

TP Séries Temporelles – Mini-challenge

On considère une série mensuelle de **consommation d'électricité** (en KWh) de janvier 1980 à décembre 1994 (15 ans d'historique). On considère également une série mensuelle (notée **htdd** pour heating degree day) correspondant respectivement au cumul sur un mois de degrés-jours de chauffage (quantité conçue pour quantifier la demande d'énergie nécessaire pour chauffer un bâtiment) ainsi qu'une série (notée **cldd** pour cooling degree day) correspondant au cumul sur un mois de degré-jour de climatisation.

Le but est de construire le meilleur modèle de prévision sur un an pour la consommation d'électricité. Pour cela, on met à disposition comme ensemble d'apprentissage `data_test` l'ensemble des données sauf la dernière année (de janvier 1994 à décembre 1994).

Le critère utilisé pour évaluer ce challenge est la moyenne des carrés des erreurs de prévision :

$$\text{RMSE} = \sqrt{\text{MSE}} \text{ où } \text{MSE} = \frac{1}{12} \sum_{t=1}^{12} (c_t - \hat{c}_t)^2$$

en notant c_t les consommations réelles de l'année 1994 et \hat{c}_t leurs prévisions.

Afin d'alléger le travail à réaliser, on fournit le fichier des valeurs des deux séries `htdd` et `cldd` pour l'année 1994 (fichier `data_test`). En pratique, il faudrait également prévoir ces deux séries mais on considère ici que cette prévision pourrait être réalisée avec une grande précision.

On pourra s'aider des indications suivantes :

1. Tracer les 3 séries en question. Que remarquez-vous ?
2. Envisager une régression linéaire multiple sans se soucier de l'aspect temporel.
3. Analyser la dynamique des résidus bruts obtenus.
4. Proposer un modèle de type ARMAX

On déposera sur Campus le fichier texte correspondant à vos prévisions (fichier comportant donc une seule colonne formée de vos 12 valeurs de prévision de consommation d'électricité). Mettre ce fichier au format **VotreNOM_pred.txt**.