Estimation of climate factors for future extreme rainfall:
Comparing observations and RCM simulations

I.B. Gregersen1*, H. Madsen2 and K. Arnbjerg-Nielsen1

1Department of Environmental Engineering, Technical University of Denmark, Miljoevej, Bldg 113, DK-2800 Kgs. Lyngby, Denmark
2DHI Water • Environment • Health, DK-2970 Hørsholm
*Corresponding author (idbg@env.dtu.dk)

ABSTRACT
The application of climate factors has become more common in urban drainage design. The climate factor accounts for the expected increase in the magnitude of the extreme rainfall events during the technical lifetime of the drainage system. The present practice in Denmark is the application of climate factors of 1.2, 1.3 and 1.4 for 2-, 10- and 100-year events, respectively. These estimates are based on a comparison between the ‘control’ and ‘scenario’ period simulated by the regional climate model, HIRHAM4. The present paper presents new estimates of the climate factors on the basis of two new sources of data, focusing on a rainfall duration of 60 minutes. Temporal development of observed extreme rainfall is described by a parametric model and subsequently extrapolated, resulting in estimates of climate factors between 2.3 and 2.5. A recent transient model simulation by the regional climate model RACMO is divided into five time slices whereby their independent extreme rainfall characteristics are compared. The climate factors vary between 1.2 and 2, depending on the return period.

KEYWORDS
Climate factor, extreme rainfall, non-stationarity, regional climate models, uncertainty

INTRODUCTION
Due to the increased attention towards climate change and its potential effects on the hydrological cycle, the application of climate factors has become more common in urban drainage design. The climate factor accounts for the expected increase in the magnitude of the extreme rainfall events during the technical lifetime of the drainage system and is defined as the ratio between the best estimate of the design intensity in the future and the design intensity at present. Traditionally climate factors are found from regional climate model (RCM) simulations (Arnbjerg-Nielsen, 2011a; Jørgensen et al., 2006; Larsen et al., 2009). The advantage of this method is that the driving force of the temporal development is represented in the climate model together with the physics behind the response. It is, however, questioned if the RCM representation and simulation of extreme rainfall is adequate. One problem is the spatial resolution, which in present RCMs lies in the range from 50x50km² to 10x10km². The convective clouds that creates rainfall with very high intensities over short durations has, in comparison, a spatial extent of 1 to 10 km² (Niemczynowicz 1988). Furthermore highly important cloud processes like droplet formation, draughts and turbulent mixing takes place on a even smaller scale (Baker and Peter, 2008). This means that in the RCM output convective rainfall is averaged over a grid cell and thereby underestimated. Secondly, it suggests that the representation of the convective clouds in the model is insufficient. Another concern is the temporal resolution. The RCM simulation time step is as small as a few minutes but the results are very uncertain at this scale (Meijlgaard et al. 2008). The temporal
resolution of the output is therefore aggregated to one hour, which still is on the limits of the predictive ability of the model. The main reason for doing so is that short rainfall durations are of primary interest in urban drainage design, as most sewer systems have concentration times of one hour or less.

The RCMs constantly improve. Earlier RCM simulations where run for two periods: ‘control’, covering 1961-1990 and forced by observed greenhouse gas concentration (GHG), and ‘scenario’, covering 2071-2100 and forced by a given emission scenario (Arnbjerg-Nielsen, 2011a; Jørgensen et al., 2006; Larsen et al., 2009). In the recent ENSEMBLES project an ensemble of transient simulation where made (Linden and Mitchell 2009). From these simulations direct trends in the annual development can be computed for selected variables. This leads to an extended amount of information in comparison to what is obtained from a simple comparison of ‘control’ and ‘scenario’. Several extensive rainfall events have been observed in Denmark within the last decade. This encouraged a number of analyses of observed rainfall and a significant increase in the frequency and intensity of the extreme events was found (Gregersen et al., 2010; Madsen et al., 2009; Sadri et al., 2009). In the present paper the development in observed extreme rainfall with a duration of one hour is compared to the output from one transient RCM simulation. Climate factors are subsequently calculated from both data sets and evaluated.

**METHODS**

**Data**

The rainfall measurements originate from 70 high resolution tipping bucket rain gauge stations that have been active in the period between 1979 and 2009. The network is operated by the Danish Water Pollution Control Committee (SVK) and the Danish Meteorological Institute. The temporal resolution is one minute. When periods of rain gauge malfunction has been taken into account, the present data material represents 1428 station-years. These data have formed the basis of several other analysis of extreme rainfall made in Denmark during the last decade (Madsen et al. 2002; Madsen et al. 2009; Sadri et al. 2009).

The applied RCM rainfall simulation is with the regional climate model RACMO (Meijlgaard et al. 2008) by The Royal Netherlands Meteorological Institute, KNMI, and was created as a part of the ENSEMBLES project (Linden and Mitchell 2009). RACMO was chosen between the ensemble models due to high skill for simulation of precipitation. The temporal resolution is one hour, whereas the spatial resolution of the grid cells is 25x25km². RACMO is forced by the global climate model ECHAM5 and the simulation is transient going from 1950 to 2100. Up until 2000 the external forcing of the climate system in ECHAM5 comes from the observed GHG, beyond this year it is based on the A1B scenario (Linden and Mitchell 2009). Only data from land cells in Denmark is included in the analyses, a land cell is defined as a cell with a land-to-water ratio of more than 50% leading to the selection of 79 cells.

In the following SVK and RACMO refers to the data set of observed and simulated rainfall, respectively.

**Definition of the climate factor**

The climate factor is defined in the following way:

\[
CF_{T,\Delta t} = \frac{z_{T+\Delta t}}{z_T}
\]
Where $z_T$ is the design intensity of a given return period $T$ and $\Delta t$ the expected lifetime of the drainage system. The denominator thereby represents the design intensity in the present climate, whereas the numerator represents the design intensity of the future $\Delta t$-years from now. The climate factors recommended in Denmark today are based on a comparison between the ‘control’ and ‘scenario’ period simulated by the regional climate model, HIRHAM4 (Christensen et al. 1998). There are, however, several different estimation methods behind the recommended values, for details see Arnbjerg-Nielsen (2011a).

**Statistical methods**

Time series with observed rainfall intensities for a duration of one hour are estimated from the raw data by a moving average procedure (Gregersen et al., 2010). For both SVK and RACMO the method of Partial Duration Series (PDS) were applied to sample the extreme events. In this method either the average annual number of extreme events ($\lambda$) or the threshold level ($z_0$) is a prefixed, constant, sample parameter. A declustering procedure was furthermore used to ensure that the found extremes where independent (Gregersen et al., 2010; Larsen et al., 2009). The PDS method implies that the annual number of extreme events follows a Poisson distribution. For the Danish rainfall intensities the Generalized Pareto (GP) distribution has been found to provide the best fit (Madsen et al. 2002). The $T$-year event ($z_{T,t}$) can therefore be calculated as:

$$
 z_{T,t} = z_0(t) + \frac{\alpha_t}{\kappa_t} \left[ 1 - \left( \frac{1}{T \cdot \lambda(t)} \right)^\kappa_t \right]^{\alpha_t} 
$$

(2)

Where $\alpha_t$ and $\kappa_t$ is the scale and shape parameter in the GP distribution, respectively. The subscript $t$ indicates time dependency. $\alpha_t$ and $\kappa_t$ are estimated for the data together with either $\lambda$ or $z_0$, depending on the chosen sampling method. When $\lambda$ is prefixed $z_0$ becomes a random time depended variable and vice versa, this is reflected by the parenthesis around $t$. For parameter estimation the method of L-moments is applied, where $\alpha_t$ and $\kappa_t$ depend on the first and second L-moment (Stedinger et al. 1993).

When the process that generates the extreme values is non-stationary, either of the two sampling methods could potentially lead to extensively many/few extremes in the last part of the analyzed period. Consequently, there is a risk of violating of some of the fundamental assumptions behind the extreme value theory (Kysely et al. 2010), which lead to an increased uncertainty on the GP parameter estimates. A solution could be to divide the transient time series into smaller stationary parts, especially if the observation/simulation period is more than 100 years. A suitable time slice length would be 30 years, corresponding to the length of a climatologically standard period (WMO, 2011).

**RESULTS**

**Development in the annual number of extreme events**

PDS sampling with a fixed threshold was used to evaluate non-stationary tendencies in $\lambda$, see Figure 1. Threshold values of 2.1 μm/s and 1.1 μm/s are used for SVK and RACMO, respectively. The first corresponds directly to the threshold applied by Madsen et al. (2002), the latter was based on regional evaluations aiming at an average regional $\lambda$-value of three for the first 30 years of simulations. The left (right) part of the figure show the average development for all stations (landcells), where the parametric model of the development is based on Poisson regression (Gregersen et al., 2010). The climate model simulation predicts that the frequency of the extreme rainfall will increase significantly in the period 1950-2100, with 0.4 % pr. year. The predicted increase over these 151 years is, however, found to be smaller than the increase, which we have observed during the last 31 years (2.5 % pr. year).
The deviation from the regression line is approximately the same for the two data sets, when looking at the period 1979-2009, meaning that the inter-annual variability of the extremes in RACMO fits well with what is observed. For the RACMO data this variability is however seen to increase in the last part of the simulation period.

**Development in the extreme intensities**

Similar to Figure 1 the annual mean intensity of the exceedances (μ) of the observed extreme was computed for the SVK data, see Figure 2 (left). Note that μ corresponds to the sum of the first L-moment and z₀. A significant increase was found and described by a linear regression model with an annual increase of 0.0095 μm/s/year. The SVK time series were evaluated as being too short for identification of changes in κ, because an annual estimate of κ would be highly influenced by sampling uncertainties.

For the RACMO data a different approach is used in order to comply with the extreme value theory. The dataset is divided into five time slides and for each parameter estimates are calculated. Here λ is prefixed (as three events/year) to ensure an equal sample sizes for all land cells in all periods, thus z₀ is estimated from the data. The found μ is given in Figure 2 (right). The temporal developments of the mean intensity can be difficult to compare visually, but the linear annual increase for RACMO was found to be 0.0026 μm/s/year and thereby smaller than the increase observed during the last 31 years. The GP parameters for RACMO are given in Figure 3 (left). Looking at the mean of all cells z₀ increases in the middle of the 21st century, and α shows a slow but more constant increase. κ is on the other hand highly variable; the largest value is observed in 1980-2009, whereas the most negative value again is in the middle of the 21st century. The minimum and maximum approximately follows the same tendencies as the mean.

**T-year events and climate factors**

In order to estimate climate factors that reflect the tendency that we have seen in the last 31 years of observations, the trends in λ and μ are extrapolated.
A procedure like this requires great amounts of caution, especially here where the extrapolation extends 100 years into the future, in order to obtain values that are comparable to the climate factor applied in Denmark today. For $\kappa$, a constant value of -0.21 is used, corresponding to the regional average estimated by Madsen et al. (2009). Having present and future estimates of $\lambda$, $\mu$ and $\kappa$, Equation 2 is used followed by Equation 1 to obtain $CF$. Table 1 shows the factors corresponding to a linear development in both $\lambda$ and $\mu$. For RACMO the $T$-year events are given in Figure 3 (right). NOTE that the lowest and highest value is obtained in 1980-2009 and 2040-2069, respectively. Consequently these two periods are compared in the estimation of the climate factors, see Table 1.

All the recent climate factors based on the observed development are found to be remarkably higher than the Danish standards, which are based on old RCM simulations and other information (Arnbjerg-Nielsen, 2011a). The relative difference between the factors for the three $T$-year events are however comparable. For RACMO the climate factors for $z_2, t$ and $z_{10}, t$ are close to the Danish standards, whereas it for the 100-year event is considerably larger.

**DISCUSSION**

In the development predicted by RACMO all the design intensities shows a slight decrease followed by a large increase and at last another decrease. The initial decrease could seem like a trivial coincidence, but it is interesting because the same development characterize the design intensities applied in Danish urban drainage design (Arnbjerg-Nielsen 2011b). The end period decrease could be scenario dependent, as the greenhouse gas emissions decrease in the in A1B Scenario in the end of the 21st century (IPCC, 2000). Figure 2 also indicate that extrapolation of a short term trends e.g. the development of $z_{100}$ from 1980-2009 to 2010-2039 could lead to a large overestimation of the future intensity. On the other hands, if the two end periods blindly where selected for climate factor estimation, lower values and an underestimation of $CF$ would have resulted.
Figure 3. GP parameters (left) and T-year events (right) estimated from RACMO data for five time slices. The solid line represents the mean value of all land cells, the two broken lines minimum and maximum.

Table 1. Climate factors: The present Danish standards are compared to the values based on the SVK and the RACMO analyses for these data. The rainfall duration is 60 minutes.

<table>
<thead>
<tr>
<th></th>
<th>CF definition</th>
<th>( z_{2,1} )</th>
<th>( z_{10,1} )</th>
<th>( z_{100,1} )</th>
<th>( A_t ) [years]</th>
</tr>
</thead>
<tbody>
<tr>
<td>RACMO (Scenario A1B)</td>
<td>( \frac{Z_T(2040-2069)}{Z_T(1980-2009)} )</td>
<td>1.2</td>
<td>1.4</td>
<td>2</td>
<td>60</td>
</tr>
<tr>
<td>SVK (linear develop. ( \lambda ) and ( \mu ))</td>
<td>( \frac{Z_T(2090)}{Z_T(1990)} )</td>
<td>2.3</td>
<td>2.4</td>
<td>2.5</td>
<td>100</td>
</tr>
<tr>
<td>Present Danish standards (Scenario A2)</td>
<td>( \frac{Z_T(2070-2100)}{Z_T(1961-1990)} )</td>
<td>1.2</td>
<td>1.3</td>
<td>1.4</td>
<td>110</td>
</tr>
</tbody>
</table>
The uncertainty on the estimate of κ is high, independent of the choice of estimation method, but especially for short time series. Consequently κ often is assumed to be stationary, when a development over time is addressed (e.g. Kysely et al., 2010). A close examination of Figure 3 shows that the non-linear development of the T-years events to a large degree is controlled by the change of κ, suggesting that this perhaps is too important to ignore. The Danish standards are in fact based on a development in κ, expressed as a climate factor of 1.03 (Arnbjerg-Nielsen, 2011a). Despite its small value it is seen to have a high influence on the CF estimate for the 100 year event. The SVK climate factors are based on a constant value of κ, but here the dramatic changes in λ and μ are large enough also to affect the 100-year event. All together great precautions are necessary, when the conclusions depend on the value of κ, this can, however, both be an argument for and against the assumption of the time dependent κ.

The analysis found a remarkable difference between the recent increase, in frequency and magnitude of observed extreme rainfall, and the overall increase predicted by the climate model. There are two possible explanations; the climate model may in fact underestimate the development of extreme rainfall or a random or deterministic temporary variation may have caused the observed pattern. The article presents no means for preferring one explanation over the other, but the outcome is quite different estimates of CF. In addition, new estimates of the climate factor will continuously be found, both when other RCM simulations are analyzed and when additional years of observations becomes available. Eventually the climate models will also improve in terms of temporal and spatial downscaling and parameterization. This clearly demands a framework for selection of the climate factor, which could be some kind of a weighted average based on Bayesian updating that continuously included new information (Gregersen and Arnbjerg-Nielsen, 2011). Future work will explore this possibility and extend the analysis to include more ensemble models, longer rainfall durations and, if possible, other emission scenarios.

CONCLUSION
The climate factors for urban drainage design recommended in Denmark today are based on a comparison of control and scenario simulations ran by HIRHAM4. This paper presents an analysis of two new sources of data; trends in the properties of observed extreme rainfall and transient model simulations ran by RACMO. Focus is on a rainfall duration of 60 minutes. Both results suggest that the factors presently applied are underestimated. The trends in observed rainfall extremes points towards a climate factor above two; the time series behind this conclusion do, however, have a short time length so the validity of the applied extrapolation can be questioned. It is still intimidating that the increase in the frequency of the extreme events predicted by RACMO for a period of 151 years (0.4% pr. year) is significantly smaller than the increase that we have observed during the last 31 years (2.5 % pr. year). The temporal and spatial resolution of the RCM does for certain lead to an underestimation of the extreme rainfall; the question is if the temporal development also is underestimated. It is necessary to acknowledge that a variety of climate factor estimations exists and that more will come, a method for weighted selection should therefore be developed.

ACKNOWLEDGEMENT
The authors thank The Royal Netherlands Meteorological Institute, KNMI, and Erik van Meijlgaard who kindly provided the RACMO data in a temporal resolution of one hour, although this was outside the agreement of the ENSEMBLES project.
REFERENCES

Arnbjerg-Nielsen, K. (2011a). Quantification of climate change effects on extreme precipitation used for high resolution hydrologic design, Submitted to Urban Water Journal.


