Formulating and testing a method for perturbing precipitation time series to reflect anticipated climatic changes

Hjalte Jomo Danielsen Sørup[1,2], Stylianos Georgiadis[1,3], Ida Bülow Gregersen[4] and Karsten Arnbjerg-Nielsen[1,2]

[1]Technical University of Denmark, Global Decision Support Initiative, Lyngby, Denmark
[2]Tecnical University of Denmark, Department of Environmental Engineering, Lyngby, Denmark
[3]Tecnical University of Denmark, Department of Applied Mathematics and Computer Science, Lyngby, Denmark
[4]Ramboll Danmark A/S, Department of Climate Adaptation and Green Infrastructure, Copenhagen, Denmark

Correspondence to: Hjalte J.D. Sørup (hjds@env.dtu.dk)

Abstract. Urban water infrastructure has very long planning horizons and planning is thus very dependent on reliable estimates on the impacts of climate change. Many urban water systems are designed using time series with a high temporal resolution. To assess the impact of climate change on these systems similarly high resolution precipitation time series for future climate are necessary. Climate models cannot at their current resolutions provide these time series at the relevant scales. Known methods for stochastic downscaling of climate change to urban hydrological scales have known shortcomings in constructing realistic climate changed precipitation time series at the sub-hourly scale. In the present study we present a deterministic methodology to perturb historical precipitation time series at minute scale to reflect non-linear expectations to climate change. The methodology shows good skill in meeting the expectations to climate change of extremes at event scale when evaluated at different timescales from the minute to the daily scale. The methodology also shows good skill with respect to representing expected changes to seasonal precipitation. The methodology is very robust to the actual magnitude of the expected changes as well as the direction of the changes (increase or decrease) even for situations where the extremes are increasing for seasons that in general should have a decreasing trend in precipitation. The methodology can provide planners with valuable time series representing future climate that can be used as input to urban hydrological models and give better estimates of climate change impacts on these systems.

1 Introduction

Climate change impacts water management worldwide as the water cycle is an essential part of the climate system. The planning horizon for water infrastructure is often very long, making reliable predictions of future climate crucial (Arnbjerg-Nielsen et al., 2015b). In design of water infrastructure precipitation data is needed. Especially for urban infrastructure the time resolution of precipitation data needed for design and planning is much finer than what is provided by climate models (Berndtsson and Niemczynowicz, 1988; Schilling, 1991). Hence a lot of effort is put into giving reliable estimates of what the expected change in precipitation will be at these fine scales (Fowler et al., 2007; Kendon et al., 2014; Mayer et al., 2015). Expected changes in precipitation, however, do not translate directly into changes in floods or overflows from structures. To determine these changes, urban hydrological models have to be run, driven by the changed precipitation (Olsson et al., 2009;

Willems et al., 2012). By definition, fine resolution precipitation time series for future climates are not observable and hence a multitude of statistical approaches have been developed to enable generation of time series with properties that for a large range of metrics have the same characteristics as the expected future precipitation (Willems, 1999; Olsson and Burlando, 2002; Cowpertwait, 2006; Molnar and Burlando, 2008; Burton et al., 2010; Willems et al., 2012; Sørup et al., 2016a).

5 Expectations to precipitation at event level under climate change are often non-linear with the anticipation that changes in occurrence and size of extreme events are higher than changes in seasonal or yearly precipitation (Boberg et al., 2010). This is a problem often solved by weather generators or other similar downscaling techniques (Fowler et al., 2007; Burton et al., 2010), but these often have difficulty in presenting realistic time series at the sub-hourly to hourly time scales, relevant for urban infrastructure (Segond et al., 2006; Verhoest et al., 2010; Sørup et al., 2016a). Several studies have tested the

10 applicability of Markov models for simulation of high resolution precipitation series (Srikantan and McMahon, 1983; Thyregod et al., 1998; Ailliot et al., 2009; Gelati et al., 2010; Sørup et al., 2012). The approach has the advantage that realistic chronology is created in the output. However, for very high resolutions the sensing method of the gauge may have an impact on the signal, giving an upper bound to temporal resolution of the model, as it has been shown for e.g. tipping bucket gauges (Thyregod et al., 1998; Sørup et al., 2012).

15 In the present study, we develop and demonstrate a novel non-linear methodology that perturbs existing precipitation time series to reflect complex expectations to precipitation in a changed future climate. The method incorporates regional historical knowledge about precipitation through the use of Intensity-Frequency-Duration (IDF) relationships (WMO, 2009) and knowledge about the expected changes of these due to climate change. Thus, the method generates time series for a changed climate which are chronologically identical to the observations used as input but perturbed to reflect climate change.

20 These series can be used as input for hydraulic or hydrologic models where the climate change effect has to be assessed for all possible rain conditions.

The presented methodology is based on the assumption that precipitation can be scaled according to identified expectations to climate changes. In its simplest form, this assumption is identified as the Delta Change (DC) method (Fowler et al., 2007). The basic assumption is that relative changes in output from climate models might represent expectations to climate change

25 well even though the output itself could be wrongly scaled in absolute values. A more elaborate use of this assumption is provided by Distribution Based Scaling (DBS) presented by Yang et al. (2010). In this approach parameters are derived from regional climate model data to estimate present and future distribution functions for rainfall intensities. The relative change in the distribution parameters is applied to a similar distribution function based on observational data. Thereby, perturbation of rainfall intensities due to climate change relies on the rarity of the individual events and change markedly from average to

30 extreme events with high impact on hydrological responses of simulation models (van Roosmalen et al., 2011). Unlike the study by Yang et al. (2010), the expected changes in this study are not calculated directly using the DC method on Regional Climate Model output; they are derived from comprehensive state-of-the-art studies where the full available data basis is used to determine realistic expectations to changes to precipitation due to climate change (e.g. Giorgi, 2006; Kendon et al., 2008; Christensen et al., 2015).

2 Methodology

In urban water management, the relevant time frame to consider is most often that of the rain event (Willems, 1999). The determination of robust IDF relationships for present climate at the relevant time scales is a prerequisite. For developed countries where high resolution precipitation is generally available these two prerequisites are very often met (Arnbjerg-Nielsen et al., 2015a), making the methodology generally relevant. The general flow of the methodology is presented in Figure 1 and how to proceed with each step is presented in the following sections (2.1-2.5).

Figure 1 Flow diagram showing the general process involved in the presented methodology.

2.1 Modelling Framework

Let us consider a system S that describes precipitation over a time period. The original data are expressed as a time series of precipitation intensity over fixed time steps. This time series alternates between a dry period (no precipitation) and a rainy one. A given event is characterized as dry, *extreme* or *non-extreme* with respect to amount of precipitation during the event. We denote by E the state space of the system S, with at least three states, i.e. $|E| \geq 3$. Let also D_0, D_1 and D_2 be the non-empty sets of states of dry periods, non-extreme and extreme events, respectively, with $|D_0| = 1$, $|D_1| = d_1 \geq 1$ and $|D_2| = d_2 \geq 1$, i.e. there exists exactly one state for dry periods, $D_0 = D^{dry}$, but a different number of extreme and non-extreme states (d_1 and d_2; respectively), can be defined for both the non-extreme and extreme events. An extreme event can be further categorized according the severity of the phenomenon, expressed in terms of the return period of the measured intensity. Non-extreme events can be categorized according to the season in which they appear. Hence, the state space E is partitioned into three disjoint subsets as follows:

$$E = D_0 \cup D_1 \cup D_2, \tag{Eq. 1}$$

where $D_i \cap D_j = \emptyset$, $i \neq j$, $i, j \in \{0,1,2\}$. We link the non-extreme events to the seasonality of the phenomenon and thus $D_1 = \{D^{winter}, D^{spring}, D^{summer}, D^{fall}\}$, that is $d_1 = 4$. D_2 can likewise be partitioned into one or several states appropriate to describe extreme precipitation which may have different return periods or different hydro-climatic origin. In this study, we use a partition based on return periods with $D_2 = \{D^2, D^{10}, D^{100}\}$, referring to states that classifies the extremes as either 2, 10 and 100-year events based on return level.

By definition there is always a dry period between two events and we assume that there is no dependence between consecutive events. We define the following processes that describe the evolution of a semi-Markov system (Barbu and Limnios, 2008):

- $\mathcal{J} := (\mathcal{J}_n)_{n \in N}$ is a Markov chain with state space E, where $\mathcal{J}_n$ is the state of the system at the n-th event;

- $U := (U_n)_{n \in N}$ is the sequence of jump times between states with state space $\mathbb{N}$ and $U_0 = 0$;
- $Z := (Z_k)_{k \in N}$ is a discrete-time process with states on E, with Z_k to be the state of the system at a time step k.

The processes $\mathcal{J}$ and Z are related through the formula

$$Z_k = \mathcal{J}_{N(k)}, k \in N, \tag{Eq. 2}$$

where $N(k)$ is the discrete-time counting process of events in $[1, k] \subset N$, i.e.

$$N(k) := max\{n \in N : U_n \leq k\}. \tag{Eq. 3}$$

The corresponding transition matrix of the chain $\mathcal{J}$ is very simple to be written. Figure 2b illustrates the evolution of the stochastic system described above.
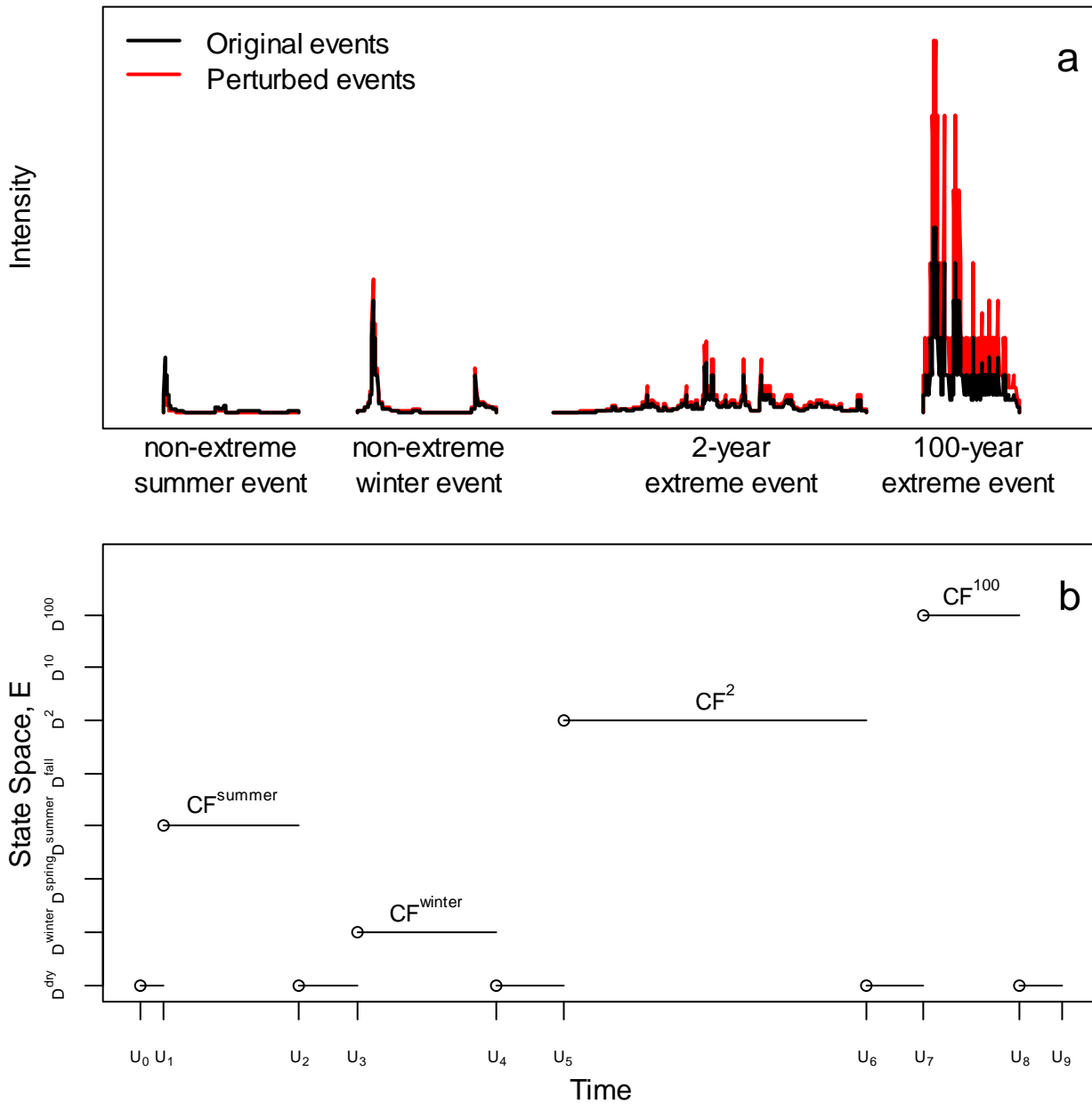
Figure 2 a) Illustration of the magnitude of perturbation of events for non-extreme summer and winter events as well as 2- and 100-year extreme events, with summer events being perturbed with a factor below one and factors for the winter and the extremes being above one. Factors for extremes are higher than for the winter events, and factors for the very extreme is higher than for the more moderate extreme. b) Illustration of the states associated with the different events if they were to happen in the shown chronology, the dry state, D^{dry}, is present between all wet states.

2.2 Framework for determining state of individual events

There is no unique way to assign a state to an extreme event. In the literature some studies apply hydro-climatic regimes for this classification (Gelati et al., 2010; Svoboda et al., 2016), while others apply event statistics (Madsen et al., 2009; Sørup et al., 2016). For any given application, the most appropriate classification depends on the data available. In this paper, various methods based on the maximum mean intensities are used to define the event state. For all investigated methods the changes to extremes are evaluated by calculation of IDF curves based on return levels, z_i's, at event level for a selection of return periods, I (WMO, 2009). The return period (T) of individual events across all intensities is determined using the median plotting position (Rosbjerg, 1988):

$$T_{median} = \frac{T_{total} + 0.4}{rank - 0.3},$$
(Eq. 4)

where T_{total} is the length of the time series and *rank* is the rank number of the individual event.

Using data with observations every minute and a minimum dry weather separation between events of 60 minutes, the mean maximum intensities over 5, 10, 30, 60, 180, 360 and 720 minutes are calculated for each event. At shorter timeframes, e.g. one minute, the variability of the observed extremes are expected to be very large due to the inherent sampling error (Fankhauser, 1998) and at very long timeframes, e.g. one day (i.e. 1440 minutes) the extremes are often consisting of several events following one another and a different event definition would be necessary to ensure that the real extremes are identified (Madsen et al., 2009). A representative return period for the event is derived based on a mathematical comparison to regional IDF estimates (Figure 3). This return period is then in turn used to define the state of the event. We test four different selection criteria which define the state of extreme events as either D^2, D^{10} or D^{100}. The selection criteria are listed in Section 3.3.
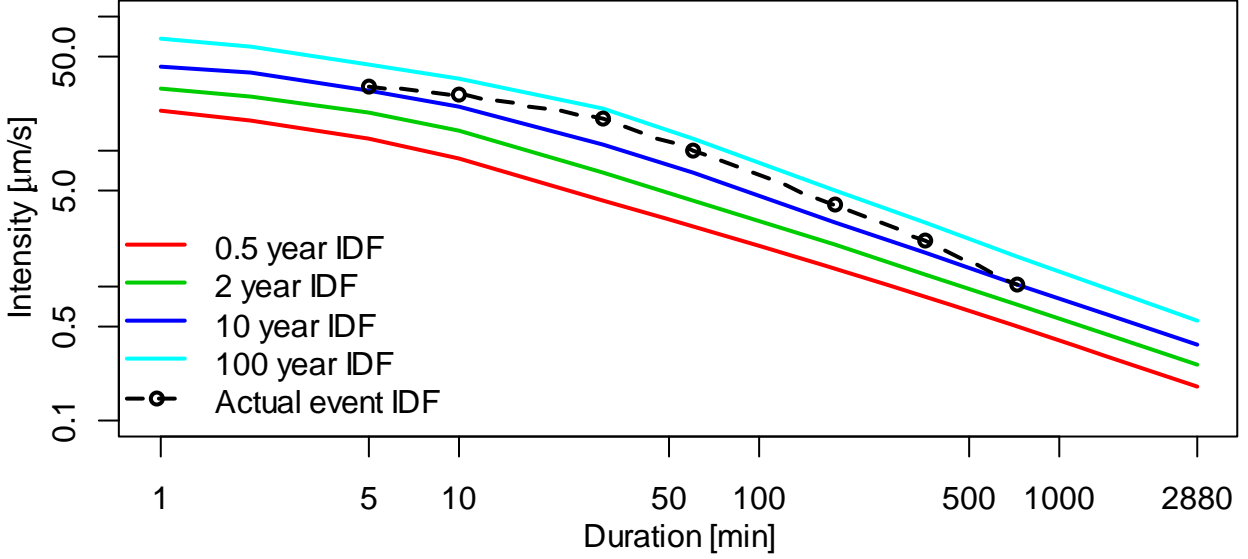
Figure 3 The IDF curve for an extreme event in comparison to the regional IDF curves for 0.5, 2, 10 and 100 year return periods respectively (based on Madsen et al. (submitted)).

2.3 Perturbation and change factor

With each event of a time series classified according to a state, the time series can be perturbed using the following methodology linking the time series to the states of the individual events.

Let R_k, $k \in N$, be the precipitation intensity at time step k and $R := (R_k)_{k \in N}$ the corresponding process describing these intensities. The process of perturbed precipitation in each time step k is denoted by $R^* := (R_k^*)_{k \in N}$.

Similarly to the state space E, we introduce the state space of the change factors, denoted by E_{CF}, $|E_{CF}| = |E|$. We can then write

$$E_{CF} = C_0 \cup C_1 \cup C_2, \tag{Eq. 5}$$

with $|C_0| = 1$, $|C_1| = d_1$ and $|C_2| = d_2$.

We consider the process $CF := (CF_n)_{n \in N}$ with state space E_{CF}, where CF_n is the change factor at the n-th event. Let $W :=$ $(W_k)_{k \in N}$ be the chain, with state space E_{CF}, of change factors in time steps $k \in N$, that is

$$W_k = CF_{N(k)}, \tag{Eq. 6}$$

with $N(k)$ to be the counting process defined in (Eq. 3). Under the above notation, the original and perturbed sequences of precipitation, R_k and R_k^*, are written as

$$R_k^* = W_k R_k. \tag{Eq. 7}$$

This means that, for a sequence of events, some events will be perturbed more than others and for extreme cases some might be reduced while others are increased depending on the local expectations to climate change. Figure 2a shows an example where a non-extreme summer event is perturbed to a lesser volume than original while a winter non-extreme is increased marginally and both 2 and 100 year extremes are increased considerably more (both in absolute numbers as well as in relative percentages). Figure 2b shows how the state space changes if these four events were to happen chronologically in time with the state jump times marked at the x-axis.

2.4 Volume correction based on seasonal dependence of extremes

The extreme part of precipitation is only expected to constitute a smaller fraction of the total precipitation volume on an annual basis (Sørup et al., 2016b) but as extreme precipitation is often associated with a particular season (see e.g. Sørup et al., 2012) the volumetric part of the extremes might be higher for sub-annual considerations. This implies that situations where the expectations for changes to the extremes are very different from the expectations to changes to seasonal precipitation has to be handled through volumetric corrections in order to accommodate that both expectations to changes in extremes and overall seasonal changes are correct. How to do this best will be very much dependent on the local conditions. In our case this is described in Section 3.4.

2.5 Evaluation of Perturbed Time Series

The evaluation of the perturbed time series is done against the original time series and against the expected changes.

The average percentwise difference between the perturbed return levels, $z_{i,j,m}^*$, of the modelled time series, R_k^*, perturbed with the time dependent change factors, W_k, against the same return levels, $z_{i,j,m}$, of the original time series, R_k, multiplied with the target change factor, CF_e^j, can be defined as:

$$\Phi_{i,j,m} = \left(1 - \frac{z_{i,j,m}^*}{z_{i,j,m}CF_e^j}\right)100\% ,$$ (Eq. 8)

across all IDF points, i, all extremity levels and seasonality, j, and all perturbed time series, m. A combined skill score, Φ, across all considered metrics that describe the average deviance from the expectations can then be defined as:

$$\Phi = \sum_{i\in I}\sum_{j\in J}\sum_{m\in M}\frac{|1-\Phi_{i,j,m}|}{|I|\,|J|\,|M|} ,$$ (Eq. 9)

with $|I|\,|J|\,|M|$ being the product of the total number of IDF points, I, the total number of extreme levels considered plus seasonality, J, and the total number of time series perturbed , M, as a normalization factor.

2.6 Sensitivity Analysis

The robustness of the methodology is tested by evaluating its sensitivity to the actual magnitude of the target parameters for both extreme and seasonal changes. Low (L), mean (M) and high (H) scenarios are constructed and paired in all possible combinations to assess both the individual and combined influence of these (Table 1). As this increases the number of

scenarios with which the precipitation time series substantially are perturbed, this is not done until after an initial evaluation of the state selection criteria.

Table 1 Tested combinations of extreme and seasonal changes.

Seasonality	Extremes		
	Low expected change	Mean expected change	High expected change
Low expected change	LL	ML	HL
Mean expected change	LM	MM	HM
High expected change	LH	MH	HH

3 Case study: Denmark

To showcase the methodology it is applied to Danish conditions where the situation is that complex non-linear changes are expected with respect to precipitation in a changed climate.

3.1 Data

10 **3.1.1 Observational Data**

Precipitation data from the Danish SVK rain gauge network is used in this study (Mikkelsen et al., 1998; Madsen et al., 2002). For this study 10 time series from different parts of Denmark with lengths of approximately 33 years between 1979 and 2012 are used. To distinguish individual events a dry weather period between individual events of at least 60 minutes is applied.

15 **3.1.2 IDF Curves**

For present climate IDF curves are extracted from a regional model for extremes originally developed by Madsen et al. (1998) and updated by Madsen et al. (2009) and Madsen et al. (In Review). The IDF curves vary across Denmark but a single mean regional curve is chosen for this study independent of the location of the gauge considered. Table 2 summarizes the IDF values used.

Table 2 IDF intensities (μm/s) for various return periods for Denmark extracted from the model presented by Madsen et al. (In review).

Return Period (years)	Duration (min)						
	5	10	30	60	180	360	720
T=100	43.67	34.80	20.63	12.47	5.21	3.11	1.72
T=10	28.62	21.43	11.37	6.95	3.09	1.86	1.09
T=2	19.54	14.08	7.08	4.38	2.04	1.25	0.75
T=0.5	12.40	8.73	4.33	2.75	1.33	0.84	0.51

3.1.3 Expectations to Climate Change

The official recommendations regarding climate change for urban infrastructure in Denmark was determined by Gregersen et al. (2014) on the basis of the ENSEMBLES data set (van der Linden and Mitchell, 2009), with the addition of a few simulations using high-end scenarios. The data set indicates that in general precipitation amounts and intensities will increase and that extremes will increase more than the expected mean increases for Denmark. Furthermore, the results show that it is very likely that increases will be more pronounced for the very rare extremes compared to the more frequent extremes. Table 3 sums up these official expectations for the three return periods that has to be assessed in Danish urban hydrological contexts.

Table 3 Expected changes in extreme precipitation for Denmark. All values from Tabel 1 of Gregersen et al. (2014).

Change factor for extreme precipitation (-)	2 year event [CF^2]	10 year event [CF^{10}]	100 year event [CF^{100}]
Low expected change	1.0	1.0	1.0
Mean expected change	1.2	1.3	1.4
High expected change	1.45	1.7	2.0

In addition, the Danish Meteorological Institute has published expectations regarding climate change on a seasonal basis (Olesen et al., 2014). The analysis is performed for a range of climate variables and focus on utilizing the data available in the best possible way to create realistic uncertainty intervals for the expected changes. The estimated change factors for precipitation are based on analysis of the RCP2.6 and the RCP8.5 scenarios (Moss et al., 2010), hence, a low-end and a high-

end emission scenario, respectively. Table 4 lists these expectations as well as a simple mean average of the two to represent the mean expected change. To match the change factors for extreme precipitation in Gregersen et al. (2014), which primarily is based on the more average emission A1B scenario (Nakicenovic et al., 2000), simple scaling of the seasonal expectations to a mid-point is applied, as scalability has been shown to be a valid assumption across most scales and most indices (Christensen et al., 2015). The A1B scenario does not lie exactly in the middle between the RCP2.6 and the RCP8.5 scenarios, but definitely somewhere between these and the original estimates from Olesen et al. (2014) are kept as low and high expected changes for the sensitivity analysis.

Table 4 Expected seasonal changes to precipitation in Denmark based on data from Tabel 5 of Olesen et al. (2014) and linear scaled midpoint values.

Change factor for seasonal precipitation (-)	Winter [CF^winter]	Spring [CF^spring]	Summer [CF^summer]	Fall [CF^fall]
Low expected change (RCP2.6)	1.0	1.0	1.0	1.0
Mean expected change	1.1	1.05	0.9	1.05
High expected change (RCP8.5)	1.2	1.1	0.8	1.1

3.2 Defining states

For Denmark the state space (Eq. 1) is defined with a total of eight states based on the expectations to climate change listed in Tables 3 and 4 with four seasonal states defined for the non-extreme events and three states for the different extreme event levels:

$$E = \{D^{dry}, D^{winter}, D^{spring}, D^{summer}, D^{fall}, D^2, D^{10}, D^{100}\}. \tag{Eq. 10}$$

Correspondingly the change factors used to perturb the time series are, as a starting point, determined based on the mean expectations listed in Tables 3 and 4.

3.3 Determining state of individual events

For the determination of the state of the individual extreme events four different selection criteria are investigated, with the purpose of defining a representative return period for each event. All points mentioned refer to the return periods of the events intensity points, $T = \{T_5, T_{10}, T_{30}, T_{60}, T_{180}, T_{360}, T_{720}\}$, shown in a situation as depicted in Figure 3:

A. The maximum return period is used to define the return period of the whole event (based on one point);

$$T_{event} = T_1^* = \max T. \tag{Eq. 11}$$

B. The mean of the three largest return periods is used to define the events (based on three points);

$$T_{event} = \frac{1}{3}\sum_{i=1}^{3} T_i^*, \tag{Eq. 12}$$

where T_2^* and T_3^* are the second and third maxima respectively, i.e. $T_2^* = \max\{\boldsymbol{T}\backslash T_1^*\}$ and $T_3^* = \max\{\boldsymbol{T}\backslash(T_1^* \cup T_2^*)\}$.

C. The mean of all the return periods is used to define the events (based on all seven points);

$$T_{event} = \overline{\boldsymbol{T}}. \tag{Eq. 13}$$

D. A customized step-wise threshold selection criterion is constructed where the event-specific IDF curve is compared to regional IDF levels.

Criterion D is important to test as this allows for construction of a criterion that is closely linked to specific knowledge on the place-specific precipitation dynamics, i.e. for how many duration points on the IDF curve a given return period has to be exceeded for it to be essential for the classification of the event.

Following these selection criteria, four different systems, $S_i, i \in \{A, B, C, D\}$, are constructed and analysed. Options S_A to S_C are straightforward based on Equations (11)-(13) but option S_D is determined specifically for the case study. Table 5 summarizes the methodology for option S_D used in this study; specifically it is reflected that for very extreme events less durations has to be extreme for the event as a whole to be considered extreme compared to the more moderate 2-year return level.

Table 5 Selection criterion S_D for choosing T_{event}s at event level.

A T_{event} is chosen of	If	Or
2 year event	At least 4 points from the event have a return period above 0.5 years	At least 2 points from the event have a return period above 2 years
10 year event	At least 3 points from the event have a return period above 2 years	At least 2 point from the event have a return period above 10 years
100 year event	At least 3 points from the event have a return period above 10 years	At least 2 point from the event have a return period above 100 years
Non-extreme event	None of the above criteria are met	

3.4 Volume correction based on seasonal dependence of extremes

In previous studies using the SVK data set, it has been shown that:

1. the extreme events account for at most 25% of the total rainwater volume on an annual basis (Sørup et al., 2016b), and

2. the extreme events occur mostly in the summer season (Sørup et al., 2012)

Furthermore, in the summer season the excepted seasonal change (-10%) differs mostly from the expected change in extremes (+20-40%), see Table 4 and Table 3, respectively. Based on this information the seasonal change factor for non-extreme summer events has to be adjusted to reach overall changes factors reported in Table 4. We estimate a partition between non-extreme and extreme events of $\{f_{non-extreme}, f_{extreme}\} = \{0.8, 0.2\}$ and the change factor for 2-year events, CF^2, is used to represent the extremes as the largest seasonal volume by far is for the more frequent extremes (Sørup et al., 2016b). In this way the change factor for summer, CF^{summer}, can be adjusted from its value listed in Table 4 (0.9) as:

$$CF_{adjusted}^{summer} = \frac{CF^{summer} - CF^2 f_{extreme}}{f_{non-extreme}} = \frac{0.9 - 1.2*0.2}{0.8} = 0.825. \tag{Eq. 14}$$

In other words the change factor for non-extreme summer events are modified from -10% to -17.5% in order to compensate for the positive change of +20-40% to the extremes occurring in the summer period. For the other seasons such an adjustment is not needed.

4 Results

4.1 4.1 Evaluation of Selection Criteria

The 10 time series are perturbed using the four different state selection criteria (S_A-S_D) and the evaluation metric is calculated using Eq. (9) with the extreme events having return periods closest to 2, 10 and 100 years (Table 6). Overall, state selection criterion S_D outperforms the other alternatives even though all selection criteria seem reasonable as all estimated deviances are below 13% of the expected changes.

Table 6 Calculated skill scores, Φ, for the four selection criteria A-D calculated using Eq. (10).

	S_A	S_B	S_C	S_D
Φ	9.3%	8.5%	12%	6.4%

In order to study the performance for each state, we construct the skill score variable of Eq. (8) and plot them against the duration for the individual extremes and against months for seasonal precipitation (Figure 4). Plotted this way 100% represent a perfect fit, 0% represent no change and everything positive represent a change in the right direction. For the 2-year return levels both state selection criteria S_B and S_D perform similarly and with a relative change close to 100 %. State selection criterion S_A overestimates the 2-year return level with approximately 10 % on average and state selection criterion S_C likewise underestimates it, which still corresponds to a positive change for the events (Figure 4a). For the 10-year return level, all state selection criteria perform similarly very well (Figure 4b). When the 100 year return level is evaluated the reason for criterion S_D's better overall performance become clear; it is the only criterion that does not systematically underestimate this return level (Figure 4c). Even so, all criteria produce results where the direction of change is correct.

Given the inherent uncertainty in estimating the actual levels of such events, obtaining close to 85% of the expected change is considered good. With respect to the seasonal behaviour all state selection criteria have approximately the same performance at a level close to 100% (Figure 4d).
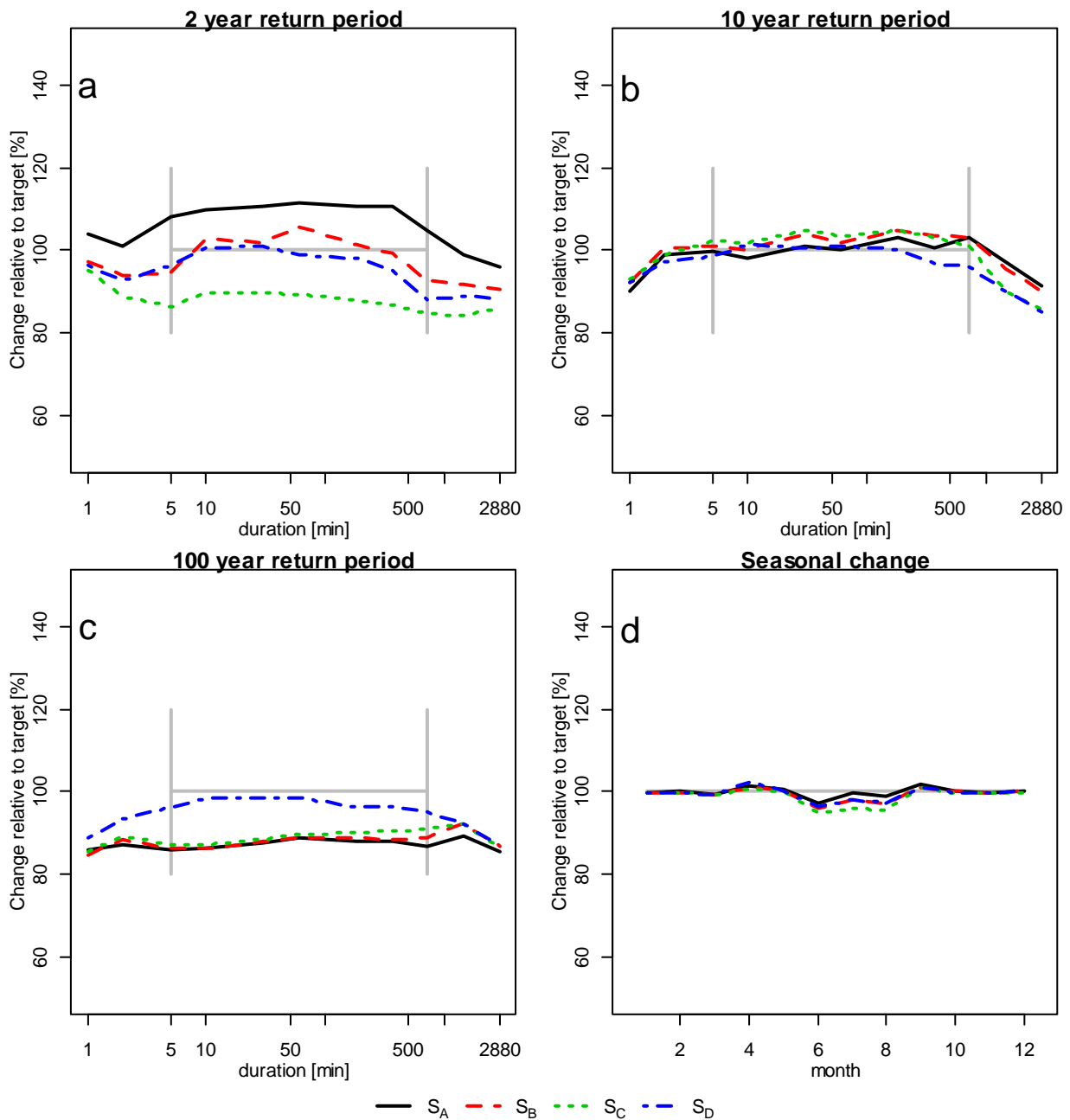
Figure 4 Performance of the different selection criteria, S_A-S_D, in producing a) 2 year extremes, b) 10 year extremes, c) 100 year extremes and d) seasonal changes according to the perturbation schemes listed in Tables 3 and 4.

The performance of all the state selection criteria drops when considering duration that are both shorter and longer than the durations used in the state selection methodology (5-720 minutes). At the minute scale, this is of minor importance but at

two days (2880 minutes) the tendency is very robust across different state selection criteria and extremity levels. This is most likely because these average extreme events are caused by several events with dry periods in between. Hence the individual events are each assessed to be non-extreme and they are adjusted towards lower volumes, even though they combined are rather extreme.

4.2 Sensitivity Analysis with Selection Criterion D

The sensitivity analysis is carried out for the best state selection criterion only, i.e. criterion S_D. The resulting skill scores for the nine individual sensitivity scenarios are listed in Table 7. The highest sensitivity is found when changing between the different extreme precipitation scenarios; with a large increase of the metric when moving from low to mean and also a notable increase when moving from mean to high scenarios. As such the performance of the methodology drops with the magnitude of the expected changes to extremes, but even for the high extremes the performance is similar to the performance of state selection criteria S_A to S_D in Table 6. The methodology, on the other hand, shows very little sensitivity to the variation in expectations to seasonal changes, not even for the combination where the difference between expectations to seasonal summer precipitation (-20%) and the extremes become very high (+45-100%).

Table 7 Calculated skill scores, Φ, for selection criterion S_D for the nine different sensitivity scenarios listed in Table 1 calculated using Eq. (9).

Φ		Extremes		
		Low	Mean	High
Seasonality	Low	0.0%	6.0%	8.6%
	Mean	1.0%	6.4%	8.8%
	High	1.2%	6.3%	8.8%

For all extreme indices (Figure 5a-c), the sensitivity of the expected change of extremes is notable and, especially for the 100 year return level, it is clear that performance drops with increased magnitude of the expected changes to extremes (Figure 5c) but only to levels comparable to that of the state selection criteria S_A-S_C as shown in Figure 4. Again the performance for two day events (2880 minutes) is worse than average as also seen in Figure 4. For seasonality (Figure 5d), the general picture is that the sensitivity of both expectations to seasonality and extremes are of less importance and at a similar level, which in general is a lower level than the one observed for the three extreme indices.
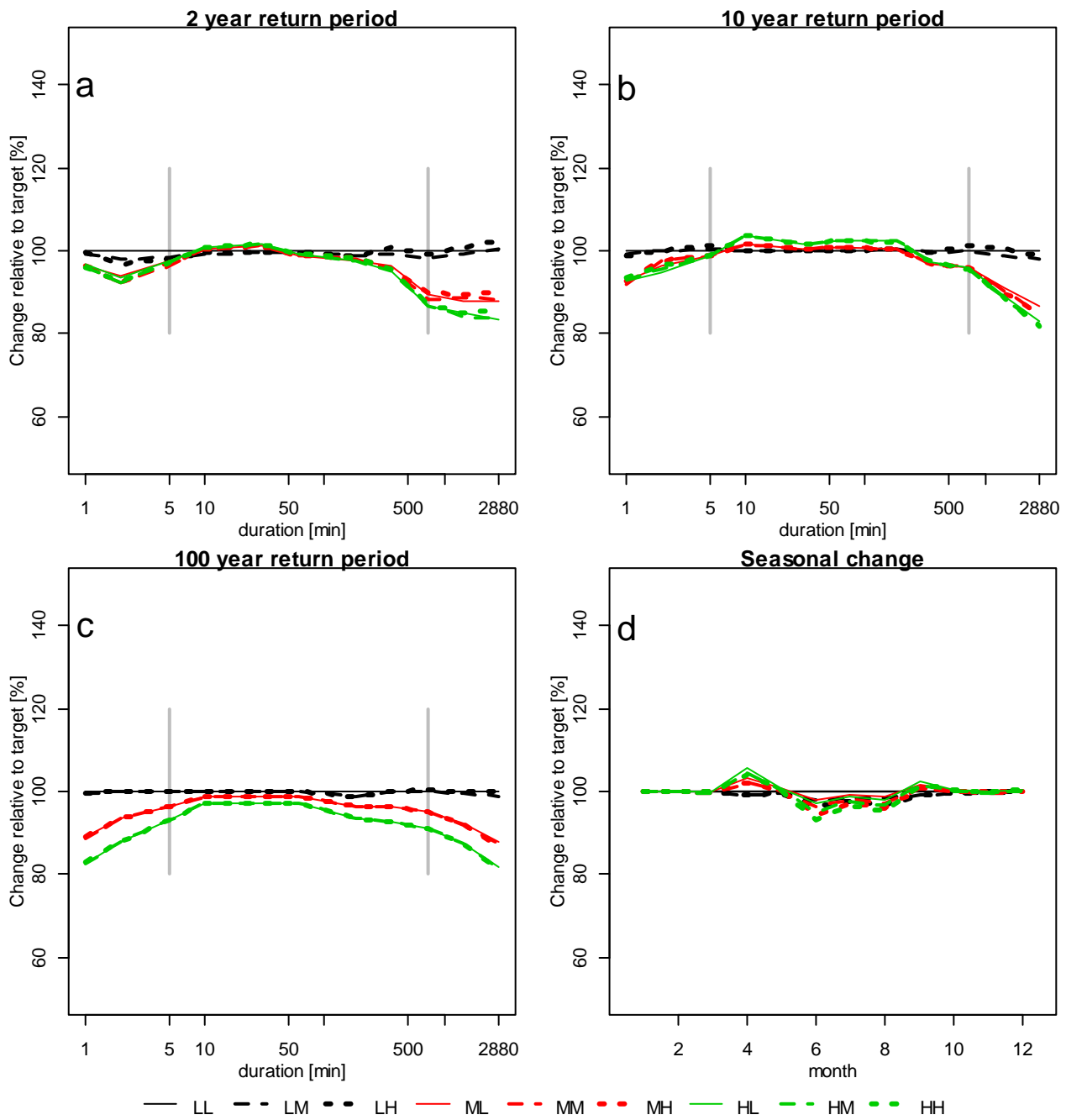
Figure 5 Performance of selection criterion S_D for different parameter values as specified in Table 1 for a) 2 year extremes, b) 10 year extremes, c) 100 year extremes and d) seasonal changes under climate change.

5 Discussion

The proposed framework is very flexible and the separation of dry, non-extreme and extreme weather makes it possible to very effectively perturb time series to reflect different changes in different categories. The presented case study uses eight states to distinguish between different levels of extremes and different season and is able to produce time series that satisfactory represent the expected changes listed in Tables 3 and 4. For other places a different number of states could be relevant and the seasonal partition could be different depending on the local climate and expectation to climate change. The proposed modelling framework fully supports these spatial variations.

Four different state selection criteria over specified event durations are tested in the present study, see Section 2.2, as these covered realistic possibilities for the data set used in this study and the focus on urban hydrology. As such, different state selection criteria for different event durations could be relevant in different contexts and could, as illustrated by state selection criterion S_D, be specified as very subjective and case specific criteria. In this study, the subjective state selection criterion S_D outperforms the other criteria, see Table 6 and Figure 2, but the superiority is mainly due to its ability to produce the largest changes for the very large, and very uncertain, extreme events. If this part of the evaluation is disregarded, criteria S_B and S_D have very similar performance pointing at criterion S_B as being a good onset for investigating data sets where no presumptions exist and no case specific criterion can be constructed.

All state selection criteria showed a drop in performance for longer duration events than the ones used in the methodology; this is likely due to the used event definition with a minimum of 60 minutes of dry weather between individual events which will mean that very long lasting extremes likely are split into several events and therefore not identified as extremes. A different event definition with longer minimum dry period between events could probably partly solve this, but it would reduce the number of events markedly and increase the chance of small events close to extremes being seen as part of the extreme with a somewhat false classification as a consequence.

The methodology is relatively sensitive to the magnitude of the perturbation factors, see Section 4.2, but the sensitivity is not very dominant and is only at the same size as the sensitivity of the different state selection criteria. Also, the methodology does not address the possibilities of changes to dry spells or changes to the occurrence rate of extremes in general. A future research direction could be to study how the state selection criteria along with the semi-Markov system applied here can be used to generate fully stochastic time series where both the inter-event time and the occurrence probability of the extreme states will be included as criteria that can be changed to meet the expectations to climate change.

6 Conclusions

The proposed methodology is a promising way of creating artificial perturbed precipitation time series, which can represent a changed climate and be used as input in hydrologic and hydraulic models. The methodology perturbs existing time series based on a semi-Markov system where precipitation time series are split into events characterized as *dry*, *extreme* or *non-extreme*. The wet events are divided into different states based on an Intensity-Duration-Frequency relationship based state

selection criterion. Of the four tested state selection criteria, the case specific show the best results, but also the more general criteria could be of use when less knowledge about the precipitation regime is available. The sensitivity of the methodology was tested against very different expectations to climate change both with respect to seasonal changes and changes to extremes and is generally very robust, also regarding seasons where the general change is negative while the expectations to

5   extremes is positive. The produced time series satisfactory reproduce changes across all seasons and across all levels of extremes relevant for urban hydrology.

References

Ailliot, P., Thompson, C. and Thomson, P.: Space-time modelling of precipitation by using a hidden Markov model and censored Gaussian distributions. *Journal of the Royal Statistical Society Series C-Applied Statistics*, 58(3), 405–426.

10   doi:10.1111/j.1467-9876.2008.00654.x. 2009.

Arnbjerg-Nielsen, K., Funder, S.G. and Madsen, H.: Identifying climate analogues for precipitation extremes for Denmark based on RCM simulations from the ENSEMBLES database. *Water Science and Technology*, 71(3), 418–425. doi:10.2166/wst.2015.001. 2015a.

Arnbjerg-Nielsen, K., Leonardsen, L. and Madsen, H.: Evaluating adaptation options for urban flooding based on new high-

15   end emission scenario regional climate model simulations. *Climate Research*, 64(1), 73–84. doi:10.3354/cr01299. 2015b.

Barbu, V. and Limnios, N.: *Semi-Markov Chains and Hidden Semi-Markov Models toward Applications: Their Use in Reliability and DNA Analysis*. Springer, New York, NY, USA. doi:10.1007/978-0-387-73173-5. 2008.

Berndtsson, R. and Niemczynowicz, J.: Spatial and temporal scales in rainfall analysis: Some aspects and future perspectives, *Journal of Hydrology*, 100, 293–313, doi:10.1016/0022-1694(88)90189-8. 1. 1988.

20   Boberg, F., Berg, P., Thejll, P., Gutowski, W.J. and Christensen, J.H.: Improved confidence in climate change projections of precipitation further evaluated using daily statistics from ENSEMBLES models. *Climate Dynamics*, 35(5), 1509–1520. doi:10.1007/s00382-009-0683-8. 2010.

Burton, A., Fowler, H.J., Blenkinsop, S., and Kilsby, C.G.: Downscaling transient climate change using a Neyman-Scott Rectangular Pulses stochastic rainfall model. *Journal of Hydrology*, 381(1-2), 18–32. doi:10.1016/j.jhydrol.2009.10.031.

25   2010.

Cowpertwait, P. S. P.: A spatial-temporal point process model of rainfall for the Thames catchment, UK, *Journal of Hydrology*, 330, 586–595, doi:10.1016/j.jhydrol.2006.04.043. 2006.

Christensen, O.B., Yang, S., Boberg, F., Maule, C.F., Thejll, P., Olesen, M., Drews, M., Sørup, H.J.D. and Christensen, J.H.:. Scalability of regional climate change in Europe for high-end scenarios. *Climate Research*, 64(1).

30   doi:10.3354/cr01286. 2015.

Fankhauser R.: Influence of systematic errors from tipping bucket rain gauges on recorded rainfall data. *Water Science and Technology*, 37(11): 121–129. doi:10.1016/S0273-1223(98)00324-2. 1998.

Fowler, H.J., Blenkinsop, S. and Tebaldi, C.: Review linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology* 27, 1547 – 1578. doi:10.1002/joc.1556. 2007.

Gelati, E., Christensen, O. B., Rasmussen, P. F. and Rosbjerg, D.: Downscaling atmospheric patterns to multi-site precipitation amounts in southern Scandinavia. *Hydrology Research*, 41(3-4), 193–210. doi:10.2166/nh.2010.114. 2010.

Giorgi, F.: Climate change hot-spots. *Geophysical Research Letters*, 33(8), L08707. doi:10.1029/2006GL025734. 2006.

Gregersen, I.B. Madsen, H., Linde, J.J. and Arnbjerg-Nielsen, K.: *Opdaterede klimafaktorer og dimensionsgivende regnintensiteter* (Updated climate factors and design rain intensities) - Spildevandskomiteen, Skrift nr. 30. The Danish Water and Wastewater Committee under the Danish Engineering Society, Copenhagen, Denmark. In Danish. Accessed 21-12-2016 at https://ida.dk/sites/prod.ida.dk/files/svk_skrift30_0.pdf. -2014.

Kendon, E.J., Rowell, D.P., Jones, R.G. and Buonomo, E.: Robustness of future changes in local precipitation extremes. *Journal of Climate*, 21(17), 4280–4297. doi:10.1175/2008JCLI2082.1. 2008.

Kendon, E.J., Roberts, N.M., Fowler, H.J., Roberts, M.J., Chan, S.C. and Senior, C.A.: Heavier summer downpours with climate change revealed by weather forecast resolution model, *Nature Climate Change*, 4, 570–576. 2014.

Madsen, H., Mikkelsen, P.S., Rosbjerg, D., and Harremoes, P.: Estimation of regional intensity-duration-frequency curves for extreme precipitation. *Water Science and Technology*, 37(11), 29–36. doi:10.1016/s0273-1223(98)00313-8. 1998.

Madsen, H., Mikkelsen, P.S., Rosbjerg, D. and Harremoes, P.: Regional estimation of rainfall intensity-duration-frequency curves using generalized least squares regression of partial duration series statistics. *Water Resources Research*, 38(11), 21–1-21–11. doi:10.1029/2001wr001125. 2002.

Madsen, H., Arnbjerg-Nielsen, K. and Mikkelsen, P.S.: Update of regional intensity-duration-frequency curves in Denmark: Tendency towards increased storm intensities. *Atmospheric Research*, 92(3), 343-349. 2009.

Madsen, H., Gregersen, I.B., Rosbjerg, D. and Arnbjerg-Nielsen, K.: Regional frequency analysis of short duration rainfall extremes in Denmark from 1979 to 2012. In Review.

Mayer, S., Maule, C.F., Sobolowski, S., Christensen, O.B., Sørup, H.J.D., Sunyer, M., Arnbjerg-Nielsen, K., and Barstad, I.: Identifying added value in high-resolution climate simulations over Scandinavia, *Tellus A*, 67, 24941, doi:10.3402/tellusa.v67.24941. 2015.

Mikkelsen, P.S., Madsen, H., Arnbjerg-Nielsen, K., Jørgensen, H.K., Rosbjerg, D. and Harremoës, P.: A rationale for using local and regional point rainfall data for design and analysis of urban storm drainage systems. *Water Science and Technology*, 37(11), 7–14. doi: 10.1016/s0273-1223(98)00310-2. 1998.

Molnar, P., and Burlando, P.: Variability in the scale properties of high-resolution precipitation data in the Alpine climate of Switzerland, Water Resources Research, 44, W10404, doi:10.1029/2007wr006142. 2008.

Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., van Vuuren, D.P., Carter, T.R., Emori, S., Kainuma, M., Kram, T., Meehl, G.A., Mitchell, J.F.B., Nakicenovic, N., Riahi, K., Smith, S.J., Stouffer, R.J., Thomson,

A.M., Weyant, J.P. and Wilbanks, T.J.: The next generation of scenarios for climate change research and assessment. *Nature*, 463(7282), 747–756. doi:10.1038/nature08823. 2010.

Nakicenovic ,N., Alcamo, J., Davis, J., de Vries, B., Fenhann, J., Gaffin, S.,Gregory, K., Grübler, A., Jung, T.Y., Kram, T., Lebre La Rovere, E., Michaelis, L., Mori, S., Morita, T., Pepper, W., Pitcher, H., Price, L., Riahi, K., Roehrl, A., Rogner, H.-H., Sankovski, A., Schlesinger, M., Shukla, P., Smith, S., Swart, R., van Rooijen, S., Victor, N. and Dadi, Z.: *Special report on emission scenarios*. A special report of Working Group III for the Intergovernmental Panel on Climate Change. Cambridge University Press, New York. 2000.

Olesen, M., Madsen, K.S., Ludwigsen, C.A, Boberg, F., Christensen, T., Cappelen, J., Christensen, O.B., Andersen, K.K. and Christensen, J.H.: *Fremtidige klimaforandringer i Danmark* (Future climate changes in Denmark). Danmarks Klimacenter rapport nr. 6 2014. Danish Meteorological Institute, Copenhagen, Denmark. In Danish. Accessed 21-12-2016 at https://www.dmi.dk/fileadmin/user_upload/Rapporter/DKC/2014/Klimaforandringer_dmi.pdf. 2014.

Olsson, J. and Burlando, P.: Reproduction of temporal scaling by a rectangular pulses rainfall model, Hydrological Processes, 16, 611–630, doi:10.1002/hyp.307. 2002.

Olsson, J., Berggren, K., Olofsson, M., and Viklander, M.: Applying climate model precipitation scenarios for urban hydrological assessment: a case study in Kalmar City, Sweden, *Atmospheric Research*, 92, 364–375, doi:10.1016/j.atmosres.2009.01.015. 2009.

Scharling, M.: *Climate Grid Denmark – Dataset for use in research and education*, Technical Report no 12-10. Danish Meteorological Institute, Copenhagen, Denmark. 2012.

Schilling, W.: Rainfall data for urban hydrology: what do we need?, *Atmospheric Research*, 27, 5–22, doi:10.1016/0169-8095(91)90003-F. 1991.

Segond, M.-L., Onof, C. and Wheater, H.S.: Spatiat-temporal disaggregation of daily rainfall from a generalized linear model. *Journal of Hydrology*, 331(3-4), 674–689. doi:10.1016/j.jhydrol.2006.06.019. 2006.

Srikanthan, R. and McMahon, T. A.: Sequential generation of short time-interval rainfall data. *Nordic Hydrology*, 14(5), 277–306. 1983.

Sunyer, M.A., Madsen, H., Rosbjerg, D. and Arnbjerg-Nielsen, K.: A Bayesian Approach for Uncertainty Quantification of Extreme Precipitation Projections Including Climate Model Interdependency and Nonstationary Bias. *Journal of Climate*, 27(18), 7113–7132. doi:10.1175/JCLI-D-13-00589.1. 2014.

Svoboda, V., Hanel, M., Máca, P and Kyselý. J.: Projected changes of rainfall event characteristics for the Czech Republic. *Journal of Hydrology and Hydromechanics*. doi: 10.1515/johh-2016-0036. 2016.

Sørup, H.J.D., Madsen, H. and Arnbjerg-Nielsen, K.: Descriptive and predictive evaluation of high resolution Markov chain precipitation models. *Environmetrics*, 23(7), 623–635. doi:10.1002/env.2173. 2012.

Sørup, H.J.D., Christensen, O.B., Arnbjerg-Nielsen, K. and Mikkelsen, P.S.: Downscaling future precipitation extremes to urban hydrology scales using a spatio-temporal Neyman–Scott weather generator. *Hydrology and Earth System Sciences*, 20, 1387-1403. doi:10.5194/hess-20-1387-2016. 2016a.

Sørup, H.J.D., Lerer, S.M., Arnbjerg-Nielsen, K., Mikkelsen, P.S. and Rygaard, M.: Efficiency of stormwater control measures under varying rain conditions: Quantifying the Three Points Approach (3PA). *Environmental Science and Policy*, 63, 19-26. doi:10.1016/j.envsci.2016.05.010. 2016b.

Thyregod, P., Arnbjerg-Nielsen, K., Madsen, H. and Carstensen, N. J.: Modelling the embedded rainfall process using tipping bucket data. *Water Science and Technology*, 37(11), 57–64. doi:10.1016/S0273-1223(98)00316-3. 1998.

van der Linden, P. and Mitchell, J.F.: *Ensembles: Climate change and its impacts: Summary of research and results from the ensembles project*, Technical Report, Met Office Hadley Centre, Exeter, UK. 2009.

van Roosmalen, L., Sonnenborg, T. O., Jensen, K. H. and Christensen, J. H.: Comparison of Hydrological Simulations of Climate Change Using Perturbation of Observations and Distribution-Based Scaling. *Vadose Zone Journal*, 10(1), 136–150. doi:10.2136/vzj2010.0112. 2011.

Verhoest, N.E.C., Vandenberghe, S., Cabus, P., Onof, C., Meca-Figueras, T. and Jameleddine, S.: Are stochastic point rainfall models able to preserve extreme flood statistics? *Hydrological Processes*, 24(23), 3439–3445. doi:10.1002/hyp.7867. 2010.

Willems, P.: Stochastic generation of spatial rainfall for urban drainage areas. *Water Science and Technology*, 39(9), 23–30. doi:10.1016/s0273-1223(99)00212-7. 1999.

Willems, P., Arnbjerg-Nielsen, K., Olsson, J., and Nguyen, V.-T.-V.: Climate change impact assessment on urban rainfall extremes and urban drainage: methods and shortcomings, *Atmospheric Research*, 103, 106–118, doi:10.1016/j.atmosres.2011.04.003. 2012.

WMO: *Guide to hydrological practices. Volume II: Management of Water Recourses and Application of hydrological practices*. WMO report 168, 6th edn. World Meteorological Organization, Geneva, Switzerland, p. 302. 2009.

Yang, W., Andreasson, J., Graham, L.P., Olsson, J., Rosberg, J. and Wetterhall, F.: Distribution-based scaling to improve usability of regional climate model projections for hydrological climate change impacts studies. *Hydrology Research*, 40(1), 211–229. doi:10.2166/nh.2010.004. 2012.