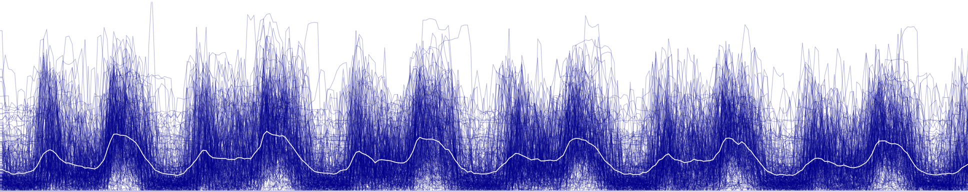
Generalized additive models for electricity demand forecasting: thinking inside the box

Matteo Fasiolo

Joint work with:

Simon N. Wood (University of Bristol, UK)
Yannig Goude (EDF R&D)
Margaux Zaffran (ENSTA Paris)
Raphaël Nedellec (Talend, formerly EDF R&D)

matteo.fasiolo@bristol.ac.uk



Talk structure

1. GAMs for electricity load forecasting
2. Probabilistic forecasting with GAMLSS and quantile GAMs
3. Current work: multi-resolution GAMs

Introduction to GAMs

We consider Generalized Additive Models (GAMs, Hastie and Tibshirani (1990)), which are used by Électricité de France to forecast demand.

GAM model structure:

$$\text{Load}_i|\mathbf{x}_i \sim \text{Distr}\{\text{Load}_i|\theta_1 = \mu(\mathbf{x}_i), \theta_2, \dots, \theta_p\},$$

where

$$\mathbb{E}(\text{Load}_i|\mathbf{x}_i) = \mu(\mathbf{x}_i) = g^{-1}\Big\{\sum_{j=1}^{m} f_j(\mathbf{x}_i)\Big\},$$

and g is the link function.

f_j's can be fixed (parametric) or smooth effects.

$\theta_2, \dots, \theta_p$ control scale and shape of distribution.

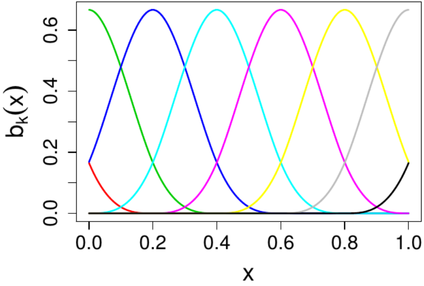
Introduction to GAMs

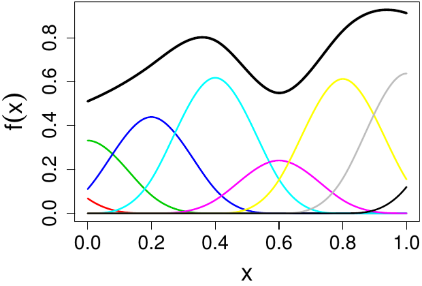
Smooth effects built using spline bases

$$f(x) = \sum_{k=1}^{r} \beta_k b_k(x)$$

where

- β_k unknown coeff
- $b_k(x)$ known spline basis functions
- smoothness of $f(x)$ controlled by smoothing prior $p(\boldsymbol{\beta}|\boldsymbol{\lambda})$.





Introduction to GAMs

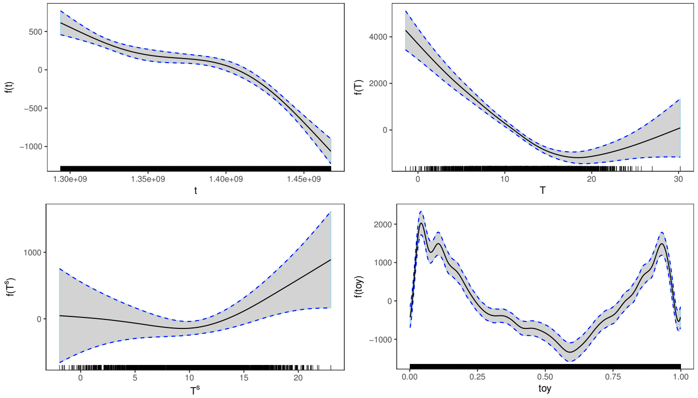
Example: a Gaussian GAM for expected load is

$$
\begin{aligned}
\mathbb{E}(\texttt{Load}_i) \;=\;& \sum_{j=1}^{7} \beta_j w_{d(i)}^{j} && \cdot \text{ Day-of-week factor} \\
+\;& \beta_8 \texttt{Load}_{i-1} && \cdot \text{ Lagged load} \\
+\;& \beta_9 \texttt{h}_i && \cdot \text{ Holiday binary} \\
+\;& f_1(t_i) && \cdot \text{ Long-term trend} \\
+\;& f_2(T_i) && \cdot \text{ Temperature} \\
+\;& f_3(T_i^s) && \cdot \text{ Smoothed temperature (for thermal inertia)} \\
+\;& f_4(\texttt{toy}_i), && \cdot \text{ Time-of-year}
\end{aligned}
$$

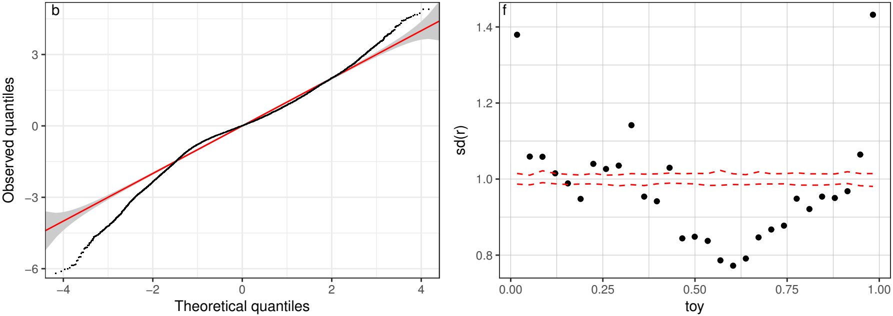
where $T_i^s = \alpha T_i + (1-\alpha) T_{i-1}^s$, with $\alpha = 0.05$.

Using `mgcv` R package (Wood, 2001):

```
fit <- gam(load ~ dow + loadLag + holy + s(time) +
                  s(temp) + s(tempSmo) + s(toy),
           family = gaussian, data = UKload)
```



Limitation: parametric assumption on $\mathrm{Distr}(y|\mathbf{x})$.



Talk structure

1. GAMs for electricity load forecasting
2. Probabilistic forecasting with GAMLSS and quantile GAMs
3. Current work: multi-resolution GAMs

From GAMs to GAMLSS

Generalized Additive Models for Location Scale and Shape (GAMLSS, Rigby and Stasinopoulos (2005)) let scale and shape change with $\mathbf{x}$.

GAMLSS model structure:

$$\text{Load}|\mathbf{x} \sim \text{Distr}\{\text{Load}|\theta_1 = \mu_1(\mathbf{x}), \theta_2 = \mu_2(\mathbf{x}), \dots, \theta_p = \mu_p(\mathbf{x})\},$$

where

$$\mu_1(\mathbf{x}) = g_1^{-1}\Big\{\sum_{j=1}^{m_1} f_j^1(\mathbf{x})\Big\},$$

$$\dots$$

$$\mu_p(\mathbf{x}) = g_p^{-1}\Big\{\sum_{j=1}^{m_p} f_j^p(\mathbf{x})\Big\},$$

and $g_1, \dots, g_p$ are link function.

From GAMs to GAMLSS

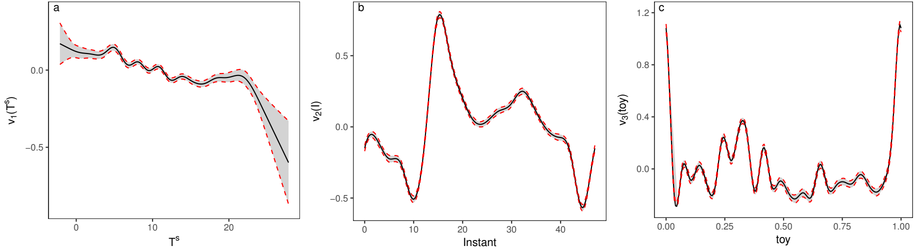
Example: Gaussian model for location and scale (see `?mgcv::gaulss`)

$$\text{Load}|\mathbf{x} \sim N\{\text{Load}|\mu(\mathbf{x}), \sigma(\mathbf{x})\}$$

where

$$\mu(\mathbf{x}) = \sum_{j=1}^{m} f_j^1(\mathbf{x}), \quad \sigma(\mathbf{x}) = \exp\left\{\sum_{j=1}^{m} f_j^2(\mathbf{x})\right\}$$

and $g_2 = \log$ to guarantee $\sigma > 0$.

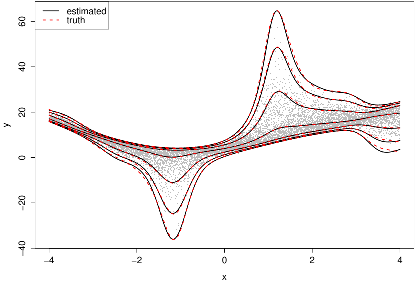


From GAMs to GAMLSS

```
fit <- gam(list(load ~ s(time) + ...,        # location
                     ~ s(temp) + ...,        # scale
                     ~ s(toy) + ...,         # skewness
                     ~ s(instant) + ...      # kurtosis)
```

Still parametric assumption on Distr(load|**x**).

Quantile regression estimates quantiles $\mu_\tau(\mathbf{x})$ for $\tau \in (0, 1)$ directly.



Example: a QGAM for daily electricity load is

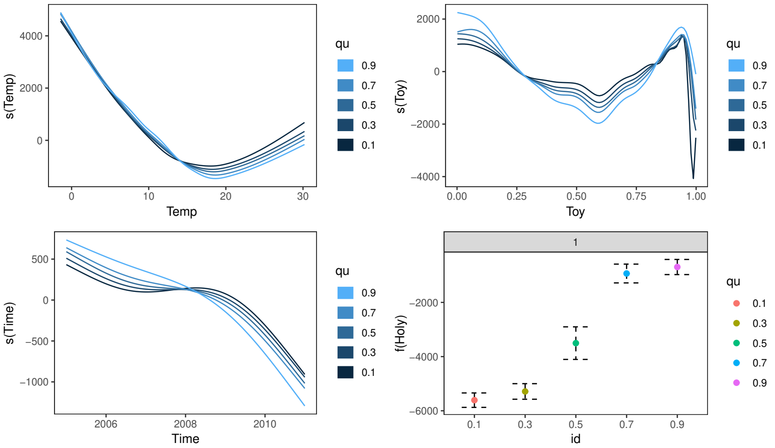
$$
\begin{aligned}
\mu_\tau(\mathbf{x}_i) \;&=\; \sum_{j=1}^{7} \beta_j w^j_{d(i)} && \cdot \text{ Day-of-week factor} \\
&+\; \beta_8 \texttt{Load}_{i-1} && \\
&+\; \beta_9 \texttt{h}_i && \cdot \text{ Holiday binary} \\
&+\; f_1(t_i) && \cdot \text{ Long-term trend} \\
&+\; \cdots &&
\end{aligned}
$$

Implemented by qgam R package (Fasiolo et al., 2018):

```
fit <- mqgam(load ~ dow + loadLag + holy + s(time) +
                    s(temp) + s(tempSmo) + s(toy),
             qu = c(0.1, 0.3, 0.5, 0.7, 0.9),
             data = UKload)
```

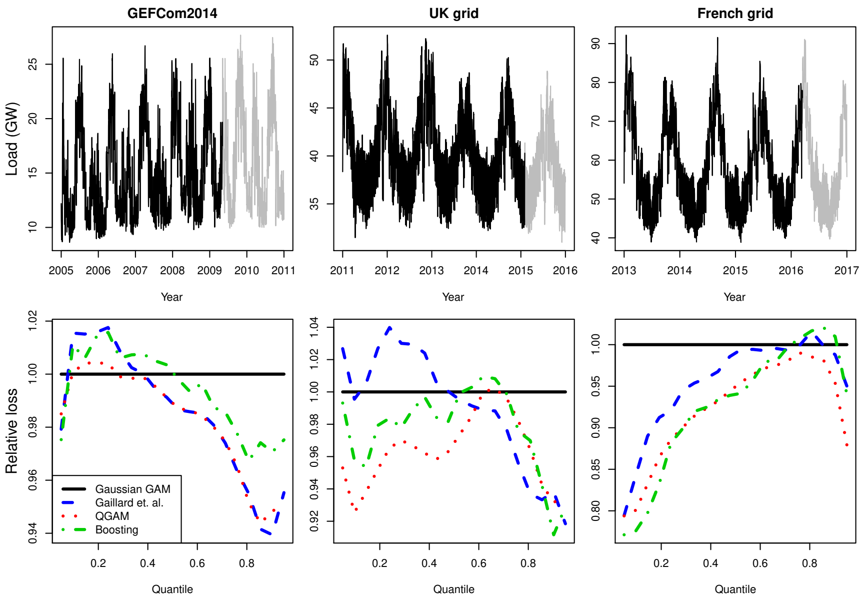
From GAMLSS to QGAM

Fit on aggregate UK demand data:



Plots produced using `mgcViz` visualization package (Fasiolo et al., 2018).

From GAMLSS to QGAM



Talk structure

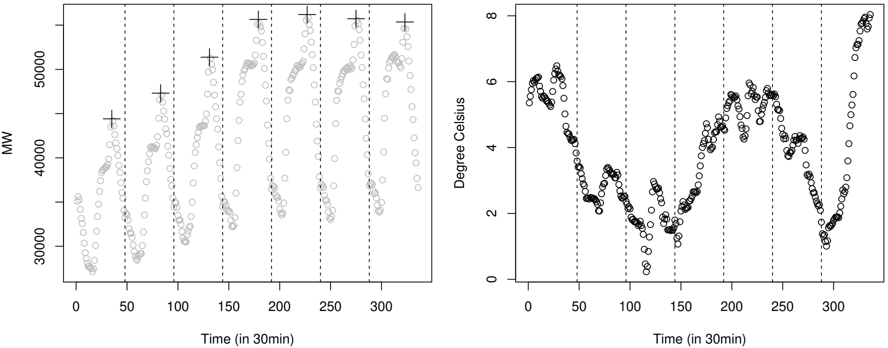
1. GAMs for electricity load forecasting
2. Probabilistic forecasting with GAMLSS and quantile GAMs
3. Current work: multi-resolution GAMs

Multi-resolution GAMs

Consider modelling max demand over time horizon.

We have n days and 30min electricity demand $L_{1:48n}$.

We want to predict y_i, the maximal demand on the i-th day.



Multi-resolution GAMs

We need to deal with data at different resolutions.

Modelling approach:

- distribution for day max y_i is Generalized Extreme Value (GEV)
- capture information at 30min resolution using functional effects

Integrating high-resolution data:

- naive approach $\mathbb{E}(y_i) = f_1(\text{Temp}_1^i) + \cdots + f_{48}(\text{Temp}_{48}^i) + \cdots$
- functional $\mathbb{E}(y_i) = \sum_{k=1}^{48} \text{te}(\text{Temp}_k^i, k) + \cdots$

Multi-resolution GAMs

Final model for daily max on UK data is $y_i \sim \text{GEV}(\mu, \sigma, \xi)$ where

$$\mu_i = \sum_{k=1}^{7} \beta_k \mathbb{I}(\text{wd}_i = k) + s_1(\text{toy}_i) + s_2(\text{t}_i) + \sum_{k=1}^{48} \text{te}_1(\text{temp}_k^i, k) + \sum_{k=1}^{48} \text{te}_2(\text{L}_k^{i-1}, k).$$

RMSE on test set (UK data):

- Multi-resolution: 773 (best)
- Big model by-instant: 965
- 48 models by-instant: 930

Multi-resolution GAMs

Note y_i does not need to be daily max:

- total demand in a day ($y_i \sim$ Normal?)
- position of daily max ($y_i \in \{1, \dots, 48\}$, $y_i \sim$ OCAT?)

and functional structure stays the same.

We can be multi-resolution across space:

$$\mathbb{E}(\text{Load}_i) = \int f\left\{\text{lon}, \text{lat}, \text{temp}(\text{lon}, \text{lat})\right\} \, d\text{lon} \, d\text{lat} + \cdots$$
$$\approx \sum_k \text{te}(\text{lon}_k, \text{lat}_k, \text{temp}_k^i) + \cdots$$

Basic functional effects are in `mgcv` (see `?linear.functional.terms`), for more methods see `refund` package (Crainiceanu et al., 2012).

Conclusion

The additive structure of GAMs offers:

- interpretability (see `mgcViz` visualization R package)
- scalability to Big Data (see Wood et al. (2017) and mgcv::bam())
- modularity

Modularity facilitates addition of new:

- response distributions (e.g. GEV)
- smooth effect types (e.g. functional terms)
- model classes (e.g. GAMLSS and quantile GAMs)

These properties, and the availability of **reliable open-source** software, should assure the competitiveness of additive models in the context of modelling future energy systems.

References I

Crainiceanu, C., P. Reiss, J. Goldsmith, L. Huang, L. Huo, F. Scheipl, S. Greven, J. Harezlak, M. Kundu, and Y. Zhao (2012). refund: Regression with functional data. *R package version 0.1-6*.

Fasiolo, M., Y. Goude, R. Nedellec, and S. N. Wood (2018). Fast calibrated additive quantile regression. *arXiv preprint arXiv:1707.03307*.

Fasiolo, M., R. Nedellec, Y. Goude, and S. N. Wood (2018). Scalable visualisation methods for modern generalized additive models. *arXiv preprint arXiv:1809.10632*.

Hastie, T. and R. Tibshirani (1990). *Generalized Additive Models*, Volume 43. CRC Press.

Hothorn, T., P. Bühlmann, T. Kneib, M. Schmid, and B. Hofner (2010). Model-based boosting 2.0. *The Journal of Machine Learning Research 11*, 2109–2113.

Koenker, R. (2005). *Quantile regression*. Number 38. Cambridge university press.

McLean, M. W., G. Hooker, A.-M. Staicu, F. Scheipl, and D. Ruppert (2014). Functional generalized additive models. *Journal of Computational and Graphical Statistics 23*(1), 249–269.

References II

Rigby, R. A. and D. M. Stasinopoulos (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 54*(3), 507–554.

Wood, S. N. (2001). mgcv: Gams and generalized ridge regression for r. *R news 1*(2), 20–25.

Wood, S. N., Z. Li, G. Shaddick, and N. H. Augustin (2017). Generalized additive models for gigadata: modeling the uk black smoke network daily data. *Journal of the American Statistical Association 112*(519), 1199–1210.