

1                    **A Quantile Generalised Additive Approach for**  
2                    **Compound Climate Extremes: Pan-Atlantic Extremes**  
3                    **as a Case Study**

4                    **Leonardo Olivetti<sup>1,2</sup>, Gabriele Messori<sup>1,2,3,4</sup>, Shaobo Jin<sup>5,6</sup>**

5                    <sup>1</sup>Department of Earth Sciences, Uppsala University

6                    <sup>2</sup>Centre of Natural Hazards and Disaster Science, Uppsala University

7                    <sup>3</sup>Department of Meteorology, Stockholm University

8                    <sup>4</sup>Bolin Centre for Climate Research, Stockholm University

9                    <sup>5</sup>Department of Statistics, Uppsala University

10                    <sup>6</sup>Department of Mathematics, Uppsala University

11                    **Key Points:**

- 12                    • Quantile general additive models (QGAMs) can model the relationship between  
13                    compound climate extremes flexibly and robustly.
- 14                    • North American cold spells are a useful predictor of subsequent wet or windy ex-  
15                    tremes in Western Europe.
- 16                    • QGAMs can predict those extremes more accurately than conventional peak-over-  
17                    threshold models.

---

Corresponding author: Leonardo Olivetti, [leonardo.olivetti@geo.uu.se](mailto:leonardo.olivetti@geo.uu.se)

**Abstract**

We present an application of quantile generalised additive models (QGAMs) to study spatially compounding climate extremes, namely extremes that occur (near-) simultaneously in geographically remote regions. We take as an example wintertime cold spells in North America and co-occurring wet or windy extremes in Western Europe, which we collectively term Pan-Atlantic compound extremes. QGAMS are largely novel in climate science applications and present a number of key advantages over conventional statistical models of weather extremes. Specifically, they remove the need for a direct identification and parametrisation of the extremes themselves, since they model all quantiles of the distributions of interest. They thus make use of all information available, and not only of a small number of extreme values. Moreover, they do not require any a priori knowledge of the functional relationship between the predictors and the dependent variable. Here, we use QGAMs to both characterise the co-occurrence statistics and investigate the role of possible dynamical drivers of the Pan-Atlantic compound extremes. We find that cold spells in North America are a useful predictor of subsequent wet or windy extremes in Western Europe, and that QGAMs can predict those extremes more accurately than conventional peak-over-threshold models.

**Plain Language Summary**

In this paper we propose a new data-driven method to study climate extremes occurring simultaneously in multiple, possibly remote, locations. Such extremes can pose a greater threat to human societies than single, isolated extremes, as their effects may exacerbate each other and lead to correlated losses. The method we suggest requires fewer assumptions than conventional extreme value statistical techniques, and can help us to identify previously unknown relationships between the extremes themselves and their possible drivers. We exemplify its use by studying the co-occurrence of periods of unusually cold weather in North America and uncommonly strong wind and abundant precipitation in Western Europe. We find that the new method has better predictive power for the European extremes than conventional statistical approaches. Furthermore, we confirm the results of previous studies suggesting an association between the wintertime extremes in North America and Western Europe.

**1 Introduction**

The statistical properties of climate extremes have been extensively studied using parametric approaches to extreme value theory (EVT; e.g. Fisher & Tippett, 1928; Gumbel, 1941; Davison & Smith, 1990; Coles, 2001; Mares et al., 2009; Elvidge & Angling, 2018; French et al., 2019), and since its inception, parametric EVT regression has been successfully used to investigate extreme event drivers (Gumbel, 1958; Pickands, 1975; Coles, 2001; Mares et al., 2009; do Nascimento et al., 2021). Parametric EVT aims to identify and characterise extreme observations by identifying their underlying distribution which, in turn, provides valuable information on the expected frequency and intensity of the extremes. Two fundamental approaches lie at the core of classic parametric EVT: the block maxima approach (BM) (Fisher & Tippett, 1928; Gumbel, 1958) and the peak over threshold (POT) approach (Pickands, 1975; Smith, 1984; Davison & Smith, 1990).

The BM approach defines  $k$  time periods (blocks) of equal length and extracts the  $n$  largest independent observations from each block (block maxima). According to the Fisher–Tippett–Gnedenko theorem (Fisher & Tippett, 1928), properly normalised block maxima converge to a distribution belonging to the generalised extreme value family of distributions (GEV). A challenge of the BM approach lies in finding a suitable size for the blocks (Ferreira & Haan, 2015). Small blocks may lead to the identification of some spurious extremes, whereas large blocks may ignore some extremes and slow down the

68 convergence to the GEV family of distributions. Furthermore, given that the Fisher–Tippett–Gnedenko  
 69 theorem is an asymptotic theorem, there is no guarantee that the appropriately normalised  
 70 block maxima belong to the GEV family of distributions when the sample size is limited.  
 71

72 The POT approach addresses these challenges by selecting a high threshold  $u$ , and  
 73 defining as extremes observations above that threshold (the maxima,  $\mu$ ). The excesses  
 74 follow approximately a generalised Pareto distribution (GPD), in accordance with Pickands’  
 75 theorem (Pickands, 1975). The challenge in this case is to define an appropriate thresh-  
 76 old: sufficiently high to isolate extremes yet sufficiently low to ensure an appropriate sam-  
 77 ple size. Technical details of the BM and POT approaches are provided in the Support-  
 78 ing Information text S1.

79 While being a widely-used and versatile tool, implementing parametric EVT for  
 80 studying climate extremes presents some challenges. The first is finding a suitable em-  
 81 pirical definition for the extremes, as discussed above and also highlighted by Passow and  
 82 Donner (2019). Second, using parametric EVT in a regression context to study extreme  
 83 event drivers requires some a priori knowledge or assumption of the functional relation-  
 84 ship between the drivers and the extremes. Defining this relationship a priori is often  
 85 challenging and may result in simplistic or unrealistic assumptions about the relation-  
 86 ship between the two. Third, parametric EVT does not make optimal use of all the sta-  
 87 tistical information available, as it ignores all that is not an extreme (see also the dis-  
 88 cussion in Passow & Donner, 2019).

89 The literature offers extensions to the classic parametric EVT theory (Lucarini et  
 90 al., 2016) as well as a number of non-parametric alternatives (Koenker & Hallock, 2001;  
 91 Yee, 2015; Fasiolo et al., 2021). The main appeal of non-parametric techniques, such as  
 92 quantile-based methods, is that they do not require an empirical definition of the extremes  
 93 in the same way as parametric EVT does. Furthermore, they make use of all statisti-  
 94 cal information available, rather than just the extreme data. This is particularly rele-  
 95 vant when the interest lies in extreme events with a relatively short return period (e.g.  
 96 windstorms recurring every 6 months or 1 year at a given location), for which the con-  
 97 vergence of the GPD parameters may first be observed at quantiles larger than the one  
 98 of interest.

99 Here, we apply quantile generalised additive models (QGAMs, Fasiolo et al., 2021;  
 100 Koenker, 2011) to the study of spatially compounding climate extremes and their drivers,  
 101 namely extremes that occur (near-)simultaneously in geographically remote regions. The  
 102 synchronised – or compound – occurrence of remote extremes is often associated with  
 103 greater impacts than those of the corresponding individual extremes, for example expos-  
 104 ing actors with international coverage to correlated losses (Mills, 2005) and imperilling  
 105 global food security (Kornhuber et al., 2020). QGAMs are a non-parametric approach  
 106 which is largely novel in the context of climate science. They are a recent extension to  
 107 generalised additive models, bearing the promise of addressing the aforementioned three  
 108 key limitations of parametric EVT. We benchmark QGAMs against other parametric  
 109 and non-parametric approaches.

110 To illustrate the methodology, we consider the repeated occurrence of wintertime  
 111 cold spells in Eastern North America and wet or windy extremes in Western Europe, which  
 112 we collectively term Pan-Atlantic compound extremes. The repeated occurrence of these  
 113 extremes in recent winters (e.g. Coumou & Rahmstorf, 2012; Lee et al., 2015; Trenary  
 114 et al., 2015; Dodet et al., 2019; Wild et al., 2015) has led to hypothesise a connection  
 115 between the two sets of extremes (Messori et al., 2016; De Luca et al., 2020; Leeding et  
 116 al., 2023; Messori & Faranda, 2023). Nonetheless, a systematic statistical characterisa-  
 117 tion of such connection has largely been limited to simple co-occurrence statistics. The  
 118 aim here is to provide a proof-of-concept for the use of QGAMs in the study of compound  
 119 climate extremes. We therefore seek to evaluate the performance of QGAMs relative to  
 120 alternative statistical models by applying them to a previously studied set of spatially  
 121 compounding climate extremes, as opposed to investigating novel extreme occurrences  
 122 and related large-scale atmospheric drivers. We further highlight that we do not aim to

123 use the regression models we present as self-standing forecasting tools. In other words,  
 124 we do not aim to forecast extreme event occurrences using only information from sev-  
 125 eral days before the extremes – as one would do in a conventional forecasting exercise.  
 126 Rather, we see these models as useful tools to robustly quantify the statistical connec-  
 127 tions between geographically remote extremes and investigate the roles of different po-  
 128 tential dynamical drivers.

129 The remainder of this paper is structured as follows: Section 2 provides a short in-  
 130 troduction to parametric EVT regression, and introduces quantile-based non-parametric  
 131 approaches as an alternative, including QGAMs. Section 3 defines the scope of this pa-  
 132 per, and discusses practical concerns related to variable selection and model formulation.  
 133 Section 4 compares the methods presented in previous sections, by studying the Pan-  
 134 Atlantic compound extremes and their possible dynamical drivers, considering specif-  
 135 ically the North Atlantic jet stream and the North Atlantic Oscillation (NAO). Section  
 136 5 concludes the paper by providing a short summary of the findings and discussing the  
 137 strengths and limitations of QGAM applications to study climate extremes.

## 138 2 Extreme Value Statistical Models

139 This section presents some parametric and non-parametric approaches to EVT re-  
 140 gression and discusses how their performance can be compared in practice.

### 141 2.1 Parametric EVT regression

142 The relationship between the extremes and their likely precursors can be described  
 143 parametrically through a generalised linear model, where the expected value of a BM  
 144 or POT extreme at a time  $t$ ,  $E(M_t)$  is a function of previous values of itself,  $M_{t-k}$  and  
 145 other factors likely to affect its strength,  $X_{t-k}$ . Then:

$$146 \quad E(M_t | M_{t-k}, X_{t-k}) = M_{t-k}\phi + X_{t-k}\beta. \quad (1)$$

147 Since the extremes are selected through the BM or the POT approach, paramet-  
 148 ric EVT regression largely shares the same strengths and limitations of these approaches.

### 149 2.2 Non-parametric EVT regression

150 Non-parametric models are a broad class of methods which do not rely on a pre-  
 151 determined functional relationship between the outcome and the predictor, but estimate  
 152 it empirically (Härdle, 1990). Quantile-based models are a subset of non-parametric mod-  
 153 els which are particularly suitable to the analysis of extremes, as they can be used to es-  
 154 timate values in the tails of the distribution. A quantile  $Q(\tau)$  is defined as the inverse  
 155 of the cumulative distribution function, uniquely identifying the value of the cumulative  
 156 distribution function corresponding to probability  $\tau$ .

#### 157 2.2.1 Quantile Regression

158 The linear relationship between a conditional quantile of an output and a predic-  
 159 tor can be estimated non-parametrically through quantile regression (Koenker & Hal-  
 160 lock, 2001; Koenker & Bassett, 1978). Quantile regression aims to estimate the condi-  
 161 tional quantile of the dependent variable  $Q_{Y|X}(\tau)$  as a function of the regressors  $X$ , so  
 162 that:

$$163 \quad Q_{Y|X}(\tau) = X\beta_\tau \quad (2)$$

164 It solves the following minimisation problem:

165 
$$\min E\{\rho_\tau(Y - X^*\beta)\}, \tag{3}$$

166 where  $\rho_\tau = (\tau - 1)zI(z < 0) + \tau zI(z \geq 0)$  is the so-called "pinball loss", punish-  
 167 ing predictions which are further away from the quantile of interest. Here,  $z$  is the resid-  
 168 ual and  $\tau$  the quantile of interest.

169 By estimating the effect of the regressors on an ensemble of extreme quantiles, it  
 170 is possible to draw some conclusions on how the regressors affect the intensity of the ex-  
 171 tremes. This is done without relying on the large sample asymptotics employed by para-  
 172 metric EVT, thus without the need of reducing sample size.

173 The main limitation of quantile regression is that it is a linear model, ignoring all  
 174 possible non-linear effects of the regressors on the quantile of interest.

175 **2.2.2 Quantile Generalised Additive Models**

176 Generalised additive models (GAMs, Wood, 2017; Hastie & Tibshirani, 1986) are  
 177 a broad family of non-parametric models describing the dependent variable as an addi-  
 178 tive function of unknown smooths of the regressors. A smooth function is a function which  
 179 is derivable up to certain order at each point throughout its domain. GAMs can be de-  
 180 scribed as follows:

181 
$$g(E(Y)) = \beta_0 + f_1(X_1) + f_2(X_2) + \dots + f_i(X_i), \tag{4}$$

182 where  $E(Y)$  is the expected value of the outcome,  $g()$  is the link function,  $\beta_0$  is an  
 183 intercept, and  $f_1(X_1) + f_2(X_2) + \dots + f_i(X_i)$  are smooth functions of the predictors.

184 GAMs do not require determining the functional relationship between the outcome  
 185 and the predictors a priori. This is instead determined empirically, through a data-driven  
 186 process testing a large number of possible combinations. First, a set of bases is chosen  
 187 for the predictors, so that the original covariates  $X_1, X_2, \dots, X_i$  are embedded into a larger  
 188 feature space  $X^*$  including higher order terms of the original covariates. Then, the best  
 189 model is chosen out of the expanded feature space by minimising a loss function of choice,  
 190  $L(X^*)$ , while penalising for excessive complexity. In the absence of a link function, a nat-  
 191 ural choice of loss function is the quadratic loss, so the best model is chosen according  
 192 to the following minimisation:

193 
$$\min E\{(Y - X^*\beta)^2 + \lambda J\}, \tag{5}$$

194 where  $(Y - X^*\beta)^2$  is the sum of squared residuals,  $\lambda$  is a smoothing parameter and  
 195  $J$  is a penalty term. A common penalty term is  $J = \int f''(x)^2 dx$ , with other forms of penalty  
 196 also being possible (James et al., 2022).  $J$  is larger when the function becomes wigglier,  
 197 punishing excessive functional complexity. Choosing larger values of the smoothing pa-  
 198 rameter  $\lambda$  pushes the model towards a simpler functional form, so that  $\lambda \rightarrow \infty$  makes  
 199 GAM equivalent to linear regression. The value of  $\lambda$  is usually determined empirically,  
 200 through generalised cross-validation or restricted maximum likelihood. For technical de-  
 201 tails, the reader is referred to Wood (2017).

202 Classic GAMs as expressed in equation 4 model the expected value of the outcome  
 203 and not of the maxima, making them unsuitable for extreme value analysis. One approach  
 204 could be to model the expected value of a series of independent and identically distributed  
 205 maxima  $(M_1, M_2, \dots, M_N)$ , by selecting the maxima through a BM or a POT approach;  
 206 however, this reintroduces into the model the limitations related to those approaches.  
 207 An alternative approach is to model a conditional quantile of the outcome as a function  
 208 of the predictors, as in quantile regression: this approach goes under the name of quan-  
 209 tile generalised additive model (QGAM).

210 QGAMs are a recent extension of GAMs and quantile regression, which model a  
 211 chosen conditional quantile of interest as a sum of unknown smooth functions of the re-  
 212 gressors (Fasiolo et al., 2021). This builds on earlier theoretical work from Koenker (2011).  
 213 QGAMs can be expressed as follows:

$$214 \quad Q_{Y|X}(\tau) = f_1(X_1) + f_2(X_2) + \dots + f_i(X_i), \quad (6)$$

215 where  $Q_{Y|X}(\tau)$  is a conditional quantile of choice of the dependent variable. They  
 216 aim to minimise a loss function similar to the one described by equation 3,

$$217 \quad \min E\{\rho_\tau^*(Y - X^*\beta)\}, \quad (7)$$

218 where  $\rho_\tau^*$  is defined as:

$$219 \quad \rho_\tau^* = (\tau - 1) \frac{z}{\sigma} I(z < 0) + \lambda \log(1 + e^{\frac{z}{\lambda\sigma}}) \quad (8)$$

220  $\rho_\tau^*$  is the extended log-f loss, which, similarly to the pinball loss, punishes predic-  
 221 tions which are further away from the quantile of interest.  $\sigma > 0$  is a scale parameter  
 222 and  $\lambda > 0$  is a penalty term, meant to prevent excessive functional complexity. As  $\lambda$   
 223 approaches 0,  $\rho_\tau^*$  becomes equivalent to  $\rho_\tau$ , the pinball loss used in quantile regression  
 224 (Fasiolo et al., 2021).

225 Similarly to quantile regression, QGAMs do not make any assumptions on the dis-  
 226 tribution of the extremes, and only require defining a set of quantiles of interest to study  
 227 the effect of the regressors on the output. Furthermore, similarly to GAMs, they model  
 228 the relationship between the output and the regressors empirically, without requiring any  
 229 previous knowledge of the functional relationship between the two.

230 A possible limitation of GAMs and QGAMs is that they are additive in nature, and  
 231 even though interactions between terms may be modelled, every interaction has a large  
 232 effect on the computational burden of the model, thus limiting de facto the number of  
 233 regressors which may be added to the model.

### 234 **2.3 Benchmarking EVT models**

235 In order to compare the performance of the EVT models presented so far, we ap-  
 236 ply them to full-complexity climate data. We specifically consider the association between  
 237 cold spells in North America and wet or windy extremes in Western Europe. Following  
 238 previous literature looking at these extremes (Messori et al., 2016; De Luca et al., 2020;  
 239 Leeding et al., 2023; Messori & Faranda, 2023), we consider surface extremes occurring  
 240 one to a few times per year (e.g. 1% - 5% extreme quantiles), rather than extreme events  
 241 with multiannual return times. A useful statistical model should be able to verify whether  
 242 any relation between surface temperature in North America and surface extremes in West-  
 243 ern Europe is present, and, if this is the case, make use of this and additional informa-  
 244 tion on the state of the North Atlantic atmospheric circulation to improve the predic-  
 245 tion of said extremes. In particular, we identify three key characteristics of a useful EVT  
 246 model in this context:

247 1. It should provide a spatially resolved prediction of extremes in Western Europe,  
 248 which is more accurate than competing models, given that similar information is pro-  
 249 vided.

250 2. It should provide a consistent estimate of the spatially resolved return levels of  
 251 the extremes in Western Europe, where consistency is defined as the property of an  
 252 estimator whose probability of being arbitrarily close to the true value tends to one for  
 253 increasing sample size.

254 3. It should improve its performance whenever relevant information is added to the  
255 model.

256 The following steps are implemented to assess the performance of the models:

257 - For objective 1, we compute and compare the pseudo  $R^2$  quantile regression good-  
258 ness of fit introduced by Koenker and Machado (1999). This is defined as:

$$259 \quad R_{pseudo}^2 = 1 - \frac{L_{complete}}{L_{baseline}}, \quad (9)$$

260 where  $L_{complete}$  is the pinball loss of the model of interest, and  $L_{baseline}$  is the pinball  
261 loss of a model using the unconditional quantile  $Q_Y(\tau)$  of the anomalies at a given grid  
262 point as a fixed prediction for all observations at that same grid point.

263 - For objective 2, we define  $\hat{PO}$  as the percentage of overpredictions in the test sam-  
264 ple. An unbiased EVT model should overpredict the value of the output a percentage  
265 of times corresponding to the probability  $\tau$  of the target conditional quantile  $Q_{Y|X}(\tau)$ ,  
266 so that  $\hat{PO} = \tau$ . An estimate of the bias of the model is then given by the absolute dif-  
267 ference between the percentage of overpredictions  $\hat{PO}$  and  $\tau$ , so that

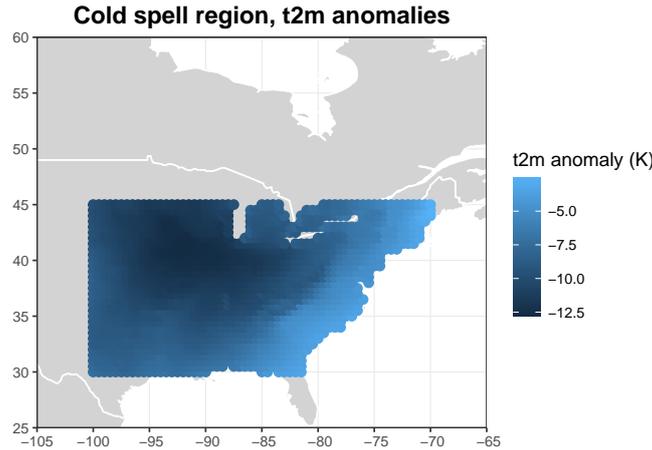
$$268 \quad \widehat{Bias} = |\hat{PO} - \tau| \quad (10)$$

269 - For objective 3, we first use the models to make spatially resolved predictions of  
270 extreme wind or precipitation events in Western Europe based on the latitude and the  
271 longitude of each location. Then, we add some information on the upstream large-scale  
272 atmospheric circulation and compare the results. A model fulfilling objective 3 is expected  
273 to progressively improve its performance, given that more information is provided and  
274 that this information has been identified as relevant to the extremes in previous stud-  
275 ies (Messori et al., 2016; Leeding et al., 2023).

276 Holdout cross-validation is performed by splitting the dataset into three parts: a  
277 training set, a validation set and a test set, containing approximately 50%, 25% and 25%  
278 of the available observations, respectively. The random split uses seasonal data blocks  
279 (we focus on the winter season, see Sect. 3) to minimise information leakage between the  
280 different sets. The models are trained using the training set, with feature selection aided  
281 by the use of the validation set in addition to previous research. All metrics of model  
282 performance are computed based on the test set.

### 283 3 Defining Pan-Atlantic Compound Extremes and Prediction Mod- 284 els

285 Atmospheric data are taken from the ERA 5 global reanalysis (Hersbach et al., 2020)  
286 with a daily time resolution and a  $0.5^\circ$  horizontal resolution. We consider the period Novem-  
287 ber 1959–January 2022, and the months of November, December, January and Febru-  
288 ary (NDJF). The daily NAO values, are taken from the NOAA online archive (NOAA/  
289 National Weather Service, 2022). Cold spells in North America are defined as days with  
290 2m temperature (t2m) anomalies relative to the daily climatology, area-averaged over  
291  $30^\circ - 45^\circ$  N,  $100^\circ - 70^\circ$  W, below the 5th quantile of the distribution. The daily clima-  
292 tology is obtained by applying a 7 days running mean and then averaging over each cal-  
293 endar day over the full time period. The domain of the North American cold spells fol-  
294 lows Messori et al. (2016) and is illustrated in Figure 1, together with the t2m anoma-  
295 lies associated with the selected cold spells. To decluster the extremes, we require that  
296 at least five days elapse between separate cold spells. Whenever several days within a  
297 five-day period meet the criteria for being classified as a cold spell, only the first day is  
298 selected. The choice of selecting the first day rather than the coldest aims to prevent the  
299 skill of the statistical models from being affected by the duration of the cold spells.



**Figure 1.** Domain of North American cold spells and mean t2m anomaly during cold spells. Darker shades are associated with larger negative t2m anomalies.

300 For Western Europe, we focus on daily mean 10m wind speed and daily precipi-  
 301 tation over Iberia and Western France: regions that Messori et al. (2016) and Messori  
 302 and Faranda (2023) highlighted as experiencing large anomalies following the North Amer-  
 303 ican cold spells. We test the statistical model predictions on the 95th and 99th quan-  
 304 tiles of the local distributions of these two variables. These correspond to return peri-  
 305 ods of approximately six times and once per winter season, respectively.

306 Based on previous work, we consider the characteristics of the Polar – or eddy-driven  
 307 – jet stream over the North Atlantic region as a possible predictor of the European ex-  
 308 tremes (Messori et al., 2016; Leeding et al., 2023). Jet stream speed is defined as the largest  
 309 zonally averaged zonal wind anomaly from a seven-day smoothed climatology at 250 hPa  
 310 over a North Atlantic domain spanning 30° - 75° N, 70° - 5° W. The location of the jet  
 311 is given by the latitude displaying the largest zonal mean zonal wind anomaly as defined  
 312 above. The choice of time lag of temperature in North America and zonal wind over the  
 313 North Atlantic to be included in the models has been aided by cross-validation. The lag  
 314 is relative to the prediction date for European extremes. Only one lag for each variable  
 315 is included in the models in order to avoid multicollinearity issues, as both t2m in North  
 316 America and the above jet indices display a high degree of autocorrelation.

317 We provide the statistical models with three levels of information for predicting t2m  
 318 and daily mean 10m wind speed in Western Europe:

319 - Basic models, making predictions as a function of time, latitude, longitude, and  
 320 a training-set based seasonal climatology, only.

321 - Cold spell models, where t2m in North America from two days prior to the pre-  
 322 diction date is added to the regressors.

323 - Cold spell and jet stream models, where the NAO and strength and location of  
 324 the Polar jet from one day prior are also added to the regressors. The NAO values are  
 325 included here as a way to control for possible confounders affecting North American sur-  
 326 face temperatures and surface weather in Western Europe, which might otherwise lead  
 327 to biased estimates of the association between the two.

328 For each information level, we build three models: a QGAM, a linear quantile re-  
 329 gression (QREG) and a POT model. Therefore, a total of nine models is estimated for  
 330 each target variable in Western Europe. The models' performance on the test data is com-  
 331 pared as outlined in Subsection 2.3. Technical details including the exact formulation  
 332 of each model can be found in Supporting Information text S2.

333  
334

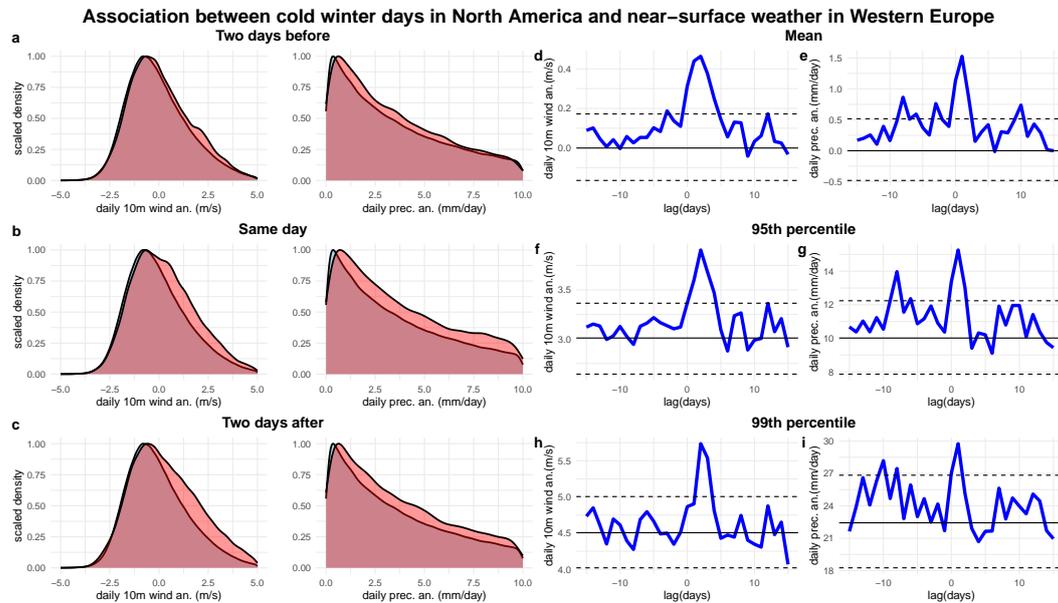
#### 4 Non-parametric EVT in practice: a QGAM approach to the study of Pan-Atlantic Compound Extremes

335  
336  
337  
338  
339  
340

This section applies the models presented in Sections 2 and 3 to study Pan-Atlantic compound extremes and their drivers. It aims to support previous analyses of the connection between cold spells in North America and wet or windy extremes in Europe (Messori et al., 2016; De Luca et al., 2020; Leeding et al., 2023; Messori & Faranda, 2023) through robust statistical estimation, and verify whether QGAMs outperform alternative statistical models.

341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351

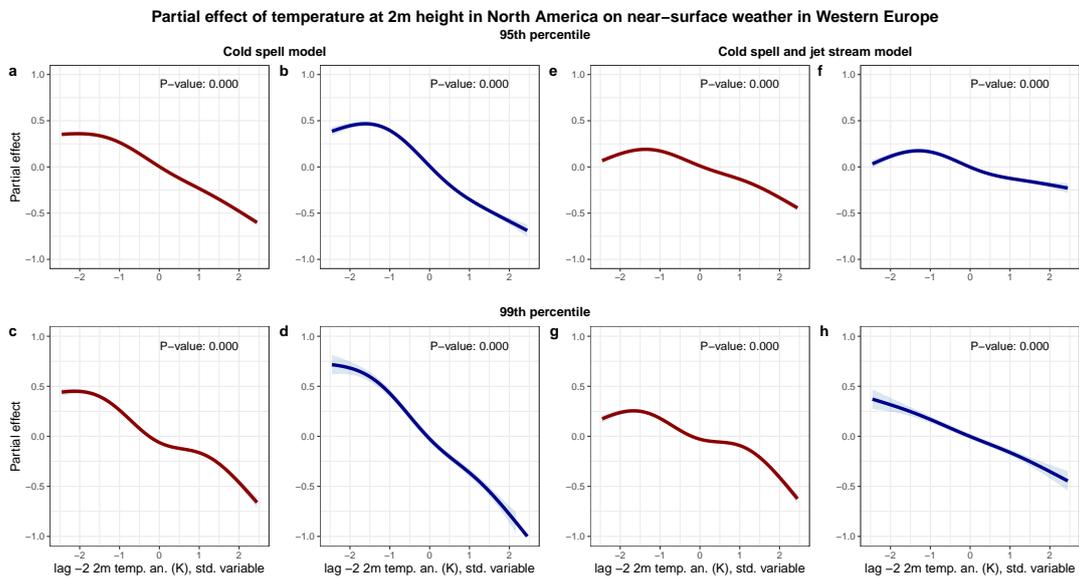
A first overview of the relation between cold spells in North America and wet or windy extremes in Western Europe can be obtained by means of a composite analysis (Fig. 2). There is a heightened frequency of positive anomalies in 10m wind speed and precipitation in Western Europe in conjunction with cold spells in North America (Fig. 2a, b, c), with the largest effect in Western Europe being observed immediately after the cold spell in North America. This is particularly true for middle-sized to large positive anomalies, pointing to the possibility of near-simultaneous extremes in the two regions. Fig. 2d, e, f corroborate this hypothesis, by showing how the mean, the 95th and the 99th quantiles of daily mean 10m wind speed and daily precipitation in Western Europe are significantly higher than usual in the aftermath of North American cold spells (Fig. 2d-f).



**Figure 2.** a-c: Kernel density estimate of daily mean 10m wind speed (left) and daily precipitation anomalies (right) in Western Europe: climatological distribution (blue) and distribution of events two days before (a), the same day (b), and two days after (c) a cold spell in North America (red). Mean (d-e), 95th quantile (f-g) and 99th quantile of (h-i) of daily mean 10m wind speed (left) and daily precipitation anomalies (right) in Western Europe 15 days before and after a cold spell in North America. The solid black lines are the overall mean/quantile, while the dashed lines mark approximate 95% significance levels, computed by means of a Monte Carlo permutation test with 20000 replications.

352 We next test whether QGAMs and the other models introduced in Section 2 can  
 353 leverage this association to make statistical forecasts of extreme events in Europe (95th  
 354 and 99th quantiles of daily mean 10m wind speed and daily precipitation at locations  
 355 in Iberia and Western France). We present the results for the three QGAM models de-  
 356 scribed in Section 3 (basic, cold spell and cold spell and jet stream), and compare their  
 357 performance to the quantile regression and POT models.

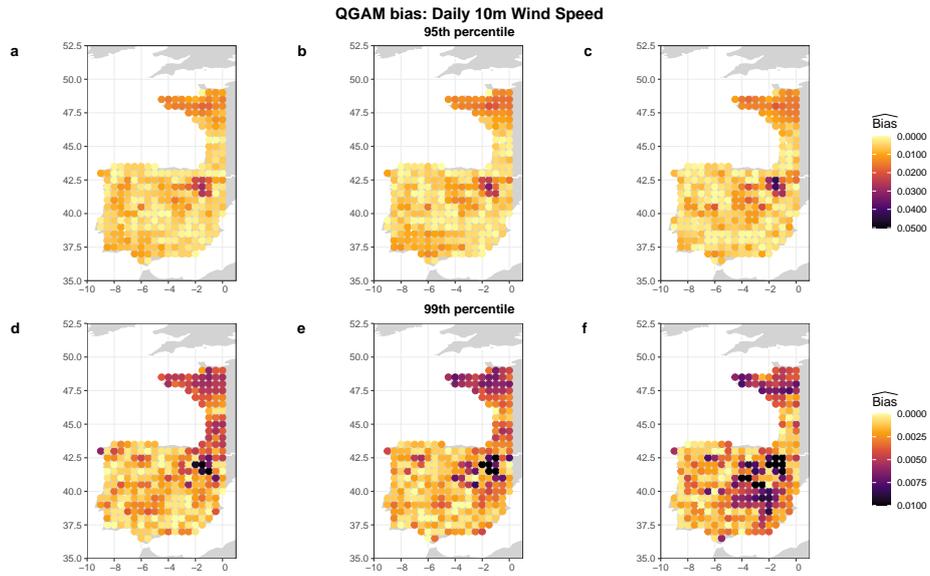
358 QGAMs can identify the relation between t2m in North America and weather ex-  
 359 tremes in Western Europe. Figure 3 shows the partial effect of t2m in North America  
 360 on near-surface weather in Western Europe, when holding all other variables included  
 361 in the model constant. Lower temperatures in North America are significantly associ-  
 362 ated with higher values of the extreme quantiles and therefore more extreme weather events  
 363 in Western Europe. This effect is at its strongest for temperature anomalies of two stan-  
 364 dard deviations below the mean, i.e. cold spells. In the model also including jet stream  
 365 and NAO information (Fig. 3e-h), the effect of North American temperatures is weaker  
 366 due to the fact that part of the effect is likely mediated by the jet and/or NAO.



**Figure 3.** Partial effect of t2m anomalies in North America at lag -2 days on the 95th and 99th quantiles of daily mean 10m wind speed (red) and daily precipitation (blue) in Western Europe generated through QGAMs (standardised variables). The effect is measured in terms of standard deviations of the dependent variable, when holding all other variables in the model fixed. **a-d:** Cold spell model, holds latitude, longitude and time-variable fixed. **e-h:** Cold spell and jet stream model, holds all of the above, NAO, and strength and location of the jet stream fixed. The p-values indicate the overall significance of the smooth term in the respective model.

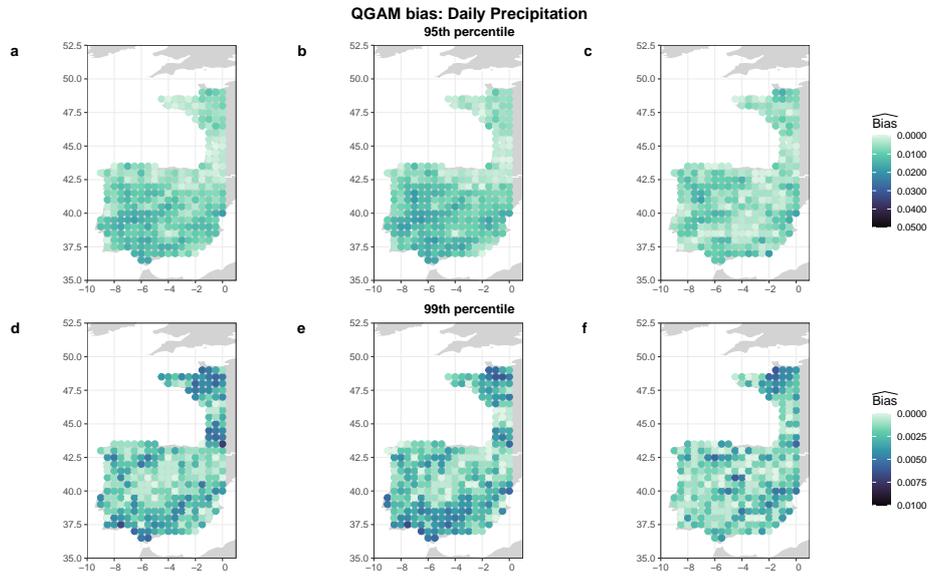
367 Figure 4 shows the bias associated with the prediction of extreme quantiles of daily  
 368 mean 10m wind speed in Western Europe through QGAMs. Ideally, the bias of our mod-  
 369 els should be held under  $1-\tau$  for most grid points, since a model with bias greater than  
 370  $1-\tau$  has a larger bias than a model providing a systematic underprediction. At the same  
 371 time, zero bias is not an aim itself due to the variance-bias trade-off, and thus some bias  
 372 is to be expected. Figure 4 suggests that QGAMs mostly have a bias lower than  $1-\tau$   
 373 when predicting daily mean 10m wind speed for most grid points. The models predicting the  
 374 95th quantile may appear to perform better than those predicting the 99th quantile, but

375 this is mostly a question of scale, since the figure colourscale is always capped at  $1-\tau$ ,  
 376 which is lower for the 99th than for the 95th quantile.



**Figure 4.** Estimated bias of QGAMs in terms of absolute distance between the percentage of overpredictions ( $\hat{P}O$ ) and the theoretical quantile ( $\tau$ ). Estimation of 95th quantile of daily mean 10m wind. Basic model (**a**), model with information on t2m in North America at lag -2 days (**b**), model with same information as above plus jet stream and NAO information at lag -1 days (**c**). **d-f**: As a-c, but for the 99th quantile of daily mean 10m wind speed. In each row, the colourscale is capped at  $1-\tau$ .

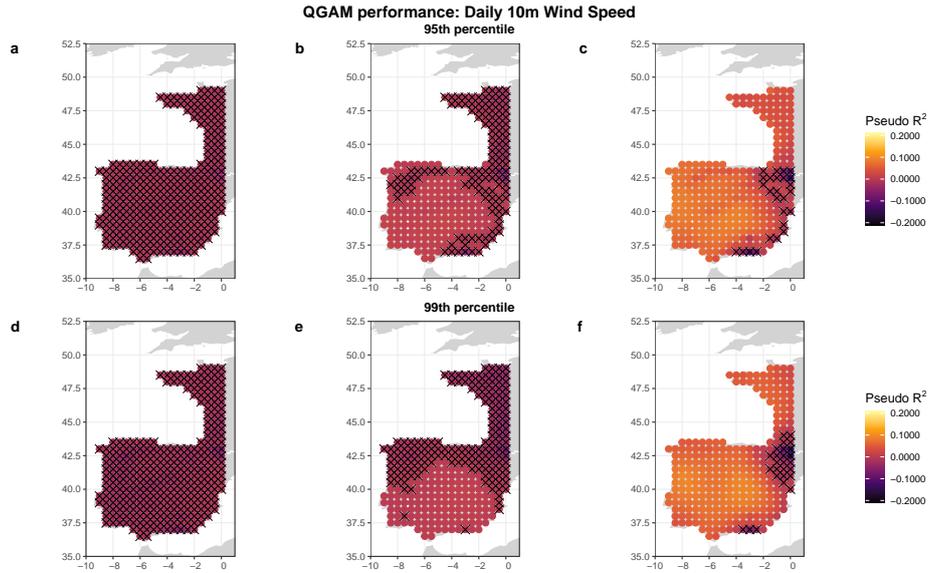
377 Similar conclusions can be drawn for daily precipitation extremes (Figure 5). The  
 378 bias of the models at most grid points is acceptable, as it is well under the  $1-\tau$  thresh-  
 379 old.



**Figure 5.** As Fig. 4 but for daily precipitation.

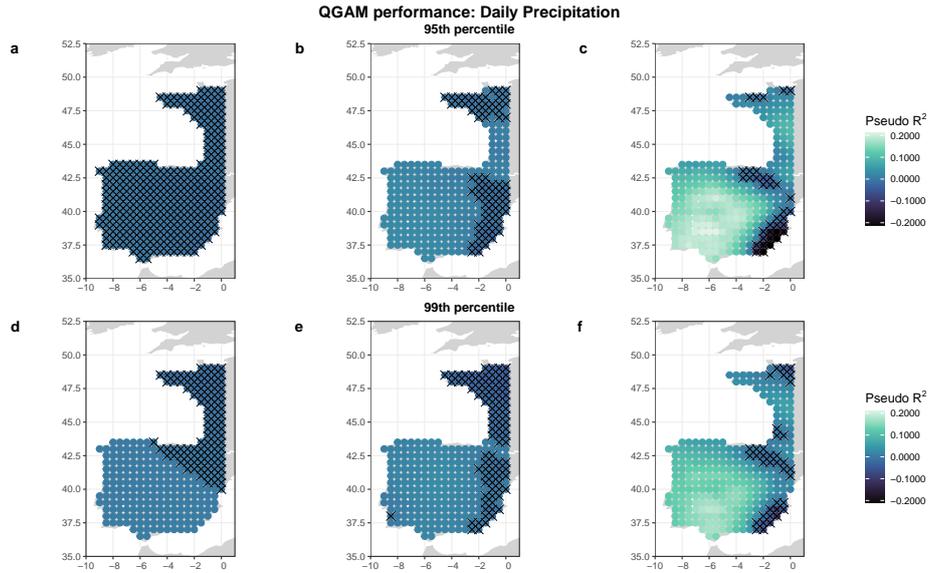
380 Figures 6 and 7 show the performance of QGAMs in terms of pseudo  $R^2$ . The pseudo  
 381  $R^2$  can take any values between one and minus one, where one represents a perfect model  
 382 and minus one the worst possible model. Here, a pseudo  $R^2$  over zero indicates that the  
 383 model is better than the seasonal climatology of the quantile of interest at the given grid  
 384 point, whereas values under zero indicate that it performs worse than the seasonal clim-  
 385 atology. Stippling is added to grid points with a pseudo  $R^2$  lower than zero.

386 All QGAMs predicting extreme quantiles of daily mean 10m wind speed appear  
 387 to gradually improve their performance as they are provided with more information on  
 388 the upstream large-scale atmospheric state. Sizeable gains can be observed already when  
 389 adding information on t2m to the models, so that cold spell models (Figures 6 b and e)  
 390 are better than the seasonal climatology alone at most grid points. However, the largest  
 391 gains occur at the last step, after the information on the state of the North Atlantic and  
 392 the Polar jet is added. Figures 6 c and f show that cold spell and jet stream models are  
 393 better than the seasonal climatology alone at all grid points apart from on the Pyrenees  
 394 and some mountainous areas in South-Eastern Spain, where the difference between the  
 395 two models is in any case small.



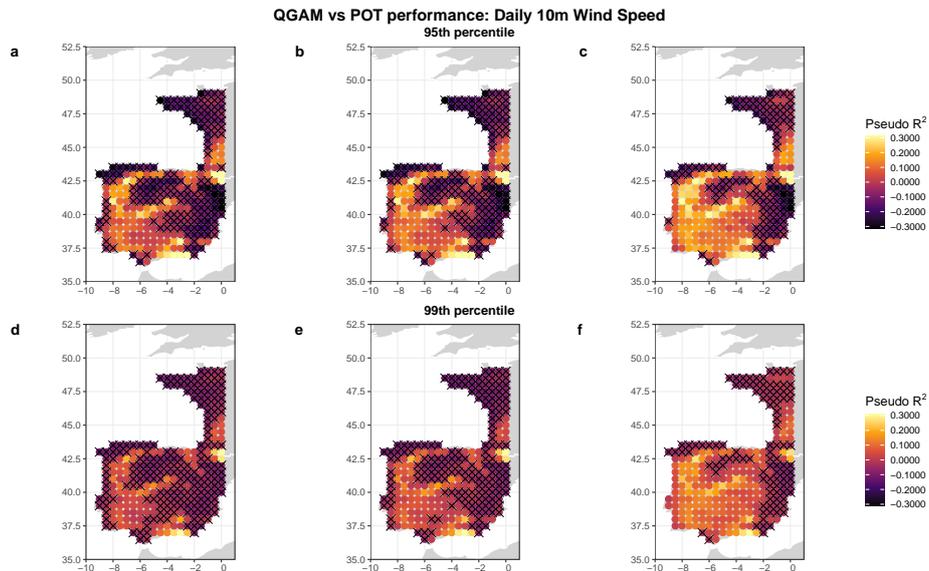
**Figure 6.** As Figure 4 but for pseudo  $R^2$ . Stippling indicates that the QGAM performs worse than the seasonal climatology of the quantile of interest.

396 The overall trend is similar for models predicting extreme quantiles of daily precipi-  
 397 tation. All the models in Figure 7 show gradual improvements as more information  
 398 is added, with the largest improvement occurring as above at the last step. This sug-  
 399 gests that information on the state of the North Atlantic atmospheric circulation is key  
 400 to the predictability of surface extremes in Western Europe following North American  
 401 cold spells. The basic model is approximately equivalent to the seasonal climatology of  
 402 the quantile, whereas the cold spell and jet stream models systematically outperform the  
 403 seasonal quantile (and the base model) for both 10m wind speed and precipitation. Even  
 404 in the case of precipitation, QGAMs perform poorly in some mountainous areas, sug-  
 405 gesting that they may have difficulties in accounting for the effect of local orographic fea-  
 406 tures.



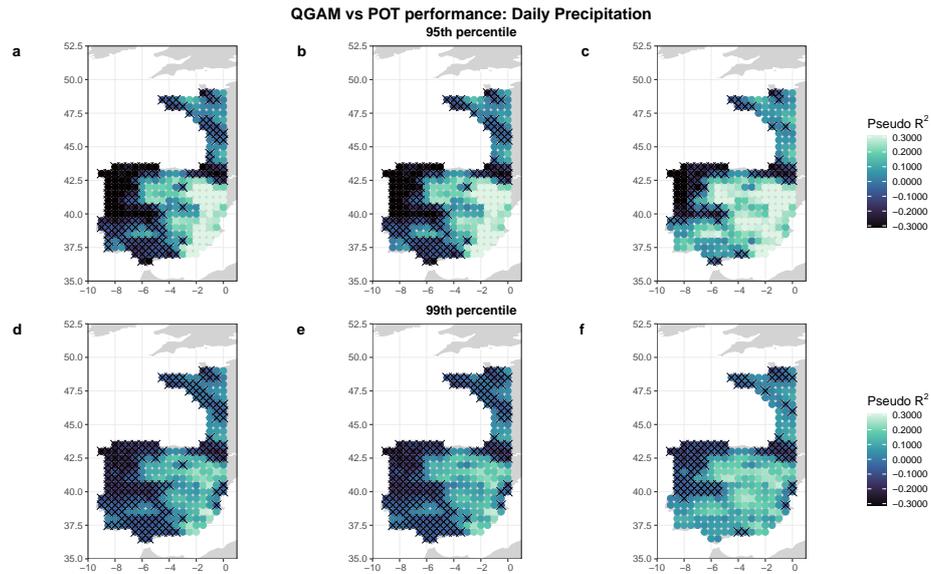
**Figure 7.** As Figure 5 but for pseudo  $R^2$ . Stippling indicates that the QGAM performs worse than the seasonal climatology of the quantile of interest.

407 Figure 8 and 9 compare the performance of QGAMs to conventional POT mod-  
 408 els. The same set of variables is used for the two models, and a new Pseudo  $R^2$  is com-  
 409 puted using the POT models as baseline. A negative pseudo  $R^2$  at a given grid point is  
 410 to be interpreted as POT models performing better for that grid point, whereas a pos-  
 411 itive pseudo  $R^2$  suggests that the QGAMs perform better.



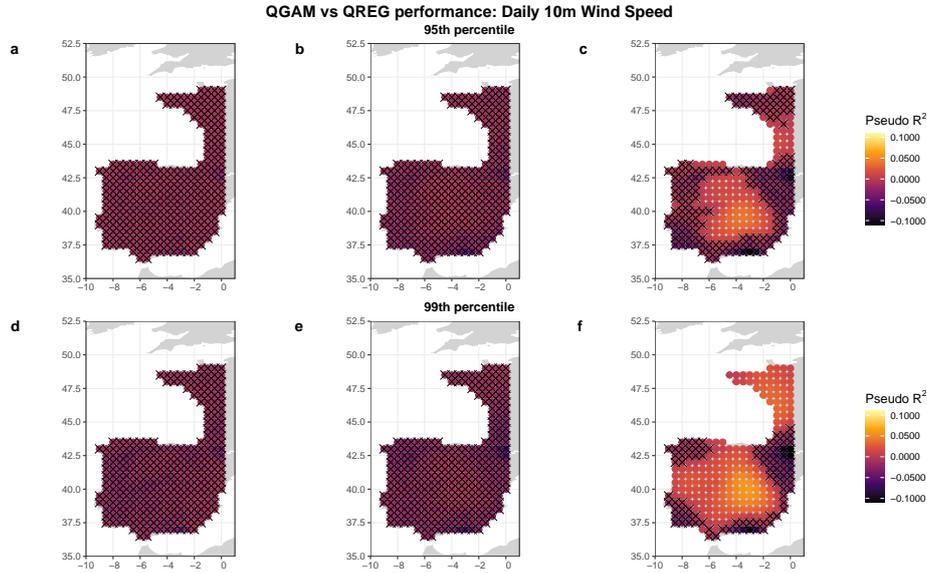
**Figure 8.** As Figure 6, but using the linear POT model as baseline for the computation of the pseudo  $R^2$ . Stippling indicates that the QGAM performs worse than the baseline model.

412 QGAMs become gradually better than POT models at most grid points as more  
 413 information is added to the models, both when predicting daily mean 10m wind speed  
 414 (Figure 8) and daily precipitation (Figure 9). This suggests that QGAMs are better than  
 415 POT models at modelling the non-linear effects of upstream atmospheric factors on the  
 416 surface extremes as a whole. However, some regional differences can be observed. In the  
 417 case wind, QGAMs are clearly superior for Western and Central Iberia, whereas the dif-  
 418 ference is smaller in Eastern Iberia and Western France. POT models outperform QGAMs  
 419 in mountainous areas and in North-Western France. In the case of precipitation, QGAMS  
 420 outperform POT models almost everywhere, with the exception of North-Western Iberia  
 421 and the Pyrenees. The fact that the QGAMs struggle in mountainous areas is consis-  
 422 tent with what found for the comparison with the seasonal climatology of the quantile  
 423 (Figures 6 and 7)



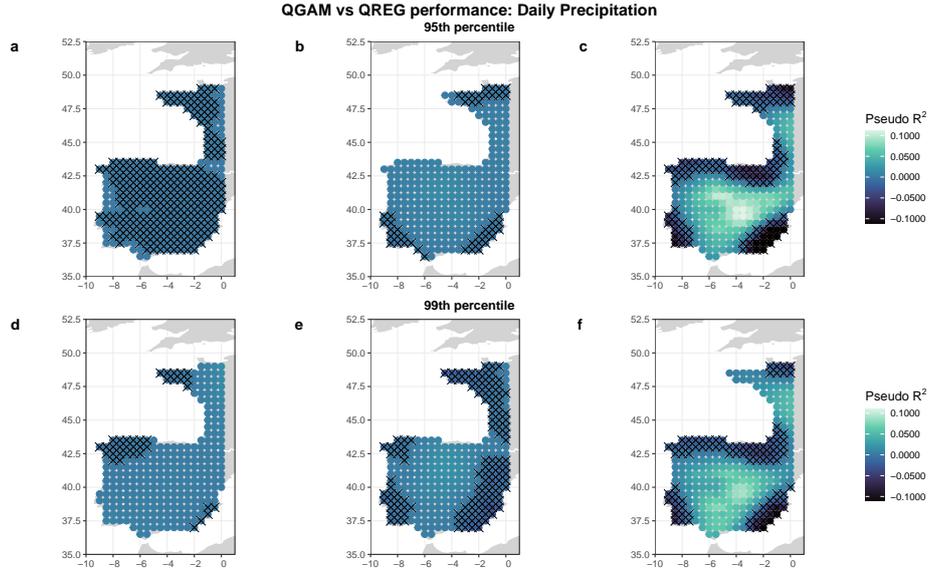
**Figure 9.** As Figure 7, but using the linear POT model as baseline for the computation of the pseudo  $R^2$ . Stippling indicates that the QGAM performs worse than the baseline model.

424 Figures 10 and 11 display a comparison between QGAMs and QREG models. The  
 425 comparison is performed similarly to the previous case, where in this case the QREG mod-  
 426 els are used as baseline for the Pseudo  $R^2$  computation.



**Figure 10.** As Figure 6, but using the quantile regression model as baseline for the computation of the pseudo  $R^2$

427 The difference between the two models is overall smaller in this case, with QREG  
 428 models generally performing better for simpler models and QGAMs improving their per-  
 429 formance relative to QREG models as more information is added (Figures 10 and 11).  
 430 A difference between the models predicting daily mean 10m wind speed and those pre-  
 431 dicting daily precipitation is that in the first case QGAMs show significant improvements  
 432 compared to QREG models only when information on both t2m in North America and  
 433 the Polar jet stream and NAO are added to the model (Figure 10), whereas in the sec-  
 434 ond case the trend is less clear, as QGAMs outperform QREG models at most locations  
 435 already when information on t2m is added to the model (Figure 11). In the cold spell  
 436 and jet stream models, QGAMs outperform QREG models in inland Iberia, perform sim-  
 437 ilarly to QREG models on the coast in North-Western France, and are outperformed in  
 438 the Pyrenees and other mountainous regions (Figures 10 c,f and 11 c,f). This difference  
 439 appears to be consistent across output variable and quantile of choice.



**Figure 11.** As Figure 7, but using the quantile regression as baseline for the computation of the pseudo  $R^2$

440 An overview of the overall performance of the cold spell and jet stream models in  
 441 terms of Pseudo  $R^2$  is given in Table 1. As suggested from previous figures, QGAMs per-  
 442 form overall better than conventional alternatives (corresponding to positive Pseudo  $R^2$   
 443 values in the table 1) when both t2m and atmospheric circulation information is provided.  
 444 This holds in all cases bar when estimating the 95th quantile of 10m wind speed, where  
 445 QREG models perform approximately at the same level.

**Table 1.** Overall pseudo  $R^2$  of cold spell and jet stream QGAMs used to forecast the 95th and 99th quantiles of daily mean 10m wind speed and daily precipitation compared to different baseline models.

Variable	Baseline model	95th quantile	99th quantile
Daily mean 10m wind speed	Quantile of seasonal climatology	0.0476	0.0495
Daily mean 10m wind speed	POT	0.0316	0.0353
Daily mean 10m wind speed	QREG	-0.0036	0.0085
Daily precipitation	Quantile of seasonal climatology	0.0982	0.0830
Daily precipitation	POT	0.0713	0.0518
Daily precipitation	QREG	0.0125	0.0155

446 In the supporting information (Figures S1-S8), we repeat the same analysis per-  
 447 formed in this section while adding the previous lag of the variable of interest to the mod-  
 448 els. This is done to show how the comparison between models changes when we correct  
 449 for the autocorrelation in the extremes. This autocorrelation is ignored here since the  
 450 models do not make use of any information on the outcome variable of interest in Eu-  
 451 rope, except for the training set-based seasonal climatology. The trends observed in the  
 452 main analysis largely hold even in the supporting materials, with the key difference be-  
 453 ing that QGAMs improve further in comparison to other models.

## 5 Discussion and Conclusions

This paper has introduced QGAMs (Fasiolo et al., 2021) as an alternative to conventional parametric methods for the analysis of spatially compounding climate extremes. Through a case study on pan-Atlantic cold spells in North America and wet or windy extremes in Europe, it has been shown that non-parametric quantile-based methods generally forecast near-surface weather extremes with a short return period more accurately than through a conventional POT approach (Tables 1 and S1). The difference between QGAMs and QREG models is relatively small, with QGAMs generally being superior when introducing information on the state of Polar jet and the North Atlantic atmosphere into the models (Figures 10 and 11 and Tables 1 and S1). This suggests that QGAMs may be recommended over other techniques when a larger number of drivers is explored. The advantage of QGAMs over alternative techniques widens when information on the autocorrelation of the extremes is added to the models (Figures S7 and S8).

Despite the overall superiority of QGAMs, some interesting regional differences could be observed, with QGAMs performing at their best in inland Iberia, and at their worst in the Pyrenees and in other mountainous regions (Figures 6 and 7). The poor performance of QGAMs in those regions is probably to be ascribed to the lack of orographic information in the models and the relatively coarse spatial resolution, which make it hard for QGAMs to reconstruct realistic spatial patterns. The fact that other models seems to suffer less from this may be due to the fact that simpler linear models give greater relative importance to the seasonal climatology compared to spatial features.

Non-parametric quantile methods have additional applications for the analysis of near-surface extremes, which we have touched upon in this paper. First, they may be used to assess the impact of a particular driver on a given quantile of a downstream atmospheric variable of interest. This is particularly useful for spatially compounding extremes, namely extremes that occur (near-)simultaneously in geographically remote regions. In this study, we used QGAMs to show that lower area-averaged 2m temperatures in North America are significantly associated with higher values of the extreme quantiles of daily mean 10m wind speed and daily precipitation in Western Europe (Figures 2 and 3). Second, non-parametric quantile-based methods may be used to provide ranges of uncertainty to deterministic numerical forecasts. In particular, Figures 4 and 5 show that QGAMs overpredict the value of the variable of interest a percentage of times close to  $\tau$ , thus displaying a good empirical coverage of the upper boundary of uncertainty when used for estimation of the range of uncertainty of the forecast.

This paper focused on a specific case study of previously studied spatially compounding extremes. We considered a limited range of conventional models for comparison to QGAMs, and only tested a small number of possible large-scale dynamical drivers of the extremes. Our work should be viewed as a proof-of-concept to show the potential of QGAMs compared to conventional parametric models for the study and understanding of spatially compounding extremes, rather than an attempt to build a statistical forecast model or to investigate novel extreme occurrences and the related large-scale atmospheric drivers. We also note that, even though QGAMs perform better than linear POT models in forecasting near-surface extremes with short return periods, it does not mean that they are equally effective for extremes with longer return periods. More research may be needed to verify the robustness of QGAMs when analysing extremes of this nature.

The statistical analysis in this paper should also be contextualised relative to previous research on Pan-Atlantic compound extremes (Messori et al., 2016; De Luca et al., 2020; Leeding et al., 2023; Messori & Faranda, 2023). Our results strengthen the hypothesis of a connection between wintertime North American cold spells and wet or windy extremes in Western Europe, by showing that introducing information on surface temperature in North America has a clear effect in the model and significantly improves the prediction of extreme quantiles of 10m wind and precipitation in Iberia and Western France (Figures 3, 6, 7). The fact that the effect of temperature in North America is weakened but still significant when adding information on the Polar jet and the state of the North Atlantic atmosphere to the models, points to the presence of teleconnections which can-

509 not be fully explained by a simple causal flow in which cold spells influence the behaviour  
 510 of the Polar jet, which in turn affects surface weather in Europe. This might suggest the  
 511 presence of more complex or multiple pathways through which Pan-Atlantic compound  
 512 extremes may be engendered.

## 513 6 Open Research

514 The ERA 5 data used in this study is freely available from the Copernicus Climate  
 515 Change service at <https://doi.org/10.24381/cds.adbb2d47> (Hersbach et al., 2020)  
 516 and <https://doi.org/10.24381/cds.bd0915c6> (Hersbach et al., 2020). The daily NAO  
 517 index is available through the NOAA online archive at [https://www.cpc.ncep.noaa](https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml)  
 518 [.gov/products/precip/CWlink/pna/nao.shtml](https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml) (NOAA/ National Weather Service,  
 519 2022). The models have been built and tested with the help of open-source R Statisti-  
 520 cal Software (R Core Team, 2021). All software is freely available through CRAN at [https://](https://cran.r-project.org/)  
 521 [cran.r-project.org/](https://cran.r-project.org/). More details can be obtained by contacting the corresponding  
 522 author L.Olivetti, [leonardo.olivetti@geo.uu.se](mailto:leonardo.olivetti@geo.uu.se).

## 523 Acknowledgments

524 L. Olivetti and G. Messori thankfully acknowledge the support of the European  
 525 Research Council (ERC) under the European Union’s Horizon 2020 research and inno-  
 526 vation programme (project CENÆ: compound Climate Extremes in North America and  
 527 Europe: from dynamics to predictability, Grant Agreement No. 948309). The authors  
 528 also acknowledge valuable discussions with R. Leeding, J. Riboldi and A. Segalini.

## 529 7 Conflict of interest

530 The authors declare no conflicts of interest relevant to this study.

## 531 References

- 532 Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*.  
 533 Coumou, D., & Rahmstorf, S. (2012, July). A decade of weather extremes. *Nature*  
 534 *Climate Change*, 2(7), 491–496. (Number: 7 Publisher: Nature Publishing  
 535 Group) doi: 10.1038/nclimate1452  
 536 Davison, A. C., & Smith, R. L. (1990). Models for Exceedances over High Thresh-  
 537 olds. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(3),  
 538 393–442. (Publisher: Royal Statistical Society, Wiley)  
 539 De Luca, P., Messori, G., Pons, F. M. E., & Faranda, D. (2020). Dynamical systems  
 540 theory sheds new light on compound climate extremes in Europe and East-  
 541 ern North America. *Quarterly Journal of the Royal Meteorological Society*,  
 542 146(729), 1636–1650. doi: 10.1002/qj.3757  
 543 Dodet, G., Castelle, B., Masselink, G., Scott, T., Davidson, M., Floc’h, F., . . .  
 544 Suarez, S. (2019). Beach recovery from extreme storm activity during the  
 545 2013–14 winter along the Atlantic coast of Europe. *Earth Surface Processes*  
 546 *and Landforms*, 44(1), 393–401. doi: 10.1002/esp.4500  
 547 do Nascimento, F. F., Azevedo, A. A., & Ferraz, V. R. d. S. (2021). Regression mod-  
 548 els to dependence for exceedance. *Journal of Applied Statistics*, 48(16), 3048–  
 549 3059. doi: 10.1080/02664763.2020.1795088  
 550 Elvidge, S., & Angling, M. (2018, April). Using Extreme Value Theory for De-  
 551 termining the Probability of Carrington-Like Solar Flares. *Space Weather*, 16,  
 552 417–421. doi: 10.1002/2017SW001727  
 553 Fasiolo, M., Wood, S. N., Zaffran, M., Nedellec, R., & Goude, Y. (2021, July). Fast  
 554 calibrated additive quantile regression. *Journal of the American Statistical As-*

- 555 *sociation*, 116(535), 1402–1412.
- 556 Ferreira, A., & Haan, L. d. (2015, February). On the block maxima method  
557 in extreme value theory: PWM estimators. *The Annals of Statistics*,  
558 43(1), 276–298. (Publisher: Institute of Mathematical Statistics) doi:  
559 10.1214/14-AOS1280
- 560 Fisher, R. A., & Tippett, L. H. C. (1928, April). Limiting forms of the frequency  
561 distribution of the largest or smallest member of a sample. *Mathematical Pro-*  
562 *ceedings of the Cambridge Philosophical Society*, 24(2), 180–190. (Publisher:  
563 Cambridge University Press) doi: 10.1017/S0305004100015681
- 564 French, J., Kokoszka, P., Stoev, S., & Hall, L. (2019, March). Quantifying the  
565 risk of heat waves using extreme value theory and spatio-temporal func-  
566 tional data. *Computational Statistics & Data Analysis*, 131, 176–193. doi:  
567 10.1016/j.csda.2018.07.004
- 568 Gumbel, E. J. (1941). The Return Period of Flood Flows. *The Annals of Mathemat-*  
569 *ical Statistics*, 12(2), 163–190. (Publisher: Institute of Mathematical Statis-  
570 tics)
- 571 Gumbel, E. J. (1958). *Statistics of Extremes*. Columbia University Press. (Publica-  
572 tion Title: Statistics of Extremes) doi: 10.7312/gumb92958
- 573 Hastie, T., & Tibshirani, R. (1986, August). Generalized Additive Models. *Statisti-*  
574 *cal Science*, 1(3), 297–310. (Publisher: Institute of Mathematical Statistics)
- 575 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J.,  
576 ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of*  
577 *the Royal Meteorological Society*, 146(730), 1999–2049. doi: 10.1002/qj.3803
- 578 Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.
- 579 James, G., Hastie, T., Witten, D., & Tibshirani, R. (2022). *An introduction to sta-*  
580 *tistical learning with applications in R* (Second ed.). New York: Springer Sci-  
581 ence.
- 582 Koenker, R. (2011, November). Additive models for quantile regression: Model selec-  
583 tion and confidence band-aids. *Brazilian Journal of Probability and Statistics*,  
584 25(3), 239–262. (Publisher: Brazilian Statistical Association) doi: 10.1214/10  
585 -BJPS131
- 586 Koenker, R., & Bassett, G. (1978). Regression Quantiles. *Econometrica*, 46(1), 33–  
587 50. (Publisher: [Wiley, Econometric Society]) doi: 10.2307/1913643
- 588 Koenker, R., & Hallock, K. F. (2001, December). Quantile Regression. *Journal of*  
589 *Economic Perspectives*, 15(4), 143–156. doi: 10.1257/jep.15.4.143
- 590 Koenker, R., & Machado, J. A. F. (1999, December). Goodness of Fit and Related  
591 Inference Processes for Quantile Regression. *Journal of the American Statisti-*  
592 *cal Association*, 94(448), 1296–1310. (Publisher: Taylor & Francis) doi: 10  
593 .1080/01621459.1999.10473882
- 594 Kornhuber, K., Coumou, D., Vogel, E., Lesk, C., Donges, J. F., Lehmann, J., &  
595 Horton, R. M. (2020, January). Amplified Rossby waves enhance risk  
596 of concurrent heatwaves in major breadbasket regions. *Nature Climate*  
597 *Change*, 10(1), 48–53. (Number: 1 Publisher: Nature Publishing Group)  
598 doi: 10.1038/s41558-019-0637-z
- 599 Lee, M.-Y., Hong, C.-C., & Hsu, H.-H. (2015). Compounding effects of warm  
600 sea surface temperature and reduced sea ice on the extreme circulation over  
601 the extratropical North Pacific and North America during the 2013–2014  
602 boreal winter. *Geophysical Research Letters*, 42(5), 1612–1618. doi:  
603 10.1002/2014GL062956
- 604 Leeding, R., Riboldi, J., & Messori, G. (2023, March). On Pan-Atlantic cold, wet  
605 and windy compound extremes. *Weather and Climate Extremes*, 39, 100524.  
606 doi: 10.1016/j.wace.2022.100524
- 607 Lucarini, V., Faranda, D., Freitas, A. C. G. M. M. d., Freitas, J. M. M. d., Holland,  
608 M., Kuna, T., ... Vienti, S. (2016). *Extremes and Recurrence in Dynamical*  
609 *Systems* (1st edition ed.). Hoboken, New Jersey: Wiley.

- 610 Mares, C., Mares, I., & Stanciu, A. (2009, March). Extreme value analysis in the  
 611 Danube lower basin discharge time series in the twentieth century. *Theoretical  
 612 and Applied Climatology*, *95*(3), 223–233. doi: 10.1007/s00704-008-0001-0
- 613 Messori, G., Caballero, R., & Gaetani, M. (2016). On cold spells in North America  
 614 and storminess in western Europe. *Geophysical Research Letters*, *43*(12), 6620–  
 615 6628.
- 616 Messori, G., & Faranda, D. (2023). On the Systematic Occurrence of Com-  
 617 pound Cold Spells in North America and Wet or Windy Extremes in  
 618 Europe. *Geophysical Research Letters*, *50*(7), e2022GL101008. doi:  
 619 10.1029/2022GL101008
- 620 Mills, E. (2005, August). Insurance in a Climate of Change. *Science*, *309*(5737),  
 621 1040–1044. (Publisher: American Association for the Advancement of Science)  
 622 doi: 10.1126/science.1112121
- 623 NOAA/ National Weather Service. (2022). *Climate Prediction Center - Telecon-*  
 624 *nections: North Atlantic Oscillation*. Retrieved 2022-12-22, from [https://www](https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml)  
 625 [.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml](https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml)
- 626 Passow, C., & Donner, R. V. (2019). A Rigorous Statistical Assessment of Recent  
 627 Trends in Intensity of Heavy Precipitation Over Germany. *Frontiers in Envi-*  
 628 *ronmental Science*, *7*.
- 629 Pickands, J. (1975). Statistical Inference Using Extreme Order Statistics. *The An-*  
 630 *nals of Statistics*, *3*(1), 119–131. (Publisher: Institute of Mathematical Statis-  
 631 tics)
- 632 R Core Team. (2021). *R: A Language and Environment for Statistical Computing*  
 633 [software]. Retrieved from <https://www.R-project.org/> (Vienna, Austria. R  
 634 Foundation for Statistical Computing)
- 635 Smith, R. L. (1984). Threshold Methods for Sample Extremes. In J. T. de  
 636 Oliveira (Ed.), *Statistical Extremes and Applications* (pp. 621–638). Dordrecht:  
 637 Springer Netherlands. (doi: 10.1007/978-94-017-3069-3\_48)
- 638 Trenary, L., DelSole, T., Tippett, M. K., & Doty, B. (2015, December). Was the  
 639 Cold Eastern Us Winter of 2014 Due to Increased Variability? *Bulletin of the*  
 640 *American Meteorological Society*, *96*(12), S15–S19. doi: 10.1175/BAMS-D-15  
 641 -00138.1
- 642 Wild, S., Befort, D. J., & Leckebusch, G. C. (2015, December). Was the Extreme  
 643 Storm Season in Winter 2013/14 Over the North Atlantic and the United  
 644 Kingdom Triggered by Changes in the West Pacific Warm Pool? *Bulletin of*  
 645 *the American Meteorological Society*, *96*(12), S29–S34. (Publisher: Ameri-  
 646 can Meteorological Society Section: Bulletin of the American Meteorological  
 647 Society) doi: 10.1175/BAMS-D-15-00118.1
- 648 Wood, S. (2017). *Generalized Additive Models: An Introduction With R* (Vol. 66).  
 649 Productivity Press.
- 650 Yee, T. W. (2015). *Vector Generalized Linear and Additive Models*. Springer.