Partially-interpretable neural networks for high-dimensional extreme quantile regression: With application to wildfires within the Mediterranean Basin

Jordan Richards[1] Raphaël Huser[1]
Emanuele Bevacqua[2] Jakob Zscheischler[2]

[1]King Abdullah University of Science and Technology (KAUST)

[2]Helmholtz Centre for Environmental Research (UFZ)

Data

- We are interested in identifying the **drivers** of wildfire occurrence and extreme spread $\Rightarrow$ logistic/parametric extreme quantile **regression**
- Spatial domain is **Mediterranean Basin** and southern Europe
- MODIS monthly **burnt area** (BA) for $0.5^\circ \times 0.5^\circ$ grid-cell
- 2001-2020, all months. 10083 locations, 116337 non-zero values

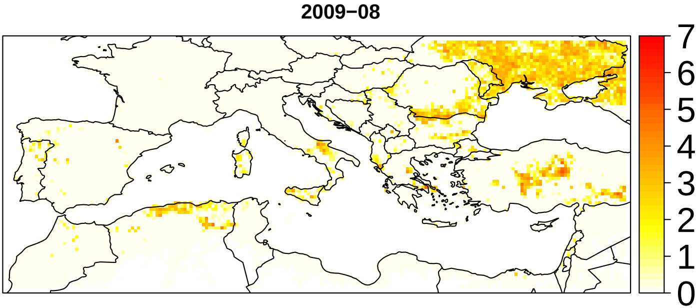


Figure: Map of $\log(1 + \mathrm{BA})$ for August 2009.

Parametric regression

The relationships between climate, fuel type/availability and wildfires is generally accepted to be very **complex**.

- [Richards and Huser, 2022] propose PINNs - Partially-interpretable Neural Networks for **conditional density estimation**
- Neural networks are "black box" in the sense that it's difficult/impossible to interpret their output - no good for understanding the drivers of risk
- The effect of some predictors is modelled using "**interpretable**" functions, whilst the rest feed a neural network - Here we use a convolutional neural network

Partially-interpretable neural networks

Let the response follow $\mathcal{F}(\boldsymbol{\theta}(\mathbf{x}))$ with parameter set $\boldsymbol{\theta}(\mathbf{x}) = (\theta_1(\mathbf{x}), \theta_2(\mathbf{x}), \dots)$. Then for all $i = 1, 2, \dots,$

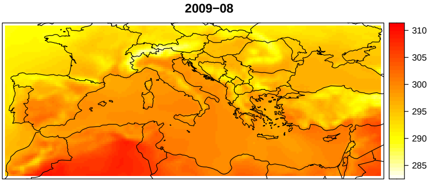
- Split predictor set $\mathbf{x}$ into two **complementary** subsets $\mathbf{x}^{(i)}_{\mathcal{I}}$ and $\mathbf{x}^{(i)}_{\mathcal{N}}$ - "interpreted" and "non-interpreted"
- Let

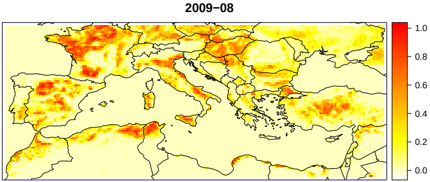
$$\theta_i(\mathbf{x}) = h_i[\eta^{(i)}_0 + m^{(i)}_{\mathcal{I}}(\mathbf{x}^{(i)}_{\mathcal{I}}) + m^{(i)}_{\mathcal{N}}(\mathbf{x}^{(i)}_{\mathcal{N}})],$$

  for constant intercept $\eta^{(i)}_0 \in \mathbb{R}$ and link $h_i : \mathbb{R} \to \mathbb{R}$
- Interpretable: $m^{(i)}_{\mathcal{I}}$, e.g., linear, **spline**. Neural network: $m^{(i)}_{\mathcal{N}}$.
- Functions and subsets **can differ** across components of $\boldsymbol{\theta}$
- Our framework applies for **any generic parametric distribution** $\mathcal{F}$, e.g., Bernoulli for occurrence, EV distributions for spread

Predictors

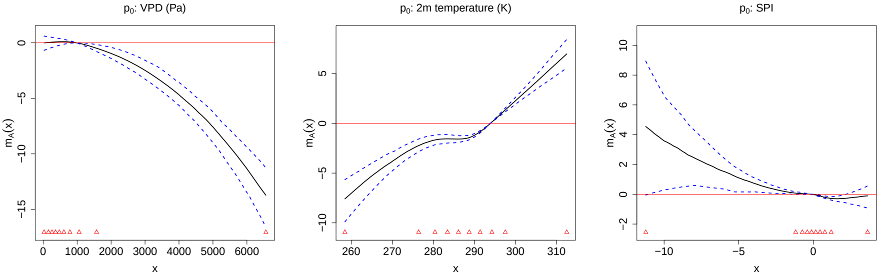
- 13 meteorological variables from **ERA-5 reanalysis**, e.g., **air temperature**, wind-speed components, evaporation, radiation
- Land cover maps (**COPERNICUS**) with proportion of grid-cell consisting of one of 21 types, e.g., tree species, urban areas, **cropland**
- Orographic: mean and s.d. altitude
- Left: 2m air temp (K). Right: cropland proportion. August 2009.
- We **interpret** the effect of VPD, 2m air temperature and 3-month SPI





Interpretable results

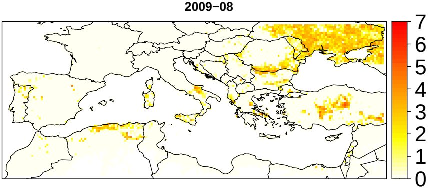
Effect of VPD, 2m temperature and 3-month SPI on *log*-odds of **occurrence probability**. Red triangles are knots, blue dashed lines are 95% confidence envelopes.

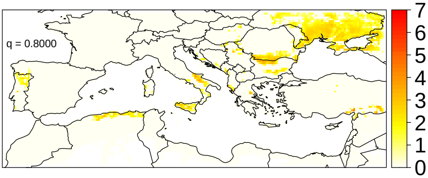


Estimated extreme quantile maps

Left: observed. Right: estimated q-quantile for $\log(1 + \text{BA})$ with $0.8 \leq q \leq 0.999$.

- Extreme quantiles are estimated using a **point process** model with three parameters
- **Location/scale** represented as functions of predictors
- Fixed **shape** parameter - $\hat{\xi} = 0.25\ (0.23, 0.27)$ - Much lighter-tailed than similar data for the U.S.

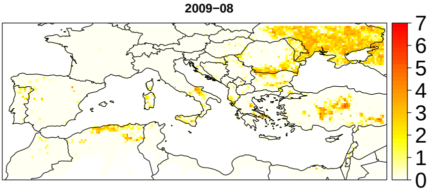


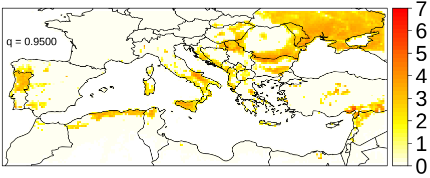


Estimated extreme quantile maps

Left: observed. Right: estimated 0.95-quantile for $\log(1 + \text{BA})$.

- Extreme quantiles are estimated using a **point process** model with three parameters
- **Location/scale** represented as functions of predictors
- Fixed **shape** parameter - $\hat{\xi} = 0.25\ (0.23, 0.27)$ - Much lighter-tailed than similar data for the U.S.





References

Richards, J. (2022).
pinnEV: Partially-Interpretable Neural Networks for modelling of Extreme Values.
R package. Will be made available at github.com/Jbrich95/pinnEV.

Richards, J. and Huser, R. (2022).
High-dimensional extreme quantile regression using partially-interpretable neural networks: With application to U.S. wildfires.
Pre-print. Not available online.

Both will be available alongside extended slides at my website `jbrich95.github.io` (via QR code).

SCAN ME