Modeling Extreme Events in Spatial Domain by Copula Graphical Models

Hang Yu*, Zheng Choo†, Wayne Isaac T. Uy†, Justin Dauwels*, and Philip Jonathan‡

*School of Electrical and Electronics Engineering, †School of Physical and Mathematical Sciences
Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798
‡Shell Technology Centre Thornton, P.O. Box 1, Chester, UK

Abstract—We propose a new statistical model that captures the conditional dependence among extreme events in a spatial domain. This model may for instance be used to describe catastrophic events such as earthquakes, floods, or hurricanes in certain regions, and in particular to predict extreme values at unmonitored sites. The proposed model is derived as follows. The block maxima at each location are assumed to follow a Generalized Extreme Value (GEV) distribution. Spatial dependence is modeled in two complementary ways. The GEV parameters are coupled through a thin-membrane model, a specific type of Gaussian graphical model often used as smoothness prior. The extreme events, on the other hand, are coupled through a copula Gaussian graphical model with the precision matrix corresponding to a (generalized) thin-membrane model. We then derive inference and interpolation algorithms for the proposed model. The approach is validated on synthetic data as well as real data related to hurricanes in the Gulf of Mexico. Numerical results suggest that it can accurately describe extreme events in spatial domain, and can reliably interpolate extreme values at arbitrary sites.

I. INTRODUCTION

Extreme events such as floods, hurricanes, and earthquakes often have a devastating impact on our society. Statistical models may help to assess the likelihood of such extreme events [1], and the dependency of extreme events across space [1], [2], [3]. These models may be useful to quantify the risks associated with certain infrastructures and facilities exposed to extreme conditions.

Extreme value theory governs the behavior regarding the asymptotic distributions of the extreme order statistics [1]. The Fisher-Tippett-Gnedenko (FTG) theorem, often called the first theorem in extreme value theory, asserts that the block maxima (e.g., monthly or annual maxima) of i.i.d. univariate samples converge to the (GEV) distribution [4]. Of great interest are models for extreme events in *spatial* domain, since they can be used to describe catastrophic events in certain regions (e.g., extreme waves in the Gulf of Mexico). One of the most challenging issues in modeling extreme events in spatial domain is to fully capture the spatial dependence of extreme events, while constraining the marginals to be GEV distributions (motivated by the FTG Theorem).

So far, the literature on modeling high-dimensional spatial extreme events is rather limited. In the following, we provide a brief review. In [2], a procedure is proposed to compute the pairwise spatial dependence of extreme events, i.e., the probability of threshold exceedance at one site conditioned on the exceedance at another site [3]. Alternatively, Naveau et al. [5] quantified pairwise spatial dependence through the concept of a variogram. However, both models are only limited to pairwise dependency. On the other hand, it has been shown in [6] and [7] that spatial dependence can be captured to some extent by smoothing the parameters of marginal extreme value distributions. However, the extreme events at different locations are assumed to be conditionally independent after considering the parameter dependence. Recently, Sang et al. [8] mitigated the conditional independence assumption by means of a Gaussian copula, allowing the marginals to follow GEV distributions while capturing the spatial dependence through the exponential covariance function in the Gaussian latent layer. An obvious advantage of using a Gaussian latent layer is that once the mean and the covariance structure have been specified, it is straightforward to compute marginal and conditional distributions. However, Sang et al. [8] assume that the shape and scale parameters of the GEV distribution do not vary in space, which is less appropriate for phenomena that exhibit substantial fluctuations over a large spatial domain (e.g., global extreme sea states corresponding to hurricanes). A further difficulty with the model of [8] is that a standard Gaussian copula is parameterized by a dense covariance matrix. As a result, inference in such model is computationally complex and may be intractable for high-dimensional data (e.g., large grids with thousands of sites) [9].

We introduce a novel model that is similar in spirit as the approach of [8]. The main idea is to introduce Markov Random Fields (MRF) in extreme spatial model, specifically, thin-membrane models. More precisely, instead of using a covariance function to capture the spatial dependence, we use a sparse precision matrix (inverse covariance matrix) with the structure of thin-membrane model, highly reducing the computational complexity. In addition, we allow the marginal GEV parameters to vary in space: we smoothen them through a thin-membrane model [7]. The GEV parameters are estimated by Gaussian inference. The smoothness parameters of the thin-membrane models, which are hyperparameters in the overall model, are inferred from the data by expectation maximization. The resulting estimates of the shape and scale parameters tend to be more accurate and may vary systematically across space instead of being constant. We will refer to the proposed approach as copula MRF-GEV model.

We further extend the concept of thin-membrane models to

irregular grids through Delaunay triangulation [10], allowing us to handle the common situation where measurements are collected at random locations. We derive interpolation algorithms from the copula MRF-GEV model. The resulting interpolation schemes strongly resemble inverse distance weighted (IDW) interpolation [11], and are quite simple and efficient, due to the sparse thin-membrane structure.

We apply the copula MRF-GEV model to synthetic data and real data, related to extreme wave heights in the Gulf of Mexico. We benchmark the proposed model with several other spatial models: MRF-GEV model [7] (with spatially dependent GEV parameters but conditionally independent extreme values), copula GEV model (with locally fitted GEV parameters but coupled extreme events), and a thin-membrane model, directly fitted to the data without using copulas. The numerical results clearly demonstrate that incorporating both extreme-value dependence and parameter dependence across space leads to more accurate inference. Moreover, by adjusting the smoothness of GEV parameters automatically, the estimated GEV parameters are able to capture different types of spatial variations.

The rest of the paper is organized as follows. In the next section, we briefly review thin-membrane models, since those models play a central role in our approach. In Section III we discuss the GEV marginals, and describe algorithms to infer the GEV parameters. In Section IV, we describe how we incorporate dependencies among the extreme events by means of a copula Gaussian graphical model. In Section V we explain how our proposed model can be used for interpolating extreme values at sites without observations. In Section VI we assess the proposed model and benchmark it with other spatial models by means of synthetic and real data. We offer concluding remarks in Section VII.

II. THIN-MEMBRANE MODELS

We use thin-membrane models to capture the spatial dependence of the GEV parameters and the extreme events. We first briefly review Gaussian graphical models, and subsequently, the special case of thin-membrane models. Next, we elaborate on generalized thin-membrane models.

In Gaussian graphical models or Gauss-Markov random fields, a joint p-dimensional Gaussian probability distribution $N(\mu, \Sigma)$ is represented by an undirected graph $\mathcal{G}$ which consists of nodes $\mathcal{V}$ and edges $\mathcal{E}$. Each node i is associated with a random variable X_i. An edge (i, j) is absent if the corresponding two variables X_i and X_j are conditionally independent: $P(X_i, X_j | X_{\mathcal{V}|i,j}) = P(X_i | X_{\mathcal{V}|i,j}) P(X_j | X_{\mathcal{V}|i,j})$, where $\mathcal{V}|i, j$ denotes all the variables except X_i and X_j. It is well-known that for multivariate Gaussian distributions, the above property holds if and only if $K_{i,j} = 0$, where $K = \Sigma^{-1}$ is the precision matrix (inverse covariance matrix).

The thin-membrane model is a Gaussian graphical model that is commonly used as smoothness prior as it minimizes the difference between values at neighboring nodes. The thin-membrane model is usually defined for regular grids, as

illustrated in Fig. 1(a), and its pdf can be written as:

$$P(X) \propto \exp\{-\alpha \sum_{i \in V} \sum_{j \in \mathcal{N}(i)} (X_i - X_j)^2\} \qquad (1)$$

$$\propto \exp(-\alpha\, X^T K_p X), \qquad (2)$$

where $\mathcal{N}(i)$ denotes the neighboring nodes of node i, and α is the smoothness parameter. The matrix K_p is an adjacency matrix with its diagonal elements $[K_p]_{i,i}$ equal to the number of neighbors of site i, while its off-diagonal elements $[K_p]_{i,j}$ are -1 if the sites i and j are adjacent and 0 otherwise. Note that $K = \alpha K_p$ is the precision matrix of $P(X)$ (2). The parameter α controls the smoothness of the whole thin-membrane model, imposing the same extent of smoothness for all pairs of neighbors.
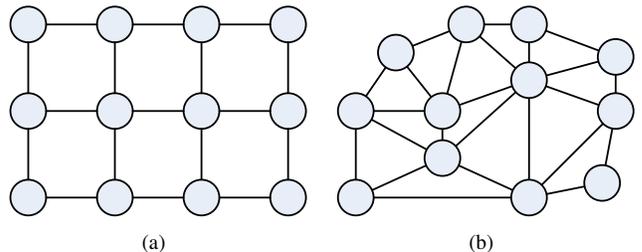
Fig. 1.   Generalized thin-membrane model: (a) Regular grid; (b) Irregular grid.

Thin-membrane models can be extended to irregular grids, as illustrated in Fig. 1(b). The adjacency structure may be generated automatically by Delaunay triangulation, cf. [10], which maximizes the minimum angle for all the triangles in the grid. In this case, $\mathcal{N}(i)$ denotes all the nodes that have direct connection with node i. As a natural extension of (2), the non-zero entries in K_p may be defined as $[K_p]_{i,j} = -1/d_{i,j}^2$, where $d_{i,j}$ is the distance between node i and j. The diagonal elements in the adjacency matrix are given by $[K_p]_{i,i} = -\sum_{j=1,j \neq i}^{p} [K_p]_{i,j}$. We refer to this model as the irregular thin-membrane model. Note that the regular thin-membrane model is a special case of the irregular thin-membrane model, where all the nodes are located on a regular grid, and all distances $d_{i,j}$ are identical.

As pointed out in [9], for some applications the off-diagonal entries $[K_p]_{i,j}$ are not necessarily related to the distance $d_{i,j}$ between node i and node j. More generally, the entries of the precision matrix K may be inferred from the data, without specifying any dependence on the distance $d_{i,j}$. However, the sparsity pattern of K is fixed, as it is specified by the (regular or irregular) grid, i.e., $K_{i,j} \neq 0$ iff edge (i, j) is present. In generalized thin-membrane models, the non-zero entries of K are learned from data, for a fixed sparsity pattern determined by the grid (cf. Fig. 1).

III. GEV MARGINALS

In this section, we describe how we infer the GEV marginal distributions at each site. Suppose that we have n samples $x_i^{(j)}$ (block maxima) at each of the p locations, where $i =$

$1, \cdots, p$ and $j = 1, \cdots, n$. Our objective is to infer the three GEV parameters with the consideration of spatial dependence. Specifically, we first fit the GEV distribution to block maxima at each location using the Probability-Weighted Moment (PWM) method [13]. Those GEV parameter estimates are then smoothed by means of thin-membrane model priors [7], both for regular and irregular grids. At last, the smoothness parameter α for the whole grid is inferred from the data by expectation maximization (EM).

A. Local Estimates of GEV Parameters

We assume that the block maxima $x_i^{(j)}$ at each location follow a Generalized Extreme Value (GEV) distribution [4]:

$$F(x_i) = \begin{cases} \exp\{-[1 + \frac{\gamma_i}{\sigma_i}(x_i - \mu_i)]^{-\frac{1}{\gamma_i}}\}, & \gamma_i \neq 0 \\ \exp\{-\exp[-\frac{1}{\sigma_i}(x_i - \mu_i)]\}, & \gamma_i = 0, \end{cases} \quad (3)$$

for $1 + \gamma_i/\sigma_i(x_i - \mu_i) \geq 0$ if $\gamma_i \neq 0$ and $x_i \in \mathbb{R}$ if $\gamma_i = 0$, where $\mu_i \in \mathbb{R}$ is the location parameter, $\sigma_i > 0$ is the scale parameter and $\gamma_i \in \mathbb{R}$ is the shape parameter.

We estimate the parameters μ_i, σ_i, and γ_i locally at each site i by the Probability-Weighted Moment (PWM) method [13], which aims to match the moments $E[X_i^t(F(X_i))^r(1 - F(X_i))^s]$ with the empirical ones, where t, r and s are real numbers. For the GEV distribution, $E[X_i(F(X_i))^r]$ (with $t = 1$ and $s = 0$) can be written as:

$$b_r = \frac{1}{r+1}\{\mu_i - \frac{\sigma_i}{\gamma_i}[1 + (r-1)^{\gamma_i}\Gamma(1 - \gamma_i)]\}, \quad (4)$$

where $\gamma_i < 1$ and $\gamma_i \neq 0$, and $\Gamma(\cdot)$ is the gamma function. The resulting PWM estimates $(\hat{\mu}_i^{\text{PWM}})$, $(\hat{\sigma}_i^{\text{PWM}})$ and $(\hat{\gamma}_i^{\text{PWM}})$ are the solution of the following system of equations:

$$\begin{cases} b_0 = \mu_i - \frac{\sigma_i}{\gamma_i}(1 - \Gamma(1 - \gamma_i)) \\ 2b_1 - b_0 = \frac{\sigma_i}{\gamma_i}\Gamma(1 - \gamma_i)(2^{\gamma_i} - 1) \\ \frac{3b_2 - b_0}{2b_1 - b_0} = \frac{3^{\gamma_i} - 1}{2^{\gamma_i} - 1}. \end{cases} \quad (5)$$

B. Spatial-Dependent Estimates of GEV Parameters

We assume that each of the three parameter vectors $\mu = (\mu_1, \cdots, \mu_p)$, $\gamma = (\gamma_1, \cdots, \gamma_p)$ and $\sigma = (\sigma_1, \cdots, \sigma_p)$ has a thin-membrane model as prior, where p is the number of sites. Since the thin-membrane models of μ, γ, and σ share the same structure and inference methods, we present the three models in a unified form.

Let $y = (y_1, y_2, \ldots, y_p)$ denote the local estimates of $z = (z_1, z_2, \ldots, z_p)$, where y is either $(\hat{\mu}^{\text{PWM}})$, $(\hat{\sigma}^{\text{PWM}})$, or $(\hat{\gamma}^{\text{PWM}})$ and z is either μ, σ, or γ after considering spatial dependence. We model the local estimates as $y = z + b$, where $b \sim N(0, R_z)$ is zero-mean Gaussian random vector with diagonal covariance matrix R_z.

As a result of this Gaussian approximation, the conditional distribution of the observed value y given the true value z is a Gaussian distribution:

$$P(y|z) \propto \exp\{-\frac{1}{2}(y - z)^T R_z^{-1}(y - z)\}. \quad (6)$$

Since we assume that the prior distribution of z is a thin-membrane model (cf. (2)), the posterior distribution is given by:

$$P(z|y) \propto \exp(-\alpha_z z^T K_p z)\exp\{-\frac{1}{2}(y - z)^T R_z^{-1}(y - z)\}$$
$$\propto \exp\{-\frac{1}{2}z^T(\alpha_z K_p + R_z^{-1})z + z^T R_z^{-1}y\}. \quad (7)$$

The maximum a posteriori estimate of z is given by:

$$\hat{z} = \text{argmax}\, P(z|y) = (\alpha_z K_p + R_z^{-1})^{-1}R_z^{-1}y. \quad (8)$$

The noise covariance R_z can be estimated by the bootstrap approach described in [7], [14].

We infer the parameter α_z by expectation maximization. In the E-step, we compute [15]:

$$Q(\alpha_z, \hat{\alpha}_z^{(k-1)}) = E_{Z|y, \hat{\alpha}_z^{(k-1)}}[\log P(y, Z|\alpha_z)] \quad (9)$$
$$= -\frac{1}{2}\alpha_z\{\text{trace}[K_p(\hat{\alpha}_z^{(k-1)}K_p + R_z^{-1})^{-1}]$$
$$+ (\hat{z}^{(k-1)})^T K_p \hat{z}^{(k-1)}\} + \frac{1}{2}\log\det(\alpha_z K_p),$$

where $\hat{z}^{(k)}$ is computed as in (8), and α_z is replaced by $\hat{\alpha}_z^{(k)}$. Note that $\hat{z}^{(k)}$ is the MAP estimate of z conditioned on $\hat{\alpha}_z^{(k)}$ and y. Since the posterior distribution in z is Gaussian, the MAP estimate $\hat{z}^{(k)}$ is also the mean of the (Gaussian) posterior of z. In the M-step, we select the value $\hat{\alpha}_z^{(k)}$ of α_z that maximizes $Q(\alpha_z, \hat{\alpha}_z^{(k-1)})$. A closed form expression of $\hat{\alpha}_z^{(k)}$ exists, cf. [15], and is given by:

$$\hat{\alpha}_z^{(k)} = \frac{p}{\text{trace}[K_p(\hat{\alpha}_z^{(k-1)}K_p + R_z^{-1})^{-1}] + (\hat{z}^{(k-1)})^T K_p \hat{z}^{(k-1)}}, \quad (10)$$

where p is the number of sites. We iterate the E-step and M-step until convergence, yielding a local extremum of the marginal posterior of α_z.

IV. COPULA GAUSSIAN GRAPHICAL MODEL

We capture the spatial dependence between the extreme values (block maxima) $x_i^{(j)}$ at the each of the p locations by means of a copula Gaussian graphical. Also for this purpose, we use a thin-membrane model. The smoothness parameter $\alpha_{i,j}$ now varies for each pair of neighbors in the grid (both regular and irregular), adding extra flexibility to the model. The sparsity structure K_p of K is fixed, as it corresponds to a thin-membrane model. The non-zero elements of K are inferred from data.

In the following, we denote the observed GEV variables and hidden Gaussian variables as $X_1, \ldots, X_P$ and $Z_1, \ldots, Z_P$ respectively. A copula Gaussian graphical model is defined as [16]:

$$Z \sim \mathcal{N}(0, K^{-1}) \quad (11)$$
$$X_i = F_i^{-1}(\Phi(Z_i)), \quad (12)$$

where K is the precision matrix whose inverse K^{-1} (covariance matrix) has normalized diagonal, Φ is the cdf of the standard Gaussian distribution, and F_i is the marginal GEV cdf of Y_i with corresponding parameters μ_i, γ_i and σ_i. Note that F_i^{-1} is the pseudo-inverse of F_k, which is defined as:

$$F^{-1}(x) = \inf_{y \in \mathcal{Y}} \{F(y) \geq x\}. \tag{13}$$

where y takes values in $\mathcal{Y}$.

According to definition [16], a copula Gaussian graphical model is determined by the marginals F_i and the precision matrix K. The marginals F_i are GEV distributions, as described in Section III. The spatial dependence among the GEV marginals F_i is captured by coupling the GEV parameters through thin-membrane models (cf. Section III).

In the following, we describe how we infer the precision matrix K. As a first step, we transform the non-Gaussian observed variables X into Gaussian distributed latent variables Z:

$$Z_i = \Phi^{-1}(F_i(X_i)). \tag{14}$$

In the second step, for given thin-membrane sparsity structure K_p, the precision matrix K is estimated from the latent Gaussian variables Z [17]:

$$\hat{K} = \underset{K \succ 0}{\mathrm{argmax}} \log \det K - \mathrm{trace}(SK), \tag{15}$$
$$s.t. \quad K_{i,j} = 0 \quad \forall (i,j) \notin K_p,$$

where S is the empirical covariance of latent variables Z. The convex optimization problem (15) can be solved efficiently by the Newton-CG primal proximal point algorithm [17] or iterative proportional fitting [18].

V. Interpolation

Here we explain how extreme values can be inferred at any location P_0 in space, including sites without observations. We assume that the surface of the extreme values is smooth across space. Since both spatial-dependent GEV parameters μ, γ, σ and the latent variables Z in (11) share a (generalized) thin-membrane structure, we will first formulate the interpolation problem in a unified form, and then describe the minor differences.

Let x represent the parameter vectors μ, γ, σ, or hidden variables Z, associated to the sites with measurements, and let x_0 denote the interpolated value at site P_0 (without measurements). The random variables x and x_0 are assumed to form a thin-membrane model with joint precision matrix:

$$K_0 = \begin{pmatrix} K_{x_0} & K_{x_0,x} \\ K_{x,x_0} & K_x \end{pmatrix}. \tag{16}$$

The conditional expected value of x_0 therefore equals:

$$E[x_0|x] = -K_{x_0}^{-1} K_{x_0,x}\, \hat{x}, \tag{17}$$

where $\hat{x}$ is the expected value of x.

Since we consider thin-membrane models, x_0 is conditionally dependent on its neighbors only. For a standard thin-membrane model (on regular or irregular grid) as illustrated
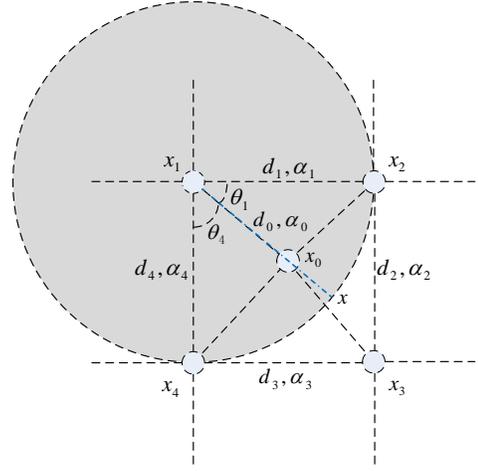
Fig. 2. Illustration of the modified interpolation method

in Fig. 2, the expected value (17) can be simplified as:

$$E[x_0|x] = \frac{K_{x_0,x_1} x_1 + K_{x_0,x_2} x_2 + K_{x_0,x_3} x_3 + K_{x_0,x_4} x_4}{K_{x_0,x_1} + K_{x_0,x_2} + K_{x_0,x_3} + K_{x_0,x_4}}. \tag{18}$$

Clearly, due to the sparse precision matrix corresponding to a thin-membrane model, the expression (18) is quite simple. It is natural to choose $K_{x_0,x_j} = -\alpha/d_{0,j}^2$ in standard thin-membrane models (both on regular and irregular grids. Consequently, the weighted average (18) boils down to inverse distance weighted (IDW) interpolation [11]. We use IDW to interpolate the GEV parameters.

For the latent Gaussian variables Z, however, the thin-membrane precision matrix is not a function of distance. On the other hand, for interpolation purposes, we need to consider the distance from sites with observations, and potentially also other parameters. Here we propose a modified inverse distance weighted (MIDW) interpolation method, where K_{x_0,x_j} is not only a function of distance but of direction as well. We assume that K_{x_0,x_j} changes linearly with direction when the distance remains unchanged and is proportional to the inverse square distance when the direction remains unchanged. As illustrated in Fig. 2, K_{x_0,x_1} for instance can be computed as:

$$\alpha_{x_0,x_1} = \frac{1}{\frac{\pi}{2}}\left(\alpha_1 \theta_4 + \alpha_4 \theta_1\right), \tag{19}$$

$$K_{x_0,x_1} = -\frac{\alpha_{x_0,x_1}}{d_{x_0,x_1}^2}. \tag{20}$$

After interpolating the parameters μ_0, γ_0, σ_0, and hidden variable Z_0 for site P_0, we obtain the GEV distributed value of site P_0 through (12).

VI. Numerical Results

In this section, we benchmark the proposed copula MRF-GEV model against the MRF-GEV model (without modeling the extreme value dependence) [7], copula GEV (without modeling the GEV parameter dependence), and thin-membrane model directly fitted to the data according to (15) (where S is the empirical covariance matrix of the observations), both on synthetic and real data sets. We compare all four models

by means of three criteria: the mean square error between the interpolated extreme value and the true value, the KL-divergence, and the number of parameters.

For synthetic data, we also compute the mean square error (MSE) for inferring the GEV parameters. Specifically, we report the MSE for (i) local PWM estimates; (ii) spatial-dependent estimates from copula MRF-GEV model; (iii) IDW interpolation from copula MRF-GEV model at unobserved sites.

A. Synthetic Data

We generate spatially dependent GEV distributed synthetic data as follows:

1) We generate the coordinates of the 256 observed and 400 unobserved sites. We consider two cases: First, the observed sites are arranged in a regular grid (e.g., wave height measuring stations in the Gulf of Mexico [19]) while unobserved sites are randomly distributed across the grid. Second, both observed and unobserved sites are randomly distributed in the same spatial domain (e.g., precipitation measuring stations in South Africa [8]).

2) We generate 315 samples from a zero-mean multivariate Gaussian distribution, both at observed and unobserved sites. The covariance matrix of that distribution is defined as $\Sigma_{i,j} = \exp(-d_{i,j}^2/\phi)$, where ϕ is the range parameter.

3) We select GEV parameters for each site. All GEV parameters vary smoothly across space. For both case studies, the location parameter surface is a quadratic Legendre polynomial, as shown in Fig. 3(a). The other parameters are chosen differently in each case, as we will explain later.

4) We transform the Gaussian samples generated in Step 2 to GEV distributed samples with the GEV parameters chosen in Step 3 using (12).
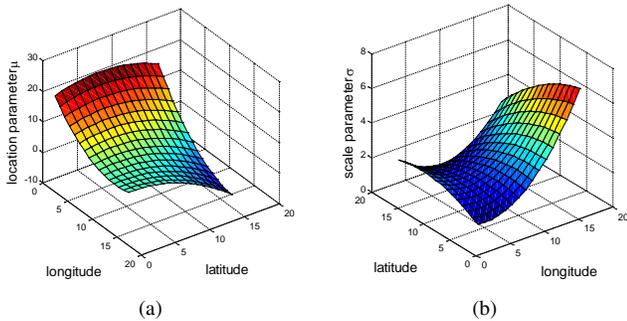
Fig. 3. True GEV parameters for synthetic data: (a) Location parameter surface; (b) Scale parameter surface.

1) Case Study 1: The observed sites are located on a 16x16 grid, whereas the observed sites are randomly distributed, as shown in Fig. 4. The shape and scale parameters γ and σ are chosen to be constant, and are equal to 0.4 and 2 respectively. The results for GEV parameter estimation are summarized in Table I. As mentioned earlier, we report the MSE for (i) local PWM estimates; (ii) spatial-dependent estimates from copula
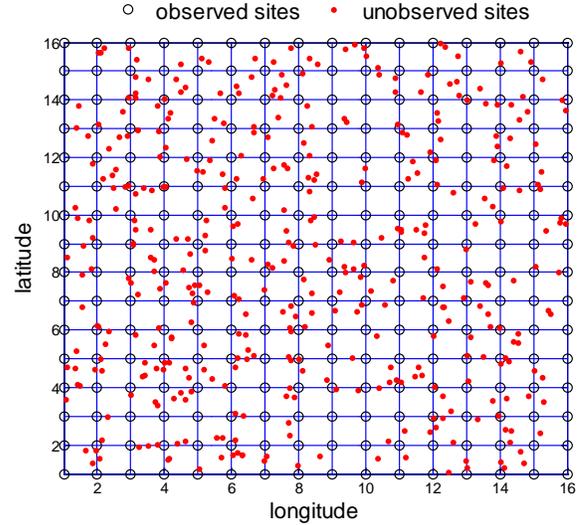
Fig. 4. Coordinates of observed and unobserved sites for Case Study 1.

MRF-GEV model; (iii) IDW interpolation from copula MRF-GEV model at unobserved sites. The GEV estimates by the

TABLE I
MEAN SQUARE ESTIMATION ERROR FOR GEV PARAMETERS IN CASE STUDY 1

GEV parameter	PWM	copula MRF-GEV	IDW interpolation
shape parameter γ	0.0013	9.0334×10^{-4}	9.0334×10^{-4}
scale parameter σ	0.0093	0.0084	0.0084
location parameter μ	0.0114	0.0113	0.0260

copula MRF-GP model are more accurate than the local PWM estimates, although the difference is minor. The PWM method often results in accurate estimates of GEV parameters [13]. The corresponding smoothness parameter $\alpha_\mu = 0.3370$ while α_γ and α_σ converge to infinity. As a consequence, the shape and scale parameter do not depend on location, in agreement with the true parameter values. As can also be seen from Table I, IDW parameter interpolation, based on estimates from the copula MRF-GP model, generates accurate estimates of the GEV parameters at unobserved sites.

TABLE II
COMPARISON FOR CASE STUDY 1

Models	MSE	KL-divergence	No. of Parameters
copula MRF-GEV	0.0497	301.8436	738
MRF-GEV	81.0440	646.8706	258
copula GEV	116.0777	302.3067	1248
Gaussian	0.7025	804.4967	736

Table II summarizes the performance of the four methods for inferring extreme values. The proposed copula MRF-GEV model has the smallest MSE and KL divergence. Both MRF-GEV and copula GEV fail to describe the spatial extreme values suitably, since they only capture one of the two types of spatial dependence, motivating our approach to model the spatial dependence both for the parameter and extreme values.

Interestingly, the Gaussian model has the second best performance in terms of MSE, probably due to the smooth nature of the extreme value surface. However the KL divergence for the Gaussian model is large compared to the other models, since Gaussian models are not capable of capturing extreme events; such models mostly describe fluctuations around the mean value. Compared with the copula GEV model, the copula MRF-GEV model achieves a smaller KL divergence with fewer parameters, suggesting that it is beneficial to model the spatial dependence of the GEV parameters.

We further set the scale parameter surface to be quadratic instead of constant, as shown in Fig. 3(b). The results are qualitatively similar. The only difference is that α_σ is now finite (4.9831), implying that the estimates of σ are no longer independent of location. This is not surprising since the true parameter σ follows a quadratic surface. In other words, by inferring smoothness parameters, the smoothness of the GEV parameters can be automatically and suitably adjusted.

2) Case Study 2: In the second scenario, both the observed and unobserved sites are randomly distributed in space, as shown in Fig. 5. The adjacency structure of the corresponding irregular thin-membrane model is generated automatically using Delaunay triangulation [10], indicated by the blue lines in the figure.
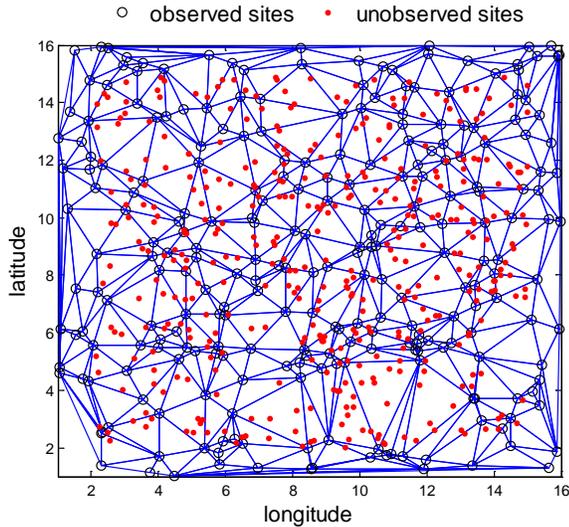
Fig. 5. Coordinates of observed and unobserved sites for Case Study 2.

As in the first scenario, we first set the shape and scale parameters γ and σ to be constant, and equal to 0.4 and 2 respectively. The results of parameter estimation are presented in Table III. The proposed copula MRF-GP model yields smaller estimation error than the PWM local estimates. The smoothness parameters α_γ and α_σ converge to infinity while α_μ is finite (0.2188), in agreement with the true underlying model.

Table IV summarizes the comparison of the four models. It can be seen that the proposed copula MRF-GEV outperforms the other methods both in terms of MSE and KL divergence,

TABLE III
MEAN SQUARE ESTIMATION ERROR FOR GEV PARAMETERS IN CASE STUDY 2

GEV parameter	PWM	copula MRF-GEV	IDW interpolation
shape parameter γ	2.8527×10^{-4}	9.9447×10^{-5}	9.9447×10^{-5}
scale parameter σ	0.0126	0.0126	0.0084
location parameter μ	0.0157	0.0155	0.0650

TABLE IV
COMPARISON FOR CASE STUDY 2

Models	MSE	KL-divergence	No. of Parameters
copula MRF-GEV	0.1869	150.7595	1009
MRF-GEV	95.0509	652.2792	258
copula GEV	99.1437	151.0429	1519
Gaussian	0.2283	690.6725	1007

which suggests that the method is also suitable for irregular grids.

Comparing the results with Case Study 1, we notice that by introducing more parameters, the KL-divergence between the copula MRF-GEV model and the data is reduced. Meanwhile, due to the random location of the observed sites, there is a lack of information in some areas compared to the regular grid, resulting in a larger mean square error of interpolation.

Next, as in the first scenario, we set the scale parameter surface to be a quadratic Legendre polynomial as shown in Fig. 3(b). Consistent with the true model, the resulting α_σ is finite (2.8716).

B. Real Data

In this section, we consider the GOMOS (Gulf of Mexico Oceanographic Study) data [19], which consists of 315 maximum peak wave height values; each corresponds to a hurricane event in the Gulf of Mexico. The distance between each pair of neighbors is $0.125°$ (approximately 14km).
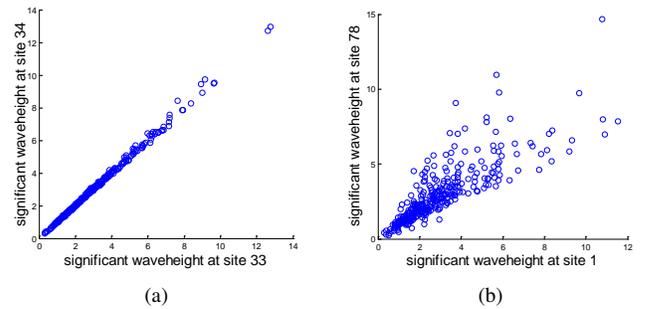
Fig. 6. Scatter plots of wave height at pairs of sites: (a) two distant sites; (b) two nearby sites.

As can be seen from Fig. 6, strong spatial dependence exists between two nearby sites, and even between sites that are far apart. Fig. 7 shows that the maximum peak wave heights vary smoothly over space.

First we analyze a 31x31 lattice in a central region of the Gulf of Mexico (see Fig. 8(b)). The nodes of a 16x16 regular sublattice are chosen as locations with observations
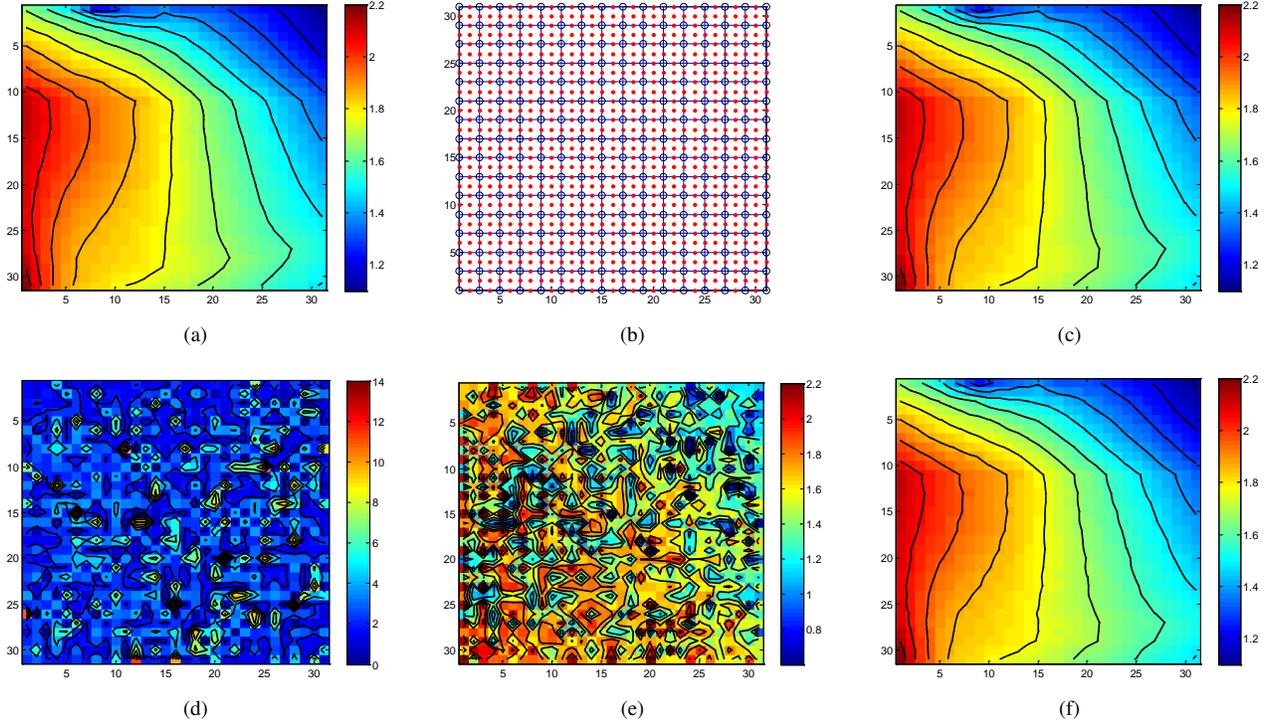
Fig. 8. Interpolation of the maximum wave heights caused by a storm (small regular subgrid). (a) True graph; (b) Observed sites (black) and unobserved sites (red) on the grid; Interpolation by (c) the copula MRF-GEV model; (d) MRF-GEV model; (e) copula GEV model; (f) Gaussian model.
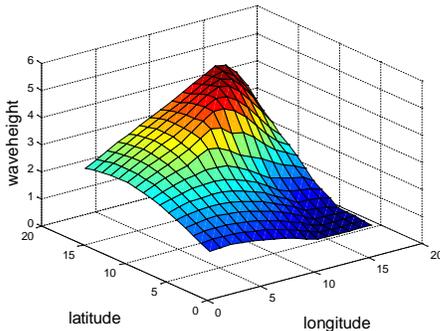
Fig. 7. The maximum wave height corresponding to one storm for 256 sites located on 16x16 lattice.

TABLE V
COMPARISON FOR REGULAR GRID IN THE GULF OF MEXICO

Models	MSE	KL-divergence	No. of Parameters
copula MRF-GEV	0.0011	238.3892	993
MRF-GEV	17.0022	513.9034	258
copula GEV	1.3863	240.3537	1248
Gaussian	0.0023	506.9120	736

sites in the Gulf of Mexico. We randomly select 1000 sites as the locations with measurements, while the remaining sites are regarded as locations without measurements. Fig. 9 shows the interpolation by all four methods, for one storm.

Again, the copula MRF-GEV model seems to yield the lowest interpolation error. The Gaussian model also performs well but its contour plot fluctuates more than the copula MRF-GEV model, suggesting that the estimates are less reliable. On the other hand, the other two methods cannot correctly interpolate the values for the unobserved sites, since they fail to capture the spatial dependencies among the GEV parameters and the wave heights at different locations.

TABLE VI
COMPARISON FOR IRREGULAR GRID IN THE GULF OF MEXICO

Models	MSE	KL-divergence	No. of Parameters
copula MRF-GEV	0.0130	524.4927	5944
MRF-GEV	7.6933	1.7159×10^3	3000
copula GEV	1.6094	529.2659	5944
Gaussian	0.0254	791.1372	3944

(indicated by black circles in the figure), while the other nodes in the 31x31 lattice are treated as sites without observations. The interpolation results for one storm event are shown in Fig. 8. The figure suggests that the copula MRF-GEV model outperforms the other 3 models in terms of interpolation accuracy. Results for all 315 storms are summarized in Table V, listing the interpolation mean square error (MSE), KL divergence, and number of parameters for each model. The proposed copula MRF-GEV model achieves the smallest KL divergence and MSE with only relatively few parameters. The resulting smoothness parameters $\alpha_\gamma = \infty$, $\alpha_\sigma = 127.1875$ and $\alpha_\mu = 78.5047$. The parameter γ is constant, whereas both σ and μ fluctuate across space.

Now we consider the irregular grid with all 4363 measuring

Table VI shows quantitative results of all four models, com-
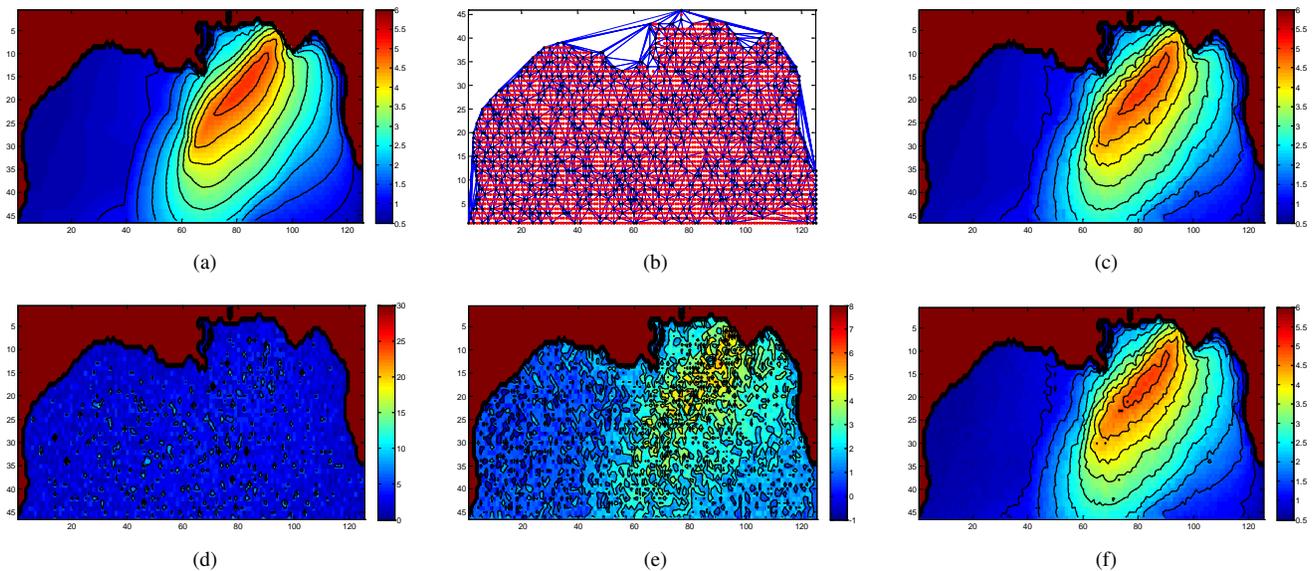
Fig. 9. Interpolation of the maximum wave heights caused by a storm (irregular grid covering the Gulf of Mexico). (a) True graph; (b) Observed sites (black) and unobserved sites (red) on the grid; Interpolation by (c) the copula MRF-GEV model; (d) MRF-GEV model; (e) copula GEV model; (f) Gaussian model.

puted from all hurricane events. The corresponding smoothness parameters for the copula MRF-GEV model are $\alpha_\gamma = 4.8860 \times 10^3$, $\alpha_\sigma = 148.0747$ and $\alpha_\mu = 83.9102$. None of the smoothness parameters converge to infinity, suggesting significant spatial variations in the GEV parameters across the Gulf of Mexico.

VII. CONCLUSIONS

In this study, we have presented a new model for quantifying dependencies among spatial extreme events. The dependency of the extreme events across space is modeled through a thin-membrane Gaussian copula. Also the parameters of the GEV marginals are coupled in space through thin-membrane models.

Numerical results show that the proposed model not only provides accurate estimates and predictions, it is also computationally efficient and practical for both regular and irregular grids, even for large regions such as the entire Gulf of Mexico.

The present model has several limitations. The monoscale thin-membrane model cannot capture the long-range dependencies effectively in a large spatial domain. In future work, we will attempt to utilize multiscale models instead. In addition, the model does not consider other covariate effects (e.g., direction, season) and this will be further explored in detail.

REFERENCES

[1] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling extremal events for insurance and finance*, Berlin: Spring Verlag, 1997.
[2] J. E. Heffernan and J. A. Tawn, "A conditional approach for multivariate extreme value", *J.R. Statist. Soc.* B 66, pp.1–34, 2004.
[3] C. Keef, C. Svensson, and J. A. Tawn, "Spatial dependence in extreme river flows and precipitation for Great Britain," *Journal of Hydrology*, vol. 378, pp. 240–252, 2009.
[4] S. G. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer, London, 2001.
[5] P. Naveau, A. Guillou, D. Cooley, and J. Diebolt, "Modelling pairwise dependence of maxima in space," *Biometrika*, vol. 96(1), pp. 1-17, 2009.
[6] P. J. Northrop and P. Jonathan, "Threshold modelling of spatially dependent non-stationary extremes with application to hurricane-induced wave heights," *Environmetrics*, vol. 22, pp. 799–809, 2011.
[7] H. Yu, Z. Choo, J. Dauwels, P. Jonathan, and Q. Zhou, "Modeling Spatial Extreme Events using Markov Random Field Priors," submitted to *ISIT 2012*.
[8] H. Sang and A. E. Gelfand, "Continuous Spatial Process Models for Extreme Values", *Journal of Agricultural, Biological and Environmental Statistics*, 15 ,49-65, 2010.
[9] F. Lindgren and H. Rue, "An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach," *J.R. Statist. Soc.*, ser. B, vol. 73, pp. 423-498, 2011.
[10] Ø. Hjelle and M. Dæhlen, *Triangulations and Applications*, Berlin: Springer, 2006.
[11] L. Mitas and H. Mitasova, "Spatial interpolation," *In: Longley, P., Goodchild, M., Maguire, D., Rhind, D., (Eds.), Geographical Information Systems: Principles, Techniques, Management and Applications*, vol. 1, pp. 481-492, 1999.
[12] E. Casson and S. Coles, "Spatial regression models for extremes," *Extremes* vol. 1 pp. 449C468, 1999.
[13] J.R.M. Hosking, J. R. Wallis and E. F. Wood, "Estimation of the generalized extreme-value distribution by the method of probability-weighted moments", *Technometrics*, vol. 27, pp. 251-261, 1985.
[14] P. Jonathan and K. Ewans, "Uncertainties in extreme wave height estimates for hurricane-dominated regions," *Journal of offshore mechanics and arctic engineering*, vol. 129(4), pp. 300-305, 2007.
[15] M. J. Choi, V. Chandrasekaran, D. M. Malioutov, J. K. Johnson, and A. S. Willsky, "Multiscale stochastic modeling for tractable inference and data assimilation", Comput. Methods Appl. Mech. Engrg. Vol. 197, pp. 3492-3515, 2008.
[16] A. Dobra and A. Lenkoski, "Copula Gaussian graphical models and their application to modeling functional disability data," *Annals of Applied Statistics,* vol. 5, No. 2A, pp. 969-993, 2011.
[17] C. Wang, D. Sun, and K. C. Toh, "Solving log-determinant optimization problems by a Newton-CG primal proximal point algorithm," *Society for Industrial and Applied Mathematics,* vol. 20, pp. 2994-3013, 2009.
[18] P. F. Xu, J. Guo, and X. He, "An Improved Iterative Proportional Scaling Procedure for Gaussian Graphical Model," *Journal of Computational and Graphical Statistics*, vol. 20, pp. 417-431, 2011.
[19] Oceanweather Inc. *GOMOS - USA Gulf of Mexico Oceanographic Study, Northen Gulf of Mexico Archive*, 2005.