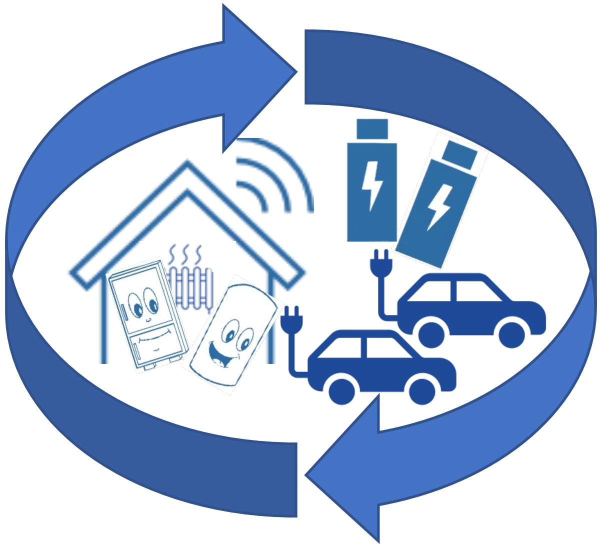
EDF R&D

Ateliers Fime 13 et 14 septembre 2023

Session "Distributed optimization : Machine learning & Mean Field games"

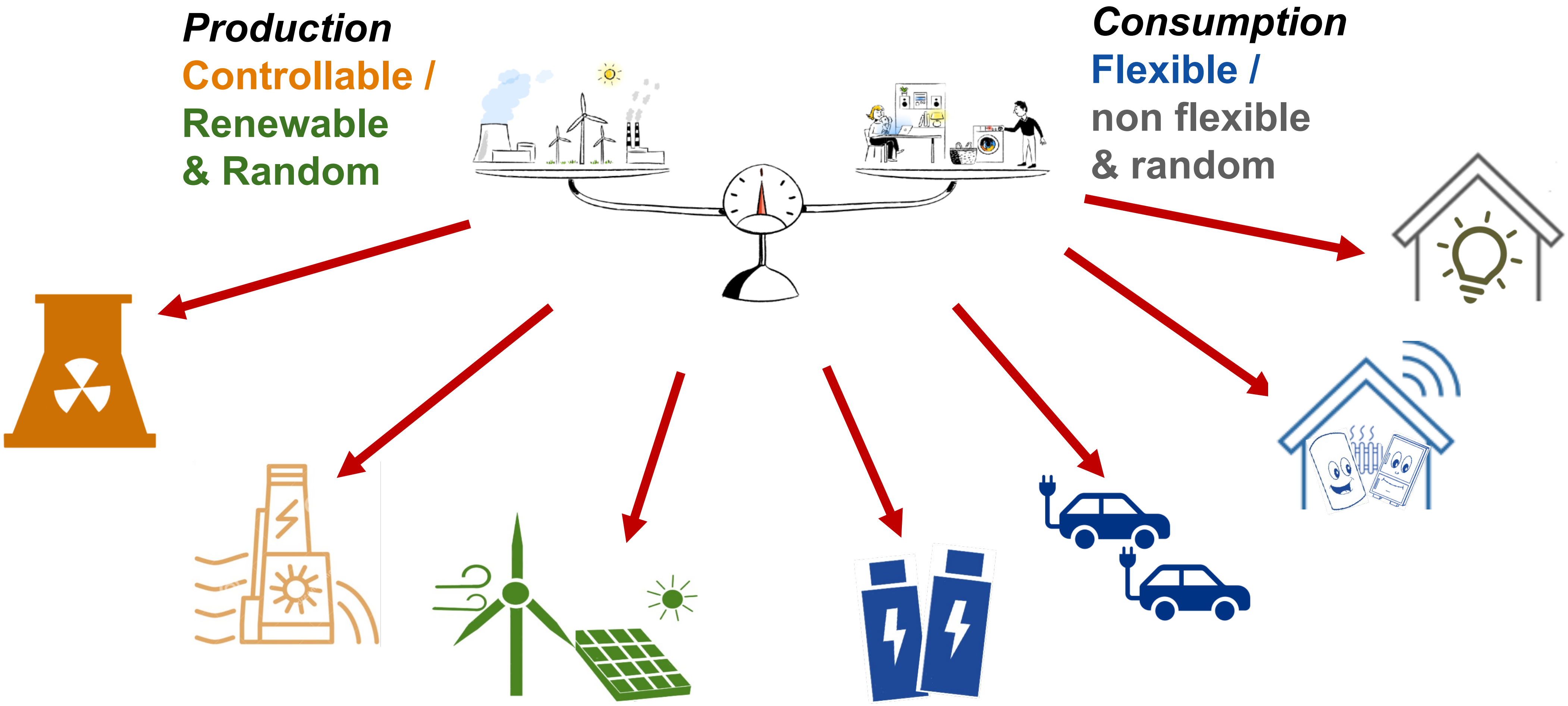
*Forecast & **control** in decentralized electrical systems*



*Yannig Goude and **Nadia Oudjane** - Ateliers Fime 13 septembre 2023*

CONTROL OF FLEXIBILITIES IN NEW ELECTRICAL SYSTEMS

Goal : balance production & consumption while **minimizing costs** at each instant



FORMULATION AND ISSUES

- Goal

Control= Controllable Production or - Flexible consumption

$$\min_{\forall i \ u_i \in U_i} J(u)$$

Controlled state

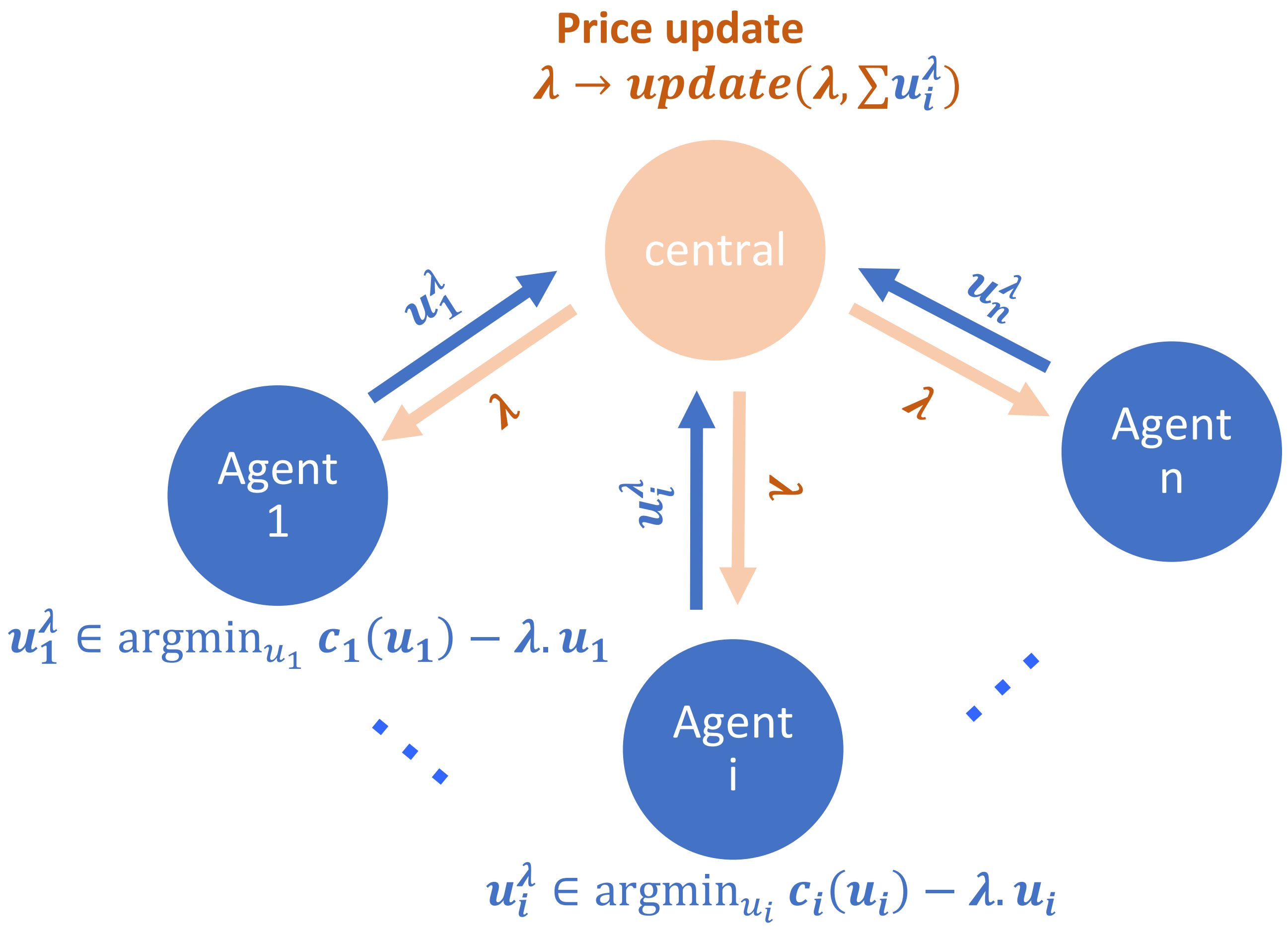
$$J(u) = \underbrace{E[c_0(\frac{1}{n}\sum_{i=1}^{n} u^i)]}_{\textit{Global cost}} + \frac{1}{n}\sum_{i=1}^{n} \underbrace{E[c_i\left(u^i, X^{i,u^i}\right)]}_{\textit{Local costs}}$$

- Issues
  - very large scale
  - heterogeneous population
  - unknown local models or data
  - non convex
  - stochastic
  - privacy issues

DISTRIBUTED APPROACH

Franck Wolfe, Projected Subgradient, Proximal Gradient, Alternating minimization, Lagrangian decomposition, …

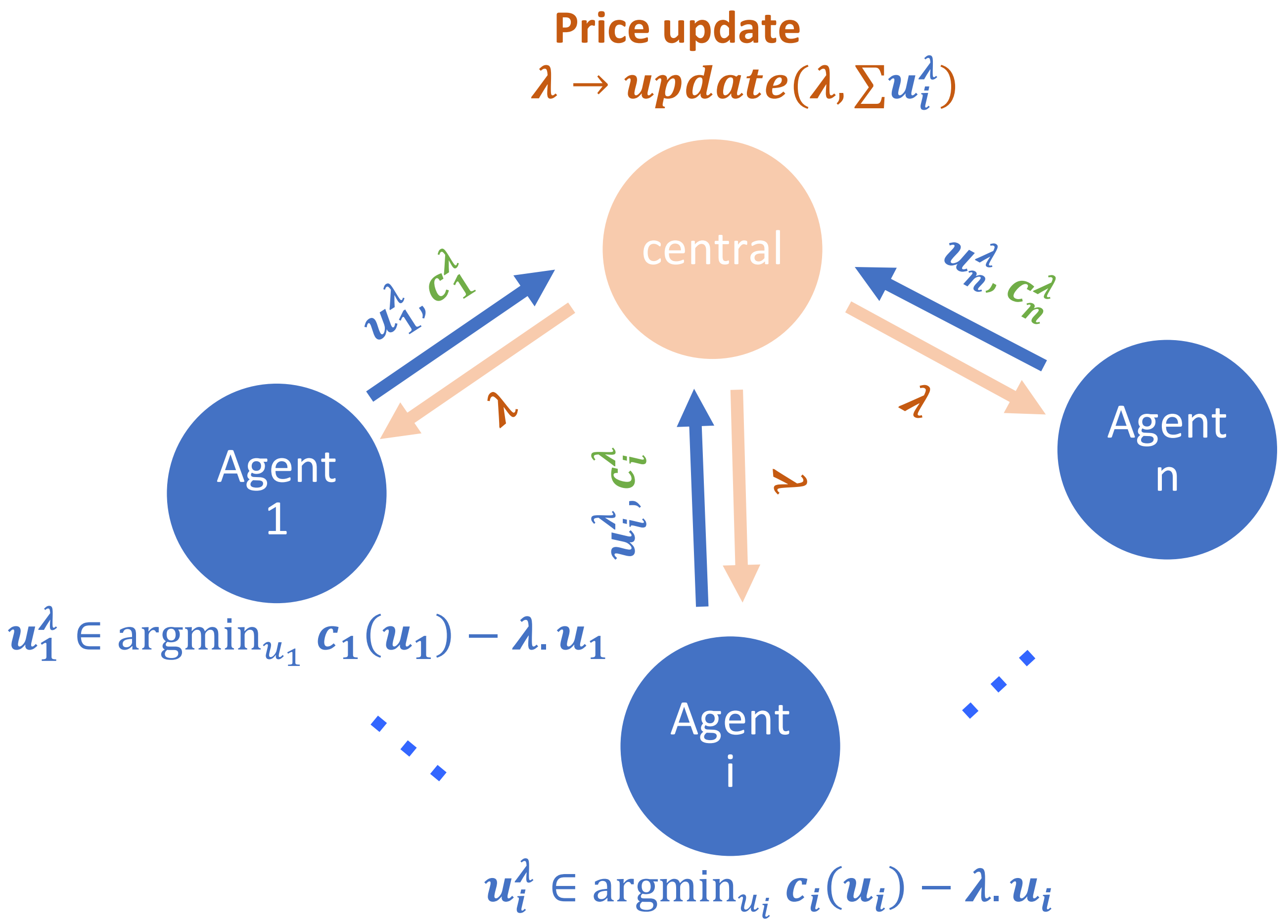
Example : Lagrangian decomposition in a ***deterministic & convex*** *setting*



DISTRIBUTED APPROACH

Franck Wolfe, Projected Subgradient, Proximal Gradient, Alternating minimization, Lagrangian decomposition, …

Example : Lagrangian decomposition in a ***deterministic & convex*** *setting*

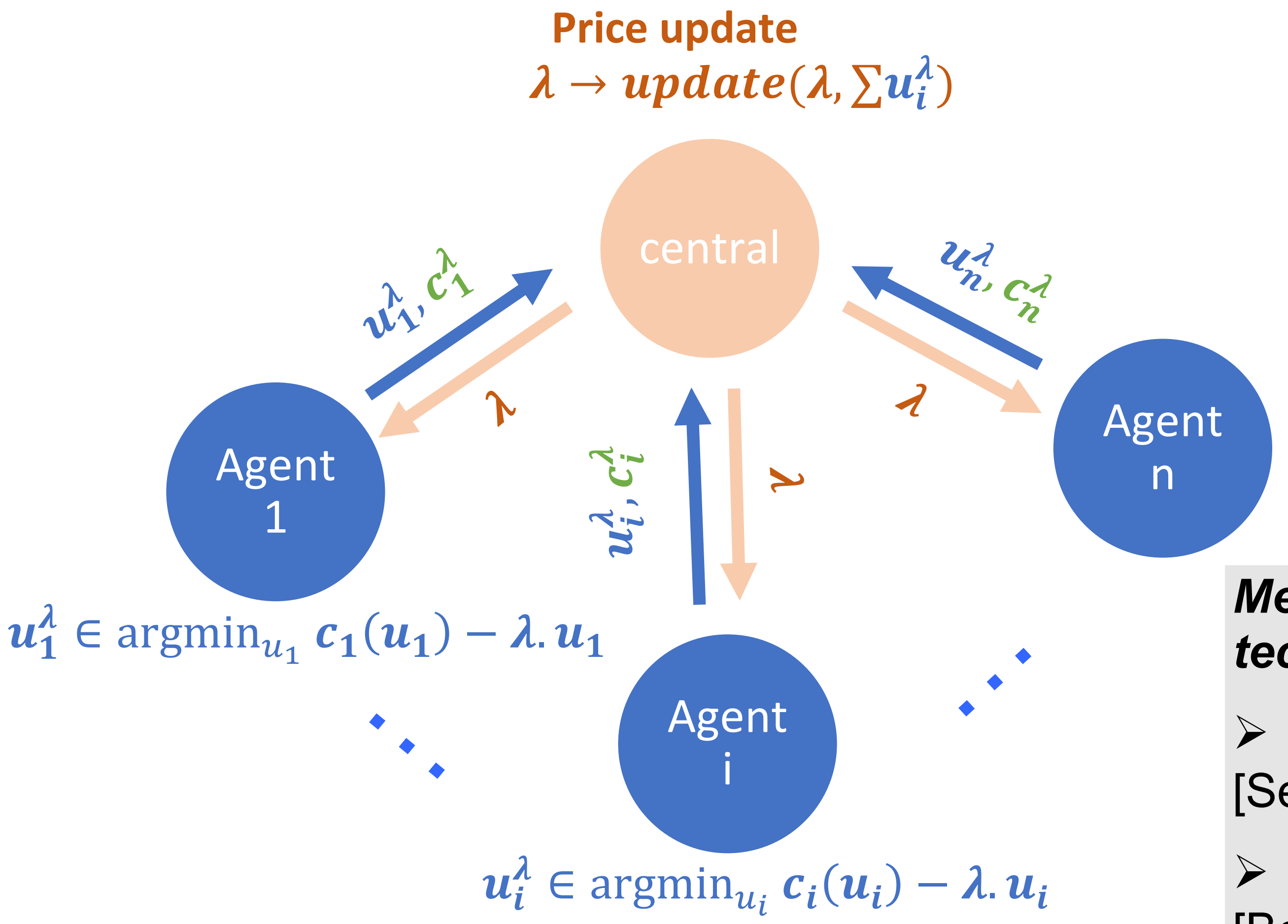


- **Many interactions (iterations)**
- **Computationally costly at the local level**
- **Preserving local models privacy**
- **Revealing local profiles or gradients**

DISTRIBUTED APPROACH

Franck Wolfe, Projected Subgradient, Proximal Gradient, Alternating minimization, Lagrangian decomposition, …

Example : Lagrangian decomposition in a ***deterministic & convex*** *setting*



- **Many interactions (iterations)**
- **Computationally costly at the local level**
- **Preserving local models privacy**
- **Revealing local profiles or gradients**

Mean-field approx to extend distributed otpimization techniques to

- Large scale convex **stochastic control** problems [SeguretEtal2023]
- Large scale **non convex** problems [BonnansEtal2022]

AGGREGATED APPROACH (IN TWO STEPS)

- **Low level of uncertainties on the aggregate**

Target profile

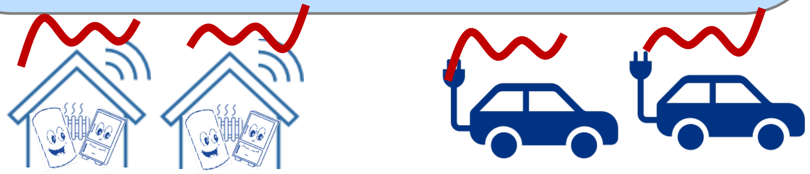
- **High level of uncertainties on each local agent**

1) Optimization

an aggregated model of flexibilities $\mathcal{M}$ to generate a target profile for the aggregate

2) Real time dispatch

Disaggregation of the target profile among the local agents



AGGREGATED APPROACH (IN TWO STEPS)

❖ **Low level of uncertainties on the aggregate**

❖ **High level of uncertainties on each local agent**

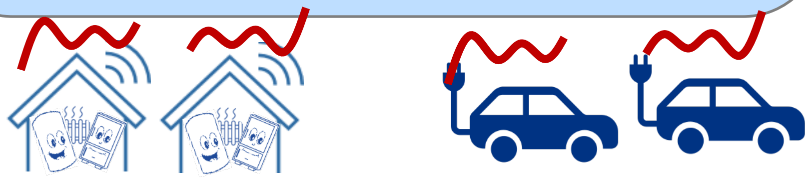
1) Optimization

an aggregated model of flexibilities $\mathcal{M}$ to generate a target profile for the aggregate

- **Few interactions**
- **Computationally cheap at the local level**
- **Preserving local models**

2) Real time dispatch

Disaggregation of the target profile among the local agents



AGGREGATED APPROACH (IN TWO STEPS)

Target profile

- ❖ **Low level of uncertainties on the aggregate**
- ❖ **High level of uncertainties on each local agent**

1) Optimization

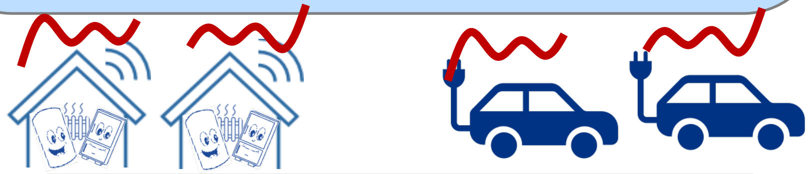
an aggregated model of flexibilities $\mathcal{M}$ to generate a target profile for the aggregate

- ➢**Few interactions**
- ➢**Computationally cheap at the local level**
- ➢**Preserving local models**

2) Real time dispatch

Disaggregation of the target profile among the local agents

Building&Optimizing the Aggregated model
[JacquotEtal2020], [GobetEtal2021], [SeguretEtal2021]
[MarinEtal2022], [BourdaisEtal 2023]



Operating Real time dispatch
[BusicEtal2016], [BendottiEtal2022],
[GobetEtal2021]

AGGREGATED APPROACH : OPTIMIZATION, DISPATCH, LEARNING

- ❖ **Low level of uncertainties on the aggregate**

Target profile

- ❖ **High level of uncertainties on each local agent**

Optimization & Learning

an aggregated model of flexibilities $\mathcal{M}$ **continuously updated** to generate a target profile for the aggregate

- ➢ **Few interactions**
- ➢ **Computationally cheap at the local level**
- ➢ **Preserving local models**

Real time dispatch

Disaggregation of the target profile among the local agents

Tracking error

SOME PERSPECTIVES

❖ Considering **network constraints** ?

❖**Learning the mean-field model** while optimizing

[Phd Bianca Marin Moreno in progress]

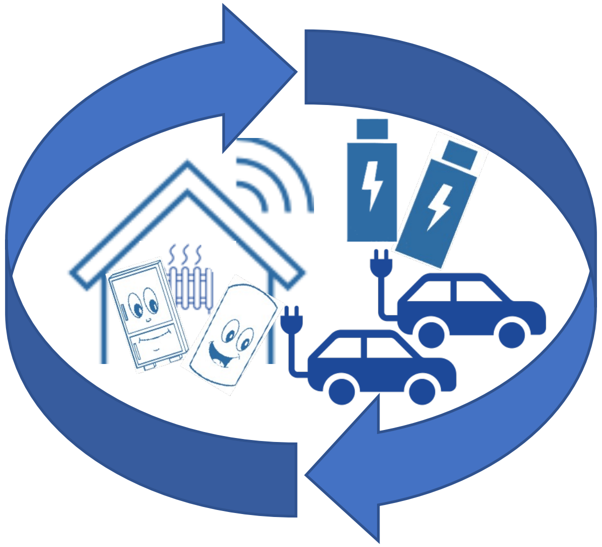
❖**Incentives design and games**
[JacquotEtal2017, AidEtal2018, ElieEtal2020, AusselEta2020, LiuEtal20, AlasseurEtal2022]

❖**Sharing data** while optimizing

❖Mean-field control on networks

Session “Distributed optimization : Machine learning & Mean Field games”

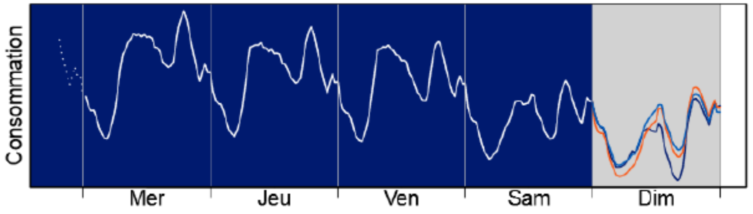
***Forecast** & control in decentralized electrical systems*



__Yannig Goude__ and Nadia Oudjane - Ateliers Fime 13 septembre 2023

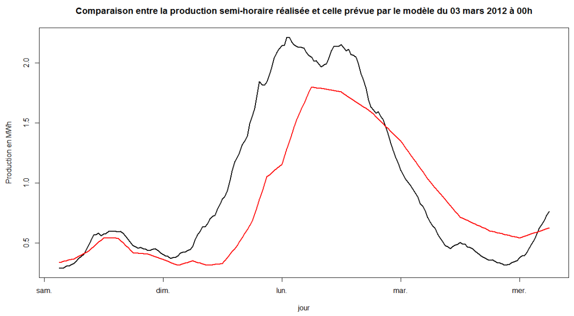
- Time series forecasting is an important issue for EDF, particularly for electricity markets:
  - Production/consumption planning: optimisation of the production fleet, demand response
  - Trading: buy/sell on electricity markets
  - Grid management

Electricity consumption

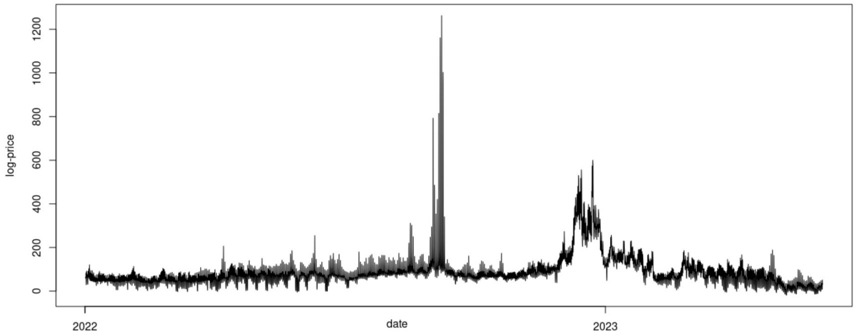


- Availability of data at a low level of aggregation (consumers, assets), high temporal resolution
- Development of advanced analytical approach to get value from such data.

Renewable production

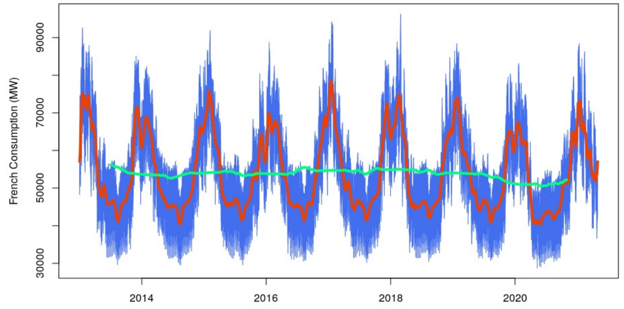


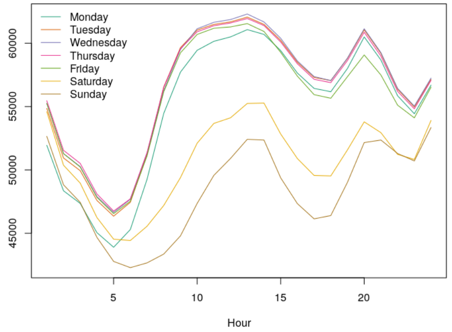
Electricity Prices

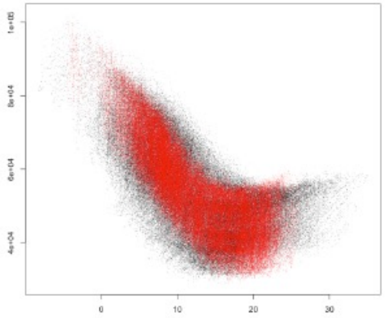


DATA ASSETS

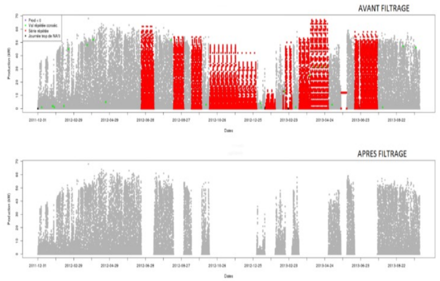
Global electricity load

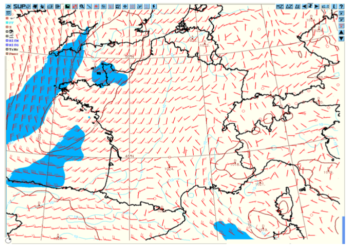




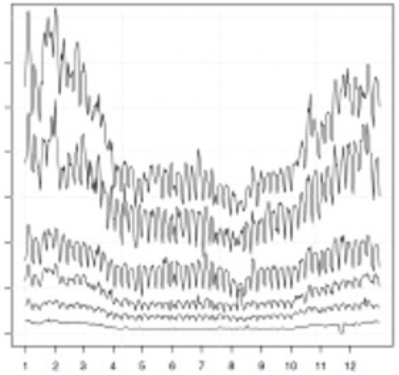


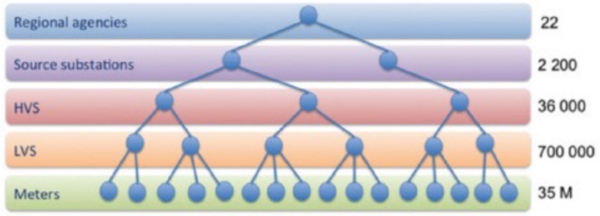
Renewable production





Local electricity load

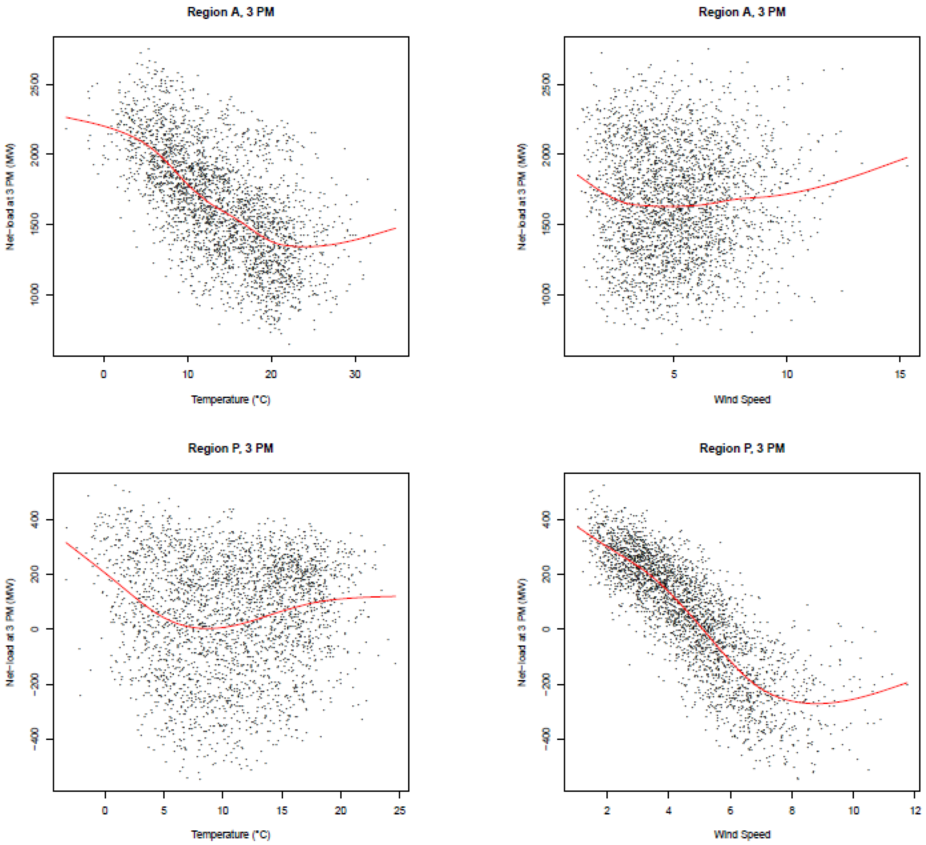




FORECASTING MODELS

$$y_t = \sum_{j=1}^{d} f_j(x_{t,j}) + \varepsilon_t\,, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma^2)$$

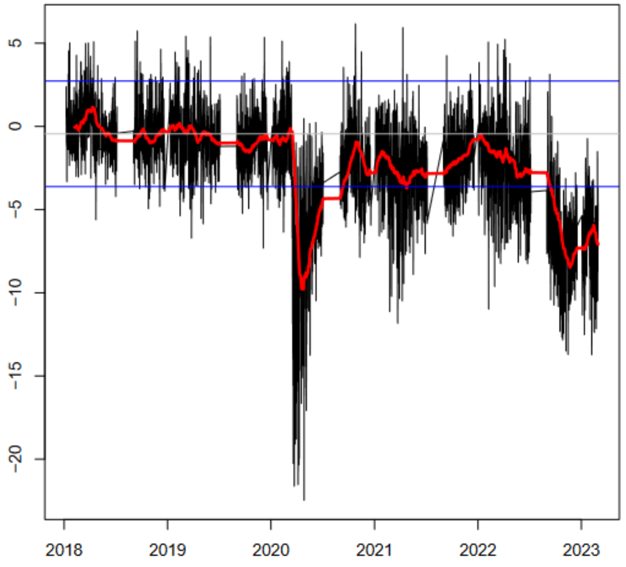
- Statistical models have to **perform well** (arround 1% error for France) and be ***explainable***
- assumption of data stationarity



Wood, S. N., Goude, Y., & Shaw, S. (2015). Generalized additive models for large data sets. Journal of the Royal Statistical Society Series C: Applied Statistics, 64(1), 139-155.

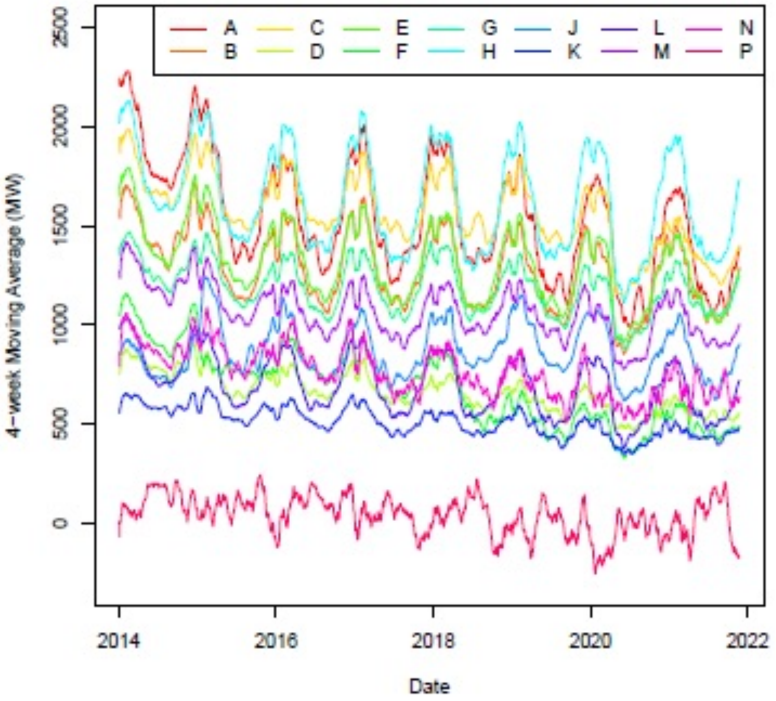
ENERGY MARKETS ARE CHANGING

Unexpected Events



New Usages

Development of Renewables



SEQUENTIAL LEARNING

adaptive models and ***sequential learning***

At each round $t = 1, \ldots, n$, the learner

- observes an input $x_t \in \mathcal{X} \subset \mathbb{R}^d$
- makes a prediction $\hat{y}_t = \langle x_t, \theta_t \rangle$ with a linear model
- observes the true output $y_t \in \mathcal{Y}$
- measures loss by $\ell_t(\theta_t) = (\hat{y}_t - y_t)^2$
- updates his prediction rule $\theta_t \to \theta_{t+1}$

$$y_t = \boldsymbol{\theta}_t^\top f(\boldsymbol{x}_t) + \varepsilon_t \,,$$
$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \boldsymbol{\eta}_t \,,$$

(ε_t) and $(\boldsymbol{\eta}_t)$ are gaussian white noises

variance / covariance σ^2 and Q

Algorithm 1: Kalman Filter

Initialization: the prior $\boldsymbol{\theta}_1 \sim \mathcal{N}(\hat{\boldsymbol{\theta}}_1, P_1)$ where $P_1 \in \mathbb{R}^{d \times d}$ is positive definite and $\hat{\boldsymbol{\theta}}_1 \in \mathbb{R}^d$.

Recursion: at each time step $t = 1, 2, \ldots$

1) Prediction:

$$\mathbb{E}\left[y_t \mid (\boldsymbol{x}_s, y_s)_{s<t}, \boldsymbol{x}_t\right] = \hat{\boldsymbol{\theta}}_t^\top f(\boldsymbol{x}_t) \,,$$
$$Var\left[y_t \mid (\boldsymbol{x}_s, y_s)_{s<t}, \boldsymbol{x}_t\right] = \sigma^2 + f(\boldsymbol{x}_t)^\top P_t f(\boldsymbol{x}_t) \,.$$

2) Estimation:

$$\hat{\boldsymbol{\theta}}_{t+1} = \hat{\boldsymbol{\theta}}_t + \frac{P_t f(\boldsymbol{x}_t)}{f(\boldsymbol{x}_t)^\top P_t f(\boldsymbol{x}_t) + \sigma^2}(y_t - \hat{\boldsymbol{\theta}}_t^\top f(\boldsymbol{x}_t)) \,,$$
$$P_{t+1} = P_t - \frac{P_t f(\boldsymbol{x}_t) f(\boldsymbol{x}_t)^\top P_t}{f(\boldsymbol{x}_t)^\top P_t f(\boldsymbol{x}_t) + \sigma^2} + Q \,.$$

- high dimension/ online variable selection
- extremes, out of distribution forecasts

Obst, D., De Vilmarest, J., & Goude, Y. (2021). Adaptive methods for short-term electricity load forecasting during COVID-19 lockdown in France. IEEE transactions on power systems, 36(5), 4754-4763.

VARIANCE TRACKING

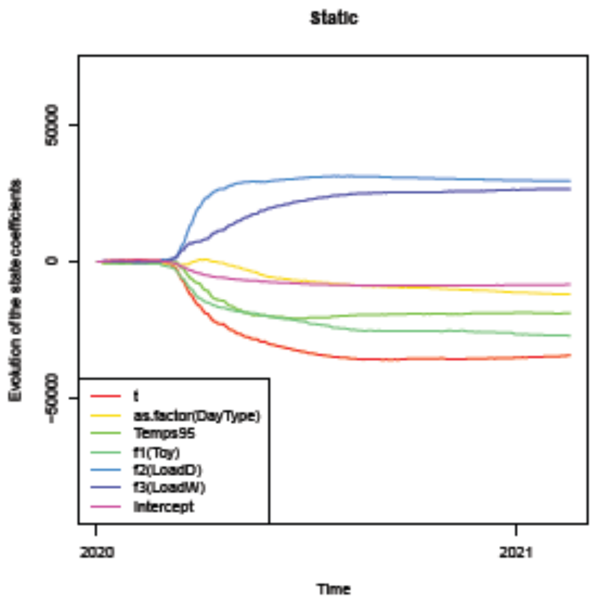
$$P_{t|t} = P_t - \frac{P_t f(x_t) f(x_t)^\top P_t}{f(x_t)^\top P_t f(x_t) + \sigma_t^2}, \qquad P_{t+1} = P_{t|t} + Q_{t+1},$$

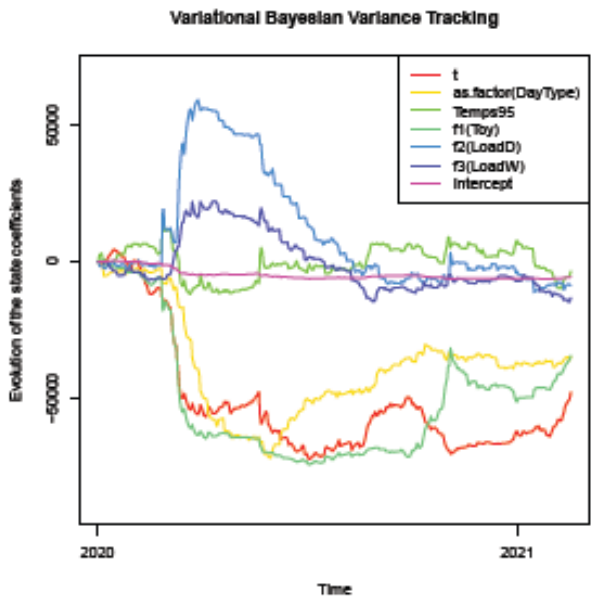
$$\hat{\theta}_{t+1} = \hat{\theta}_t - \frac{P_{t|t}}{\sigma_t^2}\left(f(x_t)(\hat{\theta}_t^\top f(x_t) - y_t)\right).$$

σ and Q can also be updated online (latent variable)

Evolution of θ in function of time

static - **Viking**

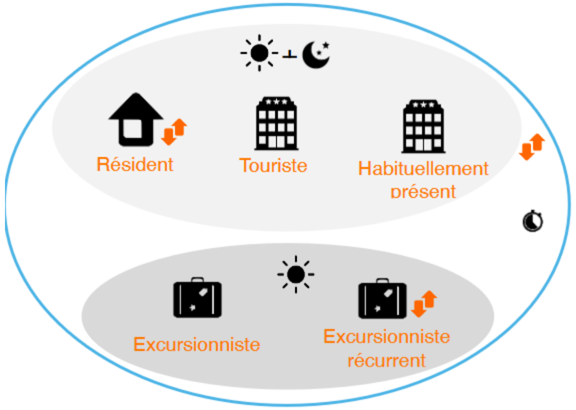


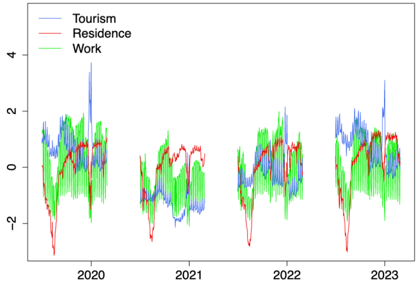


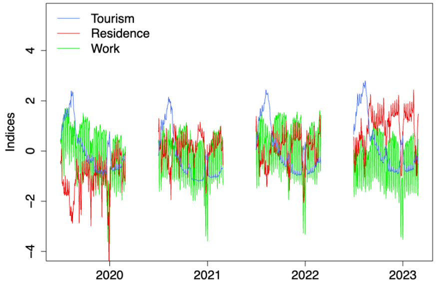
De Vilmarest, J., & Goude, Y. (2022). State-Space Models for Online Post-Covid Electricity Load Forecasting Competition. IEEE Open Access Journal of Power and Energy, 9, 192-201.

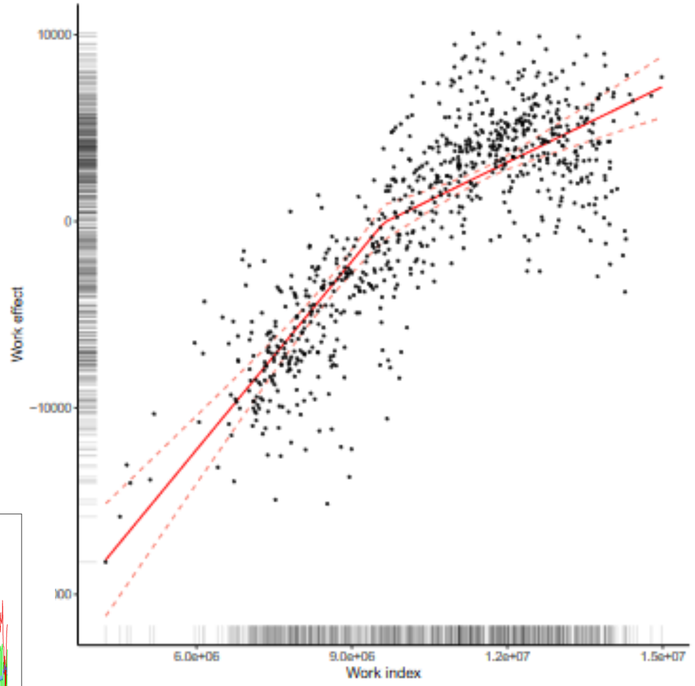
FORECASTING WITH NEW DATA

101 geographical areas
Anonymised presence data









	Without mobility data		With mobility data	
	RMSE (GW)	MAPE (%)	RMSE (GW)	MAPE (%)
Statistical model				
Persistence (1 day)	4.0 ± 0.2	5.5 ± 0.3	N.A.	N.A.
SARIMA	2.4 ± 0.2	**3.1** ± 0.2	N.A.	N.A.
GAM	2.34 ± 0.1	3.54 ± 0.2	**2.17** ± 0.08	3.3 ± 0.1
Data assimilation technique				
Static Kalman filter	2.1 ± 0.1	3.1 ± 0.2	1.72 ± 0.08	2.5 ± 0.1
Dynamic Kalman filter	1.4 ± 0.1	1.9 ± 0.1	1.20 ± 0.08	1.7 ± 0.1
Viking	1.5 ± 0.1	1.8 ± 0.1	1.24 ± 0.07	1.7 ± 0.1
Aggregation of experts	1.4 ± 0.1	1.8 ± 0.1	**1.16** ± 0.07	**1.6** ± 0.1
Machine learning				
GAM boosting	2.6 ± 0.2	3.7 ± 0.2	2.4 ± 0.1	3.5 ± 0.2
Random forests	2.5 ± 0.2	3.5 ± 0.2	**2.0** ± 0.1	**2.7** ± 0.2
Random forests + bootstrap	2.2 ± 0.2	3.0 ± 0.2	**2.0** ± 0.1	**2.7** ± 0.2

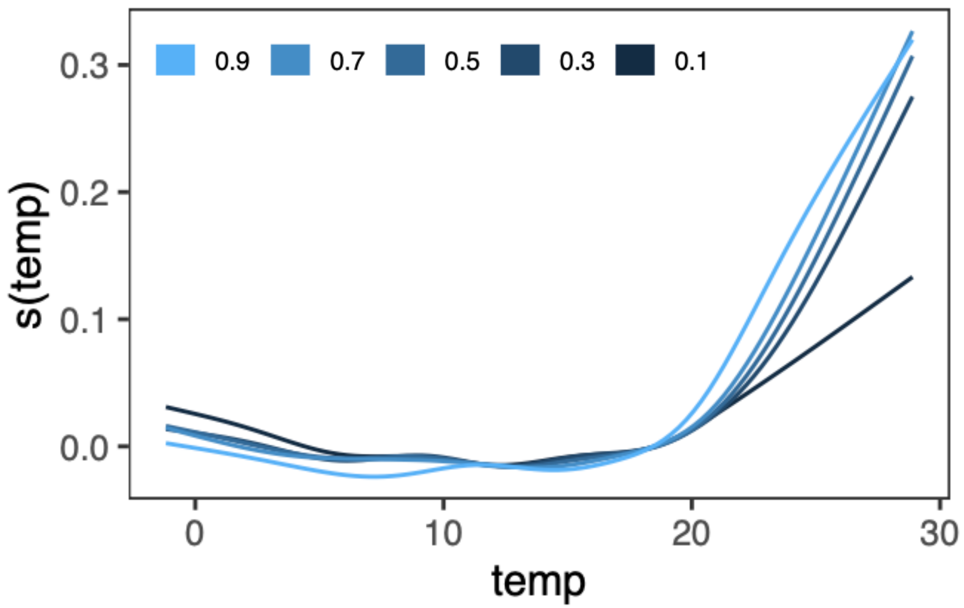
- Costly data: bandit…
- Complex time/space dependencies (drifts, breaks)

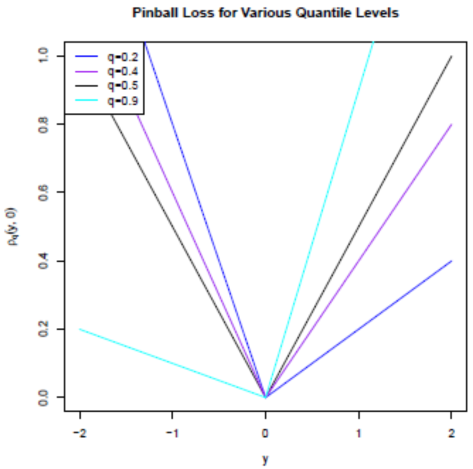
Human spatial dynamics for electricity demand forecast: the case of France during the 2022 energy crisis, Nathan Doumèche, Yann Allioux, Yannig Goude, Stefania Rubrichi, preprint.

(ONLINE) PROBABILISTIC FORECAST

- **Mean** forecast: $\hat{y}_t = \mathbb{E}[y_t \mid x_t]$.
  Equivalent to the minimum of $\mathbb{E}[(y_t - \hat{y}_t)^2 \mid x_t]$.
- **Probabilistic** forecast: estimation of $\mathcal{L}(y_t \mid x_t)$.
  For $0 < q < 1$, we find $\hat{y}_{t,q}$ such that $\mathbb{P}(y_t \le \hat{y}_{t,q} \mid x_t) = q$.
  Equivalent to the minimum of $\mathbb{E}[\rho_q(y_t, \hat{y}_t) \mid x_t]$:

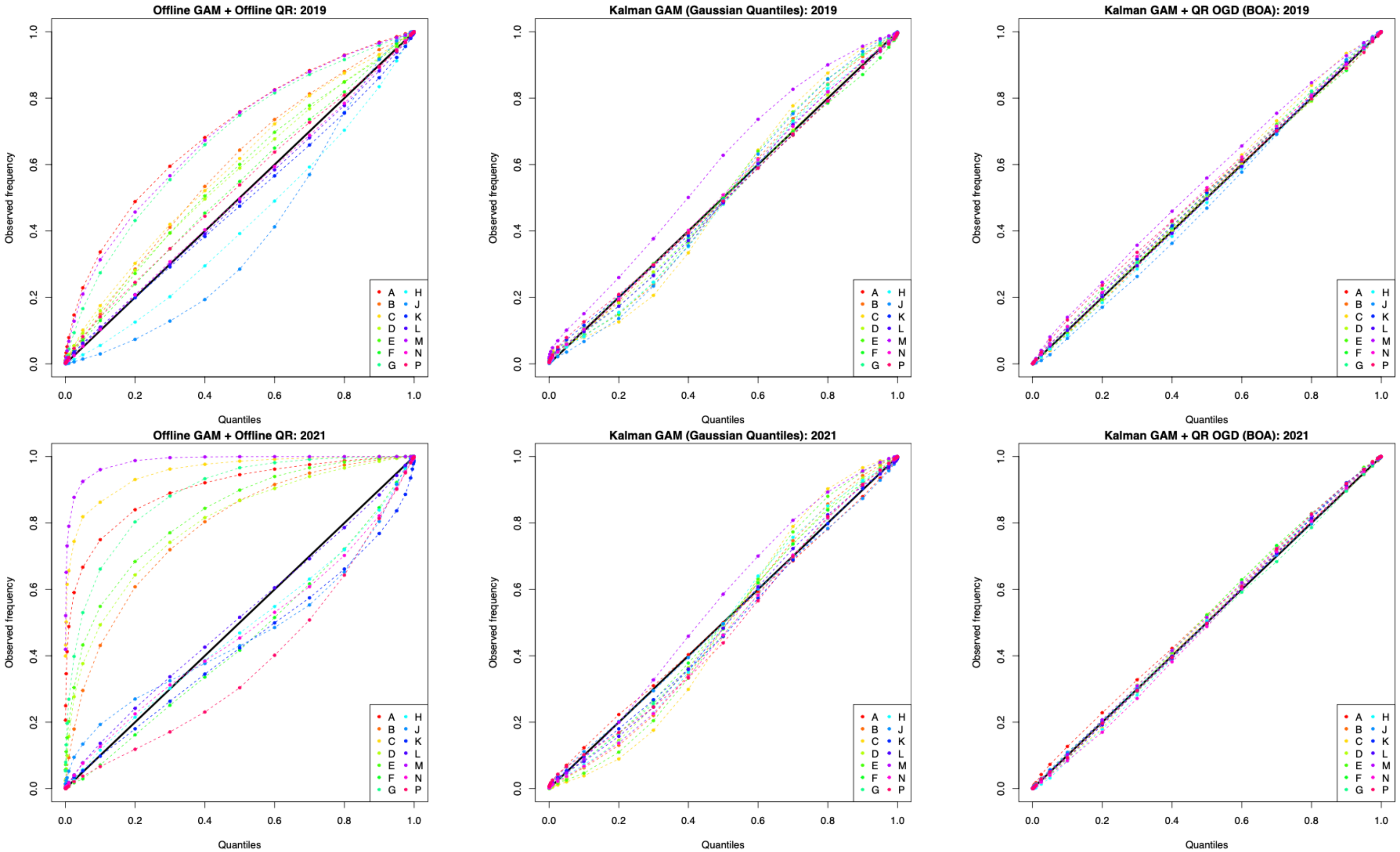
$$\rho_\tau(z) = (\tau - 1)z\mathbb{1}(z < 0) + \tau z\mathbb{1}(z \geq 0)$$

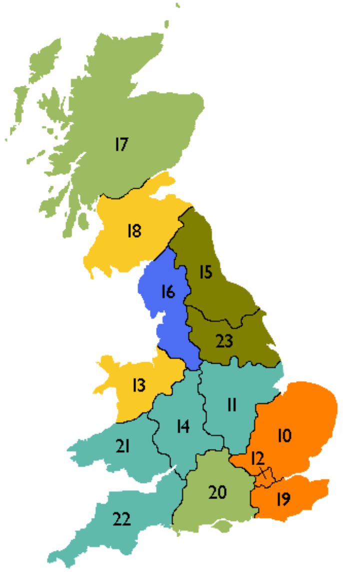


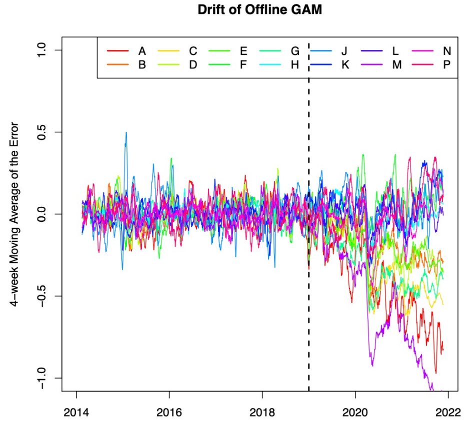


Fasiolo, M., Wood, S. N., Zaffran, M., Nedellec, R., & Goude, Y. (2021). Fast calibrated additive quantile regression. *Journal of the American Statistical Association*, *116*(535), 1402-1412.

(ONLINE) PROBABILISTIC FORECAST: REGIONAL UK RESIDUAL DEMAND

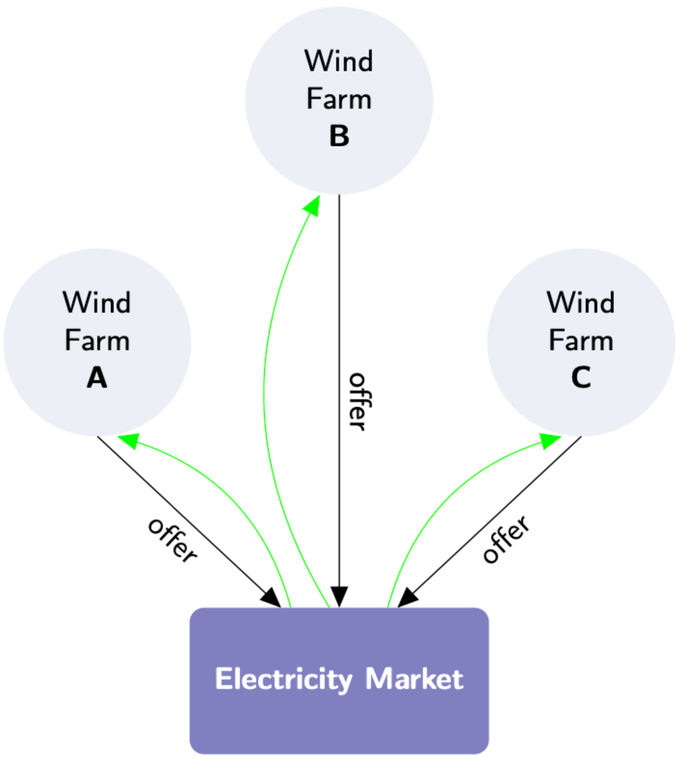


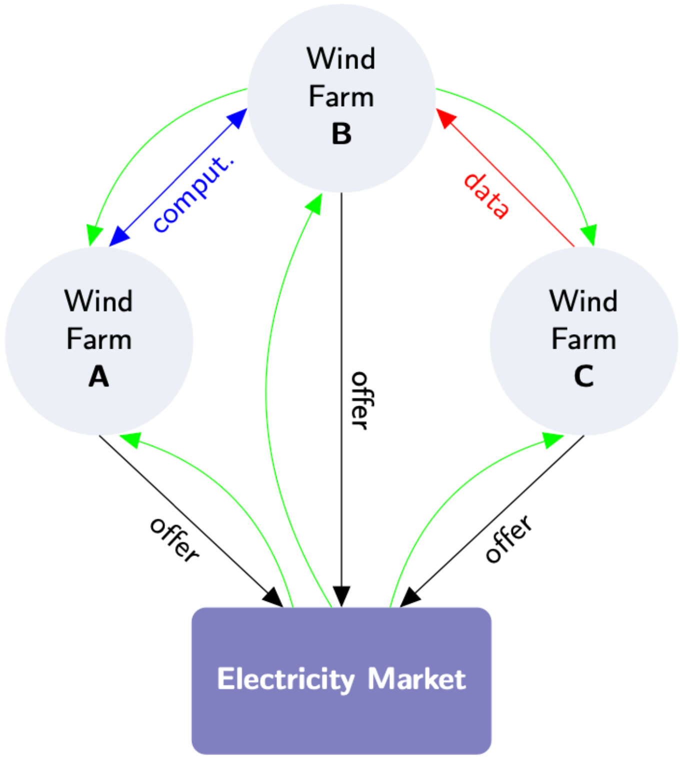




De Vilmarest, J., Browell, J., Fasiolo, M., Goude, Y., & Wintenberger, O. (2023). Adaptive Probabilistic Forecasting of Electricity (Net-) Load, to appear in IEEE Transactions on Power Systems.

DATA MARKETS FOR ELECTRICITY FORECASTING





- Wind farms offer in electricity markets based on their individual (probabilistic) forecasts and private information
- Their revenue is affected by their (lack of) forecast accuracy

Opportunity: All *could* benefit from some form of collaboration (e.g., information sharing)

Challenge: They have no interest in doing so

Proposal: Design a framework allowing for all agents to collaborate and benefit from it

Source: P. Pinson, Towards markets for data and analytics, keynote ENBIS, Valencia 2023.

DATA MARKETS FOR ELECTRICITY FORECASTING

- Consider a *central agent* (“**Forecaster**”) with a regression problem, e.g., as a basis to forecast renewable power generation for a given site (y_{t+k})
- **Forecaster** owns a set ω of m features, $\omega = \{x_1, \ldots, x_m\}$

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}}\, S_\omega(\boldsymbol{\beta})$$

- Two *support agents* **Good Data** and **Useful Features** may bring in additional features z_1 and z_2, to be remunerated
- The overall set of features now is $\Omega = \omega \cup \{z_1, z_2\}$

$$\hat{\boldsymbol{\beta}}^+ = \underset{\boldsymbol{\beta}^+}{\operatorname{argmin}}\, S_\Omega(\boldsymbol{\beta}^+),$$

- If z_1 and/or z_2 are informative features, one expects $S^*_\Omega < S^*_\omega$

Allocation policy based on feature valuation (can be obtained with, e.g., leave-one-out or shapley-based allocation)
Application: Wind farm production forecasting in North Carolina

Source: P. Pinson, Towards markets for data and analytics, keynote ENBIS, Valencia 2023.

- Privacy
- Competition between agents
- Ethics
- Connection with online learning

P. Pinson, L. Han, J. Kazempour (2022) Regression markets and application to energy forecasting. TOP 30: 533–573.

DEFI INRIA / EDF (2023-2027)

4 years collaboration between EDF/OSIRIS and 10 INRIA teams including **5 post-doc and 4 Phd thesis** related to the **management of new electrical systems** such as :

- Phd Thesis 2024 : ***Aggregated modelization of flexibilities in long-term optimization models***

with Stéphane GAUBERT (Tropical), Nicolas GAST and Bruno GAUJAL (Polaris)

- Phd Thesis 2024 : ***Future onshore/offshore wind generation in a context of energy transition***

with Claire MONTELEONi (Arches)

- Phd Thesis 2024 : ***Adaptive hierarchical probabilistic revision of time series, application to power system management***

with Gilles STOLTZ (Celeste), Jean-Michel POGGI (Celeste)

- Post-Doc 2024 : ***Distributed optimization techniques for managing demand flexibilities for large scale non convex systems including network constraints***

with Laurent PFEIFFER (Disco), Francis BACH (Sierra)

- Post-Doc 2025 : ***Real time learning methods for demand management on the electrical network***

with Pierre GAILLARD (Thoth), Hadrien HENDRIKX (Thoth), Gilles STOLTZ (Celeste)

Chaire MARCHÉS ET APPRENTISSAGE avec la fondation INRIA

Objectif :
Dans un contexte de marchés de l'énergie volatils (variations de prix, de consommation, de production d'ENR), l'objectif est de développer nos compétences/outils à l'intersection de l'**économie et l'Intelligence Artificielle**.

Etude des mécanismes de marchés dans lesquels il existe une interaction entre acteurs via des échanges de données (ex: système de recommandation).

L'IA permet d'apprendre les comportements des acteurs et d'envoyer des incitations (charge de VE, autoconso...) pour parvenir à un optimum économique pour l'ensemble des acteurs.

Partenaires : **INRIA - SNCF, BNP-Paribas, Orange, Air Liquide**

Thèmes : IA, algorithmes distribués et marchés, incitation économique et apprentissage pour l'optimisation de flexibilités diffuses, prévision exploitant des données individuelles/locales

Collaboration des chercheurs d'EDF avec **Michael Jordan** et l'équipe INRIA Sierra (Apprentissage statistique) dirigée par **Francis Bach**.

Agenda: En cours de finalisation

References

- **[BonnansEtal2022]** Bonnans, J. F., Liu, K., Oudjane, N., Pfeiffer, L., & Wan, C. (2022). Large-scale nonconvex optimization: randomization, gap estimation, and numerical resolution. arXiv preprint arXiv:2204.02366.
- **[LiuEtal2023]** Liu, K., Oudjane, N., & Pfeiffer, L. (2023). Decomposed resolution of finite-state aggregative optimal control problems. In 2023 Proceedings of the Conference on Control and its Applications (CT) (pp. 56-63). Society for Industrial and Applied Mathematics.
- **[SeguretEtal2023]** Seguret, A., Alasseur, C., Bonnans, J. F., De Paola, A., Oudjane, N., & Trovato, V. (2023). Decomposition of convex high dimensional aggregative stochastic control problems. Applied Mathematics & Optimization, 88(1), 8.
- **[BusicEtal16]** Busic A. and Meyn S. Distributed randomized control for demand dispatch In IEEE 55th Conference on Decision and Control (CDC), Dec. 2016, pp. 6964 6971
- **[GobetEetal2021]** Gobet, E., & Grangereau, M. (2021). Federated stochastic control of numerous heterogeneous energy storage systems.
- **[Jacquot2020]** Jacquot P., Beaude, O., Benchimol P., Gaubert, S. and Oudjane N. A Privacy-preserving Method to Optimize Distributed Resource Allocation SIAM J. Optim., 30(3), 2303-2336, Aug. 2020
- **[Brevet2021]** Demande de Brevet PCT/EP2021/086343 16 décembre 2021, Programmation conjointe de flexibilités de production et de consommation
- **[BendottiEtal2022]** Bendotti P., Oudjane N. and Wan C., Distribued control of flexible loads by mean-field inversion work submitted, 2022
- **[MarinEtal2023]** Marin Moreno, B., Brégère, M., Gaillard, P., & Oudjane, N. (2023). A mirror descent approach for Mean Field Control applied to Demande-Side management. arXiv e-prints, arXiv-2302.
- **[BourdaisEtal2023]** Bourdais, T., Oudjane, N., & Russo, F. (2023). An entropy penalized approach for stochastic control problems. Complete version. arXiv preprint arXiv:2309.01534.
- **[AlasseurEtal2022]** Alasseur, C., Bayraktar, E., Dumitrescu, R., & Jacquet, Q. (2022). A Rank-Based Reward between a Principal and a Field of Agents: Application to Energy Savings. arXiv preprint arXiv:2209.03588.
- **[SeguretEtal2021]** Seguret, A., Wan, C., & Alasseur, C. (2021, October). A mean field control approach for smart charging with aggregate power demand constraints. In 2021 IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe) (pp. 01-05). IEEE.

Merci

Incentives design and games

[JacquotEtal2017, AidEtal2018, ElieEtal2020, AusselEta2020, LiuEtal20, AlasseurEtal2022]

- **How to incentive consummers to participate** in balancing the system ?

$\Rightarrow$**Design Agent bills or tarifs** such that the resulting game equilibrium is good for the whole system

Each agent/player minimizes its bill

$$\min_{u^i} \varphi_i(u^i, X^{i,u^i}, \frac{1}{n}\sum_j u^j)$$

=>Principal/Agent, Stackelberg Game, bilevel optimization, ….

2 steps approach : *D-1* then *real time*

1. Forecast (D-1) : « Flexibility Commitment Problem » (FCP) : compute a reachable target (r_k)

Centralized flexibilities model

- ✓ **technical constraints**
- ✓ **costs**

Decisions $(u_i)_{i \in I_{central}}$

Distributed flexibilities model

- ✓ **technical const., QoS, uses**
- ✓ **costs**

Decisions $(u_i)_{i \in I_{distrib}}$

Short-term Optimization

(ex : daily time horizon adapted to seasonality of uses, time step 15mn)

Centralized flexibilities schedule

$(u_i)_{i \in I_{central}}$

Reference profile for distributed flexiblilities

$$r = (r_t) = \frac{1}{\# I_{ditrib}} \sum_{i \in I_{distrib}} u_i$$

2. Real time dispatch « meanfield Inversion »

Real time Dispatch

[BusicEal2016, BendottiEtal2022]

PI (Proportionnel Intégral), LQR, meanfield Inversion …