



# Stage de recherche: Agrégation de scores pour la prévision probabiliste de séries temporelles

**Durée** 3 à 6 mois

**Environnement** laboratoire LPSM de Sorbonne Université, encadré par Y. Goude (EDF) et O. Wintenberger (Sorbonne).

**Rémunération** dédommagement de stage selon la grille des rémunérations SU (environ 500 euros/mois)

**Sujet de stage** Le nombre de modèles numériques pour prévoir des observations futures, à des échelles temporelles variables et des échelles spatiales complexes est en augmentation. Un point essentiel est de savoir déterminer quels modèles ont le meilleur pouvoir prédictif. Ce besoin devient complexe si on doit prendre en compte plusieurs aspects de la prédiction. Par exemple, les prévisions de consommation électriques effectuées par EDF pour ajuster la production sont évaluées selon divers critères, des critères statistiques mais également des critères plus métiers en lien avec les marchés (gains ou pertes financière, coûts marginaux...).

A ce problème de modélisation statistique multivariée s'ajoute la question de comparer des prévisions qui ne sont ni stationnaires dans l'espace, ni dans le temps. Une autre difficulté est de savoir qui doit choisir le score pour départager différents modèles de prévisions. Ainsi dans la gestion d'un réseau électrique, la prise en compte du risque extrême est liée aux prévisions pouvant être effectuées à différents nœuds du réseau incluant diverses aléas (température pour les consommateurs, vent et rayonnement pour la production renouvelable, forts de moyens de productions, événements climatiques extrêmes) mais aussi aux moyens de pilotage du réseau nécessitant des calculs relativement complexes.

En résumé, l'évaluation d'une prévision probabiliste est un problème complexe et a donné lieu à la création de multiples scores ou tests pour évaluer et/ou valider ces prévisions. Parallèlement à cela, des méthodes d'agrégation d'experts ont été développées et sont devenues très populaires dans le domaine de la prévision de séries temporelles (voir par exemple *Stoltz, G. (2010)* pour une présentation générale et des applications à la prévision de consommation électrique et à la qualité de l'air, *O. Wintenberger (2017)* pour des résultats récents concernant l'agrégation en ligne d'experts). Ces méthodes se sont révélées pertinentes pour la prévision probabiliste de consommation électrique et de prix de marché (voir *P. Gaillard, Y. Goude and R. Nedellec (2016)*).

Le stage consiste à exploiter la variété des indicateurs et scores développés dans le domaine de la prévision probabiliste par l'intermédiaire de l'agrégation d'experts. Nous proposons d'étendre les algorithmes proposés dans *Gaillard, P. (2011)*, *O. Wintenberger (2017)* et *Thorey, J., Mallet, V., and Baudin, P. (2017)* aux différents scores et scores composites existant pour la prévision probabiliste. Un axe important sera le développement de méthodes d'agrégation respectant l'ordre des quantiles (quantile crossing problem, c.f. *Bondell et al. (2010)*).

**Déroulement** Le premier mois de stage sera consacré à la revue de la littérature, le second à l'étude spécifique de l'agrégation de scores, en particulier de type coûts quantiles. Le troisième mois sera dédié à la rédaction du mémoire et à l'implémentation de la méthode sur données EDF en utilisant le package OPERA sous R.