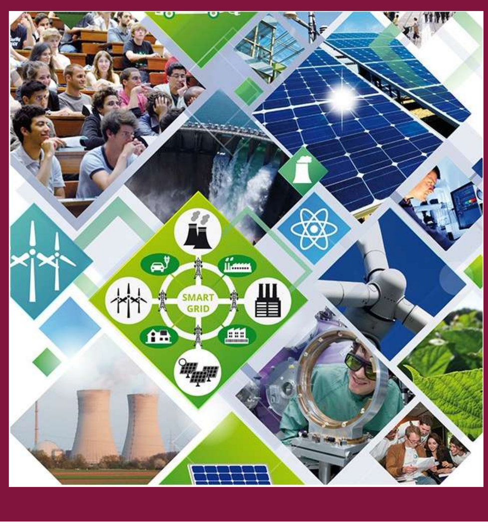


Bandit Algorithms for Power Consumption Control

Margaux Brégère

EDF R&D, Université Paris-Sud, Inria

with Gilles Stoltz, Yannig Goude and Pierre Gaillard

Introduction

Electricity cannot be stored
☛ Equilibrium production/consumption

How to optimize the transmission of these control signals?

Today EDF adjusts its power production to the expected consumption

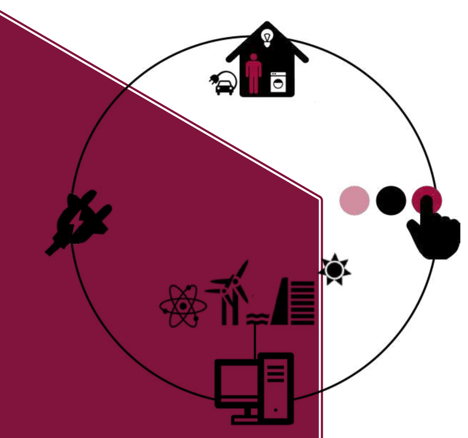
Sending incentive signals (electricity tariff variations) to influence consumption

Renewable energies are subject to climate
☛ Difficulties in adjusting production

New communication tools (smart meters)
☛ Data access & instantaneous communication with consumers

Modelling

In order to influence power consumption, K tariffs are available
The vector $\nu = (\nu_1, \dots, \nu_K)$ of the consumption laws associated with the different tariffs is unknown



At $t = 1, 2, \dots$

- A target consumption c_t is given
- The tariff $I_t \in \{1, \dots, K\}$ (discrete setting) or the population proportions receiving the different tariffs $p_t \in \mathcal{P} = \{p \in [0,1]^K \mid \sum_{k=1}^K p_k = 1\}$ (continuous setting) are chosen
- The consumption Y_t is observed. Y_t is a random variable with $Y_t \sim \nu_{I_t}$ (discrete setting) or $Y_t = p_t^T Z$, $Z \sim \nu$ (continuous setting)
- The loss $\ell(Y_t, c_t)$ is suffered, with $\ell(\cdot, c_t)$ a convex function of minimum c_t

Application of bandit theory

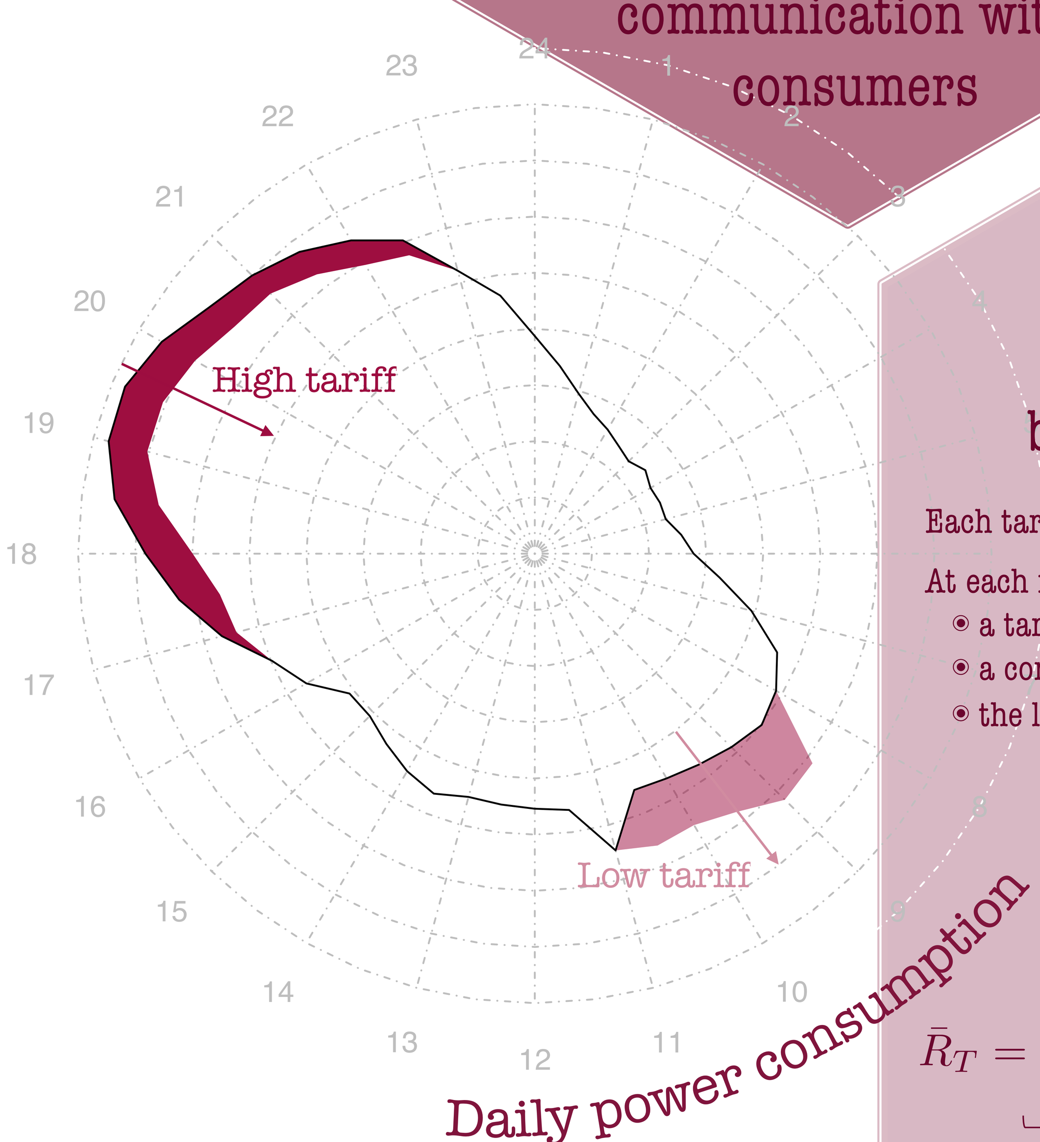
Learn client behaviour & Optimize chosen tariffs
Estimate consumptions & Pick tariffs accordingly

Exploration - Exploitation trade-off and Sequential Learning

Application of bandit algorithms for power consumption control:
Tariffs = Slot machines

How to simulate and test bandit algorithms?

Analyse data, model it and simulate new data to be in full information setting and be able to evaluate algorithms performance



Bandit Models

In a multi-armed bandit problem, a gambler faces a row of K slot machines (also called "one-armed bandits") and has to decide which ones to play in order to maximize her reward

Stochastic multi-armed bandit for target tracking

Each tariff k is defined by an unknown distribution ν_k

At each round $t = 1, \dots, T$

- a tariff $I_t \in \{1, \dots, K\}$ is picked according to c_t
- a consumption Y_t is observed, with $Y_t | I_t \sim \nu_{I_t}$
- the loss $\ell(Y_t, c_t)$ is suffered

Pseudo regret

$$\bar{R}_T = \underbrace{\sum_{t=1}^{T} \mathbb{E}[\ell(Y_t, c_t)]}_{\text{Expectation of the global loss of the strategy}} - \overbrace{\sum_{t=1}^{T} \min_k \ell_{k,t}}^{\text{Expectation of the global loss of the best strategy: at each time } t \text{ the optimal tariff is picked}}$$

$$\ell_{k,t} = \mathbb{E}[\ell(Y, c_t)], \quad Y \sim \nu_k$$

Tracking - Upper Confidence Bound (UCB) algorithm

For $t = 1, \dots, K$

$$I_t = t$$

For $t = K+1, \dots$

$$I_t \in \operatorname*{Argmin}_k \left\{ \hat{\ell}_{k,t} - \overbrace{\sqrt{\frac{2 \log t}{N_{k,t-1}}}}^{\text{Exploration term } \searrow \text{ with } N_{k,t-1}, \nearrow \text{ with } t} \right\}$$

Empirical loss of the tariff I_t at t:

$$\hat{\ell}_{k,t} = \frac{1}{N_{k,t-1}} \sum_{s=1}^{t-1} \ell(Y_s, c_t) \mathbf{1}_{I_s = k}$$

Number of times the tariff I_t has been picked:

$$N_{k,t-1} = \sum_{s=1}^{t-1} \mathbf{1}_{I_s = k}$$

For Tracking - UCB strategy:

$$\bar{R}_T \leq 3K + 4\sqrt{2KT \log T}$$

Data

"Smart Meter Energy Consumption Data in London Households"
Low Carbon London - UK Power Networks
Consumption at half-an-hour intervals of 1 100 clients subjected to Dynamic Time of Use energy prices
3 tariffs: Low, Normal, High

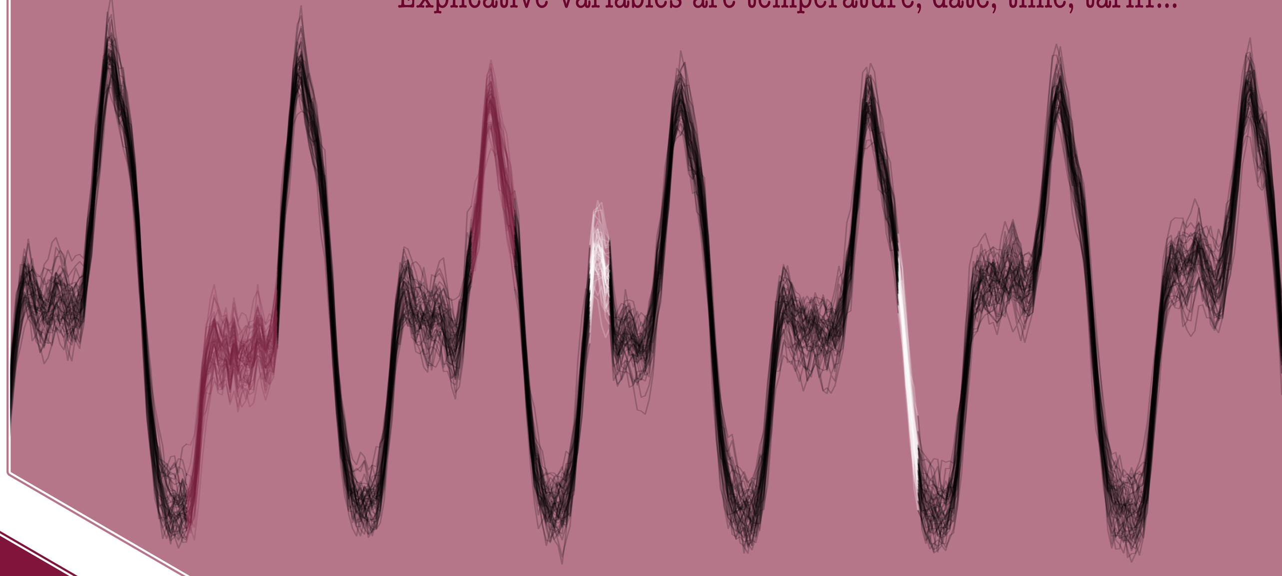
A generalized additive model for consumption Y:

With $(X_i)_i$ some explicative variables, $\mathbb{E}[Y] = \sum_i f_i(X_i)$

- if X_i is a discrete variable with m modalities: $f_i(x_i) = \sum_{j=1}^{m} \alpha_j \mathbf{1}_{x_i = j}$
- if X_i is a continuous variable, its effect is modelled with spline:

$$f_i(x_i) = \sum_{\ell=1}^{k} P_\ell(x_i) \mathbf{1}_{x^{\ell-1} \leq x_i \leq x^\ell} \text{ with } P_\ell(x_i) = a_{\ell,0} + a_{\ell,1} x_i + \dots + a_{\ell,d_i} x_i^d \text{ and } P_\ell(x^{\ell-1}) = P_{\ell+1}(x^\ell)$$

Explicative variables are temperature, date, time, tariff...



With the data provided, consumption law is estimated and thanks to this model, for any temperature, date, time, tariff... a realistic random consumption is generated. Thus, performance of bandit algorithms can be evaluated.

Results

Assumptions

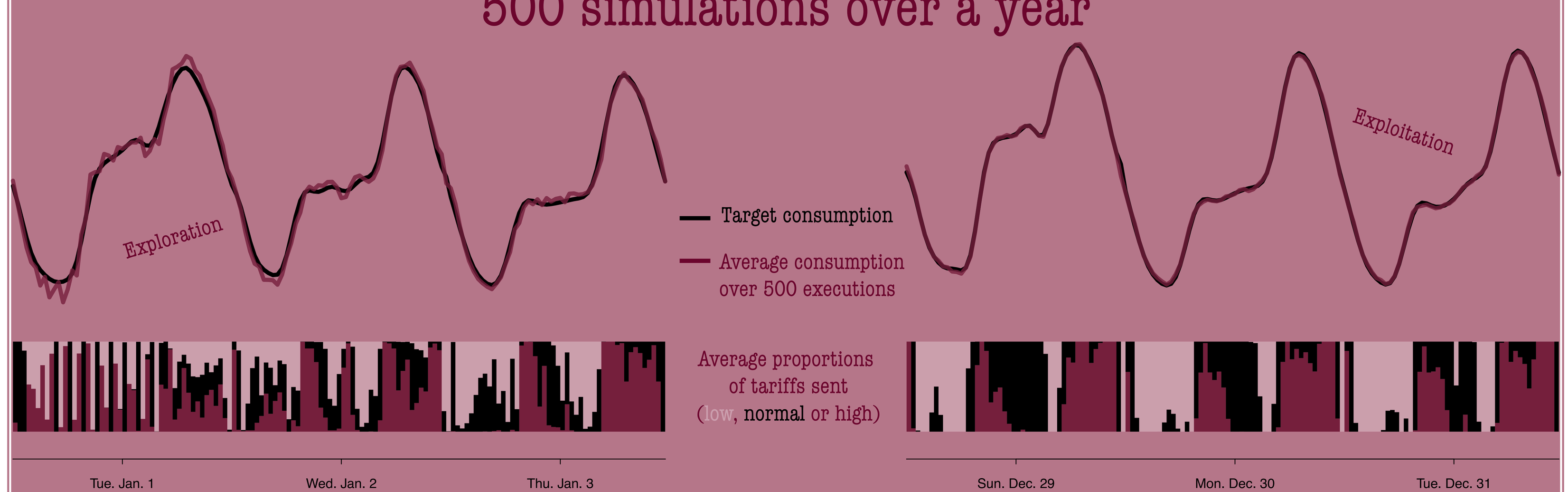
- Quadratic loss: $\ell(Y_t, c_t) = (Y_t - c_t)^2$
- Consumption model: $Y_t = p_t^T \theta + \varepsilon_t$ (variance of noise independent of chosen tariffs)
- Only two tariffs given at a same time: $\mathcal{P} = \left\{ p \in [0,1]^K \mid \Sigma_{k=1}^K p_k = 1 \text{ and } \Sigma_{k=1}^K \mathbf{1}_{p_k = 0} \geq K - 2 \right\}$
- Attainable target: $\forall t, \ \exists \ p \in \mathcal{P} \mid p^T \theta = c_t$

Algorithm (continuous setting)

- Parameters: $\lambda, (\beta_t)_t$
- Initialisation: $V = \lambda I_K$ and $\hat{\theta} = \mathbf{0}$
- For $t = 1, \dots$

$$p_t \in \operatorname*{Argmin}_{p \in \mathcal{P}} \left\{ \left(p^T \hat{\theta} - c_t\right)^2 - \beta_t \ p^T V^{-1} p \right\}$$
$$V = V + p_t p_t^T$$
$$\hat{\theta} = V^{-1} \sum_{s=1}^{t} Y_s p_s \text{ (Ridge estimator)}$$

500 simulations over a year



Conclusion and References

Conclusion and prospects

- ☑ Design, implement and test an efficient algorithm with theoretical guarantees to track a target consumption under basic assumptions
- ☐ Deal with more complex models: noise variance depending on tariff and eventually on explicative variables (temperature, date, time...), any loss function...
- ☐ Create client clusters to send personalized signals and improve power consumption control

References

- S. Bubeck and N. Cesa-Bianchi. Regret analysis of stochastic an nonstochastic multi- armed bandit problems. Foundations and Trends in Machine Learning, 2012.
- V. Dani, T. Hayes and S. Kakade. Stochastic Linear Optimization under Bandit Feedback. COLT, 2008.
- S. Filippi, O. Cappe, A. Garivier, and C. Szepesvári. Parametric bandits: The generalized linear case. NIPS, 2010.
- J. Schoeld, R. Carmichael, S. Tindemans, M. Woolf, M. Bilton, and G. Strbac. Residential consumer responsiveness to time-varying pricing. Electrical and Electronic Engineering, 2014.
- S. Wood. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, 2006.