

PRÉVISION DE LA CONSOMMATION D'ÉLECTRICITÉ À L'ÉCHELLE DES MÉNAGES

Fatima Fahs ¹ & Myriam Maumy-Bertrand ² & Frédéric Bertrand³ & Céline Caldini-Queiros ⁴

¹IRMA, 7 rue René Descartes, 67084 Strasbourg Cedex, fatimeh.fahs@gmail.com

²IRMA, 7 rue René Descartes, 67084 Strasbourg Cedex, mmaumy@math.unistra.fr

³IRMA, 7 rue René Descartes, 67084 Strasbourg Cedex, fbertran@math.unistra.fr

⁴IRMA, 7 rue René Descartes, 67084 Strasbourg Cedex, celine.caldini-queiros@unistra.fr

Résumé. La maîtrise de la demande en énergie (MDE) a provoqué depuis ces trente dernières années des travaux de type *DSR* (*Demand Side Response* : réduire la demande de consommation électrique) qui regroupent les techniques permettant de diminuer la consommation d'énergie électrique d'un ménage, d'un bâtiment, d'un territoire, d'un pays, et de réaliser ainsi des économies en ressources fossiles et de diminuer la pollution par le CO_2 . Les travaux de l'EPRI (*Electric Power Research Institute*) ont montré qu'informer en temps réel le client de sa consommation permet de la réduire, donc confronter le consommateur à un modèle de sa demande peut mener à bien cet enjeu. Notre objectif principal est de développer et d'exploiter des modèles statistiques et de machine learning permettant de donner une analyse prévisionnelle complète de la consommation électrique le jour J à l'échelle des ménages à partir des historiques de consommation et des données météorologiques afin d'alerter le consommateur en cas d'anomalie de consommation (sur-consommation, fuite d'électricité, chutes anormales de puissance, ...).

Mots-clés. Prévision court terme, consommation d'électricité, ménages, anomalies, modèles statistiques, *machine learning*.

Abstract. The control of the energy demand (MDE) provoked for the last thirty years DSR-type work (Demand Side Response: smart way to save on total energy costs) that combine techniques to reduce electric energy consumption of a household, a building, a territory, a country, and thus to achieve savings in fossil fuels and reduce CO_2 pollution. The work of the EPRI (*Electric Power Research Institute*) has shown that informing the consumer in real time about its consumption reduces it, so confronting the consumer with a model of its demand can lead to an amelioration of this issue. Our main goal is to develop and explore statistical and *machine learning* models to provide a reliable and interpretable forecast day-to-day of electricity consumption at household-level using consumption histories and weather data, in order to allow consumer to control its electricity consumption and to be alerted in risky cases such as power leaks, over-consumption, and abnormal drop of power demands.

Keywords. Short-term forecast, electricity consumption, household level, anomalies, statistical models, *machine learning*.

1 Introduction

Le terme "Maîtrise de la Demande d'Énergie" (MDE) désigne le regroupement d'actions d'économies d'énergie mises en place pour le consommateur final et non pour le producteur d'énergie. Introduite dans les années 1990, l'objectif de la MDE vise à diminuer la consommation générale d'énergie par le biais de la demande plutôt que par l'offre. La MDE peut également viser des acteurs spécifiques ou une zone géographique particulière.

La MDE a été introduite avec le soutien de l'AFME devenue l'ADEME (l'Agence De l'Environnement et de la Maîtrise de l'Énergie) et répond à plusieurs enjeux dont la diminution du gaspillage énergétique, des émissions de GES (Gaz à Effet de Serre), la dépendance énergétique tant au niveau local que national, mais également la limitation des risques liés au nucléaire et la diminution de la précarité énergétique.

Dans l'objectif de répondre à ces défis, la MDE implique une maîtrise globale de la demande énergétique, une amélioration de la maîtrise des usages dédiés à la croissance de pointe ainsi qu'un développement de l'effacement de la consommation électrique.

De nombreuses techniques existent pour maîtriser la consommation de l'énergie, nous distinguons deux grandes familles de techniques dans ce domaine. La première famille que nous pouvons citer est le pilotage direct de la demande (voir par exemple le pilotage par le réseau électrique des pompes des piscines privées ou des climatiseurs aux États-Unis) qui a démontré son efficacité pour limiter la demande dans les périodes de pointe. Dans un deuxième temps, nous citons les travaux de l'EPRI (*Electric Power Research Institute*) qui ont montré qu'informer en temps réel le client de sa consommation permet de réduire de 10 à 15% sa consommation.

Dans ce dernier cas, un des axes d'amélioration serait de confronter le consommateur à un modèle de sa demande. Nous nous concentrons dans ce papier sur la problématique de prévision à court terme de la consommation de l'énergie du consommateur à partir de l'historique de ses propres consommations, de la température qui est la principale variable climatique exogène et des événements particuliers liés aux jours fériés qui affectent l'organisation économique et agissent ainsi directement sur la consommation d'énergie.

2 État de l'art

Différentes techniques ont été développées pour modéliser la charge d'électricité au niveau national et régional tant avec les outils classiques de l'économétrie des séries temporelles [1] (par exemple SARMAX, TBATS), qu'avec les méthodes de l'intelligence artificielle tel que les réseaux de neurones [2]. En revanche, l'application de ces modèles à la prévision d'une courbes de charge au niveau d'un ménage est une tâche difficile en raison de sa volatilité extrême. Les charges résidentielles sont souvent influencées par un certain nombre de facteurs, tels que les caractéristiques opérationnelles des appareils, les comportements des utilisateurs, les facteurs économiques, l'heure, le jour de la semaine, les vacances, les conditions météorologiques, les modèles géographiques et d'autres effets

aléatoires.

2.1 Approche statistique basée sur la famille ARMA des modèles autorégressifs

2.1.1 Processus autorégressif AR(p)

Les premiers processus autorégressifs ont été introduits par George Udny Yule dans [3]. Un processus autorégressif est un processus où l'on écrit une observation au temps t comme une combinaison linéaire des observations passées plus un certain bruit blanc.

Définition: On dit que la suite $X_t : t \geq 0$ est un processus autorégressif d'ordre p ($p > 0$) s'il peut s'écrire sous la forme suivante :

$$X_t = \sum_{k=1}^p \phi_k X_{t-k} + \epsilon_t, \text{ où } \epsilon_t \sim N(0, \sigma_\epsilon^2).$$

Les $\phi_k (K = 1, \dots, p)$ constituent les paramètres du modèle.

Dans ce cas, on note $X_t \sim AR(p)$, de la même façon on peut réécrire un processus $AR(p)$ avec un polynôme $\phi(B)$ qui multipliera X_t cette fois-ci :

$$\phi(B)X_t = \epsilon_t \text{ avec } \phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p.$$

Exemple: Un processus AR(1) prend la forme suivante:

$$X_t = \phi X_{t-1} + \epsilon_t \text{ où } \epsilon_t \sim N(0, \sigma_\epsilon^2).$$

2.1.2 Processus à moyenne mobile MA(q)

C'est Eugen Slutsky qui, en 1927, dans son article [4], a introduit pour la première fois les processus à moyenne mobile. La définition suivante présente ce processus.

Définition: On dit que la suite $X_t : t \geq 0$ est un processus à moyenne mobile d'ordre q ($q > 0$) si celui-ci peut s'écrire sous la forme suivante:

$$X_t = \sum_{k=1}^q \theta_k \epsilon_{t-k} + \epsilon_t, \text{ où } \epsilon_t \sim N(0, \sigma_\epsilon^2).$$

Les $\theta_k (K = 1, \dots, q)$ constituent les paramètres du modèle.

Dans ce cas, on note $X_t \sim MA(q)$.

Exemple: Un processus MA(1) prend la forme suivante:

$$X_t = \theta \epsilon_{t-1} + \epsilon_t \text{ où } \epsilon_t \sim N(0, \sigma_\epsilon^2).$$

On peut utiliser l'opérateur de retard B pour écrire ce processus sous une autre forme. On aura donc un polynôme en B qui multipliera ϵ_t :

$$X_t = \theta(B)\epsilon_t \text{ avec } \theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q.$$

2.1.3 Modèle ARMA(p,q)

Les modèles ARMA consistent à avoir une partie autorégressive et une partie à moyenne mobile. Herman Wold [5] a montré que les processus ARMA pouvaient être utilisés pour modéliser n'importe quelle série pour autant que les ordres p et q soient bien choisis. Box et Jenkins [6] ont travaillé pour développer une méthodologie pour l'estimation du modèle d'une série chronologique.

Définition: Un processus est dit ARMA(p,q) s'il existe des suites réelles ϕ_k et θ_k telles que

$$X_t - \sum_{k=1}^p \phi_k X_{t-k} = \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}, \text{ avec } \epsilon_t \sim N(0, \sigma_\epsilon^2).$$

On peut aussi utiliser les polynômes $\phi(B)$ et $\theta(B)$ pour réécrire ce modèle sous la forme :

$$\phi(B)X_t = \theta(B)\epsilon_t.$$

avec

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \text{ et } \theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q.$$

2.1.4 Modèle SARMA(p,q)(P,Q)

Il existe une version saisonnière du modèle ARMA(p,q), qui permet aux composants AR et MA de s'adapter aux cycles observés (semaine, mois, saison). On dénote ses modèles par **SARMA(p,q)(P,Q)** où p et q sont les paramètres habituels de notre ARMA non saisonnier et P, Q sont des paramètres spécifiques à la composante saisonnière. Pour un modèle **SARMAX**, la composante **X** du modèle indique que le processus dépend de variables explicatives exogènes.

2.2 L'approche basée sur l'intelligence artificielle

L'intelligence artificielle connaît actuellement un développement très important. Même si tous les problèmes ne peuvent pas être résolus avec ces méthodes, il n'en reste pas moins qu'un grand nombre de domaines a largement évolué ces dernières années grâce à leurs apports [7] et [8]. Par exemple, dans le domaine de la prévision, il est possible de résoudre des problèmes pratiques comme la réalisation de systèmes de prévision de la demande/production en utilisant des données historiques concernant l'usage de certaines ressources et en faisant des corrélations avec d'autres jeux de données [9]. La particularité de ces méthodes est qu'elles tiennent compte de relations complexes et non linéaires entre les variables endogènes et exogènes.

3 Application à la prédiction de la consommation d'électricité à l'échelle des ménages

3.1 Les données de l'étude

Un millier de séries temporelles sont simulées pour cette étude au pas de 30 minutes de la consommation d'électricité domestique suivant des processus autorégressifs comportant une composante saisonnière, composante thermo-dépendante et du bruit sur une période de huit mois (de Novembre à Juin) et la température à chaque demi-heure de la journée. La consommation d'électricité est mesurée en Watts (W). La simulation est réalisée avec le logiciel libre R. Notre objectif est alors d'implémenter un algorithme automatisé capable de donner une analyse prédictive fiable pour chaque série temporelle parmi les mille simulées.

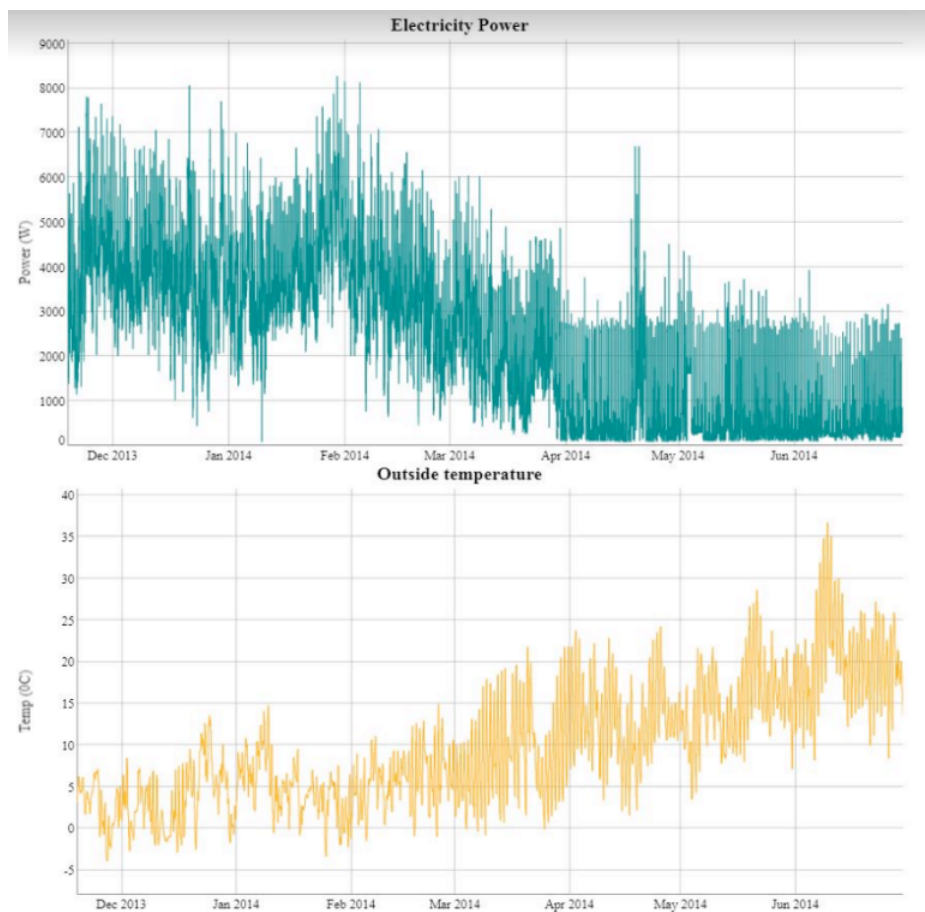


Figure 1: Une courbe de charge simulée pour un ménage et les données de la température extérieure pour la période de Novembre à Juin.

3.2 Les résultats obtenus

Plusieurs modèles standards en analyse des séries temporelles ont été appliqués au jeu de données tels que le modèle TBATS (state space model)[10], le modèle ARMA saisonnier (SARMA), le modèle ARMA saisonnier avec variables exogènes (SARMAX). Nous les comparerons à un modèle d'intelligence artificielle (RNN). La performance prédictive intra et hors échantillon de chaque modèle est évaluée avec différents indicateurs d'ajustement. L'horizon temporel hors-échantillon est testé à 24 heures (un jour). Le modèle SARIMAX se révèle le plus performant hors-échantillon sur un horizon temporel de 24h au pas demi horaires. La qualité des prévisions est tout de même globalement satisfaisante, avec l'erreur MAPE (pourcentage moyenne de l'erreur) qui varie selon le profil consommateur de 20% à 100%.

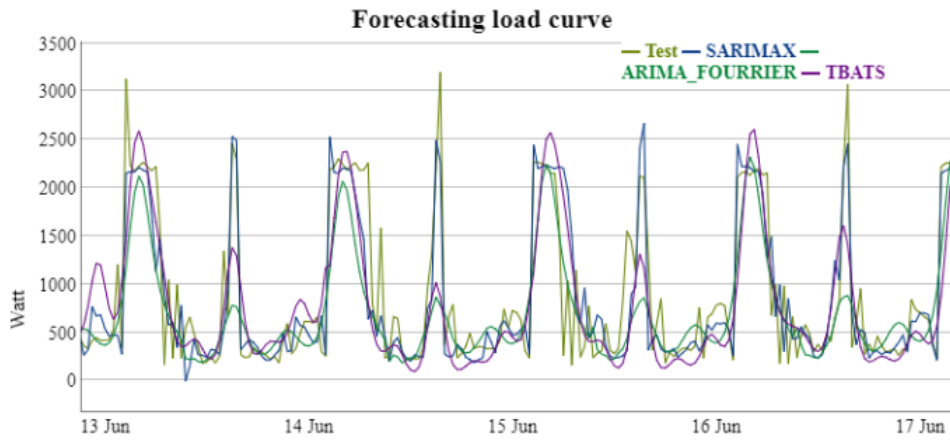


Figure 2: Exemples des courbes prédictives d'une courbe de charge simulée d'un ménage A (en vert) par différents modèles statistiques SARMA (en bleu) et TBATS [10] (en violet).

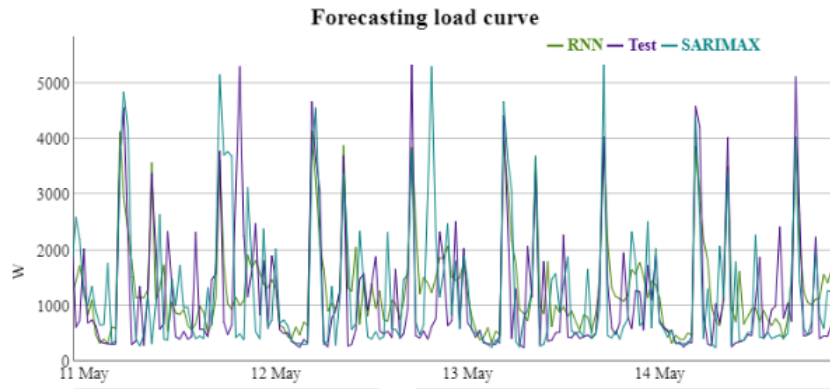


Figure 3: Comparaison des courbes prédictives d’une courbe de charge simulée d’un ménage B (en violet) par un modèle statistique (SARMAX) en bleu et un réseau de neurone (RNN) en vert.

3.3 Conclusion

La prédiction de la consommation d’énergie à l’échelle d’un ménage représente une particularité liée à la haute volatilité des séries chronologiques individuelles et un niveau de bruit élevé lié aux comportements aléatoires non réguliers, qui sont difficiles à interpréter et ne sont pas modélisables par des facteurs exogènes. Dans cette étude, le modèle obtenu donne une prévision fiable et interprétable de la consommation électrique d’un ménage à partir des données simulées par des processus autorégressifs et des données météorologiques en tenant compte des autres facteurs exogènes comme les jours de la semaine et les jours fériés.

Bibliographie

1. **Erkan Erdogdu**, *Electricity demand analysis using cointegration and ARIMA modelling: A case study of Turkey*, 2006.
2. **Maryam Jamela Ismail, Rosdiazli Ibrahim, Idris Ismail**, *Development of Neural Network Prediction Model of Energy Consumption*, 2011.
3. **George Udny Yule**, method of investigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society*, 1927, p.226.
4. **E. Slutsky**, The summation of random causes as the source of cyclic processes. *Econometrica*, 1927, p.5:105–146.
5. **H. Wold**, A study in the Analysis of Stationary Time Series, 1938.
6. **G. Box et G. Jenkins**, Time series analysis : Forecasting and Daviscontrol. *Holden Day*, 1970.
7. **Fulin Fan, Kwoklun Lo, Gao Gao**, Comparison of ARIMA and ANN Models, 2017.
8. **Mohammad Azhar Mat Daut**, Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods, 2017.
9. **Sajal Gosh**, Univariate time-series forecasting of monthly peak demand of electricity in northern India, 2008.
10. **Alysha M De Livera, Rob J Hyndman and Ralph D Snyder**, Forecasting time series with complex seasonal patterns using exponential smoothing, 2010.
11. **Hamid R. Khosravani, María Del Mar Castilla, Manuel Berenguel**, A Comparison of Energy Consumption Prediction Models Based on Neural Networks of a Bioclimatic Building, 2015.
12. **Alexis Gerossier, Robin Girard, Georges Kariniotakis, Andrea Michiorri**, Probabilistic Day-Ahead Forecasting of Household Electricity Demand, 2017.
13. **Rong Chen, John.L.Harris**, A Semi-Parametric Time Series Approach in Modeling Hourly Electricity Loads, 2016.