Texte 3 :

**Short-term load forecasting based on a semi-parametric additive model**

On décrit une méthode d’estimation de la demande, consommation électrique du Australian National Electricity Market.

On veut estimer les demandes futures par tranches de demie heures jusqu’à 7 jours en avance.

Pour cela on fera un modèle additif semi-paramétrique.

Variables explicatives :

* Calendar variables
* Lagged actual demand observations.
* Historical and forecast temperature for one or more sites.

On obtient ainsi des Point forecats (pérédiction des moyennes et médianes de al distribution de la demande future).

On complète cela par une prédiction de la densité de distribution afin de déterminer des IC. Méthode : bootdtrap modifié pour l’adéquation avec la saisonnalité de la demande d’électricté.

Jusqu’à présent :

* Modèle de régressions linéaires
* Modèle de processus stochastiques.
* Exponential smoothing and Arima Models.
* Mais aussi des modèles qui prennent en compte la relation non-linéaires entre demande d’électricité et variables exogènes (températures) : ANN, machine learning.

**Le papier propose une méthode par régression mais STLF : le but est de prendre en compte les termes non-linéires et non-paramétriques dans le modèle de régression.**

La validité du modèle sera testée par une comparaison hors échatillon avec des données réelles. Et ca a produit des prédictions correctes en Victoria et Australie du Sud.

**Méthodologie**:

Notre modèle inclut des termes non linéires et non paramétriques et des serially correlated errors.

Comme la consommation d’électricité varie selon le moment de la journée, on créé 48 modèles séparés pour chaque demie-heure.

La variables expliquée est le log naturel de la consommation d’électricé.

Variable explicatives :

* effet de calendrier (donc prend en compte saisonnalité) : on considère le fait d’être à un jour en prticulier, à un moment donné de l’année grâce à une smooth fonction , en vacances ou non.
* effet de température : les températures récentes ont un effet sur la demande. Mais comme les températures entre deux sites sont corrélées, on prend la moyenne et la différence des deux.
* Effet de demande lagged : en considérant les demandes récentes, on capture la corrélation serialles dans les séries temporelles de demande.
* Terme d’erreur : le terme d’erruer ne sera pas serially correlated au sein de chaque modèle, grâce à la D lagged, mais les termes d’erreur de chaque modèle sont légèrement corrélés.

B) Prédiction de la distribution :

La prédiciton de la distribution permet de résoudre les risques dûs à la vlatilité de la demande d’électricté, notamment par l’estimation d’IC et de la validté de la prédiciton.

On a des méthodes paramétriques ou non-paramétriques.

Pb des méthodes paramétriques : on fati l’hypothèse que les bruits sont blancs, ce qui est rejeté par le test de Anderson-Darling