



CNRM, UMR 3589

SEMINAIRE CNRM
N° 2018_19

mercredi 12 décembre 2018 à 10h

**APPRENTISSAGE STATISTIQUE SUR DES DONNÉES
SPATIALES DE CONSOMMATION ÉLECTRIQUE**

par Benjamin DUBOIS

(ENPC/RTE)

en salle Joël Noilhan

Résumé :

Dans cette présentation, nous nous intéressons à la prévision J+1 de la consommation électrique désagrégée en France. Notre travail a pour but de modéliser les séries temporelles non-stationnaires des consommations aux différents points de livraison en fonction de variables explicatives endogènes (passé des séries) et exogènes (températures, dates ...). En plus de la précision du modèle que différentes mesures de performances permettent d'évaluer, nous tenons d'une part à ce que notre modèle soit facilement interprétable et d'autre part à maîtriser le temps de calcul de l'étape apprentissage.

Pour cela, nous avons choisi de construire un unique modèle multivarié avec une seule étape d'optimisation favorisant un partage d'information entre les différents sites de consommation. Les similarités entre les différentes séries semblent en effet justifier un tel apprentissage par transfert. Formellement, nous cherchons un sous-espace de faible dimension au sein d'un ensemble de splines multivariées qui permet une représentation parcimonieuse des différentes consommations. Cette approche nous permet d'intégrer les interactions entre les variables d'entrée tout en conservant une complexité raisonnable. Afin d'en extraire un sous-ensemble pertinent, nous avons étudié la régression de rang faible avec différents termes de pénalisation dont le but est d'effectuer de la sélection de variables et de favoriser la régularité de notre modèle de prévision. Cela est essentielle compte tenu des importantes variations au sein de la base de données au fil du temps. La représentation graphique des résultats obtenus nous permet de souligner les difficultés liées à ce problème spécifique de prévision ainsi que le besoin d'aller plus loin qu'un modèle générique de prévision. En particulier, nous proposons d'affiner l'intégration de l'information récente aux entrées du modèle ainsi que la modélisation de leurs interactions.