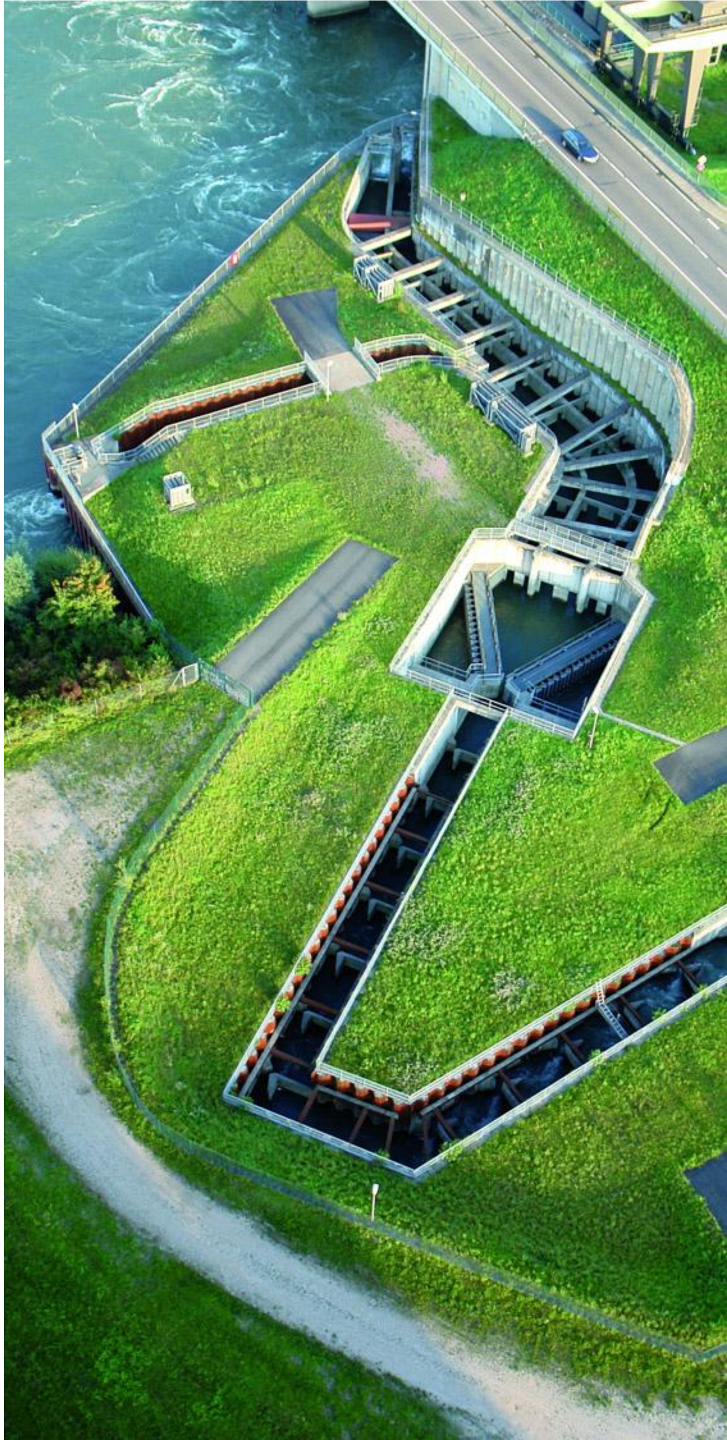
PROBABILISTIC LOAD FORECASTS AT EDF

SAMba ITT: skype meeting, 23rd april 2015

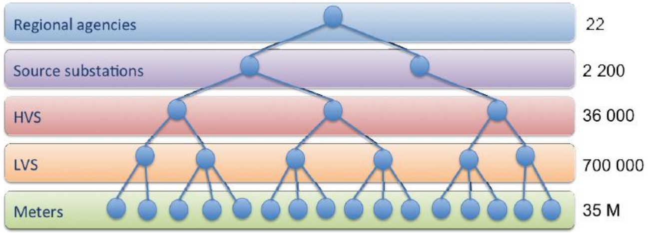
Yannig Goude & Raphaël Nédellec

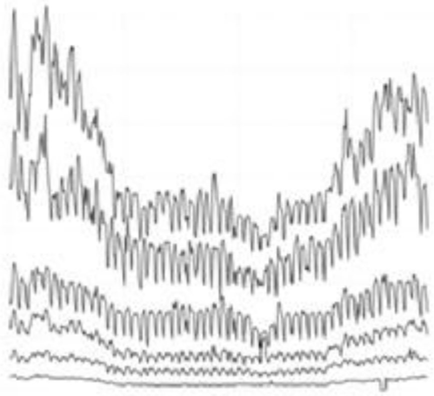
EDF R&D

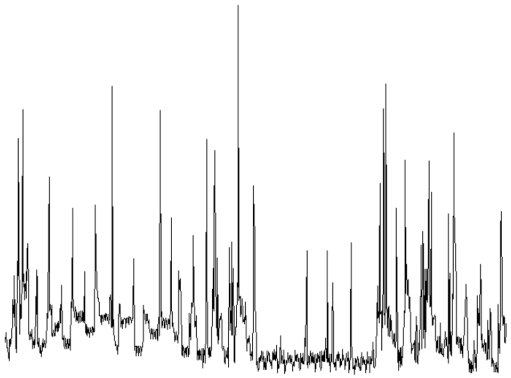


INDUSTRIAL MOTIVATION

- **Renewables energy development**
  - A « probabilistic world »: *Wind power generation is a direct function of the meteorological conditions, which we humans have no control of, and is hence highly fluctuating* Pr. P.Pinson
- **Opening of the electricity market:**
  - Losses/gains of customers:
    - scenario based forecasts, adaptivity, time varying parameters
    - Bottom up forecasts
- **Local forecasting**
  - Low agregation level
  - Wide variety of consumers
  - Covariate selection
  - Noisy data



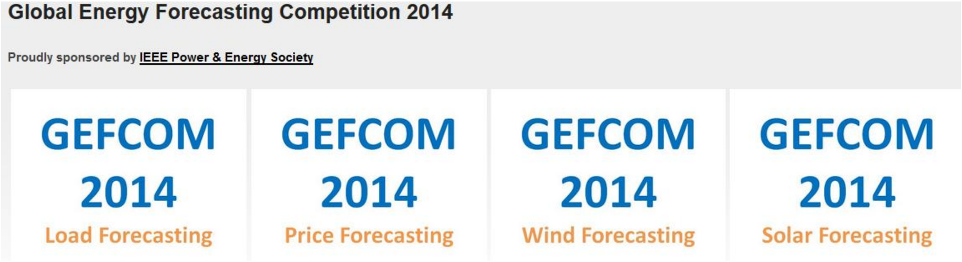




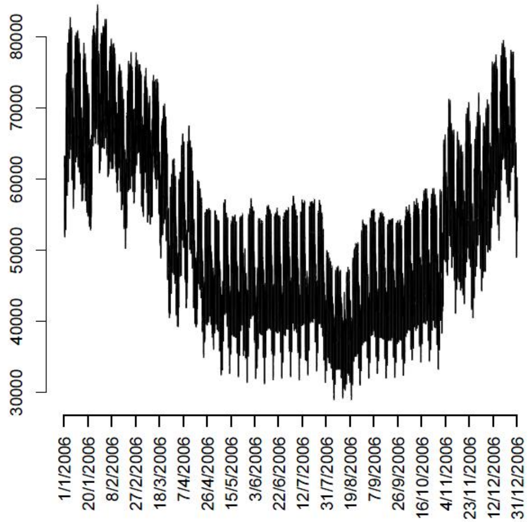
PROBABILISTIC FORECAST: A HOT TOPIC

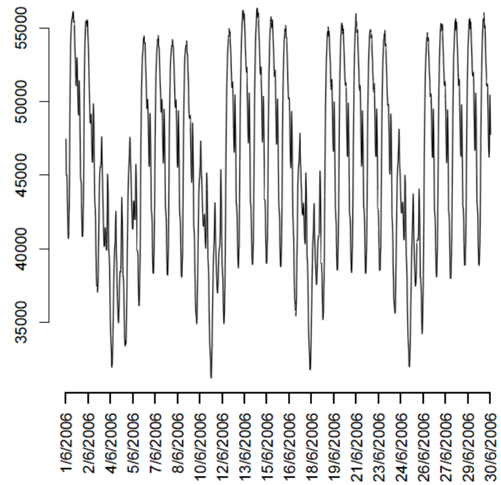
- **A special issue of International Journal of Forecasting**

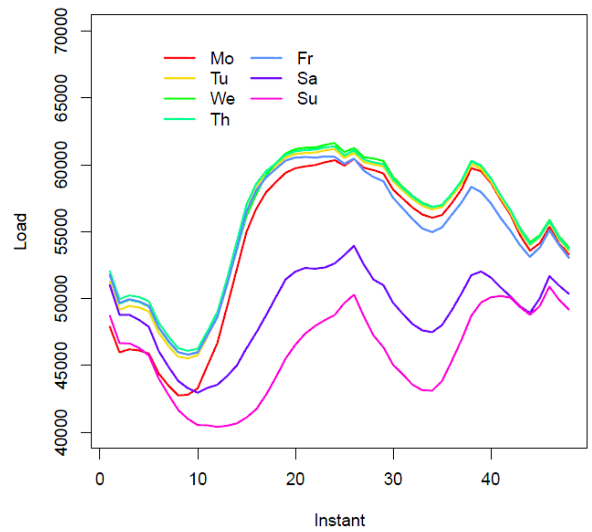
- **GEFcom 2014 competition:**

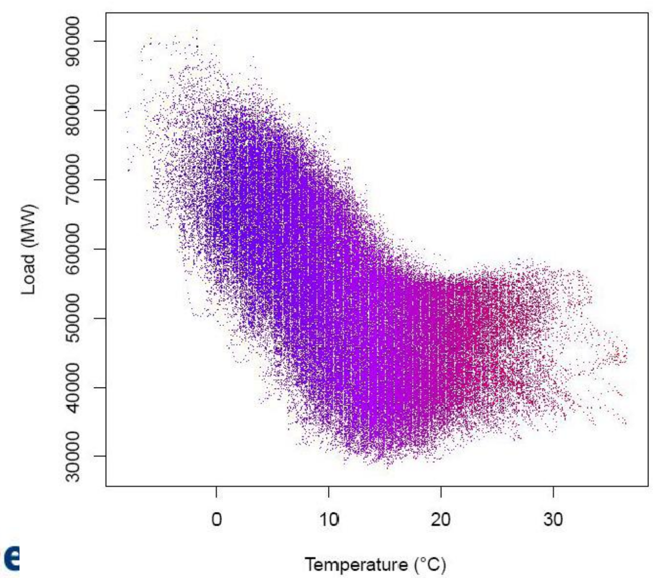


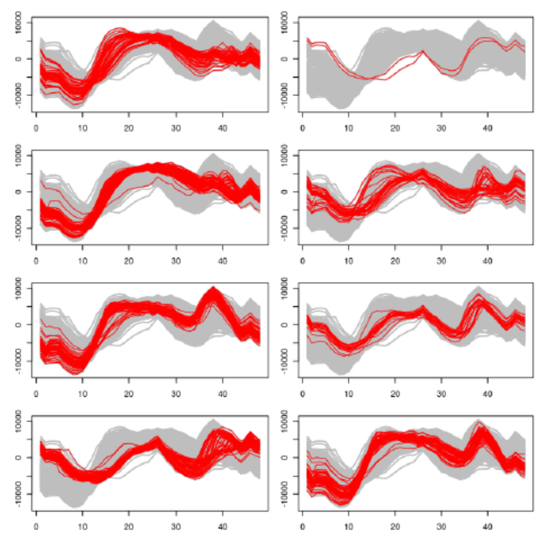
ELECTRICITY LOAD DATA











EDF WORK ON LOAD FORECASTING

- **Models for the conditional expectation:**

  - Parametric models:
    - Metehore equation (Premis, Eventail, Kheops)

  - GAM models
    - French load, EDF load
    - Substations, Feeders

  - Functionnal models
    - Kernel Wavelet Functional
    - Curve Linear Regression

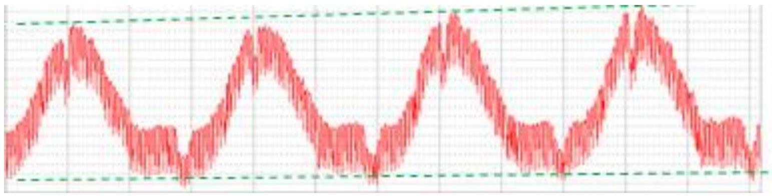
  - Agregation of experts

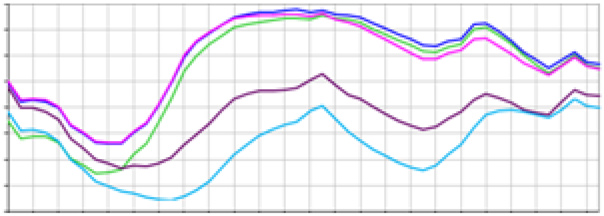
EDF WORK ON LOAD FORECASTING: METEHORE

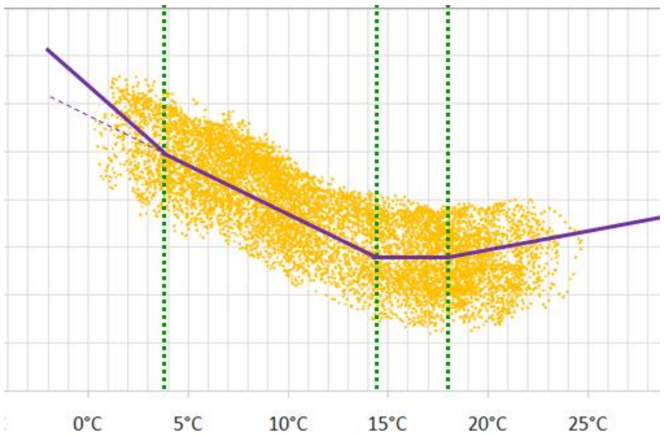
- **Metehore model**
  - Historical model at EDF
  - A high dimension non-linear regression model

$$y_t = S_t + Wd_t + Tf_t + \varepsilon_t$$

- S_t: weather independent seasonal variations of the load: daily, weekly, yearly
- Wd_t: weather dependent load (heating and cooling)
- Tf_t: Tarif effects
- ε_t: an AR process





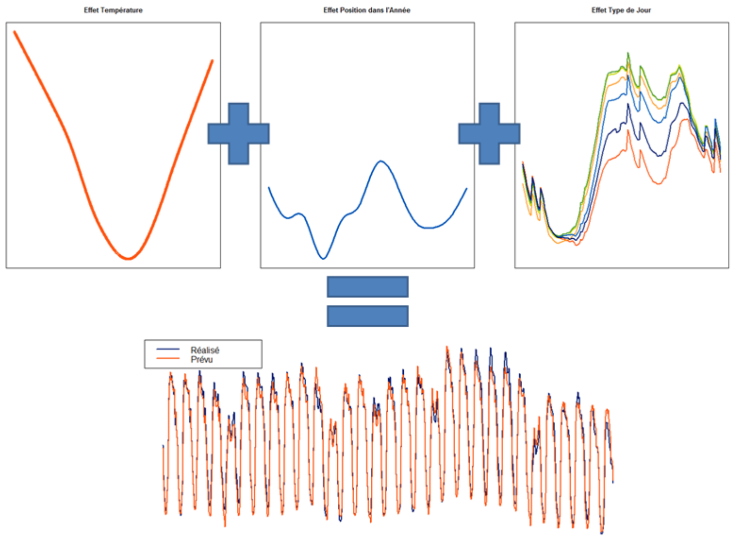


EDF WORK ON LOAD FORECASTING: GAM

- **A good trade-off complexity/adaptivity**

$$y_t = f_1(x_t^1) + f_2(x_t^2) + \dots + f(x_t^3, x_t^4) + \dots + \varepsilon_t$$

$$min_{\beta, f_j} ||y - f_1(x_1) - f_2(x_2) - \dots||^2 + \lambda_1 \int f_1''(x)^2 dx + \lambda_2 \int f_2''(x)^2 dx + \dots$$

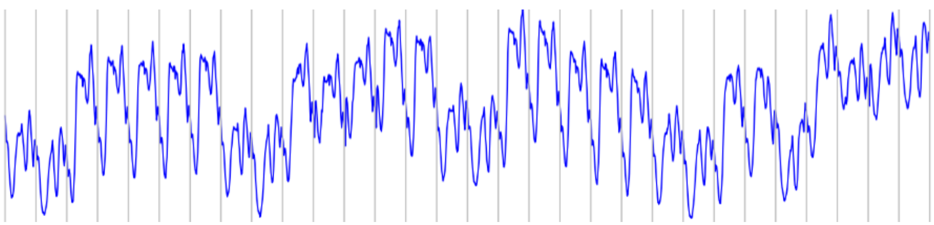


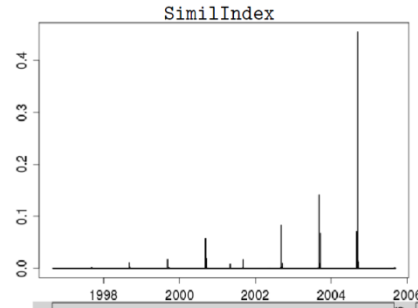
- A. Pierrot and Y. Goude, Short-Term Electricity Load Forecasting With Generalized Additive Models **Proceedings of ISAP power**, pp 593-600, 2011.
- R. Nédellec, J. Cugliari and Y. Goude, GEFCom2012: Electricity Load Forecasting and Backcasting with Semi-Parametric Models, **International Journal of Forecasting** , 2014, 30, 375 - 381.
- Goude, Y., Nédellec, R. and Kong, N., Local Short and Middle term Electricity Load Forecasting with semi-parametric additive models, **IEEE transactions on smart grid**, 2013, 5 , Issue: 1, 440 – 446.

EDF WORK ON LOAD FORECASTING: KWF

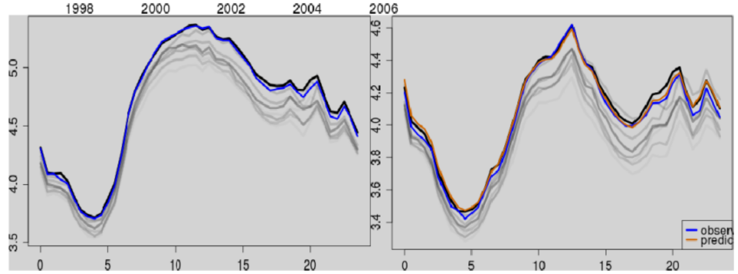
- **Kernel Wavelet Functional models:**

$$\widehat{\mathrm{Load}}_{n+1}(t) = \sum_{m=1}^{n-1} \mathrm{SimilIndex}_m \times \mathrm{Load}_{m+1}(t)$$





date	SimilIndex
2004-09-10	0.455
2003-09-05	0.141
2002-09-06	0.083
2004-09-03	0.070
2003-09-19	0.068
2000-09-08	0.058
2000-09-15	0.019
1999-09-10	0.017

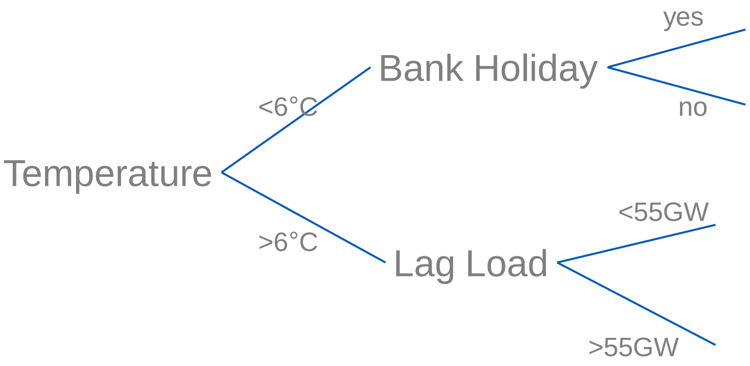


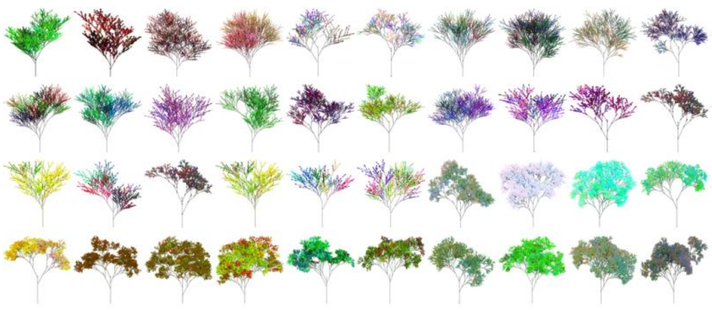
- Similarity index: distance based on wavelet decomposition day by day of the load

- Antoniadis, A., Brossat, X., Cugliari, J., Poggi, J.: Clustering functional data using wavelets. In: **Proceedings of the Nineteenth International Conference on Computational Statistics(COMPSTAT),** 2010.
- Antoniadis, A., Paparoditis, E., Sapatinas, T.: A functional wavelet–kernel approach for time series prediction. **Journal of the Royal Statistical Society: Series B** 68(5), 837–857, 2006.
- J. Cugliari, Prévision non paramétrique de processus à valeurs fonctionnelles. Application à la consommation d'électricité, thèse Université Paris-Sud, 2011.

EDF WORK ON LOAD FORECASTING: RANDOM FOREST

- **Random forest: a popular machine learning method for classification/regression**
  - Breiman, L., . Random Forests, **Machine Learning,** 45 (1), 2001.





http://luc.devroye.org/BRUCE/brucepics.html

EDF WORK ON PROBABILISTIC LOAD FORECASTING

- **Statistical models on the residuals**
  - Parametric models
  - GAM models
  - bootstrap
- **Functionnal models**
- **Bayesian models**
- **Quantile regression**
  - Linear regression
  - Quantile random forest
  - GAM quantile
- **Combining algorithms**

STATISTICAL MODELS ON RESIDUALS

- **2 interships at EDF R&D:**
  - P. Grente ,G. Adde 2009, *Probabilisation de la prévision de consommation électrique*, Technical Report
  - B. Finas ,G. Adde 2010, *Incertitude autour de la prévision de consommation électrique*, Technical Report

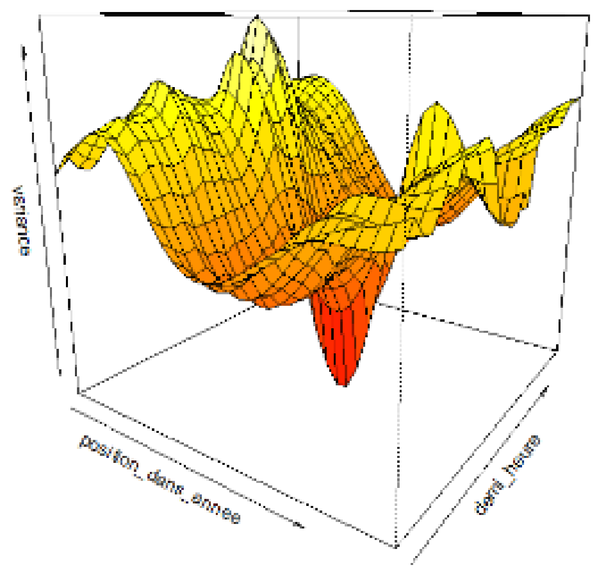
- **GAM models on Eventail residuals:**

  - Bias correction:

$$\varepsilon_t = b_t + u_t$$
$$b_t = f(Toy_t) + \sum_{k=1}^{K} g(h_t)\,\mathrm{I}_{d_t=k}$$

  - Standardization:

$$u_t = \sigma_t w_t$$
$$\sigma_t = f(h_t, Toy_t) + g(h_t, T_t) + \xi_t$$



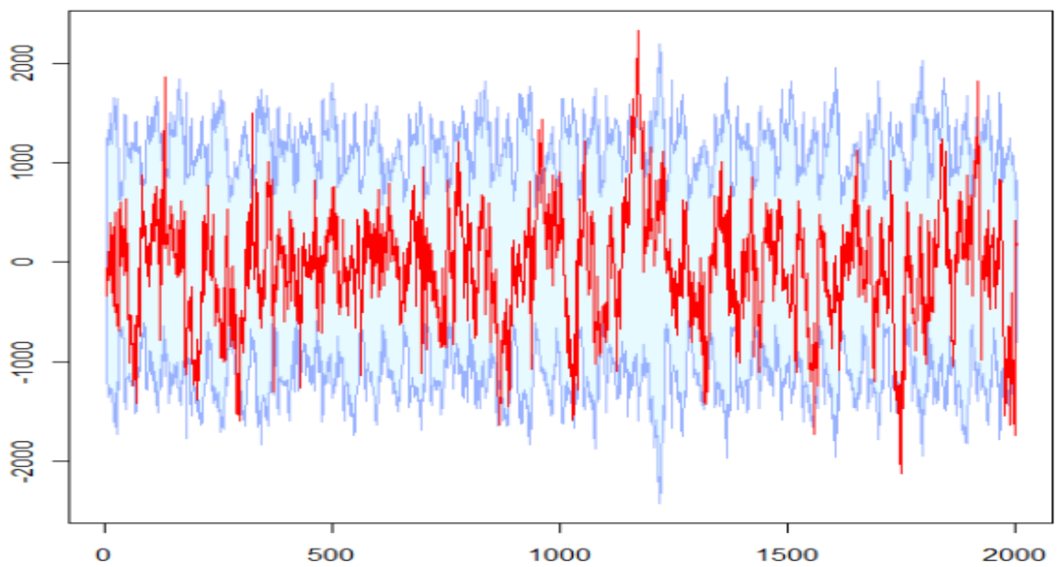
Quantile estimation: empirical quantile on the corrected residuals

Other approach: Kernel density estimation on clusters of residuals, AR models by clusters

Short term correction: Garch(1,1) on the short term (intraday) corrected residuals

STATISTICAL MODELS ON RESIDUALS

040809	3336BBF	336BBFArch	BNapr	BNaprArch
rmse	731	731	715	715
taux en %	7.8	6.1	8.3	6.9
largeur	2460	2444	2358	2322
tauxHaut en %	53	50	49	46
duree	3.23	2.52	3.11	2.53
intensite	658	487	591	461
varTauxIJ	7.5	1.5	7.4	2.2
asymetrieIS	10.9	18.5	4.5	15.6
giniJour	0.84	0.85	0.83	0.83



STATISTICAL MODELS ON RESIDUALS

- **Residuals on GAM models:**
  - A. Pierrot 2011, *Intervalle de prévision à partir des résidus d’estimation du modèle GAM France*, Technical Report.

- **GAM model on French Data:**

$$\boxed{\begin{aligned} \mathrm{P}_{j,h} = {} & f_1(\mathrm{P}_{j-1,h}).1_{day\_type(j)} + f_2(\mathrm{T}_{j,h}) + f_3(\mathrm{T}_{j-1,h}) + f_4(\mathrm{T}_{j-2,h}) \\ & + f_5(\underline{\mathrm{T}_{23_{j,h}}}) + f_6(\overline{\mathrm{T}_{23}}_{\,j,h}) + f_7(nebu_{j,h}) + f_8(posan_{j,h}).1_{day\_type2(j)} \\ & + \lambda_1.1_{day\_type(j)} + \lambda_2.trend_{j,h} + C \end{aligned}}$$

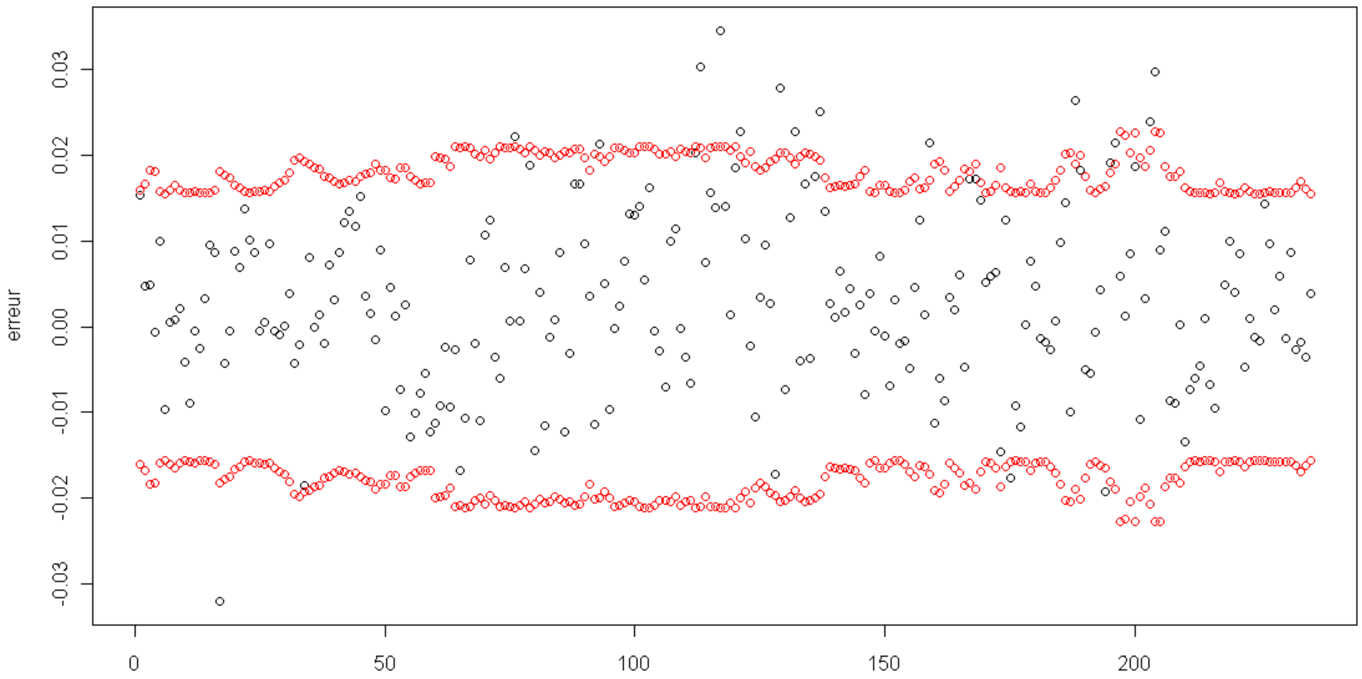
- **Test of different quantile estimation of the residuals:**
  - Bootstrap of the (standardized) residuals
    - Standardization: (hourly) GAM model on the quadratic residuals:

$$\varepsilon_t^2 = \alpha_1 \, \mathrm{I}_{summer} + \alpha_2 \, \mathrm{I}_{winter} + f(T_t) + \xi_t$$

  - Normality asumption with standart deviation by clusters (summer/winter)

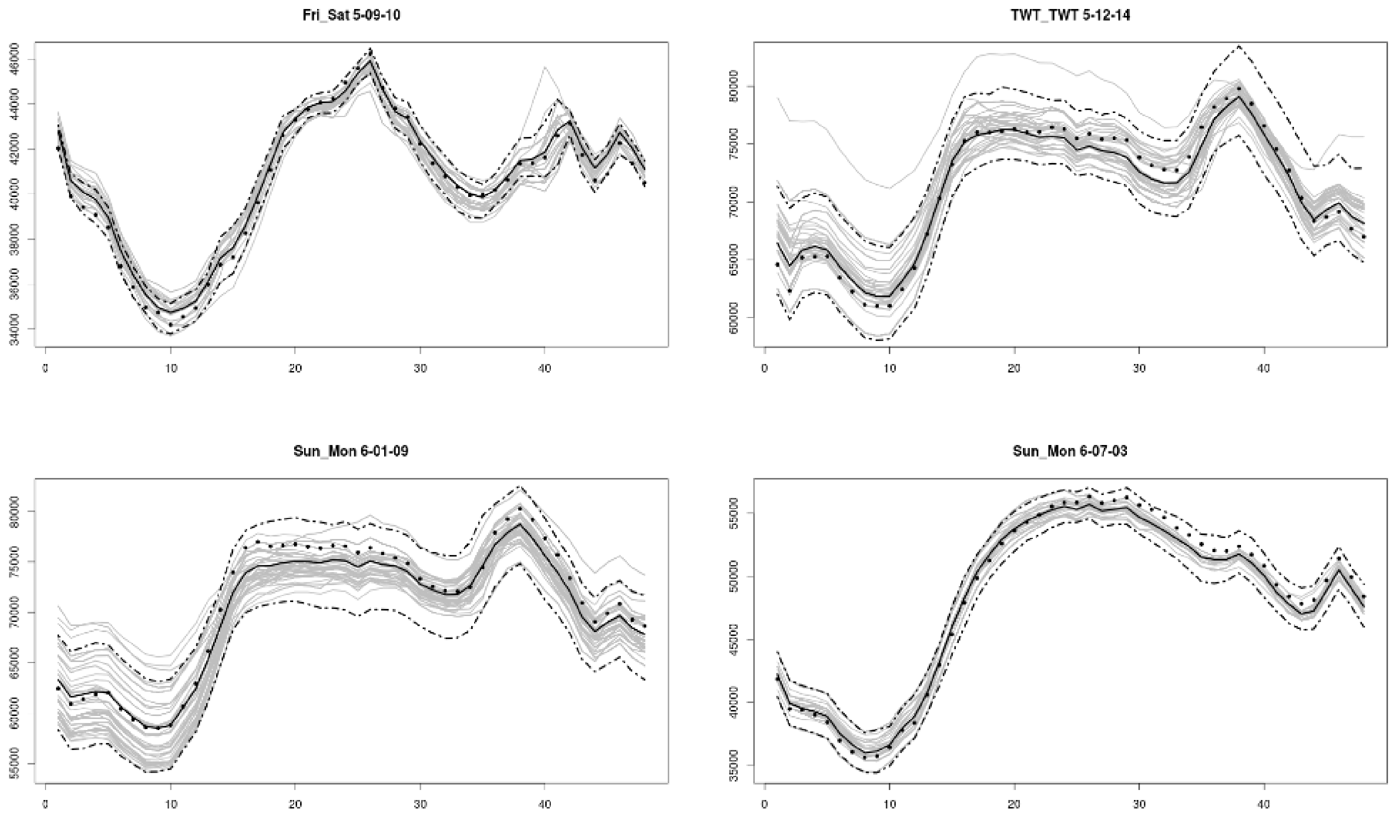
STATISTICAL MODELS ON RESIDUALS

Tx	7.7%
Width	2660



BOOTSTRAP KWF

- **Main (among many) idea: use similarity weights to resample load each day**

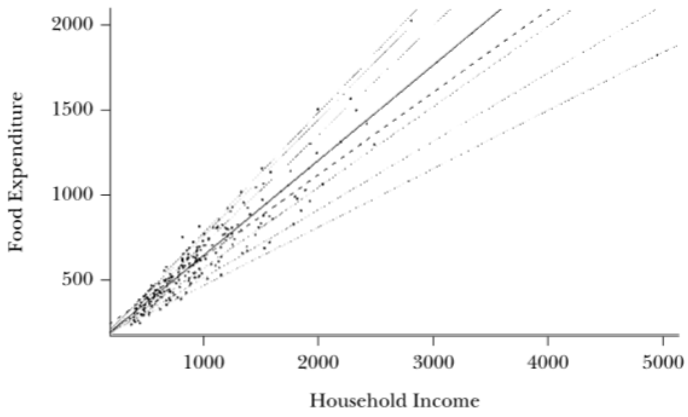


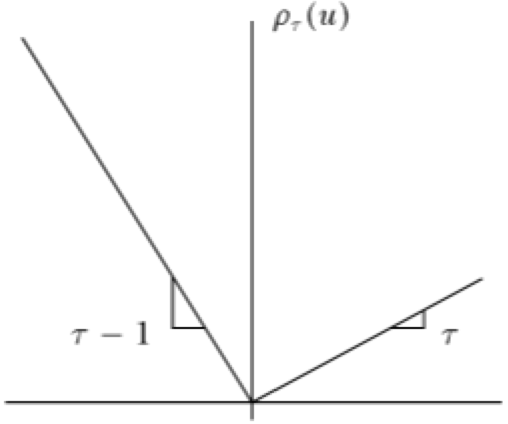
- A. Antoniadis, X. Brossat, J. Cugliari, J.-M. Poggi A prediction interval for a function-valued forecast model, submitted, 2014.

QUANTILE REGRESSION

- **2 studies at EDF R&D:**
  - J. Collet, 2006, *Prévision par intervalles de la consommation d'électricité nationale*, Tech. Report.
  - G. Agoua, R. Nédellec, 2014, *Prévision probabiliste de la consommation électrique,* Tech. Report.

- **Linear Quantile Regression:**
  - Solution of a linear programming problem (R.Koenker; G.Bassett, Jr., *Regression Quantile*, Econometrica, Vol. 46, No. 1. (Jan., 1978), pp. 33-50.)
  - Package R, (SAS)





$$\beta_\tau = \arg\, min_\beta \sum \rho_\tau(Y_i - X_i\beta)$$

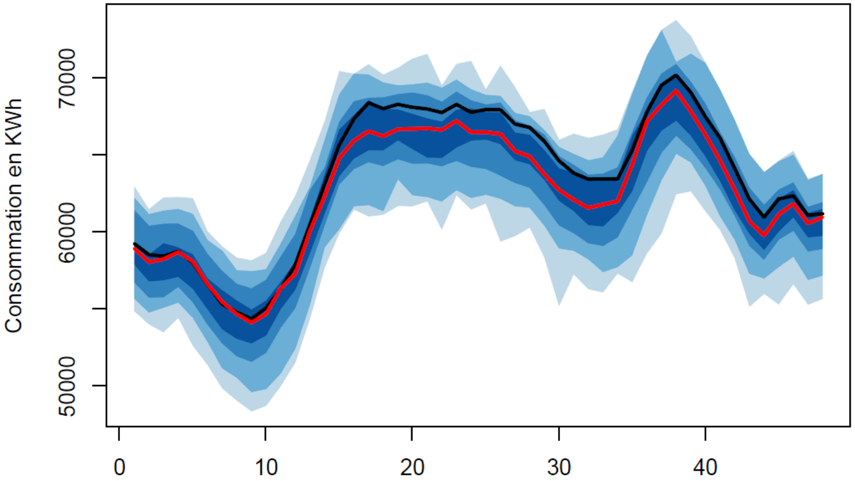
QUANTILE NON-LINEAR REGRESSION

- **G. Agoua and R. Nédellec:**
  - GAM model on the mean
  - Quantile estimation with linear adjustment: $\min_\beta \rho_\tau(Y - \beta f(X))$
- **Quantile Regression Forest:**
  - Machine learning method based on random forest: N. Meinshausen, *Quantile Regression Forests* Journal of Machine Learning Research 7 (2006), pp. 983-999.

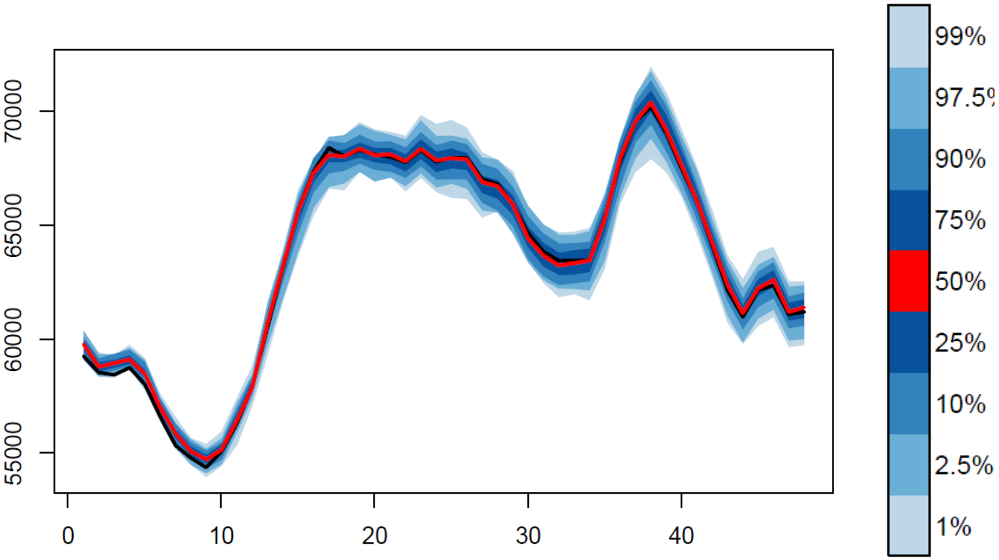
Critère	M1	M2	M3
$RMSE_{\text{prévision}}$	611.71	536.06	513.73
MAPE	1.08	0.95	0.90
largeur	2193.80	1926.75	1836.61
score moyen	173.09	151.12	144.35

Critère	QRF1	QRF2	QRF3
$RMSE_{\text{prévision}}$	1640.85	1530.25	1511.48
MAPE	2.71	2.43	2.42
largeur	7333.78	6965.80	6995.83
score moyen	336.91	302.67	299.40

QUANTILE NON-LINEAR REGRESSION



Quantile Regression Forest



GAM Quantile

AGGREGATION OF EXPERTS

- **a dynamic field of research in the machine learning community**
- **empirical literature is large and diverse**
- **massive development of new forecasting methods**
  - implementation in open source softwares
  - Easier access to a large variety of forecasts
- **aggregating them is a natural ambition**
- **in many recent forecasting challenges aggregation is a key point that often makes the difference:**
  - the energy forecasting competition GEFCOM12, Hong, T.; Pinson, P. & Fan, S. *Global Energy Forecasting Competition 2012* **International Journal of Forecasting** , 2014, 30, 357 - 363
  - netflix competition Paterek, *A. Predicting movie ratings and recommender systems* - **a monograph** 2012

$$\widehat{y}_t = \sum_{i=1}^{N} \widehat{p}_{i,t} x_{i,t} \qquad \widehat{p}_{k,t} = \frac{e^{-\eta \sum_{s=1}^{t-1} \ell_s(x_{k,s})}}{\sum_{i=1}^{K} e^{-\eta \sum_{s=1}^{t-1} \ell_s(x_{i,s})}}$$

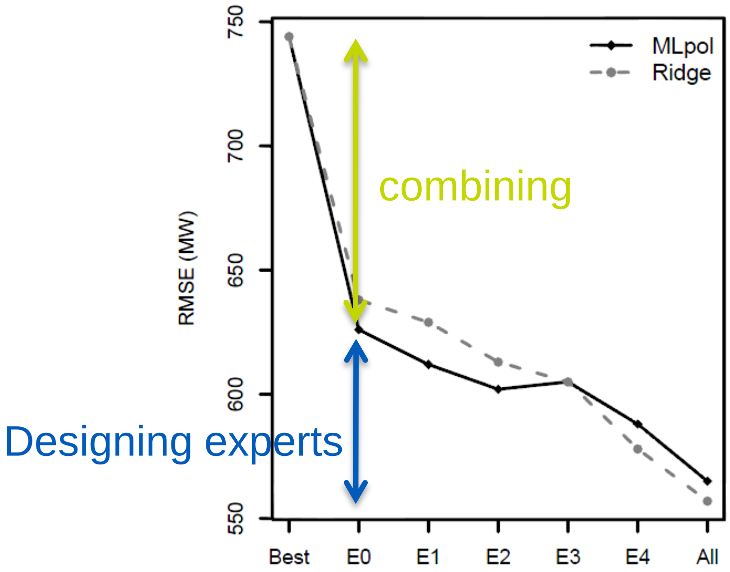
$$\ell_t(x_{k,t}) = (y_t - x_{k,t})^2 \longrightarrow \underbrace{\sum_{t=1}^{T} (\widehat{y}_t - y_t)^2}_{\text{Our loss}} = \underbrace{\min_{i=1,\dots,N} \sum_{t=1}^{T} (x_{i,t} - y_t)^2}_{\text{Loss of the best expert}} + \underbrace{R_T}_{\text{Estimation error}}$$

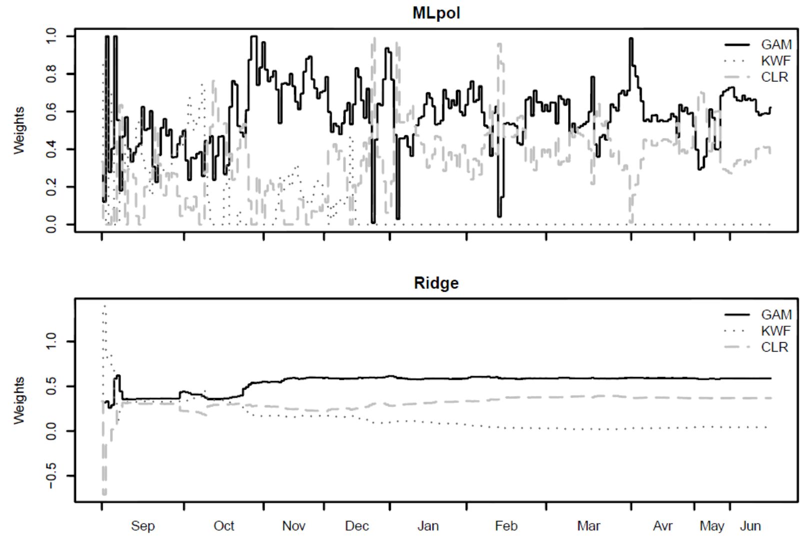
Good set of experts

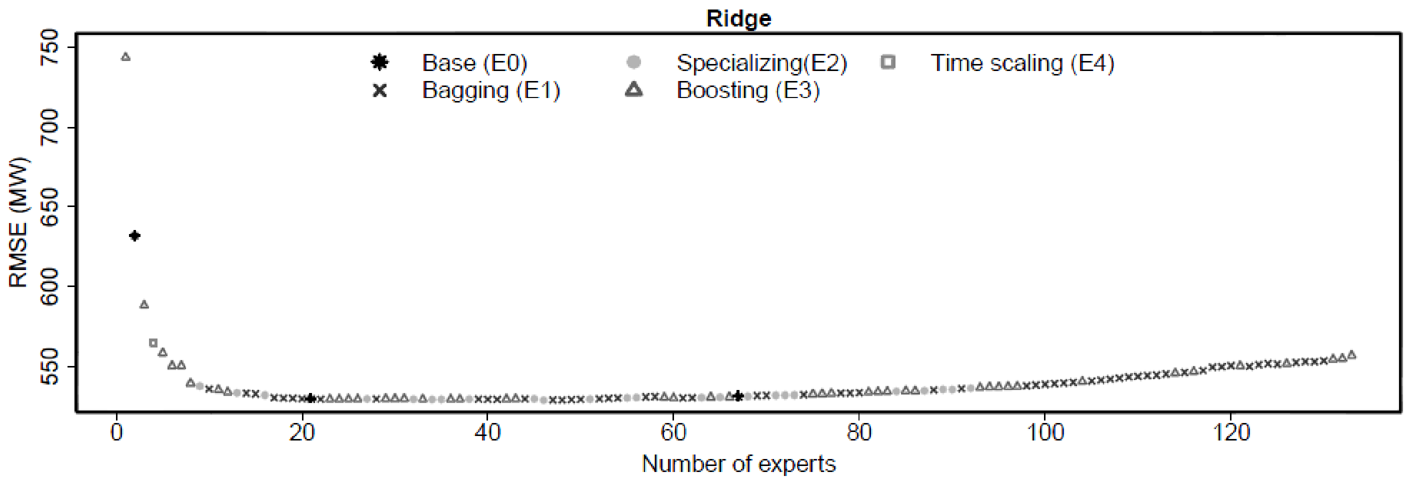
Good aggregating algorithm

- **Publications**
  - M. Devaine, P. Gaillard, Y. Goude & G. Stoltz, Forecasting electricity consumption by aggregating specialized experts - A review of the sequential aggregation of specialized experts, with an application to Slovakian and French country-wide one-day-ahead (half-)hourly predictions **Machine Learning**, 2013, 90, 231-260.
  - Gaillard, P. & Goude, Y., Forecasting electricity consumption by aggregating experts; how to design a good set of experts to appear in **Lecture Notes in Statistics: Modeling and Stochastic Learning for Forecasting in High Dimension**, 2013.

COMBINING FORECASTS





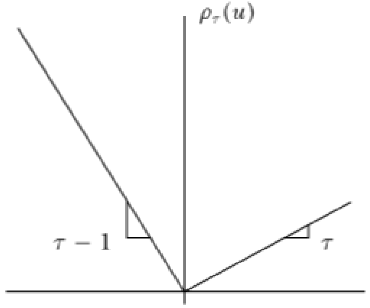


COMBINING FORECASTS FOR QUANTILE FORECASTING

$$\widehat{y}_t = \sum_{i=1}^{N} \widehat{p}_{i,t} x_{i,t}$$

$$\widehat{p}_{k,t} = \frac{e^{-\eta \sum_{s=1}^{t-1} \ell_s(x_{k,s})}}{\sum_{i=1}^{K} e^{-\eta \sum_{s=1}^{t-1} \ell_s(x_{i,s})}}$$

$$l_t(x_{k,t}) = \rho_\tau(y_t - x_{k,t}) \longrightarrow$$



- Consider a set of experts producing quantile forecasts, set the loss function to be the pinball loss results in quantile forecasting
- Application on load/price forecasting: work in progress with Pierre Gaillard (PhD student EDF)

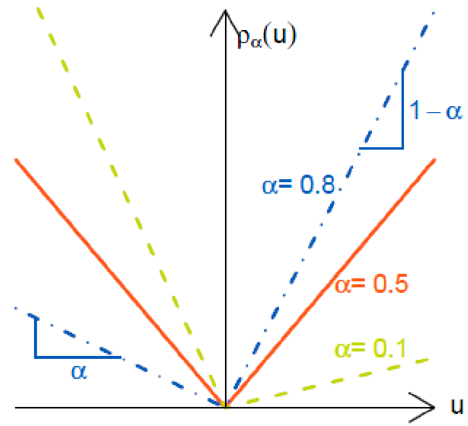
- Biau, G. & Patra, B. Sequential Quantile Prediction of Time Series**, IEEE Transactions on Information Theory**, 2011, 57, 1664-1674.

PROB. ELECTRICITY PRICE FORECASTING

GEFCOM14

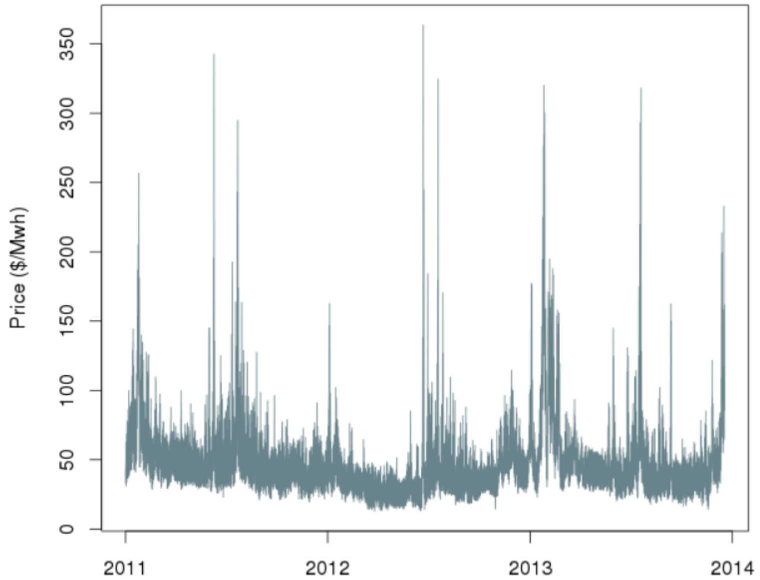
- **competition GEFCOM 2014, sponsored by IEEE Power and Energy Society**
  - Online forecasting of 15 days
  - Performance evaluation: pin-ball loss

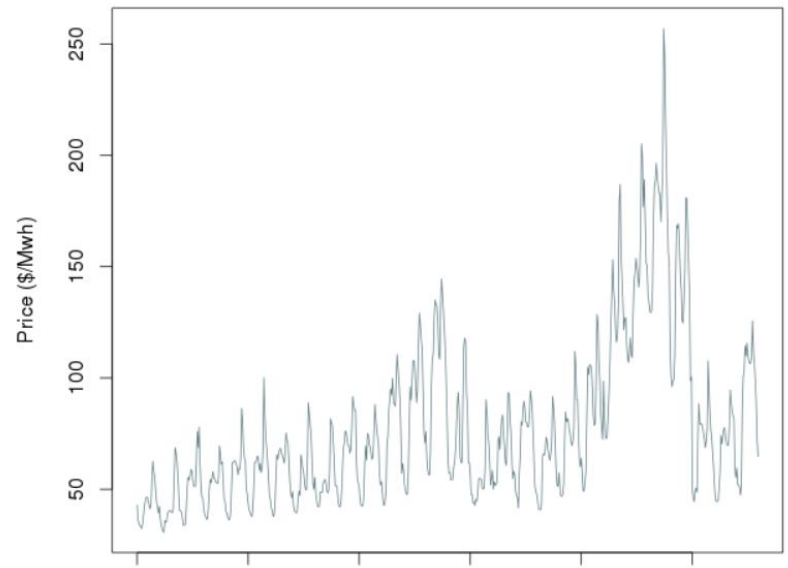
$$L(q_a, y) = \begin{cases} (1 - a/100)(q_a - y), & \text{if } y < q_a; \\ a/100(y - q_a), & \text{if } y \geq q_a; \end{cases}$$

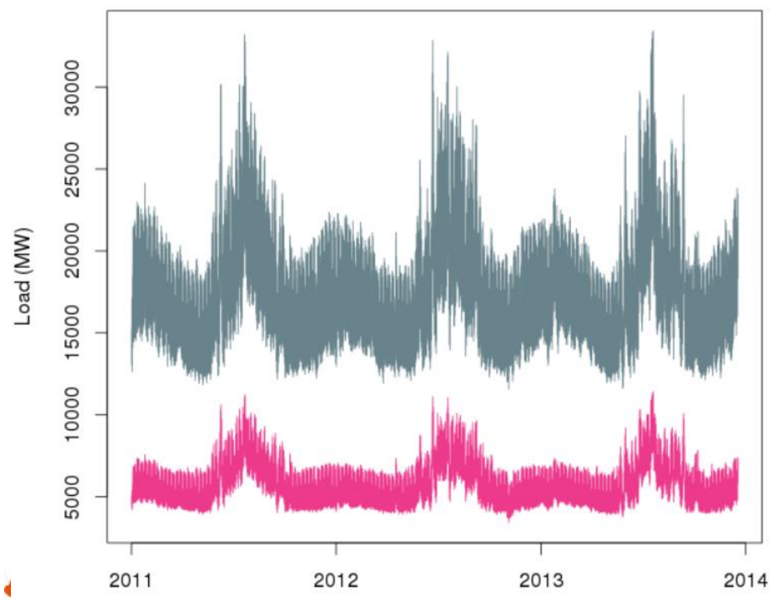


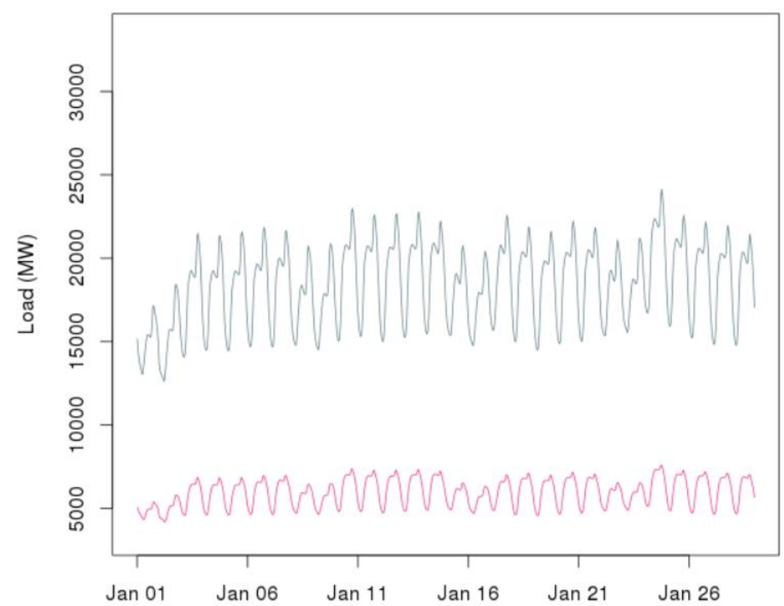
- **We proposed 3 methods:**
  - **Aggregation of 13 experts**
  - Non linear quantile regression: GAM quantile
  - Quantile Kernel lasso selection

ELECTRICITY PRICE DATA









PROB. ELECTRICITY PRICE FORECASTING

GEFCOM14

- **Aggregation of 13 experts:**
  - autoregressive model (AR)

$$\log(\mathrm{P}_t) = \alpha_1 \log(\mathrm{P}_{t-24}) + \alpha_2 \log(\mathrm{P}_{t-48}) + \alpha_3 \log(\mathrm{P}_{t-168}) + \alpha_4 \log(\mathrm{P.min}_{t-24}) + h(\mathrm{DayType}_t) + \varepsilon_t$$

  - An autoregressive model with forecasted electric loads as additional covariates (ARX).
  - A threshold autoregressive model TAR defined as an extension of AR to two regimes depending of the variation of the mean price between a day and eight days ago.
  - TARX the extension of ARX to the two regimes model.
  - Spike pre-processed autoregressive model PAR
  - PARX similar to PAR, but ARX is fitted with pre-processed prices.

  inspire from Weron, R., Misiorek, A., 2008. *Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models*. **International Journal of Forecasting** 24 (4), 744 – 763

  - 2 linear regressions
  - 2 GAMS
  - 2 random forests
  - GBM

$$\log(\mathrm{P}_t) = \alpha_1 \log(\mathrm{P}_{t-24}) + \alpha_2 \log(\mathrm{P}_{t-48}) + \alpha_3 \log(\mathrm{P}_{t-168}) + \alpha_4 \log(\mathrm{P.max}_t) + \alpha_5 \mathrm{FZL}_t^{(0.95)} + \alpha_6 \mathrm{FTL}_t^{(0.95)} + \alpha_7 \mathrm{FZL}_t^{(0.8)} + \alpha_8 \mathrm{FTL}_t^{(0.8)} + h(\mathrm{DayType}_t) + \varepsilon_t$$

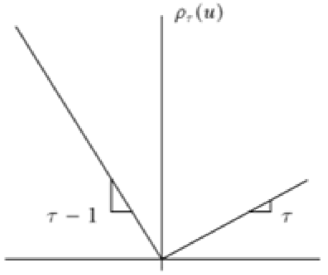
PROB. ELECTRICITY PRICE FORECASTING

GEFCOM14

- **(Convex) Aggregation with pin-ball loss:**

$$\widehat{y}_t = \sum_{i=1}^{N} \widehat{p}_{i,t} x_{i,t} \qquad \widehat{p}_{k,t} = \frac{e^{-\eta \sum_{s=1}^{t-1} \ell_s(x_{k,s})}}{\sum_{i=1}^{K} e^{-\eta \sum_{s=1}^{t-1} \ell_s(x_{i,s})}}$$

$$l_t(x_{k,t}) = \rho_\tau(y_t - x_{k,t}) \longrightarrow$$

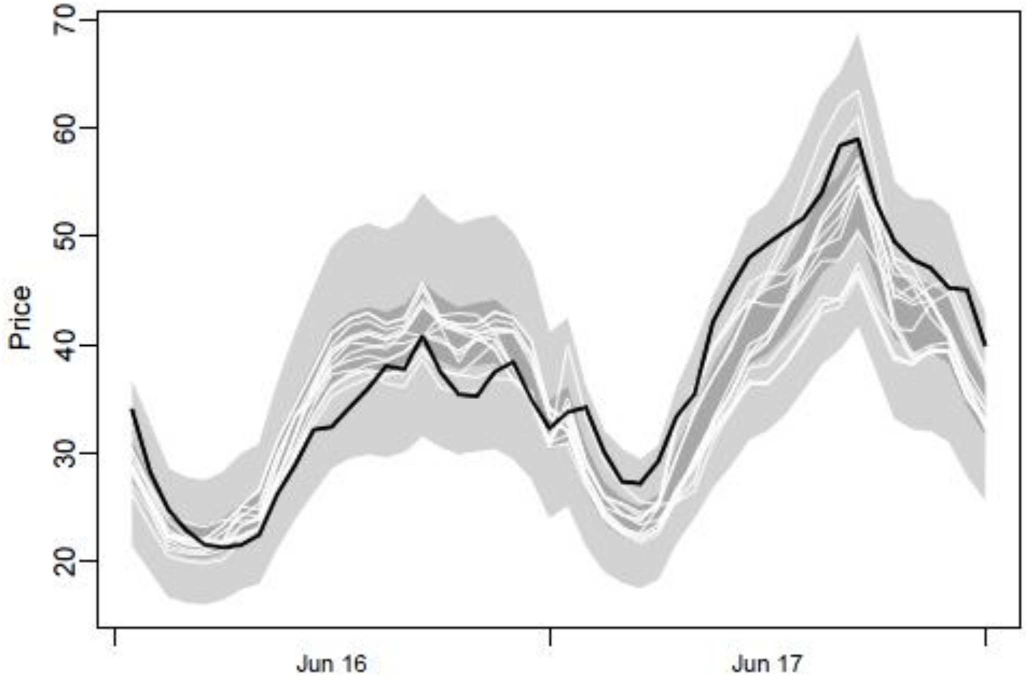


- **Extension to linear aggregation:**
  - substitute to original experts $\beta x_{1,t}, \ldots, \beta x_{K,t}, -\beta x_{1,t}, \ldots, -\beta x_{K,t}$

PROB. ELECTRICITY PRICE FORECASTING

GEFCOM14

- **Results**



PROB. ELECTRICITY PRICE FORECASTING

GEFCOM14

- **Results: 1st rank of the competition**

Task	TOLOLO	quantGAM	quantMixt	quantGLM
Jun. 06	XX	0.72	0.85	1.87
Jun. 17	1.06	1.15	1.37	0.71
Jun. 24	1.91	1.31	1.58	3.05
Jul. 04	1.71	2.06	1.27	1.59
Jul. 09	1.45	0.99	3.31	1.57
Jul. 13	1.10	2.23	1.20	1.18
Jul. 16	2.01	2.63	2.28	5.02
Jul. 18	9.15	5.13	7.90	11.72
Jul. 19	4.68	4.80	6.45	13.27
Jul. 20	1.59	1.90	2.35	2.80
Jul. 24	0.75	0.75	1.78	1.42
Jul. 25	2.46	2.30	0.84	2.12
Dec. 06	2.96	0.82	1.03	0.86
Dec. 07	1.35	3.63	3.23	3.22
Dec. 17	3.56	3.83	4.26	2.87

	Load		Price
Ranking	Team	Rating	Team
1	Tololo	50,0%	Tololo
2	Adada	49,0%	Team Poland
3	Jingrui (Rain) Xie	48,0%	GMD
4	OxMath	47,6%	C3 Green Team
5	E.S. Mangalova	45,4%	pat1

CONCLUSION

- **Quantile GAM provides a unify and quite efficient way to derive quantile forecasts from our models in operation**
- **Improvement are coming from agregation of experts (specially on short horizon)**
  - combining « specialized » GAM for proba. forecasts (benchmark on the GEFCOM14)
- **Still a lot of work to do:**
  - Mid-term horizon: how can we merge model uncertainty and weather scenarios/portfolio evolution?
  - Short-term: **work on real quantile GAM**, not only linear adjustement of the median model
  - Local forecasting
  - Focus on **intraday forecasts**
  - **Improve machine learning method** (Random forest, gradient boosting machine)
  - Net demand: connexion with wind/solar power probabilistic forecasts (work on gefcom 2014 data?)
  - **Adaptive** (with time) forecasts to deal with breaks, data flow (related to BAM)

THANKS!