Author's response

Review of HESS discussion paper reference no. hess-2017-437

Censored rainfall modelling for estimation of fine–scale extremes

David Cross, Christian Onof, Hugo Winter, and Pietro Bernardara

We are very grateful to two anonymous referees for their constructive comments and suggestions on our manuscript. We set out here our responses to those and list our changes to the manuscript.

Section 1. point-by-point response to the reviewers' comments.

Section 2. list of all relevant changes in the manuscript.

Section 3. marked-up manuscript version showing track changes.

1. Responses to reviewers' comments

1.1 Anonymous Referee #1

Received and published: 25 September 2017

General comments

The aim of this paper is to improve the ability of mechanistic rainfall models extremes to reproduce rainfall extremes. This is achieved by fitting these models to the amounts by which rainfall totals exceed a certain threshold level, or censor. Totals below this level are censored: their value is taken to be zero when the models are fitted. Applying a threshold to data is a standard approach in extreme value modelling. It's use here is a novel and interesting idea. If modelling rainfall extremes is the primary goal then using the censor to reduce the influence of small totals is sensible. Discussion paper Of course, a key issue is the choice of censor: a low censor may not achieve the desired objective but as the censor is raised the precision of estimation reduces and, in the current context, may exacerbate the parameter identifiability problems to which the fitting of these models are prone. This is analogous to the choice of threshold in an extreme value analysis and therefore it is unsurprising that in Figure 13 they consider informing this choice using a graphical approach that is common in extreme value Interactive modelling. It may be productive to explore other methods proposed in the extreme comment value threshold selection literature.

Overall I am positive about this paper. My main criticism is that better reproduction of rainfall extremes is achieved by tuning various things: censor; model parameterisation; fitting properties; perhaps even the model itself, in order to achieve this objective. There is further work to be done to provide methodology to make these choices.

Specific comments

page 11, line 288. This isn't quite correct. These weights are not optimal. However, in practice they are close to being optimal and are easier to estimate than the weights that are optimal. On that note: how are these weights estimated?

The estimation of weights follows the theory and procedure set out in the MOMFIT manual (Chandler et al. 2010). First, the covariance matrix of mean summary statistics Σ is estimated using Equation 1,

Formatted: Font: 11 pt, Not Italic

Deleted: Equation 1

$$\hat{\Sigma} = \frac{1}{n(n-1)} \sum_{i=1}^{n} (T_i - \bar{T})(T_i - \bar{T})'$$

Equation 1

T_i is the vector of summary statistics for the ith period. $\bar{T}$ is the vector of mean summary statistics and n is the total number of periods which in practice is the length of the dataset in years. The optimal choice of weights W would be closely approximated by the inverse of the covariance matrix given in Equation 2.

Formatted: Font: 11 pt, Not Italic

Deleted: Equation 2

$$W = \Sigma^{-1}$$

Equation 2

However, as reported by Chandler et al. (2010), this results in highly inaccurate standard errors because the unique elements of Σ are estimated separately by Equation 1. To overcome this, a reasonable approximation for the weights is to take the inverse of the observed variance where t_i is the vector of diagonal elements of Σ.

Deleted: Equation 1

Formatted: Font: 11 pt, Not Italic

$$\omega_i = 1/var(t_i)$$

We can update the manuscript to indicate that the estimation of weights is not optimal, but provides a robust estimate.

page 12, lines 305-308. This isn't quite correct. The sampling distribution of the GMM estimators is approximated by a MVN distribution, i.e. there is an approximation involved and the result is for the rule that is used to calculate the estimates, rather than the estimates themselves. Line 306: Hessian of what? Lines 307-308. This sentence isn't clear. Presumably the point is that the calculation of confidence intervals can fail in some cases. Perhaps it would be sufficient to reverse the ordering of the sentence to make the causation clearer.

Line 305: We can amend the text to highlight that the distribution of the estimator resulting from the minimisation routine is approximately multivariate normal.

Line 306: The covariance matrix is estimated with the Hessian of the least squared objective function S given on line 280. We can revise the text to make this clearer.

Lines 307-308: Indeed, the sentence was provided to highlight that the calculation of confidence intervals can fail. We will reverse the sentence as below.

On occasions that the model parameters are poorly identified, it may not be possible to calculate the Hessian of the objective function preventing the estimation of parameter uncertainty.

page 13, line 345. If you are interested in the 1000 year return level why not simulate 100 realisations of 1000 years duration?

Extreme value estimation up to the 1000-year return level is provided to indicate the potential magnitude of rarer events. Estimation up to the 1000-year return level is performed for the optimum parameter set, therefore there is no sampling from the MVN distribution of model parameters. To ensure these estimates are reasonably stable up to 1000 years, simulations have been extended to 10k years. Furthermore, simulation bands for 100 simulations of 1000 years would differ from the 100 year simulation bands shown and make the already busy figures more crowded.

For validation that the observed data at each site are sampled from the sampling distribution, it makes sense to derive simulation bands for a length of data that is of the same order of magnitude as that of the observations. Therefore, the duration of 100 years was chosen to cover the range of the data at both sites and to enable comparison between sites.

pages 14-16, Figures 4-6. There seems to be a slight upward curvature in the lines based on the 10000 year simulation. In the context of an extreme value analysis this is consistent with the shape parameter discussed on page 4 (line 111) being positive. This might be worth a brief comment. Is there any work that examines how the extreme value properties of this type of model relate to the model parameters and therefore provide a link to the theoretical basis that underpins extrapolation from extreme value models?

This is a good suggestion and we will comment on this as suggested in the discussion. That said, we are not aware of any such work, but we will briefly review the literature again to investigate.

page 17, lines 385-386. I disagree. I suppose that it depends what you mean by "poorly identified". However, it is to be expected that as the censor is increased un- certainty about model parameters increases. If we think that we need a larger censor, because otherwise there is systematic underestimation of extreme rainfall totals, then we need to accept higher levels of parameter uncertainty.

This is a good point, although we have observed the same with uncensored models. Overall, our choice of summary statistics gives rise to apparently very well identified model parameters at sub-hourly scales. The comments here relate specifically to estimation at the hourly scale.

Our statement that the parameters are poorly identified may be overly strong. We can change this to state that confidence in the estimation is reducing.

page 18, Section 6.2. Is this level of tinkering with the choice of censor justified? Having a different censor for different levels of data aggregation feels like cherry-picking. Also, in the previous section an argument was made against a censor of 0.6mm for Atherstone but now it is being used.

The model is fitted separately for each temporal resolution which explains why we have different censors. Because of the effect of aggregation, we cannot use a model censored at one resolution to estimate rainfall extremes at a coarser one. Therefore, censored model parameters are scale dependent which is explained in section 3. Furthermore, we believe that the use of different censors at different levels of aggregation is justified on the basis that the distribution of rainfall amounts differs between aggregation levels.
The research as presented is exploratory and intended to investigate the potential for censoring in estimating extremes. In this context, a range of censors have been investigated and shown to be effective at improving the extreme value estimation of the models, and a choice had to be made for validation. Considering that, we agree that it would have been more consistent to select 0.2 mm for validation of the Atherstone hourly censored model and we will make that change.

page 19, Figure 8 caption (and elsewhere). "optimal censors" seems like a bold claim given the difficultly of choosing the censor.

We agree and will refer instead to selected censors.

page 28, line 516. The independence criterion isn't a requirement in extreme value modelling. See Fawcett and Walshaw (2012) Estimating return levels from serially dependent extremes. Environmetrics 23(3), 272-283.

Thank you for the reference, we were unaware of this research and appreciate its relevance in the context of ours.

We note that the methods set out in Fawcett and Walshaw (2012) requires the estimation of the extremal index which appears to be subjective and potentially non-trivial. Therefore, we feel that our comparison with the standard peaks over threshold approach in which independent cluster peaks are identified is still valid. We will highlight this in the discussion with reference to this research.

Technical corrections

page 3, line 81. At this point, or perhaps even in the abstract, it is worth explaining briefly the nature of the censoring. At the moment we need to wait until page 7 for this.

We will make the following change to the text.

To test our hypothesis, a simple approach is proposed in which low observations for fine–scale data are censored from the models in calibration. For a given temporal resolution, a censor amount is set. Rainfall below the censor is set to zero and rainfall over the censor is reduced by the censor amount.

page 4, line 118. "behavioural parameterizations". Given that you use this term later it would be worth explaining (somewhere) what this means in the context of the current paper.

This was also highlighted by Anonymous Referee #2 therefore the response below is the same as that provided to referee #2.

We have used the term "behavioural parameters" by analogy with Beven and Binley (1992). We have used the term to refer to well identified models. We have found that for well identified parameters with narrow 95% confidence intervals, simulation bands on the extreme value estimates are correspondingly narrow. As the parameters become less well identified, their 95% confidence intervals increase giving rise to extreme value estimates which deviate significantly from the observations, which in turn results in significant deviation of the simulation band upper limit. This effect is shown in Fig.11 resulting from the very large parameter uncertainty shown in Fig.12.

We will remove the reference to "behavioural parameterizations" in the context of this research and change all references to well identified parameters.

page 6, line 160. "Lower variability" may be better than "less variance".

Agreed. We will make this change in the manuscript.

page 10, line 262. Presumably the reason for the missing data in 1974-75 was political rather than environmental. It might be worth noting that the fact that the data are missing is not expected to be informative about rainfall totals.

We do not know why the data are missing, but we can certainly highlight that this is not expected to affect the results. We will make the following change to the text.

Atherstone is a tipping bucket rain gauge (TBR) operated and maintained by the Environment Agency of England. The record duration is 48 years from 1967 to 2015, with one notable period of missing data from January 1974 to March 1975. The reason for the missing data is unknown, although it is not expected to affect model fitting and the estimation of extremes.

page 12, line 313. "ldots extreme values continued to be underestimated . . . " might be better.

We will change this sentence to the following.

While good model fits were obtained for some low censors, extreme value estimation continued to be underestimated.

pages 14-16, Figures 4-6. The plots would be clearer if the scale on the lower horizontal axis was return level in years. The AEP on the upper horizontal axis would then be unnecessary. The scale of the Gumbel reduced variate adds no information in itself. These plots are quite crowded and

We agree that the Gumbel reduced variate adds no additional information given that the AEP is provided on the secondary x axis. We also agree that removing this and moving the AEP to the primary x axis will simplify the plots. However, we propose to keep the content of the plots unchanged as they show the convergence in estimation for all AEPs up to 0.001 for increasing censors.

page 17, line 358. I'm not sure that I would use "confidence intervals" here. Perhaps "simulation bands"? ... and say explicitly what this means, i.e. how the lines in the plot are calculated.

This was also highlighted by Anonymous Referee #2 therefore the response below is the same as that provided to referee #2. We agree with the suggestion and will change all occurrences in the manuscript.

There are in fact two issues here: if we were doing very long simulations with practically no random noise (so that another simulation would yield practically the same result), then we would have identified approximate confidence intervals. But with the shorter simulation length, both parameter uncertainty and the randomness of the model are combined in the spread we observe in the simulated statistics, so that 'simulation bands' is indeed a better descriptor.

page 17, line 379. I'm not sure what this sentence means. Are we supposed to be looking at Figure 7 for evidence of this?

The reader should be looking at Fig.6 for evidence of this divergence in estimation. This is explained with the aid of Fig.7. We will make the following change to the text.

The mean of the MVN realisations for the BL1M model at Atherstone with the 0.6 and 0.8 mm censors (see Fig.6) diverges from the optimum because of the generation of unrealistic extremes. This divergence is also observable in the larger spread of 95% simulation intervals over 100 realisations.

pages 20-21, Sections 6.2.1, 6.2.2 and 6.2.3. I don't see the point of including these sections. Section 6.2.1 shows exactly what we expect: by excluding properties that are difficult to reproduce we are able to reproduce well the properties that are not excluded. The comparison in Section 6.2.2 is unsatisfactory because we cannot compare like with like, owing to the truncation of the data but not the model. Section 6.2.3 just shows that there are clear local minima in the objective function but we can't expect to search too far in the search for confidence limits.

It is not always the case that the fitted parameters well reproduce the summary statistics used in fitting. The purpose of these checks is to ensure they do given that the data are censored. That said, given that the ability of the models to reproduce the summary statistics used in fitting at both sites is equally good, we could reduce this section by only showing plots for one site. We could then state that comparable performance is achieved at both sites.

We take the point about checking skewness. Given that we've already highlighted that skewness is not expected to be well reproduced because of the truncation of the data we are happy to remove these plots. However, we feel there is still validity in checking the proportion of dry periods as this property is strongly affected by removing low observation.

We would be happy to remove the profile objective function plots given that there is other evidence of good parameter identifiability with high confidence on the parameter estimates.

page 27, Figure 14. I don't think that these figures add much to the statistics concerning
the proportions of totals lying below the censors, with the possible exception of the visualisation of the resolution of the Atherstone data.

These figures were included to demonstrate graphically how much data is removed by censoring. Given that the methodology for selecting a censor presented in this research is based on a graphical approach, these figures are useful to understand the rainfall quantiles which have given rise to well parameterised censored models. Until an alternative method is developed to optimise the censor, we feel these plots will aid other practitioners in estimating rainfall extremes with censoring. Therefore, we propose to keep these plots in the manuscript.

page 28, line 514. The rule to try to create independent peaks needs to be given earlier: before the concept appears in Table 3.

Noted. We will bring this forward so that it is highlighted before Table 3.

pages 28-30. Do we need both "Further discussion" and "Conclusions"?

We can look at combining these into one section possibly called *Further discussion and conclusions* or just *Conclusions*.

page 31, line 591. Is the first inequality sign the wrong way round?

Yes, it is. Thank you for highlighting this. We will correct this in the manuscript.

page 32, Figure A.2. Below and to the right of the plot is says that 1/L has an exponential distribution, which, according to the description of the models on page 8, isn't true.

Thank you for highlighting this inconsistency. We will correct this and update Fig.A.2. We also notice that the parameterizations for X and L are listed wrongly in line 225. We will revise this as follows.

Both X and L are assumed to be independent of each other and follow exponential distributions with parameters $1/\mu_x$ *and* η *respectively.*

References

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrol. Process., 6, 279-298, 1992.

Chandler, R., Lourmas, G. and Jesus, J.: MOMFIT Software for moment-based fitting of single-site stochastic rainfall model fitting, User guide, Department of Statistical Science, University College London, London, 2010.

Fawcett, L. and Walshaw, D.: Estimating return levels from serially dependent extremes, Environmetrics, 23, 272-283, 2012.

1.2 Anonymous Referee #2

Received and published: 9 October 2017

General comments

The research work presented in the manuscript develops a new methodology to estimate fine-scale rainfall extremes. Although there has been a substantial amount of work done, by many authors over the years, on stochastic point process models for rainfall, most of the models proposed tend to underestimate the rainfall extremes at fine-scales. Estimation or reproduction of extreme rainfall at hourly and sub-hourly scales is a well-known problem. In this context, this paper attempts to address this problem by using a censored approach to model rainfall extremes. This is in a way similar to the Excess Over Threshold (EOT) method commonly used in extreme value modelling, but here a stochastic mechanistic model is used along with this idea. Application of this novel idea of censured modelling approach is illustrated in the estimation of fine-scale rainfall extremes from two different regions to provide an improved representation of extremes.

The paper gives an excellent coverage of the history of work carried out in this area to convey the rationale for the need to study or explore alternative methods for fine-scale extremes.

The success of this new approach, of course depends heavily on the choice of the censor level and, hence, emphasis was placed on finding appropriate value of the censor for the application. If the estimation of extreme rainfall is the main objective of the study then using this censored approach is certainly a useful tool and worthwhile addition to the existing methods.

One drawback in the proposed approach might be the amount of fine-tuning required to get the best set of potential estimates for the extreme rainfall with respect to model, its parameterisation, censor, statistics used in fitting as well as aggregation levels. This level of tweaking or fine-tuning might prove to be a lot to generate sufficient interest amongst practitioners. The rationale behind the need to make these choices, however, has been explained in the manuscript though.

Specific comments.

Line 257: Would be useful to give a reason for the assumption of rain cells starting at the storm origin.

In the Bartlett-Lewis rectangular pulse (BL) models, it is assumed that rain cells start at the storm origin largely for mathematical convenience. Because the BL cluster mechanism is defined by the interval between successive cells, a starting point is required. Therefore, it is convenient to assume that rain cells start at the storm origin. In contrast, the Neyman-Scott rectangular pulse (NS) cluster process is defined by the temporal distance between storm and cell origins and typically assumes that rain cells do not start at the storm origin. Again, this is largely for mathematical convenience.

In the case of the BL models, the assumption of rain cells starting at the storm origin prevents the simulation of empty storms which can occur if the first rain cell starts after the end of the storm. This issue does not arise in NS models because the number of cells per storm is a model parameter to be fitted, therefore a minimum value of one can be specified thus preventing the simulation of empty storms.

Line 316-319: Can appreciate the reason given for the choice of fitting statistics used for model calibration, but the question now is that how do the parameter estimates compare when the same fitting stats are used for uncensored fitting? Has this been explored?

This is an interesting question but one that we haven't explored. We will look into this although we don't feel that it will change the analysis presented in this paper.

Line 358: Perhaps you need to explain what you meant by behavioural parameters for the readers.

This was also highlighted by Anonymous Referee #1 therefore the response below is the same as that provided to referee #1.

We have used the term "behavioural parameters" by analogy with Beven and Binley (1992). We have used the term to refer to well identified models. We have found that for well identified parameters with narrow 95% confidence intervals, simulation bands on the extreme value estimates are correspondingly narrow. As the parameters become less well identified, their 95% confidence intervals increase giving rise to extreme value estimates which deviate significantly from the observations, which in turn results in significant deviation of the simulation band upper limit. This effect is shown in Fig.11 resulting from the very large parameter uncertainty shown in Fig.12.

We will remove the reference to "behavioural parameterizations" in the context of this research and change all references to well identified parameters.

Line 358: 95% confidence intervals: unless you are using the standard errors of the estimates, I am not sure whether "confidence" interval is the appropriate terminology here. Simulation bands?

This was also highlighted by Anonymous Referee #1 therefore the response below is the same as that provided to referee #1. We agree with the suggestion and will change all occurrences in the manuscript.

There are in fact two issues here: if we were doing very long simulations with practically no random noise (so that another simulation would yield practically the same result), then we would have identified approximate confidence intervals. But with the shorter simulation length, both parameter uncertainty and the randomness of the model are combined in the spread we observe in the simulated statistics, so that 'simulation bands' is indeed a better descriptor.

Line 396: Not sure why your validation for Atherstone was based on 0.6mm censor which seem to contradict your statement on lines 375-380. Some insight/explanation would be useful to the readers.

We agree that it would have been more consistent to select 0.2 mm for validation of the Atherstone hourly censored model. We will change the selection in Table.2 to include 0.2 mm for the hourly resolution at Atherstone (as below). We will then revise the validation plots to suit.

Table 2 Censor selection for model validation.

	5 minutes	15 minutes	60 minutes
Bochum	0.5 mm	1.0 mm	1.0 mm
Atherstone	0.6 mm	0.6 mm	0.2 mm

Line 396: Table 2. Different censor for different sites is understandable. However, why do you need to use the same sensor at 3 different aggregation level for Atherstone while using different censors for the 3 levels of aggregation for Bochum?

The model is fitted separately for each temporal resolution which explains why we have different censors. Because of the effect of aggregation, we cannot use a model censored at one resolution to estimate rainfall extremes at a coarser one. Therefore, censored model parameters are scale dependent which is explained in section 3.

The censors given in Table.2 were chosen for validation. Our analyses show that there are a range of censors that could be applied giving improved estimation of extremes. In the case of the two sites investigated, the gauge resolution at Atherstone is much coarser than that at Bochum. We note on lines 504-7 (page 28) that that a censor of 0.5 mm for 15 minute rainfall at Atherstone gives very similar extreme value estimation to the selected 0.6 mm censor, implying that it may be sufficient at this site to limit the censor to the gauge resolution. At Bochum, the finer gauge resolution will capture rainfall amounts with greater accuracy than at Atherstone. Therefore, we expect that there is greater capacity for the Bochum models to give improved estimation of extremes with increasing censors hence the different censors selected in Table.2.

Fig 8: row 2. The nice seasonal pattern observed in the mean rainfall for Atherstone at 5 and 15 minutes has become less prominent or disappeared at 60 minutes. Can you comment on why? No observation or comment was made about this.

The plots in Fig.8 show the summary statistics for censored rainfall with different censors applied in each column. While the censors chosen for validation in Table.2 are the same (0.6 mm), their effect on model fitting is different because they are applied to each temporal scale. Hence, when we look at the mean monthly rainfall in validation, we are looking at the seasonal variation in the rainfall after censoring.

Without censoring, the seasonal variation in mean monthly rainfall will only change in magnitude between scales. For a constant censor between scales as shown in panels d, e and f, the seasonal variation in mean monthly rainfall will vary between scales because there is a higher proportion of low observations at short temporal scales removed by the censors. The greater prominence in seasonal variation shown in plots d and e indicates that the summer months (approx. Apr - Oct) are more prone to short intense bursts of rain, and the winter months longer periods of low rainfall intensity. This is consistent with there being more convective rainfall in the summer, and stratiform rainfall in the winter.

References

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrol. Process., 6, 279-298, 1992.

2. Changes to the manuscript

2.1 In response to Anonymous Referee #1 (AR#1)

1. In response to AR#1, **page 11, line 288**:

New page 12, lines 291-4 We have update the manuscript with the following text.

The i-th summary statistic is weighted according to the inverse of its observed variance ωi = 1/var(ti) where ti is the vector of diagonal elements of the estimated covariance matrix of the mean summary statistics, Σ̂. While this weighting is not optimal, it provides a reasonable approximation to the optimal weights for the GMM giving robust estimation of the parameter standard errors (R. Chandler et al. 2010).

2. In response to AR#1, **page 12, lines 305-308**:

New page 12, Line 309 onwards: We have update the manuscript with the following text.

The sampling distribution of the estimators resulting from the GMM minimisation routine are approximately multivariate normal (MVN). The optimal parameter set is estimated by the mean of this distribution, and the covariance matrix is estimated from the Hessian of the least squares objective function S (Wheater et al. 2007b). The MVN distribution of model parameter estimators is used to estimate 95% confidence intervals for the parameter estimates. On occasions that the model parameters are poorly identified, it may not be possible to calculate the Hessian of the objective function preventing the estimation of parameter uncertainty.

3. In response to AR#1, **page 13, line 345**:
New page 14, Line 352 onwards: We have update the manuscript with the following text.

Extreme value estimation up to the 1000-year return level is also provided to indicate the potential magnitude of rarer events. For this extrapolation, extremes are estimated from one realisation using the mean of the MVN distribution of parameter estimators (hereafter referred to as the optimal estimates). To ensure stability of the extreme value estimates up to approximately the 1000-year return level, simulations have been extended to 10,000 years.

4. In response to AR#1, **pages 14-16, Figures 4-6**:
New page 28, Line 559 onwards: We have update the manuscript with the following text and included two new references.

In all three models, there is a slight upward curvature in the Gumbel plotting of extremes which is consistent with the GEV and GP distributions taking a positive shape parameter ($\xi > 0$). This curvature is more pronounced for the BL1M model which would be consistent with a higher positive shape parameter. While extreme value theory encompasses a range of distributions characterised by the sign of the shape parameter, Koutsoyiannis (2004a) argues that rainfall extremes naturally follow the Fréchet distribution for annual maxima (equivalent to the GEV with $\xi > 0$), supported with empirical evidence in Koutsoyiannis (2004b). The positive growth in extremes observed in our results is consistent with this hypothesis, and suggests that important information about the distribution of extremes is captured in the full storm profile hyetograph over the low censor. Futher research is required to investigate the theoretical link between mechanistic model parameters and their extreme value performance.

Koutsoyiannis, D.: Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: I. Recherche théorique, Hydrological sciences journal, 49, 575-590, 2004a.

Koutsoyiannis, D.: Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations, Hydrol. Sci. J. -J. Sci. Hydrol., 49, 591-610, 2004b.

5. In response to AR#1, **page 17, lines 385-386**:
New page 18, Line 392 onwards: We have update the manuscript with the following text.

These large confidence intervals indicate that the confidence in parameter estimation is reducing with higher censors and consequently the model error is too large for the reliable simulation of extremes.

6. In response to AR#1, **page 18, Section 6.2**:
New page 18, Line 395 onwards: We have changed the selected censor for validation of hourly extremes at Atherstone to 0.2 mm. This is reflected in Table 2, Fig. 8 and generally within the text.

7. In response to AR#1, **page 19, Figure 8 caption (and elsewhere)**:
New page 19, Fig.8 and elsewhere: We have changed the reference from "optimal censors" to "selected censors".

8. In response to AR#1, **page 28, line 516**:
New page 27, Line 524 onwards: We have changed the text as follows and added a new reference.

The actual number of peaks retained for fitting the Bartlett-Lewis models is much greater than this because serial dependence in the rainfall time-series is simulated with mechanistic modelling. While it is possible to estimate return levels for serially dependent extremes using extreme value theory, the analysis set out in Fawcett and Walshaw (2012) demonstrate that estimating the extremal index is non-trivial can be subjective.

Fawcett, L. and Walshaw, D.: Estimating return levels from serially dependent extremes, Environmetrics, 23, 272-283, 2012.

9. In response to AR#1, **page 3, line 81**:
New page 3, Line 80 onwards: We have changed the text as follows.

To test our hypothesis, a simple approach is proposed in which low observations for fine–scale data are censored from the models in calibration. For a given temporal resolution, a censor amount is set. Rainfall below the censor is set to zero and rainfall over the censor is reduced by the censor amount.

10. In response to AR#1, **page 4, line 118. "behavioural parameterizations"**:
The following changes have been made:

1. *Line 361 - Results presented are limited to the 4 highest censors with well identified model parameters, together with 95% confidence intervals derived from MVN realisations.*
2. *Line 446 - Figure 10 shows the change in 95% confidence intervals and the mean of the MVN realisations obtained with censored models with well and poorly identified parameters for 15 minute data at Bochum and Atherstone.*
3. *Line 465 - When sampling from the multivariate normal distribution for model parameters in simulation, these large uncertainties give rise to poor extreme value estimation.*
4. *Figure 10 Change in 95% confidence intervals and mean of the MVN realisations for Bochum and Atherstone 15 minute data with well identified (> 1.0 mm and > 0.6 mm) and poorly identified (> 1.5 mm and > 1.0 mm) censored model parameters.*

11. In response to AR#1, **page 6, line 160**:
New page 6, Line 160: We have changed the text as follows.

Despite this, it is likely that monthly extremes will have lower variability than those taken from the whole year, and hence model performance is likely to be better.

12. In response to AR#1, **page 10, line 262**:
New page 10, Line 264 onwards: We have changed the text as follows.

Atherstone is a tipping bucket rain gauge (TBR) operated and maintained by the Environment Agency of England. The record duration is 48 years from 1967 to 2015, with one notable period of missing data from January 1974 to March 1975. The reason for the missing data is unknown, although it is not expected to affect model fitting and the estimation of extremes.

13. In response to AR#1, **page 12, line 313**:
New page 10, Line 318 onwards: We have changed the text as follows.

While good model fits were obtained for some low censors, extreme value estimation continued to be underestimated.

14. In response to AR#1, **pages 14-16, Figures 4-6**:
New page 15-17, Figs.4-6: We have changed the figures as indicated and changed all references to annual exceedance probability or AEP to Return Period.

New page 22 and 24, Figs.10 and 12: We have changed these figures too in the same way.

15. In response to AR#1, **page 17, line 358**:
New page 14, Line 360 onwards: We have changed the text as follows. We have also changed references to simulation bands in Figs. 4-6 and their captions.

Upper limits on censoring were identified when model parameterization noticeably deteriorated resulting in the mean of the MVN realisations to deviate away from the optimal. Results presented are limited to the 4 highest censors with well identified model parameters, together with 95% simulation bands. The simulation bands show *the range of extreme value estimation between the 2.5 and 97.5 quantiles of the 100 MVN realisations for each simulated data point.*

New page 19, Line 408 onwards: We have also changed the text as follows in relation to the presentation of validation statistics as the simulation bands here are estimated in a similar way.

The plots show the estimated summary statistics calculated using the optimum parameter estimates, together with 95% simulation bands obtained by randomly sampling 100 parameter sets from the multivariate normal distribution of model parameters, estimating the summary statistics under the model and displaying the range of estimates between the 2.5 and 97.5 quantiles.

16. In response to AR#1, **page 17, line 379**:
New page 18, Line 386 onwards: We have changed the text as follows.

The mean of the MVN realisations for the BL1M model at Atherstone with the 0.6 and 0.8 mm censors (see Fig.6) diverges from the optimum because of the generation of unrealistic extremes.

17. In response to AR#1, **pages 20-21, Sections 6.2.1, 6.2.2 and 6.2.3**:
New page 18-21, Sections 6.2.1, 6.2.2 and 6.2.3: We have made the following changes.

1. We have removed calibration plots for Bochum, but retained those for Atherstone.
2. We have updated the hourly plots for Atherstone to reflect the changed censor selection in Table 2.
3. We have removed the validation plots for the coefficient of skewness at both sited, but retained the plots for proportion of wet periods at both sites.
4. We have removed Section 6.2.3.

18. In response to AR#1, **page 27, Figure 14**:
New page 26, Fig.13: We have retained this figure and renumbered it following the removal of the profile objective function plots.

19. In response to AR#1, **page 28, line 514**:
New page 26, Line 514 onwards: We have introduced the rule to create independent peaks before Table 3. We have also removed the quantiles of the rainfall amounts from Table.3 as they are provided in Fig.13. The following change to the text is made.

While the proportion of rainfall observations removed prior to model fitting is large - over 90% and 80% for 5 and 15 minute rainfall from Bochum and Atherstone respectively - comparison with the maximum rainfall amounts and an assessment of the number of independent peaks over the censor demonstrate that the censors are low. Table 3 shows the proportion of maximum rainfall and the number of independent peaks per year for the selected censors given in Table. 2. The number of peaks over the censors are estimated using a temporal separation of 48 hours to define independence.

20. In response to AR#1, **pages 28-30**:
New page 27-9, Section 8: The manuscript has been updated as follows:

1. Section 8 has been renamed “Further discussion and conclusions”
2. Section 9 has been removed and combined with section 8.
3. The first sentence from the old Section 9 has been removed because it repeats the first sentence of Section 8. “Censored rainfall synthesis using mechanistic pulse based models appears to offer an alternative approach to estimation of rainfall extremes and to frequency estimation techniques”
4. The following changes have been made to the subsequent two sentences: *“The results presented in this paper show that the method has worked for single site data from two very different locations, and recorded using different gauging techniques. Consistency in the relative magnitude of selected censors identified at each location, and the stability of estimation across a range of censors gives confidence in the approach and supports the original hypothesis.”*

5. Discussion on the upward curvature in the Gumbel extreme value plots is included.

21. In response to AR#1, **page 31, line 591**:
New page 30, Line 608: The inequality is corrected.

22. In response to AR#1, **page 32, Figure A.2**:
New page 31, Fig.A2: The figure is corrected.

New page 9, Line 228: We have changed the text as follows.

Both X and L are assumed to be independent of each other and follow exponential distributions with parameters 1/μ_x and η respectively.

2.2 In response to Anonymous Referee #2 (AR#2)

1. In response to AR#2, **Line 257**:
New page 10, lines 258 onwards: We have update the manuscript with the following text.

For the models used in this study, it is assumed that rain cells start at the storm origin to prevent the simulation of empty storms which can occur with the Bartlett-Lewis clustering mechanism if the first rain cell starts after the end of the storm.

2. In response to AR#2, **Line 316-319**:
We have not update the manuscript with respect to this comment.

3. In response to AR#2, **Line 358**: (as per AR#1, comment no. 10)
The following changes have been made:

1. *Line 361 - Results presented are limited to the 4 highest censors with well identified model parameters, together with 95% confidence intervals derived from MVN realisations.*
2. *Line 446 - Figure 10 shows the change in 95% confidence intervals and the mean of the MVN realisations obtained with censored models with well and poorly identified parameters for 15 minute data at Bochum and Atherstone.*
3. *Line 465 - When sampling from the multivariate normal distribution for model parameters in simulation, these large uncertainties give rise to poor extreme value estimation.*
4. *Figure 10 Change in 95% confidence intervals and mean of the MVN realisations for Bochum and Atherstone 15 minute data with well identified (> 1.0 mm and > 0.6 mm) and poorly identified (> 1.5 mm and > 1.0 mm) censored model parameters.*

4. In response to AR#2, **Line 358**: (as per AR#1, comment no. 15)
New page 14, Line 360 onwards: We have changed the text as follows. We have also changed references to simulation bands in Figs. 4-6 and their captions.

Upper limits on censoring were identified when model parameterization noticeably deteriorated resulting in the mean of the MVN realisations to deviate away from the optimal. Results presented are limited to the 4 highest censors with well identified model parameters, together with 95% simulation bands. The simulation bands show the range of extreme value estimation between the 2.5 and 97.5 quantiles of the 100 MVN realisations for each simulated data point.

New page 19, Line 408 onwards: We have also changed the text as follows in relation to the presentation of validation statistics as the simulation bands here are estimated in a similar way.

The plots show the estimated summary statistics calculated using the optimum parameter estimates, together with 95% simulation bands obtained by randomly sampling 100 parameter sets from the multivariate normal *distribution of model parameters, estimating the summary statistics under the model and displaying the range of estimates between the 2.5 and 97.5 quantiles.*

5. In response to AR#2, **Line 396**: (as per AR#1, comment no. 6)

New page 18, Line 395 onwards: We have changed the selected censor for validation of hourly extremes at Atherstone to 0.2 mm. This is reflected in Table 2, Fig. 8 and generally within the text.

6. In response to AR#2, **Line 396: Table 2:**

We have not update the manuscript with respect to this comment.

7. In response to AR#2, **Fig 8: row 2:**

Section 6.2 has substantially changed following comments from both reviewers. The following changes listed in response to AR#1 comment no. 17 are relevant to this comment by AR#2.

New page 18-21, Sections 6.2.1, 6.2.2 and 6.2.3: We have made the following changes.

1. We have removed calibration plots for Bochum, but retained those for Atherstone.
2. We have updated the hourly plots for Atherstone to reflect the changed censor selection in Table 2.
3. We have removed the validation plots for the coefficient of skewness at both sited, but retained the plots for proportion of wet periods at both sites.
4. We have removed Section 6.2.3.

New page 20, lines 415 onwards: We have update the manuscript with the following commentary as suggested.

All models perform very well with respect to replicating the summary statistics used in fitting with the 95% simulation bands comfortably bracketing the observations. The estimated summary statistics are very close to the observed with all models performing equally well. The seasonal variation in mean monthly rainfall varies between scales because there is a higher proportion of low observations at short temporal scales removed by the censors. The greater prominence in seasonal variation shown in plots a and b indicates that the summer months (approx. Apr - Oct) are more prone to short intense bursts of rain, and the winter months longer periods of low rainfall intensity. This is consistent with there being more convective rainfall in the summer, and stratiform rainfall in the winter. The plots in Fig. 8 demonstrate that the models are able to reproduce the censored fitting statistics, confirming reliability of the process.

New page 21, lines 431 onwards: We have update the manuscript with the following additional commentary on the ability of the models to reproduce the proportion of wet periods.

The ability of the models to reproduce the proportion of wet periods is generally good, although there is a tendency for all models to overestimate this statistic at both sites. At the 5 minute resolution for Bochum, the 95% simulation bands comfortably bracket the observations between the months of May and October, although there is over-estimation in the other months and for all months at the 15 and 60 minute scales. At Atherstone, there is good representation of the proportion of wet periods at the 15 minute scale, although over-estimation at the 5 and 60 minute scales. Generally, there is very slightly better agreement in the summer months which, as highlighted in Sect. 6.2.1, may be more prone to short intense downpours at fine temporal scales. This suggests that the censored models may be more effective at simulating the heavier short duration rainfall characteristic of summer convective storms, than the longer duration low intensity rainfall characteristic of winter storms.

3. Track changes to the manuscript

Manuscript commences on the next page.

Censored rainfall modelling for estimation of fine–scale extremes

David Cross[1], Christian Onof[1], Hugo Winter[2], Pietro Bernardara[2]

[1]Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, UK

[2]EDF Energy R&D UK Centre, London SW1E 5JL, UK

Correspondence to: David Cross (david.cross12@imperial.ac.uk)

Abstract. Reliable estimation of rainfall extremes is essential for drainage system design, flood mitigation and risk quantification. However, traditional techniques lack physical realism and extrapolation can be highly uncertain. In a warming climate, the moisture holding capacity of the atmosphere is greater which increases the potential for short duration high intensity storm events. In this study, we improve the physical basis for short duration extreme rainfall estimation by simulating the heavy portion of the rainfall record mechanistically using the Bartlett-Lewis rectangular pulse model. Mechanistic rainfall models have had a tendency to underestimate rainfall extremes at fine temporal scales. Despite this, the simple process representation of rectangular pulse models is appealing in the context of extreme rainfall estimation because it emulates the known phenomenology of rainfall generation. A censored approach to Bartlett-Lewis model calibration is proposed and performed for single site rainfall from two gauges in the UK and Germany. Extreme rainfall estimation is performed for each gauge at the 5, 15 and 60 minute resolutions, and considerations for censor selection discussed.

Deleted: is

1. Introduction

With growing evidence that the frequency and intensity of short duration rainfall extremes are increasing with climate change (Stocker et al. 2013, Westra et al. 2014, Kendon et al. 2014), the need for reliable extreme value estimation techniques is becoming more pressing. Extreme rainfall estimation is required for numerous applications in diverse disciplines ranging from engineering and hydrology to agriculture, ecology and insurance. It facilitates the planning, design and operation of key municipal infrastructure such as drainage and flood defences, as well as scenario analysis for climate impact assessment, and hazard risk modelling. Extremes are usually estimated using frequency techniques and intensity duration frequency curves. However, these methods are highly dependent on the observed rainfall record which may not be characteristic of the extreme behaviour.

In this study we improve the physical basis of short duration extreme rainfall estimation by simulating the heavy portion of the observed rainfall time-series. Traditional approaches to extreme value estimation rely on sampling extremes from the observed record. However, rainfall observations present various problems for the practitioner. They are often not available at the location of interest, they are typically short in duration, and they may not be available at the temporal scale appropriate for the intended use. These difficulties, together with the necessity to obtain perturbed time-series representative of future rainfall, have motivated the development of stochastic rainfall generators since the earliest such statistical models developed by Gabriel and Neumann (1962). The reader is referred to Waymire and Gupta (1981), Wilks and Wilby (1999) and Srikanthan and McMahon (2001) for detailed reviews of early developments in rainfall simulation.

The principle of rainfall simulation is to replicate statistical properties of the observed record such that multiple realisations of statistically identical rainfall may be synthesized (Richardson 1981). Various methods of simulation exist, and there have been several attempts in the literature to categorize the different approaches. Aside from dynamic methods used in numerical weather prediction models, Cox and Isham (1994) suggest that statistical simulation methods may be broadly categorized as either purely statistical or stochastic, while Onof et al. (2000) further categorize stochastic methods into either multi-scaling or mechanistic. The latter of these differ from other statistical approaches because rainfall synthesis follows a simplified representation of the physical rainfall generating mechanism. Through the clustering of rain cells in storms, the unobserved continuous-time rainfall is constructed by superposition, enabling the synthetic rainfall hyetograph to be aggregated to whatever scale is desired (Kaczmarska et al. 2014). Because of this simplified process representation, mechanistic model parameters have physical meaning which makes this class of model particularly appealing in the context of extreme value estimation.

When no likelihood function can be formulated (Rodriguez-Iturbe et al. 1988, R. E. Chandler 1997), mechanistic models are typically calibrated using a generalised method of moments (Wheater et al. 2007a) with key summary statistics at a range of temporal scales such as the mean, variance, autocorrelation and proportion of dry periods. Performance is assessed on the ability of the models to reproduce the calibration statistics as well as others not used in calibration including central moments and extremes. Since their inception in the late 1980s by Rodriguez-Iturbe et al. (1987, 1988), numerous studies have demonstrated the ability of these models to satisfactorily reproduce observed summary statistics [see Cowpertwait et al. (1996), N. Verhoest et al. (1997), Cameron et al. (2000a, 2000b), Kaczmarska et al. (Kaczmarska et al. 2014), Wasko and Sharma (2017) and Onof et al. (2000) for a review]. However, these studies have also shown that mechanistic models tend to underestimate rainfall extremes at the hourly and sub-hourly scales which limits their usefulness [see Verhoest et al. (2010) and references therein].

We hypothesize that stochastic mechanistic pulse-based models may be poor at estimating fine–scale extremes because the training data, and calibration method, are dominated by low intensity observations. Mechanistic stochastic models are fitted to the whole rainfall hyetograph, including zeroes, aggregated to a range of temporal scales. Typically, the range of scales used varies from hourly to daily, although implicit in most studies is the assumption that scales required in simulation should be within the range of scales used in calibration. Hence, if the intention of the model is to simulate 15 minute rainfall the training data should include 15 minute observations. As the temporal resolution of rainfall data becomes finer, the distribution of rainfall amounts becomes more positively skewed. Primarily, this is because of the increased proportion of dry periods, but also the higher proportion of low intensity events characteristic of fine–scale rainfall. Because the calibration method uses

Deleted: &
Deleted: &
Deleted: &
Deleted: &
Deleted: &
Deleted: (Rodriguez-Iturbe et al. 1988, Chandler 1997)
Deleted: Rodriguez-Iturbe, Cox & Isham
Field Code Changed
Deleted: (1987, 1988)
Field Code Changed
Deleted: (1996)
Deleted: Verhoest, Troch & De Troch
Field Code Changed
Deleted: (1997)
Deleted: Cameron, Beven & Tawn
Field Code Changed
Deleted: (2000a, 2000b)
Deleted: Kaczmarska, Isham & Onof
Field Code Changed
Deleted: (2014)
Deleted: &

central moments to fit model parameters, the greater skewness at finer temporal scales makes it difficult to obtain a good fit to extremes at these scales.

In addition to the dominance of low observations, the estimation of fine–scale extremes may be further undermined by operation and sampling errors. This is particularly true of tipping bucket gauges where measurement precision at fine temporal scales is limited to the bucket volume, typically 0.2 or 0.5 mm. Fine–scale rainfall is highly intermittent (starting and stopping with high frequency), yet a tipping bucket gauge can only make a recording when the bucket is full. The limitations of tipping bucket measurements at fine temporal scales have long been understood (Goldhirsh et al. 1992, Nystuen et al. 1996, Yu et al. 1997), although the first formal estimation of sampling error was performed by Habib et al. (2001). In this study, the authors investigate the ability of tipping bucket gauges to capture the temporal variability of fine–scale rainfall at 1, 5 and 15 minute scales using tipping bucket measurements simulated from high resolution optical rain gauge observations. The authors show that for the lowest rainfall intensities (< 5 mm/h) the mean relative error of the tipping bucket gauge at the 5 minute resolution is + 3.5%, with corresponding standard deviation just under + 30% for a bucket volume of 0.254 mm. Larger errors are obtained for the 1 minute resolution. They also show that increasing the bucket volume to 0.5 mm significantly increases the spread of the sampling error for low observations at the 5 minute resolution. The observed record comprised mainly convective storm events which are typical for Iowa in the US where the data were collected, although the error estimates are significant and present compelling evidence of the impact of sampling error on fine–scale low intensity rainfall observations.

Significant effort has been made since the late 1980s to improve the performance of mechanistic rainfall models through structural developments, with substantial focus on the improved representation of fine–scale extremes (see Sect. 2 for a review). Despite this, little progress has been achieved. To test our hypothesis, a simple approach is proposed in which low observations for fine–scale data are censored from the models in calibration. For a given temporal resolution, a censor amount is set. Rainfall below the censor is set to zero and rainfall over the censor is reduced by the censor amount. This focusses model fitting on the heavier portion of the rainfall record at fine temporal scales, and reduces rainfall intensity at coarser scales. The aim is to investigate if existing mechanistic models can be used as simulators of fine–scale storm events by changing the data and not the model, thereby reducing the impact of low observations and sampling error on fine–scale extreme rainfall estimation.

The choice of models is limited to those within the Bartlett-Lewis family of models which conform to the original concept of rectangular pulses developed by Rodriguez-Iturbe et al. (1987). Preference is given to the most parsimonious model variants on the basis that having fewer parameters improves parameter identifiability and reduces uncertainty. The Neyman-Scott family of models is excluded on the understanding that the clustering mechanisms of both model types perform equally well (Wheater et al. 2007a), and there is no evidence that randomisation of the Neyman-Scott model (Entekhabi et al. 1989) has any advantage over its Bartlett-Lewis counterpart.

Deleted: Habib, Krajewski & Kruger

Field Code Changed

Deleted: (2001)

Deleted: Rodriguez-Iturbe, Cox & Isham

In Sect. 2, we outline the main mechanistic model developments for improved representation of extremes. The Censored modelling approach for the estimation of fine–scale extremes is described in Sect. 3. Model structure and selection is explained in Sect. 4, and the data and fitting methodology are presented in Sect. 5. Results are given in Sect. 6 together with validation analysis. Discussion on censor selection and the results are given in Sects. 7 and 8. Section 9 outlines our main conclusions and thoughts for future research.

2. Mechanistic model developments

Attempts to improve the estimation of fine–scale extremes for point (single-site) rainfall using mechanistic models have focused on changing the model structure. Several authors have cited significant improvement (P. Cowpertwait 1994, Cameron et al. 2000b, Evin and Favre 2008), although increased parameterization and limited verification with real data have meant that most changes have not been widely adopted. An early criticism of the original mechanistic models presented by Rodriguez-Iturbe et al. (1987) is that the exponential distribution applied to rainfall intensities is light-tailed. This choice is consistent with the observation that rainfall amounts, which in the model are obtained through the superposition of such cells, are approximately Gamma distributed (Katz 1977, Stern and Coe 1984).

On the basis that the Gamma distribution gives more flexibility in generating rain cell intensities, Onof and Wheater (1994b) reformulate the modified (random η) Bartlett-Lewis (MBL) model (Rodriguez-Iturbe et al. 1987) with the Gamma distribution to improve the estimation of extremes. Despite the good fit to hourly extremes cited by Onof and Wheater (1994b), subsequent studies have continued to show underestimation at hourly and sub-hourly scales (N. Verhoest et al. 1997, Cameron et al. 2000a, Kaczmarska et al. 2014).

In an extension of this approach, Cameron, Beven and Tawn (2000b) replace the exponential distribution in the MBL model with the Generalized Pareto (GP) distribution for rain cells over a high threshold. Depending on the value of the shape parameter (ξ), the GP converges to one of three forms: upper-bounded (ξ < 0), exponential (ξ = 0) and Pareto (ξ > 0). In the last case, we have a distribution with a heavier tail than the exponential or the gamma. Because the GP is a model for threshold exceedance, the authors specify a threshold below which the MBL model with exponential intensity distribution is used to simulate rain cell depth, and above which the Pareto distribution is used. This is justified on the assumption that the rain cell depth may be of either high or low intensity.

The authors present a calibration strategy in which they first fit the MBL model with exponential cell depths to the whole rainfall record using the method of moments from Onof and Wheater (1994b). Generalised Likelihood Uncertainty Estimation (Beven and Binley 1992) is then used to find behavioural parameterizations of the Pareto scale and shape parameters for rain cell depths over the threshold – the location parameter being fixed at the threshold value. The central assumption of this model is that the Pareto scale and shape parameters for cells depths over the threshold will have “minimal impact on the standard statistics of the simulated continuous rainfall time-series” (Cameron et al. 2000b, p.206). The validity of this assumption is

Deleted: (Cowpertwait 1994, Cameron et al. 2000b, Evin and Favre 2008)

Deleted: Rodriguez-Iturbe, Cox & Isham

Deleted: &

Deleted: &

Deleted: (Verhoest et al. 1997, Cameron et al. 2000a, Kaczmarska et al. 2014)

Deleted: &

Deleted: D

Deleted: (GPD)

Deleted: D

Deleted: D

Deleted: &

disputed by Wheater et al. (2007a) who argue that the MBL model should be fitted to rainfall coincident with rain cells below the threshold, but point out that this is "impossible since cell intensities are not observed" (Wheater et al. 2007a, p.16).

The model framework of Cameron, Beven and Tawn (2000b) differs from that of the MBL Gamma model of Onof and Wheater (1994a) and is essentially the nesting of two models. The authors present significant improvement in the estimation of hourly extremes and show good agreement with Generalized Extreme Value (GEV) estimates. However, because the underlying process of continuous-time rainfall is unobserved, the authors are forced to implement a calibration strategy which limits the impact on standard rainfall statistics – an approach which is undesirable (Wheater et al. 2007a). Furthermore, the framework appears to be an analogue of the N-cell rectangular pulse model structure initially developed by P. Cowpertwait (1994) for the Neyman-Scott model, and later incorporated into the Bartlett-Lewis models by Wheater et al. (2007a). Regardless of their relative performance, the large number of parameters required for these models is undesirable on the basis that more parameters reduces parameter identifiability and increases parameter uncertainty.

In an earlier study, P. Cowpertwait (1994) differentiates between light and heavy rain cells in a modified version of the original (fixed η) Neyman-Scott Rectangular Pulse (NSRP) model (Rodriguez-Iturbe et al. 1987) by allowing rain cell intensity and duration to be drawn from more than one pair of exponential distributions. The new model termed the Generalised NSRP model (GNSRP) leads to a significant increase in parameterization over the original NSRP model, although the author presents an intelligent way to simplify calibration by relating model parameters to harmonic signals. While improvement is achieved in the fit to hourly extremes, the performance of the model in replicating other important statistics is not presented, in particular autocorrelation and the proportion of dry periods. Both of these properties are addressed by Rodriguez-Iturbe et al. (1987, 1988) for the Bartlett-Lewis model with the inclusion of a "high frequency jitter" and randomisation of the rain cell duration parameter η. Entekhabi et al. (1989) present a randomised version of the Neyman-Scott model with significant improvement in the fit to dry periods. However, because no analytical expression was available for the proportion of dry periods this statistic was not used in model fitting, and other model parameters were not allowed to vary from storm to storm with randomisation. Consequently, while the MBL and the GNSRP models each allow rain cell intensity and duration to be drawn from more than one pair of distributions, the MBL structure is preferred because it has fewer parameters.

In a later study, P. Cowpertwait (1998) hypothesised that including higher-order statistics in the fitting routine for mechanistic rainfall models would give a better fit to the tail of the empirical distribution for rainfall amounts. Focussing on the original (fixed η) NSRP model, analytical equations for skewness of the aggregated rainfall depth are presented and used in fitting the models. Empirical analysis showed that including skewness in the fitting statistics improved the estimation of Gumbel distribution parameters from simulated maxima when compared with parameters obtained from observed annual maxima.

A criticism of the rectangular pulse model structure by Evin and Favre (2008) is that it assumes independence between rain cell intensity and duration. Following previous attempts to link the two variables [Kakou (1997), De Michele and Salvadori (2003), Kim and Kavvas (2006)], Evin and Favre (2008) present a new NSRP model in which the dependence between rain

Deleted: &

Deleted: &

Deleted: Cowpertwait

Field Code Changed

Deleted: (1994)

Field Code Changed

Deleted: (1994)

Deleted: Rodriguez-Iturbe, Cox & Isham

Deleted: , Rodriguez-Iturbe & Eagleson

Deleted: &

Field Code Changed

Deleted: (1997)

Deleted: &

Deleted: &

Deleted: &

cell depth and duration is explicitly modelled using a selection of copulas. While the authors are not primarily motivated to improve the estimation of rainfall extremes, good estimation of fine–scale extremes is achieved. However, the manner in which the results are presented makes interpretation and comparison with other studies difficult. In the first instance, the extreme performance of all models is almost entirely indistinguishable indicating that no overall improvement is achieved. Secondly, monthly annual extremes are presented at hourly and daily scales but without clearly stating which month in the year. Despite this, it is likely that monthly extremes will have lower variability than those taken from the whole year, and hence model performance is likely to be better. On the basis of the results presented, it is not clear that explicitly modelling dependence between rain cell depth and duration with copulas offers any discernable benefit over the original model structure.

Deleted: less variance

Theoretically, copulas offer an attractive framework for modelling the dependence structure between rainfall intensity and duration. However, the obvious mechanism for building copula dependence into mechanistic rainfall models is at the rain cell level as per Evin and Favre (2008). This approach draws upon the intuition that, just as for the rainfall amounts of storm events, rain cell amounts may be correlated with their duration. Such intuition follows earlier studies into the dependence structure between rainfall intensity and duration (Bacchi et al. 1994, Kurothe et al. 1997) – although as stated by Vandenberghe et al. (2011, p.14) *"it is not very clear in which way this modelled dependence at cell level alters the dependence between the duration and mean intensity of the total storm"*.

Deleted: &

In recent years, renewed focus on estimating rainfall extremes at hourly and sub-hourly scales has led to the development of a new type of mechanistic rainfall model based on instantaneous pulses (P. Cowpertwait et al. 2007, Kaczmarska 2011). In this model structure, rectangular pulses are replaced with a point process of instantaneous pulses with depth X and zero duration, the summation of pulses giving the aggregated time rainfall intensity. Considered initially to offer a more suitable representation of rainfall at sub-hourly scales than rectangular pulses, Kaczmarska et al. (2014) found that the best performing Bartlett-Lewis Instantaneous Pulse (BLIP) model effectively generated rectangular pulses when depth X was kept constant, and cell duration η was randomized. Because of the very large number of pulses generated within cells, the authors noted that this model structure imposes the *"most extreme form of dependence"* Kaczmarska et al. (2014, p.1977). Consequently, the authors developed a new rectangular pulse model in which both η and μ_x are randomized ($BLRPR_X$) which was found to perform equally as well as the randomised version of the BLIP model but without additional parameterization.

Deleted: (Cowpertwait et al. 2007, Kaczmarska 2011)

Deleted: , Isham & Onof

Deleted: , Isham & Onof

3. Censored modelling for fine–scale extremes

Despite the model improvements outlined in Sect. 2, there is an on-going tendency for stochastic mechanistic models to underestimate extremes at hourly and sub-hourly scales, requiring the practitioner to employ additional methods for better extreme value performance including disaggregation (Koutsoyiannis and Onof 2000, Koutsoyiannis and Onof 2001, Onof et al. 2005, Onof and Arnbjerg-Nielsen 2009, Kossieris et al. 2013). We propose a censored approach to mechanistic rainfall modelling for improved estimation fine–scale extremes by focussing model fitting on the heavy portion of the rainfall time-series. The aim of this research is to investigate if mechanistic models can be used as simulators of fine–scale design storm events to reduce the impact of low observations on the estimation of fine–scale extremes. In this approach, rainfall below a low censor is set to zero and rainfall over the censor is reduced by the censor amount. The effect is to generate a time-series of heavy rainfall based on the observed record in which the proportion of dry periods is increased, and rainfall amounts are reduced.

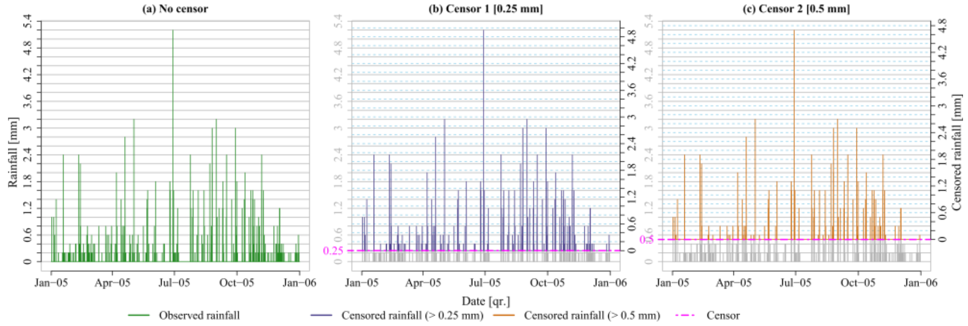

Figure 1 Example censoring applied to 15 minute rainfall data at Atherstone in 2005. Arbitrary censors of 0.25 mm and 0.5 mm are applied to demonstrate the effect of censoring on the rainfall record.

Deleted: 1

Figure 1 shows two arbitrary censors applied to 15 minute data at Atherstone in 2005 (refer to Sect. 5.1 for a description of the data). The left plot shows the uncensored rainfall, and the two right plots the change in rainfall with increasing censors. The reduced rainfall amounts are shown on the secondary y axes. It can be seen from these plots that the minimum recorded rainfall is 0.2 mm which corresponds to the tip volume of the tipping bucket rain gauge. Compared with higher rainfall amounts this volume is recorded with very high frequency throughout the year at the 15 minute resolution.

Censored rainfall synthesis is a method for estimating sub-hourly to hourly extremes. Because observations below the censor are omitted from model fitting, censored model parameters are scale dependent and can only be used to simulate storm profiles above the censor at the same scale as the training data. It is the ability to simulate the heavy portion of storm profiles which enables extreme rainfall estimation. The basic procedure is as follows:

1. For the chosen temporal resolution, select a suitable censor [mm] and apply it to the observed rainfall time-series by setting rainfall amounts below the censor to zero, and reducing rainfall amounts over the censor by the censor amount.
2. Fit the mechanistic rainfall model to the censored rainfall by aggregating the censored time-series to a range of temporal scales and calculating summary statistics as necessary for model fitting.
3. Simulate synthetic rainfall time-series at the same resolution as the training data in Step 1 and sample annual maxima.
4. Restore the censor to the simulated annual maxima and plot against the observed.

4. Model structure and selection

Mechanistic point process rainfall models, first developed by Rodriguez-Iturbe et al. (1987) exist in various forms, although all models are formulated around two key assumptions about the rainfall generating process. Firstly, rainfall is assumed to arrive in rain cells following a clustering mechanism within storms. Secondly, the total rainfall within cells is represented by a pre-specified rainfall pattern which describes the rain cell duration and amount. The continuous time rainfall is the summation of all rainfall amounts in time Δt. Most models assume rectangular pulses to describe rainfall amount and duration, although alternative patterns have included a Gaussian distribution (Northrop and Stone 2005) and instantaneous pulses (P. Cowpertwait et al. 2007, P. S. P. Cowpertwait et al. 2011, Kaczmarska et al. 2014). In this latter formulation, pulses are assumed to arrive according to a Poisson process within cells, with each pulse representing an amount with zero duration. The continuous-time rainfall is therefore the summation of all pulse amounts in time Δt. Figure 2 shows a schematic representation of the rainfall generating process for rectangular pulse models.

Deleted: Rodriguez-Iturbe, Cox & Isham

Deleted: (Cowpertwait et al. 2007, Cowpertwait et al. 2011, Kaczmarska et al. 2014)

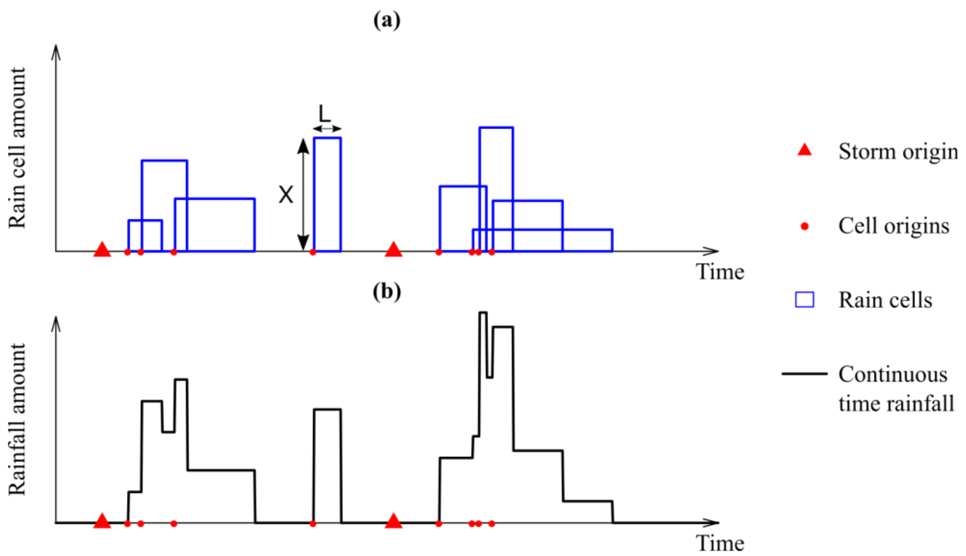

Figure 2 Rainfall generation mechanism for mechanistic stochastic models with rectangular pulses. Panel (a) shows the arrival of storms and cells. Raincell intensity defines the height of each cell (X), and duration the length (L). Panel (b) shows the unobserved continuous-time rainfall time-series derived from the superposition of cells shown in (a).

Deleted: 3

In the original form of the model, storms arrive according to a Poisson process with rate λ, and terminate after an exponentially distributed period with rate γ. The arrival of rain cells within storms follows a clustering mechanism which defines a secondary Poisson process with rate β. Two clustering mechanisms are specified by Rodriguez-Iturbe et al. (1987): the first is the Neyman-Scott mechanism in which the time intervals between storm and cell origins are assumed to be independent and identically distributed random variables; the second is the Bartlett-Lewis mechanism in which the time intervals between successive cell origins are independent and identically distributed random variables. In each case, the time intervals are

Deleted: Rodriguez-Iturbe, Cox & Isham assumed to be exponentially distributed. Rain cell profiles are rectangular with heights X for amounts, and lengths L for durations. Both X and L are assumed to be independent of each other and follow exponential distributions with parameters $1/\mu_x$ and η respectively. See Fig. 2 for a graphical illustration of the continuous-time rainfall generation process.

Deleted: and

The original Neyman-Scott and Bartlett-Lewis rectangular pulse models (NSRP and BLRP respectively) with exponential cell depth distributions are the most parsimonious models, each having only 5 parameters (see Table 1). A limitation of these models is that their simplicity implies all rainfall – stratiform, convective, and orographic - has the same statistical properties. On the assumption that rainfall may derive from different storm types, in particular convective and stratiform, it is physically

Table 1 Model parameters for original and randomized BLRP models and the original NSRP model.

	Units	BLRP	NSRP	BLRPR$_\eta$	BLRPR$_X$
Storm arrival rate	hr^{-1}	λ	λ	λ	λ
Cell arrival rate	hr^{-1}	β	β	{β}$^{(1)}$	{β}
Ratio of cell arrival rate to cell duration	-	-	-	κ = β / η	κ = β / η
Mean cell depth	mm hr^{-1}	μ_x	μ_x	μ_x	{μ_x}
Ratio of mean cell depth to cell duration	mm	-	-	-	ι = μ_x / η
Ratio of standard deviation to the mean cell depth	-	r = σ_x / μ_x	r = σ_x / μ_x	r = σ_x / μ_x	r = σ_x / μ_x
• Expected square of the cell depth $^{(2)}$	mm^2 hr^{-2}	{μ_{x^2}}	{μ_{x^2}}	{μ_{x^2}}	{μ_{x^2}}
• Expected cube of the cell depth for inclusion of skewness in the objective function $^{(2)}$	mm^3 hr^{-3}	{μ_{x^3}}	{μ_{x^3}}	{μ_{x^3}}	{μ_{x^3}}
Cell duration parameter	hr^{-1}	η	η	{η}	{η}
• Gamma scale parameter for η	-	-	-	ν	ν
• Gamma shape parameter for η	hr			α	α
Storm duration parameter	hr^{-1}	γ	-	{γ}	{γ}
Ratio of storm duration to cell duration	-			φ = γ / η	φ = γ / η
Mean number of cells per storm	-	-	μ_c	-	-
Number of parameters: exponential cell intensity	-	5	5	6	6
Number of parameters: gamma cell intensity	-	6	6	7	7

NOTES:

1. Parameters in curly brackets {} are not included in the objective function (see Sect. 5.2).
2. For the two parameter gamma cell depth distribution, the expected square and cube of the cell depth (μ_{x^2} and μ_{x^3}) are calculated from the standard deviation (σ_x) and mean (μ_x) of the cell depth. In practice it is the ratio of these (r) which is parameterized enabling calculation of μ_{x^2} and μ_{x^3}. For both the exponential and gamma distributions, $\mu_{x^2} = f_1 \mu_X{}^2$ and $\mu_{x^3} = f_2 \mu_X{}^3$ where $f_1 = 1 + r^2$ and $f_2 = 1 + 3r^2 + 2r^4$. Because the exponential distribution is a special case of the gamma distribution where r is equal to 1, $\mu_{x^2} = 2\mu_x{}^2$ and $\mu_{x^3} = 6\mu_x{}^3$. Therefore it is not necessary to parameterize r for the exponential distribution, meaning the exponential versions of these models require 1 parameter less with r set to 1 in calibration.

more appealing to allow the statistical composition of rainfall models to vary between storms.

Two different approaches have been developed to accommodate the simulation of different rainfall types with rectangular pulses. For the Neyman-Scott model, concurrent and superposed process have been developed in generalised (P. Cowpertwait 1994) and mixed (P. Cowpertwait 2004) rectangular pulse models respectively. Both models enable explicit simulation of multiple storm types, although their increased parameterization and consequent impact on parameter identifiability means that

Deleted: (Cowpertwait 1994)

Deleted: (Cowpertwait 2004)

it is undesirable to simulate more than two storm types. For the Bartlett-Lewis model, randomization of the rain cell duration parameter η (Rodriguez-Iturbe et al. 1988, Onof and Wheater 1993, 1994b) with a Gamma distribution allows all storms to be drawn from different distributions. Because rain cell durations are assumed to be exponentially distributed, rain cells with high values of η are more likely to be shorter in duration, and those with low values of η will typically have longer durations. Additionally the rate at which rain cells arrive, and the storm durations, are defined in proportion to η by keeping the ratios β/η and γ/η constant (equal to κ and φ respectively). This means that typically, shorter storms will comprise shorter rain cells with shorter rates of arrival and the opposite for longer storms, which is characteristic of the differences between convective and stratiform rainfall.

The modified (random η) Bartlett-Lewis model (see $BLRPR_{\eta}$ in Table 1) of Onof and Wheater (1993, 1994b) is the most parsimonious of the model structures able to accommodate multiple storm types comprising a minimum of 6 parameters for the exponential version. The modified (random η) Neyman-Scott model has the same number of parameters as the modified Bartlett-Lewis model, but because there is no evidence that it has any advantage over the latter it is excluded from this study. The updated random η Bartlett-Lewis model with mean cell depth μ_x also randomised (see $BLRPR_X$ in Table 1) requires fewer parameters than its instantaneous pulse counterpart and the same number of parameters as the modified $BLRPR_{\eta}$ model. Structurally, it is identical to the modified model, although μ_x is also allowed to vary randomly between storms by keeping the ratio $\iota = \mu_x/\eta$ constant.

Deleted: &

Because the Neyman Scott and Bartlett Lewis clustering mechanisms are considered to perform equally well, model selection is limited to the most parsimonious model structures within the Bartlett-Lewis family of models: the original model (BLRP), the linear random parameter model (BLRPRη) and the linear random parameter model with randomized μ_x ($BLRPR_X$). Hereafter, these models are referred to as BL0, BL1 and BL1M respectively. For the models used in this study, it is assumed that rain cells start at the storm origin to prevent the simulation of empty storms which can occur with the Bartlett-Lewis clustering mechanism if the first rain cell starts after the end of the storm.

5. Data and model fitting

5.1 Data selection

Estimation of fine–scale extremes with censored rainfall simulation is performed on two gauges: Atherstone in the UK and Bochum in Germany. Atherstone is a tipping bucket raingauge (TBR) operated and maintained by the Environment Agency of England. The record duration is 48 years from 1967 to 2015, with one notable period of missing data from January 1974 to March 1975. The reason for the missing data is unknown, although it is not expected to affect model fitting and the estimation of extremes. This site was selected from all TBRs for the Environment Agency's Midlands Region on the basis that the number of Environment Agency quality flags highlighted as "good" in the record is greater than 90%, and the number of "suspect" flags less than 10% (92.3% and 6.7% respectively). Between the 8th February 1981 and after 20th November 2003 the gauge resolution is 0.5 mm. Before and after this period it is 0.2 mm. In the period before the 8th February 1981, the TBR record includes a number of observations of 0.1 mm at precisely 09:00:00. It is assumed that these are manual observations to correct the rain gauge totals to match with check gauge totals following quality checks of the data.

Bochum is a Hellmann raingauge operated and maintained by the German Meteorological Service. It uses a floating pen mechanism to record rainfall on a drum or band recorder with a minimum gauge resolution of 0.01 mm. The duration is 69 years from 1931 to 1999, and the data are aggregated to a minimum temporal resolution of 5 minutes. These sites are selected to represent rainfall in different geographical regions obtained using different measurement techniques. Figure 3 shows the locations of these two gauges.

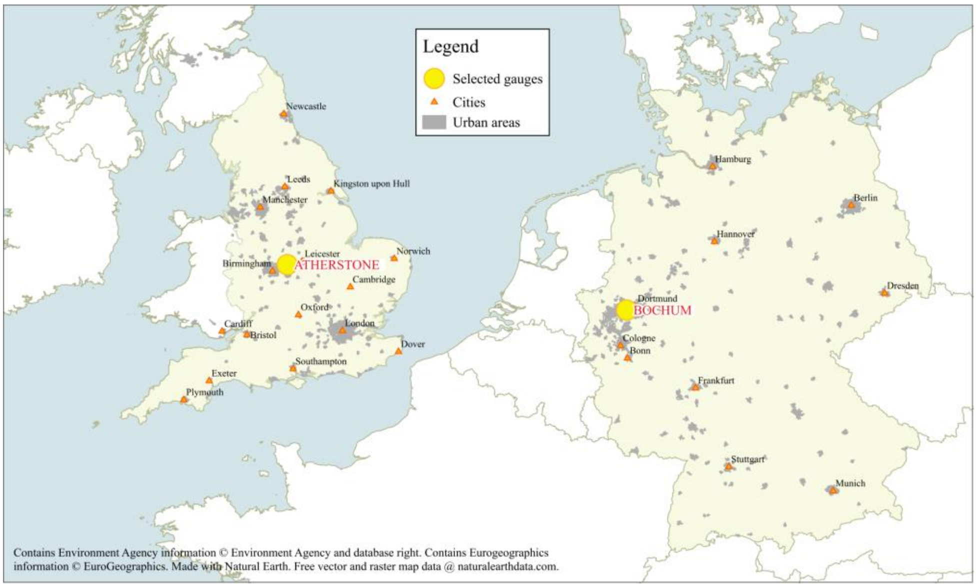

Figure 3 Plan showing the location of the UK and German rain gauges used in this study.

Deleted: 4

Deleted: 3

5.2 Parameter estimation

Model fitting is performed in the R programming environment (R Core Team 2017) using an updated version of the MOMFIT software developed by Chandler et al. (2010) for the UK Government Department for the Environment, Food and Rural Affairs (DEFRA) research and development project FD2105 (Wheater et al. 2007a, 2007b). In this software, parameter estimation is performed using the generalised method of moments (GMM) with weighted least squares objective function: $S(\theta|T) = \sum_{i=1}^{k} \omega_i [t_i - \tau_i(\theta)]^2$. The reader is referred to Wheater et al. (2007b, Appendix A) for a detailed explanation of the fitting methodology.

Deleted: (R Core Team 2016)

Field Code Changed

Deleted: , Lourmas & Jesus

The GMM is preferred for mechanistic rainfall models because the complex dependency structure and marginal distribution of aggregated time-series makes it very difficult to obtain a useful likelihood function (Rodriguez-Iturbe et al. 1988). In this procedure, the difference between observed and expected summary statistics of the rainfall time-series at a range of temporal scales is minimised giving an optimal parameter set θ where: $\boldsymbol{t} = (t_1 \ldots t_k)'$ is a vector of k observed summary statistics, $\boldsymbol{\tau(\theta)} = (\tau_1(\theta)\ldots \tau_k(\theta))'$ is a vector of k expected summary statistics which are functions of $\boldsymbol{\theta} = (\theta_1\ldots\theta_p)'$, i.e. of the vector of p model parameters for which analytical expression are available. The i-th summary statistic is weighted according to the inverse of its observed variance $\omega_i = 1/\mathrm{var}(t_i)$ where t_i is the vector of diagonal elements of the estimated covariance matrix of the mean summary statistics, $\hat{\Sigma}$. While this weighting is not optimal, it provides a reasonable approximation to the optimal weights for the GMM giving robust estimation of the parameter standard errors (R. Chandler et al. 2010). Other weights can be applied allowing the user to influence the dominance of specific rainfall properties, although for unbiased estimates of the summary statistics the weights must be independent of the model parameters and the data (Wheater et al. 2007b).

Typically, the vector of observed summary statistics T comprises the mean, variance, auto-correlation and proportion of dry periods for temporal scales between 1 and 24 hours. Prior to model fitting and to allow for seasonality, summary statistics are calculated for each month over the record length and pooled between months. For each month, the pooled statistics are used to estimate the covariance matrix of model parameters required for parameter uncertainty estimation, and the mean of the monthly statistics. Therefore 12 parameter sets are obtained for the whole year.

Model parameters are estimated using two minimisation routines. First, Nelder-Mead optimisations are performed on random perturbations around user-supplied parameter values to identify promising regions of the parameter space. Following a series of heuristics to identify the best performing parameter set, random perturbations around these values are used as new starting points for subsequent Newton type optimisations. The parameter set with the lowest objective function is the best performing and selected for that month. Following the approach employed by Kaczmarska (2013) to obtain smoothly changing parameters throughout the year, this two-step optimisation is only applied to one month. Subsequent parameter estimation is based on a single Newton type optimisation using the previous month's estimate as the starting point. Testing of this approach has shown that when the parameters are well identified the same seasonal variation is achieved regardless of the starting month. The sampling distribution of the estimators resulting from the GMM minimisation routine are approximately multivariate normal (MVN). The optimal parameter set is estimated by the mean of this distribution, and the covariance matrix is estimated from the Hessian of the least squares objective function S (Wheater et al. 2007b). The MVN distribution of model parameter estimators is used to estimate 95% confidence intervals for the parameter estimates. On occasions that the model parameters are poorly identified, it may not be possible to calculate the Hessian of the objective function preventing the estimation of parameter uncertainty.

5.3 Experimental design

Initial experiments with the coefficient of skewness and proportion of dry periods included in model fitting for censored data were limited by the inability to obtain well identified parameters for some or all months. While good model fits were obtained for some low censors, extreme value estimation continued to be underestimated. On the basis that censoring is a new approach

Formatted: Subscript

Deleted: which has been shown to provide

Deleted: (Chandler et al. 2010)

Deleted: GMM

Deleted: parameter estimates are

Deleted: ly

Deleted: distributed where t

Deleted: computed by the optimisation routine

Deleted: on

Formatted: Font: Not Italic

Deleted: If parameter uncertainty is not estimated in model fitting it is indicative that the parameters are poorly identified

Deleted: understated

to enhance the estimation of rainfall extremes, skewness is not considered to be an important fitting statistic for censored simulations. Furthermore, because censored models cannot be used to generate continuous time-series of the sort which may be used for hydrological modelling, the proportion of dry periods is also considered to be unimportant for censoring. Consequently for censored model calibration, the choice of fitting statistics is reduced to the 1 hour mean, the coefficient of variation and lag-1 autocorrelation of the rainfall depths at the censor resolution, and the 6 and 24 hour resolutions. Again, to ensure well identified model parameters for the Atherstone dataset, it was necessary to extent the choice of fitting statistics to include the 1 hour statistics for 5 minute simulations. This was neither necessary for 15 and 60 minute simulations at Atherstone, nor the Bochum dataset.

For all simulations the fitting window is widened to 3 months, hence for any given month the models are fitted to data for that month, together with the preceding and following months. This approach is used to increase the data available for fitting the models when censoring on the basis that censoring removes data which would otherwise be used in fitting. Tests have shown that widening the fitting window from one to three months has the effect of smoothing the seasonal variation in model parameters and improving parameter identifiability. There is also negligible impact on the estimation of summary statistics and extremes under the model parameters.

For the two randomized models, BL1 and BL1M, the Gamma shape parameter α is constrained to a fixed value in calibration and simulation. The Gamma shape parameter α is an insensitive model parameter and can take any value within a very large range without significant impact on the estimation of summary statistics or extremes (see Appendix A). For the BL1 model, parameterization without an upper bound on α often results in poor identifiability with parameter estimates in the thousands to tens of thousands. For the BL1M model, α is typically better identified than for BL1 with a tendency to move towards the lower boundary. In order to avoid having infinite skewness, α must be greater than 4 for the BL1 model and 1 for the BL1M model (see Kaczmarska et al. (2014) and references therein for discussion on these criteria). Therefore, by fixing α at 100 for the BL1 model and 5 for the BL1M model, the number of parameters to be identified for these models is reduced by one. All models are fitted using the exponential distribution for mean cell depth. This further reduces the number of model parameters to be fitted for both uncensored and censored models, therefore in all cases the ratio of standard deviation to the mean cell depth ($r = \sigma_x/\mu_x$) is fixed at 1. Fitted model parameters are presented in Appendix B for 5 and 15 minute rainfall at both sites for uncensored and censored rainfall using censors selected in Sect. 6.2 (Table 2).

Deleted: , Isham & Onof

6. Results

6.1 Extreme value estimation

Rainfall extremes are estimated from the models by sampling annual maxima directly from simulations. For each model fitted to uncensored data, 100 realisations of 100 years duration are simulated using parameters randomly sampled from the multivariate normal (MVN) distribution of model parameters. This allows model parameter uncertainty to be represented in the spread of the MVN extreme value estimates (hereafter referred to as MVN realisations), covering the full range of observations. Extreme value estimation up to the 1000-year return level is also provided to indicate the potential magnitude of rarer events. For this extrapolation, extremes are estimated from one realisation using the mean of the MVN distribution of parameter estimators (hereafter referred to as the optimal estimates). To ensure stability of the extreme value estimates up to approximately the 1000-year return level, simulations have been extended to 10,000 years.

Deleted: rainfall

Deleted: also

Deleted: of 10,000 years duration simulated

Deleted: parameter

Deleted: By extending this simulation to 10,000 years duration, extreme value estimation up to approximately the 0.001 AEP (1000 year return level) may reasonably be expected to be stable.

Extreme value estimation for the censored calibrations is shown in Figs. 4, 5 and 6 for 5, 15 and 60 minutes temporal resolutions respectively. The top three plots in each figure show the results for Bochum, and the bottom three plots the results for Atherstone, with observed and simulated annual maxima plotted using the Gringorten plotting positions. All plots show the equivalent extreme value estimates obtained without censoring obtained by simulating one realisation of 10,000 years duration with the optimal parameter set. Upper limits on censoring were identified when model parameterization noticeably deteriorated resulting in the mean of the MVN realisations to deviate away from the optimal. Results presented are limited to the 4 highest censors with well identified model parameters, together with 95% simulation bands. The simulation bands show the range of extreme value estimation between the 2.5 and 97.5 quantiles of the 100 MVN realisations for each simulated data point.

All censored models have significantly improved the estimation of extremes at each site and scale with very good estimation by all three model variants particularly at the 5 and 15 minute scales. At these scales, the estimation of extremes with the 4 censors presented has approximately converged on the observations. At the 60 minute scale there is notable improvement in the estimation of extremes with some convergence in estimation with increasing censors, although there is continued underestimation of the observed. The 95% confidence intervals by all censored models broadly bracket the observations and are largely unvaried with increasing censors, other than with the BL1M model at the 60 minute resolution.

At the 5 minute scale, estimation has converged on the observations with censors between 0.5 and 0.65 mm at Bochum, and between 0.6 and 0.75 mm at Atherstone. For all three models there is slight underestimation of extremes higher than approximately the 10 year return period, although the BL1M model accurately estimates the highest observed extreme at both sites. At the 15 minute scale, convergence at Bochum has occurred for censors between 1.0 mm and 1.3 mm, while at Atherstone convergence has occurred for censors between 0.6 mm and 0.9 mm. As for the 5 minute resolution models, the BLIM model appears to perform slightly better than the BL0 and BL1 models, resulting in improved estimation of the highest observed extremes and elevated estimates of the 1000 year return period rainfall at both sites. At the 60 minute resolution, there is good convergence in estimation for all three models at Bochum, and the BL1M model at Atherstone. However, extreme value estimation with the BL0 and BL1 models at Atherstone is more widely spread across the applied censors. For the BL0 and BL1 models, the 0.2 mm censor results in much lower estimates than the three higher censors, although the mean of the MVN realisations for the 0.6 and 0.8 mm censors are starting to deviate away from the optimum realisation. For the BL1M model, there is good convergence between the optimal realisations with each censor, although the mean of the MVN estimates for the 0.6 and 0.8 mm censors have significantly deviated from the optimum.

Deleted: 0.1 annual exceedance probability (AEP)

Deleted: 0.001 AEP

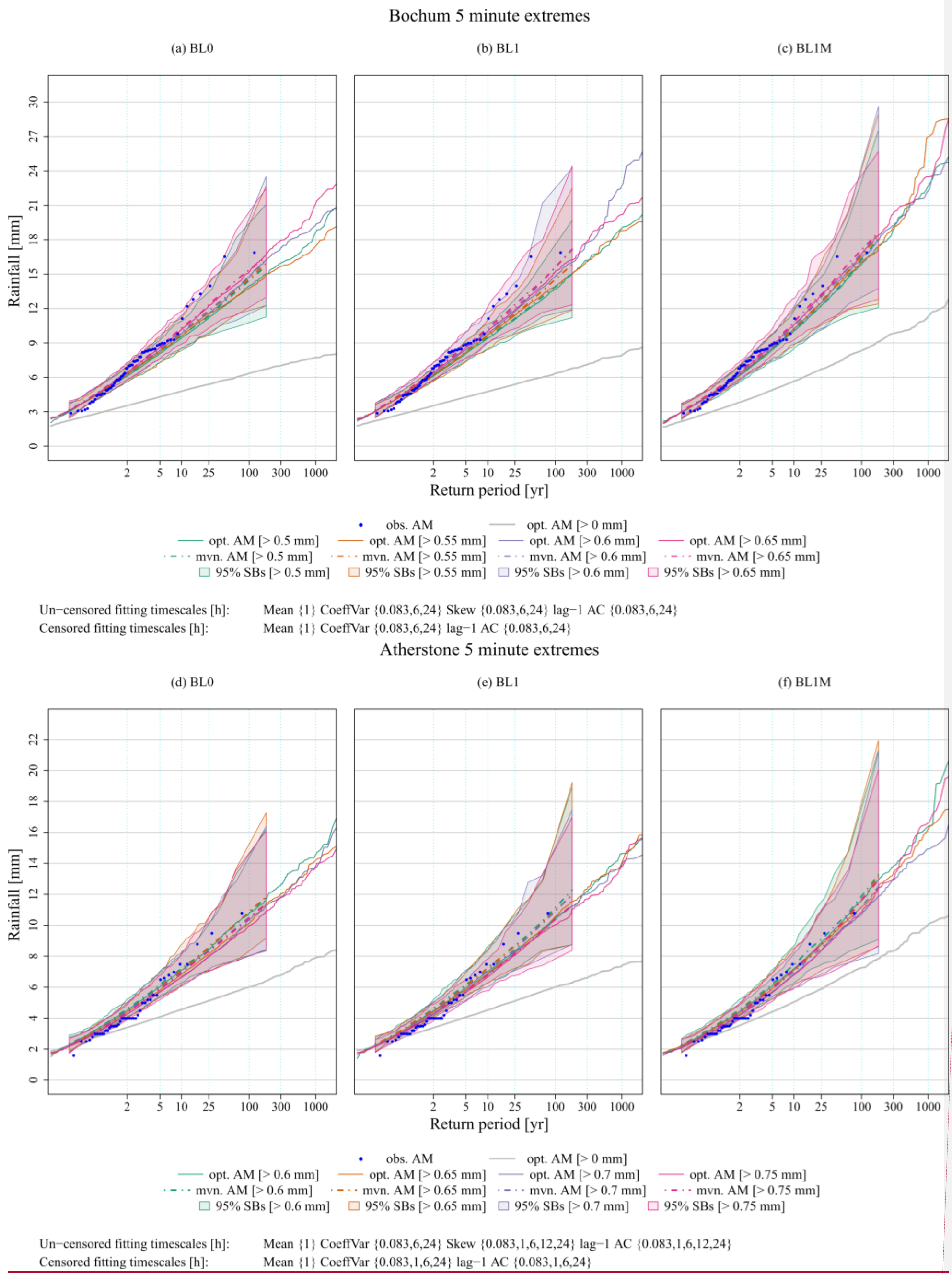

Figure 4 Extreme value estimation at 5 minute resolution. Optimal realisations (opt. AM) are shown with solid lines and the mean of the MVN realisation (mvn. AM) are shown with dashed lines.

Deleted: ¶

Deleted: 5

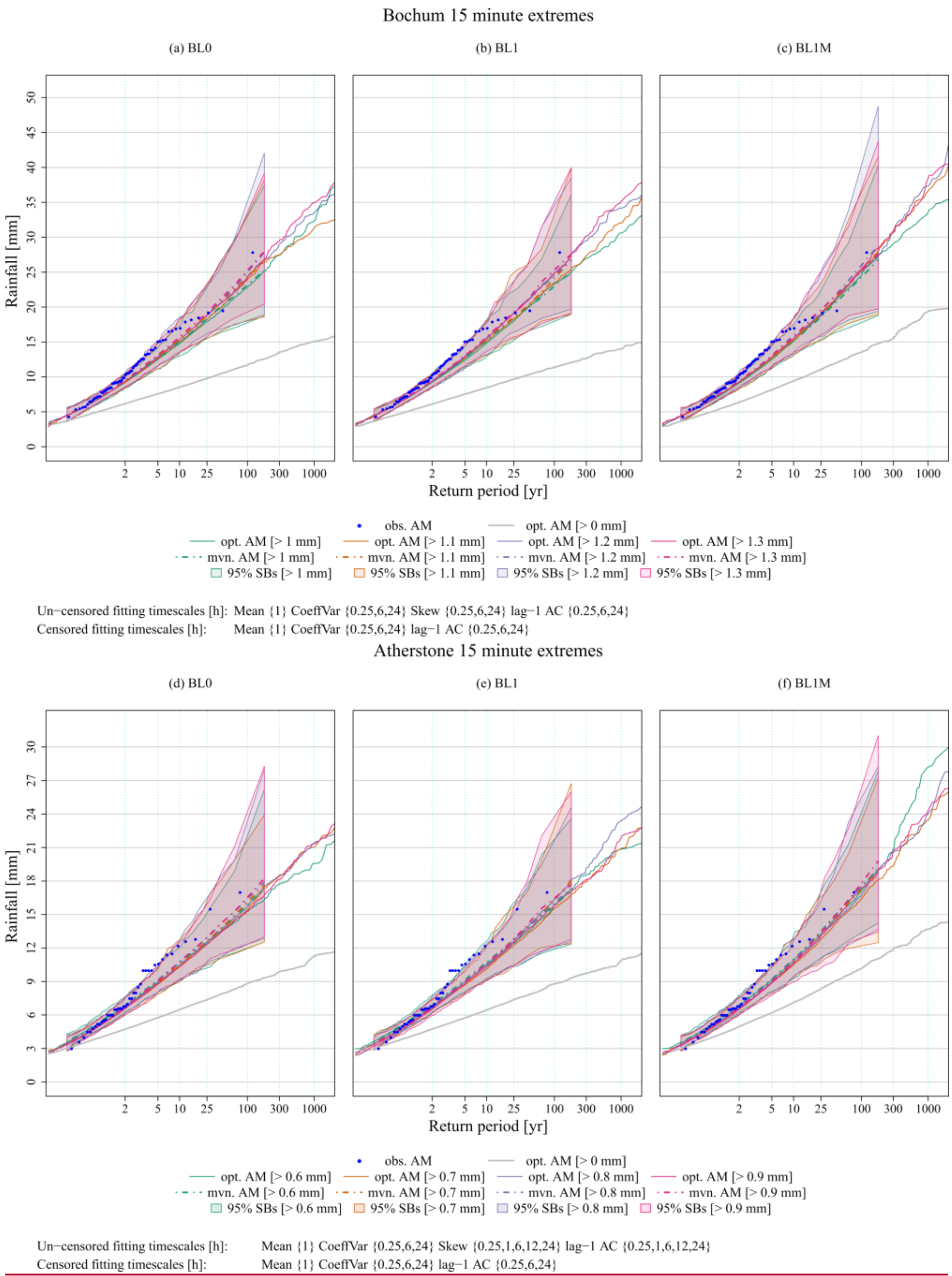

Figure 5. Extreme value estimation at 15 minute resolution. Optimal realisations (opt. AM) are shown with solid lines and the mean of the MVN realisation (mvn. AM) are shown with dashed lines.

Deleted: 6

Deleted: 5

Deleted: ¶

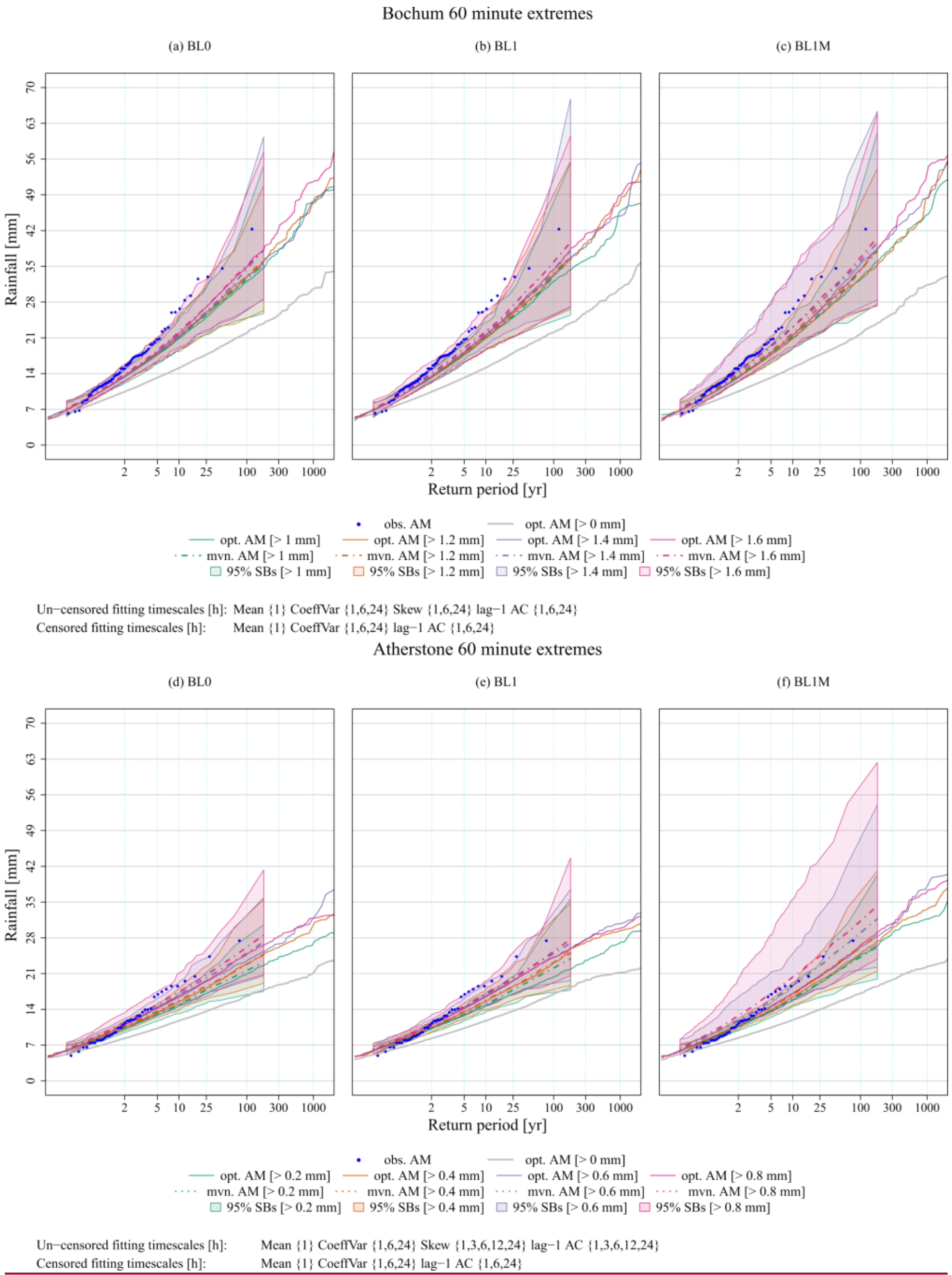

Figure 6 Extreme value estimation at 60 minute resolution. Optimal realisations (opt. AM) are shown with solid lines and the mean of the MVN realisation (mvn. AM) are shown with dashed lines.

Deleted: 7

Deleted: 6

The mean of the MVN realisations for the BL1M model at Atherstone with the 0.6 and 0.8 mm censors (see Fig.6) diverges from the optimum because of the generation of unrealistic extremes. This divergence is also observable in the larger spread of 95% simulation bands over 100 realisations. While it has been possible to fit the model, Fig. 7 shows that as censoring has increased to 0.8 mm, confidence intervals on model parameters have widened for several months of the year, notably January, February and June. When sampling from the MVN distribution in simulation, these large confidence intervals mean that there is a high chance of sampling parameters which deviate significantly from the mean of the distribution thereby giving rise to a wide spread in extreme value estimates. These large confidence intervals indicate that the confidence in parameter estimation is reducing with higher censors and consequently the model error is too large for the reliable simulation of extremes.

Formatted: Font: Not Italic

Deleted: confidence intervals

Deleted: s

Deleted: are poorly identified

Deleted: therefore

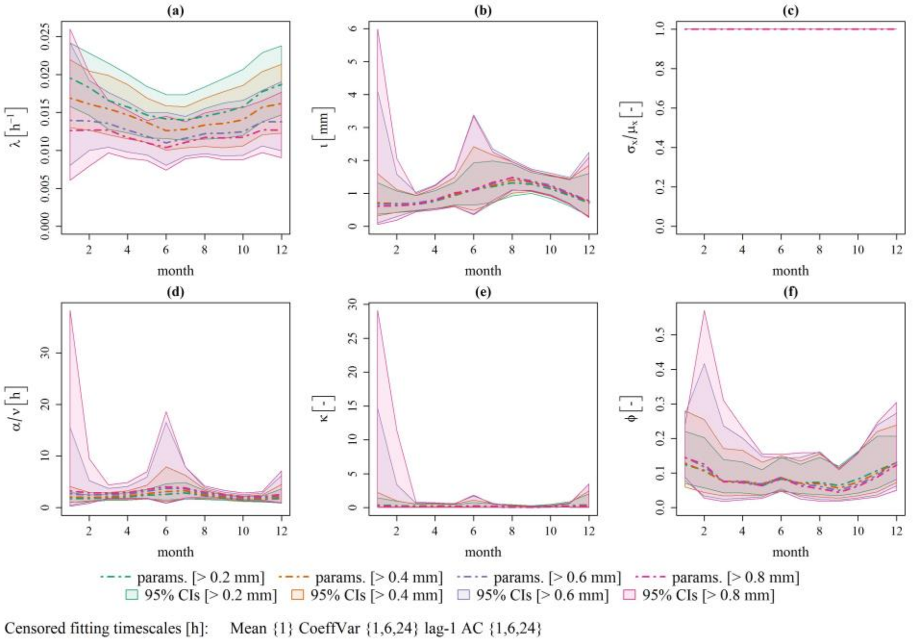

Figure 7 Comparison of censored BL1M model parameters for Atherstone 60 minute data. Optimal parameter estimates (params.) are shown with dot-dashed lines, and parameter uncertainty is represented with 95% confidence intervals.

Deleted: 8

Deleted: 7

6.2 Validation

The rainfall extremes presented in Sect. 6.1 have been generated mechanistically using model parameters derived from central moments of the censored rainfall time-series. While censored models cannot be used to simulate the whole rainfall hyetograph, it is important to ensure that the process by which the extremes are estimated is reliable. Therefore, model performance is validated in the usual way for this class of model by comparing the analytical summary statistics under the model parameters with the observations – here the observations are censored. The lowest censors presented in Figs. 4, 5 and 6 are selected for validation. No distinction is made between models in this choice, although it is recognised that there is some variation in the extreme value performance of specific censors between model types. See Table 2 for censor selection at each site and scale.

Deleted: Given the spread of estimates at Atherstone for the 60 minute resolution, validation is based on the 0.6 mm censor instead of the lowest which would be 0.2 mm.

Table 2 Censor selection for model validation.

	5 minutes	15 minutes	60 minutes
Bochum	0.5 mm	1.0 mm	1.0 mm
Atherstone	0.6 mm	0.6 mm	0.2 mm

Deleted: 6

6.2.1 Replication of fitting statistics

Deleted: ¶

Formatted: Normal, Line spacing: single, No bullets or numbering

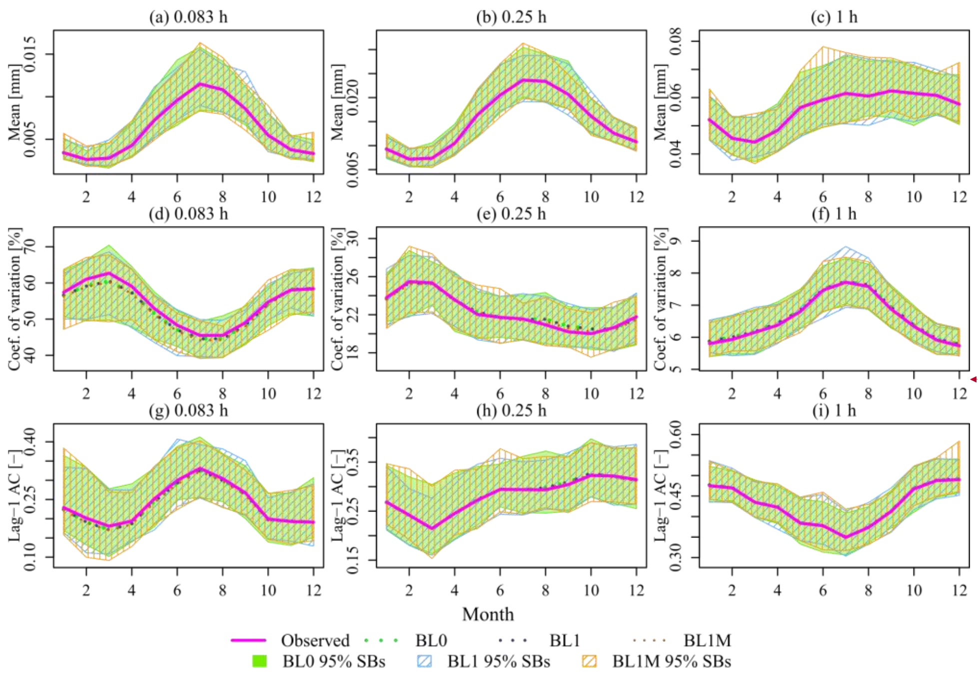

Figure 8 Seasonal variation in mean, coefficient of variation and lag-1 autocorrelation for selected censors at Atherstone, observed vs. estimated.

Deleted: 9

Deleted: 8

Figure 8 shows the seasonal variation in mean, coefficient of variation and lag-1 autocorrelation for all three models at Atherstone with the selected censors in Table 2. Comparable performance is achieve with the models for Bochum and hence these results are not presented. The plots show the estimated summary statistics calculated using the optimum parameter estimates, together with 95% simulation bands obtained by randomly sampling 100 parameter sets from the multivariate normal distribution of model parameters, estimating the summary statistics under the model and displaying the range of estimates between the 2.5 and 97.5 quantiles. Because models are fitted over 3 monthly moving windows, estimated summary statistics are compared with summary statistics for censored observations for the same periods. Fitting statistics for the 6 and

Deleted:

Deleted: confidence intervals

Deleted: for

24 hour scales are not shown. The limits on the vertical Y axes are optimized at each site and scale, therefore the reader is advised to pay careful attention to the scales when comparing summary statistics.

All models perform very well with respect to replicating the summary statistics used in fitting with the 95% simulation bands comfortably bracketing the observations. The estimated summary statistics are very close to the observed with all models performing equally well. The seasonal variation in mean monthly rainfall varies between scales because there is a higher proportion of low observations at short temporal scales removed by the censors. The greater prominence in seasonal variation shown in plots a and b indicates that the summer months (approx. Apr - Oct) are more prone to short intense bursts of rain, and the winter months longer periods of low rainfall intensity. This is consistent with there being more convective rainfall in the summer, and stratiform rainfall in the winter. The plots in Fig. 8 demonstrate that the models are able to reproduce the censored fitting statistics, confirming reliability of the process.

6.2.2 Replication of statistics not used in fitting

A consequency of censoring is that it truncates the thin tail of the rainfall amounts distribution which significantly changes it's shape. Because this truncation is not replicated in the analytical equations of the models used in this study, the models are not expected to be able to reproduce this statistic well. Therefore this statistic is excluded from validation. Conversely, censoring is not expected to significantly impact the ability of the models to estimate the proportion of wet periods. Despite this, censoring significantly changes this statistic at fine temporal scales. Figure 9 shows the seasonal variation in the proportion of wet periods for all three models at both sites with the selected censors in Table 2.

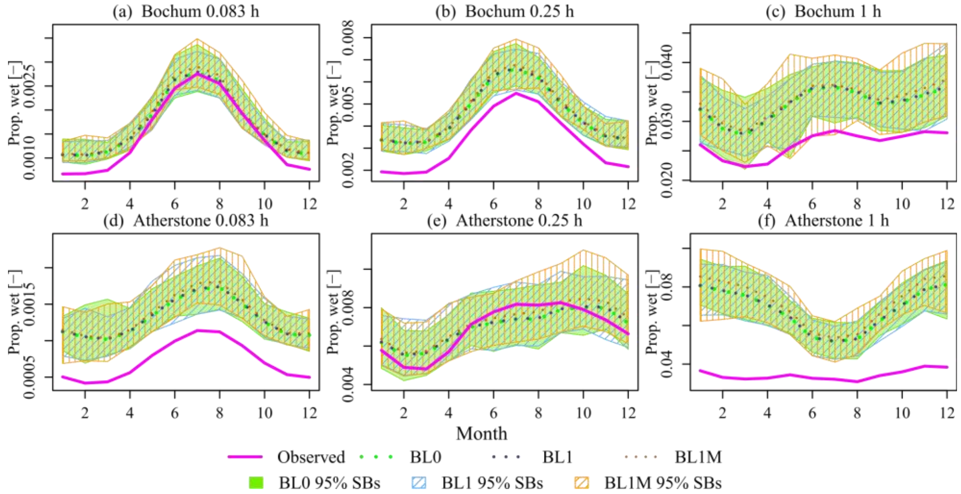

Figure 9 Seasonal variation in skewness coefficient and proportion of wet periods for selected censors, observed vs. estimated.

Deleted: coefficient of skewness and

Deleted: Both of these statistics were excluded from model fitting for censored simulations, although are generally considered to be important fitting statistics.

Deleted: 10

Deleted: 9

The ability of the models to reproduce the proportion of wet periods is generally good, although there is a tendency for all models to overestimate this statistic at both sites. At the 5 minute resolution for Bochum, the 95% simulation bands comfortably bracket the observations between the months of May and October, although there is over-estimation in the other months and for all months at the 15 and 60 minute scales. At Atherstone, there is good representation of the proportion of wet periods at the 15 minute scale, although over-estimation at the 5 and 60 minute scales. Generally, there is very slightly better agreement in the summer months which, as highlighted in Sect. 6.2.1, may be more prone to short intense downpours at fine temporal scales. This suggests that the censored models may be more effective at simulating the heavier short duration rainfall characteristic of summer convective storms, than the longer duration low intensity rainfall characteristic of winter storms.

Deleted: All the models over-estimate the skewness coefficient with observations for all months falling outside the 95% confidence intervals. This is not an unexpected result given that censoring effectively truncates the thin tail of the rainfall amounts distribution which will significantly change its shape. Because this truncation is not replicated in the analytical equations the models are unlikely to be able to replicate this behaviour using the observations provided. ¶

Deleted: better than their ability to reproduce skewness

Deleted: confidence intervals

Deleted: and 60

Deleted: s

7. Discussion on censor selection

The censors selected for validation in Table 2 were chosen based on their extreme value performance. For the estimation of extremes at other locations, it would be preferable to have a set of heuristics to guide censor selection. The following discussion of extreme value estimation performed in this study is intended to guide practitioners in the application of censored modelling.

7.1 Stability of confidence intervals

Upper limits on censoring were identified where model parameters were either poorly identified or the mean of the MVN realisations deviated significantly from the observations. The onset of this effect was observed in Fig. 6 for estimation of hourly extremes at Atherstone with the BL1M model. Figure 10 shows the change in 95% confidence intervals and the mean of the MVN realisations obtained with censored models with well and poorly identified parameters for 15 minute data at Bochum and Atherstone. The comparison is made between extremes for the selected censors given in Table 2 (1.0 mm and 0.6 mm respectively) and extremes from higher censors (1.5 mm and 1.0 mm respectively).

Deleted: 1

Deleted: behavioural and non-behavioural

Confidence intervals on extreme value estimates for Bochum 15 minute rainfall obtained with censors from 1.0–1.3 mm, and for Atherstone with censors from 0.6–0.9 mm (Fig. 5), are broadly stable and unchanging. This is indicative that parameterization across each model variant and censor is good enabling robust estimation of extremes. As the censor at Bochum is increased to 1.5 mm (Fig. 10, panels a–c), there is a noticeable increase in the upper confidence bound and the mean of the MVN realisations has diverged leading to over-estimation of the extremes. Increasing the censor at Atherstone to 1.0 mm has resulted in very significant widening of the confidence intervals and divergence of the mean of the MVN realisation (Fig. 10, panels d–f). In each case, this divergence results from the generation of unrealistic extreme value realisations which are shown in Fig. 10 (light grey lines).

Deleted: 1

Deleted: 1

Deleted: 1

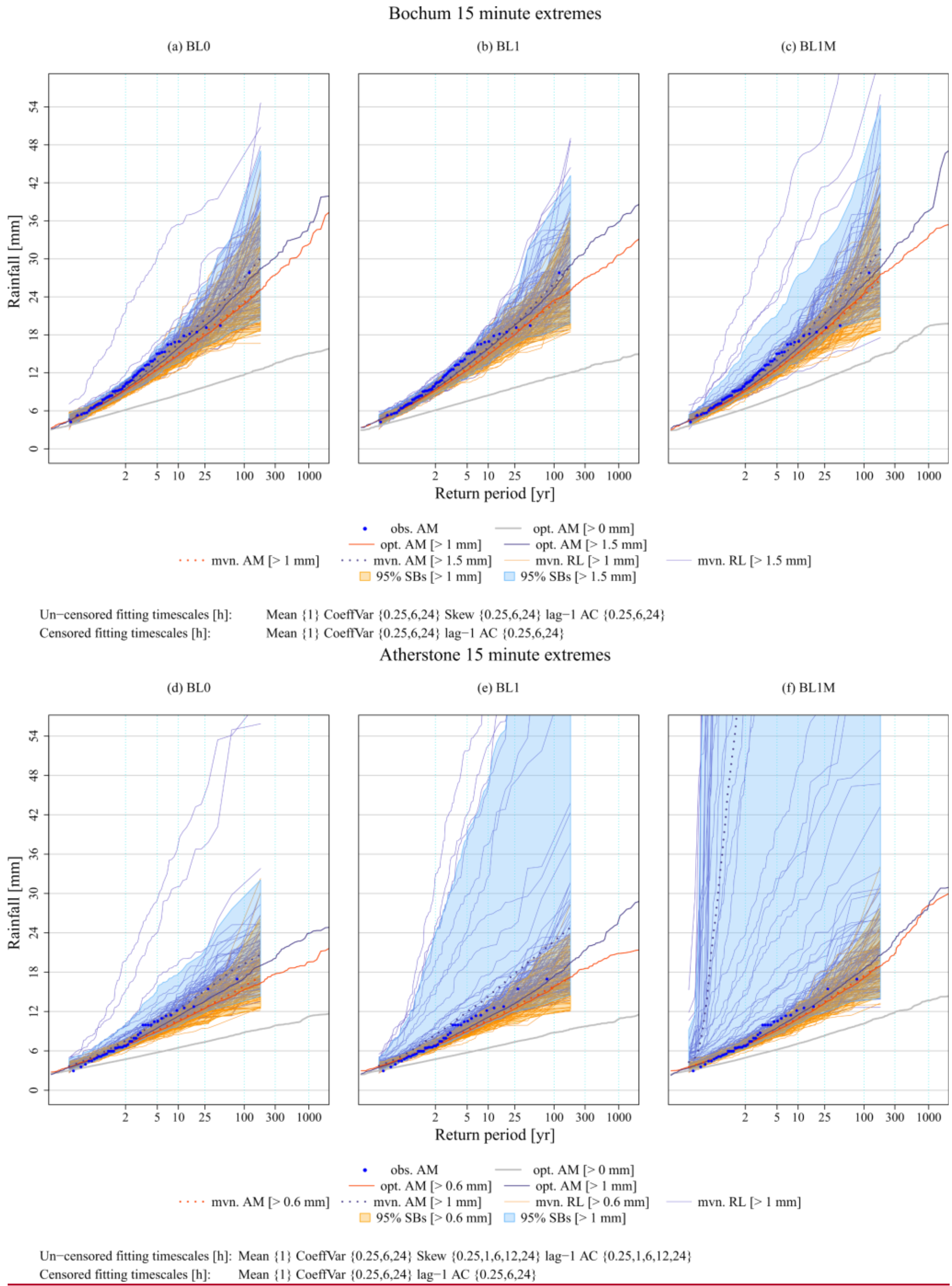

Figure 10 Change in 95% confidence intervals and mean of the MVN realisations for Bochum and Atherstone 15 minute data with well identified (> 1.0 mm and > 0.6 mm) and poorly identified (> 1.5 mm and > 1.0 mm) censored model parameters.

Deleted: 11

Deleted: Figure Figure 1411

Deleted: behavioural

Deleted: non-behavioural

Deleted: censors

While it has been possible to fit models to data with these high censors, examination of the parameter estimates and associated uncertainty reveals that parameter identifiability is reducing. Figure 11 shows the seasonal variation in estimates for the BL1M model parameters α/ν, κ and φ fitted to Bochum 15 minute data with a 1.5 mm censor. Parameters λ and ι are well identified with tight confidence brackets around the optimum, while r and α are fixed, therefore these parameters are not shown. Confidence intervals on α/ν, κ and φ are very large in the winter months indicating that identifiability of these parameters has deteriorated. When sampling from the multivariate normal distribution for model parameters in simulation, these large uncertainties give rise to poor extreme value estimation. The same behaviour was observed for the BL1M model at Atherstone for 60 minute data as shown in Fig. 7.

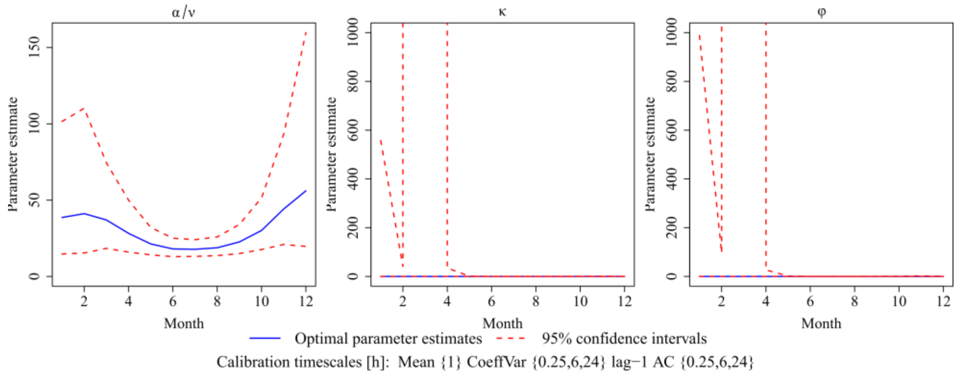

Figure 11 Fitted model parameters for the BL1M model with 1.5 mm censor applied to Bochum 15 minute data.

With the upper bound on censoring identified, the obvious question is how to identify a lower bound? The results presented in Figs. 4, 5 and 6 suggest that there is convergence in the estimation of extremes with increasing censors. If so, when is the onset of convergence? Figure 12 shows the change in extreme value estimation with censor for 15 minute rainfall at Bochum (top plots) and Atherstone (bottom plots) for 10 and 25 year return periods.

At both locations, divergence in the mean of the MVN realisations and confidence intervals shown in Fig. 12 is easily identified with the very large box-plot whiskers at 1.5 mm and 1.0 mm censors for Bochum and Atherstone respectively. The plots for Bochum also show a large spread in the extreme realisations with a 1.4 mm censor for the BL1M model suggesting that parameter identifiability is deteriorating at this censor.

At Atherstone, there is clear evidence of convergence in estimation between censors 0.5–0.9 mm. However, convergence is less obvious at Bochum. At Bochum, there is continual improvement in extreme value estimation with the increasing censors, although there is a perceptible reduction in improvement with each successive increase in censor. For censors of 0.7 mm and above, all model realisations bracket the observed extremes (horizontal dashed blue line), which is also true for censors above 0.5 mm at Atherstone. Therefore, ranges may be identified at both sites for censors which may be considered to give satisfactory estimation of extremes: 0.7–1.3 mm at Bochum and 0.5–0.9 mm at Atherstone.

Deleted: 2

Deleted: non-behavioural

Deleted: Figure 12

Deleted: 1211

Deleted: 3

Deleted: 0.1

Deleted:

Deleted: 0.04 AEPs

Deleted: 3

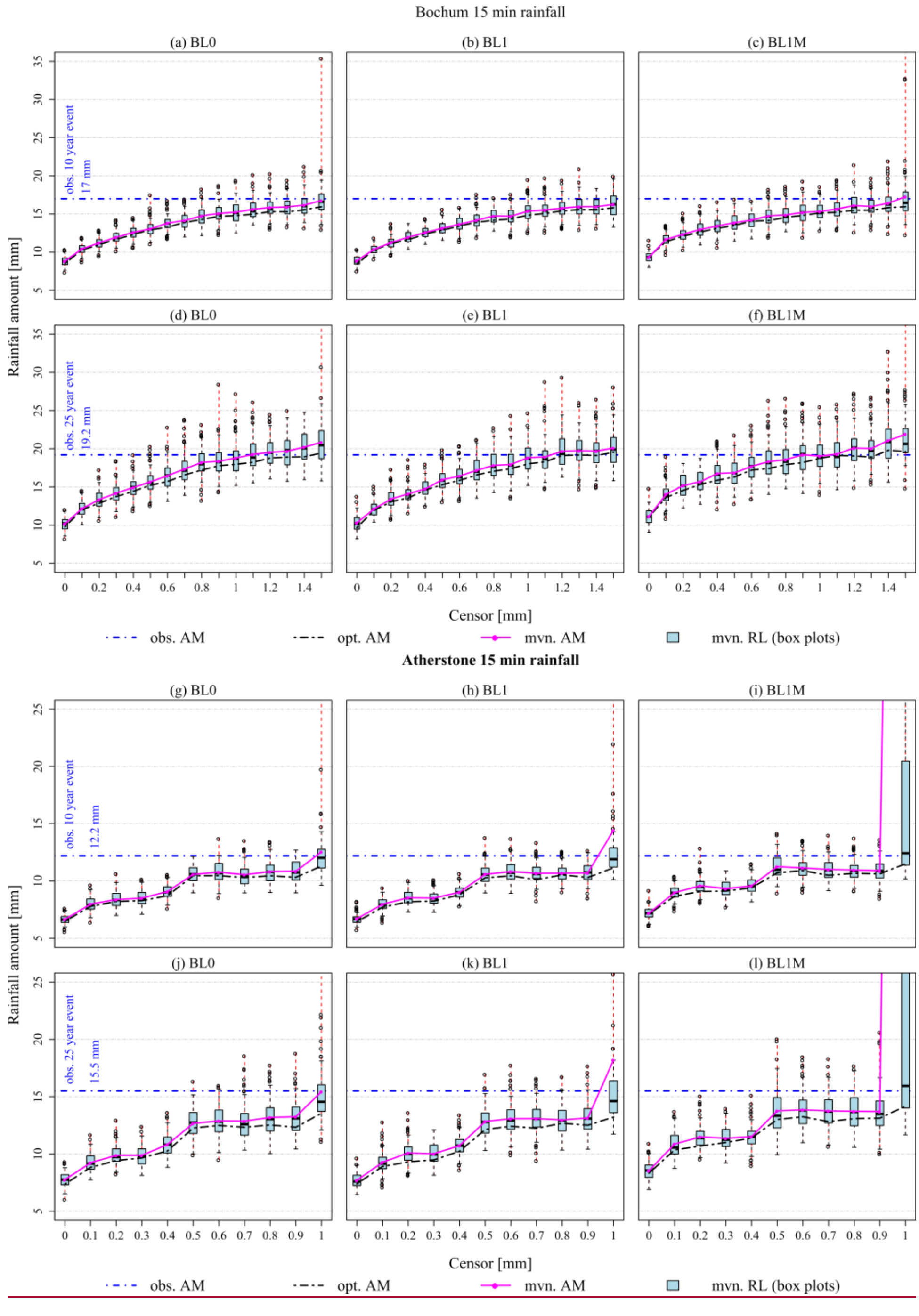

Figure 12 Variation in extreme value estimation with censor for 15 minute data at Bochum and Atherstone for two annual return periods: 10 and 25 years.

Deleted: ¶

Deleted: 13

Deleted: Figure 13

Deleted: exceedance probabilities

Deleted: 0.1

Deleted: 0.04

7.2 How much rainfall to censor?

In Sect. 7.1 we identify plausible censor ranges on the basis parameter stability and convergence of extreme value estimation. However, this doesn't address the question of how much rainfall to censor? Because extremes are generated mechanistically, we want to simulate the storm event hyetograph therefore it is in our interest to keep the censor low in relation to the rainfall depth profile. The most basic check is that the minimum observed extreme (here designated as the smallest annual maxima) is greater than the censor being used. This is true for all the sites and scales investigated in this study, with the lowest observed annual maxima of 1.6 mm occurring at the 5 minute scale in Atherstone. However, this significantly exceeds the maximum censor applied to 5 minute data at Atherstone, 0.75 mm (see Fig. 4), therefore it's unlikely that a well parameterized model would be achieved.

Figure 13 shows the empirical cumulative distribution function (ECDF) plots for the above zero rainfall records at Bochum and Atherstone aggregated to 5 and 15 minute resolutions. All the censors used for the estimation of fine–scale extremes in Figs. 4, 5 and 6 are shown, with the top three censors highlighted magenta. The censors selected for model validation (Table 2) are highlighted blue, and the lower limits on censors identified in Sect. 7.1 for 15 minute rainfall are shown and highlighted green. The ECDF plots are truncated at the 99th percentile to aid comparison of the applied censors, therefore the maximum rainfall is highlighted in red text on the right of each plot. For all censors, their rainfall quantile values are shown with the colour matching the plotted lines.

It can be seen from Fig. 13 that a substantial proportion of the above zero rainfall record is masked from the models with censoring. At the 5 minute scale, the selected censor of 0.5 and 0.6 mm removes in excess of 98% and 96% of the above zero rainfall from Bochum and Atherstone respectively. At the 15 minute scale, the selected censors of 1.0 and 0.6 mm remove in excess of 96% and 81% respectively. These quantiles are high and support the hypothesis that mechanistic models may be poor at estimating fine–scale extremes because the training data are dominated by low observations.

A striking difference in the ECDF plots for the two locations is the smoothness of the curves. The stepped nature of the Atherstone plots is very pronounced and reflects the resolution of the gauge: 0.5 mm between 1982 and 2003, and 0.2 mm before and after these dates. The stepped nature of the plots at Atherstone highlights that the selected censor quantiles (blue) are just greater than the 0.5 mm quantiles. We also know from Fig. 12 that a censor of 0.5 mm for 15 minute rainfall at Atherstone would give very similar extreme value estimation to the selected 0.6 mm censor (highlighted in green on the ECDF plot, Fig. 13). This implies that to improve the estimation of fine–scale extremes at Atherstone, it has been necessary to remove all observations which correspond with the gauge resolution.

Deleted: 4

Deleted: For the 15 minute plots,

Deleted: their quantiles listed separately in Table 3. ¶

Deleted: 4

Deleted: 3

Deleted: 4

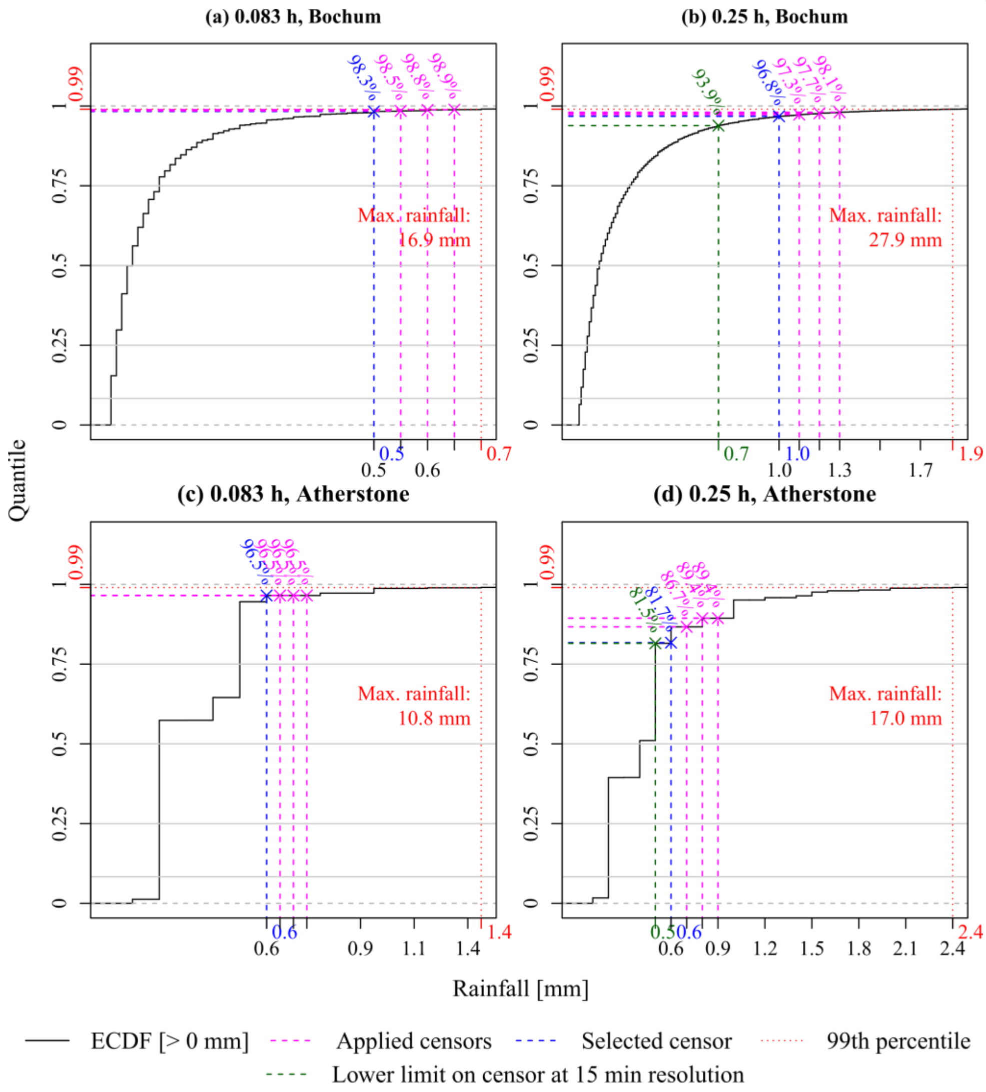

Figure 13 Empirical Cumulative Distribution Function plots for Bochum and Atherstone rainfall aggregated to 5 and 15 minute temporal resolutions. The plots are limited to the 99^{th} percentile rainfall and show the rainfall quantiles corresponding to the optimum censors used in the estimation of extremes in Figs. 4, 5, and 6.

While the proportion of rainfall observations removed prior to model fitting is large - over 90% and 80% for 5 and 15 minute rainfall from Bochum and Atherstone respectively - comparison with the maximum rainfall amounts and an assessment of the number of independent peaks over the censor demonstrate that the censors are low. Table 3 shows the proportion of maximum rainfall and the number of independent peaks per year for the selected censors given in Table. 2. The number of peaks over the censors are estimated using a temporal separation of 48 hours to define independence.

Deleted: 14

Deleted: Figure 14

Deleted: (see Table 5)

Deleted: ximim

Table 3 Proportion of maximum rainfall and number of independent peaks per year for the selected censors given in Table 2.

	Scale [mins]	Bochum	Atherstone
Proportion of maximum rainfall	5	3.0%	5.7%
	15	3.6%	3.5%
Number of independent peaks / year	5	53	27
	15	46	65

The proportion of the maximum observed rainfall is less than 6% in all cases which is very low considering that the maximum recorded rainfall across both sites and scales is just 27.9 mm for Bochum 15 minute rainfall. For a standard Peaks over Threshold extreme value analysis, the threshold is typically set so that between 3 and 5 independent peaks per year remain in the partial duration series. Using a temporal separation of 48 hours to define independence, the number of peaks per year retained after censoring is between 27 and 65 (Table 3). The actual number of peaks retained for fitting the Bartlett-Lewis models is much greater than this because serial dependence in the rainfall time-series is simulated with mechanistic modelling. While it is possible to estimate return levels for serially dependent extremes using extreme value theory, the analysis set out in Fawcett and Walshaw (2012) demonstrate that estimating the extremal index is non-trivial can be subjective.

8. Further discussion and conclusions

The estimation of rainfall extremes presented in this study using censored rainfall simulation is highly promising and offers an alternative to frequency techniques. The estimation of extremes at sub-hourly scales has far exceeded expectation with all three models giving a very high level of accuracy across a range of censors. However, censoring uses rainfall models in a way they were never previously intended. Rainfall models have invariably been used for simulation of long duration time-series across a range of scales for input into hydrological and hydrodynamic models. Censored rainfall synthesis cannot be used in this way because only the heavy portion of the hyetograph is simulated.

The success of this research is to broaden the scope of mechanistic rainfall modelling and ask new questions of it. Mechanistic models and related weather generators are very powerful at simulating key summary statistics for a range of environmental variables. An area where these models have consistently underperformed is the estimation of fine–scale extremes. Efforts to improve extreme value estimation at fine temporal scale have focussed on structural developments. But those developments have always been undertaken in the context of rainfall time-series generation. Continued underestimation at fine temporal scales has given rise to the notion that rectangular pulse models are potentially *"unsuitable for fine–scale data"* (Kaczmarska et al. 2014, p.1985).

For effective scenario planning with hydrological models, good reproduction of rainfall time-series is necessary, with accurate estimation of key summary statistics. However, for assessment of extremes and estimation of storm profiles, good replication of rainfall central moments is arguably less important. The ability of the censored models to adequately reproduce the central moments used in calibration was checked to ensure that the process by which the extremes are constructed is reliable. Because

Deleted: Quantile of the above zero rainfall amounts,
Deleted: p
Formatted Table
Deleted: Quantile of rainfall amounts
Deleted: 5
Deleted: 98.3%
Deleted: 96.5%
Deleted:
Deleted: 15
Deleted: 96.8%
Deleted: 81.7%
Deleted: Proportion of maximum rainfall
Deleted: 5
Deleted: 3.0%
Deleted: 5.7%
Deleted: 15
Deleted: 3.6%
Deleted: 3.5%
Deleted: Number of independent peaks / year
Deleted: there is no requirement to meet the independence criteria with
Deleted:

rainfall over the censor is by definition coincident with rainfall below the censor, the censored models can be used to estimate uncensored extremes by simply restoring the censor to the estimates.

Extreme rainfall estimation with censoring across all models, scales and sites is significantly improved on that without censoring as shown in Figs. 4, 5 and 6. Up to approximately the 25 year return period, estimation is broadly equivalent across all models. For rarer events, the BL1M model appears to perform better than the other two at the 5 and 15 minute scales at Bochum and Atherstone by accurately estimating the highest observations at those scale. This improvement over the BL0 and BL1 models is significant in the event that extreme rainfall estimation is required beyond the range of observations. This is demonstrated in all 4 cases (5 and 15 minute scales at Bochum and Atherstone) with the higher estimation of extremes at the 1000 year return level by the BL1M model compared with the other two. Below approximately the 25 year return period the differences in extreme rainfall estimation are so small that it is not possible to single out any one model as having the best overall performance, although for increasingly rare events the results suggest a preference for the BL1M model. This result supports the findings reported by Kaczmarska et al. (2014) that the dependence structure between rain cell amounts and duration in the BL1M model is beneficial in estimating fine–scale extremes.

In all three models, there is a slight upward curvature in the Gumbel plotting of extremes which is consistent with the GEV and GP distributions taking a positive shape parameter ($\xi > 0$). This curvature is more pronounced for the BL1M model which would be consistent with a higher positive shape parameter. While extreme value theory encompasses a range of distributions characterised by the sign of the shape parameter, Koutsoyiannis (2004a) argues that rainfall extremes naturally follow the Fréchet distribution for annual maxima (equivalent to the GEV with $\xi > 0$), supported with empirical evidence in Koutsoyiannis (2004b). The positive growth in extremes observed in our results is consistent with this hypothesis, and suggests that important information about the distribution of extremes is captured in the full storm profile hyetograph over the low censor. Futher research is required to investigate the theoretical link between mechanistic model parameters and their extreme value performance.

The results presented in this paper show that the method has worked for single site data from two very different locations, and recorded using different gauging techniques. Consistency in the relative magnitude of selected censors identified at each location, and the stability of estimation across a range of censors gives confidence in the approach and supports the original hypothesis. It is an obvious limitation of censoring that it cannot be used to obtain time-series of synthetic rainfall as is the principal application of mechanistic rainfall models. However, the intention of this research was to investigate if mechanistic models could be used for estimation of fine–scale extremes as an alternative to frequency techniques. The accuracy of estimates for sub-hourly rainfall extremes using all three model variants is very good, although the BL1M model appears to outperform the other two models at both sites for the 5 and 15 minute scales by accurately predicting the highest observed extreme.

Reducing parameterization by fixing the Gamma shape parameter α in the randomised models, and increasing the data for parameterization by widening the fitting window to 3 months has enabled the models to be fitted successfully to censored

Deleted: 0.04 AEP

Deleted: 0.001 AEP

Deleted: 0.04 AEP

Deleted: lower probability

Deleted: , Isham & Onof

Deleted: <#>**Conclusions** ¶
<#>Censored rainfall synthesis using mechanistic pulse based models appears to offer an alternative approach to estimation of rainfall extremes and to frequency estimation techniques.

Deleted: <#>which has been collected

Deleted: <#>optimal

Deleted: <#>for

Deleted: <#>¶

observations. It is likely that these aides to parameterization are necessary because censoring truncates the statistical distribution of the training data. The analytical solutions in the models do not make this assumption, therefore a mismatch between the training data and the models arises with censoring. At low censors, truncation is minor and the analytical solutions in the models are able to make reasonable estimates of the fitting statistics. However, as the censor increases and the mismatch grows a point is reached at which the analytical solutions are no longer able to estimate the fitting statistics causing deterioration in parameter identifiability.

A principal goal of this research was to improve the physical basis of short duration extreme rainfall estimation. This has been achieved by simulating storm profiles mechanistically in a way which mimics the phenomenology of rainfall generation. This has given rise to extreme rainfall estimation which may be described as a function of a set of model parameters with physical meaning, e.g., the extreme rainfall quantile $z = F\{\lambda, \mu_x, \delta_x, \delta_c, \mu_c, \delta_s\}$ for the original Bartlett-Lewis model (See Appendix A for definitions of mechanistic model parameters). Future research is required to link model parameters to environmental covariates and spatial information. The latter may follow the regionalisation methodology of Kim et al. (2013).

Further research is also required to investigate the potential for incorporating censored modelling into a multi-model approach for synthetic rainfall generation. This may take the form of simulating the rainfall below the censor using a bootstrapping approach similar to that in Costa et al. (2015), or continuous simulation of uncensored rainfall with replacement of storms simulated using the censoring approach.

Deleted: , Fernandes & Naghettini

Data availability

The Atherstone tipping bucket raingauge dataset was obtained directly from the Environment Agency for England, UK. The data are not publicly accessible because they are used by the Environment Agency for operational purposes, but can be obtained for non-commercial purposes on request. The Bochum dataset was obtained directly from Deutsche Montan Technologie and was recorded by the Emschergenossenschaft / Lippeverband in Germany. The data are not publicly accessible because they belong to the Emschergenossenschaft and Lippeverband public German water boards and are used for operational purposes.

Appendix A: Bartlett-Lewis model parameter sensitivity and impact on extreme value estimation.

To demonstrate the insensitivity of α for the randomised Bartlett-Lewis models, the BL1 and BL1M models were fitted to Bochum 15 minute rainfall with changing constraints on α. The models were fitted using the 1 hour mean and the 0.25, 6 and 24 hour coefficient of variation, skewness coefficient and lag-1 autocorrelation. For the BL1 model, α is constrained between 4.1 (lower bound) and 5, 10, 25, 50, 75 and 100 (upper bounds). For the BL1M model, α is constrained between 5, 10, 25, 50, 75 and 100 (lower bounds) and infinity (upper bound). For the BL1 and BL1M models, α converges on the upper and lower bounds respectively, although because α is not held fixed parameter uncertainty is estimated. Parameter ranges are presented in the parallel coordinate plots in Fig. A1 by sampling 1000 parameter sets from the multivariate normal distribution of model parameters for $4.1 \leq \alpha < 1e+06$. The parameter sets corresponding to $\alpha = 100$ and $\alpha = 5$ are shown for the BL1 and BL1M models respectively with dashed magenta lines.

Deleted: >

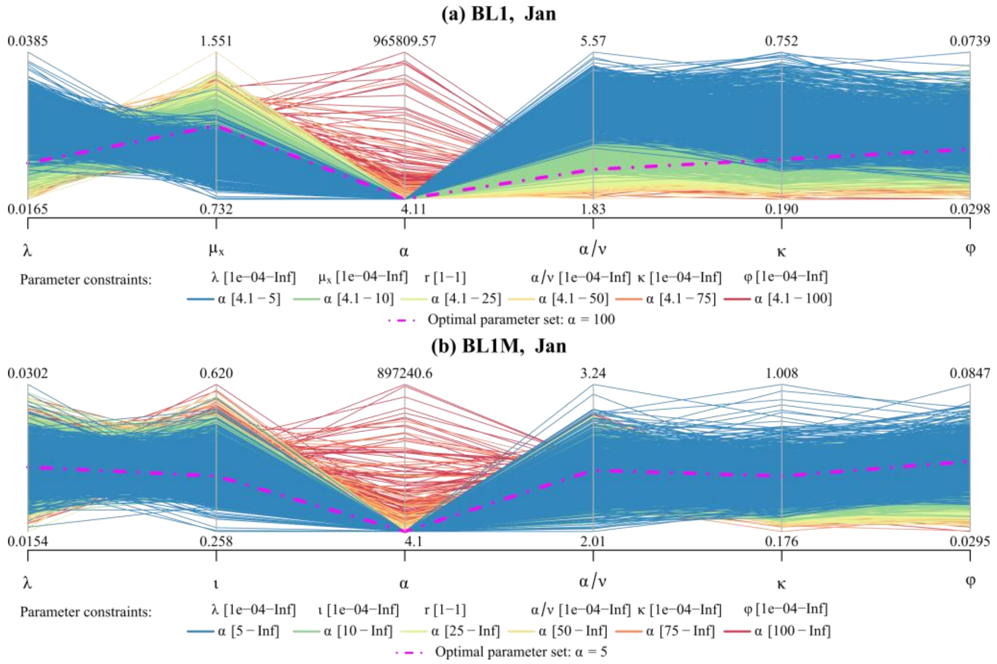

Figure A1 Parallel coordinate plots for the two randomised Bartlett-Lewis rectangular pulse models, BL1 and BL1M. Plots show the range of January parameter values for uncensored models fitted to Bochum 15 minute rainfall. The dashed magenta lines show the parameter sets corresponding to $\alpha = 100$ and $\alpha = 5$ for the BL1 and BL1M models respectively.

The parallel coordinate plots clearly show the insensitivity of α compared with the other model parameters. When α is constrained with upper and lower bounds of between 25-50 for the BL1 and BL1M models respectively, α is poorly identified and can take any value over a very large range (see Fig. A1). When α is constrained with upper and lower bounds of less than 25 for the BL1 and BL1M models respectively, identifiability of α is improved. This insensitivity results from the shape of the fitted Gamma distribution used to sample η shown in Fig. A2.

As α increases the Gamma distribution converges on the Normal distribution and becomes increasingly flat. Therefore, for high values of α, the probability of randomly sampling anywhere within the distribution is greater compared with low values of α. For $\alpha \geq 50$, the Gamma distribution is approximately normal and the range of η values which may be randomly sampled by both models is always large resulting in a narrow range of potential Exponential distributions from which to sample L where L is the cell duration. This impacts the estimation of extremes as shown in Fig. A3. Figure A3 shows extreme rainfall estimates from the BL1 and BL1M models with α fixed at 5, 50 and 100. For $\alpha \geq 50$, extreme rainfall estimation by both models is identical. For $\alpha = 5$, the BL1 model estimates lower extremes than with higher α values, while the BL1M model gives improved

Deleted: 1/

estimation of the growth curve of extremes. Because of this combination of parameter insensitivity and relative performance in the extremes, α is fixed at 100 and 5 for the BL1 and BL1M models respectively.

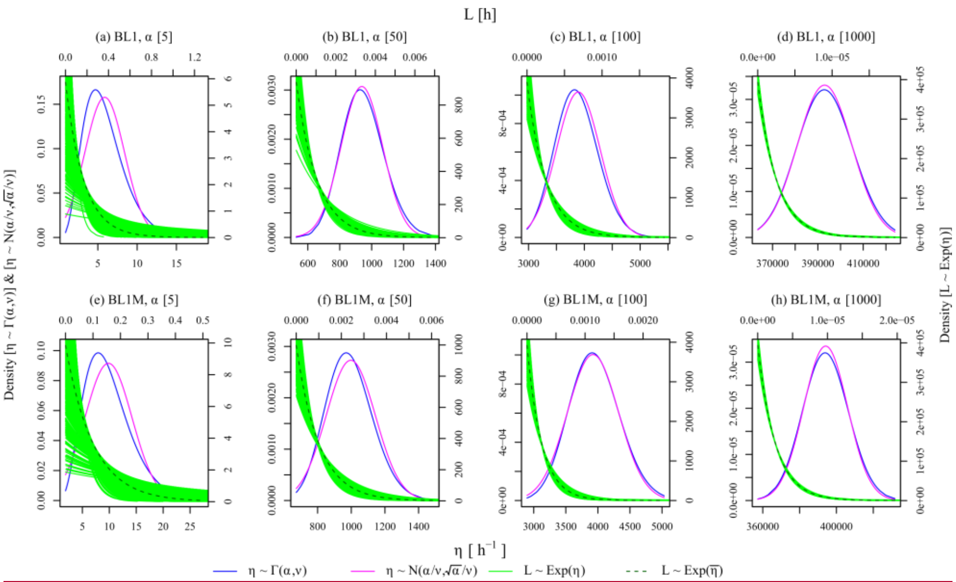

Figure A2. Fitted Gamma distributions for the cell duration parameter η for the BL1 and BL1M models with α = 5, 50, 100 and 1000. Plots show the equivalent Normal distributions fitted to the mean and standard deviation of the Gamma distributions. The range of Exponential distributions for the cell duration parameter η are obtained by sampling 500 η values from the fitted Gamma distributions. The Exponential distributions for the mean of the fitted Gamma distributions are also shown.

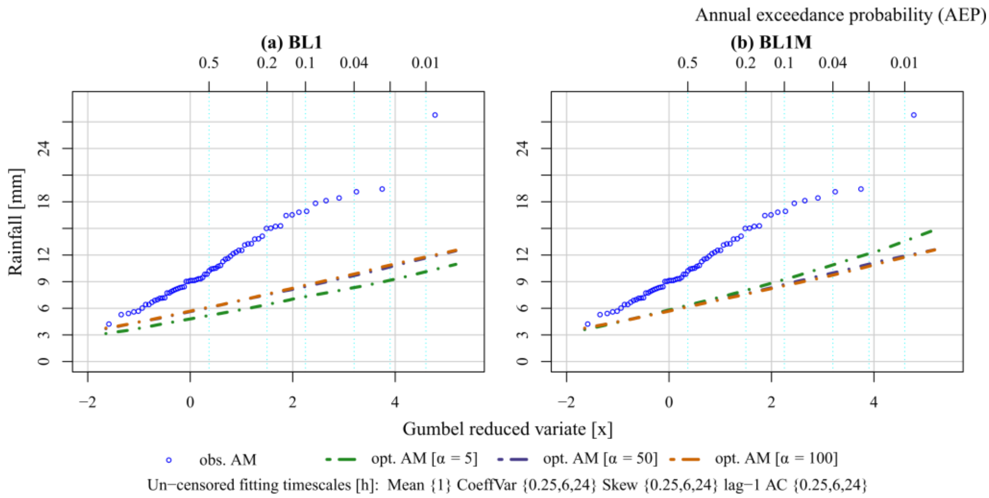

Figure A3 Sensitivity of extreme value estimation to choice of α for the randomised Bartlett-Lewis models.

Appendix B: Fitted model parameters

Tables B1–4 show fitted model parameters for the BL1M model (BLRPR$_X$ in Table 1) for 5 and 15 minute rainfall at Bochum and Atherstone with uncensored and censored data. Censored model parameters correspond to the censors selected in Table 2. Additionally, Tables B1–4 show the objective function value, S_{min}, for the fitted parameter set, as well as mechanistic model parameters defined by Wheater et al. (2007b) which are listed below.

Mean number of cells per storm:	$\mu_c = 1 + \frac{\kappa}{\varphi}$	[-]
Mean cell duration:	$\delta_c = \frac{\nu}{\alpha - 1}$	[h]
Mean duration of storm activity:	$\delta_c = \frac{\nu}{(\alpha - 1)\varphi}$	[h]

Table B1 BL1M model parameters for the Bochum 5 minute data

	λ [hr^{-1}]	ι [mm]	α [hr]	α/ν [hr]	κ [-]	φ [-]	S_{min} [-]	μ_c [-]	δ_c [min]	δ_s [hr]
Uncensored [> 0 mm]										
Jan	0.022	0.318	4.100	3.788	0.469	0.042	22.4	12.2	20.9	8.3
Feb	0.021	0.326	4.100	4.052	0.387	0.038	25.2	11.2	19.6	8.6
Mar	0.021	0.350	4.100	4.788	0.300	0.034	27.4	9.8	16.6	8.1
Apr	0.022	0.423	4.100	5.943	0.211	0.029	41.2	8.3	13.4	7.7
May	0.024	0.510	4.100	7.594	0.205	0.032	48.4	7.4	10.4	5.4
Jun	0.024	0.682	4.100	9.082	0.164	0.032	55.8	6.1	8.7	4.6
Jul	0.024	0.766	4.100	9.839	0.152	0.032	57.0	5.8	8.1	4.2
Aug	0.023	0.786	4.100	9.294	0.133	0.029	54.8	5.6	8.5	4.9
Sep	0.021	0.626	4.100	7.743	0.175	0.029	46.8	7.0	10.2	5.9
Oct	0.021	0.506	4.100	6.008	0.226	0.030	32.1	8.5	13.2	7.3
Nov	0.021	0.380	4.100	4.697	0.359	0.036	28.1	11.0	16.9	7.8
Dec	0.022	0.332	4.100	3.984	0.435	0.039	29.5	12.2	19.9	8.5
Censored [> 0.5 mm]										
Jan	0.007	0.288	5.000	42.472	0.003	0.007	0.5	1.4	1.8	4.2
Feb	0.007	0.302	5.000	47.366	0.002	0.005	0.2	1.4	1.6	5.3
Mar	0.008	0.310	5.000	47.877	0.003	0.008	0.4	1.4	1.6	3.3
Apr	0.009	0.500	5.000	36.524	0.005	0.011	0.0	1.5	2.1	3.1
May	0.010	0.831	5.000	27.921	0.020	0.036	0.6	1.6	2.7	1.2
Jun	0.011	1.056	5.000	21.226	0.030	0.039	1.6	1.8	3.5	1.5
Jul	0.012	1.215	5.000	21.060	0.029	0.040	0.9	1.7	3.6	1.5
Aug	0.011	1.177	5.000	21.399	0.025	0.033	0.7	1.8	3.5	1.8
Sep	0.011	1.033	5.000	26.439	0.009	0.021	0.3	1.4	2.8	2.3
Oct	0.009	0.716	5.000	31.356	0.003	0.010	0.1	1.3	2.4	4.0
Nov	0.008	0.377	5.000	39.141	0.001	0.003	0.0	1.3	1.9	10.6
Dec	0.007	0.317	5.000	41.561	0.001	0.005	0.0	1.2	1.8	6.0

Table B2 BL1M model parameters for the Atherstone 5 minute data

	λ [hr^{-1}]	ι [mm]	α [hr]	α/ν [hr]	κ [-]	φ [-]	S_{min} [-]	μ_c [-]	δ_c [min]	δ_s [hr]
Uncensored [> 0 mm]										
Jan	0.023	0.095	4.100	79.758	0.157	0.005	49.4	32.4	1.0	3.3
Feb	0.022	0.083	4.100	110.403	0.117	0.004	54.4	30.3	0.7	3.0
Mar	0.022	0.097	4.100	64.565	0.139	0.005	61.7	28.8	1.2	4.1
Apr	0.020	0.137	4.100	46.008	0.161	0.007	41.1	24.0	1.7	4.1
May	0.018	0.233	4.100	28.827	0.172	0.011	28.3	16.6	2.8	4.2
Jun	0.017	0.328	4.100	22.831	0.195	0.016	20.5	13.2	3.5	3.6
Jul	0.017	0.395	4.400	19.700	0.186	0.018	20.0	11.3	3.9	3.6
Aug	0.017	0.338	4.400	22.385	0.209	0.018	21.5	12.6	3.5	3.2
Sep	0.018	0.253	4.100	26.711	0.224	0.014	28.2	17.0	3.0	3.5

Oct	0.019	0.164	4.100	37.406	0.245	0.010	37.6	25.5	2.1	3.5
Nov	0.021	0.115	4.100	52.659	0.234	0.007	47.6	34.4	1.5	3.6
Dec	0.022	0.091	4.100	84.400	0.185	0.005	49.1	38.0	0.9	3.1
					Censored [> 0.6 mm]					
Jan	0.007	0.316	5.000	43.290	0.029	0.057	5.5	1.5	1.7	0.5
Feb	0.007	0.250	5.000	52.351	0.025	0.045	6.7	1.6	1.4	0.5
Mar	0.007	0.275	5.000	56.244	0.014	0.028	4.5	1.5	1.3	0.8
Apr	0.007	0.392	5.000	50.130	0.012	0.020	1.9	1.6	1.5	1.2
May	0.007	0.594	5.000	37.073	0.014	0.022	0.9	1.6	2.0	1.5
Jun	0.008	0.695	5.000	31.026	0.029	0.036	1.9	1.8	2.4	1.1
Jul	0.008	0.805	5.000	26.653	0.027	0.034	0.6	1.8	2.8	1.4
Aug	0.008	0.719	5.000	29.868	0.027	0.032	0.5	1.8	2.5	1.3
Sep	0.008	0.644	5.000	33.789	0.014	0.023	1.5	1.6	2.2	1.6
Oct	0.008	0.463	5.000	46.623	0.009	0.018	0.2	1.5	1.6	1.5
Nov	0.008	0.369	5.000	46.777	0.007	0.021	1.6	1.3	1.6	1.3
Dec	0.007	0.318	5.000	49.550	0.015	0.037	5.2	1.4	1.5	0.7

Table B3 BL1M model parameters for the Bochum 15 minute data

	λ [hr^{-1}]	ι [mm]	α [hr]	α/ν [hr]	κ [-]	φ [-]	S_{min} [-]	μ_c [-]	δ_c [min]	δ_s [hr]
					Uncensored [> 0 mm]					
Jan	0.022	0.373	4.100	2.562	0.545	0.059	18.1	10.2	31.0	8.7
Feb	0.021	0.375	4.100	2.752	0.458	0.052	20.7	9.8	28.8	9.2
Mar	0.021	0.376	4.100	3.251	0.404	0.049	23.6	9.2	24.4	8.3
Apr	0.022	0.448	4.100	4.023	0.292	0.043	34.4	7.8	19.7	7.6
May	0.024	0.518	4.100	5.323	0.307	0.049	40.3	7.3	14.9	5.1
Jun	0.024	0.665	4.100	6.799	0.254	0.048	46.7	6.3	11.7	4.1
Jul	0.024	0.738	4.100	7.496	0.241	0.048	46.8	6.0	10.6	3.7
Aug	0.023	0.749	4.100	7.054	0.219	0.044	43.5	6.0	11.2	4.3
Sep	0.021	0.624	4.100	5.539	0.268	0.045	35.0	7.0	14.3	5.3
Oct	0.021	0.529	4.100	4.079	0.321	0.045	23.7	8.1	19.5	7.2
Nov	0.021	0.434	4.100	3.117	0.439	0.051	19.3	9.6	25.5	8.3
Dec	0.022	0.411	4.100	2.593	0.463	0.052	20.1	9.9	30.6	9.8
					Censored [> 1.0 mm]					
Jan	0.008	0.340	5.000	21.917	0.005	0.010	0.4	1.5	3.4	5.7
Feb	0.008	0.355	5.000	23.830	0.004	0.010	0.3	1.4	3.1	5.2
Mar	0.008	0.401	5.000	22.721	0.005	0.014	0.7	1.4	3.3	3.9
Apr	0.009	0.629	5.000	18.092	0.007	0.016	0.6	1.4	4.1	4.3
May	0.010	0.987	5.000	15.213	0.014	0.026	1.8	1.5	4.9	3.2
Jun	0.012	1.240	5.000	13.109	0.017	0.026	1.5	1.7	5.7	3.7
Jul	0.012	1.397	5.000	13.518	0.017	0.025	1.2	1.7	5.5	3.7
Aug	0.012	1.372	5.000	13.857	0.013	0.020	0.4	1.7	5.4	4.5
Sep	0.010	1.141	5.000	15.219	0.010	0.019	0.0	1.5	4.9	4.3
Oct	0.009	0.809	5.000	16.864	0.005	0.011	0.0	1.5	4.4	6.7
Nov	0.008	0.442	5.000	18.891	0.004	0.008	0.1	1.5	4.0	8.3
Dec	0.007	0.385	5.000	21.324	0.003	0.007	0.1	1.4	3.5	8.4

Table B4 BL1M model parameters for the Atherstone 15 minute data

	λ [hr^{-1}]	ι [mm]	α [hr]	α/ν [hr]	κ [-]	φ [-]	S_{min} [-]	μ_c [-]	δ_c [min]	δ_s [hr]
					Uncensored [> 0 mm]					
Jan	0.022	0.147	4.100	14.734	0.505	0.025	17.6	21.2	5.4	3.6
Feb	0.022	0.129	4.100	17.115	0.445	0.021	18.7	22.2	4.6	3.7
Mar	0.022	0.141	4.100	16.373	0.354	0.019	24.5	19.6	4.8	4.3
Apr	0.020	0.184	4.100	14.038	0.373	0.022	20.0	18.0	5.7	4.3
May	0.018	0.302	4.100	10.715	0.327	0.027	19.1	13.1	7.4	4.6
Jun	0.017	0.410	4.100	9.692	0.335	0.035	19.4	10.6	8.2	3.9
Jul	0.017	0.482	4.100	9.013	0.304	0.038	21.9	9.0	8.8	3.9
Aug	0.018	0.408	4.100	9.896	0.366	0.040	22.4	10.2	8.0	3.3
Sep	0.019	0.335	4.100	9.388	0.440	0.039	24.0	12.3	8.5	3.6
Oct	0.019	0.243	4.100	9.798	0.569	0.036	22.6	16.8	8.1	3.7
Nov	0.021	0.199	4.100	10.078	0.620	0.034	20.4	19.2	7.9	3.9

Dec	0.021	0.164	4.100	11.699	0.640	0.031	15.9	21.6	6.8	3.6
Censored [> 0.6 mm]										
Jan	0.010	0.472	5.000	12.186	0.047	0.048	0.5	2.0	6.2	2.1
Feb	0.010	0.400	5.000	13.782	0.041	0.046	0.9	1.9	5.4	2.0
Mar	0.010	0.399	5.000	15.383	0.029	0.038	0.3	1.8	4.9	2.1
Apr	0.010	0.501	5.000	13.827	0.046	0.039	0.2	2.2	5.4	2.3
May	0.010	0.780	5.000	11.245	0.031	0.028	0.1	2.1	6.7	4.0
Jun	0.009	0.904	5.000	11.080	0.058	0.039	0.0	2.5	6.8	2.9
Jul	0.010	1.056	5.000	10.606	0.041	0.033	0.1	2.2	7.1	3.6
Aug	0.011	1.082	5.000	9.943	0.025	0.024	0.2	2.0	7.5	5.2
Sep	0.011	0.924	5.000	9.330	0.019	0.017	0.2	2.1	8.0	7.9
Oct	0.010	0.711	5.000	8.875	0.029	0.024	0.3	2.2	8.5	5.9
Nov	0.010	0.522	5.000	9.761	0.057	0.040	0.5	2.4	7.7	3.2
Dec	0.010	0.484	5.000	10.229	0.064	0.048	0.8	2.3	7.3	2.5

Author contribution

David Cross designed the experiments, carried them out and prepared the manuscript. Christian Onof, Hugo Winter and Pietro Bernardara supervised the work and reviewed the manuscript preparation.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

David Cross is grateful for the award of an Industrial Case Studentship from the Engineering and Physical Sciences Research Council and EDF Energy. The Environment Agency of England are gratefully acknowledged for providing the UK rainfall data, and Deutsche Montan Technologie and Emschergenossenschaft / Lippeverband in Germany are gratefully acknowledged for providing the Bochum data.

References

Bacchi, B., Becciu, G. and Kottegoda, N. T.: Bivariate exponential model applied to intensities and durations of extreme rainfall, Journal of Hydrology, 155, 225-236, http://dx.doi.org/10.1016/0022-1694(94)90166-X, 1994.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrol. Process., 6, 279-298, 1992.

Cameron, D., Beven, K. and Tawn, J.: An evaluation of three stochastic rainfall models, Journal of Hydrology, 228, 130-149, 2000a.

Cameron, D., Beven, K. and Tawn, J.: Modelling extreme rainfalls using a modified random pulse Bartlett–Lewis stochastic rainfall model (with uncertainty), Adv. Water Resour., 24, 203-211, 2000b.

Chandler, R., Lourmas, G. and Jesus, J.: MOMFIT Software for moment-based fitting of single-site stochastic rainfall model fitting, User guide, Department of Statistical Science, University College London, London, 2010.

Deleted: ¶

Formatted: Heading 1, Line spacing: Double

Formatted: Font: 10 pt

Formatted: Font: 10 pt

Formatted: Font: 10 pt

Chandler, R. E.: A Spectral Method for Estimating Parameters in Rainfall Models, Bernoulli, 3, 301-322, 1997.

Costa, V., Fernandes, W. and Naghettini, M.: A Bayesian model for stochastic generation of daily precipitation using an upper-bounded distribution function, Stochastic Environmental Research and Risk Assessment, 29, 563-576, 2015.

Cowpertwait, P.: A generalized point process model for rainfall, Proceedings of the Royal Society of London.Series A: Mathematical and Physical Sciences, 447, 23-37, 1994.

Cowpertwait, P., Isham, V. and Onof, C.: Point process models of rainfall: developments for fine-scale structure, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science, 463, 2569-2587, 2007.

Cowpertwait, P., O'Connell, P., Metcalfe, A. and Mawdsley, J.: Stochastic point process modelling of rainfall. I. Single-site fitting and validation, Journal of Hydrology, 175, 17-46, 1996.

Cowpertwait, P. S. P.: A Poisson-cluster model of rainfall: high-order moments and extreme values, Proceedings of the Royal Society A-Mathematical Physical and Engineering Sciences, 454, 885-898, 1998.

Cowpertwait, P. S. P., Xie, G., Isham, V., Onof, C. and Walsh, D. C. I.: A fine-scale point process model of rainfall with dependent pulse depths within cells, Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 56, 1110-1117, 2011.

Cowpertwait, P.: Mixed rectangular pulses models of rainfall, Hydrol. Earth Syst. Sci., 8, 993-1000, 2004.

Cox, D. and Isham, V.: Stochastic models of precipitation, Statistics for the Environment, 2, 3-18, 1994.

De Michele, C. and Salvadori, G.: A Generalized Pareto intensity-duration model of storm rainfall exploiting 2-Copulas, Journal of Geophysical Research: Atmospheres, 108, n/a-n/a, 2003.

Entekhabi, D., Rodriguez‐Iturbe, I. and Eagleson, P. S.: Probabilistic representation of the temporal rainfall process by a modified Neyman‐Scott Rectangular Pulses Model: Parameter estimation and validation, Water Resour. Res., 25, 295-302, 1989.

Evin, G. and Favre, A. -.: A new rainfall model based on the Neyman-Scott process using cubic copulas, Water Resour. Res., 44, W03433, 2008.

Fawcett, L. and Walshaw, D.: Estimating return levels from serially dependent extremes, Environmetrics, 23, 272-283, 2012.

Gabriel, K. R. and Neumann, J.: A Markov chain model for daily rainfall occurrence at Tel Aviv, Q. J. R. Meteorol. Soc., 88, 90-95, 1962.

Goldhirsh, J., Krichevsky, V. and Gebo, N. E.: Rain rate statistics and fade distributions at 20 and 30 GHz derived from a network of rain gauges in the mid-Atlantic coast over a five year period, IEEE transactions on antennas and propagation, 40, 1408-1415, 1992.

Habib, E., Krajewski, W. and Kruger, A.: Sampling errors of tipping-bucket rain gauge measurements, J. Hydrol. Eng., 6, 159-166, 2001.

Kaczmarska, J.: Single-site point process-based rainfall models in a nonstationary climate, PhD thesis/masters, University College London (University of London), 2013.

Kaczmarska, J.: Further development of Bartlett–Lewis models for fine-resolution rainfall, Department of Statistical Science, University College London, London, 2011.

Kaczmarska, J., Isham, V. and Onof, C.: Point process models for fine- resolution rainfall, Hydrological Sciences Journal, 59, 1972-1991, 2014.

Kakou, A.: Point process based models for rainfall., Doctor of Philosophy thesis/masters, University College London (University of London), 1997.

Katz, R. W.: Precipitation as a Chain-Dependent Process, J. Appl. Meteor., 16, 671-676, 1977.

Formatted: Font: (Default) Cambria Math, 10 pt

Formatted: Font: 10 pt

Formatted: Font: (Default) Cambria Math, 10 pt

Formatted: Font: 10 pt

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

Kim, S. and Kavvas, M. L.: Stochastic point rainfall modeling for correlated rain cell intensity and duration, J. Hydrol. Eng., 11, 29-36, 2006.

Kim, D., Olivera, F., Cho, H. and Socolofsky, S. A.: Regionalization of the Modified Bartlett- Lewis Rectangular Pulse Stochastic Rainfall Model, Terrestrial Atmospheric and Oceanic Sciences, 24, 421-436, 2013.

Kossieris, P., Tyralis, w. H., Koutsoyiannis, D. and Efstratiadis., A.: HyetosR: A package for temporal stochastic simulation of rainfall at fine time scales, 2013.

Koutsoyiannis, D.: Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: I. Recherche théorique, Hydrological sciences journal, 49, 575-590, 2004a.

Koutsoyiannis, D. and Onof, C.: Rainfall disaggregation using adjusting procedures on a Poisson cluster model, Journal of Hydrology, 246, 109-122, 2001.

A computer program for temporal rainfall disaggregation using adjusting procedures (HYETOS): https://www.itia.ntua.gr/en/docinfo/59/, 2.

Koutsoyiannis, D.: Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations, Hydrol. Sci. J. -J. Sci. Hydrol., 49, 591-610, 2004b.

Kurothe, R., Goel, N. and Mathur, B.: Derived flood frequency distribution for negatively correlated rainfall intensity and duration, Water Resour. Res., 33, 2103-2107, 1997.

Northrop, P. J. and Stone, T. M.: A point process model for rainfall with truncated Gaussian rain cells, Department of Statistical Science, University College London, London, 1-15 pp., 2005.

Nystuen, J. A., Proni, J. R., Black, P. G. and Wilkerson, J. C.: A Comparison of Automatic Rain Gauges, J. Atmos. Oceanic Technol., 13, 62-73, 1996.

Onof, C. and Arnbjerg-Nielsen, K.: Quantification of anticipated future changes in high resolution design rainfall for urban areas, Atmos. Res., 92, 350-363, 2009.

Onof, C., Chandler, R., Kakou, A., Northrop, P., Wheater, H. and Isham, V.: Rainfall modelling using Poisson-cluster processes: a review of developments, Stochastic Environmental Research and Risk Assessment, 14, 384-411, 2000.

Onof, C., Townend, J. and Kee, R.: Comparison of two hourly to 5-min rainfall disaggregators, Atmos. Res., 77, 176-187, 2005.

Onof, C. and Wheater, H. S.: Improved fitting of the Bartlett-Lewis Rectangular Pulse Model for hourly rainfall, Hydrological sciences journal, 39, 663-680, 1994a.

Onof, C. and Wheater, H. S.: Improvements to the modelling of British rainfall using a modified random parameter Bartlett-Lewis rectangular pulse model, Journal of Hydrology, 157, 177-195, 1994b.

Onof, C. and Wheater, H. S.: Modelling of British rainfall using a random parameter Bartlett-Lewis rectangular pulse model, Journal of Hydrology, 149, 67-95, 1993.

R Core Team: R: A Language and Environment for Statistical Computing, 2017.

Richardson, C. W.: Stochastic simulation of daily precipitation, temperature, and solar radiation, Water Resour. Res., 17, 182-190, 1981.

Rodriguez-Iturbe, I., Cox, D. and Isham, V.: A point process model for rainfall: further developments, Proceedings of the Royal Society of London.A.Mathematical and Physical Sciences, 417, 283-298, 1988.

Rodriguez-Iturbe, I., Cox, D. and Isham, V.: Some models for rainfall based on stochastic point processes, Proceedings of the Royal Society of London.A.Mathematical and Physical Sciences, 410, 269-288, 1987.

Srikanthan, R. and McMahon, T.: Stochastic generation of annual, monthly and daily climate data: A review, Hydrology and Earth System Sciences Discussions, 5, 653-670, 2001.

Stern, R. and Coe, R.: A model fitting analysis of daily rainfall data, Journal of the Royal Statistical Society.Series A (General), 1-34, 1984.

Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, B. and Midgley, B.: IPCC, 2013: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1-33, 2013.

Vandenberghe, S., Verhoest, N. E. C., Onof, C. and De Baets, B.: A comparative copula-based bivariate frequency analysis of observed and simulated storm events: A case study on Bartlett-Lewis modeled rainfall, Water Resour. Res., 47, W07529, 2011.

Verhoest, N. E., Vandenberghe, S., Cabus, P., Onof, C., Meca‐Figueras, T. and Jameleddine, S.: Are stochastic point rainfall models able to preserve extreme flood statistics?, Hydrol. Process., 24, 3439-3445, 2010.

Verhoest, N., Troch, P. A. and De Troch, F. P.: On the applicability of Bartlett–Lewis rectangular pulses models in the modeling of design storms at a point, Journal of Hydrology, 202, 108-120, 1997.

Wasko, C. and Sharma, A.: Continuous rainfall generation for a warmer climate using observed temperature sensitivities, Journal of Hydrology, 544, 575-590, 2017.

Waymire, E. and Gupta, V. K.: The Mathematical Structure of Rainfall Representations .1. a Review of the Stochastic Rainfall Models, Water Resour. Res., 17, 1261-1272, 1981.

Westra, S., Fowler, H., Evans, J., Alexander, L., Berg, P., Johnson, F., Kendon, E., Lenderink, G. and Roberts, N.: Future changes to the intensity and frequency of short‐duration extreme rainfall, Rev. Geophys., 52, 522-555, 2014.

Wheater, H. S., Isham, V. S., Chandler, R. E., Onof, C. J. and Stewart, E. J.: Improved methods for national spatial-temporal rainfall and evaporation modelling for BSM, Department for Environment, Food and Rural Affairs (DEFRA), Flood management Division, London, 2007a.

Wheater, H. S., Isham, V. S., Chandler, R. E., Onof, C. J. and Stewart, E. J.: Improved methods for national spatial-temporal rainfall and evaporatoin modelling for BSM: Appendices, Department for Environment, Food and Rural Affairs (DEFRA), Flood management Division, London, 2007b.

Wilks, D. S. and Wilby, R. L.: The weather generation game: a review of stochastic weather models, Prog. Phys. Geogr., 23, 329-357, 1999.

Yu, B., Ciesiolka, C., Rose, C. and Coughlan, K.: A note on sampling errors in the rainfall and runoff data collected using tipping bucket technology, Trans. ASAE, 40, 1305-1309, 1997.

Formatted: Font: (Default) Cambria Math, 10 pt

Formatted: Font: 10 pt

Formatted: Font: (Default) Cambria Math, 10 pt

Formatted: Font: 10 pt

Formatted: Justified, Line spacing: 1.5 lines