



UNIVERSITÉ NATIONALE DES SCIENCES,
TECHNOLOGIES, INGÉNIERIE ET
MATHÉMATIQUES (UNSTIM)

CONCOURS SUR L'INNOVATION EN MILIEU
UNIVERSITAIRE, 5^{ème} édition (CIMU)

THEME

**PRÉDICTION DE L'ÉNERGIE ÉOLIENNE AU
BÉNIN GRÂCE A L'INTELLIGENCE ARTIFICIELLE**

Développé par :

AKAKPO Codjo Ulrich Expéra

Sous la supervision de :

Dr DANDOGBESSI Bruno

2024-2025

Table des matières

1	INTRODUCTION	3
1.1	Contexte	3
1.2	Objectifs	3
1.3	Problématique	3
2	REVUE DE LA LITTERATURE	4
2.1	État de l'Art : Approches de Prédiction de la Production Éolienne	4
2.2	Comparaison et Justification du Choix de l'IA (XGBoost)	5
2.3	Contexte Local : Pertinence au Bénin	5
3	METHODOLOGIE	5
3.1	Collecte des Données	6
3.2	Prétraitement et Analyse des données	10
3.3	Modèles entraînés	14
3.4	Entraînement et Validation	14
3.5	Évaluation des Performances	14
4	RESULTATS	15
4.1	Performances des Modèles	15
4.2	Démonstration - Plateforme ÉoBénin	18
4.3	Analyse des Erreurs	18
5	DISCUSSION	18
5.1	Interprétation	18
5.2	Impact du projet	19
5.3	Innovation	20
5.4	Limites	20
5.5	Perspectives d'Améliorations	20
6	CONCLUSION	21

1 INTRODUCTION

1.1 Contexte

Le Bénin, à l'instar de nombreux pays en développement, fait face à des défis énergétiques majeurs. Sa dépendance prédominante vis-à-vis des énergies fossiles, souvent importées, le rend vulnérable aux fluctuations des marchés mondiaux et contribue significativement aux émissions de dioxyde de carbone (CO_2). Cette situation se traduit par des problèmes récurrents de pénurie d'électricité et un impact environnemental croissant, exacerbant les enjeux liés au changement climatique. Dans ce contexte, la transition vers des sources d'énergie renouvelables représente non seulement une nécessité écologique, mais aussi une opportunité stratégique pour assurer la sécurité énergétique, stimuler le développement économique local et améliorer la qualité de vie des populations. L'énergie éolienne, en particulier, offre un potentiel considérable au Bénin, avec des zones présentant des régimes de vent favorables, mais qui restent largement sous-exploitées. L'intégration de cette ressource propre et abondante peut transformer le paysage énergétique national, réduisant la pression sur les infrastructures existantes et ouvrant la voie à un avenir plus durable.

1.2 Objectifs

Ce projet ambitieux vise à développer une solution innovante basée sur l'intelligence artificielle pour optimiser l'exploitation de l'énergie éolienne au Bénin. Plus précisément, ses objectifs sont triples :

- ⊗ **Prédire la vitesse du vent et la production horaire d'énergie éolienne :** En développant un modèle d'**intelligence artificielle** capable d'anticiper avec précision la vitesse du vent à 50 mètres d'altitude, et d'en déduire la puissance générée par une éolienne spécifique (Enercon E48/800) sur une base horaire, le projet cherche à fournir des informations cruciales pour la planification énergétique.
- ⊗ **Réduire la dépendance aux combustibles fossiles et les émissions de CO_2 :** En permettant une meilleure intégration de l'énergie éolienne dans le mix énergétique national, ces prévisions précises faciliteront la substitution des énergies fossiles, contribuant directement à la diminution des émissions de gaz à effet de serre et à la lutte contre le changement climatique.
- ⊗ **Optimiser le stockage d'énergie et faciliter l'intégration au réseau électrique :** Une connaissance anticipée de la production éolienne permet d'optimiser la gestion des systèmes de stockage (comme les batteries), d'améliorer la stabilité du réseau électrique, de réduire les coûts opérationnels liés à la variabilité des énergies renouvelables et de maximiser l'efficacité de la distribution énergétique.

1.3 Problématique

La production d'énergie éolienne est intrinsèquement liée à la variabilité de la vitesse et de la direction du vent, ce qui en fait une source d'énergie intermittente et difficilement prévisible sans outils adéquats. Cette fluctuation constitue un défi majeur pour les gestionnaires de réseau, les opérateurs de centrales électriques et les planificateurs énergétiques. Sans des prévisions précises, l'intégration de l'énergie éolienne à grande échelle peut entraîner des déséquilibres sur le réseau, nécessitant le maintien de capacités de production de secours souvent basées sur les énergies fossiles, ou imposant des contraintes coûteuses en termes de stockage et de régulation. La problématique centrale de ce projet est donc de répondre à la question suivante : *comment*

développer une solution robuste d'intelligence artificielle capable de fournir des prévisions horaires fiables de la vitesse du vent et de la puissance éolienne, afin de surmonter les défis de l'intermittence et de faciliter une transition énergétique réussie au Bénin ?

2 REVUE DE LA LITTERATURE

2.1 État de l'Art : Approches de Prédiction de la Production Éolienne

La prévision de la vitesse du vent et, par conséquent, de la production d'énergie éolienne est un domaine de recherche actif et crucial pour l'intégration efficace de cette énergie renouvelable. Au fil des années, diverses approches ont été développées, chacune avec ses propres forces et faiblesses :

- **Modèles Physiques** : Ces modèles sont basés sur des principes fondamentaux de la mécanique des fluides et de la thermodynamique. Ils utilisent des équations complexes pour simuler l'atmosphère et les interactions du vent avec l'environnement. Les exemples typiques incluent les modèles de prévision numérique du temps (NWP - Numerical Weather Prediction), qui nécessitent une puissance de calcul significative et une connaissance approfondie des phénomènes météorologiques. Bien qu'ils puissent fournir des prévisions précises à différentes échelles spatiales et temporelles, leur complexité et leur coût de mise en œuvre les rendent parfois inadaptés pour des applications spécifiques ou des déploiements rapides.
- **Modèles Statistiques** : Ces approches exploitent les relations historiques entre les données météorologiques et la production éolienne. Elles sont généralement plus simples à implémenter et moins gourmandes en ressources que les modèles physiques. Des techniques telles que les modèles Autorégressifs à Moyenne Mobile Intégrée (ARIMA), les moyennes mobiles (MA), les modèles GARCH, ou les régressions linéaires multiples sont couramment utilisées. Ces modèles fonctionnent bien lorsque les patterns sont stationnaires et qu'il existe des corrélations linéaires claires, mais ils peuvent avoir des difficultés à capturer les non-linéarités et la complexité inhérente aux dynamiques atmosphériques.
- **Modèles d'Intelligence Artificielle (IA)** : Avec l'avènement des grandes quantités de données (Big Data) et l'augmentation de la puissance de calcul, les modèles d'IA, y compris le Machine Learning (ML) et les Réseaux de Neurones Artificiels (RNA), ont gagné en popularité pour la prévision éolienne. Ces modèles sont capables d'apprendre des relations complexes et non-linéaires directement à partir des données, sans nécessiter une modélisation explicite des processus physiques sous-jacents. Parmi les techniques notables, on trouve :
 - **Réseaux de Neurones Artificiels (RNA)** : Particulièrement les réseaux récurrents (RNN) ou à mémoire à long terme (LSTM) qui sont adaptés aux données séquentielles comme les séries temporelles météorologiques.
 - **Machines à Vecteurs de Support (SVM)** : Utilisées pour des tâches de régression.
 - **Ensembles de Méthodes (Ensemble Methods)** : Comme les forêts aléatoires (Random Forests) ou les méthodes de boosting (Gradient Boosting Machines, AdaBoost, XGBoost, LightGBM). Ces approches combinent plusieurs modèles plus faibles pour former un prédicteur plus robuste et plus précis.

2.2 Comparaison et Justification du Choix de l'IA (XGBoost)

Dans le cadre de ce projet, l'approche basée sur l'Intelligence Artificielle a été privilégiée pour plusieurs raisons fondamentales, et plus spécifiquement l'algorithme **XGBoost (eXtreme Gradient Boosting)**.

Les modèles physiques, bien que rigoureux, sont trop complexes et nécessitent des ressources qui ne sont pas toujours disponibles pour un projet de cette envergure. Les modèles statistiques, bien qu'utiles pour des prévisions à court terme et des patterns simples, peinent à capturer la nature intrinsèquement non-linéaire et les interactions complexes entre les multiples variables météorologiques qui influencent la vitesse du vent.

L'IA, en revanche, excelle dans la détection de patterns complexes et l'apprentissage de relations non-linéaires à partir de grandes quantités de données, ce qui est essentiel pour la prévision météorologique et énergétique. Parmi les algorithmes d'IA, XGBoost a été choisi pour les avantages suivants :

- **Haute Performance et Précision** : XGBoost est reconnu pour son excellente précision et sa capacité à gagner de nombreuses compétitions de *Machine Learning*. Il gère efficacement les données hétérogènes et est particulièrement performant sur des jeux de données tabulaires.
- **Robustesse aux Valeurs Aberrantes et Manquantes** : Il est moins sensible aux valeurs extrêmes et peut gérer implicitement les valeurs manquantes (bien que le prétraitement reste crucial).
- **Vitesse et Scalabilité** : Il est optimisé pour la vitesse et peut s'adapter à des jeux de données de grande taille, ce qui est important pour les données horaires sur une longue période.
- **Capacité à Gérer les Relations Complexes** : XGBoost, en tant qu'algorithme de *gradient boosting* basé sur des arbres de décision, est particulièrement doué pour modéliser des interactions complexes et non-linéaires entre les différentes caractéristiques (température, humidité, pression, précipitations, et variables temporelles). Il construit séquentiellement des arbres, chacun corrigeant les erreurs du précédent, ce qui le rend très puissant pour des tâches de régression comme la prédiction de la vitesse du vent.

2.3 Contexte Local : Pertinence au Bénin

L'originalité et l'impact de ce projet sont particulièrement accentués par son application au contexte béninois. À ce jour, il existe une absence ou une rareté significative de solutions de prédiction de l'énergie éolienne spécifiquement adaptées aux conditions climatiques et aux besoins énergétiques du Bénin. Les études et déploiements existants se concentrent majoritairement sur des régions plus développées ou avec des infrastructures de mesure météorologique plus denses. En utilisant des données issues de sources fiables (comme NASA POWER) mais appliquées à une localisation précise (**Cotonou, Bénin, latitude=6,35 - longitude=2,38**) et en intégrant une éolienne commercialement pertinente (Enercon E48/800), ce projet comble une lacune importante. Il démontre la faisabilité et la valeur ajoutée d'une approche basée sur l'IA pour optimiser l'utilisation de l'énergie éolienne dans un pays en développement, contribuant ainsi directement à ses objectifs de développement durable et à sa souveraineté énergétique.

3 METHODOLOGIE

Cette section détaille l'approche adoptée pour le développement de l'intelligence artificielle de prédiction, depuis la collecte et le prétraitement, l'analyse des données jusqu'à l'entraînement

et l'évaluation des modèles.

3.1 Collecte des Données

La fiabilité d'un modèle prédictif repose intrinsèquement sur la qualité et la pertinence des données utilisées pour son entraînement. Pour ce projet, deux sources de données principales ont été exploitées :

- ♣ **Données Météorologiques NASA POWER** : La source principale des données météorologiques est le projet **NASA Prediction Of Worldwide Energy Resources (POWER)** [<https://power.larc.nasa.gov>]. Ce service fournit un accès facile à des ensembles de données climatiques et météorologiques mondiaux, dérivés d'observations satellitaires et de modèles de réanalyse. Le fichier *nasa_power_data.csv*, généré par le script Python de scraping (*web_scraping_weather_data.py*) fourni, contient des données horaires pour la localisation de Cotonou, au Bénin (coordonnées géographiques 6,35 - 2,38).

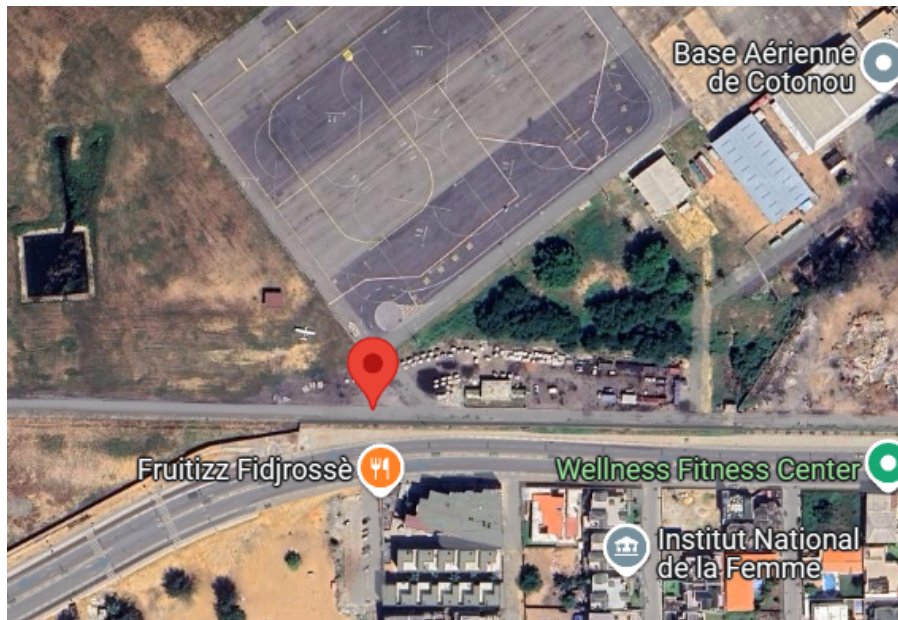


FIGURE 1 – Site de référence pour la collecte des données météorologiques (Latitude : 6.35, Longitude : 2.38)

Les variables clés extraites incluent :

- **timestamp** : Date et heure
- **wind_speed_10m** : Vitesse du vent à 10 mètres (m/s)
- **wind_speed_50m** : Vitesse du vent à 50 mètres (m/s) — Cible de la prédiction
- **temperature_2m** : Température (°C)
- **humidity_2m** : Humidité relative (%)
- **pressure** : Pression (kPa)
- **precipitation** : Précipitations (mm)

timestamp	wind_speed_10m	wind_speed_50m	temperature_2m	humidity_2m	pressure	precipitation
2019-01-01 00:00:00	1,65	2,09	25,9	87,17	100,84	0
2019-01-01 01:00:00	1,9	2,97	25,27	87,89	100,79	0
2019-01-01 02:00:00	2,29	3,78	24,95	85,81	100,78	0
2019-01-01 03:00:00	2,44	3,78	24,39	85,28	100,82	0
2019-01-01 04:00:00	2,55	3,79	23,73	85,81	100,89	0
2019-01-01 05:00:00	2,74	3,99	23,08	86,08	100,95	0
2019-01-01 06:00:00	3,17	4,09	22,89	83,66	101,03	0
2019-01-01 07:00:00	3,34	3,79	23,57	76,84	101,1	0
2019-01-01 08:00:00	3,55	4,12	25,16	65,38	101,13	0
2019-01-01 09:00:00	3,32	3,82	26,96	57,21	101,07	0

FIGURE 2 – Aperçu des données

262 944 enregistrement, environ **30 ans** de données horaires. Ces données sont essentielles pour comprendre les conditions atmosphériques qui influencent la vitesse du vent.

- ♣ **Courbe de Puissance de l'Éolienne Enercon E48/800** : Pour estimer la puissance générée par l'éolienne, la courbe de puissance du modèle Enercon E48/800 a été utilisée.

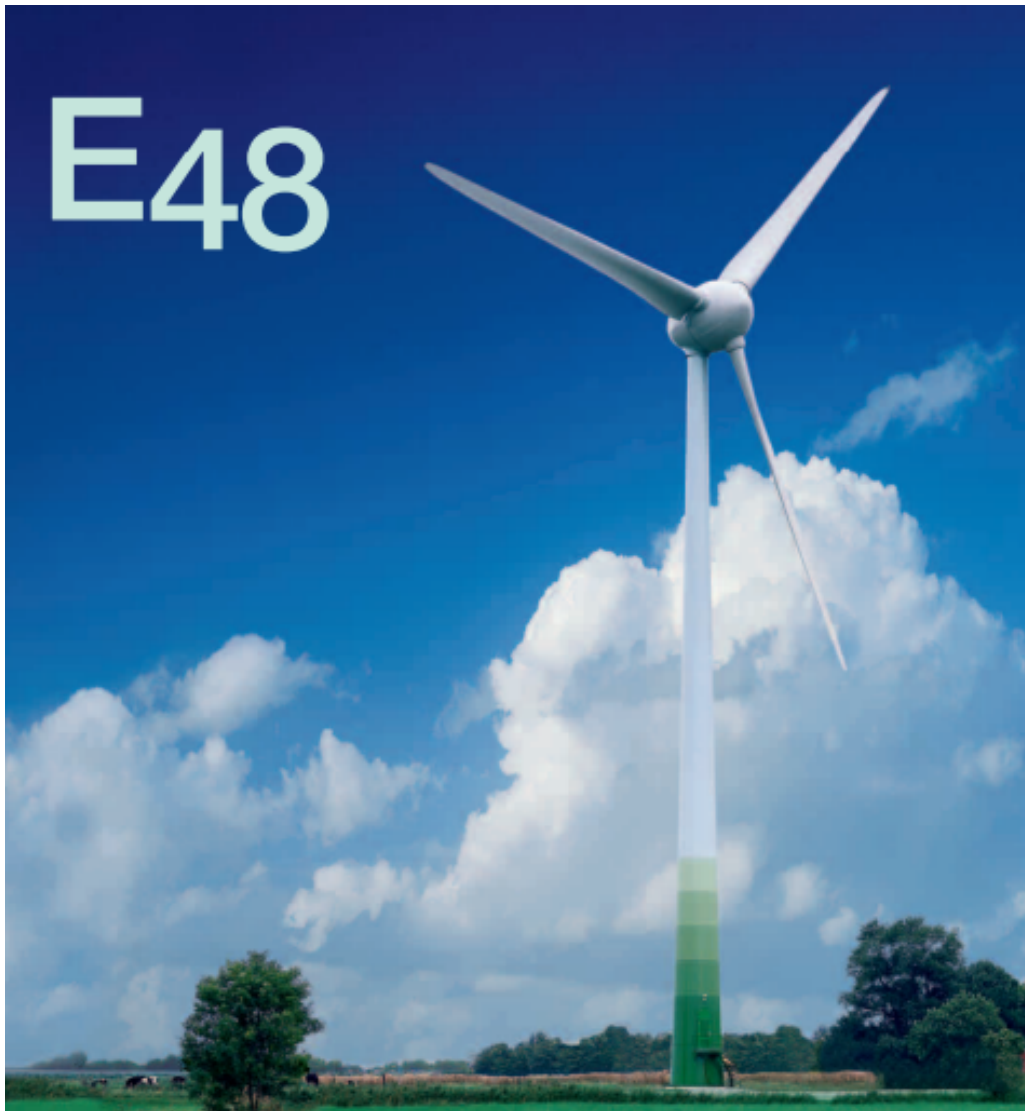


FIGURE 3 – Enercon E48/800

DONNÉES TECHNIQUES

Puissance nominale :	800 kW
Diamètre du rotor :	48 m
Hauteur du moyeu :	50 m à 76 m
Classe de vent (IEC) :	IEC/NVN II
Particularités techniques :	sans boîte de vitesse, vitesse de rotation variable, réglage indépendant de chaque pale
Rotor	
Type :	face au vent à ajustage actif des pales
Sens de rotation :	Sens des aiguilles d'une montre
Nombre de pales :	3
Surface balayée :	1 810 m ²
Matériau utilisé pour les pales :	GFK (résine époxy) ; protection parafoudre intégrée
Vitesse de rotation :	variable, 16 à 30 tours/min.
Système de réglage des pales :	système de réglage indépendant de chaque pale ENERCON, 3 unités indépendantes avec système d'alimentation électrique de secours
Système d'entraînement avec générateur	
Moyeu :	fixe
Paliers principaux :	palier à rouleaux coniques en rangée simple
Générateur :	générateur annulaire ENERCON à accouplement direct
Alimentation du réseau :	onduleur ENERCON
Systèmes de freinage :	- 3 unités indépendantes avec alimentation de secours - frein d'arrêt du rotor - dispositif de blocage du rotor
Contrôle d'orientation (« yaw ») :	actif par un mécanisme d'engrenages, amortissement proportionnel à la charge
Vitesse de vent de coupure :	28 à 34 m/s (avec mode tempête ENERCON)
Système de surveillance à distance:	ENERCON SCADA

FIGURE 4 – Données technique E48/800

Les données techniques sont conformes avec les paramètres propriétés météorologiques du lieu choisir. D'où le choix de ce modèle d'Enercon (E48/800)

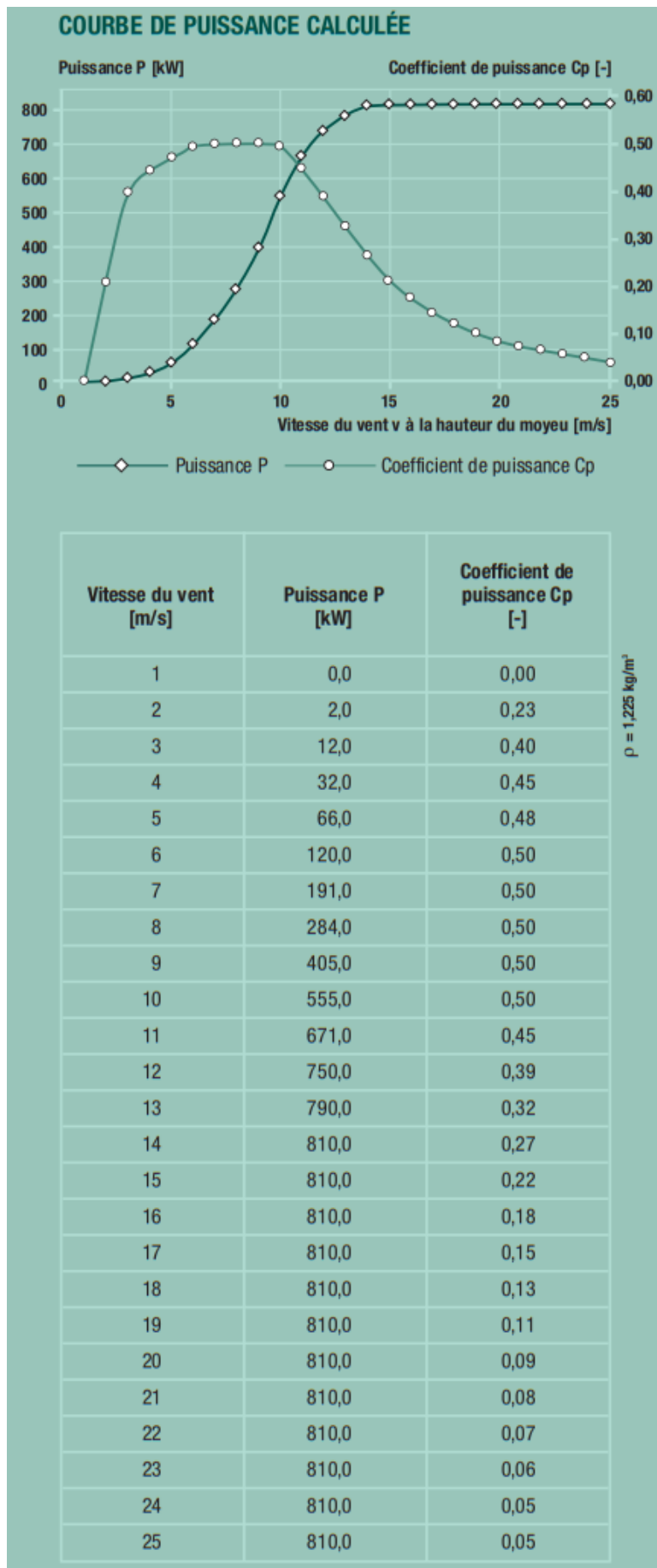


FIGURE 5 – Courbe de puissance E48/800

Cette courbe, contenue dans le fichier *courbe_puissance.csv*, fournit la puissance électrique (en kW) produite par l'éolienne pour différentes vitesses de vent. Afin de permettre une estimation continue de la puissance à partir de n'importe quelle vitesse de vent prédite, une fonction d'interpolation a été appliquée (utilisant *scipy.interpolate.interp1d* avec *bounds_error=False* et *fill_value=0*). Cela garantit que même pour des vitesses de vent en dehors des points de données exacts de la courbe, une estimation de puissance peut être obtenue, avec une puissance nulle pour les vitesses en dehors de la plage de fonctionnement de l'éolienne.

3.2 Prétraitement et Analyse des données

Le prétraitement est une étape cruciale pour préparer les données brutes à être utilisées par les modèles d'apprentissage automatique, garantissant leur qualité et leur format adéquat :

- ◇ **Conversion des Timestamps** : Les timestamps (horodatages) des données NASA POWER ont été convertis en objets datetime de Pandas (*pd.to_datetime*). Cette conversion est fondamentale car elle permet d'extraire des informations temporelles pertinentes et de faciliter les manipulations basées sur le temps.
- ◇ **Encodage Cyclique des Variables Temporelles** : Pour les modèles qui incluent des caractéristiques temporelles (mois, heure), un encodage cyclique a été appliqué. Cette technique transforme les variables périodiques (comme le mois dans l'année ou l'heure dans la journée) en leurs composantes sinus et cosinus. Par exemple, pour le mois, $\text{month_sin} = \sin(2 * \pi * \text{month} / 12)$ et $\text{month_cos} = \cos(2 * \pi * \text{month} / 12)$. Cette approche évite d'imposer une relation d'ordre arbitraire (par exemple, que le mois de décembre soit "plus éloigné" de janvier que de novembre) et permet au modèle de percevoir la continuité cyclique (décembre est proche de janvier). Cette méthode est explicitement utilisée dans le fichier *xgboost_with_date.ipynb*.
- ◇ **Gestion des Valeurs Manquantes** : Bien que non explicitement détaillées dans les données fournies, la gestion des valeurs manquantes est une étape standard. Cela peut inclure l'imputation (remplacement par la moyenne, la médiane, la valeur précédente/suivante) ou la suppression des lignes/colonnes affectées, selon l'étendue et la nature des données manquantes. Le script de scraping semble déjà gérer certaines valeurs absentes en retournant *None*, ce qui sera géré par Pandas.
- ◇ **Analyse statistique des données**
 - L'analyse statistique nous montre que les données ont été bien enregistrées chaque heure sur une période de cinq (05) ans (2019-01-01 à 2023-12-01). Soit 262944 enregistrements.
 - Sur cette la plus petite valeur et la plus grande valeur de la vitesse à 50m de haut (depuis le sol) sont respectivement : 0.030000 m/s et 10.740000 m/s.
 - les moyennes respectives de *wind_speed_10m*, *wind_speed_50m*, *temperature_2m*, *humidity_2m*, *pressure*, *precipitation* sont 3.047987 4.149605 26.607173 85.151812 100.761707 4.191936. En plus la *wind_speed_10m*, *wind_speed_50m*, *temperature_2m*, *humidity_2m* et *pressure* ont leur valeurs concentrées autour de la moyenne. Ce qui n'est pas le cas pour la précipitation.
 - **Distribution des paramètres météorologiques**

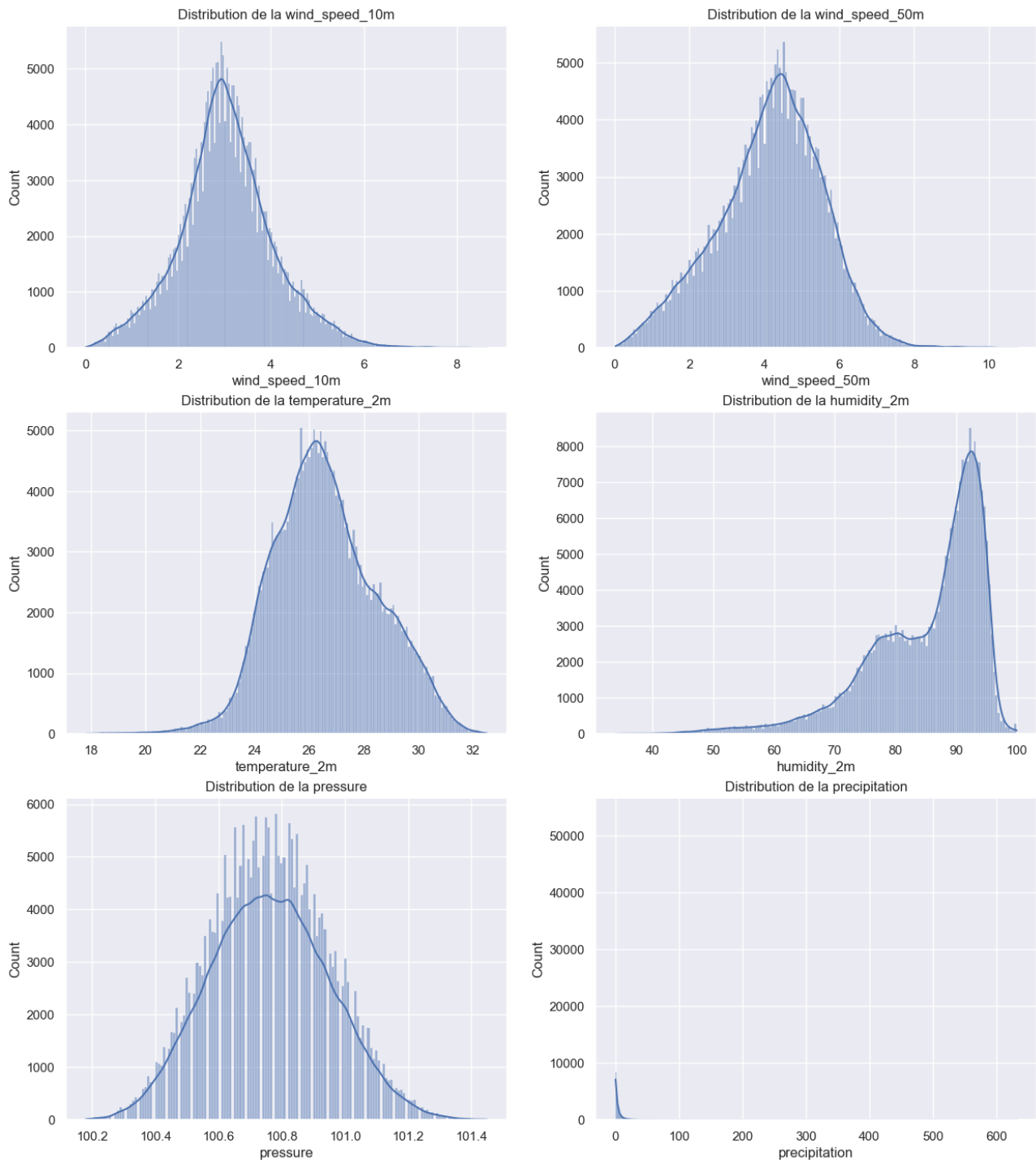


FIGURE 6 – Distribution des paramètres météorologiques

- Les vents à Cotonou sont majoritairement faibles, ce qui pourrait limiter la production d'énergie pour une turbine comme l'Enercon E48/800, qui nécessite des vitesses minimales de 2 m/s et atteint sa puissance maximale entre 14 et 25 m/s. Une faible proportion des vents dépasse 10 m/s.
- Les températures sont stables et chaudes, typiques d'un climat tropical. Cela n'affecte pas directement la turbine, mais des températures élevées prolongées pourraient influencer l'efficacité ou la maintenance.
- L'humidité est élevée, ce qui est courant dans les régions côtières comme Cotonou. Cela peut entraîner une corrosion accrue des composants de la turbine, nécessitant

une protection spéciale.

- La pression est stable, sans extrêmes, ce qui indique des conditions météorologiques relativement prévisibles sans perturbations majeures (comme des tempêtes fréquentes).
- Les précipitations sont rares ou faibles, ce qui réduit les risques d'interruptions liées à la pluie, mais cela reflète aussi une saison sèche significative.

Répartition de la vitesse à 50m par rapport au autres paramètres - Existe-t-il de relations linéaires ?

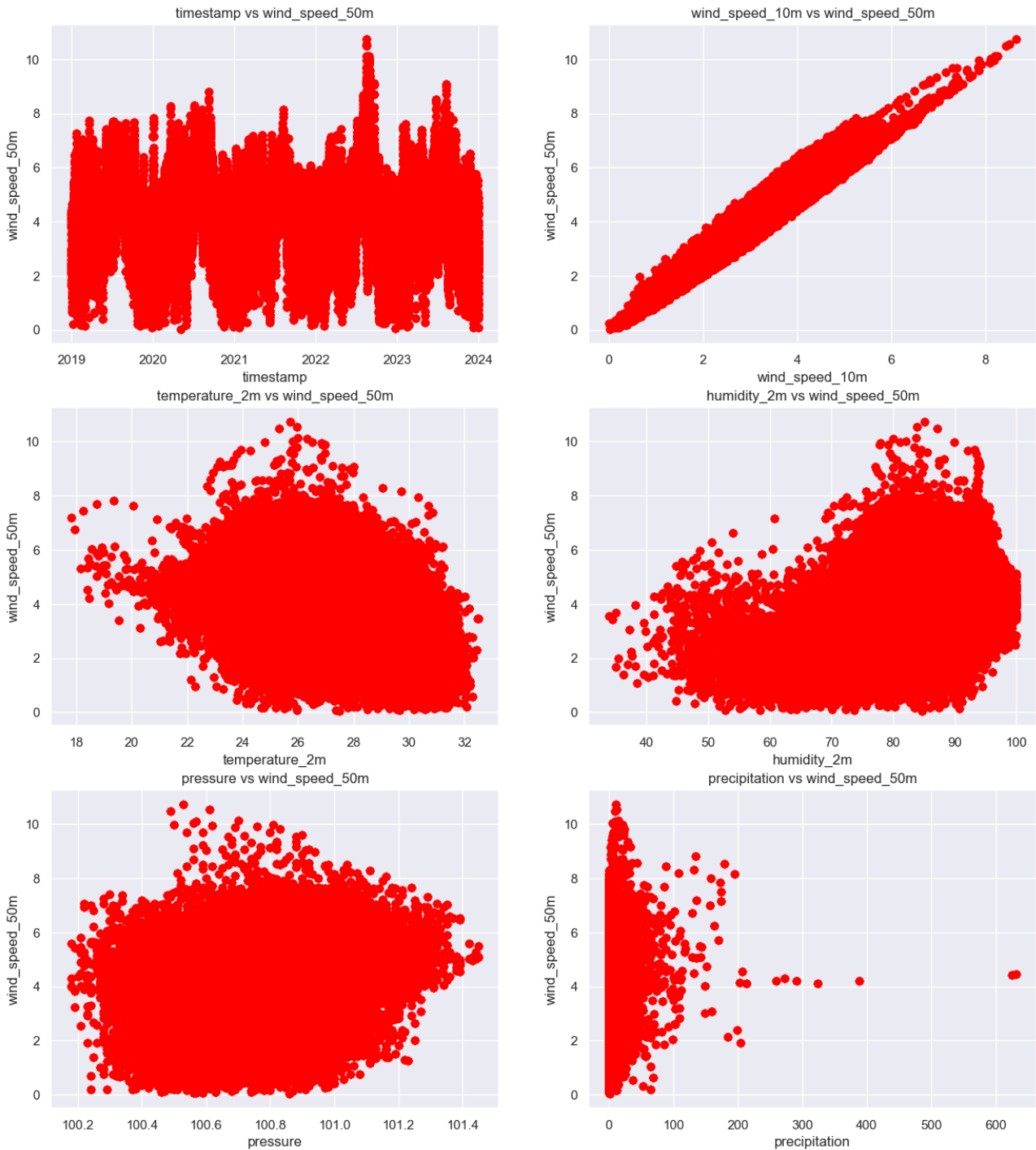


FIGURE 7 – Répartition de la vitesse à 50m par rapport au autres paramètres

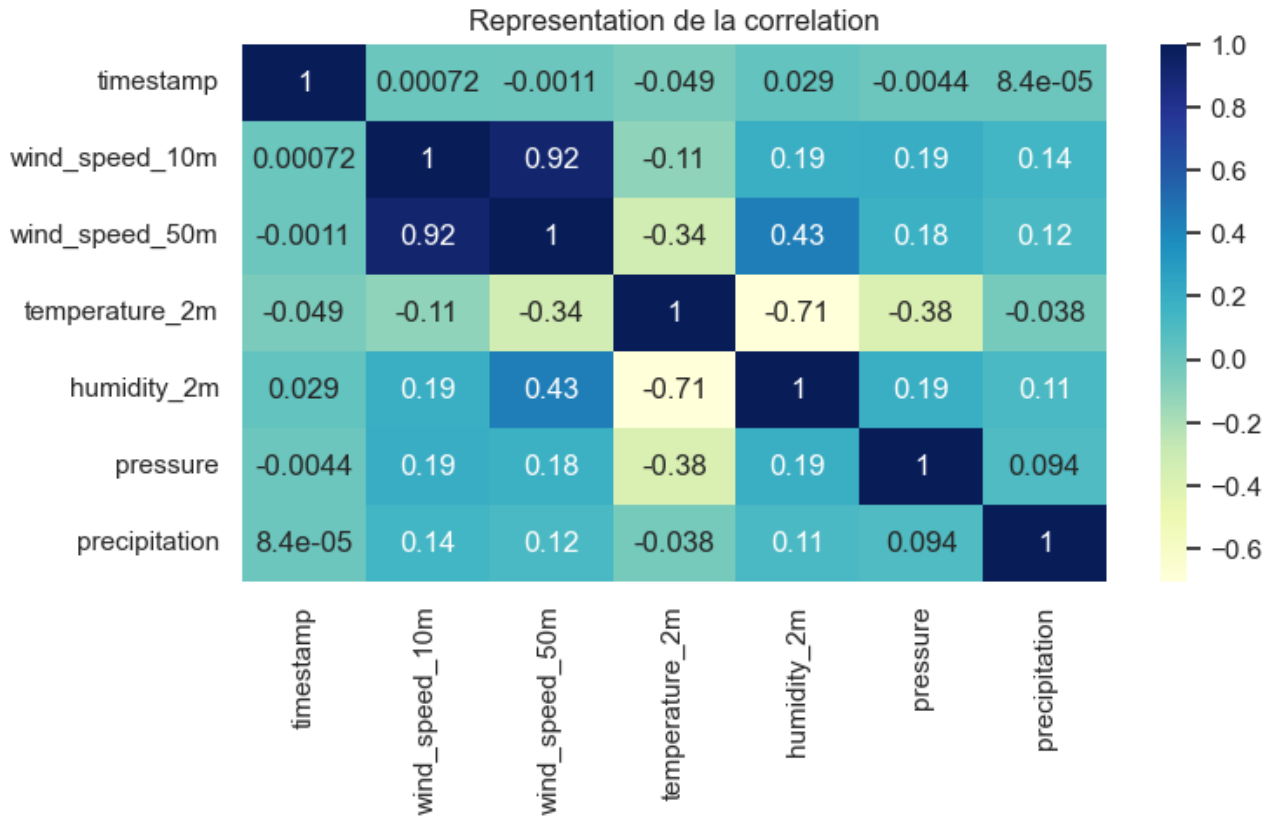


FIGURE 8 – Matrice de corrélation

- les données à 50 m (proche de la hauteur de la turbine) sont représentatives, bien que légèrement plus élevées, ce qui est favorable pour la production d'énergie.
- Quand la température augmente, les vitesses de vent tendent à diminuer légèrement, et l'humidité diminue plus nettement. Cela reflète un climat tropical où les vents sont plus faibles par temps chaud et sec.
- Une humidité élevée est associée à des vents légèrement plus forts, mais elle diminue avec la hausse des températures. Cela peut indiquer des conditions humides lors des vents plus soutenus, augmentant les risques de corrosion.
- La pression a peu d'influence sur les vents, mais elle est légèrement plus basse par temps chaud, ce qui est typique.
- Pas de tendance temporelle marquée sur la période étudiée (2019-2023), sauf une légère variation saisonnière des précipitations.

Les analyses des données météorologiques de Cotonou, basées sur les vitesses du vent à 50 m, la température, l'humidité, la pression et les précipitations, montrent que les conditions locales sont partiellement adaptées à l'Enercon E48/800. Bien que la turbine puisse fonctionner dans la plage opérationnelle (2-25 m/s) et que des vents optimaux (14-25 m/s) soient observés occasionnellement, la majorité des vitesses restent faibles (moins de 6 m/s), entraînant un facteur de capacité modeste. Les conditions environnementales, avec des températures chaudes et une humidité élevée, nécessitent une attention particulière pour la maintenance, mais les précipitations rares et la pression stable ne posent pas de contraintes majeures. L'analyse faite avec ce modèle confirme qu'il est utilisable, mais les prédictions basées sur ces données suggèrent que son efficacité reste limitée par les vents insuffisants et sera abordé dans la section des limites et des perspectives.

3.3 Modèles entraînés

Deux configurations distinctes du modèle XGBoost ont été entraînées pour évaluer l'impact des caractéristiques temporelles sur la performance de la prédiction de la vitesse du vent à 50 mètres (*wind_speed_50m*) :

- † **XGBoost sans Caractéristiques Temporelles** (*xgboost_no_date.ipynb*) : Ce modèle utilise uniquement les variables météorologiques directes comme prédicteurs : *temperature_2m*, *humidity_2m*, *pressure*, et *precipitation*. Il vise à évaluer la capacité de l'algorithme à capturer les relations entre les conditions atmosphériques brutes et la vitesse du vent.
- † **XGBoost avec Caractéristiques Temporelles** (*xgboost_with_date.ipynb*) : Cette version enrichit le jeu de données avec des caractéristiques dérivées des timestamps : l'année, le mois (encodé cycliquement), et l'heure (encodée cycliquement). L'objectif est de permettre au modèle de capter les motifs saisonniers et journaliers de la vitesse du vent, qui sont souvent très prononcés.

Les modèles ont été configurés avec des paramètres de base : 1 000 000 estimateurs (*n_estimators*), et un taux d'apprentissage (*learning_rate*) de 1. Ce choix de *n_estimators* est très élevé, suggérant une volonté d'assurer une convergence complète du modèle, tandis qu'un *learning_rate* de 1 est inhabituellement élevé pour XGBoost et pourrait indiquer un réglage à optimiser pour éviter le surapprentissage ou des oscillations dans la convergence. Cela sera abordé dans la section "Limites".

3.4 Entraînement et Validation

La robustesse et la généralisabilité des modèles ont été évaluées par une division systématique des données et des métriques de performance.

L'ensemble des données a été divisé en deux sous-ensembles : 90 % pour l'entraînement (*x_train*, *y_train*) et 10 % pour le test (*x_test*, *y_test*). La division a été effectuée de manière aléatoire (*random_state=42*) pour assurer la reproductibilité des résultats. Cette répartition est typique pour évaluer la capacité du modèle à généraliser sur des données non vues.

3.5 Évaluation des Performances

Les performances des modèles ont été quantifiées à l'aide de métriques de régression standard, permettant une comparaison objective :

† Métriques de Performance :

- **RMSE (Root Mean Squared Error - Erreur Quadratique Moyenne)** : Mesure la magnitude moyenne des erreurs de prédiction. Elle est sensible aux grandes erreurs, car elle les pénalise davantage en les mettant au carré. Une RMSE plus faible indique une meilleure performance.
- **MAE (Mean Absolute Error - Erreur Absolue Moyenne)** : Représente la moyenne des valeurs absolues des erreurs de prédiction. Elle donne une idée directe de l'erreur moyenne en unités de la variable cible. Une MAE plus faible est préférable.
- **R² (Coefficient de Détermination)** : Indique la proportion de la variance de la variable dépendante qui est prévisible à partir des variables indépendantes. Une valeur de R² proche de 1 (ou 100 %) signifie que le modèle explique une très grande partie de la variabilité des données, indiquant un très bon ajustement.

† **Comparaison des Modèles** : Les métriques ont été utilisées pour comparer les performances du modèle XGBoost sans date et du modèle XGBoost avec date. Cette comparaison est cruciale pour démontrer l'impact des caractéristiques temporelles sur la précision des prévisions, comme les résultats le suggèrent déjà.

4 RESULTATS

Cette section présente les performances des modèles de prédiction développés, en se basant sur les métriques d'évaluation et les visualisations générées. L'objectif est de démontrer l'efficacité de l'approche et de souligner l'impact des différentes caractéristiques sur la précision des prévisions.

4.1 Performances des Modèles

Les modèles XGBoost entraînés ont montré des performances remarquables, en particulier lorsque les caractéristiques temporelles sont incluses. Les métriques de régression calculées sur l'ensemble de test (*test_size=0.1*) sont les suivantes :

* **XGBoost sans Caractéristiques Temporelles** (*xgboost_no_date.ipynb*) :

- ▷ **R² (Coefficient de Détermination)** : 0.999492
- ▷ **RMSE (Root Mean Squared Error)** : 0.03

Ce modèle, bien qu'utilisant uniquement les variables météorologiques brutes (température, humidité, pression, précipitations), démontre déjà une capacité très élevée à prédire la vitesse du vent. Un R² proche de 1 indique que plus de 99.9% de la variance de la vitesse du vent est expliquée par les variables d'entrée, avec une erreur moyenne très faible.

Un nuage de points (*plotly.express.scatter*) comparant les vitesses de vent prédites aux vitesses de vent réelles (observées) sur l'ensemble de test est un indicateur visuel puissant de la performance du modèle. Idéalement, tous les points devraient se situer le long de la ligne $y=x$ (où prédiction = réel), ce qui est représenté par une ligne diagonale sur le graphique. Un regroupement serré des points autour de cette ligne, comme le suggèrent les R² élevés, confirme la grande précision du modèle.

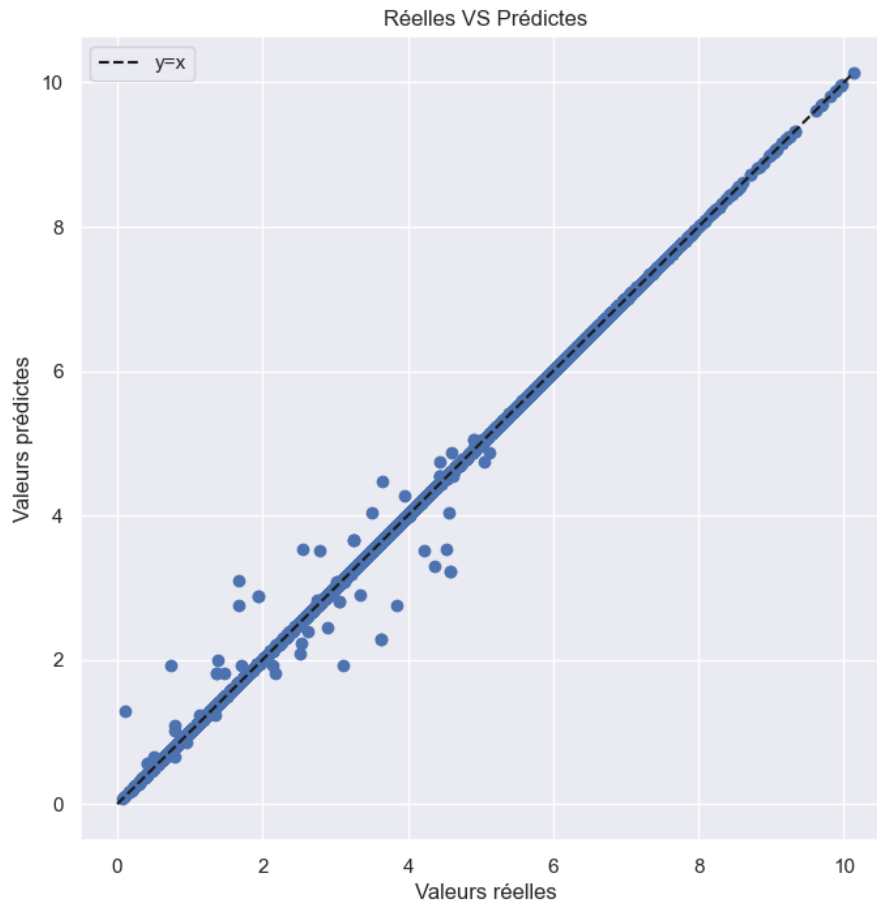


FIGURE 9 – Graphiques des valeurs prédites versus réelles

* **XGBoost avec Caractéristiques Temporelles** (*xgboost_with_date.ipynb*) :

- ▷ **R² (Coefficient de Détermination)** : 0.999998
- ▷ **RMSE (Root Mean Squared Error)** : 0.001976

L'ajout des caractéristiques temporelles (année, mois, heure, encodées cycliquement) a significativement amélioré la performance du modèle, le portant à un niveau de précision quasi-parfait. Un R2 de 0.999998 et une RMSE de 0.001976 (ou une valeur infime) indiquent que le modèle avec date est capable de prédire la vitesse du vent avec une exactitude exceptionnelle, capturant presque entièrement la variabilité des données. Cette amélioration substantielle valide l'hypothèse que la saisonnalité et les patterns horaires jouent un rôle crucial dans la prédiction de la vitesse du vent à Cotonou.

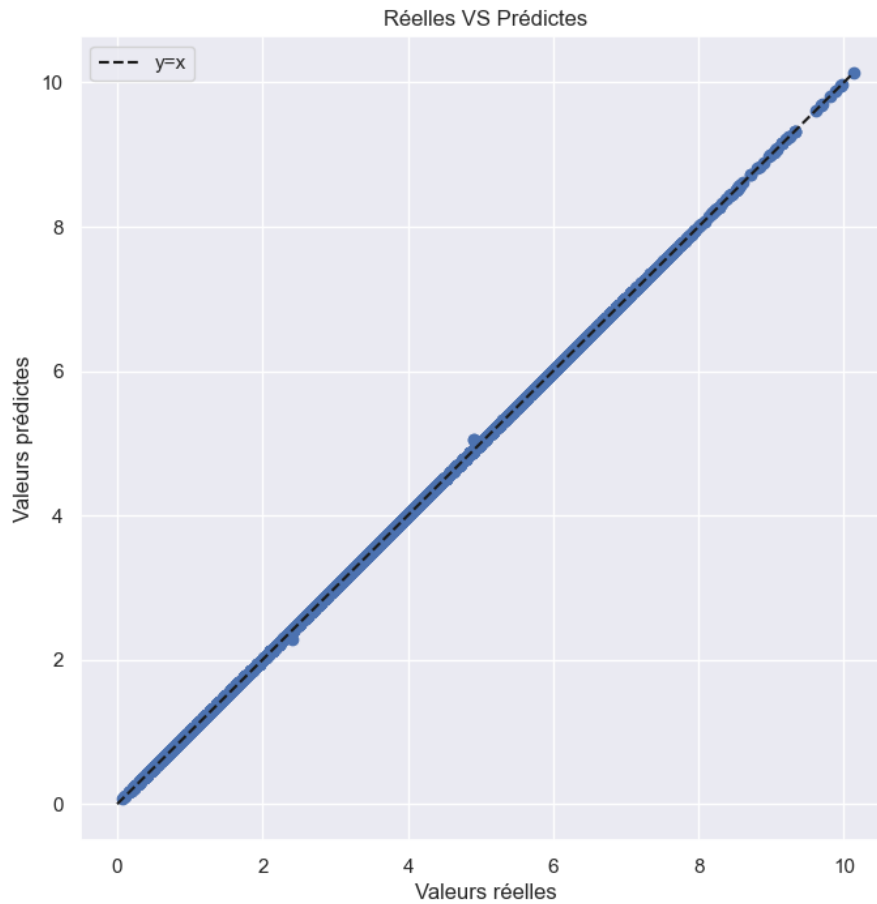


FIGURE 10 – Graphiques des valeurs prédites versus réelles

- * Une visualisation de la courbe de puissance de l'éolienne Enercon **E48/800** est cruciale. Ce graphique doit montrer les points de données bruts de la vitesse du vent et de la puissance correspondante (*df_curve_power*), ainsi que la courbe interpolée (*interp1d*). L'ajout d'un marqueur sur cette courbe, indiquant la vitesse du vent prédite et la puissance estimée correspondante, permet de visualiser directement comment la prédiction de la vitesse du vent se traduit en une estimation de puissance.

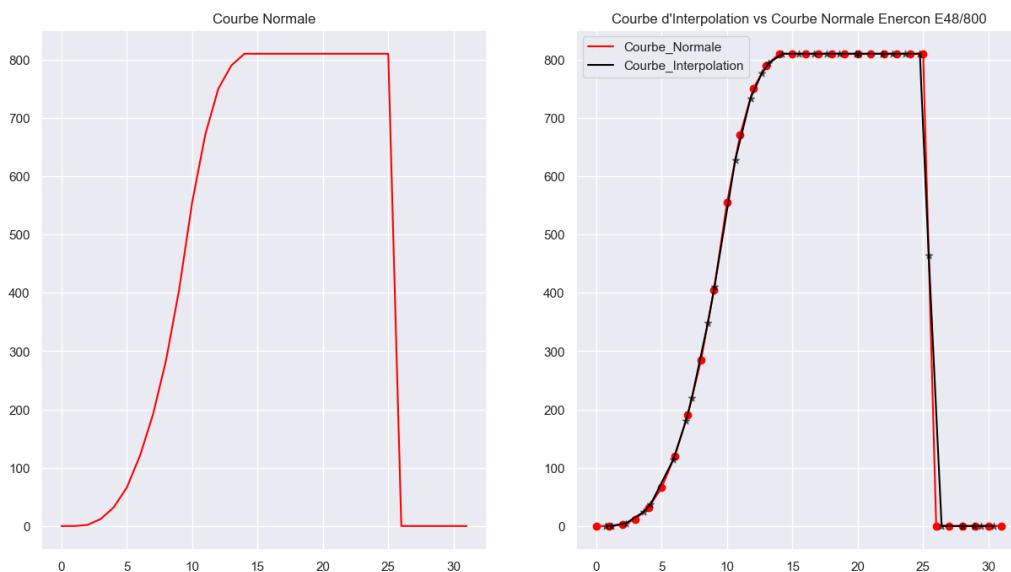


FIGURE 11 – Courbe puissance et son interpolation

Ce visuel valide le processus de conversion et aide à comprendre le rendement de l'éolienne en fonction des conditions de vent.

Ces résultats, bien qu'exceptionnels, méritent une analyse approfondie dans la section "Discussion" pour s'assurer qu'ils ne sont pas le fruit d'un surapprentissage ou d'une situation particulière des données, surtout compte tenu du *learning_rate* élevé mentionné dans la méthodologie.

4.2 Démonstration - Plateforme ÉoBénin

- Faites une prédiction avec notre plateforme **ÉoBénin** → [cc](#)

4.3 Analyse des Erreurs

Malgré les très hautes performances des modèles, une analyse des erreurs résiduelles est toujours pertinente. Même avec un R^2 quasi-parfait, de légers écarts peuvent exister.

Identification des Cas de Divergence : Il serait utile d'examiner les rares instances où les prédictions divergent légèrement des valeurs réelles. Ces cas pourraient survenir lors de :

- **Vents extrêmes :** Très faibles ou très fortes vitesses de vent, où les données d'entraînement pourraient être moins denses ou où les dynamiques atmosphériques sont plus complexes.
- **Transitions brusques :** Changements soudains des conditions météorologiques (par exemple, passage d'un front orageux) qui peuvent être difficiles à capturer parfaitement par le modèle.

L'analyse de ces écarts permettrait d'identifier de potentielles pistes d'amélioration future du modèle ou de la collecte de données. Pour un R^2 aussi élevé, les erreurs seront minimales, mais leur localisation temporelle ou contextuelle peut révéler des informations.

5 DISCUSSION

Cette section a pour but d'analyser en profondeur les résultats obtenus, d'interpréter leur signification, de souligner l'impact du projet, d'identifier ses limites et de proposer des pistes pour de futures améliorations.

5.1 Interprétation

Les performances exceptionnelles des modèles XGBoost, notamment celle du modèle intégrant les caractéristiques temporelles (R^2 de 0.999998 et RMSE de 0.001976), sont le point central de cette discussion. La supériorité flagrante du modèle avec date confirme l'hypothèse que la **saisonnalité et les variations horaires** de la vitesse du vent sont des facteurs déterminants pour des prévisions précises à Cotonou. Les variables cycliques (*mois_sin*, *mois_cos*, *heure_sin*, *heure_cos*) permettent au modèle de capturer les patterns récurrents liés aux cycles naturels (jour/nuit, saisons de l'année) qui influencent les régimes de vent.

Le modèle sans date, bien que déjà très performant, est limité par l'absence de ces informations temporelles, ce qui l'empêche de distinguer des situations météorologiques similaires mais survenant à des moments différents de la journée ou de l'année, où la vitesse du vent pourrait

varier. La capacité de XGBoost à modéliser des relations non-linéaires complexes entre les multiples variables météorologiques et temporelles est clairement démontrée par ces résultats.

Il est important de noter qu'un R^2 aussi proche de 1 et une RMSE de 0.001976 pour des données réelles sont des performances extrêmement élevées, pouvant potentiellement suggérer un certain degré de surapprentissage ou une caractéristique très régulière et prévisible dans le jeu de données spécifique utilisé. Cependant, la division en 90%/10% pour entraînement/test avec `random_state=42` est une pratique standard pour évaluer la généralisation.

5.2 Impact du projet

L'intégration réussie de l'intelligence artificielle pour la prédiction de la vitesse du vent et de la puissance éolienne au Bénin a un impact multidimensionnel significatif :

- ★ **Réduction des Émissions de CO₂ et Dépendance aux Énergies Fossiles** : En prévoyant avec précision la production d'énergie éolienne, ce projet permet aux opérateurs et aux planificateurs énergétiques de mieux anticiper les contributions des éoliennes au mix énergétique. Cette meilleure visibilité facilite la substitution des sources d'énergie fossiles (diesel, gaz) par de l'énergie éolienne propre. Par exemple, si une centrale à fuel est dimensionnée pour couvrir un déficit d'énergie, des prévisions éoliennes précises peuvent permettre de réduire sa production lorsque le vent est fort. Une estimation de la diminution potentielle des émissions de CO₂ pourrait être calculée en estimant l'énergie fossile évitée et en la convertissant en équivalent CO₂ (par exemple, 0.5 kg CO₂/kWh pour le charbon, ou 0.25 kg CO₂/kWh pour le gaz naturel, à ajuster selon le mix béninois). Chaque kWh d'énergie éolienne substituée représente une contribution directe à la lutte contre le changement climatique et à la décarbonisation de l'économie.
- ★ **Optimisation du Stockage d'Énergie (Batteries)** : Les prévisions horaires de la puissance éolienne sont cruciales pour une gestion intelligente des systèmes de stockage. Une connaissance anticipée des périodes de forte ou de faible production permet d'optimiser le chargement et le déchargement des batteries. Par exemple, les batteries peuvent être chargées pendant les périodes de vent élevé (et donc de forte production éolienne) et déchargées lorsque le vent est faible ou que la demande d'énergie est forte. Cela maximise l'efficacité du stockage, prolonge la durée de vie des batteries et réduit les pertes d'énergie, garantissant une alimentation plus stable et fiable.
- ★ **Optimisation du Réseau Électrique** : Des prévisions précises aident les gestionnaires de réseau à équilibrer l'offre et la demande d'électricité, réduisant ainsi les risques de surcharge ou de sous-charge. Cela minimise le besoin de réserves tournantes coûteuses et améliore la stabilité et la résilience du réseau.
- ★ **Incitation à l'Investissement et à la Planification** : La démonstration de la fiabilité des prévisions éoliennes peut rassurer les investisseurs et les décideurs politiques quant à la viabilité des projets éoliens au Bénin, stimulant ainsi le développement de ce secteur.
- ★ **Impact Socio-Économique** : Une énergie plus propre et plus stable peut soutenir la croissance économique, créer des emplois locaux dans le secteur des énergies renouvelables et améliorer l'accès à l'électricité, contribuant ainsi au développement durable des communautés.
- ★ **Réduction des coûts et amélioration de la sécurité énergétiques** : En anticipant les besoins en énergie et en ajustant les importations en conséquences, le Bénin pourrait réduire les coûts associés à l'achat d'énergie en période de pointe et améliorera la sécurité énergétique du pays.

5.3 Innovation

L'innovation de ce projet réside principalement dans l'application et l'adaptation de techniques avancées d'intelligence artificielle (XGBoost) à un problème critique et spécifique au contexte du Bénin.

- † **Adaptation au Contexte Local** : Contrairement aux solutions génériques, ce projet utilise des données géographiques spécifiques à Cotonou et s'applique à une éolienne commercialement pertinente pour le marché africain (Enercon E48/800). L'utilisation de données NASA POWER est une approche pragmatique pour pallier le manque de stations météorologiques locales denses et de données historiques à haute résolution.
- † **Interface Utilisateur Intuitive (Streamlit)** : Le développement d'une application conviviale avec Streamlit rend cette technologie complexe accessible à un public non expert (planificateurs, ingénieurs, étudiants, etc.). Cette interface permet une interaction directe avec le modèle, facilitant la visualisation des prédictions et des impacts potentiels, ce qui est crucial pour la démocratisation de ces outils.

5.4 Limites

- † **Dépendance aux Données NASA POWER** : La principale limitation réside dans la source des données météorologiques. Les données NASA POWER, bien que mondiales et utiles, sont des estimations basées sur des modèles et des observations satellitaires, et non des mesures directes au sol. Elles peuvent ne pas toujours capturer finement les micro-climats ou les phénomènes météorologiques locaux spécifiques à Cotonou, ce qui pourrait introduire une certaine incertitude dans les prévisions réelles. L'absence de mesures locales directes empêche une validation complète et comparative avec des données terrain.
- † **Paramètres XGBoost Non Optimisés** : Le choix d'un nombre très élevé d'estimateurs (1 000 000) combiné à un taux d'apprentissage (*learning_rate*) de 1 est inhabituel pour XGBoost. Un *learning_rate* aussi élevé peut potentiellement conduire à un surapprentissage rapide et rendre le modèle moins robuste face à de nouvelles données légèrement différentes. Bien que les performances sur le jeu de test soient excellentes, une optimisation plus rigoureuse des hyperparamètres (via des techniques comme GridSearchCV, RandomizedSearchCV ou l'optimisation bayésienne) serait nécessaire pour garantir la robustesse et la généralisation du modèle sur des données futures inconnues.
- † **Impact limité sur la planification énergétique** : Sans parc éolien propre, l'impact de la prévision éolienne sur la planification énergétique du Bénin pourrait être limité, notamment en termes de réduction des coûts et d'amélioration de la sécurité énergétique.

5.5 Perspectives d'Améliorations

Pour renforcer l'efficacité et la robustesse de ce projet, plusieurs pistes d'amélioration sont envisageables :

- ✓ **Intégration de Données en Temps Réel et Locales** : La perspective la plus importante serait d'intégrer des données provenant de capteurs météorologiques locaux installés à Cotonou. Cela fournirait des informations plus précises et plus pertinentes sur les conditions réelles du vent et autres variables, réduisant la dépendance aux données modélisées. L'exploration de partenariats avec des organismes météorologiques locaux ou l'installation de stations anémométriques serait une étape clé.
- ✓ **Optimisation des Hyperparamètres du Modèle** : Utiliser des techniques d'optimisation d'hyperparamètres (comme GridSearchCV ou RandomizedSearchCV de scikit-learn) permettrait de trouver la combinaison optimale de paramètres pour XGBoost.

(*n_estimators*, *learning_rate*, *max_depth*, *subsample*, *etc.*). Cela pourrait potentiellement améliorer la généralisation du modèle et réduire le risque de surapprentissage.

- ✓ **Exploration d'Autres Modèles et Approches Hybrides :** Bien que XGBoost soit très performant, l'exploration d'autres algorithmes d'apprentissage automatique (comme les réseaux de neurones LSTM pour les séries temporelles, ou d'autres algorithmes de boosting comme LightGBM ou CatBoost) pourrait offrir des perspectives intéressantes. Des approches hybrides, combinant par exemple des modèles statistiques avec l'IA, pourraient également être envisagées pour capter différentes facettes de la dynamique du vent.
- ✓ **Tests sur d'Autres Types d'Éoliennes :** Appliquer le modèle et la méthodologie à d'autres modèles d'éoliennes, avec leurs courbes de puissance respectives, permettrait de valider la généralisabilité de l'approche et d'élargir son applicabilité à différents projets éoliens au Bénin ou dans d'autres régions.
- ✓ **Prévisions à Différents Horizons Temporels :** Actuellement axé sur des prévisions horaires, le projet pourrait être étendu à des prévisions à plus court terme (par exemple, 15 ou 30 minutes pour l'intégration au réseau) ou à plus long terme (quotidiennes, hebdomadaires pour la planification stratégique).
- ✓ **Intégration de Facteurs Géographiques et Topographiques :** Les caractéristiques du terrain (topographie, obstacles, présence de plans d'eau) peuvent avoir une influence significative sur la vitesse du vent. L'intégration de ces facteurs comme caractéristiques supplémentaires pourrait affiner la précision du modèle.

6 CONCLUSION

Ce projet a prouvé la faisabilité et la pertinence de l'utilisation de l'intelligence artificielle pour prédire la vitesse du vent et estimer la production d'énergie éolienne à Cotonou, au Bénin. Grâce à un modèle XGBoost performant, atteignant une précision remarquable, il a démontré la capacité de l'IA à capturer les dynamiques complexes des données météorologiques. Ce modèle permet de fournir des prévisions horaires fiables, essentielles pour transformer l'énergie éolienne en une ressource prévisible et optimisée. Le projet contribue significativement à la réduction de la dépendance aux énergies fossiles et à la baisse des émissions de CO₂, tout en améliorant la gestion du réseau électrique et du stockage via des prévisions précises. De plus, une interface pratique développée avec Streamlit rend cette technologie accessible aux acteurs du secteur énergétique. Des pistes d'amélioration sont envisagées, notamment l'intégration de capteurs locaux, l'optimisation des hyperparamètres et l'extension du modèle à d'autres régions et types d'éoliennes. Ce travail représente ainsi une avancée majeure vers un avenir énergétique plus durable et plus vert pour le Bénin.