

1 **Simulating multimodal seasonality in extreme daily precipitation occurrence**

2 Mari R. Tye^{1, 2*}, Stephen Blenkinsop², Hayley J. Fowler², David B. Stephenson^{3,4},

3 Christopher G. Kilsby^{2,4}

4 ¹ National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO, 80307, USA

5 ² School of Civil Engineering and Geosciences, Newcastle University, Newcastle-

6 upon-Tyne, UK

7 ³ Exeter Climate Systems, Mathematics Research Institute, University of Exeter,

8 Exeter, UK

9 ⁴ Willis Research Network, 51 Lime Street, London, EC3M 7DQ, UK

10 Corresponding Author: Mari R. Tye: maritye@ucar.edu

11 **Highlights:**

- 12 • The frequency of extreme daily precipitation occurrence displays a distinctive non-
- 13 uniform seasonal pattern
- 14 • The seasonal distribution is simulated with Generalized Additive Models
- 15 • A strong dependence on atmospheric driving conditions is replicated.
- 16 • Statistical simulations indicate a future shift toward more frequent autumnal extreme
- 17 precipitation.

18

Abstract

19 Floods pose multi-dimensional hazards to critical infrastructure and society and
20 these hazards may increase under climate change. While flood conditions are dependent
21 on catchment type and soil conditions, seasonal precipitation extremes also play an
22 important role. The extreme precipitation events driving flood occurrence may arrive
23 non-uniformly in time. In addition, their seasonal and inter-annual patterns may also
24 cause sequences of several events and enhance likely flood responses.

25 Spatial and temporal patterns of extreme daily precipitation occurrence are
26 characterized across the UK. Extreme and very heavy daily precipitation is not uniformly
27 distributed throughout the year, but exhibits spatial differences, arising from the relative
28 proximity to the North Atlantic Ocean or North Sea. Periods of weeks or months are
29 identified during which extreme daily precipitation occurrences are most likely to occur,
30 with some regions of the UK displaying multimodal seasonality.

31 A Generalized Additive Model is employed to simulate extreme daily
32 precipitation occurrences over the UK from 1901-2010 and to allow robust statistical
33 testing of temporal changes in the seasonal distribution. Simulations show that
34 seasonality has the strongest correlation with intra-annual variations in extreme event
35 occurrence, while Sea Surface Temperature (SST) and Mean Sea Level Pressure (MSLP)
36 have the strongest correlation with inter-annual variations. The north and west of the UK
37 are dominated by MSLP in the mid-North Atlantic and the south and east are dominated
38 by local SST.

39 All regions now have a higher likelihood of autumnal extreme daily precipitation
40 than earlier in the twentieth century. This equates to extreme daily precipitation occurring

41 earlier in the autumn in the north and west, and later in the autumn in the south and east.
42 The change in timing is accompanied by increases in the probability of extreme daily
43 precipitation occurrences during the autumn, and in the number of days with a very high
44 probability of an extreme event. These results indicate a higher probability of several
45 extreme occurrences in succession and a potential increase in flooding.

46 **Key Words:** Precipitation, extreme, Generalized Additive Model, seasonality,
47 UK

48 **Abbreviations:**

49 EA East Anglia (Extreme Region)

50 ES East Scotland (Extreme Region)

51 FOR Forth (Extreme Region)

52 GLM Generalized Linear Model

53 GAM Generalized Additive Model

54 HUM Humber (Extreme Region)

55 MSLP Mean Sea Level Pressure

56 MW Mid Wales (Extreme Region)

57 NAO North Atlantic Oscillation

58 NE North East (Extreme Region)

59 NHI North Highlands and Islands (Extreme Region)

60 NI Northern Ireland (Extreme Region)

- 61 NW North West (Extreme Region)
- 62 SE South England (Extreme Region)
- 63 SH South Highlands (Extreme Region)
- 64 SOL Solway (Extreme Region)
- 65 SST Sea Surface Temperature
- 66 SW South West (Extreme Region)
- 67 WC West Country (Extreme Region)

68 1 Introduction

69 Despite considerable evidence for increases in the number and intensity of
70 extreme daily precipitation (e.g. Jones et al. 2014; Westra et al. 2014), there has been
71 little focus on *when* and *how often* these events occur within the year (Pal et al., 2013;
72 Dhakal et al., 2015). Yet, there is strong evidence that UK extreme and very heavy daily
73 precipitation events do not occur uniformly. Instead, they follow a seasonal pattern that
74 may differ substantially from seasonal cycles in monthly total, or mean wet day,
75 precipitation (Rodda et al., 2009; Maraun et al., 2009). This article is about the
76 occurrence of extreme daily precipitation exceeding a high threshold, and considers very
77 heavy and extreme daily precipitation (Alexander et al., 2006) as “extreme”. Extreme
78 daily precipitation in the UK tends to occur in the late summer and autumn (August-
79 November), with very few occurrences during spring months (March-May) (Allan et al.,
80 2015). Many parts of the UK exhibit more than one peak in the observed frequency of
81 extreme daily precipitation during the year, or multi-modal seasonality (Rodda et al.,
82 2009). There is also strong evidence of year-to-year variability in the number and time
83 interval between extreme precipitation events, arising from fluctuations in atmospheric
84 conditions (Tramblay et al., 2011).

85 While Zheng et al. (2015) focussed on seasonal changes in extreme precipitation
86 intensities, they highlighted that the changes can differ at one location dependent on both
87 the season and the duration of the event. Maraun et al. (2009) employed a seasonally
88 varying statistical model to examine monthly block maxima of extreme daily
89 precipitation intensity and frequency over multiple years, but did not explore the
90 frequency distribution within the year. Changes in extreme precipitation occurrences do

91 not necessarily equate with changes in flooding (Stephens et al., 2015), as flooding is
92 more dependent on the antecedent wetness conditions than extreme precipitation (Pui et
93 al., 2011). There is evidence for increases in the duration of wet periods, particularly
94 within autumn and winter, with associated increases in the intensity of longer wet spells
95 (Zolina, 2014; Zolina et al. 2013). If seasonal extreme occurrences coincide with longer
96 wet periods (e.g. autumn and winter) the potential for flooding increases. Changes in the
97 time interval between extreme daily precipitation occurrence will have impacts on runoff
98 response by either improving the recovery period (where the interval between maxima
99 increases) or increasing flooding if the interval decreases (and increases catchment
100 wetness). Unimodal sinusoidal patterns have often been adopted by previous analyses to
101 replicate the annual frequency distribution (e.g. Rust et al., 2009).

102 This work demonstrates that simple sinusoidal forms do not adequately reflect the
103 observed annual frequency distribution of extreme daily precipitation occurrence across
104 the UK. Flooding also exhibits considerable interannual variability, with a tendency to
105 cluster in “flood rich” and “flood poor” periods (Huntingford et al., 2014). In common
106 with extreme precipitation frequency, these fluctuations respond to atmospheric
107 oscillations such as the North Atlantic Oscillation or Arctic Oscillation (Bouwer et al.,
108 2008; Hannaford & Marsh, 2010; Huntingford et al., 2014). Thus, identifying natural
109 cycles and multi-year patterns is a crucial element in determining the current risks posed
110 by extreme daily precipitation and can help identify vulnerabilities in risk management
111 strategies (Wilby and Keenan, 2012). Statistical models play an essential role in
112 identifying these changes and their significance.

113 Generalized Linear Models (GLM) have been widely applied to simulate
114 seasonally varying daily precipitation occurrence and intensity (e.g. Chandler and
115 Wheeler, 2002; Verdin et al., 2014) and to examine long term responses to atmospheric
116 and oceanic oscillations (Chandler, 2005; Sapiano et al., 2006). However, more flexibility
117 is required for non-linear complex responses, such as the seasonal frequency distribution
118 of extreme daily precipitation occurrence, to external drivers. Generalized Additive
119 Models (GAM) are an enhancement to the GLM that employ smoothly varying non-
120 linear functions (Hastie and Tibshirani, 1990) to achieve flexibility in the modelled
121 response. Beuchat et al. (2012) demonstrated that explicit representation of temporally
122 varying processes with a GAM is more efficient than examining each month separately
123 (e.g. Maraun et al., 2009). Serinaldi and Kilsby (2014) employed vector GLMs and
124 GAMs to successfully simulate precipitation processes over the Danube Basin on a 0.25°
125 $\times 0.25^\circ$ grid. However, their model includes seasonality as a sinusoidal measure and
126 differs from the observed frequency of extreme daily precipitation occurrences.

127 GLMs and GAMs are useful for understanding long term changes in climatic
128 variables (Chandler, 2005; Westra et al., 2013), and to obtain realistic estimates of future
129 changes in precipitation through statistical downscaling (e.g. Kallache et al., 2011;
130 Schindler et al., 2012). Other researchers have employed GAMs to assess the time- and
131 atmospheric dependence in precipitation data (Hyndman and Grunwald, 2000) or trends
132 in daily precipitation (Underwood, 2009). Likely changes in the future temporal
133 distribution of sub-daily precipitation intensity were explored by Westra et al. (2013),
134 premising a GAM on current responses to atmospheric variables. With a few exceptions
135 (e.g. Chavez-Demoulin and Davison, 2005; Yee and Stephenson 2007; Maraun et al.,

136 2009; Westra et al., 2013; among others) there are not many examples of extreme daily
137 precipitation response analyses using GLMs or GAMs.

138 In this paper we use GAMs to explore variability in the seasonal frequency
139 distribution of extreme daily precipitation, evidence for long-term changes, and
140 dependence on oceanic and atmospheric phenomena. Multiple simulations from a time-
141 varying Binomial distribution are used to examine changes in the frequency distribution
142 between 1901-2010 and to compare their spatial differences across the UK.

143 **2 Data**

144 **2.1 Precipitation**

145 Daily precipitation observations compiled and described by Jones et al. (2013,
146 2014) for 199 observation stations across the UK are used in this analysis and illustrated
147 in Figure 1. All stations have >40 years observations, commencing between 1856 and
148 1961 and extending to 2010. Months were rejected when >3 days were missing, or years
149 rejected when >10% observations were missing. Stations were removed from the analysis
150 if >5 consecutive years were removed due to these criteria. Of the available 223 stations,
151 only 199 had adequate consecutive observation years to be suitable for a peak-over-
152 threshold analysis. Prior to 1961 and after 2000 there is considerable variability in the
153 number of operational stations in each region. Model construction is based on the period
154 where all observations are available (1961-2000) to remove the effects of this variability.

155 The analysis considers the upper tail of the wet day distribution, where wet days
156 at all stations are those recording ≥ 1 mm precipitation. Daily precipitation totals
157 exceeding a high threshold, u , of wet days are considered “extreme”. The threshold, u , is
158 defined as $\geq 95\%$ of the wet day distribution (i.e. the wettest 5%), encompassing very

159 heavy and extreme daily precipitation (Alexander et al., 2006). The 95th percentile gives a
160 different value of u at each station, but a similar mean annual frequency of extreme
161 occurrences enabling comparison across sites. Where extreme occurrences occur on two
162 consecutive days, the event with the highest precipitation total is selected as the day of
163 occurrence (Bernardara et al., 2011) to maintain the independence between events.
164 Excesses are recorded as 1, or 0 for no maxima, with no associated precipitation intensity
165 used in this analysis. For brevity, these excesses are referred to as “extreme occurrences”.

166 **2.2 Air Temperature**

167 Atmospheric moisture holding capacity is governed by temperature (e.g.
168 Trenberth, 2011), with recent changes in extreme daily precipitation associated with
169 enhanced water vapour availability (Westra et al., 2014). Evidently the relationship
170 between air temperature and extreme precipitation varies both by season and time of day,
171 as well as by the “extremity” of the precipitation (Wasko & Sharma, 2014). Changes in
172 extreme daily precipitation intensity and frequency are also associated with other
173 processes such as changes in event type (e.g. Berg et al. 2013, Prein et al., 2016) or
174 moisture transportation (Lenderink et al., 2008). The latter respond to large scale
175 atmospheric processes, represented by the covariates discussed below; while the former
176 may be well represented by air temperature and sea surface temperature (SST).

177 To maintain the assumption of identically distributed variables when considering
178 non-stationarity, covariates should vary on approximately the same time scale as the
179 model data. Precipitation occurrences appear to respond to monthly measures of air and
180 sea surface temperature (e.g. Philips & McGregor, 2002; King et al., 2014). As a result,
181 we examine the relationship between extreme occurrences and monthly maximum and

182 minimum air temperatures. Temperature data are from the CRU TS3.1 0.5° x 0.5°
183 gridded global indices of land surface monthly climate variations over the period 1901-
184 2011 (Harris et al., 2014). The selected 14 grid box time series correspond with the
185 centroid of each extreme region (refer to Figure 1).

186 **2.3 Sea Surface Temperature**

187 Increasing North Atlantic SST will likely enhance the hydrological cycle, with
188 increased oceanic evaporation generating positive moisture anomalies moving over the
189 UK and Western Europe from the Atlantic (Wang and Dong, 2010). Moisture transported
190 from the North Atlantic by wintertime extra-tropical cyclones provides a clear link
191 between SST and winter extreme daily precipitation (Lavers et al., 2013a). Although
192 convective precipitation is not directly related to North Atlantic evaporation (Lavers et
193 al., 2013a), research indicates a strong relationship between lagged North Atlantic SST
194 and summer precipitation in south and southeast England (Wilby et al., 2004; Neal and
195 Phillips, 2009). Previous research has established stronger correlations between extreme
196 precipitation and localised SST than with the wider North Atlantic basin (e.g. Phillips and
197 McGregor, 2002). We examine the relationship between extreme occurrences, and
198 concurrent and lagged monthly SST anomalies (from the mean annual cycle over 1880-
199 2010) from the HadSST2 (Rayner et al., 2005) 5° x 5° gridded monthly averages. Nine
200 grid boxes are used in this analysis, covering the UK and surrounding coastal waters.

201 **2.4 Mean Sea Level Pressure**

202 Both the North Atlantic Oscillation (NAO) and mean sea level pressure (MSLP)
203 were considered as candidate variables for the statistical model and tested for the
204 strongest relationship. The NAO is considered one of the major drivers of atmospheric
205 moisture redistribution over Europe; its positive phase is strongly correlated with extreme

206 precipitation (Fowler and Kilsby, 2002). We present a brief comparison of the
207 relationship between extreme occurrences and a principal components analysis of
208 monthly values of the NAO index (Hurrell, 1995) in Section 3.

209 Connections between MSLP, or derived indices, and temperature and
210 precipitation patterns are well established (e.g. Lavers et al. 2013a). Recent research
211 suggests extreme precipitation has greater correlation with monthly MSLP than it does
212 with the NAO index for Atlantic facing regions (Lavers et al., 2013a). Furthermore,
213 normalised MSLP over the northern Atlantic is more reliably represented in climate
214 models due to biases in climatology affecting the calculation of the NAO index (Flato et
215 al., 2013). This connection to MSLP may facilitate the use of the current research for
216 statistical downscaling analyses of future extreme daily precipitation occurrences.

217 Lavers et al. (2013a) found that MSLP centred over 63°-73°N, 10°W-5°E had the
218 strongest significant correlation with precipitation across the UK in all seasons, despite
219 fluctuations in the centre of action (i.e. the centre of a dipole in high and low pressure
220 that does not necessarily coincide with the standard Azores-Iceland definition used for
221 the NAO). Two measures of MSLP are examined here for their influence on UK extreme
222 occurrences: a monthly time series for nine individual grid boxes covering the UK and
223 adjacent oceans, and a monthly value averaged time series for the region over the larger
224 ocean area (MSLPall). The Met Office Hadley Centre's MSLP data set, HadSLP2r from
225 1901 to the present day (Allan and Ansell, 2006) on a 5° x 5° global grid is used in this
226 analysis. Nine grid boxes, corresponding with those for SST, are used here.

227 **3 Statistical Methods**

228 For a series of $t = 1, \dots, T$ days, we consider an indicator variable $J_t = 1$ when the
 229 daily precipitation exceeds a high threshold, u ; $J_t = 0$ otherwise. u is defined individually
 230 for each station using the 95th percentile of the wet day distribution. The extreme
 231 occurrences are Binomially distributed $J_t \sim \text{Binomial}(n, \pi_t)$, with $n=1$ trial per day.
 232 $\pi_t = \Pr(J_t=1)$ is the probability of a precipitation maximum, and each occurrence $J_t = 1$ is
 233 independent of previous maxima.

234 Time-varying extreme occurrences could be represented as a logistic regression
 235 model with covariates, derived from a Generalized Linear Model (GLM). That is, where
 236 the mean response variable, $E(Y_t) = \mu_t$, is described by a smooth monotonic link function,
 237 $g(\mu_t)$, of linear predictors $\beta_i x_i$ (Dobson, 2002):

238
$$g(\mu_t) = \eta_t = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n \quad (1)$$

239 and

240
$$\eta_t = \text{logit}(\pi_t) = \log\left(\frac{\pi_t}{(1 - \pi_t)}\right) \quad (2)$$

241 However, extreme occurrences are distributed non-homogeneously within the
 242 year and vary inter-annually, necessitating flexible statistical parameter estimates.
 243 Predictor terms, $g(\Phi_i)$, in the Generalized Additive Model (GAM; Hastie and Tibshirani,
 244 1990) are represented with a summation of smooth flexible non-linear functions f_y of the
 245 covariates, where η_t is the logit function presented in Equation (2).

246
$$g(\Phi_t) = \eta_t = \mathbf{X}_t \beta = \beta_0 + f_1(x_1) + \dots + f_n(x_n) \quad (3)$$

247 The relationship between β_t and \mathbf{X}_t is constrained to be smooth by a non-negative
248 smoothing parameter. Model smoothness functions, f_i , are transformed into individual
249 linear models via linear basis functions, b , for each parameter of the explanatory
250 variables (β_1, \dots, β_j).

$$251 \quad f(x) = \sum_{j=1}^q b_j(x) \beta_j \quad (4)$$

252 The simplest basis function is the polynomial basis (Equation 4). For example, a
253 four-dimension polynomial ($q=4$ or $f(x) = \beta_1 + \beta_2x + \beta_3x^2 + \beta_4x^3$) is a GLM with four
254 degrees of freedom and readily estimated by iterated re-weighted least squares (IRLS;
255 Dobson, 2002). However polynomial bases are numerically unstable and have poor
256 approximation theoretic properties so are often a poor choice for GAMs. In contrast,
257 cubic spline bases are popular as they have good approximation theoretic properties and
258 are easily reduced to a linear form (Wood and Augustin, 2002). In practice, cubic splines
259 require knot locations to be specified which can lead to model over-fitting. Using
260 penalized regression splines applies a “wiggleness” penalty so that a large number of
261 knots can be used to reduce sensitivity to their selected location, while minimising over-
262 fitting (Wood and Augustin, 2002; Wood, 2006). The penalty term is referred to as the
263 smoothing parameter, λ . As $\lambda \rightarrow \infty$ the smooth will tend towards a straight line, while
264 $\lambda \rightarrow 0$ leads to a fully interpolated spline. Penalized maximum likelihood estimation is an
265 efficient method to optimise the model complexity where multiple dimensions and
266 interactions are present (Wood, 2000).

267 There are several GAM fitting algorithms that are implemented in statistical
268 software packages (e.g. Hastie and Tibshirani, 1990; Wood, 2006). The *R* statistical

269 software environment (R Core Team, 2014) and package *mgcv* (Wood, 2006) are used
270 here. We use the default *mgcv* functions, except where there is strong motivation to
271 choose an alternative (such as cyclic splines for calendar day), to remove the subjectivity
272 in choosing an appropriate degree of model smoothness (Faraway, 2006). Variables are
273 included in the GAMs in increasing combinations in the fitting process to determine their
274 relative importance or redundancy as interactive terms. The aim is to achieve the best
275 data representation with the greatest degree of parsimony tested by measures such as the
276 deviance statistic, Generalized Cross Validation or Akaike Information Criterion
277 (Akaike, 1974). Model simplifications can also be achieved through combinations of
278 linear or quadratic terms in addition to flexible smoothness (Wood, 2006).

279 The Poisson distribution is a special case of the Binomial distribution, with $n \rightarrow \infty$
280 and $\pi_t \rightarrow 0$. As extreme occurrences are intermittent with a small sample per year, the
281 annual frequency of extreme occurrences is Poisson distributed. Consequently, the
282 statistical model can also be extended to examine long-term changes in the annual
283 frequency of extreme occurrences. We make use of this relationship in Section 4.2 to
284 identify potential covariates for the Binomial GAM using a log linear regression, or
285 Poisson regression model. The Poisson regression takes the form of a GLM, Equation (1),
286 to describe the log of the Poisson mean with nonlinear functions of the predictors. That
287 is, $\ln(\mu_t) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n$.

288 All extreme occurrence time series ($J_t = 1$ or 0) are collated into a single pooled
289 time series for each extreme region (Jones et al., 2014) shown in Figure 1. The pooled
290 time series improves estimates of the frequency distribution of extreme occurrences and
291 extends the effective time series for model validation. As a large synoptic event could

292 generate extreme daily precipitation at several stations on the same day or over two days,
293 repeated extremes are discarded leaving a single event per day. To maintain
294 independence, and following the same criteria used for individual stations, the highest
295 precipitation total of two consecutive days is considered as the day of occurrence.

296 **4 Exploratory Analysis**

297 Exploratory analyses presented here inform the statistical model development and
298 are used to test the significance of any observed changes. The exploratory analyses
299 examine whether there is a seasonal pattern in extreme occurrences, and the correlation
300 with likely atmospheric and oceanic driving conditions. The remainder of this article
301 focuses on the Southeast (SE) region, with the remaining results in the supplementary
302 information.

303 **4.1 Seasonality**

304 Figure 2 shows frequency density distributions of regional extreme occurrences
305 throughout the year for each extreme region. Frequency density is estimated for each
306 region using the pooled record of extreme occurrences to obtain a summation of events
307 per calendar day, which is then divided by the total number of events and a smooth spline
308 applied. The extremes occur mostly in summer (June-August) or autumn (September-
309 November) months with very few during spring (March-April). Rodda et al. (2009; Figures
310 5 and 6) presented a similar bimodal frequency distribution of UK extreme daily
311 precipitation occurrences, peaking in mid-summer and late autumn. However, Rodda et al.
312 (2009) included extreme occurrences from all UK observation stations in a single analysis,
313 potentially creating an unrealistic multimodal distribution arising from geographically
314 different extreme precipitation climatologies. The apparent multimodality may be also be

315 influenced by sampling variability, where natural fluctuations appear more important
316 than the climatological response (Kundzewicz and Robson, 2000). However, it
317 demonstrates that a simple sinusoidal form is unlikely to describe the annual frequency
318 distribution of extreme daily precipitation occurrences in the UK adequately.

319 Multimodal seasonality, with peaks in the frequency distribution in summer and
320 again in late autumn, is most evident in regions along the UK east coast (e.g. NE and EA).
321 In contrast, exposed western regions (notably NHI, NI, MW and SOL) exhibit a unimodal
322 frequency distribution of extreme occurrences encompassing autumn and/or winter.
323 These regional differences are representative of the dominant weather systems around the
324 UK. Extreme daily precipitation in the northwest is mainly caused by westerly airflow
325 from the Atlantic, while extreme precipitation in the southeast is dominated by easterly
326 flows from the North Sea (Maraun et al., 2009).

327 The frequency distribution of extreme occurrences by day and year for South
328 England (SE) is shown in Figure 3; a similar pattern is produced by all regions. Each
329 colour band reflects the relative frequency density of extreme occurrences within a 15-
330 day/15-year smooth. “Edge effects” are mitigated by plotting only the period where all
331 SE region observation stations were effective (1950-1999). The highest frequency density
332 of extreme occurrences is shown in pink/purple and is centred around July-November.
333 This figure illustrates the multi-decadal fluctuations in both the annual distribution and
334 total number of events, while highlighting the broader seasonal period when extreme
335 occurrences are most likely to occur. There is a suggestion of longer-term changes in the
336 duration and timing of the peak in extreme occurrences. This figure appears to show that
337 most extreme events occurred earlier in the autumn during the 1990s than during the

338 1960s. The period when they are most likely is also several days shorter, i.e. an
339 increasing number of extreme occurrences are more likely to be clustered in time.

340 **4.2 Identification of potential covariates**

341 Poisson regression models are used to examine the relationship between extreme
342 occurrences and the candidate covariates. Relationships are tested at the 5% significance
343 level against the χ^2 distribution. We consider covariates to have field significance when
344 the correlation between a variable and extreme occurrences is significant at >40 stations.
345 Figure 4(a-d) presents the results of the most important correlations between extreme
346 occurrences and the coincident monthly variables. These results are indicative only; the
347 relationship between extreme occurrences and covariates is explored in greater detail
348 with the GAM to identify the best model configuration.

349 While concurrent and lagged monthly SST anomalies for the nearest grid box are
350 correlated with the annual number of extreme occurrences, the most significant
351 correlations are found for the previous month (i.e. lag 1; SST1 henceforth). A strongly
352 negative correlation across most of the UK is illustrated in Figure 4a; the few positive
353 stations are only weakly positive, although the correlation is still significant. Monthly
354 MSLP over the larger Atlantic region (63°-73°N, 10°W-5°E) has the most significant
355 correlation with extreme occurrences (Figure 4b) and is the inverse of the monthly NAO
356 correlation pattern (Figure 4c), i.e. positive MSLP in the northwest and negative MSLP in
357 the southeast.

358 Monthly maximum and minimum air temperatures and maximum diurnal
359 temperature range are also examined for correlation and have significant relationships.
360 Although the correlation with extreme occurrences is not significant, monthly minimum

361 air temperature has the greatest correlation (Figure 4d) and is explored further with the
362 GAM. The correlation between the annual count of extreme occurrences and average
363 number of days (interval) between extreme occurrences presents a strong exponential
364 relationship (Figure 4e). Conditional statistical analyses require that the covariates and
365 the data are identically distributed and vary at the same scale. Using monthly variables
366 ensures that the data vary at the same time-scales. Confirming that both the distribution
367 of the extreme occurrences and their annual frequency are from the exponential family
368 validates the assumption of identical distributions.

369 Figure 5 presents pair plot correlations between the same atmospheric and oceanic
370 variables against the calendar day of extreme occurrences for South East England (SE).
371 All other extreme regions display similar patterns in their paired correlations (not shown).
372 SST1 and monthly minimum daily air temperature have the expected seasonal patterns
373 with minima in either April (SST1) or January (air temperature).

374 Monthly MSLP over the wider Atlantic domain (MSLPall) is linearly correlated
375 with the NAO index, confirming that only one of these measures is required in the final
376 GAM. Both measures of MSLP are better correlated with extreme occurrences than
377 NAO. However, monthly MSLPall, over the wider Atlantic domain, shows greater
378 coherence with extreme occurrences than individual grid box measures, and is selected as
379 a potential covariate. This is supported by recent research identifying that sea level
380 pressures to the west of Scotland explain more of the variance in precipitation (~80%) for
381 a larger proportion of the UK than the NAO index (~5.4% ; van Oldenborgh et al., 2015).
382 As our aim is to identify a covariate that is reliably reproduced by most global circulation
383 models, the NAO index is not used in later analyses. Standard tests for multi-collinearity

384 in the remaining variables (e.g. Faraway, Ch9.4 2006) did not identify that other
385 covariates should be removed.

386 **5 Statistical Analysis of Long-term Changes**

387 Non-stationary Binomial distributions, that is where the distribution parameters
388 vary with time, are fitted to the regional time series of extreme occurrences. A subset of
389 1961-2000 is selected to fit the model, while the extended time series to 2010 is used to
390 validate the statistical model. Finally, covariates from 1901-2011 are used to test for
391 statistical evidence of longer-term changes in behaviour. The variables examined for
392 importance in the regional GAMs are listed in Table 1. Explicit representation of
393 seasonality is tested in addition to seasonal fluctuations represented by temperature and
394 MSLP covariates.

395 **5.1 Model construction**

396 Three general methods are suggested by Wood (2006) for incorporating smooth
397 functions of the covariates:

- 398 • All terms are additive of the form $f_1(x) + f_2(y)$;
- 399 • Variable coefficients $z + zf(x)$ where x varies smoothly, but its effect is modified by
400 the constant covariate z (Hastie and Tibshirani, 1990);
- 401 • Bivariate smoothing $f(x,y)$ where x and y vary jointly to represent a more complex
402 interaction of covariates.

403 *Mgcv* has several fitting options for smoothing parameter functions, i.e. f from
404 above, including cubic and polynomial regression splines; the default are thin plate

405 regression splines. We employ cubic regression splines fitted by penalized likelihood as
406 these outperform the default option in efficiency and efficacy.

407 Covariates are tested individually in the Binomial family GAM to confirm their
408 significance, before testing combinations of multiple covariates. Linear and non-linear
409 functions of the smoothers are also examined for the best representation of extreme
410 occurrences. Linear combinations would improve the parsimony and efficiency of the
411 statistical model, but may not be fully representative. Each model combination is tested
412 objectively using the AIC and deviance statistics, and subjectively with the distribution of
413 the fitted model residuals using quantile plots and histograms (Wood, 2006). Covariates
414 are included in the model in the order of greatest explanation of data variance, identified
415 from the individual testing. To balance the overall model goodness of fit against
416 parsimony a term is dropped from the model when all three of the following criteria are
417 satisfied (Wood and Augustin, 2002):

- 418 • the term effective degree of freedom is close to its lower limit;
- 419 • the confidence region for the smooth contains zero everywhere;
- 420 • removing the term reduces the deviance statistic, or other relative comparison
421 measure such as AIC.

422 Stepwise iteration to reject unnecessary covariates (Hastie and Tibshirani, 1990)
423 is time consuming, particularly when many covariates are involved. A modified approach
424 where only the significant terms from less complex models are included in later model
425 iterations is used here.

426 The final Binomial distribution takes the form in Equation 5, with covariate
427 smooth terms presented in Table S1 in the Supplement. All regional models follow this
428 form, only the coefficients (β_i) and smoothing parameters (f_i) differ between regions.

$$429 \quad \text{logit } \pi(t) = \beta_0 + f_1(d_t) + f_2(P_t) + f_3(\Theta_{Nt}) + f_4(ST1_t) \quad (5)$$

430 Figure 6 is an example of the GAM smoothing functions for the SE region; ± 2
431 standard error bars are indicated by dashed lines. The day of occurrence (d_t) is the most
432 important covariate in the SE region, followed by monthly minimum air temperature
433 (Θ_{Nt}). Seasonality arises from the interaction between d_t and Θ_{Nt} , modifying the
434 sinusoidal shape of d_t to develop the characteristic bimodal seasonality in extreme
435 occurrences. There is a strong positive correlation with Θ_{Nt} related to summer convection
436 that tails off at higher air temperatures, consistent with changes in Clausius-Clapeyron
437 scaling (Westra et al., 2014). In contrast, both lagged monthly SST anomalies ($ST1_t$) and
438 monthly MSLP (P_t) are negatively correlated; however the relationship with $ST1_t$ is not
439 significant in SE.

440 Similar GAM smoothing functions were developed for all regions (not illustrated).
441 P_t and d_t , are the most important covariates for most regions and models, explaining
442 between 34-89% of the variability in extreme occurrence frequency; the next most
443 important variable is monthly minimum air temperature (Θ_{Nt}). $ST1_t$ and P_t differ in
444 importance between regions, and display a north-south divide with the northwest UK
445 dominated by MSLP and southeast by lagged SST anomalies. The relationship with Θ_{Nt}
446 is also variable by region: in the north and west there is a strong negative correlation
447 reflecting the dominance of winter frontal systems, with a reversal around the freezing
448 point. In contrast, southern and eastern regions have a positive correlation similar to that

449 in SE. Air temperature could be a response to the precipitation rather than a driver,
450 however the two are strongly interlinked and it is difficult to elicit which is the
451 precursor.

452 **5.2 Model validation**

453 Each regional model is validated objectively using a leave-one-out cross
454 validation to estimate the parameters. Median scale parameters from the cross validation
455 are very close to the original model estimates (not shown). A more subjective test of
456 model validity is to compare the observed extreme occurrence annual frequency
457 distribution against model samples. Visual validation compares 1901-2010 observations
458 against 100 samples per day from the regional Binomial distributions. Figure 7a
459 illustrates the frequency density of extreme occurrences by calendar day for the
460 observations in red and samples in grey. There is a good match between the observations
461 and the sampled distributions, without excessive reproduction of day-to-day variability or
462 implied over-fitting (Wood, 2006). Quantile-quantile plots (not shown) of the same
463 samples also indicate good correspondence between the observed and simulated
464 distribution quantiles for the Binomial distribution.

$$465 \quad N_{rel,i} = \frac{N_i}{S_i} \quad (6)$$

466 The annual relative frequency of simulated extreme occurrences is also evaluated
467 and shown in Figure 7b. Relative frequency is the number of extreme occurrences at one
468 location, and calculated as the ratio of extreme occurrences (N) in the i^{th} year to the
469 number of observation stations (S) in the same year (Equation 6). The simulations are
470 derived from all available stations in 1961-2000, while observations may have fewer
471 active stations prior to 1961 or after 2000. Results for all regions are included in

472 Supplemental Figures S1 and S2. There is generally good correspondence between the
473 observed and simulated results.

474 **5.3 Model results**

475 Monthly observed covariate data (1901-2010) are applied to the regional
476 Binomial distributions to predict the daily probability of an extreme occurrence for 109
477 years. We are most concerned with the period of consecutive days to weeks with the
478 highest probability of extreme occurrences; that is the 0.95 quantile of the regional
479 probability distribution. Figure 8 illustrates changes in extreme occurrences for South
480 East England; the same figures for all regions are included in Supplemental Figures S3 to
481 S6.

482 Figure 8a shows the approximate calendar day with the highest probability of an
483 extreme occurrence (red squares) for 1901-2010 (i.e. ≥ 0.95 quantile of the distribution of
484 probabilities). Confidence intervals are derived from the range of dates at which the
485 highest probability \pm standard errors exceeded the same threshold. Horizontal lines
486 denote the first day of the month from July to February. A smoothed running mean (black
487 dash) indicates that while the period with the highest likelihood of an extreme occurrence
488 is variable, there is no discernible change in seasonality in the SE region. Similar
489 responses can be seen for all regions, with nine regions also showing changes in the date
490 when the probability of an extreme occurrence is greatest (Figure S3). Over the 1901-
491 2010 period, SH, SOL and NW show changes to significantly earlier dates when the
492 probability of an extreme precipitation occurrence is highest; NHI, NE and HUM all
493 show significant changes to a later date (for the F-statistic at a 5% significance level). In

494 almost all cases, the changes equate to the highest probability of extreme occurrences
495 between September and November.

496 Figure 8b illustrates the number of days where the daily probability of an extreme
497 precipitation occurrence exceeds the 0.95 quantile of annual probabilities. We consider
498 the total number of days as the duration of very high probability of an extreme
499 occurrence. In the SE region the threshold probability is $p \geq 0.16$. Red circles indicate the
500 most likely duration; confidence intervals are in grey, calculated as for the dates in Figure
501 8a. Zero duration indicates that the daily probability does not exceed $p=0.16$ during that
502 year. The black dashed 5-year running mean displays an overall decrease in mean
503 duration of very high probability of an extreme occurrence by about 1 week. Other
504 regions (Figure S4) display more variability in the duration of very high probability, with
505 seven others (NHI, ES, FOR, SOL, NW, MW, WC and HUM) showing significant (at
506 5% significance) increases in the duration.

507 Figure 8c shows the highest daily probability of an extreme occurrence each year,
508 confidence intervals are derived from the distribution standard errors. Higher
509 probabilities (e.g. 1920s, mid-1980s, and early 2000s) corroborate records of very wet
510 periods with widespread flooding. A smoothed running mean (black dash) indicates a
511 periodic fluctuation in the probability of extreme occurrences. The 0.95 quantile
512 threshold of the distribution of probabilities for other regions lies in a range between 0.07
513 (Mid Wales) and 0.27 (Humber). Eight of the regions indicate increases in the highest
514 daily probability of extreme occurrences (Figure S5) which, where combined with a
515 longer duration of the highest probability, could lead to increased flooding. Considerable

516 variability indicates that the highest probability of extreme occurrences is most sensitive
517 to multi-annual fluctuations in large-scale processes.

518 The mean daily probabilities of extreme occurrences over 20-year time-slices are
519 compared in Figure 8d. Estimating probability over a longer timescale focuses on the
520 climatological responses rather than inter-annual fluctuations. There is little change in the
521 period of year with the lowest probability of extreme occurrences (F-M-A-M). However,
522 the summer (J-A-S) and winter (N-D-J) bi-modal peak when extreme occurrences are
523 most likely has shifted towards a single autumnal peak in the highest probability. A
524 similar pattern of changes in the timing, and disappearance of the bi-modal peak, is
525 apparent in all other regions (Figure S6). In some regions such as SH, the changes are
526 very slight and difficult to discern; while in others, such as NHI and SE, they are more
527 definitive. All regions indicate a change in the season with the highest probability of
528 extreme occurrences toward the autumn (S-O-N).

529 Many patterns present in the observations are replicated by the simulation results,
530 confirming that they did not arise solely through sampling variability. The time series of
531 highest daily probability of extreme occurrences also confirms reported increases in the
532 frequency of late summer and autumn extreme precipitation (Jones et al., 2013). Several
533 devastating floods in the UK during the recent decade arose after a sequence of heavy
534 precipitation days occurred in short succession (e.g. Lavers et al., 2011). The modelled
535 results indicate that fluctuating years with shorter or longer durations of very high
536 probability of extreme occurrences correspond with known natural fluctuations in wet
537 and dry years. The simulated results also indicate a longer-term change in the time when
538 extreme occurrences are most likely.

539 **6 Discussion and Conclusions**

540 This article examined the annual frequency distribution of extreme daily
541 precipitation occurrences, or seasonal pattern, to characterize spatial and temporal
542 differences across the UK. It was shown that extreme and very heavy daily precipitation
543 does not occur uniformly throughout the year. There are also considerable regional
544 differences, arising from the relative proximity to the North Atlantic Ocean or North Sea
545 and dominating westerly or easterly air flows (Maraun et al., 2009). There are definite
546 weeks or months during which extreme daily precipitation is most likely to occur, with
547 some regions displaying multimodal seasonality. Oceanic and atmospheric variables were
548 shown to have significant correlations with extreme occurrences and so were valuable
549 proxy information for the probability of extremes.

550 The seasonality of the highest probability of extreme daily precipitation varies
551 across the UK, contrasting with mean wet day occurrences which are more likely to occur
552 over autumn and winter months for all regions (Rodda et al., 2009). The most intense
553 events along the north and west of the UK are associated with large wintertime synoptic
554 systems (multi-day), which generate both frequent wet days and high intensity
555 precipitation. In contrast, short duration (sub-daily) summer convective storms are more
556 common in the south and east of the UK (e.g. van Delden et al., 2001). Multimodal
557 seasonality in event probability likely arises where both large-scale synoptic systems and
558 convective systems regularly dominate regional weather. Recent research suggests that
559 Atmospheric Rivers have a role in the development of winter floods over the UK but
560 have little influence over summer extreme precipitation (Lavers et al., 2011, Champion et
561 al., 2015). While such events usually extend over several days, and this analysis focussed

562 on single day extremes, this supports the likely cause of multimodal seasonality.
563 Projections of future increases in North Atlantic winter atmospheric rivers (Lavers et al.,
564 2013b) correspond with the shift towards late autumn extreme occurrences identified here.

565 Allowances for non-stationarity in regional extreme precipitation or flood
566 estimates is not common practice (Jakob et al., 2011). Most analyses have focused on
567 traditional extreme value analyses to estimate the likely return frequency of specific
568 events (Ghil et al., 2011). Very few have considered the frequency distribution of
569 extreme daily precipitation occurrences within the year. This article presented a statistical
570 model that accommodates non-stationarity in the probability of extreme daily
571 precipitation by allowing distribution parameters to vary in response to external
572 covariates. Statistical model results indicated that the probability of several extreme
573 events occurring in a few weeks has increased in most regions from 1901-2010.

574 A Generalized Additive Model (GAM) was constructed using observation data
575 (1961-2000) and validated using all observation data (>1901-2010). Covariate data for
576 1901-2010 were used to test the changes in the daily probability of extreme occurrences.
577 The statistical model reflected natural climate variability from year to year in the calendar
578 days when the probability of an extreme occurrence is at its highest. Smoothing over
579 several years to decades illustrated that all regions now have a higher likelihood of
580 autumnal extreme precipitation. This equates to extreme occurrences earlier in the north
581 and west (formerly winter dominated) and later in the south and east (formerly late
582 summer and autumn). This change in timing was accompanied in eight regions with
583 significant increases in the probability of extreme occurrences during the autumn. The
584 same regions also had significant increases in the number of days with a very high

585 probability of extreme precipitation, or an extended duration of the peak season. These
586 results indicate a higher probability of several extreme occurrences in succession, or a
587 potential clustering of events. While intensity was not included in this analysis, the
588 combination of a change in seasonality to coincide with the wetter period of the year and
589 saturated ground conditions indicates a potential increase in flood likelihood.

590 Seasonality, represented in the model jointly by monthly minimum temperature
591 and calendar day, is the principal driver of extreme daily precipitation event frequency in
592 the UK. SST and MSLP were found to differ in importance between regions, displaying a
593 north-south divide with the north and west of the UK dominated by MSLP and the south
594 and east dominated by SST. Monthly minimum air temperature had a positive
595 relationship with extreme occurrences; this decreased beyond a high temperature, similar
596 to results from other research on changes in Clausius-Clapeyron scaling (e.g. Westra et
597 al., 2014).

598 The statistical models presented here could be used in combination with covariates
599 derived from climate model projections to improve future estimates of changes in the
600 seasonal distribution of extreme daily precipitation, complementing other recent research
601 (Schindler et al., 2012). Further refinements to the model could be made to simulate the
602 sequences of extreme occurrences and their dependence structure. For instance, a Markov
603 Chain or a Cox-Process model could be used to simulate the interval between extreme
604 occurrences more effectively (e.g. Smith and Karr, 1986). Alternatively, an inhomogeneous
605 Poisson point process model could enable examination of the intensity of extreme daily
606 precipitation as part of the same analysis (e.g. Katz, 2010).

607 While the changes in the characteristics of extreme daily precipitation occurrences
608 are significant, they are also fairly small. A potential follow on study could assess the
609 strength of the link between changes in extreme occurrences to changes in flood
610 characteristics. For example the extended model would examine the dependence of large
611 floods on the monthly minimum air temperature, monthly MSLP and lagged SST.

612 **Acknowledgments**

613 The R analysis software (*R Development Core Team*, 2011) and packages mgcv
614 (Wood, 2006), and plotrix (*Lemon*, 2006) were used for analyses and figures in this paper.
615 Covariate data are available from the KNMI Climate Explorer portal
616 <http://www.climexp.knmi.nl> and the NCAR Climate Data Guide
617 <https://www.climatedataguide.ucar.edu>. R-scripts and extreme daily precipitation
618 occurrence data can be made available on request for non-commercial use.

619 NCAR is sponsored by the National Science Foundation. M.R.T. was partially
620 supported by NSF EASM grant S1048841, the NCAR Weather and Climate Assessment
621 Science Program and a NERC funded Postgraduate Research Studentship NE/G523498/1
622 (2008-2012). H.J.F. was supported by a NERC Postdoctoral Fellowship Award
623 NE/D009588/1 (2006–2010) and is now funded by the Wolfson Foundation and the
624 Royal Society as a Royal Society Wolfson Research Merit Award holder (WM140025).

625 **References**

626 Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans.*
627 *Automat. Contr.* 19, 716–723.

628 Alexander, L. V, Zhang, X., Peterson, T.C., Caesar, J., Gleason, B., Klein Tank,
629 A.M.G., Haylock, M., Collins, D., Trewin, B., Rahimzadeh, F., Tagipour, A., Rupa

630 Kumar, K., Revadekar, J., Griffiths, G., Vincent, L., Stephenson, D.B., Burn, J., Aguilar,
631 E., Brunet, M., Taylor, M., New, M., Zhai, P., Rusticucci, M., Vazquez-Aguirre, J.L.,
632 2006. Global observed changes in daily climate extremes of temperature and
633 precipitation. *J. Geophys. Res.* 111, D05109. doi:10.1029/2005jd006290

634 Allan, R., Ansell, T., 2006. A New Globally Complete Monthly Historical
635 Gridded Mean Sea Level Pressure Dataset (HadSLP2): 1850–2004. *J. Clim.* 19, 5816–
636 5842. doi:10.1175/JCLI3937.1

637 Bernardara, P., Andreewsky, M., Benoit, M., 2011. Application of regional
638 frequency analysis to the estimation of extreme storm surges. *J. Geophys. Res.* 116,
639 C02008. doi:10.1029/2010JC006229

640 Beuchat, X., Schaepli, B., Soutter, M., Mermoud, A., 2012. A robust framework
641 for probabilistic precipitations downscaling from an ensemble of climate predictions
642 applied to Switzerland. *J. Geophys. Res.* 117, D03115. doi:10.1029/2011JD016449

643 Bouwer, L.M., Vermaat, J.E., Aerts, J.C.J.H., 2008. Regional sensitivities of
644 mean and peak river discharge to climate variability in Europe. *J. Geophys. Res.* 113,
645 D19103. doi:10.1029/2008JD010301

646 Champion, A.J., Allan, R.P., Lavers, D. A., 2015. Atmospheric rivers do not
647 explain UK summer extreme rainfall. *J. Geophys. Res. Atmos.* 120, 6731–6741.
648 doi:10.1002/2014JD022863

649 Chandler, R.E., 2005. On the use of generalized linear models for interpreting
650 climate variability. *Environmetrics* 16, 699–715. doi:10.1002/env.731

651 Chandler, R.E., Wheater, H.S., 2002. Analysis of rainfall variability using
652 generalized linear models: A case study from the west of Ireland. *Water Resour. Res.* 38,
653 10–1–10–11. doi:10.1029/2001WR000906

654 Chavez-Demoulin, V., Davison, A.C., 2005. Generalized additive modelling of
655 sample extremes. *J. R. Stat. Soc. Ser. C (Applied Stat.* 54, 207–222. doi:10.1111/j.1467-
656 9876.2005.00479.x

657 Dhakal, N., Jain, S., Gray, A., Dandy, M., Stancioff, E., 2015. Nonstationarity in
658 seasonality of extreme precipitation: A nonparametric circular statistical approach and its
659 application. *Water Resour. Res.* 51, 4499–4515. doi:10.1002/2014WR016399

660 Dobson, A.J., 2002. An introduction to generalized linear models. Chapman &
661 Hall/CRC.

662 Faraway, J.J., 2006. Extending the linear model with R: generalized linear, mixed
663 effects and nonparametric regression models. Chapman & Hall/CRC.

664 Flato, G., J., Marotzke, Abiodun, B., Braconnot, P., Chou, S.C., Collins, W., Cox,
665 P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C.,
666 Kattsov, V., Reason, C., Rummukainen, M., 2013. Evaluation of Climate Models, in:
667 Stocker, T.F., D. Qin, Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A.,
668 Xia, Y., Bex, V., Midgley, P.M. (Eds.), *Climate Change 2013: The Physical Science*
669 *Basis. Contribution of Working Group I to the Fifth Assessment Report of the*
670 *Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge,
671 UK, pp. 741–866. doi:10.1017/CBO9781107415324.020

672 Fowler, H.J., Kilsby, C.G., 2002. Precipitation and the North Atlantic Oscillation:
673 a study of climatic variability in northern England. *Int. J. Climatol.* 22, 843–866.

674 Ghil, M., Yiou, P., Hallegatte, S., Malamud, B.D., Naveau, P., Soloviev, A.,
675 Friederichs, P., Keilis-Borok, V., Kondrashov, D., Kossobokov, V., Mestre, O., Nicolis,
676 C., Rust, H.W., Shebalin, P., Vrac, M., Witt, A., Zaliapin, I., 2011. Extreme events:
677 dynamics, statistics and prediction. *Nonlinear Process. Geophys.* 18, 295–350.
678 doi:10.5194/npg-18-295-2011

679 Hannaford, J., Marsh, T.J., 2008. High-flow and flood trends in a network of
680 undisturbed catchments in the UK. *Int. J. Climatol.* 28, 1325–1338. doi:10.1002/joc.1643

681 Harris, I., Jones, P.D., Osborn, T.J., Lister, D.H., 2014. Updated high-resolution
682 grids of monthly climatic observations - the CRU TS3.10 Dataset. *Int. J. Climatol.* 34,
683 623–642. doi:10.1002/joc.3711

684 Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Taylor and
685 Francis.

686 Huntingford, C., Marsh, T., Scaife, A.A., Kendon, E.J., Hannaford, J., Kay, A.L.,
687 Lockwood, M., Prudhomme, C., Reynard, N.S., Parry, S., Lowe, J.A., Screen, J.A.,
688 Ward, H.C., Roberts, M., Stott, P.A., Bell, V.A., Bailey, M., Jenkins, A., Legg, T., Otto,
689 F.E.L., Massey, N., Schaller, N., Slingo, J., Allen, M.R., 2014. Potential influences on the
690 United Kingdom's floods of winter 2013/14. *Nat. Clim. Chang.* 4, 769–777.
691 doi:10.1038/nclimate2314

692 Hurrell, J.W., 1995. Decadal Trends in the North Atlantic Oscillation: Regional
693 Temperatures and Precipitation. *Science* (80-.). 269, 676–679.

694 Hyndman, R.J., Grunwald, G.K., 2000. Applications: Generalized Additive
695 Modelling of Mixed Distribution Markov Models with Application to Melbourne's
696 Rainfall. *Aust. N. Z. J. Stat.* 42, 145–158. doi:10.1111/1467-842x.00115

697 Jakob, D., Karoly, D.J., Seed, A., 2011. Non-stationarity in daily and sub-daily
698 intense rainfall - Part 2: Regional assessment for sites in south-east Australia. *Nat.*
699 *Hazards Earth Syst. Sci.* 11, 2273–2284. doi:10.5194/nhess-11-2273-2011

700 Jones, M.R., Blenkinsop, S., Fowler, H.J., Kilsby, C.G., 2014. Objective
701 classification of extreme rainfall regions for the UK and updated estimates of trends in
702 regional extreme rainfall. *Int. J. Climatol.* 34, 751–765. doi:10.1002/joc.3720

703 Jones, M.R., Fowler, H.J., Kilsby, C.G., Blenkinsop, S., 2013. An assessment of
704 changes in seasonal and annual extreme rainfall in the UK between 1961 and 2009. *Int. J.*
705 *Climatol.* 33, 1178–1194. doi:10.1002/joc.3503

706 Kallache, M., Vrac, M., Naveau, P., Michelangeli, P., 2011. Nonstationary
707 probabilistic downscaling of extreme precipitation. *J. Geophys. Res.* 116, 1–15.
708 doi:10.1029/2010JD014892

709 Katz, R.W., 2010. Statistics of extremes in climate change. *Clim. Change* 100,
710 71–76. doi:10.1007/s10584-010-9834-5

711 King, A.D., Klingaman, N.P., Alexander, L. V., Donat, M.G., Jourdain, N.C.,
712 Maher, P., 2014. Extreme rainfall variability in Australia: Patterns, drivers, and
713 predictability. *J. Clim.* 27, 6035–6050. doi:10.1175/JCLI-D-13-00715.1

714 Kundzewicz, Z.W., Robson, A., 2000. Detecting trend and other changes in
715 hydrological data. World Meteorological Organization.

716 Lavers, D., Allan, R.P., Villarini, G., Lloyd-Hughes, B., Brayshaw, D.J., Wade,
717 A.J., 2013b. Future changes in atmospheric rivers and their implications for winter
718 flooding in Britain. *Environ. Res. Lett.* 8, 34010. doi:10.1088/1748-9326/8/3/034010

719 Lavers, D., Prudhomme, C., Hannah, D.M., 2013a. European precipitation
720 connections with large-scale mean sea-level pressure (MSLP) fields. *Hydrol. Sci. J.* 58,
721 310–327. doi:10.1080/02626667.2012.754545

722 Lavers, D.A., Allan, R.P., Wood, E.F., Villarini, G., Brayshaw, D.J., Wade, A.J.,
723 2011. Winter floods in Britain are connected to atmospheric rivers. *Geophys. Res. Lett.*
724 38, L23803. doi:10.1029/2011gl049783

725 Maraun, D., Osborn, T.J., Rust, H.W., 2009. The influence of synoptic airflow on
726 UK daily precipitation extremes. Part I: Observed spatio-temporal relationships. *Clim.*
727 *Dyn.* 36, 261–275. doi:10.1007/s00382-009-0710-9

728 Neal, R.A., Phillips, I.D., 2009. Summer daily precipitation variability over the
729 East Anglian region of Great Britain. *Int. J. Climatol.* 29, 1661–1679.
730 doi:10.1002/joc.1826

731 Pal, I., Anderson, B.T., Salvucci, G.D., Gianotti, D.J., 2013. Shifting seasonality
732 and increasing frequency of precipitation in wet and dry seasons across the U.S.
733 *Geophys. Res. Lett.* 40, 4030–4035. doi:10.1002/grl.50760

734 Phillips, I.D., McGregor, G.R., 2002. The relationship between monthly and
735 seasonal South-west England rainfall anomalies and concurrent North Atlantic sea
736 surface temperatures. *Int. J. Climatol.* 22, 197–217. doi:10.1002/joc.726

737 Pui, A., Lal, A., Sharma, A., 2011. How does the Interdecadal Pacific Oscillation
738 affect design floods in Australia? *Water Resour. Res.* 47. doi:10.1029/2010WR009420

739 R Core Team, 2014. R: A Language and Environment for Statistical Computing.

740 Rayner, N.A., Brohan, P., Parker, D., Folland, C.K., Kennedy, J.J., Vanicek, M.,
741 Ansell, T.J., Tett, S., 2005. Improved analyses of changes and uncertainties in sea surface
742 temperature measured in situ since the mid-nineteenth century: The HadSST2 Dataset. *J.*
743 *Clim.* 19, 446–469.

744 Rodda, H.J.E., Little, M.A., Wood, R.G., MacDougall, N., McSharry, P.E., 2009.
745 A digital archive of extreme rainfalls in the British Isles from 1866 to 1968 based on
746 British Rainfall. *Weather* 64, 71–75. doi:10.1002/wea.354

747 Rust, H.W., Maraun, D., Osborn, T.J., 2009. Modelling seasonality in extreme
748 precipitation. *Eur. Phys. J. - Spec. Top.* 174, 99–111.

749 Sapiano, M.R.P., Stephenson, D.B., Grubb, H.J., Arkin, P.A., 2006. Diagnosis of
750 Variability and Trends in a Global Precipitation Dataset Using a Physically Motivated
751 Statistical Model. *J. Clim.* 19, 4154–4166. doi:10.1175/JCLI3849.1

752 Schindler, A., Maraun, D., Toreti, A., Luterbacher, J., 2012. Changes in the
753 annual cycle of heavy precipitation across the British Isles within the 21st century.
754 *Environ. Res. Lett.* 7, 044029. doi:10.1088/1748-9326/7/4/044029

755 Serinaldi, F., Kilsby, C.G., 2014. Simulating daily rainfall fields over large areas
756 for collective risk estimation. *J. Hydrol.* 512, 285–302. doi:10.1016/j.jhydrol.2014.02.043

757 Smith, J.A., Karr, A.F., 1986. Flood Frequency Analysis Using the Cox
758 Regression Model. *Water Resour. Res.* 22, 890–896. doi:10.1029/WR022i006p00890

759 Stephens, E., Day, J.J., Pappenberger, F., Cloke, H., 2015. Precipitation and
760 floodiness. *Geophys. Res. Lett.* 42, 10,316–10,323. doi:10.1002/2015GL066779

761 Trambly, Y., Neppel, L., Carreau, J., 2011. Brief communication “Climatic
762 covariates for the frequency analysis of heavy rainfall in the Mediterranean region.” *Nat.*
763 *Hazards Earth Syst. Sci.* 11, 2463–2468. doi:10.5194/nhess-11-2463-2011

764 Trenberth, K.E., 2011. Changes in precipitation with climate change. *Clim. Res.*
765 47, 123–138.

766 Underwood, F.M., 2009. Describing long-term trends in precipitation using
767 generalized additive models. *J. Hydrol.* 364, 285–297. doi:10.1016/j.jhydrol.2008.11.003

768 van Delden, A., 2001. The synoptic setting of thunderstorms in western Europe.
769 *Atmos. Res.* 56, 89–110. doi:10.1016/S0169-8095(00)00092-2

770 van Oldenborgh, G.J., Stephenson, D.B., Sterl, A., Vautard, R., Yiou, P.,
771 Drijfhout, S.S., Storch, H. Von, Dool, H. Van Den, 2015. Drivers of the 2013 / 14 winter
772 floods in the UK. *Nat. Clim.* 5, 490–491. doi:10.1038/nclimate2612

773 Verdin, A., Rajagopalan, B., Kleiber, W., Katz, R.W., 2014. Coupled stochastic
774 weather generation using spatial and generalized linear models. *Stoch. Environ. Res. Risk*
775 *Assess.* 29, 347–356. doi:10.1007/s00477-014-0911-6

776 Wang, C., Dong, S., 2010. Is the basin-wide warming in the North Atlantic Ocean
777 related to atmospheric carbon dioxide and global warming? *Geophys. Res. Lett.* 37,
778 L08707. doi:10.1029/2010GL042743

779 Wasko, C., Sharma, A., 2014. Quantile regression for investigating scaling of
780 extreme precipitation with temperature. *Water Resour. Res.* 50, 3608–3614.
781 doi:10.1002/2013WR015194

782 Westra, S., Evans, J.P., Mehrotra, R., Sharma, A., 2013. A conditional
783 disaggregation algorithm for generating fine time-scale rainfall data in a warmer climate.
784 *J. Hydrol.* 479, 86–99. doi:10.1016/j.jhydrol.2012.11.033

785 Westra, S., Fowler, H.J., Evans, J.P., Alexander, L. V., Berg, P., Johnson, F.,
786 Kendon, E.J., Lenderink, G., Roberts, N.M., 2014. Future changes to the intensity and
787 frequency of short-duration extreme rainfall. *Rev. Geophys.* 52, 522–555.
788 doi:10.1002/2014RG000464

789 Wilby, R.L., Keenan, R., 2012. Adapting to flood risk under climate change.
790 *Prog. Phys. Geogr.* 36, 348–378. doi:10.1177/0309133312438908

791 Wilby, R.L., Wedgbrow, C.S., Fox, H.R., 2004. Seasonal predictability of the
792 summer hydrometeorology of the River Thames, UK. *J. Hydrol.* 295, 1–16.
793 doi:10.1016/j.jhydrol.2004.02.015

794 Wood, S.N., 2006. Generalized additive models: an introduction with R.
795 Chapman & Hall/CRC.

796 Wood, S.N., 2000. Modelling and smoothing parameter estimation with multiple
797 quadratic penalties. *J. R. Stat. Soc. Ser. B* 62, 413.

798 Wood, S.N., Augustin, N.H., 2002. GAMs with integrated model selection using
799 penalized regression splines and applications to environmental modelling. *Ecol. Modell.*
800 157, 157–177. doi:10.1016/S0304-3800(02)00193-X

801 Yee, T., Stephenson, A., 2007. Vector generalized linear and additive extreme
802 value models. *Extremes* 10, 1–19. doi:10.1007/s10687-007-0032-4

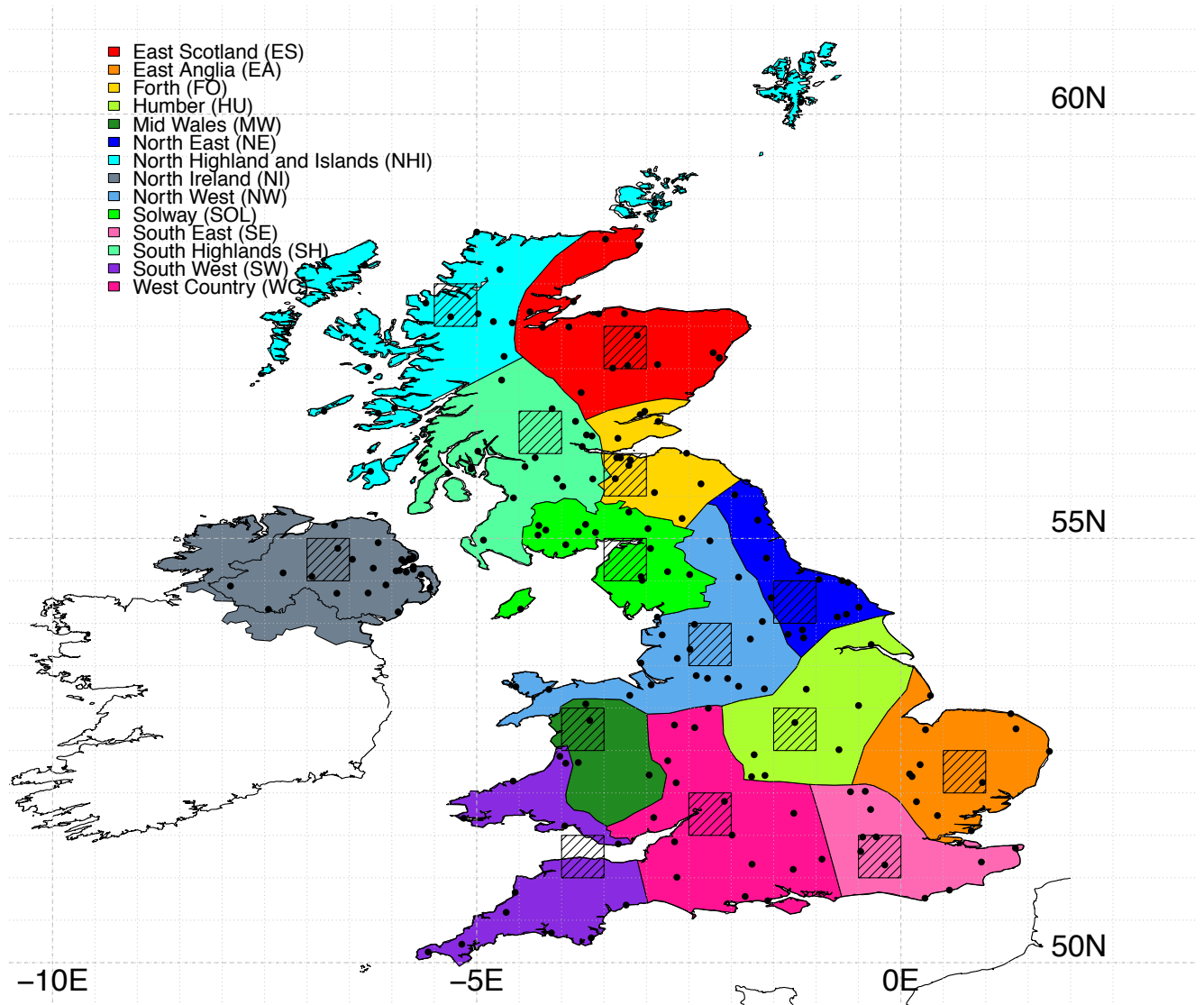
803 Zheng, F., Westra, S., Leonard, M., 2015. Opposing local precipitation extremes.
804 *Nat. Clim. Chang.* 5, 389–390. doi:10.1038/nclimate2579

805 Zolina, O., 2014. Multidecadal trends in the duration of wet spells and associated
806 intensity of precipitation as revealed by a very dense observational German network.
807 *Environ. Res. Lett.* 9, 025003. doi:10.1088/1748-9326/9/2/025003

808 Zolina, O., Simmer, C., Belyaev, K., Gulev, S.K., Koltermann, P., 2013. Changes
809 in the duration of European wet and dry spells during the last 60 years. *J. Clim.* 26, 2022–
810 2047. doi:10.1175/JCLI-D-11-00498.1

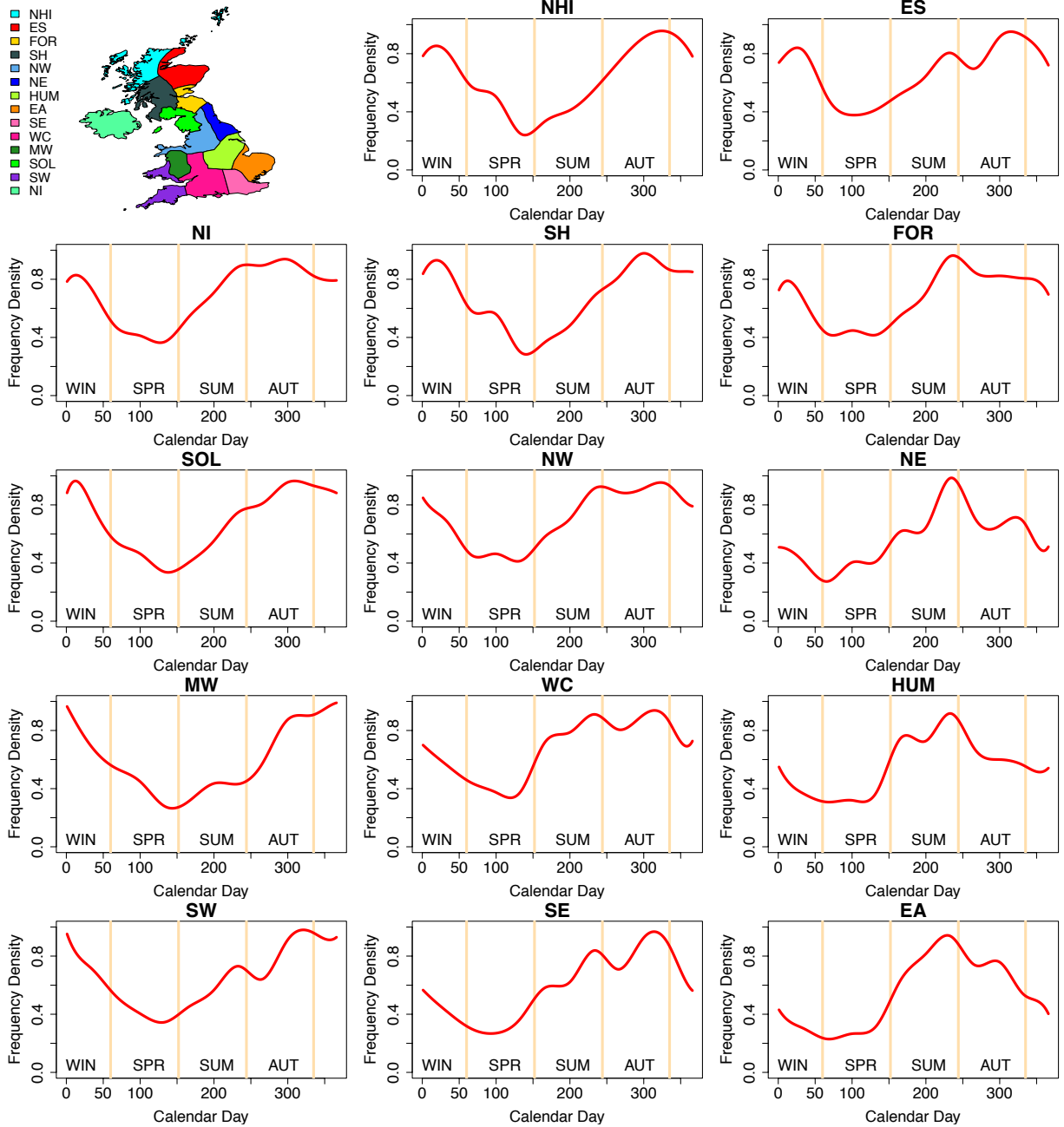
Table 1: Terms used in the Generalized Additive Models

Covariate	Term	Definition
Day of the year	d_t	where t is the event
Linear seasonality: sine	S_{kt}	$\sin(dt')$ $t' = 2\pi t/365.25$
Linear seasonality: cosine	C_{kt}	$\cos(dt')$
Event occurrence	y_t	$y \in [0,1]$ $y=1$ when the daily total exceeds a high threshold
Sea surface temperature	$ST1_t$	1 month lagged grid box average
Normalised mean sea level pressure	P_t	Monthly average over North Atlantic
Monthly min air temperature	Θ_{Nt}	Grid box average minimum



813

814 **Figure 1 : Location of gauging stations in relation to the 14 extreme rainfall regions (Jones et al.,**
 815 **2014). North Highlands and Islands (NHI), East Scotland (ES), Forth (FOR), South Highlands (SH),**
 816 **North West (NW), North East (NE), North Ireland (NI), Solway (SOL), Humber (HU), South West**
 817 **(SW), Mid Wales (MW), West Country (WC), Southern England (SE), East Anglia (EA). Dashed**
 818 **grid lines indicate the 5°x5° boxes corresponding to SST and MSLP measurements; hatched boxes**
 819 **indicate 0.5°x0.5° corresponding to minimum monthly air measurements.**



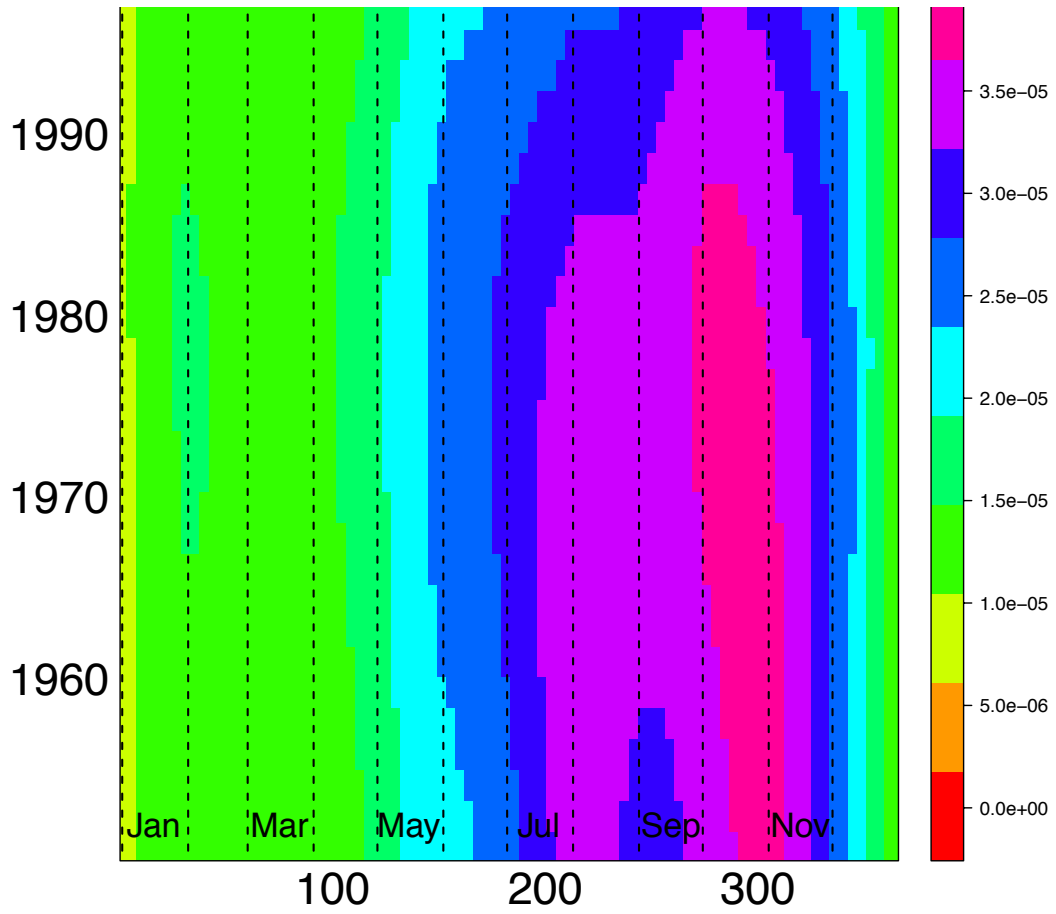
820

821 **Figure 2 : Frequency density distributions of 1-day extreme daily precipitation occurrences by month**

822

in each extreme rainfall region (Region names as Figure 1).

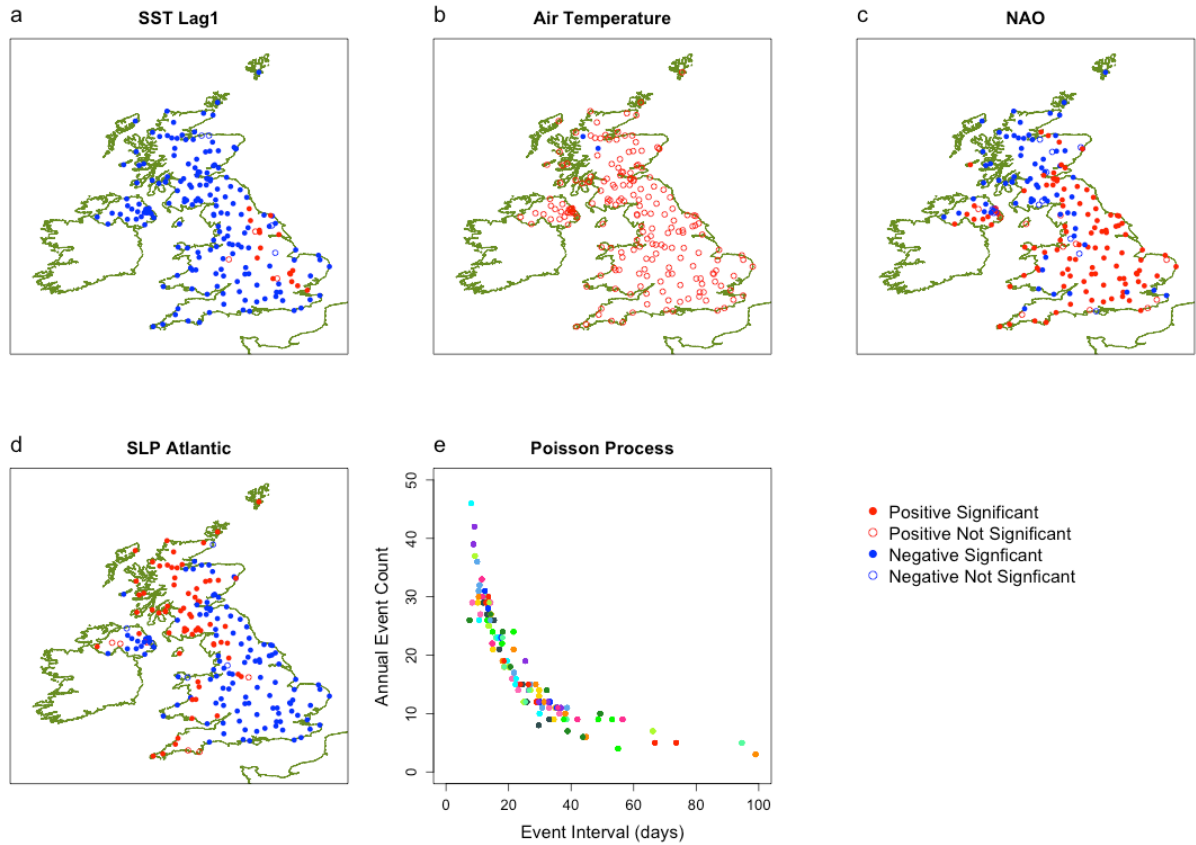
823



824

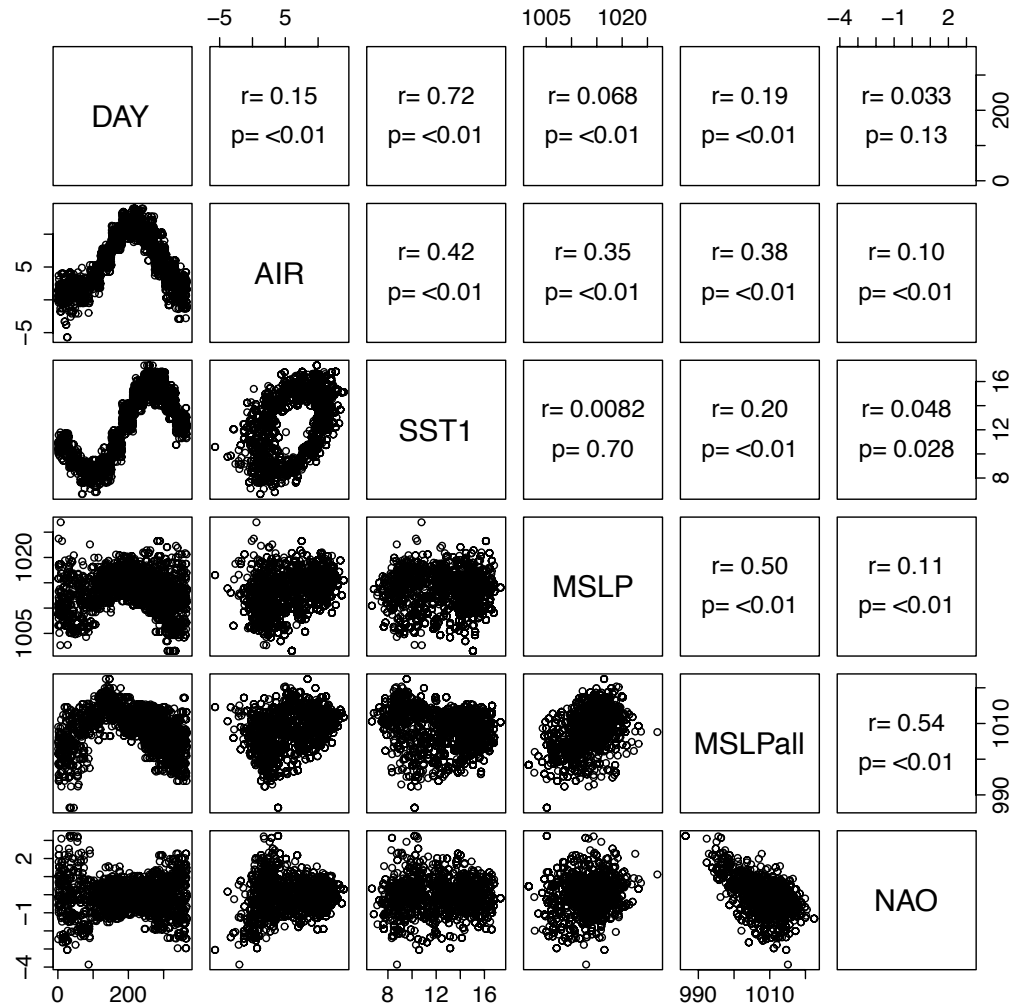
825 **Figure 3 : Smoothed relative frequency distribution of extreme occurrences per day of year each year**
 826 **for SE England. Data is regionally pooled extreme occurrences for 1950-1999. Shading indicates**
 827 **relative “density” of occurrences, pink/purple shades indicate more extreme occurrences in a 15**
 828 **day/15 year smoothing window. Black lines indicate first day of month.**

829



830

831 **Figure 4 : Significance (at 95%) of external covariates on Poisson regression model for annual counts**
 832 **of extreme occurrences from left to right top row: (a) Lagged Sea Surface Temperature for the 9**
 833 **adjacent grid boxes; (b) Monthly Minimum Air Temperature per gauge in each extreme region; (c)**
 834 **Monthly NAO index; bottom row: (d) Sea Level Pressure over Atlantic Region; (e) Exponential**
 835 **distribution of annual count of extreme occurrences with respect to time between events colours**
 836 **relate to the extreme regions. Significant correlations are filled circles (can be positive or negative)**

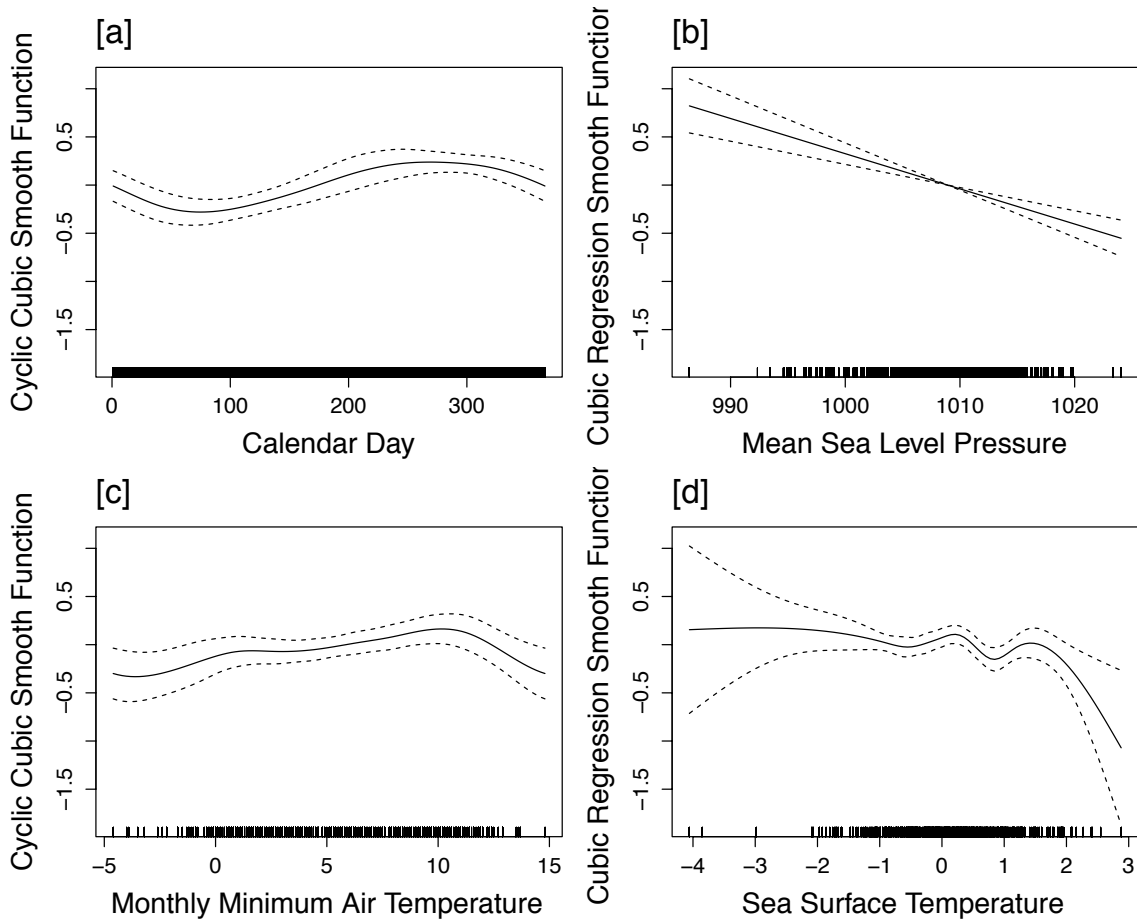


837

838 **Figure 5: Pair plots of correlations between the calendar day of extreme occurrences (DAY) in South**
 839 **East England region and: monthly minimum air temperature (AIR); monthly lagged Sea Surface**
 840 **Temperature (SST1); concurrent monthly Sea Level Pressure in adjacent grid box (MSLP);**
 841 **concurrent monthly Sea Level Pressure over an Atlantic domain (MSLPall); North Atlantic**
 842 **Oscillation Index (NAO)**

843

844



845

846

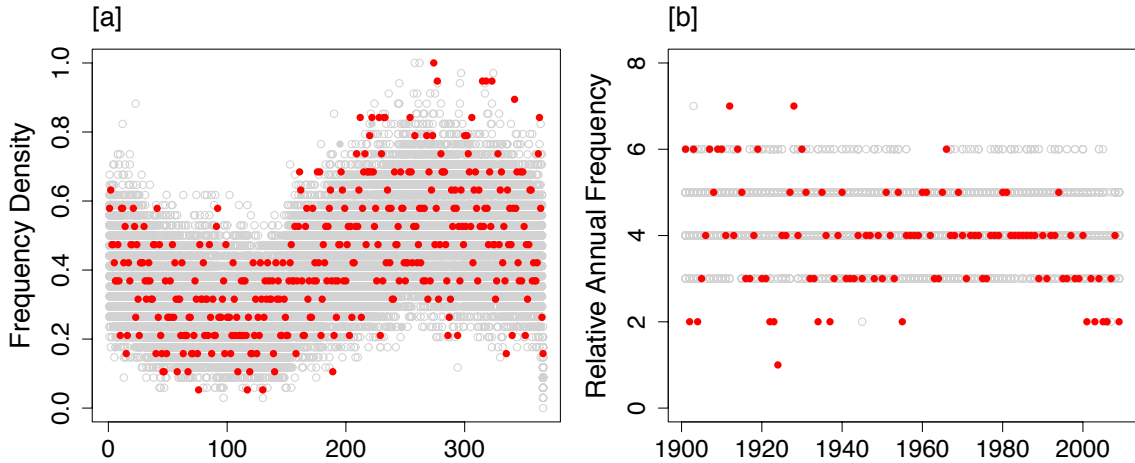
847

848

849

850

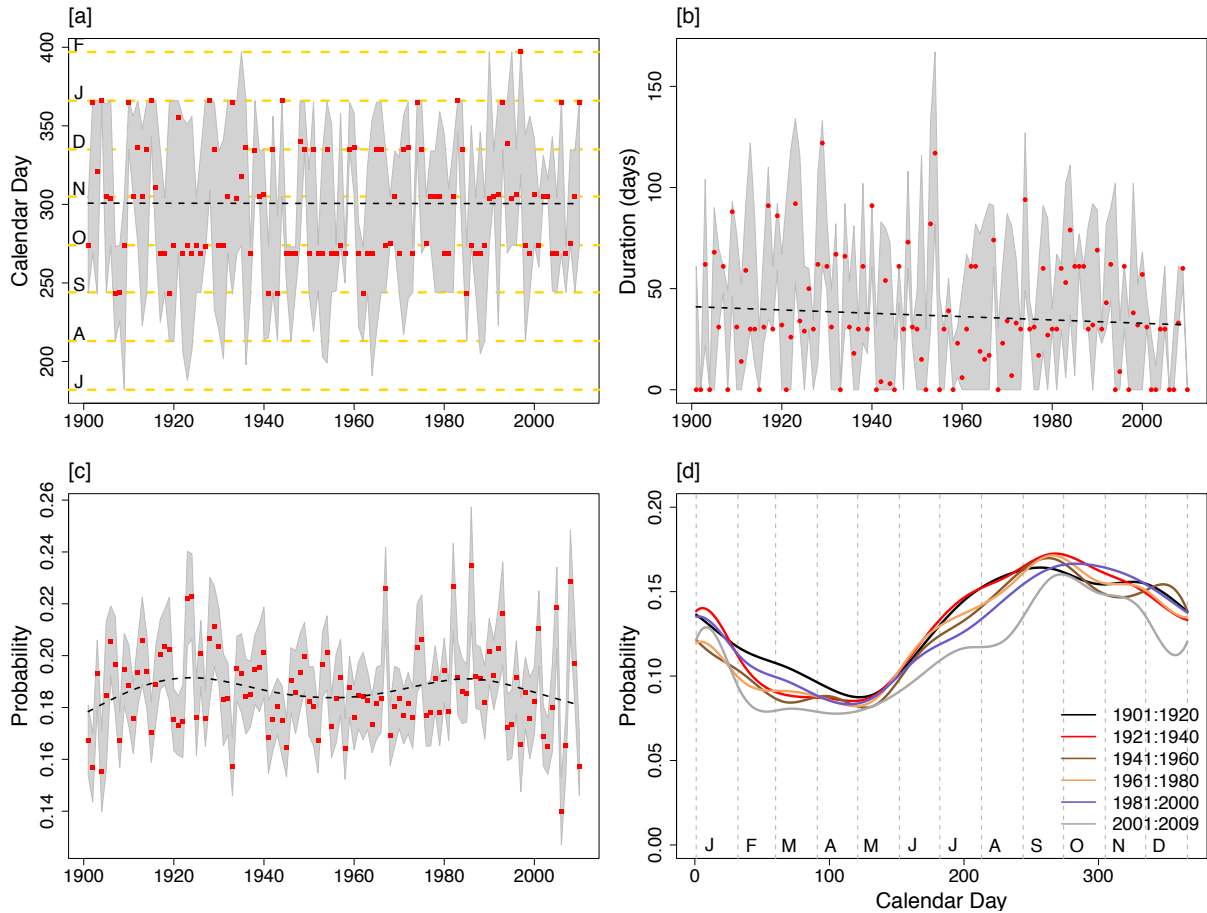
Figure 6: Smoothing parameters for South East England (SE) GAM showing: (a) seasonality as calendar day; (b) Monthly Sea Level Pressure over North Atlantic; (c) Monthly minimum daily air temperature in adjacent grid box; and (d) Lagged monthly sea surface temperature.



851

852 **Figure 7: Comparison of observed and simulated Extreme occurrences for the validation data (1901-**
 853 **2010) for SE region for (a) occurrences per day; and (b) annual relative frequency extreme**
 854 **occurrences from 100 samples per day from the Binomial distribution. Relative frequency is the**
 855 **number of extreme occurrences per year at one location in the region.**

856



857
858

859 **Figure 8: Changes in the probability of extreme occurrences for SE region showing: (a) mean day of**
 860 **occurrence of highest 5% probability; (b) number of days per year with the highest 5% probability;**
 861 **(c) highest probability of extreme occurrences per year; and (d) 20 year mean daily probability.**
 862 **Confidence intervals for a-c shown in grey, derived from distribution standard errors, smoothed 5-**
 863 **year running mean in black dash, first day of month in yellow dash (a).**