ICE Publishing

Badr and Venables
ISBN 978-0-7277-6525-3
https://doi.org/10.1680/oicwe.65253.383


# QUANTITATIVE MODELLING OF HYDROCLIMATIC EXTREMES

*Y. Xuan[1]* and H. Wang[1]*

[1] *College of Engineering, Swansea University, Swansea, UK*
**Corresponding author*

**ABSTRACT** Design of many civil engineering structures relies on understanding of the dynamics and behaviour of hydroclimatic extremes, such as extreme storms, floods. Conventionally this is done by applying frequency analysis over the events that happened in the past. Yet, climate change and further global change have already changed the dynamics of the extremes and therefore there is an urgent need to study how the changing hydroclimatic extremes affect the engineering design in a quantitative way. In this paper, modelling of the spatial-temporal variations of the hydroclimatic extremes are demonstrated using high-resolution rainfall observation and climatic projections over the UK and Australia. Results show that: (1) non-stationary modelling is preferred to capture the time varying distribution of rainfall extremes; (2) spatial distribution of the rainfall extremes is highly heterogeneous but demonstrates coherent patterns of the parameterisation over the characteristics of the sampling area; (3) the current level of climate model projections are able to represent the average time-varying trend of the extremes but do not yet come to closely reveal the overall variations of the extremes.

## Notation

| | |
|---|---|
| GEV | Generalised Extreme Value |
| $\mu$ | Location parameter |
| $\sigma$ | Scale parameter |
| $\xi$ | Shape parameter |
| $sp$ | Spatial index |
| AMDR | Annual maximum daily rainfall |
| ROI | Region of interest |
| GB | Great Britain |
| AU | Australia |

## 1. Introduction

Design of many civil engineering structures relies on understanding of the dynamics and behaviour of hydroclimatic extremes, such as extreme storms, floods (Hall *et al.*, 2014; Merz *et al.*, 2014). Conventionally this is done by applying frequency analysis over the events that happened in the past. Yet, climate change and further global change have already changed the dynamics of the extremes and therefore this is an urgent need to study how the changing hydroclimatic extremes affect the engineering design in a quantitative way. For example, flood defence structures are designed to withhold floods up to certain threshold, e.g. flood size, which in turn is determined by fitting historical extremes such as annual maximum flood peaks using probability distribution whose parameters are assumed to be stationary (Coles and Tawn, 1996; Lazoglou and Anagnostopoulou, 2017; Mannshardt-Shamseldin *et al.*, 2010; Shukla *et al.*, 2012; Yoon *et al.*, 2015).

As has become increasingly clear, climate change has already altered the environment and hence the flooding process. Such a stationary view needs to be changed (Assani and Guerfi, 2017; Herring *et al.*, 2018). Further, it has also been widely observed that the impact from climate change is not homogeneous: while certain areas may suffer more severe flooding, others simply observed more water deficiencies. Understanding the spatial difference of the changing hydroclimatic extremes is another important dimension to be explored. Lastly, climate projections from climate models remain as the main source of information regarding future hydroclimatic extremes (Sillmann *et al.*, 2013). However, most aforementioned studies have been concerned with the average trend variation and barely any discussion about how reliably they can reproduce the required hydroclimatic extremes by engineering design. The present paper gives a brief summary about the recent quantitative modelling efforts in these directions. Cutting-edge, high resolution and long-term rainfall data are used alongside the recent climate projections. The study focuses on two areas, Great Britain (GB) and Australia (AU). The rest of the paper is organised in a way to cover the three questions mentioned above.

## 2. Study areas and datasets

### *2.1 The GEAR rainfall dataset for Great Britain*

The GEAR dataset is an archive of the daily gridded rainfall over Great Britain (GB) with a spatial resolution of 1km × 1km from 1890 to 2010 (Tanguy *et al.*, 2016). It is derived from the UK Meteorological Office national database of observed precipitation collected by the UK rain gauge network. The daily time window is defined as 24 hours between 9am on the day until 9am on the following day.

### *2.2 The ADAM rainfall dataset for Australia*

The ADAM data (the Australian Data Archive for Meteorology) are also a daily gridded rainfall dataset but over Australia (AU) from 1900 to 2018. It is generated using a sophisticated analysis technique described in Jones *et al.* (2009). It uses an optimised Barnes successive correction technique that applies a weighted averaging process to the station data. Each grid represents an approximately square area with sides of about 5 km (0.05 degrees). The data are recorded as 24-hour total rainfall from 9am (local time) the day before, to 9am of the current day.

### *2.3 The UK climate projection*

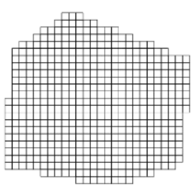
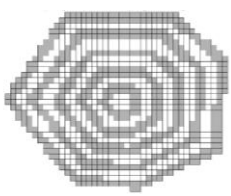
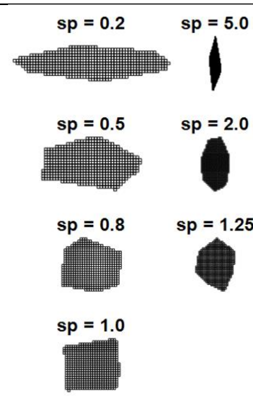
The ERA20CM data, is derived from an ensemble climatic projection with 10 ensemble members over the area of a $0.4^{o} \times 3.15^{o}$ (latitude × longitude) grid with a 3-hour spatial and temporal resolution from 1900 to 2010 (Hersbach *et al.*, 2015), provided by the European Centre for Medium-Range Weather Forecasts (ECMWF).

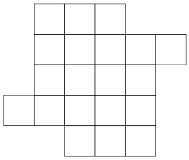
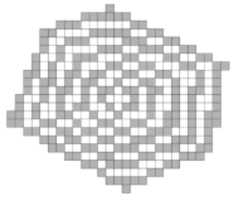
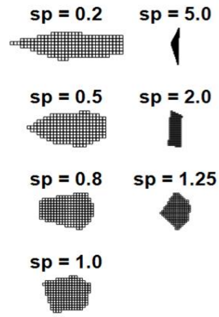
# 3. Methodology

## *3.1 Random spatial sampling of rainfall data*

In order to explore and generalise the spatial patterns and characteristics of rainfall extremes of grid-based datasets, a toolbox of Spatial Random Sampling for Grid-based Data Analysis (SRS-GDA) was developed by Wang and Xuan (2020) to generate arbitrary regions of interest (ROI) with predefined or non-predefined spatial variables i.e., location, size and shape with dominant orientation, from any grid-based dataset automatically. In this study we applied SRS-GDA toolbox for obtaining a large number of ROIs with different location, size and shape to reveal the spatial patterns of annual maximum daily rainfall (AMDR) from two grid-based datasets with different resolutions in GB and Australia (AU). The details of these ROIs are shown in Table 1.

Table 1 ROIs for analysing the spatial patterns in GB and AU

| Sampling areas | Spatial index | Changing with location | Changing with size (Ten in one group) | Changing with shape (Seven in one group) |
|---|---|---|---|---|
| | | Coordinates of centroid of ROI (x- and y-index) | ROI size | sp[a] |
| GB | ROIs of 1x1 km grid |  |  |  |
| | Size | 500 km$^2$ | from 10 to 1025 km$^2$ | 500 km$^2$ each ROI |
| | No. | 88 in total | 81×10=810 in total | 74×7=518 in total |

| | | | | |
|---|---|---|---|---|
| AU | ROIs of 5x5 km grid |  |  |  |
| | Size | 500 km$^2$ | from 125 to 9900 km$^2$ | 5000 km$^2$ each ROI |
| | No. | 679 in total | 627×10=6270 in total | 378×7=2646 in total |

[a]sp defines the ratio of north-south dimension over west-east dimension of the sample.

## *3.2 Bayesian Markov-Chain Monte-Carlo for non-stationary GEV distribution*

The AMDR of each ROI is fitted by a non-stationary GEV distribution (Equation 1) assuming that the location and scale parameters change over elapsing time.

$$F_t\,(x_t;\sigma_t,\mu_t,\xi) = exp[-(1+\xi(\frac{x_t-\mu_t}{\sigma_t}))^{-1/\xi}] \tag{1}$$

where $F_t$ is non-stationary GEV cumulative distribution function of three parameters while the subscript *t* indicates that the location-scale parameters are considered as a linear monotonic varying with time (Equation 2). The shape parameter $\xi$ indicates the type of GEV: Gumbel ($\xi = 0$), Fréchet ($\xi > 0$), Weibull ($\xi < 0$), which is constant under this assumption.

$$\begin{cases}\mu_t = \mu_0 + \mu_1 \times t \\ \sigma_t = \sigma_0 + \sigma_1 \times t\end{cases} \tag{2}$$

Therefore, the non-stationary GEV model has five parameters $\theta = \{\sigma_t = (\sigma_0,\sigma_1), \mu_t = (\mu_0,\mu_1)$ and $\xi\}$. A Bayesian inference (Coles and Tawn, 1996) is introduced to estimate the parameters and starts with the prior knowledge of parameters who are estimated by stationary GEV and Equation (3) shows the transformation from prior distribution (uniform distribution with prior knowledge) to posterior distribution by multiplying the likelihood of time series.

$$p(\boldsymbol{\theta}|x,t) \propto p(x|\boldsymbol{\theta},t) \times p(\boldsymbol{\theta}|t) = \prod_{t=1}^{n=113 \text{ or } 129} p(x_t|\boldsymbol{\theta}_t,t) \times p(\boldsymbol{\theta}|t), \tag{3}$$

where $p(x|\theta,t) \propto L(x;\theta,t)$ and $p(\theta|t)$ is the prior probability distribution of parameters. As it cannot be solved analytically, a numerical iteration procedure is carried out by Markov-Chain Monte Carlo (MCMC) simulation method, which also can be used to reveal the uncertainty of nonstationary GEV model with linear assumption.

### *3.3 Quantification on spatial variants of rainfall extremes*

According to the analysis result on location-scale parameters in GB and AU which show a clear spatial pattern from west to east (details in Section 4.3.1), the meridional averages of these two parameters are calculated over the ROIs located along the same x-index. Several linear models (LM) were selected to quantify THE spatial variations on meridional average parameters such as Polynomial, Gaussian and Power models. Among them, the quadratic polynomial model was tested to be most proper one to describe the trend of spatial changes of location-scale parameters in GB and AU by comparing $R^2$ and root mean square error (RMSE) which are widely-used indicators for testing and selecting LMs. $R^2$ is the coefficient of determination for testing how the data are fitted by the assumed linear model and normally ranges from 0 to 1. An $R^2$ indicates that the model perfectly fit the data. RMSE is used measure of the difference between data (the parameters) and those predicted by linear models. The smaller the RMSE is, the better fitted can be concluded.

The LM for quantifying the meridional average parameters changing in EW direction are:

$$f(\mu \ or \ \sigma) = p1 \times x^2 + p2 \times x + p3 \tag{4}$$

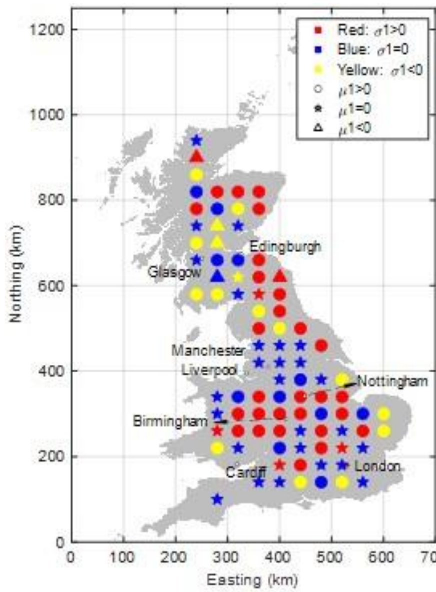
where $x$ is the x-index of ROI and it increases eastward. The three coefficients are estimated to compare among different size and shape in two countries.

## 4. Results and discussion

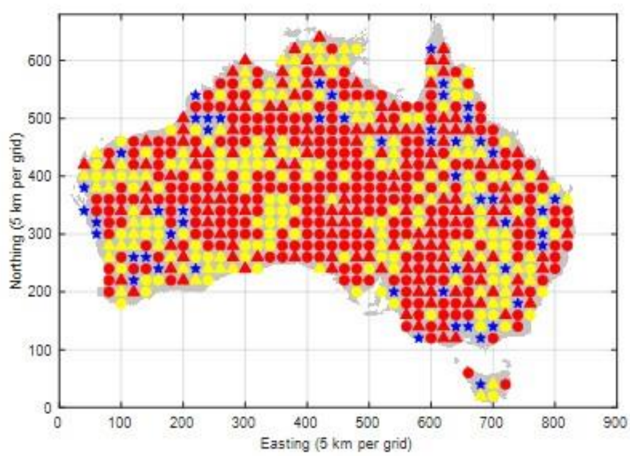
### *4.1 Nonstationary GEV modelling of the AMDR in GB and AU*

Figure 1 presents the parameters of nonstationary GEV modelling in GB and AU. The results show some similarities between two countries: 1) The majority of ROIs have an increasing $\sigma$ (marked as red) which indicates a stretch on GEV probability distribution curves where the occurrence probability of rainfall extremes is increased; 2) More than half of ROIs in GB and the majority in AU show an increased $\mu$ (marked as circle) which indicates an increased rainfall value occurring most frequently over years; 3) The parameter $\mu$ of more than half of ROIs are further accompanied by an increase on $\sigma$ (marked as red circle), which leads to an overall dropping of the return level of rainfall extremes from 1-in-50-year to 1-in-20 year over the study period. It means that not only do the most frequent events become more extreme, the extreme events also become more frequent; 4) Around 3% ROIs in GB and around 10% of ROIs in AU show a decrease on ;location-scale parameters (marked as yellow triangle) who are estimated to become drier. 5) Only a little ROIs keep a stationary GEV distribution (marked as blue star) estimated by our model which may inspire a re-consideration of the current practice of designing storms.

Figure 1 Spatial distribution the changes of location-scale parameters of nonstationary GEV model in: (a) GB and (b) AU




(a)

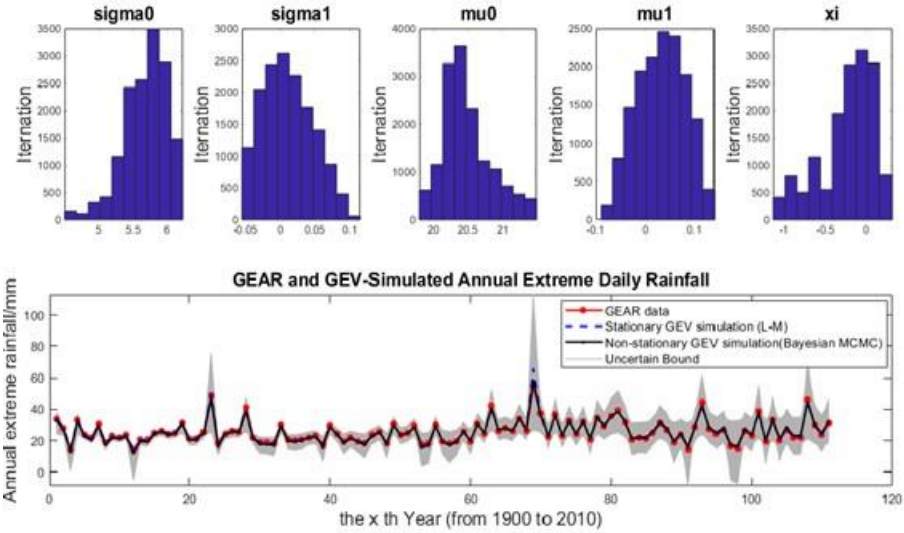



(b)

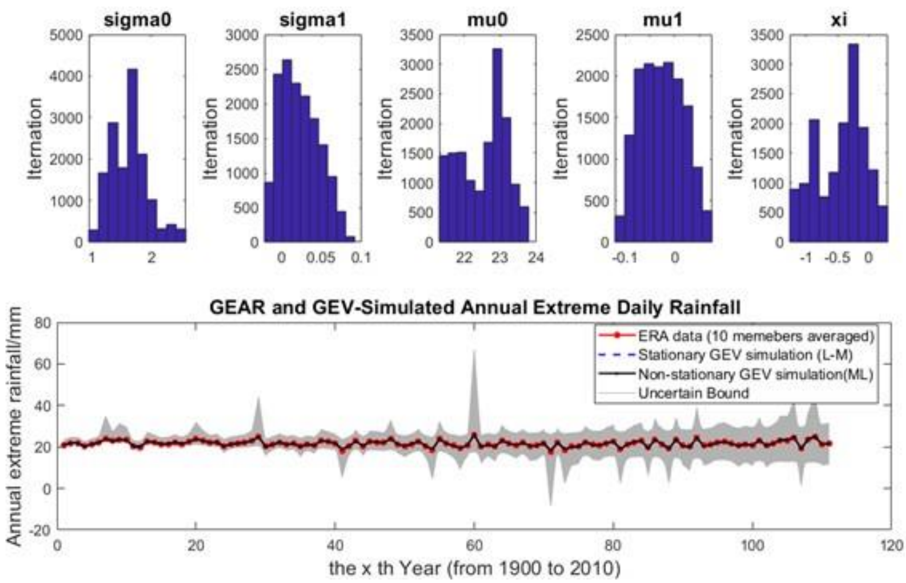
*4.2 Performance of the UK climate projection*

Figure 2 presents the five parameter distributions for 15000 iterations and the comparison between AMDR simulation results by observed GEAR data (Figure 2(a)) and ERA20CM dataset (Figure 2(b)) of middle England from the year of 1900 to 2010.

Figure 2 Stationary and non-stationary GEV simulation with parameter distribution in England with: (a) GEAR data and (b) ERA20CM data




(a)




(b)

The parameters estimated by both stationary and nonstationary GEV models are shown in Table 2. The results can be concluded as:

1) Non-stationary GEV simulation with the assumption of location-scale parameters linearly changing over time fits better with GEAR data but worse with climate projection data.

2) The simulation on the average of annual extreme rainfall in England by ensemble climate modelling is relative more accurate corresponding to the in-situ observation but the difference between the simulated rainfall by GEV and the average annual extreme rainfall (reflects on $\sigma$) is much smaller than the difference generated from GEV with observed GEAR data. And in the analysis of GEAR data, the occurrence probability of extreme rainfall slightly dropped but the magnitude of extreme rainfall was kept increasing in the period of 111 years.

Table 2 GEV distribution analysis on the AMDR with GEAR and ERA20CM data of middle England (Wang and Xuan, 2019).

| GEV Model Simulation | $\sigma_0(\sigma)$ | $\sigma_1$ | $\mu_0(\mu)$ | $\mu_1$ | $\xi$ | RMSE (mm) |
|---|---|---|---|---|---|---|
| Stationary (GEAR) | (5.25) | - | (22.41) | - | 0.053 | 1.2 |
| Non-stationary (GEAR) | 5.78 | -0.003 | 21.08 | 0.024 | -0.025 | 1.0 |
| Stationary (ERA20CM) | (1.38) | - | (21.05) | - | -0.242 | 0.2 |
| Non-stationary (ERA20CM) | 1.24 | 0.004 | 21.49 | -0.006 | -0.296 | 0.3 |

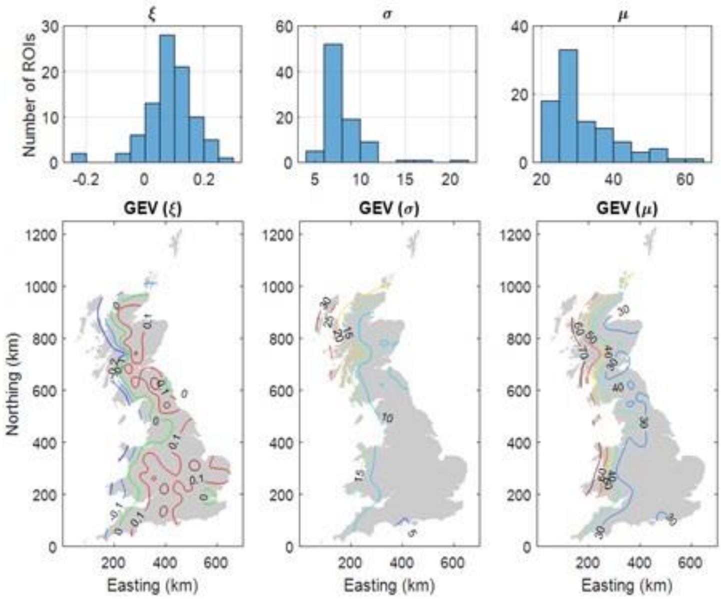
### *4.3 Spatial variance of location-scale parameters in GB and AU*

*4.3.1 Variance with change of location:* The spatial patterns of extremes reflected by the GEV model can be concluded into the following aspects observed in Figure 3:

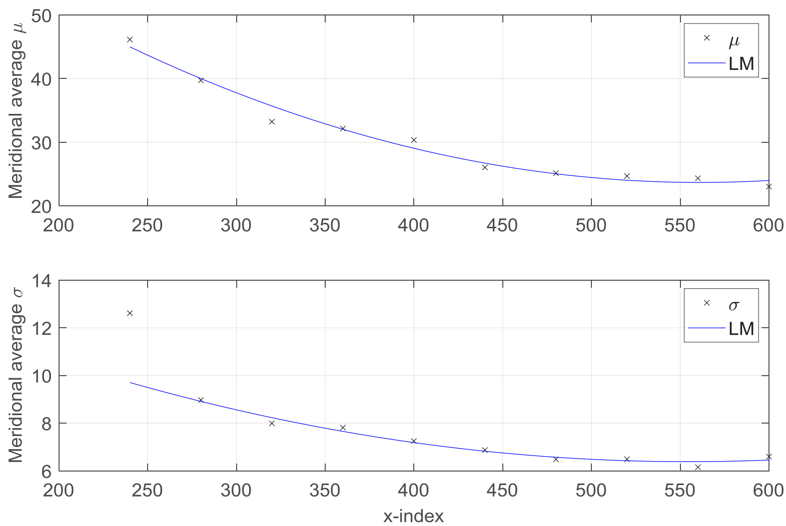
1) Majority of ROIs favour Weibull distribution ($\xi > 0$) in both GB and AU;
2) $\sigma$ and $\mu$ present a similar spatial pattern, and a higher $\mu$ is usually accompanied by a higher $\sigma$.

In GB for the same y-index (latitude), the location-scale parameter of the western region, especially in the coastal area, are much greater than those in eastern region with a gradient ascent and such ascent mainly appears in western region. However there is no remarkable changing trend at the same x-index (longitude) in the north-south direction and the only difference can be observed that $\sigma$ and $\mu$ of the ROIs in Scotland are higher than Wales and England is the lowest.

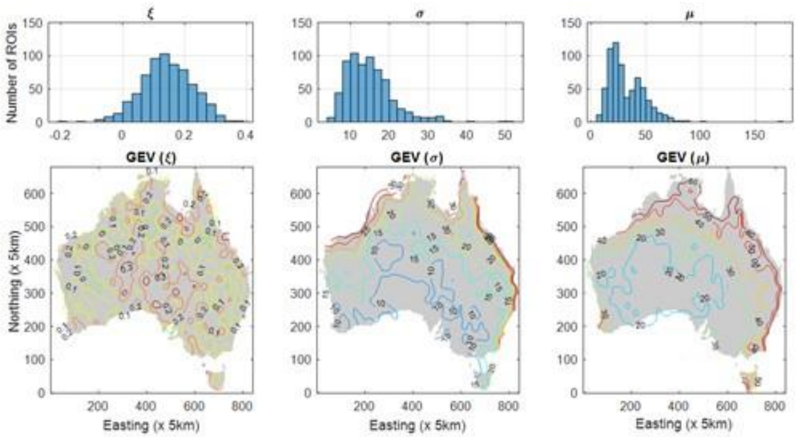
Figure 3 Spatial variation of three parameters of GEV distribution and the quantifications on location-scale parameters changing from west to east (an increasing x-index) in: (a and b) GB and (c and d) AU



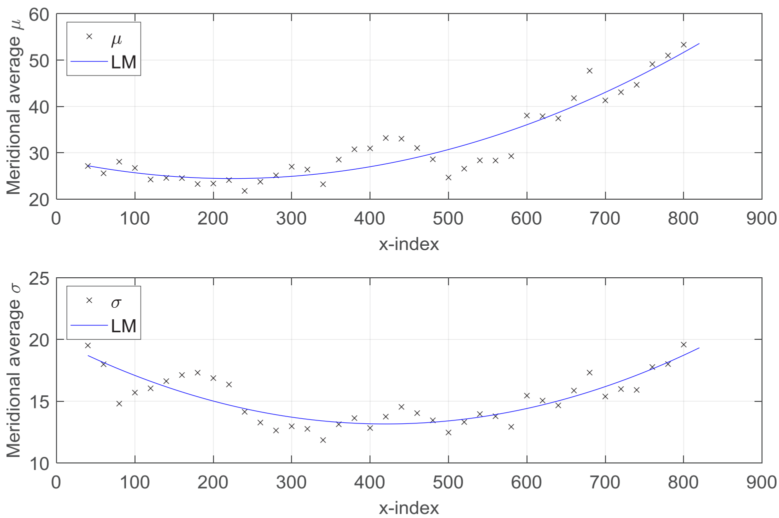

(a)




(b)




(c)



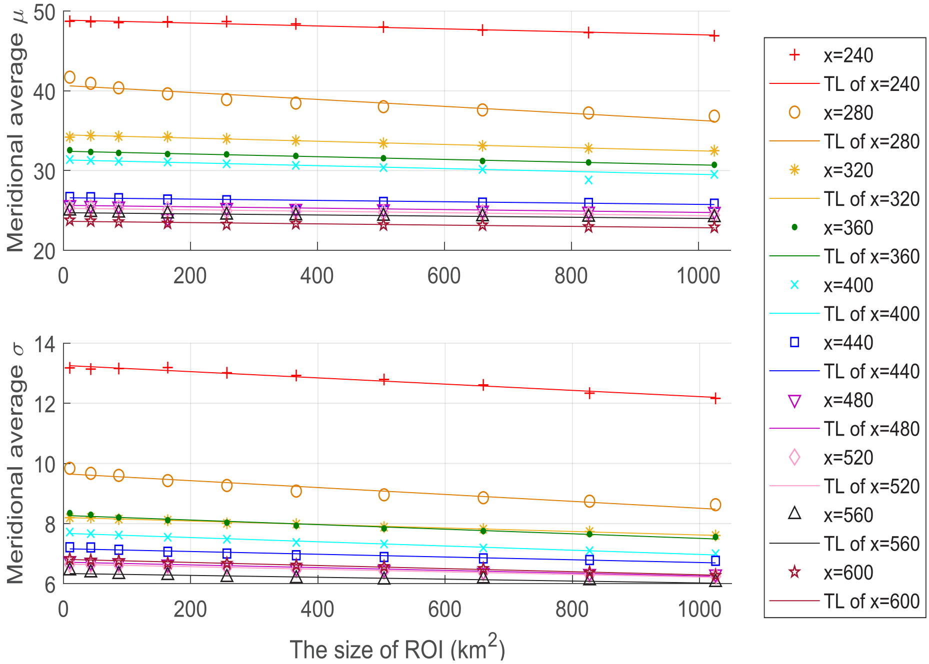

(d)

The location-scale parameter in AU can be clearly observed to have an increase gradient from the south/middle zone to the northeast coastal region and this spatial pattern can be seen as a series of concentric circles with gradient increasing rapidly when they reach the shore of Northeast.

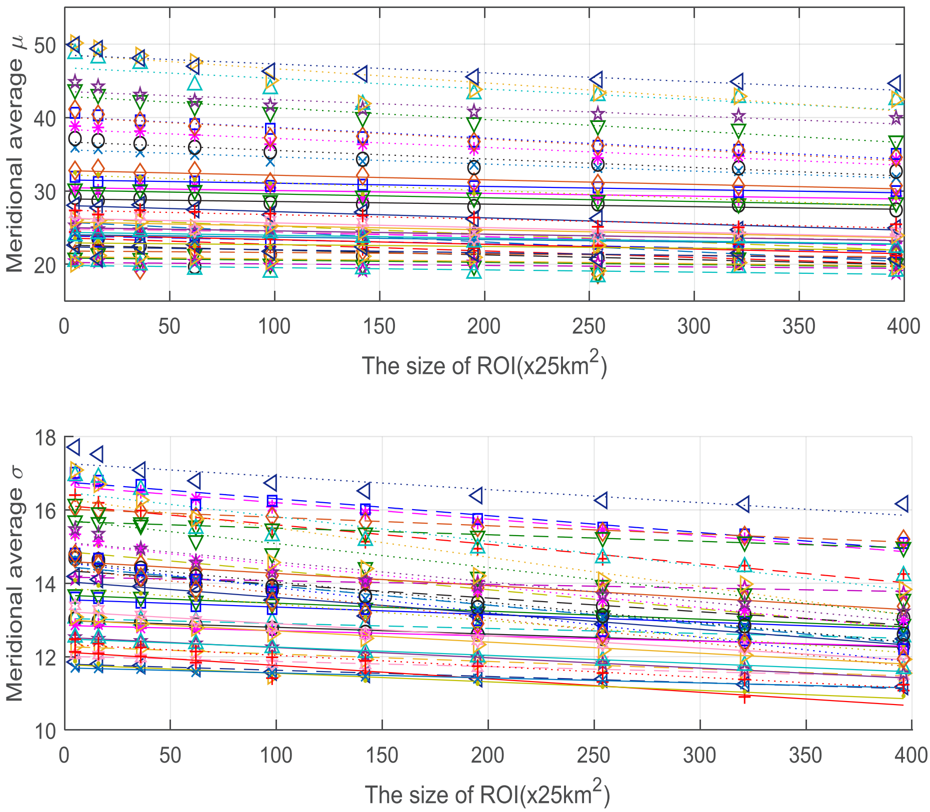
*4.3.2 Variance with the change of size:* Figure 4 shows the meridional averaged location-scale parameters changing with the ROIs' size by a series of fitted trend lines in GB (Figure 4(a)) and AU (Figure 4(b)). It can be observed that regardless of the effect of ROI location, all ROIs in GB and AU show a declining trend of both location and scale parameters with proportionally enlarging the ROI's size.

Although both GB and AU follow generally the descending trend in that "location-scale parameters decrease when the ROI size is increased", the degrees of such reduction are affected by geographic locations. In GB, the negative gradient of trend line for both two parameters of ROIs located in the west are greater than east. In physically with the increase of ROI size, the AMDR with highest frequency is decreased and the occurrence probability of extreme events are also decreased. It can be attributed to the heterogeneity of gridded-based extreme rainfall in western GB which is more significant than that in the east. However, in the east, the grid-based rainfall is more homogeneous and relatively small so increasing size does not affect as much as it does to the parameters of distribution in the west. In AU this negative gradient of ROIs located in the centre of desert zone (x-index from 280km to 360km) is relatively small as enlarging the size of these ROIs only includes more grids of same low AMDR which will not affect the areal AMDR's distribution. However, the GEV distribution of ROIs located in the eastern region and the coastal areas have more negative effects from the increased size.

Figure 4 Location-scale parameters changing with the ROIs' size in: (a) GB and (b) AU where different colours indicate different locations.




(a)

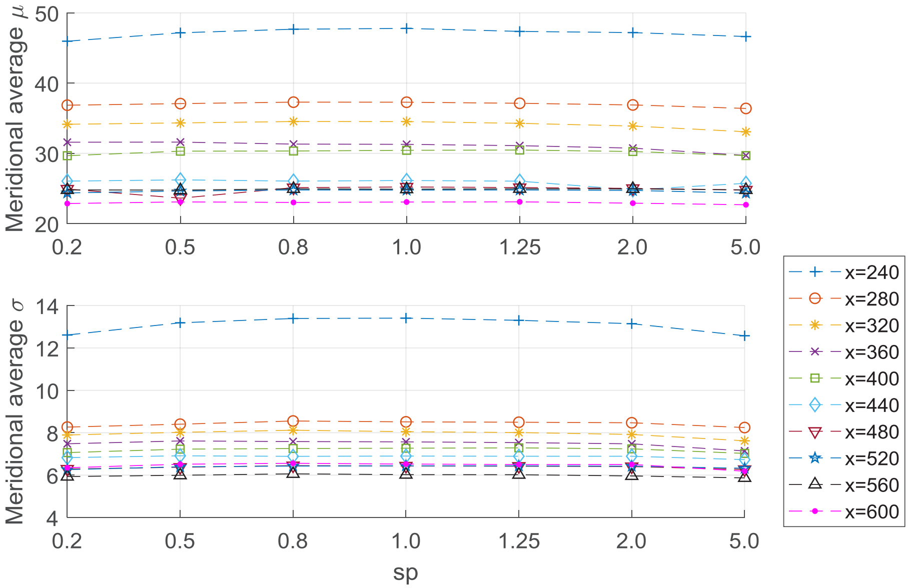



(b)

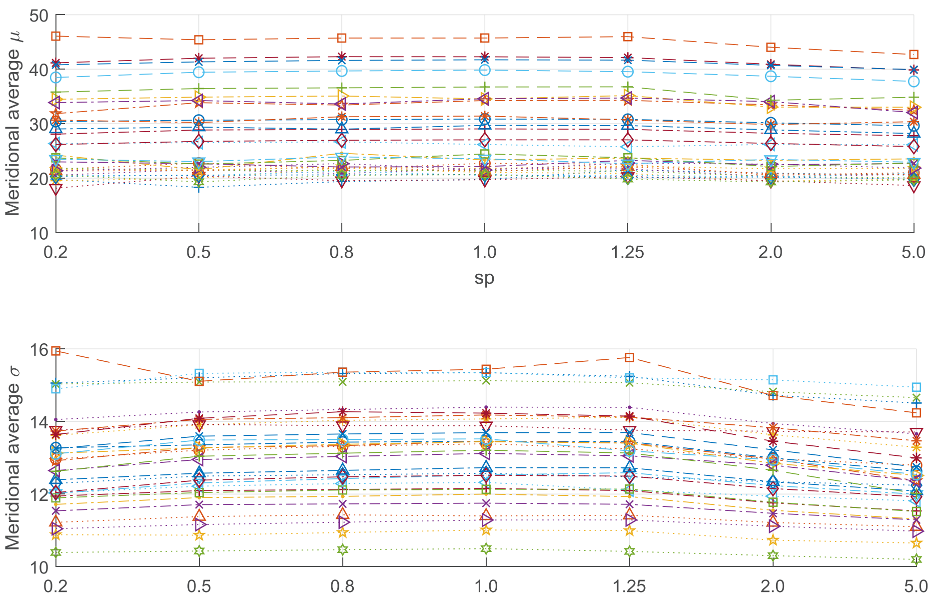
*4.3.3 Variation with the change of shape:* A symmetric pattern of the meridional averaged location-scale parameters around *sp*=1.0 with regards to the sample shape in GB and AU is shown Figure 5. In GB, it implies that the ROIs with slight elongation in north-south direction are expected to have a higher amount of rainfall than those spread more in east-west direction at given frequency/return period. For samples with same size and location, there is a remarkable difference of areal averaged rainfall between more elongated (e.g. *sp* = 0.2 or 5.0) and rounded shape (e.g. *sp* = 1.0) which can be attributed to heterogeneity of the grid rainfall distribution that cannot compensate to the areal average.

In AU, more than 60% of the ROI groups show a higher amount of AMDR of samples with east-west orientation than those with north-south orientation. And around 73% of sample group show the AMDR in the more elongated sample is lower than that in more rounded shape.

Figure 5 Location-scale parameters changing with the ROIs' shape in: (a) GB and (b) AU where different colours indicate different locations.




(a)




(b)

## 5. Conclusion

This paper gives a brief account of the latest efforts in modelling hydroclimatic extremes in the context of climate change. It focuses on the nonstationary nature of the extreme events as well as their spatial dependency including the reliance on the sampling strategy. A preliminary study on how climate projection can capture the large variation of the hydroclimatic variable is also presented. Results show that: (1) non-stationary modelling is preferred to capture the time varying distribution of rainfall extremes; (2) spatial distribution of the rainfall extremes is highly heterogeneous but demonstrates coherent patterns of the parameterisation over the characteristics of the sampling area; (3) the current level of climate model projections are able to represent the average time-varying trend of the extremes but do not yet come to closely reveal the overall variations of the extremes. It should be noted that the study still focuses on very limited geographical areas and further study/data will need to be carried out to refine the findings. In addition, it is also expected that merging the natural variance of observed hydroclimatic variables with model projections should be a promising way forward and certainly warrants more efforts in this direction.

## 6. Acknowledgement

The authors would like to thank the support provided by the OICWE2020 organisational committee. The co-author H. Wang is supported by the PhD grant offered by the China Scholarship Council. In addition, Xuan's contribution is partly supported by the HEFCW GCRF grant (REF: RIG1032-114) offered by the Swansea University. The two datasets GEAR and ADAM are made freely available to public domain. All supports are gratefully acknowledged.

## References

Assani A and Guerfi N (2017) Analysis of the Joint Link between Extreme Temperatures, Precipitation and Climate Indices in Winter in the Three Hydroclimate Regions of Southern Quebec. *Atmosphere* 8(4): 75.

Coles SG and Tawn JA (1996) A Bayesian analysis of extreme rainfall data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 45: 463–478, doi: 10.2307/2986068.

Hall J et al. (2014) Understanding flood regime changes in Europe: A state of the art assessment. *Hydrology and Earth System Sciences* 18: 2735–2772.

Herring SC et al. (2018) Explaining extreme events of 2016 from a climate perspective. *Bulletin of the American Meteorological Society* 99(1): S1–S157.

Hersbach H et al. (2015) ERA-20CM: a twentieth-century atmospheric model ensemble. *Quarterly Journal of the Royal Meteorological Society* 141(691): 2350–2375.

Jones DA et al. (2009) High-quality spatial climate data-sets for Australia. *Australian Meteorological and Oceanographic Journal* 58(4): 233.

Lazoglou G and Anagnostopoulou C (2017) An overview of statistical methods for studying the extreme rainfalls in Mediterranean. *In Multidisciplinary Digital Publishing Institute Proceedings* 1(5): 681.

Mannshardt-Shamseldin EC et al. (2010) Downscaling extremes: A comparison of extreme value distributions in point-source and gridded precipitation data. *The Annals of Applied Statistics* 4(1): 484–502.

Merz B et al. (2014) Floods and climate: emerging perspectives for flood risk assessment and management. *Natural Hazards and Earth System Sciences* 14(7): 1921–1942.

Shukla RK et al. (2012) On the proficient use of GEV distribution: a case study of subtropical monsoon region in India. *Annals Computer Science Series* 8(1): 81–92.

Sillmann J et al. (2013) Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *Journal of Geophysical Research: Atmospheres* 118(6): 2473–2493.

Tanguy M et al. (2016) *Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890–2015) [CEH-GEAR]*. NERC Environmental Information Data Centre, Swindon, UK.

Wang H and Xuan Y (2019) A Multi-scale Analysis of the Correlation in Representing Flood-Inducing Extreme Precipitations from Observation and Climate Model Simulation. *Proceedings of the 4th IMA International Conference on Flood Risk*. Swansea, UK.

Wang H and Xuan Y (2020) SRS-GDA: A spatial random sampling toolbox for grid-based hydro-climatic data analysis in environmental change studies. *Environmental Modelling & Software* 124: 104598.

Yoon S et al. (2015) Spatial Modelling of Extreme Rainfall in Northeast Thailand. *Procedia Environmental Sciences* 26: 45–48.