Consistent Large-Scale Response of Hourly Extreme Precipitation to Temperature Variation Over Land

Haider Ali1, Hayley J. Fowler1, Geert Lenderink2, Elizabeth Lewis3, and David Pritchard1

1School of Engineering, Newcastle University, Newcastle upon Tyne, UK, 2Royal Netherlands Meteorological Institute, De Bilt, Netherlands

Abstract Hourly precipitation extremes can intensify with temperature at higher rates than expected from thermodynamic increases explained by the Clausius-Clapeyron (CC) relationship (~6.5%/K), but local scaling with surface air temperature is highly variable. Here we use daily dew point temperature, a direct proxy of absolute humidity, to estimate at-gauge local scaling across six macro-regions for a global data set of over 7,000 hourly precipitation gauges. We find scaling rates from CC to 2 × CC at more than 60% of gauges, peaking in the tropics at a median rate of ~1.5CC. Moreover, regional scaling rates show surprisingly universal behavior at around CC, with higher scaling in Europe. Importantly for impacts, hourly scaling is persistently higher than scaling for daily extreme precipitation. Our results indicate greater consistency in global scaling than previous work, usually at or above CC, with positive scaling in the (sub)tropics. This demonstrates the relevance of DPT scaling to understanding future changes.

Plain Language Summary Globally, hourly precipitation extremes have increased during recent decades and are projected to intensify more in the future. Therefore, understanding the relationship of precipitation extremes with temperature may be helpful for risk assessment and in decision making for climate adaptions. Using station-based observations of hourly precipitation and daily dewpoint temperature, we show that precipitation extremes increase with the increase in dewpoint temperature consistently at a regional scale, and even with a higher rate at the gauge-level. This rate of increment is much higher than that for daily extreme precipitation. Our results have implications for understanding drivers of changes to extreme precipitation at different spatial scales.

1. Introduction

Short-duration precipitation extremes can cause flash floods, landslides, and debris flows with little or no warning, leading to serious socioeconomic impacts (Fadhel et al., 2018). Hence, they are highly relevant in the design of water resource and stormwater drainage infrastructure systems, especially in urban areas due to surface impermeability (Ali & Mishra, 2017). Insight into the impacts of warming on subdaily (particularly hourly to 3-hourly) extreme precipitation may, therefore, be crucial for societal decision-making for climate adaptation (Bader et al., 2018). Studying changes to historical precipitation extremes can aid the understanding of potential future changes to extreme precipitation (Allen & Ingram, 2002). Previous studies have reported that climate warming and urbanization are intensifying observed daily precipitation extremes globally (Alexander, 2016; Ali et al., 2018; R. Barbero et al., 2017; Westra et al., 2014) with some evidence suggesting enhanced increases in observed extreme hourly precipitation intensities (Guerreiro et al., 2018; Li et al., 2020).

The intensification of precipitation extremes with warming can be explained thermodynamically from the increase in humidity, or water vapor, of the atmosphere. The Clausius-Clapeyron (CC) relationship has therefore been used as a benchmark to interpret change to extreme precipitation (Fischer & Knutti, 2016; O’Gorman, 2015). This relation links air temperature and atmospheric humidity when the air is saturated, giving an increase in humidity of 6%–7% per degree warming (Trenberth et al., 2003). Under the assumption of constant relative humidity, which is partly supported by climate modeling results and physical reasoning at least over (relatively) wet surfaces, the actual humidity of the air increases at the same rate.

Recent studies have confirmed an approximately CC rate of increase in long-term trends in observations and projections of daily extreme rainfall (Fischer & Knutti, 2016; Guerreiro et al., 2018; Rajczak & Schär, 2017;
Scherrer et al., 2016; W. Zhang et al., 2019) when averaged globally or over large regions and mostly using large-scale temperature rise or even global temperature rise. To investigate the relationship between warming and the intensification of rainfall extremes, a common approach uses scaling between surface air temperature (SAT) and precipitation extremes. This approach, here called “apparent” scaling following the terminology introduced by Bao et al. (2017), employs short-term variations in temperature and precipitation, mostly caused by day-to-day synoptic variability up to seasonal variations, to derive dependencies of daily precipitation extremes on temperature.

Results of studies using apparent scaling have shown a wide range of behavior. Often, behavior close to the CC rate is obtained (Ali et al., 2018; Gao et al., 2018; C. Wasko et al., 2016). However, for some regions signs of superCC behavior, exceeding the CC rate, have been found for subdaily (mostly hourly) precipitation (Berg et al., 2013; Lenderink & Van Meijgaard, 2010; Lenderink et al., 2011; Park & Min, 2017). On the other hand, negative scaling rates, signifying decreases in precipitation intensity with warming have been found for subtropical and tropical regions (Hardwick Jones et al., 2010; Vittal et al., 2016). Where scaling rates with SAT deviate significantly from CC (Ali et al., 2018; Hardwick Jones et al., 2010; X. Zhang et al., 2017) this has been shown to be a result of confounding factors such as local cooling effects (Ali & Mishra, 2017; Bao et al., 2017), moisture limitations at higher temperatures (Barbero et al., 2018; Gao et al., 2018; Lenderink et al., 2018; Trenberth & Shea, 2005), temperature seasonality (Ali et al., 2018; X. Zhang et al., 2017), statistical methods and inappropriate modeling assumptions (Pumo et al., 2019; Wasko et al., 2015) and mixing of different rainfall types (P. Berg & Haerter, 2013; Molnar et al., 2015). The intermittent nature of precipitation can also be responsible for the deviation in scaling (Schleiss, 2018) and Visser et al. (2020) argued that this could be resolved by using dry-bulb temperature prior to the storm. Moreover, the localized effects of large-scale circulation patterns enhanced local moisture availability through upward motions and moisture convergence, and local-scale dynamics can influence scaling rates (Ali & Mishra, 2018; Guerreiro et al., 2018; G. Lenderink et al., 2017; Magan et al., 2020; Mishra et al., 2012; Pfahl et al., 2017).

This diversity of behavior, and the complexity of the physical processes involved, has led to a large debate in the literature on the potential use of apparent scaling, and how it could be related to climate change. Here, we show that a considerable part of this controversy can be resolved when using dew point temperature (DPT), which measures the actual humidity of the air instead of the humidity at saturation, extending several recent studies (Ali & Mishra, 2017; Barbero et al., 2017, 2018; Gao et al., 2018; Lenderink et al., 2018; X. Zhang et al., 2017), statistical methods and inappropriate modeling assumptions (Pumo et al., 2019; Wasko et al., 2015) and mixing of different rainfall types (P. Berg & Haerter, 2013; Molnar et al., 2015). The results obtained by W. Zhang et al. (2019) indicated consistent global scaling results at the CC rate for DPT but removal of seasonal effects on DPT and precipitation intensity has the best results (Ali et al., 2018; X. Zhang et al., 2017).

Therefore, in this paper, we focus on the following questions. Are negative scaling rates, in particular for subtropical and tropical areas, an artifact of the use of SAT in most scaling studies? Do we find more universal behavior across the globe using DPT, and how much could this behavior deviate from CC rates? How widespread is superCC behavior in hourly extremes, and do scaling rates of hourly extremes exceed those of daily extremes?

Here we use gauge observations of hourly precipitation (PPT) from the global subdaily rainfall (GSDR) data set (Lewis et al., 2019) and daily DPT from HadISD (R. J. Dunn et al., 2016) to establish, for the first time, the scaling relationship between extreme hourly precipitation and daily DPT at a global scale. We examine six main regions: the USA, Australia, Europe, Japan, India, and Malaysia with a total of 7,088 gauges which have at least 12 years of data (Barbero et al., 2019b) (start and end year varies between 1979 and 2014). We estimate scaling rates across these selected regions at different spatial scales. For every gauge, we estimate the scaling using the classic binning method (BM, Lenderink & Van Meijgaard, 2008) and check the consistency of our results against other scaling methods (quantile regression, QR; Wasko et al., 2014) and removing the seasonality in DPT (X. Zhang et al., 2017) (ZM)). The Methods section provides details on the data sets, their quality control, and the methods used to estimate scaling relationships at these different spatial scales.
2. Data and Methods

2.1. PPT and DPT Data

We obtained hourly precipitation data (PPT) from the GSDR data set (Lewis et al., 2019) which was compiled under the Global Water and Energy Exchanges (GEWEX) Hydroclimatology Panel INTENSE project (Blenkinsop et al., 2018) and has been used in many recent studies (Barbero et al., 2017, 2019a, 2019b; Guerreiro et al., 2018; Li et al., 2020; Moron et al., 2019). The GSDR data has been quality-controlled using 25 different checks to identify and remove a range of errors, such as physical and spatial consistency issues, spikes and flat lines (streaks) (Blenkinsop et al., 2017; Lewis et al., 2018). We selected six macro-regions where data was available: The United States of America (USA), Australia, Europe, Japan, India, and Malaysia to provide a comprehensive global study covering different climate zones with large latitudinal and elevation ranges. The gauges in the USA were of mixed-precision (0.25 and 2.54 mm), therefore, all the gauges in the USA were explicitly processed to have a consistent 2.54 mm precision (Barbero et al., 2019a). Although the spatial and temporal coverage of the GSDR data set is not uniform, we ensured a sufficient length of precipitation data for estimating scaling. Therefore, we only considered PPT stations which have at least 12 years of data with less than 20% missing hours in any given year (start and end year varying between 1979 and 2014; their location is shown in Figure S1).

We obtained daily DPT data from the Met Office Hadley Center observations data set: HadISD (version 2.0.2.2017f) (Dunn et al., 2019; Lewis et al., 2019). This is a global data set (8,103 stations) spanning January 1, 1931 to December 12, 2017 and is based on the Integrated Surface Data Set (ISD) from the National Oceanic and Atmospheric Administration’s (NOAA’s) National Climatic Data Center (NCDC).

The quality control of DPT data, pairing of PPT-DPT pairs and pooling for three neighboring locations are explained in the supplemental information.

2.2. Methods

We used three methods for estimating scaling: (a) BM, (b) QR, and (c) X. Zhang et al. (2017) method (ZM).

To apply the BM, we first considered all wet hours (with precipitation \( \geq 0.1 \) mm, except for the USA and Japan for which the minimum amount was 2.5 and 1 mm, respectively) for each station’s PPT-DPT pair. We then placed data into 12 bins of equal size, sorted from the lowest to highest DPT, and estimated the 99th percentile of PPT (P99) and the mean DPT for each temperature bin. We excluded the first and last bins from the scaling estimate to avoid influence of the specific atmospheric circulation causing very high or very low DPT values. We fitted a linear regression on the logarithm of P99 and mean DPT for the second bin to the second last bin, which is given by:

\[
\log(P99) = \alpha + \beta T
\]

(1)

Then scaling \((d(P99)/dT)\) was estimated using an exponential transformation of the regression coefficient \((\beta)\) given by:

\[
\frac{d(P99)}{dT} = 100 \left( e^\beta - 1 \right)
\]

(2)

QR is similar to BM except there is no assumption of the number and size of bins (Wasko et al., 2014). The scaling was estimated using Equations 1 and 2 for all wet hours for paired PPT-DPT data for 99th percentile using the “quantreg” package in statistical programming language R (Koenker, 2013). The important difference between BM and QR is that instead of minimizing the sum of the squared errors in linear regression in BM, in QR the absolute-deviation of the errors is minimized with a penalty term as explained in Koenker and Bassett (1978) and Wasko et al. (2014). To ensure that the increase in extreme precipitation with DPT is not dominated by seasonal trends in DPT, we also removed the seasonality from the DPT data to estimate the scaling (Ali et al., 2018; X. Zhang et al., 2017). In the ZM method, we applied the following procedure on each location. We first identified the 4 months in a year, which receive the highest precipitation. For these 4 months for each year, we estimated DPT anomalies and 1-h monthly precipitation maxima (four values for each in a given year). We next normalized the time series of 1-h monthly precipitation maxima.
by its median. Then to estimate scaling, the normalized time series of 1-h monthly precipitation maxima was fitted to a generalized extreme value distribution (GEV) with fixed scale and shape parameters (equal to 1) and location parameter changing linearly with the DPT anomalies, using the “extRemes” package in R (https://cran.r-project.org/web/packages/extRemes/index.html). For more information on the methodologies please refer to X. Zhang et al. (2017) and Ali et al. (2018). We used all wet hours in BM and QR and hourly monthly maxima in ZM thus avoiding the potential bias due to pairing different hourly PPT to the same DPT values. Since our scaling results based on ZM and BM at the hourly scale are consistent, we suggest that this effect is not strong.

We also examined the latitudinal variation in scaling using 5-degree width latitudinal bands (see supplemental information), and the difference in scaling between dry and wet regions (defined in Donat et al., 2016) and constructed scaling curves for larger regions by pooling all locations within similar climate zones, based on the Koppen-Geiger classification system. We distributed the available 7,088 gauges into 5-degree width latitudinal bands and estimated the median scaling for each band. Moreover, we grouped the available gauges into dry and wet regions following the classification of Donat et al. (2016). Donat et al. (2016) calculated precipitation indices (annual maximum precipitation, Rx1day; and total precipitation, PRCTOT) for each grid cell and normalized them by dividing by the average of the base period (1951–1980). The grid cells with the 30 per cent lowest normalized precipitation index values were labeled as dry and the 30% highest values were labeled wet, respectively.

Finally, we constructed scaling curves for the selected regions by pooling all stations with an elevation greater than and less than 400 m within a Climate Zone. The choice of the 400 m threshold is somewhat arbitrary but is motivated to avoid the mixing of DPT from different altitudes (Lenderink et al., 2021). High altitudes have lower DPT than lower altitudes and the mixing of DPT may artificially flatten the scaling curves. All the wet hour PPT-DPT pairs were placed in 2°C wide bins, ranging from 0°C up to 28°C. From this binned data we then computed the 95th, 99th, and 99.9th percentile of the distribution of the wet events (as G. Lenderink et al., 2017).

3. Results

3.1. Extreme Precipitation Scaling at Station-Level and Pooled Regions

We first examine the scaling rates at individual gauges. We find that the relationship between hourly precipitation and DPT is generally consistent with CC (6.5%/K) scaling at most locations across the selected regions. Although spatial variability in scaling is high across regions, a small majority of 60% of gauges show scaling at greater than the CC rate (Figure 1). The highest regional median scaling rates are observed for Malaysia (11.8%/K) and Australia (8.5%/K). More noteworthy is that around 10% (22%) of locations in Australia (Malaysia) show greater than 2 × CC scaling. Using other methods to estimate scaling rates gives consistent results (Figures S2 and S3). With the exception of Malaysia, the scaling rates are somewhat higher for the warm season (June to August, except for December to February for Australia) in particular for Europe and India (Figure S4).

We now move on to assess whether the record length for an individual hourly precipitation gauge may lead to any bias in scaling rate, by pooling P-DPT pairs for three neighboring gauges within 30 km distance with elevation difference no greater than 50 m (Figure S1). Note that there is a chance that the same DPT station observations are paired to two or more precipitation gauges. Notwithstanding this limitation, we observe that the pooled median scaling rates are still higher than CC for all regions (Figure S5), although slightly lower than those for individual gauges (Figure 1). These show consistently strong relationships between hourly precipitation extremes and DPT.

Accumulating the hourly precipitation data to daily totals produces scaling rates consistent with those of Ali et al. (2018) (Figure 2), mostly lower than CC or even negative for the tropics (Malaysia and India). Tropical regions show relatively low-temperature variability, e.g., all hourly precipitation extremes in Malaysia occur in a DPT range of just 20°C–26°C (Figure S7). Since BM and QR methods can be affected by seasonality, in particular, if the DPT range is small, ZM provides a better method to estimate scaling in the tropics (Ali et al., 2018) and this produces consistently higher scaling rates in Malaysia and India (Figure S6), but still consistently lower than the scaling for hourly extremes.
We now examine the latitudinal distribution of scaling from hourly precipitation using 5-degree width latitudinal bands, except for 10–20 degrees North (with only 3% of total 7,088 gauges), with the distribution of gauges in each latitudinal band shown in Figure 3a. We find that the median scaling is mostly at (or slightly above) the CC rate for all latitudes (Figure 3b), and that scaling peaks in the Tropics at over 1.5CC. We also examine the difference in overall scaling between wet and dry regions, as classified by Donat et al. (2016). It is important to note that their classification is based only on precipitation (extreme precipitation: annual maximum precipitation, Rx1day; or total precipitation, PRCTOT) and does not account for differences in temperature, humidity, etc. Based on the Rx1day extreme precipitation index and using the 30% highest/lowest gauges to define wet (extreme; blue)/dry (less extreme; red) (Donat et al., 2016), the scaling for the less extreme region (median 9.11%/K) is greater than for the more extreme region (median 7.9%/K). However, classifying based on total precipitation (PRCTOT), we find the scaling is higher for the wet region. The differences between the classification based on the two indices are mainly due to gauges in Europe (which show generally higher scaling rates) falling into different regions in each case. The results are the same when a higher threshold (40%) is used for the classification into wet and dry regions (Figure S8).

3.2. Scaling Curves for Selected Regions

We now examine the scaling relation for different climate zones by constructing scaling curves for selected regions. Scaling curves can help visualize the scaling relationship within the full DPT range, unlike scaling rates (Figures 1 and 2) which are the single values of linear slope (coefficient) across the DPT range. We first split the data into locations higher and lower than 400 m altitude (to avoid effects of differences in DPT with altitude), and then pool PPT-DPT pairs within the same climate zone, based on the Koppen-Geiger classification system (Kottek et al., 2006) (Table S2). The absolute number of stations pooled within each climate
Geophysical Research Letters

zone are presented at the top of each upper panel in Figure 4. We use the same methodology as previously to construct scaling curves for each regional climate zone (Figure 4, upper panels) and compare these to the distribution of scaling rates at individual gauges from Figure 1 (Figure 4, lower panels). For the Temperate climate zone C (common to five from six regions where hourly precipitation gauge data is available), scaling curves follow at least the CC rate in all regions (Figure 4). To summarize, scaling curves for hourly precipitation over large climate regions tend to the CC rate or above, although scaling curves are flatter for gauges at greater height (>400 m; dashed lines/darker colors in upper panel in Figure 4) compared to lower altitude gauges (<400 m; solid lines/brighter colors). Strikingly, we find that scaling curves for Europe follow 2 × CC beyond 12°C (Figure 4c) which supports the findings of Lenderink and Meijgaard (2008) for the Netherlands. What is also remarkable is that scaling curves are almost universal for the different regions; that is, for a given DPT the values of the different percentiles are very close. The scaling rates derived from these pooled data are also similar, at close to the CC rate or slightly above, with exceptions for the lowest and highest dew point temperatures. However, there are some systematic differences, such as the relatively low intensities in the low DPT range (<10°C) in Europe.

Examining the distribution of scaling rates estimated at gauge-level, most stations exceed the CC rate, and a small fraction of stations even exceed 2 × CC in all four regions for the C climate zone (Figures 4b, 4d, 4f, and 4h). The at-gauge scaling distribution within the C climate zone is similar for locations higher and lower than 400 m MSL for Australia and Japan (Figures 4d and 4f). On the contrary, there is a significant (p < 0.05) difference in the at-gauge scaling for Europe, where lower altitude gauges (<400 m MSL) show much higher scaling than gauges above 400 m MSL (Figure 4d). In tropical Malaysia (A climate zone), with a very low range in DPT (∼21°C–26°C), scaling curves follow at least the CC rate (Figure S7). Scaling curves for the US (C and D climate zones) also follow the CC rate beyond 12°C (Figure S9). Since the measurement precision for rain gauges in the USA is much coarser (2.54 mm) we did not include these results in the main figures.

Figure 2. Scaling rates (% K⁻¹) estimated using daily precipitation totals (PPT) from the GSDR data set (Lewis et al., 2019) and daily DPT from the HadISD data set (Dunn et al., 2019). The scaling is estimated using the BM at the 99th percentile for 7,088 locations, which have at least 12 years of daily precipitation data. The number in blue indicates the number of gauges (NS) in each region and the number in black indicates the median scaling (% K⁻¹) for each region. The numbers below each panel indicate the percentage of gauges within each region, which show scaling rates ranging from 0-0.5CC (green), 0.5CC-CC (yellow), CC-1.5CC (orange), 1.5CC-2CC (pink), and greater than 2CC (red) respectively, where CC is 6.5%/K.
4. Conclusions and Discussion

In this study, we have used observed hourly precipitation and daily DPT to estimate scaling over six macro-regions. At gauge level, we found that scaling rates ranged between CC and $2 \times CC$ for more than 60% of gauges. We note that comparatively lower scaling rates at US gauges as compared to Mishra et al. (2012) may be due to their coarse measurement resolution and the different scaling methods used. The scaling results for Australia are similar to those presented in C. Wasko et al. (2018). To remove spatial variability in at-gauge estimates of scaling, from short record lengths and local modes of variability, we assessed various pooling methods. Pooling data for three neighboring gauges reduced scaling rates, but still with a scaling range between CC and $2 \times CC$ for more than 50% of gauges. Moreover, the median scaling is greater than CC for gauges in wet and dry regions. When we further pooled gauges across selected regions within the same Koppen-Geiger climate zone, we found that regional scaling curves consistently follow the CC rate. The exception to this is for Europe where regional scaling for climate zone C is significantly higher than the CC rate for temperatures above 12°C, consistent with findings for the Netherlands (Lenderink & Van Meijgaard, 2008). Our results suggest that by pooling data the influence of local dynamics producing superCC behavior is averaged out, resulting in lower scaling rates from the pooled data analyses than from the individual gauges. These local dynamics produce higher superCC sensitivities in most of the gauges analyzed globally. For instance, excess latent heat released during intense short-duration rainfall may enhance scaling by intensifying upward within-cloud motions (Loriaux et al., 2013), and increases in large-scale moisture convergence producing larger storms (G. Lenderink et al., 2017; Pfahl et al., 2017) but see Fowler et al. (2021) for a comprehensive review. Our results highlight the importance of understanding the thermodynamic and dynamic processes governing precipitation extremes at different spatial scales when estimating future changes.

In summary, we have shown that the scaling of hourly extreme precipitation consistently follows at least the CC rate at a regional scale, and often a superCC rate at the gauge-level, across regions where hourly data is available. This is a much stronger scaling than that for daily extreme precipitation and adds critical information to the debate on how precipitation extremes may change in the warming climate, particularly pertinent for impacts. Of interest is whether we expect changes to be regionally or locally constrained or enhanced, and why regions like Europe show superCC scaling rates. It is an open question whether these observed scaling rates are indicative of rates of future change with warming but some evidence from high-resolution convection-permitting modeling now suggests that they may be (Lenderink et al., 2021). Of course, scaling rates may further increase (or decrease) due to other dynamical processes, including changes to large-scale circulation patterns (Pfahl et al., 2017), cloud size and the spatial extent of rainfall events (Lochbihler et al., 2017), storm type (Molnar et al., 2015), and changes to long-term moisture transport patterns (Pfahl et al., 2017). We emphasize that these factors have not been considered explicitly in our scaling approach and further studies are still needed both to understand the processes governing precipitation extremes at different temporal and spatial scales and the potential future changes to these processes. Meanwhile, the observed strong and the surprisingly universal relationship between hourly precipitation extremes and DPT has implications for the design of stormwater infrastructure systems, and perhaps provides a way of updating such design estimates.

**Author Contributions**

HA and GL carried out the analysis. HJF and HA contributed to the design of the methodology. EL and DP produced the quality controlled hourly precipitation data. All authors discussed the results and contributed to writing the paper.
Data Availability Statement

The GSDR data set will shortly be hosted by the Global Precipitation Climatology Centre at Deutsche Wetterdienst and available through the Copernicus Climate Change Service Climate Data Store. Until then, the data can be obtained from the authors. For the GSDR data sources, see Table A1 of https://doi.org/10.1175/JCLI-D-18-0143.1. HadISD DPT data is freely available at https://www.metoffice.gov.uk/hadobs/hadisd/.

Competing Financial Interests

The authors declare no competing financial interests.

Acknowledgments

This work was supported by the INTENSE project. INTENSE is supported by the European Research Council (grant ERC-2013-CoG-617329). Hayley Fowler is funded by the Wolfson Foundation and the Royal Society as a Royal Society Wolfson Research Merit Award holder (grant WM140025). We thank Markus Donat (University of New South Wales) for providing us with the shape files of wet and dry regions.

References


