

Problématique

- × **Demande imprévisible** : Consommation électrique hautement variable selon l'heure, la saison et le climat [3].
- × **Limites des méthodes classiques** : ARIMA et régression échouent face aux dépendances temporelles non linéaires [1].
- × **Instabilité du réseau** : L'intégration des renouvelables intermittentes complique la planification énergétique [2].
- × **Manque d'application industrielle** : Les modèles restent rarement transposés vers des outils concrets de gestion [4].

Consommation horaire moyenne — une variabilité considérable

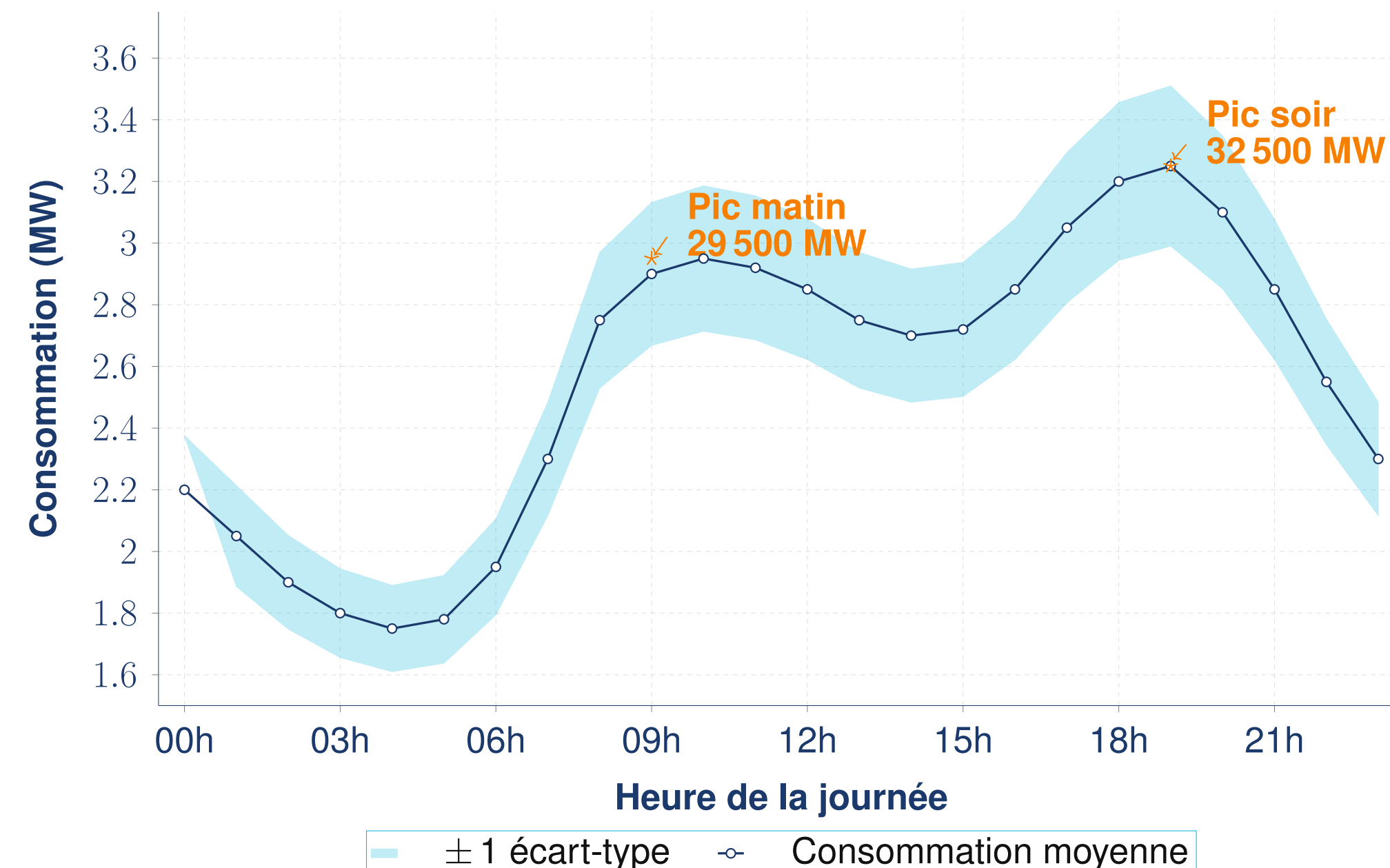


FIGURE 1. Consommation horaire moyenne, variabilité journalière (Espagne 2024).

Objectifs

Développer un système de prévision et d'optimisation de la consommation énergétique par l'IA, des données nationales au cas industriel.

- Collecter et structurer un dataset multi-sources (REE, météo, calendrier) à résolution horaire.
- Comparer des modèles ML : Random Forest, XGBoost, LSTM et CNN-LSTM.
- Évaluer les performances selon les métriques MAE, RMSE et MAPE.
- Transposer vers un cas industriel simulé pour optimiser la gestion énergétique.

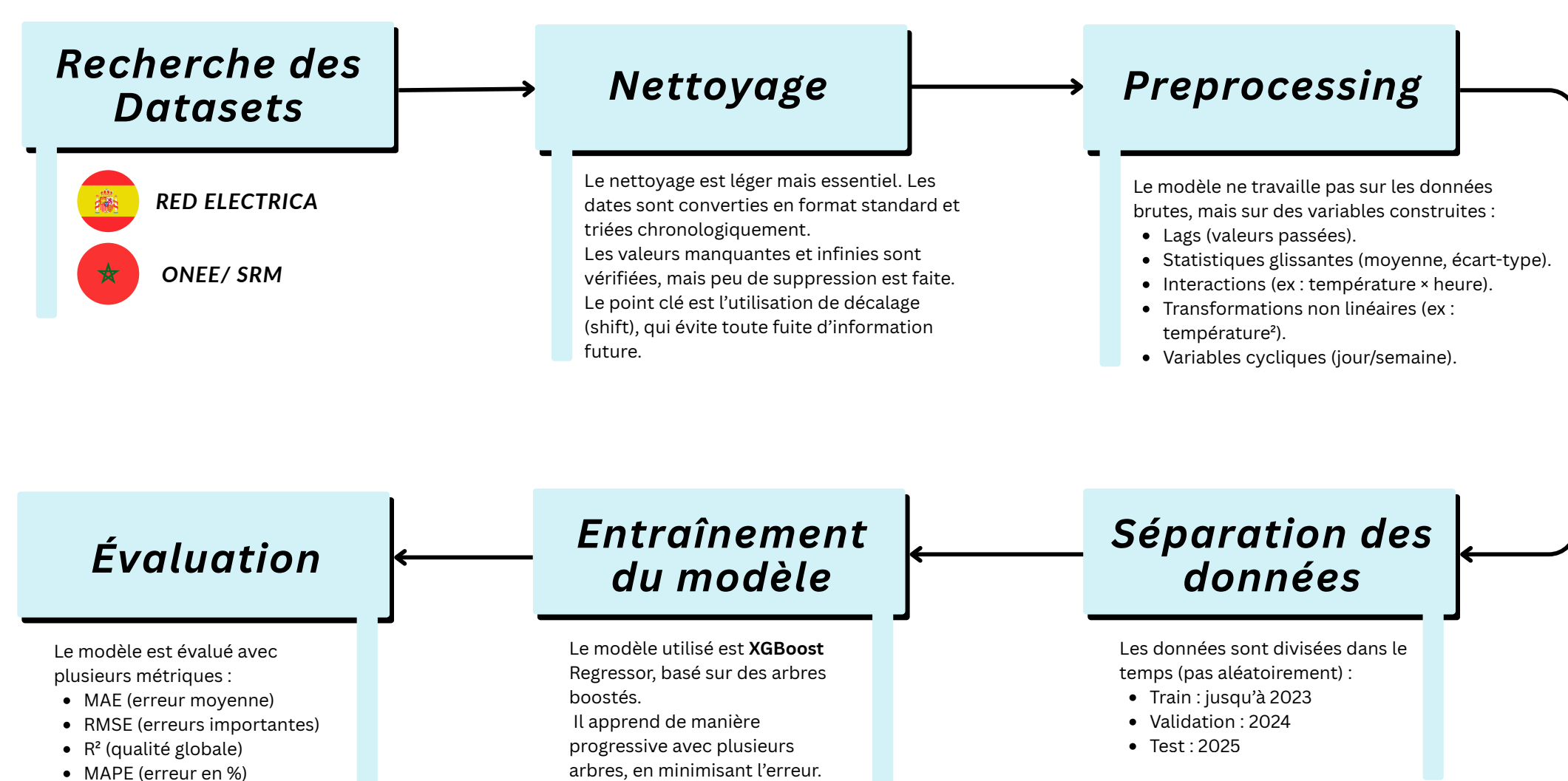


FIGURE 2. Pipeline méthodologique de la collecte à l'évaluation.

Méthodologie suivie

I- Récupération et préparation des données :

- **Source principale** : données horaires de consommation électrique nationale issues de **Red Electrica de España (REE/ESIOS)**, couvrant la période 2021–2025 à résolution horaire (~55 000 points de données).
- **Variables exogènes** : température, humidité, vitesse du vent et nébulosité issus de **OpenWeather API**, ainsi que les jours fériés du calendrier espagnol.
- **Prétraitement** : nettoyage des valeurs manquantes, tri chronologique, normalisation des séries temporelles et décalage temporel (*shift*) pour éviter toute fuite d'information future.

II- Ingénierie des caractéristiques :

Construction de variables temporelles (heure, jour de la semaine, mois, saison, jours fériés) ainsi que de variables décalées (*lags* : 1h, 24h, 48h, 168h) capturant les dépendances historiques, statistiques glissantes (moyenne, écart-type) et interactions météo × temps.

III- Classification et comparaison des modèles :

Les modèles évalués sont répartis en deux familles :

Modèles statistiques et ML :

- ARIMA : référence classique (baseline).
- **XGBoost** : gradient boosting sur arbres.
- LightGBM : boosting optimisé en mémoire.

Modèles de Deep Learning :

- LSTM, BiLSTM : réseaux récurrents.
- GRU : version allégée du LSTM.
- Transformer, CNN-LSTM : architectures hybrides.

→ **Critère de sélection** : comparaison sur le jeu de test selon MAE, RMSE, R^2 et MAPE. **XGBoost** est retenu pour ses meilleures performances en précision, robustesse et temps de calcul.

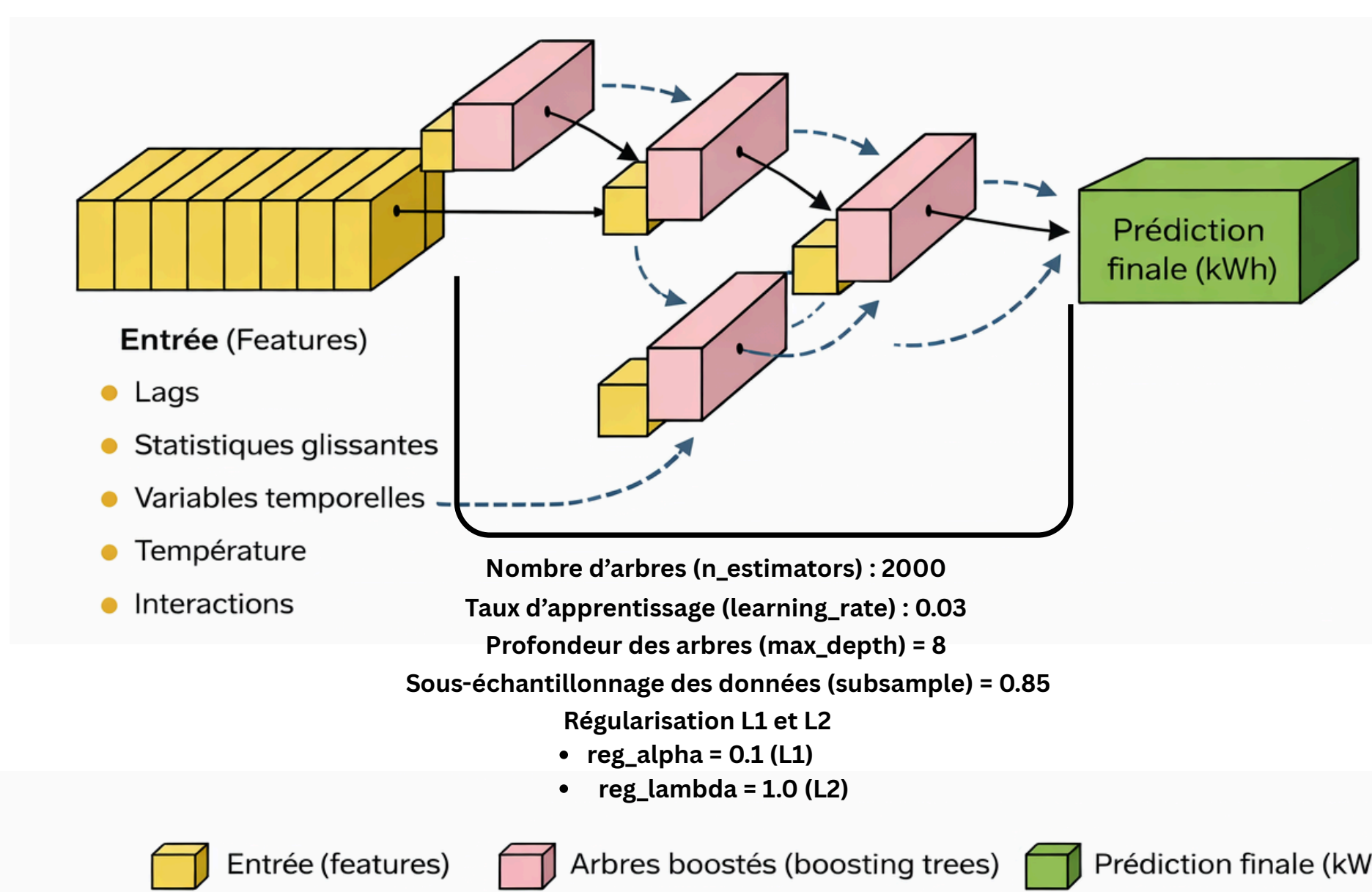


FIGURE 3. Architecture XGBoost → entrées, arbres boostés et prédiction finale.

→ Optimisation des hyperparamètres par validation croisée temporelle (*time-series split*) et évaluation selon les métriques MAE, RMSE et MAPE.

IV- Application industrielle simulée :

→ Transposition du modèle XGBoost vers un cas d'étude industriel simulé : planification de la consommation d'une usine, détection des pics de charge et proposition de stratégies énergétique.

Résultats

Les résultats obtenus sur le jeu de test montrent que les modèles de machine learning, en particulier **XGBoost**, se distinguent nettement avec un R^2 de **0.9867** et un **MAPE de seulement 1.30 %**, surpassant non seulement LightGBM mais également les modèles de deep learning tels que LSTM, BiLSTM, GRU, Transformer et CNN-LSTM en termes de précision et de stabilité.

Modèle	Type	MAE	RMSE	R^2	MAPE%
XGBoost	ML	349.55	488.37	0.9867	1.30
LightGBM	ML	368.95	511.54	0.9854	1.37
GRU	DL	681.14	911.06	0.9536	2.49
Transformer	DL	740.95	999.05	0.9442	2.75
CNN-LSTM	DL	841.42	1116.49	0.9303	3.06
BiLSTM	DL	867.13	1145.92	0.9265	3.17
LSTM	DL	916.04	1216.24	0.9173	3.36

TABLE 1. Tableau comparatif des performances des modèles sur le jeu de test.

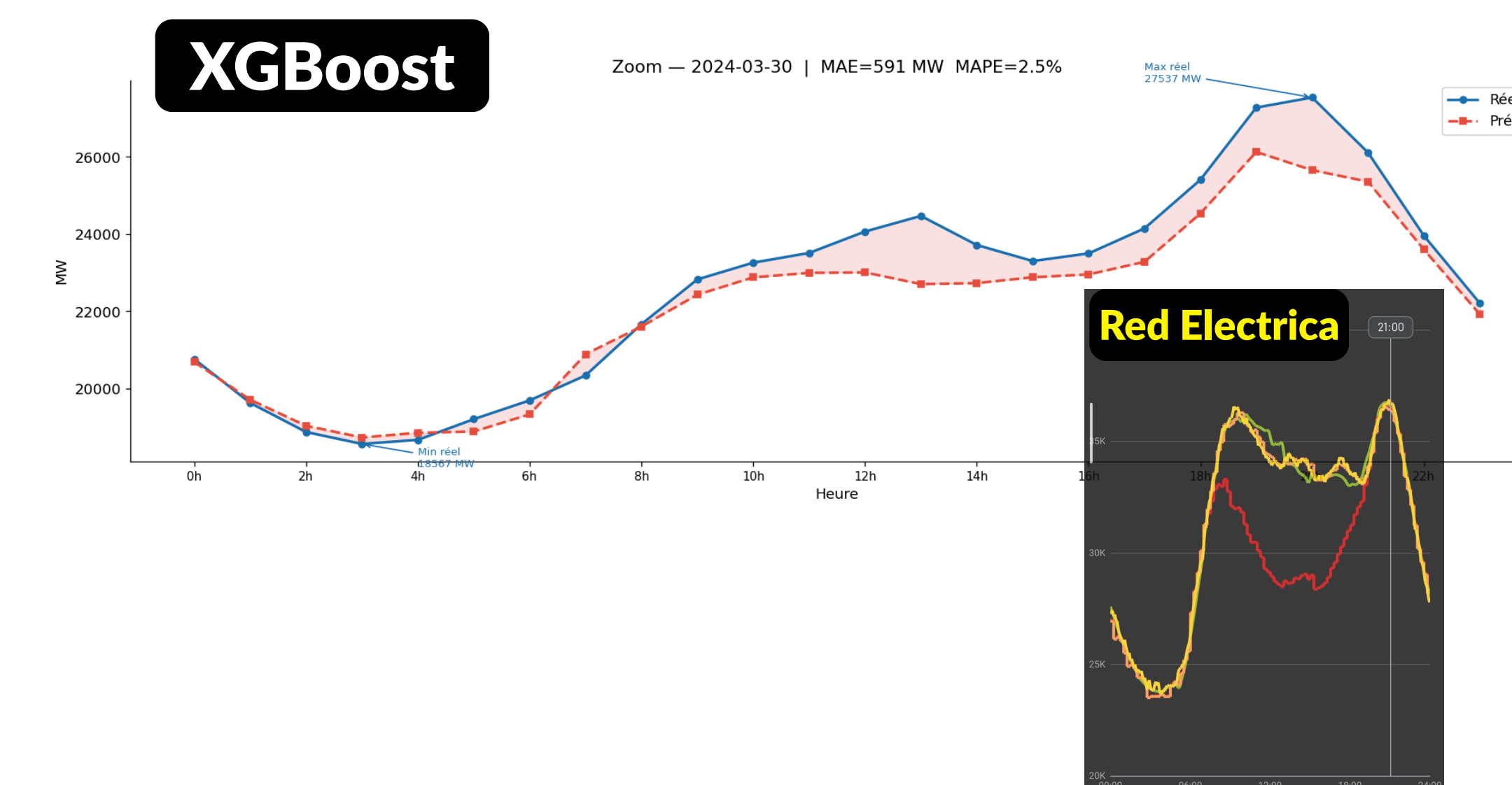


FIGURE 4. Comparaison XGBoost vs. données Red Electrica.

Les résultats montrent que, malgré la puissance des modèles deep learning, les modèles de type boosting comme XGBoost restent les plus efficaces dans un contexte de données tabulaires enrichies.

Conclusion et perspectives

XGBoost ($R^2 = 0.9867$, MAPE = 1.30 %) surpasse tous les modèles testés, confirmant l'apport des variables météorologiques et calendaires pour la prévision énergétique.

Ces résultats ouvrent la voie à une extension au réseau marocain (ONEE – UM6P), un **déploiement IoT** en temps réel et une optimisation industrielle visant une réduction de 10–15 % des coûts énergétiques.

Références

- [1] Enrique Caro and Jesús Juan. Short-term load forecasting for spanish insular electric systems. *Energies*, 13 :3645, 2020.
- [2] Chukwuebuka Joseph Ejayi et al. Comprehensive review of artificial intelligence applications in renewable energy systems : current implementations and emerging trends. *Journal of Big Data*, 12 :169, 2025.
- [3] Roman V. Klyuev et al. Methods of forecasting electric energy consumption : a literature review. *Energies*, 15 :8919, 2022.
- [4] Othmane Laayati et al. Smart energy management system : Design of a monitoring and peak load forecasting system for an experimental open-pit mine. *Applied System Innovation*, 5 :18, 2022.