Event-based stochastic point rainfall resampling for statistical replication and climate projection of historical rainfall series
Søren Thorndahl¹, Aske Korup Andersen², Anders Badsberg Larsen²
¹Department of Civil Engineering, Aalborg University, Aalborg, DK9220, Denmark
²Niras A/S, Aalborg, DK-9000, Denmark

Correspondence to: Søren Thorndahl (st@civil.aau.dk)

Abstract. Continuous and long rainfall series are a necessity in rural and urban hydrology for analysis and design purposes. Local historical point rainfall series often cover several decades which makes it possible to estimate rainfall means at different time scales, and to assess return periods of extreme events. Due to climate change, however, these series are most likely not representative for future rainfall. There is therefore a demand for climate projected long rainfall series, which can represent a specific region and rainfall pattern as well as fulfil requirements of long rainfall series which includes climate changes projected to a specific future period.

This paper presents a framework for resampling of historical point rainfall series in order to generate synthetic rainfall series, which has the same statistical properties as an original series. Using a number of key target predictions for the future climate, such as winter and summer precipitation, representation of extreme events, the resampled historical series are projected to represent rainfall properties in a future climate. Climate projected rainfall series are simulated by brute force randomization of model parameters which leads to a large number of projected series. In order to evaluate and select the rainfall series with matching statistical properties as the key target projections, an extensive evaluation procedure is developed.

1 Introduction

In design of new and analysis of existing storm water drainage systems valid rainfall statistics are crucial. With climate changes anticipated to impact precipitation patterns, the historical rainfall statistics upon which the traditional design is based, is no longer valid for future design. There is therefore a need for climate projection of the rainfall statistics in order for these to represent the future loads on storm water drainage systems.

Traditionally many simple urban drainage systems are designed with Intensity-Duration-Frequency (IDF) –relationships, or types of design storms (e.g. Unit Hydrograph: Sherman, 1932; Chicago Design Storm, CDS: Keifer and Chu, 1957; SCS: NRCS, 1986) which represent statistics for rain with specific return periods. Climate projection of these types of design methods can be relatively simple e.g. by multiplying the design rain by a bias climate factor (e.g. Semadeni-Davies et al., 2008; Olsson et al., 2009; Willems et al., 2012a; Willems, 2013; Shahabul Alam and Elshorbagy, 2015) assuming that extreme rainfall events for a specific return period will be increased linearly with a given factor as a function of time. The most recognized approaches for estimating climate factors is downscaling of Global Circulation Models (GCMs) and/or Regional Climate Models (RCMs), (e.g. Wilby and Wigley, 1997; Fowler et al., 2007).
In general, statistical downscaling determines a statistical relationship between a large- and a local-scale climate variable based on historical records. The relationship can be used on a GCM/RCM to obtain local variables for a specific domain in a given time frame of climate projection. (e.g. Wilby et al., 2002; Nguyen et al., 2007; Willems and Vrac, 2011; Willems et al., 2012b; Arnbjerg-Nielsen, 2012; Sunyer et al., 2015). The statistical downscaling approach requires long historical records of observations in order to establish the necessary statistical relationships. Based on various types of statistical downscaling assumptions and methods, climate factors for urban drainage design purposes (e.g for multiplication on IDF-relationships) can be derived by statistically comparing contemporary climate conditions with projected future rainfall with regards to specific return periods, and aggregation levels (durations) or rainfall (e.g. Mailhot et al., 2007; Larsen et al., 2009; Madsen et al., 2009; Nguyen et al., 2009; Nguyen et al., 2010; Willems and Vrac, 2011; Olsson et al., 2012; Willems, 2013).

Whereas a large proportion of the recent research described above has been conducted on estimating climate factors for design purposes, there is also a significant need - not only to describe future extremes e.g. in the form of IDF-relationships, but also to be able to project climate changes to continuous rainfall time series. Basically, simple design methods assume agreement between the return period of the rain intensity (for a given duration), and on the other hand the return period of the critical load in the drainage system (water level, flow, basin storage, etc.). Multiplying climate factors to simple design storms is sufficient for many applications of urban drainage design, however for more complex drainage systems with non-linear rainfall runoff response the simple design methods falls short. That is, for complex systems the return periods of the rainfall duration and intensity are not in agreement with the return periods of the corresponding drainage system state.

Therefore, historical rainfall series (or climate projected rainfall series) are required for complex systems, in order to estimate maximum water levels in manholes, flooding, to estimate the return periods, and other loads on the drainage system such as outlet to recipient, inlet to wastewater treatment plants, combined sewer overflow, outlet flow, pollutants loads, etc. in the future climate (e.g. Schaarup-Jensen et al., 2009; Thorndahl, 2009; Thorndahl et al., 2015).

According to Willems et al., (2012a, 2012b) there are generally two methods that produce continuous climate projected time series either by 1) Stochastic rainfall generators which generate locally representative synthetic rainfall conditioned on climate variables in present and future climate; or 2) Statistical approaches to downscaling such as change factor, resampling or weather typing -methods, in which future local rainfall is sought in historical rainfall records under equivalent historical climate conditions as projected in the future, or modified to represent future climate conditions.

In the literature, the most acknowledged methods for stochastically generating synthetic rainfall series are based on Poisson clusters processes and rectangular pulse models such as Bartlett-Lewis (Koutsoyiannis and Onof, 2001; Onof and Wheater, 1994, 1993; Segond et al., 2007; Onof and Arnbjerg-Nielsen, 2009; Paschalis et al. 2014; Kossieris et al. 2016) or Neyman-Scott (e.g. Entekhabi et al., 1989; Cowpertwait, 1991; Cowpertwait et al., 2002; Fowler et al., 2005; Burton et al., 2008; Cowpertwait, 2010; Paschalis et al., 2014; Sørup et al., 2016). Calibration of the generators is typically performed comparing generated series to observed series and adjusting relevant parameters prior to climate projection. Methods for estimating
point rainfall (e.g. Cowpertwait et al., 1996; Marani and Zanetti, 2007; Onof and Arnbjerg-Nielsen, 2009) and spatially distributed rainfall or multi-site generators with spatial dependency (e.g. Kilsby et al., 2007; Burton et al., 2008; Sørup et al., 2015) have been applied. These methods have been shown to provide valid results for hourly or daily time steps, but also have significant shortcomings in terms of modelling rainfall at a finer temporal resolution. For urban hydrological applications with fast rainfall-response, a temporal resolution of input data down to 1 to 10 minutes is required (e.g. Schilling, 1991; Willems, 2000; Thorndahl et al., 2008; Thorndahl et al., 2016a, 2016b). Since we are interested in maintaining the fine temporal resolution of observed rainfall series, generation of synthetic rainfall series using Poisson clusters is rejected here as an applicable method.

Change factor, resampling or weather typing -methods (Willems et al., 2012a, 2012b) of statistical downscaling outcomes of RCMs/GCMs, can provide data in the required temporal resolution, since directly based upon historical records. Arnbjerg-Nielsen (2012) applied historical rain series originating from another geographical region, which had a climate analogue to the projected climate in order to obtain continuous representative rainfall series for future climate conditions. Zorita and Von Storch (1999), Olsson et al. (2009), Willems and Vrac (2011), and Ntegeka et al. (2014) used historical records of rain and modified these records to represent climate-representative continuous climate projected rain series. Ntegeka et al. (2014) alternated the number of dry and wet days and used quantile perturbation (an advanced delta change method) to modify rainfall intensities. Olsson et al. (2009) applied the delta change method to multiply historical records with bias climate factors depending on rainfall intensity levels in order to fit projections of extreme, seasonal and annual precipitation. This approach, however, was implemented without alternating the temporal variability and the seasonal distribution of events of the rain series and maintaining the chronology of the original series. This particular shortcoming might be problematic in order to project the frequency of extreme events sufficiently.

The approach presented in this paper is different from the methods presented above, although it can be considered as a variation of resampling combined with stochastic generation. Whereas other methods use other climate variables, e.g. pressure, temperature etc. as climate predictors, this approach aims at fitting statistical properties of climate projected precipitation directly. In this case, these properties are derived from other studies of RCM projection (see section 2 for details). The validity of the method therefore depends on whether the climate projected target variables are comprehensive and detailed enough to project the future rainfall upon. The aim is to develop a generally adaptive method which can be applied to an arbitrary rainfall series and with different climate scenarios and projection period. On the contrary to studies described above, climate projected time series are generated directly for urban drainage modelling purposes. The objective has been to develop a generally applicable method that can be used directly by practitioners and scientist within the field of urban drainage, who do not necessarily have detailed knowledge on climate projection, RCM’s, downscaling, etc.

The procedure is divided in two major parts: 1) Resampling of a historical point rainfall time series (Method development: section 3.1 and Results and evaluation: 4.1); and 2) Climate projection of resampled time series (Method development: section 3.2 and Results and evaluation: 4.2);
The essential concept of the method is, to stochastically generate a large number of either resampled historical series or climate projected series, and to evaluate the statistical properties of the generated series against a number of key target variables. Rather than optimizing for the best parameter fit, the basic concept is to sample parameters from broad uniform distribution functions for each parameter and to either accept or reject each stochastically simulated series using a specified criteria. Repeating, this procedure for a large number of realizations of rainfall series, it is possible to select a number of rainfall series which has a satisfying statistical representativeness in comparison with historical series or climate projection targets. The evaluation procedure is inspired by the GLUE method (Beven and Binley 1992, Thorndahl et al. 2008) and is presented in detail in section 3.3.

The method assumptions and subjectivity is discussed in section 5 and in section 6 conclusions on this approach to climate projection of single point historical rainfall series are provided.

2 Data

The development of the model is based on rain gauge data from Denmark and projection of Danish climate conditions, but could easily be extended to other regions/countries of interest.

Specific statistical properties for the future precipitation in Denmark are necessary in order to climate project the resampled rainfall series. In Olesen et al. (2014) the Danish Meteorological Institute has collected and processed data from the ENSEMBLES project (http://www.ensembles-eu.org/, http://ensemblesr3.dmi.dk/, Van der Linden and Mitchell, 2009; Boberg et al., 2010; Maule et al. 2013). The report includes projection of weather extremes (including precipitation) using the SRES A1B-scenario (IPCC, 2007) and is produced from an ensemble of 14 regional climate models in the ENSEMBLES project. The RCM’s are simulated for 1961-1990, 2021-2050 and 2071-2100, but in this case only the first and last time interval are applied. Table 1 presents annual and seasonal precipitation increment (expressed as a climate change factor) in 2071-2100 compared to the reference in 1961-1990. Furthermore, the report specifies changes in other climate indices. In the context of precipitation, the variables number of events above 10 mm, number of events above 20 mm, and max. daily precipitation are relevant (table 1). In this paper these three variables are used to climate-project the resampled rainfall series, as they are considered important with regards to urban drainage modelling. Since the data from Olesen et al. (2014) represent the SRES scenarios (IPCC, 2007), new data representing the Representative Concentration Pathways, (RCP)-scenarios (IPCC, 2013, Christensen et al. 2015) is developed for this paper. Daily Regional climate model simulations from an ensemble of 14 models has been derived over Denmark from the Euro-CORDEX database (Casanueva et al., 2016; Jacob et al., 2014; Prein et al., 2016) and statistically processed by the same variables as in Olesen et al. (2014). Derived values are provided in table 1. For the climate projections in this paper the RCP4.5 scenario is chosen throughout the paper, but could easily also have been presented with other SRES or RCP scenarios.

The Water Pollution Committee of The Society of Danish Engineers has published a reports (guideline no. 29 and 30), with recommendations for design of drainage systems considering climate change (WPC, 2008, WPC, 2014, background report:
Based also on the climate simulations of the ENSEMBLES project, the climate factors for drainage system design in Denmark are recommended (table 2). Design rainfall, e.g. IDF-relationships, with a specified return period is recommended to be multiplied by these climate factors. The values are derived for rainfall intensities over 1 hour, but also recommended for other durations (up to 3 hours). In this paper these values are used to certify a correct representation of extreme events.

The rainfall series which are applied in this study has its origin in the rain gauge network of the Water Pollution Committee (WPC) of The Society of Danish Engineers. At present, the network consists of 145 tipping bucket rain gauges (DMI, 2014). The rain gauge no. 5047 located in Sulsted, North Jutland (lat: 57.17, long: 9.96) is applied since this is a station with a long recording time and few errors compared to other gauge records. The gauge has been in operation over a period of 34 years from 1979 to 2014, but due to minor interruptions in the dataset, the effective length of the series is 32 full years. The interruptions do not affect the statistical calculations as these are excluded from the data before the calculations are performed. The time series of 1 min. values for the Sulsted rain gauge is shown in figure 1.

In the WPC rain gauge network the temporal resolution of data is 1 minute. The start time of an event is determined at the minute of the first tip of 0.2 mm. All events therefore have initial values equivalent a multiple of 0.2 mm/min (12 mm/h). These initial values are easily identifiable in figure 1. The end of an event is specified when there is not registered a tip within an hour. Using this definition of events, the minimum inter-event time (time between events) will be 1 hour.

Using Danish rainfall data on a daily scale Gregersen et al. (2014) have been able to identify multidecadal climate oscillations (Ntegeka and Willems, 2008; Willems, 2013a) as well as climate related changes in precipitation patterns over the past 140 years. Nevertheless, since this paper is based on evidently shorter rainfall series, it is assumed that no significant trends or climate changes in this period are present. The historical records from the Sulsted series are therefore assumed to be stationary in terms of climate properties.

3 Method development

The procedure of the method is divided in two sections: the resampling of historical rainfall series (section 3.1) and the stochastic climate projection of resampled historical rainfall series (section 3.2). Since both methods involve random selection of events and brute force randomization of parameters there is a need for a unique method to evaluate the generated series against target values. This evaluation method is inspired by the GLUE methodology (Beven and Binley, 1992). The basic concept is to generate a large number of rainfall series and evaluate whether each generated series should be accepted or rejected based on an empirical likelihood (performance) measure based on individual criteria for each target value. For the accepted generated rainfall series a combined performance measure for each realization is calculated in order to find the rainfall series realization which in general fit the target values the best. This method is described in detail in section 3.3.
3.1 Historical rainfall series resampling

The objective is to create synthetic rainfall series resampled stochastically from a historical series such that the synthetic and the historical series have the same statistical properties. The first step is to divide the historical rainfall series into smaller parts in order to describe variability of intensities, event duration and time between events over the year. It is chosen to divide the series into four seasons (winter: DJF; spring: MAM; summer: JJA; autumn: SON), although a finer division (e.g. monthly) could have been implemented. Since the target projections (table 1) are implemented in seasons, this is the one used. The summer precipitation in the synthetic rainfall series is thus generated based on statistics calculated for every summer period’s precipitation in the historical rainfall series and correspondingly for the other seasons.

The stochastic generation (resampling) is based on:

1. Statistics of the inter-event time (also referred to as rainfall intermittency, e.g. by Molini et al., 2001 and Schleiss et al., 2011) using the definition of events presented in section 2.
2. Sampling of rainfall events including original event durations and intensities randomly from the pool of historical rain events for each season.

The concept is outlined in Figure 2.

The inter-event times ($t_{ie}$) for each season are approximated by a two-component mixed exponential probability density function:

$$f(t_{ie}) = p[\lambda_{a,ie}\exp(\lambda_{a,ie}t_{ie})] + (1 - p)[\lambda_{b,ie}\exp(\lambda_{b,ie}t_{ie})]$$

(1)

Where $\lambda_{a,ie}$ and $\lambda_{b,ie}$ are the rate parameters for two populations, $a$ and $b$, with different exponential distributions and $p$ is the weight of population $a$. This mixed distribution function was also applied by Rossi et al. (1984) and Willems (2000).

Willems (2000) applied the distribution for fitting rainfall intensities arguing that the two distributions originated from two different types of storms (convective thunder storms and frontal storms respectively). The same rationale is applied here. The approximation to inter-event times for each season thus require approximation of three parameters, $p$, $\lambda_{a,ie}$, and $\lambda_{b,ie}$.

Molini et al. (2001) applied a Weibull distribution to describe the inter-event time of rainfall events. Weibull distributions along with exponential, gamma and generalized pareto distributions was also investigated for this paper, but was however outperformed by the mixed exponential distribution especially in fitting both ends of the distribution.

As opposed to other rainfall generators which use a fixed time scale, (e.g. Furrer and Katz, 2008), the time is sampled discontinuously in this case.

The sampling of the events is an automated process with random selection of events from the pool of historical rainfall events for each season. When sampling a specific event, the intensity sequence and consequently also the duration is maintained. Synthetic resampled time series are, therefore produced by random alternating sampling of the inter-event times and historical events from a specific season. It is possible to sample the same event more than once. The procedure is repeated until the length of the generated series corresponds to the length of the historical series or any other specified length.
shorter than the total length of the original series. The number and the chronology of events are therefore different from season to season and from year to year.

A vital assumption here is that events from the historical series can be sampled independently. Depending on the meteorological conditions at the time of a specific event there might potentially, be some correlation to prior and posterior events due to short inter-event times. Extreme event statistics and development of IDF-relationships from partial duration series, in Denmark is also produced assuming independent events (Mikkelsen et al., 1998; Madsen et al., 2009) so in order to preserve this methodology, no inter-correlation between events has been implemented in the presented approach.

3.2 Climate projection and stochastic resampling of rainfall series

The climate projected rainfall series is generated in three steps:

1. The inter-event time for each season is sampled using the same procedure as described in the previous section, however, the parameters of the mixed exponential distribution for each season is implemented as stochastic variables and thus sampled randomly from a uniform distribution with fixed upper and lower boundaries. This allows for different distributions of inter-event times than the ones used in the resampling of historical series. In the climate projected series, it is thus possible to accommodate for climate changes in seasonal precipitation and the distribution between small and large events, by changing the number of events per season. As an example the method is able to accommodate a moderate increase of total summer precipitation, and at the same time a considerable increase in frequency and intensity of extreme events, with generally a lower number of total events in summer as a result.

2. Rainfall events are sampled from the pool of historical events for each season in the same way as described in section 3.1. The duration of each event is not alternated under impact of climate change, since there is presently no evidence that single events will become shorter or longer in the future. This is obviously a crucial assumption, but nonetheless the best current estimate, which also has been applied by e.g. Olsson et al. (2009). The sampling of events is therefore done without alternating the events from the pool, other than multiplying by different change factors as presented below.

3. The climate projection of the generated time series is inspired by the delta change method. However unlike Olsson, the change factors are implemented as random variables. The change factor for a given rainfall intensity, $i$, is derived using the probability, $F(i)$, of that the intensity being less than or equal to $i$. For each season, the rainfall intensities from the original historical rain series are fitted to the same type of mixed exponential distribution (Willems, 2000a) as applied for fitting the inter-event times (eq. 1):

$$F(i) = p[1 - \exp(\lambda_a i)] + (1 - p)[1 - \exp(\lambda_b i)]$$

Where $\lambda_a$ and $\lambda_b$ are rate parameters for two populations, $a$ and $b$ and $p$ is the weight given to population $a$. $F(i)$ has a range from 0 to 1.
For each season change factors are multiplied to intensities on the minute scale. The change factor as a function of intensity, \( c(i) \), is thus calculated for each season by a linear function:

\[
    c(i) = aF(i) + \beta
\]

Where \( a \) and \( \beta \) are random variables samples from uniform distributions with fixed limits.

For each projected rainfall series there is a different value of \( a \) and \( \beta \) for each season. During the development of the procedure, the limits of the uniform distribution of \( a \) and \( \beta \) for each season was empirically selected starting with broad intervals which was reduced by discarding non-accepted runs (see below).

The total number of random variables for generating climate projected stochastic rain series in the current setup with four yearly seasons is: 20 (2x4 for the change factor plus 3x4 for the mixed exponential distributions).

### 3.3 Evaluation and optimization procedure

The governing assumption behind the resampling procedure is that the resampled rainfall series should have the equivalent statistical characteristics as the historical series on a number of key target variables. The climate projected resampled series should therefore also have the equivalent statistical characteristics by means of a number of key target climate projections (as the ones presented in tables 1-2). It is not a necessity that the same target variables are used to evaluate resampled historical rainfall series and the climate projected series, but it is chosen to do so in this paper, in order to keep the evaluation procedures the same regardless of generating series which should statistically represent historical series or climate projected series. The key target variables are described in detail below:

1. **Annual precipitation (ap).** This target variable is included as it is a measure of the total “mass”-balance. Since the individual years of the resampled and historical series are not directly comparable year by year, the mean of all years is applied as target variable.

2. **Seasonal precipitation (sp).** The mean seasonal precipitation is applied as a target variable in order to ensure same distribution between seasons in the resampled series. The four target parameters are labeled \( spwi, spsp, spsu, spau \) corresponding to winter, spring, summer, and autumn precipitation respectively.

3. **Number of events above 10 mm per day (n10mm).** This target variable provides a measure of the representation of extreme events.

4. **Number of events above 20 mm per day (n20mm).** Same as no 3.

5. **Maximum daily precipitation (mdp, as a mean of the maximum day for all years).** This target variable also certifies the representation of extreme events.

6. **IDF-relationships.** The intensity-duration-frequency relationships are traditionally applied in design of urban drainage systems and are therefore relevant to include as a target variable. In accordance with table 2, it is chosen to use the mean rain intensity over a duration of 60 minutes for return periods of 2 and 10 years respectively as a target value. The two values are labeled \( d60T2 \) and \( d60T10 \) respectively.
The performance of each individual target variable is estimated using a simple ratio measure between the target value and the corresponding modeled value:

$$P_{i,j} = 1 - \frac{|T_i - M_{i,j}|}{T_i}$$  \hspace{1cm} (4)

$P_{i,j}$ is the individual performance parameter for target variable $i$ (as presented above) corresponding to realization $j$; $T_i$ is the target value; and $M_{i,j}$ is the modeled value of the target variable of the $j$’th realization. For the evaluation of the resampled series against the historical series, $T_i = H_i$, where $H_i$ is the value of the target variable of the historical series. With respect to the evaluation of the climate projected rainfall series, where the target value is given by a climate factor ($cf$) multiplied by the target variable of the historical series:

$$T_i = cf_i \cdot H_i$$  \hspace{1cm} (5)

thus the performance measure is:

$$P_{i,j} = 1 - \frac{|cf_i \cdot H_i - M_{i,j}|}{cf_i \cdot H_i}$$  \hspace{1cm} (6)

$P$ can vary between 0 and 1, where $P=1$ corresponds to a perfect fit.

In order for a simulated rainfall series to be accepted $P_{i,j}$ has to be larger than a specified threshold. For the resampled historical the acceptance criteria for the individual performance measures is fixed and have been chosen to $P_{crit,i} = 0.90$, hence all ten individual performance measures should exceed this value in order for the realization to be accepted (table 4). This means, that if a target value of just one of the 10 target values deviates more than 10 % from the value of the historical series, the realization is rejected.

For the climate projected series, it is possible to estimate individual values of the performance using the standard deviations of the climate factors ($cf$) given in tables 1 and 2:

$$P_{crit,i} = 1 - \frac{2\sigma_{cf_i}}{cf_i}$$  \hspace{1cm} (7)

Assuming Gaussian distributed target variables, we will thus accept values which are within the 95 % confidence intervals of the distribution of each target variable. The acceptance criteria of the performance measure will thus be different for each target variable depending on the uncertainty (standard deviation) related to that specific climate projection (cf. table 1 and 2). The acceptance criteria for the performance of each target variable are presented in table 6 along with climate factors and standard deviation for each variable.

The combined performance measure $P_j$ of each realization series ($j$) is estimated as:

$$P_j = \sum_{i=1}^{I} w_i P_{i,j}$$  \hspace{1cm} (8)

where $w_i$ is the weights of the individual performance measures, $\sum w_i = 1$, and $I$ is the total numbers of individual performance parameters.
The individual weights are presented in section 4.2 and table 6. One could argue that each season should be given the same weight, however since summer precipitation tends to be more important in terms of extreme events in Denmark this is given a higher weight. Moreover, since winter precipitation might be associated with larger measurement errors due to poor measurement of solid precipitation, this is given a smaller weight.

5 Results and evaluation

4.1 Historical rainfall series resampler

The synthetic resampled series are generated with the same total length as the original historical series. In this case 32 years. The inter-event times for each season are sampled from the mixed exponential distribution as detailed in section 3.1. The estimated parameters are presented in table 3. By comparing the parameters, it is evident that there is a significant difference for each season. Therefore, it is important that the inter-event times are sampled individually for each season to ensure a representative number of events in the resampled rainfall series compared to the historical rainfall series. Figure 3 exemplifies empirical cumulative distribution functions for summer inter-events times for the historical series and for the fitted mixed exponential distribution of summer inter-event times. Furthermore, the empirical distribution from the resampled series with the best combined performance measure is presented ($P_j=0.98$). Using the mixed exponential distribution, there is small underestimation of inter event times between 1 and 6 hours and an equivalent overestimation between 6 and 24 hours. This is however insignificant in comparison to other fitted distribution functions and thus not considered a problem in random sampling of inter-event times from these distributions.

There is a stringent dependency between inter-event times and number of events in the rainfall series. In order to generate a valid and representational resampled rain series, the number of events series should correspond somewhat to the number of events in the historical rainfall series. Table 4 therefore includes the mean and standard deviation of the number of events per year even though the number of events is not used a target variable for estimating the individual performances.

The resampling of the observed rainfall series is performed generating 5,000 different resampled rainfall series and assessing the performance of each generated series using the method described in section 3.3. Out of the 5,000 realizations of simulated series, 275 (5.5 %) are accepted using the criteria of a minimum individual performance measure ($P_{\text{crit},i}$) of 0.90. The fact that all 10 individual performance measures have to be larger than the acceptance criteria have shown to be a tough criterion to fulfill. Often one or two of the 10 has a slightly lower value and the realization is thus rejected. On average the accepted realizations has a combined performance measure ($P_j$) of 0.95 (ranging between 0.92 and 0.98). Figure 4 presents a bar plot (blue shades) of each of the target variables for the historical series, the one resampled series with the highest combined performance measure, as well as the mean of the accepted resampled series (with uncertainty bounds indicating the minimum and maximum of the accepted series).

Generally there is a good agreement between the historical series and the accepted series on the target parameters with the highest weights, i.e. the seasonal precipitation. This is actually the case for the majority of the 5,000 realizations, however
the performance measures becomes rather low if the extreme events are not represented correctly in the resampled series and is in that case rejected. The variability between the resampled series is only due to the randomness assembling events and \textit{inter-event times} from the historical series since the mixed exponential parameters for each season are fixed corresponding to the fits (table 3). The rejection of resampled series is therefore often due to either sampling of too few or too many “extreme” events within a season.

In many situations, only the one resampled series with the highest performance measure is of interest. Table 4, therefore, lists target values of the historical series and the resampled series with the highest performance measure (best fit). Besides the best combined performance measure of $P_{j}=0.98$, the individual performance measures are given in the right column. In order not only to compare series on mean values, table 4 also presents standard deviations describing the year to year variability over the total length of the series. Generally there is a satisfactory agreement (below 10%) on both mean and standard deviations between the historical series and the “best” accepted resampled rainfall series.

To verify the representativeness of extreme rainfall, figure 5 (left) presents $IDF$-relationships (from 10 min to 360 min durations) for the historical and “best” resampled series for return periods of 2 and 10 years respectively. Grey areas represent the variability in all the accepted realizations. Generally, there is a acceptable agreement between the curves which verifies the resampling method. There is however a minor divergence for short durations of the 10-year return period. In general the longer the return period the larger divergence between the curves must be expected as a result of the random sampling of historical events in the generated series. Figure 6 shows the time series of the “best fit” resampled time series.

The overall assessment of the previous evaluation indicates that the rainfall resampler can represent the historical rainfall series well based on the selected performance parameters. Due to the stochasticity of the sampling of \textit{inter-event times} and rainfall events, there is obviously some variability from year to year and from series to series, but since none of the target variables are significantly biased, the overall performance of the resampler is accepted. Since it is possible to produce resampled rainfall series with the same statistics as the corresponding original historical series, the resampling algorithm will be applied to generate climate projected rainfall series in the following section.

\textbf{4.2 Climate projected rainfall series}

Figure 4 and table 6 provides results for the climate projected rainfall series. The target variables (\textit{Climate projected historical}) are estimated using eq. 5 and is thus the mean values of the historical series of table 4 multiplied by the climate factors specified in table 6. In addition figure 4 provides an uncertainty estimate on the target values obtained from the standard deviations of table 1 and 2.

Since the climate projection of rainfall series involves randomization of not only the event assembling, but also randomization of \textit{mixed exponential distribution} parameters and change factors as function of intensity for each season, the generation of rainfall series requires a larger quantity of realizations compared to the resampling of series described in the previous section. Therefore a total of 10,000 climate projected rainfall series are generated. The acceptance criteria implemented are however slightly different compared to the ones detailed in section 4.1. In the evaluation of climate
projected rainfall series an individual acceptance criterion for each target variable is estimated using eq. 7. For the ten target variables the acceptance criteria range between 0.59 \((n20mm)\) and 0.89 \((spwi)\) as presented in table 6. The total number of accepted realizations is 721 (7.2 %). The reason that a larger percentage is accepted here than in the previous section is that the acceptance criteria is somewhat softer encountering the uncertainty of climate factors. On average the accepted realizations has a combined performance measure \(P_j\) of 0.90 (ranging between 0.81 and 0.97).

Table 5 presents the range of mixed exponential distribution parameters as well as ranges of change factor parameters for the accepted climate projected realizations for each season. Comparing to table 3 (in which the parameter assessment is based on fitting the historical data) it is clear that the parameter values obtained by random sampling have a broader range, indicating that an accepted realization with a high performance value can be obtained from a broad range of parameter values. Scatter plotting the performance values as a function of parameter values (not shown) shows flat tops indicating that an equal performance can be obtained equally from low and high values within the range (uniform distribution). This means that there is a dependency between inter-event time parameters and chance factor parameters.

As seen in table 5, the change factor is allowed to be both smaller and larger than 1. This allows for both decrease and increase in precipitation amounts in each seasons. The climate projected precipitation can thus be obtained from an insignificant change in seasonal precipitation, but a rather large increase in extreme precipitation.

Generally there is an acceptable agreement of climate projected target variables \((Climate\ projected\ historical)\) and corresponding values for the climate projected resampled series (red shades in figure 4). There is, however, slightly more deviation compared to the present-time simulations and larger ranges of target variable values. This is as expected since the climate projection includes more random parameters and complexity as well as a broader acceptance criteria. For the accepted realization with the highest performance measure, \(P=0.97\) (figure 4 and table 6) there is a tendency for the target variables related to the extreme values to be marginally underestimated. This is inevitably a result of high weights given to the target values related to seasonal precipitation, especially summer precipitation. By changing the weights it would be possible to obtain more equal extreme values, however potentially at the expense of a poorer fit of the accumulated precipitation values.

In figure 5 (right) the \(IDF\)-curves for the climate projected series are shown. There is a slight underestimation of extremes for the 10-year return period, however an overestimation of the 2-year return periods on low durations. Since the total length of the series is 32 years, return periods larger than 10 years are not presented well, since they the associated with large uncertainties (see e.g. Thorndahl, 2009). The uncertainty bands (grey areas) however cover the climate projected intensities. Figure 7 shows the time series of the “best fit” resampled time series.

The overall performance of the climate projection of resampled rainfall series is considered to be acceptable within the range of uncertainties related to the climate projections. The introduction of 20 random variables and the random assembling of rain events, obviously require many realizations in order to produce accepted rainfall series which has a satisfactory degree of agreement on all target parameters.
5 Discussion

The developed procedure obviously involves a large degree of subjectivity in the choice of processes and parameters to include. This section will discuss and argue for some of these choices.

The target variables have been chosen to represent both annual and seasonal precipitation as well as more extreme values. The choice of the ten specific target variables is closely connected to the fact that this is what is currently available for Danish future climate conditions. However, when other, and maybe more detailed target variables becomes available, it would be possible to redo the generation of climate projected rainfall series with other target variables. It was initially decided only to present values from the RCP4.5 climate scenario; however the implementation of the method could just as well have been implemented with another RCP or SRES scenario. Another possibility could be to implement other durations and return periods than for 60 minute durations for 2 and 10 years respectively, in order to emphasize specific extremes further.

It is of utmost importance that the chosen target variables are representative for the future precipitation patterns and that they are comprehensive in that way that they cover both annual/seasonal variations as well as single events and the statistics related to these. In this paper, it is chosen only to include mainly yearly mean values of target parameters (except the target variables related to return periods), but it could also be relevant to apply the year to year variability as a target in itself, in order to certify the correct representativeness of the resampled series in comparison with the original historical series.

The weights applied in estimating the overall performance of resampled series are chosen in order to emphasize the accumulated precipitation values, but on the other hand not neglect the extremes. Other weights could have been applied. One could imagine that the weights were chosen according to the purpose of use of the resampled and climate projected series. If, for example, the series should be used as an input to an urban drainage model simulating overflow from combined sewer systems to a recipient, it would probably be most important to have a good representation of the precipitation (event) totals. On the other hand if the purpose was simulating surcharge or flooding of a drainage system, the representation of extremes would be more important.

In the present approach a linear function and the probability of a given rainfall intensity for a given season is applied to derive the change factor as a function of intensity. The choice of parameters allows change factors to be both smaller and larger than 1. This might entail that the lowest fraction of intensities is allowed to be smaller in a future climate while the highest fraction of intensities will increase. Other continuous functions, rather than the applied linear function, might be an objective of future studies.

The proposed method applies two major assumptions which are relevant to discuss here. The first assumption is that events are sampled independently for each season. With inter-event times down to 1 hour, this might constitute a problem in hydrological applications where the response time of the system in question is larger than 1 hour. Hence, coupled events might impact the hydrological system response. The second assumption is that the duration of events does not change under
changed climate signals. It has presently not been possible to find evidence for this contention in the scientific literature on climate change. Both of the assumptions are subject to further investigations.

6 Conclusion
This paper presented a procedure to generate both statistically representative resampled rainfall series from original historical rainfall series as well as climate projected rainfall series, which includes the advantages in local historical rainfall series as well as projections on changes in rain patterns in the future climate.
The simulated rainfall series can represent the climate projected target variables and it is shown possible to produce rainfall series which does not only project accumulated seasonal precipitation, but also extremes in correspondence with the climate projection of the RCP4.5 scenario. The procedure is generic so if other climate scenarios and potentially other target variables for further precipitation patterns are available, the method will be able to adapt to these as well.
The procedure for generating resampled and climate projected rainfall series fulfills a need for having local representative rainfall series which are valid both for the present and future climate. The series can be applied directly as inputs to urban drainage models in order to analyze the loads on a drainage system, e.g. combined sewer overflow, surcharge, storage filling, flooding in the present and future climate.

7 Acknowledgements
The authors would like to acknowledge Cathrine Fox Maule at the Danish Meteorological Institute for providing RCM data from EURO-CORDEX database.

References


Molini, A., La Barbera, P., Lanza, L. G. and Stagi, L.: Rainfall intermittency and the sampling error of tipping-bucket rain gauges, Physics and Chemistry of the Earth, Part C: Solar, Terrestrial & Planetary Science, 26(10), 737–742,


Thorndahl, S., Schaarup-Jensen, K. and Rasmussen, M. R.: On hydraulic and pollution effects of converting combined sewer
WPC: Forventede ændringer i ekstremregn som følge af klimaændringer, Skrift nr. 29 (Anticipated changes in extrem precipitation as a result of climate change, Guideline no. 29), The Water Pollution Committee of the Society of Danish Engineers, Copenhagen, Denmark (in Danish), 2008.
WPC: Opdaterede klimafaktorer og dimensionsgivende regnintensiteter, Skrift nr. 30 (Updated climate factors and rain intensities for design, Guideline no. 30), The Water Pollution Committee of the Society of Danish Engineers, Copenhagen, Denmark (in Danish), 2014.
Figure 1: Measured time series of the Sulsted rain gauge. The temporal resolution of rainfall data is 1 min.

Figure 2: Sketch of the construction of the synthetic (resampled) rainfall series.
Figure 3: Example of cumulative distribution functions for summer inter-event times
Figure 4: Target variables and their values for comparing historical series and resampled series (in blue shades) and climate projected historical series and climate projected and resampled series (in red shades). For the climate projected target (deep red) the uncertainty bounds (black lines) represent two times the standard deviation of tables 1 and 2. For the resampled series the uncertainty bounds reprint the total range of the accepted realizations.
Figure 5: *IDF*-curves for historical and resampled rainfall series (left) and climate projected historical and resampled (right).

Figure 6: Time series example of resampled rainfall series. The temporal resolution of rainfall data is 1 min.
Figure 7: Time series example of climate projected rainfall series. The temporal resolution of rainfall data is 1 min.

Table 1: The calculated Danish climate changes in annual and seasonal precipitation as well as extremes. The values are expressed as a multiplicative climate factor describing the difference from the reference period 1961-1990 to 2071-2100. The A1B scenario is presented in (Olesen et al., 2014) and represents 14 regional climate model runs from the ENSEMBLES project. The climate factors from the two RCP scenarios are previously unpublished, but derived from the Euro-CORDEX-11 database (Jacob et al., 2014) and processed statistically for this paper. The standard deviation is listed in parenthesis. The indices marked with bold are the ones used in this paper.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual precipitation</td>
<td>1.14 (± 0.06)</td>
<td>1.08 (± 0.06)</td>
<td>1.14 (± 0.07)</td>
</tr>
<tr>
<td>Winter precipitation (DJF)</td>
<td>1.25 (± 0.06)</td>
<td>1.12 (± 0.06)</td>
<td>1.24 (± 0.07)</td>
</tr>
<tr>
<td>Spring precipitation (MAM)</td>
<td>1.13 (± 0.06)</td>
<td>1.13 (± 0.08)</td>
<td>1.23 (± 0.11)</td>
</tr>
<tr>
<td>Summer precipitation (JJA)</td>
<td>1.05 (± 0.08)</td>
<td>1.06 (± 0.18)</td>
<td>1.03 (± 0.21)</td>
</tr>
<tr>
<td>Fall precipitation (SON)</td>
<td>1.13 (± 0.06)</td>
<td>1.05 (± 0.07)</td>
<td>1.09 (± 0.13)</td>
</tr>
<tr>
<td>Events above 10 mm (days/year)</td>
<td>1.37 (±0.12)</td>
<td>1.20 (± 0.13)</td>
<td>1.35 (± 0.14)</td>
</tr>
<tr>
<td>Events above 20 mm (days/year)</td>
<td>2.50 (±0.14)</td>
<td>1.41 (± 0.30)</td>
<td>1.80 (± 0.40)</td>
</tr>
<tr>
<td>Max. daily precipitation (mm)</td>
<td>1.16 (±0.12)</td>
<td>1.12 (± 0.09)</td>
<td>1.24 (± 0.11)</td>
</tr>
</tbody>
</table>
Table 2: Recommended climate factors for design of drainage systems in Denmark according to WPC (2008), WPC (2014), and Gregersen et al. (2014). The climate factors are valid for a duration of 1 hour, but also recommended for other durations up to 3 hours. The indices marked with bold are the ones used in this paper. The standard deviations are not provided directly in the references, but estimated from tables and figures.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.20 (±0.1)</td>
<td>1.20 (±0.1)</td>
<td>1.45 (±0.1)</td>
</tr>
<tr>
<td>10</td>
<td>1.30 (±0.2)</td>
<td>1.30 (±0.2)</td>
<td>1.70 (±0.2)</td>
</tr>
<tr>
<td>100</td>
<td>1.40 (±0.3)</td>
<td>1.40 (±0.3)</td>
<td>2.00 (±0.3)</td>
</tr>
</tbody>
</table>

Table 3. The fitted rate and weight parameters for the mixed exponential distribution specified for each season.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Autumn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate, population $a$, $\lambda_{a,ie}$ (days)</td>
<td>0.38</td>
<td>0.33</td>
<td>0.24</td>
<td>0.26</td>
</tr>
<tr>
<td>Rate, population $b$, $\lambda_{b,ie}$ (days)</td>
<td>4.87</td>
<td>4.46</td>
<td>3.00</td>
<td>2.90</td>
</tr>
<tr>
<td>Weight population $a$, $p$ (-)</td>
<td>0.69</td>
<td>0.56</td>
<td>0.55</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Table 4. Target variables (mean and standard deviation) and performance measures for the historical series and the one resampled series with the highest performance measure.

<table>
<thead>
<tr>
<th>Target variable</th>
<th>Acceptance criteria and weights</th>
<th>Historical series (target)</th>
<th>“Best fit” resampled series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual no. of events</td>
<td>$P_{crit, i}$, $w_i$</td>
<td>Mean Std. dev.</td>
<td>Mean Std. dev.</td>
</tr>
<tr>
<td>Annual precipitation</td>
<td>$ap$ (mm)</td>
<td>200.1 39.4</td>
<td>218.2 45.7</td>
</tr>
<tr>
<td>Seasonal precipitation, winter</td>
<td>$spwi$ (mm)</td>
<td>0.90 576.1 122.3</td>
<td>586.5 140.6 0.96</td>
</tr>
<tr>
<td>Seasonal precipitation, spring</td>
<td>$spsp$ (mm)</td>
<td>0.90 0.05 90.9 36.5</td>
<td>101.6 26.7 0.99</td>
</tr>
<tr>
<td>Seasonal precipitation, summer</td>
<td>$spsu$ (mm)</td>
<td>0.90 0.10 86.5 43.1</td>
<td>84.9 29.3 0.98</td>
</tr>
<tr>
<td>Seasonal precipitation, autumn</td>
<td>$spau$ (mm)</td>
<td>0.90 0.25 213.0 57.0</td>
<td>209.1 75.3 0.98</td>
</tr>
<tr>
<td>Annual number of events above 10 mm per day</td>
<td>$n10mm$ (%)</td>
<td>0.90 0.17 16.0 4.5</td>
<td>15.8 5.4 0.99</td>
</tr>
<tr>
<td>Annual number of events above 20 mm per day</td>
<td>$n20mm$ (%)</td>
<td>0.90 0.08 3.3 2.1</td>
<td>3.3 2.3 0.99</td>
</tr>
<tr>
<td>Annual Maximum daily precipitation</td>
<td>$mdp$ (mm)</td>
<td>0.90 0.08 35.2 12.7</td>
<td>32.0 12.0 0.99</td>
</tr>
<tr>
<td>Rain intensity for 60 min, T=2 years</td>
<td>$d60T2$ (mm/h)</td>
<td>0.90 0.08 15.7</td>
<td>16.4 0.95</td>
</tr>
<tr>
<td>Rain intensity for 60 min, T=10 years</td>
<td>$d60T10$ (mm/h)</td>
<td>0.90 0.08 28.4</td>
<td>28.7 0.99</td>
</tr>
<tr>
<td>Combined performance measure</td>
<td>$P$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Ranges of accepted parameter values for the mixed exponential distribution applied to sampling inter-event times and for the linear function applied to sample change factors for each season.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Autumn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate, population $a$, $\lambda_{a,ie}$ (days)</td>
<td>0.32 0.44</td>
<td>0.27 0.39</td>
<td>0.20 0.28</td>
<td>0.23 0.29</td>
</tr>
<tr>
<td>Rate, population $b$, $\lambda_{b,ie}$ (days)</td>
<td>4.10 5.60</td>
<td>3.90 5.00</td>
<td>2.70 3.30</td>
<td>2.60 3.20</td>
</tr>
<tr>
<td>Weight population $a$, $p$ (-)</td>
<td>0.63 0.74</td>
<td>0.50 0.61</td>
<td>0.51 0.60</td>
<td>0.60 0.68</td>
</tr>
<tr>
<td>Change factor slope, $\alpha$ (-)</td>
<td>0.000 0.050</td>
<td>0.000 0.050</td>
<td>0.000 0.025</td>
<td>0.000 0.049</td>
</tr>
<tr>
<td>Change factor intercept, $\beta$ (-)</td>
<td>0.80 1.20</td>
<td>0.80 1.20</td>
<td>0.81 1.20</td>
<td>0.86 1.20</td>
</tr>
</tbody>
</table>
Table 6 Climate factors of target variables and minimum acceptance criteria of the individual performance parameters $P_{ij}$: empirical combined performance measure weights ($w_i$); Climate projected target variables and the corresponding values (± standard deviation) of the best fit climate projected series.

<table>
<thead>
<tr>
<th>Target variable</th>
<th>Climate factors</th>
<th>Acceptance criteria and weights</th>
<th>Climate proj. hist. series (target)</th>
<th>“Best fit” climate proj. series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual no. of events</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual precipitation</td>
<td>$ap$ (mm)</td>
<td>1.08 (± 0.06)</td>
<td>0.89</td>
<td>599.6</td>
</tr>
<tr>
<td>Seasonal precipitation, winter</td>
<td>$spwi$ (mm)</td>
<td>1.12 (± 0.06)</td>
<td>0.89 0.05</td>
<td>101.8</td>
</tr>
<tr>
<td>Seasonal precipitation, spring</td>
<td>$spsp$ (mm)</td>
<td>1.13 (± 0.08)</td>
<td>0.86 0.10</td>
<td>97.7</td>
</tr>
<tr>
<td>Seasonal precipitation, summer</td>
<td>$spsu$ (mm)</td>
<td>1.06 (± 0.18)</td>
<td>0.66 0.25</td>
<td>225.8</td>
</tr>
<tr>
<td>Seasonal precipitation, autumn</td>
<td>$spau$ (mm)</td>
<td>1.05 (± 0.07)</td>
<td>0.87 0.10</td>
<td>195.0</td>
</tr>
<tr>
<td>Annual number of events above 10 mm per day</td>
<td>$n_{10mm}$ (#)</td>
<td>1.20 (± 0.13)</td>
<td>0.78 0.17</td>
<td>19.2</td>
</tr>
<tr>
<td>Annual number of events above 20 mm per day</td>
<td>$n_{20mm}$ (#)</td>
<td>1.41 (± 0.30)</td>
<td>0.57 0.08</td>
<td>4.7</td>
</tr>
<tr>
<td>Annual Maximum daily precipitation</td>
<td>$mdp$ (mm)</td>
<td>1.12 (± 0.09)</td>
<td>0.84 0.08</td>
<td>39.4</td>
</tr>
<tr>
<td>Rain intensity for 60 min, T=2 years</td>
<td>$d_{60T2}$ (mm/h)</td>
<td>1.20 (±0.10)</td>
<td>0.83 0.08</td>
<td>18.8</td>
</tr>
<tr>
<td>Rain intensity for 60 min, T=10 years</td>
<td>$d_{60T10}$ (mm/h)</td>
<td>1.30 (±0.20)</td>
<td>0.69 0.08</td>
<td>36.9</td>
</tr>
<tr>
<td>Combined performance measure</td>
<td>$P$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>