

TECHNICAL  
REPORTS: METHODS

10.1029/2019EA000665

## Key Points:

- Piecewise linear quantile regression is used to quantify the transition from CC scaling to super-CC scaling
- Large uncertainties exist for the scaling rates and change points of observed precipitation extremes
- Goodness-of-fit measures indicate that the dew point temperature is a better scaling predictor than temperature

## Supporting Information:

- Supporting Information S1

## Correspondence to:

H. Van de Vyver,  
hvjiver@meteo.be

## Citation:

Van de Vyver, H., Van Schaeybroeck, B., De Troch, R., Hamdi, R., & Termonia, P. (2019). Modeling the scaling of short-duration precipitation extremes with temperature. *Earth and Space Science*, 6, 2031–2041. <https://doi.org/10.1029/2019EA000665>

Received 11 APR 2019

Accepted 24 AUG 2019

Accepted article online 29 AUG 2019

Published online 16 OCT 2019

Modeling the Scaling of Short-Duration Precipitation  
Extremes With TemperatureHans Van de Vyver<sup>1</sup> , Bert Van Schaeybroeck<sup>1</sup>, Rozemien De Troch<sup>1</sup>, Rafiq Hamdi<sup>1</sup>, and Piet Termonia<sup>1,2</sup> <sup>1</sup>Royal Meteorological Institute of Belgium, Brussels, Belgium, <sup>2</sup>Department of Physics and Astronomy, Ghent University, Ghent, Belgium

**Abstract** The Clausius-Clapeyron (CC) relation expresses the exponential increase in the moisture-holding capacity of air of approximately 7%/°C. Earlier studies show that extreme hourly precipitation increases with daily mean temperature, consistent with the CC relation. Recent studies at specific locations found that for temperatures higher than around 12 °C, hourly precipitation extremes scale at rates higher than the CC scaling, a phenomenon that is often referred to as “super-CC scaling.” These scalings are typically estimated by collecting rainfall data in temperature bins, followed by a linear fit or a visual inspection of the precipitation quantiles in each bin. In this study, a piecewise linear quantile regression model is presented for a more flexible, and robust estimation of the scaling parameters, and their associated uncertainties. Moreover, we use associated information criteria to prove statistically whether or not a pronounced super-CC scaling exists. The techniques were tested on stochastically simulated data and, when applied to hourly station data across Western Europe and Scandinavia, revealed large uncertainties in the scaling rates. Finally, goodness-of-fit measures indicated that the dew point temperature is a better scaling predictor than temperature.

## 1. Introduction

Short-duration extreme precipitation events pose a serious threat to society and economy as they may cause flash floods (Fadhel et al., 2018; Hettiarachchi et al., 2018; Pall et al., 2011). This is especially true in an urban environment due to the strong surface impermeability and dense population. For this reason, there is a large demand from city authorities to know the impact of climate change on the frequency and intensity of extreme rainfall that causes such events (Bader et al., 2018). Such knowledge could allow adequate decision making in the context of adaptation measures for urban planning or the design of hydraulic infrastructure. An intensification of extreme daily rainfall at the global scale is already observed and is expected to continue as a consequence of anthropogenic climate change (Barbero et al., 2017; Flato et al., 2013; Min et al., 2011; Seneviratne et al., 2012; Westra et al., 2013).

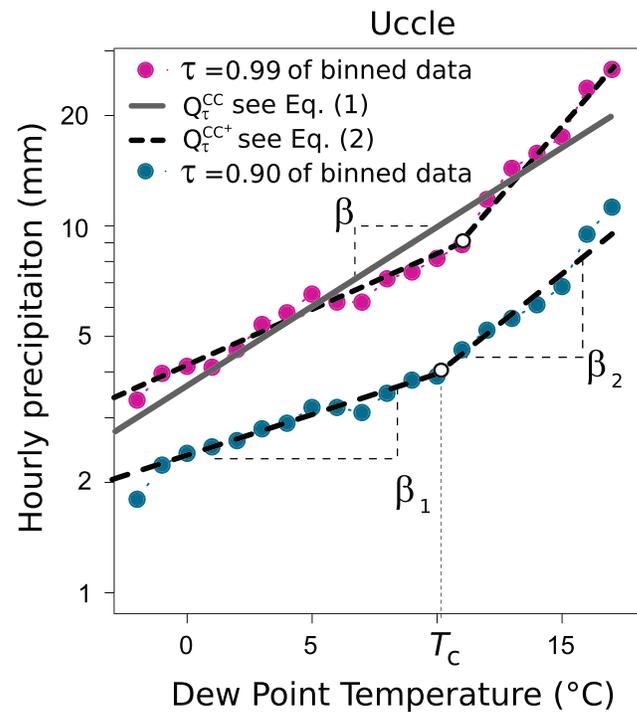
Intense precipitation during storm events is known to increase exponentially with temperature, as is traditionally explained on the basis of the Clausius-Clapeyron (CC) relationship, which states that the saturation water vapor pressure increases with temperature at a rate of 6%/°C at the surface, or 7–8%/°C in the column integral (Bao et al., 2017). The CC relation constrains the change of extreme precipitation from daily to seasonal scale (Pall et al., 2011) and is consistent with observations at midlatitudes (Utsumi et al., 2011).

However, Lenderink and van Meijgaard (2008) revealed for a single station location in the Netherlands that for temperatures above a threshold of around 12 °C, the rate of exponential growth for short-duration events exceeds the CC scaling rate. Such increased scaling is often referred to as “super-CC scaling,” and its existence has been confirmed for different locations worldwide (Lenderink et al., 2011; Mishra et al., 2012; Panthou et al., 2014; Schroeder & Kirchengast, 2018).

Where evidenced, the existence of super-CC scaling has particular societal implications, as it implies that in a warming world, the increase of short-duration precipitation extremes would exceed the currently expected CC-like growth and, consequently, would have an increased impact on human society. Westra et al. (2014) outlined the state-of-the-art research on the expected intensification of subdaily extreme rainfall due to climate change.

©2019. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.



**Figure 1.** The  $\tau$ -quantiles of hourly precipitation as a function of the daily mean dew temperature, for  $\tau=0.9$  and  $\tau=0.99$ . Dotted lines: estimated with binning. Solid gray line: estimated with linear quantile regression, equation (1). Black dashed line: estimated with piecewise linear quantile regression, equation (2).

Note that, apart from super-CC scaling, subdaily rainfall extremes also exhibit sub-CC scaling, mostly present at the highest temperature range, but sometimes also extending over a wide temperature range (Drobinski et al., 2016). Negative scaling relationships between extreme daily precipitation and (dry-bulb) temperature were found for a great number of stations in the tropics (Ali et al., 2018). These authors also argue that CC scalings is mostly consistent with respect to dew point temperature. Moreover, it was argued that in Ali et al. (2018), Barbero et al. (2018), Lenderink and van Meijgaard (2010), Roderick et al. (2019), and Wasko et al. (2018), dew point temperature is a better predictor for precipitation extremes than temperature.

To date, the physical cause of super-CC scaling behavior is still under discussion (Barbero et al., 2018; Berg et al., 2013; Hardwick-Jones et al., 2010; Loriaux et al., 2013). Potential physical drivers underlying the super-CC scaling include changes to moisture availability and atmospheric circulation and rainfall types (convective vs. stratiform). Rather than focusing on the physical interpretation, a methodology is proposed here for inference of scaling models, which also enables to discriminate the best proxies for extreme rainfall.

To investigate the scaling of extreme precipitation with (dew point) temperature, a binning approach is commonly used, where the data are divided in bins of 1 or 2°C. Next, for each bin, the high quantiles of the distribution of observed precipitation are then computed. Figure 1 shows the binned 0.90- and 0.99-quantiles of hourly precipitation against daily mean dew point temperature (dotted lines) for Uccle (Belgium). The scaling rate clearly increases for dew point temperatures above the change point  $T_c \approx 12$  °C, as was reported in Lenderink and van Meijgaard (2010). Bins with dew point temperatures above 19–20 °C include very few data points but show a downturn of the percentiles. This downturn is stronger upon use of the subdaily dew point temperature and might be attributed to moisture limitations under warmer anticyclonic conditions. Hence, one must be cautious not to extrapolate the super-CC scaling to the highest temperature regimes, for instance, in the context of a warming climate, as this may severely overestimate precipitation extremes (Prein et al., 2017; Zhang et al., 2017).

Wasko and Sharma (2014) have shown that, in case of a constant scaling across a wide temperature range, the use of linear quantile regression (Koenker, 2005) is superior to the binning approach for extracting the scaling properties. In particular, the quantile regression estimator is asymptotically unbiased (Koenker &

Basset, 1978), in contrast to the binning approach. Moreover, a proper statistical framework is necessary given the lack of long and reliable subdaily time series (Li et al., 2019; Westra et al., 2014).

For locations exhibiting super-CC scaling, applying an additional piecewise linear quantile regression to both ranges ( $T \leq T_c$  and  $T_c < T \leq 19$  °C) turns out to be problematic: first, if the change point is not known in advance, the regression lines may show a discontinuity at the change point. Second, linear quantile regression provides uncertainty estimates of the scaling rates, but the uncertainty in the change point cannot be obtained.

In this paper, we apply the piecewise linear quantile regression framework of Li et al. (2011) by simultaneously estimating the scaling rates and the change point. In Figure 1, the quantiles (dashed lines) provided by the change point model of Li et al. (2011) have been added, which is made up of two different lines with slopes  $\beta_1$  and  $\beta_2$  instead of a single slope  $\beta$ . The work of Wasko and Sharma (2014) is extended here in the sense that we model two scaling regimes and, in addition, propose a more complete inference, including uncertainty estimation, model selection with information criteria, and predictor selection with goodness-of-fit measures.

## 2. Methodology

### 2.1. Quantile Regression With a Change Point Model

In the present study, nonzero precipitation is denoted by  $P$  and, unless mentioned otherwise,  $T$  corresponds to the daily mean dew point temperature. Below, different models are introduced for  $Q_\tau(T)$ , which is the  $\tau$ -quantile of the conditional distribution of  $\log(P)$ , given a certain  $T$ .

*The CC model.* The scaling of high quantiles of  $\log(P)$  with the predictor variable,  $T$ , is calculated using a linear model (Wasko & Sharma, 2014):

$$Q_\tau^{CC}(T) = \alpha + \beta T, \quad (1)$$

where  $\beta$  is the slope which is used to estimate the CC scaling rate.

*The CC<sup>+</sup> model.* The “change point model” uses piecewise linear quantile regression (Li et al., 2011):

$$Q_\tau^{CC^+}(T) = \begin{cases} \alpha_1 + \beta_1 T & \text{for } T \leq T_c, \\ \alpha_2 + \beta_2 T & \text{for } T > T_c. \end{cases} \text{ with change point } T_c, \quad (2)$$

where  $\beta_1$  is supposed to correspond with the CC scaling rate and  $\beta_2$  may correspond to the scaling rate of super-CC ( $>7\%/^\circ\text{C}$ ), sub-CC ( $<7\%/^\circ\text{C}$ , and  $>0\%/^\circ\text{C}$ ), or even negative scaling ( $<0\%/^\circ\text{C}$ ), but the latter case is rather exceptional (Ali et al., 2018). Note that given the continuity of the quantile regression lines at the change point  $T_c$ , the CC<sup>+</sup> model has effectively four unknown parameters. We choose the free parameters,  $\alpha_1, \beta_1, \beta_2$ , and  $T_c$ , and use the following relation:  $\alpha_2 = \alpha_1 + (\beta_1 - \beta_2) T_c$ .

Although the change point model, equation (2), might be a valuable tool for quantifying changes in the extremes, other quantile regression functions may fit the data as well. Obvious other choices, including a quadratic model, a piecewise linear-quadratic model, and a piecewise linear-(powered) exponential model, were studied, but none of them were found to be systematically superior to the linear or piecewise linear function (see Text S1 in the supporting information).

The extrapolation to higher  $\tau$ -quantiles ( $\tau > 0.99$ ) is possible by applying the peak-over-nonstationary threshold model (Beirlant et al., 2004; Coles, 2001). However, this requires a number of additional assumptions, especially concerning the dependence of the change point on  $\tau$ .

### 2.2. Inference for Quantile Regression Models

For a given set of  $n$  data pairs  $(T_i, P_i)$ , the parameter estimates of the quantile regression models are obtained as the solution of the minimization problem (Koenker, 2005):

$$\text{CC model : } \hat{\mathcal{S}}^{\text{CC}} = \min_{(\alpha, \beta)} \sum_{i=1}^n \rho_{\tau}(\log(P_i) - Q_{\tau}^{\text{CC}}(T_i)) \quad , \quad (3)$$

$$\text{CC}^+ \text{ model : } \hat{\mathcal{S}}^{\text{CC}^+} = \min_{(\alpha_1, \beta_1, \beta_2, T_c)} \sum_{i=1}^n \rho_{\tau}(\log(P_i) - Q_{\tau}^{\text{CC}^+}(T_i)) \quad ,$$

with  $\rho_{\tau}(u) = u(\tau - 1_{\{u < 0\}})$ , the loss function. Estimation of the CC model is straightforward and can be performed using the R-package “quantreg” (Koenker, 2018; Wasko & Sharma, 2014). Efficient estimation of the CC<sup>+</sup> model can be achieved by minimizing over a range of possible  $T_c$  values and then selecting the value of  $T_c$  at which the minimum is reached.

Among the different models (linear, piecewise linear, or others), the best model is found by means of the Bayesian information criterion (BIC), which assesses goodness-of-fit but avoids overfitting by penalizing additional degrees of freedom. Let  $\hat{L}$  be the maximum value of the likelihood function  $\mathcal{L}$  for a statistical model, then

$$\text{BIC} = -2 \log \hat{\mathcal{L}} + \ln(n)p, \quad (4)$$

with  $p$  the number of estimated parameters and  $n$  the number of data points (for the CC model,  $p=2$ , while  $p=4$  for the CC<sup>+</sup> model). The model with the lowest BIC value is chosen. The approach of Yu and Moyeed (2001) is followed by considering a likelihood function based on an asymmetric Laplace distribution for the error term. Maximization of this likelihood function is equivalent to the minimization of equation (3).

In order to examine the influence of the predictor  $T$  to the precipitation data, one quantifies the relative success of a quantile regression model against the unconditional quantiles,  $Q_{\tau}^{(0)}$ , of the climatological distribution. These are obtained by minimizing

$$\hat{\mathcal{S}}_0 = \min_{Q_{\tau}^{(0)}} \sum_{i=1}^n \rho_{\tau}(\log(P_i) - Q_{\tau}^{(0)}). \quad (5)$$

Koenker and Machado (1999) defined the goodness-of-fit criterion for a particular quantile as follows:

$$R = 1 - \hat{\mathcal{S}} / \hat{\mathcal{S}}_0, \quad (6)$$

where  $\hat{\mathcal{S}}$  stands for either  $\hat{\mathcal{S}}^{\text{CC}}$  or  $\hat{\mathcal{S}}^{\text{CC}^+}$ .  $R$  obtains values between 0 and 1 and the closer  $R$  is to 1, the better the quantile regression model (CC or CC<sup>+</sup>) for a certain  $\tau$ -quantile.

### 3. Simulation Study

Prior to introducing observational data, stochastic simulations were used to gain insight into the performance of the proposed estimator. On the one hand, we assessed the ability of BIC to uncover the underlying model, and on the other hand, we examined two factors that may impact the inference: (i) the ratio of the scaling parameters,  $\beta_2/\beta_1$ , and (ii) the position of the change point  $T_c$  in the support of the distribution of  $T$ .

#### 3.1. Setup

We have drawn bivariate pairs  $(T_i, \log(P_i))$  from the following random process:

$$\log(P_i) = g(T_i) + \varepsilon_i, \quad \text{with } T_i \sim \mathcal{N}(0, 1), \quad (7)$$

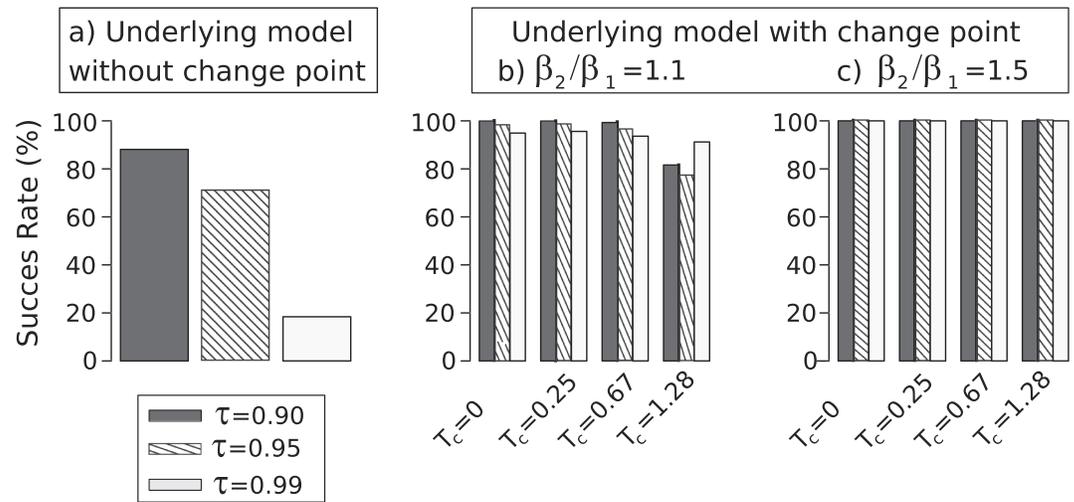
with independent and identically distributed errors  $\varepsilon_i$ . We assumed that  $\varepsilon_i$  follows the Laplace distribution,

$$F(x) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right), \quad (8)$$

where we put  $\sigma=0.014$ . To mimic the different types of scaling behavior, we assumed that the function  $g(T)$  is either linear (CC) or continuous piecewise linear (CC<sup>+</sup>) with change point  $T_c$ :

$$g(T_i) = \begin{cases} \alpha + \beta T_i, & \text{(CC simulation),} \\ \mathbf{1}_{\{T_i \leq T_c\}}(\alpha_1 + \beta_1 T_i) + \mathbf{1}_{\{T_i > T_c\}}(\alpha_2 + \beta_2 T_i), & \text{(CC}^+ \text{ simulation),} \end{cases} \quad (9)$$

with  $\beta=\beta_1=0.07$ , the well-known CC scaling rate, and for simplicity, we took  $\alpha=\alpha_1=0$ . In the CC<sup>+</sup>



**Figure 2.** Success rate of Bayesian information criterion-based model selection, based on 1,000 simulations. For each simulation, data were drawn from the random process, equations (7) and (8). (a) Linear model. (b and c) Piecewise linear model, where the “true” parameter values are indicated in the title ( $>\beta_2/\beta_1$ ), and the x axis ( $T_c$ ).

simulation, two free parameters,  $\beta_2$  and  $T_c$  remain and they should be chosen in accordance with the observations. As we have found by fitting on observed precipitation extremes (see section 4 for data description), the normalized change point ranges between 0.5 and 0.9 for different station locations. Consequently, we chose  $\beta_2=1.1, \dots, 2$ , and for  $T_c$  we chose the 0.50, 0.60, 0.75, and 0.90-quantiles of the distribution of  $T$  (i.e., standard normal). For each model, bivariate time series with length  $n=10^4$  were generated, which corresponds more or less to a realistic number of observational events spanning a period of around 50–60 years.

### 3.2. Simulation Results

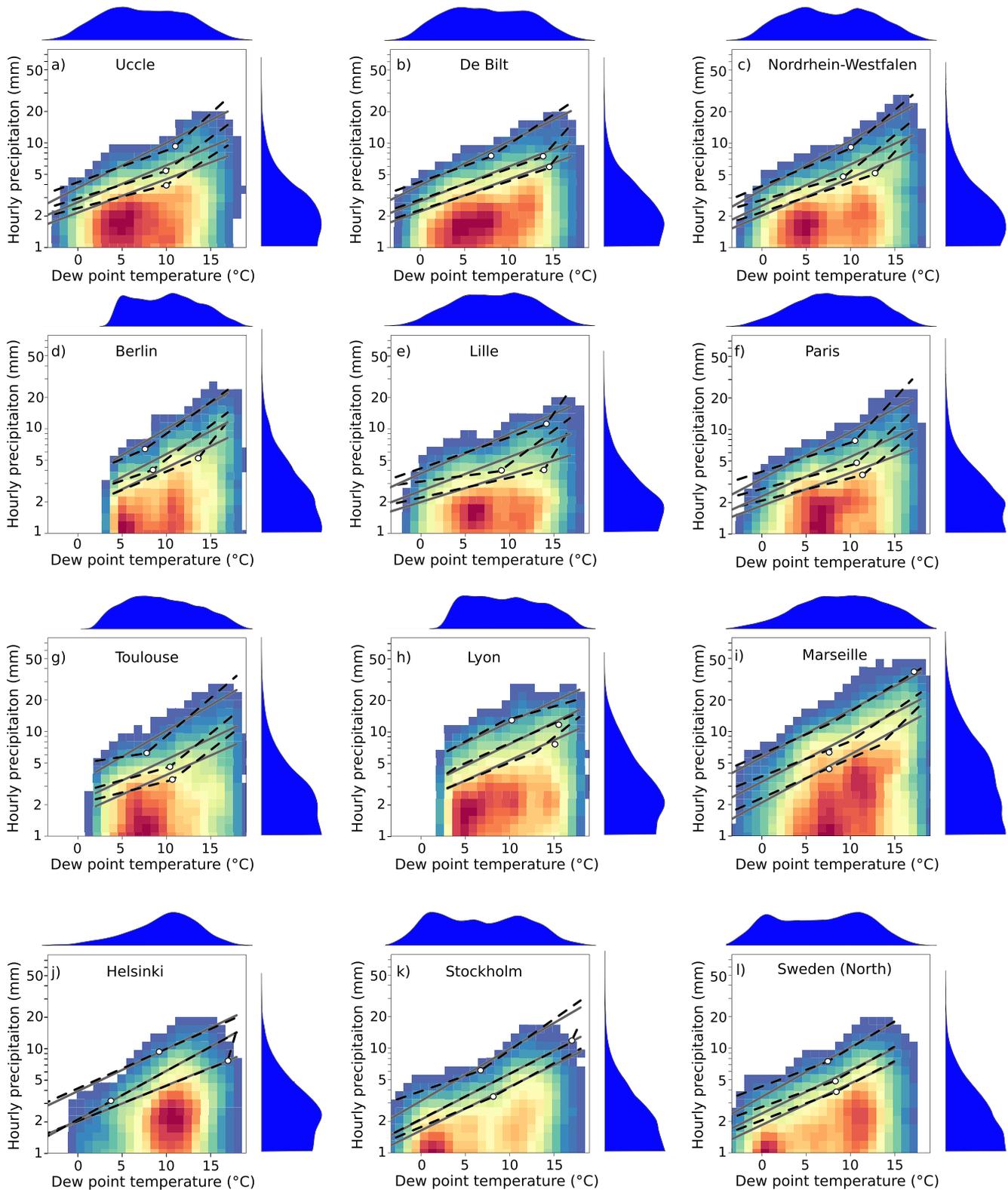
For each of  $10^3$  simulation replicates, both the CC and CC<sup>+</sup> models were fitted and the corresponding BIC values were compared. The model with the lowest BIC value should be the underlying model. Figure 2 shows the percentage of cases that the correct model was selected. For  $\tau=0.9$ , BIC correctly identified the CC model in 90% of the cases (see Figure 2a). However, the success rate strongly decreases for higher values of  $\tau$ . This result may agree with Marshall et al. (2005) where it was emphasized that BIC model selection may not be consistent for smaller data sets (higher  $\tau$  values imply less data above the regression lines). From Figures 2b and 2c, it can be seen that the success rate to correctly identify the underlying CC<sup>+</sup> model is the lowest (around 80%) for  $\beta_2/\beta_1=1.1$ , and effectively 100% for  $\beta_2/\beta_1=1.5$ . Another commonly used information criterion, the Akaike Information Criterion, was outperformed by BIC (not shown).

Finally, other random processes, including different  $\sigma$  values, or assuming a normal distribution for  $\epsilon_i$ , were found to yield similar results.

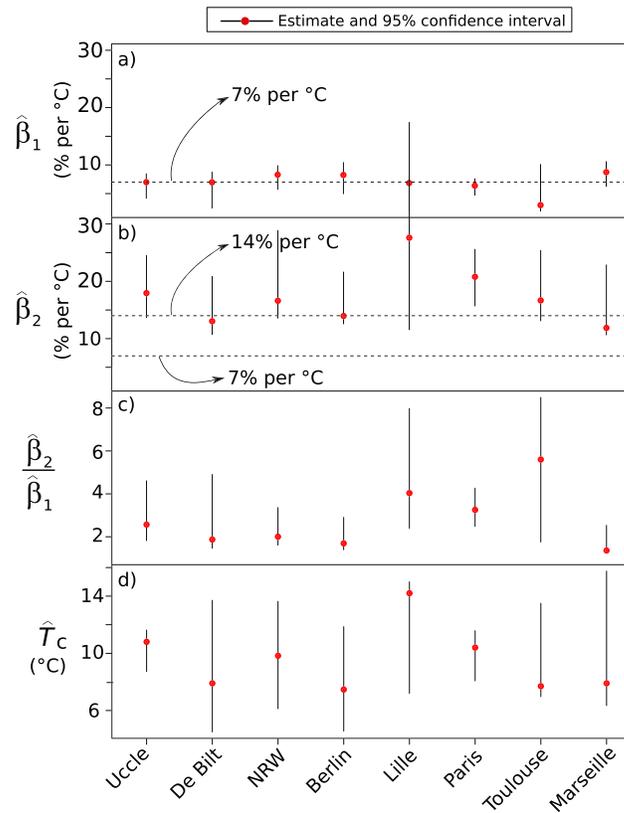
In sum, BIC-based model selection was seen to have reasonable success rates for  $\tau \leq 0.9$ . For the observations, therefore, we further base our model selection on the use of BIC at the 0.9-quantile, but one may rightfully question whether this is representative for more extreme precipitation. We will return to this in section 5.

Figure S1 shows the mean and the 2.5th and 97.5th percentiles of the estimates of  $\beta_2/\beta_1$  and  $T_c$  from the  $10^3$  replicates, for the  $\tau=0.99$  quantile (which is more representative for extremes than  $\tau=0.9$  or  $\tau=0.95$ ). For comparison, the true values of the parameters were added to the plots as blue dots. The following conclusions can be drawn:

- For the case  $\beta_2/\beta_1=1.1$  (Figures S1a and S1d), the uncertainty in the estimation of the change point  $T_c$  is very large. A large value range for change point estimates, in turn, gives rise to a wide range of estimates of the scaling factors. In addition, it can be seen that the estimator is biased in such a case. Yet, it was pointed



**Figure 3.** The  $\tau$ -quantile estimates of hourly precipitation (mm), with  $\tau=0.90, 0.95, 0.99$ . Solid lines: linear quantile regression lines (CC). Dashed lines: piecewise linear quantiles regression lines (CC<sup>+</sup>). The shaded areas represent the two-dimensional histogram. The probability distribution of  $\log P$  (right bar) is truncated at  $P=1$ . (a–l) Different locations.



**Figure 4.** Inference results for the  $CC^+$  model, equation (2), for different locations, and  $\tau=0.99$ . (NRW: Nordrhein-Westfalen): (a)  $\hat{\beta}_1$ , (b)  $\hat{\beta}_2$ , (c)  $\hat{\beta}_2/\hat{\beta}_1$ , and (d)  $\hat{T}_c$ .

out in Li et al. (2011) that the estimator is asymptotically unbiased. Clearly, convergence is not achieved for the present sample size ( $n=10^4$ ). Increasing the ratio  $\beta_2/\beta_1$  strongly reduces the uncertainty and removes the bias to a certain extent.

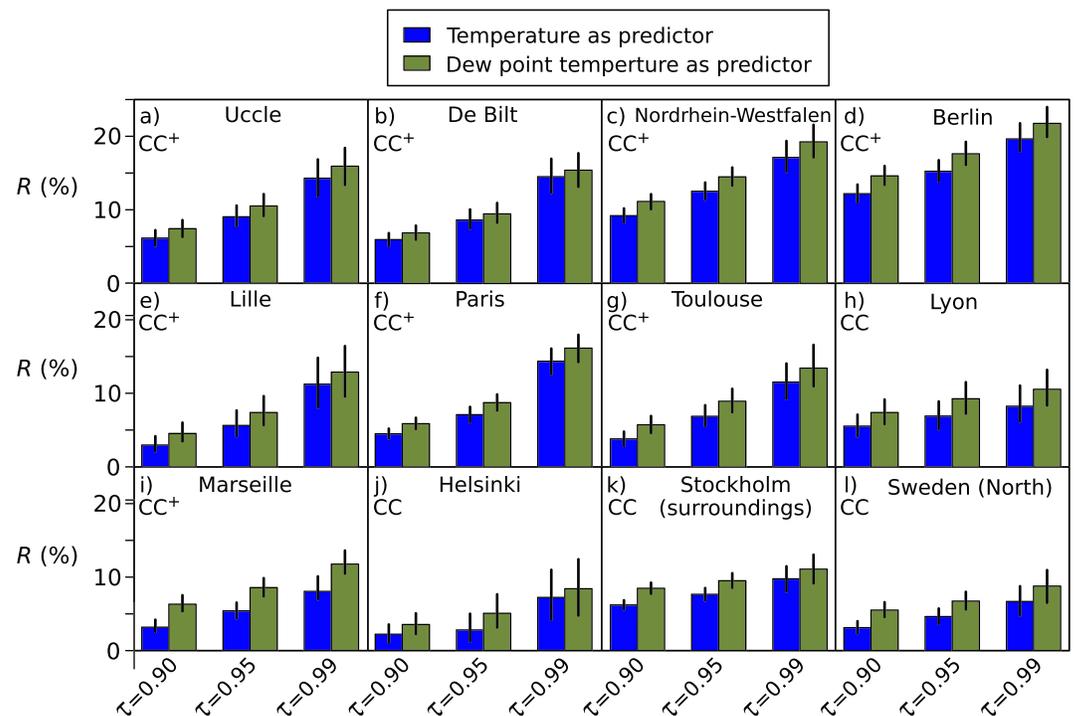
- The estimation uncertainty increases when the change point moves to the upper range of the distribution of  $T$ , where less data are available (Figures S1b, S1c, S1e, and S1f).

#### 4. Data

Long-time quality-controlled series of hourly observed precipitation, temperature, and dew point temperature were collected for different locations in Western Europe and Scandinavia (see Table S1). These include Belgium (Uccle), The Netherlands (De Bilt), France (Paris, Lille, Toulouse, Lyon, and Marseille), Germany (Nordrhein-Westfalen and Berlin), Sweden (Stockholm surrounding area, and Northern Sweden), and Finland (Helsinki; see Kilpeläinen et al., 2008). In the statistical analysis, neighboring station data were treated as one single data set (see Table S1), as is commonly done to improve the estimation of extremes (Buishand, 1991; Davison et al., 2012; Hosking & Wallis, 1997).

As in the original approach of Lenderink and van Meijgaard (2010), we computed the daily mean dew point temperature. Data points were excluded (i) when precipitation observations are equal to, or less than, 0.1 mm; (ii) when hourly instantaneous temperatures are below 0 °C (to avoid snow); (iii) for events associated with the downturn in precipitation extremes (mostly dew point temperatures above 18–20 °C), as this is likely due to a lack of moisture content; and (iv) when the daily mean temperature exceeds 24 °C because Hardwick-Jones et al. (2010) found a reduction in relative humidity in such a case, which may affect the scaling relationship. As in Wasko and Sharma (2014), rainfall events were separated by 5 hr of no precipitation, and we withheld the maximum precipitation depth within each event.

Note that the maximum hourly precipitation depth underestimates the peak 1-hr precipitation because the latter likely occurs across hourly measurements.



**Figure 5.** The goodness-of-fit criterion  $R$ , equation (6). For each station, the Bayesian information criterion-selected model was considered (either CC or CC<sup>+</sup>). The vertical lines indicate the 95% confidence intervals of  $R$ . (a–l) Different locations.

## 5. Results

The application of the quantile regression models to the observational time series is demonstrated in Table S2. The corresponding quantile lines, for different  $\tau$  values, are plotted in Figure 3. Based on the BIC values for the 0.9-quantile (Table S2), the CC<sup>+</sup> model is better than the CC model for Uccle, De Bilt, Nordrhein-Westfalen, Berlin, Lille, Paris, Toulouse, and Marseille. For the remaining locations (e.g., Lyon, Stockholm surrounding area, Northern Sweden, and Helsinki), the lowest BIC values at the 0.9-quantile were obtained by the CC model.

In Helsinki, it appears that the optimization of the CC<sup>+</sup> model is problematic, as the estimated change points  $T_c$  are included at the right end of the range of possible  $T_c$  values. For Stockholm surrounding area, the same is true to a certain extent at the 0.95-quantile. Based on the BIC analysis, a change point is clearly apparent for the 0.95 and 0.99-quantiles, opposed to the case  $\tau=0.90$ . This is in line with section 3.2, where it was concluded that BIC-based model selection can become unreliable for higher quantiles, especially when the underlying model is linear (see Figure 2a). For Northern Sweden, the change in the scaling rates is too low to be considered as significant, as the BIC values show. The results for Lyon are anomalous in the sense that it is the only location with *sub*-CC scaling for  $\tau=0.99$ , but the BIC analysis might be, again, unreliable for such high  $\tau$  values.

Figure S2 shows the parameter estimates for Uccle, together with the 95% confidence intervals, obtained by the bootstrap method. As expected, the uncertainty in  $\beta_2$  is much higher than that of  $\beta_1$ . The confidence intervals of the change in the slope,  $\beta_2/\beta_1$ , range approximately between 2 and 4 (for  $\tau=0.90, 0.95$ ), and between 2 and 5 (for  $\tau=0.99$ ). The confidence intervals of  $T_c$  are large and extend up to about 7 °C. The lower bounds of the confidence intervals of  $\beta_2/\beta_1$  are larger than 1, indicating that the CC<sup>+</sup> model is significant, even for the 0.99-quantile. Although, we had less confidence in the BIC analysis for higher  $\tau$  values, the foregoing lends additional support to the CC<sup>+</sup> model selection.

In the same way, Figure 4 shows the inference results for all the locations where the CC<sup>+</sup> model is significant, for the particular choice  $\tau=0.99$ . The confidence intervals of  $\beta_1$  (Figure 4a) cover the range 5–10%/°C and are thus compatible with the well-known CC rate of 7%/°C. The estimation of the super-CC scaling

rate in De Bilt (around 14%/°C) agrees well with the results shown in Figure 2c of Lenderink and van Meijgaard (2010), giving extra confidence to our results. Most likely, the scaling rates seem to change by a factor of more than 2 (Figure 4c), although, due to the large estimation uncertainties, assessing potential regional differences in the scaling rates is difficult. Confidence intervals for Lille and Marseille are particularly large, but the estimation was based on less than 1,000 data pairs. As in Figure S2c, the lower bounds of the confidence intervals of  $\beta_2/\beta_1$  are larger than 1.

Inference results for locations without change point (Scandinavian stations and Lyon) indicate that the confidence intervals are also large and the bootstrap distributions regularly appear to be very skewed (not shown). Furthermore, the lower bounds of the confidence intervals of  $\beta_2/\beta_1$  are smaller than 1. Additionally to the BIC analysis, this provides extra support for the conservative CC model selection. Overall, it is therefore concluded that the existence of a change point could not be demonstrated for the Scandinavian stations and for Lyon.

A BIC-based model selection was done between the CC and CC<sup>+</sup> models and a few alternative models, as described in Text S1. For the 0.9-quantile, the CC<sup>+</sup> model is the most suitable model for Uccle, De Bilt, Nordrhein-Westfalen, Toulouse, and Marseille. The piecewise linear-quadratic model was selected in Paris, Lille, and Berlin, although the difference in the BIC values of both quantile models are very small. As expected, the CC model is the most suitable for the Scandinavian stations and Lyon. The quantile regression lines of the alternative models were displayed in Figure S3 (Uccle). As the differences with the CC<sup>+</sup> model are rather small, it can be concluded that the CC<sup>+</sup> model is sufficiently good for practical purposes.

Finally, the scaling was tested for different predictors by means of the goodness-of-fit criterion, equation (6). As potential candidate predictors, the temperature and the dew point temperature were compared (Figure 5). The predictive skill of the dew point temperature is slightly, but systematically higher than that of the temperature, which is physically plausible. Note also that, irrespective of the predictor, the predictive skill at locations with a change point is significantly higher than at locations with no change point (Scandinavian stations and Lyon).

## 6. Conclusions and Outlook

A statistical methodology was outlined to study the scaling properties of subdaily extreme precipitation as a function of (dew point) temperature. More specifically, piecewise linear quantile regression (i.e., the change point model) was used to model the transition of CC-like scaling to a regime with different scaling. Compared to the binning approach, which involves a visual check only, the biggest advantages of quantile regression are (i) asymptotic unbiasedness (Li et al., 2011); (ii) the ability to estimate the scaling rates and the change point, together with the associated uncertainties; (iii) the use of information criteria for selecting the most suitable statistical model from a set of candidate models; and (iv) the ability to select the best predictor to characterize the scaling of extreme precipitation, using goodness-of-fit measures. More specifically, the following were found:

1. Simulations with simple stochastic models showed that, for a realistic sample size of  $n=10^4$ , the estimator is fairly unbiased and has a reasonable uncertainty, unless (i) the scaling rates differ only slightly and, to a lesser extent, (ii) the change point temperature is at the upper percentiles of its distribution.
2. Simulations with simple stochastic models showed that BIC-based inference is useful in detecting the existence of a change point. However, when there is no change point, the success rate at the 0.9-quantile is acceptable but decreases at increasing quantiles.
3. The results show a strong evidence for the change point model in Western Europe. Results at Marseille suggest also a change point, but the change in the scaling rates is smaller than for Western Europe. On the contrary, evidence lacks for the change point model in the Scandinavian stations and in Lyon.
4. Although deviations from linear scaling are evidenced at multiple locations, the associated estimations for change points and scaling rates are highly uncertain. More specifically, the factorial change in the scaling coefficients ranges between 2 and 5, while the change point estimates ranges between 5 and 15 °C.
5. In view of the recent controversy regarding using air temperature/dew point temperature as proxies for extreme precipitation, an approach is presented to discriminate the best predictor. More specifically, at

all observational locations, dew point temperature is slightly superior to temperature as a predictor for extreme precipitation. Moreover, locations with a change point show larger overall explanatory skill than locations without a change point.

Our analysis highlighted that statistically proving deviations from linear scaling is not trivial. Even for time series of more than 60 years, the values for the scaling rates and change points remain highly uncertain.

The study on subdaily rainfall extremes could be extended by using alternative rainfall data sets, that is, other than rain gauge measurements, but they may possibly entail other difficulties. For instance, radar data mostly feature high temporal frequencies and cover a wide area, but the available time period rarely exceeds one decade (Berg et al., 2013; Goudenhoofd & Delobbe, 2016; Overeem et al., 2009). In addition, this requires a proper declustering method for spatial precipitation extremes.

The methodology at hand could serve to study the impact of climate change on the scaling behavior of extreme short-duration precipitation and to estimate the reliability of these results. The most reliable modeled rainfall extremes are the ones from regional climate models at convection-permitting scales; however, these are scarce (Kendon et al., 2014; Tabari et al., 2016; Termonia et al., 2018). The results derived from this work may also serve for in-depth regional studies (Schroerer & Kirchengast, 2018), or the simulation of extreme precipitation events in a warmer climate (Hazeleger et al., 2015; Manola et al., 2018).

Different extensions of this work could be envisioned. The quantile regression model could be extended to multiple predictors (Li et al., 2011; Wasko & Sharma, 2014) including, for example, time-delayed (dew point) temperatures (Bao et al., 2017), and the predictive skill could be compared as a function of the rainfall duration (Berg et al., 2013; Hardwick-Jones et al., 2010; Lenderink & van Meijgaard, 2010; Utsumi et al., 2011).

#### Acknowledgments

We acknowledge Antti Mäkelä (FMI) and Valéry Masson (Météo-France) for providing observed rainfall time series. Peter Berg (SMHI) and Jules Beersma (KNMI) are acknowledged for their help with obtaining their national data. We acknowledge the DWD, KNMI, SMHI, FMI, and Météo-France for their observation data. The Swedish, Dutch, and German data are available through [opendata-download-metobs.smhi.se/explore/](https://opendata-download-metobs.smhi.se/explore/), [projects.knmi.nl/klimatologie/uurgegevens](https://projects.knmi.nl/klimatologie/uurgegevens), and [opendata.dwd.de](https://opendata.dwd.de) websites, respectively. This work is supported by URCLIM and INDECIS that have received funding from EU's H2020 Research and Innovation Program under Grant Agreement 690462.

#### References

- Ali, H., Fowler, H. J., & Mishra, V. (2018). Global observational evidence of strong linkage between dew point temperature and precipitation extremes. *Geophysical Research Letters*, *45*, 12,320–12,330. <https://doi.org/10.1029/2018gl080557>
- Bader, D. A., Blake, R. A., Grimm, A., Hamdi, R., Kim, Y., Horton, R. M., et al. (2018). Urban climate science. In C. Rosenzweig, W. Solecki, P. Romero-Lankao, S. Mehrotra, S. Dhakal, & S. Ali Ibrahim (Eds.), *Climate change and cities: Second assessment report of the urban climate change research network (ARC3.2)* (pp. 27–60). Cambridge: Cambridge University Press.
- Bao, J., Sherwood, S. C., Alexander, L. V., & Evans, J. P. (2017). Future increases in extreme precipitation exceed observed scaling rates. *Nature Climate Change*, *7*(2), 128–132.
- Barbero, R., Fowler, H. J., Lenderink, G., & Blenkinsop, G. (2017). Is the intensification of precipitation extremes with global warming better detected at hourly than daily resolutions? *Geophysical Research Letters*, *44*, 974–983. <https://doi.org/10.1002/2016GL071917>
- Barbero, R., Westra, S., Lenderink, G., & Fowler, H. J. (2018). Temperature-extreme precipitation scaling: A two-way causality? *International Journal of Climatology*, *38*(S1), e1274–e1279.
- Beirlant, J., Goegebeur, Y., Segers, J., & Teugels, J. (2004). *Statistics of extremes*. Chichester, England: John Wiley & Sons, Ltd.
- Berg, P., Moseley, C., & Haerter, J. O. (2013). Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience*, *6*(3), 181–185.
- Buishand, T. A. (1991). Extreme rainfall estimation by combining data from several sites. *Hydrological Sciences Journal*, *36*(4), 345–365.
- Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London Ltd: Springer-Verlag.
- Davison, A. C., Padoan, S. A., & Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statistical Science*, *27*(2), 161–186.
- Drobinski, P., Alonzo, B., Bastin, S., Silva, N. D., & Muller, C. (2016). Scaling of precipitation extremes with temperature in the French Mediterranean region: What explains the hook shape? *Journal of Geophysical Research: Atmospheres*, *121*, 3100–3119. <https://doi.org/10.1002/2015jd023497>
- Fadhel, S., Rico-Ramirez, M. A., & Hana, D. (2018). Sensitivity of peak flow to the change of rainfall temporal pattern due to warmer climate. *Journal of Hydrology*, *560*, 546–559.
- Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., et al. (2013). Evaluation of climate models. In Stocker, T. F., et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 741–866). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press.
- Goudenhoofd, E., & Delobbe, L. (2016). Generation and verification of rainfall estimates from 10-yr volumetric weather radar measurements. *Journal of Hydrometeorology*, *17*(4), 1223–1242.
- Hardwick-Jones, R., Westra, S., & Sharma, A. (2010). Observed relationships between extreme sub-daily precipitation, surface temperature, and relative humidity. *Geophysical Research Letters*, *37*, L22805. <https://doi.org/10.1029/2010gl045081>
- Hazeleger, W., van den Hurk, B. J. J. M., Min, E., van Oldenborgh, G. J., Petersen, A. C., Stainforth, D. A., et al. (2015). Tales of future weather. *Nature Climate Change*, *5*, 107–113.
- Hettiarachchi, S., Wasko, C., & Sharma, A. (2018). Increase in flood risk resulting from climate change in a developed urban watershed—The role of storm temporal patterns. *Hydrology and Earth System Sciences*, *22*, 2041–2056.
- Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis: An approach based on L-moments*. Cambridge, UK: Cambridge University Press.
- Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., & Senior, C. A. (2014). Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change*, *4*(7), 570–576.
- Kilpeläinen, T., Tuomenvirta, H., & Jylhä, K. (2008). Climatological characteristics of summer precipitation in Helsinki during the period 1951–2000. *Boreal Environment Research*, *13*, 67–80.

- Koenker, R. (2005). *Quantile regression*. Cambridge: Cambridge University Press.
- Koenker, R. (2018). Quantreg: Quantile regression.
- Koenker, R., & Basset, G. (1978). Regression quantiles. *Econometrica*, *46*(1), 33–55.
- Koenker, R., & Machado, J. (1999). Goodness of fit and related inference processes for quantile regression. *Journal of the American Statistical Association*, *94*(448), 1296–1310.
- Lenderink, G., Mok, H. Y., Lee, T. C., & van Oldenborgh, G. J. (2011). Scaling and trends of hourly precipitation extremes in two different climate zones—Hong Kong and the Netherlands. *Hydrology and Earth System Sciences*, *15*(9), 3033–3041.
- Lenderink, G., & van Meijgaard, E. (2008). Increase in hourly extreme precipitation beyond expectations from temperature changes. *Nature Geoscience*, *1*, 511–514.
- Lenderink, G., & van Meijgaard, E. (2010). Linking increases in hourly precipitation extremes to atmospheric temperature and moisture changes. *Environmental Research Letters*, *5*, 025208.
- Li, C., Wei, Y., Chappell, R., & He, X. (2011). Bent line quantile regression with application to an allometric study of land mammals' speed and mass. *Biometrics*, *67*(1), 242–249.
- Li, C., Zwiers, F., Zhang, X., & Li, G. (2019). How much information is required to well constrain local estimates of future precipitation extremes? *Earth's Future*, *7*, 11–24. <https://doi.org/10.1029/2018EF001001>
- Loriaux, J. M., Lenderink, G., De Roode, S. R., & Siebesma, AP. (2013). Understanding convective extreme precipitation scaling using observations and an entraining plume model. *Journal of the Atmospheric Sciences*, *70*(11), 3641–3655.
- Manola, I., van den Hurk, B., De Moel, Hans, & Aerts, J. C. (2018). Future extreme precipitation intensities based on a historic event. *Hydrology and Earth System Sciences*, *22*(7), 3777–3788.
- Marshall, L., Nott, D., & Sharma, A. (2005). Hydrological model selection: A Bayesian alternative. *Water Resources Research*, *41*, W10422. <https://doi.org/10.1029/2004wr003719>
- Min, S.-K., Zhang, X., Zwiers, F. W., & Hegerl, G. C. (2011). Human contribution to more-intense precipitation extremes. *Nature*, *470*(7334), 378–381.
- Mishra, V., Wallace, J. M., & Lettenmaier, D. P. (2012). Relationship between hourly extreme precipitation and local air temperature in the United States. *Geophysical Research Letters*, *39*, L16403. <https://doi.org/10.1029/2012gl052790>
- Overeem, A., Holleman, I., & Buishand, A. (2009). Derivation of a 10-year radar-based climatology of rainfall. *Journal of Applied Meteorology and Climatology*, *48*(7), 1448–1463.
- Pall, P., Aina, T., Stone, D. A., Stott, P. A., Nozawa, T., Hilberts, A. G., et al. (2011). Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature*, *470*(7334), 382–385.
- Panthou, G., Mailhot, A., Laurence, E., & Talbot, G. (2014). Relationship between surface temperature and extreme rainfalls: A multi-time-scale and event-based analysis. *Journal of Hydrometeorology*, *15*(5), 1999–2011.
- Prein, A. F., Rasmussen, R. M., Ikeda, K., Liu, C., Clark, M. P., & Holland, G. J. (2017). The future intensification of hourly precipitation extremes. *Nature Climate Change*, *7*, 48–52.
- Roderick, T., Wasko, C., & Sharma, A. (2019). Atmospheric moisture measurements explain increases in tropical rainfall extremes. *Geophysical Research Letters*, *46*, 1375–1382. <https://doi.org/10.1029/2018gl080833>
- Schroeder, K., & Kirchengast, G. (2018). Sensitivity of extreme precipitation to temperature: The variability of scaling factors from a regional to local perspective. *Climate Dynamics*, *50*(11–12), 3981–3994.
- Seneviratne, S. I., Nicholls, N., Easterling, D., Goodess, C. M., Kanae, S., Kossin, J., et al. (2012). Changes in climate extremes and their impacts on the natural physical environment. In C. B. Field, et al. (Eds.), *Managing the risks of extreme events and disasters to advance climate change adaptation. A special report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)* (pp. 109–230). Cambridge, UK and New York, NY, USA: Cambridge University Press.
- Tabari, H., De Troch, R., Giot, O., Hamdi, R., Termonia, P., Saeed, S., et al. (2016). Local impact analysis of climate change on precipitation extremes: Are high-resolution climate models needed for realistic simulations? *Hydrology and Earth System Sciences*, *20*(9), 3843–3857.
- Termonia, P., Van Schaeybroeck, B., De Cruz, L., De Troch, R., Caluwaerts, S., Giot, O., et al. (2018). The CORDEX.be initiative as a foundation for climate services in Belgium. *Climate Services*, *11*, 49–61.
- Utsumi, N., Seto, S., Kanae, S., Maeda, E. E., & Oki, T. (2011). Does higher surface temperature intensify extreme precipitation? *Geophysical Research Letters*, *38*, L16708. <https://doi.org/10.1029/2011GL048426>
- Wasko, C., & Sharma, A. (2014). Quantile regression for investigating scaling of extreme precipitation with temperature. *Water Resources Research*, *50*, 3608–3614. <https://doi.org/10.1002/2013WR015194>
- Wasko, C., Tang Lu, W., & Mehrotra, R. (2018). Relationship of extreme precipitation, dry-bulb temperature, and dew point temperature across Australia. *Environmental Research Letters*, *13*(7), 074031.
- Westra, S., Alexander, L. V., & Zwiers, F. W. (2013). Global increasing trends in annual maximum daily precipitation. *Journal of Climate*, *26*(11), 3904–3918.
- Westra, S., Fowler, H. J., Evans, J. P., Alexander, L. V., Berg, P., Johnson, F., et al. (2014). Future changes to the intensity and frequency of short-duration extreme rainfall. *Reviews of Geophysics*, *52*, 522–555. <https://doi.org/10.1002/2014RG000464>
- Yu, K., & Moyeed, R. A. (2001). Bayesian quantile regression. *Statistics & Probability Letters*, *54*(4), 437–447.
- Zhang, X., Zwiers, F. W., Li, G., Wan, H., & Cannon, A. J. (2017). Complexity in estimating past and future extreme short-duration rainfall. *Nature Geoscience*, *10*, 255–259.