Research Article Environmetrics

Received: 10 February 2012, Revised: 31 August 2012, Accepted: 13 September 2012, Published online in Wiley Online Library: 14 October 2012

(wileyonlinelibrary.com) DOI: 10.1002/env.2173

Descriptive and predictive evaluation of high resolution Markov chain precipitation models

H. J. D. Sørup[a]*, H. Madsen[b] and K. Arnbjerg-Nielsen[a]

A time series of tipping bucket recordings of very high temporal and volumetric resolution precipitation is modelled using Markov chain models. Both first and second-order Markov models as well as seasonal and diurnal models are investigated and evaluated using likelihood based techniques. The first-order Markov model seems to capture most of the properties of precipitation, but inclusion of seasonal and diurnal variation improves the model. Including a second-order Markov Chain component does improve the descriptive capabilities of the model, but is very expensive in its parameter use. Continuous modelling of the Markov process proved attractive because of a marked decrease in the number of parameters. Inclusion of seasonality into the continuous Markov chain model proved difficult. Monte Carlo simulations with the models show that it is very difficult for all the model formulations to reproduce the time series on event level. Extreme events with short (10 min), medium (60 min) and long (12 h) durations were investigated because of their importance in urban hydrology. Both the descriptive likelihood based statistics and the predictive Monte Carlo simulation based statistics are valuable and necessary tools when evaluating model fit and performance. **Copyright © 2012 John Wiley & Sons, Ltd.**

Keywords: Box–Cox transformation; Monte Carlo simulation; seasonal variation; tipping bucket rain gauges; waiting times

1. INTRODUCTION

Modelling of precipitation is of major interest on many scales, ranging from large-scale temporal-spatial climate driven models (Gelati *et al*., 2010; Mehrotra and Sharma, 2010) over local models modelling daily precipitation amounts (Ailliot *et al*., 2009, Burton *et al*., 2010) down to temporal extremely high resolution models modelling observed point precipitation (Srikantan and McMahon, 1983; Arnbjerg-Nielsen *et al*., 1998; Thyregod *et al*., 1999; Gaume *et al*., 2007). In an urban hydrological context, these high-resolution models are of key interest, given that the typical hydrological response times in urban drainage systems is 10–60 min. Different kinds of models are used to model rainfall with a high temporal resolution. Weather generators using random pulses to simulate rainfall intensity time series are popular, but lack the short-term memory at very small time steps (Wilby *et al*., 2003; Onof and Arnbjerg-Nielsen, 2009; Burton *et al*., 2010). Markov chain models more closely related to classic time series analysis have been used for fine-scale time series generation and forecasting (Srikantan and McMahon, 1983; Arnbjerg-Nielsen *et al*., 1998; Thyregod *et al*., 1998). Furthermore, statistical models have been widely used to model only the precipitation extremes (Madsen *et al*., 2002; Koutsoyiannis, 2004a, 2004b; Betrò *et al*., 2008).

Markov chain models are very flexible and can also be used to generate models of time series covariates (Cox and Miller, 1965). Furthermore, Markov chain models are very efficient for simulating time series using Monte Carlo algorithms (Negra *et al*., 2008), but can also be used for downscaling of precipitation from climate models (Gelati *et al*., 2010), possibly down to scales useful in urban drainage.

At the minute scale, pulse models come short because of the previously mentioned lack of short-term memory at this scale. This is a logical consequence of the structure of the models and is not solved easily by scaling (Olsson and Burlando, 2002; Marani, 2003). The Markov chain model seems an attractive approach due to its apparent ability to model the chronology of intensities within rainfall events at the minute scale (Thyregod *et al*., 1998). Furthermore, Arnbjerg-Nielsen *et al*. (1998) show that instead of modelling the occurrence of tipping bucket tips as a poison process, which is a classical representation of the Markov chain model, waiting times between recordings of equal volume of rainfall from tipping buckets can be modelled based on previous states of waiting times in a Markov Chain model. Waiting times between equal volumes is a surrogate measure of the mean intensity while the volume was collected. In this study, a stochastic description of rainfall waiting times is done by Markov chain models, and descriptive and predictive qualities of these models are evaluated. Within the Markovian framework, also seasonal and diurnal differences are investigated. The first-order Markov chain model can be described as (Arnbjerg-Nielsen *et al*., 1998; Thyregod *et al*., 1998):

* *Correspondence to: H. J. D. Sørup, Department of Environmental Engineering, Technical University of Denmark, Miljøvej Building 113, DK-2800 Lyngby, Denmark. E-mail: hjds@env.dtu.dk*

a *Technical University of Denmark, Department of Environmental Engineering, Miljøvej Building 113, DK-2800 Lyngby Denmark*

b *Technical University of Denmark, Department of Informatics and Mathematical Modelling, Richard Petersens Plads Building 321, DK-2800 Lyngby Denmark*

 Copyright © 2012 John Wiley & Sons, Ltd.

$$p(w_i|w_{i-1}, w_{i-2}, \cdots w_0; \theta) = p(w_i|w_{i-1}; \theta) \quad (1)$$

where:

w_i is the waiting time in minutes between tips at index i.
w_{i-1}, w_{i-2}, ... , w_0. are the waiting times at index $i-1$, $i-2$ up until the first one at index 0.
θ is the parameter set defining the model.

meaning that the probability density function for w at index i is only dependent on the last value of w at index $i-1$. Likewise, the second-order Markov chain model is described by also including w_{i-2} on the right hand side of Equation (1).

The evaluation of the descriptive capability of the models is carried out by computing likelihood statistics. However, these statistics tend to focus too much on reducing the number of parameters based on average properties of rainfall at the cost of obtaining a good description of the precipitation extremes. The extremes are of key interest in urban hydrology. Therefore, the models should also be evaluated on the basis of their ability to reconstruct the extreme properties of precipitation. This will be done by performing Markov chain Monte Carlo (MCMC) simulations with the identified models and comparing extreme statistics between the observed and simulated series.

The aim of this paper is to create parsimonious Markov chain models for precipitation at the minute scale and assess whether likelihood based techniques can be a fast way of identifying which models perform better than others when long synthetic time series are generated with the models.

2. DATA

The data used in the modelling is 12 years of tipping bucket data from a Danish measuring station in Kolding, Denmark (Mikkelsen *et al*., 1998). The tipping bucket rain gauge records precipitation at a volumetric resolution of 0.2 mm, and the raw data set is given as total depth for those minutes where tips have been recorded. The data set is quality assured to avoid erroneous data following the approach by Fankhauser (1998) and Jørgensen *et al*. (1998) and subsequently converted to waiting times between consecutive tips. The conversion implies that the time series is transformed from equal time steps to equal volumetric steps. Because the data set does not contain any information below 1 min in the temporal domain, minutes where several tips have been recorded will be denoted as 1/(the number of tips within the minute). Data contained no recording of more than 10 tips/min after quality assurance. Thus, data consists of discrete realizations: {1/(10 tips/min), 1/(9 tips/min), ... ,1/(2 tips/min), 1 tip/min, 2 min between tips, ... , k min between tips}, as indicated in Figure 1. The primary advantage of this treatment of data is the creation of data that contain all the information in the observed data but avoid zero recordings. The waiting times span six orders of magnitude, from rainfall intensities in the order of 2 mm/min (10 tips/min) up to dry spells of weeks (>10 000 min between tips).

3. METHODS

3.1. Markov chain models

Markov chain models are characterized by describing the state of observation at a given time as a function of the previous state only (Equation (1)). Both Arnbjerg-Nielsen *et al*. (1998) and Thyregod *et al*. (1998) have used this approach to model precipitation. Arnbjerg-Nielsen *et al*. (1998) constructed discrete Markov chain models with good descriptive and predictive statistics, but uses a vast amount of parameters by allowing almost all realizations to represent an individual discrete state. These models describe the precipitation process very well, but tend to be over parameterized. More coarse models have been developed with fever parameters, still with a good fit. Thyregod *et al*. (1998), on the other hand, used the Markov chain models to simulate precipitation events. The models satisfactorily reproduced the chronology of intensities observed for both convective and frontal precipitation events.

To obtain a more parsimonious model for an entire rain series, a substantial reduction of the number of parameters is needed. Initial tests have shown that after a simple Box–Cox transformation, the numerical values of the states can be approximated by a normal distribution. Another feature of the Box–Cox transformation is that the per definition positive waiting times follow an unlimited normal distribution in the transformed space. A standard Box–Cox transformation is applied as follows (Box and Cox, 1964; Katz, 1999):

$$w^{(\lambda)} = \frac{w^{\lambda} - 1}{\lambda} \quad (2)$$

where:

w is the original waiting time.
$w^{(\lambda)}$ is the Box–Cox transformed waiting times.
λ is the transformation parameter.

The obtained maximum likelihood estimate for the transformation parameter is $\hat{\lambda} = -0.1755$. This is used to transform the data set. Figure 2 shows data before and after transformation. The fit is not perfect, but the discrepancies are primarily for the very long waiting times (Figure 2). This means that the models using the transformed data will be poorer at describing the dry periods between the rain events in comparisons with the chronology of waiting times within events.

As a consequence of the transformation, it may be possible to construct models where each state is characterized by a specific normal distribution. Normal distributions have advantages over other distributions. They have attractive statistical properties. The distribution parameters are independent of each other, providing the opportunity to model them independently (Box and Cox, 1965). The model is of the type defined in Equation (1), but operates in a transformed space where the waiting times at index i of the time series is described by the state at index $i-1$, $S(i-1)$. It can be expressed as follows:

 Copyright © 2012 John Wiley & Sons, Ltd.

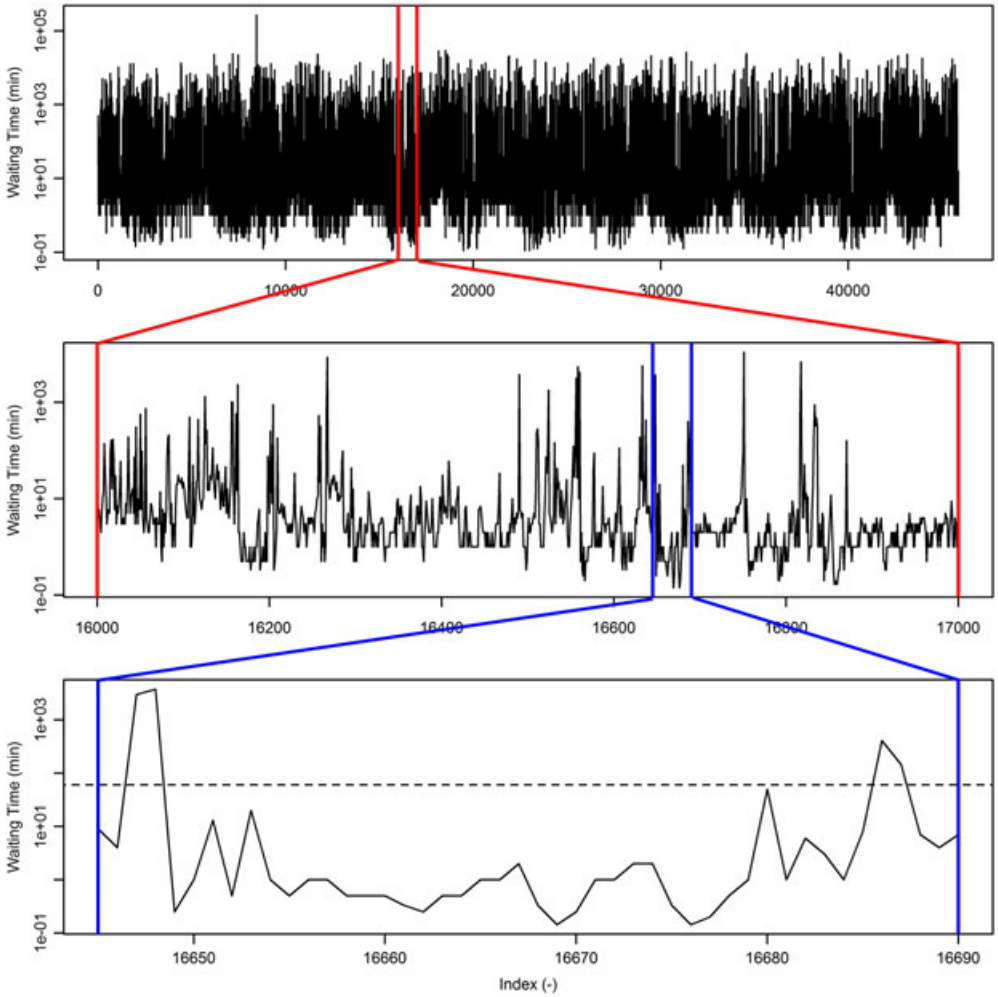

Figure 1. Visualization of data. (Top) The whole time series measured as minutes between tips. (Middle) The red subsection of the top time series. (Bottom) The blue subsection of the middle figure – A typical course of a single rain event. The dashed line represents waiting times of 1 h, which is used as the shift between dry and wet weathers (Mikkelsen *et al*., 1998)

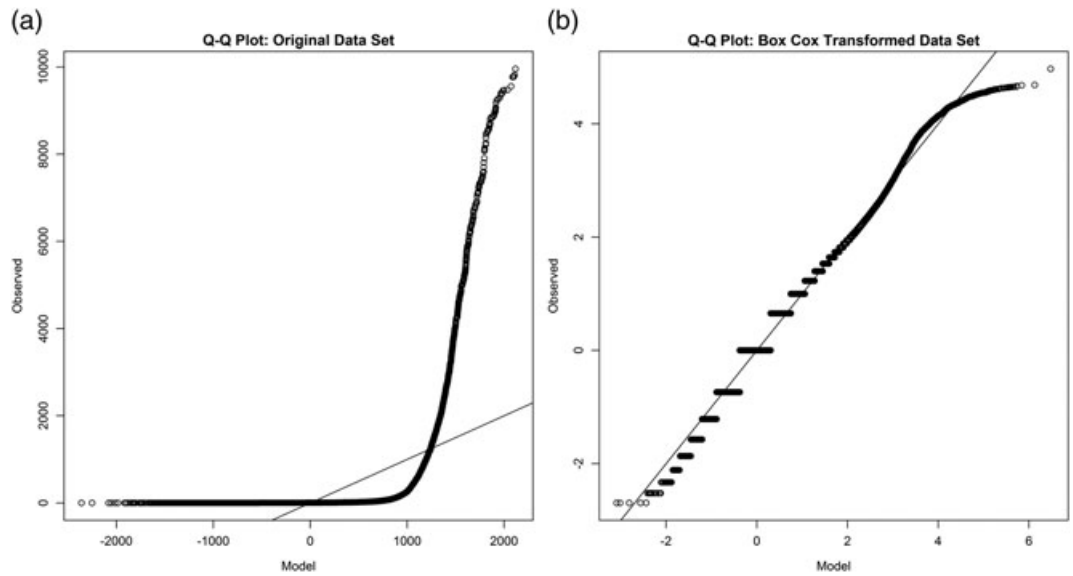

Figure 2. (a) Quantile–quantile plots of the original and (b) Box–Cox transformed data sets against a normal distribution

 Copyright © 2012 John Wiley & Sons, Ltd.

$$w_i^{(\lambda)}(S(i-1)) = \begin{cases} \mu_{S_1} + \sigma_{S_1} \in_{S_1 i} & \text{for } S(i-1) = S_1 \\ \vdots \quad ; & \vdots \\ \mu_{S_n} + \sigma_{S_n} \in_{S_n i} & \text{for } S(i-1) = S_n \end{cases} \tag{3}$$

where:

$w_i^{(\lambda)}$ is the Box–Cox transformed waiting time at index i.
$S(i-1)$ is the state at index $i-1$.
μ_{S_k}'s and σ_{S_k}'s are the means and standard deviations characterising the normal distributions defining the states of a given model; $k = 1, 2, \ldots, n$, with n being the number of states.
$\varepsilon_{S_k i}$'s are standard normally distributed random quantities defining the error in the model at the given state and index.

The states can be defined by the waiting times themselves and/or by external explanatory variables such as season or time of day. Initially, three different model state dependence will be investigated; a simple first order Markov model (named Markov(1)), a seasonal model depending only on the month (named Seasonal), and a diurnal model depending on the hour of the day (named Diurnal). The Seasonal and Diurnal models are not Markov chain models but provide useful information that is later combined with the Markov properties.

To create the Markov(1) model, the states are set up to include a larger proportion of the waiting times the longer waiting times considered. This gives a very fine description of the precipitation for the very short waiting times and an increasingly coarser model for longer waiting times. The aggregated states are believed to be scaled fine enough to capture the precipitation processes sufficiently well. Furthermore, this aggregation secures that there is sufficient data in each state.

The states definitions for the three initial discrete models are given by the following:

$$\bar{S}_{\text{Markov(1)}} = \begin{cases} S_1 & \text{for } w_{i-1} \leq \frac{1}{6} \\ S_2 & \text{for } w_{i-1} = \frac{1}{5} \\ \vdots & \vdots \\ S_5 & \text{for } w_{i-1} = \frac{1}{2} \\ S_6 & \text{for } w_{i-1} = 1 \\ \vdots & \vdots \\ S_{14} & \text{for } w_{i-1} = 9 \\ S_{15} & \text{for } 10 \leq w_{i-1} \leq 14 \\ \vdots & \vdots \\ S_{24} & \text{for } 55 \leq w_{i-1} \leq 59 \\ S_{25} & \text{for } 60 \leq w_{i-1} \leq 89 \\ \vdots & \vdots \\ S_{28} & \text{for } 150 \leq w_{i-1} \leq 179 \\ S_{29} & \text{for } w_{i-1} \geq 180 \end{cases} \tag{4a}$$

$$\bar{S}_{\text{Seasonal}} = \begin{cases} S_1 & \text{for month} = \text{January} \\ \vdots & \vdots \\ S_{12} & \text{for month} = \text{December} \end{cases} \tag{4b}$$

$$\bar{S}_{\text{Diurnal}} = \begin{cases} S_1 & \text{for hour of the day} = 00 \\ \vdots & \vdots \\ S_{24} & \text{for hour of the day} = 23 \end{cases} \tag{4c}$$

with 29 states for the Markov(1) model, one for each defined bin of waiting times; 12 states for the Seasonal model, one for each month; and 24 states for the Diurnal model, one for each hour of the day. According to Equation (3), each of the states uses two parameters for determining the probability distribution function for the next waiting time, and each model needs one extra parameter to account for the Box–Cox transformation (Equation (2)). Several models are developed from the original ones. For the Markov chain model, two different alternative model formulations are examined; a second-order Markov model that includes information on the two previous waiting times, and a continuous model where the normal distribution parameters are modelled across states using continuous functions. The seasonal and diurnal models are regarded as having a cyclic nature and the means (μ) and standard deviations (σ) are modelled across states using a three-parameter single harmonic sine function, $g_x(S)$ defined as follows:

$$g_x(S) = a_x + b_x \sin\left(c_x + \frac{2\pi \cdot S}{S_{max}}\right) \tag{5}$$

where:

x is the original parameter that should vary according to the sine function (e.g. μ or σ).
a, b, and c are the three parameters used to fit the harmonic variation.
S is the numeric representation of the state (e.g. one for January).
S_{max} is the number of states (e.g. 12 for the seasonal variation over months).

 Copyright © 2012 John Wiley & Sons, Ltd.

This approach reduces the number of parameters used to describe the variation and introduces a smooth evolution in the modelled variables over the considered states.

3.2. Statistics for comparison of models

Likelihood based statistics are used to determine the descriptive nature of the models and simulations in combination with extreme statistics to determine the predictive strength.

3.2.1. Likelihood based statistics

The models are compared directly using likelihood based statistics. The likelihood is an objective measure of the models fit to data, and is, as such, ideal for comparison of nested models including different variables but describing the same data set (Madsen and Thyregod, 2011). All simple models are tested against their more complex counterparts using the likelihood ratio test:

$$-2\ \log\left(\frac{\text{maximum likelihood of model with } r+p \text{ parameters}}{\text{maximum likelihood of sub model with } p \text{ parameters}}\right) \sim \chi^2(r) \tag{6}$$

This test specifically tests whether a change in the number of parameters (from $r+p$ to p) can be justified given the change in likelihood using a χ^2 distribution with r degrees of freedom (Madsen and Thyregod, 2011).

The Bayesian information criteria (BIC) is calculated for each model for comparison. The BIC is chosen because it favours simple models with few parameters. The BIC is calculated as (Yang, 2005):

$$BIC = \ -2 \cdot l(\theta|x) + \log(np) \cdot np \tag{7}$$

where:

$l(\theta|x)$ is the maximum log-likelihood.
np is the number of parameters in the model.

3.2.2. Monte Carlo simulations boundary conditions

Monte Carlo simulations are carried out for all models. The Box–Cox transformation results in good model representation of the waiting times in the chosen interval (Figure 2), but the unbound normal distributions can result in unrealistic extremes, and boundary conditions are introduced in the simulations to avoid this. The boundary and initial conditions are set equal for all simulations with all models. The very long waiting times are removed as a physically based constrain because dry periods of more than 2 weeks are unexpected at the modelled location. Finally, a modification of the boundary condition regarding the very short waiting times is included; these are important when modelling the short-time extremes, but the data basis is very limited, and it is impossible to formulate a reasonably model. The constrains are expected to result in the models to slightly overestimate the mean annual precipitation (*MAP*), due to the lack of very long waiting times, and to underestimate the short time peaks slightly due to the lack of descriptive data in this area. The boundary conditions are specified as follows:

- All models cover the range of waiting times from 10 s to 2 weeks; that is $w \in \left[\frac{1}{6}; 20\,160\right]$ with the waiting times expressed in minutes.
- If values outside this range are produced, they are corrected to the boundary value, and the state is determined with the boundary value as input.
- The only exception to the previously stated rules is for the very extreme precipitation represented by values between six and 10 s $\left(w \in \left[\frac{1}{10}; \frac{1}{6}\right]\right)$, where the value is kept, but the state is determined as if the value were $w = \frac{1}{6}$.

All models are initiated with the first value of the observed time series, except the second-order Markov model, which takes the two first values as initial conditions. For each model, 100 time series are simulated, each with 200 000 waiting times. This corresponds to approximately 50 years in the time domain if the *MAP* is satisfactorily reproduced.

3.2.3. Extreme statistics on simulations

A number of statistics are extracted from the simulated time series to evaluate the model performance. These are as follows:

- The *MAP*. This parameter is evaluated to ensure that the model produce the right amount of precipitation on the large (annual) scale.
- Seasonality of extreme events. The top 300 extreme events are extracted for events with 10 min, 60 min and 12 h durations and compared with the observed seasonality of extreme events.
- Magnitude of extreme events. The magnitude of the extreme events with 10 min, 60 min and 12 h durations are compared with the observed ones, and to an extreme value distribution fitted to the observations. The comparison is based on the estimated at-site confidence limits for a partial duration series model of the extremes as introduced by Madsen *et al.* (2002, 2009).

 Copyright © 2012 John Wiley & Sons, Ltd.

The return period of the extreme events is calculated using the California plotting formula with the length of the time series defining the return period of the most extreme event (Chow *et al*., 1988) and the threshold values identifying extreme values for the Generalized Pareto distribution (GPD) for different event durations are adopted from Madsen *et al*. (2009). A GPD is fitted to the observations (Madsen *et al*., 2002), and from this, confidence intervals of the observations are estimated.

4. IDENTIFICATION OF MODEL COMPONENTS

Initially, three different model components are investigated. The short-term memory is investigated using the simple first-order Markov chain model named Markov(1). Seasonality is investigated by the creation of a monthly model named Seasonal, and diurnal variation by an hourly model named Diurnal. Equations (4a)–(4c) defines the states of the three models. All the models are compared with a single state model named Base. This is a model consisting of only one normal distribution fitted to all data after the Box–Cox transformation.

Figure 3 shows the marginal distribution of four states of the Markov(1) model. The distributions seem reasonable. Table 1 lists the log-likelihood of all models, together with the calculated likelihood ratio test conclusions. The results indicate that extensions with all three types of models lead to a significantly better description than using just the Base model. The Markov(1) model is seen to improve the value of the log-likelihood most, but also, the seasonal variation seem to be important. The diurnal variation seems to have only limited influence compared with the other two models. Only the Markov and seasonal properties will be studied further.

4.1. Optimization of the model components

As the Markov process seems to be the most important, this is the one first optimized. Two different alternative formulations are investigated: (i) a second-order Markov model (denoted Markov(2)) and (ii) a first-order Markov model with continuous state space (denoted Markov

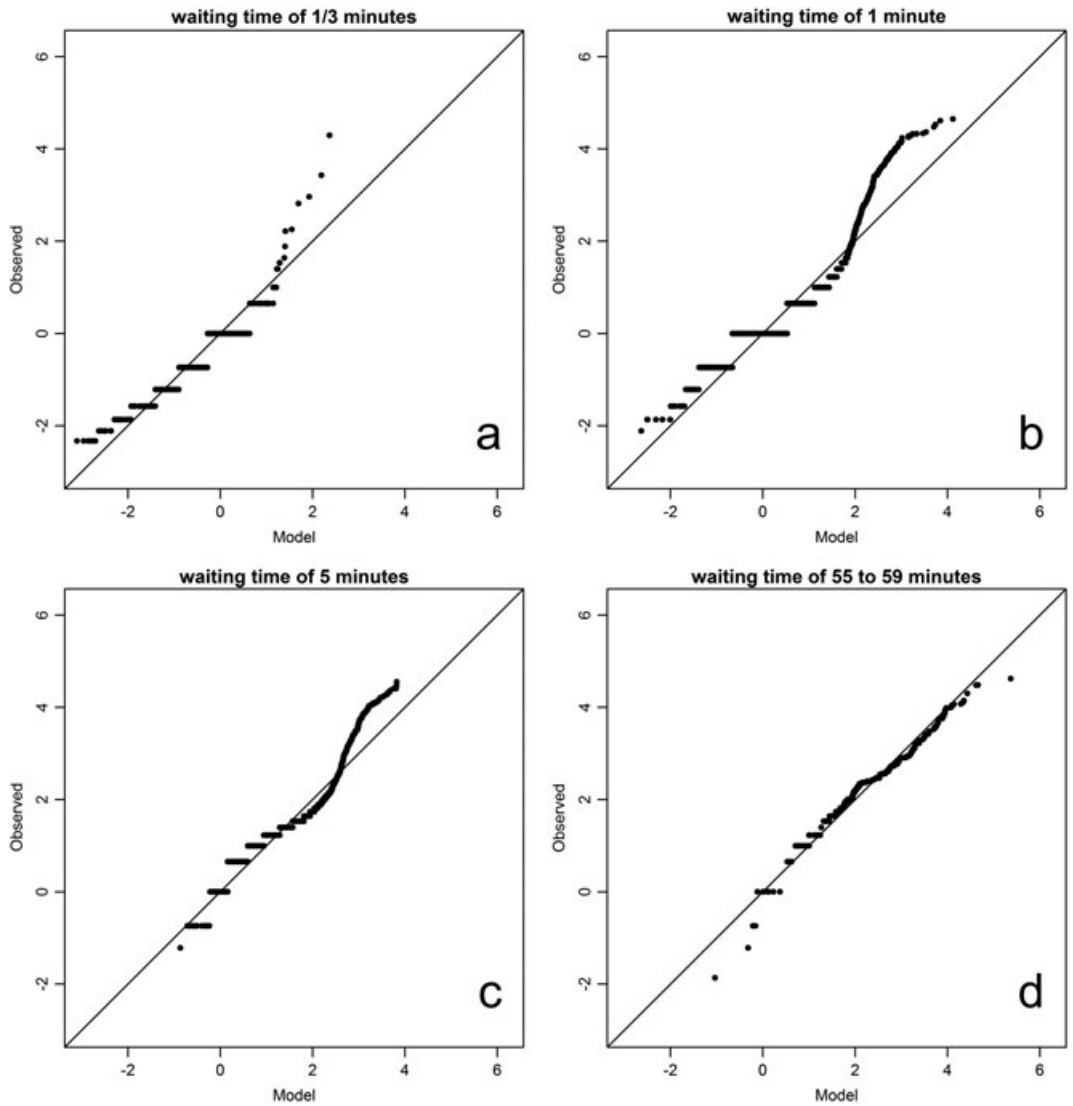

Figure 3. The quantile–quantile plots for four states of the Markov(1) model. Box–Cox transformed values

 Copyright © 2012 John Wiley & Sons, Ltd.

Table 1. The likelihood based test statistics of the models

Model	Number of parameters	Log-likelihood	Likelihood ratio test conclusions	BIC	BIC rank
Base	3	−70 551	—	141 103	8
Markov(1)	59	−56 993	The models with the most	114 090	2
Markov(2)	1683	−55 168	parameters are always favoured	115 765	4
Seasonal	25	−68 021		136 077	5
Seasonal(sine)	7	−68 172		136 350	6
Diurnal	49	−70 265		140 613	7
Markov(continuous)	10	−57 136		114 282	3
Markov(seasonal, continuous)	28	−56 060		112 161	1

Likelihood ratio test probabilities of bold italic models are calculated against the bold model.
For italic models, it is calculated against the preceding bold italic model.

(continuous)) to emphasize that the precipitation process can be described continuously even though precipitation at very short time steps by nature is discrete (Ignaccolo and de Michele, 2011). The continuous state space model should solve some of the problems that arise from splitting data into discrete states. These two models have inverse impacts on the number of parameters in the model. The Markov(2) model will increase the number of parameters dramatically whereas the Markov(continuous) will reduce it markedly. The Markov(2) model states are modified from the Markov(1) states as follows:

$$\bar{S}_{\text{Markov(2)}} = \begin{cases} S_1 & \text{for } w_{i-1} \le \frac{1}{6} \text{ and } w_{i-2} \le \frac{1}{6} \\ \vdots & \vdots \\ \vdots & \vdots \\ S_{29} & \text{for } w_{i-1} \ge 180 \text{ and } w_{i-2} \le \frac{1}{6} \\ S_{30} & \text{for } w_{i-1} \le \frac{1}{6} \text{ and } w_{i-2} = \frac{1}{5} \\ \vdots & \vdots \\ S_{841} & \text{for } w_{i-1} \ge 180 \text{ and } w_{i-2} \ge 180 \end{cases} \tag{8}$$

equal to $\bar{S}_{\text{Markov(2)}} = \bar{S}_{\text{Markov(1)}} \times \bar{S}_{\text{Markov(1)}}$, giving 841 states in total.

The Markov(continuous) model uses a two compartment model where the normal distribution parameters are described as continuous functions for waiting times below 60 min (subscript r for rain) and as constants above (subscript n for no rain). This is a simple way of estimating the embedded process that leads to the pulses of 0.2 mm rainfall that is observed by the rain gauge. The Markov(continuous) is defined as a variant of Equation (3) as:

$$w_i^{(\lambda)} = \mu(w_{i-1}) + \sigma(w_{i-1})\varepsilon_i \tag{9a}$$

where:

$$\mu(w_{i-1}) = \begin{cases} a_{\mu_r} + b_{\mu_r} \log(w_{i-1}) + c_{\mu_r} \log(w_{i-1})^2 & , 0 < w_{i-1} < 60 \\ \mu_n & , w_{i-1} \ge 60 \end{cases} \tag{9b}$$

$$\sigma(w_{i-1}) = \begin{cases} a_{\sigma_r} + b_{\sigma_r} \log(w_{i-1}) + c_{\sigma_r} \log(w_{i-1})^2 + d_{\sigma_r} \log(w_{i-1})^3 & , 0 < w_{i-1} < 60 \\ \sigma_n & , w_{i-1} \ge 60 \end{cases} \tag{9c}$$

The model described by Equations (9a)–(9c) uses a mixture of transformed and untransformed waiting times; this was done to keep close connection to the physical properties of precipitation and the distinction between within and between events.

Table 1 shows the main results of estimating the proposed model extensions. The Markov(2) model improves the log-likelihood markedly, and the likelihood ratio test show that it is in fact significantly better than the Markov(1) model. However, it uses so many parameters that is not considered useful in this context, and no further work will be done with the Markov(2) model. The Markov(continuous) and Markov(1) models are not nested models and cannot be compared directly using the likelihood ratio test. The BICs in Table 1 show that the Markov(1) model is the best model. However, the Markov(continuous) model only decreases the log-likelihood slightly, even though it decreases the number of parameters used markedly. From the results in Table 1, it is concluded that the Markov(continuous) model hold the largest potential for further development given a need for few parameters; this is the approach that will be used to describe the Markovian nature in the following combined model.

 Copyright © 2012 John Wiley & Sons, Ltd.

The approach of modelling parameters across states is also used in the optimization of the seasonal model. Here, a sine function, see Equation (5), is used to estimate parameters that vary over the course of the year. Table 1 shows the likelihoods of the seasonal models. The decrease in log-likelihood is not big even though the likelihood ratio test concludes that the simple seasonal model cannot be used. Despite this result, the simplified seasonality formulation will be used because it drastically decreases the number of parameters in a model.

A model is created using elements from the Markov(continuous) and Seasonal(sine) models. The model, Markov(seasonal, continuous), is created by sine varying the parameters from the Markov(continuous) model. This gives the following model formulation for Markov (seasonal, continuous):

$$w_i^{(\lambda)} = \mu(w_{i-1}) + \sigma(w_{i-1})\varepsilon_i \tag{10a}$$

where:

$$\mu(w_{i-1}) = \begin{cases} g_{a_{\mu_r}}(m) + g_{b_{\mu_r}}(m)\log(w_{i-1}) + g_{c_{\mu_r}}(m)\log(w_{i-1})^2 & , 0 < w_{i-1} < 60 \\ g_{\mu_n}(m) & , w_{i-1} \geq 60 \end{cases} \tag{10b}$$

$$\sigma(w_{t-1}) = \begin{cases} g_{a_{\sigma_r}}(m) + g_{b_{\sigma_r}}(m)\log(w_{i-1}) + g_{c_{\sigma_r}}(m)\log(w_{i-1})^2 + g_{d_{\sigma_r}}(m)\log(w_{i-1})^3 & , 0 < w_{i-1} < 60 \\ g_{\sigma_n}(m) & , w_{i-1} \geq 60 \end{cases} \tag{10c}$$

The *g(m)* functions mean that the parameter vary with month according to Equations (5).

Table 1 shows the log-likelihood, the likelihood ratio test conclusion and the BIC of the Markov(seasonal, continuous) model in comparison with the Markov(continuous) model. Inclusion of seasonality significantly improves the model and leads to a marked increase in the log-likelihood. Most interesting is that the Markov(seasonal, continuous) model with 28 parameters better describes data than the Markov(1) model with 59 parameters. This further strengthens our belief in the usefulness of the continuous description of the Markov process.

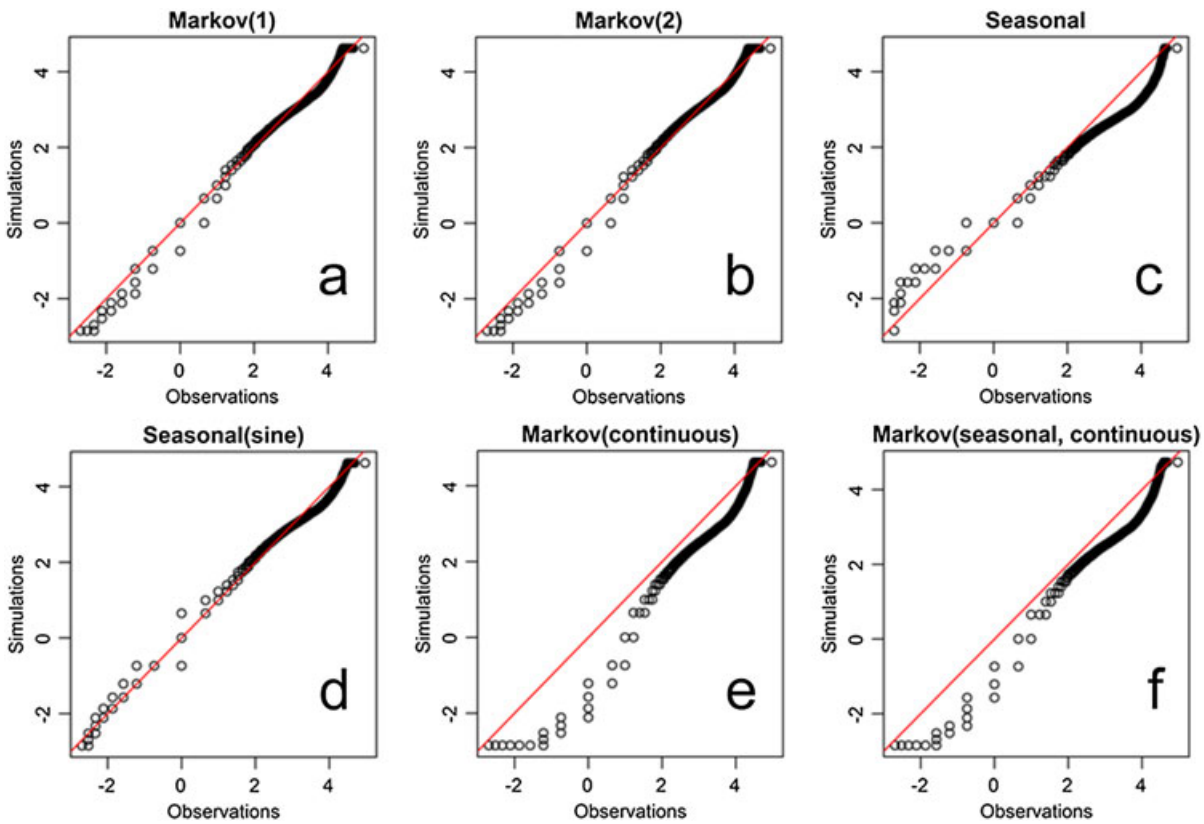

Figure 4. Quantile-quantile plots for the (a) Markov(1), (b) Markov(2), (c) Seasonal, (d) Seasonal(sine), (e) Markov(continuous) and (f) Markov(seasonal, continuous) models. Box–Cox transformed values

 Copyright © 2012 John Wiley & Sons, Ltd.

5. MONTE CARLO SIMULATIONS

5.1. Marginal distribution

Figure 4 shows the quantile–quantile plots for the six models chosen for simulation. Generally, they all show the same trend. The points are rather scattered for the low waiting times, with the two continuous models systematically producing too low values (Figure 4(e) and (f)). All models show a consistent shape for the high waiting times where the upper simulation bound is clearly marked and the spread is very narrow. Again, the continuous models show systematic discrepancy from the observations (see Figure 4(e) and (f)). The clear marking of the upper simulation bound suggests that all models could have problems with simulating dry weather; resulting in a too short mean waiting time between events.

5.2. Mean annual precipitation

The *MAP* of the individual simulations is presented in Figure 5. The Markov(1) and Markov(2) models both produce less precipitation than observed on an annual basis indicating that the mean waiting time simulated by these models is somewhat higher than the observed.

The Seasonal model vastly overestimates the *MAP*. This indicates that for some months, the model simulates very intense precipitation, and also that the model does not capture the full dynamics of the time series. The Seasonal(sine) model, on the other hand, reproduces the *MAP* fairly well, only overestimating it a bit; somehow, the dampening effect of using sine curves to represent annual variation result in much more realistic waiting times overall. The Markov(continuous) and Markov(seasonal, continuous) models both overestimate the *MAP* with approximately 70% which is in line with the discrepancies observed in the quantile–quantile plots on Figure 4(e) and (f).

5.3. Seasonality

The seasonal distribution of the extreme events for event lengths of 10 min, 60 min and 12 h are shown on Figure 6(a), (b) and (c), respectively. Figure 6(a) and (b) show that for the short events, there is a clear seasonality and only the three seasonal models (Seasonal, Seasonal(sine) and Markov(Seasonal, continuous)) seem to reproduce it (with varying precision though). For the 10-min events, the seasonal variation is very pronounced and quite well reproduced by the seasonal models. For the 60-min events, where the seasonality is less pronounced, the seasonal models seem to retain the strong seasonality, and as a consequence, they all overestimate the seasonality. For the 12-h event, no real seasonality is observed (although there is a slight indication of more events in the fall and less in spring), but the seasonal models all have a clear seasonal variation similar to the one for the shorter events. This indicates that the extreme 10-min events in general are the peaks of the extreme 60-min events, which again are the peaks of the extreme 12-h events for the seasonal models. This is not what is expected from the observations, and it means the models are not able to reproduce the long lasting low intensity rains of the fall well.

From Figures 4–6, it seems the Seasonal(sine) model is the best model; however, this model has no mechanism to order the chronology of waiting times and is thus expected to perform poorly on event level.

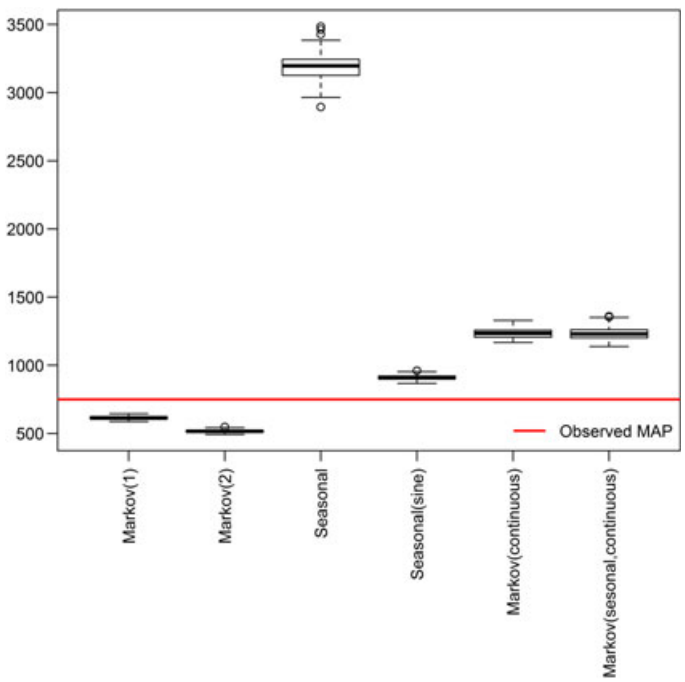

Figure 5. The mean annual precipitation (*MAP*) produced by the simulations with the six models compared to the observed *MAP* of the recorded precipitation time series

 Copyright © 2012 John Wiley & Sons, Ltd.

5.4. Extreme events statistics

The simulations result in 100 realizations of the individual models and confidence intervals are constructed from these, see Figure 7 for examples. Figure 7(a)–(c) show the observations, the fitted GPD and its 95% confidence interval along with the mean of the Markov (seasonal, continuous) model and its 95% confidence interval for 10 min, 60 min and 12 h extreme events, respectively. The other models have similar sized confidence intervals (not shown in this paper). Figure 7 illustrates that not only do the models result in uncertainty but also the observations themselves are to be regarded just as uncertain in the sense that the true return period of the most extreme events are effectively unknown.

Figure 8(a)–(c) show the mean of the simulations with the six models along with the observations and the GPD fitted to the observations for 10 min, 60 min and 12 h extreme events, respectively. The extreme statistics in Figure 8(a) illustrates the difficulty of reproducing the observed time series on event level. Furthermore, difficulties in reproducing realistic dry weather periods are reflected in the assigned return periods for the simulated events, adding even further, unquantifiable, uncertainty to the figures. All the models seem to have problems capturing the shape of the curve resulting in a flattened curve. For the continuous model, this result in too high intensities for the short return periods and too low intensities for the long return period. The Markov(1) and Markov(2) models fairly well reproduce the short return periods but again underestimates the long ones. The seasonal models simply lack clustering of the very high intensity waiting times, and thus vastly underestimate this statistic. For event durations of 60 min, Figure 8(b) clearly shows that the Markov(1) and Markov(2) models very well capture the pattern, whereas the continuous models, as discussed earlier, include the same events as for the extreme 10-min events and thus seriously overestimates the statistics. Again, the seasonal models lack clustering of the waiting

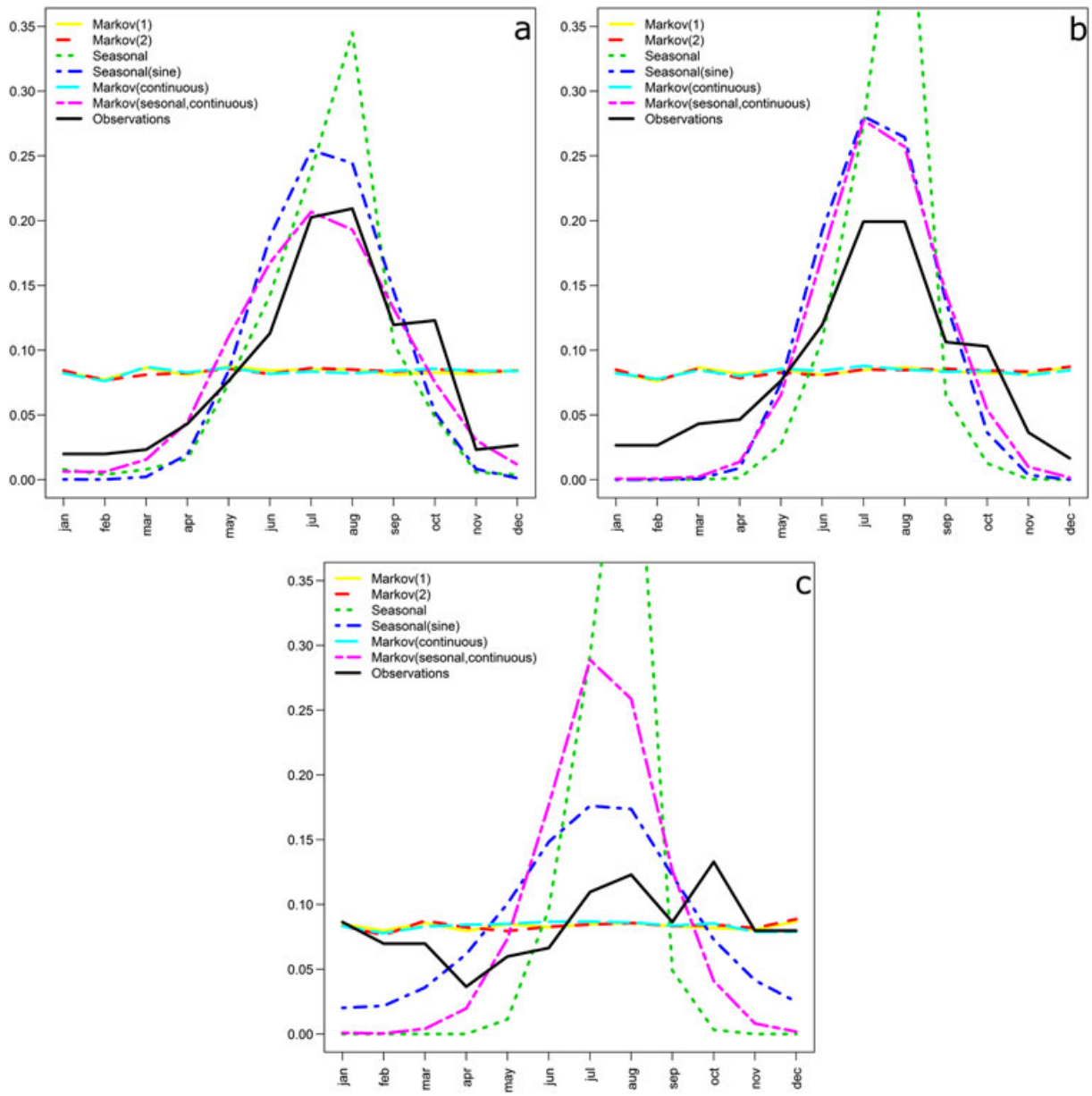

Figure 6. Seasonal variation of occurrence of the (300 most) extreme events. (a) 10 min extreme events; (b) 60 min extreme events; and (c) 12 h extreme events

 Copyright © 2012 John Wiley & Sons, Ltd.

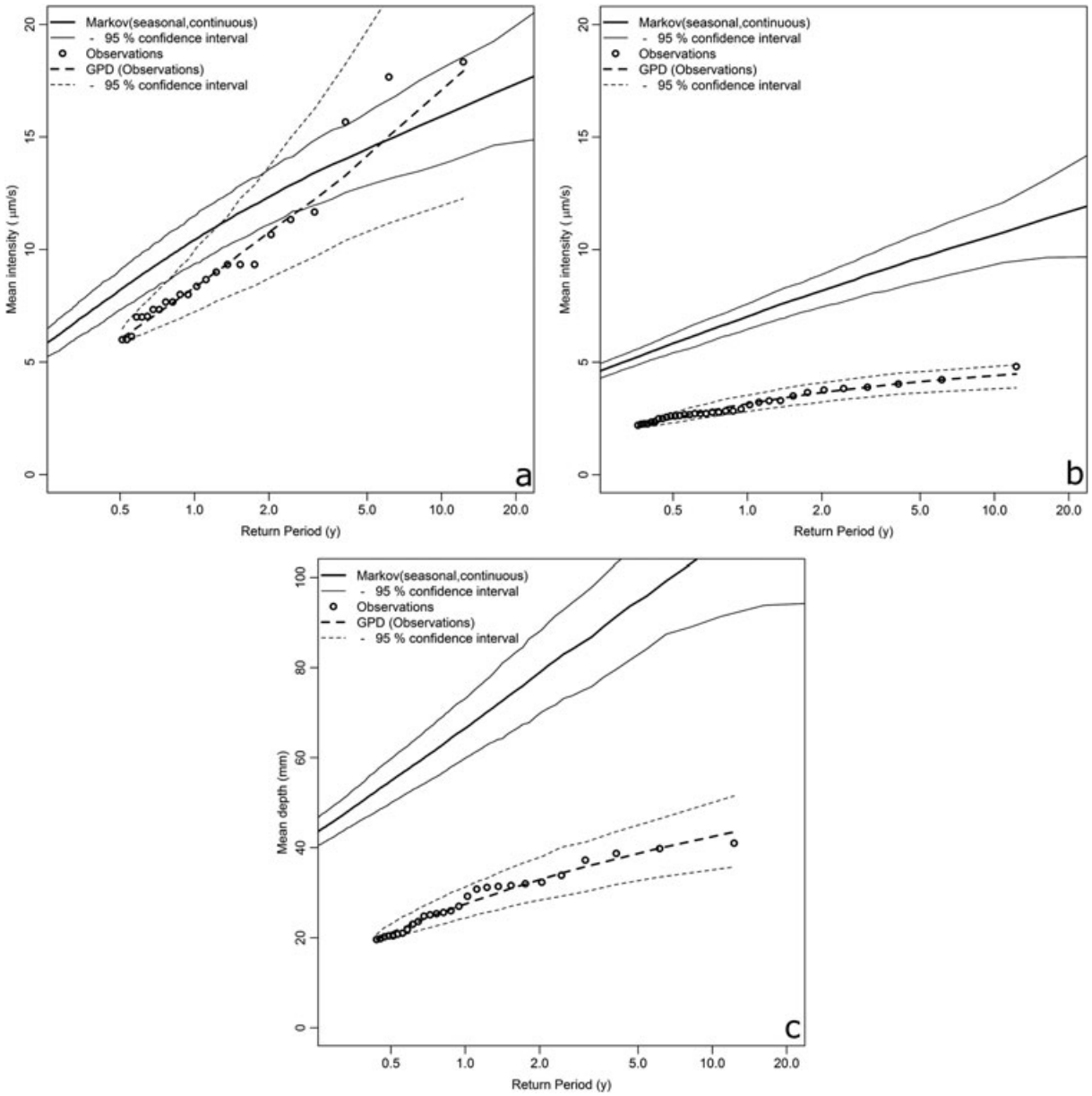

Figure 7. The uncertainty associated with both the Monte Carlo simulations and the observations, both 95% confidence intervals. Examples (a)–(c) showing only the Markov(seasonal, continuous) model and the observations; other models not shown

times and underestimate the intensities. The 12 hours total depth shown on Figure 8(c) again substantiates the assumption that it is the same event that produces the extreme statistics for all the event durations. Furthermore, the Markov(1) and Markov(2) models seem to have problems reproducing the long duration extreme events, a problem also recognized by Arnbjerg-Nielsen *et al*. (1998). The Seasonal model is actually the best performing model with respect to this statistic, but the wrong seasonality of the model (Figure 6(c)), leads to the conclusion that it is a result without physical relations to the observed time series.

6. CONCLUSIONS

Stochastic modelling of precipitation is a useful tool. The Markov chain models used in this paper provide a fairly simple way of describing changing intensity as an inherent property of the precipitation itself. To do this, the use of waiting times, instead of intensities, is a valuable tool. Seasonality is also shown to have importance when describing this process, but the ideal incorporation into the model is not identified. On the event level, the simulations show that even the models with the best likelihood scores have problems in describing the events. Both descriptive and predictive statistics will have to be used when evaluation of such models are to be carried out; the descriptive statistics is a quick tool to select among very similar models, but for most time series, where chronology matters, the descriptive statistics derived from the simulations are essential in evaluating the model performance.

 Copyright © 2012 John Wiley & Sons, Ltd.

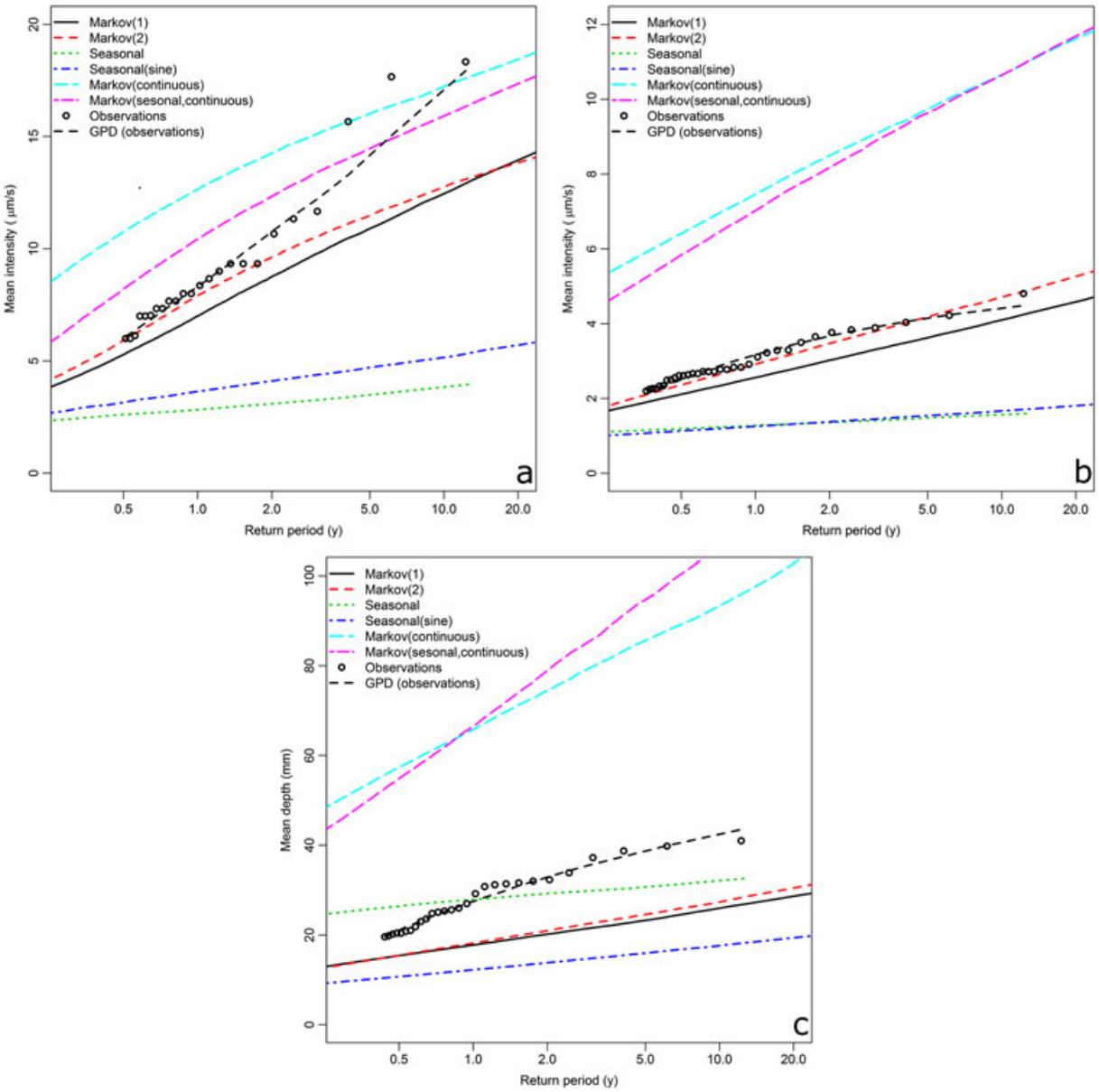

Figure 8. Simulation results of all six models compared with the observations and the fitted generalized Pareto distribution. (a) 10 min extreme events; (b) 60 min extreme events; and (c) 12 h extreme events

Acknowledgement

Data were kindly provided by the Water Pollution Committee of the Society of Danish Engineers.

REFERENCES

Ailliot P, Thomson C, Thomson P. 2009. Space-time modelling of precipitation by using a hidden Markov model and censored Gaussian distributions. *Journal of the Royal Statistical Society* **58**(3): 405–426.

Arnbjerg-Nielsen K, Madsen H, Harremoës P. 1998. Formulating and testing a rain series generator based on tipping bucket gauges. *Water Science and Technology* **37**(11): 47–55.

Betrò B, Bodini A, Cossu QA. 2008. Using a hidden Markov model to analyse extreme rainfall events in Central-East Sardinia. *Environmetrics* **19**(7): 702–713.

Box GEP, Cox D. 1964. An analysis of transformations. *Journal of the Royal Statistical Society* **26**(2): 211–252.

Burton A, Fowler HJ, Blenkinsop S, Kilsby CG. 2010. Downscaling transient climate change using a Neyman–Scott Rectangular Pulses stochastic rainfall model. *Journal of Hydrology* **381**(1–2): 18–32.

Chow VT, Maidment DR, Mays LW. 1988. Applied Hydrology. McGraw-Hill: New York.

Cox D, Miller H. 1965. The Theory of Stochastic Processes. 1st ed. Chapman & Hall: London.

Fankhauser R. 1998. Influence of systematic errors from tipping bucket rain gauges on recorded rainfall data. *Water Science and Technology* **37**(11): 121–129.

 Copyright © 2012 John Wiley & Sons, Ltd.

Gaume E, Mouhous N, Andrieu H. 2007. Rainfall stochastic disaggregation models: calibration and validation of a multiplicative cascade model. *Advances in Water Resources* **30**(5): 1301–1319.
Gelati E, Rosbjerg D, Christensen OB, Rasmussen PF. 2010. Downscaling atmospheric patterns to multi-site precipitations amounts in southern Scandinavia. *Hydrology Research* **41**(3–4): 193–210.
Ignaccolo M, de Michele C. 2011. The discrete charm of rain. *Physics Today* **64**(1): 68–69.
Jørgensen HK, Rosenørn S, Madsen H, Mikkelsen PS. 1998. Quality control of rain data used for urban runoff systems. *Water Science and Technology* **37**(11): 113–120.
Katz RW. 1999. Moments of power transformed time series. *Environmetrics* **10**(3): 301–307.
Koutsoyiannis D. 2004a. Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation. *Hydrological Science Journal* **49**(4): 575–590.
Koutsoyiannis D. 2004b. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. *Hydrological Science Journal* **49**(4): 591–610.
Madsen H, Thyregod P. 2011. An Introduction to General and Generalized Linear Models. 1st ed. Taylor & Francis Ltd.: Boca Raton, FL.
Madsen H, Mikkelsen PS, Rosbjerg D, Harremoës P. 2002. Regional estimation of rainfall intensity–duration–frequency curves using generalized least squares regression of partial duration series statistics. *Water Resource Research* **38**(11): Art. no 21.
Madsen H, Arnbjerg-Nielsen K, Mikkelsen PS. 2009. Update of regional intensity–duration–frequency curves in Denmark: tendency towards increased storm intensities. *Atmospheric Research* **92**(3): 343–349.
Marani M. 2003. On the correlation structure of continuous and discrete point rainfall. *Water Resources Research* **39**(5) Art. No. 1128.
Mehrotra R, Sharma A. 2010. Development and application of a multisite rainfall stochastic downscaled framework for climate change impact assessment. *Water Resource Research* **46**(7): W07526.
Mikkelsen PS, Madsen H, Arnbjerg-Nielsen K, Jørgensen HK, Rosbjerg D, Harremoës P. 1998. 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.
Negra NB, Holmstrøm O, Bak-Jensen B, Sørensen P. 2008. Model of a synthetic wind speed time series generator. *Wind Energy* **11**(2): 193–209.
Olsson J, Burlando P. 2002. Reproduction of temporal scaling by a rectangular pulses rainfall model. *Hydrological Processes* **16**(3): 611–630.
Onof C, Arnbjerg-Nielsen K. 2009. Quantification of anticipated future changes in high resolution design rainfall for urban areas. *Atmospheric Research* **92**(3): 350–363.
Srikantan R, McMahon TA. 1983. Sequential generation of short time-interval rainfall data. *Nordic Hydrology* **14**(5): 277–306.
Thyregod P, Arnbjerg-Nielsen K, Madsen H, Carstensen J. 1998. Modelling the embedded rainfall process using tipping bucket data. *Water Science and Technology* **37**(11): 57–64.
Thyregod P, Madsen H, Carstensen J, Arnbjerg-Nielsen K. 1999. Integer valued autoregressive models for tipping bucket rainfall measurements. *Environmetrics* **10**(4): 395–411.
Wilby RL, Tomlinson OJ, Dawson CW. 2003. Multi-site simulation of precipitation by conditional resampling. *Climate Research* **23**(3): 183–194.
Yang Y. 2005. Can the strengths of AIC and BIC be shared? A conflict between model identification and regression estimation. *Biometrika* **92**(4): 937–950.

 Copyright © 2012 John Wiley & Sons, Ltd.