AIMS Mathematics, 11(3): 5322–5339.
DOI: 10.3934/math.2026219
Received: 12 January 2026
Revised: 14 February 2026
Accepted: 26 February 2026
Published: 02 March 2026

https://www.aimspress.com/journal/Math

Rsearch article

Structure learning for multivariate extremes: A comparative study of regional UK rainfall

Jeongjin Lee[1] **and Yongku Kim**[2,3,*]

[1] School of Mathematical Sciences, Lancaster University, Fylde College, Lancaster, LA1 4YF, United Kingdom

[2] Department of Statistics, Kyungpook National University, Daegu, Korea

[3] KNU G-LAMP Research Center, Institute of Basic Sciences, Kyungpook National University, Daegu, Korea

* **Correspondence:** Email: kim.1252@knu.ac.kr.

Abstract: Characterizing extremal structural relationships between sets of variables is central to the development of parsimonious models in extreme value analysis, particularly as statistical modeling in high dimensions remains challenging. In this study, we considered recently proposed statistical methods for learning the dependence structure of multivariate variables, with a focus on their ability to capture relationships at extreme levels. We considered complementary approaches that differed in their underlying modeling assumptions. One approach was less model-based and relied on the notion of partial tail correlation to assess extremal dependence between pairs of variables given the others. The other methods were rooted in graphical modeling frameworks, which provided a flexible means of representing complex dependence patterns and facilitated the investigation of higher-order extremal dependencies. We applied the methods to extreme rainfall data from the Lancashire region of the United Kingdom. The resulting dependence structures revealed some spatial heterogeneity, with distinct clustering behavior observed between northern and southern subregions. In particular, evidence of stronger higher-order dependence was concentrated in the southeastern area. These findings suggested that the effectiveness of flood defense and mitigation strategies may vary across subregions, highlighting the importance of accounting for extremal dependence structure in regional risk assessment and infrastructure planning.

Keywords: extremal dependence; graphical models; multivariate extremes; extreme rainfall
Mathematics Subject Classification: 62H05

1. Introduction

Characterizing extreme weather events is a critical component in the design and planning of effective defense and mitigation systems. To this end, statistical models grounded in extreme value theory have been widely used to quantify the frequency, magnitude, and dependence of rare events (see, e.g., de Haan and Ferreira [1]). Despite their strong theoretical foundations, extending these methods to high-dimensional settings remains a substantial challenge, due to the complexity of extremal dependence and the scarcity of joint tail observations. In response, there has been growing interest within the extremes community in recent years in identifying and characterizing structural relationships among sets of variables at extreme levels. Such structures provide insight into the underlying dependence mechanisms and offer a pathway toward more tractable modeling framework for high-dimensional extremes.

From a practical perspective, understanding these extremal dependence structures can inform governmental and policy decision-making by identifying regions or components where investments in defense infrastructure may be most effective. Motivated by these considerations, we apply several recently developed statistical methodologies to investigate extremal dependence structures and to assess their practical implications. While the proposed approaches are broadly applicable across a wide range of data types, the study here focuses on extreme rainfall events that have caused substantial damage and economic losses in the Lancashire region of the United Kingdom. Our primary objective is to characterize extremal relationships between groups of rainfall stations and to examine how these relationships vary across the region.

The first methodology considered is relatively less model-based and relies on the notion of *partial tail correlation*, introduced by Kim and Lee [2], which serves as an extreme analogue of partial correlation in non-extreme settings. This measure quantifies the strength of extremal association between pairs of variables conditional on the remaining variables, without imposing specific distributional or constructional assumptions. We apply the associated hypothesis testing procedure proposed in Kim and Lee [2] to identify statistically significant partial tail correlations between pairs of stations, thereby uncovering an underlying extremal dependence structure.

In addition, we consider a class of graphical models for extremes that impose both a density assumption and a graphical structure on variables. We first consider regular vine tree sequences (Bedford and Cooke [3, 4]), which have been widely used in dependence modeling due to their flexibility in constructing high-dimensional models from collections of arbitrary bivariate parametric families. Vine models also allow for parsimony and sparsity through the use of independent copulas for selected components. Building on this framework, the recently developed *X-vine* model of Kiriliouk et al. [5] extends vine-based constructions to settings involving extremal dependence, enabling flexible modeling of exponent measure densities associated with multivariate extreme value distributions.

Another approach, introduced by Engelke and Hitz [6], considers *extremal graphical models* for multivariate Pareto distributions (Rootzén and Tajvidi [7]). This framework formalizes the notion of extremal conditional independence in a density-based setting, allowing for a clearer interpretation of conditional independence in the tail. In particular, under the Hüsler-Reiss model in Hüsler and Reiss [8], Engelke and Hitz [6] show that sparsity in the resulting graphical structure can be induced via the inverse covariance matrix of this model, yielding interpretable and computationally tractable

graphical representations.

Taken together, these methodologies enable a comparative analysis of structural relationships among variables at extreme levels. By applying them to rainfall extremes, we gain insight into the dependence mechanisms governing regional extreme rainfall behavior and assess the extent to which different modeling frameworks reveal consistent or complementary structural features.

The remainder of the paper is organized as follows. In Section 2, we provide background and review the statistical methodologies used to investigate extremal dependence structures. In Section 3, we apply these methods to an extreme rainfall dataset from the Lancashire region of the United Kingdom and compare their implications. We conclude in Section 4.

2. Statistical methods for exploring extremal structures

The modeling framework underlying partial tail correlation is based on *multivariate regular variation*, a cornerstone of extreme value theory whose characterization is closely linked to the asymptotic dependence structure of multivariate extreme value distributions (de Haan and Ferreira [1], Chapter 6). A comprehensive treatment of regular variation can be found in Resnick [9].

2.1. Partial tail correlation

Let $\mathbf{X} = (X_1, \ldots, X_p)^\top$ be a random vector taking values in $\mathbb{R}_+^p$. For a given norm $\|\cdot\|$, the random vector $\mathbf{X} \in RV_+^p(\alpha)$ is *regularly varying* with tail index $\alpha > 0$ if there exist a normalizing function $b(s) \to \infty$ and a finite measure $H_{\mathbf{X}}$ on the positive unit sphere $\Theta_+^{p-1} = \{\mathbf{x} \in \mathbb{R}_+^p : \|\mathbf{x}\| = 1\}$ such that,

$$s\,\mathbb{P}\left\{\left(\frac{\|\mathbf{X}\|}{b(s)}, \frac{\mathbf{X}}{\|\mathbf{X}\|}\right) \in \cdot\right\} \xrightarrow{v} \nu_\alpha \times H_{\mathbf{X}}, \qquad s \to \infty, \tag{2.1}$$

where $\xrightarrow{v}$ denotes vague convergence and $\nu_\alpha(x, \infty] = x^{-\alpha}$; see, e.g., Theorem 6.1 of Resnick [9]. The measure $H_{\mathbf{X}}$, referred to as the *angular measure*, fully characterizes the extremal dependence structure of $\mathbf{X}$, whose modeling or estimation remains challenging in high dimensions.

To circumvent these challenges, we instead consider a summary of extremal dependence given by the *tail pairwise dependence matrix* (TPDM), introduced by Cooley and Thibaud [10]. Intuitively, the TPDM can be viewed as an extremal analogue of the covariance matrix, as it shares several analogous properties (see Cooley and Thibaud [10]). While the covariance matrix summarizes linear dependence around the mean, the TPDM captures the strength of pairwise extremal dependence in the upper tail, thereby providing a concise and useful summary information of tail dependence. The TPDM is defined as

$$\Sigma_{\mathbf{X}} = [\sigma_{ij}]_{i,j\in[p]}, \qquad \sigma_{ij} = \int_{\Theta_+^{p-1}} w_i w_j \, H_{\mathbf{X}}(\boldsymbol{w}), \tag{2.2}$$

where $[p] := \{1, \ldots, p\}$ and $\mathbf{W} := \mathbf{X}/\|\mathbf{X}\|$. The (i, j)th element σ_{ij} quantifies the strength of extremal dependence between components X_i and X_j. In particular, large values of σ_{ij} indicate a stronger tendency for extreme values of X_i and X_j to occur simultaneously. In particular, $\sigma_{ij} > 0$ corresponds to asymptotic dependence between the variables, whereas $\sigma_{ij} = 0$ indicates asymptotic independence.

To investigate the tail dependence between two components of $\mathbf{X}$ after accounting for the influence of all remaining variables, consider the partition $[\mathbf{X}_K^\top, \mathbf{X}_L^\top]^\top$, where $\mathbf{X}_K = [X_i, X_j]^\top$, $i \neq j \in [p]$ and

$\mathbf{X}_L = \mathbf{X}_{\setminus(i,j)}$. Under this partition, an extreme analogue of the conditional covariance matrix is then given by the *tail conditional covariance matrix*, which is derived from the associated tail residuals (Kim and Lee [2], Proposition 4.1),

$$\Sigma_{K|L} := \Sigma_{KK} - \Sigma_{KL}\Sigma_{LL}^{-1}\Sigma_{LK},$$

where Σ_{KL} denotes the submatrix of $\Sigma_{\mathbf{X}}$ indexed by K and L.

Definition 2.1 (Kim and Lee [2])**.** *The partial tail correlation between X_i and X_j given $\mathbf{X}_L$ is defined as*

$$\rho_{ij|L} := \frac{[\Sigma_{K|L}]_{ij}}{\sqrt{[\Sigma_{K|L}]_{ii}[\Sigma_{K|L}]_{jj}}},$$

where $[\Sigma_{K|L}]_{ij}$ denotes the off-diagonal element of $\Sigma_{K|L}$, and $\rho_{ij|L} \in [-1, 1]$.

A zero partial tail correlation indicates that X_i provides no additional information on X_j in the tail once the remaining variables $\mathbf{X}_L$ have been accounted for. To assess the significance of the extremal relationship between X_i and X_j given $\mathbf{X}_L$, we conduct a hypothesis testing procedure based on the partial tail correlation. Since $\rho_{ij|L} = 0$ is equivalent to testing

$$H_0 : [\Sigma_{K|L}]_{ij} = 0, \tag{2.3}$$

we base our inference on an estimator of $[\Sigma_{K|L}]_{ij}$ and its asymptotic normality under the null hypothesis (2.3).

2.1.1. Estimation procedure

Let $\mathbf{X}_t = [\mathbf{X}_{tK}^\top, \mathbf{X}_{tL}^\top]^\top$ for $t = 1, \ldots, n$, be i.i.d. copies of $\mathbf{X}$. For a given norm, define the radial and angular components of $\mathbf{X}_t$ by $(R_t, \mathbf{W}_t) = (\|\mathbf{X}_t\|, \mathbf{X}_t/\|\mathbf{X}_t\|)$. We begin by estimating the TPDM $\Sigma_{\mathbf{X}}$ using its empirical estimator based on exceedances of the radial component above a high threshold,

$$\widehat{\Sigma}_{\mathbf{X}}(n, k) = \frac{m}{k}\sum_{t=1}^{n} \mathbf{W}_t\mathbf{W}_t^\top \mathbb{I}[R_t \geq R_{(k)}], \tag{2.4}$$

where $m := H_{\mathbf{X}}(\Theta_+^{p-1})$ denotes the total mass of the angular measure, $\mathbb{I}$ is the indicator function, and $R_{(k)}$ is the kth largest order statistic of $\{R_t\}_{t=1}^n$. Under a tail index $\alpha = 2$ and with the L_2-norm used to define Θ_+^{p-1}, the total mass reduces to the trace of the TPDM $\Sigma_{\mathbf{X}}$, that is, $m = \sum_{j=1}^p \sigma_{jj}$ (Cooley and Thibaud [10]). In particular, if the data is preprocessed to have unit marginal scales, the total mass simplifies to $m = p$. This known value removes the need to estimate m and consequently reduces uncertainty in the estimation of the TPDM.

Given the estimator $\widehat{\Sigma}_{\mathbf{X}}(n, k)$ in (2.4), the estimator of the tail conditional covariance matrix $\widehat{\Sigma}_{K|L}$ is obtained from the corresponding block decomposition of $\widehat{\Sigma}_{\mathbf{X}}(n, k)$. Specifically,

$$\widehat{\Sigma}_{K|L}(n, k) := \widehat{\Sigma}_{KK} - \widehat{\Sigma}_{KL}\widehat{\Sigma}_{LL}^{-1}\widehat{\Sigma}_{LK}. \tag{2.5}$$

Kim and Lee [2] establish the asymptotic normality of the off-diagonal element $[\widehat{\Sigma}_{K|L}(n, k)]_{ij}$ under the null hypothesis of zero partial tail correlation. This result is derived under the second-order regular variation condition; see, e.g., Resnick and C. Stărică [11, 12].

Theorem 2.1 (Kim and Lee [2])**.** *Under the second-order regular variation condition and the null hypothesis* $[\Sigma_{K|L}]_{ij} = 0$*, the following asymptotic normality holds:*

$$\sqrt{k}\,[\widehat{\Sigma}_{K|L}(n,k)]_{ij} \;\Rightarrow\; \mathcal{N}(0,\tau^2),$$

where

$$\tau^2 = m^2 Var[\mathbf{c}_1^\top \mathbf{W}^\star \mathbf{W}^{\star\top} \mathbf{c}_2].$$

Here, $\mathbf{W}^\star$ *is a random vector taking values on* Θ_+^{p-1} *with distribution* $\mathbb{P}[\mathbf{W}^\star \in \cdot] = m^{-1} H_{\mathbf{X}}$*, and the column vectors* $\mathbf{c}_1$ *and* $\mathbf{c}_2$ *are defined through*

$$C := \begin{bmatrix}\mathbf{c}_1 & \mathbf{c}_2\end{bmatrix} = \begin{bmatrix} I_{2\times 2} & -\Sigma_{KL}\Sigma_{LL}^{-1}\end{bmatrix}^\top \in \mathbb{R}^{p\times 2},$$

Full details on the construction of the estimator, its asymptotic properties, and the associated testing procedure are provided in Kim and Lee [2].

2.2. X-vine models

We next consider the *X-vine* model (Kiriliouk et al. [5]), which provides a flexible graphical framework for modeling extremal dependence. Our focus is on the tail dependence structure of $\mathbf{X} = (X_1, \ldots, X_p)^\top$ after standardizing its marginal distribution F_j to a common scale, a standard preprocessing step in multivariate extreme value analysis. The *X-vine* framework is built on the *tail copula measure*, which characterizes extremal dependence and is closely linked to the asymptotic dependence structure of a wide class of extreme value models.

Specifically, let C denote the joint distribution function of the random vector $\mathbf{U} = (U_1, \ldots, U_p)^\top$, where $U_j = 1 - F_j(X_j)$ for $j = 1, \ldots, p$. Since our interest lies in extreme values of X_j, attention is focused on the lower tail behavior of $\mathbf{U}$.

Definition 2.2 (Kiriliouk et al. [5], Tail copula measure)**.** *The (lower) tail copula* R *associated with* C *is defined on the set* $\mathbb{E} = (0,\infty]^p \setminus \{\infty\}$ *by*

$$R(\mathbf{x}) = \lim_{t \searrow 0} t^{-1} C(t\mathbf{x}), \qquad \mathbf{x} \in \mathbb{E},$$

provided the limit exists. The corresponding tail copula measure is the Borel measure on $\mathbb{E}$ *determined by the relation* $R((\mathbf{0}, \mathbf{x}]) = R(\mathbf{x})$ *for* $\mathbf{x} \in \mathbb{E}$*.*

We assume that the tail copula measure R is supported on $(0,\infty)^p$, which is suitable for modeling joint asymptotic dependence. Moreover, R is assumed to be absolutely continuous with respect to the p-dimensional Lebesgue measure, with a continuous tail copula density $r : (0,\infty)^p \to [0,\infty)$ (Schmidt and Stadtmüller [13]). That is, $R(B) = \int_B r(\mathbf{x})\,\mathrm{d}\mathbf{x}$ for all Borel sets $B \subseteq (0,\infty)^p$. Further properties of tail copula measures and their densities, as well as their connections to alternative concepts in multivariate extreme value theory, are discussed in Section 2.1 of Kiriliouk et al. [5].

Let $G = (N, E)$ denote an undirected graph with node set N and edge set E. A vine structure is defined as a sequence of connected trees, denoted by $\mathcal{V} = \{T_j\}_{j=1}^{p-1}$. The *X-vine* construction adapts regular vine representations to the setting of multivariate extremes. Specifically, each edge in the first tree T_1 is equipped with a bivariate tail copula density, capturing primary extremal dependence between pairs of variables. Edges in the higher-order trees $T_j = (N_j, E_j)$, $j \in \{2, \ldots, p-1\}$, are

assigned bivariate copula densities that need not belong to the class of extreme value models, allowing for substantial modeling flexibility.

Figure 1 illustrates an example of a five dimensional *X-vine* model. The corresponding vine structure consists of four trees, with the edges of each tree forming the nodes of the subsequent tree. For each pair, the leading pair of variables is displayed before the semicolon, while the indices following the semicolon denote the associated conditioning set. The first tree represents a Markov tree and captures the primary extremal dependence structure. In the illustrated example, T_1 consists of four bivariate tail copula densities, including Hüsler-Reiss, negative logistic, logistic, and Dirichlet models. The higher-order trees $\{T_j\}_{j=2}^4$ encode more complex conditional dependence relationships and are specified through six bivariate copula densities, such as Clayton, Gumbel, and Gaussian copulas. The *X-vine* framework allows these bivariate building blocks to be chosen independently, without imposing additional structural constraints, resulting in a flexible modeling approach.

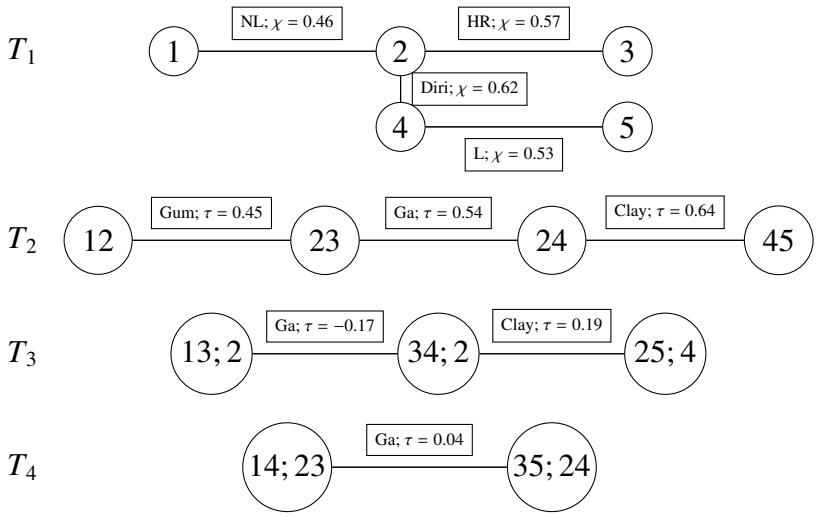

Figure 1. A five-dimensional X-vine specification. In T_1, the bivariate tail copula densities are chosen from the negative logistic (NL), logistic (L), Hüsler-Reiss (HR), and Dirichlet (Diri) parametric families, each with specified tail dependence coefficients χ. In higher-order trees $\{T_j\}_{j=2}^4$, the bivariate copula densities are selected from the Clayton (Clay), Gaussian (Ga), and Gumbel (Gum) copula families, each with specified Kendall's tau τ.

Model fitting is carried out via a sequence of recursive algorithms and consists of three main steps. For completeness, we briefly summarize the overall procedure, which is described in detail in Section 7 of Kiriliouk et al. [5].

2.2.1. Estimation of the sequential vine structure

The first step is to learn and estimate the regular vine structure from the data. Selection of the regular vine structure can be formulated as a sequential optimization problem. We adopt a nonparametric approach based on maximum spanning tree algorithms (Prim [14]; Kruskal [15]), in which edge weights are assigned using suitable measures of dependence. This greedy forward procedure constructs the vine level by level by selecting, at each stage, the maximum spanning tree that captures the strongest remaining dependencies.

The structure learning procedure proceeds as follows. For the first tree T_1, which captures extremal dependence between pairs of variables, we use an extremal measure of pairwise dependence as the edge weight. In principle, any such measure may be employed; here, we adopt the tail dependence coefficient χ_{ab} (Sibuya [16]) as the edge weight, which is widely used in extreme value analysis. This coefficient is defined as

$$\chi_{ab} = \lim_{u \to 1} \mathbb{P}(F_a(X_a) > u \mid F_b(X_b) > u) \in [0, 1], \tag{2.6}$$

where a positive value $\chi_{ab} > 0$ implies asymptotic dependence between X_a and X_b, while $\chi_{ab} = 0$ corresponds to asymptotic independence; see e.g., Coles et al. [17]. The selection procedure then begins by constructing a complete graph whose nodes represent the random variables $X_1, \ldots, X_p$. Each unordered pair $\{a, b\}$, where $a \neq b \in N_1$, is assigned an edge weight $\chi_{a,b}$. The first tree is then selected as the maximum spanning tree that maximizes the sum of these extremal dependence measures over all spanning trees

$$T_1 = \arg\max_{\text{spanning tree } T=(N_1,E)} \sum_{e=\{a,b\}\in E} \chi_e.$$

Subsequent trees T_j, for $j = 2, \ldots, p-1$, are constructed conditionally on the previously selected trees $\{T_k\}_{k=1}^{j-1}$. Each T_j, $j \geq 2$, is estimated as a maximum spanning tree on the node set N_j, with candidate edges restricted to pairs satisfying the *proximity condition*. Specifically, for any edge $\{a, b\} \in E$, the condition $|a \cap b| = 1$ must hold, meaning that two nodes in T_j may be connected only if they share a common node. For example, the edges labeled (13; 2) and (34; 2) are connected in T_4 since the leading pairs share the common node 3. A comprehensive overview of regular vine constructions is provided in Czado [18]. For these higher-order trees, Kendall's tau τ can be used as the dependence measure to define edge weights. Formally, the higher-order trees are identified as the maximum spanning trees that maximize the sum of these dependence measures over all spanning trees whose edges satisfy the *proximity condition*

$$T_j = \arg\max_{\text{spanning tree } T=(N_j,E) \text{ with } E \subseteq E_{j,prox}} \sum_{e\in E} |\tau_e|,$$

where τ_e is the edge weight associated with the edge e, and $E_{j,prox}$ denotes the set of candidate edges satisfying the *proximity condition*. To induce parsimony, the vine structure may be truncated at a selected tree level. Kriliouk et al. [5] adapted a modified Bayesian information (mBIC) to determine the truncation level, inspired by the approach proposed for regular vine copulas in Nagler et al. [19].

2.2.2. Selection of copula families given a regular vine sequence

Given the estimated vine tree sequence obtained in Section 2.2.1, we select an appropriate bivariate (tail) copula family for each edge from a specified set of candidate families. Model selection is performed using information criteria such as the Akaike information criterion (AIC) or the Bayesian information criterion (BIC).

2.2.3. Estimation of (tail) copula parameters

Given the selected vine structure and the chosen bivariate (tail) copula family for each edge, the associated parameter(s) are estimated via maximum-likelihood. In contrast to standard vine copula estimation procedures in non-extreme settings, the pseudo-likelihoods in the *X-vine* model depend

on an effective sample size for the conditioning variables, as the procedure is based on the threshold exceedances that approximate the (inverted) multivariate Pareto distribution.

2.3. Extremal graphical models

Engelke and Hitz [6] introduce the notion of *extremal conditional independence* within the framework of multivariate Pareto distributions (MPDs; Rootzén and Tajvidi [7]) and propose corresponding modeling and estimation procedure for *extremal graphical models*. After marginally standardizing the original random vector $\mathbf{X} = (X_1, \ldots, X_p)^\top$ to have standard Pareto margins where $U_j = (1 - F_j(X_j))^{-1}$, the MPD, denoted $\mathbf{Y}$, arises as the limiting distribution of threshold exceedances of $\mathbf{X}$, providing a probabilistic representation of multivariate extremes.

An *extremal graphical model* is defined on an undirected graph $G = (N, E)$ via the pairwise Markov property: for any pair of nodes (a, b) such that $(a, b) \notin E$, the corresponding variables satisfy $Y_a \perp_e Y_b \mid \mathbf{Y}_{\setminus(a,b)}$, where $\perp_e$ denotes extremal conditional independence. This condition encodes extremal conditional independence between Y_a and Y_b given all remaining variables. A key distinction is that, within the MPD framework, the density assumption permits extremal conditional independence between Y_a and Y_b, given $\mathbf{Y}_{\setminus(a,b)}$ to be characterized directly, in contrast to the partial tail correlation approach described in Section 2.1.1. However, when the underlying MPD admits a density, the associated graph G must be connected (Engelke and Hitz [6]). By contrast, the hypothesis testing procedure for zero partial tail correlation imposes no structural assumptions and thus allows for disconnected graphs.

Similar to the fitting procedure for the *X-vine* model, the underlying conditional independence structure is typically unknown in practice and must first be inferred from the data. A particularly simple and parsimonious class of extremal graphical models is obtained by restricting the graph G to a tree. Tree-structured extremal graphical models are attractive from an applied perspective, as they admit nonparametric estimation procedures and are computationally efficient. Engelke and Volgushev [20] show that, if $\mathbf{Y}$ follows an extremal graphical model on an unknown tree T, then the minimum spanning tree T_{mst}, constructed from suitable extremal dependence weights, is unique and recovers the true structure, i.e., $T_{mst} = T$. The minimum spanning tree is defined as

$$T_{mst} = \arg\min_{\text{spanning tree T=(N,E)}} \sum_{e=\{a,b\}\in E} w_e.$$

Parametric MPDs can also be specified on tree-structured graphs. In particular, Engelke and Volgushev [20] focus on bivariate Hüsler-Reiss models along the edges of the tree, parametrized by a variogram matrix Γ. When empirical estimates of the extremal variogram $\widehat{\Gamma}_{ab}$ are used as edge weights, the resulting minimum spanning tree provides a consistent estimator of the underlying tree structure.

For a given tree and a parametric specification of the MPD on that tree, the parameters associated with each edge can be estimated using, for example, the (censored) maximum likelihood method. Notably, the Hüsler-Reiss model is especially well suited for characterizing sparse extremal graphs, as zero elements in the precision matrix derived from Γ directly encode extremal conditional independence relationships.

To accommodate more general structure learning beyond trees or decomposable graphs, Engelke et al. [21] proposed the *eglearn* method for high-dimensional settings. Given an estimate of the

variogram matrix $\widehat{\Gamma}$, the corresponding covariance matrix can be computed. Sparsity in the estimated extremal graph is then induced by incorporating an l_1-penalty on the precision matrix, controlled by a tuning parameter $\rho \geq 0$, through graphical lasso-type algorithms. The tuning parameter is selected by minimizing the Bayesian information criterion (BIC) across the fitted models, providing a data-driven balance between model fit and complexity.

2.4. Summary of modeling assumptions

Table 1 summarizes the key implications of the modeling assumptions underlying the three methods described above. *Partial tail correlation* (PTC) requires neither density assumptions nor a prescribed graphical structure, making it less model-dependent and enabling formal hypothesis testing without restricting attention to specific graphical classes. However, PTC does not imply extremal conditional independence. In contrast, graphical approaches such as *X-vines* and *extremal graphical models* rely on density and structure assumptions, which allow for extremal conditional independence, but necessarily induce connected graphs and exclude disconnected structures.

Table 1. Comparison of modeling assumptions and implications across the three approaches. PTC and EGM indicate partial tail correlation and extremal graphical models, respectively.

	PTC	X-vine	EGM
Density assumptions	Not required	Required	Required
Graphical structure	Not required	Regular vine	Tree or block graphs
Hypothesis testing	Available	Not available	Not available
Extremal conditional independence	Not implied	Implied	Implied
Disconnected graphs	Allowed	Not allowed	Not allowed

3. Application: Lancashire rainfall extremes

We apply the three methodologies introduced in Section 2 to a rainfall dataset from the Lancashire region of the United Kingdom, with the aim of investigating and comparing extremal dependence structures among rainfall stations. This application provides a setting for assessing how different structure learning approaches characterize extremal relationships and their practical implications.

3.1. Data description and exploratory analysis

The original data is obtained from Met Office [22] and consists of daily rainfall measurement recorded at Met Office observation stations across the United Kingdom. Our analysis focuses on the period from 1960 to 2024. As is typical for rainfall data, seasonality is present, with winter precipitation exhibiting heavier extremes than summer rainfall. As an illustration, Figure 2 displays box-plots comparing summer and winter rainfall for the station with location 'ID 12068_entwistle-resr' and 'ID 00507_slaidburn'. Similar heavier-tailed behavior is observed at other locations. To mitigate the influence of seasonal effects, we therefore restrict attention to observations from the winter months. The Mann-Kendall test (Mann [23]; Kendall [24]) detected no statistically significant trends in the annual maxima across the stations at the 5% level.

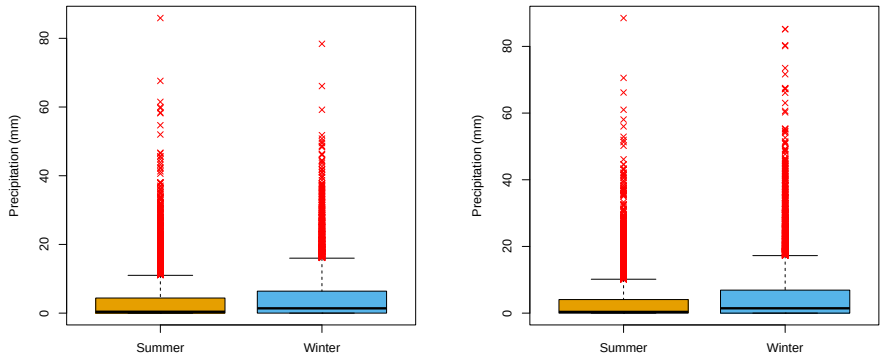

Figure 2. Boxplots comparing summer and winter rainfall at the station with location 'ID 12068_entwistle-resr' (left) and 'ID 00507_slaidburn' (right).

The raw data contain missing observations. To ensure data consistency across stations, we retain only the 30 gauge stations for which the proportion of missing values is below 0.1. We further restrict the dataset to days on which observations are available at all retained stations (see Figure 3).

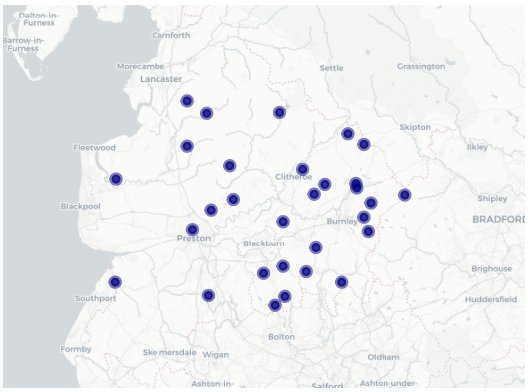

Figure 3. Geographical map of the Lancashire region of the United Kingdom, with observation stations marked by blue circles. Note: Map tiles by CartoDB (CC BY 3.0); Data by OpenStreetMap.

To assess marginal tail behavior, we estimate the shape parameter at each station by fitting a generalized Pareto distribution to threshold exceedances. Specifically, let X_t denote the daily rainfall on day t, and suppose $\{X_t\}$ forms a stationary process. Extremes of $\{X_t\}$ are defined as exceedances above a high threshold u. Pickands [25] showed that the distribution of excesses, $X_t - u$, conditional on $X_t > u$, converges to a nondegenerate limiting distribution, the generalized Pareto distribution (GPD). The conditional distribution of the excesses, denoted by GPD(ψ_u, ξ), is given for $x > 0$ by

$$\mathbb{P}(X > x + u \mid X > u) = \left(1 + \frac{\xi}{\psi_u} x\right)_+^{-1/\xi}, \tag{3.1}$$

where $s_+ = \max(s, 0)$, and ψ_u and ξ denote the scale and shape parameters, respectively. Negative, zero, and positive values of the shape parameter correspond to bounded, light-tailed, and heavy-tailed

distributions, respectively. The average estimated shape parameter across stations is 0.063, with 63% of the stations exhibiting positive shape parameter estimates.

3.2. *Hypothesis test for zero partial tail correlation*

Let $\mathbf{X}$ denote the random vector of daily rainfall amounts, with $\mathbf{X}_t$, $t = 1, \ldots, n$, representing i.i.d. replicates of $\mathbf{X}$. As a standard preprocessing step in multivariate extreme value analysis, we apply a rank-based marginal transformation to achieve tail equivalence with tail index $\alpha = 2$; see, e.g., Theorem 6.5 of Resnick [9]. Specifically, for each station $j \in \{1, \ldots, p\}$, define $U_{t,j} = R_{t,j}/(n+1)$, where $R_{t,j} = \sum_{s=1}^{n} I_{(X_{s,j} \le X_{t,j})}$ denotes the rank of $X_{t,j}$ among $X_{1,j}, \ldots, X_{n,j}$. We then apply the marginal transformation $Z_{t,j} = (1 - U_{t,j})^{-1/2}$, yielding transformed observations with standard Pareto margins and common tail index $\alpha = 2$ at each gauge station.

Using the transformed data, we estimate the 30×30 TPDM $\Sigma_{\mathbf{Z}}$ via the estimator (2.4), with the 95% quantile of the radial component $\|\mathbf{Z}_t\|$ used as the threshold. The resulting TPDM estimates are observed to be stable across a range of high quantile choices. The left panel of Figure 4 displays the estimated pairwise extremal dependence between a reference station (labeled 1), selected for its low proportion of missing values, and all remaining stations across Lancashire. As expected, extremal dependence is stronger among stations near the reference station and decays with increasing distance. This spatial pattern is consistently observed across other reference stations.

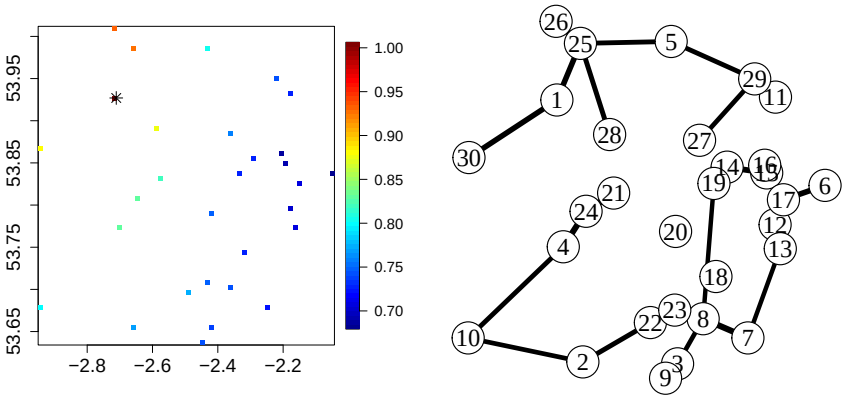

Figure 4. Left: Estimated elements of the TPDM between the reference station (marked by $*$) and the remaining stations. Right: Graph induced by statistically significant partial tail correlations.

To assess whether extremal dependence between two stations persists after accounting for the influence of all remaining stations, we consider the partition $\mathbf{X}_t = [\mathbf{X}_{tK}^\top, \mathbf{X}_{tL}^\top]^\top$ for $t = 1, \ldots, n$, where $\mathbf{X}_K = [X_i, X_j]^\top$ for distinct $i, j \in [p]$ and $\mathbf{X}_L = \mathbf{X}_{\setminus(i,j)}$. Under the assumption $\mathbf{X} \in RV_+^p(2)$, we test the null hypothesis $H_0 : [\Sigma_{K|L}]_{ij} = 0$. Inference is based on an estimator of the tail conditional covariance matrix $\widehat{\Sigma}_{K|L}(n,k)$ in (2.5) and the corresponding z-statistic, $z_{n,k} := \sqrt{k}\hat{\tau}^{-1}[\widehat{\Sigma}_{K|L}(n,k)]_{ij}$ derived from its asymptotic normality under the null hypothesis from Theorem 2.1; full details are provided in Kim and Lee [2]. Values of the z-statistic close to zero indicate weak residual tail association between the station pair after conditioning on the remaining stations. This testing procedure is applied to all

unordered pairs of stations.

To control the family-wise Type 1 error rate arising from multiple testing, we apply a Bonferroni adjustment, resulting in a critical value $z_{crit} = 3.86$. The null hypothesis is not rejected if $|z_{n,k}| < 3.86$, in which case the corresponding pair of stations is disconnected in the inferred graph. For visualization, we construct an *undirected* graph over the 30 stations, shown in the right panel of Figure 4. Edge thickness is proportional to the magnitude of the z-statistic, reflecting the strength of residual extremal dependence after conditioning. The resulting graph contains 25 edges out of a possible 435 and provides a concise summary of the dominant extremal dependence structure in the rainfall data. Notably, the inferred network exhibits two distinct clusters corresponding broadly to northern and southern subregions of Lancashire, suggesting spatial heterogeneity in extremal rainfall behavior.

3.3. X-vine model fit

We next apply X-vine models to explore extremal dependence among extreme rainfall observations across the Lancashire region. Let $\widehat{\mathbf{U}}_i = (\widehat{U}_{i,1}, \ldots, \widehat{U}_{i,p})$, for $i = 1, \ldots, n$, where $\widehat{U}_{i,j} = 1 - (R_{i,j} - 0.5)/n$, and $R_{i,j} = \sum_{s=1}^{n} \mathbb{I}_{X_{s,j} \le X_{i,j}}$ is the rank of $X_{i,j}$ among $\{X_{t,j}\}_{t=1}^{n}$ for each $j \in \{1, \ldots, p\}$. For large $t > 0$, the rescaled points $t\widehat{\mathbf{U}}_i$ such that $\min \widehat{\mathbf{U}}_i < 1/t$ can be reviewed as pseudo-observations from a distribution approximating the *inverted MPD*.

Let $\widehat{\mathbf{U}}_i$ denote the pseudo-uniform observations from the joint distribution C defined in Definition 2.2. We define $t = n/k$, where $k \in \{1, \ldots, n\}$ is chosen to be large while ensuring that the threshold ratio k/n remains small. In the standard asymptotic framework of extreme value theory, this corresponds to a sequence $k = k_n$ satisfying $k_n \to \infty$ and $k_n/n \to 0$. We construct the subsamples $\widehat{\mathbf{Z}}_i = (n/k)\widehat{\mathbf{U}}_i$ by retaining observations for which $\min \widehat{\mathbf{U}}_i < k/n$, where k/n is the threshold applied to the uniform margin $\hat{u}_{i,j}$.

Choosing an appropriate k requires balancing the need for a sufficiently large effective sample size against the bias-variance trade-off inherent in threshold-based extreme value models. As is standard in extremes, larger values of k reduce variance at the expense of increased bias, and vice versa. We consider threshold ratios $k/n \in \{0.1, 0.06, 0.04, 0.02\}$ and select k such that $k/n = 0.06$. This choice is supported by the goodness-of-fit assessment based on the χ-plot shown in the left panel of Figure 8, which indicates relatively low bias and variability while maintaining a sufficient effective sample size.

To first estimate the vine structure, we select the trees sequentially using a maximum spanning tree approach as described in Section 2.2.1. Edge weights are defined using the empirical tail dependence coefficient $\widehat{\chi}_e$ for edges $e \in E_1$ in T_1, and empirical Kendall's tau $\widehat{\tau}_e$ for edges $e \in E_j$, $j \ge 2$, in higher-order trees $\{T_j\}_{j=2}^{29}$. The resulting first maximum spanning tree is shown in the top-left panel of Figure 5.

Among the $p - 1 = 29$ edges in T_1, the selected bivariate tail copula families consist of negative logistic (8 edges), logistic models (20 edges), and Dirichlet models (1 edge). As the tree level increases, we observe that residual dependence tends to weaken, leading to a growing number of independence copulas in higher-order trees. To further investigate model parsimony, we consider truncated X-vine models. The truncation level is selected using a modified Bayesian information criterion (mBIC) as in Kiriliouk et al. [5]. The mBIC values shown in the left panel of Figure 6 indicate an optimal truncation level of $q^* = 6$, implying all pair copulas from tree level 7 onward are set to the independence copula. In other words, the vine is truncated at level $q = 6$.

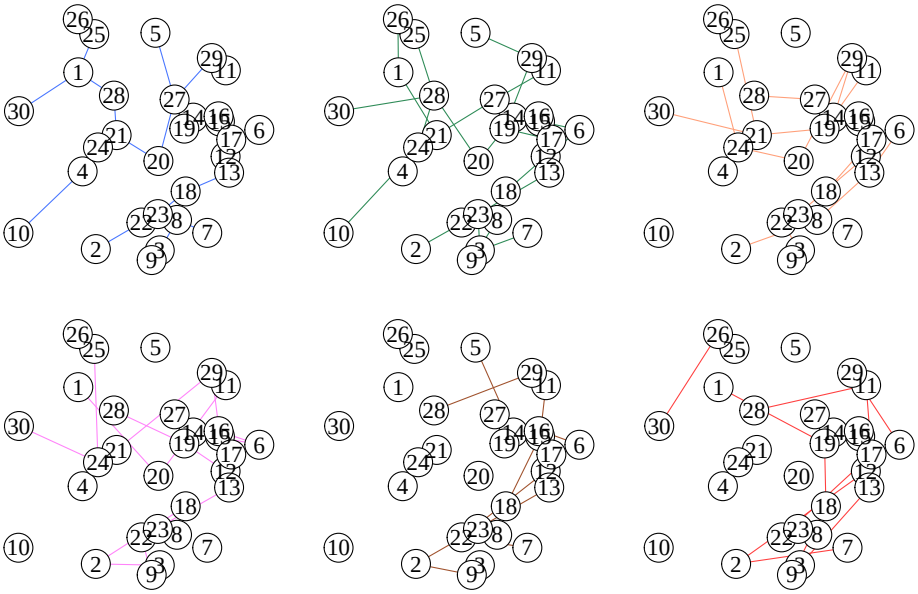

Figure 5. The estimated vine tree sequence (T_1 to T_6), ordered from top to bottom and left to right.

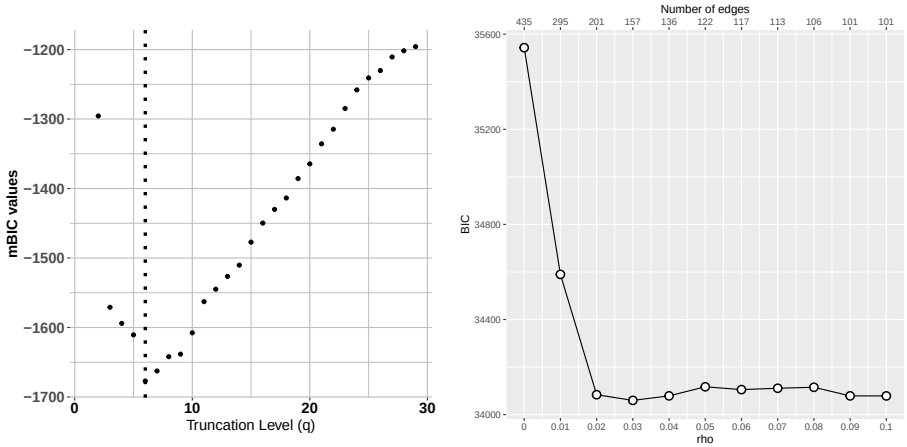

Figure 6. Left: mBIC values plotted against the truncation level q with a minimum at $q^* = 6$. Right: BIC values plotted against the tuning parameter value ρ, attaining a minimum at $\rho^* = 0.03$.

Under this truncated X-vine specification, 48 dependence copulas are retained out of 159 possible edges from T_2 to T_6. These include Gaussian (7), Clayton (1), Gumbel (23), Frank (20), Joe (13), Survival Clayton (10), Survival Gumbel (7), and Survival Joe (1) copulas. The diversity of selected bivariate building blocks highlights the flexibility of the X-vine framework. For visual clarity, Figure 5 displays the sequence T_1 through T_6, illustrating the progression toward independence as the tree level increases. Model fitting is performed using the `Xvine` R package, available at `https://github.com/JeongjinLee88/Xvine`.

3.4. Extremal graphical model fit

As in the *X-vine* analysis, we begin by applying a rank-based marginal transformation to the rainfall data. Let $\widehat{\mathbf{U}}_i = (\widehat{U}_{i,1}, \ldots, \widehat{U}_{i,p})$, for $i = 1, \ldots, n$, where $\widehat{U}_{i,j} = (1 - R_{i,j}/n)^{-1}$ and $R_{i,j} = \sum_{s=1}^{n} \mathbb{I}_{X_{s,j} \le X_{i,j}}$ is the rank of $X_{i,j}$ among $\{X_{t,j}\}_{t=1}^{n}$ for each $j \in \{1, \ldots, p\}$. For large $t > 0$, the rescaled points $\widehat{\mathbf{U}}_i/t$ such that $\max \widehat{\mathbf{U}}_i > t$ may be regarded as pseudo-observations that approximate the MPD. We construct the subsamples $\widehat{\mathbf{Z}}_i = (k/n)\widehat{\mathbf{U}}_i$ by retaining observations for which $\max \widehat{\mathbf{U}}_i > n/k$, where k/n is the threshold applied to the uniform margin $\hat{u}_{i,j}$. For comparability with the *X-vine* model, we select k such that $k/n = 0.94$, ensuring a large effective sample size relative to the number of variables in the extremal graphical model.

We consider extremal graphical models based on the Hüsler-Reiss distribution. To first examine the dominant extremal dependence structure, we estimate an extremal tree model using the empirical extremal variogram matrix $\widehat{\Gamma}_{ab}$ as edge weights. In this construction, a *minimum* spanning tree is formed, with smaller entries of $\widehat{\Gamma}_{ab}$ indicating stronger extremal dependence. The resulting extremal tree, shown in the left panel of Figure 7, is fully connected by construction, reflecting the assumptions of the model. Note that if the empirical tail dependence coefficient $\widehat{\chi}_{ab}$ were used as the edge weight instead, the resulting first tree would coincide with that obtained from the X-vine analysis.

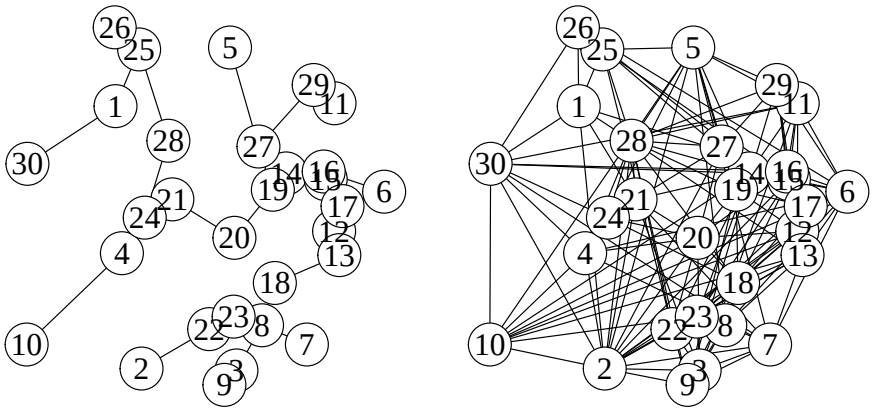

Figure 7. Left: Estimated extremal tree structure. Right: Estimated extremal graphical structure via `eglearn` with the selected optimal tuning parameter value $\rho^* = 0.03$, resulting in 157 edges.

To allow for sparser dependence structures beyond trees, we further consider the structure learning approach proposed by Engelke et al. [21], implemented via the EGlearn algorithm. This method introduces an ℓ_1-penalty on the precision matrix of the Hüsler-Reiss model, controlled by a tuning parameter $\rho \ge 0$, to induce sparsity in the estimated extremal graph. We select the optimal tuning parameter by evaluating the Hüsler-Reiss log-likelihood values over a grid of $\rho \in \{0, 0.01, \ldots, 0.1\}$ values, yielding the optimal value $\rho^* = 0.03$.

The resulting sparse extremal graph at $\rho^* = 0.03$ contains 157 edges and is displayed in the right panel of Figure 7. Estimation is carried out using the R package `graphicalExtremes` (Engelke et al. [26]), available at `https://CRAN.R-project.org/package=graphicalExtremes`. We apply the extremal graphical lasso using the `eglearn` function.

3.5. Goodness-of-fit assessment for graphical models

We evaluate goodness-of-fit by comparing empirical tail dependence coefficients with those derived from the fitted graphical models; see Appendix E.2 in the supplement of Kiriliouk et al. [5] for details. Figure 8 presents the corresponding χ-plots, with the X-vine model shown in the left panel and the extremal Hüsler-Reiss graphical model in the right panel. Both plots show close alignment along the diagonal, indicating a satisfactory overall fit. Due to the recursive nature of sequential parameter estimation, uncertainty tends to accumulate across tree levels in the X-vine model, resulting in relatively low bias but increased variability. In contrast, the extremal Hüsler-Reiss graphical model displays lower variability, albeit with somewhat higher bias for higher tail dependence coefficients.

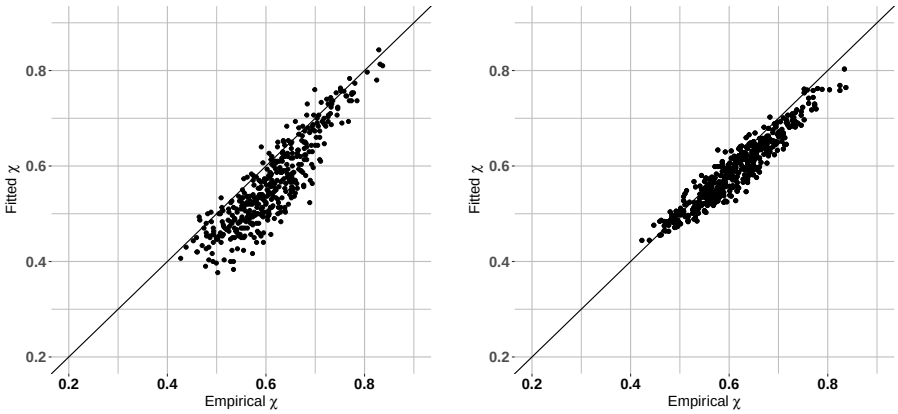

Figure 8. Left: χ-plot of empirical tail dependence coefficients plotted against fitted coefficients from the X-vine model. Right: χ-plot of empirical coefficients versus fitted values from the Hüsler-Reiss extremal graphical model.

3.6. Comparative interpretations

A key finding is that the graph induced by the hypothesis testing procedure for partial tail correlations exhibits two distinct clusters corresponding broadly to the northern and southern subregions of Lancashire. This separation indicates a lack of residual extremal association between the two subregions. In contrast, graphical models that rely on density or constructional assumptions necessarily yield connected structures by design and hence do not recover this complete separation.

Despite this structural constraint, the two graphical models exhibit substantial agreement in their representation of dominant extremal dependence patterns with their corresponding extremal trees sharing similar large-scale features. When the dependence information across multiple trees is superimposed, both graphical modeling approaches indicate the presence of clustered extremal dependence within Lancashire.

Further insight is provided by the sequential nature of the X-vine model. Extremal dependence in the first tree T_1 is strong along the selected edges, with estimated tail dependence coefficients ranging from 0.66 to 0.81. In subsequent trees, higher-order dependence exhibits a clear tendency to strengthen in the southeastern region as the order of conditioning increases. Notably, as the tree level grows, higher-order dependence, measured by the $\hat{\tau}$ estimate, tends to persist around a sub-cluster of stations in the southeastern area (e.g., stations $6, 12, 13, 14, 15, 16, 17, 18, 19$) across successive trees.

This persistence indicates that extremal dependence within this sub-cluster remains strong even after conditioning on other stations. Overall, the dominant source of higher-order dependence arises from the southeastern sub-cluster. This heterogeneity in extremal rainfall behavior, together with the consistently stronger dependence in the southeastern region, suggests that particular attention to risk mitigation in this area is warranted.

This pattern is compatible with the separation observed in the partial tail correlation graph and may suggest that extremal dependence in the southeastern is relatively more persistent under conditioning. Taken together, the three methodologies provide complementary perspectives: partial tail correlation highlights separability, while graphical models reveal the internal structure and hierarchical organization of extremal dependence within connected regions.

4. Conclusions

In this study, we applied three complementary statistical methodologies to investigate extremal dependence structures in extreme rainfall data from the Lancashire region of the United Kingdom.

The fitted graphical models reveal an overall clustered pattern of extremal dependence across the region, suggesting that spatial extreme models incorporating structured dependence may be well suited for quantifying rare rainfall events in this region while accounting for spatial extent. The partial tail correlation analysis uncovers a distinct separation between northern and southern subregions. This form of separation is not imposed by, and therefore cannot be directly recovered from, density-based graphical modeling approaches.

From a practical perspective, these findings point to meaningful heterogeneity in extremal rainfall behavior across Lancashire. The presence of stronger and higher-order extremal dependence in the southeastern area suggests that mitigation strategies and investments in defense infrastructure may have differential impacts across subregions. Accounting for such structural differences in extremal dependence can therefore play an important role in informing region-specific risk assessment and decision-making for extreme weather events.

Author contributions

Jeongjin Lee: Conceptualization, Methodology, Formal analysis, Writing original draft; Yongku Kim: Conceptualization, Writing review & editing. All authors have read and approved the final version of the manuscript for publication.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

Jeongjin Lee's research was partially supported by UK Engineering and Physical Sciences Research Council (EPSRC) grant EP/X010449/1 and Yongku Kim's research was supported by Global-Learning & Academic research institution for Master's·PhD students, and Postdocs (G-LAMP) Program of the

National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (No. RS-2023-00301914).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. L. de Haan, A. Ferreira, *Extreme Value Theory: An Introduction*, New York: Springer, 2006.
2. M. Kim, J. Lee, Hypothesis testing for partial tail correlation inmultivariate extremes, preprint paper, 2025. http://doi.org/10.48550/arXiv.2210.02048
3. T. Bedford, R. M. Cooke, Probability density decomposition for conditionally dependent random variables modeled by vines, *Ann. Math. Artific. Intell.*, **32** (2001), 245–268. https://doi.org/10.1023/A:1016725902970
4. T. Bedford, R. M. Cooke, Vines–a new graphical model for dependent random variables, *Ann. Statist.*, **30** (2002), 1031–1068. https://doi.org/10.1214/aos/1031689016
5. A. Kiriliouk, J. Lee, J. Segers, X-vine models for multivariate extremes, *J. Royal Stat. Soc. Ser. B: Stat. Methodol.*, **87** (2025), 579–602. https://doi.org/10.1093/jrsssb/qkae105
6. S. Engelke, A. S. Hitz, Graphical models for extremes, *J. Royal Stat. Soc. Ser. B: Stat. Methodol.*, **82** (2020), 871–932. https://doi.org/10.1111/rssb.12355
7. H. Rootzén, N. Tajvidi, Multivariate generalized Pareto distributions, *Bernoulli*, **12** (2006), 917–930. https://doi.org/10.3150/bj/1161614952
8. J. Hüsler, R. D. Reiss, Maxima of normal random vectors: Between independence and complete dependence, *Stat. Probab. Lett.*, **7** (1989), 283–286.
9. S. I. Resnick, *Heavy-tail Phenomena: Probabilistic and Statistical Modeling*, New York: Springer, 2007.
10. D. Cooley, E. Thibaud, Decompositions of dependence for high-dimensional extremes, *Biometrika*, **106** (2019), 587–604. https://doi.org/10.1093/biomet/asz028
11. S. I. Resnick, C. Stărică, Asymptotic behavior of Hill's estimator for autoregressive data, *Commun. Statist. Stochast. Models*, **13** (1997a), 703–721. https://doi.org/10.1080/15326349708807448
12. S. I. Resnick, C. Stărică, Smoothing the Hill estimator, *Adv. Appl. Probab.*, **29** (1997), 271–293. https://doi.org/10.2307/1427870
13. R. Schmidt, U. Stadtmüller, Non-parametric estimation of tail dependence, *Scandinavian J. Statist.*, **33** (2006), 307–335. https://doi.org/10.1111/j.1467-9469.2005.00483.x
14. R. C. Prim, Shortest connection networks and some generalizations, *Bell Syst. Techn. J.*, **36** (1957), 1389–1401. https://doi.org/10.1002/j.1538-7305.1957.tb01515.x
15. J. B. Kruskal, On the shortest spanning subtree of a graph and the traveling salesman problem, *Proc. Amer. Math. Soc.*, **7** (1956), 48–50. https://doi.org/10.2307/2033241

16. M. Sibuya, Bivariate extreme statistics, *Ann. Inst. Statist. Math.*, **11** (1960), 195–210.
17. S. Coles, J. Bawa, L. Trenner, P. Dorazio, *An Introduction to Statistical Modeling of Extreme Values*, London: Springer, 2001.
18. C. Claudia, *Analyzing Dependent Data with Vine Copulas*, Switzerland: Springer, 2019. https://doi.org/10.1007/978-3-030-13785-4
19. T. Nagler, C. Bumann, C. Czado, Model selection in sparse high-dimensional vine copula models with an application to portfolio risk, *J. Multivar. Anal.*, **172** (2019), 180–192. https://doi.org/10.1016/j.jmva.2019.03.004
20. S. Engelke, S. Volgushev, Structure learning for extremal tree models, *J. Royal Statist. Soc.: Ser. B Statist. Methodol.*, **84** (2022), 2055–2087. https://doi.org/10.1111/rssb.12556
21. S. Engelke, M. Lalancette, S. Volgushev, Learning extremal graphical structures in high dimensions, preprint paper, 2021. https://doi.org/10.48550/arXiv.2111.00840
22. Met Office, MIDAS open: UK daily rainfall data, v202507, 2025.
23. H. Mann, Nonparametric tests against trend, *Econometrica*, **13** (1945), 245–259. https://doi.org/10.2307/1907187
24. M. Kendall, *Rank Correlation Methods*, London: Charles Griffin, 1975.
25. J. Pickands III, Statistical inference using extreme order statistics, *Ann. Statist.*, **3** (1975), 119–131.
26. S. Engelke, A. Hitz, G. Nicola, M. Hentschel, GraphicalExtremes: Statistical methodology for graphical extreme value models, *R package version 0.3.4*, 2022.

AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0)