**Rapport sur l’article :**

**Decentralized IoB for Influencing IoT-based**

**Systems Behavior**

par Kévin TANG

# **1 Synthèse de l’article**

## 1.1 Introduction

Les progrès dans les domaines des communications et de l'informatique ont permis l'intégration de plusieurs technologies émergentes dans diverses applications et services. L'Internet des Objets (IoT) a particulièrement bénéficié de ces avancées. On prévoit une croissance significative du marché de l'IoT industriel, avec un nombre estimé de 22,48 à 29,3 milliards de dispositifs intelligents connectés à internet d'ici 2023.

Les appareils IoT ont évolué pour devenir des dispositifs intelligents, capables d'effectuer des traitements en temps réel, de surveiller les données et de prendre des décisions cruciales. Leur omniprésence, combinée à leur capacité de traitement instantané et à l'utilisation de l'Intelligence Artificielle (IA), leur confère la particularité de pouvoir également suivre les comportements humains. Ceci correspond à un nouveau concept appelé l'Internet des Comportements (IoB), annoncé par Gote Nyman en 2012.

L’IoB repose sur l’utilisation de l’IA, du calcul distribué, de l’IoT et la coopération de l’utilisateur dans ses activités. Chaque appareil IoT représente un comportement de l’utilisateur et chaque IoB aura certaines caractéristiques basées sur différents critères psychologiques. Étant donné que plusieurs appareils IoT (par exemple les smartphones, les caméras domestiques, les appareils de surveillance générale) peuvent appartenir à un même utilisateur final, ces appareils ont tendance à transmettre des caractéristiques similaires à celles de l'utilisateur. En appliquant cinq critères de personnalité aux appareils IoT (cognition, émotion, comportement, personnalité et intercommunication), un système IoB peut être développé pour évaluer et contrôler les comportements d’un utilisateur.

Cet article propose une structure IoB décentralisée pour atteindre la durabilité énergétique en suivant, analysant et influençant le comportement des appareils IoT. L'objectif de ce cadre est de transformer le comportement des consommateurs en un comportement écologique afin de réduire la consommation d'énergie. Ce cadre intègre des techniques d'IoT, d'IA, d'analyse de données et de sciences comportementales pour obtenir des avantages pour les utilisateurs et les entreprises. La structure a été testé dans un environnement décentralisé, comparativement à un environnement centralisé, et montre comment l'influence sur les comportements des utilisateurs améliore la consommation d'énergie et les coûts.

## 1.2 Travaux connexes

Influencer le comportement des utilisateurs par l'intermédiaire des appareils IoT peut sembler anodin, mais cela peut en réalité aller de l'influence sur les préférences individuelles jusqu'à exercer un contrôle sur l'ensemble d'une foule.

Un exemple récent est celui de la pandémie mondiale de COVID-19, où l'utilisation de technologies de suivi des déplacements et de reconnaissance du port du masque a été déployée pour encourager le respect de cette mesure dans les endroits fréquentés.

Tenter d’influencer et de modifier le comportement des utilisateurs est délicat, car cela peut se heurter à des résistances et à d’autres facteurs psychologiques liés au confort et à la confiance mais certaines technologies comme l’IA explicative (XAI) peuvent aider les utilisateurs à comprendre comment les IA fonctionnent dans le but de donner une meilleure perception du système et faciliter les opérations de suivi, d’analyse et d’influence du comportement.

Dans leurs études sur la prédiction de la consommation d'énergie, Syed a développé un cadre comprenant des étapes de nettoyage des données et de construction de modèles alors que Kim et Lee ont cherché à accélérer le déploiement des modèles d'apprentissage en profondeur en utilisant une structure edge-cloud. D'autres auteurs ont exploré des applications d'algorithmes d'apprentissage par renforcement pour optimiser la consommation d'énergie, tandis que certains se sont penchés sur la structure des modèles de renforcement en proposant des cadres de jeu stochastiques pour contrôler les interactions entre la consommation d'énergie et les ménages.

## 1.3 La structure

Cette section propose un cadre IoB décentralisé visant à influencer le comportement des appareils électriques intelligents IoT afin de contrôler la consommation excessive d'énergie, dans le but d'atteindre la durabilité énergétique grâce à des techniques stables et hautement tolérantes aux pannes.

Le cadre architectural illustré dans la Figure 1 montre comment chaque appareil sera connecté à un compteur intelligent IoT, relié à un service cloud via un réseau IoT. Les données de consommation d'énergie seront collectées par les compteurs intelligents, puis stockées, prétraitées par un dispositif en périphérie, et enfin utilisées par un modèle d'IA pour prédire la consommation d'énergie dans l’heure à venir.

Un sous-système de contrôle utilisera cette prédiction pour déterminer la quantité d’énergie à consommer pour chaque appareil. En cas de prédiction de surconsommation d’un appareil, le contrôleur envoie une alerte au compteur intelligent associé via le réseau, permettant à celui-ci de limiter la consommation des appareils connectés.

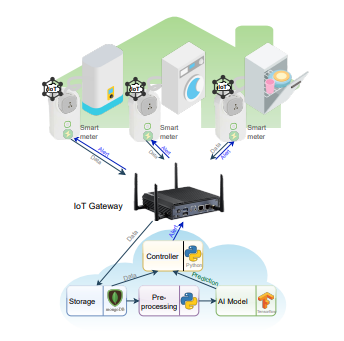


Fig. 1. Framework Architecture

## 1.4 Mise en oeuvre expérimentale

Une expérience a été menée pour mettre en œuvre le cadre proposé en construisant un système IoB distribué visant à améliorer l'efficacité énergétique en influençant le comportement des appareils IoT. Le processus comprenait le prétraitement du jeu de données, le développement d'un modèle de prédiction de séries temporelles, l'implémentation d'un contrôleur système pour analyser et contrôler le comportement des appareils IoT, et enfin la création de quatre modèles de prédiction de séries temporelles pour une structure de système IoB centralisée, afin de comparer leurs performances avec celles du modèle de prédiction décentralisé.

### 1.4.1 Les données

L'expérience a utilisé des données de consommation électrique domestique françaises, échantillonnées à la minute sur une période d'environ quatre ans. Les données comprennent des mesures telles que la date, puissance active globale, la puissance réactive globale, la tension, l'intensité globale et trois sous-compteurs correspondant à différentes utilisations dans la maison : la consommation électrique de la cuisine (sous-compteur 1), de la buanderie (sous-compteur 2) et du chauffe-eau électrique et de la climatisation (sous-compteur 3).

Pour adapter les données au fonctionnement horaire du cadre expérimental, elles ont été ré-échantillonnées à une fréquence horaire. Les valeurs nulles ont été remplacées par la moyenne des valeurs et les données ont été normalisées avec MinMaxScaler. Finalement, le jeu de données a été divisé en un ensemble d'entraînement de 26 280 heures et un ensemble de test de 8 308 heures.

### 1.4.2 Autoencoder

Un modèle d’apprentissage profond approprié doit être construit pour prendre en charge l’architecture distribuée du système. Le modèle d'IA recevra les valeurs de chaque compteur intelligent individuellement, calculera la puissance active totale pour soutenir le processus d'apprentissage du modèle, et prédira la consommation d'énergie de chaque compteur intelligent connecté. Pour cela, une architecture encodeur-décodeur séquence à séquence a été utilisée, avec des cellules LSTM, une fonction d'activation tanh, et une couche dense distribuée dans le temps pour décoder la séquence de sortie. Le modèle a été construit en utilisant TensorFlow et la bibliothèque Keras, et entraîné avec une fonction de perte de moindre carrés moyens et une fonction d'optimisation Adam sur 20 époques d'entraînement.

### 1.4.3 Contrôleur système

Le contrôleur reçoit la prédiction de consommation d’énergie de chaque appareil. Si la prédiction de consommation d’un appareil est supérieure à la consommation moyenne historique du même appareil au cours des dernières années au cours de la même heure, du même mois et du même jour de la semaine, alors le contrôleur envoie une alerte au compteur associé à l’appareil pour limiter sa consommation d’énergie.

### 1.4.4 Système IoB centralisé

Pour évaluer l’efficacité de la structure proposée, le système Iob distribué a été comparé à un système IoB centralisé. Quatre modèles ont été proposés pour prédire les valeurs de la puissance active globale et des 3 compteurs avec les mêmes paramètres que précédemment.

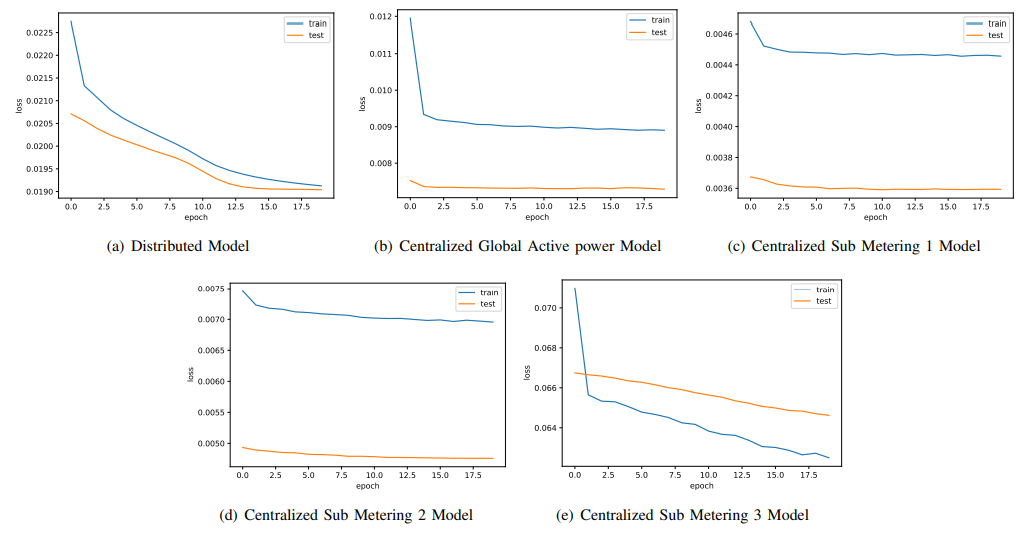
## 1.5 Evaluation et résultats

Cette section évalue la comparaison entre un système IoB distribué et un système IoB centralisé en fonction des performances de leur modèle de prédiction, des taux d'erreur, de la précision des alertes système et de la quantité totale d'énergie économisée résultant de l'influence du comportement de l'IoB.

### 1.5.1 Performance du modèle : fonction de perte

La fonction de perte utilisée pour optimiser le modèle est l'erreur quadratique moyenne. Il calcule les différences quadratiques moyennes entre les valeurs attendues et les valeurs réelles.

Dans le graphique (a), on voit que les courbes convergent, ce qui indique une bonne performance du modèle. Dans les graphiques (b), (c) et (d), les deux courbes sont très éloignées l’une de l’autre, suggérant ainsi un problème de sous-ajustement tandis que le graphique (e) présente une perte plus importante sur les données de test par rapport aux données d'entraînement, montrant un problème de sur-ajustement.

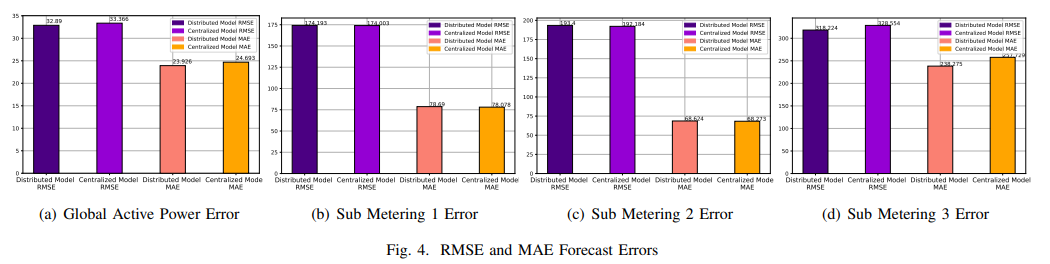


### 1.5.2 : Erreurs de prévision

La différence de résultat entre les valeurs réelles et les valeurs prédites est mesuré à l’aide de deux indicateurs : l’erreur quadratique moyenne (RMSE) et l’erreur absolue moyenne (MAE).

Comme le montre les graphiques 4(a) et 4(d), le modèle centralisé obtient un résultat d’erreurs plus élevé concernant la puissance active globale ainsi que les compteurs 1 et 3.

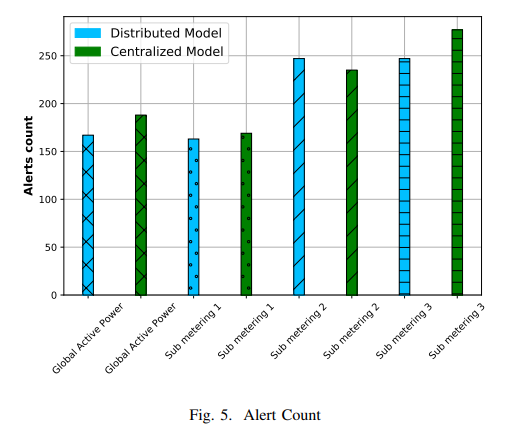
Tandis que les erreurs se produisent légèrement plus souvent dans le système centralisé par rapport au système distribué pour les compteurs 1 et 2, d’après les graphiques 4(b) et 4(c).



### 1.5.3 Les alertes

Comme dit précédemment, une alerte est émise à chaque fois que le modèle prédit une sur-consommation d’un appareil.

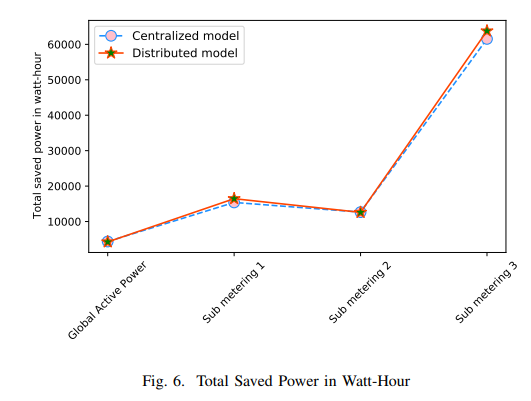
Selon les observations de la figure 5, le système centralisé génère un nombre plus élevé d'alertes, ce qui suggère une précision accrue et un nombre plus faible d’erreurs par rapport au système décentralisé.



### 1.5.4 Energie totale économisée

Grâce au système d’alerte, il est donc possible de limiter la consommation d’énergie au niveau de la moyenne historique pour éviter la sur-consommation. Cela nous permet donc de calculer l’énergie totale qui peut être économisée avec ce système.

On peut voir que le montant d’énergie économisé est pratiquement identique, que ce soit dans le système centralisé que le système distribué. Il y a seulement deux légères différences concernant les compteurs 1 et 3.



En conclusion de cette partie, on peut dire que le système distribué possède de meilleures performances que le système centralisé car le taux de perte converge plus rapidement entre les données d’entraînement et de test, les erreurs sont moins fréquentes et cela lui permet d’économiser plus d’énergie.

## 1.6 Conclusion

Dans le but de promouvoir la durabilité énergétique grâce à la technologie IoB (Internet of Behavior) qui influe sur le comportement de l'IoT (Internet des objets) pour atteindre certains objectifs, cet article présente un nouveau cadre IoB décentralisé pour la consommation électrique domestique. De plus, il propose une mise en œuvre expérimentale du cadre IoB décentralisé proposé et une évaluation approfondie comparée à un système IoB centralisé.

Les résultats démontrent que le système IoB décentralisé offre de meilleures performances en termes de précision de prédiction, réduisant ainsi les alertes de 3,5 % et économisant plus d'énergie de 3,4 % pour 3 sous-compteurs par rapport au système centralisé.

Quelques points à investiguer à l’avenir :

* Évaluer l’efficacité du réseau sur un plus grand nombre d’appareils connectés
* Évaluer l'architecture du modèle de prédiction sur un plus grand nombre de fonctionnalités d’entrée
* Intégrer différentes techniques de support à la décentralisation telles que l'apprentissage fédéré et le Fog Computing.

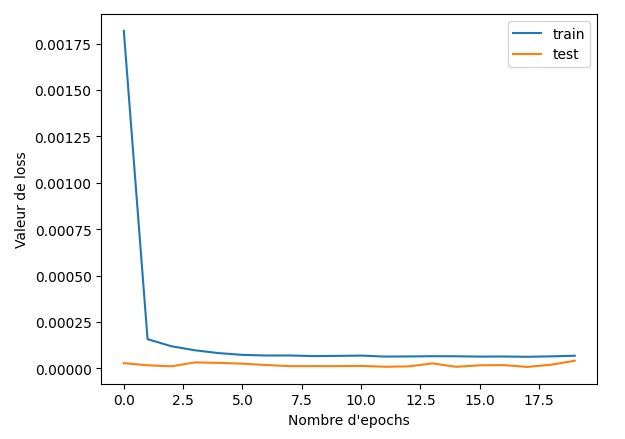
# **2 Expérimentation**

## 2.1 Résultats

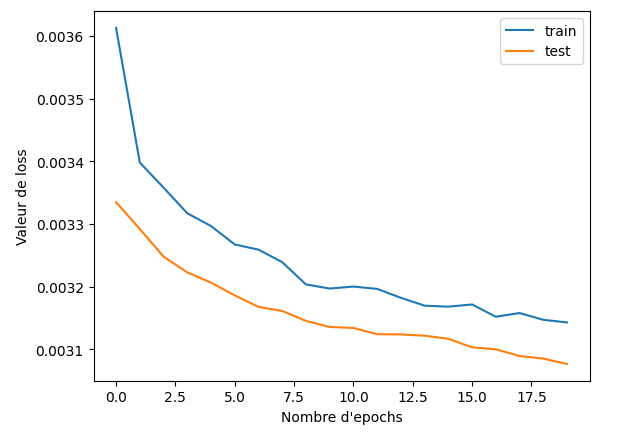
### 2.1.1 Fonction de perte

Je ne retrouve pas exactement les mêmes résultats pour le système centralisé. Cependant, je remarque que les graphiques obtenus pour chaque caractéristique du système distribué sont à peu près similaires à ceux attendus pour le système centralisé. Je me demande si cette observation résulte d'une erreur de ma part ou si elle est liée à une inexactitude dans l'article.

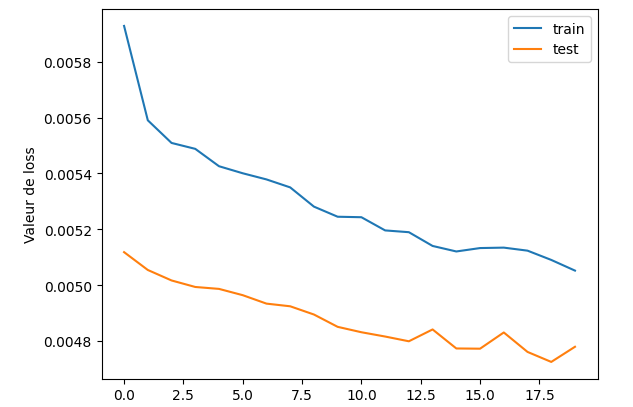
* Centralized Global Active Power Model :



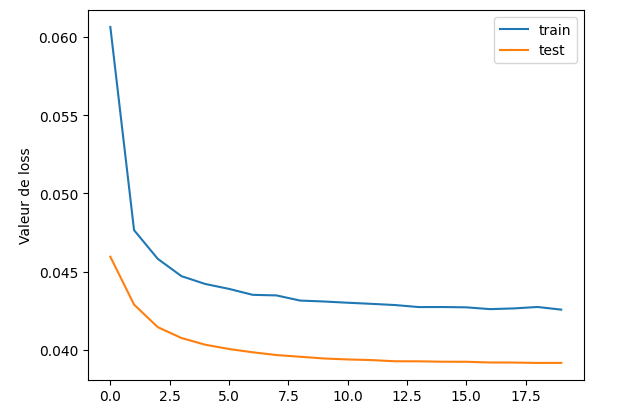
* Centralized Sub Metering 1 Model :



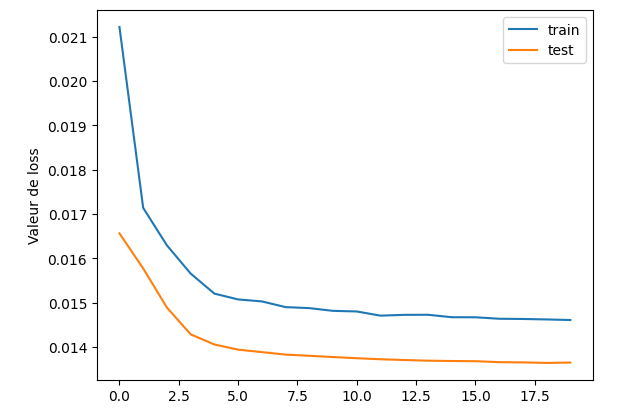
* Centralized Sub Metering 2 Model :



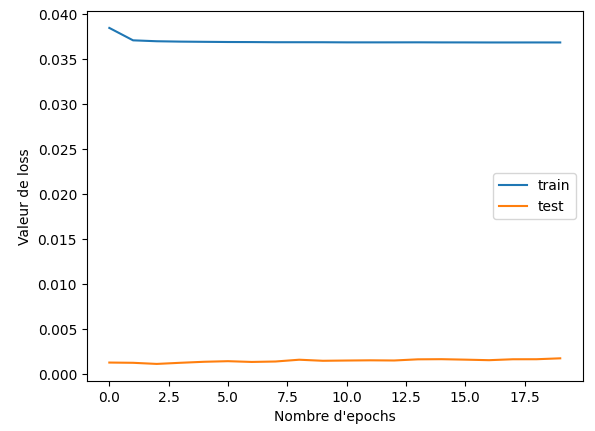
* Centralized Sub Metering 3 Model :



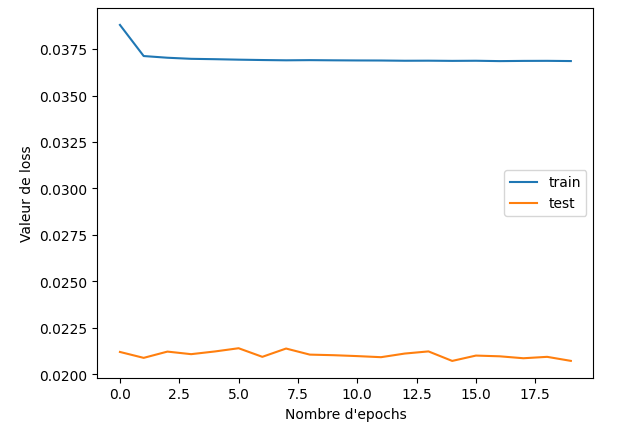
* Distributed Model :



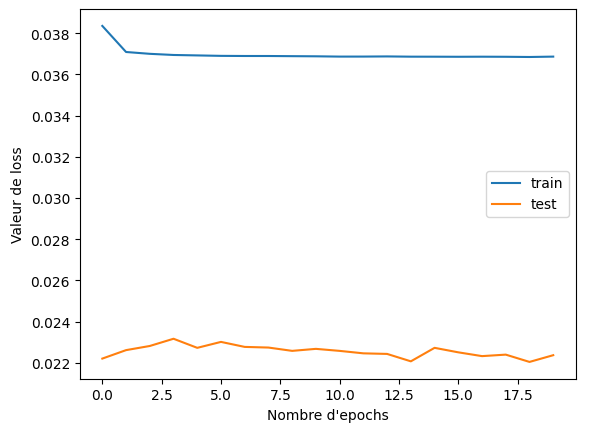
* Distributed Global Active Power Model :



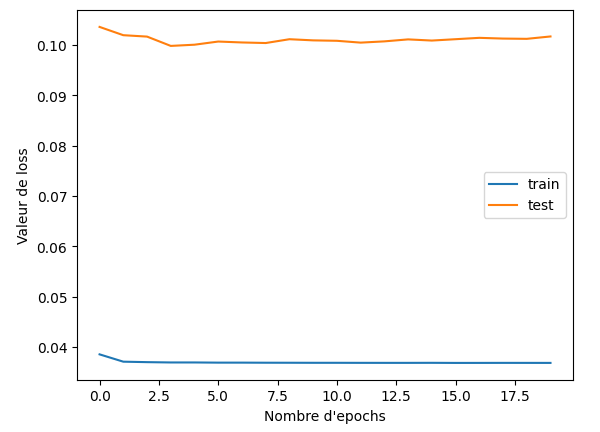
* Distributed Sub Metering 1 Model :



* Distributed Sub Metering 2 Model :

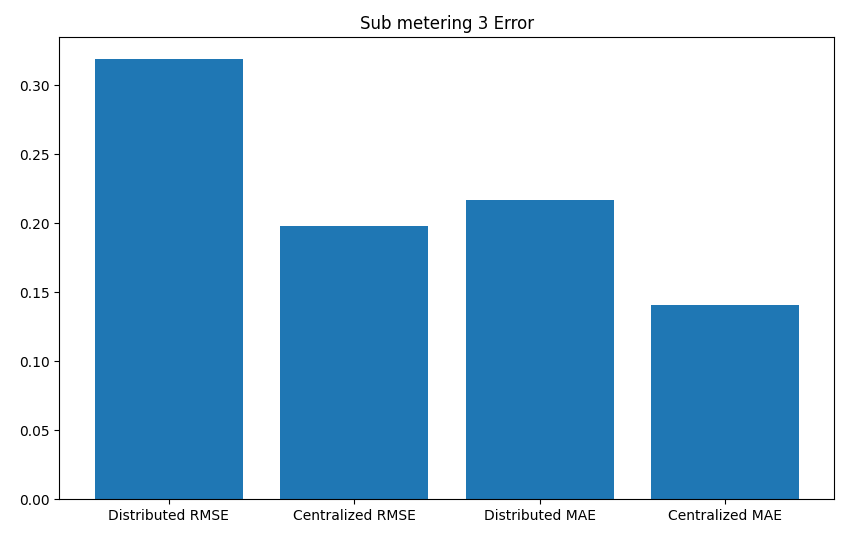
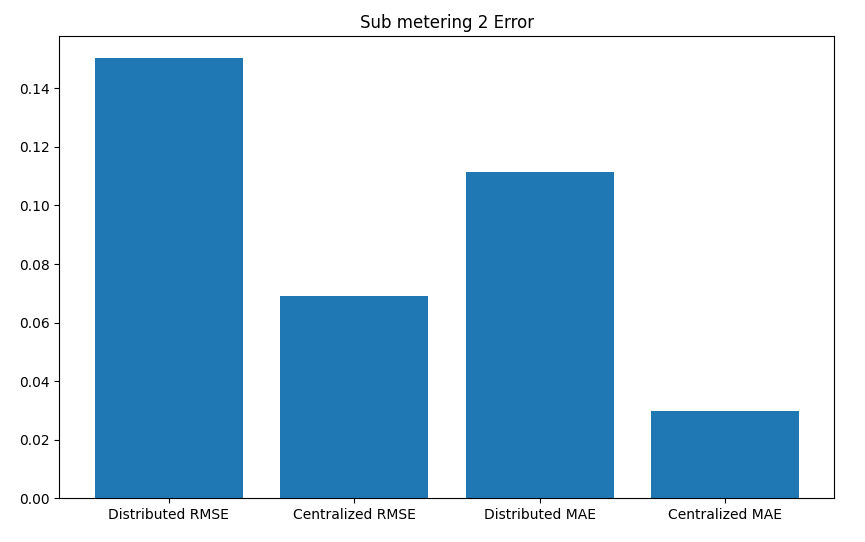
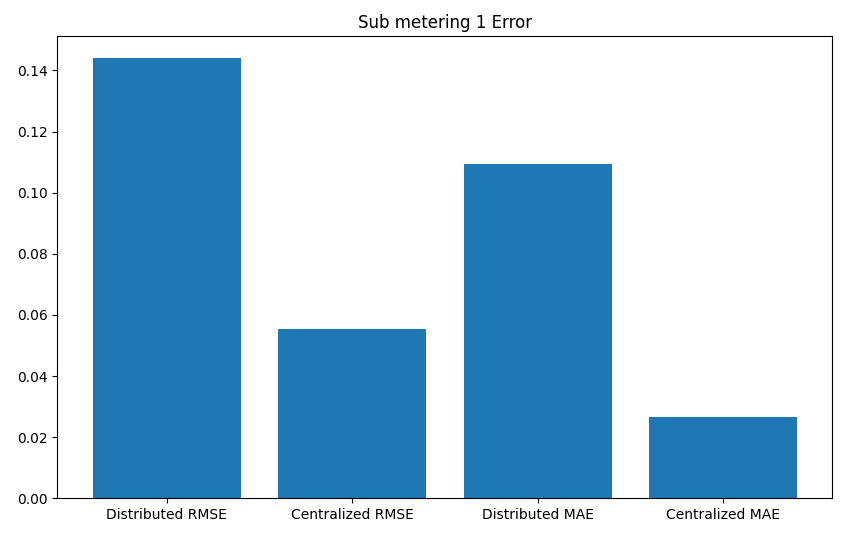
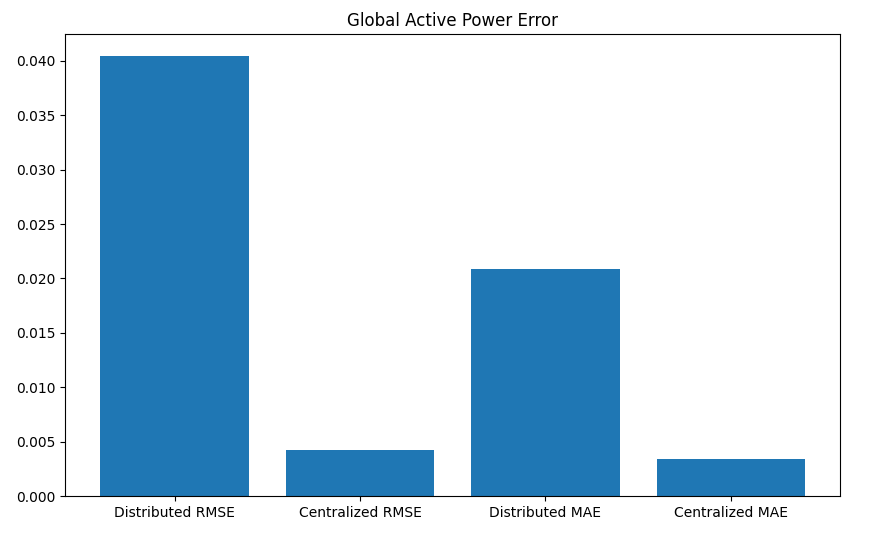


* Distributed Sub Metering 3 Model :



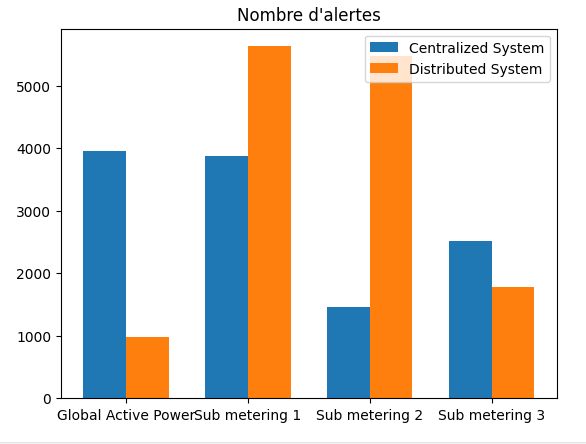
### 2.1.2 Erreurs de prévision

D’après mon expérimentation, les résultats obtenus sont contraires à ceux obtenus dans l’article car j’obtiens un nombre plus important d’erreurs de la part du système distribué et avec une proportion largement supérieure.



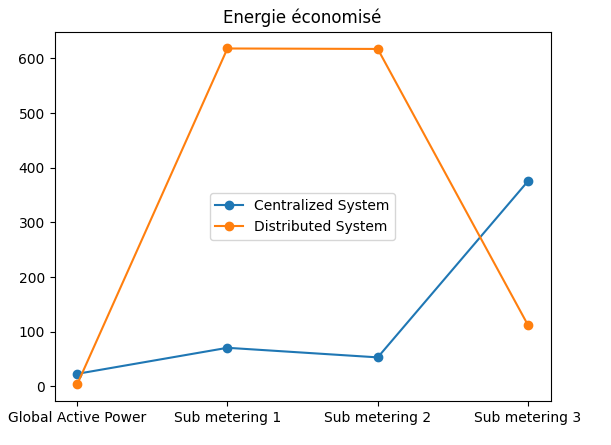
### 2.1.3 Les alertes

En ce qui concerne le nombre d'alertes, les conclusions restent les mêmes quant à savoir si le système distribué ou centralisé reçoit le plus d'alertes, bien que les proportions diffèrent.



### 2.1.4 Energie économisée

L’énergie économisée par le système centralisée est parfaitement conforme à l’expérimentation faite dans l’article, ce qui n’est pas du tout le cas pour le système distribué.



## 2.2 Conclusion

Conformément aux informations fournies dans l'article, suite à l'étape de chargement des données et à la phase de prétraitement, il est relevé que approximativement 1.25% des lignes présentent des valeurs nulles.

Le dataset a ensuite bien été séparé pour former 26250 heures de données d’entraînement et 8308 heures de données de test, à une heure près.

Ensuite, je me suis interrogé sur la meilleure façon de simuler un système distribué, notamment en me demandant s'il était nécessaire de modifier la taille des lots (batchs) pour y parvenir. J'ai conclu que chaque appareil traitait ses propres données de manière indépendante dans son propre contrôleur, sans accéder aux données des autres appareils. Par conséquent, j'ai procédé à la séparation des données correspondantes afin de simuler l'absence d'informations provenant des autres appareils. Ce n’est peut-être pas la bonne conclusion mais c’est sur cette hypothèse que j’ai poursuivi l’expérimentation.

C’est peut-être ceci qui a causé des différences de résultats entre le système distribué de mon expérimentation et celui de l’article.