EXTREME-VALUE GRAPHICAL MODELS WITH MULTIPLE COVARIATES

Hang Yu, Jingjing Cheng, and Justin Dauwels

School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore

ABSTRACT

To assess the risk of extreme events such as hurricanes and floods, it is crucial to develop accurate extreme-value statistical models. Extreme events often display heterogeneity, varying continuously with a number of covariates. Previous studies have suggested that models considering covariate effects lead to reliable estimates of extreme value distributions. In this paper, we develop a novel model to incorporate the effects of multiple covariates. Specifically, we analyze as an example the extreme sea states in the Gulf of Mexico, where the distribution of extreme wave heights changes systematically with location and wind direction. The block maxima at each location and sector of wind direction are assumed to follow the Generalized Extreme Value (GEV) distribution. The GEV parameters are coupled across the spatio-directional domain through a graphical model, particularly, a multidimensional thin-membrane model. Efficient learning and inference algorithms are then developed based on the special characteristics of the thin-membrane model. Numerical results for both synthetic and real data indicate that the proposed model can accurately describe marginal behavior of extreme events.

Index Terms— extreme events modeling, graphical model, covariates, Laplacian matrix, Kronecker product

1. INTRODUCTION

Extreme events, such as heat waves, cold snaps, tropical cyclones, hurricanes, heavy precipitation and floods, droughts and wild fires, have possibly tremendous impact on people's life and properties. Furthermore, both observational data and computer climate models suggest that the occurrence and sizes of such catastrophes will increase in the future [1]. It is therefore imperative to model such events, assess the risk, and further take precaution measurements.

Extreme-value theory governs the statistical behavior of extremes, such as block maxima (e.g., monthly or annually) and peaks over a high enough threshold [2]. More specifically, the theory provides closed-form distribution functions for extreme values. The main challenge in fitting such distributions is the lack of data, as extreme events are by definition very rare. The problem may be addressed by assuming that all the collected data (e.g., extreme wave heights at different measuring sites [3]) follow the same distribution, resulting in sufficiently large sample size. However, there usually exists clear heterogeneity in the extreme value data. Extreme temperature, for instance, is greatly influenced by the altitude of

the measuring site. The latter is regarded as a covariate. Accommodating the heterogeneity in the model is essential since the estimated models will be unreliable otherwise [4]. A fruitful approach is to assume that the parameters of the extreme value distributions vary smoothly with covariates; such prior knowledge can significantly help to improve the fitting.

Extreme-values models have been developed with single [5, 6, 7, 8, 9] and multiple covariates [10, 11]. The first group usually treats the parameters of the marginal distributions as a function of the covariate and apply regression methods [6, 7, 8] or Bayesian fitting [9] to learn the model. However, such procedures are computationally complex and can be prohibitive for large-scale systems. On the other hand, few attempts have been made to model multiple covariates. A standard method is to predefine the distribution parameters as a function of all covariates [10]. Unfortunately, only linear or log-linear models have been considered so far, which are rather limited. As an alternative, Jonathan *et al.* [11] proposed to process two covariates individually. However, capturing the two covariates independently fails to consider the possible correlation between them.

The aforementioned shortcomings spark our interest in exploiting graphical models to incorporate multiple covariates in extreme value models. The interdependencies between extreme values with different covariates are often highly structured, and thus, can be leveraged by the graphical model framework to yield efficient algorithms [12, 13, 14, 15]. As an example, we model the storm-wise maxima of significant wave heights in the Gulf of Mexico (see Fig. 3), where the covariates are longitude, latitude, and wind direction. The extreme events are assumed to follow Generalized Extreme Value (GEV) distributions [2]. The parameters of those GEV distributions are assumed to depend smoothly on the covariates. To facilitate the use of graphical models, we discretize the continuous covariates within a finite range. In the example of extreme-wave heights in the Gulf of Mexico, space is discretized as a finite homogeneous two-dimensional lattice, and the wind direction is discretized in a finite number of equal-sized sectors. More generally, the GEV distributed variables (and hence also the GEV parameters) are defined on a finite number of points indexed by the (discretized) covariates. We characterize the dependence between the GEV parameters through a graphical model prior, in particular, a multidimensional thin-membrane model where edges are

only present between pairs of neighboring points (see Fig. 1). We demonstrate that the multidimensional model can be constructed flexibly from one-dimensional thin-membrane models for each covariate. The proposed model can therefore easily be extended to an arbitrary number of covariates. Efficient learning and inference algorithms are proposed based on the special pattern of the eigenvalues and eigenvectors of the one-dimensional thin-membrane models. Our numerical results suggest that the proposed model indeed accurately captures the effect of covariates on the statistics of extreme events. Moreover, the proposed model can flexibly accommodate a large variety of (smooth) dependencies on covariates, as the smoothness parameters of the thin-membrane model are inferred from data.

The paper is organized as follows. We briefly review thinmembrane models in Section 2. We introduce our proposed extreme-value graphical models with multiple covariates in Section 3. Results for both synthetic and real data are presented in 4. Lastly, we offer concluding remarks in Section 5.

2. THIN-MEMBRANE MODELS

In this section, we first give a brief introduction to graphical models, and then consider two specific cases of thinmembrane models, which serve as building blocks in the proposed extreme-value graphical model.

In an undirected graphical model, the probability distribution 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 Z_i . An edge (i, j) is absent if the corresponding two variables Z_i and Z_j are conditional independent: $P(Z_i, Z_j | Z_{\mathcal{V}|i,j}) = P(Z_i | Z_{\mathcal{V}|i,j})P(Z_j | Z_{\mathcal{V}|i,j})$, where $\mathcal{V}|i, j$ denotes all the variables except Z_i and Z_j . In particular, for Gaussian distributed Z, the graph \mathcal{G} is characterized by the inverse covariance matrix (precision matrix) K, i.e., $K(i, j) \neq 0$ if and only if the edge $(i, j) \in \mathcal{E}$ [16].

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:

$$P(Z) \propto \exp\{-\alpha \sum_{i \in \mathcal{V}} \sum_{j \in N(i)} (Z_i - Z_j)^2\} \propto \exp\{-\alpha Z^T K_p Z\}$$

where K_p is a graph Laplacian matrix [17], and α is the smoothness parameter which controls the smoothness across the domain defined by the thin-membrane model. Since K_p is a Laplacian matrix, the determinant of the precision matrix $K = \alpha K_p$ is 0. The thin-membrane model is a partially informative normal prior, commonly used in spline smoothing [18]. To make the distribution well-defined, the improper density function is usually applied in practice [18]:

$$P(Z) \propto |K|^{0.5}_{+} \exp\{-Z^T K Z\},$$
 (1)

where $|K|_+$ denotes the product of nonzero eigenvalues of K. We now turn our attention to the thin-membrane models of the chain and the circular graph as shown in Fig. 1a and Fig. 1b respectively. The former can well characterize the dependence structure of nonperiodic covariates (e.g., longitude

and latitude), while the latter is suitable for periodic covariates (e.g., direction and season). The corresponding Laplacian matrices are denoted as K_B and K_C respectively. The eigenvalues of K_B and K_C are $\lambda_{B\,k} = 2 - 2\cos(k\pi/P)$ and $\lambda_{C\,k} = 2 - 2\cos(2k\pi/P)$ respectively [19], where P is the dimension of K_B and K_C . Moreover, let V_B and V_C be the eigenvector matrix of K_B and K_C and x be a $P \times 1$ column vector, then $V_B x$ and $V_C x$ amount to discrete cosine transform and discrete Fourier transform respectively of x [19]. The nicely structured eigendecomposition motivates us to use thin-membrane models in the proposed model.

3. EXTREME-VALUE MODELS WITH MULTIPLE COVARIATES

In this section, we introduce our extreme-value graphical model with multiple covariates. We will explain it by the example of extreme wave heights in the Gulf of Mexico (GoM), specifically, (monthly or annual) block maxima of the wave heights. Inspired by extreme-value theory, we assume that the extreme values follow the Generalized Extreme Value (GEV) distribution [2]. The GEV variables are assumed to depend smoothly on several covariates. In our example of the GoM, we assume that the extreme wave heights depend smoothly on location and wind direction. For the sake of simplicity, the covariates are discretized within a finite range; the location is discretized as a finite two-dimensional grid (spanning the GoM), and the wind direction is divided in equal-size sectors. Each location is indexed by its longitude and latitude (i, j) $(i = 1, \dots, P, j = 1, \dots, Q)$, and each directional sector by $k \ (k = 1, \dots, D)$, where $P \times Q$ is the size of the twodimensional grid, and D is the number of directional sectors. We have N samples $x_{ijk}^{(n)}$ (block maxima; $n = 1, \dots, N$) for each location (i, j) and direction k. The GEV parameters, just like the GEV variables, are indexed by (i, j, k). We couple the GEV parameters through thin-membrane models, encoding that the GEV parameters depend smoothly on the covariates. Consequently, the GEV parameters are similar at nearby locations and for similar wind directions. In the following, we provide more details on this construction.

3.1. Local estimates of GEV parameters

We assume that the block maxima $x_{ijk}^{(n)}$ follow a Generalized Extreme Value (GEV) distribution [2]:

$$F(x_{ijk}) = \exp\{-\left[1 + \frac{\gamma_{ijk}}{\sigma_{ijk}}(x_{ijk} - \mu_{ijk})\right]^{-\frac{1}{\gamma_{ijk}}}\}, \quad (2)$$

with shape parameter γ_{ijk} , scale parameter σ_{ijk} , and location parameter μ_{ijk} . The probability-weighted moment method [20] is employed here to infer the local estimates of the GEV parameters ($\hat{\mu}_{ijk}^{\text{PWM}}$), ($\hat{\sigma}_{ijk}^{\text{PWM}}$) and ($\hat{\gamma}_{ijk}^{\text{PWM}}$).

3.2. Prior distribution

We assume that the shape parameter γ_{ijk} , scale parameter σ_{ijk} , and location parameter μ_{ijk} depend smoothly on the covariates, i.e., location (i, j) and wind direction k. For each of the three parameter vectors $\mu = [\mu_{ijk}]$, $\gamma = [\gamma_{ijk}]$ and $\sigma = [\sigma_{ijk}]$, we choose a three-dimensional (3D) thin-membrane model (see Fig. 1d) as prior. Since the thin-membrane models of the three GEV parameters μ , γ , and σ share the same structure and inference methods, we present the three models in a unified form. Let z denote the true GEV parameters, that is, z is either μ , σ , or γ . We next illustrate how to construct the 3D thin-membrane model priors from the fundamental chain and circular graphs. As a first step, we build a regular lattice (see Fig. 1c) from chains. Since both the longitude and the latitude of the measurements are nonperiodic, they can be characterized by the chain graph as shown in Fig. 1a. Let K_{B_x} and K_{B_y} denote the Laplacian matrices of the longitude and the latitude respectively. We further assume that the smoothness across the longitude and the latitude are the same, thus, they can share one common smoothness parameter α_z . The resulting precision matrix of the regular lattice is given by:

 $K_L = (\alpha_z K_{B_x}) \oplus (\alpha_z K_{B_y}) = \alpha_z (K_{B_x} \oplus K_{B_y}),$ (3) where \oplus denotes Kronecker sum, and the identity matrix I_* has the same dimension as K_* . In the second step, we accommodate the effect of wind direction. We discretize the wind direction into D sectors and assume that the true GEV parameters are constant in each sector. The periodic directional dependence comprises a circular graph (Fig. 1b), whose Laplacian matrix is K_C , and another smoothness parameter β_z is introduced to dictate the smoothness across directions. We next integrate the lattice and the circle, and hence, the final precision matrix corresponding to the 3D model (Fig. 1d) is:

 $K_{\text{prior}} = K_L \oplus (\beta_z K_C) = \alpha_z K_s + \beta_z K_d.$ (4) Interestingly, $K_s = K_{B_x} \otimes I_{B_y} \otimes I_C + I_{B_x} \otimes K_{B_y} \otimes I_C$ and $K_d = I_{B_x} \otimes I_{B_y} \otimes K_C$ corresponds to lattices and circles in the graph respectively, suggesting that the former only characterizes the spatial dependence while the latter the directional dependence. Note that \otimes denotes Kronecker product. Based on the property of Kronecker sum, the eigenvalue matrix of K_{prior} is $\Lambda_{\text{prior}} = \alpha_z \Lambda_s + \beta_z \Lambda_d$, where $\Lambda_s = \Lambda_{B_x} \otimes I_{B_y} \otimes I_C + I_{B_x} \otimes \Lambda_{B_y} \otimes I_C$ and $\Lambda_d = I_{B_x} \otimes I_{B_y} \otimes \Lambda_C$, and the eigenvector matrix $V_{\text{prior}} = V_C \otimes V_{B_y} \otimes V_{B_x}$. Following from (1), the density function of the 3D model equals:

$$P(z) \propto |K_{\text{prior}}|_{+}^{0.5} \exp\{-z^T K_{\text{prior}} z\}$$
(5)

$$= |\alpha_z K_s + \beta_z K_d|_+^{0.5} \exp\{-Z^T (\alpha_z K_s + \beta_z K_d)Z\}.$$
 (6)
Note that for an arbitrary number of covariates, the resulting precision matrix can be generalized as:

 $K_{\text{prior}} = (\alpha K_a) \oplus (\beta K_b) \oplus (\gamma K_c) \oplus \cdots,$ (7) where $K_i (i \in \{a, b, c, \cdots\}) \in \{K_B, K_C\}$ are the Laplacian matrices associated with the dependence structures of covariate *i*, and α , β , and γ are corresponding smoothness parameters. The eigenvalue and eigenvector matrix of K_{prior} are:

$$\Lambda_{\text{prior}} = (\alpha \Lambda_a) \oplus (\beta \Lambda_b) \oplus (\gamma \Lambda_c) \oplus \cdots$$
(8)

$$= \alpha \Lambda_a \otimes I_b \otimes I_c \otimes \dots + \beta I_a \otimes \Lambda_b \otimes I_c \otimes \dots$$

$$+\gamma I_a \otimes I_b \otimes \Lambda_c \otimes \dots + \cdots, \qquad (9)$$

$$= \alpha \Lambda_a + \beta \Lambda_b + \gamma \Lambda_c + \cdots, \qquad (10)$$

$$V_{\text{prior}} = \dots \otimes V_c \otimes V_b \otimes V_a.$$
⁽¹¹⁾



Fig. 1: Thin-membrane models: (a) chain graph; (b) circle graph; (c) lattice; (d) spatio-directional model.

3.3. Posterior distribution and inference

Let y denote the local estimates of z, where y is either ($\hat{\mu}^{\text{PWM}}$), $(\hat{\sigma}^{\text{PWM}})$, or ($\hat{\gamma}^{\text{PWM}}$). We model the local estimates as y = z+b, where $b \sim N(0, R_z)$ is zero-mean Gaussian random vector (noise) with diagonal covariance matrix R_z . R_z can be estimated by the parametric bootstrapping [21]. The conditional distribution P(y|z) equals:

$$P(y|z) \propto \exp\{-\frac{1}{2}(y-z)^T R_z^{-1}(y-z)\}.$$
 (12)

Since we assume that the prior distribution of z is the 3D thinmembrane model (6), the posterior distribution is given by:

$$P(z|y) \propto |K_{\text{prior}}|_{+}^{0.5} \exp\{-\frac{1}{2}z^{T}(K_{\text{prior}} + R_{z}^{-1})z + z^{T}R_{z}^{-1}y\}$$

Given the smoothness parameters α_z and β_z , the maximum a posteriori estimate of z is given by:

$$\hat{z} = \operatorname{argmax} P(z|y) = K_{\text{post}}^{-1} R_z^{-1} y, \qquad (13)$$

where $K_{\text{post}} = K_{\text{prior}} + R_z^{-1}$. Inverting K_{prior} is intractable due to the complexity of order $\mathcal{O}(M^3)$. Alternatively, we propose to an efficient algorithm. When the diagonal matrix R_z can be well approximated as a scaled identity matrix cI_z , we have:

 $K_{
m post} \approx K_{
m prior} + cI_z = V_{
m prior}^T (\alpha_z \Lambda_s + \beta_z \Lambda_d + cI_z) V_{
m prior}.$ As a result, z can be computed as

 $z = V_{\text{prior}}^T (\alpha_z \Lambda_s + \beta_z \Lambda_d + cI_z)^{-1} V_{\text{prior}} R_z^{-1} y.$ (14) Note that $V_{\text{prior}} = V_{B_x} \otimes V_{B_y} \otimes V_C$, and thus $V_{\text{prior}} R_z^{-1} y$ is equivalent to reformulating the vector $R_z^{-1} y$ into a $P \times Q \times D$ matrix with the three dimensions representing longitude, latitude, and direction respectively, and then performing the fast cosine transform (FCT) in the first and second dimension while the fast Fourier transform (FFT) in the third one. The resulting computational complexity is $\mathcal{O}(M \log(M))$. Similar algorithm can be easily designed for the case of multiple covariates. When cI_z is not a good approximation to R_z , the Richardson iteration [22] will be performed.

In the case α_z and β_z are unknown, we aim to infer them from the local estimates y. However, since z is unknown, directly inferring α_z and β_z is impossible, and instead we apply Expectation Maximization (EM):

$$(\hat{\alpha}_{z}^{(\kappa)}, \hat{\beta}_{z}^{(\kappa)}) = \operatorname{argmax} Q(\alpha_{z}, \beta_{z}; \hat{\alpha}_{z}^{(\kappa-1)}, \hat{\beta}_{z}^{(\kappa-1)}), \quad (15)$$

where

$$Q(\alpha_z, \beta_z; \hat{\alpha}_z^{(\kappa-1)}, \hat{\beta}_z^{(\kappa-1)}) \propto \log |K_{\text{prior}}|_+ - \operatorname{tr} \left(K_{\text{prior}} \left(K_{\text{post}}^{(\kappa-1)} \right)^{-1} \right) - \left(z^{(\kappa)} \right)^T K_{\text{prior}} z^{(\kappa)},$$
(16)

and $z^{(\kappa)}$ is computed as in (13) with $\hat{\alpha}_z^{(\kappa-1)}$ and $\hat{\beta}_z^{(\kappa-1)}$ obtained from the previous iteration. Since K_{piror} can be

regarded as a generalized Laplacian matrix, $|K_{\text{prior}}|_+$ = $M \det S(K_{\text{prior}})$, where $S(K_{\text{prior}})$ denotes the first M-1rows and columns of the $M \times M$ matrix K_{prior} [17]. By setting the partial derivative of Q function w.r.t. α_z and β_z to 0 and solving the equations, (15) can be simplified as

$$(\alpha_z^{(\kappa)}, \beta_z^{(\kappa)}) = \operatorname{argmax} \log \det S(K_{\text{prior}}),$$
(17)
s.t. $c_1 \alpha_z + c_2 \beta_z = M - 1, \ \alpha_z \ge 0, \ \beta_z \ge 0,$

$$\text{t.} \quad c_1\alpha_z + c_2\beta_z = M - 1, \ \alpha_z \ge 0, \ \beta_z \ge 0$$

and c_1 and c_2 are known constants. Note that $\log \det S(K_{\text{prior}})$ $= \log \det(\alpha_z S(\Lambda_s) + \beta_z S(\Lambda_d)) = \sum_k \log(\alpha_z \lambda_{s\,k} + \beta_z \lambda_{d\,k}),$ thus, (17) can be solved efficiently via bisection method [23].

4. NUMERICAL RESULTS

In this section, we test the proposed spatio-directional model on both synthetic and real data against the locally fit model, the spatial model (only considering the covariate of location), and directional model (only considering direction).

4.1. Synthetic Data

Here we draw samples from GEV distributions with parameters depending on (two-dimensional) location and direction (angle). Concretely, we select 100 sites arranged on a 10×10 homogeneous lattice and discretize the direction in 8 equalsize sectors. We associate GEV parameters with each of the 100 sites and each of the 8 directions. In particular, the shape parameter γ is chosen to be constant across the space whereas varying smoothly across different directions. The scale parameter σ depends on the location according to a quadratic polynomial function, and remains constant for all the directions. The location parameter μ changes smoothly with regard to both location and direction. We then randomly generate 300 GEV distributed samples for each site and direction.



Fig. 2: Estimates of GEV parameters for different directions.

Table 1: (Quantitative	comparison	of	different	models.
------------	--------------	------------	----	-----------	---------

Models	Mean So	quare Erro	BIC		
Widdels	γ	σ	μ	DIC	
Locally fit model	0.0556	0.2160	0.0483	1.1753×10^{6}	
Spatial model	0.0505	0.1471	0.7979	1.2189×10^{6}	
Directional model	0.0503	0.7630	0.1683	1.1709×10^{6}	
Spatio-directional model	0.0233	0.0607	0.0112	1.1649×10^{6}	

Fig. 2 shows the results of GEV parameters estimated by the aforementioned four models across directions. We can see that the estimates resulting from the spatio-directional model follows the ground truth very well. On the contrary, the local estimates are quite noisy for the shape and scale parameters due to the small number of extreme-value samples in each site and wind direction. In addition, the directional model overestimates the location parameters and underestimates the scale

parameters, whereas the spatial model (see Fig.2 (right)) mistakenly ignores the directional variation of the shape and location parameters, suggesting the necessity of capturing all the significant covariates; otherwise, the estimates are inaccurate. Table 1 summarizes the overall mean square error (MSE) for each GEV parameter and the Bayesian Information Criteria (BIC) of model fitting. Apparently, the proposed model yields the smallest MSE while fitting the data the best after considering the penalty of the effective number of parameters. Interestingly, the smoothness parameters α_{γ} and β_{σ} converge to infinity in the proposed spatio-directional model, exactly consistent with the ground truth that γ and σ are constant w.r.t. location and direction respectively.

4.2. Real Data

We now analyze the GOMOS (Gulf of Mexico Oceanographic Study) data [24], which consists of 315 maximum peak wave height values; each corresponds to a hurricane event in the Gulf of Mexico. We select 256 sites arranged on a 16×16 lattice with spacing 0.25° (approximately 28km). As shown in Fig. 3), the extreme wave heights clearly depend on location and direction. Thus, we expect the proposed model is well suited for this data set.



Fig. 3: Non-stationarity in GOMOS data: (a) Histogram of four randomly selected sites; (b) Wave height w.r.t to direction.

We next test the four models. For the proposed model and the directional model, we discretize the wind direction in 6 sectors of 60° . The resulting BIC values for the locally fit model, the spatial model, the directional model and the spatiodirectional model are 3.4979×10^5 , 3.5960×10^5 , 3.3475×10^5 10^5 and 3.3402×10^5 . Again, the spatio-directional model achieves the best BIC score. The directional model also performs well but it ignores the spatial variation which is essential to model the extreme wave heights in the Gulf of Mexico [8]. On the other hand, the spatial model does not properly consider the strong directional variation (see Fig. 3b), while the locally fit model overfits the data by introducing too many parameters. Obviously, the proposed model is preferable.

5. CONCLUSION

We proposed a novel model to efficiently accommodate the covariate effects of extreme events, significantly improving the modeling accuracy. The extension to an arbitrary number of covariates is straightforward.

6. REFERENCES

- M. L. Parry, O. F. Canziani, J. P. Palutikof, van der Linden, and C. E. Hanson, Eds., IPCC, 2007: summary for policymakers, In: *Climate Change 2007: Impacts, Adaptation, and Vulnerability*. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, 2007.
- [2] S. G. Coles, An Introduction to Statistical Modeling of Extreme Values, Springer, London, 2001.
- [3] 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.
- [4] P. Jonathan, K. Ewans, and G. Forristall, "Statistical estimation of extreme ocean environments: the requirement for modelling directionality and other covariate effects," *Ocean Engineering*, vol. 35, pp. 1211–1225, 2008.
- [5] V. Chavez-Demoulin and A. C. Davison. "Generalized additive modelling of sample extremes," *J. Roy. Statist. Soc. Series C (Applied Statistics)*, vol. 54, pp. 207–222, 2005.
- [6] K. C. Ewans and P. Jonathan, "The effect of directionality on Northern North Sea extreme wave design criteria," *J. Offshore Mechanics Arctic Engineering*, vol. 130, 041604, 2008.
- [7] P. Jonathan, K. Ewans, "Modelling the seasonality of extreme waves in the Gulf of Mexico," J. Offshore Mechanics Arctic Engineering, vol. 133, 021104, 2011.
- [8] 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.
- [9] H. Sang, and A. E. Gelfand, "Hierarchical modeling for extreme values observed over space and time," *Environmental and Ecological Statistics* vol. 16, pp. 407–426, 2009.
- [10] E. F. Eastoe and J. A. Tawn, "Modelling non-stationary extremes with application to surface level ozone," *Journal of the Royal Statistical Society*, vol. 58, pp. 25–45, 2009.
- [11] P. Jonathan, K. Ewans, "A spatio-directional model for extreme waves in the Gulf of Mexico," J. Offshore Mechanics Arctic Engineering, vol. 133, 011601, 2011.
- [12] A. T. Ihler, S. Krishner, M. Ghil, A. W. Robertson, and P. Smyth, "Graphical Models for Statistical Inference and Data Assimilation," *Physica D* vol. 230, pp. 72–87, 2007.

- [13] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, P. Li, and F. Kschischang, "The factor graph approach to model-based signal processing," *Proceedings of the IEEE* 95(6), pp. 1295–1322, 2007.
- [14] H. Yu, Z. Choo, J. Dauwels, P. Jonathan, and Q. Zhou, "Modeling Spatial Extreme Events using Markov Random Field Priors," *Proc. of 2012 IEEE International Symposium on Information Theory Proceedings (ISIT)*, pp. 1453–1457, 2012.
- [15] H. Yu, Z. Choo, W. I. T. Uy, J. Dauwels, and P. Jonathan, "Modeling Extreme Events in Spatial Domain by Copula Graphical Models," *Proceedings of 15th International Conference on Information Fusion (Fusion)*, pp. 1761– 1768, 2012.
- [16] J. M. F. Moura and N. Balram, "Recursive Structure of Noncausal Gauss Markov Random Fields," *IEEE Transactions on Information Theory*, IT-38(2):334–354, March 1992.
- [17] X. D. Zhang, "The Laplacian eigenvalues of graphs: a survey," *Preprint, arXiv:1111.2897v1*, 2011.
- [18] P. L. Speckman and D. C. Sun, "Fully Bayesian spline smoothing and intrinsic autoregressive priors," *Biometrika*, vol. 90, pp. 289–302, 2003.
- [19] G. Strang, *Computational Science and Engineering*, Wellesley-Cambridge Press, 2007.
- [20] 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.
- [21] S. Amiri, D. von Rosen, and S. Zwanzig, "On the comparison of parametric and nonparametric bootstrap," U.U.D.M. Report 2008, 15: Uppsala Univ, 2008.
- [22] Y. Saad, Iterative Methods for Sparse Linear Systems, second edn, SIAM, Philadelphia, PA, 2003.
- [23] M. T. Heath, Scientific Computing: An Introductory Survey, McGraw Hill, New York, 2002.
- [24] Oceanweather Inc. GOMOS USA Gulf of Mexico Oceanographic Study, Northen Gulf of Mexico Archive, 2005.