Forecasting Uncertainty in Electricity Demand

Our general model for electricity demand:

\[ Y_t = \mu(x_t) + \sigma(x_t) \varepsilon_t \]

- \( y_t \) is the demand at time \( t \)
- \( x_t \) is vector of covariates
- \( \varepsilon_t \) is the error at time \( t \)

In practice, \( \mu(\cdot) \) and \( \sigma^2(\cdot) \) are unknown!

Estimate \( \mu(\cdot) \) and \( \sigma^2(\cdot) \) from empirical data
No bootstrapping, no ensemble

1. **Estimating the conditional mean \( \mu(\cdot) \)**

   - Use Generalized Additive Model (GAM):
     \[ F(\mu(x_t)) = \sum_{i=1}^{l} f_i(x_t) \]

     - link function, e.g., identity or logarithm
     - transfer function
     - basis functions

   - Fit a GAM to \( y_t \) to get an estimate \( \hat{\mu}(\cdot) \) of \( \mu(\cdot) \)

   \[ \log(\mathbb{E}[Y_t]) = \beta_0 + \beta_1 \text{DayType}_t + \sum_{j=1}^{5} \mathbf{1}(\text{DayType}_t=j) f_{1,j}(\text{HourOfDay}_t) + f_2(\text{TimeOfYear}_t) + f_3(\text{TempC}_t, \text{Humidity}_t) + f_4(\text{TempC}_t, \text{HourOfDay}_t) + \sum_{j=1}^{5} \mathbf{1}(\text{DayType}_t=j) f_{6,j}(\text{LagLoad}_t) \]

2. **Estimating the conditional variance \( \sigma^2(\cdot) \)**

   - We use another GAM:
     \[ G(\sigma^2(x_t)) = \sum_{j=1}^{l} g_j(x_t) \]

   - Let \( z_t = \sigma(x_t) \varepsilon_t \)
     \[ \hat{z}_t = \hat{\gamma}_t - y_t \]

   - Fit a GAM to \( \hat{z}_t^2 \) to get an estimate \( \hat{\sigma}^2(\cdot) \) of \( \sigma^2(\cdot) \)

   \[ \log(\mathbb{E}[\hat{z}_t^2]) = \beta_0 + \beta_1 \text{DayType}_t + \sum_{j=1}^{5} \mathbf{1}(\text{DayType}_t=j) f_{1,j}(\text{HourOfDay}_t) + f_2(\text{TimeOfYear}_t) + f_3(\text{TempC}_t, \text{Humidity}_t) + f_4(\text{TempC}_t, \text{HourOfDay}_t) \]

3. **Computing prediction intervals**

   - \( q(\cdot) \) is the quantile function of the standard normal dist.

   - **two-sided**
     \[ \phi^2_t(p) = \hat{\mu}(x_t) \pm q\left(1 - \frac{p}{2}\right) \cdot \hat{\sigma}(x_t) \]

   - **one-sided**
     \[ \phi^1_t(p) = \hat{\mu}(x_t) + q(p) \cdot \hat{\sigma}(x_t) \]

4. **Online learning**

   - Adaptive learning for smoothing functions (Ba et al., NIPS, 2012)

   - Adaptive construction of prediction intervals
     \[ c = \text{current empirical coverage} \]
     \[ \hat{c} = \frac{p - ac}{1 - \alpha} \]
     \[ \phi^1_t(p) = \hat{\mu}(x_t) + q(\hat{c}) \cdot \hat{\sigma}(x_t) \]
     \[ \phi^2_t(p) = \hat{\mu}(x_t) \pm q\left(1 - \frac{\hat{c}}{2}\right) \cdot \hat{\sigma}(x_t) \]

The dataset is available at https://github.com/tritritri/uncertainty