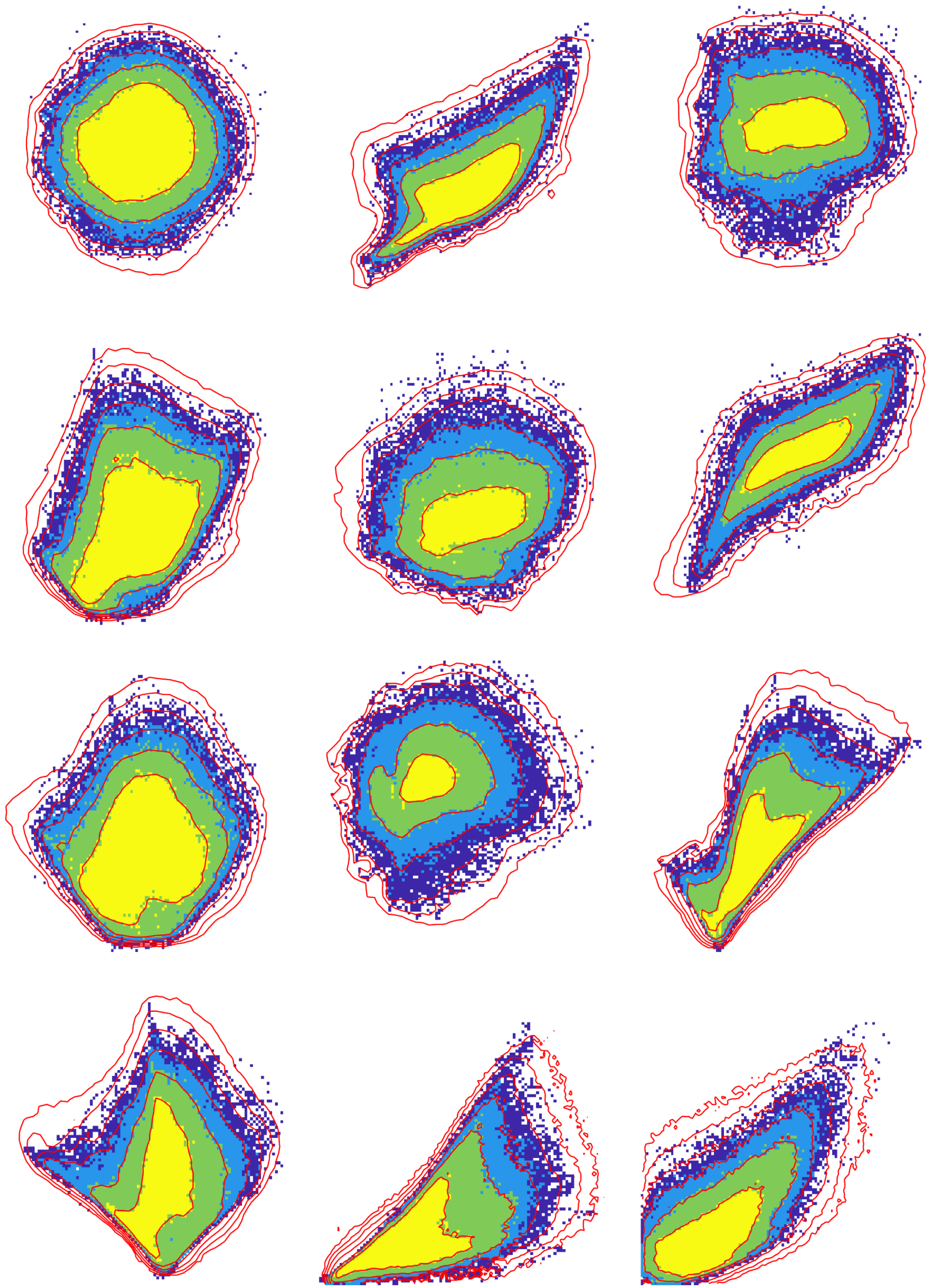
University of Exeter

Updated version

Deep Learning Joint Extremes of Metocean Variables using the SPAR Model



E. Mackay, C.J.R. Murphy-Barltrop, J. Richards, P. Jonathan

OMAE2025-155197, Vancouver

25 June 2025

Overview

University of Exeter

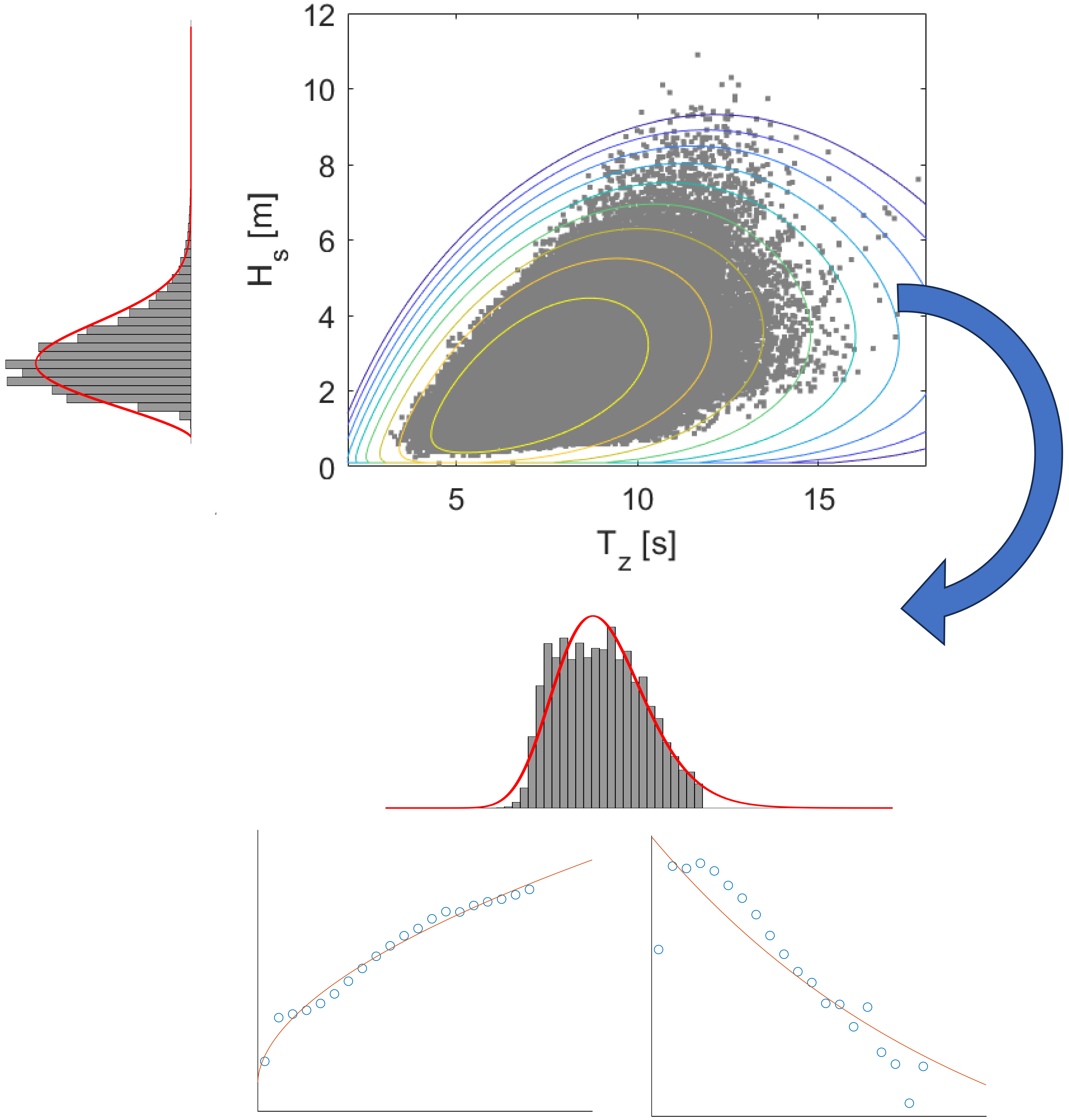
- Brief review of existing approaches and limitations
- The SPAR model
  - Definition
  - Inference
- Five-dimensional example
- Conclusions

Updated version:
https://arxiv.org/abs/2412.15808

Global hierarchical models

- Suppose random vector $\mathbf{X} = (X_1, \dots, X_d) \in \mathbb{R}^d$ has joint density $f_{\mathbf{X}}$ and marginal densities $f_{X_1}, \dots, f_{X_d}$ and distribution functions $F_{X_1}, \dots, F_{X_d}$

- Approach:
  1. Write joint density as:
     $f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1) f_{X_2|X_1}(x_2|x_1) \cdots f_{X_d|X_1,\dots,X_{d-1}}(x_d|x_1, \dots, x_{d-1})$
  2. Choose parametric models for $f_{X_1}, f_{X_2|X_1}, \dots,$ $f_{X_d|X_1,\dots,X_{d-1}}$
  3. Estimate relations between the parameters of the conditional densities and the conditioning variables



Limitations:
- No a priori reason for choosing one model over another
- Fitting a global model does not guarantee of good fit in the tail
- Model for the conditional dependence structure is ad hoc and does not provide a rationale for extrapolation
- Often a poor fit to observations

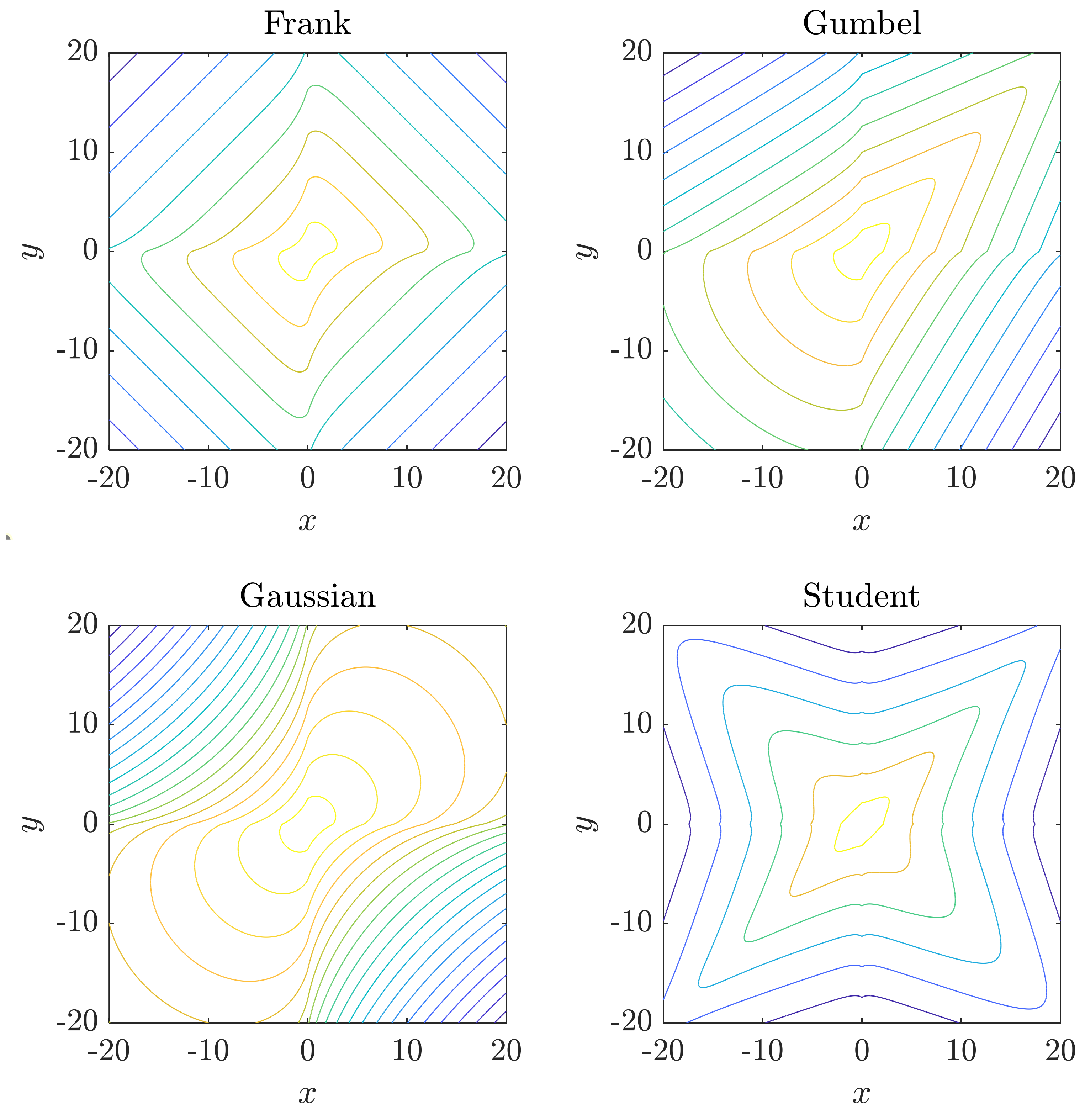
Copula modelling

Approach

1. Use Sklar's theorem to write:
   $$f_{\mathbf{X}}(\mathbf{x}) = f_{X_1}(x_1) \cdots f_{X_d}(x_d) c\left(F_{X_1}(x_1), \dots, F_{X_d}(x_d)\right)$$
   where c is the copula density of $\mathbf{X}$
2. Choose parametric model for margins and copula

Limitations:

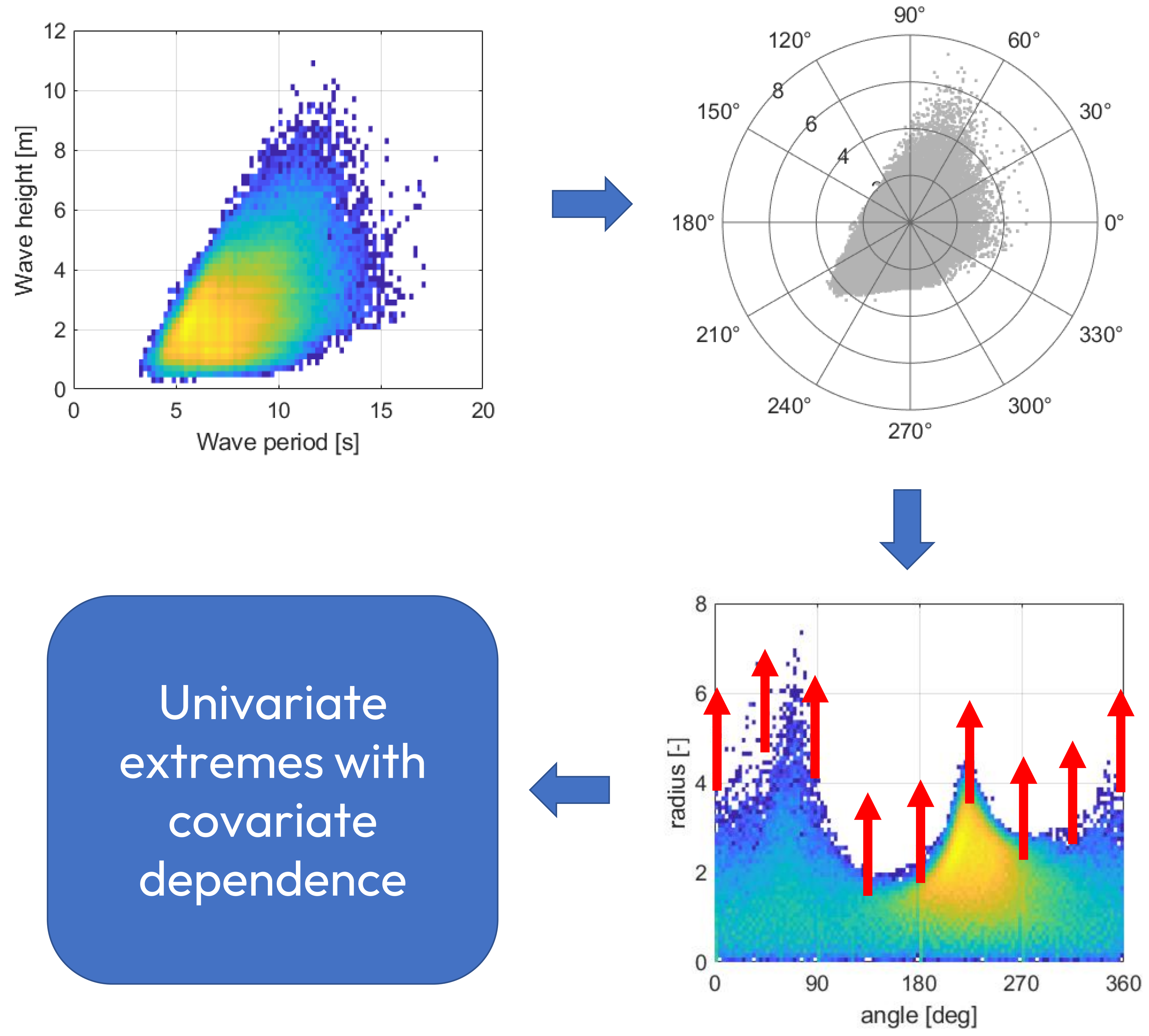
- Global model for margins and copula does not guarantee of good fit in the tail
- No rationale for choose marginal or copula models
- Different choices of copula can lead to large differences in joint tail behaviour



Copulas on Laplace margins all with correlation coefficient $\rho = 0.6$

The SPAR model

- Semi-Parametric Angular-Radial (SPAR) model
- Introduced at OMAE 2024
- Reframes multivariate extremes as an intuitive extension of univariate theory, with angular dependence
- Inference is standard univariate peaks-over-threshold (POT) problem with covariate dependence
- Many existing methods for multivariate extremes are special cases of SPAR
- SPAR is more flexible than current methods



Transformation to angular-radial coordinates

University of Exeter

- Let $\mathbf{X} = (X_1, \dots, X_d) \in \mathbb{R}^d$ have density $f_{\mathbf{X}}$

- Define radial and angular variables:

$$R = \|\mathbf{X}\|_2$$
$$\mathbf{W} = \mathbf{X}/R$$

where $\|\mathbf{X}\|_2 = \sqrt{X_1^2 + \cdots + X_d^2}$ is the Euclidean (L2) norm

- Note that
    - R is radial variable
    - $\mathbf{W}$ is a pseudo-angle corresponding to a point on the unit hypersphere

- Then $(R, \mathbf{W})$ has density

$$f_{R,\mathbf{W}}(r, \mathbf{w}) = r^{d-1} f_{\mathbf{X}}(r\mathbf{w})$$

SPAR model statement

University of Exeter

- Write joint density in conditional form:

$$f_{R,\mathbf{W}}(r,\mathbf{w}) = f_{\mathbf{W}}(\mathbf{w}) f_{R|\mathbf{W}}(r|\mathbf{w})$$

- Assume standard univariate POT model for tail of $f_{R|\mathbf{W}}(r|\mathbf{w})$
- Define threshold function $u(\mathbf{w})$ as conditional quantile of $R|(\mathbf{W}=\mathbf{w})$ at exceedance probability ζ
- SPAR model for the joint density is

$$f_{R,\mathbf{W}}(r,\mathbf{w}) = \zeta\, f_{\mathbf{W}}(\mathbf{w})\, f_{\mathrm{GP}}\big(r-u(\mathbf{w});\xi(\mathbf{w}),\sigma(\mathbf{w})\big), \qquad r > u(\mathbf{w})$$

$f_{\mathrm{GP}}(r;\xi,\sigma)$ is a generalised Pareto density function with shape parameter ξ and scale parameter σ

Inference

- Model:

$$f_{R,\mathbf{W}}(r,\mathbf{w}) = \zeta\, f_{\mathbf{W}}(\mathbf{w})\, f_{\mathrm{GP}}\big(r - u(\mathbf{w}); \xi(\mathbf{w}), \sigma(\mathbf{w})\big)$$

- Need to estimate:
  - Angular density: $f_{\mathbf{W}}(\mathbf{w})$
  - Threshold function: $u(\mathbf{w})$
  - GP shape and scale parameter functions: $\xi(\mathbf{w}), \sigma(\mathbf{w})$

- Problems are separable:
  - Angular density can be estimated independently of GP threshold and parameter functions

- Many existing methods for these problems

Angular modelling

University of Exeter

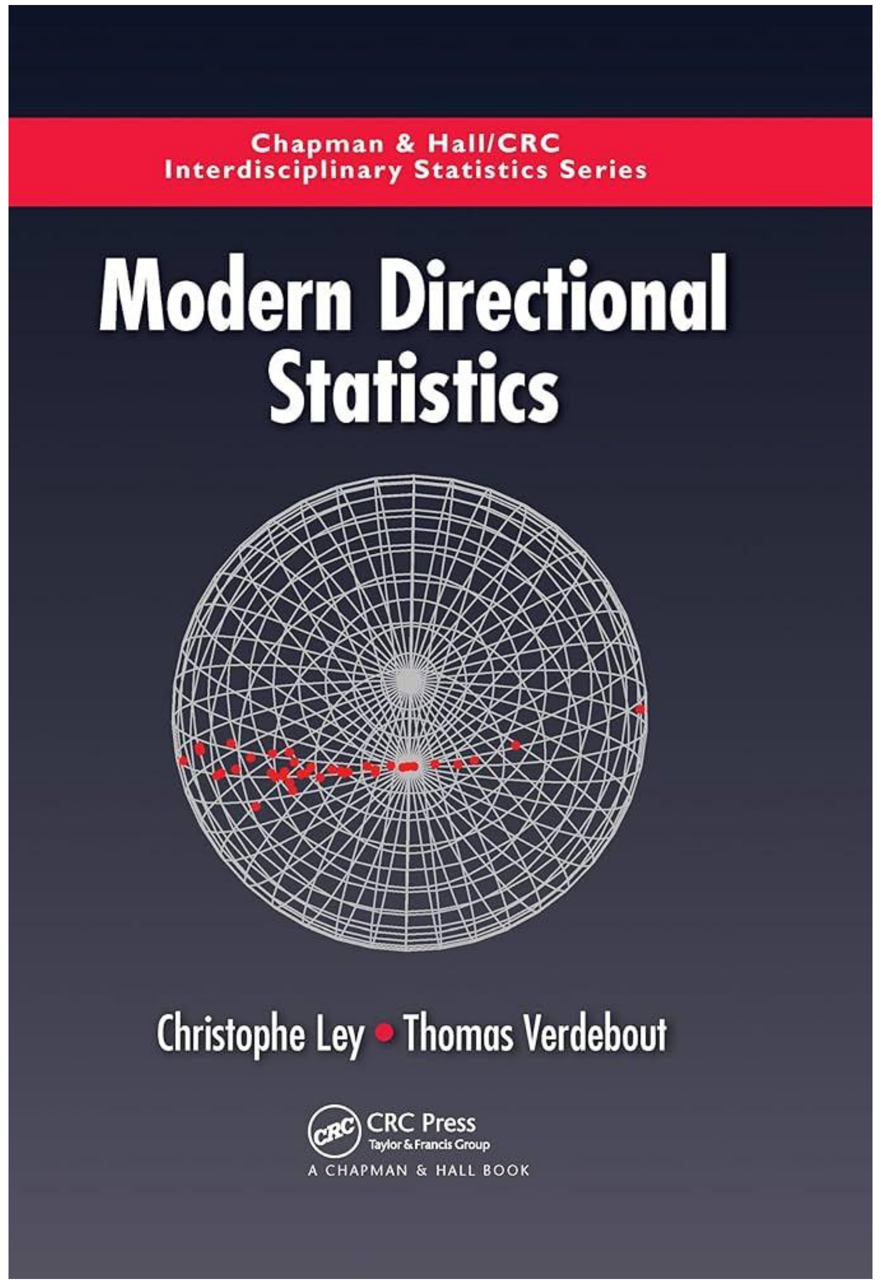
- Problem of estimating a density on the hypersphere is one of **directional statistics**

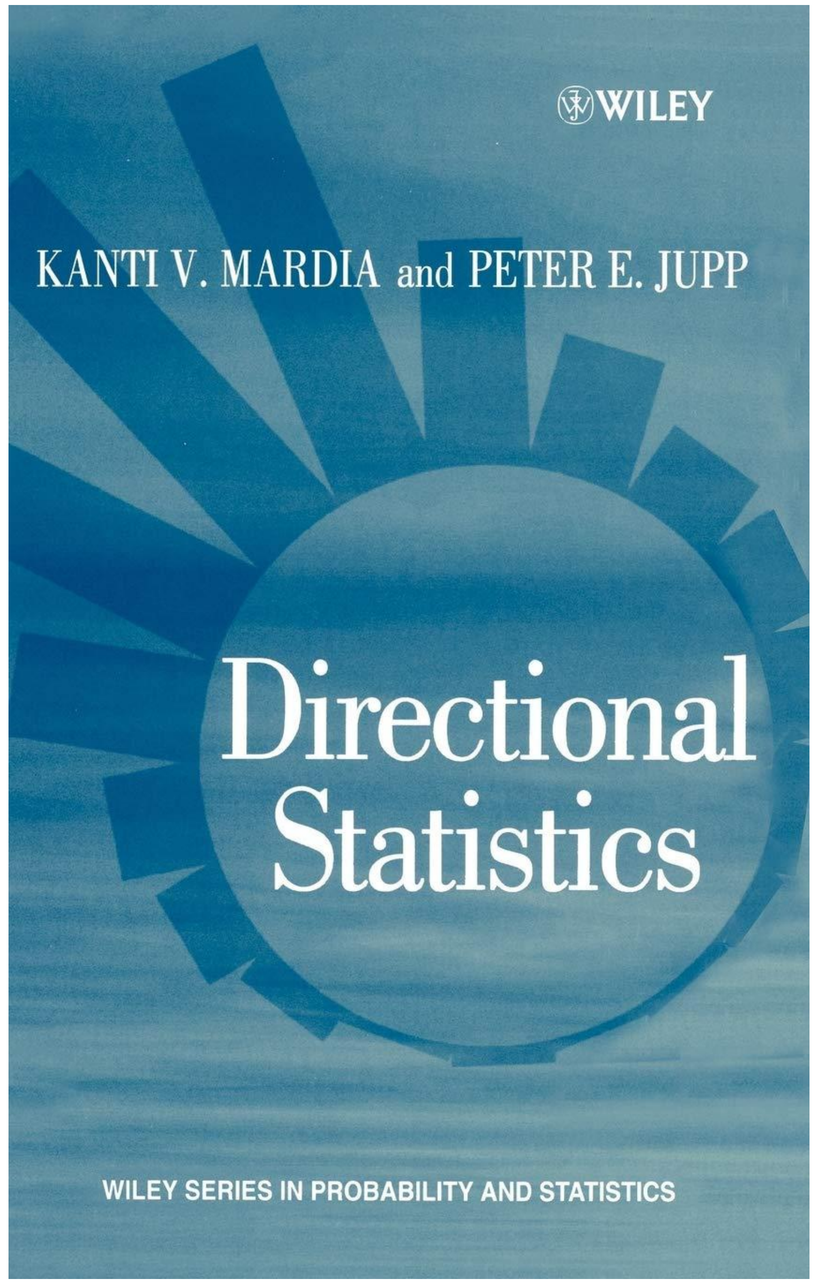
- Many ‘off-the-shelf’ methods available:
  - Kernel density estimation (KDE)
  - Mixture modelling
  - Spline-based models

- KDE used in this example:
  - Simple to implement
  - Very flexible
  - Fast to simulate from model

- Bandwidth optimised using cross-validation scheme

- Only difference from ‘standard’ KDE in $\mathbb{R}^d$ is that kernels are density functions on the sphere





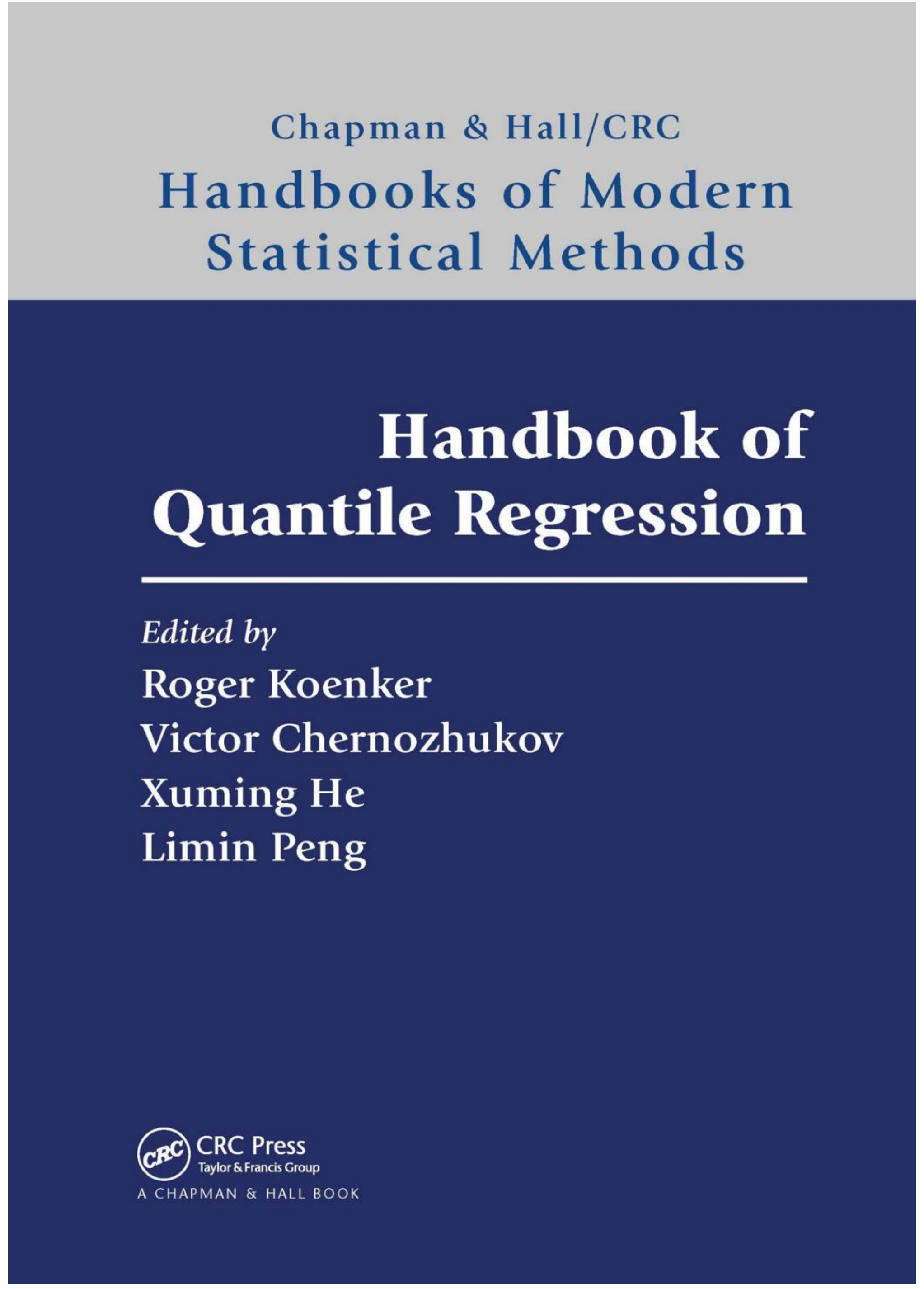
Conditional radial modelling

- Two problems:
  - Threshold function $u(\mathbf{w})$
  - GP shape and scale parameter functions: $\xi(\mathbf{w}), \sigma(\mathbf{w})$

- Threshold function :
  - Defined as a conditional quantile at fixed exceedance probability
  - Process of estimating conditional quantiles is known as **quantile regression**

- GP parameter functions:
  - Inference is non-stationary peaks-over-threshold (POT) with angle as covariate
  - Many existing methods for this

- Key consideration is choosing suitable parametric form for threshold and parameter functions



Covariate representations

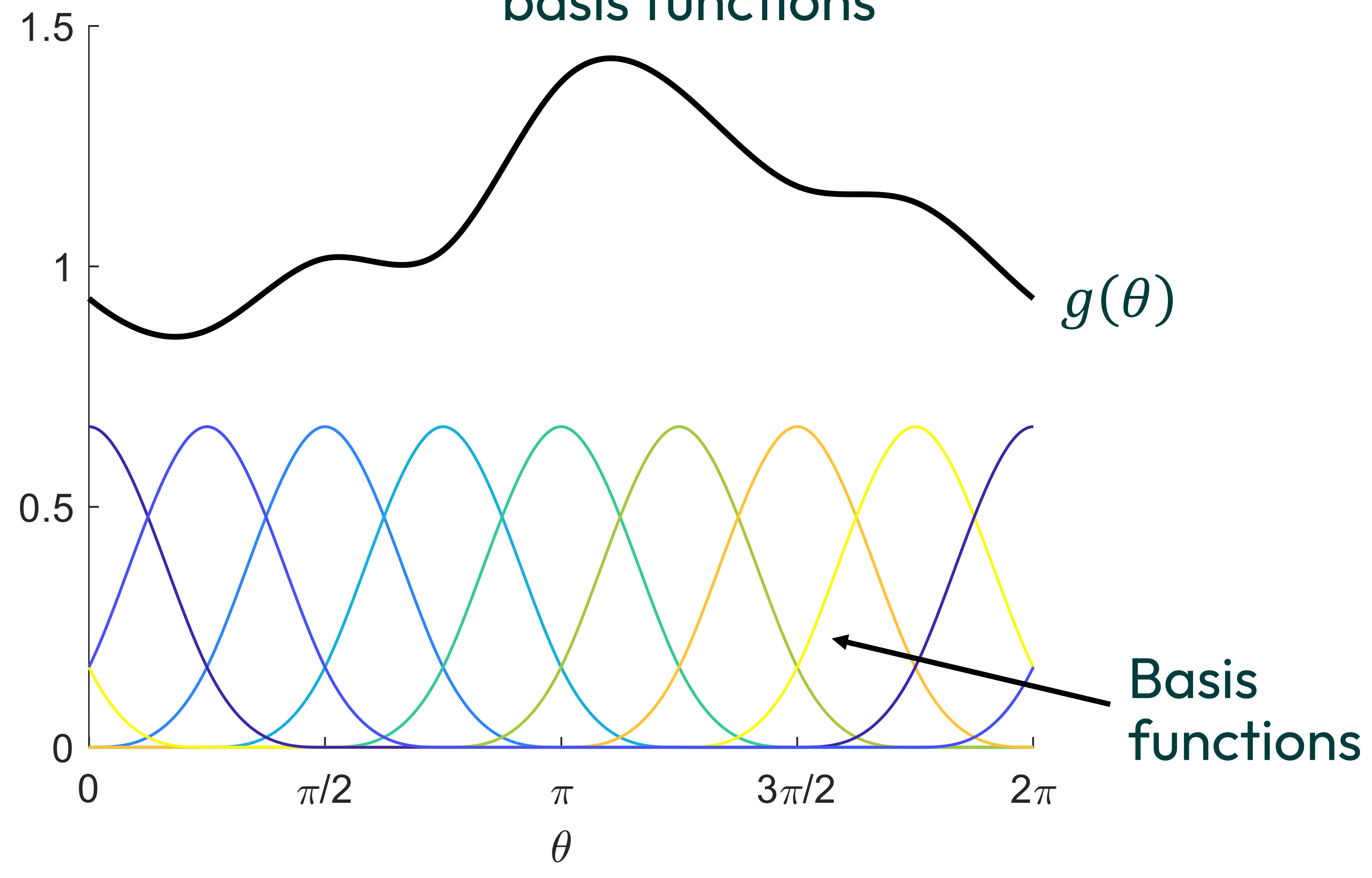
- Various ways of representing covariate functions: e.g.
  - Generalised additive models (GAMs), e.g. splines
  - Gaussian processes
  - Artificial neural networks (ANNs)

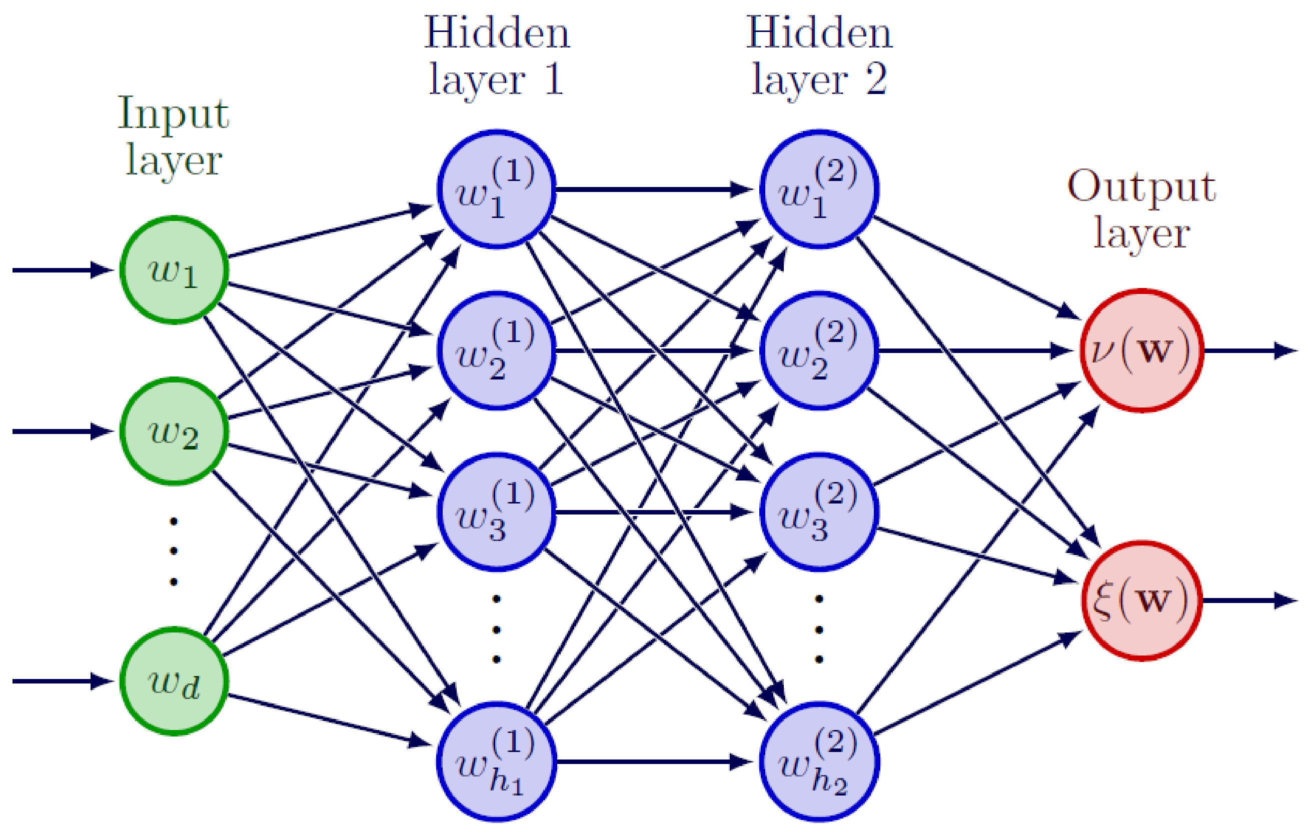
- All provide flexible function approximation

- Neural networks have computational advantages over GAMs in higher dimensions
  - Computationally-efficient training algorithms for ANN parameter optimisation and regularisation

- Parameter estimation method for ANN and GAMs both use maximum likelihood

GAM representation of a function as a sum of basis functions





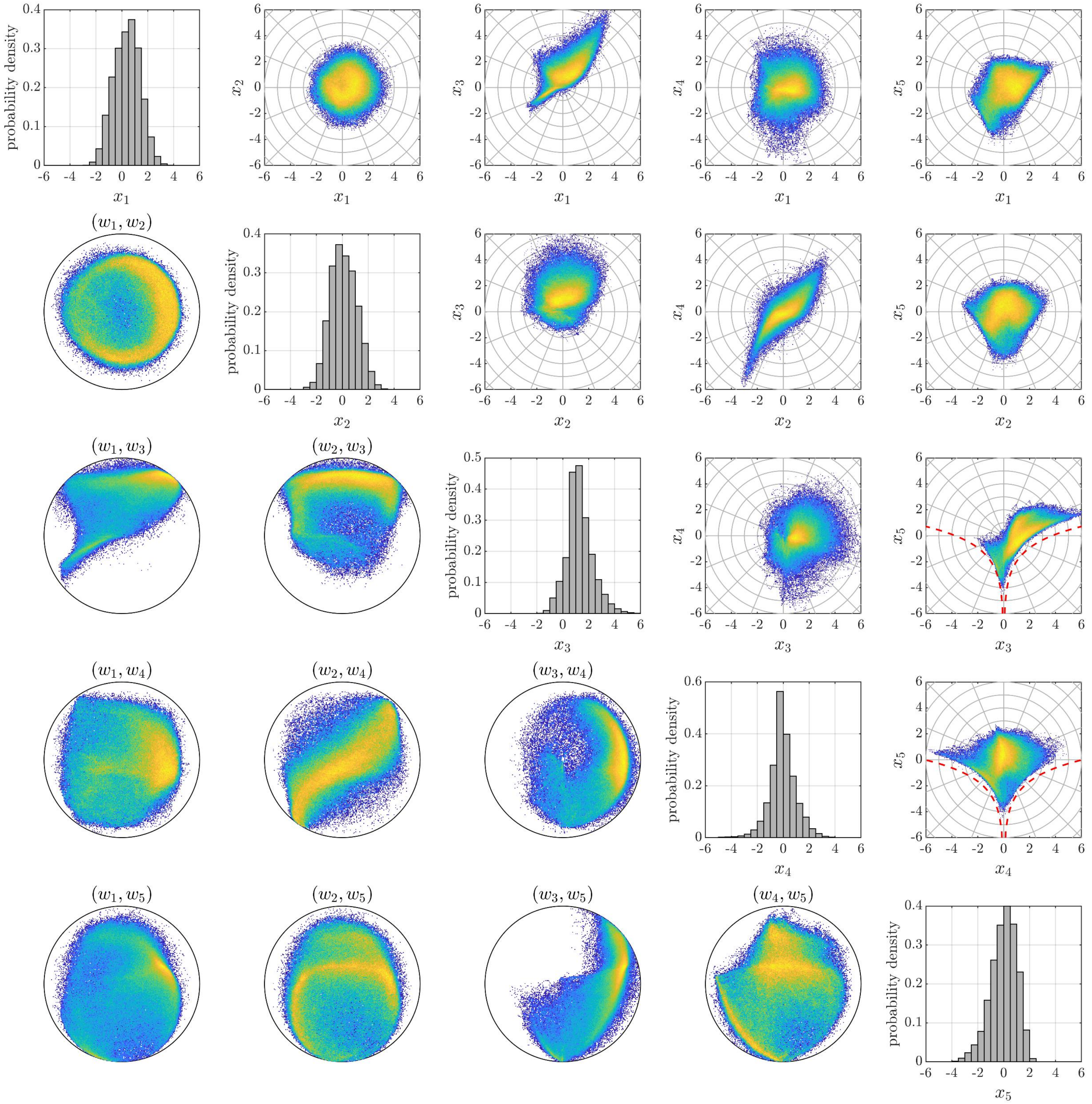
Schematic of ANN with two hidden layers

Five-dimensional application

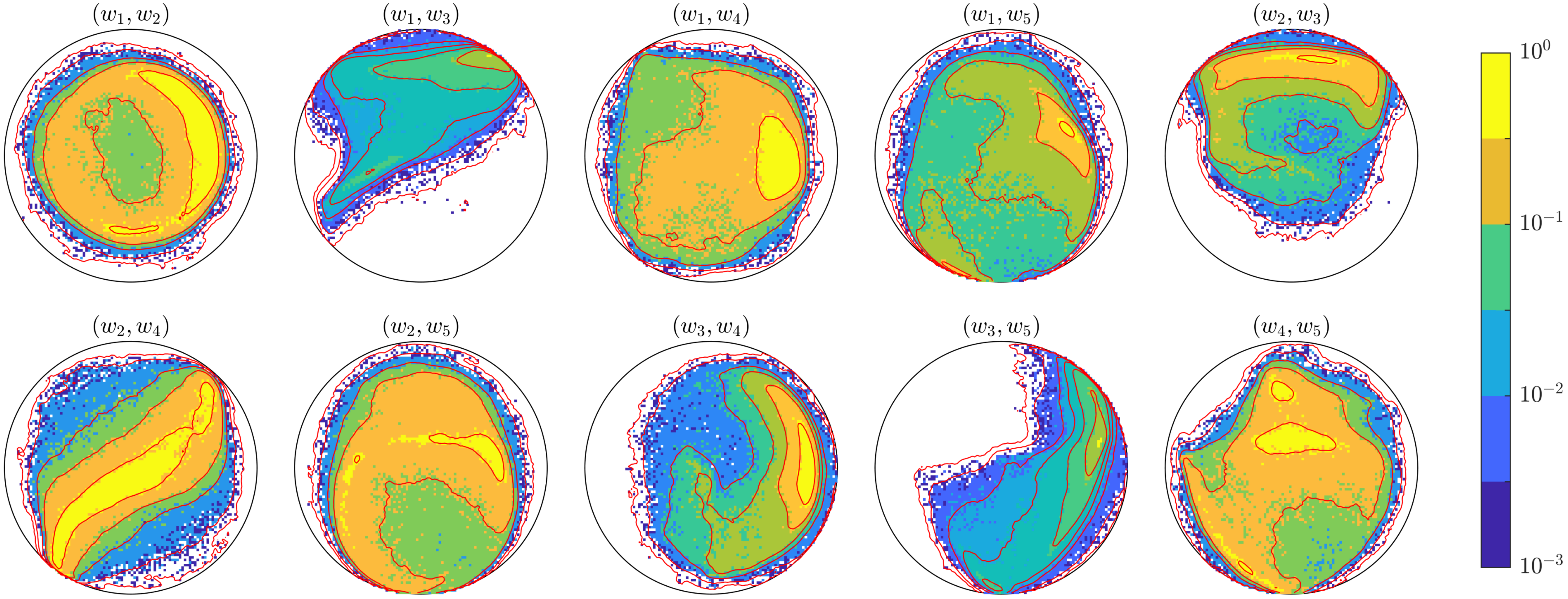
- 31-year hindcast of hourly values for a site in Celtic Sea

- Variables
  - Wind speed & direction
  - Wave height, period & direction

- Transformed to:
  - East-North components of winds and waves
  - Logarithm of wave period
  - All normalised to have unit variance



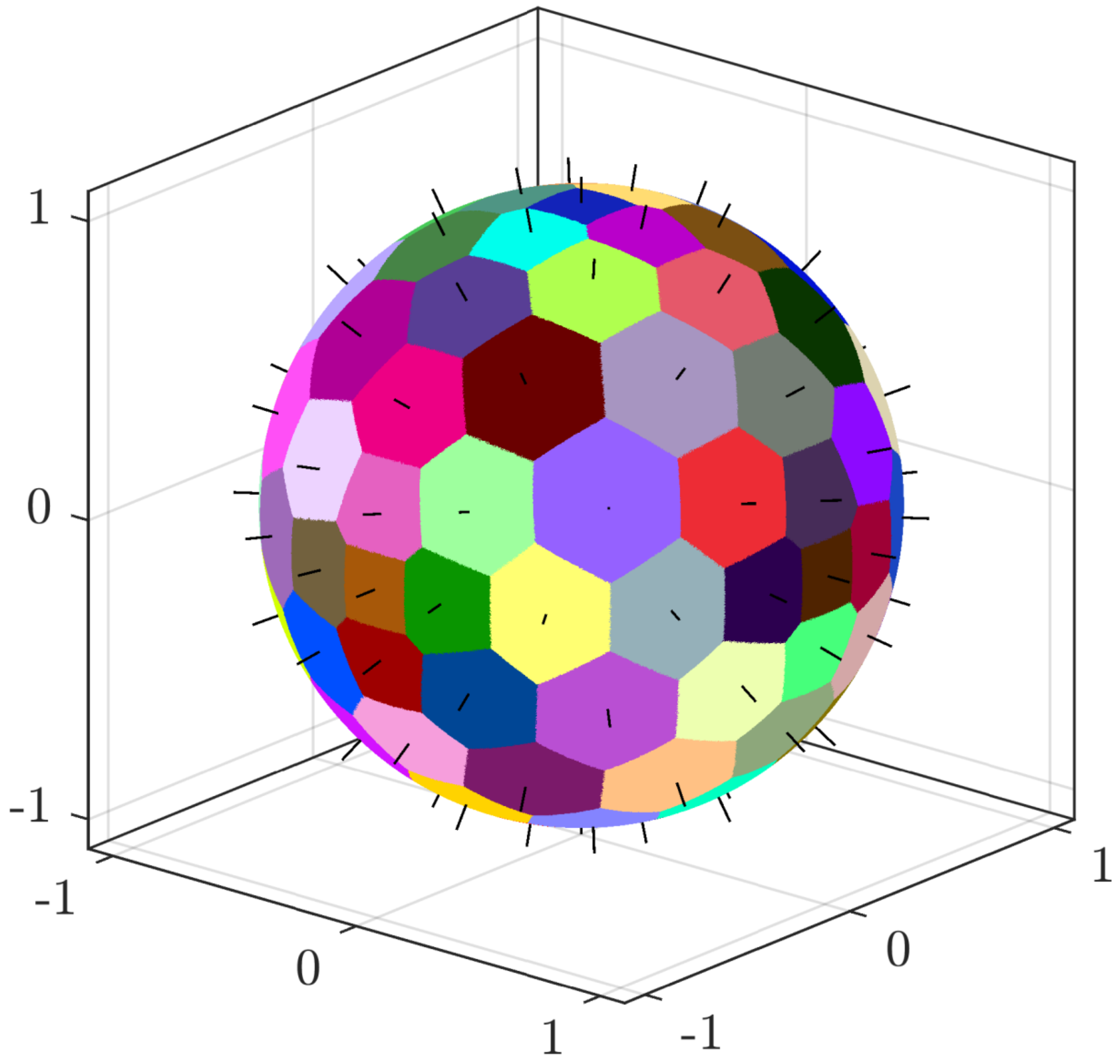
Assessment of angular model



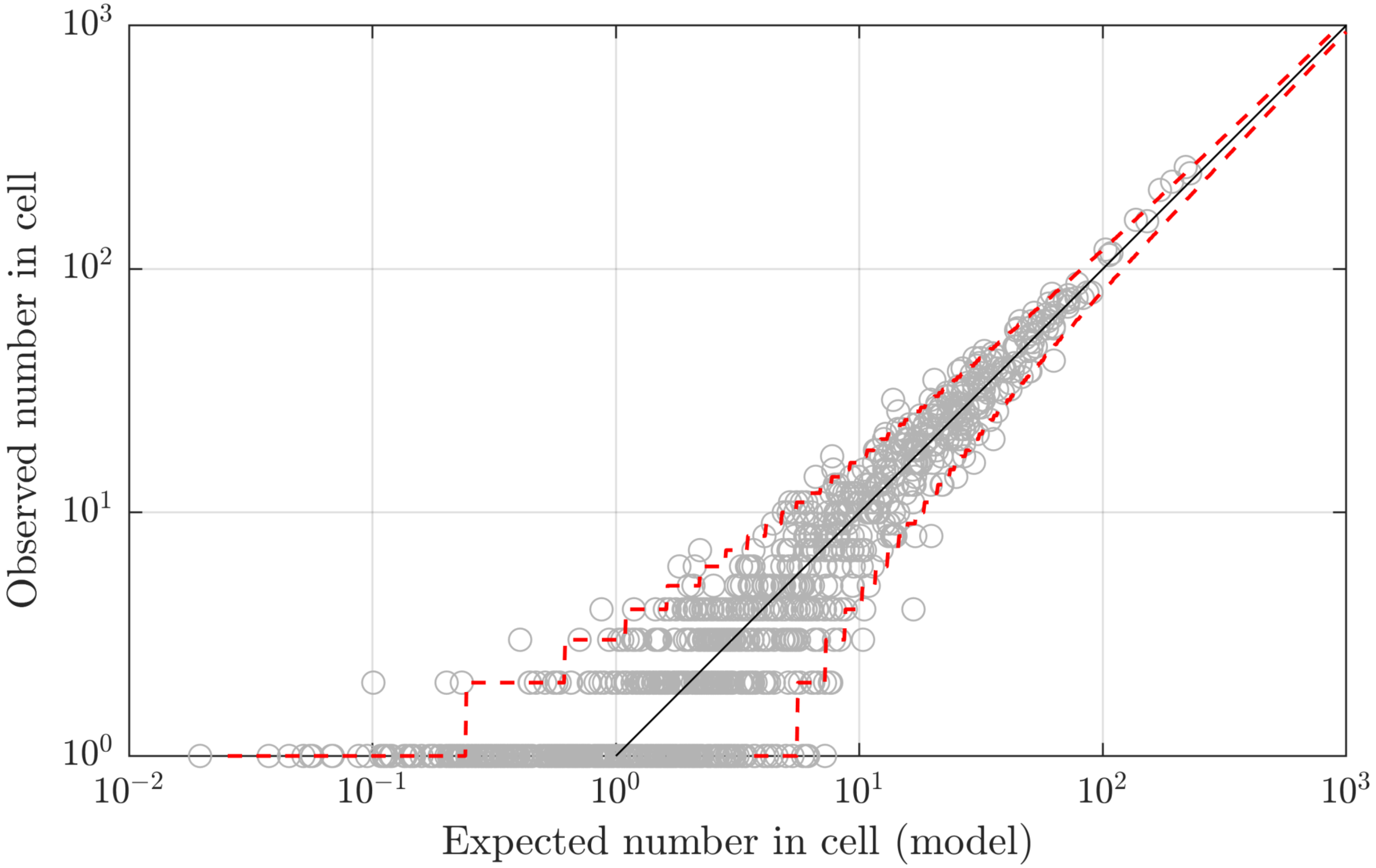
Coloured plots: Empirical joint densities of pairs of angular components

Red lines: Contours of the joint densities from a sample from the KD model

Local assessment

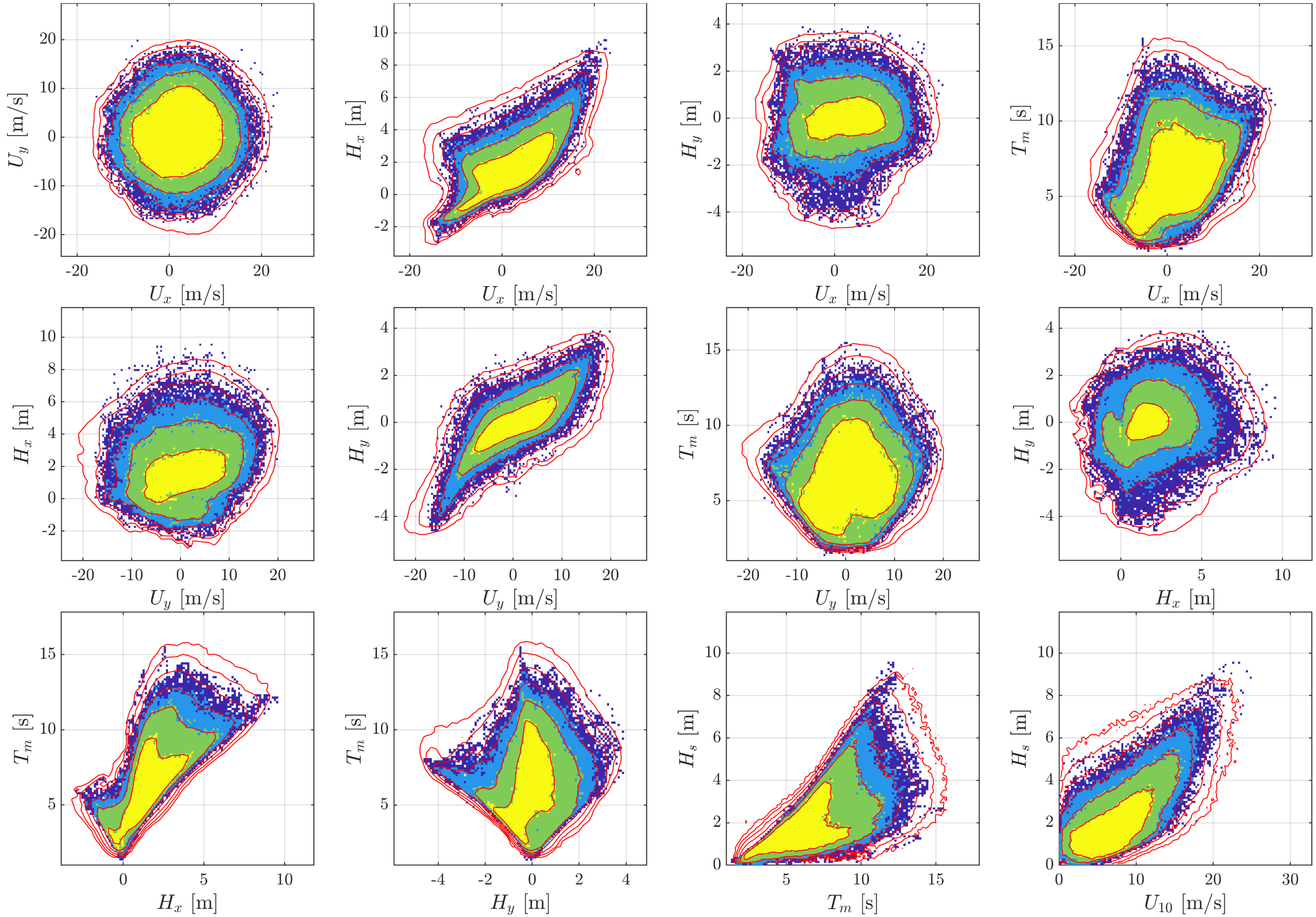


Example of Voronoi partition of sphere into non-overlapping cells



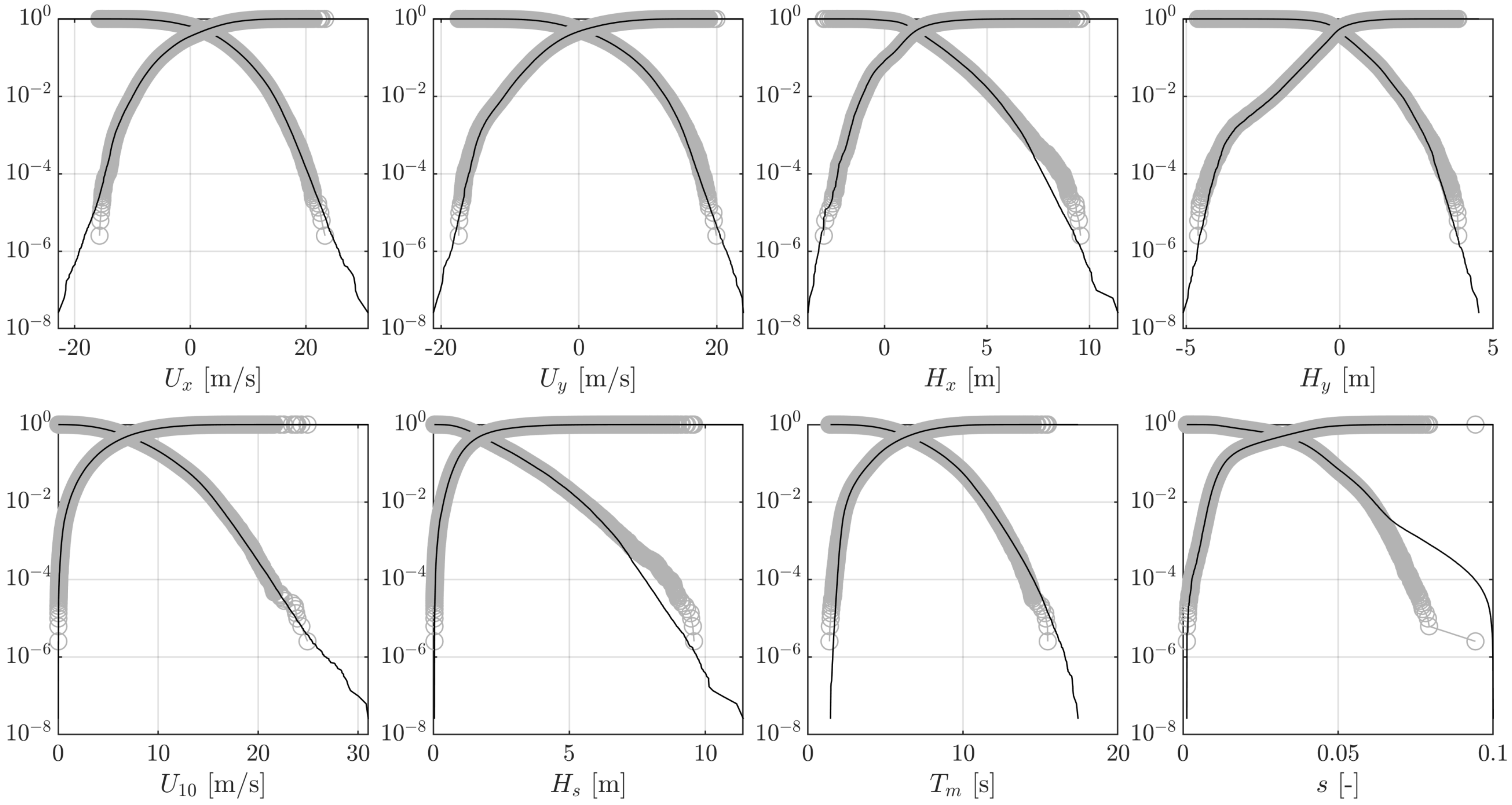
Comparison of observed to expected number of observations in each cell.

Red lines are 95% CI for sampling effects

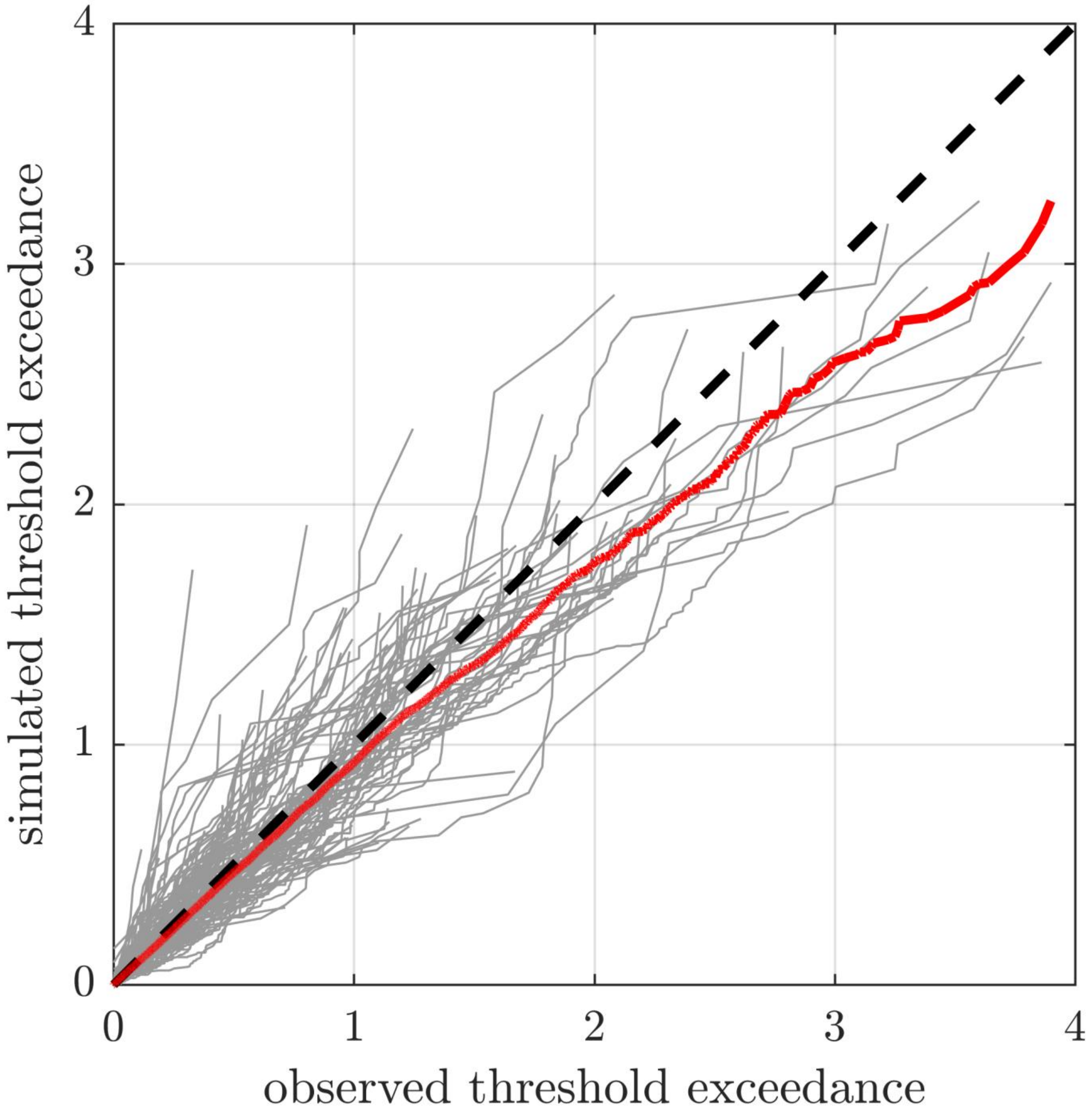


U_y [m/s]
U_x [m/s]
H_x [m]
U_x [m/s]
H_y [m]
U_x [m/s]
T_m [s]
U_x [m/s]
H_x [m]
U_y [m/s]
H_y [m]
U_y [m/s]
T_m [s]
U_y [m/s]
H_y [m]
H_x [m]
T_m [s]
H_x [m]
T_m [s]
H_y [m]
H_s [m]
T_m [s]
H_s [m]
U_{10} [m/s]

Radial model diagnostics



Marginal lower & upper tail exceedance plots



Local tail QQ plots

Summary

- Existing ‘engineering’ methods for modelling multivariate extremes have strong limitations
  - Rely on ad-hoc choices
  - Poor fit in practice
- SPAR model
  - Provides mathematically-justified model for multivariate extremes
  - Requires no strong assumptions about form of margins or dependence structure
  - Inference built on existing practices
  - Shown to provide good fit to metocean datasets
- Future work
  - Detailed comparison with existing inference methods
  - Effect of choice of margins
  - Effect of choice of origin
  - Optimisation of neural network inference

Updated version:
https://arxiv.org/abs/2412.15808

Questions?

e.mackay@exeter.ac.uk