Environmental impact on short-term mortality

Jens Robben, Katrien Antonio, Torsten Kleinow

VSAE Actuarial Congress, Amsterdam - March 4, 2025

University of Amsterdam

RCLR Research Centre for Longevity Risk

Introduction

Background and motivation

- (Extreme) environmental events pose an increasing challenge to the field of mortality modelling.

Background and motivation

- (Extreme) environmental events pose an increasing challenge to the field of mortality modelling.
- Epidemiological studies have unveiled (short-term) associations between mortality and
  - temperature, e.g., Keatinge et al. [2000] and Basu and Samet [2002],
  - cold spells and heat waves, e.g., Braga et al. [2001] and Pattenden et al. [2003],
  - air pollution, e.g., Pascal et al. [2014] for PM10 and PM2.5 and Orellano et al. [2020] for ozone and nitrogen dioxide.

Background and motivation

- (Extreme) environmental events pose an increasing challenge to the field of mortality modelling.

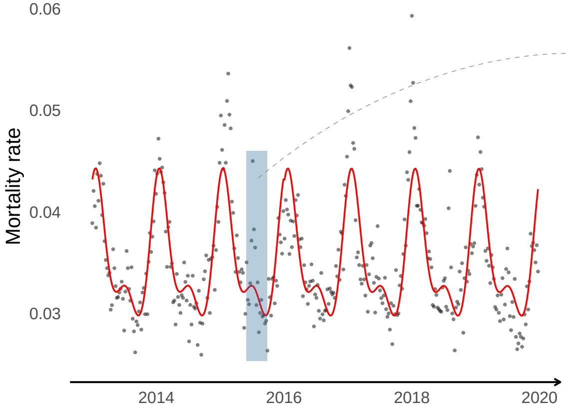
- Epidemiological studies have unveiled (short-term) associations between mortality and
  - temperature, e.g., Keatinge et al. [2000] and Basu and Samet [2002],
  - cold spells and heat waves, e.g., Braga et al. [2001] and Pattenden et al. [2003],
  - air pollution, e.g., Pascal et al. [2014] for PM10 and PM2.5 and Orellano et al. [2020] for ozone and nitrogen dioxide.

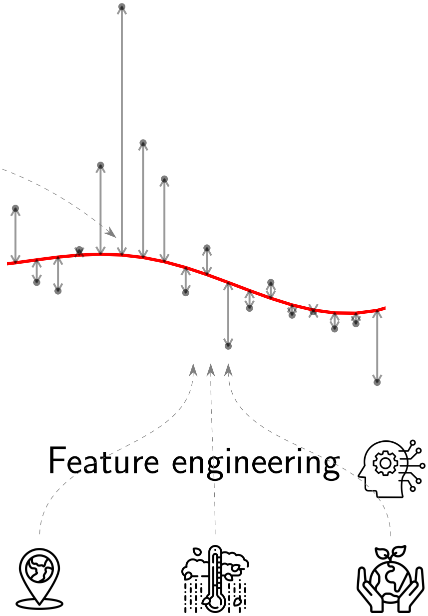
  Various methodologies have been proposed:
  - Poisson regression models, e.g., Armstrong [2006] and Braga et al. [2002],
  - Distributed Lag (Non-Linear) Models, e.g., Schwartz [2000] and Gasparrini et al. [2010],
  - Extreme value analysis, e.g., Li and Tang [2022].

Learning outcomes

In this session, we will:

- try to explain weekly death counts across European regions
- with a baseline mortality model (e.g., alike EuroMoMo)
- combined with a (high-dimensional) set of weather and air pollution features
- constructed from publicly available data sources (e.g., Eurostat, CDS, CAMS, NASA's EarthData).

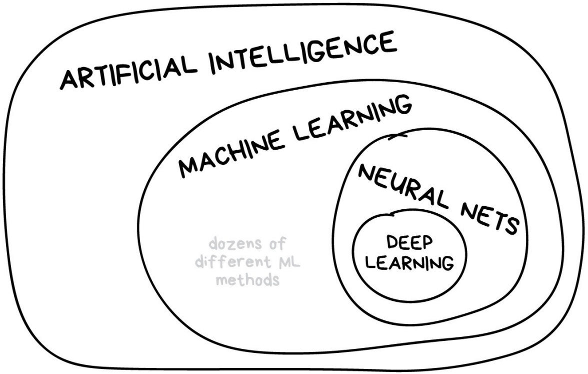

0.06
0.05
0.04
0.03
Mortality rate
2014
2016
2018
2020


Feature engineering

Machine learning and mortality modelling

We will make use of machine learning methods to find associations between mortality and environmental data:

- death counts $D_{x,t,w}^{(r)}$ under Poisson assumption, in the presence of risk factors or covariates $\boldsymbol{z}_{x,t,w}^{(r)}$
- with techniques such as:
  - Random Forests (RFs)
  - Gradient Boosting Machines (GBM, XGBoost, LightGBM, ...)
  - Neural Networks (CANNs, ANNs, RNNs, ...).



Picture taken from Machine learning for everyone. In simple words. With real-world examples. Yes, again.

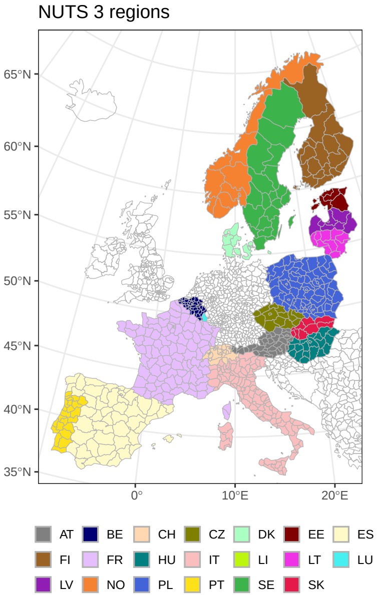
- Identify the primary environmental factors contributing to the estimation of mortality deviations from the baseline.
- Investigate the marginal impact of an environmental factor on deviations from the mortality baseline.
- Study how environmental factors interact when modelling mortality rates. Are there harvesting effects present?
- Demonstrate how to make short-term mortality projections with the model.

Data

Death counts

Eurostat: deaths by week, sex, 5-year age group and NUTS 3 region from 20 European countries throughout the years 2013-2019 ($>$ 500 regions).

Focus on old age group 65+.

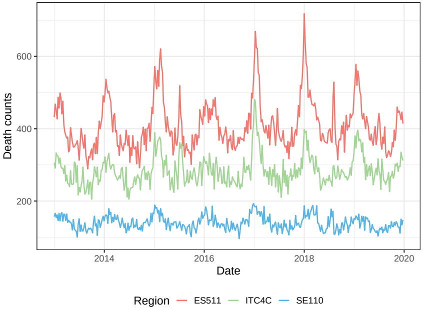


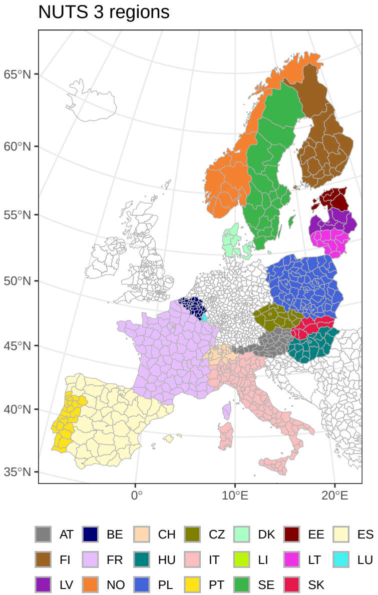
Death counts

Eurostat: deaths by week, sex, 5-year age group and NUTS 3 region from 20 European countries throughout the years 2013-2019 ($>$ 500 regions).

Focus on old age group 65+.

Seasonal trend:





Weather data

E-OBS land-only, gridded meteorological data for Europe from the Copernicus Climate Data Store.

Daily, high-resolution gridded dataset, defined on a grid with a spatial resolution of 0.10° (≈ 9 km).

Weather factors:

- `Tmax`: daily maximum temperature.
- `Tmin`: daily minimum temperature.
- `Hum`: daily average relative humidity.
- `Rain`: total daily precipitation.
- `Wind`: daily average wind speed.

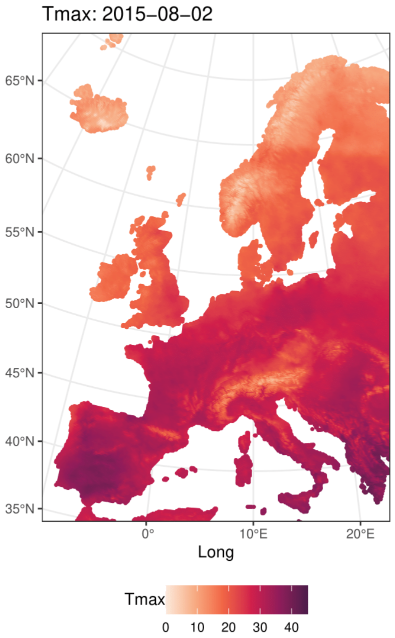
Weather data

E-OBS land-only, gridded meteorological data for Europe from the Copernicus Climate Data Store.

Daily, high-resolution gridded dataset, defined on a grid with a spatial resolution of 0.10° ($\approx$ 9 km).

Weather factors:

- `Tmax`: daily maximum temperature.
- `Tmin`: daily minimum temperature.
- `Hum`: daily average relative humidity.
- `Rain`: total daily precipitation.
- `Wind`: daily average wind speed.



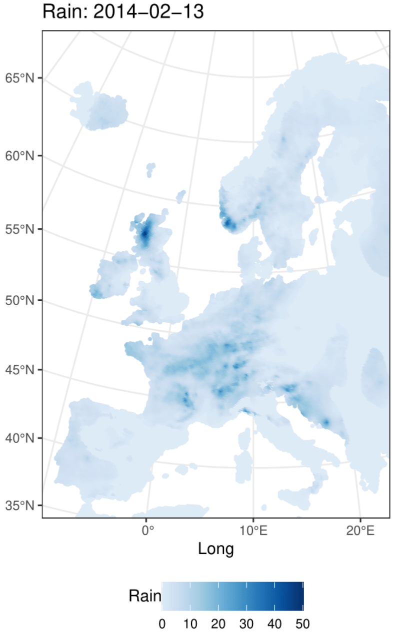
Weather data

E-OBS land-only, gridded meteorological data for Europe from the Copernicus Climate Data Store.

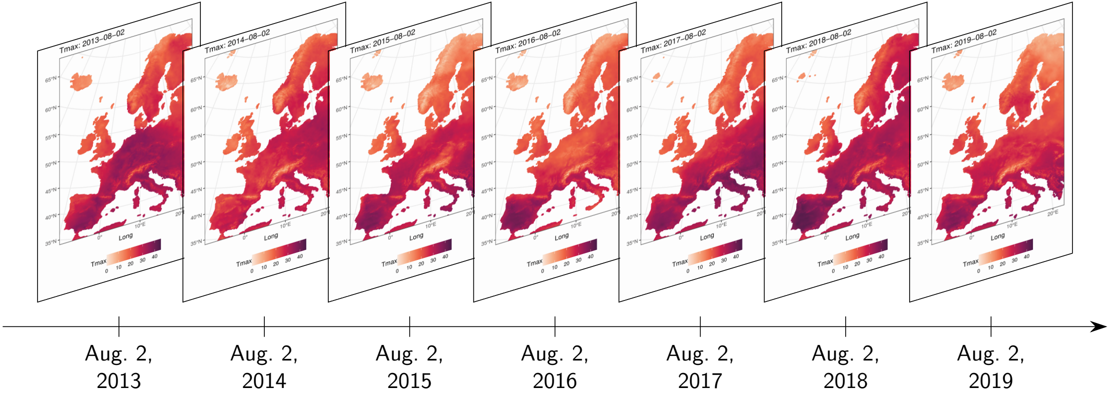
Daily, high-resolution gridded dataset, defined on a grid with a spatial resolution of 0.10° ($\approx$ 9 km).

Weather factors:

- `Tmax`: daily maximum temperature.
- `Tmin`: daily minimum temperature.
- `Hum`: daily average relative humidity.
- `Rain`: total daily precipitation.
- `Wind`: daily average wind speed.



Weather data



Air pollution data

CAMS European air quality reanalyses dataset from the Copernicus Atmosphere Monitoring Service (land + sea).

Hourly, high-resolution air quality reanalyses, defined on a grid with a spatial resolution of 0.10° (≈ 9 km).

Air pollutants ($\mu g/m^3$):

- `O3`: hourly ozone levels.
- `NO2`: hourly nitrogen dioxide levels.
- `PM10`: hourly particular matter (10 microns wide).
- `PM2.5`: hourly particular matter (2.5 microns wide).

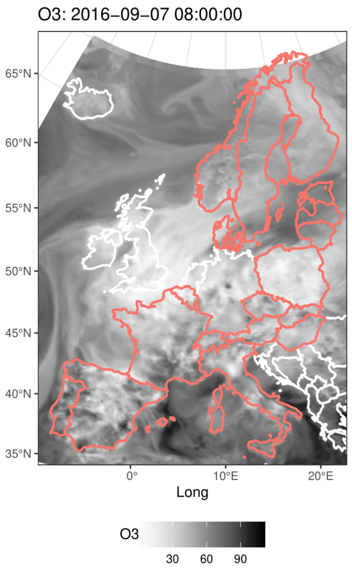
Air pollution data

CAMS European air quality reanalyses dataset from the Copernicus Atmosphere Monitoring Service (land + sea).

Hourly, high-resolution air quality reanalyses, defined on a grid with a spatial resolution of 0.10° (≈ 9 km).

Air pollutants ($\mu g/m^3$):

- `O3`: hourly ozone levels.
- `NO2`: hourly nitrogen dioxide levels.
- `PM10`: hourly particular matter (10 microns wide).
- `PM2.5`: hourly particular matter (2.5 microns wide).



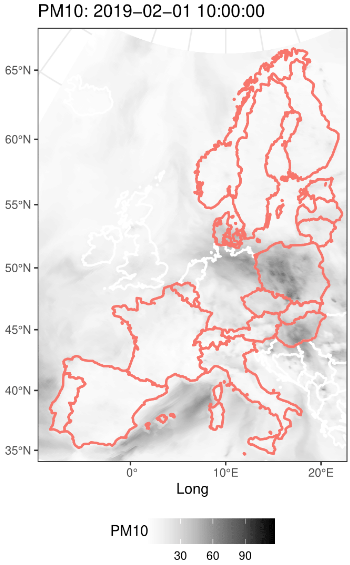
Air pollution data

CAMS European air quality reanalyses dataset from the Copernicus Atmosphere Monitoring Service (land + sea).

Hourly, high-resolution air quality reanalyses, defined on a grid with a spatial resolution of 0.10° (≈ 9 km).

Air pollutants ($\mu g/m^3$):

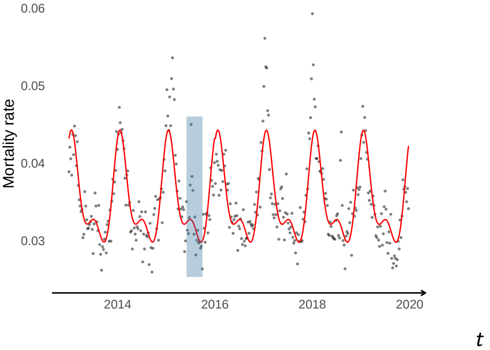
- `O3`: hourly ozone levels.
- `NO2`: hourly nitrogen dioxide levels.
- `PM10`: hourly particular matter (10 microns wide).
- `PM2.5`: hourly particular matter (2.5 microns wide).



Model specification

Weekly, region-specific baseline mortality model

A weekly, region-specific baseline mortality model to capture overall seasonal trends across all regions.



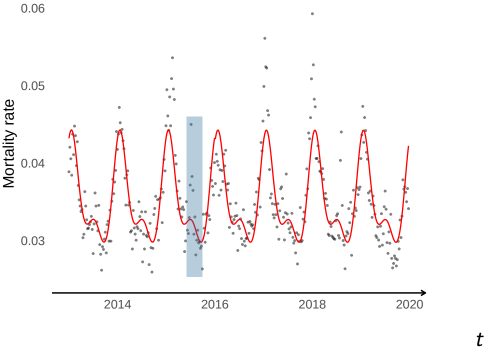
Weekly, region-specific baseline mortality model

A weekly, region-specific baseline mortality model to capture overall seasonal trends across all regions.

Incorporate seasonality through Fourier terms Serfling [1963]:

$$D_{t,w}^{(r)} \sim \text{Poisson}\left(E_{t,w}^{(r)} \cdot \mu_{t,w}^{(r)}\right),$$

$$\log \mu_{t,w}^{(r)} = \beta_0^{(r)} + \beta_1^{(r)} t + \beta_2^{(r)} \sin\left(\frac{2\pi w}{52}\right) + \beta_3^{(r)} \cos\left(\frac{2\pi w}{52}\right) + \beta_4^{(r)} \sin\left(\frac{2\pi w}{26}\right) + \beta_5^{(r)} \cos\left(\frac{2\pi w}{26}\right).$$



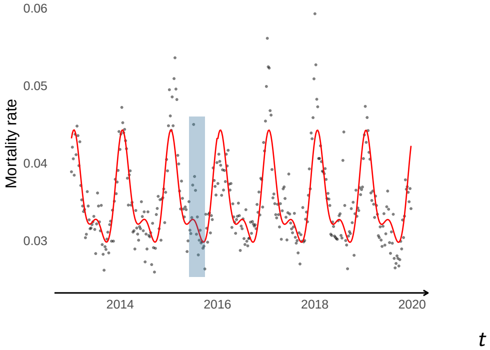
Weekly, region-specific baseline mortality model

A weekly, region-specific baseline mortality model to capture overall seasonal trends across all regions.

Incorporate seasonality through Fourier terms Serfling [1963]:

$$D_{t,w}^{(r)} \sim \text{Poisson}\left(E_{t,w}^{(r)} \cdot \mu_{t,w}^{(r)}\right),$$

$$\log \mu_{t,w}^{(r)} = \beta_0^{(r)} + \beta_1^{(r)} t + \beta_2^{(r)} \sin\left(\frac{2\pi w}{52}\right) + \beta_3^{(r)} \cos\left(\frac{2\pi w}{52}\right) + \beta_4^{(r)} \sin\left(\frac{2\pi w}{26}\right) + \beta_5^{(r)} \cos\left(\frac{2\pi w}{26}\right).$$



Region-specific population exposures $E_{t,w}^{(r)}$ from Eurostat.

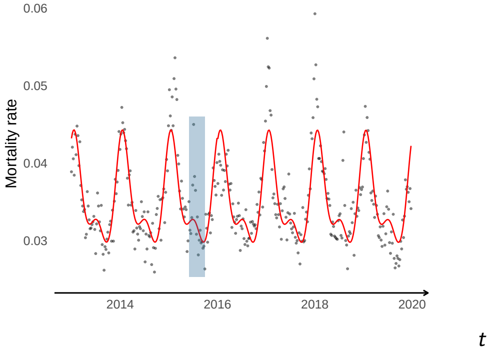
Weekly, region-specific baseline mortality model

A weekly, region-specific baseline mortality model to capture overall seasonal trends across all regions.

Incorporate seasonality through Fourier terms Serfling [1963]:

$$D_{t,w}^{(r)} \sim \text{Poisson}\left(E_{t,w}^{(r)} \cdot \mu_{t,w}^{(r)}\right),$$

$$\log \mu_{t,w}^{(r)} = \beta_0^{(r)} + \beta_1^{(r)} t + \beta_2^{(r)} \sin\left(\frac{2\pi w}{52}\right) + \beta_3^{(r)} \cos\left(\frac{2\pi w}{52}\right) + \beta_4^{(r)} \sin\left(\frac{2\pi w}{26}\right) + \beta_5^{(r)} \cos\left(\frac{2\pi w}{26}\right).$$

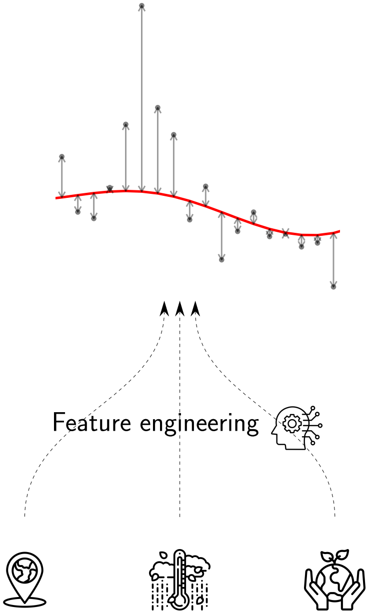


Region-specific population exposures $E_{t,w}^{(r)}$ from Eurostat.

Estimated baseline death counts: $\hat{b}_{t,w}^{(r)} := E_{t,w}^{(r)} \cdot \hat{\mu}_{t,w}^{(r)}$.

Modelling deviations from the baseline model

Explain observed deviations from the baseline deaths using region-specific environmental features.



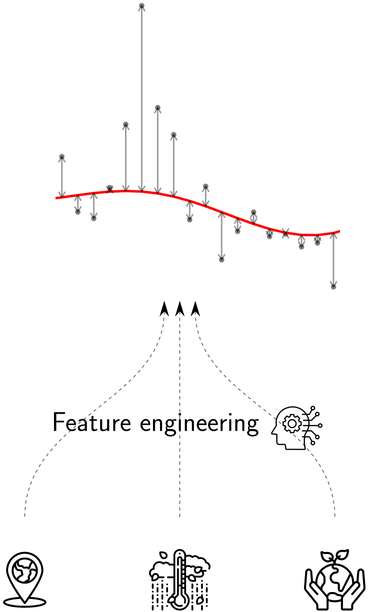
Modelling deviations from the baseline model

Explain observed deviations from the baseline deaths using region-specific environmental features.

Fix estimated baseline deaths and impose distributional assumption:

$$D_{t,w}^{(r)} \sim \text{Poisson}\left(\hat{b}_{t,w}^{(r)}\, \phi_{t,w}^{(r)}\right),$$

$$\phi_{t,w}^{(r)} = f\Big(\texttt{long}^{(r)}, \texttt{lat}^{(r)}, \texttt{season}_{t,w}, \boldsymbol{c}_{t,w}^{(r)}, \boldsymbol{e}_{t,w}^{(r)}, l^{1}\left(\boldsymbol{c}_{t,w}^{(r)}\right), l^{1}\left(\boldsymbol{e}_{t,w}^{(r)}\right), \ldots, l^{s}\left(\boldsymbol{c}_{t,w}^{(r)}\right), l^{s}\left(\boldsymbol{e}_{t,w}^{(r)}\right)\Big).$$



Modelling deviations from the baseline model

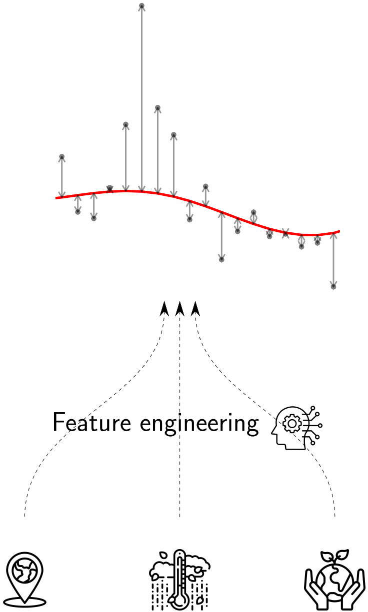
Explain observed deviations from the baseline deaths using region-specific environmental features.

Fix estimated baseline deaths and impose distributional assumption:

$$D_{t,w}^{(r)} \sim \text{Poisson}\left(\hat{b}_{t,w}^{(r)}\,\phi_{t,w}^{(r)}\right),$$

$$\phi_{t,w}^{(r)} = f\Big(\texttt{long}^{(r)},\, \texttt{lat}^{(r)},\, \texttt{season}_{t,w},\, \boldsymbol{c}_{t,w}^{(r)},\, \boldsymbol{e}_{t,w}^{(r)},\; l^{1}\left(\boldsymbol{c}_{t,w}^{(r)}\right),\, l^{1}\left(\boldsymbol{e}_{t,w}^{(r)}\right),\, \ldots,\, l^{s}\left(\boldsymbol{c}_{t,w}^{(r)}\right),\, l^{s}\left(\boldsymbol{e}_{t,w}^{(r)}\right)\Big).$$

$f(\cdot)$ is a selected predictive modelling technique.



Modelling deviations from the baseline model

Explain observed deviations from the baseline deaths using region-specific environmental features.

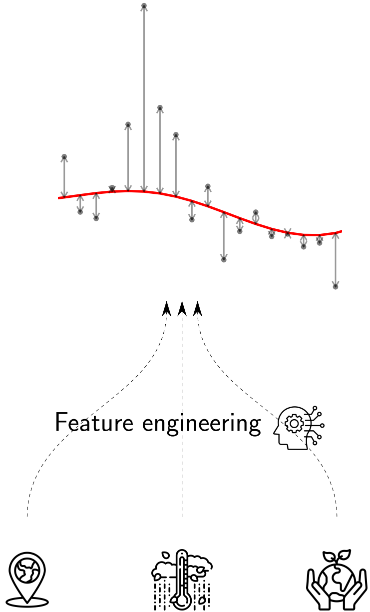
Fix estimated baseline deaths and impose distributional assumption:

$$D_{t,w}^{(r)} \sim \text{Poisson}\left(\hat{b}_{t,w}^{(r)}\, \phi_{t,w}^{(r)}\right),$$

$$\phi_{t,w}^{(r)} = f\Big(\texttt{long}^{(r)},\ \texttt{lat}^{(r)},\ \texttt{season}_{t,w},\ \boldsymbol{c}_{t,w}^{(r)},\ \boldsymbol{e}_{t,w}^{(r)},$$
$$l^1\left(\boldsymbol{c}_{t,w}^{(r)}\right),\ l^1\left(\boldsymbol{e}_{t,w}^{(r)}\right),\ \ldots,\ l^s\left(\boldsymbol{c}_{t,w}^{(r)}\right),\ l^s\left(\boldsymbol{e}_{t,w}^{(r)}\right)\Big).$$

$f(\cdot)$ is a selected predictive modelling technique.

Choice for machine learning model to identify non-linear relationships and potential interaction effects among environmental features.



Model calibration

Calibrating the baseline model

Fit one Poisson GLM jointly on all regions, and add a penalty term to obtain smooth variations in the estimated parameters $\hat{\beta}_p^{(r)}$ across neighbouring regions:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \left(-l_P(\boldsymbol{\beta}) + \sum_{p=0}^{5} \lambda_p \boldsymbol{\beta}_p^T \boldsymbol{S} \boldsymbol{\beta}_p \right),$$

Calibrating the baseline model

Fit one Poisson GLM jointly on all regions, and add a penalty term to obtain smooth variations in the estimated parameters $\hat{\beta}_p^{(r)}$ across neighbouring regions:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \left(-l_P(\boldsymbol{\beta}) + \sum_{p=0}^{5} \lambda_p \boldsymbol{\beta}_p^T \boldsymbol{S} \boldsymbol{\beta}_p \right),$$

where:

- $\boldsymbol{\beta}$: parameter vector,
- $l_P(\boldsymbol{\beta})$: Poisson log-likelihood,
- $\boldsymbol{\beta}_p^T \boldsymbol{S} \boldsymbol{\beta}_p$: quadratic penalty term,
- λ_p: smoothing or penalty parameter.

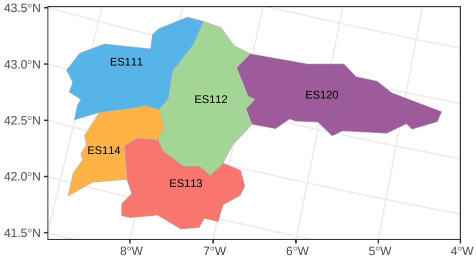
Fit one Poisson GLM jointly on all regions, and add a penalty term to obtain smooth variations in the estimated parameters $\hat{\beta}_p^{(r)}$ across neighbouring regions:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \left(-l_P(\boldsymbol{\beta}) + \sum_{p=0}^{5} \lambda_p \boldsymbol{\beta}_p^T \boldsymbol{S} \boldsymbol{\beta}_p \right),$$

where:

- $\boldsymbol{\beta}$: parameter vector,
- $l_P(\boldsymbol{\beta})$: Poisson log-likelihood,
- $\boldsymbol{\beta}_p^T \boldsymbol{S} \boldsymbol{\beta}_p$: quadratic penalty term,
- λ_p: smoothing or penalty parameter.

Example (5 Spanish NUTS 3 regions):



Calibrating the baseline model

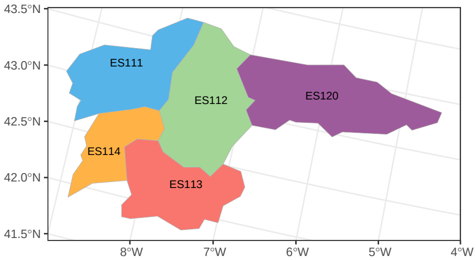
Fit one Poisson GLM jointly on all regions, and add a penalty term to obtain smooth variations in the estimated parameters $\hat{\beta}_p^{(r)}$ across neighbouring regions:

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \left(-l_P(\boldsymbol{\beta}) + \sum_{p=0}^{5} \lambda_p \boldsymbol{\beta}_p^T \boldsymbol{S} \boldsymbol{\beta}_p \right),$$

where:

- $\boldsymbol{\beta}$: parameter vector,
- $l_P(\boldsymbol{\beta})$: Poisson log-likelihood,
- $\boldsymbol{\beta}_p^T \boldsymbol{S} \boldsymbol{\beta}_p$: quadratic penalty term,
- λ_p: smoothing or penalty parameter.

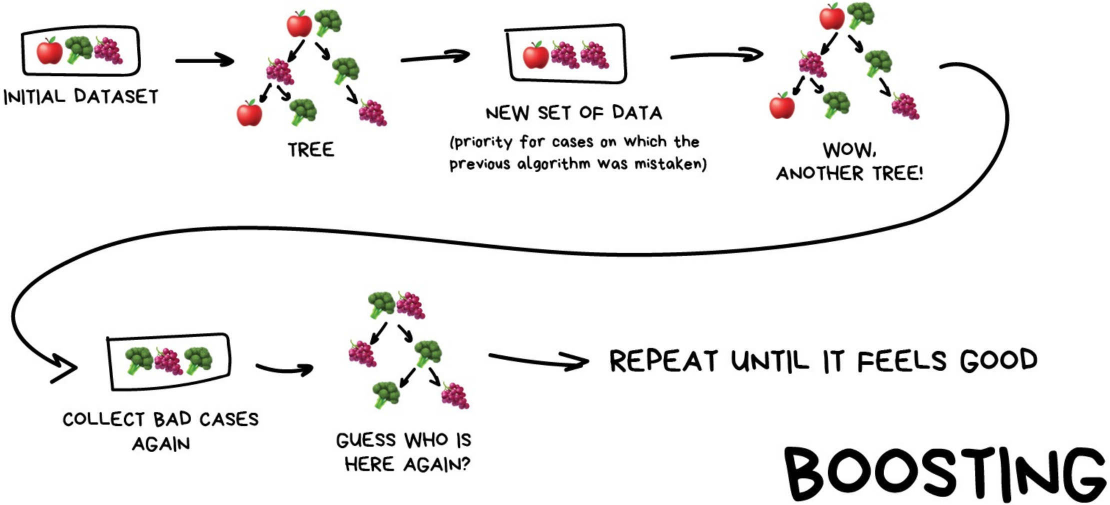
Example (5 Spanish NUTS 3 regions):



Penalty matrix $\boldsymbol{S}$:

$$\begin{array}{c} \\ ES111 \\ ES112 \\ ES113 \\ ES114 \\ ES120 \end{array} \begin{array}{c} \begin{array}{ccccc} ES111 & ES112 & ES113 & ES114 & ES120 \end{array} \\ \begin{pmatrix} 2 & -1 & 0 & -1 & 0 \\ -1 & 4 & -1 & -1 & -1 \\ 0 & -1 & 2 & -1 & 0 \\ -1 & -1 & -1 & 3 & 0 \\ 0 & -1 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

Calibrating the mortality deviations via gradient boosting



Picture taken from Machine learning for everyone. In simple words. With real-world examples. Yes, again.

Parameter configurations

XGBoost: flexible and efficient implementation of gradient boosting.

Parameter configurations

XGBoost: flexible and efficient implementation of gradient boosting.

Tuning parameters:

`nrounds`:	number of boosting iterations.
`eta`:	learning rate.
`max_depth`:	the maximum depth of a tree.
`subsample`:	subsample ratio of the training data.
`colsample_bytree`:	subsample ratio of the features.

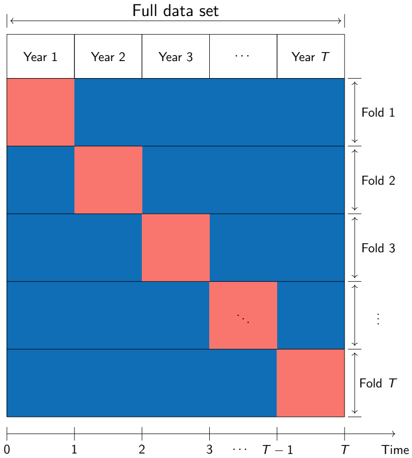
Parameter configurations

XGBoost: flexible and efficient implementation of gradient boosting.

Tuning parameters:

`nrounds`:	number of boosting iterations.
`eta`:	learning rate.
`max_depth`:	the maximum depth of a tree.
`subsample`:	subsample ratio of the training data.
`colsample_bytree`:	subsample ratio of the features.

Parameter tuning with T-fold cross-validation.



Parameter configurations

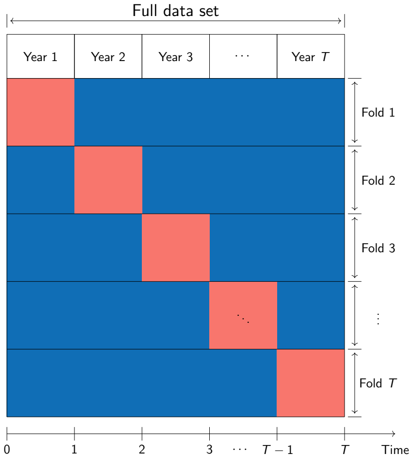
XGBoost: flexible and efficient implementation of gradient boosting.

Tuning parameters:

`nrounds`:	number of boosting iterations.
`eta`:	learning rate.
`max_depth`:	the maximum depth of a tree.
`subsample`:	subsample ratio of the training data.
`colsample_bytree`:	subsample ratio of the features.

Parameter tuning with T-fold cross-validation.

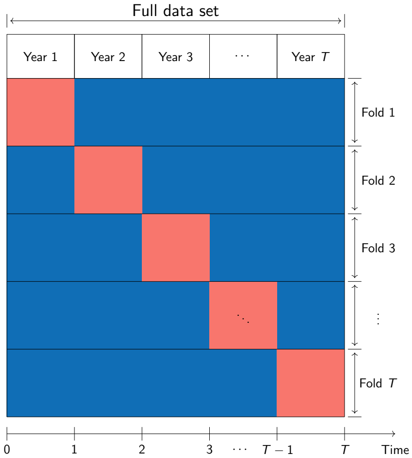
Calibrate XGBoost model on entire training data with optimal parameter configuration.



XGBoost: flexible and efficient implementation of gradient boosting.

Tuning parameters:

`nrounds`:	number of boosting iterations.
`eta`:	learning rate.
`max_depth`:	the maximum depth of a tree.
`subsample`:	subsample ratio of the training data.
`colsample_bytree`:	subsample ratio of the features.



Parameter tuning with T-fold cross-validation.

Calibrate XGBoost model on entire training data with optimal parameter configuration.

Interpretation tools to gain insights: VIP, ALE.

Case study: feature engineering

Feature engineering

Motivation

Difference in spatial and temporal dimension:

- deaths data: weekly, NUTS 3 scale.
- environmental data: hourly or daily time scale, spatial grid.

Feature engineering

Motivation

Difference in spatial and temporal dimension:

- deaths data: weekly, NUTS 3 scale.
- environmental data: hourly or daily time scale, spatial grid.

Goal of feature engineering:

- convert the temporal and spatial dimensions of the environmental data into aggregated features on a weekly, NUTS 3 scale.

Feature engineering

Motivation

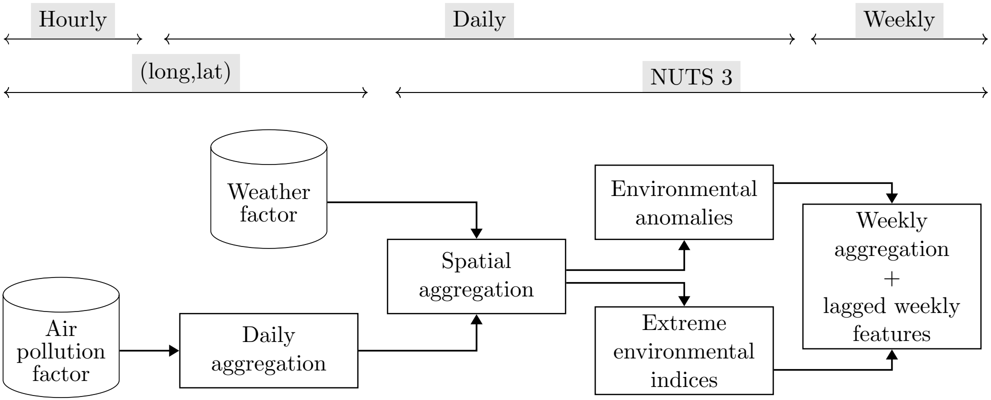
Difference in spatial and temporal dimension:

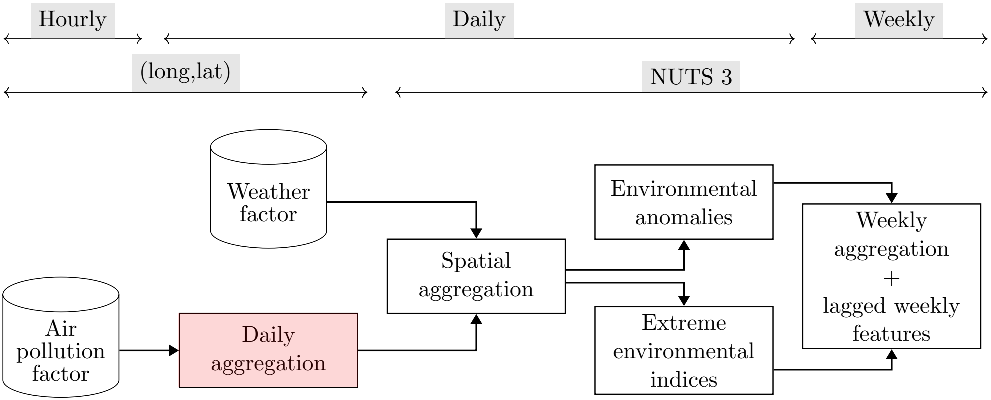
- deaths data: weekly, NUTS 3 scale.
- environmental data: hourly or daily time scale, spatial grid.

Goal of feature engineering:

- convert the temporal and spatial dimensions of the environmental data into aggregated features on a weekly, NUTS 3 scale.
- create features that measure deviations from baseline conditions from environmental data to explain excess or deficit mortality.

Flow chart


Hourly
Daily
Weekly
(long,lat)
NUTS 3
Weather factor
Air pollution factor
Daily aggregation
Spatial aggregation
Environmental anomalies
Extreme environmental indices
Weekly aggregation + lagged weekly features


Hourly
Daily
Weekly
(long,lat)
NUTS 3
Weather factor
Air pollution factor
Daily aggregation
Spatial aggregation
Environmental anomalies
Extreme environmental indices
Weekly aggregation + lagged weekly features

Daily aggregation

Consider an air pollution factor and denote its concentration at hour h of day d in week w of year t and located at longitude-latitude coordinates (long,lat) as $x_{t,w,d,h}^{(\text{long,lat})}$.

Daily aggregation

Consider an air pollution factor and denote its concentration at hour h of day d in week w of year t and located at longitude-latitude coordinates (long,lat) as $x_{t,w,d,h}^{(\text{long,lat})}$.

Compute the daily minimum, average, and maximum concentrations of the air pollutant, measured at the coordinates (long,lat) as:

$$\hat{x}{}_{t,w,d}^{(\text{long,lat})} = \min\left\{x_{t,w,d,h}^{(\text{long,lat})} \mid h = 0, 1, ..., 23\right\}$$

$$\overline{x}_{t,w,d}^{(\text{long,lat})} = \text{avg}\left\{x_{t,w,d,h}^{(\text{long,lat})} \mid h = 0, 1, ..., 23\right\}$$

$$\overset{\vee}{x}{}_{t,w,d}^{(\text{long,lat})} = \max\left\{x_{t,w,d,h}^{(\text{long,lat})} \mid h = 0, 1, ..., 23\right\}.$$

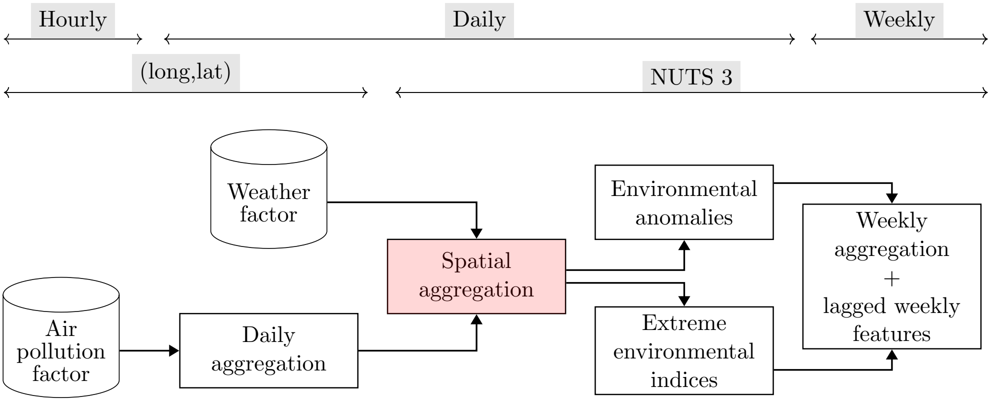
Consider an air pollution factor and denote its concentration at hour h of day d in week w of year t and located at longitude-latitude coordinates (long,lat) as $x_{t,w,d,h}^{(\text{long,lat})}$.

Compute the daily minimum, average, and maximum concentrations of the air pollutant, measured at the coordinates (long,lat) as:

$$\begin{aligned}
\hat{x}{}_{t,w,d}^{(\text{long,lat})} &= \min\left\{x_{t,w,d,h}^{(\text{long,lat})} \mid h = 0, 1, ..., 23\right\} \\
\overline{x}_{t,w,d}^{(\text{long,lat})} &= \operatorname{avg}\left\{x_{t,w,d,h}^{(\text{long,lat})} \mid h = 0, 1, ..., 23\right\} \\
\overset{\vee}{x}{}_{t,w,d}^{(\text{long,lat})} &= \max\left\{x_{t,w,d,h}^{(\text{long,lat})} \mid h = 0, 1, ..., 23\right\}.
\end{aligned}$$

Weather factors already available at the daily level (no need for daily aggregation).

Flow chart



Hourly
Daily
Weekly
(long,lat)
NUTS 3
Weather factor
Air pollution factor
Daily aggregation
Spatial aggregation
Environmental anomalies
Extreme environmental indices
Weekly aggregation + lagged weekly features

Spatial aggregation

$\tilde{x}_{t,w,d}^{(\text{long,lat})}$: daily level of a specific environmental feature at coordinates (long, lat) for year t, week w, and day d.

Spatial aggregation

$\tilde{x}^{(\text{long,lat})}_{t,w,d}$: daily level of a specific environmental feature at coordinates (long, lat) for year t, week w, and day d.

Construct feature on NUTS 3 scale:

$$\tilde{x}^{(r)}_{t,w,d} = \sum_{(\text{long,lat})\in\mathcal{I}_1(r)} \omega_{(\text{long,lat})} \cdot \tilde{x}^{(\text{long,lat})}_{t,w,d},$$

Spatial aggregation

$\tilde{x}_{t,w,d}^{(\text{long},\text{lat})}$: daily level of a specific environmental feature at coordinates (long, lat) for year t, week w, and day d.

Construct feature on NUTS 3 scale:

$$\tilde{x}_{t,w,d}^{(r)} = \sum_{(\text{long},\text{lat}) \in \mathcal{I}_1(r)} \omega_{(\text{long},\text{lat})} \cdot \tilde{x}_{t,w,d}^{(\text{long},\text{lat})},$$

where:

- $\omega_{(\text{long},\text{lat})}$: population weights using gridded population data from the Socioeconomic Data and Applications Center,

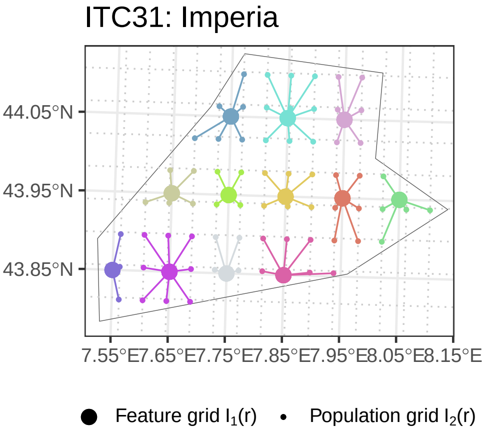
$\tilde{x}_{t,w,d}^{(\text{long},\text{lat})}$: daily level of a specific environmental feature at coordinates (long, lat) for year t, week w, and day d.

Construct feature on NUTS 3 scale:

$$\tilde{x}_{t,w,d}^{(r)} = \sum_{(\text{long},\text{lat}) \in \mathcal{I}_1(r)} \omega_{(\text{long},\text{lat})} \cdot \tilde{x}_{t,w,d}^{(\text{long},\text{lat})},$$

where:

- $\omega_{(\text{long},\text{lat})}$: population weights using gridded population data from the Socioeconomic Data and Applications Center,



Spatial aggregation

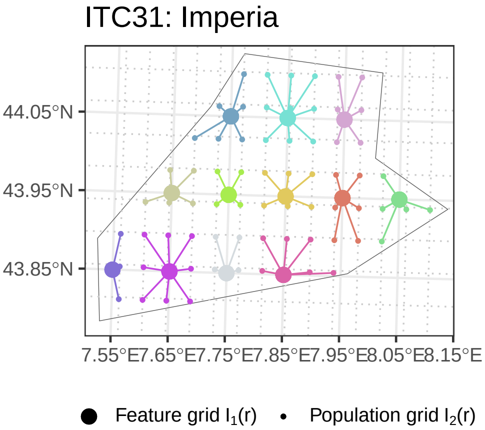
$\tilde{x}_{t,w,d}^{(\text{long,lat})}$: daily level of a specific environmental feature at coordinates (long, lat) for year t, week w, and day d.

Construct feature on NUTS 3 scale:

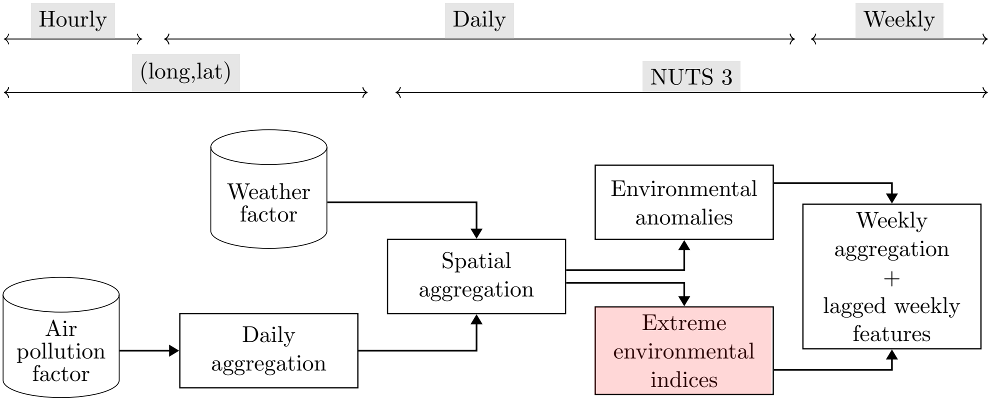
$$\tilde{x}_{t,w,d}^{(r)} = \sum_{(\text{long,lat}) \in \mathcal{I}_1(r)} \omega_{(\text{long,lat})} \cdot \tilde{x}_{t,w,d}^{(\text{long,lat})},$$

where:

- $\omega_{(\text{long,lat})}$: population weights using gridded population data from the Socioeconomic Data and Applications Center,
- $\mathcal{I}_1(r)$: feature grid restricted to region r.



Flow chart



Hourly
Daily
Weekly
(long,lat)
NUTS 3
Weather factor
Air pollution factor
Daily aggregation
Spatial aggregation
Environmental anomalies
Extreme environmental indices
Weekly aggregation + lagged weekly features

Extreme environmental indices

Aim: to capture the effects of extreme environmental conditions on mortality baseline deviations.

Extreme environmental indices

Aim: to capture the effects of extreme environmental conditions on mortality baseline deviations.

Calculate region-specific 5% and 95% quantiles of the daily historical temperature or air pollution observations over the years 2013-2019.

Extreme environmental indices

Aim: to capture the effects of extreme environmental conditions on mortality baseline deviations.

Calculate region-specific 5% and 95% quantiles of the daily historical temperature or air pollution observations over the years 2013-2019.

Define extreme high temperature index (hot-day index):

$$\text{T.ind}_{t,w,d}^{(r,95\%)} = \mathbb{1}\left\{\texttt{Tmax}_{t,w,d}^{(r)} \geq q_{\texttt{Tmax}}^{(r,95\%)}\right\} + \mathbb{1}\left\{\texttt{Tavg}_{t,w,d}^{(r)} \geq q_{\texttt{Tavg}}^{(r,95\%)}\right\} + \mathbb{1}\left\{\texttt{Tmin}_{t,w,d}^{(r)} \geq q_{\texttt{Tmin}}^{(r,95\%)}\right\}.$$

Extreme environmental indices

Aim: to capture the effects of extreme environmental conditions on mortality baseline deviations.

Calculate region-specific 5% and 95% quantiles of the daily historical temperature or air pollution observations over the years 2013-2019.

Define extreme high temperature index (hot-day index):

$$\text{T.ind}_{t,w,d}^{(r,95\%)} = \mathbb{1}\left\{\texttt{Tmax}_{t,w,d}^{(r)} \geq q_{\texttt{Tmax}}^{(r,95\%)}\right\} + \mathbb{1}\left\{\texttt{Tavg}_{t,w,d}^{(r)} \geq q_{\texttt{Tavg}}^{(r,95\%)}\right\} + \mathbb{1}\left\{\texttt{Tmin}_{t,w,d}^{(r)} \geq q_{\texttt{Tmin}}^{(r,95\%)}\right\}.$$

Index values: 0-3, indicating the severity of hot days.

Aim: to capture the effects of extreme environmental conditions on mortality baseline deviations.

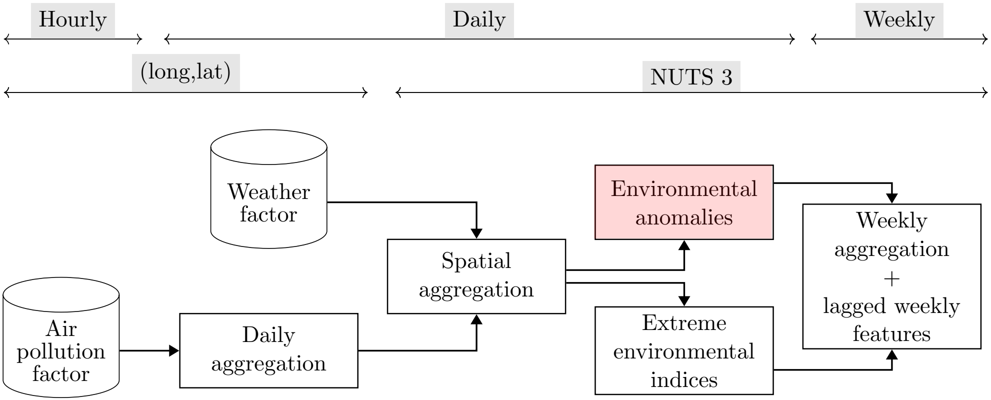
Calculate region-specific 5% and 95% quantiles of the daily historical temperature or air pollution observations over the years 2013-2019.

Define extreme high temperature index (hot-day index):

$$\text{T.ind}_{t,w,d}^{(r,95\%)} = \mathbb{1}\left\{\texttt{Tmax}_{t,w,d}^{(r)} \geq q_{\texttt{Tmax}}^{(r,95\%)}\right\} + \mathbb{1}\left\{\texttt{Tavg}_{t,w,d}^{(r)} \geq q_{\texttt{Tavg}}^{(r,95\%)}\right\} + \mathbb{1}\left\{\texttt{Tmin}_{t,w,d}^{(r)} \geq q_{\texttt{Tmin}}^{(r,95\%)}\right\}.$$

Index values: 0-3, indicating the severity of hot days.

Similar extreme indices are created for the remaining daily weather and air pollution factors.



Hourly
Daily
Weekly
(long,lat)
NUTS 3
Weather factor
Air pollution factor
Daily aggregation
Spatial aggregation
Environmental anomalies
Extreme environmental indices
Weekly aggregation + lagged weekly features

Environmental anomalies

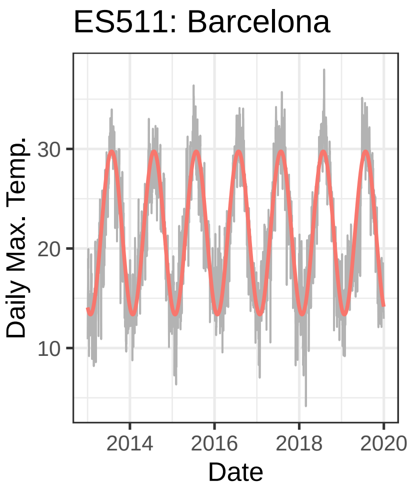
Environmental anomalies

Create features that quantify deviations from typical, baseline conditions for each day throughout the year.

Create features that quantify deviations from typical, baseline conditions for each day throughout the year.

Robust linear regression to capture baseline:

$$\tilde{x}_{t,w,d}^{(r)} = \alpha_0^{(r)} + \alpha_1^{(r)} \sin\left(\frac{2\pi w}{365.25}\right) + \alpha_2^{(r)} \cos\left(\frac{2\pi w}{365.25}\right) + \epsilon_{t,w,d}^{(r)},$$



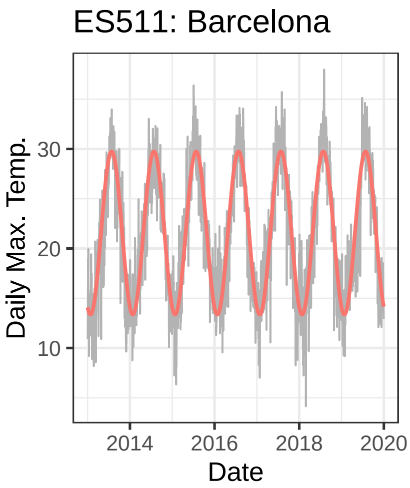
Create features that quantify deviations from typical, baseline conditions for each day throughout the year.

Robust linear regression to capture baseline:

$$\tilde{x}^{(r)}_{t,w,d} = \alpha^{(r)}_0 + \alpha^{(r)}_1 \sin\left(\frac{2\pi w}{365.25}\right) + \alpha^{(r)}_2 \cos\left(\frac{2\pi w}{365.25}\right) + \epsilon^{(r)}_{t,w,d},$$

In the paper, we work with excesses or deviations from the baseline (anomalies):

$$\tilde{x}^{(r)}_{t,w,d} - \hat{\tilde{x}}^{(r)}_{t,w,d}$$



Environmental anomalies

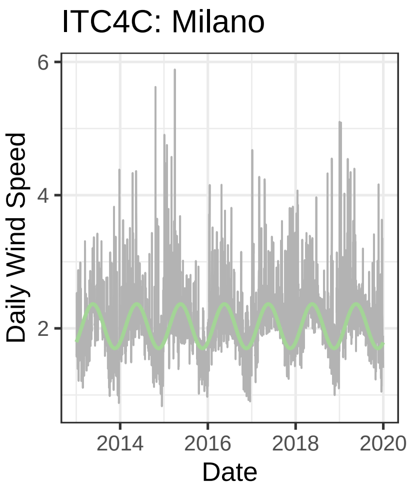
Create features that quantify deviations from typical, baseline conditions for each day throughout the year.

Robust linear regression to capture baseline:

$$\tilde{x}^{(r)}_{t,w,d} = \alpha^{(r)}_0 + \alpha^{(r)}_1 \sin\left(\frac{2\pi w}{365.25}\right) + \alpha^{(r)}_2 \cos\left(\frac{2\pi w}{365.25}\right) + \epsilon^{(r)}_{t,w,d},$$

In the paper, we work with excesses or deviations from the baseline (anomalies):

$$\tilde{x}^{(r)}_{t,w,d} - \hat{\tilde{x}}^{(r)}_{t,w,d}$$



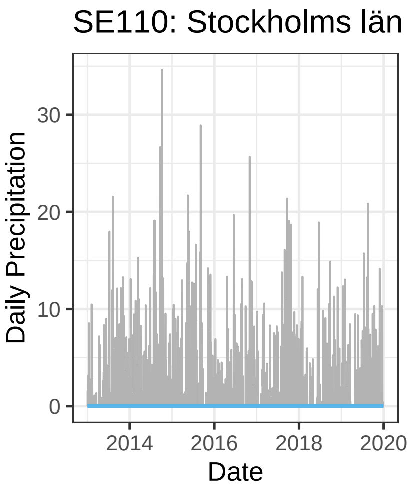
Create features that quantify deviations from typical, baseline conditions for each day throughout the year.

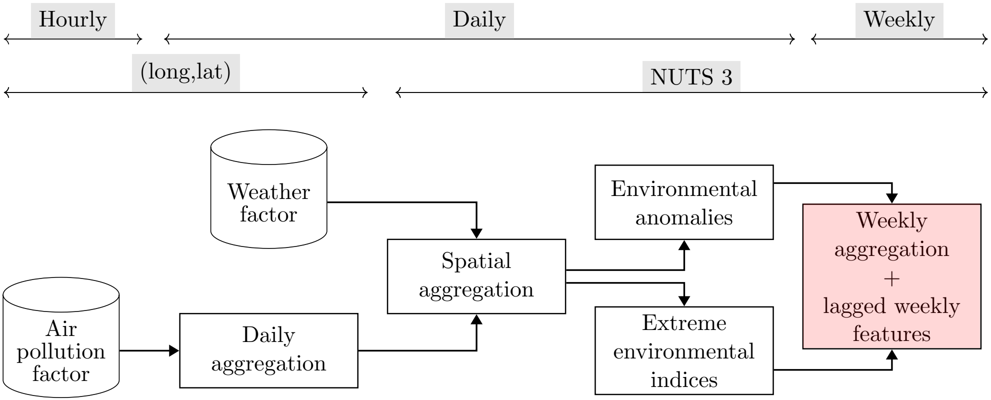
Robust linear regression to capture baseline:

$$\tilde{x}^{(r)}_{t,w,d} = \alpha^{(r)}_0 + \alpha^{(r)}_1 \sin\left(\frac{2\pi w}{365.25}\right) + \alpha^{(r)}_2 \cos\left(\frac{2\pi w}{365.25}\right) + \epsilon^{(r)}_{t,w,d},$$

In the paper, we work with excesses or deviations from the baseline (anomalies):

$$\tilde{x}^{(r)}_{t,w,d} - \hat{\tilde{x}}^{(r)}_{t,w,d}$$





Hourly
Daily
Weekly
(long,lat)
NUTS 3
Weather factor
Air pollution factor
Daily aggregation
Spatial aggregation
Environmental anomalies
Extreme environmental indices
Weekly aggregation + lagged weekly features

Weekly aggregation

Up to now: feature anomalies on daily time scale.

Up to now: feature anomalies on daily time scale.

Various weekly aggregation techniques for each region,
e.g., for temperature:

Up to now: feature anomalies on daily time scale.

Various weekly aggregation techniques for each region, e.g., for temperature:

- weekly average of daily minimum/maximum temperature anomalies,

Weekly aggregation

Up to now: feature anomalies on daily time scale.

Various weekly aggregation techniques for each region, e.g., for temperature:

- weekly average of daily minimum/maximum temperature anomalies,
- weekly average of daily hot-day index.

Up to now: feature anomalies on daily time scale.

Various weekly aggregation techniques for each region, e.g., for temperature:

- weekly average of daily minimum/maximum temperature anomalies,
- weekly average of daily hot-day index.

Similar weekly aggregation techniques for remaining environmental anomalies and extreme environmental indices.

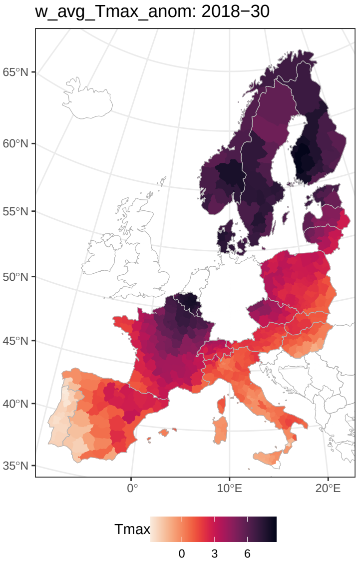
Weekly aggregation

Up to now: feature anomalies on daily time scale.

Various weekly aggregation techniques for each region, e.g., for temperature:

- weekly average of daily minimum/maximum temperature anomalies,
- weekly average of daily hot-day index.

Similar weekly aggregation techniques for remaining environmental anomalies and extreme environmental indices.



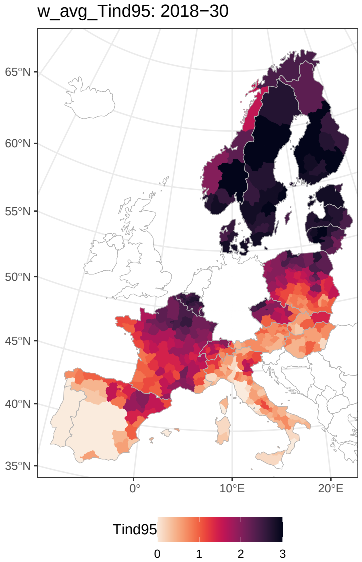
Weekly aggregation

Up to now: feature anomalies on daily time scale.

Various weekly aggregation techniques for each region, e.g., for temperature:

- weekly average of daily minimum/maximum temperature anomalies,
- weekly average of daily hot-day index.

Similar weekly aggregation techniques for remaining environmental anomalies and extreme environmental indices.



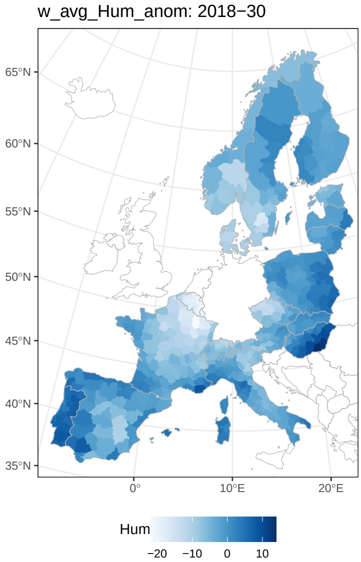
Weekly aggregation

Up to now: feature anomalies on daily time scale.

Various weekly aggregation techniques for each region, e.g., for temperature:

- weekly average of daily minimum/maximum temperature anomalies,
- weekly average of daily hot-day index.

Similar weekly aggregation techniques for remaining environmental anomalies and extreme environmental indices.



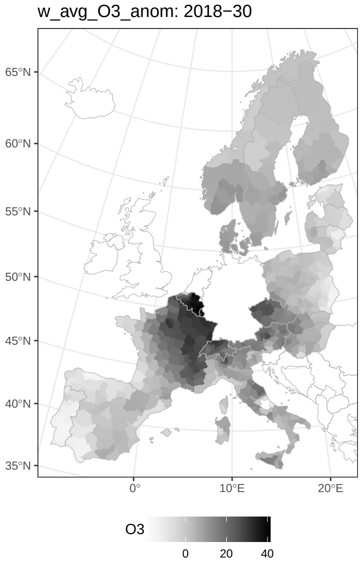
Weekly aggregation

Up to now: feature anomalies on daily time scale.

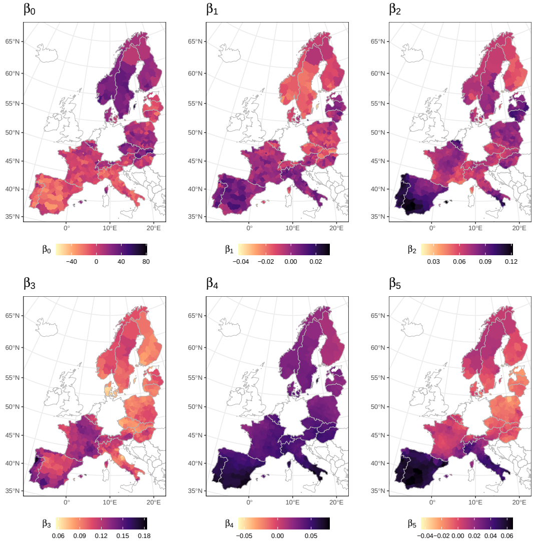
Various weekly aggregation techniques for each region, e.g., for temperature:

- weekly average of daily minimum/maximum temperature anomalies,
- weekly average of daily hot-day index.

Similar weekly aggregation techniques for remaining environmental anomalies and extreme environmental indices.



Case study: calibration results



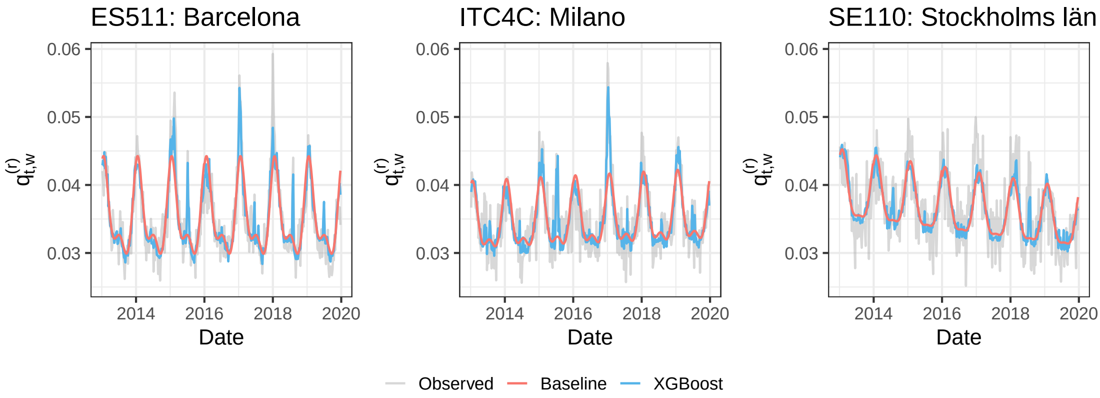
Input features: longitude-latitude coordinates, season, (one-week lagged) environmental anomalies and extreme indices.

Tuning by 7-fold cross validation over the years 2013-2019 using an extensive tuning grid.

Tuning parameters: `nrounds` (490), `eta` (0.01), `min_child_weight` (1000), `max.depth` (7), `subsample` (0.75), `colsample_bytree` (0.50).

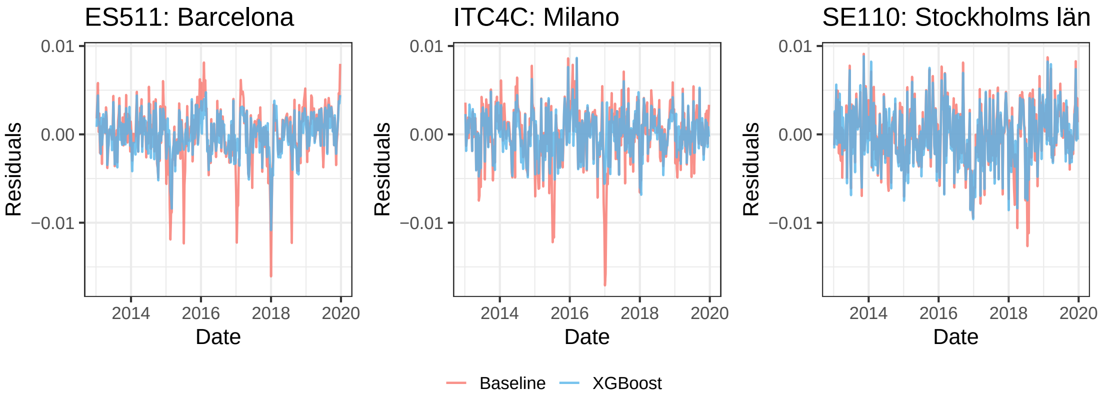
Insights in the machine-learning model

Observed and estimated mortality rates (baseline + XGBoost):



In-sample fit and model performance

Residuals of the estimated weekly mortality rates (baseline + XGBoost):



Feature importance

Machine learning techniques perform automatic feature selection.

Feature importance

Machine learning techniques perform automatic feature selection.

Which features do significantly contribute to the predictions?

Machine learning techniques perform automatic feature selection.

Which features do significantly contribute to the predictions?

We calculate the feature importance of each feature X_l as:

$$\mathcal{V}_{\text{imp}}(X_l) = \frac{1}{\texttt{nrounds}} \sum_{n=1}^{\texttt{nrounds}} \Delta\mathcal{L}_n(X_l),$$

with $\Delta\mathcal{L}_n(X_l)$ the total reduction in the Poisson loss function, caused by splits associated to feature X_l in the tree built during iteration n of the XGBoost algorithm.

Machine learning techniques perform automatic feature selection.

Which features do significantly contribute to the predictions?

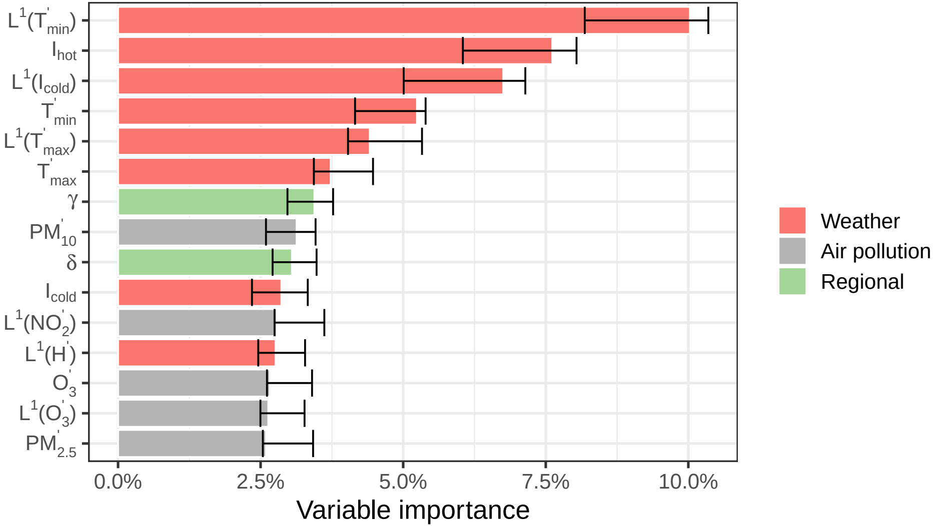
We calculate the feature importance of each feature X_l as:

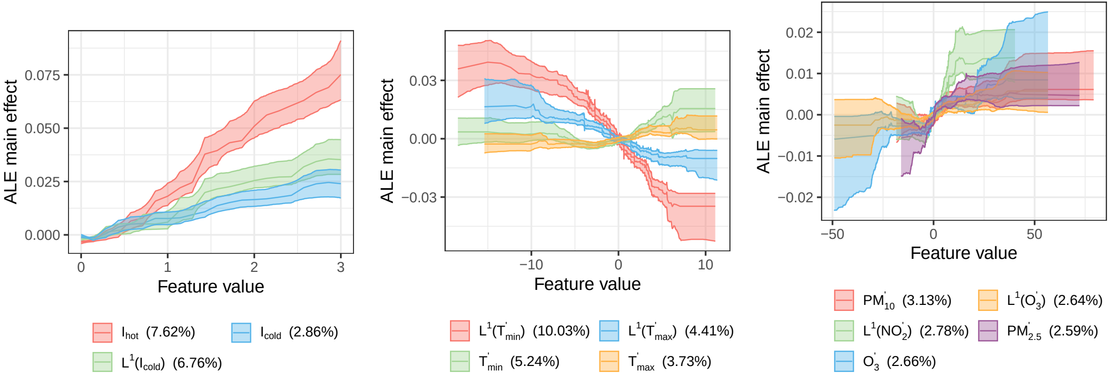
$$\mathcal{V}_{\text{imp}}(X_l) = \frac{1}{\texttt{nrounds}} \sum_{n=1}^{\texttt{nrounds}} \Delta\mathcal{L}_n(X_l),$$

with $\Delta\mathcal{L}_n(X_l)$ the total reduction in the Poisson loss function, caused by splits associated to feature X_l in the tree built during iteration n of the XGBoost algorithm.

Features with a high importance appear often and high in the tree.

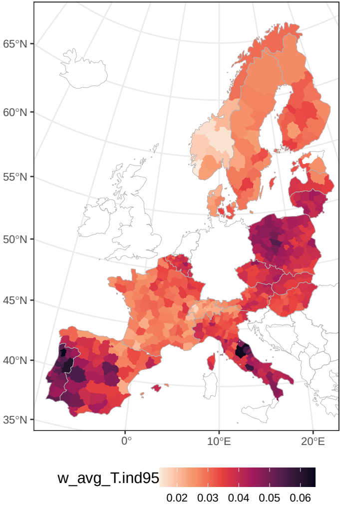
Feature importance



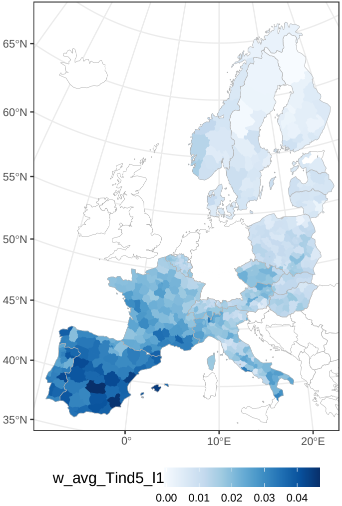

ALE main effect
0.075
0.050
0.025
0.000
0
1
2
3
Feature value
I_{hot} (7.62%)
I_{cold} (2.86%)
$L^1(I_{cold})$ (6.76%)
ALE main effect
0.03
0.00
−0.03
−10
0
10
Feature value
$L^1(T'_{min})$ (10.03%)
$L^1(T'_{max})$ (4.41%)
T'_{min} (5.24%)
T'_{max} (3.73%)
ALE main effect
0.02
0.01
0.00
−0.01
−0.02
−50
0
50
Feature value
PM'_{10} (3.13%)
$L^1(O'_3)$ (2.64%)
$L^1(NO'_2)$ (2.78%)
$PM'_{2.5}$ (2.59%)
O'_3 (2.66%)

ALE regional effects

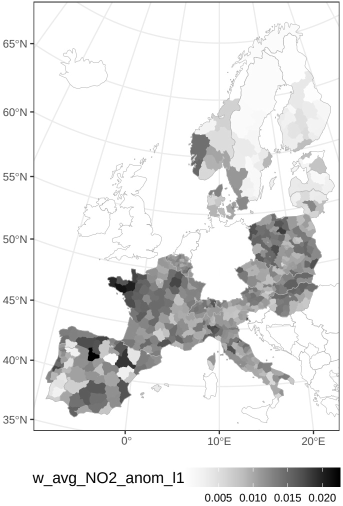
ALE regional effect



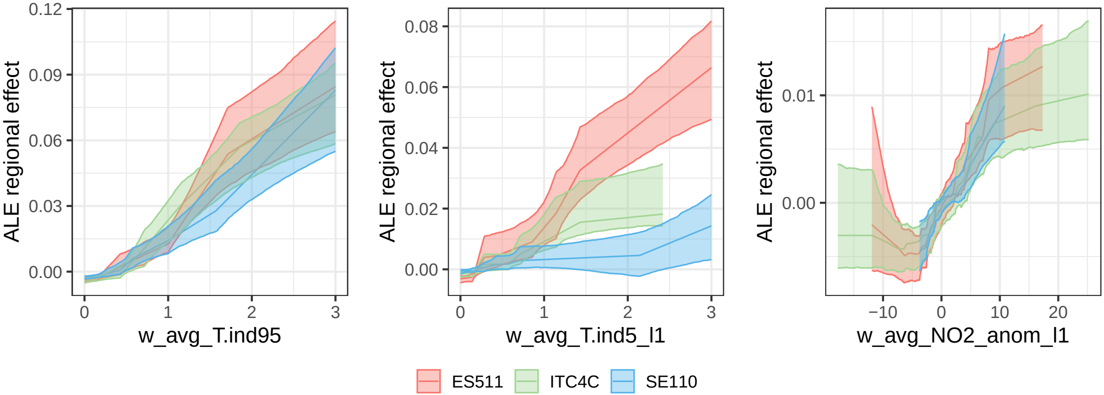
ALE regional effect



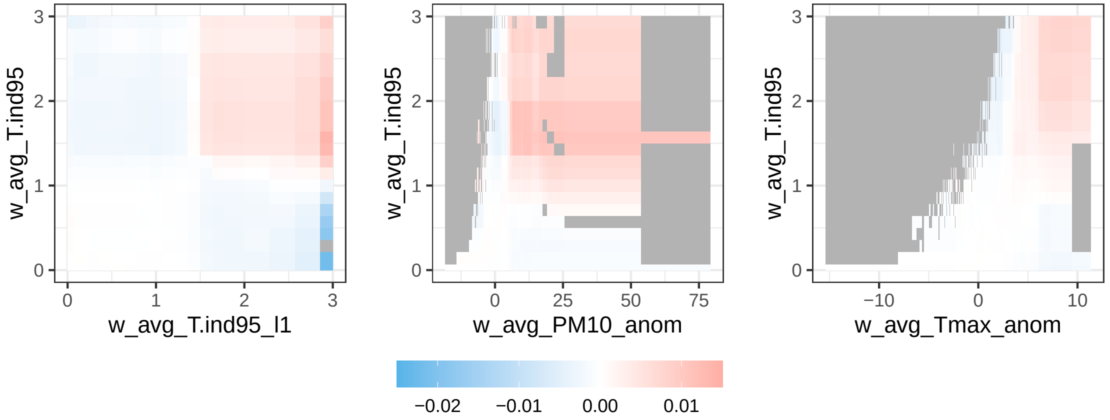
ALE regional effect



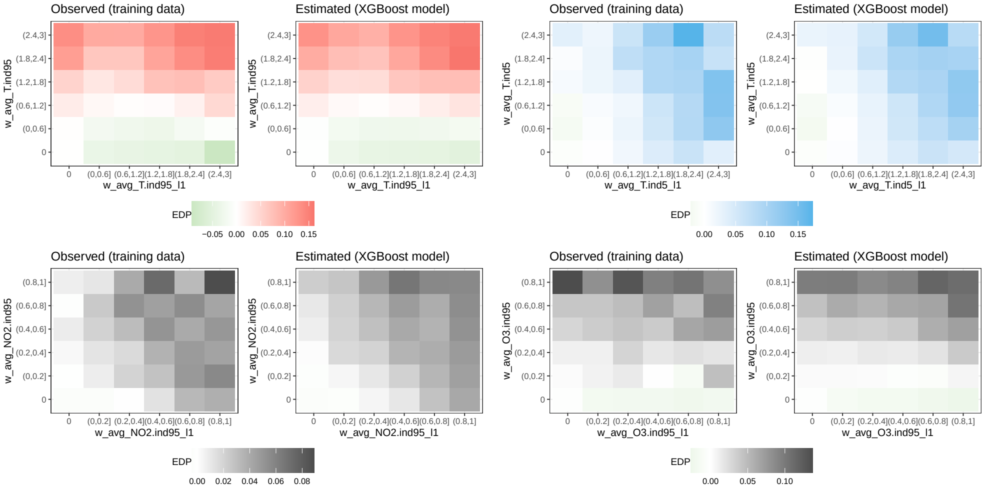
ALE regional effects



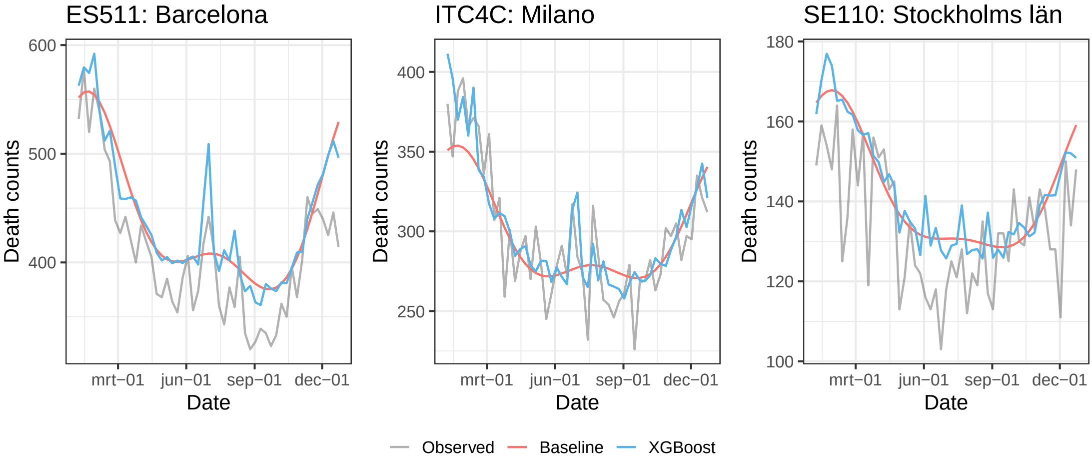
ALE interaction effects



Harvesting effects



Backtest


ES511: Barcelona
ITC4C: Milano
SE110: Stockholms län
Death counts
Date
600
500
400
350
300
250
180
160
140
120
100
mrt−01
jun−01
sep−01
dec−01
Observed
Baseline
XGBoost

We conduct some additional analyses in our paper '*The short-term association between environmental variables and mortality: evidence from Europe*' ([link], under revision at JRSS-A):

We conduct some additional analyses in our paper '*The short-term association between environmental variables and mortality: evidence from Europe*' ([link], under revision at JRSS-A):

1. We compared our method against two other tree-based approaches: random forest and LightGBM. The results were in line with XGBoost, which strengthens the robustness of our findings.

We conduct some additional analyses in our paper '*The short-term association between environmental variables and mortality: evidence from Europe*' ([link], under revision at JRSS-A):

1. We compared our method against two other tree-based approaches: random forest and LightGBM. The results were in line with XGBoost, which strengthens the robustness of our findings.

2. We highlight the advantage of incorporating the baseline number of death counts as an offset in the model. It makes our predictions more stable, robust, and interpretable, especially regarding statements about excess mortality.

Ben Armstrong. Models for the relationship between ambient temperature and daily mortality. *Epidemiology*, pages 624–631, 2006.

Rupa Basu and Jonathan M Samet. Relation between elevated ambient temperature and mortality: a review of the epidemiologic evidence. *Epidemiologic reviews*, 24(2):190–202, 2002.

Alfésio LF Braga, Antonella Zanobetti, and Joel Schwartz. The effect of weather on respiratory and cardiovascular deaths in 12 us cities. *Environmental health perspectives*, 110(9): 859–863, 2002.

Alfésio Luís Ferreira Braga, Antonella Zanobetti, and Joel Schwartz. The time course of weather-related deaths. *Epidemiology*, 12(6):662–667, 2001.

Antonio Gasparrini, Ben Armstrong, and Mike G Kenward. Distributed lag non-linear models. *Statistics in medicine*, 29(21):2224–2234, 2010.

William R Keatinge, Gavin C Donaldson, Elvira Cordioli, Martina Martinelli, Anton E Kunst, Johan P Mackenbach, Simo Nayha, and Ilkka Vuori. Heat related mortality in warm and cold regions of europe: observational study. *Bmj*, 321(7262):670–673, 2000.

Han Li and Qihe Tang. Joint extremes in temperature and mortality: A bivariate pot approach. *North American Actuarial Journal*, 26(1):43–63, 2022.

Pablo Orellano, Julieta Reynoso, Nancy Quaranta, Ariel Bardach, and Agustin Ciapponi. Short-term exposure to particulate matter (pm10 and pm2. 5), nitrogen dioxide (no2), and ozone (o3) and all-cause and cause-specific mortality: Systematic review and meta-analysis. *Environment international*, 142:105876, 2020. doi: 10.1016/j.envint.2020.105876.

Mathilde Pascal, Grégoire Falq, Vérène Wagner, Edouard Chatignoux, Magali Corso, Myriam Blanchard, Sabine Host, Laurence Pascal, and Sophie Larrieu. Short-term impacts of particulate matter (pm10, pm10–2.5, pm2. 5) on mortality in nine french cities. *Atmospheric Environment*, 95:175–184, 2014. doi: 10.1016/j.atmosenv.2014.06.030.

S Pattenden, B Nikiforov, and B Armstrong. Mortality and temperature in sofia and london. *Journal of Epidemiology and Community health*, 57(8):628, 2003.

Joel Schwartz. The distributed lag between air pollution and daily deaths. *Epidemiology*, 11(3): 320–326, 2000.

Robert E Serfling. Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public health reports*, 78(6):494, 1963.