Benchmarking study

The conditional exceedance model

To provide further benchmarking, we compare our results against the conditional exceedance model of Heffernan and Tawn (2004). The method uses a semiparametric model for the marginals and disentangles the marginal distributions from the dependence structure by transforming all margins to unit Gumbel distributions prior to learning the dependence structure, similar to our approach.

Since we use similar methods to fit the marginal distributions, we only compare our model to the Heffernan and Tawn (2004) model for its ability to model the dependence structure between variables. To do this, we supply the model with the training dataset of storm events, already transformed to a uniform distribution as described in the main text. We modify the *marginalDependence* and *predixt.mex* functions from the R *texmex* package to work with uniform-distributed input and output data. *I.e.*, we remove the generalised Pareto distribution fitting and transformations from the functions. The method models the conditional distribution of a set of variables, conditional that the conditioning variable exceeds some extreme threshold. The objective function is,

$$Q_{|i}(\theta_{|i}, \lambda_{|i}) = -\sum_{j\neq i}\sum_{k=1}^{n}\left[\log\left(\sigma_{j|i}(y_{i|i,k})\right) + \frac{1}{2}\left(\frac{y_{j|i,k} - \mu_{j|i}(y_{i|i})}{\sigma_{j|i}(y_{i|i})}\right)^2\right], \tag{S1}$$

with

$$\begin{aligned}
\mu_{j|i}(y) &= a_{j|i}(y) + \mu_{j|i} b_{j|i}(y)\\
\sigma_{j|i}(y) &= \sigma_{j|i} b_{j|i}(y)\\
a_{j|i}(y) &= a_{j|i} y + I_{\left[a_{j|i}=0, b_{j|i}<0\right]} \cdot \left[c_{j|i} - d_{j|i}\log(y)\right]\\
b_{j|i}(y) &= y^{b_{j|i}},
\end{aligned}$$

where y are unit Gumbel distributed, i is the index of the conditioning variable, j is the index of the dependent variable and $|i$ denotes all the dependent variables. The model is fit for each conditioning variable y_i by optimising Q via the parameters $\boldsymbol{\lambda}_{|i} = (\boldsymbol{\mu}_{|i}, \boldsymbol{\sigma}_{|i})$ and $\boldsymbol{\theta}_{|i} = (\mathbf{a}_{|i}, \mathbf{b}_{|i}, \mathbf{c}_{|i}, \mathbf{d}_{|i})$. See Heffernan and Tawn (2004) for a full discussion of the method.

Benchmarking method

Since the hazGAN model was trained on a subset of the 150 storms with peak wind speed anomaly exceeding 15 ms^{-1}, we compare results for the dependence fits on this subset only.

To benchmark, we randomly sample 1,000 locations from the training data domain. For each point, we fit a conditional dependence model between wind speed, precipitation, and sea-level pressure, conditioning on the wind speed. From the fitted model, we sample 914 points and calculate the tail dependence coefficient between all pairs of hazard variables. We compare this to the tail dependence coefficients between variables for the training data and the hazGAN-generated data.

We assess the spatial dependence similarly. For each of the 1,000 sampled locations, we randomly sample a second point on the domain with which to compare the spatial dependence fits. We fit a conditional dependence model between each pair of points for each of the three hazard variables, conditioning on the original point, and simulate 914 samples from each of the fitted models. We compare the tail dependence metrics of data simulated by the conditional exceedance model to the training data and to those generated by hazGAN.

Note that in each case with the conditional exceedance model we have fitted a two- or three-variable multivariate model, while the hazGAN model was trained once on all variables simultaneously.

Benchmarking results

Figure shows the results for the multi-hazard benchmarking. A scatter plot of the empirical tail dependence coefficient estimates ($\hat{\lambda}$) calculated from training data is plotted against coefficients calculated from 914 synthetic samples, for each of the 1,000 benchmarking locations. For the wind versus precipitation and wind versus sea-level pressure (negated), we see good agreement between all models, with the conditional exceedance model displaying higher variance in the fits. The conditional model shows some bias towards underestimating the tail dependence while hazGAN model shows some bias towards overestimating it. For data points with lower tail dependence between hazard variables ($\hat{\lambda} < 0.6$), the hazGAN model displays some overconfidence, and the spread of points does not cross the perfect fit line (red). Some further investigation may be warranted here, to determine the source of this overestimation of inter-variable dependence.

The left bottom subplot shows the simulated values of total precipitation and sea-level pressure when conditioned on wind speed. In this case, no dependence between the variables was modelled, so the poorer performance is expected. To model the dependence structure between precipitation and sea-level pressure we would need to fit another model, conditional on one of the two variables.

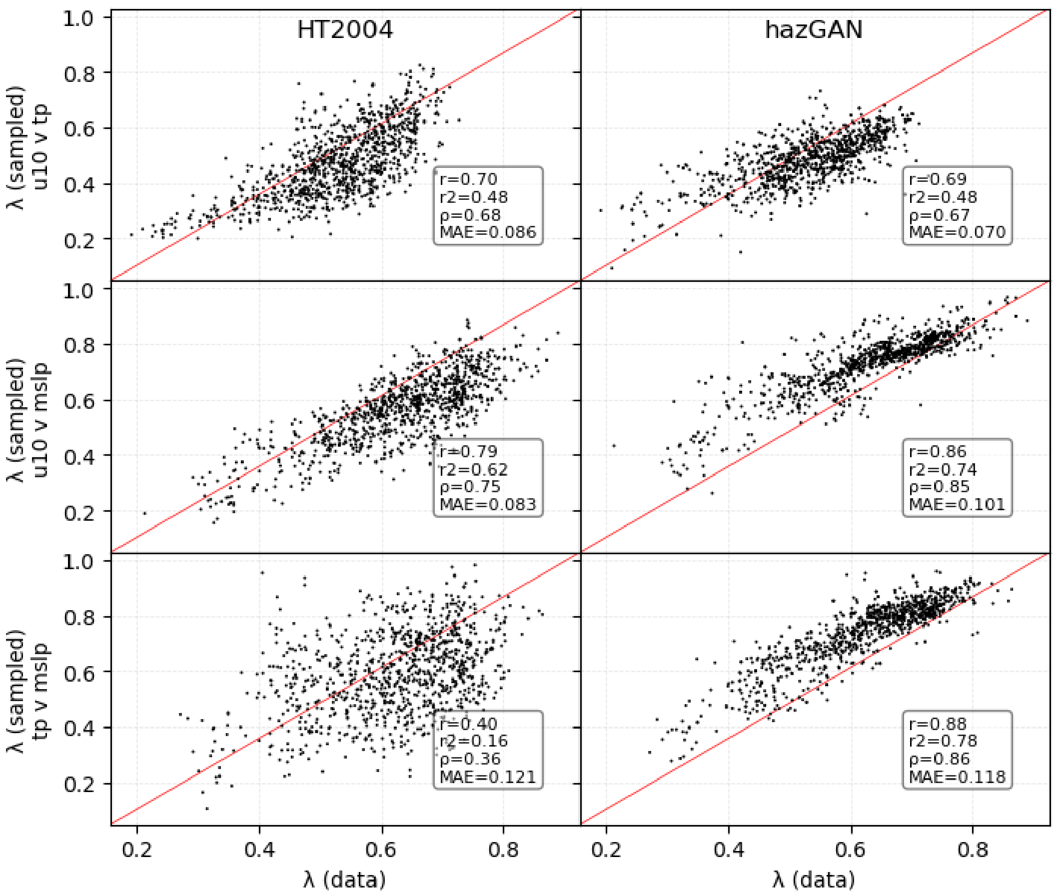

Figure S1. Comparison of multi-hazard fits for the conditional exceedance model and hazGAN. Scatterplot shows the multi-hazard tail dependence metrics at 1,000 points in the domain between the training data (x-axis) and the synthetic data (y-axis). Metrics: Correlation coefficient (r), R^2 metric (r2), Spearman correlation coefficient (ρ), and mean-absolute-error (MAE).

Figure compares the same metrics for the spatial extremal dependence for 1,000 pairs of points in the domain. These are calculated between two different locations for the same variables, conditioning the second point on the first. All models display good fits, with correlation coefficients exceeding 0.9 and mean-absolute-errors below 1%. Again, the hazGAN shows marginally lower variance than the conditional exceedance model.

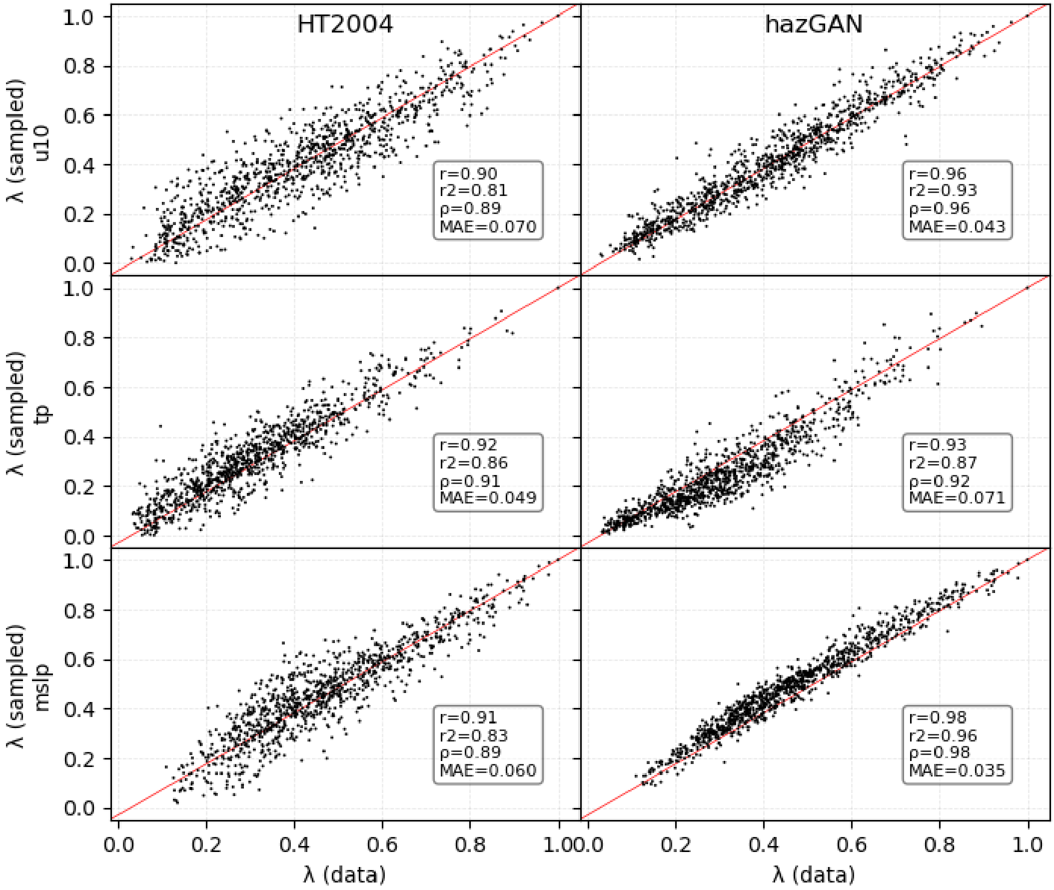

Figure S2. Comparison of spatial fits for the conditional exceedance model and hazGAN. Scatterplot shows the spatial tail dependence metrics between 1,000 pairs of points in the domain between the training data (x-axis) and the sampled data (y-axis). Metrics: Correlation coefficient (r), R^2 metric (r2), Spearman correlation coefficient (ρ), and mean-absolute-error (MAE).

Overall, this benchmarking provides empirical evidence towards the ability of hazGAN approach to learn the extremal structure between many variables and produce results that are competitive with existing multivariate dependence models.

References

Heffernan, J. E. and Tawn, J. A.: A conditional approach for multivariate extreme values (with discussion), Journal of the Royal Statistical Society Series B: Statistical Methodology, 66, 497–546, 2004.