detailed reviews about these arguments.
Figure 3. Similar as Fig. 2, but for the $P_e$–$Td$ scaling relationship.

5. Scaling of precipitation extremes with CAPE

This section focuses on the scaling relationship between precipitation extremes and CAPE ($P_e$–CAPE). Compared to the large volume of research on $P_e$–$Ta$ scaling, observational studies on scaling between precipitation extremes and CAPE are relatively sparse. The calculation of CAPE requires reliable temperature profiles, which are more difficult to obtain than those of surface variables. Early studies made use of reanalysis products to probe the $P_e$–CAPE scaling. Lepore et al. (2015) interpolated the 12-h predicted CAPE field from ERA-Interim (Dee et al., 2011) to investigate its relationship with precipitation extremes. A few follow-up studies used radiosonde observation to calculate CAPE (Dong et al., 2019; Guo et al., 2020), while others tried to retrieve it from satellite observation (Gartzke et al., 2017; Murali Krishna et al., 2019). The ERA-Interim reanalysis product was found to substantially underestimate the sounding-based measurements of CAPE across Europe (Taszarek et al., 2018). But some studies found that rainfall intensity was better correlated with CAPE from ERA-Interim reanalysis than those computed from atmospheric soundings (Barbero et al., 2018; Lepore et al., 2015). They argued that ERA-Interim reanalysis is more proximal to the rainfall measurement given its simultaneity with these station observations than the available nearby sounding-based measurements. Nevertheless, results based on these studies differ.
Figure 4. Similar as Fig. 2, but for the dependence of natural logarithms of precipitation on natural logarithms of CAPE. The upper right panel shows the CAPE values when a peak-like structure is detected. $\beta$ is the regression coefficient between the natural logarithm of precipitation and CAPE.

Given that precipitation is linked to the cloud-scale vertical velocity ($W$) in updrafts and CAPE provides an upper bound for $\frac{1}{2} W^2$ (Lepore et al., 2015; Muller et al., 2011; Muller and Takayabu, 2020), one may expect precipitation extremes scale with the square root of the CAPE (CAPE$^{\frac{1}{2}}$). This relationship is based on the assumption that the conversion of CAPE to kinetic energy is strong in the lower part of the clouds since there is not much condensation left in the upper part of the clouds. However, the real fraction of CAPE converted to upward velocity varies largely by locations and seasons (Dong et al., 2019; Guo et al., 2020; Lepore et al., 2015; Lepore et al., 2016). Following these studies, we first calculate the natural logarithm of the precipitation and CAPE and then use the following equation $\ln P \sim \beta \ln CAPE$ to examine their relationship. The advantage of this form is that $\beta$ could directly reflect the fraction of CAPE converted to upward velocity. A larger $\beta$ indicates more CAPE is converted to the kinetic energy of a rising air parcel and a theoretical value of 0.5 is the upper bound when all the CAPE is acquired by the air parcel based on the idealized parcel theory (Lepore et al., 2015; North and Erukhimova, 2009).

We re-examine the relationship between precipitation extremes and CAPE using the up-to-date global dataset. The number of available stations is further reduced to 263, which is mainly caused by
the intermittent instrument failures of radiosonde. Different from the $P_e$–$T_a$/$P_e$–$T_d$ relationship, there are no clear regional-scale differences in the $P_e$–CAPE relationship (Fig. 4). But it shares the similarity of the three major patterns of the $P_e$–$T_a$ scaling relationship. Most of the value of $\beta$ falls in the range of $-0.2$ to $0.2$, with larger positive values located over eastern North America, Europe, East Asia, and Australia. The distribution is consistent with previous studies analyzed over North America (Lepore et al., 2015; Lepore et al., 2016) and East Asia (Dong et al., 2019; Guo et al., 2020). Negative values are observed over coastal stations or stations over the oceanic island. The deviation from the theoretical value of 0.5 (assume all the CAPE is converted to kinetic energy) may arise from a combination of different influential factors. The value of $\beta$ was found to depend on the environmental humidity of the entrained air, vertical wind shear, and other factors (Derbyshire et al., 2004; Lepore et al., 2015), which varies substantially between different geographical locations. In terms of the negative $P_e$–CAPE scaling, some studies argued that convective inhibition (CIN) should be taken into consideration (Meredith et al., 2019). CIN is a measure of the energy barrier inhibiting an air parcel from rising from the surface to the level of free convection, which can undermine the relationship between precipitation extremes and CAPE (Barbero et al., 2018; Chen et al., 2020; Dong et al., 2019; Kirkpatrick et al., 2011). If CIN is small, even a modest amount of CAPE can produce updrafts strong enough for precipitation particles to coalesce effectively. Conversely, large values of CIN can suppress the occurrence of updrafts even in the presence of large values of CAPE. Rasmussen et al. (2020) pointed out that over North America if the CIN $<-200 \text{ J kg}^{-1}$, the strength of the capping inversion is often too much to overcome and convection is suppressed. Additionally, higher CAPE with strong vertical wind shear or strong steering flow can lead to fast-moving deep convective systems, which may result in less local precipitation accumulation (Barbero et al., 2019). Moreover, the daily data used here may not be able to represent the phase relationship between CAPE and precipitation (Subrahmanyam et al., 2015). The most relevant values of CAPE for a given precipitation event are those prior to the event. The daily averaged data may weaken the implied relationship between precipitation and CAPE (Berg et al., 2009; Haerter and Berg, 2009). This has been found by Guo et al. (2020) over East Asia, higher values of $\beta$ are found when radiosonde derived CAPE values more relevant to the occurrence of precipitation extremes instead of the daily means are used.
In addition to a simple $P_e$–CAPE scaling observed over North America (Lepore et al., 2015), a non-monotonic dependence of precipitation extremes on CAPE was noted over East Asia (Dong et al., 2019; Guo et al., 2020), and several sub-regions around the globe (Barbero et al. 2019). In total, there are 67 out of 263 stations showing a parabolic shape with the peak CAPE value ranging from a few hundred J kg$^{-1}$ in mid-to-high latitudes to more than 2,000 J kg$^{-1}$ in the tropics. But note that values of $\beta$ over many stations prior to or after the peak-like CAPE threshold are not significant, calling the robustness of this structure into question. And little is known about the tendency that the $P_e$–CAPE scaling levels off beyond a certain CAPE threshold. CAPE is affected by many physical processes that alter the vertical thermodynamic structure. Williams and Renno (1993) have disclosed a linear relationship between CAPE and surface wet-bulb potential temperature using tropical radiosonde data, indicating the dependence of CAPE on surface temperature and/or moisture availability may be influential factors for this peak-like structure.

**Figure 5.** Schematic plot for the different types of $P_e$–TS scaling (TS represents both the surface air temperature and the dew point temperature) reported in pioneer studies based on daily records, including positive scaling rate (upper-left), negative scaling rate (upper-right), peak-like structure (lower-left) and super-CC (lower-right) based on two precipitation regimes. Red dot in lower-left panel denotes $T_{\text{peak}}$ while red arrows in lower-left and lower-right panels denote the temperature ranges during which super-CC scaling are observed.
6. Concluding remarks

This paper has attempted to survey observational studies investigating the changing nature of precipitation extremes from the perspective of surface temperature (both air temperature and dew point temperature) and CAPE. Contributions from both factors have been found to be important for changes in precipitation extremes. Given that previous studies are analyzed separately over different regions, we re-evaluate these relationships by utilizing two up-to-date global datasets to provide a complementary test on the generality of earlier findings. Following the known scaling equations, precipitation extremes are found to respond differently to these two factors, some generalities are reasonably clear:

(1) The $P_e-T_a/P_e-T_d$ scaling is much more complex than suggested by the CC relationship. Different scaling relationships are observed across the globe, which can be generally classified into three categories: a monotonic increasing scaling, a monotonic decreasing scaling, and a peak-like structure where precipitation extremes first increase with surface temperature until reaching a maximum at a threshold temperature and then decline with a further increase in surface temperature. A schematic plot illustrating several common relationships is provided in Fig. 5. Most of the stations examined here show an intensification of daily precipitation extremes with increasing temperatures, especially over the mid- to high-latitude regions, while tropical stations show a negative scaling relationship even when dew point temperatures are used. But the number of stations characterized by negative scaling rates has been reduced from 19% based on the surface air temperature to 9% when the dew point temperature is used to calculate the scaling rates. Several factors are found to contribute to this negative scaling. Limited moisture availability at higher temperatures is one of the major reasons. Another important factor would be that precipitation extremes associated with synoptic systems are usually accompanied by drops in temperature caused by colder air mass intrusion or evaporative cooling. Besides, super-CC scaling is also noted over some stations, but this scaling is particularly prominent for high-frequency data and when analyzed over a certain temperature range. As shown in Fig. 5, super-CC scaling is often found over a temperature range prior to the $T_{peak}$ as part of the peak-like structure. While in other cases, it is also noted over a temperature range between two different precipitation regimes with small stratiform precipitation at low temperatures and large
convective precipitation at high temperatures. Many factors are found to contribute to these departures from the CC scaling and direct comparison between previous works is difficult due to the adoption of different definitions of precipitation extreme, different time periods, various temporal and spatial resolutions, the mixture of different precipitation types, local versus remote moisture availability, etc. Consequently, there is still debate in the literature as to which type of scaling is expected to be dominant in a warming climate.

(2) Though the $P_e$–CAPE scaling is comparatively weaker than the $P_e$–$T_a/P_e$–$T_d$ scaling, similar scaling patterns are identified based on the daily dataset. The schematic plot in Fig. 5 may also apply to $P_e$–CAPE scaling. Positive $P_e$–CAPE scaling is observed over North America, Europe, Australia, and East Asia. Negative values are observed over some coastal regions, suggesting that larger CAPE does not necessarily lead to larger precipitation. There are many unknowns and uncertainties associated with this scaling relationship. The CAPE prior to the precipitation extremes is more relevant and representative and should be considered in the future works. Moreover, the dependence of CAPE on surface temperature and/or moisture availability is worth further investigation.

We are aware that the depth of our understanding of the changes in precipitation extremes relies on the quality of available observations. Gauge-based datasets, like the GHCN, GSOD, and IGRA analyzed here, often suffer from systematic measurement errors and sampling issues associated with limited spatial coverage and time-varying number of weather stations. These variables, such as precipitation, CAPE, air temperature, and dew point temperature are observed merely over the land gauge, which only reflect the statistics at the sampled location. But precipitation extremes do not necessarily scale with the local land temperature considering moisture can be advected from surrounding or even remote water bodies. Alternatively, satellite-based products offer more complete spatial coverage than station records despite the overall short period of availability, their ability to reproduce the extreme precipitation is sensitive to the retrieval algorithms (Timmermans et al. 2019). While climate models may be useful to characterize precipitation globally on a longer time scale, they have limited ability to simulate small-scale convective processes associated with short-duration precipitation extremes, especially in the tropics where they largely rely on convective parameterizations to represent deep convection dynamics (Kharin et al., 2013; Kharin et al., 2007;
Meehl et al., 2000). Even when convective dynamics is resolved at convection-permitting resolutions, precipitation extremes at short durations are sensitive to the parameterization of cloud and precipitation microphysics (Bryan et al., 2003; Dwyer and O’Gorman, 2017; Lin et al., 2012). Further research is needed to quantify the relevance of these processes to the simulation of precipitation extremes. In addition, growing evidence has indicated that the intensity of sub-daily (hourly or even higher-frequency) extreme rainfall is more sensitive to climate change (Fowler et al., 2021 and references therein). These higher-frequency precipitation extremes are more relevant for flash floods and risk assessment. These events may be more suitable for examining precipitation extremes as they provide a better picture of the intermittent nature of precipitation. We therefore encourage more extensive observational and modeling studies to investigate the dominant processes that might cause precipitation extremes, especially those of short-duration, to changes in a warming climate.

In addition to the two scaling relationships summarized here, there are several other contributors playing important roles in the evolution of precipitation extremes. We want to underline another two important factors: precipitation efficiency and large-scale circulation. The impact of precipitation efficiency on precipitation extreme involves several key microphysical processes, among which the convective organization and its sensitivity to warming is one of active current research. Convective aggregation can affect the thermodynamic environment for precipitation formation. Observational studies have shown that changes in tropical precipitation were largely associated with changes in the frequency of organized convection (Dong et al., 2020a; Dong et al., 2021; Muller, 2013; Pendergrass, 2020; Tan et al., 2015). The capability of the climate model in representing these systems would provide an opportunity to achieve accurate rainfall projections in a warming climate. A detailed review of precipitation efficiency and climate sensitivity can be found in Chapter 13.

In terms of large-scale circulation impacts, considerable studies have shown that precipitation extremes are often associated with distinct synoptic and large-scale circulation patterns over different regions (Ali et al., 2021; Catto and Pfahl, 2013; Dong et al., 2018; Dong et al., 2020b; Field and Wood, 2007; Lau et al., 2008; Liu et al., 2019; Loriaux et al., 2017; Pfahl and Wernli, 2012; Sudharsan et al., 2020; Utsumi et al., 2017). Precipitation extremes are often forced by local and mesoscale factors, with the larger environment providing a favorable condition for triggering or maintaining these
smaller-scale processes. The observed relationships between the strength of large-scale circulation patterns and the occurrence of precipitation extremes are diverse. Various types of systems (tropical cyclones, extratropical cyclones, mesoscale convective systems, trough-ridge couplets, low-pressure systems, etc.) and mechanisms (fronts, atmospheric rivers, orographic ascent, etc.) have been shown to be important for precipitation extremes. Sometimes, two or more systems (e.g., extratropical cyclones and fronts) may be associated with a single precipitation extreme, with one of them acting as a trigger while the other one providing a favorable environment to sustain its development. Even the same category of a certain type of system can have multiple sub-types, which might have important regional differences. For instance, different types of mesoscale convective systems are found to be characterized with various structures, dynamical processes, and large-scale circulation patterns (Houze, 2004), leading to different impacts on the associated precipitation. Therefore, it is clear that large-scale circulation patterns are crucial for the occurrence of precipitation extremes, but detailed and process-oriented analyses are limited to regional or case studies. More analysis on these links could be a focus area for future studies. A better understanding of how different large-scale circulation patterns relate to precipitation extremes, and how these relationships vary by season and region would help to interpret any possible future changes.

Although many advances have been made in understanding changes of precipitation extremes, new challenges are emerging for both observational efforts and model development. More accurate observations for larger areas are needed to provide a coherent picture of precipitation extreme changes in response to warming. Future studies should be performed using consistent methodologies to facilitate direct comparison. Continued efforts are recommended for model development to resolve mesoscale and small-scale convective dynamics with advanced cloud and precipitation microphysics and to better simulate short-duration precipitation extremes. Intermodel comparison, including those CRM simulations with the same resolution and forcing, would be very helpful to examine whether the changes in scaling rate depend on the underlying model dynamics and physics or the imposed forcing.
Reference


70. Maeda, E. E., N. Utsumi, and T. Oki (2012), Decreasing precipitation extremes at higher temperatures in tropical regions, *Natural Hazards, 64*(1), 935-941.


