Joint Generalized Neural Models and Censored Spatial Copulas for Probabilistic Rainfall Forecasting

David Huk, Rilwan Adewoyin and Ritabrata Dutta

Department of Statistics, University of Warwick, UK

Introduction

Want a model for future **joint distributions** of rainfall across space. Use the copula trick: (1) modelling marginal distributions at each locations separately and (2) uniting them with copulas to infuse a spatial dependence. For predictors $\boldsymbol{\mathcal{X}}_{i,t} = (\mathcal{X}_{i,t}^{(1)}, \ldots, \mathcal{X}_{i,t}^{(6)})$ and target rainfall $\boldsymbol{\mathcal{Y}}_t = (y_{1,t}, \ldots, y_{n,t})$ at times $t \in \mathbb{Z}$ and locations i, combined into present information $\boldsymbol{\mathcal{P}}_{t_0} = \{\boldsymbol{\mathcal{X}}_{i,t}, \boldsymbol{\mathcal{Y}}_t : t < t_0\}$, we write our model for future $s \geq t_0$ through Sklar's theorem as:

$$\underbrace{\mathcal{M}\left(\boldsymbol{\mathcal{Y}}_s | \boldsymbol{\mathcal{P}}_{t_0}\right)}_{\text{Joint Density}} = \underbrace{f_1\left(y_{1,s} | \boldsymbol{\mathcal{P}}_{t_0}\right) \cdot \ldots \cdot f_n\left(y_{n,s} | \boldsymbol{\mathcal{P}}_{t_0}\right)}_{\text{JGNM}} \cdot \underbrace{\mathbf{c}\left(F_1\left(y_{1,s} | \boldsymbol{\mathcal{P}}_{t_0}\right), \ldots, F_n\left((y_{n,s} | \boldsymbol{\mathcal{P}}_{t_0}\right)\right)}_{\text{Censored Copula}}$$

Now the question is: "How to model marginal **densities** f_i and how to model the **copula c** to capture the **spatio-temporal** nature of the data ?". We propose novel solutions for both parts.

1. Joint Generalised Neural Models for Marginals

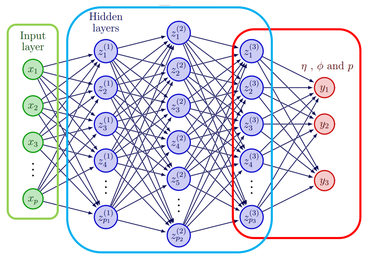
Have to take particularities of rainfall data into account when modelling $f_i\left(y_{i,s} | \boldsymbol{\mathcal{P}}_{t_0}\right)$.

- Rainfall data is **zero-inflated** (over-representation of 0-valued observations).
- Rainfall can be > 0 with a **continuous** probability density function on $\mathbb{R}_{>0}$.
- Rainfall can also never occur, resulting in a **discrete** observation with mass at 0.

We use a mixture model, giving seperate probability to 0 and positive rain as:

$$f_i\left(y | \boldsymbol{\mathcal{P}}\right) = [1 - p] \cdot \delta_0\left(y\right) + p \cdot \mathcal{G}\left(y; \mu, \phi\right) = f_i(y | \mu, \phi, p).$$

Expand on existing GLM and NN literature to create JGNMs- a model to learn parametric densities with the combined strengths of both approaches:



Green box: Initial Input data, a subset of $\boldsymbol{\mathcal{P}}_{t_0}$ relevant to location i.
Blue box: NN to refine the data into new predictors, captures non-linear trends.
Red box: Simultaneous GLMs for parameters of model, using the refined predictors.

2. Censored Spatial Copula

Assume **c** is a Gaussian copula, inducing a known form for its density and allowing us to learn the spatial dependence. But we have to take into account the zero-inflation of data, so we modify the classical Gaussian copula through **censoring** :

1. Transform all observations to a common Gaussian scale using probability integral transforms with their JGNM marginals giving $x_{i,t} = \Phi^{-1}(u_{i,t}) = \Phi^{-1}(F(y_{i,t} | \mu_{i,t}, \phi_{i,t}, p_{i,t}))$.
2. All 0-valued observations $y_{i,t}$ get mapped to the censoring levels $d_{i,t} = \Phi^{-1}(1 - p_{i,t})$.
3. Modify the copula density **c** to reflect said censoring.

The resulting density has to integrate over the censoring levels of dry locations, yielding

$$\mathbf{c}(x_{1,t}, ..., x_{n,t}) = \frac{\int_{-\infty}^{d_{i_1,t}} \ldots \int_{-\infty}^{d_{i_k,t}} \phi_n(z_{1,t}, ..., z_{n,t}; \mathbf{0}, \boldsymbol{\Sigma})\, dx_{i_1,t} \ldots dx_{i_k,t}}{\prod_{\{i:y_{i,t}>0\}} \phi(x_{i,t}) \cdot \prod_{\{j:y_{j,t}=0\}} \Phi(d_{j,t})}$$

where $\boldsymbol{\Sigma}$ is a covariance matrix that we have to estimate. Construct $\boldsymbol{\Sigma}$ using GPKs, relying on parameter $\boldsymbol{\theta}$ and distance $\mathcal{D}$: $\boldsymbol{\Sigma}_{(i,j)} = k(\mathcal{D}_{(i,j)} | \boldsymbol{\theta}) = \frac{2^{1-\nu}}{\Gamma(\nu)} \cdot \left(\sqrt{2\nu}\frac{\mathcal{D}_{(i,j)}}{\boldsymbol{\theta}}\right)^{\nu} \cdot K_\nu \left(\sqrt{2\nu}\frac{\mathcal{D}_{(i,j)}}{\boldsymbol{\theta}}\right)$.
We specify $\mathcal{D}$ with relevant spatial information and only need to estimate $\boldsymbol{\theta}$.

3. Copula Estimation via Scoring Rules

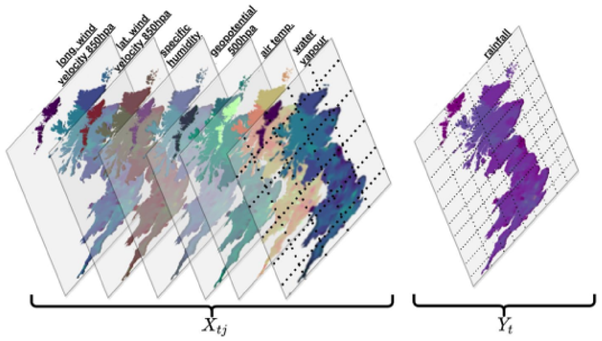
Estimation of Σ via MLE becomes expensive due to integration in **c**. Instead rely on **likelihood-free inference via Scoring Rules**. Optimise Energy Score as divergence between simulations and observations:

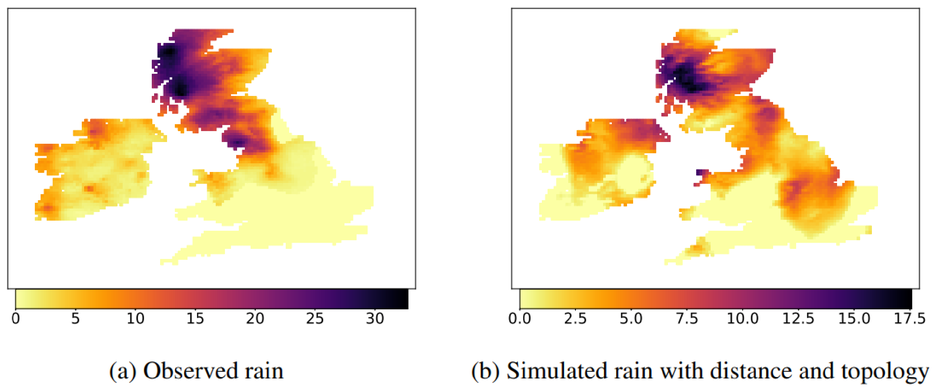
$$SR = 2 \cdot \mathbb{E}_{\mathbf{X}' \sim \mathbf{c}_\theta} \|\mathbf{X}' - \mathbf{x}\|_2 - \mathbb{E}_{\mathbf{X}'_1, \mathbf{X}'_2 \sim \mathbf{c}_\theta} \|\mathbf{X}'_1 - \mathbf{X}'_2\|_2$$

For selected $\boldsymbol{\theta}$ construct Σ to simulate data from $\phi_n(.|\mathbf{0}, \Sigma)$ and censor with $d_{i,t}$; these are simulations from $\mathbf{c}_{\boldsymbol{\theta}}$. Choose $\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} SR(\boldsymbol{\theta} | \{x_{i,t} : i \in \{1, \ldots, n\}, t < t_0\})$.

4. Application to UK Rainfall

- Data is on a 100*140 grid covering the UK with 7 predictors and target rain for 40 years with a 20 years/20 years train/test split.



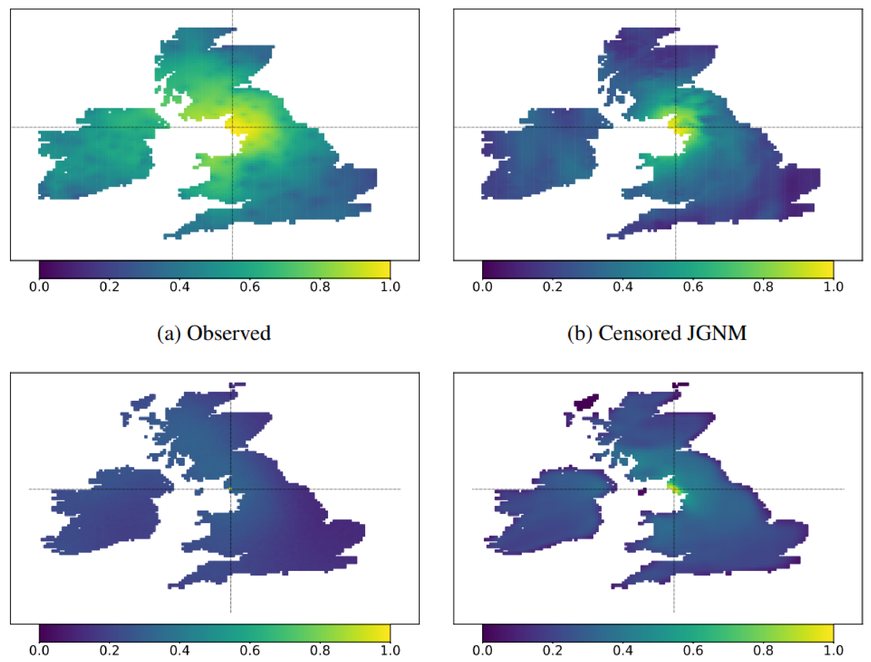


(a) Observed rain (b) Simulated rain with distance and topology

4. Application to UK Rainfall

- Compare our model **Cens-JGNM** against benchmark **ConvCNP** and **VAE-GAN**.
- Assess **calibration** (a), **spatial-coherence** (b) and **precision** (c) of forecasts.

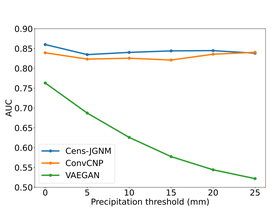
Diagnostic Metrics			
Model	Cens-JGNM	ConvCNP	VAEGAN
CRPS (a), (c)	**1.1613**	1.6536	4.5059
Energy Score (a), (c), (b)	**2.6184**	3.1003	3.2127
Variogram Score (b)	**2.9 e+6**	7.9 e+6	9.8 e+6
RMSB (c)	**3.8642**	4.3073	4.8861
MAB (c)	**1.839**	3.0266	1.9865



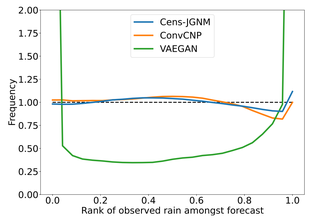
(a) Observed (b) Censored JGNM

(c) ConvCNP (d) VAE GAN

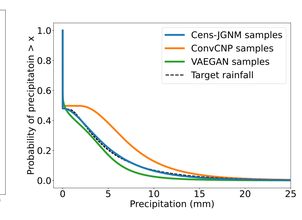
Cross-correlation maps, (b) .



Area Under Curve for calibration task. (a), (c) .



Rank histograms of marginal densities. (a).



Observed and simulated frequencies of rainfall. (a).

Conclusion

Contribution 1: JGNM which unite two powerful frameworks into a reliable learner of parametric distributions with complex data patterns such as rainfall.
Contribution 2: Censored Gaussian Copulas as a rigorous treatment of zero-inflated joint density modelling, or any data with censoring.
Contribution 3: Main contribution of this work, a methodology for the estimation of copulas without relying on the likelihood, much cheaper in most cases especially in high dimensions.