

Prévision de la consommation d'électricité par correction itérative du biais

Pierre-André Cornillon, Nicolas Hengartner, Vincent Lefieux, Eric
Matzner-Løber

► **To cite this version:**

Pierre-André Cornillon, Nicolas Hengartner, Vincent Lefieux, Eric Matzner-Løber. Prévision de la consommation d'électricité par correction itérative du biais. 41èmes Journées de Statistique, SFdS, Bordeaux, 2009, Bordeaux, France, France. 2009. <inria-00386629>

HAL Id: inria-00386629

<https://hal.inria.fr/inria-00386629>

Submitted on 22 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prévision de la consommation d'électricité par correction itérative du biais

Pierre-André Cornillon¹, Nicolas Hengartner², Vincent Lefieux³ & Eric Matzner-Løber⁴

(1) UMR ASB - Montpellier SupAgro, pierre-andre.cornillon@supagro.inra.fr

(2) Los Alamos National Laboratory - NW - USA, nickh@lanl.gov

(3) RTE - Département Méthodes et Appui, vincent.lefieux@rte-france.com

(4) IRMAR UMR 6625 - Université Rennes 2, eml@uhb.fr

Résumé

Une prévision correcte de la consommation d'électricité est fondamentale pour le bon fonctionnement du réseau électrique français, dont Réseau de Transport d'Electricité a la charge. Les prévisions utilisées quotidiennement par RTE sont issues d'un modèle alliant une régression paramétrique non linéaire et un modèle SARIMA. Dans l'idée d'obtenir un modèle de prévision adaptatif, des méthodes de prévision non-paramétriques ont déjà été testées sans succès véritable. On sait notamment que la qualité d'un prédicteur non-paramétrique résiste mal à un grand nombre de variables explicatives, ce qu'on appelle communément le fléau de la dimension. Nous utilisons ici une méthode de correction itérative du biais, en lissant les résidus obtenus à chaque étape. Nous présentons dans cette communication un exemple d'application sur la consommation d'électricité française pour lequel cette méthode se révèle performante.

Abstract

Réseau de Transport d'Electricité (RTE), in charge of operating the French electric transportation grid, needs an accurate forecast of the electricity consumption in order to operate it correctly. The forecasts used everyday result from a model combining a nonlinear parametric regression and a SARIMA model. In order to obtain an adaptive forecasting model, nonparametric forecasting methods have already been tested without real success. In particular, it is known that a nonparametric predictor behaves badly with a great number of explanatory variables, what is commonly called the curse of dimensionality. We use a method which iteratively corrects the bias initial estimator by an estimate of the latter obtained by smoothing the residuals. We apply the method to the french electricity consumption and show that the method compares favourably with existing procedures.

Mots clés

Régression non paramétrique, Méthode itérative, Lissage

Keywords

Nonparametric regression, Iterative procedure, Smoother

1 Problématique

L'électricité ne se stockant pas, la production globale sur le réseau électrique français doit, à tout moment, être strictement égale à la consommation. Toute modification de la demande ou de la production d'électricité en un point du réseau se répercute instantanément sur tout le système électrique. Celui-ci doit donc s'adapter en permanence pour satisfaire l'équilibre offre-demande.

La prévision à court terme est un facteur essentiel pour assurer l'équilibre entre la production et la consommation durant la journée. La prévision faite par RTE permet de garantir l'équilibre physique global entre l'offre et la demande à l'échelle du pays. De sa qualité dépendent le niveau des marges d'exploitation nécessaires et plus largement la sûreté du système.

La solution actuellement en place s'explique historiquement. Les séries temporelles ont connu un fort essor, notamment grâce à Box et Jenkins qui ont développé leurs modèles SARIMA. Leur ouvrage a constitué une référence théorique, mais également un guide pratique assez riche, et ce particulièrement dans le domaine électrique. Choisir SARIMA pour effectuer les prévisions était donc a priori un choix techniquement pertinent et quasiment naturel. Malheureusement, ces modèles n'intègrent pas de variables exogènes et il a donc fallu trouver un biais en leur adjoignant un modèle de régression (liant la puissance consommée avec les variables climatiques et les signaux de prix).

Au final, la prévision journalière s'effectue en trois étapes :

- On corrige la courbe de charge de l'impact du climat et des prix, afin de se ramener à une série temporelle ne dépendant plus des variables explicatives. On utilise à cet effet un modèle de régression, qu'on surnomme communément dans notre cas modèle de "correction".
- On utilise un modèle SARIMA sur la série ainsi "corrigée" ; la seule dynamique de la série permet de la prévoir.
- On réutilise le modèle de correction afin d'ajouter à la courbe "corrigée" prévue l'impact futur du climat et des prix.

Même si les résultats opérationnels sont satisfaisants la plupart du temps, cette méthodologie présente un certain nombre d'inconvénients.

Le modèle actuellement en place est complexe et il est très difficile de modifier sa structure. Si on veut tester par exemple l'incorporation d'une nouvelle variable explicative, il est peu probable d'y parvenir facilement, la juxtaposition du modèle de régression et du modèle SARIMA ne permettant pas d'isoler les causes de l'erreur finale de prévision.

De plus, une prévision même en temps qu'outil d'aide à la décision, devrait fournir un intervalle de confiance sur la prévision, ce qui est impossible pour l'instant ; les méthodes

de bootstrap pourraient être envisagées mais les dépendances temporelles complexes et les multiples variables exogènes rendent la tâche complexe.

Le non-paramétrique constitue une alternative intéressante, en ce sens qu’il ne spécifie pas la fonction de lien. On peut a priori incorporer de nouvelles variables explicatives et simplement réestimer une fonction de lien de dimension plus élevée. Il est également possible d’élaborer des intervalles de confiance en se basant sur l’historique des erreurs de modélisation. L’inconvénient majeur de ces techniques réside dans le fléau de la dimension ; on sait en effet que des problèmes de convergence se posent si la dimension du modèle s’accroît (en pratique, ce problème de convergence va se traduire par des estimateurs de piètre qualité). Citons les travaux de Poggi [3] sur la prévision de la consommation d’électricité corrigées des aléas climatiques et tarifaires.

Des alternatives semi-paramétriques existent, que ce soit l’approximation fonctionnelle ou la réduction de dimension. La méthode employée ici reste non-paramétrique, et utilise en fait le fléau de la dimension. En effet une autre façon d’interpréter le fléau de la dimension consiste à voir que les estimateurs non paramétriques en dimension élevée sont tous biaisés ; nous appliquons donc une méthode itérative de réduction du biais à la prévision non-paramétrique de séries temporelles.

2 Présentation de la méthode

On se place dans un contexte de régression et on désigne par $\{(X_i, Y_i)\}_{i=1}^n$ n observations.

Supposons que les données $(X_i, Y_i) \in \mathbb{R}^d \times \mathbb{R}$ sont liées via le modèle de régression

$$Y_i = m(X_i) + \varepsilon_i, \quad i = 1, \dots, n,$$

où $m(\cdot)$ est la fonction inconnue et les erreurs sont indépendantes, de moyenne nulle et de variance σ^2 . Ce modèle peut s’écrire sous forme matricielle

$$Y = m + \varepsilon.$$

Un lisseur linéaire s’écrit simplement sous la forme

$$\widehat{m}_1 = SY,$$

où S est matrice de lissage de taille $n \times n$ et $\widehat{m} = \widehat{Y} = (\widehat{Y}_1, \dots, \widehat{Y}_n)^t$, sont les valeurs ajustées. En général, un lisseur est paramétré par un paramètre de lissage λ (la fenêtre pour un lisseur à noyau, la pénalité pour les splines de lissage...). Nous écrirons le lisseur de façon à qu’une valeur élevée de λ corresponde à un estimateur très lisse. Classiquement, il faut ensuite sélectionner le paramètre λ .

Nous partons d'un estimateur lisse (λ grand), et calculons les résidus à cette première étape et nous corrigeons l'estimateur initial ou pilote en enlevant les résidus lissés (par simplicité nous conserverons le même lisseur à chaque étape), cela donne

$$\begin{aligned} R_1 &= Y - \hat{m}_1 = (I - S_1)Y \\ \hat{m}_2 &= \hat{m}_1 + S_1 R_1 \\ &= (S_1 + S_1(I - S_1))Y. \end{aligned}$$

Nous itérons cette procédure et l'estimateur à l'étape k peut s'écrire

$$\begin{aligned} \hat{m}_k &= S_1 Y + S_1(I - S_1)Y + \dots + S_1(I - S_1)^{k-1}Y \\ &= S_1[Y + (I - S_1)Y + \dots + (I - S_1)^{k-1}Y] \\ &= S_1 \hat{\beta}_k. \end{aligned}$$

Cette formulation permet une évaluation de la fonction en tout point x via

$$\hat{m}_k(x) = S_1(x) \hat{\beta}_k.$$

On peut montrer que le comportement de l'estimateur dépend des valeurs singulières de $I - S_1$. Cependant cet estimateur n'a un intérêt que si nous pouvons choisir le nombre d'itération k . Ce choix est obtenu en utilisant le critère de validation croisée généralisée

$$\hat{k}_{GCV} = \arg \min_{k \in \mathcal{K}} \left\{ \log \hat{\sigma}^2 - 2 \log \left(1 - \frac{\text{trace}(S_k)}{n} \right) \right\}$$

Les résultats théoriques concernant cet estimateur se trouvent dans [1].

3 Mise en œuvre pratique

Les premières données sur lesquelles nous avons mis en œuvre cette méthode, sont les consommations horaires d'électricité corrigées des aléas climatique et tarifaire. L'objectif consiste à prévoir une journée entière de charge à partir des 2 semaines précédentes, ce qui constitue un historique relativement faible, propice à introduire un biais important sur le prédicteur non-paramétrique.

Dans notre problème nous souhaitons prévoir une série chronologique $\{Z_t\}_{t=1}^n$. Pour prévoir le futur de cette série sans autre information, nous avons décidé d'estimer un modèle autorégressif non-linéaire de mémoire 24, le modèle général s'écrit donc

$$Z_H = f(Z_{H-1}, Z_{H-2}, \dots, Z_{H-24}) + \varepsilon_H$$

où $f : \mathbb{R}^{24} \mapsto \mathbb{R}$ est une fonction inconnue. Pour estimer cette fonction, nous utilisons 14 jours d'historique. Nous considérons un estimateur de cette fonction, sans préjuger de

l'existence d'éventuelles interactions entre les heures. Pour cela, nous allons utiliser un estimateur non-paramétrique par la méthode des noyaux.

Appelons Y le vecteur à $n = 336$ mesures contenant les charges de 14 jours. Construisons les vecteurs à $n = 336$ mesures X_1, X_2, \dots, X_{24} contenant les consommations de ces 14 jours (et plus) décalées d'une heure, de deux heures, etc. La matrice regroupant ces 24 vecteurs est appelée X et contient sur la ligne i les charges horaires décalées d'une heure à chaque fois. Sur les données dont nous disposons, le modèle s'écrit sous forme vectorielle comme suit :

$$Y = f(X_1, \dots, X_{24}) + \varepsilon$$

Nous utilisons le noyau produit gaussien. Pour $x \in \mathbb{R}^{24}$, nous avons :

$$K_h(x - X_i) = \prod_{j=1}^{24} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{(x_j - X_{ij})^2}{h} \right\}$$

où X_i représente la i -ème ligne de la matrice X et $h \in \mathbb{R}^+$ est la largeur de fenêtre (à déterminer).

L'ajustement du vecteur des observations Y par la méthode des noyaux est alors :

$$\hat{Y} = \mathbb{K}Y$$

où \mathbb{K} est la matrice $n \times n$ suivante :

$$\mathbb{K} = \begin{bmatrix} K_h(0)/w_1 & K_h(X_1 - X_2)/w_1 & K_h(X_1 - X_3)/w_1 & \dots \\ K_h(X_1 - X_2)/w_2 & K_h(0)/w_2 & K_h(X_2 - X_3)/w_2 & \dots \\ \vdots & & \ddots & \dots \end{bmatrix}$$

avec $w_i = \sum_{j=1}^n K_h(x - X_i)$.

Pour appliquer la méthode itérative de réduction du biais, nous partons d'un estimateur biaisé obtenu à partir des largeurs de fenêtre $\{h_j\}_{j=1}^{24}$ trop larges. Dans la seconde étape, les résidus $Y - \hat{Y}$ sont alors réajustés par la même méthode. Comme les largeurs de fenêtre sont trop élevées, les résidus de cette seconde étape contiennent encore du signal et nous pouvons lors d'une troisième étape les réajuster par la même méthode. Nous itérons donc ce procédé et nous nous arrêtons à l'étape K , choisie par validation croisée généralisée.

Le choix du nombre d'étapes k est conduit grâce au critère de validation croisée généralisée. Ce modèle permet donc en fonction des charges de 0 à 23h d'une journée j de prévoir la charge à 0h de la journée $j + 1$. Cette prévision notée \hat{Z}_0 est simplement :

$$\hat{Z}_0 = \hat{f}(Z_{23}, Z_{22}, \dots, Z_0)$$

Pour prévoir la charge du jour $j + 1$ de 1h à 23h, il suffit d'utiliser l'estimation \hat{f} de f inconnue et de calculer

$$\hat{Z}_1 = \hat{f}(\hat{Z}_0, Z_{23}, \dots, Z_1)$$

jusqu'à :

$$\hat{Z}_{23} = \hat{f}(\hat{Z}_{22}, \hat{Z}_{21}, \dots, \hat{Z}_0, Z_{23})$$

Nous pouvons alors calculer pour chaque jour des critères de qualité de prévision comme le MAPE ou le RMSE. Nous donnerons les résultats numériques obtenus et la comparaison avec les méthodes SARIMA, du noyau classique et MARS. Nous montrerons ainsi les apports de la méthode sur notre problématique.

Enfin, nous présenterons les premiers résultats sur la consommation d'électricité globale (non corrigée des aléas climatique et tarifaire) qui nécessite la prise en compte de variables exogènes, la température entre autre.

Bibliographie

- [1] Cornillon P. A., Hengartner N., Matzner-Løber E. (2009) Recursive Bias Estimation for high dimensional regression smoothers, soumis.
- [2] Friedman, J. (1991) Multivariate adaptive regression splines, *The Annals of Statistics*, Vol. 19 337–407.
- [3] Poggi J. M. (1994) Prévision non-paramétrique de la consommation électrique, *Revue de Statistique Appliquée*, Vol. 42 N°4 83-98.