Assessment of trends in hydrological extremes using regional magnification factors

[1]Thomas Rodding Kjeldsen and [2]Ilaria Prosdocimi

[1]Department of Architecture and Civil Engineering, University of Bath, Bath, BA2 7AY, United Kingdom

[2]Dipartimento di Scienze Ambientali, Informatica e Statistica, Università Ca' Foscari Venezia, Italia

Abstract

Detection and attribution of trends in individual at-site series of hydrological extremes is routinely undertaken using simple linear regression-based models. However, the available records are often too short to allow a consistent assessment of trends across different stations in a region. The theoretical developments presented in this paper propose a new method for estimating a regional regression slope parameter across a region, or pooling group, of catchment considered hydrologically similar, and where annual maximum events at different sites are cross-correlated. Assuming annual maximum events to follow a two-parameter log-normal distribution, a series of Monte Carlo simulations demonstrate the ability of the new framework to accurately identify the regional slope, and provide estimates with a reduced sampling variability as compared to the equivalent at-site estimates, thereby enhancing the statistical power of the trend test. This regionally-based trend estimates would allow for a clear characterization of changes across several stations in a region. Finally, the new method is applied to national dataset of annual maximum series of peak flow from 662 gauging sites located across the United Kingdom. The results show that the regional slope estimates are significantly positive ($p<0.05$) consistently in the west and north of the country, while mostly not significant in the east and south. This translate into a corresponding increase in design flood (as measured by regional magnification factors) of up-to 50% for time horizon of 50-years into the future.

1. Introduction

Detection and attribution of trend in series of hydrological and climatological extremes is an important endeavour both as early warning of changes in the environment and to ensure that hydrological design is appropriate via non-stationary extreme value models where distribution parameters are allowed to change as a function of one or more covariates (Salas et al., 2018). Studies investigating the impact of climate and catchment change on large-scale databases of observed records of environmental extremes, such as high and low flow, rainfall or temperature, typically rely on relatively simple at-site methods such as the Mann-Kendall test or linear regression models (Faulkner et al., 2020, Vicente-Serrano et al. (2020), Mohan et al., 2020). However, as discussed by Prosdocimi et al. (2014) and Brady et al. (2019) these tests have a low statistical power when applied to the short annual maximum series (AMS) derived from station measurements typically characterised by the high year-to-year variability of the extremes. Consequently, it is often difficult to discern compelling evidence of consistent regional trends when these tests are applied on a site by site basis (e.g. Vilarini et al., 2009; Archfield et al., 2016; Mediero et al., 2015).

Attempts to provide more consistent regional assessment of trend have been published by, for example, Bloeschl et al. (2019) who smoothed at-site estimates of trend obtained from observed flood records located across Europe to provide more coherent spatial patterns. Another approach is to develop regional trend tests analogue to regional frequency analysis, where the sampling variability of at-site estimators is overcome by trading space for time and forming a regional (or pooled) estimate of trend at a particular site. Examples of this approach include Douglas et al. (2000) who presented a regional version of the Mann-Kendall test, including consideration of the variance inflation caused by the existence of cross-correlation between the observations at different sites. Renard et al. (2008) compared the performance of four different methods for assessing field significance and investigated the existence of trend in regional flood datasets from France. Prosdocimi et al. (2019) presented a spatial hierarchical model allowing an assessment of consistent and significant trend signals in regions of the United Kingdom. These studies exemplify how it is possible to derive regional estimators with a generally lower variance than the corresponding tests based on at-site data only, but of course at the costs of having to evoke more complicated statistical models and estimation techniques. A second issue is the need to impose simplifying model assumptions that might not reflect the data; for example, classifying a region as being

'homogeneous' thereby stating that the hydrological characteristics of all sites in the region are identical. Depending on the assumptions made in the modelling building phase, these assumptions can inflate either bias or variance of the regional estimator.

Vogel et al. (2011) proposed the use of magnification factors to quantify the effect of change in the observed records on the magnitude of design floods over a pre-defined time period. The concept initially relied on the use of a two-parameter log-normal distribution (LN2) and have subsequently been applied by, for example, Prosdocimi et al. (2014) and Zhang et al. (2015). Prosdocimi and Kjeldsen (2020) extended the concept for use with a more general class of extreme value distributions (Kappa, GEV, GLO and Pareto distribution). The concept of a magnification factor is a convenient method for linking trend studies to non-stationary flood frequency analysis, while at the same time also providing an intuitive means of communicating the effects of change on design floods. Despite these obvious advantages from an engineering hydrology perspective, the current applications of magnification factors cited above are still subject to the limitations of at-site trend estimation as discussed above. Furthermore, magnification factors can only be estimated for sites where observed data are available and there is little understanding of how to use these estimated magnification factors to inform/correct the estimated design events for locations at which no measurements are available, i.e. ungauged location (Rosbjerg et al. 2013). The aim of this study is to present a new regional flood frequency model that will allow the concept of magnification factors to be applied using regional (or pooled) information based on an assumption that annual maximum peak flow events follows a LN2 distribution. This can then be extended to the case in which the regional method is used to estimate the frequency of extreme events at ungauged site. Noticeable, the regional trend estimator can be applied more generally to assess the significance of a regional trend coefficients without the need to evoke magnification factors. First, the theoretical background will be presented followed by a case study based on a national database of AMS of peak flow from the United Kingdom.

2. Flood Magnification factors

The concept of a magnification factor was introduced by Vogel et al. (2011) and is based on a description of the annual maximum series (AMS) as an iid sample following a LN2 distribution where the location parameter is itself a linear function of one of more covariates. For a LN2 distribution, the quantile function Q_p associated with an exceedance probability p is defined as

$$Q_p = exp(\mu_y + \sigma_y z_p) \quad (1)$$

where μ_y and σ_y are the mean and standard deviation of the log transformed annual maximum event $y_t = lnQ_t$, and z_p is the p'th quantile of a standard normal distribution. A non-stationary (or change-permitting) version of the quantile function in Eq. (1) can be obtained by imposing a simple linear relationship between the location parameter μ_y and a covariate x as

$$\mu_y = \beta_0 + \beta_1 x \quad (2)$$

where β_0 and β_1 are model parameters which need to be estimated. In many cases (e.g. Vogel et al, 2011, Zadeh et al. 2020) the covariate is chosen to be water-year, i.e. the year in which the annual maximum peak occurred. This relationship can also be cast as a linear regression model linking the log transformed annual maximum series $(Q_t, t = 1, \cdots, n)$ to a covariate $(x_t, t = 1, \cdots, n)$ as:

$$lnQ_t = y_t = \beta_0 + \beta_1 x_t + \varepsilon_t \quad (3)$$

Where the residuals ε_t are assumed normally distributed with mean zero and variance σ^2.

A statistical assessment of the evidence against the no-trend hypothesis can be conducted by testing if the slope β_1 is equal to zero or not. The statistical power of this test is closely linked to the sampling variance of the estimated slope and to the sample size as discussed in details by Prosdocimi et al. (2014) in the context of AMS of peak flow.

Finally, a non-dimensional magnification factor, M, is derived by combining Eq. (1) and (2) and considering the ratio between a design event at time $x + \Delta x$ and at the initial time x

$$M = \frac{exp\left(\beta_0 + \beta_1[x + \Delta x] + \sigma_y z_p\right)}{exp\left(\beta_0 + \beta_1 x + \sigma_y z_p\right)} = exp(\beta_1 \Delta x) \quad (4)$$

Notably, for the LN2 distribution the magnification factor depends only on the slope of the regression line and the chosen value of Δx. For example, Vogel et al. (2011) fixed Δx at 10-years to define a 'decadal magnification factor' communicating the expected change in a design flood of any rarity over a decade. Vogel et al. (2011), Prosdocimi et al. (2014) and Zhang et al. (2015) all demonstrated application of the magnification factor to large-scale datasets by considering each site in isolation. In the next section we will develop an analytical framework allowing a regional (or pooled) estimate of the slope to be derived based on considerations similar to those commonly adopted in regional frequency analysis, such as the index flood (Hosking and Wallis, 1997).

3. A regional model for the regression slope

Consider a region, or pooling group, consisting of $i = 1, \cdots, N$ sites. At a site i the relationship between the log-transformed annual maximum peak flow, $lnQ_{i,t}$, and the covariate $x_{i,t}$ is described by a linear relationship as:

$$lnQ_{i,t} = y_{i,t} = \beta_{i,0} + \beta_1 x_{i,t} + \varepsilon_{i,t} \quad (5)$$

where the residuals are assumed normal distributed with mean zero and a site-specific standard deviation σ_i. Notice that although t is used to index the observation of year t the covariate x_t could potentially be any physically relevant variable, and does not necessarily need to be water-year. For a pair of stations i and j, and at observations t and k, the cross-correlation between the residuals is given as:

$$cor\left(\varepsilon_{i,t}, \varepsilon_{j,k}\right) = \begin{cases} 1 & i = j \quad t = k \\ 0 & i = j \quad t \neq k \\ \rho & i \neq j \quad t = k \\ 0 & i \neq j \quad t \neq k \end{cases} \quad (6)$$

It is assumed here that the slope β_1 is constant across all sites in the region. This can be interpreted as the region, or pooling group, experiencing a consistent level of change simultaneously at every catchment in the region. This modelling assumption is similar to one of a, similarly to a fixed slope is a fixed and random effects model. A possible future model extension could consider the slope as a random effect, but this was not pursued further here. In contrast, the intercept $\beta_{i,0}$ would be expected to vary from site to site representing the general magnitude of events from each catchment. This variation in intercept could arise from a region consisting of a mix of small and large catchments. As the magnification factor defined in Eq. (4) considers only the slope parameter, this study will not focus further on characterisation and estimation of the intercepts. However, a model extension to allow prediction of non-stationary design floods in ungauged catchments would need a procedure for linking the intercept to the target catchment.

For each individual site in the region the intercept and slope can be estimated using simple ordinary least square as:

$$\hat{\beta}_{i,1} = \frac{\sum_{t=1}^{n_i}(x_t - \overline{x})y_t}{\sum_{t=1}^{n_i}(x_t - \overline{x})^2} \qquad (7)$$

$$\hat{\beta}_{i,0} = \overline{y} - \hat{\beta}_1 \overline{x}$$

Analogue to the index flood method, a regional estimate of the slope β_1^R can be derived as a weighted average of the at-site slope estimate from each of the N sites in the region. As the basic assumption is that the slope is identical at each site, the weighted average is, by definition, unbiased and has a lower sampling variance than the individual estimates. As discussed by Douglas et al. (2000) in the context of trend testing, the existence of cross-correlation between the data at different sites will inflate the variance of the regional estimate. It is therefore necessary to carefully consider here how cross-correlation will affect the variance of the regional slope estimator.

The weighted average of the at-site estimates, β_1^R, of slope is defined in a matrix form as

$$\hat{\beta}_1^R = \sum_{i=1}^{N} \omega_i \hat{\beta}_{i,1} = \boldsymbol{\omega}^T \hat{\boldsymbol{\beta}}_1 \qquad (8)$$

where ω is a $N \times 1$ vector of weights and $\hat{\boldsymbol{\beta}}_{\mathbf{1}}$ is a $N \times 1$ vector whose elements are the at-site estimates of slope from each site according to Eq. (7). The weights $\boldsymbol{\omega}$ are chosen to minimise the variance of the regional estimator, which can be derived directly from Eq, (8) as

$$var\left(\hat{\beta}_1^R\right) = \boldsymbol{\omega}^{\boldsymbol{T}}\boldsymbol{\Sigma}\boldsymbol{\omega} \tag{9}$$

where $\boldsymbol{\Sigma}$ is the $N \times N$ covariance matrix of the at-site slope estimates, the element of which are derived in Appendix A and defined for two sites i and j as:

$$cov\left(\hat{\beta}_{i,1}, \hat{\beta}_{j,1}\right) = \sigma_i \sigma_j \rho \frac{SS_{x,\delta}}{SS_{x,i} SS_{x,j}} \tag{10}$$

where σ_i and σ_j are the standard deviation of the residuals at the two sites, and ρ is the correlation coefficient between the residuals at the two sites. The sum of squares over the covariate at each site is denoted $SS_{x,i}$ and $SS_{x,j}$ while the sum of squares over the covariates covering the overlapping periods at both sites is denoted $SS_{x,\delta}$. For the diagonal elements of the covariance matrix ($i = j$), the covariance expression in Eq. (10) reduces to the variance of the at-site slope estimate as in this case: $\sigma_i = \sigma_j$, $\rho = 1$, and $SS_{x,i} = SS_{x,j} = SS_{x,\delta}$.

Using Lagrange multiplier for constraint optimisation (Kjeldsen and Prosdocimi, 2015), the set of weights minimising the variance in Eq. (9) can be derived as:

$$\boldsymbol{\omega} = \boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{i}\left(\boldsymbol{i}^{\boldsymbol{T}}\boldsymbol{\Sigma}^{-\mathbf{1}}\boldsymbol{i}\right)^{-\mathbf{1}} \tag{11}$$

where i is a $N \times 1$ vector where all elements equal one.

Initial testing of the method via Monte Carlo simulation showed that using the raw sampling estimates of the variance and covariance elements of the covariance matrix in Eq. (10) sometimes resulted in estimates of weights $\boldsymbol{\omega}$ inflating the regional variance in excess of the at-site variance. A more robust estimate of the

covariance matrix was introduced based on idea by Madsen and Rosbjerg (1997) replacing the values of σ_i and ρ in Eq.(10) with singular regional average values rather than the raw at-site sample values.

4. Model verification using Monte Carlo simulations

Two sets of Monte Carlo experiments were conducted to verify the proposed modelling strategy. First, the expression of covariance of the regional slope estimators from correlated sites (Eq. 10) was verified followed by an assessment of the ability of the regional slope estimator to provide credible estimates of the underlying regression slope parameter in a homogeneous region.

It is particularly important to consider how the procedure performs when AMS from different sites in the region are correlated, and how this will affect the ability to detect change when it exits. The Monte Carlo simulations used in this study were based on the regional model specified in Eqs. (5) and (6).

- Consider a homogeneous region consisting of $i = 1, \dots, N$ sites, each with $t = 1, \dots, n_i$ years of record and where the marginal distribution at each site is a 2-parameter log-normal distribution.
- First, generate $max(n_i)$ correlated realisation ε_t at each site from a N-dimensional multivariate standardises normal distribution with a correlation matrix $\boldsymbol{R}$ (scaled by the appropriate standard deviations σ_i to obtain the covariance matrix Σ). The off-diagonal elements of this $N \times N$ correlation matrix are all set to ρ representing the cross-correlation, while the diagonal elements are all set to one.
- At each site the $n_i$realisations of ε_t are transformed into a site-specific log-transformed annual maximum peak flow values using Eq. (5).

The Monte Carlo simulations therefore require specification of the model parameters at each of the N sites which include: two regression parameters, $\beta_{i,0}$ and β_1, as well as the standard deviation of the residuals at each site σ_i. In addition, the number of sites N, the record length at each site n_i and the degree of cross-correlation ρ all need to be specified.

4.1 Correlation between regression estimators

The analytical formula for the covariance between the estimated regression slope parameter β_1 at different sites is verified by considering a set of Monte Carlo experiments simulating AMS of log-normal distributed peak flow from two sites as outlined above.

The combinations of regression parameters are selected to illustrate the verification across a range of cross-correlations ($\rho = 0.0, \cdots, 0.80$, step by 0.10). Notably, the regional covariance estimator in Eq. (10) considers only the variance of the residuals and properties of the covariates, but are independent of the actual regression parameters. Hence in both cases, the regression parameters are specified as ($\beta_{1,0} = \beta_{2,0} = 0.0$, $\beta_{1,1} = \beta_{2,1} = 0.0050$) which are representative values from fitting OLS regression models to 662 AMS from UK catchments. A slope value of 0.0050 corresponds to a decadal magnification factor of exp(0.0050 * 10) ≈1.05, so a 5% increase in any design event over a 10 year period.

In the first example the standard deviation of the residuals at both sites are identical; $\sigma_1 = \sigma_2 = 0.20$ and with a common record-length of $n = 30$ years. In the second example the residual standard deviations are different; $\sigma_1 = 0.20$, $\sigma_2 = 0.40$. A total of 10,000 joint AMS at the two sites were generated using Monte Carlo simulations, and the mean of the estimated covariances compared to the theoretical expression in Figure 1.

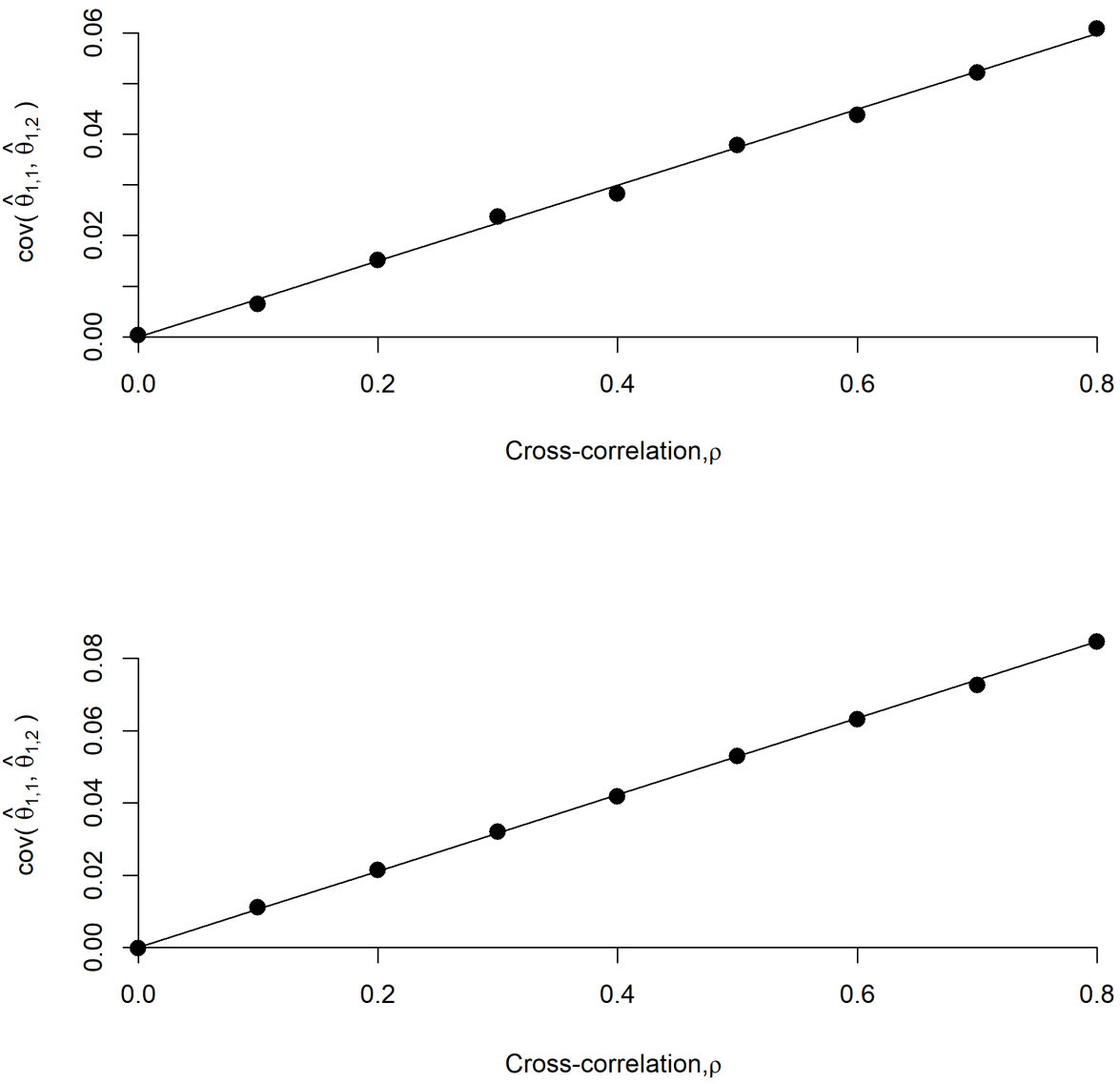

Figure 1: Comparison of analytical expressions and Monte Carlo simulations of covariance of slope of regression model at two sites (1 and 2) with correlated data. Record-length $n = 30$. Top: $\sigma_1 = \sigma_2 = 0.20$. Bottom: $\sigma_1 = 0.20$, $\sigma_2 = 0.40$.

The close correspondence between the outcome of the Monte Carlo simulations and the theoretical expression in Figure 1 is indicative that the derived expressions can be used as estimators of the covariances across a range of cross-correlations.

4.2 Performance of pooling group

To check the robustness of the estimation procedure, a homogenous region consisting of $M = 21$ sites was defined using the same basic model structure outlined in Eqs.(5) and (6) with $\beta_{i,0} = 1$ and $\beta_1 = 0.005$ (indicative of values observed in real data series where the covariate is the raw water year, i.e. not normalised). A total of 1000 regional data sets were generated across a range of cross-correlation as before ($\rho = 0.0, \cdots, 0.80$, step by 0.10). For each regional data set, the weights and the corresponding regional estimates of slope were calculated using the procedure outlined in this paper. At the same time, direct at-site estimation of the slope was also conducted at one of the 21 sites, thus allowing a comparison of the effect of using a regional procedure vs at-site trend estimation. Figure 2 shows the mean plus/minus two times the standard deviations for the regional and the at-site estimates of β_1 as a function of cross-correlation as derived from the 1000 Monte Carlo simulations.

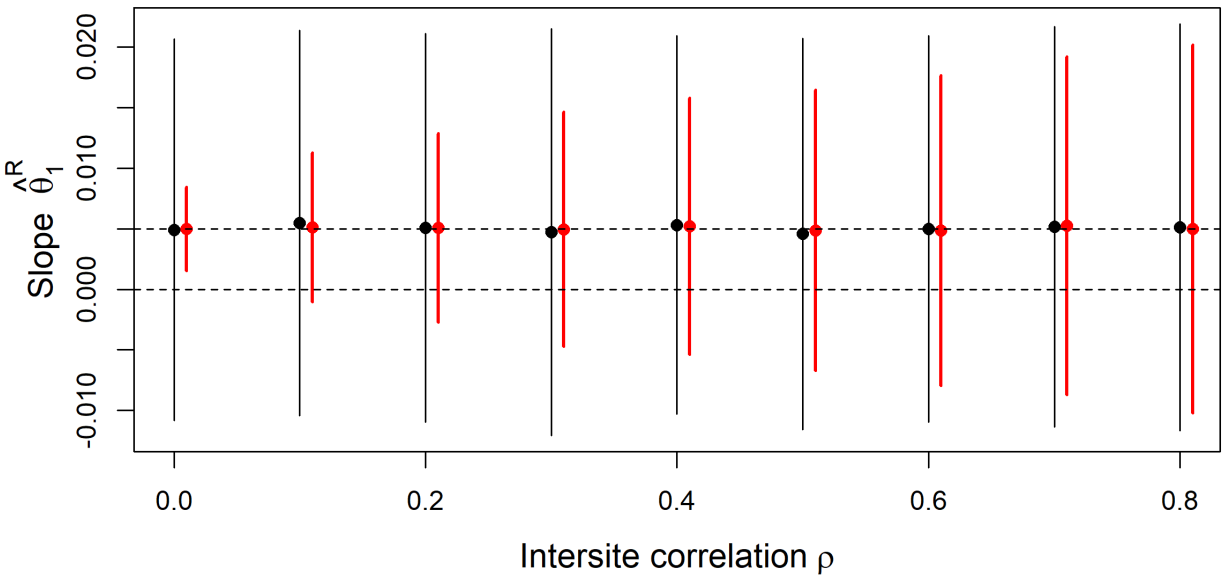

Figure 2: Comparison of the intervals covered by plus/minus two times the standard deviation for at-site (black lines) and pooled (red lines) results. The dashed line indicates the true value of the regional slope, and the points represent the mean values of the trend estimates obtained from the 10,000 Monte Carlo replications.

The results show a good correspondence between the regional estimator of β_1 and the underlying population value, and that the sampling uncertainty of the regional estimator is greatly reduced for low to moderate levels of cross-correlation when compared to the at-site estimates. However, for high degrees of cross-correlation the variance of the regional model approaches the variance of the at-site estimator. These results confirm that the proposed regional estimator is capable of providing unbiased estimates of the population parameter values, and that the sampling variance of the estimators behave as expected.

5. Case Study: United Kingdom peak flow data

The proposed model introduced and validated in the preceding sections have been applied to pooling groups formed using annual maximum series of peak flow from gauging station located through-out the United Kingdom. Version 9 of the National River Flow Archive (NRFA) peak flow dataset contains AMS of peak flow data from 935 catchments containing data up-to and including water year 2017 (i.e. ending September 2018). In this study, only catchment that fulfil the following criteria were included: (1) rating considered suitable up-to at-least bankfull, (2) a minimum record length of 20 years, and (3) considered essential rural (defined as having less than 5% urban land-use, as measured by the $URBEXT_{2000}$ catchment descriptor), resulting in a total of 662 catchments.

5.1 Formation of pooling groups

A pooling group was formed for each site in turn using the method outlined by Kjeldsen and Jones (2009), and enforcing the minimum number of peak flow events in the pooling group to be 500 years. The pooling groups are formed by selecting gauged catchments among the 662 available that are considered hydrologically similar to the target site with regards to: catchment area, standardised annual average rainfall as measured from 1961-1990 (SAAR), and the potential influence of lakes and reservoirs (FARL) and flood plain extend (FPEXT). The similarity is judged based on Euclidian distance in a four-dimensional space determined by the set of catchment descriptors associated with each catchment. This resulted in a total of 662 pooling groups (one for each site) containing on average 11.5 gauging stations, ranging from a

minimum of 6 to a maximum of 15 gauging stations. Note that individual sites can be members of more than one pooling group.

At each site a log-linear model was fitted directly to the AMS and subsequently the pooled slope $\hat{\beta}_1^R$ estimated for each pooling group using Eq. (8). Prosdocimi et al. (2014) and Laio et al. (2009) discussed the appropriateness of using the LN2 distribution to model annual maximum peak flow from rivers in the United Kingdom, and both studies found it to be a reasonable model based on application of goodness-of-fit tests applied to national data sets.

While the pooling groups are formed based on an assessment of hydrological similarity, there was no explicit consideration of similarity in term of trend or not in the data series. Studies from other regions in the World attempting to form regions (or pooling groups) for the purpose of regional flood frequency analysis include e.g. in Canada, Chunderlik and Burn (2003) and O'Brien and Burn (2014). Both studies used seasonality of the annual maximum floods as the basis for forming pooling groups, while Leclerc and Ouarda (2007) used canonical correlation analysis based on catchment area, rainfall, location (lat/long) and temperature. As the formation of the UK pooling does not explicitly consider spatial proximity, the use of standard average annual rainfall (SAAR) as one of the key features ensures some spatial consistency due to the strong east-west gradient in rainfall patterns across the UK.

5.2 Testing trend hypothesis

A recognised benefit of regional flood frequency analysis is the reduction in the sampling variance of the regionalised (pooled) parameter; while this typically apply to the estimation of parameters of the stationary model, this advantage applies in this case to the estimate of slope $\hat{\beta}_1^R$. In this case, the pooled estimate of the slope will feed into a hypothesis test determining if a trend in the annual maximum peak flow data is evident or not across the pooling group.

A null-hypothesis H_0 is defined as $\hat{\beta}_1^R = 0$ against the alternative hypothesis H_1: $\hat{\beta}_1^R \neq 0$. To assess if H_0 can be rejected, a two-sided test can be made against a t-distribution using a test statistic defined as

$$\frac{\hat{\beta}_1^R - 0}{\sqrt{var(\hat{\beta}_1^R)}} \quad (12)$$

As the combined sample size used for estimating β_1^R is an order of magnitude larger than the sample size available for each at-site analysis, we approximate the t-distribution with a standard normal distribution. The same is done for the at-site analysis to enable a direct comparison of the significance of both the at-site and the regional slope estimates.

Figure 3 shows the boxplots of the slope estimates for each site from the at-site and pooled analysis. Evidently, the range of estimates obtained using the at-site estimates is large while the pooled estimates are more constrained. Notably, the number of negative estimates of slope is large for the at-site analysis (33.5%) but only 19.2% for the pooled estimates.

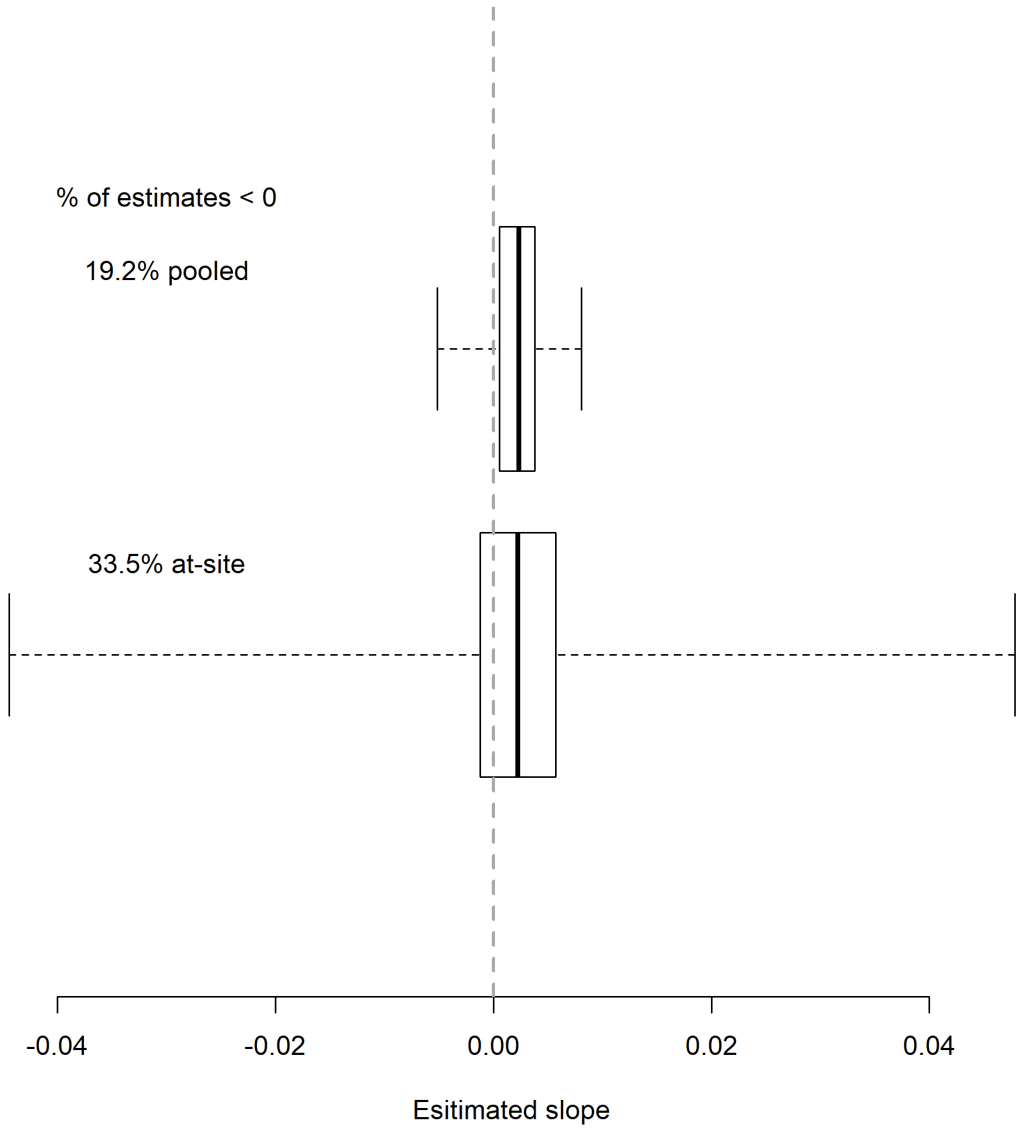

Figure 3: Box plot showing the range of slope estimates obtained from the regional model (top) and at-site analysis (bottom).

While individual pooling groups can contain a mix of sites with positive and negative at-site estimates of slope, the results in Figure 3 suggest that, in general, the AMS of peak flow are characterised by a positive slope and less variability between sites.

A comparison of the level of statistical significance in the slope estimates derived from the at-site and pooled slope estimates are shown on maps of the UK in Figure 4 where each symbol represents the location of the target site of the pooling group.

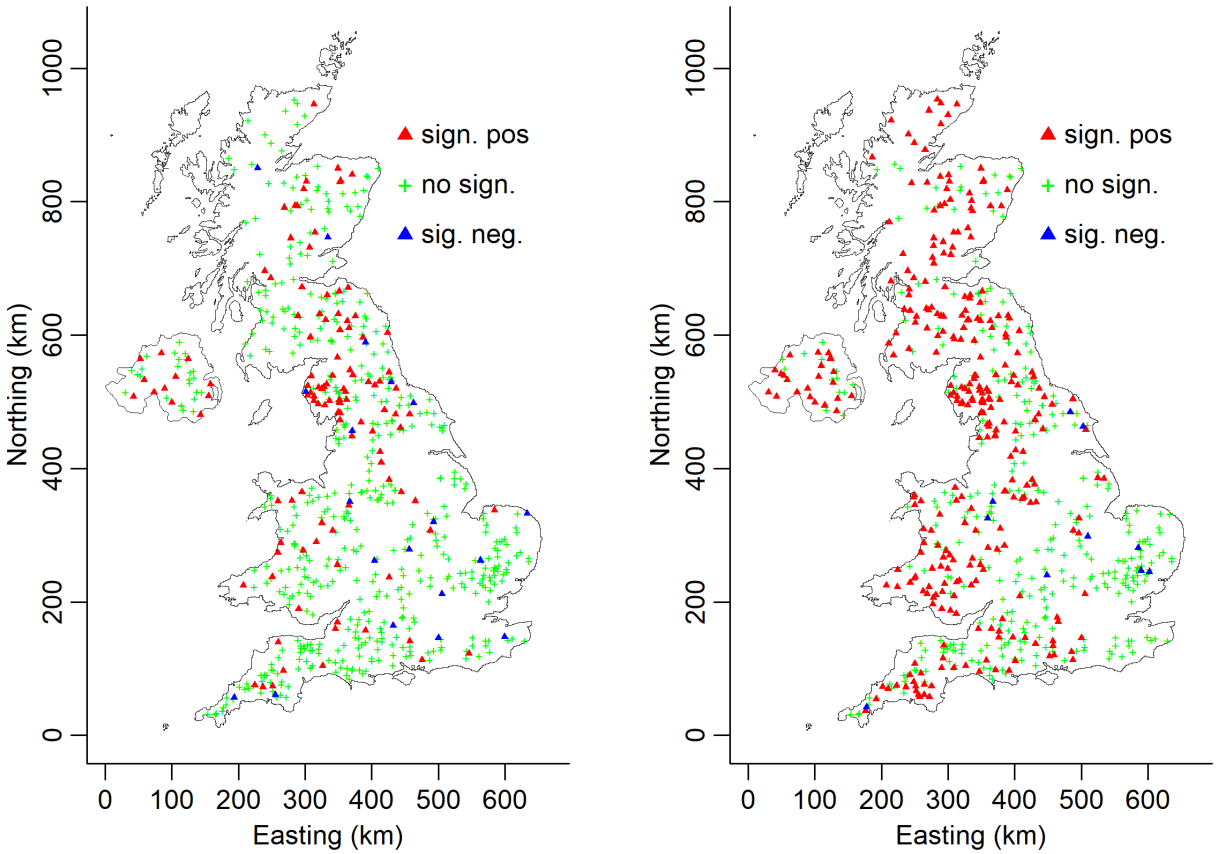

Figure 4: Statistical significance of slope estimate obtained using at-site analysis (left) and the regional model (right). green cross, non-significant trend; red triangle, significant positive slope, blue triangle, significant negative slope . Significance level: 5%.

The spatial distribution of sites with significant upwards trend, as shown in Figure 5, indicate that the western and north western parts of the UK are showing the strongest signals of upward trend in the flood data, while the East and South Eastern region shows little or no sign of a change in any direction. These results are consistent with the results of recent national flood trend studies presented by Prosdocimi et al. (2019) as well as Faulkner et al. (2020).

To further explore a possible link between catchment types and a propensity for trend in the flood data, Figure 5 shows the location of each target site as a function of catchment area and standard annual average rainfall (SAAR); the colours of the points indicate whether a significant trend, positive (red) or negative (blue) was identified or not (green).

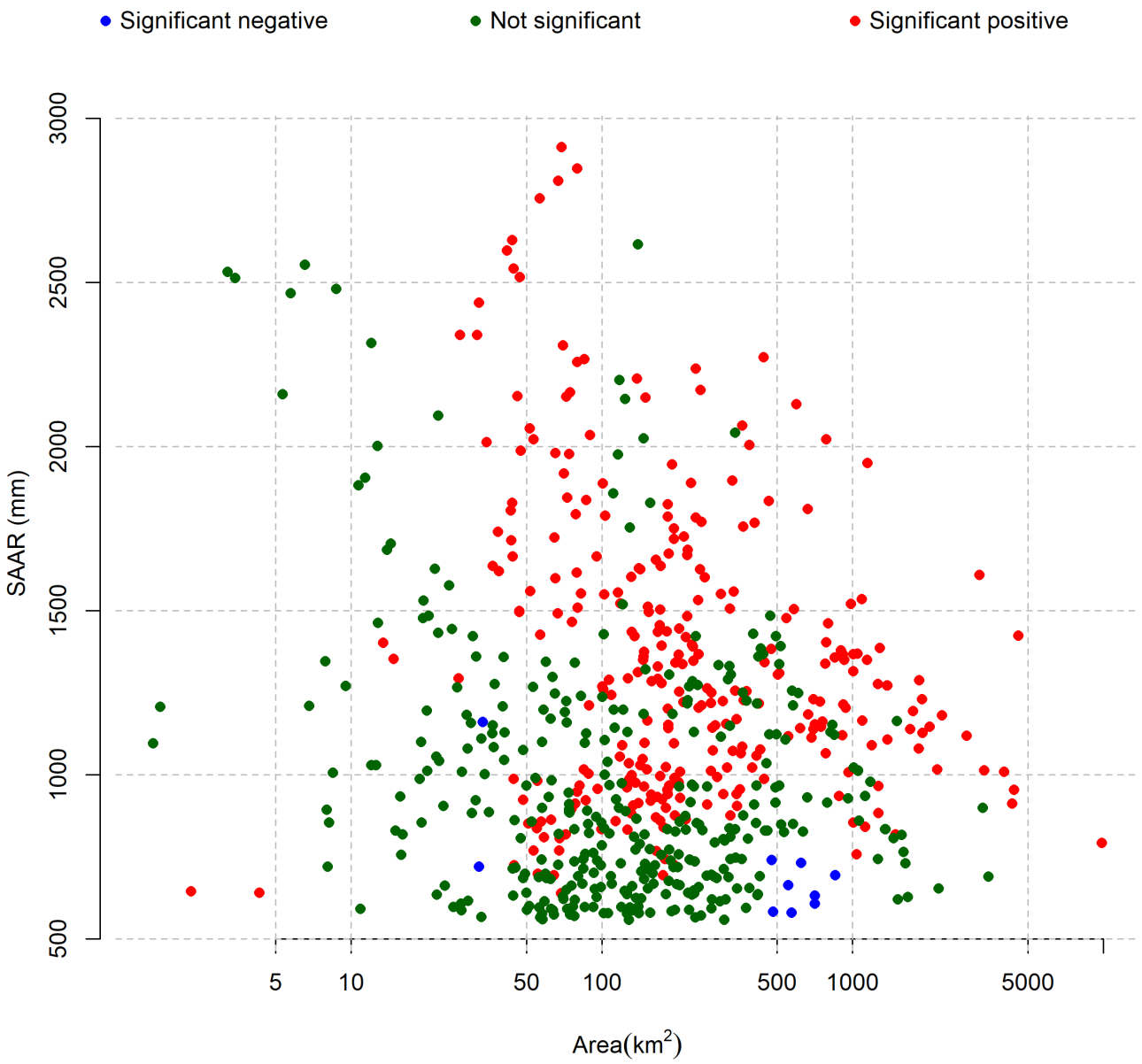

Figure 5: Catchment descriptors and significance of the regional slope estimate, $\hat{\beta}_1^R$, for each of the 655 sites.

The results in Figure 5 do not reveal any particular catchment type to be more susceptible to change than other. However, Figure 5 does show that the majority of gauged catchments are medium sized (catchment area between 25 km^2 and 500 km^2) and relatively dry (SAAR of less than 1500mm). As the pooling groups are formed predominantly based on similarity of area and rainfall, the pooling groups formed for sites with values of catchment area and SAAR located on the edge of the dataset will be rather similar as they will draw from the same limited pool of available data. The effect of this can be observed, for example, in the cluster of sites with decrease trend located in the lower right of the graph, i.e. low values of SAAR and catchment area in the vicinity of 500 km^2. It is also notable that the sites with significant upward trend (red dots) are generally associated with higher values of SAAR than the sites where no significant trend (green dots) was detected. This is reflecting the patterns also observed on the maps in Figure 4 where sites with significant upward trend are located predominantly in the wetter western and northern parts of the country, while the drier eastern and southern parts are dominated by sites where no significant trend was detected.

5.3 Impact of non-stationarity on design floods

Magnification factors (eq. 4) measure the relative increase of design flood, assuming a LN2 distribution, in response to a change in the covariate Δx. Having obtained estimates of the slope β_1, as well as the associated statistical significance, for each of the 662 catchments using both the at-site and regional method, it is possible to assess the impact of using the regional method on the expected change in design floods across the country. Figure 6 shows the evolution of the magnification factor as a function of Δx (i.e. water years into the future) for each target site when the slope β_1 is estimated using traditional at-site trend analysis (left) and the new regional trend estimator (right).

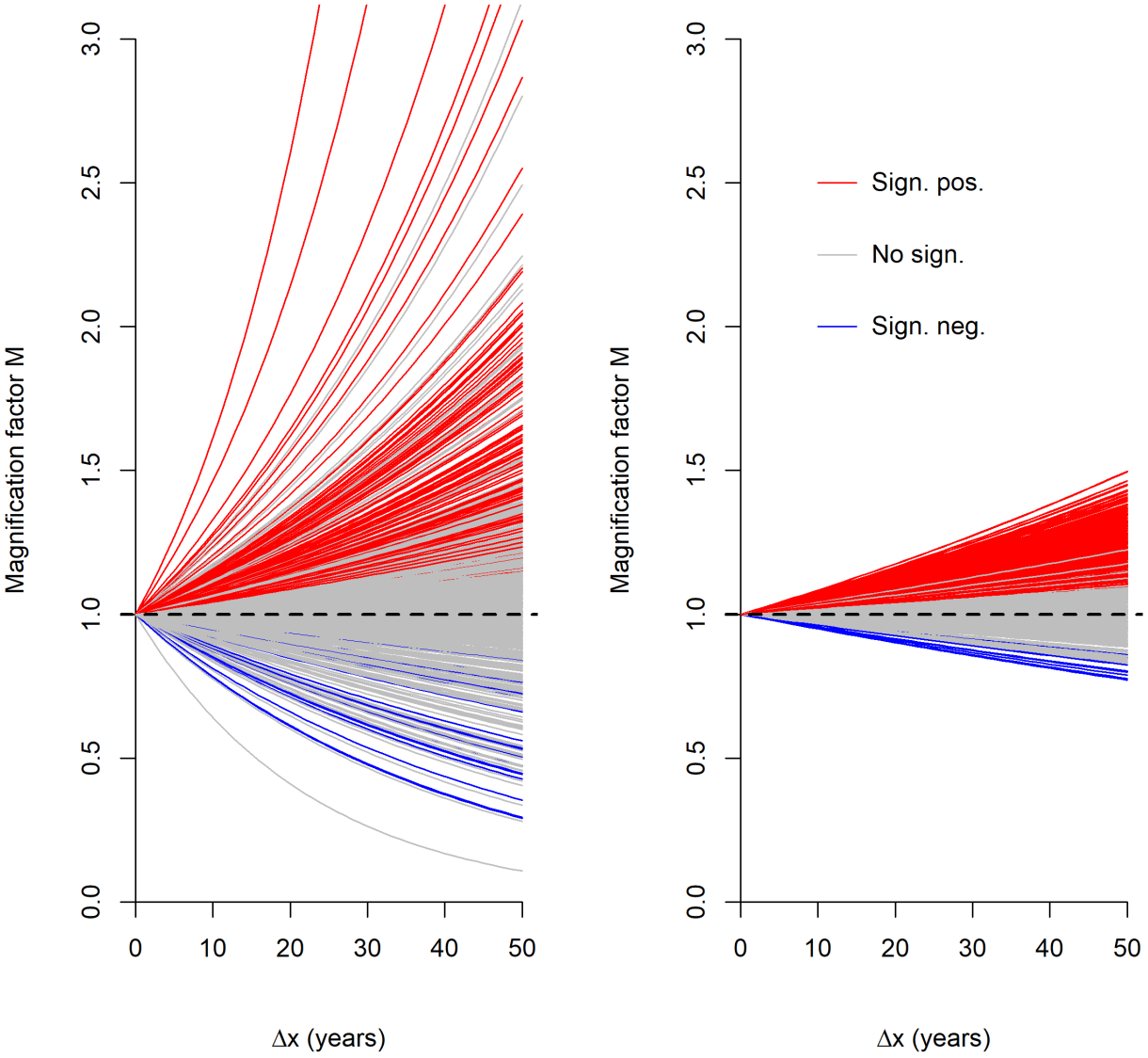

Figure 6: Magnification factors M plotted as a function of time horizon into the future Δx when basing extrapolations on the slope estimated from at-site data (left) and the regional model (right). Based on a 5% significance level, the red lines indicate significant positive trend, grey lines are non-significant, and blue lines represent significant negative trend.

There is a notable difference in the behaviour of the magnification factor when estimates are based on at-site and regional analysis. In particular, the variability between sites is comprehensively reduced when the assessment is based on regional data, restricting the increase in design floods to about 50%. In contrast,

the site-to-site variation is substantial when relying on at-site analysis, including more sites showing a significant reduction in design flood (a significant negative slope) as signified by the blue lines on Figure 6, as well as increases in design floods of up-to 200% over a 50-year period in some catchments.

6. Discussion

The new methodological developments and results presented in this study has demonstrated how the concept of magnification factors (Vogel et al., 2011) can be extended to allow more robust assessment of trend in regional series of hydrological extremes and changes in design floods across a specified region or pooling group. However, moving from at-site to regional trend analysis naturally raises a number of issues of both theoretical and practical issue. Firstly, this study has made the assumption that the pooling groups formed on the basis of hydrological similarity when considering: catchment area, annual average rainfall, and the potential influence of lakes, reservoirs and flood plains (Kjeldsen and Jones, 2009). While this is a convenient method for forming regions, it is possible that hydrological regions formed differently might provide a more suitable basis for regional trend detection. It should also be noted that the same site can be a member of several pooling groups.

Secondly, the time period covered by the observed records might vary from site to site, and the resulting at-site trend estimates might be representative of different flood rich or flood poor periods. Consequently, the weighted average of these different at-site estimates would represent an overall trend across all these periods. It might be that more consistent results could be obtained if a common data period was selected for each site, even if this would potentially result in a high degree of data loss. It would also be possible to use more sophisticated hierarchical models such as demonstrated by Brady et al. (2019) to better represent more complex dependence structures.

Thirdly, the trend model considered in this study is very simple, considering a blanket linear trend over time for each site and thus ignoring the potential influence of physical drivers of change such as, for example:

urbanisation (Hecht and Vogel, 2020), change in precipitation (Šraj et al. 2016), change in large-scale climatic indicators such as the North Atlantic Oscillation (Brady et al. 2019), or long-terms variation in flood rich and flood poor periods (Macdonald and Sangster, 2017).

Finally, a simple two-sided test was used to establish if a trend is significant or not. A more sophisticated testing methodology could be introduced where the default position is the existence of trend rather than starting from an assumption of no-trend as discussed in details by Vogel et al. (2013) and Prosdocimi et al. (2014).

7. Conclusions

This study has presented a new procedure for estimating a regional slope parameter from a collection of annual maximum series considered to form a hydrological region of pooling group. The method explicitly considers the effect of cross-correlation on the variance of the estimated slope. The method was successfully combined with the concept of a magnification factor developed by Vogel et al. (2011) to allow a statistical assessment of the impact of change on future design floods. The results of the regional analysis shows that there is substantial evidence of a significant increase in annual maximum peak flow series in the West and North West of the United Kingdom. Extrapolating the existing trends into the future shows a positive trend in the majorities of pooling groups and a predicted change in design floods of up to about 50% for a time horizon of 50 years into the future.

Acknowledgements: The authors would like to thank the National River Flow Archive (NRFA) for making the peak flow data series and the catchment descriptors available. Helpful comments from two reviewers helped to improve the manuscript and are acknowledged The data was accessed and processed using the winfapReader R-package (Prosdocimi and Shaw, 2020).

References

Archfield, S.A., Hirsch, R.M., Viglione, A. and Blöschl, G. (2016) Fragmented patterns of flood change across the United States. Geophysical Research Letters, 43(19), pp.10,232-10,239.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R.A., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G.T., Bilibashi, A. and Boháč, M. et al. (2019) Changing climate both increases and decreases European river floods. *Nature*, 573(7772), pp.108-111.

Brady, A., Faraway, J. and Prosdocimi, I. (2019) Attribution of long-term changes in peak river flows in Great Britain. *Hydrological Sciences Journal*, 64(10), pp.1159-1170.

Cunderlik, J.M. and Burn, D.H. (2003) Non-stationary pooled flood frequency analysis. *Journal of Hydrology*, 276(1-4), pp.210-223.

Douglas, E.M., Vogel, R.M. and Kroll, C.N. (2000) Trends in floods and low flows in the United States: impact of spatial correlation. *Journal of hydrology*, 240(1-2), pp.90-105.

Faulkner, D., Warren, S., Spencer, P. and Sharkey, P. (2020) Can we still predict the future from the past? Implementing non-stationary flood frequency analysis in the UK. *Journal of Flood Risk Management*, 13(1), p.e12582.

Hecht, J.S. and Vogel, R.M. (2020) Updating urban design floods for changes in central tendency and variability using regression. *Advances in Water Resources*, 136, p.103484.

Hosking, J.R.M. and Wallis, J.R. (1997) *Regional frequency analysis: an approach based on L-moments*. Cambridge, UK: Cambridge University Press, 240.

Kjeldsen, T.R. and Prosdocimi, I. (2015) A bivariate extension of the Hosking and Wallis goodness-of-fit measure for regional distributions. *Water Resources Research*, 51(2), pp.896-907.

Kjeldsen, T.R. and Jones, D.A. (2009) A formal statistical model for pooled analysis of extreme floods. *Hydrology research*, 40(5), pp.465-480.

Kjeldsen, T.R., Ahn, H. and Prosdocimi, I., 2017. On the use of a four-parameter kappa distribution in regional frequency analysis. *Hydrological Sciences Journal*, 62(9), pp.1354-1363.

Laio, F., Di Baldassarre, G. and Montanari, A. (2009) Model selection techniques for the frequency analysis of hydrological extremes. *Water Resources Research*, 45(7).

Leclerc, M. and Ouarda, T.B. (2007) Non-stationary regional flood frequency analysis at ungauged sites. *Journal of hydrology*, 343(3-4), pp.254-265.

Macdonald, N. and Sangster, H. (2017) High-magnitude flooding across Britain since AD 1750. *Hydrology and Earth System Sciences*, 21(3), pp.1631-1650.

Madsen, H. and Rosbjerg, D. (1997) Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modeling. *Water Resources Research*, 33(4), pp.771-781.

Mediero, L., Kjeldsen, T.R., Macdonald, N., Kohnova, S., Merz, B., Vorogushyn, S., Wilson, D., Alburquerque, T., Blöschl, G., Bogdanowicz, E. and Castellarin, A. (2015) Identification of coherent flood regions across Europe by using the longest streamflow records. *Journal of Hydrology*, 528, pp.341-360.

Mohan, S., Clarke, R.M. and Chadee, X.T., 2020. Variations in extreme temperature and precipitation for a Caribbean island: Barbados (1969–2017). *Theoretical and Applied Climatology*, 140, 1277–1290

O'Brien, N.L. and Burn, D.H., 2014. A nonstationary index-flood technique for estimating extreme quantiles for annual maximum streamflow. Journal of Hydrology, 519, pp.2040-2048.

Prosdocimi, I. and Kjeldsen, T. R. (2020) Parametrisation of change-permitting extreme value models and its impact on the description of change. *Stochastic Environmental Research and Risk Assessment*, submitted.

Prosdocimi, I. and Shaw, L. (2020, October). *winfapReader*: an R package to interact with Peak Flow Data in the United Kingdom (version 0.1-2). Zenodo. https://doi.org/10.5281/zenodo.4138993

Prosdocimi, I., Dupont, E., Augustin, N.H., Kjeldsen, T.R., Simpson, D.P. and Smith, T.R. (2019) Areal models for spatially coherent trend detection: the case of British peak river flows. *Geophysical Research Letters*, 46(22), pp.13054-13061.

Prosdocimi, I., Kjeldsen, T.R. and Svensson, C. (2014) Non-stationarity in annual and seasonal series of peak flow and precipitation in the UK. *Natural Hazards and Earth System Sciences*, 14(5), pp.1125-1144.

Renard, B., Lang, M., Bois, P., Dupeyrat, A., Mestre, O., Niel, H., Sauquet, E., Prudhomme, C., Parey, S., Paquet, E. and Neppel, L. (2008) Regional methods for trend detection: Assessing field significance and regional consistency. *Water Resources Research*, 44(8).

Rosbjerg, D., Bloschl, G., Burn, D., Castellarin, A., Croke, B., Di Baldassarre, G., Iacobellis, V., Kjeldsen, T.R., Kuczera, G., Merz, R. and Montanari, A. (2013) *Prediction of floods in ungauged basins*. In Runoff prediction in ungauged basins: Synthesis across processes, places and scales (pp. 189-225). Cambridge University Press.

Salas, J.D., Obeysekera, J. and Vogel, R.M. (2018) Techniques for assessing water infrastructure for nonstationary extreme events: A review. *Hydrological Sciences Journal*, 63(3), pp.325-352.

Šraj, M., Viglione, A., Parajka, J. and Blöschl, G. (2016) The influence of non-stationarity in extreme hydrological events on flood frequency estimation. *Journal of Hydrology and hydromechanics*, 64(4), pp.426-437.

Vicente-Serrano, S.M., Domínguez-Castro, F., Murphy, C., Hannaford, J., Reig, F., Peña-Angulo, D., Tramblay, Y., Trigo, R.M., Mac Donald, N., Luna, M.Y. and Mc Carthy, M. (2020) Long-term variability and trends in meteorological droughts in Western Europe (1851–2018). *International Journal of Climatology*.

Villarini, G., Serinaldi, F., Smith, J.A. and Krajewski, W.F. (2009) On the stationarity of annual flood peaks in the continental United States during the 20th century. *Water Resources Research*, *45*(8).

Vogel, R.M., Rosner, A. and Kirshen, P.H. (2013) Brief Communication: Likelihood of societal preparedness for global change: trend detection. *Natural Hazards & Earth System Sciences*, 13(7).

Vogel, R.M., Yaindl, C. and Walter, M. (2011) Nonstationarity: flood magnification and recurrence reduction factors in the United States 1. *Journal of the American Water Resources Association*, 47(3), pp.464-474.

Zadeh, S.M., Burn, D.H. and O'Brien, N. (2020) Detection of trends in flood magnitude and frequency in Canada. *Journal of Hydrology: Regional Studies*, 28, p.100673.

Zhang, Q., Gu, X., Singh, V.P., Xiao, M. and Xu, C.Y. (2015) Flood frequency under the influence of trends in the Pearl River basin, China: changing patterns, causes and implications. *Hydrological Processes*, 29(6), pp.1406-1417.

Appendix A: Covariance of regression parameter estimates

Annual maximum series (AMS) of log-transformed flow $y_{1,t}$ and $y_{2,t}$ is available from two catchment with overlapping records $\delta \in [1{:}\, n_\delta]$. For each site, a regression model linking the log-transformed flow with the covariate x_t as

$$\begin{aligned} y_{1,t} &= \beta_{1,0} + \beta_{1,1} x_t + \varepsilon_{1,t} \\ y_{2,t} &= \beta_{2,0} + \beta_{2,1} x_t + \varepsilon_{2,t} \end{aligned} \tag{A1}$$

For each site the residuals are assumed to be normal distributed with zero mean, a site-specific standard deviation σ_i and cross-correlation ρ exists between the two series for events that occur in the same year, but not between lagged years, i.e.

$$cor\left(\varepsilon_{1,t}, \varepsilon_{2,k}\right) = \begin{cases} \rho & t = k \\ 0 & t \neq k \end{cases} \tag{A2}$$

For each site, the OLS estimate of slope can be obtained as

$$\begin{aligned} \hat{\beta}_{1,1} &= SS_{x,1}^{-1} \sum_{t=1}^{n_1} (x_t - \overline{x}) y_{1,t} \\ \hat{\beta}_{2,1} &= SS_{x,2}^{-1} \sum_{t=1}^{n_2} (x_t - \overline{x}) y_{2,t} \end{aligned} \tag{A3}$$

Where $SS_{x,1}$ and $SS_{x,2}$ are the sum of squares over the covariate for each of the two sites.

The covariance of the least square slope estimates at the two sites is found by combining Eqs. (A1), (A2) and (A3) as:

$$
\begin{aligned}
cov\left(\hat{\beta}_{1,1}, \hat{\beta}_{2,1}\right) &= cov\left(SS_{x,1}^{-1} \sum_{t=1}^{n_1} (x_t - \overline{x}) y_{1,t} \, , SS_{x,2}^{-1} \sum_{t=1}^{n_2} (x_t - \overline{x}) y_{2,t} \right) \\
&= SS_{x,1}^{-1} SS_{x,2}^{-1} cov\left(\sum_{t=1}^{n_1} (x_t - \overline{x}) y_{1,t} \, , \sum_{t=1}^{n_2} (x_t - \overline{x}) y_{2,t} \right) \\
&= SS_{x,1}^{-1} SS_{x,2}^{-1} cov\left(\sum_{t=1}^{n_1} (x_t - \overline{x}) \left(\beta_{1,0} + \beta_{1,1} x_t + \varepsilon_{1,t}\right) , \sum_{t=1}^{n_2} (x_t - \overline{x}) \left(\beta_{2,0} \right.\right. \right. \\
&\left. \left. \left. + \beta_{2,1} x_t + \varepsilon_{2,t}\right) \right) \\
&= SS_{x,1}^{-1} SS_{x,2}^{-1} cov\left(\sum_{t \in \delta} x_t \varepsilon_{1,t} - \overline{x} \sum_{t \in \delta} \varepsilon_{1,t} \, , \sum_{t \in \delta} x_t \varepsilon_{2,t} - \overline{x} \sum_{t \in \delta} \varepsilon_{2,t} \right) \\
&= \frac{\sigma_1 \sigma_2 \rho}{SS_{x,1} SS_{x,2}} \left[\sum_{t \in \delta} x_t^2 - n_\delta (\overline{x})^2 \right] = \sigma_1 \sigma_2 \rho \frac{SS_{x,\delta}}{SS_{x,1} SS_{x,2}}
\end{aligned}
\tag{A4}
$$

For the special case of $\hat{\beta}_{1,1} = \hat{\beta}_{2,1}$ where $\rho = 1$, $\sigma_1 = \sigma_2$ and $SS_{xx,1} = SS_{x,2} = SS_{x,\delta}$ the covariance estimator defaults to the variance of the at site estimator $\hat{\beta}_1$, i.e.

$$
var\left(\hat{\beta}_1\right) = \frac{\sigma^2}{SS_x} \tag{A5}
$$