

# CONTRIBUTIONS À L'AGRÉGATION SÉQUENTIELLE ROBUSTE D'EXPERTS

---

Pierre Gaillard

Prix de thèse AMIES 2016

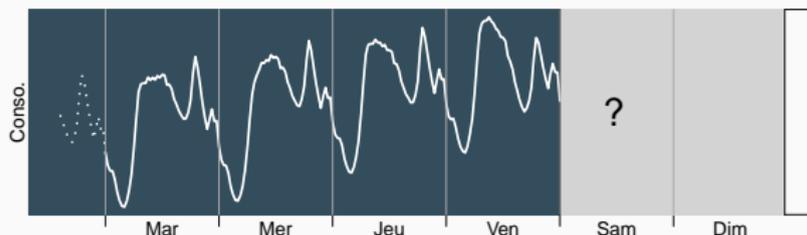
Thèse CIFRE en partenariat entre

- EDF R&D : département OSIRIS
- CNRS : laboratoire GREGHEC à Jouy-en-Josas



Encadrée par Gilles Stoltz (directeur) et Yannig Goude (responsable industriel).

Prévision à court terme (1 jour à l'avance) de la consommation électrique française



C'est un enjeu majeur pour EDF car l'électricité se stocke mal.

L'équilibre entre l'offre et la demande doit être maintenu à tout moment pour éviter

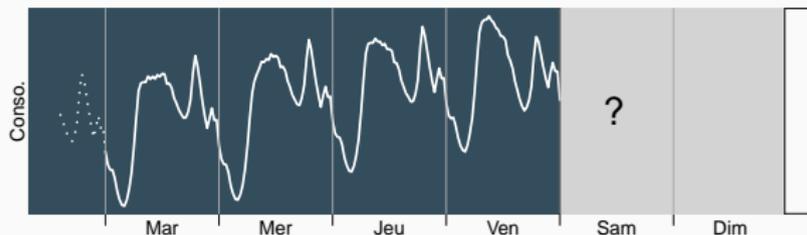
- les risques physiques : coupure d'électricité, reconfiguration de réseau
- les risques financiers.

Production



Demande

Prévision à court terme (1 jour à l'avance) de la consommation électrique française



De nombreux experts construisent des modèles :

GAM



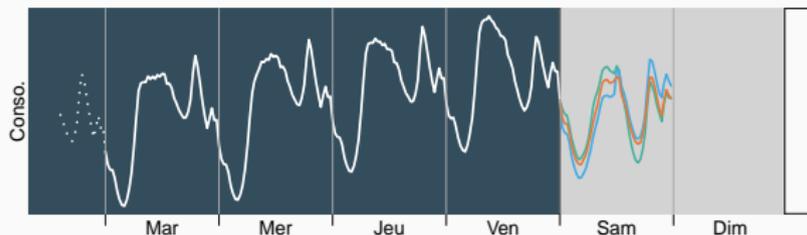
KWF



CLR



Prévision à court terme (1 jour à l'avance) de la consommation électrique française



De nombreux experts construisent des modèles :

GAM



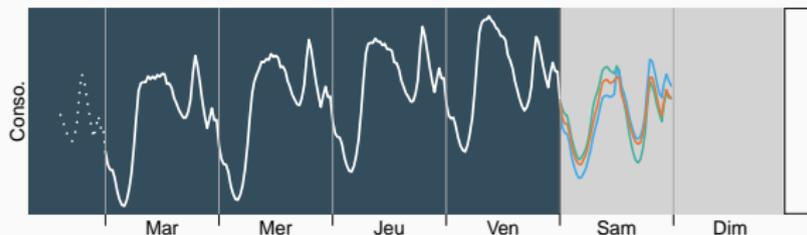
KWF



CLR



Prévision à court terme (1 jour à l'avance) de la consommation électrique française



De nombreux experts construisent des modèles :

GAM



KWF



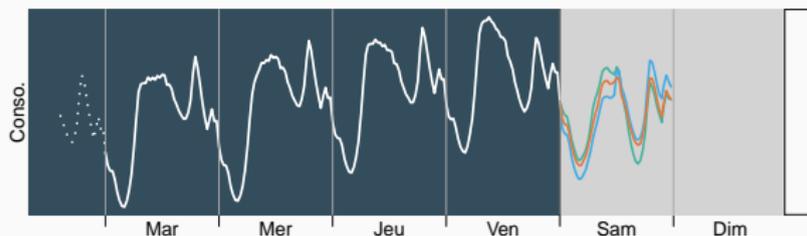
CLR



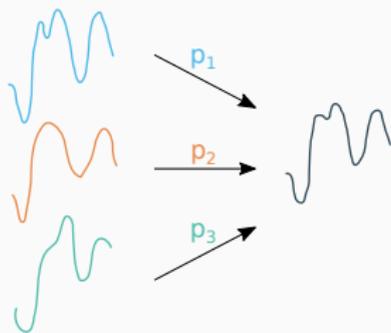
Parallèlement, le paysage électrique français évolue (voitures électriques,...)  
→ Remise en question des modèles historiques.

Quel modèle utiliser ? → les mélanger de façon adaptative.

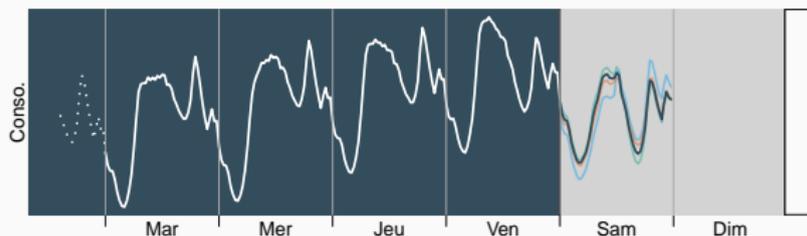
Prévision à court terme (1 jour à l'avance) de la consommation électrique française



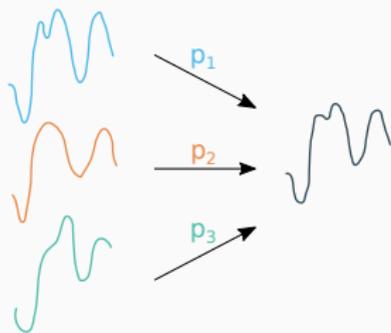
Mélanger les prévisions de façon adaptatives :



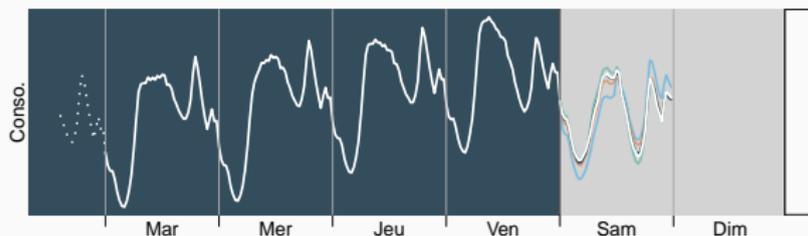
Prévision à court terme (1 jour à l'avance) de la consommation électrique française



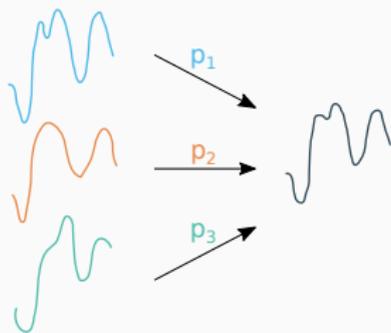
Mélanger les prévisions de façon adaptatives :



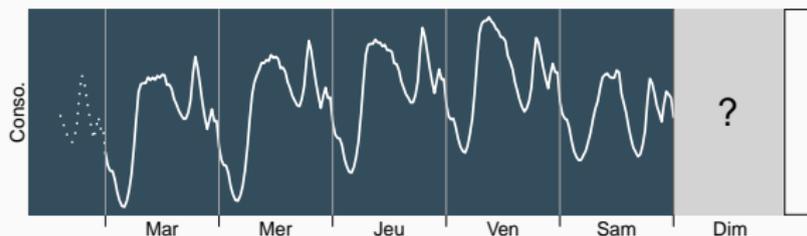
Prévision à court terme (1 jour à l'avance) de la consommation électrique française



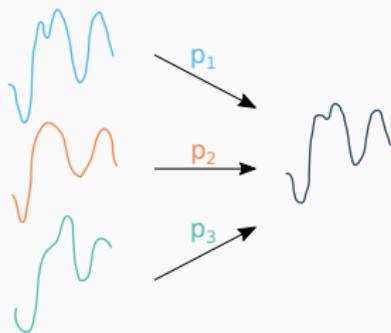
Mélanger les prévisions de façon adaptatives :



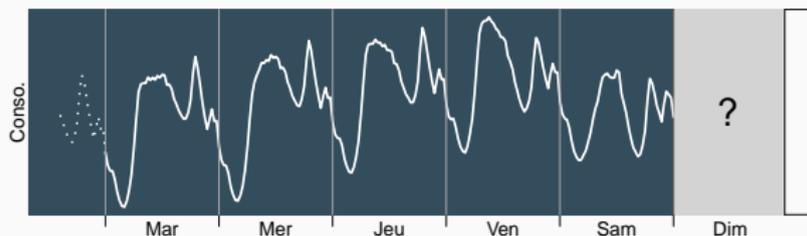
Prévision à court terme (1 jour à l'avance) de la consommation électrique française



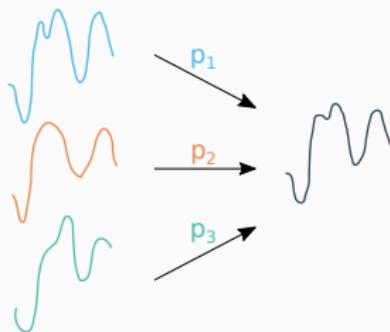
Mélanger les prévisions de façon adaptatives :



Prévision à court terme (1 jour à l'avance) de la consommation électrique française



Mélanger les prévisions de façon adaptatives :



Les poids sont mis à jour chaque jour en fonction des observations passées.

**Objectif :** se rapprocher du meilleur modèle sur le long terme

# CONTRIBUTION 1 : GARANTIES THÉORIQUES + PERFORMANCE PRATIQUE

Les modèles à mélanger reposent sur des hypothèses très différentes les uns des autres.

Réunion sous un même  
cadre difficile



→ Cadre de la prévision de suites arbitraires

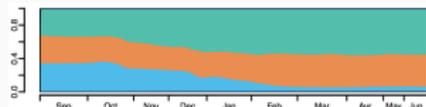


- “pas” d’hypothèse
- **robustesse** : garanties théoriques valides quoi-qu’il arrive

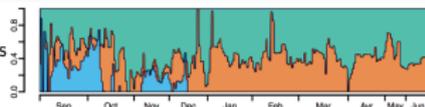


- Souvent trop conservateur et peu réaliste
- mauvaise performance des algorithmes théoriques

**Solution** : l’algorithme **ML-Poly** présente des garanties robustes mais s’adaptant à la régularité des données



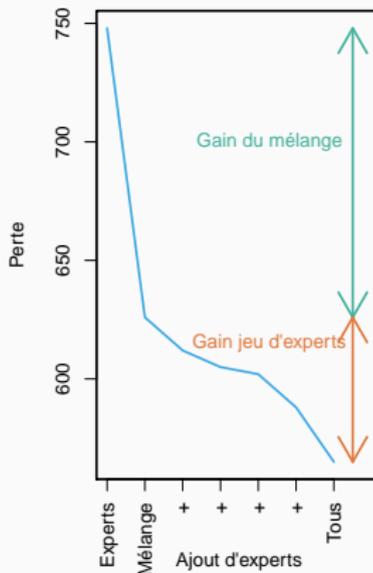
← Exp. weights  
(theory)



← ML-Poly  
(theory)

De nombreuses problématiques industrielles :

- **Non-stationnarité** : faire face aux ruptures dans les données ?
- **Information exogène** : prendre en compte la météo, l'heure de la journée, ... ?
- **Non-linéarité** : mélange non-linéaire ?
- **Jeu d'experts** : optimiser l'ensemble d'experts à inclure dans le mélange ?
- **Horizon de prévision**

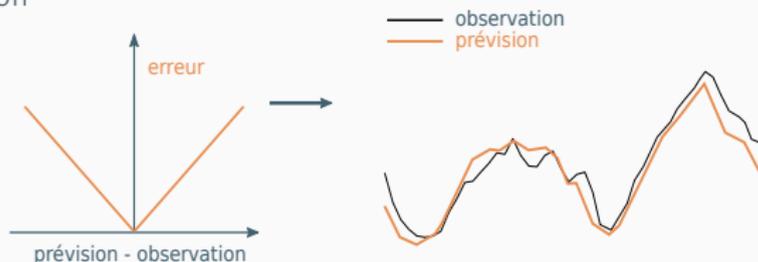


Librairie **opera** en R

Peut-on se fier à notre prévision ? Les énergies renouvelables apportent une part d'aléa.

→ **Prévision probabiliste** : prévoir la distribution possible de la consommation plutôt que juste sa moyenne.

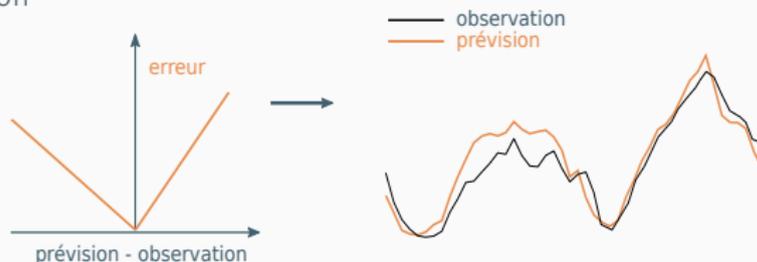
**Solution** : modifier la fonction qui mesure l'écart entre notre prévision et la consommation



Peut-on se fier à notre prévision ? Les énergies renouvelables apportent une part d'aléa.

→ **Prévision probabiliste** : prévoir la distribution possible de la consommation plutôt que juste sa moyenne.

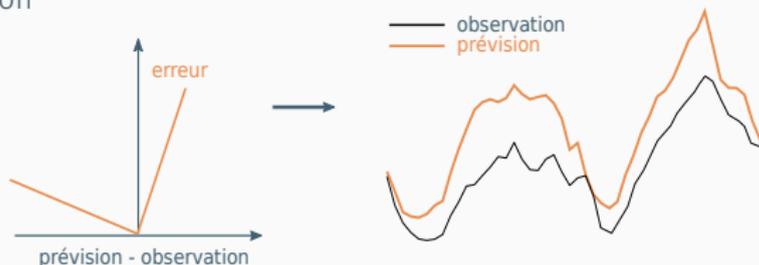
**Solution** : modifier la fonction qui mesure l'écart entre notre prévision et la consommation



Peut-on se fier à notre prévision ? Les énergies renouvelables apportent une part d'aléa.

→ **Prévision probabiliste** : prévoir la distribution possible de la consommation plutôt que juste sa moyenne.

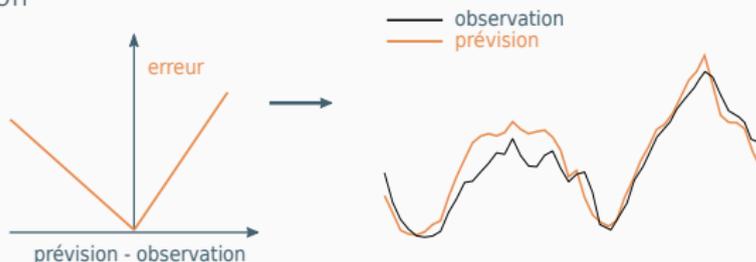
**Solution** : modifier la fonction qui mesure l'écart entre notre prévision et la consommation



Peut-on se fier à notre prévision ? Les énergies renouvelables apportent une part d'aléa.

→ **Prévision probabiliste** : prévoir la distribution possible de la consommation plutôt que juste sa moyenne.

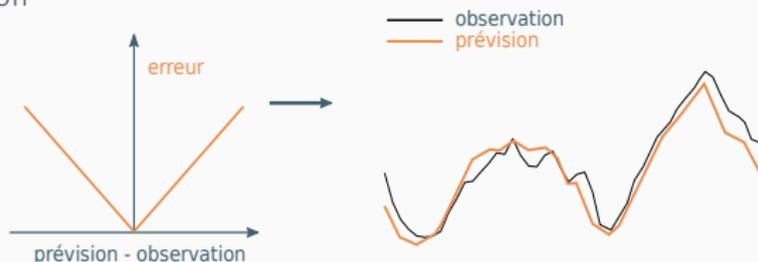
**Solution** : modifier la fonction qui mesure l'écart entre notre prévision et la consommation



Peut-on se fier à notre prévision ? Les énergies renouvelables apportent une part d'aléa.

→ **Prévision probabiliste** : prévoir la distribution possible de la consommation plutôt que juste sa moyenne.

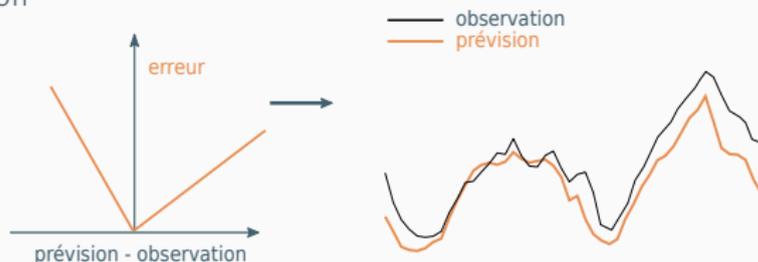
**Solution** : modifier la fonction qui mesure l'écart entre notre prévision et la consommation



Peut-on se fier à notre prévision ? Les énergies renouvelables apportent une part d'aléa.

→ **Prévision probabiliste** : prévoir la distribution possible de la consommation plutôt que juste sa moyenne.

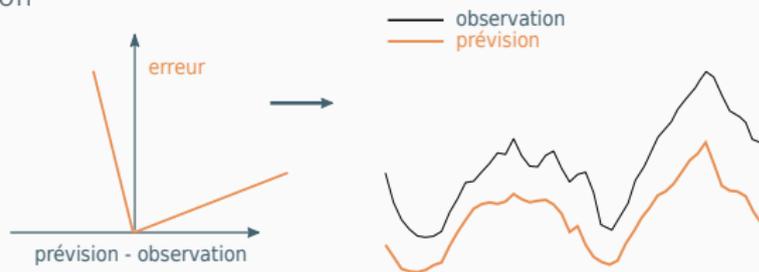
**Solution** : modifier la fonction qui mesure l'écart entre notre prévision et la consommation



Peut-on se fier à notre prévision ? Les énergies renouvelables apportent une part d'aléa.

→ **Prévision probabiliste** : prévoir la distribution possible de la consommation plutôt que juste sa moyenne.

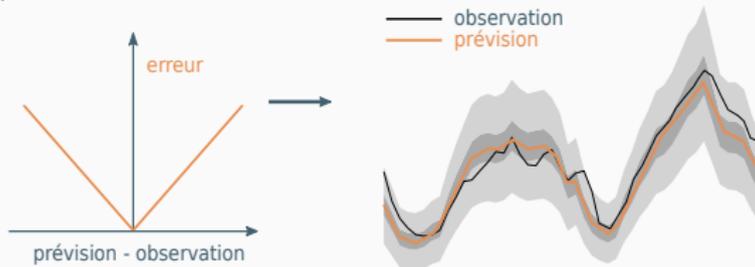
**Solution** : modifier la fonction qui mesure l'écart entre notre prévision et la consommation



Peut-on se fier à notre prévision ? Les énergies renouvelables apportent une part d'aléa.

→ **Prévision probabiliste** : prévoir la distribution possible de la consommation plutôt que juste sa moyenne.

**Solution** : modifier la fonction qui mesure l'écart entre notre prévision et la consommation



L'arrivée des compteurs Linky apporte de nouvelles perspectives industrielles :

- 35 millions de consommations individuelles
- gros volume de données disponible en "temps réel"



Développement de méthodes de mélange séquentielles et automatiques permettant

- de traiter des données en grande dimension
- de faire de la prévision spatiale

Développement d'une boîte à outils et de la librairie opera.

# MERCI !

## REFERENCES

---



N. CESA-BIANCHI, P. GAILLARD, G. LUGOSI et G. STOLTZ. “Mirror Descent Meets Fixed Share (and feels no regret)”. In : *Proceedings of NIPS’12*. 2012, p. 989–997.



M. DEVAINE, P. GAILLARD, Y. GOUDE et G. STOLTZ. “Forecasting the electricity consumption by aggregating specialized experts”. In : *Machine Learning* 90.2 (2013), p. 231–260.



M. FAURE, P. GAILLARD, B. GAUJAL et V. PERCHET. “Online Learning and Game Theory. A Quick Overview with recent results and applications”. In : *ESAIM : Proceedings*. Sous la dir. d’A. GARIVIER ET AL. To appear. EDP Sciences, 2015.



P. GAILLARD. *opera : Online Prediction by ExpeRts Aggregation*. R package version 0.01. 2015. URL : <https://github.com/Dra1liag/opera.git>.



P. GAILLARD et S. GERCHINOVITZ. “A Chaining Algorithm for Online Nonparametric Regression”. In : *Proceedings of COLT’15*. T. 40. JMLR : Workshop et Conference Proceedings, 2015, p. 764–796.



P. GAILLARD, Y. GOUDE et R. NEDELLEC. “Semi-parametric models and robust aggregation for GEFCom2014 probabilistic electric load and electricity price forecasting”. *International Journal of Forecasting*. To appear. 2015.



P. GAILLARD. “Contributions à l’agrégation séquentielle robuste d’experts : travaux sur l’erreur d’approximation et la prévision en loi. Applications à la prévision pour les marchés de l’énergie”. Thèse de doct. Université Paris-Sud 11, 2015.



P. GAILLARD et P. BAUDIN. "A consistent deterministic regression tree for non-parametric prediction of time series". arXiv :1405.1533. 2015.



P. GAILLARD et Y. GOUDE. "Forecasting electricity consumption by aggregating experts ; how to design a good set of experts". In : *Modeling and Stochastic Learning for Forecasting in High Dimensions*. Sous la dir. d'A. ANTONIADIS, X. BROSSAT et J.-M. POGGI. T. 217. Lecture Notes in Statistics. Springer, 2015, p. 95–115.



P. GAILLARD, G. STOLTZ et T. van ERVEN. "A Second-order Bound with Excess Losses". In : *Proceedings of COLT'14*. T. 35. JMLR : Workshop et Conference Proceedings, 2014, p. 176–196.