Statistical deep learning for modelling environmental extremes

Jordan Richards

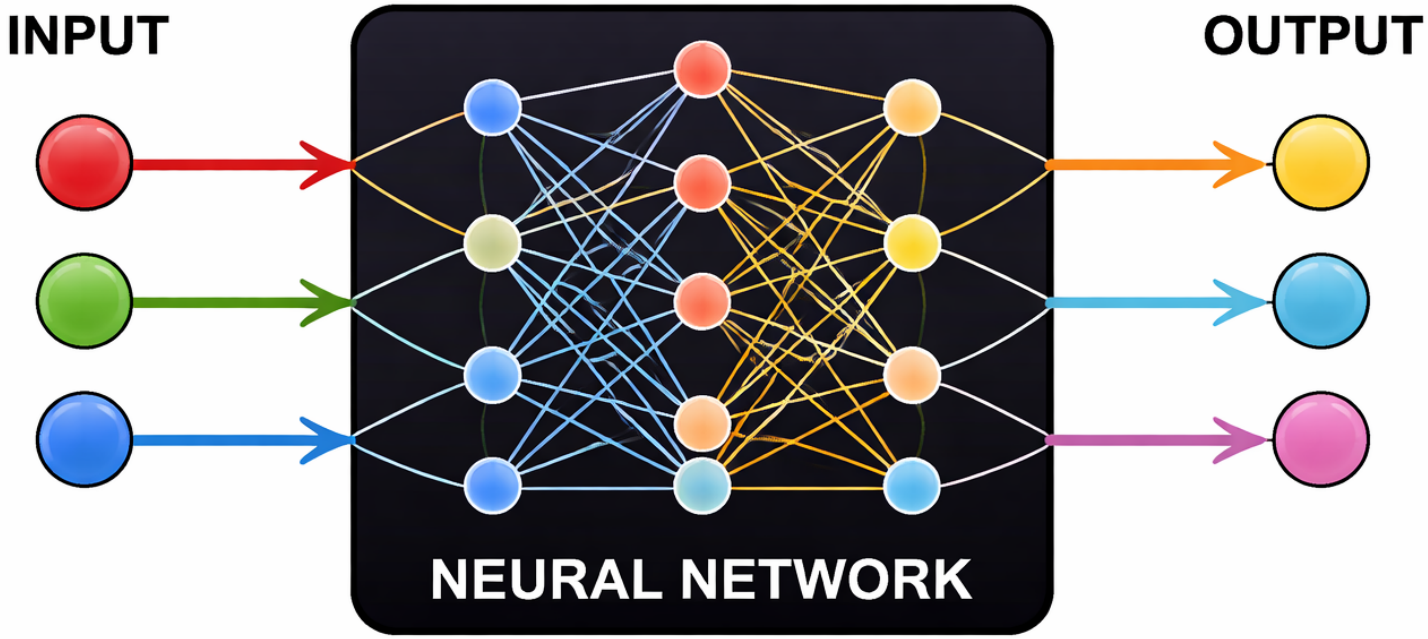
School of Mathematics and Maxwell Institute for Mathematical Sciences,
University of Edinburgh, UK

EFI 2026

Statistical Deep Learning of Environmental Extremes

Using **explainable** deep learning models for flexible and fast statistical modelling of environmental extremes

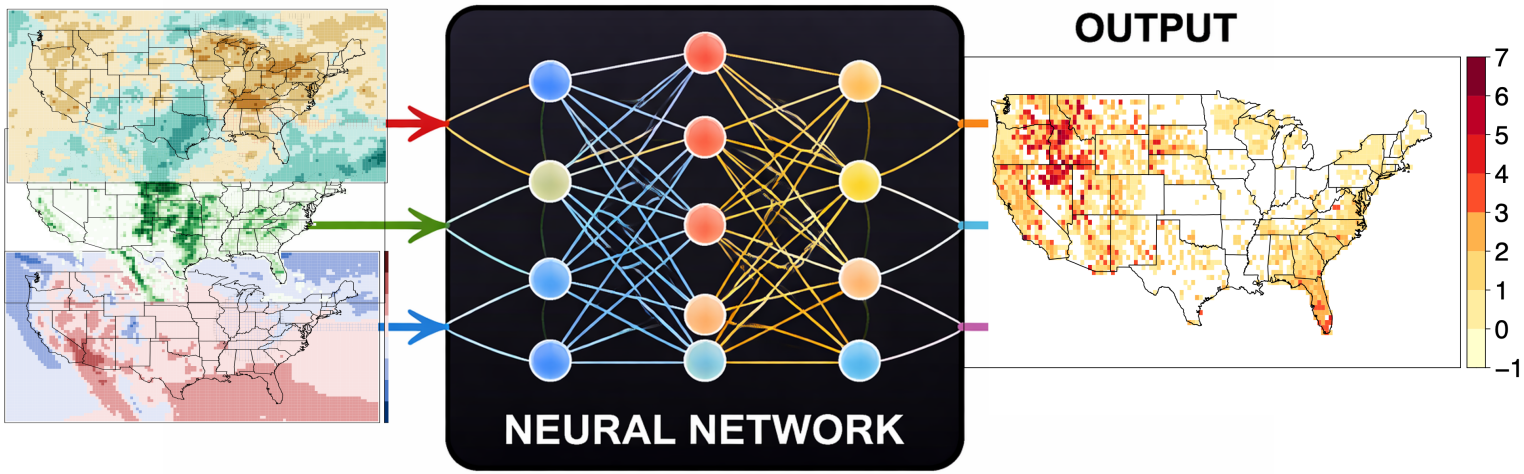
Deep learning 101 (Generated using ChatGPT)



Key ingredients:

- Some input data $\{\mathbf{x}_i\}_{i=1}^n$;
- A black-box mapping $\Rightarrow \mathcal{M} : \mathbf{x}_i \mapsto \hat{y}_i$;
- Loss function $\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}})$ - determines stats. properties of the output;

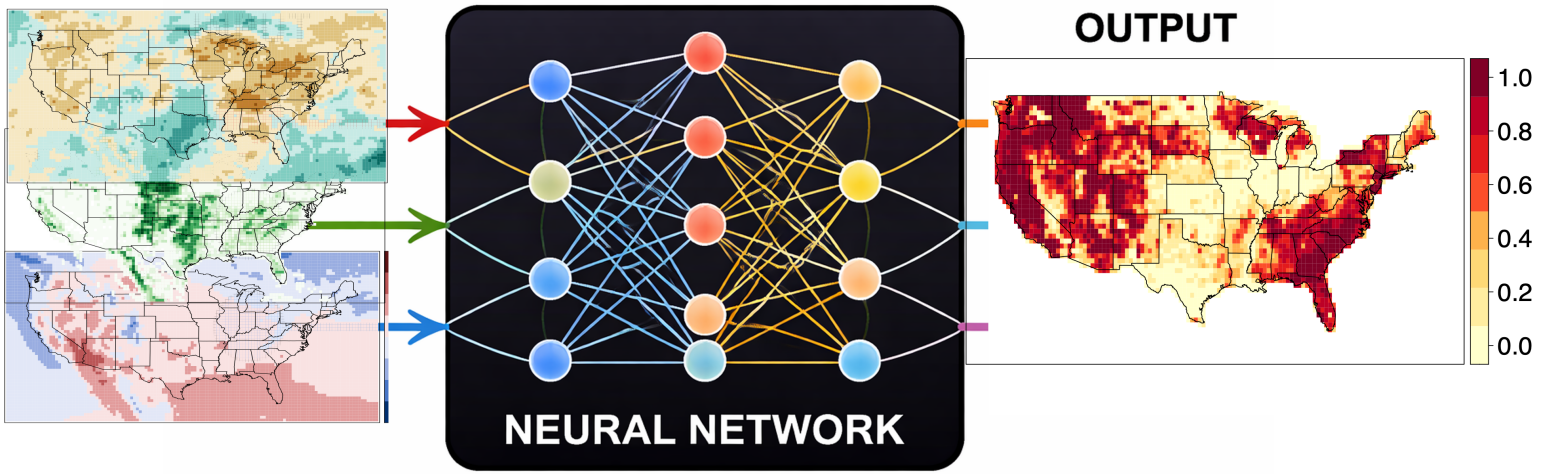
Environmental Deep learning



Key ingredients:

- $\mathbf{x}_i$: Spatial maps;
- A black-box mapping $\Rightarrow \mathcal{M} : \mathbf{x}_i \mapsto \hat{y}_i$;
- Least squares loss $\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}}) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \Rightarrow \hat{y}_i$ targets conditional expectation $\mathbb{E}[Y_i \mid \mathbf{X}_i = \mathbf{x}_i]$.

Environmental Deep learning



Key ingredients:

- $\mathbf{x}_i$: Spatial maps;
- A black-box mapping $\Rightarrow \mathcal{M} : \mathbf{x}_i \mapsto \hat{y}_i$;
- Binary cross-entropy $\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}}) = -\log\left(\prod_{i=1}^{n} \hat{y}_i^{y_i}(1-\hat{y}_i)^{1-y_i}\right)$
  $\Rightarrow \hat{y}_i$ targets conditional probabilities $\Pr[Y_i = 1 \mid \mathbf{X}_i = \mathbf{x}_i]$.

Deep Extremal Regression

Deep Extremal Regression Models

Use **EVT-compliant** loss functions so neural net targets useful risk metrics

- Risk metrics, i.e., **return levels**/quantiles/expectiles* (using asymmetric least squares loss)
- Extreme return levels and joint probabilities (using likelihood-based losses from, e.g., GPD/GEV models)

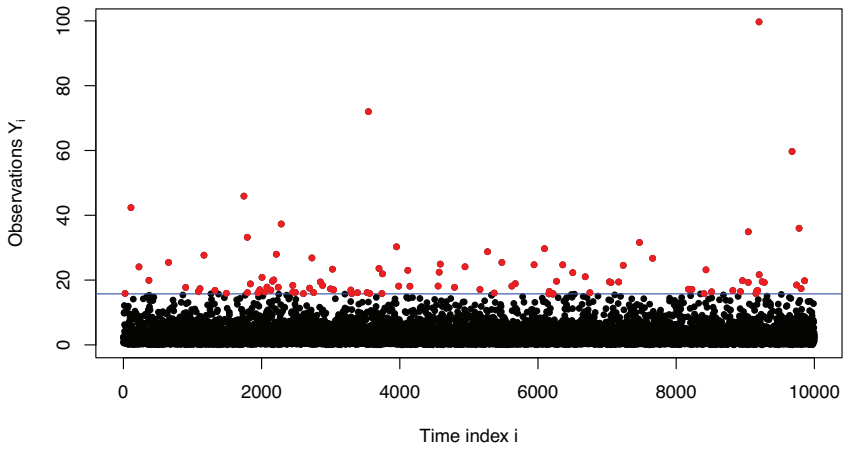
$$\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}}) = -\log\left(\prod_{i=1}^{n} f_{Y|\mathbf{X}=\mathbf{x}}(y_i|\hat{\mathbf{y}}_i)\right),$$

where $f_{Y|\mathbf{X}=\mathbf{x}}$ is the density of $Y \mid \mathbf{X}$ and $\hat{\mathbf{y}}_i := \hat{\mathbf{y}}_i(\mathbf{x}_i)$ is a parameter.

*Daouia, Girard, Richards. (2026+) Deep tail expectile regression neural networks.

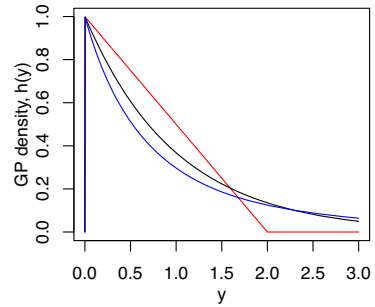
Deep Extremal Regression - Example

Threshold exceedance approach $Z_1, Z_2, \ldots, \overset{\text{iid}}{\sim} F$



Pickands–de Haan–Balkema Theorem: *For a broad range of distributions F, we have the following large-u approximation*

$$Y_i := Z_i - u \mid Z_i > u \sim \text{GPD}(\sigma, \xi).$$



Deep Extremal Regression - Example

Deep Extremal Regression Models

Use **EVT-compliant** loss functions so neural net targets useful risk metrics

- Using likelihood-based losses:

$$\mathcal{L}(\mathbf{y}, \hat{\mathbf{y}}) = -\log\left(\prod_{i=1}^{n} f_{Y|\mathbf{X}=\mathbf{x}}(y_i|\hat{\mathbf{y}}_i)\right),$$

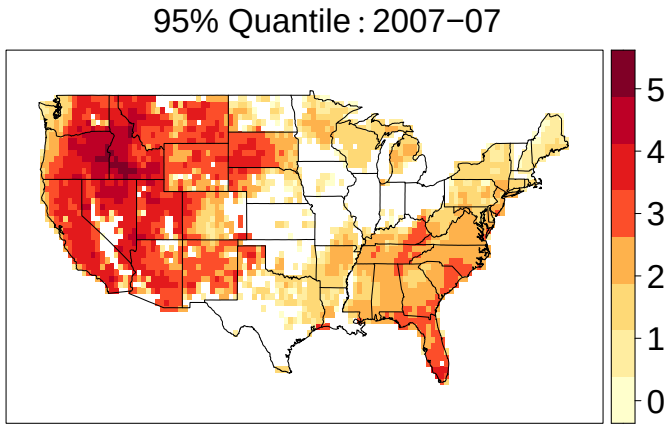
  where $f_{Y|\mathbf{X}=\mathbf{x}}$ is the density of $Y \mid \mathbf{X}$ and $\hat{\mathbf{y}}_i := \hat{\mathbf{y}}_i(\mathbf{x}_i)$ is a parameter.
- In a regression setting:

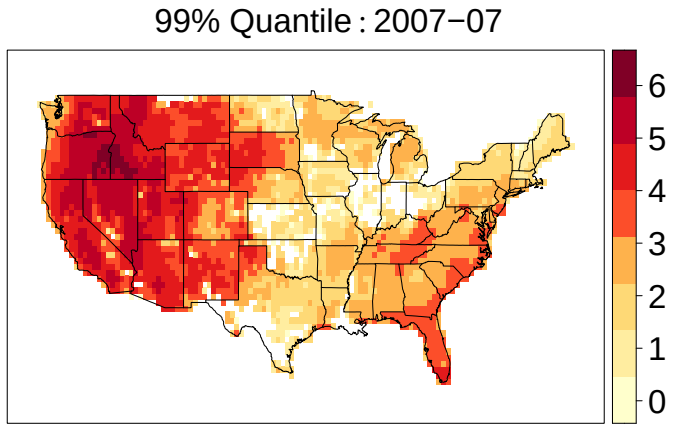
$$Y_i \mid (Y_i > 0, \mathbf{X}_i = \mathbf{x}_i) \sim \mathrm{GPD}(\sigma(\mathbf{x}_\mathrm{i}), \xi(\mathbf{x}_\mathrm{i})),$$

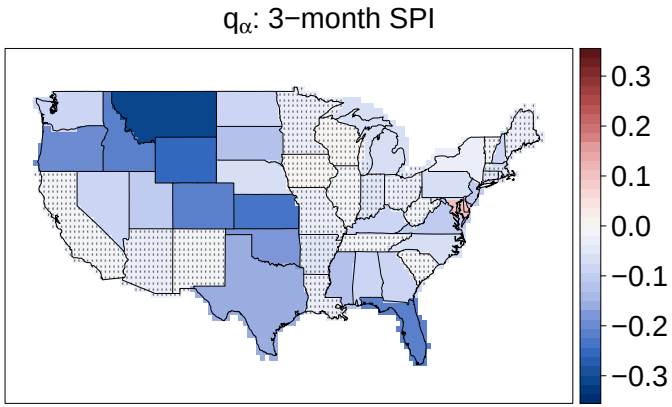
  and so $\hat{\mathbf{y}}_i := (\sigma(\mathbf{x}_i), \xi(\mathbf{x}_i))$ and $f_{Y|\mathbf{X}=\mathbf{x}}$ is the GPD density.
- Not new - just ask the experts (Carreau and Bengio, 2007)!

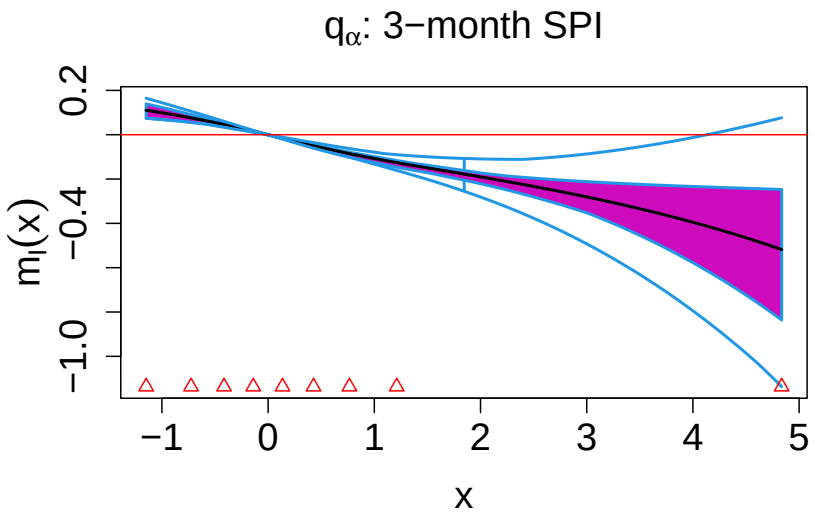
Carreau, Julie and Bengio, Yoshua. (2007) A Hybrid Pareto Model for Conditional Density Estimation of Asymmetric Fat-Tail Data. AIaS.

Extreme Wildfires - US (Richards and Huser, 2026b)



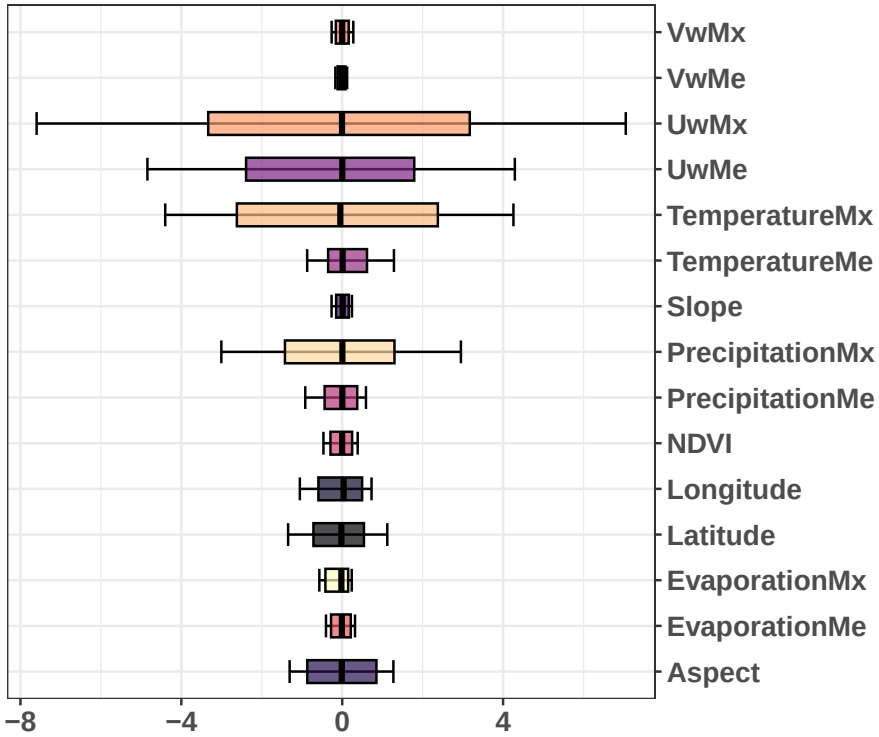






Richards, Huser (2026). Regression modelling of spatiotemporal extreme US wildfires via partially-interpretable neural networks. JCGS.

Wildfire Extremes - Australia (Cisneros et al., 2024)



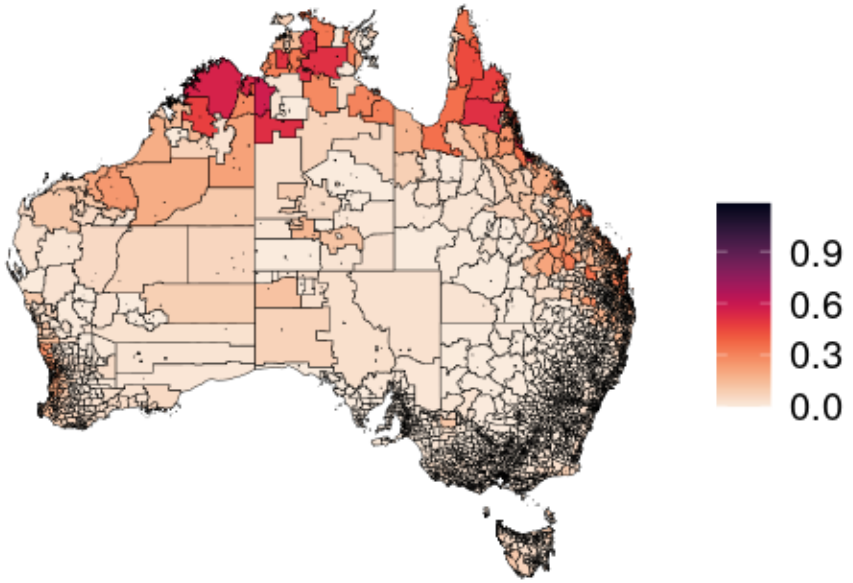


Figure: Left: boxplot of Shapley values. Right: return levels for wildfire risk.

Cisneros, Richards, Dahal, Lombardo, Huser (2024). Deep graphical regression for jointly moderate and extreme Australian wildfire. Spatial Statistics.

Wildfire Extremes - Mediterranean (Richards et al., 2023)

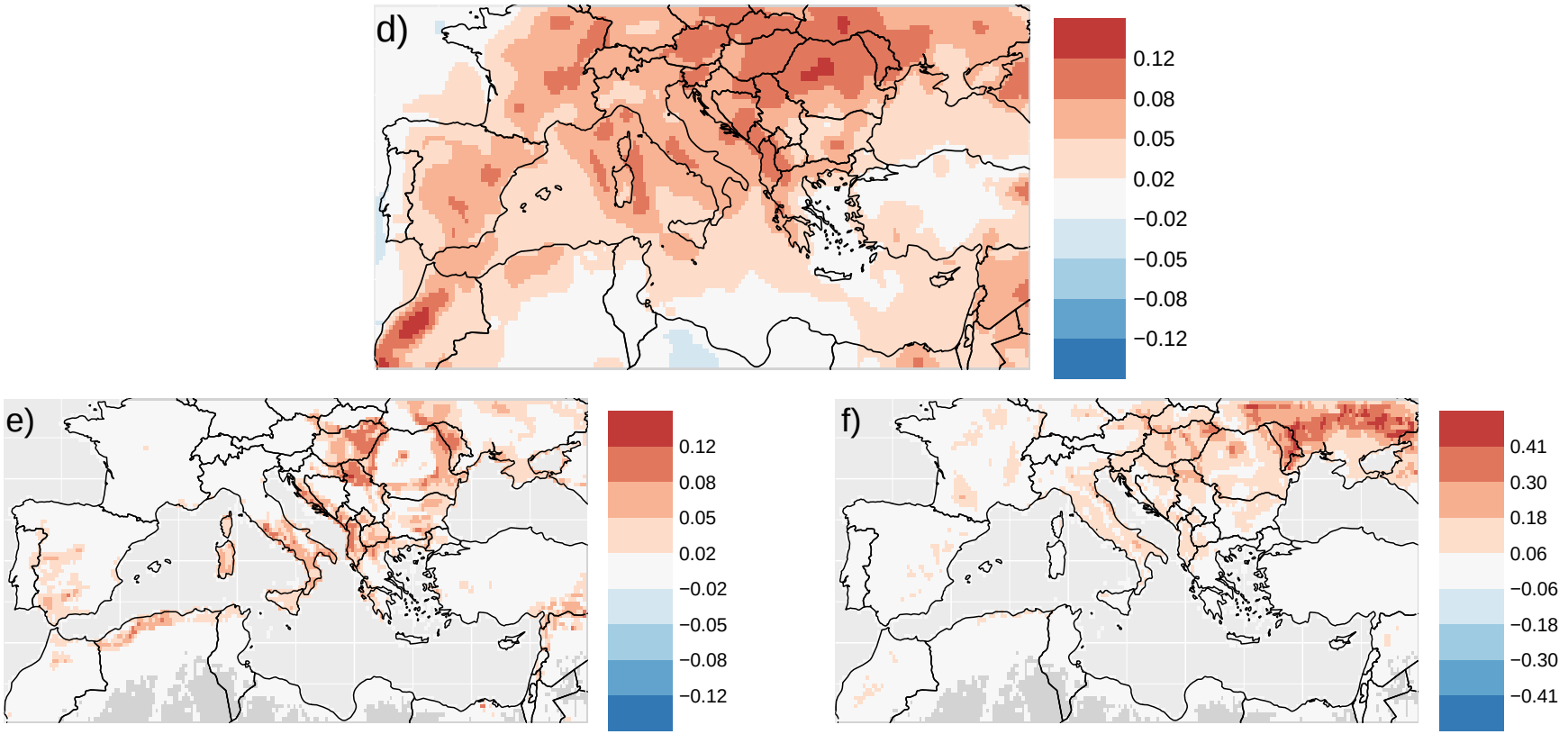


Figure: Trends in 2m air temperature [K]. Maps of changes in occurrence probability $p_0(s, t)$ and wildfire intensity for August 2001.

Richards, Huser, Bevacqua, Zscheischler, (2023). Insights into the drivers and spatio-temporal trends of extreme Mediterranean wildfires with statistical deep-learning. AIES.

Extreme UK Rainfall

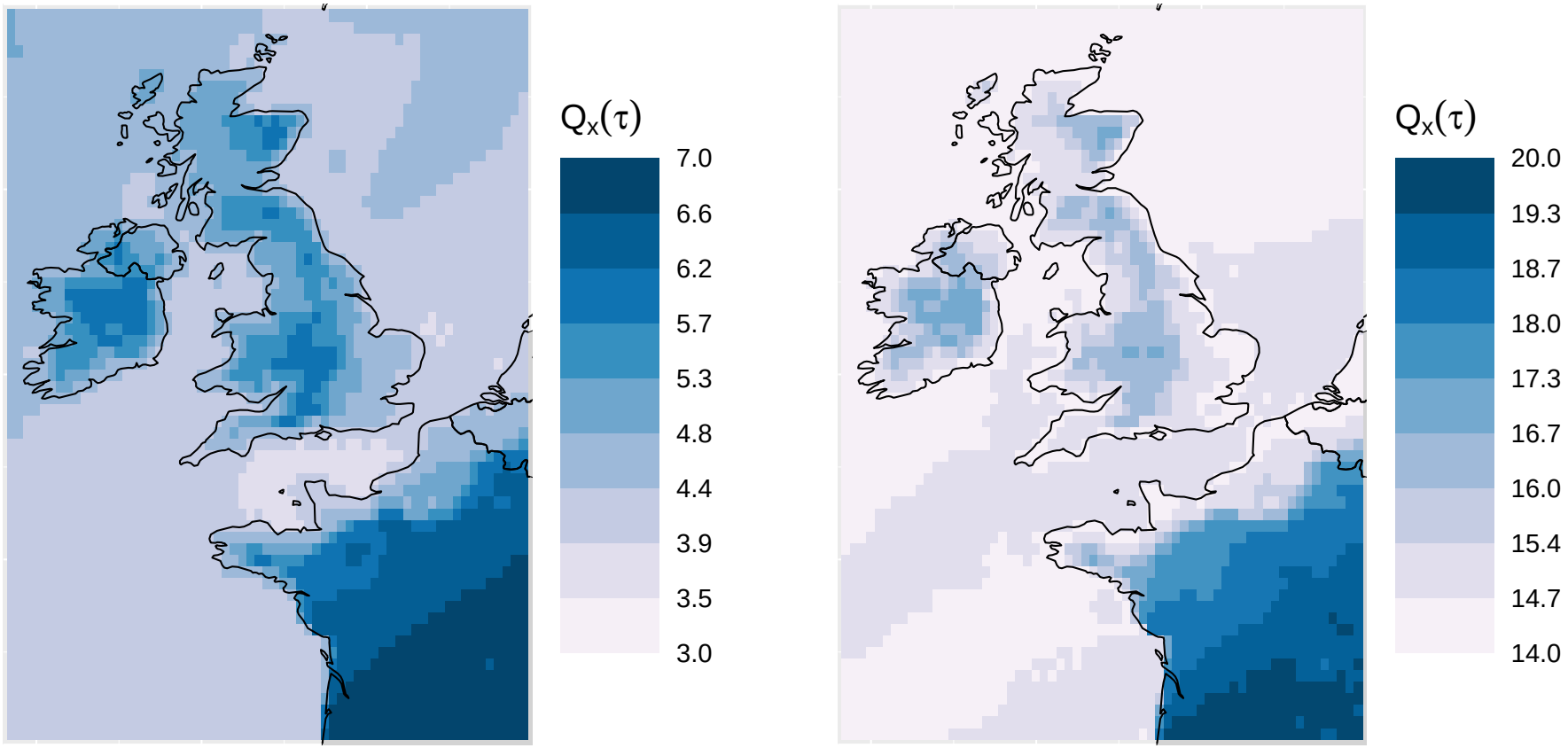
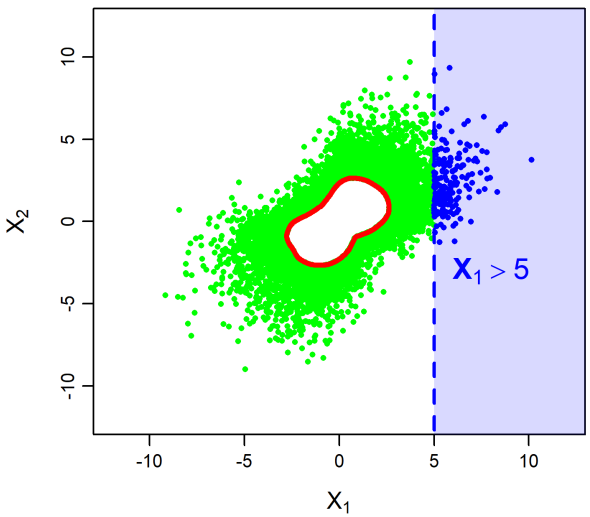
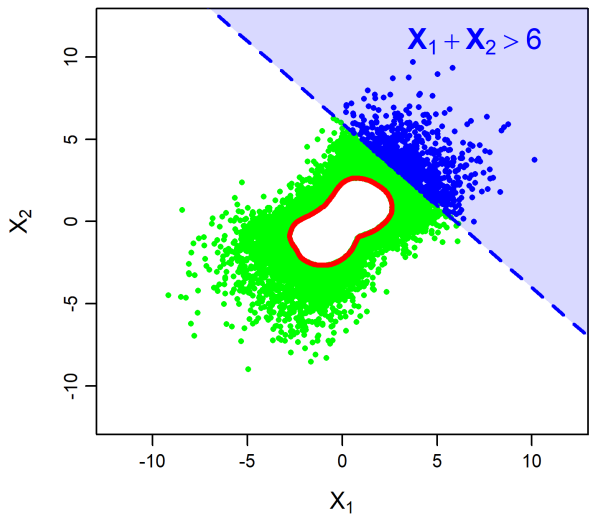


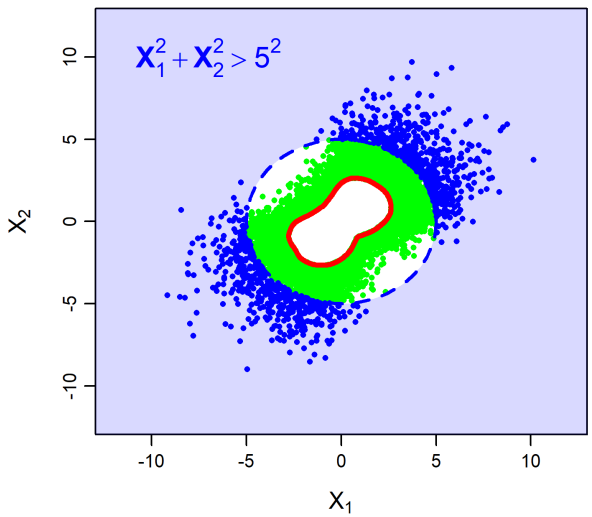
Figure: Maps of extreme conditional τ-quantiles, $Q_{\mathbf{x}}(\tau)$, of maximum hourly rainfall (mm) for July 2013, with $\tau = 0.9$ (left) and $\tau = 0.9999$ (right).

Richards, Huser (2026). Extreme Quantile Regression with Deep Learning. In Chapman and Hall/CRC Handbook on Statistics of Extremes.

Compound and multivariate extremes

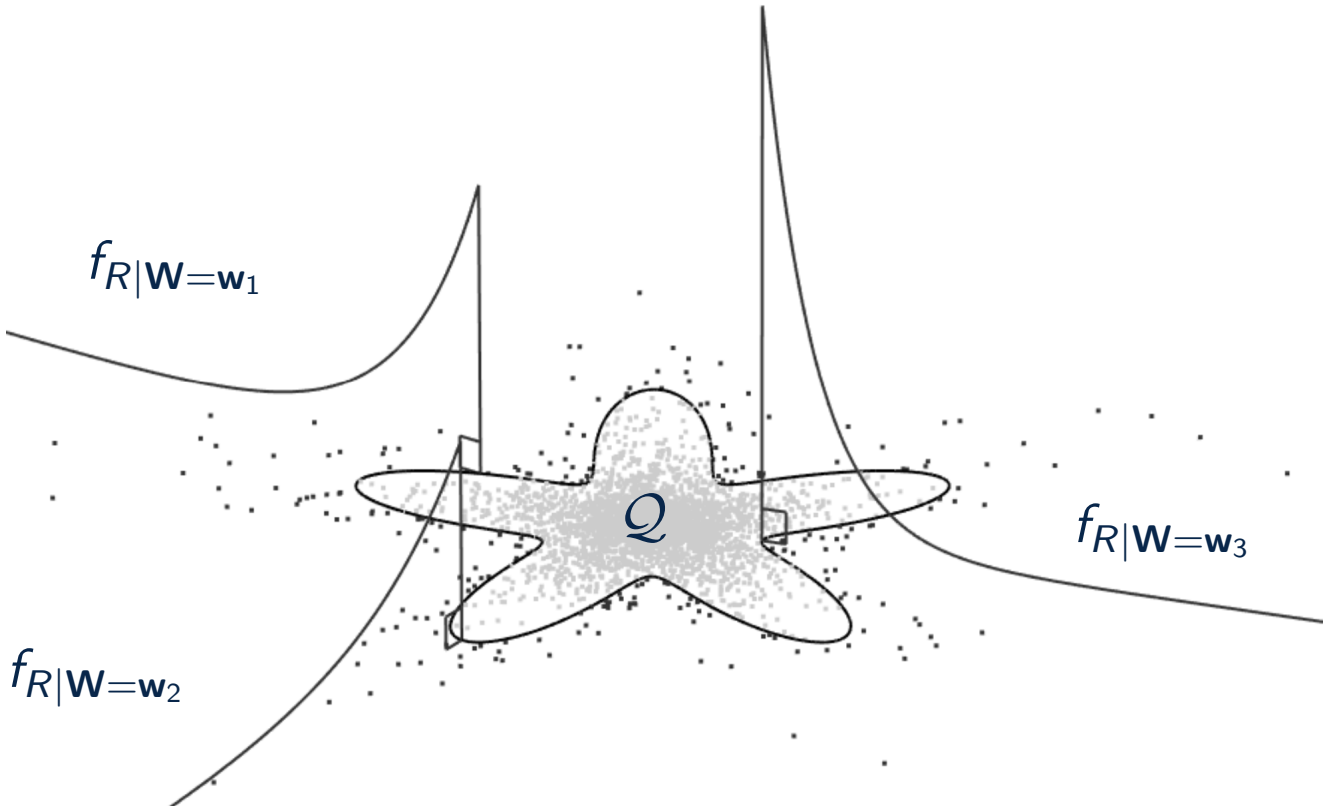






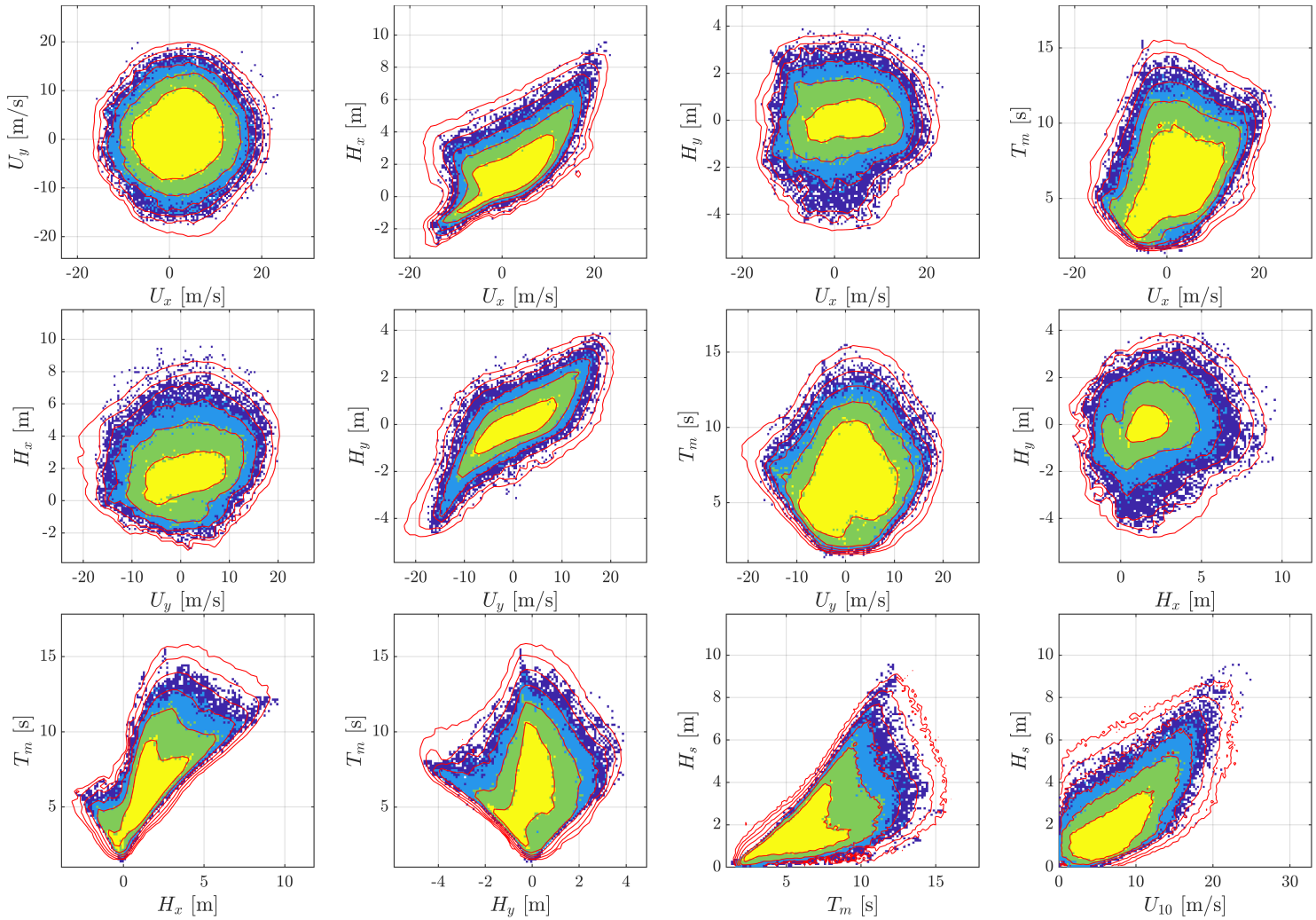
Murphy-Barltrop, Richards, Zscheischler (2026+). Exploring temporal changes in concurrent Alpine runoff extremes via statistical deep learning.

Compound and multivariate extremes



De Monte, Huser, Papastathopoulos, & Richards. (2025+). Generative modelling of multivariate geometric extremes using normalising flows. arXiv:2505.02957.

Metocean Extremes - Celtic Sea (Mackay et al., 2025)



Mackay, Murphy-Barltrop, Richards, Jonathan (2025). Deep Learning Joint Extremes of Metocean Variables Using the SPAR Model. JOMAE.

US Wind Extremes (De Monte et al., 2025)

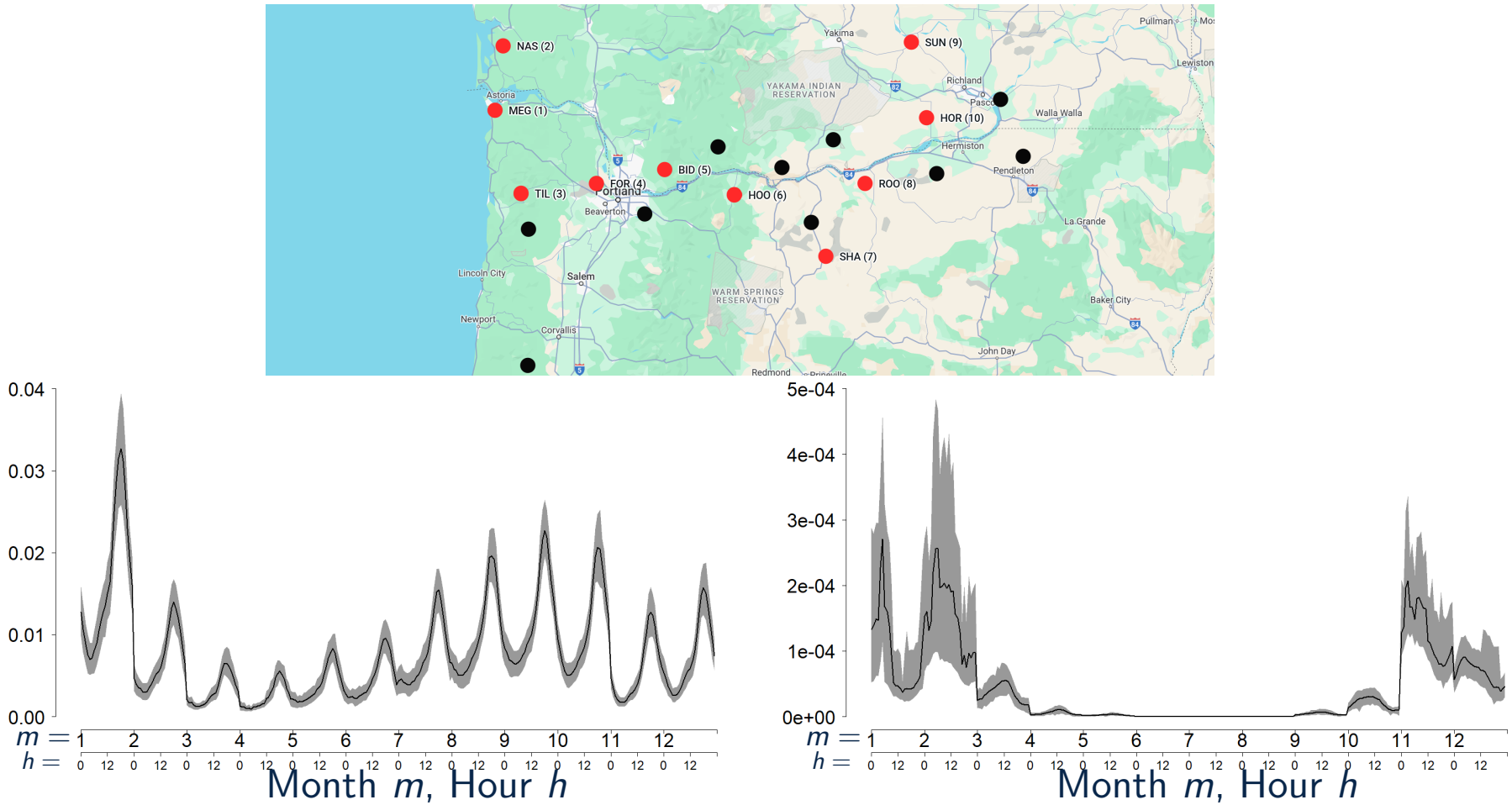
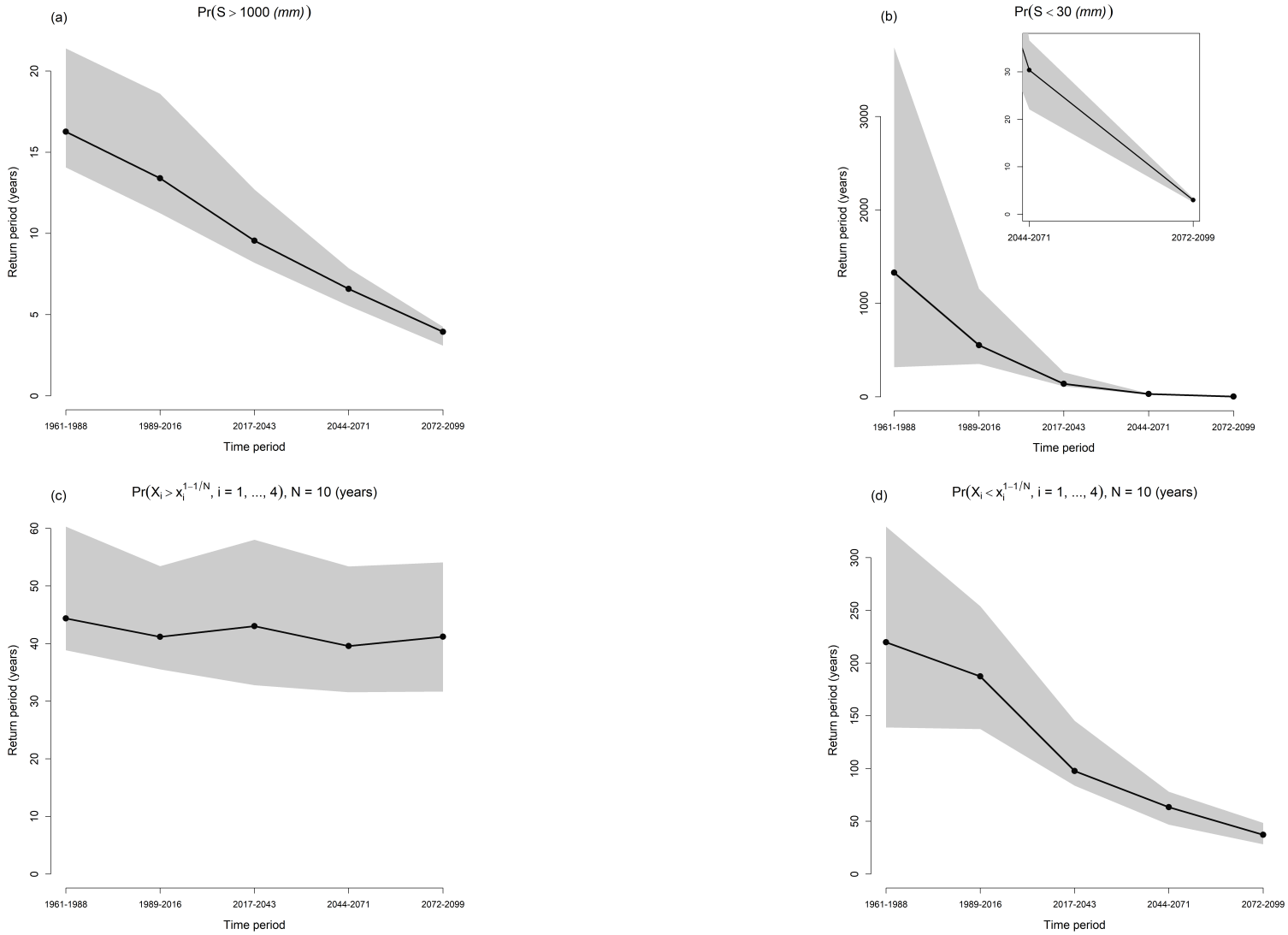


Figure: Estimated probability of no power (left) and maximum power (right).

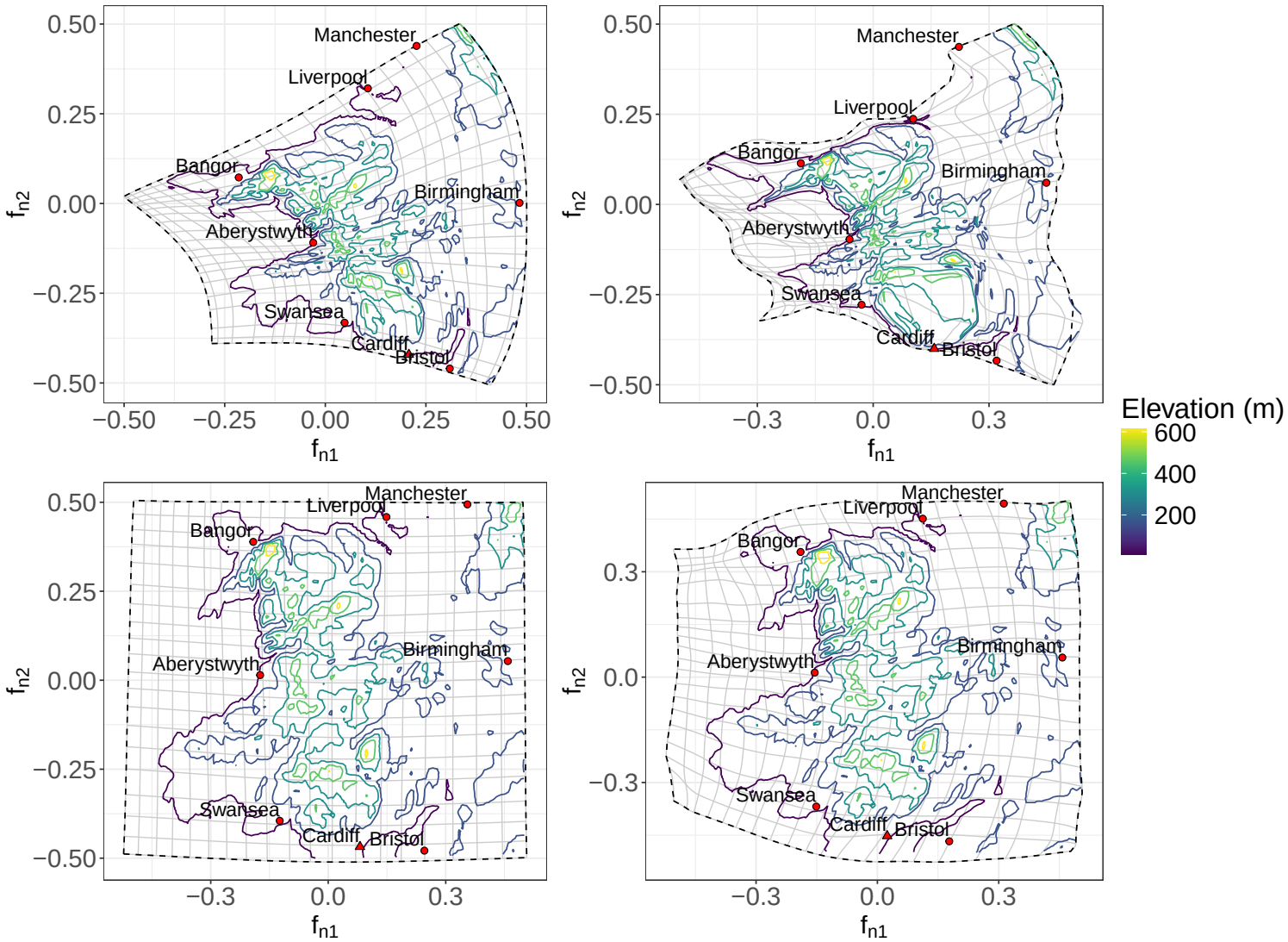
De Monte, Huser, Papastathopoulos, & Richards. (2025+). Generative modelling of multivariate geometric extremes using normalising flows. arXiv:2505.02957.

Joint Alpine runoff extremes



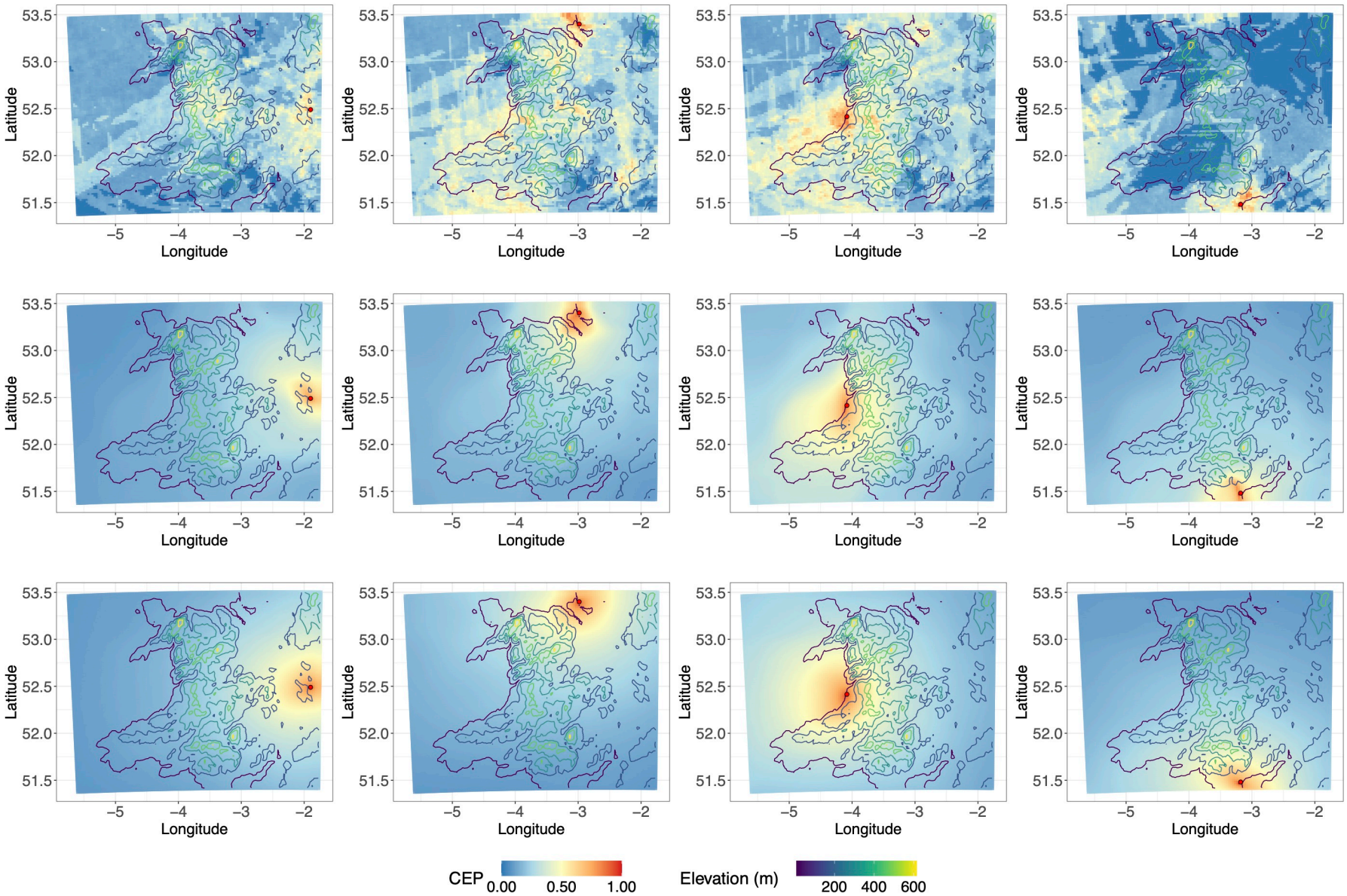
Murphy-Barltrop, Richards, Zscheischler (2026+). Exploring temporal changes in concurrent Alpine runoff extremes via statistical deep learning.

UK rainfall spatial extremal extent (Shao et al., 2025)



Shao, Richards, and Huser. (2025+). Modeling nonstationary extremal dependence via deep spatial deformations.

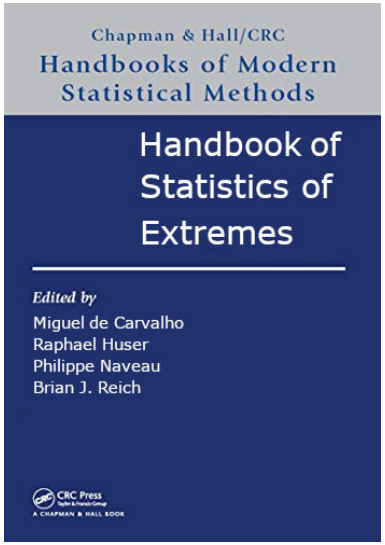
UK rainfall spatial extremal extent (Shao et al., 2025)



Conclusion

- Deep extremal regression models are conceptually simple yet powerful tools for modelling extremes - both **univariate** and **multivariate**.
- Despite their black box nature, neural networks can be used alongside statistical models to create something (atleast partially-)interpretable.
- Very useful in many environmental applications.

- Easy to implement - lots of code available on GitHub:
  - Jbrich95/pinnEV/USWildfireExtremes;
  - CASE2025_shortcourse/extQuantRegressDL/cde-RKeras-intro;
  - callumbarltrop/DeepSPAR/DeepGauge;
  - reetamm/SPQRx



Richards and Huser (2026a).
Extreme Quantile Regression with Deep Learning.
Code and short courses on GitHub - `https://github.com/Jbrich95/`

Gimeno-Sotelo, Richards, Hazra, Mhalla, and de Zea Bermudez. (2026).
A Review of Applications of Extreme Value Theory to Environmental Risk Assessment. Environmental Statistics: Innovative Methods and Applications. Taylor & Francis. To appear.

References I

Carreau, J. and Bengio, Y. (2007). A hybrid Pareto model for conditional density estimation of asymmetric fat-tail data. In *Artificial Intelligence and Statistics*, pages 51–58. PMLR.

Cisneros, D., Richards, J., Dahal, A., Lombardo, L., and Huser, R. (2024). Deep graphical regression for jointly moderate and extreme Australian wildfires. *Spatial Statistics*, 59:100811.

De Monte, L., Huser, R., Papastathopoulos, I., and Richards, J. (2025). Generative modelling of multivariate geometric extremes using normalising flows. *arXiv preprint arXiv:2505.02957*.

Mackay, E., Murphy-Barltrop, C. J. R., Richards, J., and Jonathan, P. (2025). Deep Learning Joint Extremes of Metocean Variables Using the SPAR Model. *Journal of Offshore Mechanics and Arctic Engineering*, 148(021201).

Richards, J. and Huser, R. (2026a). Extreme Quantile Regression with Deep Learning. In de Carvalho, M., Huser, R., Naveau, P., and Reich, B. J., editors, *Handbook on Statistics of Extremes*. Chapman & Hall/CRC, Boca Raton, FL.

Richards, J. and Huser, R. (2026b). Regression modelling of spatiotemporal extreme US wildfires via partially-interpretable neural networks. *JCGS, to appear*.

Richards, J., Huser, R., Bevacqua, E., and Zscheischler, J. (2023). Insights into the drivers and spatiotemporal trends of extreme mediterranean wildfires with statistical deep learning. *Artificial Intelligence for the Earth Systems*, 2(4):e220095.

Shao, X., Richards, J., and Huser, R. (2025). Modeling nonstationary extremal dependence via deep spatial deformations. *arXiv preprint arXiv:2505.12548*.

Fin.

Scan for webpage.