Supporting Information for

An Area-orientated Analysis of the Temporal Variation of Extreme Daily Rainfall in Great Britain and Australia

Han Wang [1,2,3], and Yunqing Xuan [3,*]

[1] China Institute of Water Resources and Hydropower Research, Beijing 100038, China
[2] Research Center on Flood & Drought Disaster Reduction of the Ministry of Water Resources, Beijing 100038, China
[3] Department of Civil Engineering, Swansea University Bay Campus, Fabian Way, Swansea SA1 8EN, UK
* Correspondence: (y.xuan@swansea.ac.uk)

Contents of this file

Text S1 to S3
Figures S1
Tables S1 to S2

Text S1. Maximum Likelihood (ML) and Bayesian Markov-Chain Monte-Carlo (BMCMC) Methods

- **The ML method**

The ML method (Myung, 2003) is built upon the likelihood function of the occurrence of AMDR, which is the product of the probability density function of NS-GEV distribution:

$$L(x_t;\boldsymbol{\theta}) = \prod_{t=t_0}^{t_{0+k}} f(x_t;\boldsymbol{\theta}) = \left(\frac{1}{\sigma_t}\left(1+\xi\left(\frac{x_t-\mu_t}{\sigma_t}\right)\right)^{-\frac{1}{\xi}-1} \exp\left(-\left(1+\xi\left(\frac{x_t-\mu_t}{\sigma_t}\right)\right)^{-1/\xi}\right)\right)^k$$

The set of the parameters $\boldsymbol{\theta}$ can then be estimated by maximizing the likelihood function as $\frac{\partial L(x_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0$. It usually cannot be solved analytically, thus an iterative algorithm was employed to find the minimizer of $\frac{\partial L(x_t;\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ starting with an initial guess which is based on the value of parameters estimated in stationary model S.

- **The B-MCMC method**

The transformation (see Eq. S1) from the prior distribution to posterior distribution is to multiply by its likelihood (Rasmussen and Ghahramani, 2003).

$$p(\boldsymbol{\theta}|x,t) \propto p(x|\boldsymbol{\theta},t) \times p(\boldsymbol{\theta}|t) = \prod_{t=t_0}^{t_{0+N}} p(x_t|\boldsymbol{\theta}_t,t) \times p(\boldsymbol{\theta}|t) \quad \text{(S1)}$$

where $p(x|\boldsymbol{\theta},t) \propto L(x;\boldsymbol{\theta},t)$ is the likelihood function and $p(\boldsymbol{\theta}|t)$ is the prior probability distribution of the parameters $\boldsymbol{\theta}$; t indicates the time from t_0 to t_{0+N}.

Numerical iterations for exploring the posterior distribution are carried out by using the MCMC simulation. The essence of the MCMC algorithm is to generate a trial moving from the current state of the Markov Chain with a prior probability of parameters $p(\boldsymbol{\theta}|t)$ to a next proposed state with a prior probability of the proposed parameters $p(\boldsymbol{\theta}'|t)$. In this study, to make full use of the knowledge, the estimated parameters of the stationary model were used to define the initial prior values of the nonstationary parameters which are drawn from uniform distributions using Latin Hypercube Sampling (LHS). Numerical iterations for exploring the posterior distribution are carried out by using the MCMC simulation with Metropolis within Gibbs sampling. The Metropolis ratio is calculated to accept or reject proposal status and the convergence of simulation is monitored by Gelman-Rubin diagnostic (Gelman & Rubin, 1992).

This algorithm firstly starts as a random search over the entire prior distribution ($p(\boldsymbol{\theta}|t)$ of D parameters using the LHS method then d samples are randomly assigned to N Markov chains and the sample with highest likelihood value will be selected as the starting point for each chain. To diversity the probability of the jumping direction, we broadly followed Sadegh, Ragno, and AghaKouchak (2017) to use two approaches to update the chain: some chains (N1) follow the Adaptive Metropolis (AM) approach which is effective for searching direction at the early stage of MCMC and the rest (N-N1) follow the Differential evolution (DE) approach which has a stronger potential in converging to the target distribution. The details are shown below (Sadegh et al., 2017):

1) For each chain, randomly select d samples from D parameter spaces with Gibbs sampling (Gilks, Best, & Tan, 1995).
2) For N1 chains, propose a new state S_{t+1} with a proposed set of parameters $\boldsymbol{\theta}'$ by using AM approach, i.e., $S_{t+1}(\boldsymbol{\theta}') = S_t(\boldsymbol{\theta}') + (1-\beta)N(0_{\boldsymbol{\theta}'}, \gamma_1^2 \sum_{\boldsymbol{\theta}'}) + \beta N(0_{\boldsymbol{\theta}'}, \gamma_2^2 I_{\boldsymbol{\theta}'})$ where $\sum_{\boldsymbol{\theta}'}$ is the covariance matrix of $\boldsymbol{\theta}'$.
3) For the rest N-N1 chains, the new state is $S_{t+1}(\boldsymbol{\theta}') = S_t(\boldsymbol{\theta}') + \gamma_3(S_{r_2} - S_{r_1}) + e$.
   where $\sum_{\boldsymbol{\theta}'}$ is the covariance matrix of $\boldsymbol{\theta}'$ and β is a random number in the range of 0~0.1; γ indicates the jump factors defined as $\gamma_1$is a number randomly selected from [1.2, 2.2], $\gamma_2 = 2.38/\sqrt{d}$, $\gamma_3 = 0.1/\sqrt{d}$ (Roberts & Rosenthal, 2009) and $\gamma_4 = 2.38/\sqrt{2d}$ (Ter Braak, 2006); and S_{r_1} and S_{r_2} are two samplers drawn from parameter space D just for pre-defining the chain update direction.
4) Compute the Metropolis ratio $\frac{p(x|\boldsymbol{\theta}', t)}{p(x|\boldsymbol{\theta}, t)}$; if $\min\left(1, \frac{p(x|\boldsymbol{\theta}', t)}{p(x|\boldsymbol{\theta}, t)}\right) \geq p^*$, then accept S_{t+1} and update the current chain where p^* is the random number drawn from $\boldsymbol{N}(0,1)$. If not, reject and go back to previous step to re-propose the state.
5) Check whether the iteration convergence or not by Gelman-Rubin convergence diagnostic.

Text S2. Criteria for selecting best fitted models.

In this study, we applied three criteria for selecting the best fitted model: the root mean squared error (RMSE), the Akaike Information Criterion (Eq. S2) and the Bayesian Information Criterion (Eq. S3). Small values of these criteria indicate a better model fit.

$$AIC = -\frac{2}{N} \times LL + 2 \times \frac{k}{N} \tag{S2}$$

$$BIC = -2 \times LL + \log(N) \times k, \tag{S3}$$

where N is the number of data (113 for GB case and 129 for AU case), LL is the log-likelihood of the model on these data and k is the number of parameters (e.g., 3 for stationary model S, 4 for nonstationary model NS1 and 5 for NS3 and NS4).

Text S3. Selection of stationary and nonstationary models and spatial patterns.

Table S1 presents a summary of the number of ROIs of their best-fitted model with different regional sizes and shapes in GB and AU. In GB, there are 1416 ROIs generated by the spatial random sampling for grid-based data analysis (SRS-GDA) toolbox (Wang & Xuan, 2020) while in AU, there are 9575 ROIs.

Table S1 the number of ROIs with their best-fitted model varying with regional sizes and shapes in GB and AU.

GB						AU					
ROI size (km^2)	S	NS1	NS2	NS3	Total	ROI size (km^2)	S	NS1	NS2	NS3	Total
10	32	32	12	5		125	161	300	113	51	
43	35	30	9	7		400	164	315	114	32	
87	27	39	11	4		900	162	312	115	36	
<=100	94 (38.7%)	101 (41.6%)	32 (13.2%)	16 (6.5%)	243	500	173	341	124	41	
164	32	35	10	4		<=1,000	660 (25.8%)	1268 (49.7%)	466 (18.2%)	160 (6.3%)	2554
257	30	32	11	8		1,550	159	312	107	47	
366	25	43	9	4		2,450	155	309	111	50	
500	28	43	10	7		3,550	144	305	115	61	
<=500	115 (34.7%)	153 (46.2%)	40 (12.1%)	23 (6.9%)	331	4,875	153	309	112	51	
504	28	41	7	5		<=5,000	611 (24.4%)	1235 (49.4%)	445 (17.8%)	209 (8.4%)	2500
660	32	36	6	7		6,350	150	307	107	61	
827	30	38	5	8		8,025	148	326	104	47	
1,025	30	38	8	5		9,900	137	325	107	56	

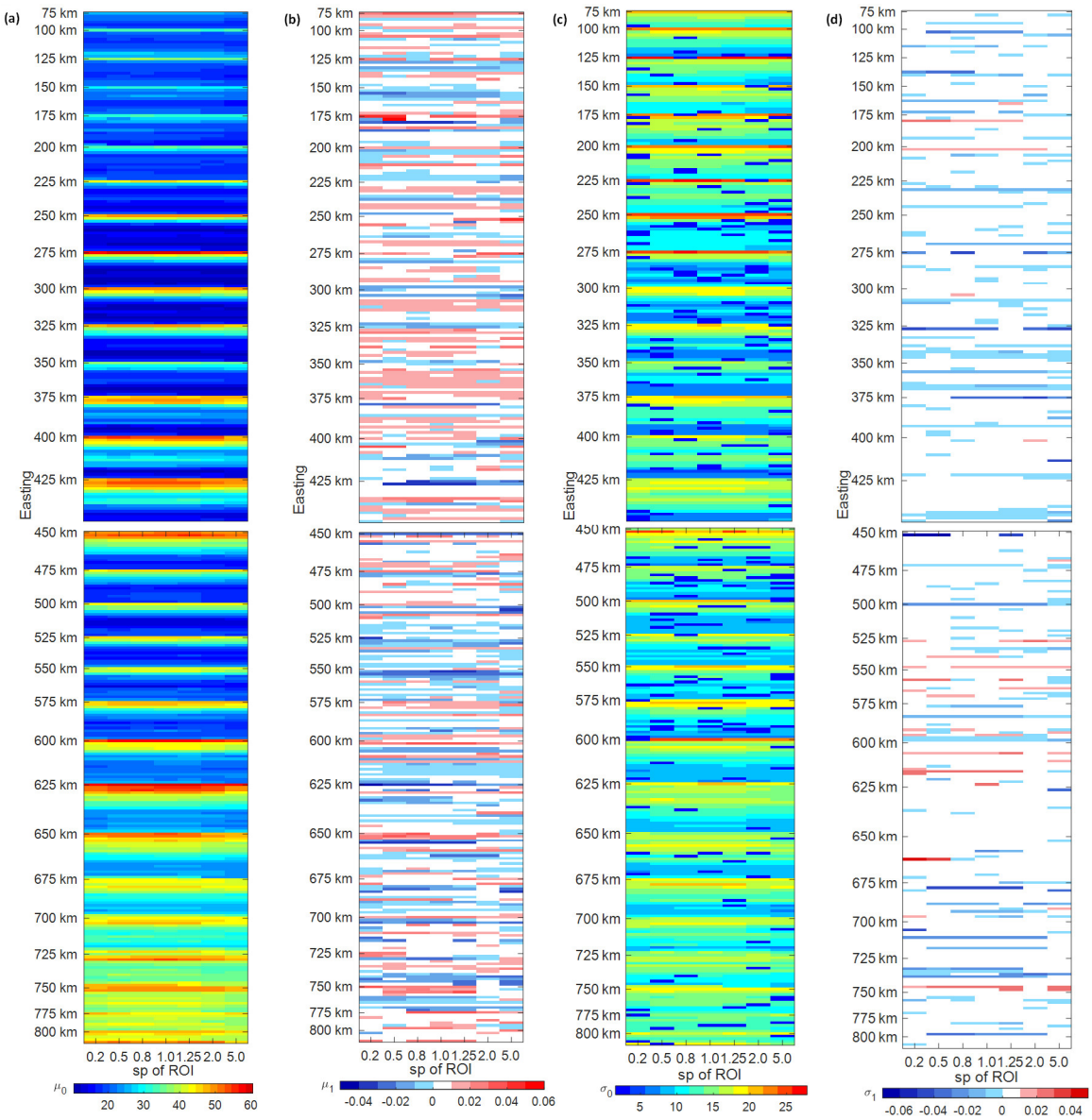
to 1,025	120 (37.0%)	153 (47.2%)	26 (8.0%)	25 (7.8%)	324	<=10,000	435 (23.2%)	958 (51.1%)	318 (17.0%)	164 (8.7%)	1875
Total	**329**	**407**	**98**	**64**	**898**	Total	**1706**	**3461**	**1229**	**533**	**6929**
ROI shape (*sp*)						ROI shape (*sp*)					
0.2	25	31	9	9	74	0.2	83	201	61	33	378
0.5	28	37	7	2	74	0.5	83	209	57	29	378
0.8	21	37	12	4	74	0.8	84	199	66	29	378
1.0	21	34	10	9	74	1.0	89	196	63	30	378
1.25	26	36	7	5	74	1.25	81	205	67	25	378
2.0	22	33	11	8	74	2.0	87	203	56	32	378
5.0	26	30	14	4	74	5.0	77	191	65	45	378
Total	**169 (32.6%)**	**238 (46.0%)**	**70 (13.5%)**	**41 (7.9%)**	**518**	Total	**584 (22.1%)**	**1404 (53.1%)**	**435 (16.4%)**	**223 (8.4%)**	**2646**

Table S2 the number of ROIs where the AMRD is fitted with best-selected GEV types in GB and AU.

GEV type	Gumbel	Fréchet	Reversed Weibull	Total
GB	68	1116	232	**1416**
AU	323	8706	546	**9575**

Out of all ROIs in GB, there are near 80% following the Fréchet distribution and around 16% following the reversed Weibull distribution. In AU, such percentage is even higher, i.e., around 90% ROIs follow the Fréchet distribution and only a very small proportion (3%) follows the Gumbel distribution.

Figure S1. Both baseline and time-varying parameters ($\mu_0, \mu_1, \sigma_0, \sigma_1$) change over the ROI shape indicated by the index of *sp* in AU. The horizontal axis indicates the location index the ROIs and colour bar shows the values of parameters.



Reference

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, *7*(4), 457-472.

Gilks, W. R., Best, N. G., & Tan, K. (1995). Adaptive rejection Metropolis sampling within Gibbs sampling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *44*(4), 455-472.

Myung, I. J. (2003). Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology*, *47*(1), 90-100.

Roberts, G. O., & Rosenthal, J. S. (2009). Examples of adaptive MCMC. *Journal of Computational and Graphical Statistics*, *18*(2), 349-367.

Sadegh, M., Ragno, E., & AghaKouchak, A. (2017). Multivariate C opula A nalysis T oolbox (MvCAT): describing dependence and underlying uncertainty using a B ayesian framework. *Water Resources Research, 53*(6), 5166-5183.

Ter Braak, C. J. (2006). A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. *Statistics and Computing, 16*(3), 239-249.

Wang, H., & 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.