Supergen ORE Hub Award FF2020-1104
Improved Models for Multivariate Metocean Extremes (IMEX)
Project Summary

Ed Mackay (PI)

University of Exeter

January 14, 2022

1 Introduction

The design of offshore renewable energy (ORE) structures requires estimates of the joint extreme values of metocean variables. For example, the design of fixed or floating offshore wind turbines requires estimates of joint (concurrent) extremes of wave heights and wind speeds. Similarly, the design of tidal turbines requires estimates of the joint extremes of wave heights and current speeds, whilst for wave energy converters the joint extremes of wave heights and periods are important.

Current design guidelines (e.g. [1, 2]) recommend statistical methods which require strong assumptions about both the dependence relations between variables and the marginal distributions of each variable. When applied to real datasets with complex dependence structures between variables, these models have been shown to be subject to large biases and uncertainties in estimating joint extreme risk sets [3]. One popular risk set is the environmental contour of the joint distribution. An illustration of environmental contours, estimated under current standards, is shown in Figure 1 for (a) wave heights and wind speeds and (b) wave heights and periods. The contours are clearly a poor match to the data. In case (a) the 50-year wave height when wind speed is around 10m/s is over-estimated. This is a key design condition for an offshore turbine, corresponding to the extreme wave conditions around the rated wind speed where the rotor experiences the highest thrust loads. In this case, the overly conservative estimates could lead to an over-engineered structure. In case (b) the highest wave conditions are under-estimated, so structures designed based on these contours could have a higher risk of failure than anticipated.

To address this issue, the IMEX project proposed to develop improved models for multivariate metocean extremes. The work would build on a recent approach, developed by the project team [4]. The novel aspect of the methodology is the modelling of the joint distribution of storm-peak values using a combination of non-stationary peaks-over-threshold (POT) models, with conditional extreme value models. The non-stationary POT component models extreme values of one variable (the 'response') conditional on other variables, or 'covariates' (e.g. wave height conditional on wind speed), where the covariate is not extreme. When both response and covariate are extreme, the conditional extremes model of Heffernan and Tawn (HT) [5] is used due to its asymptotic justification. The non-stationary POT component uses a penalised piecewise-linear (PPL) model. The model parameters are estimated at a number of covariate values ('nodes') and are assumed to vary linearly between nodes. The approach does not require prior assumptions about the functional form of the relationship between variables and is computationally efficient due to the assumption of piecewise-linear variation.

A drawback of the composite model is the need to characterise different regions of the joint distribution using separate models, and then to resolve any resulting discrepancies in overlapping regions. Moreover, the non-stationary and conditional models must be estimated for each conditioning variable. This process is clearly inefficient, and a more robust

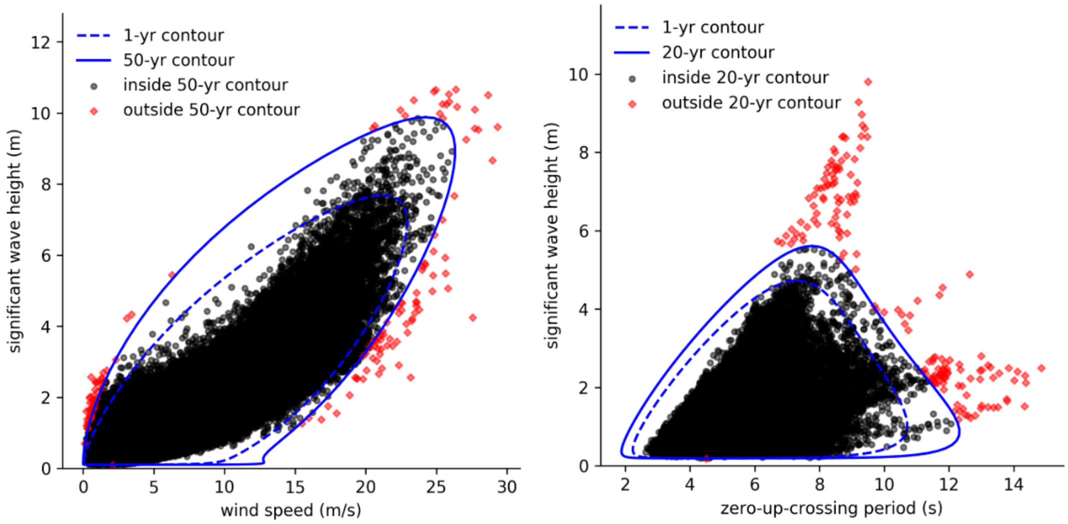

Figure 1: Examples of environmental contours for wave height with wind speed based on 25 years observations (left) and wave height and wave period based on 10 years observations (right) calculated using methods in recommended in current standards [1, 2]. Source: [3].

unified modelling approach would be preferable. Although unified approaches do exist, e.g. copula or multivariate POT models, these suffer from two drawbacks: they do not easily scale to high dimensions, and they require stronger prior assumptions about the form of the dependence structure and extremal dependence.

2 Project aims and objectives

The aim of the IMEX project was to address the research challenges described above, by (i) extending existing multivariate statistical models to create a single coherent and straightforward framework in which to estimate multivariate extremes, and (ii) developing open-source software for estimating multivariate metocean extremes, based on the methodologies developed in (i).

The objectives of the proposed research were:

1. **Extend** the existing composite model approach to higher dimensions;
2. **Develop** a novel single-model approach for multivariate extremes;
3. **Integrate** the models into open-source software for estimation of multivariate extremes;
4. **Demonstrate** the application of models to extreme loading of ORE structures.

3 Summary of work conducted

The research program was divided into four work packages, corresponding to the objectives above. The outputs from each work package are described in the subsections below.

WP1: Extension of composite approach to higher dimensions

The initial implementation of the PPL model, described in [4], was capable of modelling the extreme values of one response conditional on one covariate, restricting its application to modelling the joint distribution of two variables. In

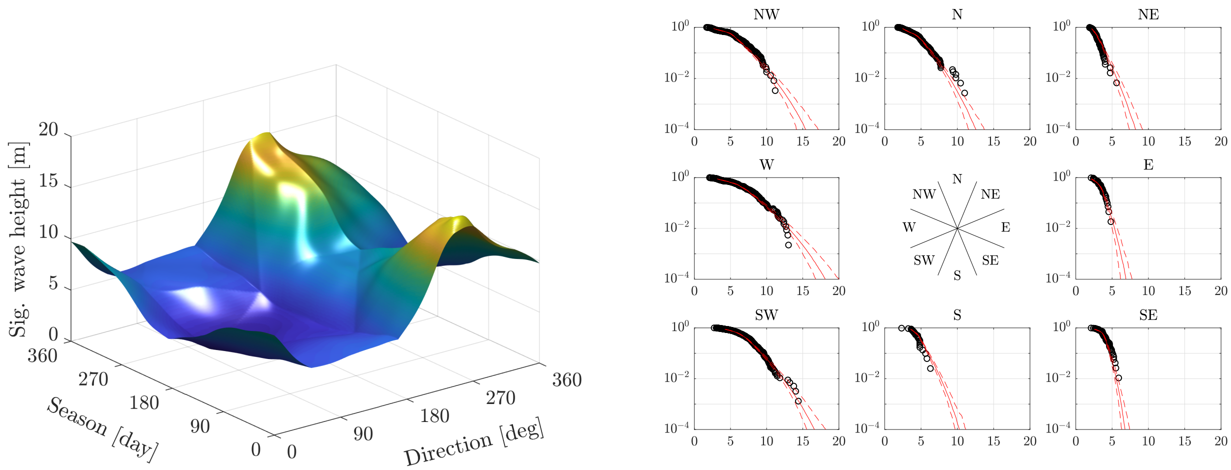

Figure 2: Examples of application of PPL model to estimate quantiles of significant wave height, conditional on season and direction. Source: [6].

WP1 we extended the capabilities of the PPL method to mode an arbitrary number of covariates. The formulation of the model in n-dimensions is presented in [6]. The application of the model for estimating extreme values of significant wave height, conditional on two covariates (season and direction) is illustrated in Figure 2.

During the project, it was found that the composite modelling approach is cumbersome to use in higher dimensions and requires a significant number of user-specified inputs. It was therefore decided that work would focus on developing single-model approach (known as the SPAR model) in WP2, which does not suffer from these drawbacks. The SPAR model provides a more rigorous and flexible framework for modelling multivariate extremes, and is a more powerful tool for metocean engineers and scientists. Focusing on the SPAR model provided a better outcome for the project than developing the composite approach for higher dimensions, whilst still meeting the overall aims of the project.

WP2: Development of single-model approach for multivariate extremes

In WP2 we explored an extension of the PPL approach to the estimation of the full joint distribution. The approach is referred to as the semi-parametric angular-radial (SPAR) model. This method removes the distinction between covariate and response. Instead, multivariate variables are transformed to a radial and angular coordinate system. The angular component is analogous to the covariate in the PPL model and the radial component is analogous to the response. The radial component is modelled by fitting the PPL model, developed in WP1, to the exceedances of an extreme threshold. The SPAR approach re-frames multivariate extremes as an intuitive extension of univariate theory, with angular dependence.

A summary of the approach is as follows. For a pair of random variables (X, Y), we write the joint density function in terms of polar coordinates (R, Θ):

$$f_{R,\Theta}(r,\theta) = r \cdot f_{X,Y}(x,y). \tag{1}$$

The joint density function in polar coordinates is then written in conditional form:

$$f_{R,\Theta}(r,\theta) = f_{\Theta}(\theta) f_{R|\Theta}(r|\theta). \tag{2}$$

For a given angle θ, the conditional density $f_{R|\Theta}(r|\theta)$ is a univariate density function, so standard results from univariate extreme value theory apply. In the SPAR model, we assume that the tail of $f_{R|\Theta}(r|\theta)$ converges to a general Pareto (GP) distribution. The SPAR model for bivariate extremes can therefore be written as

$$f_{R,\Theta}(r,\theta) = f_{\Theta}(\theta)\zeta(\theta) f_{GP}\left(r|u(\theta),\sigma(\theta),\xi(\theta)\right), \tag{3}$$

where $\zeta(\theta) = \Pr\left(r > u(\theta)\right)$ (assumed constant), and f_{GP} is the GP density function with threshold, scale and shape parameters $(u(\theta), \sigma(\theta), \xi(\theta))$, conditional on angle, θ. The angular density is simple to estimate using non-parametric

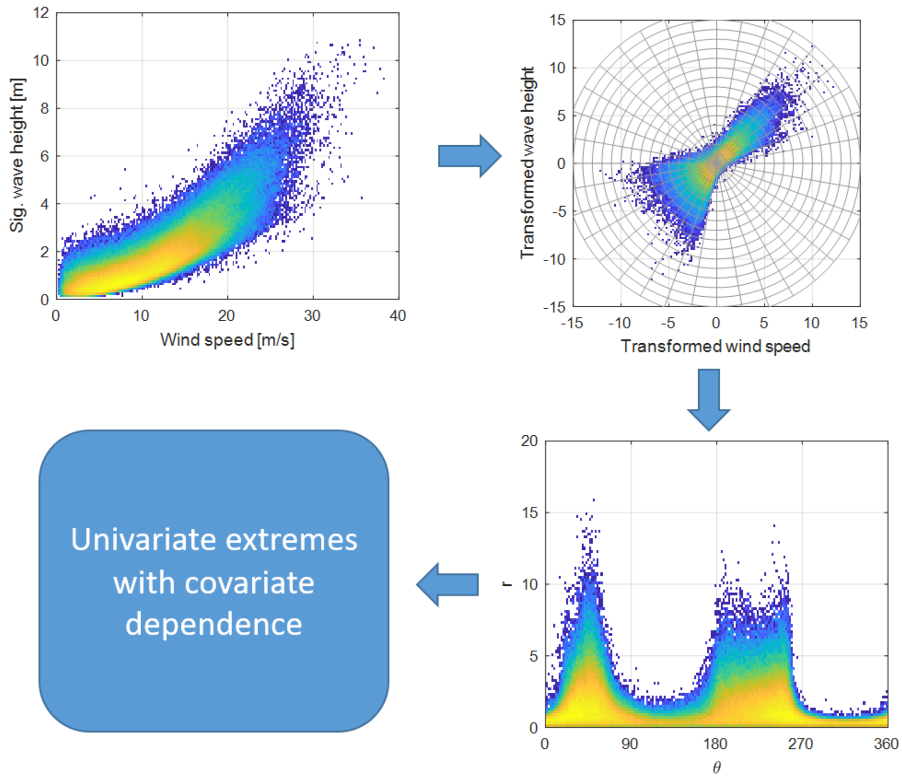

Figure 3: Illustration of steps in the application of the SPAR method.

methods, such as kernel density estimation. In the present work, we estimate the conditional GP model using the PPL method, developed in WP1. The application of the PPL model to estimate $f_\Theta(\theta)$ and $f_{GP}\left(r|u(\theta),\sigma(\theta),\xi(\theta)\right)$ is described in [6].

The method extends naturally to higher dimensions. For example, the joint distribution of three variables could be written in terms of spherical coordinates, or for more variables, hyper-spherical coordinates can be used, with the radial component modelled in the same way.

An illustration of the SPAR approach is shown in Figure 3, applied to the joint distribution of wind speed and significant wave height. In this example, the variables are first transformed to a common marginal scale (in this case, a standard Laplace scale). The Laplace variables are then converted to polar coordinates, where the extremes of the radial variable are modelled with the PPL model.

Further details of the model will be provided in [7] (currently in preparation). In this work we show that the SPAR model can be used to represent the extremes of many commonly-used parametric joint distributions. We also show that the SPAR model can represent both asymptotically dependent (AD) and asymptotically independent (AI) distributions. The property of asymptotic dependence or independence, describes whether extremes of two variables are likely to be observed concurrently or not. Somewhat surprisingly, nearly all existing methods for multivariate extremes can only model distributions which are either AI or AD [8], meaning that an analyst must choose to fit either an AI or AD distribution to the data. This can lead to very large differences in the joint tail of the fitted distribution. The SPAR model does not suffer from this drawback, making it attractive for modelling. The SPAR model is expected to be widely applicable to other problems in multivariate extremes.

WP3: Development of open-source software

In WP3 the methods developed in WP1 and WP2 were integrated into an open-source software package. The software has been made available on GitHub [9]. The software comprises MATLAB scripts and functions, together with worked examples of the application described in [6]. User manuals for the code are currently being finalised and will be made available on GitHub soon.

The software is intended to be used by engineers and researchers working in offshore renewable energy. Much of the complexity of the models is contained in functions that do not need to be edited by the user. However, the open-source code allows advanced users and researchers to create bespoke tools and develop the methods further. The software is sufficiently simple to be run by engineers who do not have a detailed statistical training.

WP4: Application of models to extreme loading of ORE structures

The methods developed as part of the project have been applied to loading of generic offshore structures, as described in publications [10, 11], and extreme loading of an offshore wind turbine has been considered in publication [12]. These publications also compare the new methods, developed in this project, with existing methods, and show that the new methods lead to reduced uncertainties in estimates of extreme loading on ORE structures.

4 Summary of outputs and dissemination

Five journal publications have been completed as part of the project, with four published already [10–13] and the fifth submitted for review and uploaded to the arXiv pre-print server [6]. A sixth journal paper [7], describing the SPAR model is currently in preparation and will be submitted to a leading statistical journal.

Publications [10] and [11] are collaborations with researchers at two of the leading certification bodies for offshore renewable energy (DNV and Bureau Veritas). These publications are expected to lead to updates of offshore design standards. Publication [10] demonstrates how the methodology developed in the IMEX project compares to existing methodologies. It shows that the model developed in IMEX gives the most accurate estimates of the joint distributions of extreme metocean parameters, compared to various methods considered as part of a recent benchmarking exercise. Publication [11] presents further theoretical and numerical results, showing why current methodologies for estimating multivariate metocean extremes can lead to bias, and how the methods developed as part of IMEX can overcome this. Publication [13] presents work on the fundamentals of how empirical estimates of extreme events should be defined. We derive various results about fundamental limits on uncertainties in extreme events, as a function of the number of observations. The results are of interest when considering extreme events in either numerical simulations, model tests or field data.

Publication [12] considers extreme loading of an offshore wind turbine. It illustrates how the methodologies developed in the IMEX project compare to methods recommended in current design standards. We quantify uncertainties in the current design standards and provide a breakdown of the uncertainties from various assumptions made in the current methods.

A conference paper was presented at the OMAE 2021 conference [14]. This paper presents an extension to an international benchmarking exercise for environmental contour methods, which motivated the work undertaken in the IMEX project. The paper is a collaboration with researchers at Sandia National Laboratories, University of Texas at Austin, and University of Bremen, which extends the reach of the initial IMEX project partners.

Finally, summary presentations of the IMEX project were presented at the Supergen ORE Annual Assemblies in January 2021 [15] and January 2022 [16].

References

[1] IEC, *Standard 61400-3-1: Wind energy systems. Part 3-1*, Apr. 2019.

[2] DNVGL, *Recommended Practice RP-C205: Environmental conditions and environmental loads*, Sep. 2019.

[3] A. F. Haselsteiner, R. G. Coe, L. Manuel, P. T. T. Nguyen, N. Martin, and A. Eckert-Gallup, "A benchmarking exercise on estimating extreme environmental conditions: Methodology & baseline results," in *38th International Conference on Ocean, Offshore and Arctic Engineering*, Glasgow, UK, 2019, OMAE2019–96 523. DOI: 10.1115/OMAE2019-96523.

[4] E. B. L. Mackay and P. Jonathan, "Estimation of environmental contours using a block resampling method," in *39th International Conference on Ocean, Offshore and Arctic Engineering*, 2020, OMAE2020–18 308. DOI: 10.1115/OMAE2020-18308.

[5] J. E. Heffernan and J. A. Tawn, "A conditional approach for multivariate extreme values," *J. R. Statist. Soc. B*, vol. 66, pp. 497–546, 2004.

[6] A. M. Barlow, E. Mackay, E. Eastoe, and P. Jonathan, *A penalised piecewise-linear model for non-stationary extreme value analysis of peaks over threshold*, 2022. arXiv: 2201.03915 [stat.ME].

[7] E. Mackay, "Multivariate peaks-over-threshold with flexible dependence class," *In preparation*, 2022.

[8] J. L. Wadsworth, J. A. Tawn, A. C. Davison, and D. M. Elton, "Modelling across extremal dependence classes," *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, vol. 79, no. 1, pp. 149–175, 2017. DOI: 10.1111/rssb.12157.

[9] E. Mackay, A. M. Barlow, E. Eastoe, and P. Jonathan, *PPL model*, 2022. [Online]. Available: https://github.com/edmackay/PPL-model/.

[10] G. de Hauteclocque, E. Mackay, and E. Vanem, "Quantitative assessment of environmental contour approaches," *Ocean Engineering*, vol. 245, p. 110 374, 2022. DOI: 10.1016/j.oceaneng.2021.110374.

[11] E. Mackay, G. de Hauteclocque, E. Vanem, and P. Jonathan, "The effect of serial correlation in environmental conditions on estimates of extreme events," vol. 242, p. 110 092, 2021. DOI: 10.1016/j.oceaneng.2021.110092.

[12] A. F. Haselsteiner, M. Frieling, E. Mackay, A. Sander, and K. D. Thoben, "Long-term extreme response of an offshore turbine: How accurate are contour-based estimates?" *Renewable Energy*, vol. 181, pp. 945–965, 2022. DOI: 10.1016/j.renene.2021.09.077.

[13] E. Mackay and P. Jonathan, "Sampling properties and empirical estimates of extreme events," *Ocean Engineering*, vol. 239, p. 109 791, 2021. DOI: 10.1016/j.oceaneng.2021.109791.

[14] E. Mackay, A. F. Haselsteiner, R. G. Coe, and L. Manuel, "A Second Benchmarking Exercise on Estimating Extreme Environmental Conditions: Methodology & Baseline Results," *Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE2021*, 2021. DOI: 10.1115/OMAE2021-64874.

[15] E. Mackay, "Improved Models for Multivariate Metocean Extremes (IMEX)," in *Supergen ORE Annual Assembly*, Jan. 2021. [Online]. Available: https://supergen-ore.net/projects/improved-models-for-multivariate-metocean-extremes-imex.

[16] ——, "Improved Models for Multivariate Metocean Extremes (IMEX)," in *Supergen ORE Annual Assembly*, Jan. 2022. [Online]. Available: https://supergen-ore.net/projects/improved-models-for-multivariate-metocean-extremes-imex.