

Precipitation extremes on multiple time scales - Bartlett-Lewis Rectangular Pulse Model and Intensity-Duration-Frequency curves

—
Authors Response to Reviewers

Christoph Ritschel, Uwe Ulbrich, Peter Névir, Henning W. Rust

August 30, 2017

We are very grateful to three anonymous reviewers for carefully reading and commenting thoroughly on our manuscript. We received highly valuable and constructive comments which very much helped to improve our work and led to new insights. We additionally got plenty ideas for further investigations.

5 In the following, we go point by point through all the comments and reply to them. Reviewers' comments are all repeated in this document, typeset in black. They are individually addressed, typeset in blue. Changes to the original manuscript as resulting from the reviewers comments are repeated here to ease the comparison with the original version; they are typeset in *blue italic*.

10 Due to some comments from the reviewers, we decided to exchange the abbreviation *BLRPM* to *OBL model* in order to distinguish the original Bartlett-Lewis model (*OBL*) from a modified version (*MBL*).

Reviewer 1:

General Comments:

15 This paper investigates the ability of the original Bartlett-Lewis model for estimating extreme rainfall at various levels of aggregation. Unfortunately, the paper is not very novel. It is already known for a long period that the Bartlett-Lewis (BL) models have problems in reproducing extremes, especially at shorter aggregation levels. It is not clear why the authors chose for the Original Bartlett-Lewis (OBL) model, while the Modified Bartlett-Lewis (MBL) model or one of the later versions (e.g. Onof and Wheather, 1994) that were further optimized for addressing the
20 problem of the undergeneration of extremes. An important part of the paper is dealing with the fact that using a short time series for calibration may have an important impact on the statistics described by the observed extremes: the highest extreme may have a much larger return period than the one estimated from the time series. This, of course, is not surprizing, and the shorter the time series used, the higher the potential becomes of facing with extremes that have true
25 return periods much larger than the length of the time series. Yet, this example may be of interest for the scientific community, especially for young researchers starting in the domain of stochastic hydrology. Therefore, I believe this part of the paper may be of interest, though not very novel.

Remarks:

30 Yet, I would like to give some suggestions that may improve this section:

Mayor (1) using the model with 12 extremes, calculate the return period of the highest extreme that was omitted (i.e. the one in year 2007) to frame how extreme this event in 2007 was?

35 We added the following sentence and table to Section 5.3: *Based on the model with parameters estimated from observations without the year 2007 (observed), we obtain return periods for the event “Kyrill” for different durations and find this event to be very rare, especially on short time scales (1-3 hours), see Tab. 3.*

Duration [h]	Probability of exceedance without Kyrill [%]	Return period without Kyrill [years]	Probability of exceedance including Kyrill [%]	Return period including Kyrill [years]
1	1.8×10^{-6}	560000	5.6×10^{-4}	1790
2	4.3×10^{-5}	23000	2.4×10^{-3}	420
3	2.2×10^{-4}	4400	5.4×10^{-3}	185
6	1.6×10^{-3}	630	1.6×10^{-2}	63
12	1.7×10^{-3}	590	2.0×10^{-2}	49
24	3.5×10^{-3}	280	3.5×10^{-2}	29
48	2.0×10^{-2}	50	9.5×10^{-2}	11

Table 1: Return period for the event *Kyrill* as estimated from the observational time series with this particular event left out and included for parameter estimation for different durations.

Mayor (2) Why not redo the same exercise with the Peak-Over-Threshold method, where the threshold is put quite low to ensure a larger number of extremes? This may reduce the uncertainty on the IDF curves as more data are used to fit the parametric model?

40 The POT approach might have given us a larger number of extremes. However, we are not sure to what extent the consistent estimation using all durations simultaneously can be performed for the GPD as it can be (and we do it here) for the GEV. Koutsoyiannis et al. [1998] suggested the duration dependent GPD as well as a model for IDF curves but explicitly states that parameter estimation would have to be carried out via annual maxima and the asymptotic equality of GEV and GPD for extremes, rendering the GPD based approach less interesting for us. Our argument
45 here is that uncertainty can be reduced due to borrowing strength from neighbouring durations by using the *duration dependent* GEV approach.

Minor

- Line 11-12: here it is not clear what is meant with a singular event. Context is not
50 sufficiently provided.

Thanks for the hint! Singular is indeed unclear here. We replace occurrences of *singular* in the text by *rare in the sense given in Tab. 3* or *...rare event (here an event with a return period larger than 1000 years on the hourly time scale) before the table is introduced in section 5.3.*

- 55 • Line 73: mention what version of the BL models is used (i.e. the Original BL model)

We added *original* and (*OBL*) to line 73 and change the notation *OBL model* instead of BLRPM throughout the text.

- Line 147: remove the footnote after the equation as it reads as if (1-p) is put to the power “1”. The text in the footnote can easily be introduced in the sentence.

60 We changed the sentence to *An IDF curve for a given return period $T = 1/(1 - p)$, with p denoting the non-exceedance probability,...*

- Lines 227-228: please introduce a figure to illustrate this.

65 Thanks for the hint. It should (and does now) read in the text *For January, IDF curves from observations and OBL model simulations ...* and not February. The figure for January is provided.

- Line 232: True, but this is a typical problem occurring for too short time series used for extreme value analysis: fitting a distribution to 13 points is questionable!

70 Here we expect that using the simultaneous fit to 9 durations makes this approach more robust. We fit one duration dependent GEV to 117 extreme values (13 years multiplied by 9 durations). They are, however, clearly not independent.

- Line 299: “which may not be reproduced by the BLRPM”: this may be reproducible! Only, its occurrence may be very low causing that this event was never modelled during the short time series generated! What is the return period of this “singular” event based on the model built from all extremes excluding this event?

75 From Tab. 3, one can see that the return period for a comparable event (for 2h duration) is several thousand years. However, we do only simulate 1000 years and probabilities of getting such a strong event in this short time period are low. We suggest a better formulation for this sentence in the introduction: *2) How are IDF curves affected by very rare extreme events which are unlikely to be reproduced with the OBL model for a reasonably long simulation?* and the conclusion *2) How are IDF curves affected by very rare extreme events which are unlikely to be reproduced with the OBL model for a reasonably long simulation? When the year 2007 is excluded from the analysis, the aforementioned discrepancy in January disappears. We conclude that an extreme event which is rare (return period of 23000 yrs) with respect to the time scales of simulation (1000×13 yrs) has the potential to influence the dd-GEV IDF curve as 1 out of 13 values per duration – i.e. one maximum per year out of a 13 years time series – does change the GEV distribution.*

- Line 330: define “relative difference”

To define this term, we changed the beginning of the paragraph to: *Figure 11 shows the relative difference*

$$\Delta = \frac{dd-GEV_{OBL} - dd-GEV_{obs}}{dd-GEV_{obs}} \cdot 100\% \quad (1)$$

90 *between IDF curves (dd-GEV) derived from the OBL model $dd-GEV_{OBL}$ and directly from the observational time series $dd-GEV_{obs}$.*

- Appendix A: please provide information to the reader of what should be learned from the figures presented in the appendix. Nor the appendix or the text sufficiently elaborates on this.

95 Appendix A is referred to twice in the text and gives an overview on estimated OBL model parameters. We consider the information in the table as necessary for reproducible research.

100 With the Figures in Appendix B, we suggest another way of looking at differences in IDF curves which aims to provide a better understanding of model deficiencies in terms of over- or underestimation of return levels. We changed the sentence referring to Appendix B in Sec. 5.2 to *The relative differences in IDF curves given in Fig. 11 (Appendix B) suggest a tendency for the OBL model to underestimate extremes, particularly for large return levels and short durations, similar to results found by, e.g. Verhoest et al. [1997] and Cameron et al. [2000].*

105 Reviewer 2:

General Comments:

This paper demonstrates the use of original Bartlett-Lewis models for simulating rainfall series having precipitation extremes on multiple time scales. I believe it is an interesting paper that confirms some of the problems already indicated for the model used. More is needed in terms of discussion and a clearer extreme-value analysis, possibly involving the examination of other cell intensity distributions and proposed a new version of the model, which they called the Modified Bartlett Lewis (MBL) model. The original Bartlett Lewis model is proved efficient to explain the rainfall characteristics at all time intervals considered (1hr to 24hr) as explained by several authors such as Rodriguez-Iturbe et al. (1988) and Onof (1992), a major deficiency is its inability to reproduce the proportion of dry periods correctly. To overcome this problem, Rodriguez-Iturbe et al. (1988) proposed a new version of the model, which they called the Modified Bartlett Lewis (MBL) model. Although several studies have pointed out limitation of the original model and suggested some improvements. Onof and Wheater (1994a), for example, introduced a two-parameter gamma distribution as opposed to the original Bartlett Lewis model which considers a single parameter exponential distribution to describe the depth of a cell in order to better capture extreme events. However, the problem of underestimation of the extreme values still persists, particularly for lower aggregation levels, as described by Verhoest et al. (1997). Vandenberghe et al. (2010) found that the models demonstrated a too severe clustering of rain events.

125 Comments:

I would recommend the paper to be published after addressing some of the following remarks. I believe that this work could be improved by better demonstrating the advantages of the original and modified models compared to other rainfall generators (for instance, rectangular pulses models better maintain statistics at different aggregation levels), but also give an overview of drawbacks of the model. For instance, Onof and Wheater (1994) introduced a gamma distribution for the depth of a cell in order to better capture extreme events. Verhoest et al. (2010) discusses that problems still remain as infeasible cells (extremely long) sometimes occur. Vandenberghe et al. (2011) found that the models demonstrated a too severe clustering of rain

135 events. Cameron et al. (2000) and Verhoest et al. (1997) found that these models generally underestimate the extreme values, especially for lower aggregation levels. Onof and Wheater (1993) reported problems for return periods greater than the length of the dataset. According to Cowpertwait (1998) this problem could be overcome if higher order properties would be included in the fitting procedure. Besides of being in mentioned above, the authors could validate whether the same problems occur for their simulations.

140 We are grateful for this comprehensive overview on the deficits associated with the original Bartlett-Lewis model (OBL) and modified versions. We used the OBL to gain an understanding of this type of stochastic precipitation models with the aim to use it in a non-stationary context in future research. Drawbacks of the OBL and also of modified versions are discussed in the literature, as mentioned by the reviewer. These deficits of the OBL might vanish (at least
145 partially) if used in a non-stationary context where model complexity is increased as parameters are linked to large scale flow variables. This is, however, not a point to be discussed here.

Besides gaining experience for our future research plans, the manuscript we presented contributes to a) the analysis of extreme precipitation over a range of time scales in a consistent way using duration-dependent IDF curves (to our knowledge, this has been only briefly touched in
150 Verhoest et al. [1997], and to b) the question whether the duration-dependent GEV is suitable to obtain IDF curves for these kind of models.

In the revised manuscript we introduce a paragraph reporting on the above mentioned issues in section 1:

*Due to the high degree of simplifications of the precipitation process, known drawbacks of
155 the OBL model include the inability to reproduce the proportion dry as reported by Rodriguez-Iturbe et al. [1988] and Onof [1992], and underestimation of extremes as found by, e.g. Verhoest et al. [1997] and Cameron et al. [2000], especially for shorter durations. Furthermore, problems occur for return levels with associated periods longer than the time series used for calibrating the model [Onof and Wheater, 1993]. Several extensions and improvements to the model have been
160 made. Rodriguez-Iturbe et al. [1988] introduced the randomised parameter Bartlett-Lewis model, allowing for different types of cells. Improvements in reproducing the probability of zero rainfall and capturing extremes have been shown for this model [Velghe et al., 1994]. A gamma-distributed intensity parameter and a jitter were introduced by Onof and Wheater [1994b] for more realistic irregular cell intensities. Nevertheless, problems still remain as Verhoest et al. [2010] discussed the occurrence of infeasible (extremely long lasting) cells and a too severe clustering of rain
165 events was found by Vandenberghe et al. [2011]. Including third-order moments in the parameter estimation showed an improvement in the Neyman-Scott models extremes [Cowpertwait, 1998]. For the Bartlett-Lewis variant Kaczmariska et al. [2014] found that a randomised parameter model shows no improvement in fit compared to the OBL model for which the skewness was included
170 in the parameter estimation. Furthermore, an inverse dependence between rainfall intensity and cell duration showed improved performance, especially for extremes at short time scales [Kaczmariska et al., 2014]. Here, we focus on the OBL model with and without the third-order moment included. This model is still part of a well-established class of precipitation models and the reduced complexity is appealing as it allows to be used in a non-stationary context [Kaczmariska
175 et al., 2015].*

As mentioned in Section 5.2 of the manuscript, we found the OBL model to underestimate extremes merely for return levels with associated return periods much longer than our observed time series. We report this result now with a reference to Onof and Wheater [1993]. We cannot confirm a significant underestimation associated with short durations as reported by Cameron
180 et al. [2000] and Verhoest et al. [1997], only the tendency is visible in Fig. 11, as is reported in Section 5.2. Small differences are present; we related those, however, to a problem of estimating

a consistent IDF for short durations, see Sect. 5.4.

In a 1000 year simulation with the OBL, we could not find any infeasible cells as mentioned by Verhoest et al. [2010]. They discovered the problem for the modified version of the BL model. In our manuscript, we report in the discussion that in our long OBL simulation, this problem does not occur.

Motivated by this reviewer comment, we included the third moment in our objective function, following Cowpertwait [1998] and using the analytical expression derived by Wheeler et al. [2006], which – as the reviewer mentions – should overcome some problems of the OBL, see Sect. 2 where we added following paragraph:

Following studies by Cowpertwait [1998] and Kaczmarek et al. [2014], we include the third moment in the parameter estimation using analytical expressions derived by Wheeler et al. [2006], replacing the probability of zero rainfall in the objective function. Thus, still 13 moments are used to calibrate the OBL model. Due to comparability with other studies most of our analyses will not include the third moment though. A comparison between IDF curves of the model calibrated with the third moment and with the probability of zero rainfall will be carried out, to discuss the effect of including the third moment.

Compared to using the known problematic probability of zero rainfall [Onof and Wheeler, 1994a], we could not find a systematic improvement related to extremes. This is discussed in the revised version in section 5.2. as follows:

Figure 8 shows the relative difference

$$\Delta = \frac{dd-GEV_{OBL} - dd-GEV_{obs}}{dd-GEV_{obs}} \cdot 100\% \quad (2)$$

between IDF curves (dd-GEV) derived from the OBL model $dd-GEV_{OBL}$ including the third moment in parameter estimation (red lines) or alternatively using the probability of zero rainfall to calibrate the model (blue lines), and directly from the observational time series $dd-GEV_{obs}$ for July and two quantiles: a) 0.5 and b) 0.99. Including the third moment in parameter estimation slightly improves the model extremes for July for all durations and both short and long return periods. Nevertheless, those promising results could not be found for all months (not shown) and thus we cannot conclude that including the third moment in parameter estimation improves extremes in the OBL model in contrast to findings for the Neyman-Scott variant [Cowpertwait, 1998].

Section 2.1) line 109: ...the weights, $(w_i; i = 1, 2, \dots, k)$ which allow more important weight to be given to fitting some sample moments relative to others. Try to give weights given by $w_i = 1/Var(T_i(y))$ where $Var(T_i(y))$ represents the i^{th} diagonal elements of the covariance matrix of the summary statistics.

Vanhaute et al. [2012] investigated different objective functions specified in the following (rewritten using a notation consistent with our manuscript):

$$Z(\boldsymbol{\theta}; \mathbf{T}) = \sum_{i=1}^k w_i [\tau_i(\boldsymbol{\theta}) - T_i]^2 \quad (OF1)$$

$$Z(\boldsymbol{\theta}; \mathbf{T}) = \sum_{i=1}^k \left\{ \left[1 - \frac{\tau_i(\boldsymbol{\theta})}{T_i} \right]^2 + \left[1 - \frac{T_i}{\tau_i(\boldsymbol{\theta})} \right]^2 \right\} \quad (OF2) \quad (3)$$

$$Z(\boldsymbol{\theta}; \mathbf{T}) = \sum_{i=1}^k 1/Var[T_i] [\tau_i(\boldsymbol{\theta}) - T_i]^2 \quad (OF3)$$

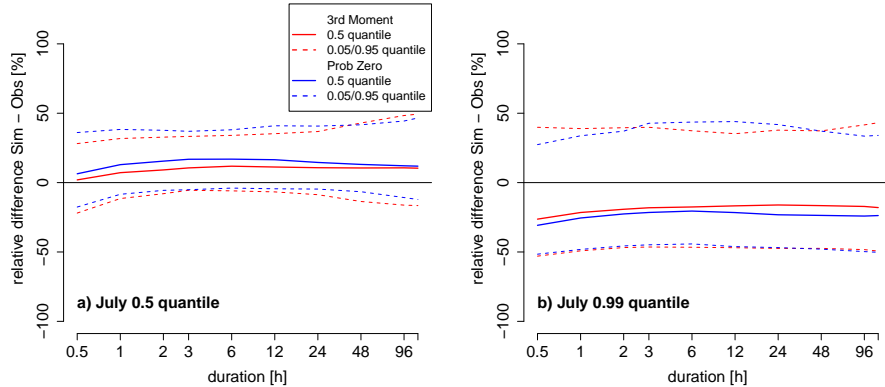


Figure 1: Relative differences between observed and simulated return levels obtained with including the third moment (red) and with using the probability of zero rainfall (blue) in parameter estimation for a) July 0.5 quantile and b) July 0.99 quantile. Dotted lines show the 0.05 and 0.95 quantile range of 1000 simulations.

with the moments $\tau_i(\boldsymbol{\theta})$ derived from model parameters $\boldsymbol{\theta}$ and the empirical moments T_i estimated from the time series.

Here, we use an objective function based on OF2, using a ratio between analytic and empirical moments. In this formulation, first and second order properties are normalised by their characteristic order of magnitudes and are thus comparable. A scaling with variances as suggested by the reviewer is thus not necessary for this particular case. Additionally, we use the weights w_i from OF1 to emphasize the first moment similarly to Cowpertwait et al. [1996a], see Sect. 2. We are, however, aware of objective functions like OF1 with weights being the variances of the moments as proposed by the reviewer and also by Kaczmarska et al. [2015] for the non-stationary setting; An approach we plan to pursue in the future.

Section 2 2) Give more info on the boundary constraints identified for the parameters of original model that contribute to the stability in the parameter estimates. For the original model, the values of λ that are only considered ranges from 0.01 to 0.05.

Thanks, that is definitively needed for reproducible research. We add the following to Appendix A:

Estimation of OBL model parameters follow the boundary constraints: For those parameter

Parameter	Lower boundary	Upper boundary
λ	0.004 [h^{-1}]	1 [h^{-1}]
γ	0.01 [h^{-1}]	10 [h^{-1}]
β	0.01 [h^{-1}]	100 [h^{-1}]
η	0.01 [h^{-1}]	100 [h^{-1}]
μ_x	1×10^{-9} [mm/h]	100 [mm/h]

Table 2: Boundary constrained used in OBL model parameter estimation.

ranges, numerical optimisation mostly converged into a global minimum. For the model variant using the third moment in the OF, no constraints are used.

230 **Section 5 Results: 1.** From results listed in Table 1, it is interesting to observe the higher number of storms with high cell intensity and this is contrary to our prior knowledge about less storm arrivals in dry periods like June. The occurrence of heavy rain in a short duration often induces flash floods in the city area. Form data, it is found the values of cell arrival based on the original model is smaller with high rainfall intensities, particularly for June. This implies that
235 there is a substantial enough cell overlap which could bring extreme rainfall events. Thus, the occurrence of these realistic rainfall cells, whereas, at the hourly time scale, the annual maxima do not generally result from this model.

Thank you for pointing us to this interesting observations, we include the following in Section 5.1: *During summer months, we observe very intensive cells ($\hat{\mu}_x$ between 4mm/h and 8mm/h).
240 However, in June and August, storm duration is relatively short ($\hat{\gamma}$ between 0.25/h and 0.35/h) which can be interpreted as short but heavy thunderstorms which are typically observed in this region in summer [Fischer et al., 2017].* This passage replaces following sentences in Section 5.1 in the original manuscript:

245 Large mean intensities $\hat{\mu}_x$ and short mean cell life-times $1/\hat{\eta}$ in summer correspond to precipitation being dominated by convective events. Similar, the mean cluster life-time $1/\hat{\gamma}$ decreases in summer, whereas the mean cell generation rate $\hat{\beta}$ increases.

2. Please check how the extreme events of the original model look like and compare this to the extremes of the historical series. From this you may conclude what is the problem rather
250 than guessing that it has to do with the nature of the rainfall (maybe it is a shortcoming of the model instead! E.g. Verhoest et al. (2010))

We checked extreme events of the OBL model and visually compare them to the extremes of the historical site. We add the following figure and text to the manuscript to section 5.2. *As an example, we show segments of time series including the maximum observed/simulated rainfall in July for durations 1h, 6h and 24h as observed (RR_{obs}) and simulated (RR_{OBL}) in Fig. 7.
255 Parts of the observed and simulated rainfall time series corresponding to the extreme events for the three different durations are shown in the left and right column, respectively. Additionally the middle column shows the simulated storms and cells generating this extreme event in the simulated time series. Note, that we show only one simulation as an example; visual inspection of several other simulated series share the main features and are not reproduced here. For all
260 durations, the extremes are a result of a single long-lasting cell with high intensity. In contrast to an analysis based on the random parameter BL model [Verhoest et al., 2010], these cells are neither unrealistic long nor have an unrealistic high intensity.*

Reviewer 3:

265 General Comments

1. The focus on IDF curves as a characteristic of mechanistic models appears to be novel and of wide relevance to hydrological modelling, climate impact assessment and risk estimation. The focus on short duration (5 minute) extremes is also of particular relevance. I therefore think this research is suitable for this publication and would be of general interest to its readership.

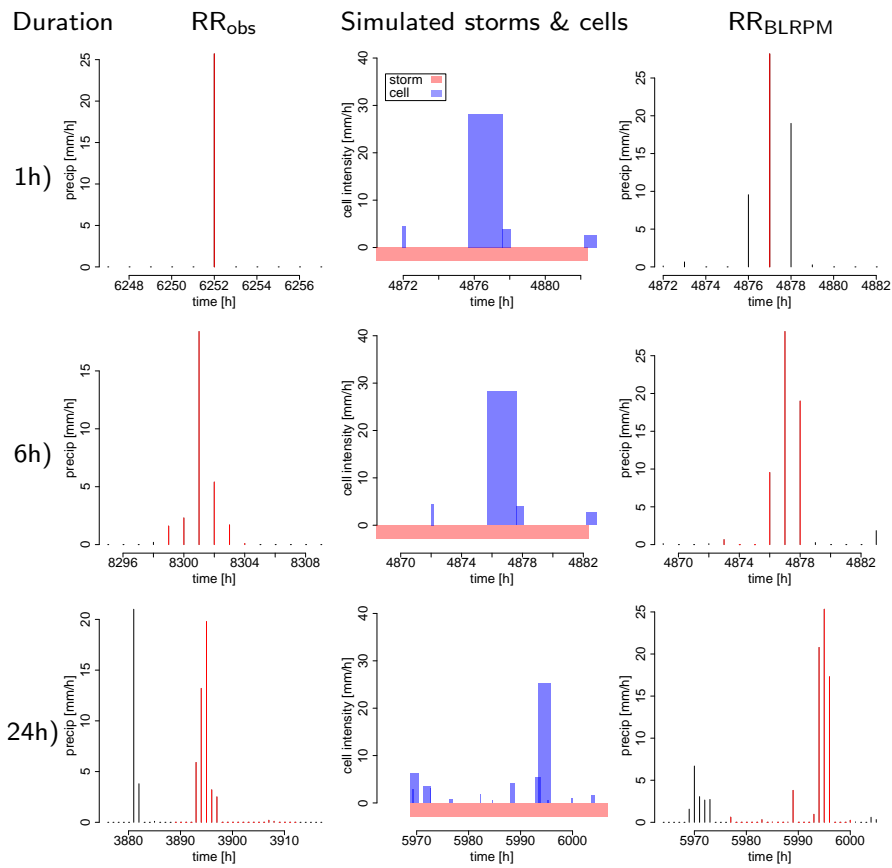


Figure 2: Visualization of July extremes as observed (RR_{obs} , left column) and simulated by the OBL model (RR_{OBL} , right column). Shown are short segments including the maximum observed/simulated rainfall (red vertical bars) at durations 1h (top row), 6h (middle row) and 24h (bottom row). Additionally, the middle column shows the simulated storms (red rectangles) and cells (blue rectangles) corresponding to the extreme event of the simulated time series.

270 2. The paper addresses three research questions which are clearly set out in the introduction. Each question is then addressed in turn in the discussion and conclusions. The questions are as follows:

I. “Is the OBL model able to reproduce the intensity-duration relationship found in observations?” The authors use a depth-dependent GEV distribution (dd-GEV) to estimate
275 extremes across different durations – it is assumed that “across different durations” means “across different temporal scales”. Optimisation of the dd-GEV parameters is performed using random sampling from a Latin-Hypercube which appears to be a new method for calibrating these models and is referred to as the depth-dependent GEV approach. This approach is used to construct IDF
280 curves from the observations, and 1000 OBL model realisations of the same length. Typically when we want to estimate extremes from a rainfall model we would sample annual maxima directly from long duration simulations without then using a second extreme value model such as GEV or GP. However, in this case it seems appropriate to apply the dd-GEV for two reasons: 1. to enable direct comparison with the IDF curves from observations, and 2. because the dd-GEV method uses extremes across different scales in fitting. That said, it is not clear from the
285 methodology set out in 5.2 at what scales rainfall has been simulated; is it the same as those used in fitting (i.e., 1, 3, 12, and 24 hrs)? This could be made clearer by the authors.

As the reviewer wrote, we use a parametric approach to obtain a consistent IDF curve based on a block-maxima approach and a duration-dependent GEV. This idea is based on work by Koutsoyiannis et al. [1998] and later taken up by Soltyk et al. [2014]. The main advantage is to
290 exploit the smoothness in the IDF curve for a more robust estimation. Parameter estimation is carried out by numerically optimising an objective function based on an approximation to the likelihood; the problem of local minima is taken care of by using a latin-hypercube resampling of initial guesses for the parameter optimisation.

From continuous cell simulation, rainfall series have been obtained by aggregating cell rainfall
295 to 1h (minimum duration) and further on to match the duration used for the observed series. We thus include the following sentences at the end of Sect. 2 and augment a sentence at the beginning of Sect. 5.2, respectively

Simulations with the OBL model are in continuous-time on the level of storms and cells. We aggregate the resulting cell rainfall series to hourly time series.

300 *Monthly block-maxima for every month in the year are drawn for various durations (1h, 3h, 6h, 12h, 24h, 48h, 72h, 96h) from the observational time series and 1000 OBL model simulations of same length.*

The authors note in Section 5.2 (lines 220-2) and in the conclusion (lines 292-3) that the OBL
305 model tends to under-estimate the extremes. The under-estimation of extremes by mechanistic rainfall models (both Bartlett-Lewis and Neyman-Scott variants), especially at fine temporal scales, is a known issue and the authors’ findings are entirely consistent with this. The discussion would be greatly improved by drawing a broader interpretation of the results with comparison with other studies that show under-estimation of extremes by mechanistic models. In particular,
310 is there something to be gained by estimating fine-scale extremes in this way?

Motivated already by the first comment of reviewer 2, we related our findings to a broader spectrum of literature, see our answer above.

For users, an IDF curve gives a broad and immediate overview about how much (intensity) rain over a period of time (duration) is likely (frequency) to fall. Previous studies mainly focus on

315 Gumbel plots in reference to extreme value analysis. Therefore we believe the presented frame-
work using consistent IDF curves based on a duration-dependent GEV together with stochastic
precipitation models can contribute to the community.

**II. “How are IDF curves affected by a singular extreme event which might not be
reproducible with the BLRPM?”** BL model parameters are estimated using central mo-
320 ments of the rainfall data therefore it is very likely that this one single extreme will not have
as much influence on the estimation of BL model parameters as it does on dd-GEV parameters
from observations. And indeed, the authors show that the problem with January disappears
when this event is taken out. The reader is however left with the impression that the implication
is that this event is treated as suspicious information, i.e. that it is fine to take out this largest
325 observation because it is so abnormally larger than any other observed hourly rainfall depth. I
don’t think that the authors meant this to be the case, but it should be clarified in the text that
the section in which this largest value is taken out does not carry the implication that it is OK
to take out the largest value because the event is in some sense ‘abnormal’.

Thanks for pointing this potential problem out! We did not intend to motivate other researchers
330 to take out a “suspicious” date as the winter storm Kyrill in January 2007. Instead we wanted
to demonstrate the OBL model’s inability to capture characteristics of an event which is much
larger in magnitude than the majority of the other events. On the other hand, we showed that
the model is generally able to reproduce extreme precipitation events if they are well represented
in the underlying data. We augment the first paragraph of Sect. 5.3: *The convective cold front*
335 *passage of Kyrill accounted for a maximum intensity of 24.8mm rainfall per hour, whereas the*
next highest value of the remaining Januaries would be 4.9mm rainfall per hour in 2002 and
thus being more than 5 times lower than for Kyrill. We construct another data set without the
extreme event due to Kyrill, i.e. without the year 2007. The intention of this experiment is not
to motivate removal of an “unsuitable” value. We rather want to show that the OBL model is in
340 *generally able to reproduce extremes; it is, however, not flexible enough to account for a single*
event with magnitude far larger than the rest of the time series. . . .

This issue brings us to an important problem with the authors’ analyses: the data set of 13
years (then reduced to 12 years) is rather short to be doing extreme-value analysis (typically,
345 a peak-over-threshold approach would normally be preferred for such a short dataset. Perhaps the
authors’ aim is to bring out the greater usefulness of making use of a rainfall model when the
data set is not long enough, in which case this should be stated.

We admit that an extreme value analysis would benefit from a longer time series, which is
unfortunately not available for this case study. With respect to the POT approach, please see
350 our answer to reviewer 1, Mayor comment 2. It is not our aim to use the OBL model as a relief
of the short data series problem. As mentioned in the last comment/answer, we also need a
long series to estimate OBL parameters in a way that extremes with a long return period are
sufficiently well reproduced.

III. “Is the parametric extension of the GEV a valid approach to obtain IDF curves?”

355 Here the authors test the validity of the dd-GEV approach to estimating IDF curves by com-
paring IDF curves obtained from 50 realizations of 1000 years duration from the BL models with
GEV estimates from the same simulations. There is an important underlying hypothesis here,
namely that the BL model has now been adopted as an accurate representation of the distribu-

tion of rainfall (in particular extremes), but we know that this is not true from the problems
360 identified in the analysis of BL’s IDF curves. So it is important to qualify the scope of this third
research question to make it clear that it is an analysis conditional upon a hypothesis that is
only approximately true.

Thanks for the hint! Indeed, we do not take the OBL as a representative for the observed rainfall
but as a tool to obtain long artificial series to be used in a model-world study. We change the
365 first sentence in Sect. 5.4 to: *In the frame of a model-world study, long time series simulated
with the OBL model can be used to investigate adequacy of the dd-GEV model conditional on the
simulated series.*

This issue also has a bearing upon the interpretation of the results. For instance, when they
370 identify an under-estimation of 10 and 100 year hourly extremes in January and July, the au-
thors conclude that this is due to poor representation of the dd-GEV IDF curves at these scales
which is described as flattening. However, this result is also consistent with the known issue
of mechanistic models under-estimating fine-scale (hourly and sub-hourly) extremes yet there is
375 no discussion to this effect. It is potentially encouraging that the estimation of fine-scale ex-
tremes with dd-GEV IDF curves from BL model simulations does not show the underestimation
ordinarily obtained from mechanistic models, therefore the authors could explore this in their
discussion.

Please note, that we now take a single fixed set of parameters to simulate 50 very long (1000yrs)
series of rainfall surrogates. Based on these series, we compare two strategies for estimating
380 return levels for different durations: the duration-dependent GEV (dd-GEV) and individual
duration GEV approach. Problems of the OBL to represent observed extremes do not play a
role here. However, we suggest that the observed effect for short durations indeed needs to be
explored in a further analysis.

A further issue potentially lies in the estimation of confidence intervals. There may be over-
385 confidence in the extreme value estimates and IDF curves presented in Figure 8. Confidence
intervals are estimated from 50 realisations from the BL models. However, GEV extreme value
estimates from each realisation would have an associated credible interval which is not shown.
It is possible that if this were, then there would be greater overlap in estimation by the two
390 methods and the marginal differences would not be statistically significant.

Here, we assume that the reviewer uses the term “credible intervals” for the statistical uncertainty
intervals, typically associated with any estimator, e.g., here for estimated GEV parameters (or
the return levels derived from them). These intervals represent sampling uncertainty, i.e. the
uncertainty due to having a particular sample and not the full population available. These
395 estimates can and will vary if another – equally likely but different – sample had been observed.
It is exactly this effect which we cover with presenting various samples – i.e. various pseudo-
observations – to the GEV estimator. We can do so only in this model-world experiment where
we have the model to generate these series. This way of presenting sampling uncertainty is
equivalent (at least in interpretation) to the uncertainty intervals based on asymptotic properties
400 of the maximum-likelihood estimator. The latter are typically associated with the GEV or other
estimators. However, these asymptotic properties do not hold for the dd-GEV approach and we

need a different approach to quantify sampling uncertainty: in this model-world study, we have the possibility to obtain more than one sample and can thus estimate the sampling uncertainty directly from different samples.

405 **Specific Comments:**

- V. The authors state on lines 44-5 that “Due to the high degree of simplification of the precipitation process, the model is known to have difficulties in the extremes.” It is not clear that this is why mechanistic models have a tendency to under-estimate short duration extremes, and many hypotheses have been put forward to address this exact problem in the literature since their inception in the late 1980s. The authors make a valid point, but it could be enhanced with some references and broader discussion.

References will be included here as discussed above in an answer to reviewer 2.

- IX. On line 73 the authors highlight that they have chosen to use the original 5 parameter BL model. It would be good to give some justification for using this model variant over the randomised versions of the models, especially given that Kaczmariska, Isham & Onof, (2014) present a new randomised model with enhanced estimation of fine-scale (sub-hourly) extremes.

We used the original BL model to gain an understanding of this type of stochastic precipitation models as we plan to use it in a non-stationary setting Kaczmariska et al. [2015], see also our answer to reviewer 2, first comment.

- XI. On line 87 the authors refer to a “time continuous step function”. Should this be “continuous-time”?

Thanks, changed.

- XII. On line 94 the authors comment that the Neyman-Scott model is “...motivated from observations of the distribution of galaxies in space”. This sounds fascinating although its relevance to rainfall simulation is perhaps somewhat removed. This statement should be reformulated with an appropriate reference.

Neyman and Scott developed a model to represent galaxies that tend to cluster. Later the very same model was found to be useful in other contexts, such as rainfall. We decide to leave this original reference in the text as it shows the origins of this model. References to Poisson-cluster models for rainfall are to be found in various places in our manuscript.

- XIII. The sentence on lines 97-9 requires further elaboration.

Similar to an answer to a comment of reviewer 2, we extended this part as follows: *Due to known drawbacks of the OBL model several improvements and extensions have been made in the past: Rodriguez-Iturbe et al. [1988] introduced the random parameter model, allowing for different type of cells, and additionally Onof and Wheeler [1994b] used a jitter and a gamma-distributed intensity parameter to account for a more realistic irregular shape of the cells. Cowpertwait et al. [2007] improved the representation of sub-hourly time scales by adding a third layer, pulses, to the model. Non-stationarity has been addressed by Salim and Pawitan [2003] and Kaczmariska et al. [2015]. Applications of these kind of models include the implementing of copulas to investigate wet and dry extremes [Vandenberghe*

et al., 2011, Pham et al., 2013], regionalisation [Cowpertwait et al., 1996a,b, Kim et al., 2013] and accounting for interannual variability [Kim et al., 2014].

- 445 • XIV. Figure 2: What is the meaning of the red? Is it the duration of the cell generating time (the time during which the storm is active)? And how does it contrast with the blue?

The top part of the figure represents a typical OBL simulation of cell clusters, drawn in red, and cells, drawn in blue. The red color corresponds to the life time of the cell cluster or usually referred to as storm. Hereby the vertical extensions of the storm has no physical meaning and only serves for better illustration. During its life time the storm generates rainfall cells (blue). Horizontally illustrated is the cell's life time and during its life time its constant intensity is illustrated by the vertical extension of the cell.

450

- XV. In Section 2 the authors introduce the BL models and their chosen calibration strategy. On lines 108-10 they highlight their choice of weights with $w_i = 100$ being applied to the first moment T_i (mean). In my experience the mean is usually very well represented by the BL model therefore it is unclear why the authors should want to up-weight this moment so much compared with the others. Given that the authors appear to be using a Generalised Method of Moments, it might be better to weight the summary statistics by the inverse of their observed variance (see)
- 455

As the same point has been risen by reviewer 2, we refer to our answer above.

- 460 • XVI. In lines 123-6 the authors discuss non-identifiability of model parameters although they don't mention if they've checked this for their own calibrations. This could be done by estimating parameter uncertainty or producing profile objective functions on model parameters.

We did check the non-identifiability and came to the conclusion the symmetrised objective function is less likely to lead the optimization algorithm into local minima. In five out of six cases the numerical optimization lead to the same (and likely the global) minimum with same parameter values. To our understanding, profile objective functions would inform about *sampling uncertainty* for the given minimum of the OF.

465

- XVII. Line 151: The notation should read $IDF_{T_2}(d) > IDF_{T_2}(d)$.

470 Thanks for the hint!

- XVIII. Line 160: What is meant by 'such a shape parameter ? Is the claim that is also independent of the scale (duration) ? Is that true?

After re-parametrising the GEV parameters to $\tilde{\mu} = \mu/\sigma_d$, $\tilde{\mu}$ and the shape parameter ξ are approximately independent of the duration d [Koutsoyiannis et al., 1998]. Please note, that these are only approximations but in the mentioned study it has been shown, that these approximations seem to be well justified.

475

- XIX. It's not clear from the information provided exactly how equation 5 is derived. If this is derived in a previous publication this should be clearly stated and referenced.

Given equations (3) and (4), one can introduce the duration dependent scale parameter σ_d into equation (3). It results:

480

$$F(x; \tilde{\mu}, \sigma_d, \xi) = \exp \left\{ - \left[1 + \xi \left(\frac{x}{\sigma_d} - \tilde{\mu} \right) \right]^{\frac{-1}{\xi}} \right\}. \quad (4)$$

Please note, that in the first version of this manuscript the tilde over μ was missing. This derivation has been made by Koutsoyiannis et al. [1998] and used, e.g. by Soltyk et al. [2014]. This is mentioned in the manuscript.

- 485
- XX. Line 164: It is not clear why there are two extra parameters. It would seem that you are placing several GEV fits (one for each scale) with 3 parameters each, by one fit with 4 parameters (?)

490

Introducing the duration-dependent scale parameter σ_d into the GEV framework leads to two additional parameters (θ and η) and a total of five parameters. These additional parameters describe the dependence of the scale parameter σ_d on the duration d . As the reviewer mentions, it is indeed possible with this formulation to estimate the IDF relationships over all durations consistently with one single model. Benefit of this approach is a) consistency, in the sense that different quantiles cannot cross along the duration axis, and b) strength in parameter estimation is borrowed from neighbouring durations.

- 495
- XXI. In Section 4 it would be useful to identify the gauge resolution. It would also be useful to provide a sentence justifying the choice of gauge location.

The gauge resolution is one minute, see Sect. 4. The location is chosen due to its vicinity to our institute and interest in local rainfall characteristics, as well as the easy data availability.

- XXIII. Line 178: explain why a data set with 13 years only was chosen

500

As mentioned we were interested in local rainfall characteristics in the vicinity of our workplace. Therefore we chose a time series from our weather station in botanical garden Berlin. Also we were interested in the question if a short time series like this can be used for this kind of studies and if it would be sufficiently long enough to gain information about its extreme value distribution. It is known that long rainfall time series with such a high temporal resolution are sparse and many stations do not have long records and thus it is an interesting problem if extreme value distributions can already be obtained from short series. Thus, this study helps in investigating this issue.

505

- XXV. In Section 5.2, line 210 the authors point the reader to a dotted line in Fig. 5 for IDF curves from observations. In the figure legend, the dotted line is for the IDF curves from BLRPM simulations. This needs to be corrected.

510

Thanks for the hint, this mistake was corrected.

- XXVI. In Section 5.2, line 227 the authors point the reader to February in their discussion of IDF curves in Fig. 5. I think the authors mean January as curves are only presented for January, April, July and October. The authors do the same on line 293 in the conclusions.

515

Yes, February was put wrongly here and January was meant. This is corrected in the revised manuscript.

- XXXII. In the conclusions on lines 314-7 the authors state that they do not find the BLRPM producing unrealistically high precipitation amounts as discussed for the random- η model by Verhoest et al., (2010). The generation of unrealistically high extremes by the modified (random- η) model is specific to that model and is therefore not relevant here as the authors have used the original 5 parameter model.
- 520

To our knowledge the occurrence of unrealistically high extremes as mentioned by Verhoest et al. (2010) was never investigated for the OBL model and thus we gave it a check. This point was also raised by Reviewer 2, see our answer above.

Marked-up manuscript version

525 Abstract

For several hydrological modelling tasks, precipitation ~~time-series~~ time series with a high (i.e. sub-daily) resolution are indispensable. This data is, however, not always available and thus model simulations are used to compensate. A canonical class of stochastic models for sub-daily precipitation are Poisson-cluster processes, with the original Bartlett-Lewis ~~rectangular-pulse~~ ~~model (BLRPM)-(OBL) model~~ as a prominent representative. The ~~BLRPM-OBL model~~ has been shown to well reproduce certain characteristics found in observations. Our focus is on intensity-duration-frequency relationship (IDF), which are of particular interest in risk assessment. Based on a high resolution precipitation ~~time-series~~ time series (5-min) from Berlin-Dahlem, ~~BLRPM-OBL model~~ parameters are estimated and IDF curves are obtained on the one hand directly from the observations and on the other hand from ~~BLRPM-OBL model~~ simulations. Comparing the resulting IDF curves suggests that the ~~BLRPM-OBL model~~ is able to reproduce main features of IDF statistics across several durations but cannot capture ~~singular-rare~~ events (here an event ~~of magnitude 5 times larger than the second largest event~~ with a return period larger than 1000 years on the hourly time scale). Here, IDF curves are estimated based on a parametric model for the duration dependence of the scale parameter in the ~~General~~ Generalised Extreme Value distribution; this allows to obtain a consistent set of curves over all durations. We use the ~~BLRPM-OBL model~~ to investigate the validity of this approach based on simulated long time series.

1 Introduction

545 Precipitation is one of the most important atmospheric variables. Large variations on spatial and temporal scales are observed, i.e. from localised thunderstorms lasting a few tens of minutes up to mesoscale hurricanes lasting for days. Precipitation on every scale ~~affect~~ affects everyday life: short but intense extreme precipitation events challenge the drainage infrastructure in urban areas or might put agricultural yields at risk; long-lasting extremes can lead to flooding [Merz et al., 2014]. Both, short intense and long-lasting large-scale rainfall can lead to costly damages, e.g. the floodings in Germany in 2002 and 2013 [Merz et al., 2014], and are therefore ~~subject of~~ the object of much research.

Risk quantification is based on an estimated frequency of occurrence for events of a given intensity and duration. This information is typically summarised in an Intensity-Duration-Frequency (IDF) relationship [e.g., Koutsoyiannis et al., 1998], also referred to as IDF curves. These curves are typically estimated from long observed precipitation ~~time-series~~ time series, mostly with a sub-daily resolution to include also short durations ~~into~~ in the IDF relationship. These are indispensable for some hydrological applications, e.g., extreme precipitation characteristics are derived from IDF curves for planning, design and operation of drainage systems, reservoirs and other hydrological structures. One way to obtain IDF curves is modelling block-maxima for a fixed duration with the generalised extreme value distribution [GEV, e.g., Coles, 2001]. Here, we employ a parametric extension to the GEV which allows a simultaneous modelling of extreme precipitation for all durations [Koutsoyiannis et al., 1998, Soltyk et al., 2014].

565 Due to a limited availability of observed high-resolution records with adequate length, simulations with stochastic precipitation models are used to generate series for subsequent studies [e.g., Khaliq and Cunnane, 1996, Smithers et al., 2002, Vandenberghé et al., 2011]. The ~~advantage~~ advantages of stochastic models are their comparably simple formulation and their low computational costs allowing to quickly generate large ensembles of long precipitation ~~time-series~~ time

series. A review of these models is given in, e.g., Onof. et al. [2000] and Wheater et al. [2005]. A canonical class of sub-daily stochastic precipitation models are Poisson cluster models with the original Bartlett-Lewis ~~rectangular pulse model (BLRPM)~~ (OBL) model as a prominent representative [Rodriguez-Iturbe et al., 1987, 1988, Onof and Wheater, 1994b, Wheater et al., 2005]. The ~~BLRPM-OBL model~~ has been shown to be able to well reproduce certain characteristics found in precipitation observations [Rodriguez-Iturbe et al., 1987].

Due to the high degree of ~~simplification~~ simplifications of the precipitation process, ~~the model is known to have difficulties in extremes~~ known drawbacks of the OBL model include the inability to reproduce the proportion dry as reported by Rodriguez-Iturbe et al. [1988] and Onof [1992], and underestimation of extremes as found by, e.g. Verhoest et al. [1997] and Cameron et al. [2000], especially for shorter durations. Furthermore, problems occur for return levels with associated periods lonter than the time series used for calibrating the model [Onof and Wheater, 1993]. Several extensions and improvements to the model have been made. Rodriguez-Iturbe et al. [1988] introduced the randomised parameter Bartlett-Lewis model, allowing for different types of cells. Improvements in reproducing the probability of zero rainfall and capturing extremes have been shown for this model [Velghe et al., 1994]. A gamma-distributed intensity parameter and a jitter were introduced by Onof and Wheater [1994b] for more realistic irregular cell intensities. Nevertheless, ~~it~~ problems still remain as Verhoest et al. [2010] discussed the occurrence of infeasible (extremely long lasting) cells and a too severe clustering of rain events was found by Vandenberghe et al. [2011]. Including third-order moments in the parameter estimation showed an improvement in the Neyman-Scott models extremes [Cowpertwait, 1998]. For the Bartlett-Lewis variant Kaczmarska et al. [2014] found that a randomised parameter model shows no improvement in fit compared to the OBL model for which the skewness was included in the parameter estimation. Furthermore, an inverse dependence between rainfall intensity and cell duration showed improved performance, especially for extremes at short time scales [Kaczmarska et al., 2014]. Here, we focus on the OBL model with and without the third-order moment included. This model is still part of a well-established class of precipitation models and ~~thus subject to investigation in this paper~~ the reduced complexity is appealing as it allows to be used in a non-stationary context [Kaczmarska et al., 2015].

The ~~BLRPM-OBL model~~ and intensity-duration-frequency relationships are of particular interest to hydrological modelling and impact assessment. In the following, we address the three research questions by means of a case study: 1) Is the ~~BLRPM-OBL model~~ able to reproduce the intensity-duration relationship found in observations?, 2) How are IDF curves affected by ~~singular very rare~~ extreme events which ~~might not be reproducible with the BLRPM~~ are unlikely to be reproduced with the OBL model for a reasonably long simulation? and 3) is the parametric extension to the GEV a valid approach to obtain IDF curves? For the class of multi-fractal rainfall models, question 1) has been addressed by Langousis and Veneziano [2007].

Section 2 gives a short overview ~~the BLRPM variant of the OBL model~~ used here, Sect. 3 briefly explains IDF curves and ~~discusses a parametric model to estimate how to obtain~~ them from precipitation time series. This is followed by a description of the data used (Sect. 4). The results section (Sect.5) starts with a description of the ~~BLRPM parameter estimates~~ estimated OBL model parameters for the case study area (Sect. 5.1), ~~investigates the capability discussed then the ability~~ of this model to reproduce IDF ~~curves~~ curves (Sect. 5.2) and ~~discusses~~ the influence of a ~~singular rare~~ extreme (Sect. 5.3) ~~It and~~ closes with a comparison of ~~singular duration GEV quantiles with~~ the duration dependent GEV approach to IDF curves with individual duration GEV quantiles (Sect. 5.4). Section 6 discusses results and concludes the paper.

615 2 Bartlett-Lewis rectangular pulse model

From early radar based observations of precipitation, a hierarchy of spatio-temporal structures was suggested by Pattison [1956] and Austin and Houze [1972]: intense rain structures (denoted as cells) tend to form in the vicinity of existing cells and thus cluster in larger structures, so called storms or cell clusters. Occurrence of these cells and storms (cell clusters) can be described using Poisson-processes. This makes the Poisson-cluster process a natural approach to stochastic precipitation modelling.

The idea of modelling rainfall with stochastic models exists since Le Cam [1961] modelled rain gauge data with a Poisson-cluster process. Later, Waymire et al. [1984] and Rodriguez-Iturbe et al. [1987] continued the development of this type of model. Poisson-cluster rainfall models are characterised by a hierarchy of two layers of Poisson processes.

Similarly to various others studies [e.g., Onof and Wheater, 1994a, Kaczmarska, 2011, Kaczmarska et al., 2015], we choose the [original Bartlett-Lewis rectangular pulse model \(BLRPM\) \(OBL\) model](#), a popular representative of Poisson-cluster models with a set of five physically interpretable parameters [Rodriguez-Iturbe et al., 1987]. At the first level cell clusters (storms) are generated according to a Poisson-process with a cluster generation rate λ and an exponentially distributed life-time with expectation $1/\gamma$. Within each cell cluster, cells are generated according to a Poisson process with cell generation rate β and exponentially distributed life-time with expectation $1/\eta$, hence the term Poisson-cluster process. Associated with each cell is a precipitation intensity being constant during cell life-time, exponentially distributed with mean μ_x . The constant precipitation intensity in one cell gave rise to the name rectangular pulse. Model parameters are summarised in the parameter vector $\theta = \{\lambda, \gamma, \beta, \eta, \mu_x\}$. [Simulations](#)

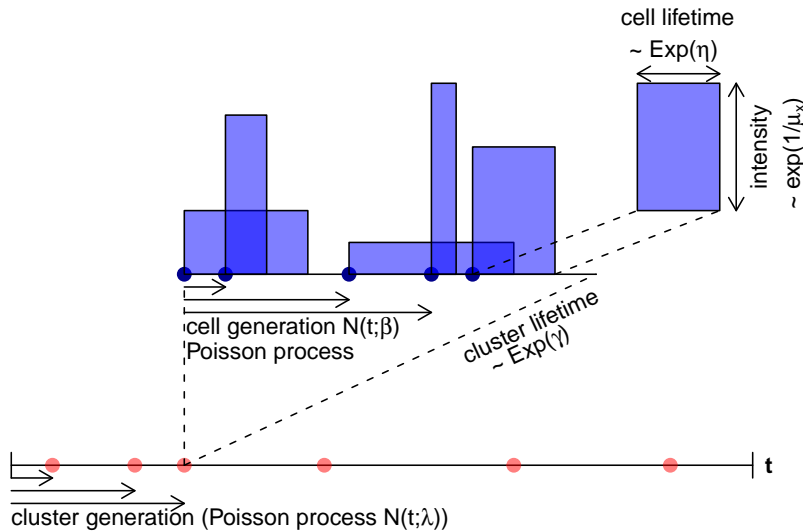


Figure 3: Scheme of the [Bartlett-Lewis rectangular pulse OBL](#) model. A similar scheme can be found in Wheater et al. [2005]

[with the OBL model are in continuous-time on the level of storms and cells. We aggregate the resulting cell rainfall series to hourly time series.](#) Figure 3 shows a sketch of the [BLRPM-OBL](#)

640 model illustrating the two levels: start of cell-clusters (storms) are shown as red dots, within
 each clusters, cells (blue rectangular pulses) are generated during the clusters' life-time starting
 from the cluster origin. The cells' lifetime is shown as horizontal and their precipitation intensity
 as vertical extension in Fig. 3; cells can overlap. This continuous-time-continuous-time
 645 model yields a sequence of pulses (cells) with associated intensity, see Fig. 4 (top panel). Adding up the
 intensity of overlapping pulses yields a time-continuous-continuous-time step-function (Fig. 4,
 middle panel). Although time continuous, this function is not continuously differentiable in
 time due to the rectangular pulses. Observational time series are typically also not continuously
 differentiable as they are discretised in time. Summing up the resulting continuous series for
 discrete time intervals makes it comparable with observations and renders unimportant the arti-
 ficial jumps from the rectangular pulses present in the continuous series, Fig. 4 (bottom layer).

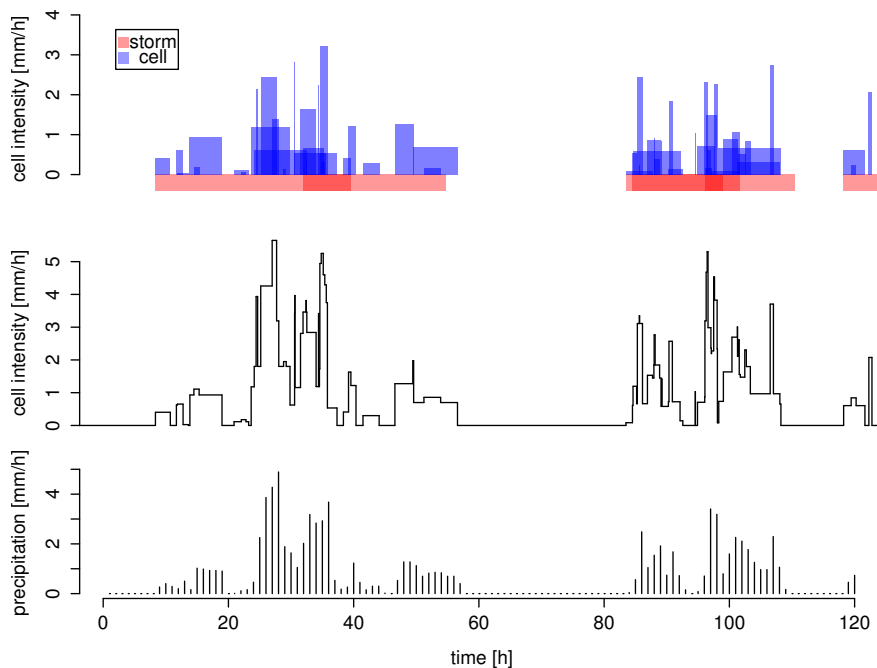


Figure 4: Example realisation of the BLRPMOBL model. Top layer shows the con-
 tinuously simulated storms and cells by them model. The plot in In the middle layer
 650 combines the cell intensities to continuous-time-are combined with a step function. The bot-
 tom layer shows the aggregated artificial precipitation series summed up for discrete
intervalseries. Parameters used for this simulationUsed parameters: $\lambda = 4/120, h^{-1}, \gamma = 1/15,$
 655 $h^{-1}, \beta = 0.4 \lambda = 4/120 h^{-1}, h^{-1}, \gamma = 1/15 h^{-1}, \eta = 0.5 \beta = 0.4 h^{-1}, h^{-1}, \eta = 0.5 h^{-1}, \mu_x = 1 \text{ mm}/h^{-1}.$

An alternative to the Bartlett-Lewis process is the Neyman-Scott process [Neyman and Scott,
 1952]. The latter is motivated from observations of the distribution of galaxies in space. In
 the Neyman-Scott process cells are distributed around the centre of a cell cluster. Both are
 prototypical models for sub-daily rainfall and are discussed in more detail in Wheeler et al.
 655 [2005].

~~Several studies already investigated~~ Due to known drawbacks of the OBL model, several

improvements and extensions have been made in the past: Rodriguez-Iturbe et al. [1988] introduced the random parameter model, allowing for different type of cells, and additionally Onof and Wheater [1994b] used a jitter and a gamma-distributed intensity parameter to account for a more realistic irregular shape of the cells. Cowpertwait et al. [2007] improved the representation of sub-hourly time scales by adding a third layer, pulses, to the model. Non-stationarity has been addressed by Salim and Pawitan [2003] and Kaczmarska et al. [2015]. Applications of these kind of precipitation models and some applications are known models include the implementing of copulas to investigate wet and dry extremes [Vandenbergh et al., 2011, Pham et al., 2013], regionalisation [Cowpertwait et al., 1996a,b, Kim et al., 2013] and accounting for interannual variability [Kim et al., 2014].

Parameter estimation for the BLRPM-OBL model is by far not trivial. The canonical approach is a method-of-moment-based estimation [Rodriguez-Iturbe et al., 1987] using the objective function

$$Z(\boldsymbol{\theta}; \mathbf{T}) = \sum_{i=1}^k w_i \left[1 - \frac{\tau_i(\boldsymbol{\theta})}{T_i} \right]^2. \quad (5)$$

This function relates moments $\tau_i(\boldsymbol{\theta})$ derived from the model with parameters $\boldsymbol{\theta}$ to empirical moments T_i from the time series. The set of k moments T_i is typically chosen from the first and second moments obtained for different aggregation times durations h . Here, we use the mean, the variances, the lag-1 auto-covariance function and the probability of zero rainfall for $h \in \{1h, 3h, 12h, 24h\}$, similar to Kim et al. [2013], and thus end up with $k = 13$ moments T_i . Their analytic counterparts $\tau_i(\boldsymbol{\theta})$ are derived from the model. The weights in the objective function were chosen to be $w_{i=1} = 100$ and $w_{i \neq 1} = 1$, similar to Cowpertwait et al. [1996a], emphasising the first moment T_1 (mean) 100 times more than the other moments.

It turns out that $Z(\boldsymbol{\theta}; \mathbf{T})$ has multiple local minima and optimization is not straightforward. To avoid those, the optimisation local minima, the optimization is repeated many times with different initial guesses for the parameters. These initial guesses are sampled from a range of feasible values in parameter space in a with Latin-Hypercube fashion [McKay et al., 1979]. As only positive model parameters are meaningful, optimisation is performed on log-transformed parameters. Similarly to Cowpertwait et al. [2007], we use a symmetric objective function

$$Z(\boldsymbol{\theta}; \mathbf{T}) = \sum_{i=1}^k w_i \left\{ \left[1 - \frac{\tau_i(\boldsymbol{\theta})}{T_i} \right]^2 + \left[1 - \frac{T_i}{\tau_i(\boldsymbol{\theta})} \right]^2 \right\}. \quad (6)$$

A few tests indicate that this objective function the symmetric version is robust and faster in convergence the sense that fewer iterations are needed to ensure convergence into the global minimum (not shown). Numerical optimisation optimization techniques based on gradient calculations, e.g. *Nelder-Mead* [Nelder and Mead, 1965] or *BFGS* [Broyden, 1970, Fletcher, 1970, Goldfarb, 1970, Shanno, 1970], are typically used. We For the current study, we use R's `optim()` function choosing *L-BFGS-B* as the underlying optimisation algorithm allowing to set additional constraints in parameter space optimization algorithm [R Core Team, 2016] and 100 different sets of initial guesses for the parameters sampled in a Latin-Hypercube way are used.

Following studies by Cowpertwait [1998] and Kaczmarska et al. [2014], we include the third moment in the parameter estimation using analytical expressions derived by Wheater et al. [2006], replacing the probability of zero rainfall in the objective function. Thus, still 13 moments are used to calibrate the OBL model. Due to comparability with other studies most of our analyses will not include the third moment though. A comparison between IDF curves of the model calibrated with the third moment and with the probability of zero rainfall will be carried out, to discuss the effect of including the third moment.

Models of this type suffer from parameter non-identifiability, meaning that qualitatively different sets of parameters lead to minima of the objective function with comparable values [Verhoest et al., 1997]. A more detailed view on global ~~optimisation~~ optimization techniques and comparisons between different objective functions is given in Vanhaute et al. [2012].

~~The~~ All computations are carried out with R – an environment for statistical computing and graphics [R Core Team, 2016]. Optimization is done with `optim()`. As additional constraints in parameter estimation we used the boundary extension of the BFGS algorithm with given parameter ranges. During this work the authors developed the R-package BLRPM ~~provides~~. The package includes functions for simulation and parameter estimation and ~~is available form the authors~~ can be obtained from the author on request.

3 Intensity-Duration-Frequency

Intensity-duration-frequency (IDF) curves show ~~return-levels~~ return levels (intensities) for given ~~return-periods~~ return periods (inverse of frequencies) as a function of rainfall duration. Their formulation goes back to Bernard [1932]. They are frequently used for supporting infrastructure risk assessment [e.g., Simonovic and Peck, 2009, Cheng and AghaKouchak, 2013]. IDF curves are an extension to classical extreme value statistics. The latter aims at better characterising the tails of a distribution by using parametric models derived from limit theorems [e.g., Embrechts et al., 1997]. There are two main approaches: modelling block-maxima (e.g., maxima out of monthly or annual blocks) with the generalised extreme value distribution (GEV) or modelling threshold excesses with the generalised Pareto distribution (GPD) [e.g., Coles, 2001, Embrechts et al., 1997]. We choose the block-maxima approach with the general extreme value distribution

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} \quad (7)$$

as parametric model for the block-maxima z . The GEV is characterised by the location parameter μ , the scale parameter σ and the shape parameter ξ . These can be estimated from block-maxima using a maximum-likelihood estimator [e.g., Coles, 2001]. Here, we use maxima from monthly blocks. To avoid mixing maxima from different seasons, a set of GEV parameters is estimated for all maxima from January, another set for all maxima from February and so on. For a given month, GEV parameters are estimated for various durations (or aggregation times), e.g. $d \in \{1h, 6h, 12h, 24h, 48h, \dots\}$. An IDF curve for a given ~~return-period~~ return period $T = 1/(1 - p)^{1/\xi}$, with p denoting the non-exceedance probability, can then be constructed from p -quantiles $Q_{p,d}$ from GEVs for different durations d by means of fitting a parametric model [e.g., Koutsoyiannis et al., 1998]. As the estimated IDF-curve $\text{IDF}_{T_1}(d)$ for ~~return-period~~ return period T_1 is independent of the estimate of another curve $\text{IDF}_{T_2}(d)$ with ~~return-period~~ return period $T_2 > T_1$, there is no constraint ensuring ~~$\text{IDF}_2(d) > \text{IDF}_1(d)$~~ $\text{IDF}_{T_2}(d) > \text{IDF}_{T_1}(d)$ for arbitrary durations d . Consequently, this approach easily leads to inconsistent (i.e. crossing) IDF-curves. For example for a given duration d , the 50-year ~~return-level~~ return level can exceed the 100-year ~~return-level~~ return level.

To overcome these problems and increase robustness in constructing IDF curves, Koutsoyiannis et al. [1998] suggested a duration-dependent scale parameter σ_d

$$\sigma_d = \frac{\sigma}{(d + \theta)^\eta}, \quad (8)$$

¹ ~~p denotes the non-exceedance probability.~~

with θ , η and σ being independent of the duration d . The parameter η quantifies the slope of the IDF curve in the main region and θ controls the deviation of the power-law ~~behaviour-~~
~~behavior~~ for short durations. Furthermore, location is ~~re-parametrised-reparametrised~~ by
740 $\tilde{\mu} = \mu/\sigma_d$ which is now independent of ~~the~~ durations d , ~~such is as well as~~ of the shape parameter ξ . This lead to ~~a-parametric-the following~~ formulation of a duration-dependent GEV distribution

$$F(x; \mu, \sigma_d, \xi) = \exp \left\{ - \left[1 + \xi \left(\frac{x}{\sigma_d} - \mu \right) \right]^{\frac{-1}{\xi}} \right\}.$$

This formulation:

$$F(x; \tilde{\mu}, \sigma_d, \xi) = \exp \left\{ - \left[1 + \xi \left(\frac{x}{\sigma_d} - \tilde{\mu} \right) \right]^{\frac{-1}{\xi}} \right\} \quad (9)$$

~~which~~ allows consistent modelling of rainfall maxima across different durations d using a single distribution at the cost of only two additional parameters. These parameters can be analogously
745 estimated by maximum-likelihood [Solyk et al., 2014]. To avoid local minima when ~~optimising~~
~~optimizing~~ the likelihood, we repeat the ~~optimisation-optimization~~ with different sets of initial guesses for the parameters, sampled ~~again~~ according to a Latin-Hypercube scheme ~~analogously~~
~~to the BLRPM parameter estimation~~. This method of constructing IDF curves is consistent in the sense that curves for different ~~return periods cannot~~ ~~return periods can not~~ cross. We refer
750 to this approach as the ~~duration-dependent~~ ~~duration dependent~~ *GEV* approach (dd-GEV).

However, the data points for different durations are dependent (as they are derived from the same underlying high-resolution data set by aggregation) and thus the i.i.d. assumptions required for maximum-likelihood estimation is not fulfilled. Consequently, confidence intervals are not readily available from asymptotic theory; they can be estimated by bootstrapping.

755 4 Data

A precipitation time series from the station *Botanical Garden* in Berlin-Dahlem, Berlin, Germany is used as a case study. A tipping-bucket ~~registers-records~~ precipitation amounts at 1-min resolution. For the analysis at hand, a 13 year time series with 1-min resolution from the years 2001-2013 is available. The series is aggregated to durations $d \in \{1h, 2h, 3h, 6h, 12h, 24h, 48h, 72h\}$
760 ~~yielding 8~~ $d \in \{1h, 2h, 3h, 6h, 12h, 24h, 48h, 72h, 96h\}$ ~~yielding 9~~ time series with different temporal resolution. IDF parameters are estimated using annual maxima for each month of the year individually using all ~~8-9~~ duration series. —

5 Results

5.1 Estimation of OBL model parameters

765 Minimising the symmetric objective function (Eq. (6)) yields ~~BLRPM-OBL model~~ parameter estimates individually for every month of the year, shown in Fig. 5 and explicitly given in Tab. 5 in Appendix A. The resulting ~~BLRPM-OBL model~~ parameters are reasonable compared to observed precipitation characteristics: ~~Large mean intensities $\hat{\mu}_x$ and short mean cell life-times $1/\hat{\eta}$ in summer correspond to precipitation being dominated by convective events.~~
770 ~~Similar, the mean cluster life-time $1/\hat{\gamma}$ decreases in summer, whereas the mean cell generation rate $\hat{\beta}$ increases~~ ~~During summer months, we observe very intensive cells ($\hat{\mu}_x$ between 4mm/h and 8mm/h). However, in June and August, storm duration is relatively short ($\hat{\gamma}$ between 0.25/h and~~

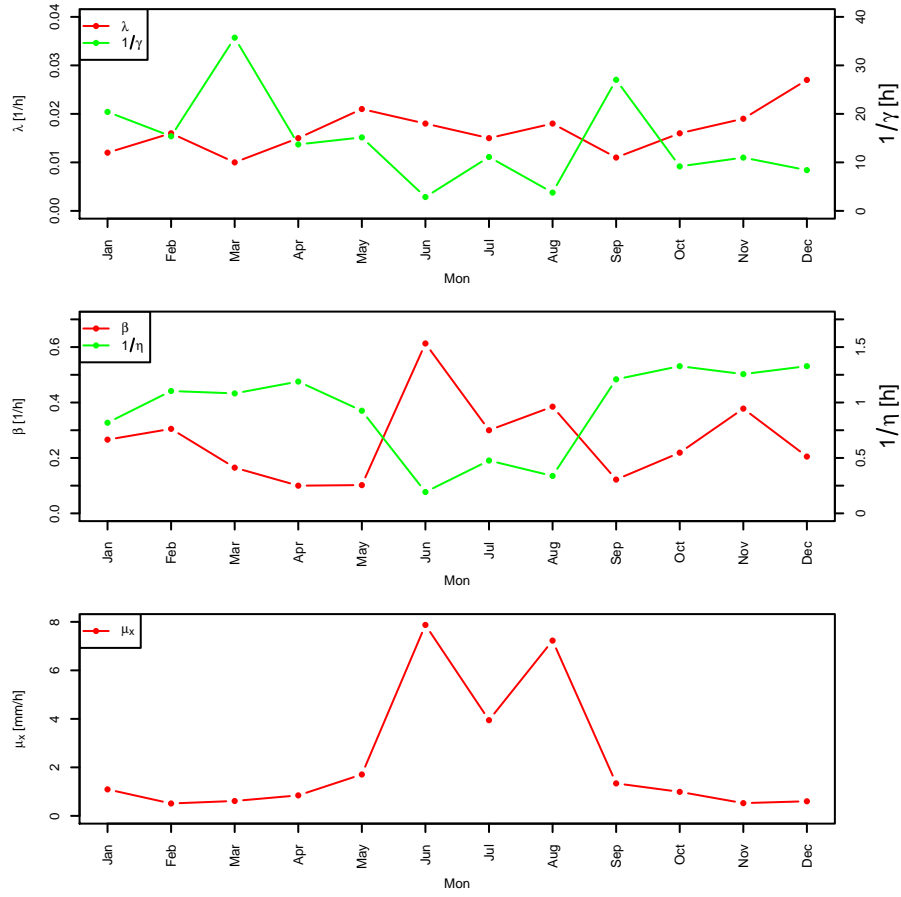


Figure 5: ~~BLRPM-OBL~~ model parameter estimates for all month of the year obtained from the Berlin-Dahlem precipitation time series. Top: cell-cluster generation rate λ and cluster lifetime $1/\gamma$; middle: cell generation rate β and cell lifetime $1/\eta$; bottom: cell mean intensities μ_x .

0.35/h) which can be interpreted as short but heavy thunderstorms which are typically observed in this region in summer [Fischer et al., 2017]. Vice versa, in winter small intensities and long storm durations correspond to stratiform precipitation patterns, typically dominating the winter precipitation in Germany. The storm generation rate λ show only a minor seasonal variation.

With the BLRPM-OBL model parameter estimates (Tab. 5, Sect. A) 1000 realisations with the same length as the observations (13 years) are generated. From both, the original precipitation series and the set of simulated time series, we derived a set of statistics for model validation. The first moment T_1 – the mean – is very well represented (not shown) as it enters the objective function with weight $w_1 = 100$ compared to weights of 1 for the other statistics. Figure 6 shows the variance for 6-hourly aggregation and the probability of zero rainfall; for all months the 6h-variances of simulated and observed series are in good agreement. This is particularly noteworthy as the 6-hourly aggregation was not used for parameter estimation. Similar to previous studies

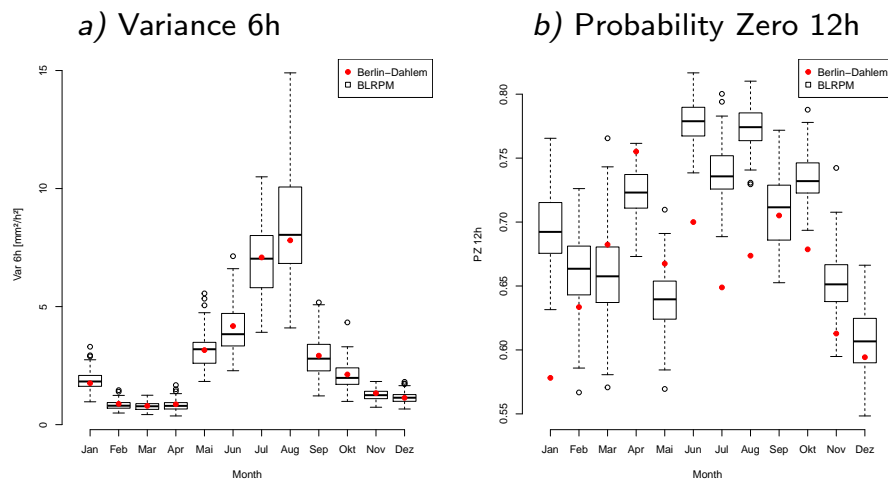


Figure 6: Comparison of statistics derived from the observational record (red dots) and 1000 simulated time-series (box plots): a) variance at 6-hourly aggregation level and b) probability of zero rainfall at 12-hourly aggregation.

[e.g., Onof and Wheater, 1994a], the model fails to reproduce the probability of zero rainfall, here exemplary for instance shown for the 12-hourly aggregation. The model mainly overestimates it and therefore has shortcomings in the representation of the time distribution of events [Rodriguez-Iturbe et al., 1987, Onof and Wheater, 1994a].

An important aspect for hydrological applications, is the model’s ability to reproduce extremes on various temporal scales. This behaviour is investigated in the next section with the construction of IDF curves.

5.2 Intensity-Duration-Frequency curves from OBL model simulations

Monthly block-maxima for every month in the year are drawn for various durations (1h, 3h, 6h, 12h, 24h, 48h, 72h, 96h) from the observational time series and 1000 BLRPM-OBL model simulations of same length. This is the basis for estimating GEV distributions for individual durations, as well as for constructing dd-GEV IDF curves.

IDF curves for Berlin-Dahlem obtained from observation are shown as dotted lines in Fig. 7 for January, April, July and October for the 0.5-quantile (2-year return-period return period,

red), 0.9-quantile (10-year ~~return-period~~return period, green) and the 0.99-quantile (100-year ~~return-period~~return period, blue). Analogously, IDF curves are derived from 1000 simulations

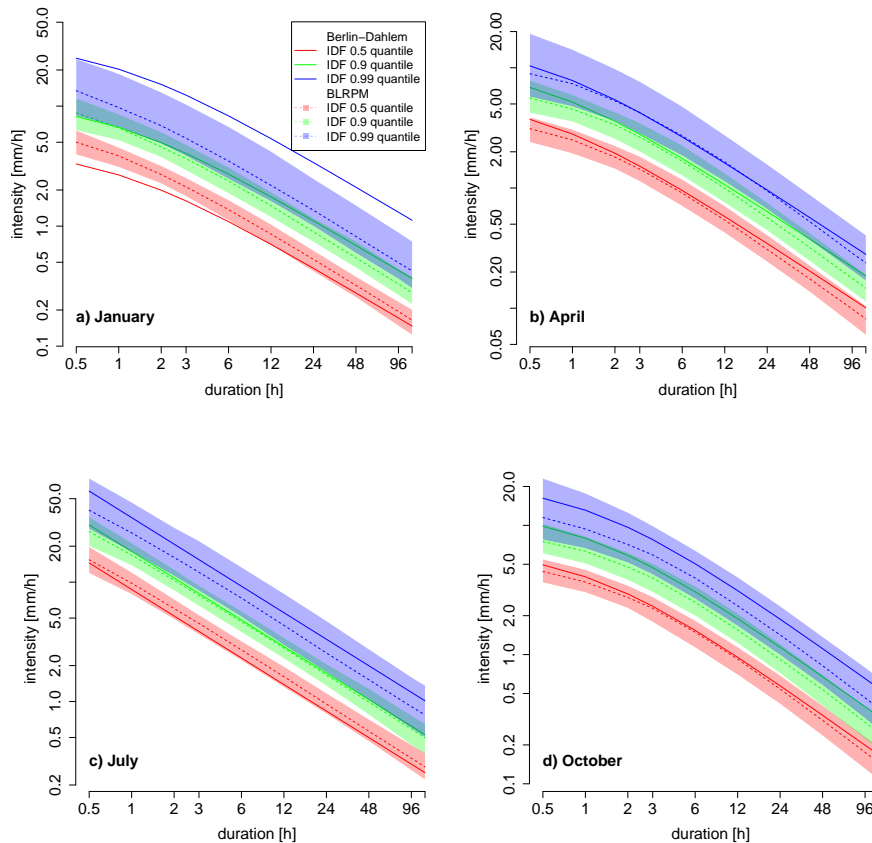


Figure 7: IDF curves obtained via dd-GEV for a) January, b) April, c) July and d) October: 0.5 (red), 0.9 (green) and 0.99-quantiles (blue) corresponding to 2-yr, 10-yr, and 100-yr ~~return-periods~~return periods, respectively. Solid lines are derived directly from the Berlin-Dahlem time series. Coloured shading mark the central 90% range of variability of IDF curves obtained in the same manner with same colour code but from 1000 ~~BLRPM-OBL model~~ simulations (Sect 5.1); the dotted lines mark the median of these curves.

800 of the ~~BLRPM-OBL model~~ precipitation series, cf. Sect. 5.1. The coloured shading in Fig. 7 give the range of variability (5% to 95%) for these 1000 curves with the median highlighted as dotted line. Except for January, the curves obtained directly from the observational series can be found within the range of variability of curves derived from the ~~BLRPMOBL model~~. The main
 805 IDF features from observations are well reproduced by the ~~BLRPMOBL model~~: the power-law-like behaviour (straight line in the double-logarithmic representation) in July extending almost across the full range of durations shown, as well as the flattening of the IDF curves for short durations for April and September. ~~There is, however, The relative differences in IDF curves given in Fig. 13 (Appendix B) suggest~~ a tendency for the ~~BLRPM-OBL model~~ to underestimate
 810 extremes, particularly for large ~~return-levels and long durations, as can be seen from plots of~~

relative differences in return levels and short durations, similar to results found by, e.g. Verhoest et al. [1997] and Cameron et al. [2000].

Figure 8 shows the relative difference

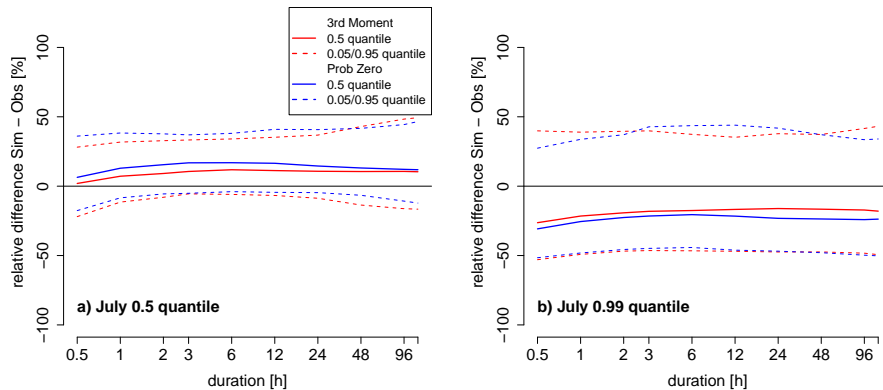


Figure 8: Relative differences between observed and simulated return levels obtained with including the third moment (red) and with using the probability of zero rainfall (blue) in parameter estimation for a) July 0.5 quantile and b) July 0.99 quantile. Dotted lines show the 0.05 and 0.95 quantile range of 1000 simulations.

$$\Delta = \frac{\text{dd-GEV}_{\text{OBL}} - \text{dd-GEV}_{\text{obs}}}{\text{dd-GEV}_{\text{obs}}} \cdot 100\% \quad (10)$$

815 between IDF curves (Fig. 13) in Appendix B dd-GEV) derived from the OBL model $\text{dd-GEV}_{\text{OBL}}$ including the third moment in parameter estimation (red lines) or alternatively using the probability of zero rainfall to calibrate the model (blue lines), and directly from the observational time series $\text{dd-GEV}_{\text{obs}}$ for July and two quantiles: a) 0.5 and b) 0.99. Including the third moment in parameter estimation slightly improves the model extremes for July for all durations and both short and long return periods. Nevertheless, those promising results could not be found for all months (not shown) and thus we cannot conclude that including the third moment in parameter estimation improves extremes in the OBL model in contrast to findings for the Neyman-Scott variant [Cowpertwait, 1998].

825 We interpret the different behaviour for short durations (flattening vs continuation of the straight line) for summer (July) and the remaining seasons as a result of different mechanisms governing extreme precipitation events: while convective events dominate in summer, frontal and thus more large scale events dominate in the other seasons.

830 For February As an example, we show segments of time series including the maximum observed/simulated rainfall in July for durations 1h, 6h and 24h as observed (RR_{obs}) and simulated (RR_{OBL}) in Fig. 9. Parts of the observed and simulated rainfall time series corresponding to the extreme events for the three different durations are shown in the left and right column, respectively. Additionally the middle column shows the simulated storms and cells generating this extreme event in the simulated time series. Note, that we show only one simulation as an example; visual inspection of several other simulated series share the main features and are not reproduced here. For all durations, the extremes are a result of a single long-lasting cell with

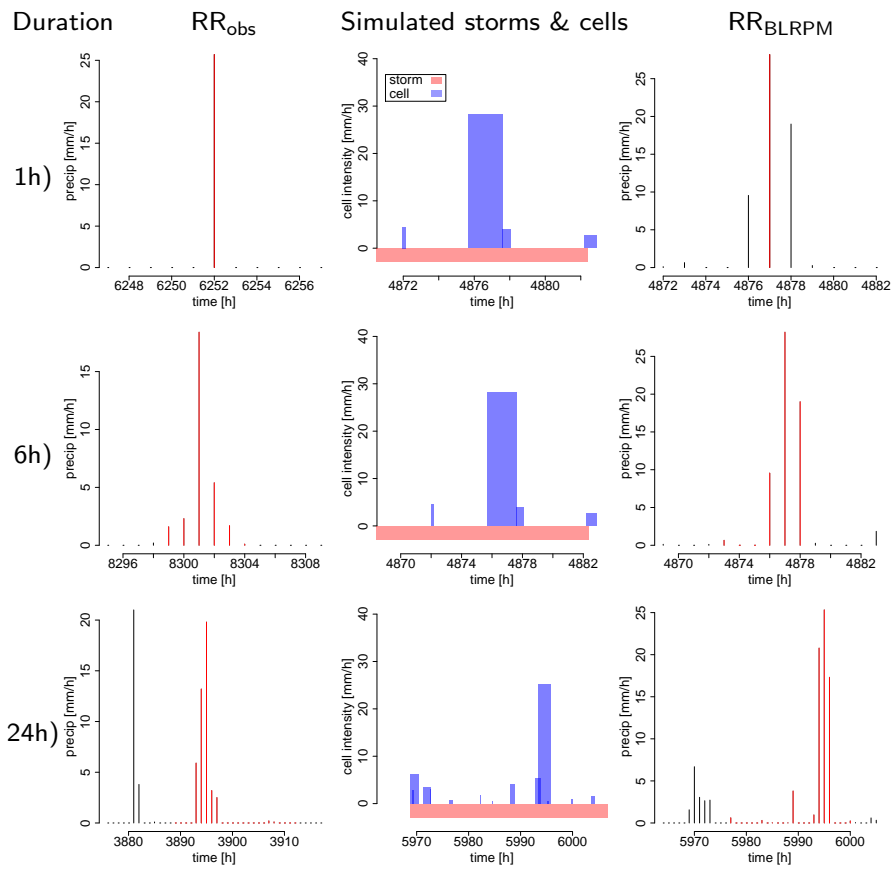


Figure 9: Visualization of July extremes as observed (RR_{obs} , left column) and simulated by the OBL model (RR_{OBL} , right column). Shown are short segments including the maximum observed/simulated rainfall (red vertical bars) at durations 1h (top row), 6h (middle row) and 24h (bottom row). Additionally, the middle column shows the simulated storms (red rectangles) and cells (blue rectangles) corresponding to the extreme event of the simulated time series.

835 ~~high intensity. In contrast to an analysis based on the random parameter BL model~~ [Verhoest
et al., 2010], ~~these cells are neither unrealistic long nor have an unrealistic high intensity.~~

For January, IDF curves from observations and ~~BLRPM-OBL model~~ simulations exhibit large
discrepancies: for all durations, the 0.99-quantile (100-yr ~~return-level~~return level) is above the
range of variability from the ~~BLRPM-OBL model~~ and the 0.5-quantile (2-yr return level) is below
840 for small durations. This implies, that the shape of the extreme value distribution characterised
by the scale σ and shape parameter ξ is differs between the two cases. This is likely due to
the winter-storm Kyrill hitting Germany and Berlin on January 18th and 19th in 2007 [Fink
et al., 2009]. We suppose that this ~~singular-rare~~ event is not sufficiently influential ~~for BLRPM~~
~~to impact OBL model~~ parameter estimation but does affect the extreme value analysis. For the
845 latter only the one maximum value per month is considered. In fact, the shape parameter ξ
estimated from the observational ~~time-series~~time series shows a large value compared to the
other months; in contrast, this value is estimated to be around zero from ~~BLRPM-OBL model~~
simulations. The following section investigates this hypothesis by excluding the precipitation
events due to Kyrill.

850 We furthermore find that the ~~BLRPM-OBL model~~ is generally able to reproduce the observed
seasonality in IDF parameters, see Fig. 10. For all parameters, the direct estimation (blue) is
mostly within the range of variability of the ~~BLRPM-OBL model~~ simulations. For $\hat{\sigma}$, $\hat{\theta}$ and $\hat{\eta}$,
the direct estimation (blue line) features a similar seasonal pattern as the median of the ~~BLRPM~~
~~OBL model~~ (red line). Whereas for $\hat{\xi}$, the direct estimation is a lot more erratic than the median
855 (red). As the GEV shape parameter is typically difficult to estimate [Coles, 2001], this erratic
behaviour is not unexpected and 11 out of 12 ~~month-months~~ stay within the expected inner 90%
range of variability.

5.3 Investigation of the impact of a rare extreme event

The convective cold front passage of Kyrill accounted for a maximum intensity of 24.8mm rainfall
860 per hour, whereas the next highest value of the remaining Januaries would be 4.9mm rainfall per
hour in 2002 and thus being more than 5 times lower than for Kyrill. We construct another data
set without the extreme event due to Kyrill, i.e. without the year 2007. ~~The intention of this~~
~~experiment is not to motivate removal of an “unsuitable” value. We rather want to show that~~
~~the OBL model is in generally able to reproduce extremes; it is, however, not flexible enough to~~
865 ~~account for a single event with magnitude far larger than the rest of the time series. Based on the~~
~~model with parameters estimated from observations with and without the year 2007 (observed),~~
~~we obtain return periods for the event “Kyrill” for different durations and find this event to~~
~~be very rare, especially on short-time scales (1-3 hours), see Tab. 3. .~~ For this data set, we
estimate the ~~BLRPM-OBL model~~ parameters and simulate again 1000 time series with these
870 new parameters. The simulated ~~time-series~~time series were also reduced in length by one year,
containing 12 years of rainfall in total. From those precipitation time series, we constructed the
dd-GEV IDF curves, see Fig. 11 (right). Without the extreme events due to Kyrill, the ~~BLRPM~~
~~OBL model~~ performs in January as well as in the other month with respect to reproducing the
IDF relations. In particular, the spread between the 0.5-quantile (2-yr ~~return-level~~return level)
875 and the 0.99-quantile (100-yr) return level is reduced and the absolute values of extreme quantiles
as well, cf. Fig. 11, left and right panel. Note the different scales for the intensity-axes.

5.4 Comparing dd-GEV IDF curves to individual duration GEV

~~Long~~ ~~In the frame of a model-world study, long~~ time series simulated with the ~~BLRPM-can~~
~~also-OBL model can~~ be used to investigate adequacy of the dd-GEV model ~~conditional on the~~

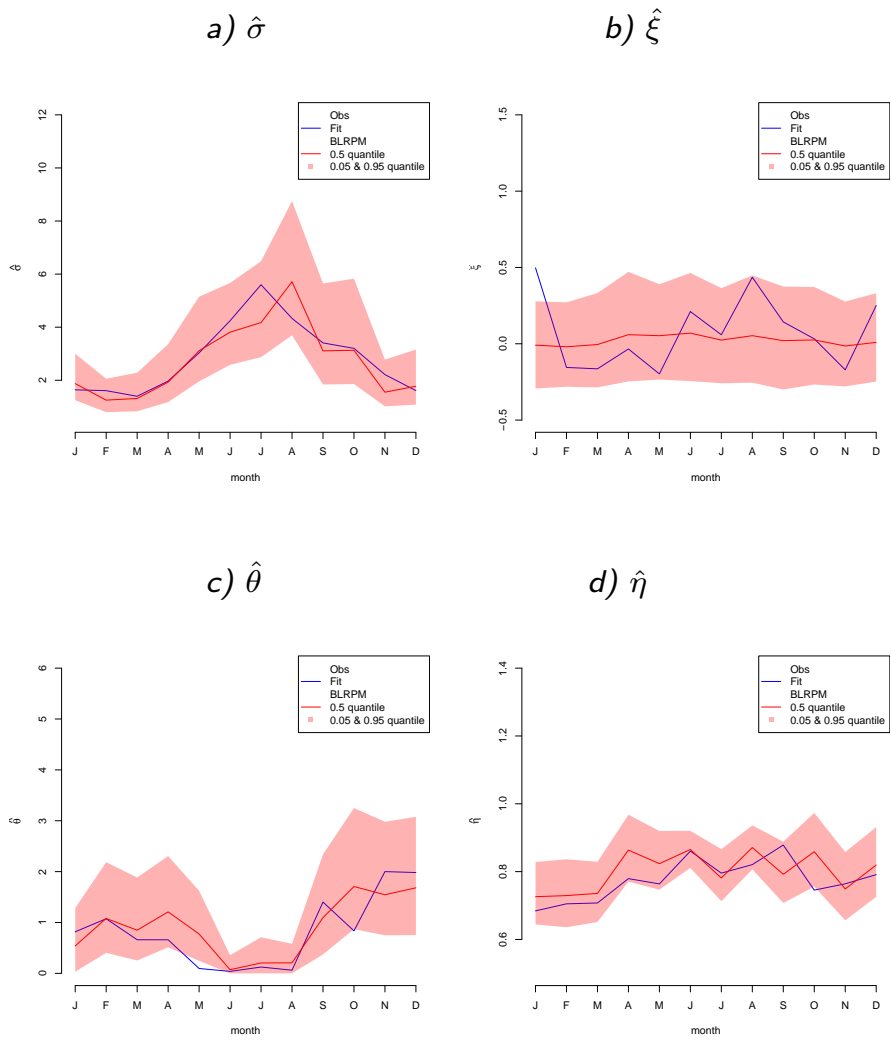


Figure 10: Seasonality of IDF model parameters estimated directly from the Berlin-Dahlem series (blue line), and estimated from 1000 **BLRPM-OBL model** simulations (red). The red shadings give the range of variability (5% to 95%) from the 1000 **Simulations simulations** with the median as solid red line.

Duration [h]	Probability of exceedance without Kyrill [%]	Return period without Kyrill [years]	Probability of exceedance including Kyrill [%]	Return period including Kyrill [years]
1	1.8×10^{-6}	560000	5.6×10^{-4}	1790
2	4.3×10^{-5}	23000	2.4×10^{-3}	420
3	2.2×10^{-4}	4400	5.4×10^{-3}	185
6	1.6×10^{-3}	630	1.6×10^{-2}	63
12	1.7×10^{-3}	590	2.0×10^{-2}	49
24	3.5×10^{-3}	280	3.5×10^{-2}	29
48	2.0×10^{-2}	50	9.5×10^{-2}	11

Table 3: Return period for the event *Kyrill* as estimated from the observational time series with this particular event left out and included for parameter estimation.

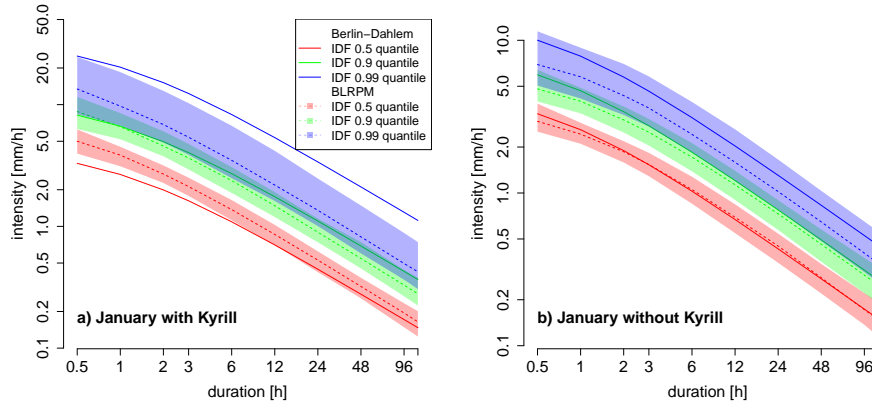


Figure 11: dd-GEV IDF curves for a) all Januarys (including 2007), b) Januarys excluding 2007 (different scaling on the intensity axis). Shown are the 0.5 (red), 0.9 (green) and 0.99 (blue) quantile from observations at Berlin-Dahlem (solid lines). The shaded areas are the respective 0.05 and 0.95 quantiles for the associated IDF curves obtained from 1000 BLRPM-OBL model simulations.

880 simulated series. To this end, we compare ~~resulting IDF curve~~ the resulting IDF curves to a GEV distribution obtained for various individual durations. The basis is a set of 1000-year ~~simulation with the BLRPM~~ simulations with the OBL model with parameters optimised for Berlin-Dahlem. For a series of this length, we expect to obtain quite accurate (low variance) results for both, the dd-GEV IDF curve and the GEV distributions for individual durations.

885 However, sampling uncertainty is quantified by repeatedly estimating the desired quantities from 50 repetitions. The resulting dd-GEV IDF curves are compared to the individual durations GEV in Fig. 12 for January (left) and July (right). For most durations in January and July, the dd-

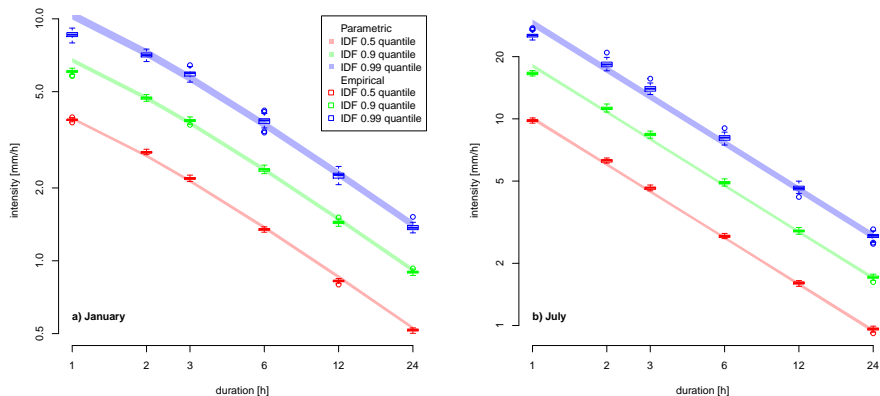


Figure 12: dd-GEV IDF curves for a) January and b) July and associated quantiles of a GEV distribution estimated for individual durations. Shown are the 0.5- (red), 0.9- (green) and 0.99-quantile (blue); shaded areas/box plots represent the variability over the 50 repetitions (5% to 95%)

IDF curves are close to the quantiles of the individual duration GEV distributions. Notable differences appear for small durations and large quantiles (~~return levels~~ return levels for long return periods); particularly in January the dd-GEV IDF model overestimates the 10-year and 100-year ~~return levels~~ return levels (duration of 1h), in July, this effect seems to be present as well but smaller in size. This is accompanied by a slight underestimation of the dd-GEV IDF for durations of 2h to 6h in July and 3h to 6h in January, most visible for the 0.99-quantile (100-yr ~~return level~~ return level). Both effects together suggest that the flattening of the dd-GEV IDF

895 for small durations is not sufficiently well represented. This could be due to deficiencies in the model for the duration dependent scale parameter (Eq. (8)) but might also be a consequence of an inadequate sampling of durations ($d \in \{1h, 6h, 12h, 24h, 48h, \dots\}$) to be used to estimate the dd-GEV IDF parameters. This is a point for further investigation.

6 Discussion and conclusions

900 ~~A~~ The original version of the Bartlett-Lewis rectangular pulse ~~model (BLRPM)~~ (OBL) model is optimized for the Berlin-Dahlem precipitation time series. Subsequently IDF curves are obtained directly from the original series and from simulation with the ~~BLRPM~~ OBL model. Basis for the IDF curves is a parametric model for the duration-dependence of the GEV scale parameter which allows a consistent estimation of one single duration-dependent GEV using all duration series

905 simultaneously (dd-GEV IDF curve). Model parameters for the ~~BLRPM~~ OBL model and the

IDF curves are estimated for all months of the ~~years~~ year and seasonality in the parameters is visible. Typical small-scale convective events in summer and large-scale stratiform precipitation patterns in winter are associated with changes in model parameters.

We show that the ~~BLRPM-OBL model~~ is able to reproduce empirical statistics used for
910 parameter estimation; Mean, variance and autocovariance of simulated ~~time-series~~ time series
are in good agreement with observational values, whereas the probability of zero rainfall is more
difficult to capture [cf. Rodriguez-Iturbe et al., 1987, Onof and Wheater, 1994a].

With respect to the first research question posed in the introduction, we investigate to what
~~extend the BLRPM~~ extent the OBL model is able to reproduce the intensity-duration relationship
915 found in observations. We show that they do reproduce the main features of the IDF curves esti-
mated directly from the original time series. However, a tendency to underestimate ~~return-levels~~
return levels associated with long ~~return-periods is observed~~ return periods is observed similar
to Onof and Wheater [1993]. Including the third moment in parameter estimation did not show
significant improvements in the OBL model's representation of extremes in contrast to findings
920 for the Neyman-Scott variant [Cowpertwait, 1998].

Furthermore, IDF curves for ~~February~~ January show a strong discrepancy between the ~~BLRPM~~
OBL model simulations and the original series. We hypothesize and investigate that this is due
to the Berlin-Dahlem precipitation series containing an extreme rainfall event associated with
the winter-storm *Kyrrill* passing over Berlin during January 18th and 19th, 2007. This event is
925 ~~singular very rare~~ in the sense that ~~the maximum hourly precipitation rate during these two days~~
~~exceeds the second largest rate found in the time series by a factor of 5.~~ on the events short time
scales it is only probable to occur once within a period larger than 1000 years. This addresses
the second research question: How are IDF curves affected by ~~singular very rare~~ extreme events
which ~~might not be reproducible with the BLRPM~~ are unlikely to be reproduced with the OBL
930 model for a reasonably long simulation? When the year 2007 is excluded from the analysis, the
aforementioned discrepancy in January disappears. We conclude that ~~such a singular extreme~~
~~event~~ an extreme event which is rare (return period of 23000 yrs) with respect to the time scales
of simulation (1000 × 13 yrs) has the potential to influence the dd-GEV IDF curve as 1 out of
13 values per duration – i.e. one maximum per year out of a 13 years time series – does change
935 the GEV distribution. However, its potential to influence mean and variance statistics used to
estimate ~~BLRPM-OBL model~~ parameters is minor.

The third questions addresses the validity of the duration dependent parametric model for
the GEV scale parameter which allows a consistent estimation of IDF curves. For a set of
long simulations (1000 years) with the ~~BLRPM~~ OBL model, the comparison of IDF curves with
940 the duration-dependent GEV approach with quantiles from a GEV estimated from individual
durations suggest a systematic discrepancy associated with the flattening of the IDF curve for
short durations. Quantiles from individual durations are smaller for small durations as the dd-
GEV IDF curves which challenges the latter modeling approach. However, instead of altering
the duration dependent formulation of the scale parameter σ_d (Eq. (8)), a different sampling
945 strategy for durations d used in the estimation of the dd-GEV parameters might alleviate the
problem. This is a topic for further investigation.

We do not find the ~~BLRPM-OBL model~~ producing unrealistically high precipitation amounts,
as discussed for the random- η model [Verhoest et al., 2010]. Nevertheless, improvements in
reproducing the observed extreme value statistics (especially large return levels) could be made
950 by adding the third moment in parameter estimation, as previous studies showed [Kaczmarska,
2011].

In summary, the ~~BLRPM-OBL model~~ is able to reproduce the general behaviour of extremes
across multiple time scales (durations) as represented by IDF curves. ~~Singular-Very rare~~ extreme
events do not have the potential to change the ~~BLRPM-OBL model~~ parameters but they do effect

955 IDF statistics and consequently modifies the previous conclusion for these cases. A duration dependent GEV is a promising approach to obtain consistent IDF curves; its behaviour at small durations needs further investigation.

A OBL model parameters

Estimation of OBL model parameters follow the boundary constraints: For those parameter

Parameter	Lower boundary	Upper boundary
λ	0.004 [h ⁻¹]	1 [h ⁻¹]
γ	0.01 [h ⁻¹]	10 [h ⁻¹]
β	0.01 [h ⁻¹]	100 [h ⁻¹]
η	0.01 [h ⁻¹]	100 [h ⁻¹]
μ_x	1×10^{-9} [mm/h]	100 [mm/h]

Table 4: Boundary constrained used in OBL model parameter estimation.

960 ranges, numerical optimisation mostly converged into a global minimum. For the model variant using the third moment in the OF, no constraints are used.

Using a Latin-Hypercube approach, we generated 100 different sets of initial guesses for the parameters used in the numerical optimization of the symmetrized objective function, Eq. (6). The estimation of **BLRPM-OBL model** parameters proved to be robust and the majority of optimisations led to the same minimum of the objective function which is then assumed to be the global minimum. Parameter estimates are given in Tab. 5.

	$\hat{\lambda}$ [h ⁻¹]	$\hat{\gamma}$ [h ⁻¹]	$\hat{\beta}$ [h ⁻¹]	$\hat{\eta}$ [h ⁻¹]	$\hat{\mu}_x$ [mm h ⁻¹]	Z _{min}
Jan	0.012	0.049	0.266	1.223	1.093	0.389
Feb	0.016	0.065	0.305	0.906	0.511	0.036
Mar	0.010	0.028	0.165	0.924	0.614	0.077
Apr	0.015	0.073	0.100	0.841	0.845	0.125
May	0.021	0.066	0.102	1.080	1.707	0.419
Jun	0.018	0.350	0.613	5.191	7.873	0.109
Jul	0.015	0.090	0.300	2.098	3.946	0.105
Aug	0.018	0.265	0.385	2.960	7.228	0.126
Sep	0.011	0.037	0.122	0.827	1.340	0.055
Oct	0.016	0.109	0.219	0.753	0.990	0.099
Nov	0.019	0.091	0.378	0.796	0.525	0.064
Dec	0.027	0.119	0.205	0.753	0.602	0.130

Table 5: Optimum of estimated **BLRPM-OBL model** parameters for individual month of the year for the Berlin-Dahlem precipitation series and corresponding value of the objective function Z

B Difference in IDF curves

Figure 13 shows the relative difference

$$\Delta = \frac{\text{dd-GEV}_{\text{OBL}} - \text{dd-GEV}_{\text{obs}}}{\text{dd-GEV}_{\text{obs}}} \cdot 100 \quad (11)$$

between IDF curves (dd-GEV) derived from the **BLRPM-OBL model** $\text{dd-GEV}_{\text{OBL}}$ and directly
970 from the observational time series $\text{dd-GEV}_{\text{obs}}$. From the four panels in Fig. 13, the discrepancies
of the **BLRPM-OBL model** can be highlighted. Apart from the large discrepancies in January
discussed in Sect. 5.3, the range of variability (colored shadows in Fig. 13 include the zero
difference line. However, the median over the 1000 **BLRPM-OBL model** simulations show general
975 tendency for the **BLRPM-OBL model** to underestimate extremes for large return periods (0.99-
quantile) by 25-50%. The best agreement is achieved for April.

Acknowledgements The project has been funded by Deutsche Forschungsgemeinschaft (DFG)
through grant CRC 1114.

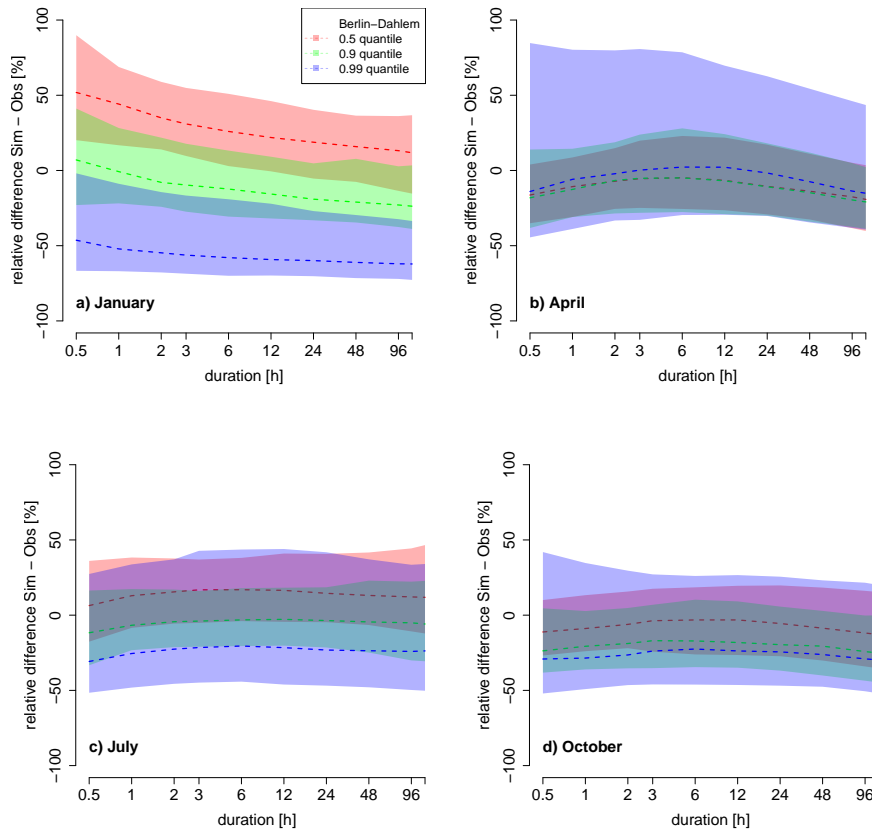


Figure 13: Relative differences (Eq. (11)) between simulated and observed IDF curves for a) January, b) April, c) July and d) October in percent relative to the observational values. Shown are the 2-yr (0.5-quantile, red), the 10-yr (0.9-quantile, green) and 100-yr ~~return-level~~ return level (0.99-quantile, blue) differences. The dashed lines denotes the median over all 1000 simulations and the surrounding coloured shading mark the range of variability (5% to 95%). Due to the usage of transparent colours, the three different colours can overlap and mix, grey shadows thus correspond to the overlapping of all three colours.

References

- 980 P.M. Austin and R.A. Houze. Analysis of the structure of precipitation patterns in new england. *J. Appl. Meteor.*, 11:926–935, 1972.
- M.M. Bernard. Formulas for rainfall intensities of long duration. *Transactions of the American Society of Civil Engineers*, 96(1):592–606, 1932.
- C.G. Broyden. The convergence of a class of double-rank minimization algorithms. *J. Instit. Math. Applications*, 6:76–90, 1970.
- 985 D. Cameron, K. Beven, and P. Naden. Flood frequency estimation by continuous simulation under climate change (with uncertainty). *Hydrology and Earth System Sciences Discussions*, 4(3):393–405, 2000.
- L. Cheng and A. AghaKouchak. Nonstationary precipitation intensity-duration-frequency curves for infrastructure design in a changing climate. *Scientific reports*, 4:7093–7093, 2013.
- 990 S. G. Coles. *An Introduction to Statistical Modelling of Extreme Values*. Springer, London, 2001.
- P.S.P. Cowpertwait. A poisson-cluster model of rainfall: high-order moments and extreme values. *Proceedings: Mathematical, Physical and Engineering Sciences*, Vol. 454, No. 1971:885–898, 1998.
- 995 P.S.P. Cowpertwait, P.E. O’Connell, A.V. Metcalfe, and J.A. Mawdsley. Stochastic point process modelling of rainfall. i. single-site fitting and validation. *J. Hydrol.*, 175:17–46, 1996a.
- P.S.P. Cowpertwait, P.E. O’Connell, A.V. Metcalfe, and J.A. Mawdsley. Stochastic point process modelling of rainfall. ii. regionalisation and disaggregation. *J. Hydrol.*, 175:47–65, 1996b.
- P.S.P. Cowpertwait, V. Isham, and C. Onof. Point process models of rainfall: Developments for fine-scale structure. Research Report 277, Departement of Statistical Science, University
1000 College London, June 2007.
- P. Embrechts, C. Klüppelberger, and T. Mikosch. *Modelling Extremal Events for Insurance and Fincance*. Springer, Berlin, 1997.
- A. H. Fink, T. Brücher, V. Ermert, A. Krüger, and J. G. Pinto. The european storm kyrill in
1005 january 2007: synoptic evolution, meteorological impacts and some considerations with respect to climate change. *Nat. Haz. Earth Syst. Sci.*, 9(2):405–423, 2009.
- M. Fischer, H. W. Rust, and U. Ulbrich. Seasonality in extreme precipitation – using extreme value statistics to describe the annual cycle in german daily precipitation. *Meteorol. Z.*, page accepted, 2017.
- 1010 R. Fletcher. A new approach to variable metric algorithms. *Computer J.*, Vol. 13, No. 3:317–322, 1970.
- D. Goldfarb. A family of variable metric updates derived by variational means. *Mathematics of Computation*, Vol. 24, No. 109:23–26, 1970.
- J.M. Kaczmarska. Further development of bartlett-lewis model for fine-resolution rainfall. Research Report 312, Departement of Statistical Science, University College London, April 2011.

- 1015 J.M. Kaczmarska, V.S. Isham, and C. Onof. Point process models for fine-resolution rainfall. *Hydrological Sciences Journal*, 59(11):1972–1991, 2014.
- J.M. Kaczmarska, V.S. Isham, and P. Northrop. Local generalised method of moments: an application to point process-based rainfall models. *Environmetrics*, 26:312–325, 2015.
- M.N. Khaliq and C. Cunnane. Modelling point rainfall occurrences with the modified bartlett-lewis rectangular pulses model. *J. Hydrol.*, 180:109–138, 1996.
- 1020 D. Kim, F. Olivera, H. Cho, and S.A. Scolofsky. Regionalization of the modified bartlett-lewis rectangular pulse stochastic rainfall model. *Terr. Atmos. Ocean Sci.*, Vol.24, No. 3:421–436, 2013.
- D. Kim, J. Kim, and Y.-S. Cho. A poisson cluster stochastic rainfall generator that accounts for the interannual variability of rainfall statistics: validation at various geographic locations across the united states. *Journal of Applied Mathematics*, 2014, 2014.
- 1025 D. Koutsoyiannis, D. Kozonis, and A. Manetas. A mathematical framework for studying rainfall intensity-duration-frequency relationships. *J. Hydrol.*, 206(1):118–135, 1998.
- A. Langousis and D. Veneziano. Intensity-duration-frequency curves from scaling representations of rainfall. *Water Resources Res.*, 43(2):n/a–n/a, 2007. W02422.
- 1030 L. Le Cam. A stochastic description of precipitation. *Proc. Fourth Berkeley Symp. on Math. Statist. and Prob.*, 3:165–186, 1961.
- M. D. McKay, R. J. Beckman, and W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245, 1979. ISSN 00401706.
- 1035 B. Merz, F. Elmer, M. Kunz, B. Mühr, K. Schröter, and S. Uhlemann-Elmer. The extreme flood in june 2013 in germany. *La Houille Blanche*, (1):5–10, 2014.
- J.A. Nelder and R. Mead. A simplex method for function minimization. *Computer J.*, 7 (4): 308–313, 1965.
- 1040 J Neyman and EL Scott. A theory of the spatial distribution of galaxies. *Astrophys. J.*, 116:144, 1952.
- C. Onof. *Stochastic modelling of British rainfall data using Poisson processes*. PhD thesis, University of London, 1992.
- C. Onof and H.S. Wheater. Modelling of british rainfall using a random parameter bartlett-lewis rectangular pulse model. *J. Hydrol.*, 149:67–95, 1993.
- 1045 C. Onof and H.S. Wheater. Improved fitting of the bartlett-lewis rectangular pulse model for hourly rainfall. *Hydrological Sciences*, Vol. 39, No. 6:663–680, 1994a.
- C. Onof and H.S. Wheater. Improvements to the modelling of british rainfall using a modified random parameter bartlett-lewis rectangular pulse model. *J. Hydrol.*, 157:177–195, 1994b.
- 1050 C. Onof., R.E. Chandler, A. Kakou, P. Northrop, H.S. Wheater, and V. Isham. Rainfall modelling using poisson-cluster processes: a review of developments. *Stochastic Environmental Research and Risk Assessment*, 14:184–411, 2000.

- A. Pattison. Synthesis of hourly rainfall data. *Water Resour. Res.*, 1 (4):489–498, 1956.
- 1055 M. T. Pham, W. Vanhaute, S. Vandenberghe, B. De Baets, and N. Verhoest. An assessment of the ability of bartlett–lewis type of rainfall models to reproduce drought statistics. *Hydrology and Earth System Sciences*, 17(12):5167–5183, 2013.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL <https://www.R-project.org/>.
- 1060 I. Rodriguez-Iturbe, D.R. Cox, F.R.S., and V. Isham. Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond.*, A 410:269–288, 1987.
- I. Rodriguez-Iturbe, D.R. Cox, and V. Isham. A point process model for rainfall: further developments. *Proc. R. Soc. Lond.*, A417:283–298, 1988.
- A. Salim and Y. Pawitan. Extensions of the bartlett-lewis model for rainfall processes. *Statistical Modelling*, 3(2):79–98, 2003.
- 1065 D.F. Shanno. Conditioning of quasi-newton methods for function minimization. *Mathematics of Computation*, Vol. 24, No. 111:647–656, 1970.
- S.P. Simonovic and A. Peck. *Updated rainfall intensity duration frequency curves for the City of London under the changing climate*. Department of Civil and Environmental Engineering, The University of Western Ontario, 2009.
- 1070 J.C. Smithers, G.G.S. Pegram, and R.E. Schulze. Design rainfall estimation in south africa using bartlett–lewis rectangular pulse rainfall models. *J. Hydrol.*, 258(1):83–99, 2002.
- S. Soltyk, M. Leonard, A. Phatak, E. Lehmann, et al. Statistical modelling of rainfall intensity-frequency-duration curves using regional frequency analysis and bayesian hierarchical modelling. In *Hydrology and Water Resources Symposium 2014*, page 302. Engineers Australia, 1075 2014.
- S. Vandenberghe, N. Verhoest, C. Onof, and B. De Baets. A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on bartlett-lewis modeled rainfall. *Water Resources Research*, 47(7), 2011.
- 1080 W. Vanhaute, S. Vandenberghe, K. Scheerlinck, B. De Baets, and N. Verhoest. Calibration of the modified bartlett-lewis model using global optimization techniques and alternative objective functions. *Hydrology and Earth System Sciences*, 16(3):873–891, 2012.
- T. Velghe, P.A. Troch, F.P. De Troch, and J. Van de Velde. Evaluation of cluster-based rectangular pulses point process models for rainfall. *Water Resour. Res.*, Vol. 30, No. 10:2847–2857, 1994.
- 1085 N. Verhoest, P.A. Troch, and F.P. De Troch. On the applicability of bartlett-lewis rectangular pulse models in the modeling of desing storms at a point. *J. Hydrol.*, 202:108–120, 1997.
- N. Verhoest, S. Vandenberghe, P. Cabus, C. Onof, T. Meca-Figueras, and S. Jameleddine. Are stochastic point rainfall models able to preserve extreme flood statistics? *Hydrological processes*, 24(23):3439–3445, 2010.
- 1090 E.C. Waymire, V.K. Gupta, and I. Rodriguez-Iturbe. A spectral theory of raifnall intensity at the meso- β scale. *Water Resour. Res.*, 20(10):1453–1465, 1984.

H.S. Wheater, R.E. Chandler, C.J. Onof, V.S. Isham, E. Bellone, C. Yang, D. Lekkas, G. Lourmas, and M.-L. Segond. Spatial-temporal rainfall model for flood risk estimation. *Stoch. Environ. Res. Risk Assess*, 19:403–416, 2005.

1095 H.S. Wheater, V.S. Isham, R.E. Chandler., C.J. Onof, and E.J. Stewart. Improved methods for national spatial-temporal rainfall and evaporation modelling for bsm. R&D Technical Report F2105/TR, Joint Defra/EA Flood and Coastal Erosion Risk Management R&D Programme, March 2006.