Uncertainty Estimation for High-dimensional Nonparametric Forecasting

Nuwani Palihawadana

Joint work with : Rob Hyndman, Xiaoqian Wang

1 July 2025

Nonparametric forecasting

$$y_t = f(\boldsymbol{a}_t, \boldsymbol{a}_{t-1}, \ldots, \boldsymbol{a}_{t-s}, y_{t-1}, \ldots, y_{t-k}) + \varepsilon_t,$$

- y_t – observed response at time t
- f – arbitrary function
- $\boldsymbol{a}_t$ – vector of exogenous variables at time t
- ε_t – stationary error with mean zero and constant variance σ^2.

Nonparametric forecasting

$$y_t = f(\boldsymbol{a}_t, \boldsymbol{a}_{t-1}, \ldots, \boldsymbol{a}_{t-s}, y_{t-1}, \ldots, y_{t-k}) + \varepsilon_t,$$

- y_t – observed response at time t
- f – arbitrary function
- $\boldsymbol{a}_t$ – vector of exogenous variables at time t
- ε_t – stationary error with mean zero and constant variance σ^2.

We also define:

- $\boldsymbol{z}_t = (y_t, \boldsymbol{x}_t)$ – observation at time t
- $\boldsymbol{x}_t = (\boldsymbol{a}_t, \boldsymbol{a}_{t-1}, \ldots, \boldsymbol{a}_{t-s}, y_{t-1}, \ldots, y_{t-k})$ – vector of all predictors at time t.
- $\hat{y}_t = \hat{f}(\boldsymbol{x}_t)$ – estimated model
- $e_t = y_t - \hat{y}_t$ – model residuals

Sparse multiple index (SMI) model

$$y_i = \beta_0 + \sum_{j=1}^{p} g_j(\boldsymbol{\alpha}_j^T \boldsymbol{x}_{ij}) + \sum_{k=1}^{d} f_k(w_{ik}) + \boldsymbol{\theta}^T \boldsymbol{u}_i + \varepsilon_i, \quad i = 1, \dots, n,$$

- y_i – univariate response
- $\boldsymbol{x}_{ij} \in \mathbb{R}^{\ell_j}$, $j = 1, \dots, p$ – p subsets of predictors entering indices
- $\boldsymbol{\alpha}_j$ – ℓ_j-dimensional vectors of index coefficients
- g_j, f_k – smooth nonlinear functions
- Additional predictors :
  - w_{ik} – nonlinear
  - $\boldsymbol{u}_i$ – linear

Sparse multiple index (SMI) model

$$y_i = \beta_0 + \sum_{j=1}^{p} g_j(\alpha_j^T \boldsymbol{x}_{ij}) + \sum_{k=1}^{d} f_k(w_{ik}) + \boldsymbol{\theta}^T \boldsymbol{u}_i + \varepsilon_i, \quad i = 1, \dots, n,$$

- y_i – univariate response
- $\boldsymbol{x}_{ij} \in \mathbb{R}^{\ell_j}$, $j = 1, \dots, p$ – p subsets of predictors entering indices
- α_j – ℓ_j-dimensional vectors of index coefficients
- g_j, f_k – smooth nonlinear functions
- Additional predictors :
  - w_{ik} – nonlinear
  - $\boldsymbol{u}_i$ – linear

Allow elements equal to zero in α_j – "Sparse"

Sparse multiple index (SMI) model

$$y_i = \beta_0 + \sum_{j=1}^{p} g_j(\boldsymbol{\alpha}_j^T \boldsymbol{x}_{ij}) + \sum_{k=1}^{d} f_k(w_{ik}) + \boldsymbol{\theta}^T \boldsymbol{u}_i + \varepsilon_i, \quad i = 1, \dots, n,$$

- y_i – univariate response
- $\boldsymbol{x}_{ij} \in \mathbb{R}^{\ell_j}$, $j = 1, \dots, p$ – p subsets of predictors entering indices
- $\boldsymbol{\alpha}_j$ – ℓ_j-dimensional vectors of index coefficients
- g_j, f_k – smooth nonlinear functions
- Additional predictors :
  - w_{ik} – nonlinear
  - $\boldsymbol{u}_i$ – linear

Both "p" and the predictor grouping among indices are unknown.

Overlapping of predictors among indices is not allowed.

Benchmarks

Forecast uncertainty

- Uncertainty of a forecast $\rightarrow$ **Prediction Interval (PI)**

Forecast uncertainty

- Uncertainty of a forecast $\rightarrow$ **Prediction Interval (PI)**
- Theoretical $100(1-\alpha)\%$ prediction interval:

$$\hat{y}_{t+h|t} \pm z_{\alpha/2} \times \hat{\sigma}_h,$$

where

  - y – time series $y_1, \ldots, y_T$
  - $\hat{y}_{t+h|t}$ – h-step-ahead point forecast for y_{t+h} given observations up to t
  - $z_{\alpha/2}$ – $\alpha/2$ quantile of standard normal distribution
  - $\hat{\sigma}_h$ – estimate of std. deviation of h-step forecast distribution

Forecast uncertainty

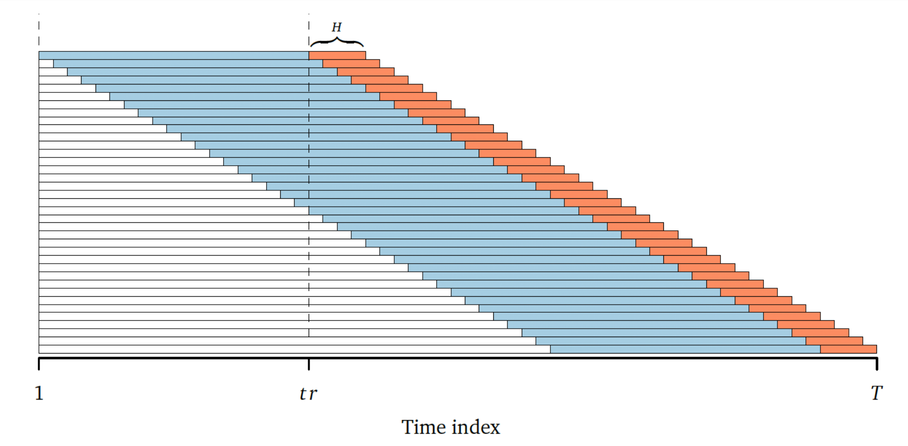
- Uncertainty of a forecast $\rightarrow$ **Prediction Interval (PI)**
- Theoretical $100(1-\alpha)\%$ prediction interval:

$$\hat{y}_{t+h|t} \pm z_{\alpha/2} \times \hat{\sigma}_h,$$

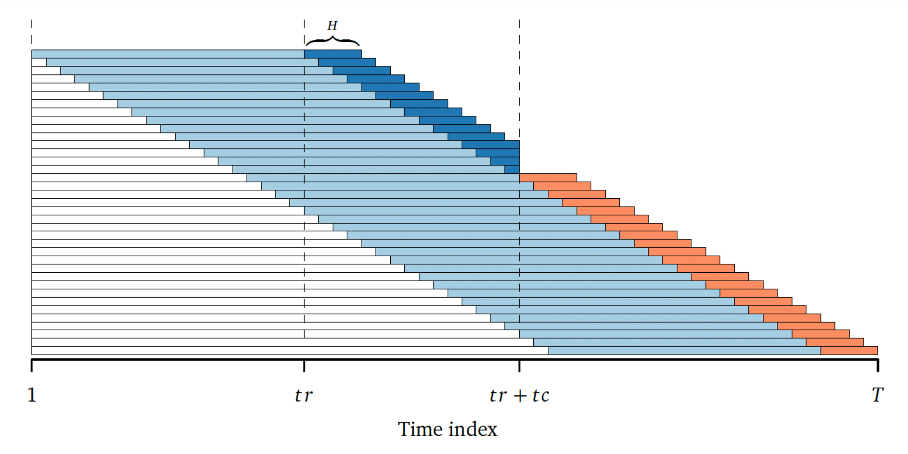
where

  - y – time series $y_1, \ldots, y_T$
  - $\hat{y}_{t+h|t}$ – h-step-ahead point forecast for y_{t+h} given observations up to t
  - $z_{\alpha/2}$ – $\alpha/2$ quantile of standard normal distribution
  - $\hat{\sigma}_h$ – estimate of std. deviation of h-step forecast distribution
- Main issue:
  - Difficult to analytically calculate h-step forecast variances for $h > 1$

Block bootstrap

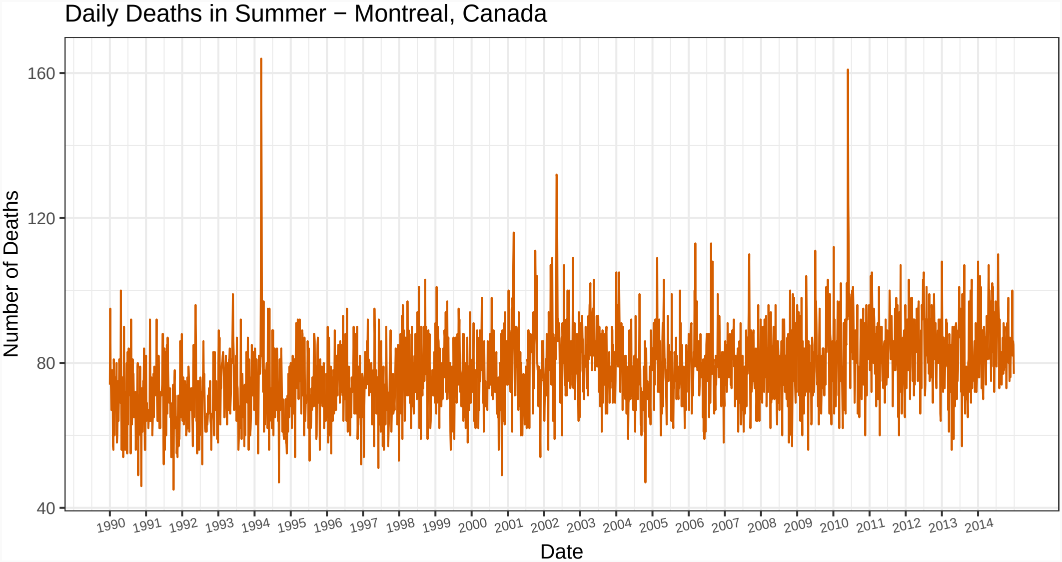


Conformal prediction



Conformal bootstrap

Forecasting heat exposure-related daily mortality



Forecasting heat exposure-related daily mortality

Data

- **Response: Daily deaths in Summer** – 1990 to 2014 – Montreal, Canada
- **Index Variables:**
  - Death lags
  - Max temperature lags
  - Min temperature lags
  - Vapor pressure lags
- **Nonlinear:** DOS (day of the season), Year

Forecasting heat exposure-related daily mortality

Data

- **Response: Daily deaths in Summer** – 1990 to 2014 – Montreal, Canada
- **Index Variables:**
  - Death lags
  - Max temperature lags
  - Min temperature lags
  - Vapor pressure lags
- **Nonlinear:** DOS (day of the season), Year

Data split

- **Training Set:** 1990 to 2007
- **Validation Set:** 2008
- **Test Set:** 2009 to 2014

Conclusion

Summary of Results (work-in-progress):

- **Block Bootstrap** – Under-coverage; too narrow
- **Conformal Prediction** – Better achieves a target coverage, with acceptable sharpness

Limitations:

- Test set is not long enough for larger forecast horizons
- Hyper-parameter choices

Find me :

nuwanipalihawadana.netlify.app

in/nuwani-palihawadana

@nuwani-palihawadana

nuwani.kodikarapalihawadana@monash.edu