*Journal of Energy Markets* **11**(3), 1–24
DOI: 10.21314/JEM.2018.179



**Research Paper**

# A latent trawl process model for extreme values

**Ragnhild C. Noven, Almut E. D. Veraart and Axel Gandy**

Department of Mathematics, Imperial College London, 180 Queen's Gate, London SW7 2AZ, UK; emails: ragnhild.noven@gmail.com, a.veraart@imperial.ac.uk, a.gandy@imperial.ac.uk

(Received December 18, 2017; accepted March 20, 2018)

## ABSTRACT

This paper presents a new model for characterizing temporal dependence in exceedances above a given threshold. Our model is based on a class of stationary, infinitely divisible stochastic processes known as trawl processes. For use with extreme values, our model is constructed by embedding a trawl process in a hierarchical framework. This ensures that the marginal distribution is a generalized Pareto, as expected from classical extreme value theory. We also consider a modified version of this model that works with a wider class of generalized Pareto distributions (GPDs) and has the advantage of separating marginal and temporal dependence properties. The model is illustrated via various applications to environmental time series; thus, we show that the model offers considerable flexibility in capturing the dependence structure of extreme value data.

**Keywords:** trawl process; peaks over threshold; generalized Pareto distribution (GPD); pairwise likelihood estimation; marginal transformation model; conditional tail dependence coefficient.

Corresponding author: A. E. D. Veraart

Print ISSN 1756-3607 | Online ISSN 1756-3615

## 1 INTRODUCTION

Modeling dependencies in extreme value data is a topic of growing importance, with applications in a number of fields such as hydrology (de Haan and de Ronde 1998), oceanography (Coles and Tawn 1991), financial risk management (Embrechts *et al* 1997; Ledford and Tawn 2003) and environmental science (Davison *et al* 2012; Heffernan and Tawn 2004).

The main contribution of this paper is a new extreme value model that can account for serial dependence in the extremes; this extends the hierarchical setup of Bortot and Gaetan (2014). The starting point for this model is the observation that the marginal distribution of exceedances converges to a generalized Pareto distribution (GPD; see Davison and Smith (1990)). Bortot and Gaetan (2014) use a decomposition of the GPD to construct a hierarchical model for exceedances that preserves this distribution marginally. We adopt this hierarchical structure and then proceed to introduce a new model incorporating the properties of so-called trawl processes. The original approach of Bortot and Gaetan (2014) involved using a Markov chain to generate dependence in the exceedances. By using a trawl process instead of a Markov chain, we obtain a more flexible dependence structure. Moreover, the trawl process framework provides a unified procedure for generating processes with a given infinitely divisible distribution and autocovariance function, whereas Bortot and Gaetan (2014) consider two different specifications for the latent Markov chain with the same marginal distribution and autocovariance function.

We also consider a modification of the new model that makes it suitable for use with any GPD, thus removing the restriction (inherited from the original model of Bortot and Gaetan (2014)) that the shape parameter of the GPD has to be positive. This means that our modified model can be used for processes with less heavy tails (often to be found in environmental applications), as illustrated by its application to air pollution data. The modification also improves the interpretability of the parameters and appears to make the estimation procedure more efficient.

We remark that our new dynamic model for environmental variables (such as precipitation and ozone levels) can also be used as a basis for designing suitable hedges through weather derivatives. Weather derivatives on rainfall have been traded in the past and there is current interest in setting up derivatives that will allow us to hedge against other climate variables.

This paper is structured as follows. Section 2 introduces the latent trawl process model for dependent extremes in a hierarchical setup. Section 3 discusses parameter estimation and inference, and develops a measure for the extremal dependence structure that is adapted to the model. Section 4 applies the model to two different examples of environmental time series (rainfall and air pollution), and Section 5 concludes. Details of our estimation procedure are presented along with the proof of our theoretical results in the online appendix.

## 2 LATENT TRAWL PROCESS MODEL

### 2.1 Basic structure

This subsection introduces the hierarchical structure used for the extreme value model, which is taken from Bortot and Gaetan (2014). Throughout the paper, we work on a probability space $(\Omega, \mathcal{F}, P)$. Consider a discrete-time stochastic process denoted by $\{Y_j\}$ that is assumed to be strictly stationary. We assume that we observe the process at times $j = 1, \dots, k$ for $k \in \mathbb{N}$, looking at extreme values of $Y_j$, meaning that $Y_j > u$ for a fixed threshold $u$.

In order to focus on extreme values only, we will consider the values and occurrence times of any exceedances. To this end, we define the exceedances $X_j$ as follows:

$$X_j := \max(Y_j - u, 0), \quad j = 1, \dots, k. \tag{2.1}$$

From standard extreme value theory (see, for example, Davison and Smith 1990; Pickands 1975), assuming $\{Y_j\}$ are in the domain of attraction of some extreme value distribution, the conditional exceedances $\{X_j \mid X_j > 0\}$ converge to a GPD for an appropriate sequence of thresholds $u_n \to \infty$. Based on this result, we will assume that the conditional distribution of $X_j$ given $X_j > 0$ can be approximated by a GPD for a sufficiently large threshold $u$. The density of the GPD is written as

$$f_{\text{GPD}}(x \mid \alpha, \beta) = \frac{\alpha}{\beta}\left(1 + \frac{x}{\beta}\right)_{+}^{-(\alpha+1)}, \quad x \geqslant 0, \ \alpha, \beta > 0,$$

where $y_+ = \max(0, y)$, which is a reparameterization of the standard density with shape parameter $\xi = 1/\alpha$ and scale parameter $\sigma = \beta/\alpha$.

Following Reiss and Thomas (2007) as well as Bortot and Gaetan (2014), the GPD can be represented as a mixture of an exponential random variable with a gamma-distributed parameter, motivating a hierarchical specification for the exceedance process $\{X_j\}$. In particular, we assume that the distribution of $X_j$ depends on the value of a latent process $\Lambda$ at time $j$, denoted by $\Lambda_j$. This latent process determines both the probability of observing an exceedance (corresponding with $X_j > 0$) and the distribution of the exceedances.

Specifically, we assume that, conditional on the latent process $\Lambda$, the random variables $X_j$ are independent and

$$X_j \mid (X_j > 0, \Lambda_j) \sim \text{Exp}(\Lambda_j). \tag{2.2}$$

To ensure the threshold exceedances follow the GPD, we can use any stationary stochastic process $\Lambda$ that has a gamma marginal law. The precise specification of $\Lambda$ will be discussed in Section 2.2.

Next, following Bortot and Gaetan (2014), we assume that

$$P(X_j > 0 \mid \Lambda_j) = \mathrm{e}^{-\kappa \Lambda_j}, \tag{2.3}$$

where the parameter $\kappa > 0$ is linked with the proportion of the occurrence of exceedances above the threshold $u$. Combining (2.2) and (2.3), we find that the conditional density of $X_j$ given $\Lambda_j = \lambda_j$ has the following functional form:

$$f(x_j \mid \lambda_j) = \begin{cases} 1 - \mathrm{e}^{-\kappa\lambda_j}, & x_j = 0, \\ \mathrm{e}^{-\kappa\lambda_j}\lambda_j \mathrm{e}^{-\lambda_j x_j}, & x_j > 0, \end{cases} \tag{2.4}$$

where the density $f$ is defined with respect to the measure $\mu(\mathrm{d}x_j) = \delta_0(\mathrm{d}x_j) + \mathrm{d}x_j$. This construction shows that the exceedance $X_j$ is generated by a two-stage process, conditional on the value of $\Lambda_j$. First, $X_j$ is set to zero with probability $1 - \mathrm{e}^{-\kappa\lambda_j}$. Second, the distribution of $X_j$ given $\{X_j > 0,\ \Lambda_j = \lambda_j\}$ is exponential with parameter $\lambda_j$. The latent process $\Lambda$ may be interpreted as an inverse intensity, as higher values of $\Lambda$ give a lower probability of exceeding the threshold $u$ and a smaller expected value of exceedances.

Since we require the observations $X_j$ to be independent for the distinct values of $j$, conditional on the corresponding values of $\Lambda$, the conditional joint density of $(X_1, \dots, X_k)$ factorizes and can be written as

$$f(x_1, \dots, x_k \mid \lambda_1, \dots, \lambda_k) = \prod_{j=1}^{k} f(x_j \mid \lambda_j). \tag{2.5}$$

This specification implies that any dependence between observations $X_1, \dots, X_k$ comes from the dependence between corresponding elements of the latent process $\Lambda$.

To complete the GPD mixture construction, the latent process requires a gamma marginal law, that is, $\Lambda_j \sim \mathrm{Gamma}(\alpha, \beta)$ for $\alpha, \beta > 0$. This implies that the corresponding density is given by

$$f_{\Lambda_j}(x) = \beta^{\alpha}\Gamma(\alpha)^{-1}x^{\alpha-1}\mathrm{e}^{-\beta x} \quad \text{for } x > 0,$$

and the characteristic function is given by

$$\mathbb{E}(\exp(\mathrm{i}u\Lambda_j)) = \exp(C(u, \Lambda_j)),$$

where $C(u, \Lambda_j) = -\alpha \log(1 - \mathrm{i}u/\beta)$ denotes the corresponding cumulant function, which is the distinguished logarithm of the characteristic function (see Sato 1999,

p. 33). This specification introduces the restriction $\alpha > 0$, meaning that the model can only capture data belonging to the Fréchet distribution class. Section 2.4 presents a modified version of the model that removes this restriction.

A straightforward computation shows that, when using the above specification, the exceedances $\{X_j : X_j > 0\}$ have a GPD$(\alpha, \beta + \kappa)$ marginal law. Further, the unconditional probability of observing an exceedance is given by

$$P(X_j > 0) = E_\Lambda[\mathrm{e}^{-\kappa\Lambda}] = \left(1 + \frac{\kappa}{\beta}\right)^{-\alpha}. \tag{2.6}$$

## 2.2 Latent trawl process

The previous subsection describes a general hierarchical model setup, using the same structure that is found in Bortot and Gaetan (2014). So far, the latent process $\Lambda$ has been specified as having a gamma marginal law only. Now we depart from the approach used in Bortot and Gaetan (2014), in which the latent process is assumed to be a Markov chain (specifically a Gaver and Lewis process (G-LP) or a Warren process (WP)). Instead, we consider a new model in which $\Lambda$ is a trawl process. In principle, any stationary process with gamma marginal law could be used in this construction. We will argue, however, that processes belonging to the "trawl" class are particularly suited to the purpose, since they allow us to model the serial correlation and the marginal distribution independently of each other.

The conditional independence assumption of the hierarchical model means that any dependence between observations comes from the latent process; hence, this process should have a flexible dependence structure. This explains our use of a trawl process capable of capturing a wide range of dependence structures (as discussed below). Using a trawl process also means that the observations $X_j$ can be seen as coming from a continuous-time process $(X_t)$, which is useful for statistical applications where there may be missing or irregularly spaced data.

Bortot and Gaetan (2014) consider two particular classes for the latent Markov chain, the G-LP and the WP, and proceed to show how these two classes result in different asymptotic properties of the extremes, even though they have the same autocorrelation function. In contrast, the latent trawl process in our model is specified by its trawl set, which corresponds with a particular autocorrelation function. As will be shown in Section 3.2, the resulting process is asymptotically independent, and the form of the dependence structure is influenced by the trawl set.

### *2.2.1 Definition and properties of the trawl process*

Let us now define the class of trawl processes and present the key properties common to all processes in this class. Trawl processes were introduced by Barndorff-Nielsen (2011) and have been further developed by Barndorff-Nielsen *et al* (2014), Shephard and Yang (2016, 2017) and Veraart (2018). They are stationary, infinitely divisible stochastic processes that are made up of two components: the Lévy basis and the trawl set. In order to define these components, we need to introduce the relevant notation first.

To this end, let $S$ be a Borel set in $\mathbb{R}^2$, with the associated Borel $\sigma$-algebra $\mathcal{S} = \mathcal{B}(S)$ and Lebesgue measure $\lambda^{\text{Leb}}$. Let $\mathcal{B}_b(S)$ be the subsets of $S$ with finite Lebesgue measure, ie, $\mathcal{B}_b(S) = \{A \in \mathcal{S} \colon \lambda^{\text{Leb}}(A) < \infty\}$. The purpose of the following definition is to set out what we mean by a homogeneous Lévy basis, which is the source of randomness in the trawl process.

Definition 2.1

(1) A random measure on $(S, \mathcal{B}(S))$ is a collection of $\mathbb{R}$-valued random variables $\{M(A) \colon A \in \mathcal{B}_b(S)\}$ such that, for any sequence $A_1, A_2, \dots$ of disjoint elements of $\mathcal{B}_b(S)$ with $\bigcup_{j=1}^{\infty} A_j \in \mathcal{B}_b(S)$, we have $M(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} M(A_j)$ almost surely.

(2) A random measure $M$ on $(S, \mathcal{S})$ is independently scattered if, for any sequence $A_1, A_2, \dots$ of disjoint elements of $\mathcal{B}_b(S)$, the random variables $M(A_1), M(A_2), \dots$ are independent.

(3) A random measure $M$ on $(S, \mathcal{S})$ is said to be infinitely divisible if, for each $n \in \mathbb{N}$, there exist $n$ independent and identically distributed (iid) random measures $Z_1^n, \dots, Z_n^n$ such that $M \overset{\text{d}}{=} Z_1^n + \cdots + Z_n^n$. In particular, infinite divisibility implies that, for any finite collection $A_1, \dots, A_n$ of elements of $\mathcal{B}_b(S)$, the random vector $(M(A_1), \dots, M(A_n))$ is infinitely divisible in $\mathbb{R}^n$.

(4) A random measure on $(S, \mathcal{S})$ is deemed stationary if, for any point $s \in S$ and finite collection $A_1, A_2, \dots, A_n \in \mathcal{B}_b(S)$ such that $A_i + \boldsymbol{s} \subset S$, we have

$$(M(A_1 + \boldsymbol{s}), M(A_2 + \boldsymbol{s}), \dots, M(A_n + \boldsymbol{s})) \overset{\text{d}}{=} (M(A_1), M(A_2), \dots, M(A_n)).$$

(5) A homogeneous Lévy basis $L$ on $(S, \mathcal{S})$ is a random measure that is independently scattered, infinitely divisible and stationary.

Let $L$ denote a homogeneous Lévy basis on $(S, \mathcal{S})$. Then, the characteristic function satisfies the fundamental relation

$$E[\exp\{\text{i}uL(A)\}] = \exp\{\lambda^{\text{Leb}}(A)K(u)\}, \quad \text{for } A \in \mathcal{B}_b(S), \tag{2.7}$$

where

$$K(u) = \text{i}u\mu - \tfrac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}} (\text{e}^{\text{i}uz} - 1 - \text{i}uz\mathbb{I}_{|z|\leqslant 1})\nu(\text{d}z) \tag{2.8}$$

for constants $\mu \in \mathbb{R}, \sigma^2 \geqslant 0$ as well as a Lévy measure $\nu$ (see, for example, Barndorff-Nielsen 2011; Rajput and Rosinski 1989). Since (2.8) takes the form of the cumulant function of an infinitely divisible random variable, we can associate a Lévy seed (denoted by $L'$) with the Lévy basis $L$, which may be defined as a random variable with law characterized by (2.8). We can then write $C(u, L') = K(u)$ for the corresponding cumulant function, leading us to conclude that

$$E[\exp\{\mathrm{i}uL(A)\}] = \exp\{\lambda^{\text{Leb}}(A)C(u, L')\}, \quad \text{for } A \in \mathcal{B}_b(S). \tag{2.9}$$

This shows that the law of $L(A)$ is fully determined by the Lévy seed $L'$ and the Lebesgue measure of the set $A$.

We can now define the class of trawl processes.

Definition 2.2 Let $A$ be any set in $\mathcal{B}_b(\mathbb{R} \times \mathbb{R})$, and define a collection of trawl sets $\{A_t\}$ by shifting $A$ along the $\mathbb{R}$-axis corresponding with the last coordinate, which represents time: $A_t = A + (0, t) := \{(a_1, a_2 + t) : (a_1, a_2) \in A\}$. Let $L$ denote a homogeneous Lévy basis. The trawl process $(\Lambda_t)_{t \in \mathbb{R}}$ is then defined by evaluating the homogeneous Lévy basis over the trawl set, ie, by setting $\Lambda_t = L(A_t)$ for $t \in \mathbb{R}$.

The trawl process definition can be written as a stochastic integral that will become useful for our calculations in the following. Specifically, we write

$$\Lambda_t = \int_{\mathbb{R}\times\mathbb{R}} \mathbb{I}_{A_t}(\xi, s) L(\mathrm{d}\xi, \mathrm{d}s) = \int_{\mathbb{R}\times\mathbb{R}} \mathbb{I}_A(\xi, s - t) L(\mathrm{d}\xi, \mathrm{d}s),$$

where points in $\mathbb{R}^2$ are denoted by $(\xi, s)$ for $\xi \in \mathbb{R}$, $s \in \mathbb{R}$, so that the last component corresponds with the time axis. We define the stochastic integral in the same sense that it is used in Rajput and Rosinski (1989) (see Barndorff-Nielsen *et al* (2015) for a review of the relevant integration theory).

From the definition of the trawl process, we can immediately deduce that the process is stationary and infinitely divisible, and that the characteristic function is given by (2.9) since $\Lambda_t \stackrel{\mathrm{d}}{=} \Lambda_0$. Moreover, the stochastic integral representation implies that the trawl process is also a so-called mixed moving-average process; it was shown in Fuchs and Stelzer (2013) that mixed moving-averge processes are mixing, so it follows that trawl processes are mixing and ergodic.

#### 2.2.1.1 A slice representation for the finite-dimensional distributions.

Next, we study the finite-dimensional distributions of a trawl process and derive what we call a slice representation for its characteristic function, which will be very useful for simulation and inference purposes later on.

To this end, consider a sequence $0 \leqslant t_1 \leqslant \cdots \leqslant t_k$ with $k \in \mathbb{N}$. Let us now derive the joint characteristic function of $(\Lambda_{t_1}, \ldots, \Lambda_{t_k})$. We write $\Lambda_j = \Lambda_{t_j}$ to simplify the exposition. Typically, we will choose $t_j = j$. We consider the union $A^{\cup,k} := \bigcup_{i=1}^{k} A_{t_i}$. Using the inclusion–exclusion principle, we construct what we call a "slice" partition $\{S_1, \dots, S_{n_k}\}$ of $A^{\cup,k}$, where $n_k$ denotes the number of elements in the partition. In addition to $\{S_1, \dots, S_{n_k}\}$ being a partition of $A^{\cup,k}$, we require the partition to be such that each trawl $A_{t_k}$ can be written as a union of elements of that partition. A further requirement is that the intersection of any number of trawl sets and trawl set complements is a union of subsets in the partition. For general trawls, one would need $n_k = 2^{k-1}$, whereas for monotonic trawls this number reduces to $k(k+1)/2$. For example, in the case where $k = 2$, a suitable slice partition of $A_{t_1} \cup A_{t_2}$ is given by $\{A_{t_1} \cap A_{t_2}, A_{t_1} \setminus A_{t_2}, A_{t_2} \setminus A_{t_1}\}$.

PROPOSITION 2.3 *For $u_1, \dots, u_k \in \mathbb{R}$, we have (using the notation introduced above) that*

$$E\left(\exp\left(\mathrm{i}\sum_{j=1}^{k} u_j \Lambda_j\right)\right) = \exp\left(\sum_{m=1}^{n_k} \lambda^{\mathrm{Leb}}(S_m) C(\theta_m^+; L')\right),$$

$$\textit{for } u_m^+ := \sum_{\substack{1 \leqslant j \leqslant k: \\ A_{t_j} \supset S_m}} u_j.$$

An immediate consequence of Proposition 2.3 is the following corollary stating the second-order properties of a trawl process.

COROLLARY 2.4 *Consider a trawl process with finite second moment. Then, for all $t \in \mathbb{R}$, $h \geqslant 0$, we have $E(\Lambda_t) = \lambda^{\mathrm{Leb}}(A) E(L')$, $\mathrm{Var}(\Lambda_t) = \lambda^{\mathrm{Leb}}(A)\,\mathrm{Var}(L')$, and*

$$\mathrm{Cor}(\Lambda_t, \Lambda_{t+h}) = \frac{\lambda^{\mathrm{Leb}}(A \cap A_h)}{\lambda^{\mathrm{Leb}}(A)}.$$

### 2.2.2 Marginal distribution

In the context of our latent trawl model, we are exclusively interested in the case of a marginal gamma law. Specifically, we fix a set $A$ in $\mathcal{B}_b(S)$ and let the Lévy seed have a normalized gamma distribution, that is,

$$L' \sim \mathrm{Gamma}\left(\frac{\alpha}{\lambda^{\mathrm{Leb}}(A)}, \beta\right). \tag{2.10}$$

Thus, the trawl process defined by $\Lambda_t = L(A_t)$ has a $\mathrm{Gamma}(\alpha, \beta)$ distribution.

Combining trawl process $\Lambda$ with the hierarchical model presented in Section 2.1, we obtain a stochastic process $(X_j)$ with finite-dimensional densities given by

$$f(x_1, \dots, x_k) = \int_{\mathbb{R}_+^k} \left(\prod_{j \in I_0} (1 - \mathrm{e}^{-\kappa\lambda_j})\right)\left(\prod_{j \in I_>} \lambda_j \mathrm{e}^{-\lambda_j(\kappa + x_j)}\right) \mathrm{d}F(\lambda_1, \dots, \lambda_k), \tag{2.11}$$

where $I_0 = \{j \in \{1, \dots, k\}\colon x_j = 0\}$ and $I_> = \{j \in \{1, \dots, k\}\colon x_j > 0\}$. These densities depend on the joint distribution $F$ of $(\Lambda_1, \dots, \Lambda_k)$, which is fully specified by the trawl set $A$ and the Lévy seed $L'$, as shown in Proposition 2.3.

### *2.2.3 Trawl set*

To complete the definition of trawl process $\Lambda$, it now remains to specify the trawl set $A$. For our model, we use the so-called exponential trawl set; this is the trawl obtained by setting

$$A = \{(\xi, s)\colon s \leqslant 0, 0 \leqslant \xi \leqslant d_{\exp}(s)\} \subset [0,1] \times (-\infty, 0], \qquad \text{for } d_{\exp}(s) = \exp(\rho s), \text{ for some } \rho > 0.$$

The resulting process is called the exponential trawl process, via which we obtain a hierarchical model with parameters $(\rho, \alpha, \beta, \kappa)$.

The autocovariance function of the trawl process is given by $\varphi(h) = \lambda^{\text{Leb}}(A \cap A_h)\operatorname{Var}(L')$ and, used with the exponential trawl, it gives $\lambda^{\text{Leb}}(A) = \rho^{-1}$, $\lambda^{\text{Leb}}(A \cap A_h) = \mathrm{e}^{-\rho h}\rho^{-1}$. Combining it with (2.10) gives $\operatorname{Var}(L') = (\alpha\rho)/\beta^2$, resulting in the autocovariance function

$$\varphi(h) = \mathrm{e}^{-\rho h}\frac{\alpha}{\beta^2} = \mathrm{e}^{-\rho h}\operatorname{Var}(\Lambda_t).$$

Thus, the autocorrelation function of the exponential trawl process has the same decay rate as the trawl function $d_{\exp}$, which is a particular property of the exponential trawl.

We can also consider a general exponential trawl set constructed from linear combinations of basic exponential trawls. To do so, we define the general exponential trawl set of order $p$ to be bounded above by the function $d_p(x) = \sum_{i=1}^{p} w_i \mathrm{e}^{\rho_i x}$, with $\sum_i w_i = 1$. The latter restriction is necessary to make the parameters identifiable, as any scaling factor in the weights $w_i$ will scale the area of the trawl and thus be canceled by the normalization in (2.10). When using the general exponential trawl set, the resulting trawl process has an autocorrelation function $r(h)$ given by the corresponding linear combination of $\mathrm{e}^{-\rho_i h}$. This construction can be seen as a special case of a superposition-type trawl where the decay parameter $\rho$ is randomized, as in Barndorff-Nielsen *et al* (2014, Section 4) (see Barndorff-Nielsen (2001) for a similar approach applied to Ornstein–Uhlenbeck-type processes). Other relevant choices for trawl sets beyond the exponential setting are discussed in Barndorff-Nielsen *et al* (2014), and we also remark that the trawl need not be restricted to an $\mathbb{R}^2$ setting but could be considered in higher dimensions if necessary.

To summarize, we have constructed a trawl process $\Lambda$ with a marginal gamma distribution and an exponentially decaying autocorrelation function. Using such a discretized trawl process as the latent process in the hierarchical structure results in a discrete-time process $(X_j)$, where the exceedances $\{X_j > 0\}$ have a GPD. Further, the model allows for dependence between observations $\{X_j\}$, which is derived directly from the dependence in the latent trawl process.

## 2.3 Autocovariance structure

We now consider the mean and the autocovariance structure of the exceedance process $(X_j)$ from the latent model, which we summarize in the following proposition.

PROPOSITION 2.5 *The mean of the exceedance process $(X_j)$ is, for all $j \in \mathbb{N} \cup \{0\}$, given by*

$$E[X] := E[X_j] = \frac{(1 + \kappa/\beta)^{-\alpha}(\beta + \kappa)}{\alpha - 1}, \quad \alpha > 1.$$

*Since $X$ is a stationary process, it has the autocovariance function $\varphi(h) = E[X_0 X_h] - E^2[X]$, where for $h \in \mathbb{N}$,*

$$E[X_0 X_h] = \int_{\kappa}^{\infty} \int_{\kappa}^{\infty} \left(1 + \frac{u_0}{\beta}\right)^{b_{0\setminus h}} \left(1 + \frac{u_0 + u_h}{\beta}\right)^{b_{0,h}} \left(1 + \frac{u_h}{\beta}\right)^{b_{h\setminus 0}} \mathrm{d}u_0 \, \mathrm{d}u_h, \tag{2.12}$$

*where $b_i = -\alpha \lambda^{\mathrm{Leb}}(B_i)/\lambda^{\mathrm{Leb}}(A)$ for $i \in \{(0\setminus h), (0, h), (h\setminus 0)\}$ with $B_{0\setminus h} = A_0 \setminus A_h$, $B_{0,h} = A_0 \cap A_h$, $B_{h\setminus 0} = A_h \setminus A_0$. Note that $b_{0\setminus h} = b_{h\setminus 0}$.*

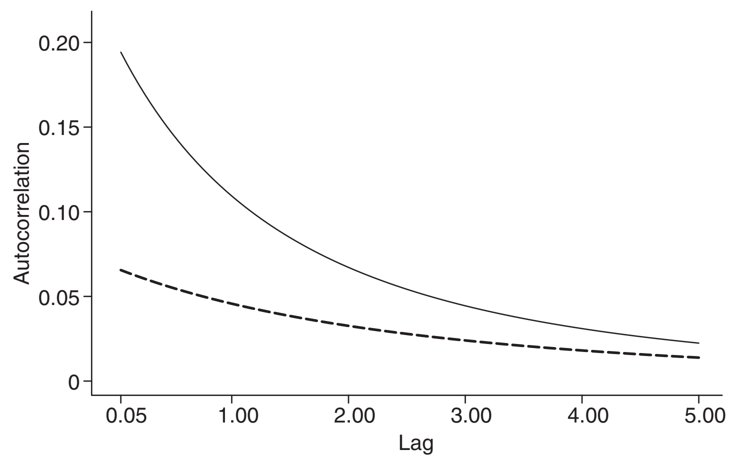
The integral in (2.12) can be computed numerically to obtain the autocovariance of $X$ for given parameters $(\alpha, \beta, \rho, \kappa)$, where the parameters $b_i$ are functions of $\rho$ and $h$.

The trawl process separates the parameters controlling the marginal and dependence properties of the model. However, this is not the case when considering the full hierarchical model, as the parameters $\alpha$, $\beta$ and $\kappa$ in the marginal distribution also influence the autocovariance structure of the process. This is illustrated in Figure 1, which shows the two different autocorrelation functions obtained by varying the parameters $\alpha$ and $\beta$ (solely for illustrative purposes, the plot is provided in a continuous-time setting). This conflation of marginal and dependence parameters motivates the model in the following subsection.

## 2.4 Marginal transformation model

This subsection considers a modification of the latent trawl model that has the effect of separating the marginal and dependence properties. This modification allows the model to have GPDs with negative shape parameter. The original restriction regarding the shape parameter, specifying positive values only, was highlighted as a potential problem in the conclusion of Bortot and Gaetan (2014), where a similar modification was suggested but not explored further.

**FIGURE 1** Autocorrelation functions of latent trawl models.




Solid line corresponds to a model with $\alpha = 4$, $\beta = 4$. Dashed line corresponds to $\alpha = 9$, $\beta = 1$. Both models set $\rho = 0.2$ and $\kappa$ such that the probability of an exceedance equals 0.05.

The model resulting from our modification is easier to interpret, as the role of each parameter is uniquely defined in terms of whether it controls

- the marginal distribution,
- the probability of exceedance, or
- the dependence properties.

This also contributes to the identifiability of the parameters. In particular, we found that the estimation procedure appears to be more efficient with the modified model.

The modified model is derived from the original model in two steps. First, we fix the parameters $\alpha, \beta$ of the latent gamma distribution such that only the parameters associated with the trawl set will influence the trawl process and thus the dependence of the exceedances. In the following, we work with $\alpha = \beta = 1$, such that the marginal law of the exceedances is given by $\text{GPD}(1, 1 + \kappa)$.

Second, we add an extra layer to the modified model, using a standard probability integral transform to give the marginals a $\text{GPD}(\xi, \sigma)$, specifically

$$Z_j = F^{-1}_{\text{GPD}(\xi,\sigma)}(F_{\text{GPD}(1,1+\kappa)}(X_j)) := g(X_j), \quad \xi \in \mathbb{R},\ \sigma > 0.$$

The above construction implies that

$$g(x) = \frac{\sigma}{\xi}\left\{\left(1 + \frac{x}{1+\kappa}\right)^{\xi} - 1\right\} \quad \text{and} \quad g^{-1}(z) = (1+\kappa)\{(1 + z\xi/\sigma)^{1/\xi} - 1\}.$$

This modified version will be called the marginal transformation (MT) model. It has parameters $(\rho, \kappa, \xi, \sigma)$, where $\xi = 1/\alpha,\ \sigma = \beta/\alpha$, and conditional density is given by

$$f(z_j \mid \lambda_j) = \begin{cases} 1 - \exp\{-\kappa\lambda_j\}, & z_j = 0, \\ J(z_j)\lambda_j \exp\{-\lambda_j(\kappa + g^{-1}(z_j))\}, & z_j > 0, \end{cases}$$

where

$$J(z_j) = \frac{f_{\mathrm{GPD}(\xi,\sigma)}(z_j)}{f_{\mathrm{GPD}(1,1+\kappa)}(g^{-1}(z_j))}$$

with respect to the measure $\mu(\mathrm{d}z_j) = \delta_0(\mathrm{d}z_j) + \mathrm{d}z_j$. This follows since we note that the transformation $g$ maps the event $\{X = 0\}$ to $\{Z = 0\}$, leaving the atom at zero unchanged. By contrast, the transformation of the continuous part on $\{X > 0\}$ introduces a standard Jacobian term.

Now let $I_0 = \{j \in \{1, \dots, k\} \colon z_j = 0\}$ and $I_> = \{j \in \{1, \dots, k\} \colon z_j > 0\}$. Then, the finite-dimensional densities of the MT model can be represented by

$$\begin{aligned} f(z_1, \dots, z_k) = \int_{\mathbb{R}_+^k} & \bigg( \prod_{j \in I_0} (1 - \exp\{-\kappa\lambda_j\}) \bigg) \\ & \times \bigg( \prod_{j \in I_>} J(z_j)\lambda_j \exp\{-\lambda_j(\kappa + g^{-1}(z_j))\} \bigg) \mathrm{d}F(\lambda_1, \dots, \lambda_k). \end{aligned}$$

When using the MT model in an application, the empirical observations of the exceedances will be described by the $(Z_j)$ and not by the $(X_j)$ as in the earlier model specification.

## 3 MODEL FITTING AND EVALUATION

### 3.1 Pairwise likelihood

We now consider parameter estimation for the latent trawl model described above. The parameter vector of interest is denoted by $\theta = (\rho, \kappa, \xi, \sigma)^{\mathrm{T}} \in \Theta$, where $\Theta \subset \mathbb{R}^4$ gives the parameter space.

As in Section 2.1, we transform a set of observations $\{Y_j\}$ from a stationary time series to obtain exceedances $X_j := \max(Y_j - u, 0),\ j = 1, \dots, k$. We assume $l$ positive observations $X_{p_1}, \dots, X_{p_l}$ as well as $m$ observations $X_{q_1}, \dots, X_{q_m}$ taking the value zero, which we will call exceedances and nonexceedances, respectively. Further, we note that $l + m = k$.

The likelihood of the observations $\{X_j\}$ under the original latent trawl model now follows from (2.11) and can be written as

$$f(x_1,\dots,x_k) = \int_{\mathbb{R}_+^k} \prod_{r=1}^{l} \lambda_{p_r} \exp\{-(x_{p_r}+\kappa)\lambda_{p_r}\} \times \prod_{s=1}^{m} (1-\exp\{-\kappa\lambda_{q_s}\})\,\mathrm{d}F(\lambda_1,\dots,\lambda_k),$$

where $F$ is the joint density function of the trawl process observations $\Lambda_1,\dots,\Lambda_k$. The above integrand can be expanded to obtain a sum with $2^m$ terms. To do so, we first define $\mathcal{S}_t$ as a collection of subsets of $\{q_1,\dots,q_m\}$ of size $t$, letting $u_r = x_{p_r} + \kappa$. This gives

$$\sum_{t=0}^{m}(-1)^t \sum_{\pi_t\in\mathcal{S}_t} \prod_{r=1}^{l} \lambda_{p_r} \exp\{-u_r\lambda_{p_r}\} \prod_{s_j\in\pi_t} \exp\{-\kappa\lambda_{s_j}\},$$

with the convention $\prod_{s_j\in\pi_0}(\cdots) = 1$. This can be rewritten as

$$\sum_{t=0}^{m} \sum_{\pi_t\in\mathcal{S}_t} (-1)^{t+l} \frac{\partial}{\partial u_1}\cdots\frac{\partial}{\partial u_l} \exp\left\{-\sum_{r=1}^{l} u_r\lambda_{p_r} - \sum_{s_j\in\pi_t} \kappa\lambda_{s_j}\right\}.$$

Using this expression and exchanging integrals and partial derivatives, we see that the full likelihood reduces to

$$\begin{aligned}
&\sum_{t=0}^{m} \sum_{\pi_t\in\mathcal{S}_t} (-1)^{t+l} \frac{\partial}{\partial u_1}\cdots\frac{\partial}{\partial u_l} \\
&\qquad \times \int_{\mathbb{R}_+^k} \exp\left\{-\sum_{r=1}^{l} u_r\lambda_{p_r} - \sum_{s_j\in\pi_t} \kappa\lambda_{s_j}\right\} \mathrm{d}F(\lambda_1,\dots,\lambda_k) \\
&\qquad\qquad = \sum_{t=0}^{m} \sum_{\pi_t\in\mathcal{S}_t} (-1)^{t+l} \frac{\partial}{\partial u_1}\cdots\frac{\partial}{\partial u_l} \\
&\qquad\qquad\qquad \times E\left[\exp\left\{-\sum_{r=1}^{l} u_r\Lambda_{p_r} - \sum_{s_j\in\pi_t} \kappa\Lambda_{s_j}\right\}\right],
\end{aligned}$$

where the expectation is with respect to the corresponding variables $\{\Lambda_{p_r}\}$ and $\{\Lambda_{s_j}\}$ determined by $\pi_t$.

The expected values in these terms are joint Laplace transforms and may be derived from the joint characteristic functions given in Proposition 2.3. They involve the parameters of the Lévy seed $L$ and the trawl intersection areas; hence, the complete likelihood reduces to a sum of the partial derivatives of the Laplace transforms.

The likelihood (as given above) is not easy to compute in practice for two reasons. First, it requires multiple numerical partial derivatives to be performed. Second, the number of nonexceedances $m$ is usually close to the number of observations $k$. This is because the latent model is defined as having a GPD, and to justify this assumption we need to consider a sufficiently high threshold for exceedances, often given by a large percentile of the observations. Now that the first two sums in the likelihood above have $2^m$ terms in total, the likelihood becomes computationally intractable for any reasonable sample size $k$.

Because of the computational issues associated with the sample size, we consider using a pairwise likelihood approach for the model fitting, that is, a particular kind of composite likelihood (Cox and Reid 2004; Varin 2008; Varin *et al* 2011). As stated in Varin (2008), composite likelihood estimators are well suited to cases when the data can be seen to consist of roughly independent blocks, that is, the autocorrelation function decays sufficiently fast. Thus, the pairwise likelihood should perform reasonably well for the latent trawl model with an exponential trawl set.

Given observations $x_1, \dots, x_k$, the pairwise likelihood $f_{\mathrm{PL}}$ for the parameter vector $\theta = (\rho, \kappa, \xi, \sigma)^{\mathrm{T}} \in \Theta \subset \mathbb{R}^4$ takes the form

$$f_{\mathrm{PL}}^{\triangle}(\theta \mid x_1, \dots, x_k) = \prod_{i=1}^{k-1} \prod_{j=i+1}^{\min(i+\triangle,k)} f(x_i, x_j),$$

where $f(\cdot)$ is the original bivariate density function and $\triangle$ denotes the maximum separation between observations. For the latent trawl model, each pairwise likelihood term $f(x_i, x_j)$ involves four terms at most from the sum above, and so these terms can be explicitly evaluated with regard to the parameters of the trawl process. There are four different cases (since $x_i$ and $x_j$ can each be an exceedance or a nonexceedance), and the explicit forms of $f(x_i, x_j)$ are given in online appendix A.1. The maximum pairwise likelihood estimator is denoted by

$$\hat{\theta} = \arg\max_{\theta} f_{\mathrm{PL}}^{\triangle}(\theta \mid x_1, \dots, x_k).$$

According to Cox and Reid (2004), the pairwise likelihood estimator is unbiased and asymptotically normal under the usual regularity conditions. When looking at the asymptotic theory in this context, we assume that we have fixed the threshold when computing the relevant exceedances. We do not allow for a double asymptotic setting, in which the threshold increases at the same time as the number of observations. A more detailed investigation of such a double asymptotic is beyond the scope of this paper.

### 3.2 Conditional tail dependence coefficient

We now consider the extremal dependence structure of our model. A common measure of dependence at high levels is the extremal index (Leadbetter *et al* 1983), which can be characterized as

$$\theta = \lim_{n\to\infty} P(X_i \leqslant u_n, 2 \leqslant i \leqslant l_n \mid X_1 > u_n),$$

where $u_n$ is an increasing sequence of thresholds and $l_n = o(n)$ (as per Ancona-Navarrete and Tawn (2000) following O'Brien (1987)).

The extremal index $\theta$ essentially describes the dependence across blocks of observations whose length tends to infinity; it can also be defined as the reciprocal mean cluster length, where a cluster represents a collection of exceedances in a block. Thus, to estimate the extremal index one has to consider very high-dimensional joint distributions, which makes the estimation analytically intractable in many cases. This is certainly an issue with our latent trawl model. For the applications detailed in the following subsections, we will consider simulation-based estimates of the extremal index instead.

There are other measures of extremal dependence that work on shorter ranges of observations than the extremal index. Coles *et al* (1999), for instance, quantify the dependence between the extreme values of two random variables $X_1$, $X_2$ in the upper tail dependence coefficient, given by[1]

$$\chi = \lim_{u\to 1} P(F(X_2) > u \mid F(X_1) > u),$$

where $F$ is the common marginal distribution of $X_1$, $X_2$. In other words, $\chi$ gives the limiting probability of $X_2$ exceeding the threshold $u$ given an exceedance $X_1$, with both variables on a uniform scale. When $X_1$, $X_2$ come from a stationary time series, $\chi$ can be seen as the probability of observing consecutive exceedances given a single exceedance. To get a broader characterization of the extremal dependence, one may also consider the complete function

$$\chi(u_1, u_2) = P(F(X_2) > u_2 \mid F(X_1) > u_1),$$

defined on $[0, 1]^2$, where $\chi(u_1, 0) = 1$ and $\chi(0, u_2) = 1 - u_2$.

We would like to use the function $\chi(u_1, u_2)$ and the limiting measure $\chi$ to evaluate the dependence between two observations $X_1$, $X_2$ from the latent trawl model as a function of the lag $t_2 - t_1$. However, the above definition is based on the assumption that $X_1$, $X_2$ are continuously valued random variables with a distribution function $F$

[1] This quantity is sometimes denoted as $\lambda_U$ and should not be confused with the "coefficient of tail dependence" defined in Ledford and Tawn (1996).

defined across their entire range. This does not fit with the exceedance framework, in which the limiting distribution can only be assumed to hold above a threshold. Certain extreme value models (see Ledford and Tawn (1996), for example) do not consider exceedances separately but specify the same probability density function $f$ for the full range of observations; any observations below the threshold $u$ are then treated as censored so that they have the probability $\int_0^u f(s)\,\mathrm{d}s$. Our model differs in that it models the occurrence of exceedances explicitly, resulting in a distribution for exceedances $X$ only with an atom at zero. Hence the standard definition of $\chi$ cannot be applied directly. While we may still construct an analog to the tail dependence coefficient by conditioning on both exceedances being positive, care must be taken to ensure that the resulting measure is uniform on the $u_2$ margin, as shown in the following definition.

DEFINITION 3.1 Consider the latent model defined in Section 2.1 and set

$$\begin{aligned} F_{2e}(x) &= P(X_0 \leqslant x \mid X_0 > 0, X_h > 0) \\ &= P(X_h \leqslant x \mid X_0 > 0, X_h > 0), \quad \text{for } h \in \mathbb{N}. \end{aligned} \tag{3.1}$$

The conditional tail dependence function $\varphi$ is defined as

$$\varphi(h, u_1, u_2) := P(F_{2e}(X_h) > u_2 \mid F_{2e}(X_0) > u_1, X_0 > 0, X_h > 0), \quad \text{for } 0 \leqslant u_1, u_2 \leqslant 1,$$

and the conditional tail dependence coefficient is defined as

$$\varphi(h) := \lim_{u \uparrow 1} \varphi(h, u, u).$$

We show in Lemma A.2 (see online appendix A.2) that the identity (3.1) holds.

The conditional tail dependence function can be calculated explicitly in terms of the parameters of the latent trawl model. Specifically, we have the following three key results (all proved in the online appendix).

PROPOSITION 3.2 *Let $h \in \mathbb{N}$, and set*

$$B_{0\setminus h} = A_0 \setminus A_h, \qquad B_{0,h} = A_0 \cap A_h, \qquad B_{h\setminus 0} = A_h \setminus A_0,$$

*and $b_i = -\alpha\lambda^{\mathrm{Leb}}(B_i)/\lambda^{\mathrm{Leb}}(A)$ for $i \in \{(0 \setminus h), (0, h), (h \setminus 0)\}$. The conditional tail dependence function for $X_0$ and $X_h$ in the latent trawl model is given by*

$$\varphi(h, u_1, u_2) = \left(1 + \frac{F_{2e}^{-1}(u_2)}{\beta + 2\kappa + F_{2e}^{-1}(u_1)}\right)^{b_{0,h}} \left(1 + \frac{F_{2e}^{-1}(u_2)}{\beta + \kappa}\right)^{b_{h\setminus 0}}, \quad \textit{for } 0 \leqslant u_1, u_2 \leqslant 1.$$

Proposition 3.3 *The conditional tail dependence function satisfies the same marginal scaling as the original tail dependence index* $\chi$*, namely,* $\varphi(h, u_1, 0) = 1$*,* $\varphi(h, 0, u_2) = 1 - u_2$ *for any* $0 \leqslant u_1, u_2 \leqslant 1$.

Theorem 3.4 *For the original latent trawl model, we have that* $\phi(h) = 0$ *for any* $h \in \mathbb{N}$*. Thus, according to this measure, the model is asymptotically tail independent.*

The speed at which the conditional tail dependence function decays to zero increases with the value of $b_{h\backslash 0}$; in other words, the larger the intersection of the trawl sets given by $X_0$, $X_h$, the slower the model decays to independence as the threshold increases.

It has been pointed out by Coles *et al* (1999) that the class of asymptotically independent distributions is of fundamental importance in multivariate extreme value theory (see also Bortot and Tawn 1998; Bruun and Tawn 1998; Ledford and Tawn 1996, 1997).

## 4 EMPIRICAL EXAMPLES

In our empirical studies, we need to decide on a suitable threshold to use when considering the exceedances. In the extreme value literature, various approaches to threshold selection have been discussed (see Wong and Li (2010) for a recent discussion on this subject). A popular graphical approach is to use a mean excess plot, since the mean excess function for the GPD is linear in the threshold. In the following, we consider precipitation data and ozone levels from London. Setting the threshold to the 95th and 97th percentile, respectively, we find the resulting exceedances to be well approximated by a GPD.

### 4.1 Heathrow data

In this subsection, we use the latent trawl model to analyze a data set consisting of daily rainfall amounts accumulated at Heathrow (in the United Kingdom) in the years between 1949 and 2012, provided by the UK Meteorological Office (2012). We set the threshold $u$ at the 95th percentile of the original data (8.9 mm), resulting in the time series of exceedance values shown in Figure 2.

We fit the latent trawl model described in Section 2.2 as well as the latent Markov chain model of Bortot and Gaetan (2014), using both the G-LP and the WP for the Markov chain. The parameter estimation was effected using pairwise likelihood (as described in Section 3.1) with the separation parameter $\triangle = 4$ (suggested by our simulation experiments); the resulting estimates are shown in Table 1. We see that the marginal parameters are similar across all the models; this seems reasonable given that all the models have a marginal GPD($\alpha, \beta + \kappa$) and in view of the fact that $\kappa$ controls the marginal exceedance probability.

**FIGURE 2** Heathrow rainfall exceedances.

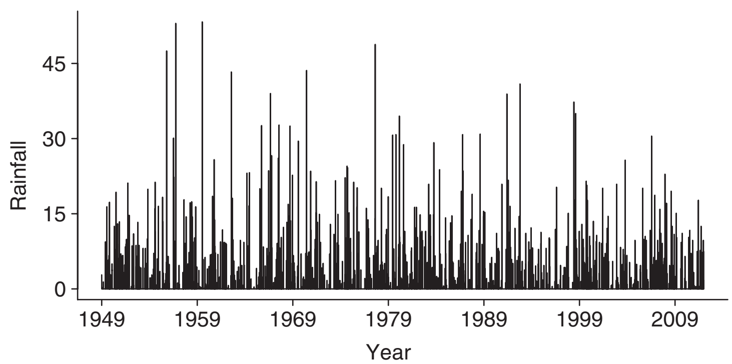



**TABLE 1** Parameter estimates for the various exceedance models.

| | $\alpha$ | $\beta$ | $\rho$ | $\kappa$ |
|---|---|---|---|---|
| Latent trawl | 6.33 | 20.12 | 0.27 | 12.18 |
| G-LP | 6.43 | 20.64 | 0.70 | 12.25 |
| WP | 6.30 | 19.94 | 0.78 | 12.15 |

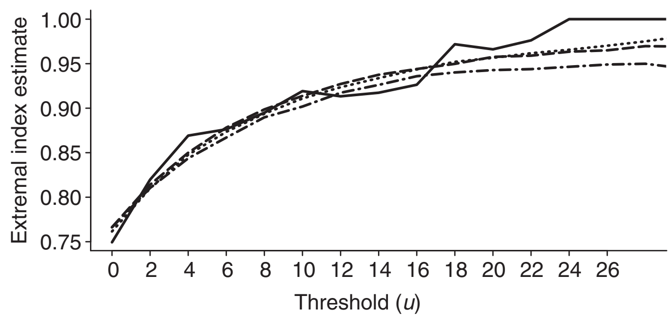
Parameters for Heathrow rainfall data estimated via the latent trawl model as well as the latent Markov chain model with G-LP and WP chains.

The parameter $\rho$ determines the latent dependence structure of the models. For the latent trawl process it is the decay parameter of the exponential trawl function, whereas for the latent Markov chains it enters in the autocorrelation function $\phi(h) = \rho^h = \exp(h \log(\rho))$. Based on this relation, we see from the fitted values of $\rho$ that all three latent processes have similar autocorrelation functions, in which the G-LP and WP have the fastest and slowest rates of decay, respectively.

Figure 3 shows estimates of the extremal index for the latent trawl and latent Markov chain models. These estimates are based on simulating time series of length 1 000 000 via the fitted models. We then estimate $\theta$ as the inverse cluster length using the R package `evd`, which defines a cluster as ending when three consecutive exceedances fall below the threshold.

These estimates indicate that the latent trawl and latent Markov chain models all manage to capture the main dependence structure in the extremes. The only visible difference occurs at higher thresholds, where the G-LP model appears to underestimate $\theta$, that is, it overestimates the dependence. This discrepancy (occurring at high levels only) could be explained by the fact that the G-LP model is asymptotically dependent,

**FIGURE 3** Estimated extremal index for empirical ozone data (solid line) as well as the latent trawl (dashed), G-LP (dash-dotted) and WP (dotted) models.




as shown by Bortot and Gaetan (2014), whereas the empirical estimates of $\theta$ indicate that the rainfall data is asymptotically independent.

## 4.2 Pollution data

We now consider a second application that illustrates the marginal transformation model introduced in Section 2.4. We use this model to analyze a set of ozone levels measured in Bloomsbury, London. Specifically, our data gives the daily maximum of the eight-hour running mean, measured in units of micrograms per cubic meter ($\mu g/m^3$). These measurements were obtained from the Environmental Research Group, King's College London (2015).

Although the original data displays evident seasonality, this effect diminishes significantly as the threshold increases. Thus, we do not adjust for seasonality in the extreme values for this application, though refinements of this approach could be considered in future works. For the threshold $u$, we choose the 97th percentile of the original data (81 $\mu g/m^3$), resulting in the exceedances shown in Figure 4.

Fitting the GPD to the data directly indicates a negative $\xi$ value (ie, a finite upper bound) that cannot be captured by the standard model. Thus, we use the transformed model described in Section 2.4. This corresponds with taking the basic hierarchical model (based on either a latent trawl or a latent Markov chain as in Bortot and Gaetan (2014)) and fixing $\alpha, \beta = 1$. The next step entails transforming the marginals to GPD$(\xi, \sigma)$.

The models were fitted using pairwise likelihood with $\triangle = 4$. Table 2 shows the resulting estimates for the marginal transformed versions of the latent trawl and the latent Markov chain models.

**FIGURE 4** Ozone level exceedances.

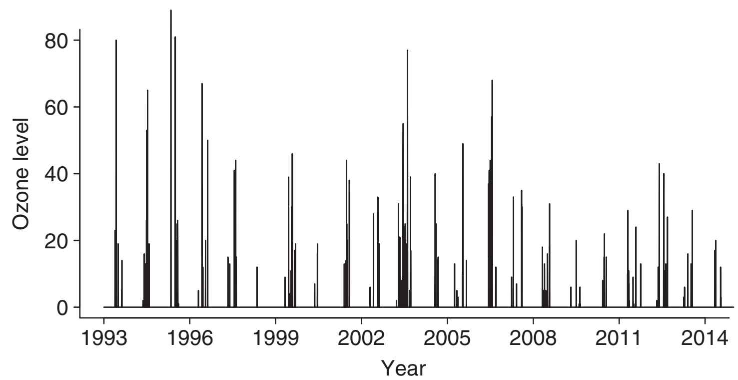



**TABLE 2** Parameter estimates for the transformed exceedance models.

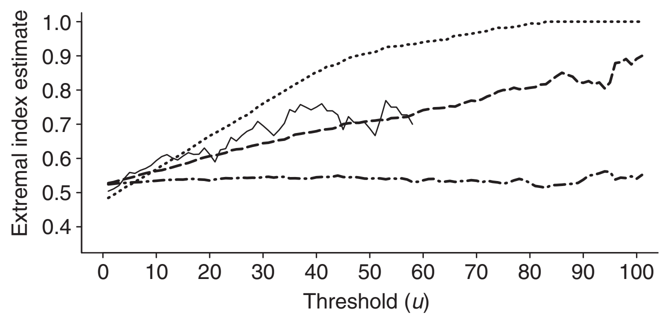
| | $\xi$ | $\sigma$ | $\rho$ | $\kappa$ |
|---|---|---|---|---|
| Latent trawl | −0.11 | 20.73 | 0.17 | 32.69 |
| G-LP | −0.04 | 21.09 | 0.56 | 32.19 |
| WP | −0.11 | 20.74 | 0.96 | 32.44 |

Estimated parameters for ozone data, generated using transformed versions of the latent trawl model as well as the latent Markov chain model with G-LP and WP chains.

Figure 5 shows estimates of the extremal index based on simulations of length 1 000 000, obtained by the same method used for the rainfall data previously. The estimates of the extremal index indicate that the transformed latent Markov chain models does not accurately capture the extremal dependence structure in this example. In particular, the WP model appears to underestimate the extremal dependence (that is, the resulting $\theta$ values are too high), whereas the opposite is the case for the G-LP model.

These results show that when the marginal parameters $\alpha, \beta$ are fixed in the latent layer, the latent Markov process model has less flexibility in the dependence structure than the latent trawl process model. This indicates that in the original hierarchical structure, the marginal parameters also have a strong influence on the dependence structure, a factor that contributes to the flexibility of the model. The transformed model has the advantage of having parameters that are clearly interpretable as contributing to either the dependence or the marginal distribution. In our experience (from our simulation studies), the parameter estimation procedure for the transformed model also appears to be more reliable.

**FIGURE 5** Estimated extremal index for empirical ozone data (solid line) as well as transformed versions of the latent trawl (dashed), G-LP (dash-dotted) and WP (dotted) models.




## 5 CONCLUSION

In this paper, we investigated a new model for time series of extremes based on trawl processes. We constructed an extreme value model that uses the trawl process framework to obtain a flexible dependence structure. This was achieved by replacing the latent Markov chain in the setup described by Bortot and Gaetan (2014) with a trawl process. In contrast with other hierarchical models, our construction has the advantage of preserving the GPD for the marginals, which is consistent with extreme value theory.

We also considered a modification to the original model structure that extends the parameter space, allowing for negative shape parameters. To evaluate the extremal dependence, we also developed an adapted version of the tail-dependence coefficient that can be evaluated analytically for the trawl process model.

Both the original and modified models were used to analyze two environmental time series, and their performance was compared with the latent Markov chain models of Bortot and Gaetan (2014). For the application of the original model, the results were very similar in terms of capturing the extremal dependence of the data. The advantage of using the latent trawl process was clearer when using the transformed models; in this application, the trawl-based model performed better due to the added flexibility in its latent dependence structure.

There are several aspects of the latent trawl model that could benefit from further investigation. For example, we have used a simple exponential trawl throughout this paper; it would be interesting to look at the results of using different parameterizations for the trawl set. Another possibility would be to consider a model where the threshold

$u$ is allowed to vary. If it were to do so, then letting the trawl set depend on $u$ should result in a wider range of dependence levels across all thresholds.

## DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

## ACKNOWLEDGEMENTS

R. C. Noven gratefully acknowledges financial support from the Grantham Institute for Climate Change, Imperial College London. A. E. D. Veraart acknowledges financial support in the form of a Marie Curie FP7 Integration Grant within the 7th European Union Framework Programme (grant agreement number PCIG11-GA-2012-3217). We thank the UK Meteorological Office, the British Atmospheric Data Centre and the London Air Quality Network for providing the data used.

## REFERENCES

Ancona-Navarrete, M. A., and Tawn, J. A. (2000). A comparison of methods for estimating the extremal index. *Extremes* **3**(1), 5–38 (https://doi.org/10.1023/A:1009993419559).

Barndorff-Nielsen, O. E. (2001). Superposition of Ornstein–Uhlenbeck type processes. *Theory of Probability & Its Applications* **45**(2), 175–194 (https://doi.org/10.1137/S0040585X97978166).

Barndorff-Nielsen, O. E. (2011). Stationary infinitely divisible processes. *Brazilian Journal of Probability and Statistics* **25**(3), 294–322 (https://doi.org/10.1214/11-BJPS140).

Barndorff-Nielsen, O. E., Lunde, A., Shephard, N., and Veraart, A. E. D. (2014). Integer-valued trawl processes: a class of stationary infinitely divisible processes. *Scandinavian Journal of Statistics* **41**(3), 693–724 (https://doi.org/10.1111/sjos.12056).

Barndorff-Nielsen, O. E., Benth, F. E., and Veraart, A. E. D. (2015). Recent advances in ambit stochastics with a view towards tempo-spatial stochastic volatility/intermittency. *Banach Center Publications* **104**, 25–60 (https://doi.org/10.4064/bc104-0-2).

Bortot, P., and Gaetan, C. (2014). A latent process model for temporal extremes. *Scandinavian Journal of Statistics* **41**(3), 606–621 (https://doi.org/10.1111/sjos.12051).

Bortot, P., and Tawn, J. A. (1998). Models for the extremes of Markov chains. *Biometrika* **85**(4), 851–867 (https://doi.org/10.1093/biomet/85.4.851).

Bruun, J. T., and Tawn, J. A. (1998). Comparison of approaches for estimating the probability of coastal flooding. *Journal of the Royal Statistical Society* C **47**(3), 405–423 (https://doi.org/10.1111/1467-9876.00118).

Coles, S. G., and Tawn, J. A. (1991). Modelling extreme multivariate events. *Journal of the Royal Statistical Society* B **53**(2), 377–392.

Coles, S. G., Heffernan, J., and Tawn, J. A. (1999). Dependence measures for extreme value analyses. *Extremes* **2**(4), 339–365 (https://doi.org/10.1023/A:1009963131610).

Cox, D. R., and Reid, N. (2004). A note on pseudolikelihood constructed from marginal densities. *Biometrika* **91**(3), 729–737 (https://doi.org/10.1093/biomet/91.3.729).

Davison, A. C., and Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society* B **52**(3), 393–442.

Davison, A. C., Padoan, S. A., and Ribatet, M. (2012). Statistical modeling of spatial extremes. *Statistical Science* **27**(2), 161–186 (https://doi.org/10.1214/11-STS376).

de Haan, L., and de Ronde, J. (1998). Sea and wind: multivariate extremes at work. *Extremes* **1**(1), 7–45 (https://doi.org/10.1023/A:1009909800311).

Embrechts, P., Mikosch, T., and Klüppelberg, C. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer (https://doi.org/10.1007/978-3-642-33483-2).

Environmental Research Group, King's College London (2015). London Air Quality Network. Database, ERG, King's College London. URL: www.londonair.org.uk.

Fuchs, F., and Stelzer, R. (2013). Mixing conditions for multivariate infinitely divisible processes with an application to mixed moving averages and the supOU stochastic volatility model. *ESAIM: Probability and Statistics* **17**, 455–471 (https://doi.org/10.1051/ps/2011158).

Heffernan, J. E., and Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society* B **66**(3), 497–546 (https://doi.org/10.1111/j.1467-9868.2004.02050.x).

Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer (https://doi.org/10.1007/978-1-4612-5449-2).

Ledford, A. W., and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83**(1), 169–187 (https://doi.org/10.1093/biomet/83.1.169).

Ledford, A. W., and Tawn, J. A. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society* B **59**(2), 475–499 (https://doi.org/10.1111/1467-9868.00080).

Ledford, A. W., and Tawn, J. A. (2003). Diagnostics for dependence within time series extremes. *Journal of the Royal Statistical Society* B **65**(2), 521–543 (https://doi.org/10.1111/1467-9868.00400).

O'Brien, G. L. (1987). Extreme values for stationary and Markov sequences. *Annals of Probability* **15**(1), 281–291 (https://doi.org/10.1214/aop/1176992270).

Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics* **3**(1), 119–131 (https://doi.org/10.1214/aos/1176343003).

Rajput, B. S., and Rosinski, J. (1989). Spectral representations of infinitely divisible processes. *Probability Theory and Related Fields* **82**(3), 451–487 (https://doi.org/10.1007/BF00339998).

Reiss, R.-D., and Thomas, M. (2007). *Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology and Other Fields*, 3rd edn. Birkhäuser, Basel (https://doi.org/10.1007/978-3-7643-7399-3).

Sato, K. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

Shephard, N., and Yang, J. J. (2016). Likelihood inference for exponential-trawl processes. In *The Fascination of Probability, Statistics and Their Applications*, Podolskij, M., Stelzer, R., Thorbjørnsen, S., and Veraart, A. E. D. (eds), pp. 251–281. Springer (https://doi.org/10.1007/978-3-319-25826-3_12).

Shephard, N., and Yang, J. J. (2017). Continuous time analysis of fleeting discrete price moves. *Journal of the American Statistical Association* **112**(519), 1090–1106 (https://doi.org/10.1080/01621459.2016.1192544).

UK Meteorological Office (2012). Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853–current). Data Archive System, NCAS British Atmospheric Data Centre. URL: https://bit.ly/2AbJrMY.

Varin, C. (2008). On composite marginal likelihoods. *AStA Advances in Statistical Analysis* **92**(1), 1–28 (https://doi.org/10.1007/s10182-008-0060-7).

Varin, C., Reid, N., and Firth, D. (2011). An overview of composite likelihood methods. *Statistica Sinica* **21**, 5–42.

Veraart, A. E. D. (2018). Modelling, simulation and inference for multivariate time series of counts using trawl processes. Working Paper, Social Science Research Network (https://doi.org/10.2139/ssrn.3100076).

Wong, T. S. T., and Li, W. K. (2010). A threshold approach for peaks-over-threshold modeling using maximum product of spacings. *Statistica Sinica* **20**(3), 1257–1272.