

Estimation spatio-temporelle de consommation électrique

Jiali Mei, Yannig Goude, and Georges Hébrail

EDF R&D

13 mars 2015

Suite au développement de moyens de production électrique décentralisée (notamment de l'énergie renouvelable), il est de plus en plus important de connaître et anticiper l'évolution de la consommation électrique à une échelle locale. Cette connaissance à une maille géographique et temporelle fine est utile à la fois pour la gestion du réseau électrique et pour la compréhension des activités économiques.

Pour prendre en compte des caractéristiques locales, il est a priori pertinent d'inclure des informations permettant d'expliquer l'hétérogénéité de la consommation électrique locale, notamment les données infra-communales de l'INSEE.

L'utilisation de différentes bases de données spatiales se heurte au problème d'incompatibilité des découpages spatiaux (Gotway and Young (2002)). Entre l'IRIS, l'entité de diffusion statistique de l'INSEE, et le niveau du réseau électrique où la consommation est mesurée à une échelle temporelle fine, les découpages sont différents. Nous proposons une méthode permettant de contourner cette difficulté.

Dans cette étude, nous considérons la consommation horaire des têtes de départ HTA (haute tension A) dans les environs de Lyon. La consommation est localisée et donc située à l'intérieur des IRIS de manière approximative par les postes HTA/BT (haute tension A/basse tension) liés aux départs HTA, les positions géographiques de ces postes HTA/BT étant connues. Nous disposons également des données de température locale et pour chaque IRIS d'un ensemble de variables socio-démographiques.

1 Méthodes

Méthode d'interpolation Ayant accès à ces informations, une méthode naïve consiste à distribuer la consommation aux IRIS proportionnellement à la population. Le poids proposé prend en compte le nombre de postes HTA/BT liés à chaque tête de départ dans chaque IRIS. Nous supposons que la population se répartie à parts égales entre les postes HTA/BT dans un IRIS et que la consommation d'une tête de départ se répartie à parts égales entre les clients qui y sont raccordés. L'estimation de la consommation se fait alors par une règle de trois.

Bien que cette méthode est simple à réaliser et propose des résultats légitimes, elle ne tient pas compte de l'aspect temporel du problème et ne se généralise pas facilement si l'on dispose d'informations socio-démographiques plus fines.

Modèle de consommation locale Orthogonalement à l'interpolation spatiale, on peut explicitement modéliser les séries temporelles de consommation électrique par des modèles additifs généralisés (GAM, Hastie and Tibshirani (1986)). Les variables explicatives qui entrent dans ces modèles sont classiquement : les variables météorologiques (température, nébulosité...) et les variables calendaires telles que le type de jour, l'heure de la journée, la position dans l'année (voir par exemple Goude et al. (2013)).

Afin d'estimer la consommation d'une zone géographique par un modèle de prévision, il faut trouver un critère permettant de lier la zone cible à un modèle de consommation. Nous proposons une approche de clustering basée sur les variables socio-démographiques. Nous attribuons à chaque

IRIS ou départ HTA une étiquette de classe. Les IRIS et les départs HTA de la même classe ont des caractéristiques socio-démographiques similaires ainsi que le même modèle de consommation.

La méthode d'estimation comprend trois étapes : dans un premier temps, nous construisons des clusters à partir des variables socio-démographiques ; ensuite, un modèle de consommation est appris par cluster, sur les données de consommation des départs HTA ; enfin, le modèle est appliqué sur les IRIS pour proposer une estimation de consommation.

2 Évaluation

Évaluation dans un contexte de prévision Les consommations des IRIS n'étant pas mesurées, l'évaluation des méthodes d'estimation ne peut être effectuée directement. Il est toutefois possible d'évaluer les modèles de consommation locale dans un contexte de prévision, sur les données de départs HTA. Pour ce faire, les dernières observations sont extraites de l'échantillon d'apprentissage, et utilisées comme données de test. De plus, la consommation d'une partie des départs HTA est entièrement exclue de l'échantillon d'apprentissage. La qualité du modèle appris est mesurée par le taux d'erreur sur l'estimation de la consommation sur les observations du jeu de test. Le taux d'erreur sur les départs HTA d'apprentissage mesure donc l'erreur de prédiction, tandis que le taux d'erreur sur les données de test est la somme d'une erreur de prédiction et d'une erreur de transportabilité du modèle. Cette dernière est une estimation de l'erreur commise par notre méthode sur la consommation des IRIS.

Jeu de données simulées Afin de compléter l'évaluation indirecte par la prévision, des courbes de consommation sont simulées pour chaque poste HTA/BT puis agrégées aux IRIS et aux départs HTA.

Les simulations sont faites à partir des profils de consommation validés par les régulateurs provenant de la méthode de profilage utilisée par ERDF pour estimer la consommation d'un groupe de clients pour la facturation des fournisseurs. Des profils de consommations, en fonction de la température et des différents cycles de consommation, sont publiés régulièrement. Dans notre méthode de simulation, à chaque poste HTA/BT est donné un type de consommation en fonction des statistiques démographiques et économiques de la zone desservie. La consommation locale est ensuite simulée en utilisant les profils.

Ce jeu de simulation permet de tester si les méthodes proposées détectent les liens entre la démographie et la consommation locale. Il est possible non seulement d'évaluer les modèles de consommation locale en prévision pour les départs HTA, comme sur les données réelles, mais aussi de comparer l'estimation de consommation des IRIS et la simulation.

3 Résultats

Les résultats seront présentés et commentés.

Bibliographie

- Gotway, C. A. and Young, L. J. (2002). Combining incompatible spatial data. *Journal of the American Statistical Association*, 97(458) :632–648.
- Goude, Y., Nedellec, R., and Kong, N. (2013). Local short and middle term electricity load forecasting with semi-parametric additive models. *IEEE transactions on smart grid*, 5 , Issue : 1 :440 – 446.
- Hastie, T. and Tibshirani, R. (1986). Generalized additive models. *Statistical science*, pages 297–310.