

Lawrence Berkeley National Laboratory

LBL Publications

Title

A nonparametric method for producing isolines of bivariate exceedance probabilities

Permalink

<https://escholarship.org/uc/item/00x0m2q9>

Journal

Extremes, 22(3)

ISSN

1386-1999

Authors

Cooley, Daniel
Thibaud, Emeric
Castillo, Federico
[et al.](#)

Publication Date

2019-09-01

DOI

10.1007/s10687-019-00348-0

Peer reviewed

A Nonparametric Method for Producing Isolines of Bivariate Exceedance Probabilities

Daniel Cooley¹, Emeric Thibaud²

Federico Castillo³, Michael F. Wehner⁴

¹Department of Statistics, Colorado State University

²Institute of Mathematics, Ecole Polytechnique Fédérale de Lausanne

³Department of Environmental Science, Policy and Management,
University of California, Berkeley

⁴Lawrence Berkeley National Laboratory

April 25, 2019

Abstract

We present a method for drawing isolines indicating regions of equal joint exceedance probability for bivariate data. The method relies on bivariate regular variation, a dependence framework widely used for extremes. The method we utilize for characterizing dependence in the tail is largely nonparametric. The extremes framework enables drawing isolines corresponding to very low exceedance probabilities and may even lie beyond the range of the data; such cases would be problematic for standard nonparametric methods. Furthermore, we extend this method to the case of asymptotic independence and propose a procedure which smooths the transition from hidden regular variation in the interior to the first-order behavior on the axes. We propose a diagnostic plot for assessing the isoline estimate and choice of smoothing, and a bootstrap procedure to visually assess uncertainty.

Keywords: Extreme Values, Multivariate, Asymptotic Independence, Regular Variation, Hidden Regular Variation.

1 Introduction

We develop a tool which will draw isolines to indicate regions of equal joint exceedance probability for bivariate data. By displaying these regions of low probability, researchers can visually assess probabilistic risk of rare bivariate extreme events. Importantly, impactful

events can arise when the combination of variables is rare even if the individual variates are not at their highest values. We employ results from multivariate extreme value (EV) theory which provide a framework for characterizing dependence in the tail of the distribution. Although our method is largely nonparametric, we are able to extrapolate to describe events more extreme than any observed in the data record.

In Figure 1 we present two motivating data sets which we will examine in this work. Details about the data are given in Sections 3 and 4. The left panel shows data related to a southern California weather regime known as the Santa Ana winds, a windy and dry weather regime conducive for wildfires. The points labeled “C” and “W” correspond to the ignition days of the Cedar and Witch Fires respectively, both of which were among the most destructive Santa Ana driven wildfires on record. The right panel shows daily temperature measurements and relative humidity measurements for Karachi, Pakistan. Here, risk is in terms of human health impacts which worsen by simultaneous hot and humid conditions. Shown in black are six successive days in June 2015 which correspond to a heat wave which is blamed for the deaths of more than 700 people (Masood et al., 2015).

Rather than using a bivariate approach, standard practice to model risk of fire conditions or heat waves is to consider univariate statistics of some combined measure of the relevant individual variables such as a burn index or a human health index. However, both examples in Figure 1 show a complex relationship between the two meteorological variables which a combined variable cannot fully capture. Instead of relying on the aforementioned indices, an understanding of bivariate extreme behavior could improve response to the crisis by allocating resources in a more efficient manner. To aid in understanding bivariate extreme behavior, we would like to draw lines to indicate how frequently events this extreme, or even more extreme, can be expected to occur.

In the univariate case there is a one-to-one correspondence between probabilities and exceedance regions (which may be respectively expressed in terms of “return periods” and “return levels” in environmental sciences (Cooley, 2013)). Given the (likely estimated) upper

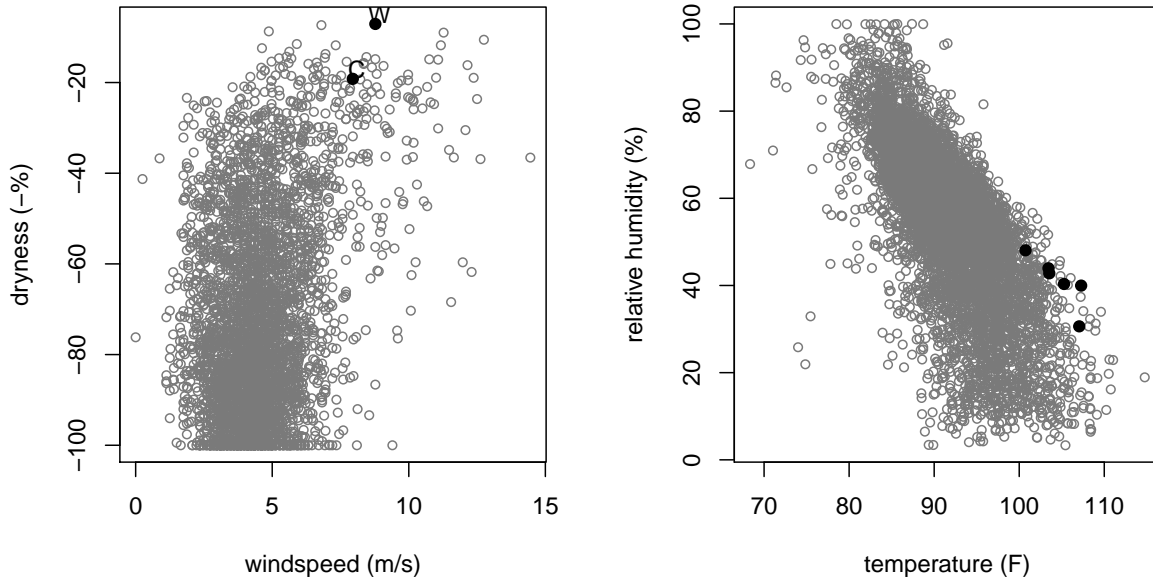


Figure 1: Left panel: windspeed and dryness from the Santa Ana dataset. Point labeled “C” corresponds to the date of the Cedar Fire, and point labeled “W” corresponds to that of the Witch Fire. Right panel: temperature and relative humidity from the Karachi data set. Dark solid circles correspond to the dates 06/18/2015-06/23/2015.

tail of a univariate distribution, one can begin with a small probability of interest and determine a threshold corresponding to the desired exceedance probability, or conversely begin with a particular high value of interest and determine the probability exceeding this value. In the bivariate case, this one-to-one relationship no longer exists. Given a risk region; that is, a region defined in terms of a specific bivariate extreme event occurring, EV methods have been devised to estimate the probability of such an event (e.g., [de Haan and de Ronde, 1998](#)). However in bivariate space, an exceedance region is not uniquely specified for a given probability.

A familiar way to visually describe bivariate data is to draw contour lines corresponding to equal values of an estimated density function. Typical methods will yield equidensity contours which form closed regions in \mathbb{R}^2 . However, equidensity contours may not be ideal for describing exceedance probabilities. First, the contours’ values correspond to density

values rather than exceedance probabilities of the contour. Calculating associated exceedance probabilities would require integration of the density function over an oddly-shaped region. More importantly, in most EV applications there is a direction of interest which is associated with impactful events. We will assume that our data are oriented such that we are concerned when the variates take on their greatest values. Notice that the data in Figure 1 reflect this orientation, and in particular the fire risk application has a “dryness” variable which was constructed by negating nighttime humidity measurements. An equidensity contour has no directional orientation.

Rather than equidensity contours, our tool will produce isolines such that the estimated survival probability of any point on the isoline is equal. That is for some exceedance probability of interest p , if $\mathbf{X} = (X_1, X_2)^T$ takes values in \mathbb{R}^2 , we seek to estimate $\ell_{\mathbf{X}}(p) := \{\mathbf{x} \in \mathbb{R}^2 : \bar{F}_{\mathbf{X}}(\mathbf{x}) = p\}$ where $\bar{F}_{\mathbf{X}}(\mathbf{x}) = P(\mathbf{X} > \mathbf{x}) = P(X_1 > x_1, X_2 > x_2)$. By defining the isoline in terms of the survival function, we orient the exceedance region in the direction of interest, and tie the line directly to the specific notion of “exceedance” given by the survival function. Others have used isolines associated with probabilities of bivariate distributions. [Salvadori and De Michele \(2004\)](#) and [Marcon et al. \(2017\)](#) draw isolines of extreme regions defined in terms of the survival function and additionally in terms of the cumulative distribution function. The function `qcbvnonpar` in the `evd` package ([Stephenson, 2002](#)) in R draws isolines associated with the bivariate cumulative distribution.

Our work relies on a dependence framework familiar to extremes, and is novel in that it is largely nonparametric. Specifically, it begins with a nonparametric estimate of an isoline at a very high level, and then uses EV results to project to more extreme levels. In contrast, [Salvadori and De Michele \(2004\)](#) use parametric copula models and both [Marcon et al. \(2017\)](#) and the `qcbvnonpar` function ([Stephenson, 2002](#)) employ a semiparametric descriptor of bivariate extremal dependence. Also importantly, we adapt our approach to draw isolines when data are determined to be asymptotically independent. Asymptotic independence, described more completely in Section 2, is not captured by standard EV dependence

frameworks, but data are frequently determined to exhibit asymptotic independence (e.g., [Huser and Wadsworth \(2018\)](#); [Das and Resnick \(2014\)](#); [Heffernan and Tawn \(2004\)](#)). To our knowledge, all previous EV-based work to draw lines characterizing bivariate extreme behavior has assumed asymptotic dependence. This includes the previously-cited work, and additionally [Cai et al. \(2011\)](#); [Einmahl et al. \(2009\)](#); [Coles and Tawn \(1994\)](#).

2 Mathematical background for approach

As we wish to draw isolines at the utmost extent of the data and beyond, we must characterize dependence for the distribution’s upper tail. EV methods typically analyze only a small extreme subset of the available data and assume these largest values are well approximated by an asymptotically-justified model. One approach is to use a subset of componentwise block (e.g., annual) maxima for which the limiting distributions are the class of multivariate extreme value distributions (MVEVDs). Because we wish to visualize isolines of the original data such as the daily data pictured in [Figure 1](#), rather than obtaining componentwise block maxima, we will use the largest values of the original data.

Our method relies on the framework of regular variation to characterize the dependence in the tail of the distribution. Informally, a bivariate regularly varying random vector is one whose joint distribution has a heavy tail, implying that the tail decays like a power function. Because the definition, given below, only describes behavior in the joint tail, and because only extreme data are used for inference, the data from the distribution’s bulk does not influence inference. More importantly, the fundamental dependence structure of multivariate regular variation can be directly linked to that of the MVEVDs ([Resnick, 1987](#), Section 5.4.2), justifying its use for extremes.

Formally, a nonnegative bivariate random vector \mathbf{Z} is regularly varying if there exists a renormalizing sequence $b_n \rightarrow \infty$ as $n \rightarrow \infty$, and a Radon measure ν on the space $\mathcal{C} =$

$[0, \infty]^2 \setminus \mathbf{0}$, such that as $n \rightarrow \infty$

$$nP\left(\frac{\mathbf{Z}}{b_n} \in A\right) \rightarrow \nu(A), \quad (1)$$

for any ν -continuity set $A \subset \mathcal{C}$. The normalizing sequence b_n is regularly varying with extreme value index $\xi > 0$; that is $b_n = n^\xi L(n)$ where $L(n)$ is a slowly varying function (Resnick, 2007). The limiting measure ν has the property such that

$$\nu(sA) = s^{-1/\xi} \nu(A), \quad (2)$$

for any scalar $s > 0$ and $A \subset \mathcal{C}$. We parametrize in terms of ξ , rather than the index of regular variation $\alpha = 1/\xi$, as readers may be more familiar with this parameter from other environmental extremes work. Larger values of ξ indicate heavier tails, and (2) is useful for extrapolating further into the tail.

Asymptotic (in)dependence is a notion that describes fundamental bivariate tail behavior. Let $\mathbf{X} = (X_1, X_2)^T$ be a bivariate vector (not necessarily regularly varying) with univariate marginal cumulative distribution functions F_{X_1} and F_{X_2} . Define

$$\chi = \lim_{u \rightarrow 1} P(F_{X_1}(X_1) > u \mid F_{X_2}(X_2) > u).$$

\mathbf{X} is deemed asymptotically independent if $\chi = 0$, and is deemed asymptotically dependent otherwise. Intuitively, asymptotic dependence implies that the two variates can obtain their largest values simultaneously.

A model will either be asymptotically dependent or independent, and it is essential for estimating joint tail probabilities that the selected model correctly captures the behavior exhibited by the data. Regular variation is a useful modeling framework for describing tail dependence under asymptotic dependence. Many multivariate models are asymptotically independent, including the Gaussian and familiar copulas like the Clayton and Frank models

(Nelsen, 2006, Section 2.4), and these will underestimate joint exceedance probabilities estimated by extrapolation into the tail if applied to data which are asymptotically dependent. However, asymptotic independence is a degenerate case for regular variation (as well as for the MVEVDs). If \mathbf{Z} is regularly varying and asymptotically independent, then for any set $A \subset \mathcal{C}$ which does not include a portion of the axes, $\nu(A) = 0$.

Ledford and Tawn (1996; 1997) were among the first to extend the regular variation framework to account for tail dependence in the asymptotically independent setting, and Resnick (2002) further formalized ideas via the concept of hidden regular variation. An intuitive explanation of asymptotic independence is that for sets $A \subset \mathcal{C}$ which do not include points on the axes, the renormalizing sequence $\{b_n\}$ in (1) grows too rapidly, and the resulting limit is 0. However, hidden regular variation obtains nontrivial convergence for such sets by using a lighter-tailed normalizing sequence $\{b_n^0\}$ with coefficient of tail dependence $0 < \eta < \xi$:

$$nP \left(\frac{\mathbf{Z}}{b_n^0} \in A \right) \rightarrow \nu_0(A). \quad (3)$$

The scaling property for sets A bounded away from the axes (i.e., $A \in \mathcal{C}$ such that $A \cap \{\mathbf{x} \in \mathcal{C} : \mathbf{x} = (x, 0) \text{ or } \mathbf{x} = (0, x) \text{ for } x > 0\} = \emptyset$) and scalar $s > 0$ is

$$\nu_0(sA) = s^{-1/\eta} \nu_0(A), \quad (4)$$

where decreasing η corresponds to weaker dependence. The scaling property (2) continues to hold for sets intersecting the axes. A model property of hidden regular variation is an abrupt transition between (1) and (3) for sets which include portions of the axes and sets which do not (c.f., Das and Resnick, 2014; Weller and Cooley, 2013).

The regular variation framework described above requires that each univariate marginal distribution be heavy-tailed with extreme value index ξ . However, when viewed as a copula, the dependence framework can be used to model data which are not heavy-tailed. Like copula modeling approaches and much extremal dependence modeling work, our approach

assumes a dependence framework after transformation to a convenient marginal. Marginal transformation can be defended by Proposition 5.10 of [Resnick \(1987\)](#) which states that the domain of attraction of a MVEVD is preserved under monotonic marginal transformation. This result can be interpreted as the extremes equivalent of Sklar’s theorem ([Nelsen, 2006](#)), which says dependence (as described by a copula) is not affected by monotonic marginal transformations. As [Ledford and Tawn \(1996\)](#), we choose to transform so that each marginal can be assumed to be regularly varying with extreme value index $\xi = 1$. In the asymptotic independent setting with $\xi = 1$, [Ledford and Tawn \(1996\)](#) classifies $1/2 < \eta < 1$ as positive association, $\eta = 1/2$ as near independence, and $0 < \eta < 1/2$ as negative association.

3 Asymptotic dependence case

3.1 Procedure

Let $\mathbf{X} = (X_1, X_2)^T$ be the random vector for which we wish to estimate $\ell_{\mathbf{X}}(p)$, and let $\mathbf{x}_t, t = 1, \dots, n$, be identically-distributed realizations. The first step in the procedure is to nonparametrically construct a “base” isoline. Let p_{base} be the selected exceedance probability for this base isoline; p_{base} should be small enough such that extreme dependence is well represented, and yet should be large enough for the nonparametric procedure to have adequate data. Our method for constructing the base isoline begins with a Gaussian-kernel based estimate of the cumulative distribution function: $\hat{F}(\mathbf{x}) = n^{-1} \sum_{t=1}^n \Phi_h(\mathbf{x} - \mathbf{x}_t)$, where Φ_h is the Gaussian cdf which assumes independence and whose variance is controlled by a bandwidth parameter h ([Liu and Yang, 2008](#)). Kernel bandwidth can either be specified by the user, or one can employ an automated bandwidth selection tool; throughout we use the `bandwidth.nrd` tool employed by the `kde2d` density estimation tool in R’s `MASS` library. The survival function’s value is estimated on a fine grid spanning the range of the data, and $\hat{F}_{\mathbf{X}}$ is monotonically decreasing by construction. The base isoline $\hat{\ell}_{\mathbf{X}}(p_{base}) = \{\mathbf{x} \in \mathbb{R}^2 : \hat{F}_{\mathbf{X}}(\mathbf{x}) = p_{base}\}$ can be drawn via standard interpolation methods.

Unless \mathbf{X} is regularly varying, marginal transformation is required to use (2) to project the estimated base isoline to more extreme levels. Let F_k denote the marginal distribution for X_k , $k = 1, 2$. Our marginal transformation procedure begins by constructing a linearly interpolated empirical cumulative distribution function $\hat{F}_k^{emp}(x)$, over the range of the data of each marginal. To allow extrapolation further into the tail, we additionally fit a generalized Pareto distribution $\hat{F}_k^{gpd}(x)$ above a high threshold $x_{thold,k}$ for each marginal. Since our isolines will take on both large and small values of each variate, we construct a smooth transition between the two marginal estimates. Define a weight function $w_k(x)$ where $w_k(x) = 0$ for $x \leq x_{thold,k}$, $w_k(x)$ is monotonically increasing from 0 to 1 in the range $x_{thold,k} < x < x_{thold+,k}$, and $w_k(x) = 1$ for $x > x_{thold+,k}$. Letting $\hat{F}_k(x) = (1 - w(x))\hat{F}_k^{emp}(x) + w(x)\hat{F}_k^{gpd}(x)$, we then construct a marginal transformation function

$$T_k(x) = -1/\log(\hat{F}_k(x)).$$

Thereby $\mathbf{Z} = T(\mathbf{X}) = (T_1(X_1), T_2(X_2))^T$ can be assumed to be regularly varying with $\xi = 1$. We employ the function $w_k(x) = (\sin(\pi(x - x_{k,thold})/(x_{k,thold+} - x_{k,thold}) - \pi/2) + 1)/2$. The monotonicity of $\hat{F}_k(x)$ is not guaranteed, but can be addressed by increasing the distance between $x_{thold+,k}$ and $x_{thold,k}$.

Projection of the base isoline to more extreme levels occurs in the transformed space. Let $\hat{\ell}_{\mathbf{Z}}(p_{base}) = T(\hat{\ell}_{\mathbf{X}}(p_{base}))$. Begin by assuming that the sample size n is fixed and large enough such that for any set $A \subset \mathcal{C}$, (1) holds approximately:

$$nP \left(\frac{\mathbf{Z}}{nL(n)} \in A \right) \approx \nu(A) \Rightarrow P(\mathbf{Z} \in A_*) \approx k\nu(A_*),$$

where $A_* = nL(n)A$, and $k = L(n)$ for the fixed n . A_* is understood to consist of large values, since A being bounded away from $\mathbf{0}$ and n being large implies $\|\mathbf{z}\|$ must be large for

all $\mathbf{z} \in A_*$. Thus, for large sets A_* and for $s > 1$, (2) implies

$$P(\mathbf{Z} \in sA_*) \approx s^{-1}P(\mathbf{Z} \in A_*). \quad (5)$$

By construction, $P(\mathbf{Z} \in [z, \infty)) \approx p_{base}$ for any $z \in \hat{\ell}_{\mathbf{Z}}(p_{base})$. From (5), setting $s = p_{base}/p_{proj}$ for any $p_{proj} < p_{base}$, then $P(\mathbf{Z} \in [sz, \infty)) \approx p_{proj}$ for any $z \in \hat{\ell}_{\mathbf{Z}}(p_{base})$. Thus on the transformed scale, we can construct $\hat{\ell}_{\mathbf{Z}}(p_{proj}) = s\hat{\ell}_{\mathbf{Z}}(p_{base})$.

To produce isolines on the original scale, simply reverse the transformation: $\hat{\ell}_{\mathbf{X}}(p_{proj}) = T^{-1}(\hat{\ell}_{\mathbf{Z}}(p_{proj}))$.

3.2 Simulation study: bivariate t -distribution

We let \mathbf{X} be a bivariate random vector with known distribution and compare selected points on the true isoline $\ell_{\mathbf{X}}(p_{proj})$ to their counterparts on the estimated isoline. Specifically, we find the value of one coordinate for a fixed value of the other coordinate. For the asymptotic dependent case, we let \mathbf{X} be a bivariate t -distribution with mean $\boldsymbol{\mu} = \mathbf{0}$, shape matrix Σ with diagonal elements of 1 and off-diagonal elements of 0.7, and 4 degrees of freedom. Although \mathbf{X} is known to be regularly varying with $\xi = 1/4$, we still estimate and transform the marginal distributions as described in Section 3.1, setting $x_{k,thold}$ and $x_{k,thold+}$ at the empirical 0.97 and 0.98 quantiles.

For each iteration, we simulate $n = 10000$ realizations from \mathbf{X} , and let $p_{proj} = 0.001$, and let $p_{base} = 0.01$. The selected true isoline points $(x_{i,1}, x_{i,2}) \in \ell_{\mathbf{X}}(p_{proj})$ appear in the top two rows of Table 1 below, with the coordinate being estimated in bold. For each simulated data set, our procedure is used to produce the isoline estimate $\hat{\ell}_{\mathbf{X}}(p_{proj})$, and the value of the estimated coordinate $\hat{x}_{i,*}$ is found at the locations of the fixed coordinate. The mean and standard deviation of these estimated coordinates from 200 simulations is reported in the third and fourth rows of Table 1. The mean values indicate a slight positive bias of the coordinate estimates. This positive bias at least partially arises from bias arising in the

| | | | | | | | | | | | | |
|------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| $x_{i,1}$ | 0.00 | 3.00 | 4.00 | 4.50 | 5.00 | 5.63 | 5.63 | 6.12 | 6.38 | 6.58 | 6.82 | 7.09 |
| $x_{i,2}$ | 7.09 | 6.82 | 6.58 | 6.38 | 6.12 | 5.63 | 5.63 | 5.00 | 4.50 | 4.00 | 3.00 | 0.00 |
| Mean $\hat{x}_{i,*}$ | 7.20 | 6.98 | 6.78 | 6.61 | 6.40 | 5.96 | 5.93 | 6.41 | 6.64 | 6.80 | 7.00 | 7.22 |
| SD $\hat{x}_{i,*}$ | 0.51 | 0.48 | 0.48 | 0.49 | 0.51 | 0.61 | 0.60 | 0.50 | 0.48 | 0.48 | 0.48 | 0.51 |
| Coverage | 0.91 | 0.91 | 0.90 | 0.89 | 0.91 | 0.92 | 0.91 | 0.92 | 0.91 | 0.91 | 0.92 | 0.93 |
| Mean $\tilde{x}_{i,*}$ | 7.26 | 7.03 | 6.78 | 6.61 | 6.36 | 5.88 | 5.90 | 6.37 | 6.62 | 6.82 | 7.09 | 7.34 |
| SD $\tilde{x}_{i,*}$ | 0.57 | 0.62 | 0.69 | 0.72 | 0.79 | 0.98 | 1.01 | 0.81 | 0.73 | 0.67 | 0.62 | 0.59 |

Table 1: Results from 200 simulations of $n = 10000$ t-distributed random vectors. Top two rows give locations of selected points along the true isoline $\ell_{\mathbf{X}}(0.001)$ with bold indicating the coordinate being estimated. Rows 3-5 give the mean estimate, standard deviation, and bootstrap confidence interval coverage rate from the proposed method. Rows 6 and 7 give the mean estimate and standard deviation from a fully-empirical estimate.

estimation of the base isoline. Each set of simulated data is resampled to produce bootstrap-based 95% confidence intervals, and the coverage of these confidence intervals is given in the fifth row. Coverage rates are slightly low, perhaps due to the bias in the mean estimate, and perhaps because we used the percentile method for the bootstrap confidence intervals (Givens and Hoeting, 2012, Section 9.3.1).

It is useful to compare results of our projection approach to a fully empirical estimate of this isoline. We follow the aforementioned procedure for fitting the base isoline at the 0.001 level, and obtain estimates $\tilde{x}_{i,*}$ for each of our selected locations. The bottom two rows of Table 1 show that the bias of the fully empirical estimate is similar to that of our approach, and that the empirical estimates have larger standard errors. Unlike an empirical approach, our method could be used to extrapolate beyond the range of the data (shown in Section 3.3). As our method is intended to produce a visual estimate of an extreme isoline, the most important difference is shown by the realized estimated isolines in Figure 2, illustrating that the fully-empirical estimated isoline is qualitatively much less smooth than the isoline estimate from our projection method. Also shown in Figure 2 are intervals indicating the middle 95% of the estimates from the 200 repeated simulations.

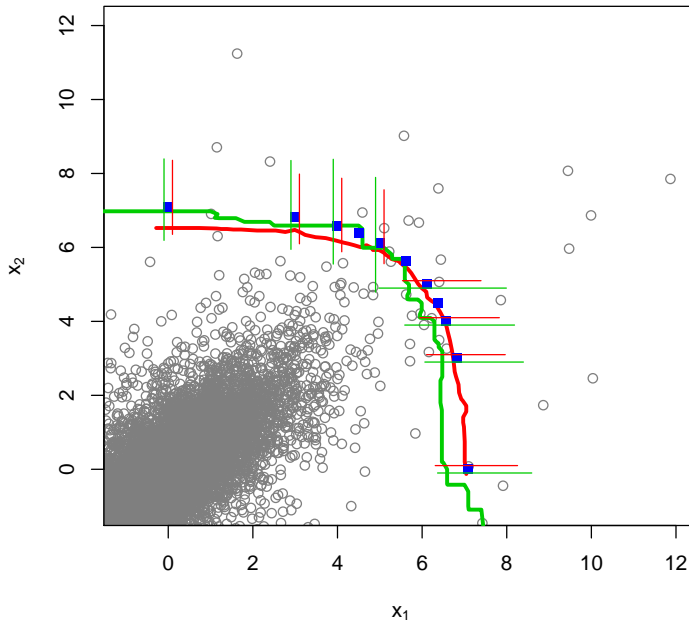


Figure 2: Illustration of the difference between the proposed method for estimating the isoline $\ell_{\mathbf{X}}(0.001)$ and a fully empirical estimate. Data are $n = 10000$ realizations of bivariate t -distributed variables. Blue squares indicate the selected points on the true isoline in the top two rows of Table 1. Red curve is the estimate from the proposed method and green curve is a fully empirical estimate. Intervals for the middle 95% of the estimates from repeated simulations for both the proposed method and the empirical method are shown by the red and green bars respectively.

3.3 Application: Santa Ana fire risk

We use the data for the Santa Ana weather regime to illustrate the isoline procedure in the case of asymptotic dependence. Hourly data were obtained from the HadISD dataset (Dunn et al., 2012) for the March AFB station¹, which lies in Riverside County, CA, and whose data record shows a correspondence with known Santa Ana events such as the Cedar and Witch fires. The data span the years 1973-2015 and we restrict our attention to the months of September, October, and November as these are months for which Santa Ana-driven fires are most prevalent. Details and motivation for construction of the daily windspeed and

¹Station number 722860-23119.

dryness time series from this hourly data are given in [Cooley et al. \(2017\)](#). The dataset contains $n = 3902$ observations. Let $\mathbf{x}_t = (x_{t,1}, x_{t,2})^T$ denote the observed windspeed and dryness on day t . Asymptotic dependence is a reasonable starting assumption for the Santa Ana data since the weather regime leads to conditions which are both very dry and windy. The `chplot` function of R's `evd` library (not shown) indicates that $\hat{\chi} \approx 0.3$, implying that conditions are at their driest levels about 30% of the time when windspeeds are at their highest levels.

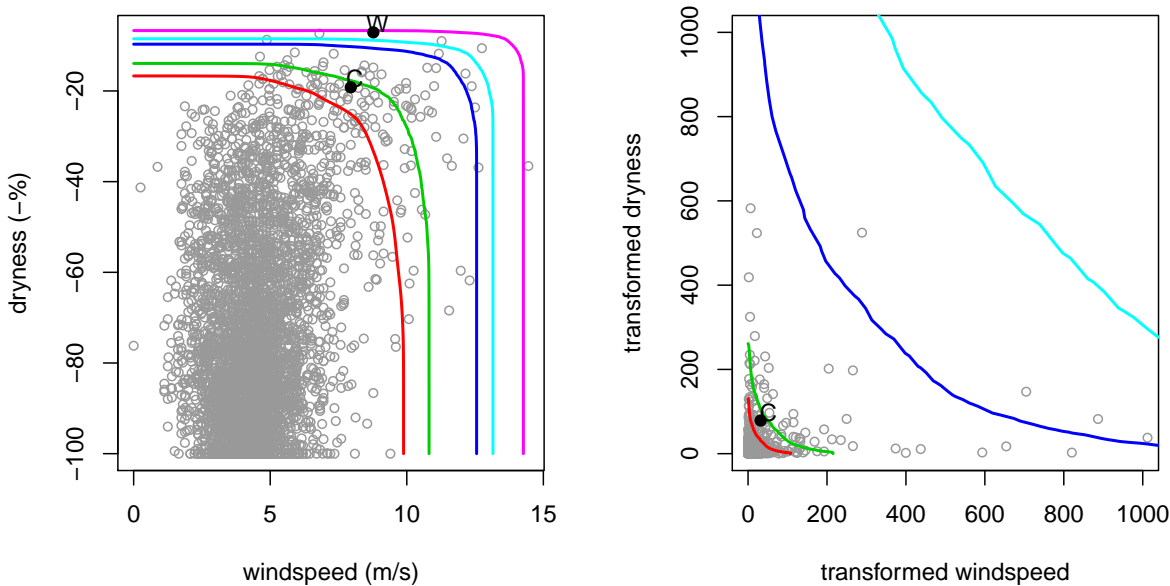


Figure 3: Plot of the Santa Ana data and estimated survival function isolines on the original scale (left), and on the transformed scale (right). Estimated isolines correspond to survival probabilities of 0.01, 0.005, 0.001, 0.0005, and 0.0001. The point corresponding to the Witch Fire is lies outside the plotting region of the right panel.

The left panel of [Figure 3](#) again shows the original data. We let $p_{base} = 0.01$, and the estimate of this base isoline is shown in red.

We next perform marginal transformation as the data itself is clearly not regularly varying; both the variables in the Santa Ana data set are found to have bounded upper tails. After checking mean residual life plots and the stability of the estimates for the tail index ([Coles,](#)

2001, Chapter 4), we set $x_{thold,k}$ to the 0.97 empirical quantile and $x_{thold+,k}$ to the 0.98 empirical quantile for both marginal distributions of the Santa Ana data and obtain $T_k, k = 1, 2$. The right panel of Figure 3 shows the data after transformation: $\mathbf{z}_t = T(\mathbf{x}_t), t = 1, \dots, n$ as well as the transformed base isoline in red. Scatterplots on this transformed scale can be difficult to interpret due to unfamiliarity with the extremely heavy tail when $\xi = 1$ (Cooley et al., 2017). Our plot window was restricted to $[0, 1000]^2$ which does not contain all of the data after transformation. Projected isolines $\hat{\ell}_{\mathbf{Z}}(p_{proj})$ were produced for $p_{proj} = 0.005, 0.001, 0.0005, 0.0001$. The first three are shown in the right panel of Figure 3, the isoline $\hat{\ell}_{\mathbf{Z}}(0.0001)$ falls outside the plotting window.

Finally, applying T_k^{-1} the projected isolines $\hat{\ell}_{\mathbf{X}}(p_{proj})$ are plotted on the original scale in the left panel of Figure 3. We see that the combination of windspeed and dryness seen at this weather station on the day corresponding to the Cedar Fire was not extremely rare, as this point is below $\hat{\ell}_{\mathbf{X}}(0.005)$. The conditions at this station on the day corresponding to the Witch Fire were quite rare as this point lies between $\hat{\ell}_{\mathbf{X}}(0.0005)$ and $\hat{\ell}_{\mathbf{X}}(0.0001)$.

4 Asymptotic independence case

4.1 Procedure

The procedure in the asymptotic independence setting begins in the same manner as before. First, a base isoline $\hat{\ell}_{\mathbf{X}}(p_{base})$ is nonparametrically estimated using the original data $\mathbf{x}_t, t = 1, \dots, n$. Then, transformations obtained from the estimated marginal distributions are applied to obtain data $\mathbf{z}_t = T(\mathbf{x}_t), t = 1, \dots, n$, and the isoline $\hat{\ell}_{\mathbf{Z}}(p_{base}) = T(\hat{\ell}_{\mathbf{X}}(p_{base}))$ on the Fréchet scale. We assume the transformed data \mathbf{z}_t arise from a random variable \mathbf{Z} which exhibits hidden regular variation on the interior of \mathcal{C} . To estimate the η , note $Z^{(min)} = \min(Z_1, Z_2)$ is regularly varying with index η (Ledford and Tawn, 1996). In Section 4.3, we will use the Hill estimator (Hill, 1975) applied to the realizations $z_t^{(min)}$ to estimate η , but there are several other widely-used methods for estimating an index of regular variation.

We next wish to use the scaling properties of regular variation and hidden regular variation to produce isolines corresponding to smaller exceedance probabilities. Because we have chosen our isolines to be based on joint survival regions, our procedure is able to utilize hidden regular variation to scale the interior. However, the interface between the first-order regular variation and hidden regular variation that occurs at the axes presents a modeling challenge. If $\mathbf{z} \in \hat{\ell}_{\mathbf{Z}}(p_{base})$ is in the interior of \mathcal{C} , the asymptotic theory says that (4) is the correct relationship to use; however, if \mathbf{z} lies on an axis, then (2) should be used. This abrupt change would result in discontinuous isolines at the axes.

Because the abrupt transition between regimes described by the asymptotic theory is not reflected in data, we propose smoothing the transition between scalings (2) and (4). Let $\mathbf{z}^{(base)} = (z_1^{(base)}, z_2^{(base)}) \in \hat{\ell}_{\mathbf{Z}}(p_{base})$. For some smoothing parameter $\beta \in [1, \infty)$, let $m_i = 1 - (z_i^{(base)} / (z_1^{(base)} + z_2^{(base)}))^\beta$, and let $\eta_i(\mathbf{z}^{(base)}) = m_i \hat{\eta} + (1 - m_i)$ for $i = 1, 2$. To construct an isoline to correspond with an exceedance probability $p_{proj} < p_{base}$, let $s = p_{base}/p_{proj}$ as before. Consider $\mathbf{z}^{(proj)} = (s^{\eta_1(\mathbf{z}^{(base)})} z_1^{(base)}, s^{\eta_2(\mathbf{z}^{(base)})} z_2^{(base)})$. If $\mathbf{z}^{(base)}$ is sufficiently away from the axes such that $m_i \approx 1$ for $i = 1, 2$, then

$$P(\mathbf{Z}_t > \mathbf{z}^{(proj)}) \approx P(\mathbf{Z}_t > s^{\hat{\eta}} \mathbf{z}^{(proj)}) \approx (s^{\hat{\eta}})^{-1/\eta} P(\mathbf{Z}_t > \mathbf{z}^{(base)}) \approx p_{proj},$$

where the second approximation comes from (4). On the other hand consider the case where $\mathbf{z}^{(base)}$ lies on the axis. Suppose $z_1^{(base)} = 0$.

$$P(\mathbf{Z}_t > \mathbf{z}^{(proj)}) = P(Z_{t,1} > 0, Z_{t,2} > s z_2^{(base)}) \approx s^{-1} P(\mathbf{Z}_t > \mathbf{z}^{(base)}) = p_{proj},$$

where the approximation follows from (2). For $\mathbf{z}^{(base)}$ near the axes, these two cases are weighted with the weight depending on the value of β selected by the investigator. As $\beta \rightarrow \infty$, one approaches the discontinuous transition between regimes and the projection of the isoline is scaled primarily by the coefficient of tail dependence η even near the axes. As β decreases, smoothing is increased, and the first-order regular variation influences behavior further from

| | | | | | | | | | | | | |
|----------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| $x_{i,1}$ | 0.00 | 1.00 | 1.50 | 2.00 | 2.50 | 2.61 | 2.61 | 2.71 | 2.96 | 3.05 | 3.08 | 3.09 |
| $x_{i,2}$ | 3.09 | 3.08 | 3.05 | 2.96 | 2.71 | 2.61 | 2.61 | 2.50 | 2.00 | 1.50 | 1.00 | 0.00 |
| Mean $\hat{x}_{i,*}$ | 3.05 | 2.99 | 2.96 | 2.91 | 2.71 | 2.61 | 2.61 | 2.70 | 2.91 | 2.96 | 2.98 | 3.04 |
| SD $\hat{x}_{i,*}$ | 0.08 | 0.08 | 0.08 | 0.09 | 0.12 | 0.15 | 0.14 | 0.12 | 0.09 | 0.08 | 0.08 | 0.08 |
| Coverage | 0.84 | 0.69 | 0.73 | 0.83 | 0.90 | 0.90 | 0.90 | 0.89 | 0.81 | 0.69 | 0.65 | 0.78 |

Table 2: Results from 200 simulations of $n = 10000$ Gaussian-distributed random vectors.

the axes. Somewhat similar to [Wadsworth and Tawn \(2013\)](#), the above smoothing procedure implies that projections are along curves rather than the straight rays of regular variation.

If $\hat{F}_{\mathbf{X}}(\mathbf{x})$ is decreasing, isolines of survival probabilities have negative slopes. We show in the appendix that projected isolines produced by the above smoothing procedure retain this property.

4.2 Simulation study: bivariate Gaussian

We assess our procedure in the asymptotic independence case via a simulation study where we let \mathbf{X} be a Gaussian-distributed bivariate random vector with mean 0, variance 1, and correlation 0.7. As before, for each iteration, we simulate $n = 10000$ realizations of \mathbf{X} , set the base isoline at the 0.01 level, and project to the 0.001 level. The asymptotic independence case requires us to choose β . Based on viewing the diagnostic plot in [Section 5](#) for a handful of simulations, a value of $\beta = 200$ seemed to perform well, and this value was used for all 200 simulations. As before, each iteration’s realized data was resampled 100 times to produce bootstrap-based confidence intervals for the estimated coordinate.

Results from the simulation study appear in [Table 2](#). Mean estimates show slight downward bias at the ‘edges’; for instance, when $x_{i,2}$ is being estimated and $x_{i,1}$ is a small value such as zero. Perhaps surprisingly given that β was chosen by visual inspection, the estimates near the ‘center’ show no bias. Coverage rates for the bootstrap-based 95% confidence intervals are low, particularly near the ‘edges’ where bias was evident.

4.3 Application: Karachi heat wave conditions

We use the Karachi heat wave data to illustrate the method in the asymptotic independence setting. The data are again a HadISD dataset (Dunn et al., 2012), this time for the Karachi Airport station.² The original data span 1973-2015 and are nominally hourly, although the recording interval was every four hours until 1992. For each day, we retain the temperature (converted to Fahrenheit) and relative humidity both at the time of the maximum heat index value which itself is a function of temperature and humidity. We examine data from the months from April to October, as these were the months during which a heat index value exceeded the 0.95 empirical quantile. Time series plots of temperature and humidity still display seasonality, as temperatures tend to increase in April and May, peak in June, dip in the monsoon months of July and August, then increase again in September before decreasing in October. Relative humidity tends to be higher in the monsoon months. This seasonality makes interpretation of the isolines more difficult, as one must think of the isolines representing exceedance probabilities of the distribution of these two variables integrated over the warm months. The data set contains $n = 8963$ data points.

The Karachi data are clearly asymptotically independent. In fact these data exhibit negative association: days with highest temperature are days with lower humidity, and conversely, the days with highest humidity are days with lower temperatures. This negative association is to be expected as moisture in the air mitigates temperature. Despite this negative relationship, there is a need to characterize the relationship in the tail of these two variables as human health is adversely affected when both temperature and humidity are high. Let $\mathbf{x}_t = (x_{t,1}, x_{t,2})$ represent the observed temperature and relative humidity from day t . The left panel of Figure 4 shows the data on its original scale and the estimated base isoline $\hat{\ell}_{\mathbf{X}}(0.01)$ in red.

The data require marginal transformation before regular variation can be assumed. Using the same procedure as described in Section 3.3, we obtain $\hat{F}_k(x)$ for $k = 1, 2$. Relative

²Station number 417800-99999.

humidity has a known upper bound of 100, and one could restrict the parameters of the fitted generalized Pareto distribution to require this upper bound. We did not choose to impose this restriction as it might adversely affect the marginal fit within the observed range, and additionally would complicate the procedure which we aim to be easily and widely applicable. From $\hat{F}_k(x)$, we obtain the marginal transformations T_k , and again let $\mathbf{z}_t = T(\mathbf{x}_t)$ which we assume to be regularly varying with index $\xi = 1$ on \mathcal{C} . The transformed data $\mathbf{z}_t = T(\mathbf{x}_t)$ and transformed base isline $\hat{\ell}_{\mathbf{Z}}(p_{base}) = T(\hat{\ell}_{\mathbf{X}}(p_{base}))$ are shown in the right panel of Figure 4. Note that this panel's range is set to $[0, 50]^2$ in order to show the behavior of points in the interior of \mathcal{C} and there are many large points near the axes which are beyond this range.

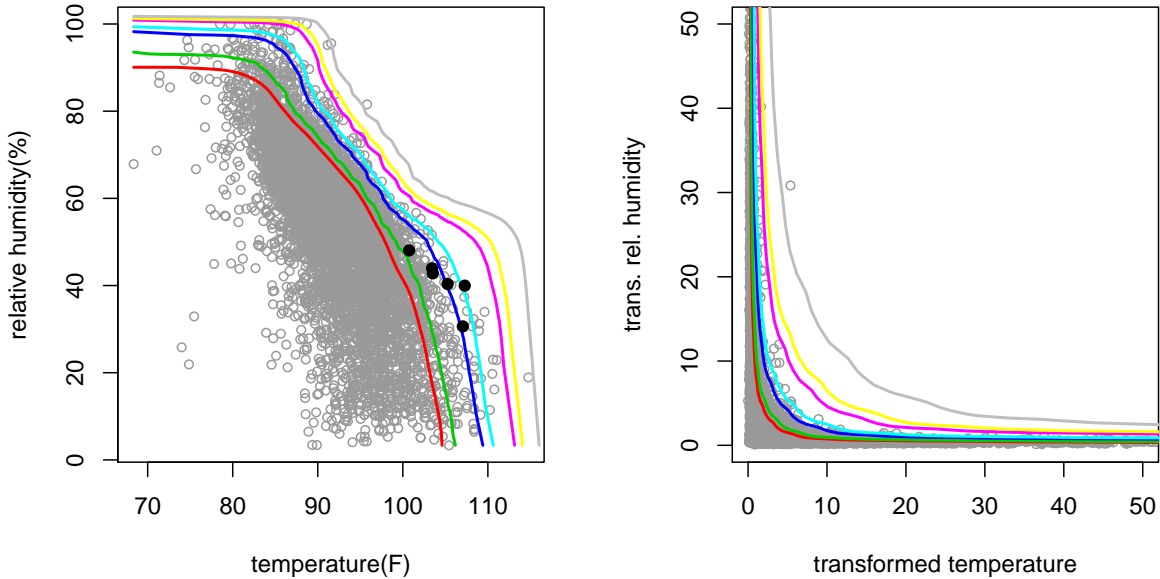


Figure 4: Plot of the Karachi data and estimated survival function isolines on the original scale (left), and on the transformed scale (right). Estimated isolines correspond to survival probabilities of 0.01, 0.005, 0.001, 0.0005, 0.0001, 0.00005, and 0.00001.

We estimate η via a Hill estimator (Hill, 1975) applied above a threshold corresponding to the 0.98 empirical quantile of $z_t^{(min)}$ of the Karachi data, and we obtain $\hat{\eta} = 0.20$. This value

indicates negative correspondence in the tail as independence between the variates would imply $\eta = 1/2$ when $\xi = 1$. A Hill plot was produced which indicated that this quantile was reasonable for estimating η .

Using diagnostics to be explained in Section 5, we selected $\beta = 200$ and produced isolines $\hat{\ell}_{\mathbf{Z}}(p_{proj})$ for $p_{proj} = 0.005, 0.001, 0.0005, 0.0001, 0.00005, 0.00001$, which are shown in the right panel of Figure 4. Inverting the marginal transformation yields the estimated isolines on the original scale shown in the left panel of Figure 4. The conditions recorded at the Karachi airport on the dates between June 18 and June 23, 2015 were all rare. The points corresponding to these dates all exceed $\hat{\ell}_{\mathbf{X}}(0.005)$, three exceed $\hat{\ell}_{\mathbf{X}}(0.001)$ and one exceeds $\hat{\ell}_{\mathbf{X}}(0.0005)$.

5 A diagnostic plot and bootstrap uncertainty

We propose a plot to assess whether a projected isoline adequately represents the survival probabilities it purports to describe. This plot can also help in selecting an appropriate smoothing parameter β . Since we presume $\hat{F}_{\mathbf{X}}(\mathbf{x}) = p$ for any $\mathbf{x} \in \hat{\ell}_{\mathbf{X}}(p)$, we propose plotting the empirical survival probability for a number of points along a projected isoline. In theory, an isoline consists of an infinite number of points, but in practice is described by a finite number of points. These points originally arise from how $\hat{\ell}_{\mathbf{X}}(p_{base})$ is stored, and then how these points are projected to create isolines with smaller exceedance probabilities. For convenience, our diagnostic plot calculates the empirical survival probabilities at these stored points.

To help quantify the expected uncertainty of these empirical probabilities, for any point $\mathbf{x} \in \hat{\ell}_{\mathbf{X}}(p)$, let $B(\mathbf{x}) = \sum_{t=1}^n \mathbb{I}(\mathbf{X}_t \in [\mathbf{x}, \infty))$. Assuming $B(\mathbf{x}) \sim \text{Binomial}(n, p)$, we find the smallest interval (n_1, n_2) , $n_i \in \mathbb{Z}$, such that $P(B(\mathbf{x}) \in [n_1, n_2]) \geq 0.95$. We then report the interval $n^{-1}[n_1, n_2]$ which can be compared to the empirical probabilities $B(\mathbf{x})/n$. For two points $\mathbf{x}_1, \mathbf{x}_2 \in \hat{\ell}_{\mathbf{X}}(p)$, $B(\mathbf{x}_1)$ and $B(\mathbf{x}_2)$ are not independent as the regions $[\mathbf{x}_1, \infty]$

and $[\mathbf{x}_2, \infty]$ will overlap, with extensive overlap if \mathbf{x}_1 and \mathbf{x}_2 are close to one another. The aforementioned interval does not account for this dependence.

Figure 5 shows empirical survival probability plots for the Santa Ana data on the left, and plots for the Karachi data for two different values of β center and right. The plotted empirical survival probabilities clearly show dependence. The plot for the Santa Ana data is for $p = 0.001$, and all empirical exceedance probabilities fall within the uncertainty interval. The Karachi plot in the center is for $\beta = 200$ and $p = 0.0005$. While some empirical probabilities fall both above and below the uncertainty interval, the values are close to the interval bounds. This is not the case for the right plot, where $\beta = 1000$. The empirical probabilities on the far right and far left clearly exceed the upper bound of the uncertainty interval by a large amount indicating a clear discrepancy of this isoline. With such a large value for β , the transition from regular variation in the interior to the first-order regular variation on the axes is too abrupt, and η overly influences the scaling near the axes, resulting in too many observed exceedances. In practice, a user will likely need to produce several diagnostic plots for different values of β to qualitatively determine which value appears to best capture the empirical probabilities. Returning to the center plot, we tried changing both β and reducing the exceedance probability of the base isoline, but were unable to qualitatively improve the performance beyond what is demonstrated in this plot.

Although we primarily view these isolines as a way to explore the extreme behavior of the data, it may be useful to produce visual measures of uncertainty for these estimated isolines. To this end, we propose a block bootstrap approach, where the block size b accounts for temporal dependence in the data. For each bootstrap iteration, we obtain a resample of blocks of data $(x_{t^*,1}, x_{t^*,2}), \dots, (x_{t^*+b,1}, x_{t^*+b,2})$ where t^* is randomly selected and where the total resample is of length n . Then, treating this resample as data, we are able to produce an isoline as before.

Figure 6 shows 200 bootstrapped isolines for the Santa Ana and Karachi data, both for an exceedance probability of $p = 0.001$. Based on autocorrelation plots, we use a block

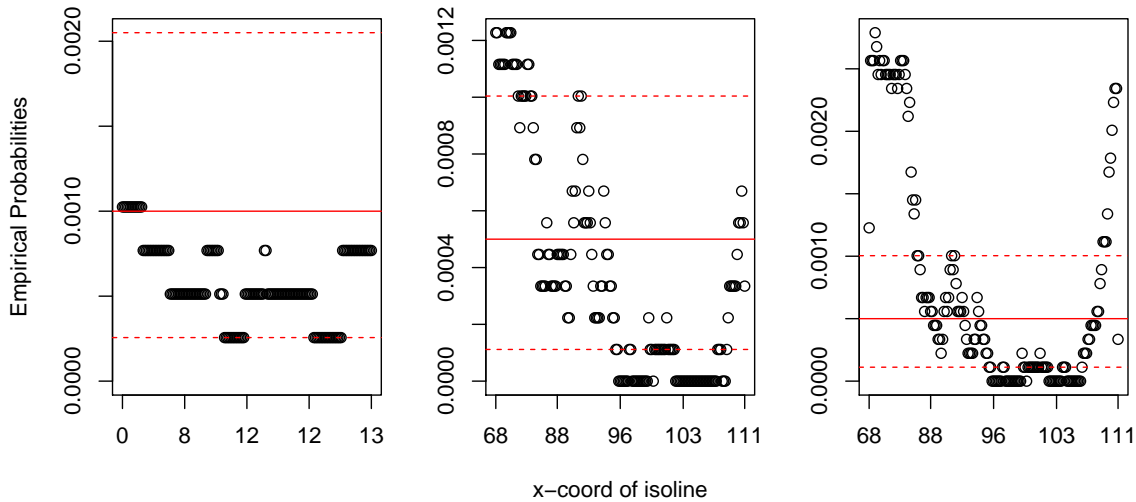


Figure 5: Plot of empirical survival probabilities for points along the estimated isoline. Left: Santa Ana data, for $\hat{\ell}_{\mathbf{X}}(0.001)$. Center: Karachi data for $\hat{\ell}_{\mathbf{X}}(0.0005)$, with $\beta = 200$. Right: Karachi data for $\hat{\ell}_{\mathbf{X}}(0.0005)$, with $\beta = 1000$. Solid line corresponds to target isoline probability and dashed lines show the boundaries of the smallest interval with at least 0.95 probability of a binomial distribution with the target probability.

length of $b = 3$ for the Santa Ana data and $b = 5$ for the Karachi data. Shown in both plots is the original data in gray, and data from a single bootstrap iteration. One notices that our simple block bootstrap technique results a subset of data with many fewer unique points, which, at a minimum, affects the GPD fit of the marginal tails and the behavior of our nonparametric method for drawing the base isoline. Nevertheless, the bootstrapped isolines do give a useful visual representation of uncertainty.

6 Conclusions and discussion

We have developed a method for producing isolines of bivariate exceedance probabilities. This method relies on a dependence structure specifically suited for describing extremal dependence and which can be used to produce isolines for which there are few or even no exceedances. Two advantages of the proposed method are that it is largely nonparametric, and that it can be extended to the asymptotic independent case. As implemented, the user is

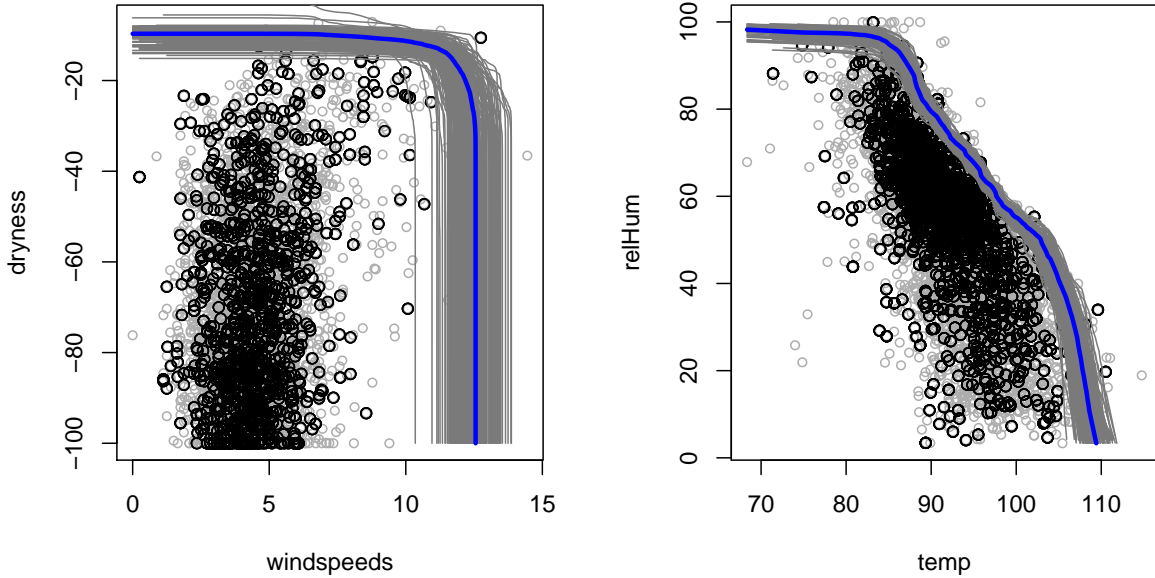


Figure 6: Bootstrapped isolines for exceedance probability $p = 0.001$ for the Santa Ana (left) and Karachi data (right). Original data shown in gray, and a single block-bootstrapped resample shown in black.

asked to choose in which asymptotic regime to model; however, conceptually the asymptotic independence approach could be used for asymptotically dependent data, as the true value of η would equal 1. We acknowledge that often it is not easy to determine if data are asymptotically dependent or independent, but our opinion is that if the data is believed to be asymptotically dependent, it is best to not estimate η .

This general approach is not necessarily limited to two dimensions. Regular variation is well-developed for higher dimensions, and hidden regular variation has been described in higher dimensions as well. However, implementing our approach in higher dimensions when the data exhibit hidden regular variation could be more complicated; for instance, different two-dimensional marginals could have different coefficients of tail dependence. More practically, higher-dimensional isolines would be difficult to visualize.

We believe this tool could be useful for researchers to explore the extremal behavior of

bivariate data to better understand potential risk. It is important to keep in mind however that the isolines denote the rarity of events, not impact. There are observations in the Karachi data, for instance rare exceptionally cool and humid days, which have equally low probability of exceedance to those corresponding to the June 2015 heat wave, but which likely have no human health impact. Hence, practical application of the technique presented here also requires additional information about which portion of the multivariate space is impactful.

For heat waves, human impact arises not only from the occurrence of extreme hot and humid conditions, but also from prolonged exposure to these conditions such as was seen in the 2015 Karachi event. Producing isolines such as we have done for Karachi can help researchers understand the rarity of combinations of conditions, but does not answer questions related to the duration of extreme conditions.

Our method assumes that the data over the study period are stationary, and it is hard to envision allowing for non-stationarity while retaining the nonparametric aspects of the method. In both the Santa Ana and Karachi examples, we do not actually believe that the data are stationary, as climate has changed over the study periods. However, over a roughly forty-year study period changes in climate, and in particular in extreme behavior, would be difficult to detect. One could imagine using the method to compare the extreme behavior between two different climates, perhaps mid-20th century to current climate, or to climate model output run under two different scenarios.

In theory, the asymptotic independence approach should cover the asymptotic dependent case since in this case $\eta = 1$. Because one is at the boundary of the parameter space, in the asymptotic dependence case the estimate $\hat{\eta}$ will be less than one, thus estimated isolines would tend to underrepresent the probabilities of extreme events in the interior. Although there is recent work which seeks to unify the two modeling paradigms ([Wadsworth et al., 2017](#)), we think there is value in forcing the researcher to consider the dependence class in which to model. Even though this choice is often not clear from the data, the researcher

could also consider the consequences of over- and under-estimating extremal probabilities when making this modeling decision.

The code and data for this project are currently posted at <http://www.stat.colostate.edu/cooleyd/Isolines/>. We are working with the maintainer of an existing R package for extremes to implement the method within this package.

Acknowledgements: Cooley, Thibaud, and Castillo received support from the project “EaSM 2: Advancing extreme value analysis of high impact climate and weather events” NSF-DMS-1243102. Wehner’s contributions to this work are supported by the Regional and Global Climate Modeling Program of the Office of Biological and Environmental Research in the Department of Energy Office of Science under contract number DE-AC02-05CH11231.

A Appendix

We show that the smoothed scaling procedure in Section 4 preserves the requirement that isolines of exceedance probabilities must have negative slopes.

Assume the estimated survival function is strictly decreasing, i.e., if $\mathbf{x}_1 = (x_{1,1}, x_{1,2}) \neq \mathbf{x}_2 = (x_{2,1}, x_{2,2})$, $x_{1,1} \leq x_{2,1}$, and $x_{1,2} \leq x_{2,2}$, then $\hat{F}_{\mathbf{X}}(\mathbf{x}_1) > \hat{F}_{\mathbf{X}}(\mathbf{x}_2)$. We first show that it follows that any isoline $\hat{\ell}_{\mathbf{X}}(p)$ must have a negative slope. Let \mathbf{x}_1 and \mathbf{x}_2 be two distinct locations on $\hat{\ell}_{\mathbf{X}}(p)$. WLOG, assume $x_{2,2} - x_{1,2} \geq 0$ and $x_{2,1} - x_{1,1} \geq 0$, implying the slope is not negative. This implies $x_{2,2} \geq x_{1,2}$ and $x_{2,1} \geq x_{1,1}$, but $\hat{F}_{\mathbf{X}}(\mathbf{x}_1) = \hat{F}_{\mathbf{X}}(\mathbf{x}_2)$, which is a contradiction.

As the transformation to Fréchet scale is monotonic, $\hat{\ell}_{\mathbf{Z}}(p)$ has negative slopes.

Let $\mathbf{z}_1^{(proj)}, \mathbf{z}_2^{(proj)}$ be any two points in $\hat{\ell}_{\mathbf{Z}}(p_{proj})$. Let $s = p_{base}/p_{proj}$. Let $\mathbf{z}_1^{(base)}, \mathbf{z}_2^{(base)}$ be the points in $\hat{\ell}_{\mathbf{Z}}(p_{base})$ such that $\mathbf{z}_i^{(proj)} = (s\eta_1(\mathbf{z}_i^{(base)}), z_{i,1}^{(base)}, s\eta_2(\mathbf{z}_i^{(base)}), z_{i,2}^{(base)})$ for $i = 1, 2$. WLOG assume $z_{2,2}^{(base)} < z_{1,2}^{(base)}$ and $z_{2,1}^{(base)} > z_{1,1}^{(base)}$.

Note that since $z_{2,1}^{(base)} > z_{1,1}^{(base)}$, $\eta_1(\mathbf{z}_2^{(base)}) > \eta_1(\mathbf{z}_1^{(base)})$, and likewise since $z_{2,2}^{(base)} < z_{1,2}^{(base)}$,

$\eta_2(\mathbf{z}_2^{(base)}) < \eta_2(\mathbf{z}_1^{(base)})$. Hence,

$$z_{2,2}^{(proj)} - z_{1,2}^{(proj)} = s^{\eta_2(\mathbf{z}_2^{(base)})} z_{2,2}^{(base)} - s^{\eta_2(\mathbf{z}_1^{(base)})} z_{1,2}^{(base)} < s^{\eta_2(\mathbf{z}_1^{(base)})} (z_{2,2}^{(base)} - z_{1,2}^{(base)}) < 0,$$

and

$$z_{2,1}^{(proj)} - z_{1,1}^{(proj)} = s^{\eta_1(\mathbf{z}_2^{(base)})} z_{2,1}^{(base)} - s^{\eta_1(\mathbf{z}_1^{(base)})} z_{1,1}^{(base)} > s^{\eta_1(\mathbf{z}_1^{(base)})} (z_{2,1}^{(base)} - z_{1,1}^{(base)}) > 0.$$

Thus, the slope between any two points on $\hat{\ell}_{\mathbf{Z}}(p_{proj})$ is negative, and since the marginal transformation is monotonic, the slope between any two points on $\hat{\ell}_{\mathbf{X}}(p_{proj})$ is negative.

References

- Cai, J.-J., Einmahl, J. H., and De Haan, L. (2011). Estimation of extreme risk regions under multivariate regular variation. *The Annals of Statistics*, 39(3):1803–1826.
- Coles, S. and Tawn, J. (1994). Statistical methods for multivariate extremes: an application to structural design. *Applied Statistics*, 43(1):1–48.
- Coles, S. G. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer-Verlag London Ltd., London.
- Cooley, D. (2013). Return periods and return levels under climate change. In *Extremes in a changing climate*, pages 97–114. Springer.
- Cooley, D., Smith, R., and Hunter, B. (2017). Univariate and multivariate extremes for the environmental sciences. In Gelfand, A. and RL, S., editors, *Environmental Statistics*, pages XX–XX. TBD, TBD.
- Das, B. and Resnick, S. (2014). Hidden regular variation: Generation and detection. *ARXIV*, 1:1–27. arXiv:1403.5774v1.

- de Haan, L. and de Ronde, J. (1998). Sea and wind: Multivariate extremes at work. *Extremes*, 1:7–45.
- Dunn, R. J., Willett, K. M., Thorne, P. W., Woolley, E. V., Durre, I., Dai, A., Parker, D. E., and Vose, R. E. (2012). HadISD: a quality-controlled global synoptic report database for selected variables at long-term stations from 1973–2011. *Climate of the Past*, 8:1649–1679.
- Einmahl, J. H., Li, J., and Liu, R. Y. (2009). Thresholding events of extreme in simultaneous monitoring of multiple risks. *Journal of the American Statistical Association*, 104(487):982–992.
- Givens, G. H. and Hoeting, J. A. (2012). *Computational statistics*, volume 710. John Wiley & Sons.
- Heffernan, J. E. and Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society, Series B*, 66:497–546.
- Hill, B. (1975). A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174.
- Huser, R. and Wadsworth, J. L. (2018). Modeling spatial processes with unknown extremal dependence class. *Journal of the American Statistical Association*, 0(0):1–11.
- Ledford, A. and Tawn, J. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society, Series B*, 59(2):475–499.
- Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, 83(1):169–187.
- Liu, R. and Yang, L. (2008). Kernel estimation of multivariate cumulative distribution function. *Journal of Nonparametric Statistics*, 20(8):661–667.
- Marcon, G., Naveau, P., and Padoan, S. A. (2017). A semi-parametric stochastic generator for bivariate extreme events. *Stat*, 6(1):184–201.

- Masood, I., Majid, Z., Sohail, S., Zia, A., and Raza, S. (2015). The deadly heat wave of Pakistan, June 2015. *The International Journal of Occupational and Environmental Medicine*, 6:672–247.
- Nelsen, R. (2006). *An Introduction to Copulas, 2nd Edition*. Lecture Notes in Statistics No. 139. Springer, New York.
- Resnick, S. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer-Verlag, New York.
- Resnick, S. (2002). Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4):303–336.
- Resnick, S. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer, New York.
- Salvadori, G. and De Michele, C. (2004). Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resources Research*, 40(12).
- Stephenson, A. G. (2002). evd: Extreme value distributions. *R News*, 2(2):31–32.
- Wadsworth, J. and Tawn, J. (2013). A new representation for multivariate tail probabilities. *Bernoulli*, 19(5B):2689–2714.
- Wadsworth, J., Tawn, J. A., Davison, A., and Elton, D. M. (2017). Modelling across extremal dependence classes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 79(1):149–175.
- Weller, G. and Cooley, D. (2013). A sum characterization of hidden regular variation with likelihood inference via expectation–maximization. *Biometrika*, 101(1):17–36.