Synthèse : Electricity consumption forecasting with outliers handling based on clustering and deep learning with application to the Algerian market

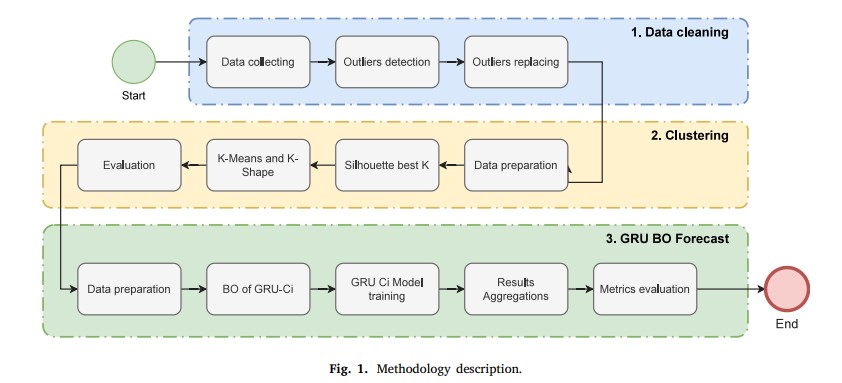
Article de 2023

3 étapes :

* Gestion des valeurs aberrantes
* Clustering avec K-Means et K-Shape
* Prédiction de chaque cluster

Dataset de 2000 consommateurs sur 14 ans à Bejaia en Algérie

Grande accuracy, faible erreur (MAPE à environ 2%)



**Dataset**

Données qui servent à la prédiction donc possèdent 13 colonnes : les 12 premières représentent les mois et servent à la prédiction de la 13e colonne.

**Remplacement des valeurs aberrantes**

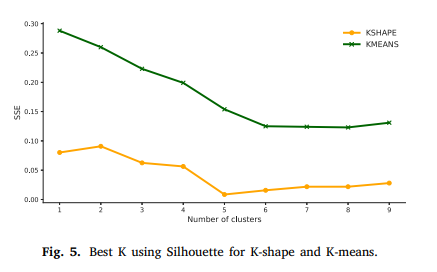
Par la méthode de Holt-Winters qui vient lisser les données et affaiblir les poids des données aberrantes

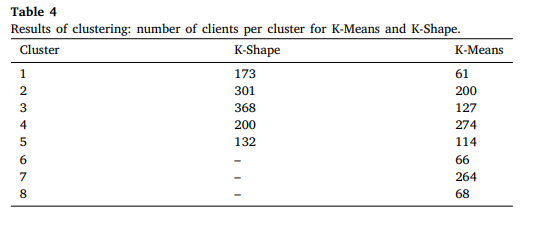
**Clustering**

Utilisation des algorithmes K-Means et K-Shape avec mesure du silhouette score.

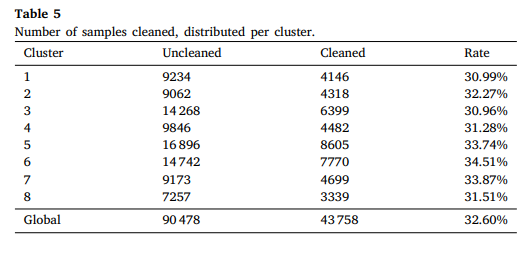
K = 5 selon K-Shape

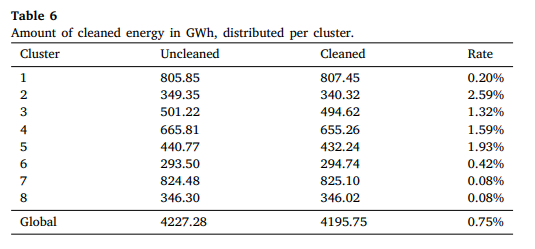
K = 8 selon K-Means





**Nettoyage**

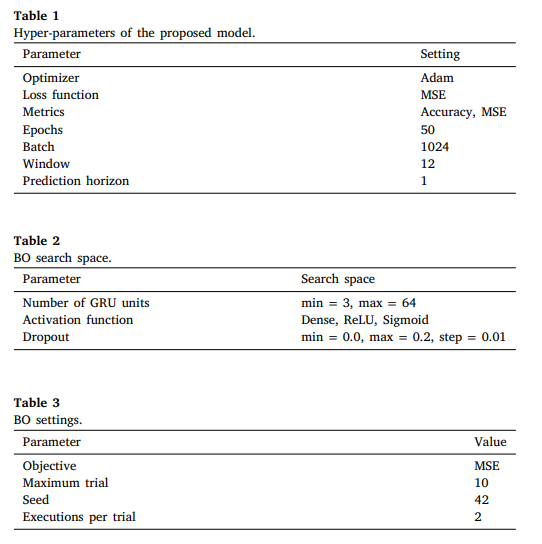




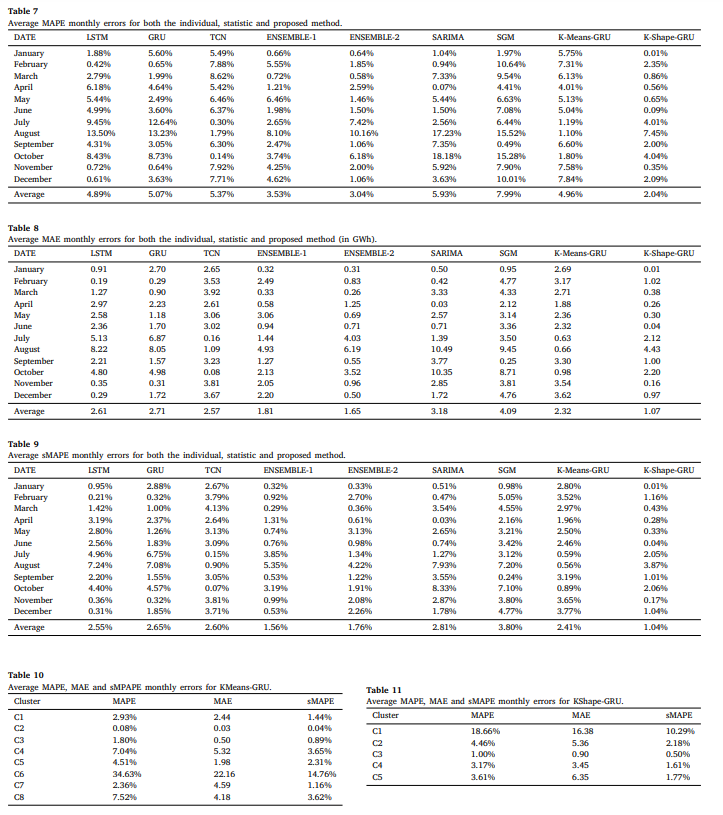
Beaucoup de données à clean mais cela représente peu d'énergie.

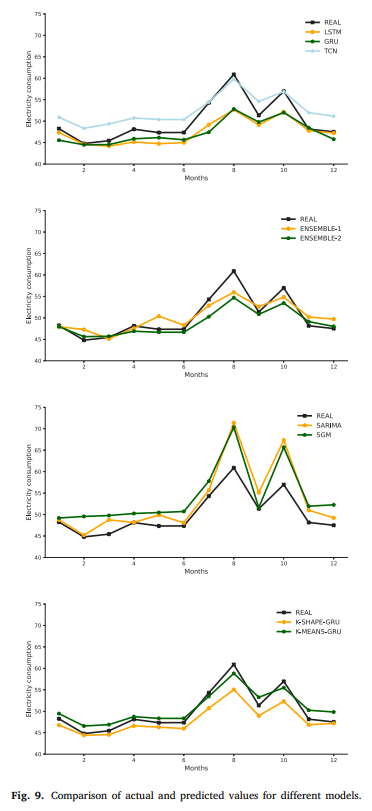
**Prédiction**

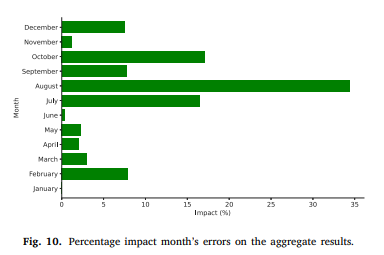
Recherche des meilleurs paramètres avec Bayesian Optimizer

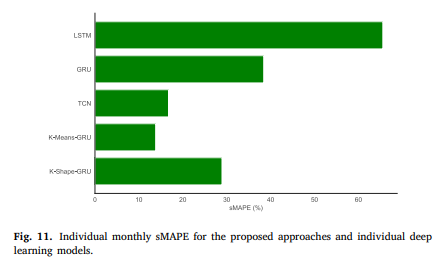


LSTM, GRU, TCN, SARIMA, SGM









Meilleures prédictions pour les entreprises de nourriture et boissons, surtout pendant le mois de ramadan.

**Limites**

Manque de données comme les facteurs météorologiques, les facteurs socio-économiques ou les évènements géopolitiques qui pourraient influencer les modèles et améliorer la précision des prédictions