Title

Methods of Tail Dependence Estimation

Permalink

https://escholarship.org/uc/item/07x6p3bk

ISBN

978-94-007-4478-3

Authors

Aghakouchak, Amir

Sellars, Scott

Sorooshian, Soroosh

Publication Date

2012-09-07

DOI

10.1007/978-94-007-4479-0_6

Chapter 6
Methods of Tail Dependence Estimation

Amir AghaKouchak, Scott Sellars, and Soroosh Sorooshian

Abstract Characterization and quantification of climate extremes and their dependencies are fundamental to the studying of natural hazards. This chapter reviews various parametric and nonparametric tail dependence coefficient estimators. The tail dependence coefficient describes the dependence (degree of association) between concurrent extremes at different locations. Accurate and reliable knowledge of the spatial characteristics of extremes can help improve the existing methods of modeling the occurrence probabilities of extreme events. This chapter will review these methods and use two case studies to demonstrate the application of tail dependence analysis.

6.1 Introduction

Weather and climate extremes are of particular importance due to their impacts on the economy, environment and human life. Understanding the spatial dependence structure of rare events is fundamental in risk assessment and decision making. Most measures of dependence (e.g., Pearson linear correlation, Spearman (1904) and Kendall (1962) correlation) are designed to describe the dependence of random variables over their distributions. Most commonly used measures are not able to correctly capture the dependence of the upper or lower parts (extremes) of the distribution (Kotz and Nadarajah 2000). For example, the Pearson correlation coefficient may not exist for random variables above a certain high (extreme) threshold (De Michele et al. 2003). The Pearson linear correlation describes how

A. AghaKouchak (✉) • S. Sellars • S. Sorooshian
Department of Civil and Environmental Engineering, University of California Irvine,
92697 Irvine, CA, USA
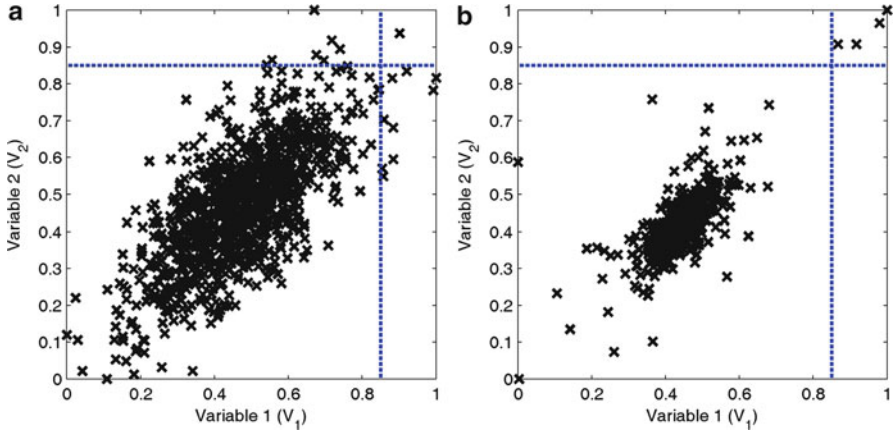e-mail: amir.a@uci.edu; scott.sellars@uci.edu; soroosh@uci.edu

Fig. 6.1 (*left*): Upper tail values (upper right quadrant – above *dotted lines*) of V_1 and V_1 are locally independent; (*right*): Upper tail values of V_1 and V_2 seem to be locally correlated (Modified after AghaKouchak et al. (2010c))

well two random variables are linearly correlated with respect to their entire distribution. However, this information cannot be used to understand how the extremes of two random variables are dependent (Serinaldi 2008).

In general, most dependence measures associate the entire distribution of two or more random variables. However, the dependence between the upper part of the distribution may be different than the mid-range and/or lower part of the distribution (Embrechts et al. 2002). For example, two random variables with low dependence between mid-range values, but strong association among high (low) values.

In extreme value analysis, the tail dependence coefficient describes the association between the upper or lower part (tail) of two or more random variables (Schmidt 2005; Frahm et al. 2005; Ledford and Tawn 1997). The tail dependence coefficient is first introduced by Sibuya (1959) as the dependence in the upper-right and lower-left quadrants of a bivariate distribution function. Put differently, in a bivariate distribution, the tail dependence refers to the limiting proportion that one variable's marginal distribution exceeds a certain threshold given the other variable's margin has already exceeded the same threshold.

Figure 6.1 explains the concept of tail dependence using an example. The figure displays two generated random variables with the same linear correlation coefficient of approximately 0.7. Figure 6.1 (left) is simulated using the bivariate normal distribution, and Fig. 6.1 (right) is generated using the bivariate t-distribution. In both cases the simulated variables are transformed to uniform [0-1] distribution. The variables (V_1 and V_2) in both figures show positive Pearson linear correlation coefficient (≈ 0.7). However, the upper right quadrant (above the dotted lines) is different in the left and right panels of Fig. 6.1. As shown, in Fig. 6.1 (left) the values in the upper right quadrant (upper tails of V_1 and V_1) are locally independent, while in Fig. 6.1 (right) the upper tail values seem to be locally correlated (compare the

upper right corners of both panels). This indicates that the probability of occurrence of V_2 above a given high threshold (e.g., dotted line in the figure), assuming V_1 exceeds the same threshold is higher in the right panel compared to the left panel in Fig. 6.1. For additional information and graphical examples, the interested reader is referred to Fisher and Switzer (2001) and Abberger (2005).

Parametric methods are frequently used for univariate extreme value analysis (e.g., Fisher and Tippett 1928; Gumbel 1958). On the other hand, in multivariate extreme value analysis, the joint probabilities of multiple random variables is considered. This includes the probability occurrence (risk) of each variable based on its univariate marginal distribution and the dependence of multiple probability occurrences. Depending on the marginal distribution of random variables and their dependence structure, a parametric model may or may not be sufficient to model the characteristics of the joint extremes. Thus far, many parametric and nonparametric methods have been developed for analysis of tail dependence of random variables (Schmidt and Stadtmüller 2006; Malevergne and Sornette 2004; Poon et al. 2004; Ledford and Tawn 2003; Malevergne and Sornette 2002; Ledford and Tawn 1996).

In the past three decades, most applications of tail dependence models have been in financial risk management and dependence analysis of between extreme assets (e.g., Schmidt (2005), Frahm et al. (2005) and Embrechts et al. (2002) and references therein). In particular, many non-parametric methods are introduced based on the concept of empirical copula. Copulas are multivariate distribution functions that can describe the dependence of two or more random variables independent of their marginal distrubtions. In recent years, multivariate copulas have been applied in numerous hydrologic applications (Nazemi and Elshorbagy 2011; AghaKouchak et al. 2010b; Bárdossy and Li 2008; Serinaldi 2009; Zhang et al. 2008; AghaKouchak et al. 2010a; Favre et al. 2004; De Michele and Salvadori 2003; Kelly and Krzysztofowicz 1997). Renard and Lang (2007) investigated the usefulness of the multivariate normal copula in extreme value analysis. With several case studies, Renard and Lang (2007) demonstrated that the multivariate normal copula can be reasonably used for extreme value analysis. However, the authors acknowledge that the low probabilities can be significantly underestimated if asymptotically dependent random variables are described by the normal copula, which is an asymptotically independent model. Serinaldi (2008) investigated the association of rainfall data using the non-parametric Kendall rank correlation. The study suggests a copula-based mixed model for modeling the dependence structure and marginal distributions of variables.

In general, the tail dependence between variables may strongly depend on the choice of model or estimation technique (Frahm et al. 2005). This chapter reviews several parametric and non-parametric tail dependence estimators. Various aspects of modeling tail dependence between variables are discussed in detail, including the choice of extreme value threshold, and advantages and disadvantages of tail dependence models. The chapter is organized into seven sections. After the introduction, the concept of tail dependence is reviewed. The third section is devoted to parametric tail dependence analysis and copulas. In section four, non-parametric

methods are discussed. Section five provides additional insight into the choice of extreme value threshold. The last section highlights two case studies using the tail dependence estimators.

6.2 Tail Dependence: Basic Definitions

Let $X_1, \ldots, X_n$ be n random variables. The the upper tail (λ_{up}) for a multivariate distribution with n random variables $X(X_1, \ldots, X_n)$ is defined as (Joe 1997; Melchiori 2003):

$$\lambda_{up} = \lim_{u \to 1^-} Pr\left(F_1(X_1) > t | F_2(X_2) > t \ldots F_n(X_n) > t\right) \tag{6.1}$$

In Eq. 6.1, $F_1, \ldots, F_n$ are the cumulative distribution functions for the random variables $X_1, \ldots, X_n$, and t is the extreme value threshold. The equation expresses the probability (Pr) of occurrence of extremes (values above the threshold t) in X_1, conditioned on the occurrence of extremes (above the same threshold) in $X_2, \ldots, X_n$. Similarly, the lower tail dependence coefficient (λ_{lo}) is described as:

$$\lambda_{lo} = \lim_{u \to 1^-} Pr(F_1(X_1) \leq t | F_2(X_2) \leq t \ldots F_n(X_n) \leq t) \tag{6.2}$$

The multivariate distribution function is said to be upper tail dependent if $0 < \lambda_{up} \leq 1$ and upper (lower) tail independent if $\lambda_{up} = 0$ ($\lambda_{lo} = 0$). For example, in Fig. 6.1 (left), the upper tail coefficient is approximately zero ($\lambda_{up} \approx 0$), while for Fig. 6.1 (right) the upper tail coefficient is approximately 0.8 ($\lambda_{up} \approx 0.8$). For a more comprehensive discussion on the theoretical concept of tail independence, the interested reader is referred to Draisma et al. (2004) and Husler and Li (2009).

6.3 Copulas and Tail Dependence

The upper (lower) tail coefficient can also be defined using copulas. Copulas are joint cumulative distribution functions that describe dependencies among variables independent of their marginals (Joe 1997; Nelsen 2006):

$$C^n(u_1, \ldots, u_n) = F(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)) \tag{6.3}$$

where C^n is an n-dimensional joint cumulative distribution function (CDF) of a multivariate random variable ($U(U_1, \ldots, U_n)$) and whoses marginals are $u[0, 1]$. Equation 6.1 can be alternatively presented as:

$$\lambda_{up} = \lim_{u \to 1^-} Pr\left(X_1 > F_{X_1}^{-1}(u) | X_2 > F_{X_2}^{-1}(u) \ldots X_n > F_{X_n}^{-1}(u)\right) \tag{6.4}$$

where $F_1^{-1}, \ldots, F_n^{-1}$ are the inverse CDF of the random variables $X_1, \ldots, X_n$. Notice that the conditional probability, given in Eq. 6.4, can be described as:

$$Pr\,(X_1 > F_{X_1}^{-1}(u)|X_2 > F_{X_2}^{-1}(u) \ldots X_n > F_{X_n}^{-1}(u))$$

$$= \frac{1 - Pr(X_1 \leq F_{X_1}^{-1}(u)) - \ldots - Pr(X_n \leq F_{X_n}^{-1}(u)) + Pr(X_1 \leq F_{X_1}^{-1}(u), \ldots, X_n \leq F_{X_n}^{-1}(u))}{1 - Pr(X_2 \leq F_{X_2}^{-1}(u)) - \ldots - Pr(X_n \leq F_{X_n}^{-1}(u))}$$

$$(6.5)$$

Substituting Eq. 6.3 into Eq. 6.5 with some algebraic manipulation yields the following formulation for the upper tail (Joe 1997; Frahm et al. 2005):

$$\lambda_{up} = \lim_{u \to 1} \frac{1 - nu + C^{(n)}(u, \ldots, u)}{1 - (n-1)u} \tag{6.6}$$

Similarly, the lower tail dependence coefficient (λ_{lo}) can be expressed as (Joe 1997):

$$\lambda_{lo} = \lim_{u \to 1} \frac{C^{(n)}(u, \ldots, u)}{(n-1)u} \tag{6.7}$$

There are various copulas families, which have been developed for different purposes. One major difference between different copula families is the upper (lower) tail association they represent. For example, copula families may differ in the upper and lower tail of the distribution, where the dependence is strongest (weakest). In this study, two elliptical copulas, namely a normal copula and t-copula, as well as a non-Gaussian (v-transformed) copula, are used for simulations. In the following section, a number of copula families and their tail dependence behavior is discussed.

6.3.1 Gaussian Copula

One of the most commonly used copula families is the multivariate Gaussian (normal) copula, which is obtained from the multivariate normal distribution (Nelsen 2006):

$$C_\rho(u_1, \ldots, u_n) = F_\rho^n(F^{-1}(u_1), \ldots, F^{-1}(u_n)) \tag{6.8}$$

Equation 6.8 describes an n-dimensional multivariate Gaussian copula with correlation matrix $\rho_{n \times n}$ whose density function is:

$$c(u_1, \ldots, u_n) = \frac{1}{\sqrt{det\rho}} exp\left(-\frac{1}{2} y(u)'(\rho^{-1} - I)y(u)\right) \tag{6.9}$$

where: F^n = Multivariate Gaussian CDF

$\quad\quad\quad y(u_i) = F^{-1}(u_i)$

6.3.2 *t-Copula*

The t-copula (alternatively known as Student copula), is an elliptical copula derived from the Student distribution:

$$C_{v,\rho}(u_1,\dots,u_n) = t_{v,\rho}^n(t_v^{-1}(u_1),\dots,t_v^{-1}(u_n)) \tag{6.10}$$

where: t^n = Multivariate Student CDF

$\quad\quad\quad \rho$ = shape matrix and

$\quad\quad\quad v$ = degrees of freedom

$$t_{v,\rho}^n(x) = \frac{1}{\sqrt{det\rho}}\frac{\Gamma\left(\frac{v+n}{2}\right)}{\Gamma\left(\frac{v}{2}\right)(\pi v)^{n/2}} \times \int_{-\infty}^{x_1}\dots\int_{-\infty}^{x_n}\frac{dx}{\left(1+\frac{x'\rho^{-1}x}{v}\right)^{(v+n)/2}} \tag{6.11}$$

For $v > 2$, the shape matrix (ρ) is proportional to the correlation matrix (Malevergne and Sornette 2003). The density function of the t-copula is expressed as (Malevergne and Sornette 2003):

$$c(u_1,\dots,u_n) = \frac{1}{\sqrt{det\rho}}\frac{\Gamma\left(\frac{v+n}{2}\right)\left(\Gamma\left(\frac{v}{2}\right)\right)^{n-1}}{\left(\Gamma\left(\frac{v+1}{2}\right)\right)^n} \times \frac{\Pi_{k=1}^n\left(1+\frac{y_k^2}{v}\right)^{(v+1)/2}}{\left(1+\frac{y'\rho^{-1}y}{v}\right)^{(v+n)/2}} \tag{6.12}$$

where: y_k = $t_v^{-1}(u_k)$

$\quad\quad\quad t_v$ =   univariate Student distribution with v degrees of freedom

Both Gaussian and t-copulas are elliptical; however, they represent different tail dependencies. The Gaussian copula is upper (lower) tail independent ($\lambda_{up} \approx 0$) regardless of the correlation coefficient among variables (Coles 2001; Renard and Lang 2007; Mikosch and Resnick 2006). This indicates that the extreme values from the different random variables occur independently, even if the random variables exhibit a high correlation. It is worth pointing out that for independent variables, one could expect $\lambda_{up} = 0$. Note that the converse is not necessarily true, meaning that $\lambda_{up} = 0$ does not indicate that the random variables are necessarily independent (Malevergne and Sornette 2003).

Contrary to the Gaussian copula, the t-copula can capture the upper (lower) tail dependence (if exists) among two or more random variables. The t-copula can capture the asymptotic dependence even when the variables are negatively (inversely) associated (Embrechts et al. 2001). In t-copula formulation, as v increases, the tail dependence weakens, and thus, the probability of occurrence of extreme values reduces. Figure 6.2a displays the tail behavior of the bivariate t-copula with $v = 1 - 10$. The Figure presents occurrences of $x > 0.8$ (percentage) in both
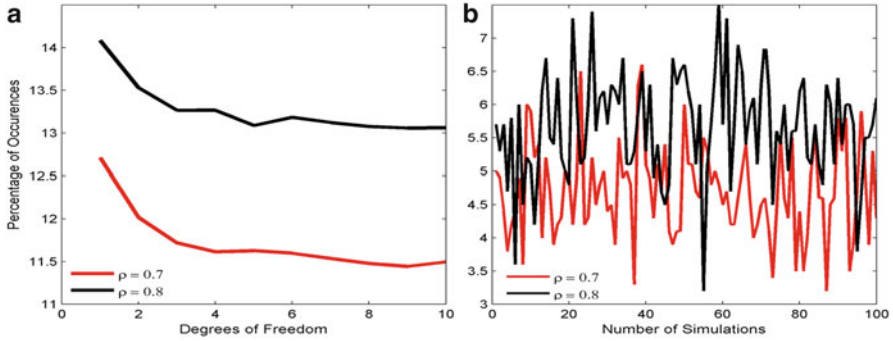
Fig. 6.2 Tail dependence behavior from bivariate random variable simulated using (**a**) t-copula and (**b**) Gaussian copula. The y-axes show the occurrences of joint extremes (percentage) in dependent random variables simulated using t-copula and Gaussian copula (modified after AghaKouchak et al. 2010b)

random vectors of the bivariate t-copula. One can see that an increase in ν results in less occurrences of extremes (values above the threshold t in Eq. 6.1). Figure 6.2b shows the tail dependence of Gaussian copula. The occurrence of joint extremes in Gaussian copulas is considerably less than t-copula (threshold: 0.8). This indicates that if strong dependence exist among multiple variables, using the Gaussian copula may not be suitable for modeling dependence of extremes. In fact, the multivariate Gaussian distribution is upper (lower) tail independent ($\lambda_{up} \approx 0$) meaning it cannot be used to describe dependencies of extremes (Coles 2001; Renard and Lang 2007). It is noted that the tail behavior of a multivariate model depends solely on the type of copula and not on the marginal distribution of individual variables. Therefore, in modeling the dependencies of extremes, the choice of copula family plays a significant role.

6.3.3 Gumbel-Hougaard Copula

In the following section, a heavy upper tailed Archimedean copula (Nelsen 2006), known as the Gumbel-Hougaard copula, is introduced. Unlike many copula families, Archimedean copulas are not derived from standard multivariate distributions. Generally, the multivariate Archimedean copulas can be expressed as:

$$C(x_1, \ldots, x_n) = \Psi^{-1}\left(\sum_{i=1}^{n} \Psi(F_i(x_i))\right) \qquad (6.13)$$

where Ψ is the so-called generator function. For the Gumbel-Hougaard copula, the generator function can be expressed as:

$$\Psi(x) = (-\ln(x))^{\theta} \qquad (6.14)$$

where: $\theta \geq 1$ = copula parameter.

By substituting Eq. 6.14 into Eq. 6.13, the general formulation of the bivariate Gumbel-Hougaard copula can be described as (Venter 2002):

$$C(u, v) = \exp\left(-\left((-\ln u)^\theta + (-\ln v)^\theta\right)^{\frac{1}{\theta}}\right)$$

(6.15)

Equation 6.15 represents the bivariate Gumbel-Hougaard copula with variables $u[0, 1]$ and $v[0, 1]$. The Gumbel-Hougaard copula is parameterized through a single parameter θ. The copula parameter θ is to be estimated based on available data. By substituting Eq. 6.15 into Eq. 6.6, the upper tail dependence coefficient for the Gumbel-Hougaard copula can be derived as: $\lambda_{up} = 2 - 2^\theta$ (Salvadori et al. 2007; Frahm et al. 2005; Nelsen 2006). A discussion on copula parameter estimation techniques is beyond the scope of this chapter. The interested reader is pointed to Genest et al. (1995), Salvadori et al. (2007) and Nelsen (2006) for more detailed discussions on parameter estimation.

6.4   Nonparametric Tail Dependence Methods

There are different nonparametric tail dependence estimators that can be used to evaluate the significance of tail behavior. The first nonparametric approach introduced here is based on the concept of the empirical copula (C^m) – AghaKouchak et al. (2010c):

$$C^m(u, v) = F_{(m)}\left(F_{(m)1}^{-1}(u), F_{(m)2}^{-1}(v)\right)$$

(6.16)

where $F_{(m)}$ refers to the empirical distribution of random variables. The tail dependence estimator $\lambda_{up}^{(1)}$ is then expressed as (Schmidt and Stadtmüller 2006):

$$\lambda_{up}^{(1)} = \frac{T}{k} C^m\left(\left(1 - \frac{k}{T}, 1\right] \times \left(1 - \frac{k}{T}, 1\right]\right)$$

(6.17)

$$= \frac{1}{k}\Sigma_{j=1}^T I(R_1^j > T - k, R_2^j > T - k)$$

where: T =   sample size
k =   threshold rank
R_1^j =   rank of u^j
R_2^j =   rank of v^j
I =   indicator function

It is worth pointing out that Eq. 6.17 is the empirical copula with the interval $(1 - \frac{k}{T}, 1] \times (1 - \frac{k}{T}, 1]$. That is, $\lambda_{up}^{(1)}$ is derived using the empirical tail-copula introduced by Genest et al. (1995).

Haung (1992) suggested another tail dependence measure (here, $\lambda_{up}^{(2)}$) based on empirical copulas and extreme value theory:

$$\lambda_{up}^{(2)} = 2 - \frac{T}{k}\left(1 - C^m\left(1 - \frac{k}{T}, 1 - \frac{k}{T}\right)\right)$$

$$= 2 - \frac{1}{k}\Sigma_{j=1}^T I(R_1^j > T - k \text{ OR } R_2^j > T - k) \qquad (6.18)$$

Coles et al. (1999) proposed a different nonparametric tail dependence measure ($\lambda_{up}^{(3)}$) as follows (Frahm et al. 2005):

$$\lambda_{up}^{(3)} = 2 - \frac{\log C^m\left(\frac{T-k}{T}, \frac{T-k}{T}\right)}{\log\left(\frac{T-k}{T}\right)} \qquad (6.19)$$

where

$$C^m(u, v) = \frac{1}{T}\Sigma_{j=1}^T I\left(\frac{R_1^j}{T} \leq u, \frac{R_2^j}{T} \leq v\right) \qquad (6.20)$$

Another nonparametric tail dependence estimator (here, $\lambda_{up}^{(4)}$) is proposed by Joe et al. (1992):

$$\lambda_{up}^{(4)} = 2 - \frac{1 - C^m\left(\frac{T-k}{T}, \frac{T-k}{T}\right)}{1 - \frac{T-k}{T}} \qquad (6.21)$$

where the term C^m is the empirical copula as described in Eq. 6.20. It should be noted that nonparametric methods of estimating tail dependence is not limited to the ones mentioned above (see for example Capéraa et al. 1997).

6.5 Extreme Value Threshold

Estimation of the extreme value threshold requires assuming a threshold above (or below) values that are considered as extreme (Frahm et al. 2005). For tail dependence analysis, one can use a fixed (e.g., above 95 % of data) or variant threshold approach. The so-called optimal threshold approach (Frahm et al. 2005; Peng 1998) uses a kernel plateau-finding algorithm to estimate the optimal extreme value threshold. In this method, the optimal plateau is estimated in four steps: (1) a kernel box with a bandwidth of b (e.g., $b = int(0.05n)$) is selected; (2) the mean values of the coefficients that fall within each box results in $n - 2b$ $\bar{\lambda}$ values; (3) for a moving plateau with a length of $l = \sqrt{n - 2b}$, the corresponding λ values are calculated ($\lambda_k, \ldots, \lambda_{k+l+1}$ where $k=1,\ldots, n-2b-m+1$); (IV) the optimal plateau (extreme value threshold) is the first one that fulfills the following condition (for more detailed description, the reader is referred to Frahm et al. (2005) and Peng (1998)):

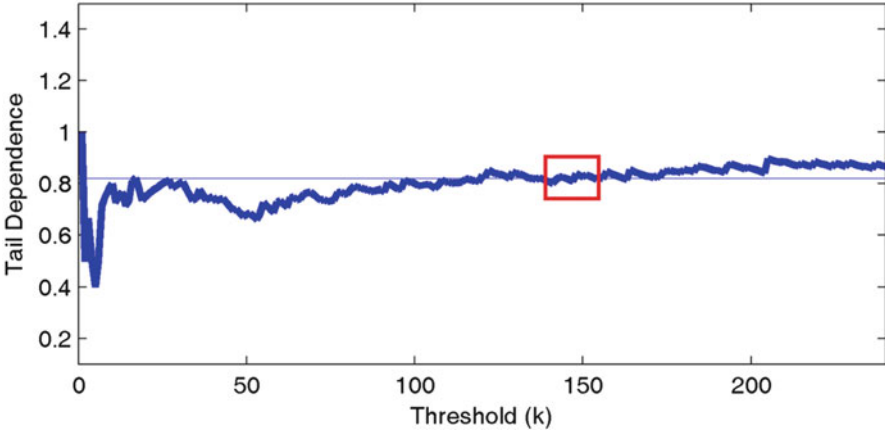$$\Sigma_{i=k+1}^{k+l-1}|\bar{\lambda}_i - \lambda_k| \leq 2\sigma \qquad (6.22)$$

Fig. 6.3 Variability of the tail dependence coefficient with respect to threshold (After AghaKouchak et al. (2010c))

where σ is the standard deviation of the $\overline{\lambda}_i$ values (means of coefficients that fall within each box). The optimal tail dependence coefficient is then expressed as:

$$\lambda = \frac{1}{T}\Sigma_{i=1}^{T}\lambda_{k+i-1} \tag{6.23}$$

Figure 6.3 displays an example of tail dependence coefficient variability versus the choice of extreme value threshold. In this figure, the box refers to the plateau that satisfies the condition mentioned above (Eq. 6.22) and its corresponding TDC. Note that the box size in Fig. 6.3 is not scaled, and the box size is placed on the figure for illustration. For other methods of extreme value threshold estimation, the interested reader is pointed to Tancredi et al. (2006).

6.6 Case Studies

Case Study A: In the following example, the tail dependence coefficient is used for analysis of anisotropy of spatial dependencies of extremes. Figure 6.4 displays the rainfall accumulations above 95 % threshold normalized to [0-1]. The precipitation data used in this example is from the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN, Sorooshian et al. 2000; Hsu et al. 1997) data set, which is an infrared-based microwave-adjusted precipitation product. The tail dependence coefficient is estimated using Eq. 6.17 ($\lambda_{up}^{(1)}$) for the two perpendicular directions shown in Fig. 6.4. In this example, the data is smoothed with a moving-average window with bandwidth of 2 pixels and 80 % overlap. Figure 6.5 indicates that heavy precipitation rates are dependent over
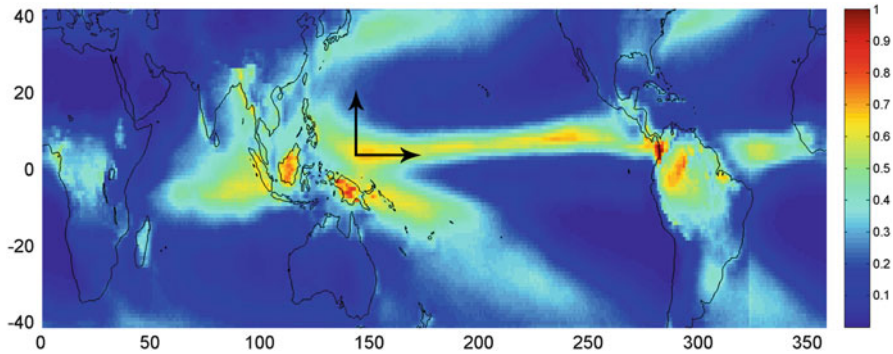
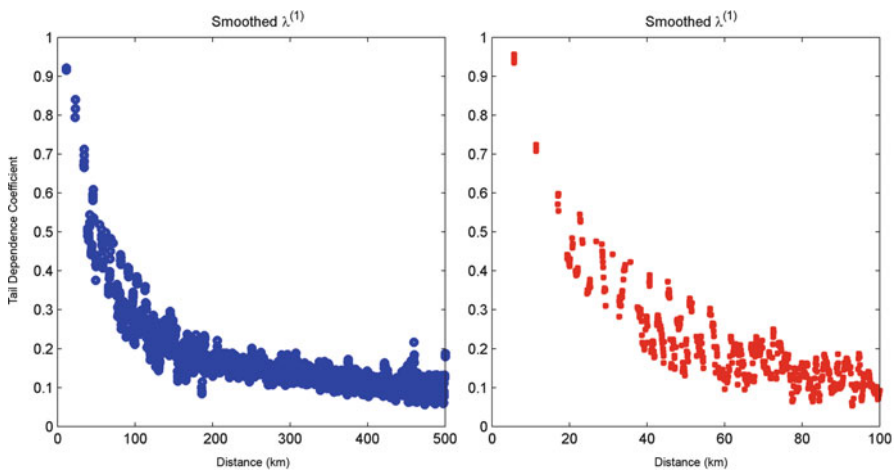**Fig. 6.4** Sum of rainfall exceedance above 95 % normalized to [0-1]

Fig. 6.5 Tail dependence behavior of rainfall data in (*left*) E-W, and (*right*) N-S directions versus distance (km)

longer distances in the horizontal direction (see Fig. 6.5 (left)) as opposed to the vertical direction (see Fig. 6.5 (right)). This indicates that the spatial dependence structure of heavy rainfall rates is asymmetrical, and in this example, heavy rainfall rates are spatially more dependent in the horizontal direction.

Case Study B: Many earth science variables provide excellent data for studying spatial dependencies of extreme events. This example demonstrates a nonparametric approach to evaluating the dependence structure of the extreme precipitation values over a region in the southern part of the United States. Understanding extreme precipitation spatial dependencies and behavior on the local, regional and global scale will provide enhanced insight in the spatial dependence structure of precipitation in different regions of the world. This information can then be used to assist in planning and decision making purposes.
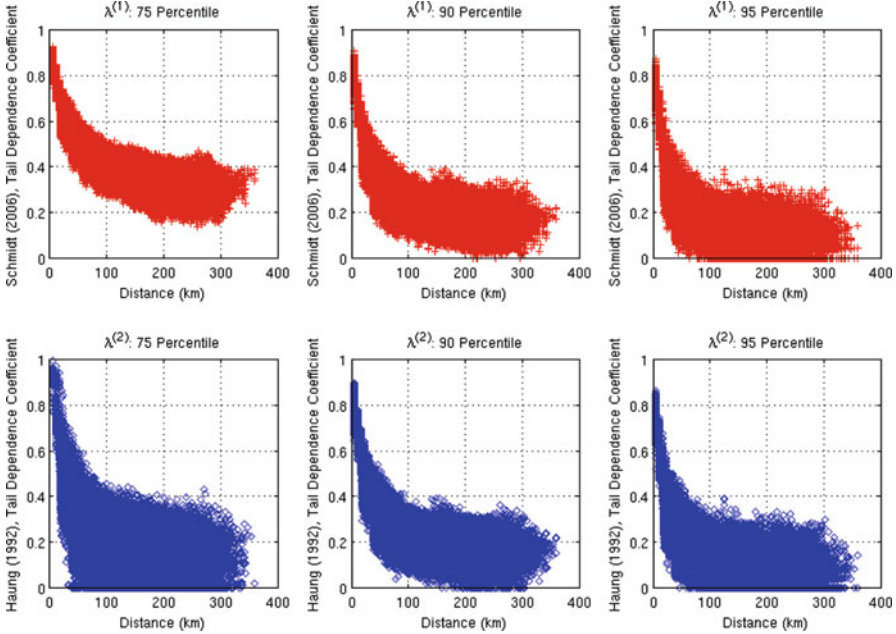
Fig. 6.6 Two nonparametric methods for calculating tail dependence coefficients. (*Top*) $\lambda_{up}^{(1)}$ for three percentile groups, 75th, 90th and 95th as described in Eq. 6.17; and (*Bottom*) $\lambda_{up}^{(2)}$ for three percentile groups, 75th, 90th and 95th as described by 6.18

Two nonparametric tail dependence methods based on empirical copulas are used to derive tail dependence estimators: (1) $\lambda_{up}^{(1)}$ introduced in Eq. 6.17; and (2) $\lambda_{up}^{(2)}$ introduced in Eq. 6.18. In both cases, the tail dependence estimator helps describe the dependent structure or degree of association between concurrent rainfall extremes at different locations. High spatial and temporal resolution precipitation data can be analyzed using these nonparametric tail dependence methods, which allows for solving for the tail dependence coefficient and thus describing the dependence structure of the extreme precipitation events.

The study region is over Mississipp, which is located in the southern part of the United States with a latitude of 38N to 35.5N and longitude of -110W to -107.5W from January 1st, 2005 to December 31st, 2008. The precipitation data used in this example is the National Center for Environmental Prediction (NCEP)'s Stage 4 mosaic multi-sensor national precipitation analysis, which has a 4 km spatial resolution and hourly temporal resolution (Lin and Mitchell 2005).

Solving for $\lambda_{up}^{(1)}$ (Eq. 6.17) and $\lambda_{up}^{(2)}$ introduced in Eq. 6.18, and smoothing the results for display purposes, one can see the dependence structure of precipitation for the data in this region. Similar to the previous example, the data is smoothed using a moving-average window with a bandwidth of 2 pixels and 80 % overlap. Figure 6.6 shows three different percentile groupings of the extreme precipitation events: 75th percentile, 90th precentile and 95th precentile for $\lambda_{up}^{(1)}$ and $\lambda_{up}^{(2)}$
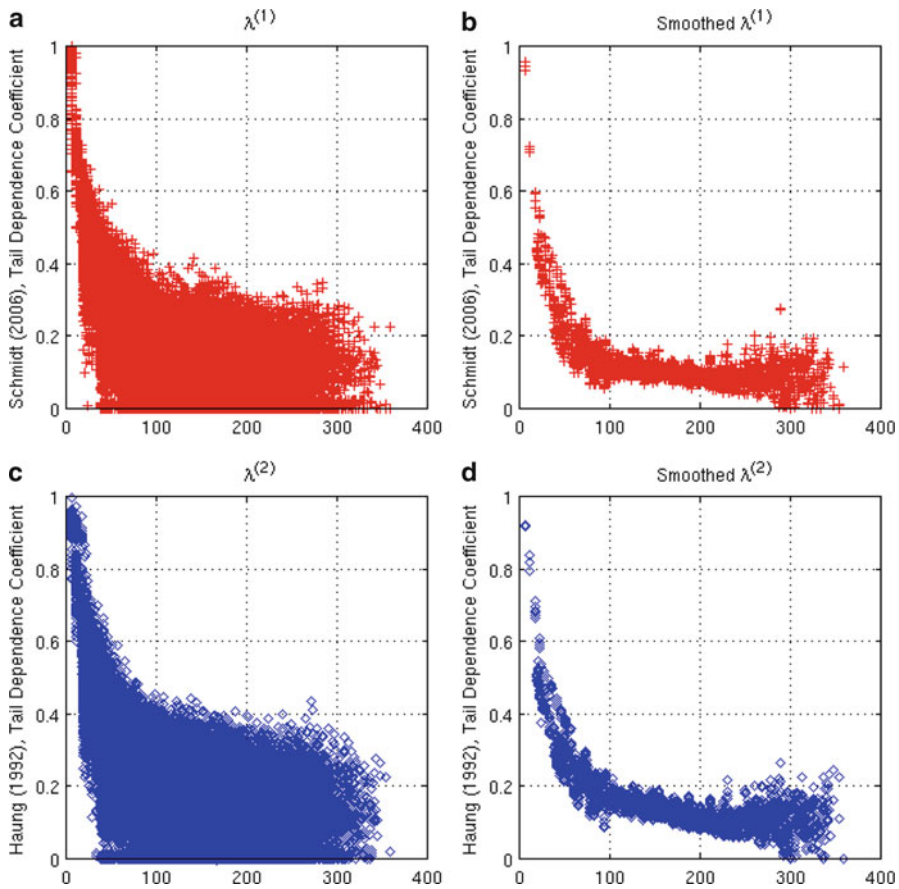
Fig. 6.7 Tail dependence analysis using $\lambda_{up}^{(1)}$ and $\lambda_{up}^{(2)}$ and based on the concept of optimal threshold introduced in Eqs. 6.22 and 6.23

tail dependence methods. Both nonparametric tail dependence methods from $\lambda_{up}^{(1)}$ (Fig. 6.6, top) and $\lambda_{up}^{(2)}$ (Fig. 6.6, bottom) are consistent with each other showing the expected decrease in dependence with distance across each of the different precentile groups. The figure indicates that the spatial dependence of extreme convective precipitation will decrease rapidly as distance increases. This is expected because with extreme precipitation events, spatial dependence is typically the highest near the region of convective activity, which produces the largest observed precipitation. This is not always the case, for example, extra-tropical cyclones can also produce extreme precipitation events and can have spatial dependence up to a few hundred kilometers.

Figure 6.7 displays $\lambda_{up}^{(1)}$ (Fig. 6.7, top) and $\lambda_{up}^{(2)}$ (Fig. 6.7, bottom) methods using the concept of optimal threshold (Eqs. 6.22 and 6.23) to determine the percentile for calculating the tail dependence coefficient as well as a smoothed version for

illustration purposes. Contrary to commonly used method (Fig. 6.6), this approach of tail dependence analysis is independent of a fixed (constant) threshold. In other words, this method provides a tail dependence analysis that does not require additional decisions regarding the choice of extreme value threshold (Frahm et al. 2005).

6.7 Summary and Conclusions

Extreme events (e.g., floods, droughts, heat waves) have varying spacial dependence structures across different geographic locations and the understanding of these dependencies is fundamental to risk assessment and decision making. Understanding the different characteristics of extreme events, including spatial dependence, will provide regional planners and policy makers with information and knowledge of extreme events that impact their the local and regional communities. Spatial characteristics of extreme events can be investigated through estimation of tail dependence coefficient for different locations. This chapter reviewed several nonparametric and parametric tail dependence coefficient estimators. The tail dependence coefficient describes the degree of association between concurrent extremes. The presented nonparametric methods are based on the concept of bivariate empirical copula of random variables, whereas the parametric approach is based on the concept of Gumbel-Hougaard Copula. The chapter also reviewed different aspects of modeling tail dependence such as the choice of extreme value threshold.

In the first case study, the tail dependence coefficient is used for analysis of anisotropy of spatial dependencies of extremes. The results showed that the spatial dependence structure of heavy rainfall rates was asymmetrical. In the second example, the tail dependence coefficient is used to investigate spatial dependencies of precipitation extremes on a local scale revealing the spatial dependence structure of the extreme convective precipitation as described by the tail dependence coefficient. Extreme precipitation impacts many aspects of human society, such as loss of property and life due to flooding and area destruction from severe storms.

In the case studies, a kernel plateau-finding algorithm is used to obtain tail dependence coefficients, avoiding a fixed extreme value threshold. The results of previous studies (e.g., AghaKouchak et al. 2010c) reveal that using the kernel plateau-finding algorithm for tail dependence is superior to the fixed threshold approaches. This method, also known as the optimal threshold approach, can obtain a measure of tail dependence that does not require additional decisions regarding the choice of extreme value threshold.

The tail dependence coefficient has numerous applications including: validation and verification of weather and climate models in reproducing extreme events; analysis of simultaneous extremes; probabilistic assessment of occurrences of extremes, and understanding climate variability. For example, by deriving tail dependence coefficients for simulations of a numerical weather prediction model or a climate model, one can evaluate whether these models produces dependencies as seen in the observations. These approaches are not limited to precipitation, but also

a wide variety of earth science variables. This study of extremes tail dependence on local, regional and global scales can assist in planning and policy making as well as validating numerical models, thus providing a valuable tool for understanding how extreme events impact society.

References

Abberger K (2005) A simple graphical method to explore tail-dependence in stock-return pairs. Appl Financ Econ Taylor Fr J 15(1):43–51

AghaKouchak A, Bárdossy A, Habib E (2010a) Conditional simulation of remotely sensed rainfall data using a non-gaussian v-transformed copula. Adv Water Resour 33(6):624–634

AghaKouchak A, Bárdossy A, Habib E (2010b) Copula-based uncertainty modeling: application to multi-sensor precipitation estimates. Hydrol Process 24(15):2111–2124

AghaKouchak A, Ciach G, Habib E (2010c) Estimation of tail dependence coefficient in rainfall accumulation fields. Adv Water Resour 33(9):1142–1149

Bárdossy A, Li J (2008) Geostatistical interpolation using copulas. Water Resour Res 44(7):W07412.1–W07412.15

Capéraa P, Fougéres A-L, Genest C (1997) A nonparametric estimation procedure for bivariate extreme value copulas. Biometrika 84(3):567–577

Coles S (2001) An introduction to statistical modeling of extreme values. Springer, London

Coles S, Heffernan J, Tawn J (1999) Dependence measures for extreme value analyses. Extremes 2:339–365

De Michele C, Salvadori J (2003) A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas. J Geophy Res Atmos 108(D2):1–11

De Michele C, Salvadori J, Canossi M, Petaccia A, Rosso R (2003) Bivariate statistical approach to check adequacy of dam spillway. J Hydrol Eng 10(1):50–57

Draisma G, Drees H, Ferreira A, de Haan L (2004) Bivariate tail estimation: dependence in asymptotic independence. Bernoulli 10:251–280

Embrechts P, Lindskog F, McNeil A (2001) Modelling dependence with copulas and applications to risk management. Department of Mathematics, ETHZ, Zurich, Switzerland

Embrechts P, Mcneil A, Straumann D (2002) Correlation and dependence in risk management: properties and pitfalls. In: Risk management: value at risk and beyond. Cambridge University Press, Cambridge, pp 176–223

Favre A-C, El Adlouni S, Perreault L, Thiémonge N, Bobée B (2004) Multivariate hydrological frequency analysis using copulas. Water Resour Res 40:1–12

Fisher N, Switzer P (2001) Graphical assessment of dependence: is a picture worth 100 tests? Am Stat 55:233–239

Fisher R, Tippett L (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proc Camb Philos Soc 24:180–190

Frahm G, Junker M, Schmidt R (2005) Estimating the tail-dependence coefficient: properties and pitfalls. Insur Math Econ 37(1):80–100

Genest C, Ghoudi K, Rivest L-P (1995) A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. Biometrika 82:543–552

Gumbel E (1958) Statistics of extremes. Columbia University Press, New York, p 23

Haung X (1992) Statistics of bivariate extreme values. Thesis Publishers, Amsterdam

Hsu K, Gao X, Sorooshian S, Gupta H (1997) Precipitation estimation from remotely sensed information using artificial neural networks. J Appl Meteorol 36:1176–1190

Husler J, Li D (2009) Testing asymptotic independence in bivariate extreme. J Stat Plan Inference 139(3):990–998

Joe H (1997) Multivariate models and dependence concepts. Chapman Hall, London

Joe H, Smith R, Weissman I (1992) Bivariate threshold models for extremes. J R Stat Soc Ser B 54:171–183

Kelly k, Krzysztofowicz R (1997) A bivariate meta-gaussian density for use in hydrology. Stoch Environ Res Risk Assess 11(1):17–31

Kendall M (1962) Rank correlation methods. Griffin, London

Kotz S, Nadarajah S (2000) Extreme value distributions: theory and applications. Imperial College Press, London. ISBN 1860942245

Ledford A, Tawn J (1996) Statistics for near independence in multivariate extreme values. Biometrika 83(1):169–187

Ledford A, Tawn J (1997). Modelling dependence within joint tail regions. J R Stat Soc Ser B-Methodol 59(2):475–499

Ledford A, Tawn J (2003) Diagnostics for dependence within time series extremes. J R Stat Soc Ser B-Stat Methodol 65(Part 2):521–543

Lin Y, Mitchell K (2005) The ncep stage ii/iv hourly precipitation analyses: development and applications. In: Preprints, 19th conference on hydrology, American Meteorological Society, San Diego, CA, 9–13 Jan 2005, Paper 1.2

Malevergne Y, Sornette D (2002) Minimising extremes. Risk 15(11):129–133

Malevergne Y, Sornette D (2003) Testing the gaussian copula hypothesis for financial assets dependences. Quant Financ Vol 3:231–250

Malevergne Y, Sornette D (2004) How to account for extreme co-movements between individual stocks and the market. Risk 6:71–116

Melchiori MR (2003) Which Archimedean Copula is the Right One? Yield Curve 37:1–20

Mikosch T, Resnick S (2006). Activity rates with very heavy tails. Stoch Process Their Appl 116(2):131–155

Nazemi A, Elshorbagy A (2011) Application of copula modelling to the performance assessment of reconstructed watersheds. Stoch Environ Res Risk Assess 1–17. doi: 10.1007/s00477-011-0467-7

Nelsen R (2006) An introduction to copulas. Springer series in statistics. Springer, New York

Peng L (1998) Second order condition and extreme value theory. Tinbergen Institute Research Series, vol 178. Thesis Publishers and Tinbergen Institute, Amsterdam

Poon S, Rockinger M, Tawn J (2004) Extreme value dependence in financial markets: diagnostics, models, and financial implications. Rev Financ Stud 17(2):581–610

Renard B, Lang M (2007) Use of a gaussian copula for multivariate extreme value analysis: some case studies in hydrology. Adv Water Resour 30:897–912

Salvadori G, DeMichele C, Kottegoda N, Rosso R (2007) Extremes in Nature: an approach using copulas. Springers, Berlin

Schmidt R (2005) Tail dependence. In: Cizek P, Häardle W, Weron R (eds) Statistical tools in finance and insurance. Springer, New York

Schmidt R, Stadtmüller U (2006) Non-parametric estimation of tail dependence. Scand J Stat 33(2):307–335

Serinaldi F (2008) Analysis of inter-gauge dependence by kendall's τ, upper tail dependence coefficient, and 2-copulas with application to rainfall fields. Stoch Environ Res Risk Assess 22:671–688

Serinaldi F (2009) Copula-based mixed models for bivariate rainfall data: an empirical study in regression perspective. Stoch Environ Res Risk Assess 23(5):677–693

Sibuya M (1959) Bivariate extreme statistics I. Ann Inst Stat Math 11(2):195–210

Sorooshian S, Hsu K, Gao X, Gupta H, Imam B, Braithwaite D (2000) Evolution of the persiann system satellite-based estimates of tropical rainfall. Bull Am Meteorol Soc 81(9):2035–2046

Spearman C (1904) The proof and measurement of association between two things. Am J Psychol 15:72–101

Tancredi A, Anderson C, O'Hagan A (2006) Accounting for threshold uncertainty in extreme value estimation. Extremes 9:87–106

Venter G (2002) Tails of copulas. In: Proceedings of the casualty actuarial society, LXXXIX 2002, Boston, MA, p 68–113

Zhang LSR, Singh V (2008) Bivariate rainfall frequency distributions using archimedean copulas. J Hydrol 32(1–2):93–109