



Projet: J1DASCI

Rapport Modèle Final

ELBAGHAZAOUI Azhar

LIU Kehan

SYLLA Nadia

XIONG Menglin



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

22/12/2023

SOMMAIRE

1. Introduction	1
2. Méthodologie	1
3. Compréhension des données	2
3.1 Collecte et description des données	2
3.2 Qualité des données	2
3.3 Visualisation des données	3
4. Préparation des données	4
5. Modèles de prédiction	4
5.1 Modèles de prédiction de la production d'énergie éolienne	4
5.2 Modèles de prédiction de la consommation d'électricité en France	5
6. Résultats et évaluations des modèles	5
6.1 Résultats et évaluation des modèles de prédiction de la production d'énergie électrique éolienne.	6
6.2 Résultats et évaluation des modèles de prédiction de la consommation d'énergie électrique en France.	7
7. Déploiement	8
7.1 Prévision de la production d'énergie éolienne totale	8
7.2 Prévision de l'équation du mix énergétique au 04/12/2023 et évaluation du modèle	8
8. Limites et pistes d'amélioration	10
9. Conclusion	11
Bibliographie	12

1. Introduction

La France dépend fortement de l'énergie nucléaire, qui représente environ 63% de sa production totale d'énergie électrique en 2022. Les 37% restants proviennent d'autres sources telles que le gaz et les énergies renouvelables [1], [2]. L'objectif est d'augmenter la part des énergies vertes dans les 37% de production d'électricité restantes, en particulier en utilisant au maximum la capacité de production d'énergie électrique éolienne pour réduire le besoin de production d'électricité au gaz. Cependant, la consommation d'énergie électrique est influencée par des facteurs tels que la météo (saison), l'heure et le jour. Il est donc intéressant d'analyser et de prévoir également la demande en électricité, pour qu'elle soit alignée à la production d'énergie électrique.

Comment alors modéliser et optimiser le mix énergétique en France en prévoyant avec précision la consommation d'électricité et la production d'énergie électrique par éoliens ? C'est la question fondamentale à laquelle cette étude ambitionne de répondre.

Développer un modèle de prédiction visant à anticiper et optimiser la production éolienne en fonction de la météo contribuerait à réduire les coûts et les pertes de production. Cette capacité de prédiction revêt une importance particulière dans le cadre de l'engagement de la France et des producteurs à réduire leurs émissions de gaz à effet de serre. De même, la possibilité de prédire la consommation d'énergie électrique permettrait de cibler la réduction de la production d'énergie au strict nécessaire et d'éviter toutes importations d'énergie (car en la prévoyant, on pourrait mettre en marche les outils pour produire et répondre à la demande). Les deux prédictions associées permettent de minimiser la production basée sur le gaz, alignant ainsi le pays sur ses objectifs environnementaux et contribuant à la stabilisation des prix de l'électricité indexés sur le gaz [3].

2. Méthodologie

L'approche proposée pour résoudre le problème énoncé repose sur une méthodologie scientifique utilisant des algorithmes de machine learning. Dans un modèle empirique, la prédiction $j+1$ du mix énergétique se base sur la prévision de la consommation et de la production d'énergie éolienne selon l'équation suivante:

$$\text{Demande d'électricité en France} = 63\% \text{ nucléaire} + 11\% \text{ hydraulique} + \text{production éolienne} + \text{gaz}$$

La première phase consiste à utiliser un modèle Random Forest [4] et d'autres modèles en apprentissage supervisé sur les données du parc éolien de St Nazaire pour prédire la production électrique du parc selon les données météorologiques. La deuxième phase étend le meilleur modèle à l'ensemble des parcs éoliens de France en tenant compte de leurs caractéristiques techniques et les données météorologiques de leur emplacement.

Une fois la prédiction de la production d'énergie éolienne établie, la démarche s'attaque à d'autres éléments de l'équation du mix énergétique en prédisant la consommation du $j+1$. Nous souhaitons utiliser les données météorologiques pour prédire la consommation d'électricité dans toute la France. Cependant, en raison des variations significatives des

données météorologiques dans chaque région, il n'est pas possible d'effectuer une prédiction simple à l'aide d'un ou de quelques modèles. Ainsi, nous avons choisi de simplifier le processus de prédiction en utilisant les variations saisonnières de la consommation au fil des années, parce que les changements saisonniers de la consommation sont suffisamment corrélés aux variations météorologiques. En analysant la saisonnalité de la consommation, un modèle LSTM, ou Sarima peut être appliqué pour prédire la consommation en fonction de ses caractéristiques saisonnières et temporelles [cf cours Série temporelles].

3. Compréhension des données

3.1 Collecte et description des données

Afin de mettre en place notre approche pour la résolution du problème, nous avons d'abord collecté des données de consommation électrique heure par heure en France du 1/01/2021 au 01/11/2023 via le site RTE [5]. Sous forme de csv, ces données collectées ont des colonnes principales que sont : **'Date'** au format jj/mm/aaaa, **'Heure'** au format 'hh:mm:ss' et enfin **'Consommation d'énergie électrique'** en MW.

Ensuite, nous avons récupéré la liste des parcs éoliens terrestres, leurs emplacements et leurs puissances de production via le site du gouvernement [8]. Par le biais de recherches, il n'a pas été nécessaire de chercher une liste des parcs éoliens en mer car un seul est en fonctionnement en France et il s'agit de celui de Saint nazaire.

Par la suite, nous avons récupéré par le biais d'un API et d'un code python, les données météorologiques des différents parcs éoliens terrestres et en mer qui nous intéressent [6]. Ces données météorologiques possèdent les attributs (colonnes) intéressantes pour notre études que sont : **'Location'** (l'emplacement du parc), **'Date'** (information complète heure et date de prise), **'température'** (la température au parc), **'précipitation'** (en mm) , **'windgust'** (rafale de vent, kph), **'windspeed'** (la vitesse de vent, kph) et la **cloudcover** (la fraction du ciel couverte par les nuages en ce moment, %), **'winddir'** (la direction de vent, degré, mesurée par rapport au nord, dans le sens des aiguilles d'une montre. 360 degré est généralement associé au nord, 90 degrés correspond à l'est, 180 degrés au sud, et 270 degrés à l'ouest.) et **'humidity'** (la quantité de vapeur d'eau présente dans l'air, %).

Enfin, à l'aide d'un API également, nous avons téléchargé au format csv, la production du parc éolien de St Nazaire heure par heure [7]. Ce csv comprend la **'Date et l'heure'** et la **Production** du Parc pour les dates et heures données.

3.2 Qualité des données

Étant donné que les données importantes dans les différentes bases de données sont toutes numériques, nous avons utilisé des box-plots, qui permettent de voir que les données météorologiques et les données de consommation sont bonnes et qu'il n'y a globalement pas de données aberrantes. Pour la production d'énergie, il y a assez de données *Null*, du

fait que les enregistrements sont présentés dans la dataset par groupe d'éolienne sur le parc. Un traitement sur ces données a permis de regrouper la production des groupes sur les mêmes dates dans la base de données et donc d'éliminer ces valeurs *Null*.

3.3 Visualisation des données

Durant l'étude, nous avons effectué plusieurs visualisations pour comprendre les données et vérifier qu'elles permettent de répondre à la problématique. Parmi ces visualisations, certaines visualisations (ci-dessous) nous ont notamment permis de comprendre les relations entre données à prédire et données d'entrée.

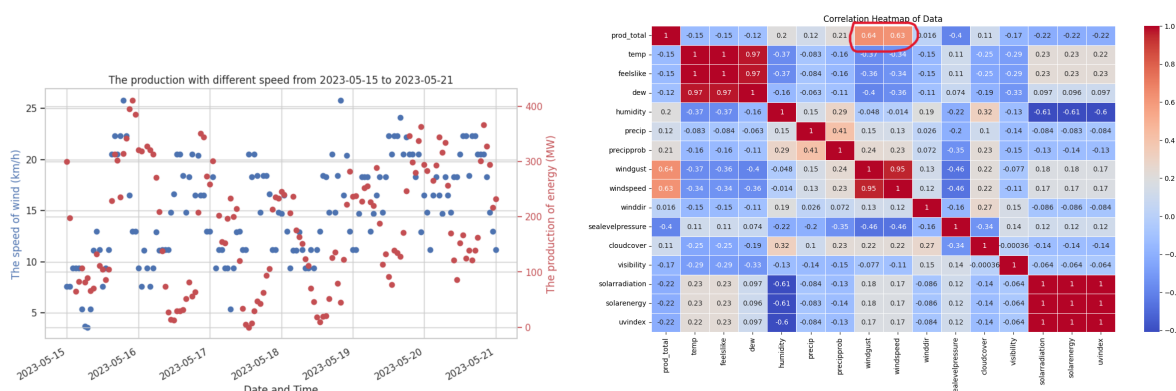


Figure 1: Production éolienne et vitesse de vent à St Nazaire. Figure 2: matrice de corrélation de Pearson entre les paramètres météorologiques et la production à St Nazaire

La figure 1 permet de voir qu'il n'y a pas de relation évidente entre la production d'électricité au parc éolien de St Nazaire et la météo sur le parc de St Nazaire. La matrice de corrélation de Pearson quant à elle, permet de voir que parmi les différentes variables dans la data météo, la vitesse du vent et des rafales sont les variables les plus corrélées à la production d'électricité sur le parc.

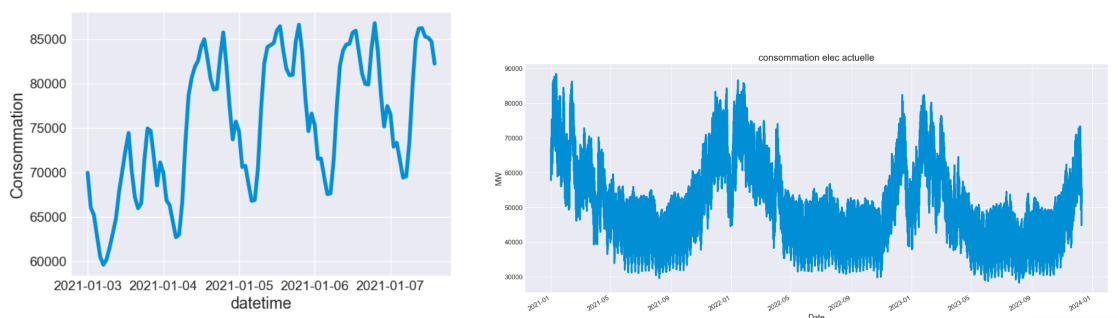


Figure 3: Consommation d'électricité en France sur 4 jours. Figure 4: Consommation d'électricité en France sur 3 ans

La figure 3 et 4, permettent de voir qu'il y a une saisonnalité journalière et annuelle dans la consommation d'électricité en France. Un troisième graphique hebdomadaire permettant de visualiser 3 semaines de consommation a été effectué mais ne permet pas de constater une saisonnalité hebdomadaire dans la consommation d'électricité en France.

4. Préparation des données

Dans cette section, nous avons manipulé les données en effectuant des opérations de sélection, de nettoyage, de construction, d'intégration et de formatage. Notre datasets contient un large éventail d'informations sur les conditions météorologiques et la production d'énergie éolienne pour le parc éolien de Saint-Nazaire, ainsi que la puissance maximale et le nom de la ville où chaque parc éolien est situé en France.

Au cours de l'analyse des données, nous avons initialement effectué une sélection de données en identifiant et en supprimant des colonnes non pertinentes après un examen approfondi et en utilisant une matrice de corrélation pour la sélection de variables.

Dans la phase de nettoyage des données, nous avons traité les valeurs manquantes et les anomalies, résolvant l'incohérence des données en fusionnant les productions des différents groupes. Ensuite, en ajustant et en formant la colonne de dates, nous avons créé une base temporelle homogène pour faciliter l'intégration future des données et les analyses ultérieures. L'étape d'intégration a été réalisée en utilisant la fonction de fusion de pandas, combinant les données de production éolienne, météorologique et de consommation en un nouveau Data Frame consolidé.

Pour préparer les données à l'apprentissage de l'algorithme, nous avons structuré les ensembles d'entraînement et de test à l'aide de la fonction 'train_test_split' de sklearn, attribuant 80% des données à l'ensemble d'entraînement (X_train, Y_train) et 20% à l'ensemble de test (X_test, Y_test), où X représente les données météorologiques en entrée et Y les données de production éolienne à prédire. Enfin, pour le formatage des données, nous avons créé une nouvelle colonne représentant l'énergie totale produite dans l'ensemble du parc éolien, consolidant ainsi les informations sur la production énergétique.

Concernant la préparation des données de consommation, nous avons dû décomposer la colonne 'Date et heure' en plusieurs colonnes (une colonne pour le mois, le jour de la semaine, l'heure, l'année, etc.) afin de pouvoir utiliser le modèle Random Forest. En effet, en faisant cela, le modèle pourra comprendre et analyser toutes les relations entre les différents éléments d'une date dans l'année et la consommation d'électricité en France. Pour les autres modèles de prédictions utilisés, cette décomposition n'a pas été faite sur les données car ils peuvent directement comprendre la colonne 'Date et heure' de type DateTime pour faire les prédictions.

5. Modèles de prédiction

5.1 Modèles de prédiction de la production d'énergie éolienne

Pour la prédiction de la production d'énergie électrique éolienne à partir des données météorologiques, nous avons utilisé plusieurs modèles d'apprentissage supervisé :

- La régression linéaire est une méthode classique qui cherche à établir une relation linéaire entre les variables explicatives et la variable cible.
- Le modèle de régression RandomForest est une technique basée sur l'agrégation d'arbres de décision, offrant une bonne capacité à capturer les relations non linéaires.
- La régression KNN (K-Nearest Neighbors) est une méthode qui se base sur la similarité entre les observations.
- Le Gradient Boosting est une technique d'ensemble qui combine plusieurs modèles simples pour créer un modèle prédictif plus puissant.
- SVR (Support Vector Regression) est basée sur les vecteurs de support, adaptée pour traiter des données complexes et non linéaires.

```
# Initialisation des modèles de régression
models = {
    'Linear Regression': LinearRegression(),
    'Random Forest Regression': RandomForestRegressor(),
    'KNN Regression': KNeighborsRegressor(),
    'Gradient Boosting Regression': GradientBoostingRegressor(),
    'SVR': SVR()
}
```

Figure 5: Les modèles utilisés pour la prédiction de la production de l'énergie éolienne

Ces modèles ont été choisis pour leur capacité à traiter des données complexes et à capturer des relations non linéaires entre la prédiction de l'énergie électrique éolienne et les variables météorologiques.

5.2 Modèles de prédiction de la consommation d'électricité en France

Pour la prédiction de la consommation d'électricité en France, nous avons utilisé 2 modèles qui permettent de capturer des relations complexes, temporelles et sur du long terme dans des séries temporelles:

- Le modèle de régression Random Forest est une technique basée sur l'agrégation d'arbres de décision, offrant une bonne capacité à capturer les relations non linéaires.
- Le modèle LSTM (Long Short-Term Memory) est une technique basée sur le modèle du réseau de neurones récurrents classiques qui gère en plus les dépendances à long terme dans les séquences de données.

NB: Pour appliquer le RandomForest à la série temporelle de consommation, nous avons dû décomposer la date en plusieurs variables (cf, voir partie data preparation).

6. Résultats et évaluations des modèles

L'évaluation des différents modèles de prédiction est importante pour garantir la fiabilité des prévisions et la comparaison des modèles entre eux pour déterminer le meilleur modèle. À

cette fin, les modèles choisis ont été évalués à l'aide de différentes métriques de performance suivantes:

- MSE (Mean Squared Error) : Mesure de la moyenne des carrés des différences entre les vraies valeurs et les valeurs prédites.
- RMSE (Root Mean Squared Error) : Racine carrée de la MSE, offrant une mesure de l'erreur moyenne.
- MAE (Mean Absolute Error) : Moyenne des valeurs absolues des écarts entre les valeurs prédites et les vraies valeurs.
- R^2 (Coefficient de détermination) : Indique la proportion de variance dans la variable dépendante qui est prévisible à partir de la variable indépendante.
- NMAE (Normalized Mean Absolute Error) : MAE normalisé pour prendre en compte l'échelle des données.

6.1 Résultats et évaluation des modèles de prédiction de la production d'énergie électrique éolienne.

<pre>Modèle: KNN Regression Temps d'exécution: 0.0147 secondes MSE: 6061.1382550335575 RMSE: 77.85331242171753 MAE: 51.43892617449664 R²: 0.7805647693186992 NMAE: 0.0539757882208779</pre>	<pre>Modèle: Linear Regression Temps d'exécution: 0.0432 secondes MSE: 12101.398650964815 RMSE: 110.00635732067859 MAE: 80.09529112606013 R²: 0.5618853930059142 NMAE: 0.08404542615536216</pre>
<pre>Modèle: Gradient Boosting Regression Temps d'exécution: 1.5493 secondes MSE: 6117.952133729596 RMSE: 78.2173390862256 MAE: 55.098677767566606 R²: 0.778507900451357 NMAE: 0.05761603123564177</pre>	<pre>Modèle: Random Forest Regression Temps d'exécution: 7.6430 secondes MSE: 5010.10581852349 RMSE: 70.7821009756244 MAE: 47.09331543624162 R²: 0.8186159629154768 NMAE: 0.04941586089846969</pre>
<pre>Modèle: SVR Temps d'exécution: 1.7364 secondes MSE: 17151.523167291205 RMSE: 130.96382388771033 MAE: 94.02396042590985 R²: 0.3790525363513924 NMAE: 0.09866102877849932</pre>	

Figure 6: Résultats de performance des modèles de prédiction de la consommation

Les résultats des métriques d'évaluation ci-dessus ont révélé des variations significatives dans les capacités prédictives des différents modèles. Le modèle de Random Forest Régression s'est démarqué en affichant une meilleure performance globale avec des valeurs de MSE et RMSE inférieures par rapport aux autres modèles, témoignant d'une précision accrue dans ses prédictions.

En revanche, le modèle SVR (Support Vector Regression) a présenté une performance relativement inférieure avec des valeurs plus élevées de MSE et RMSE, indiquant une moindre précision dans ses prédictions. Les modèles de KNN, Gradient Boosting et Linear Regression ont affiché des performances intermédiaires, chacun avec ses forces et ses limitations respectives en termes de précision et d'ajustement aux données.

En tenant compte de ces performances, le modèle de Random Forest Régression a été choisi comme le plus performant pour la prédiction de la production d'énergie éolienne, en raison de sa précision supérieure et de sa capacité à mieux déterminer les relations complexes entre les données prédites et d'entrée.

6.2 Résultats et évaluation des modèles de prédiction de la consommation d'énergie électrique en France.

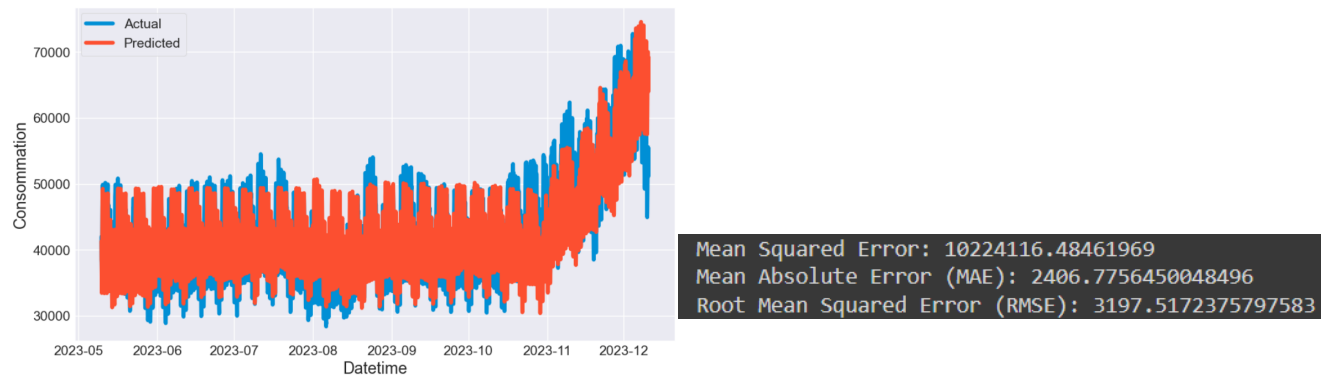


Figure 7: Résultats de performance du modèle Random Forest

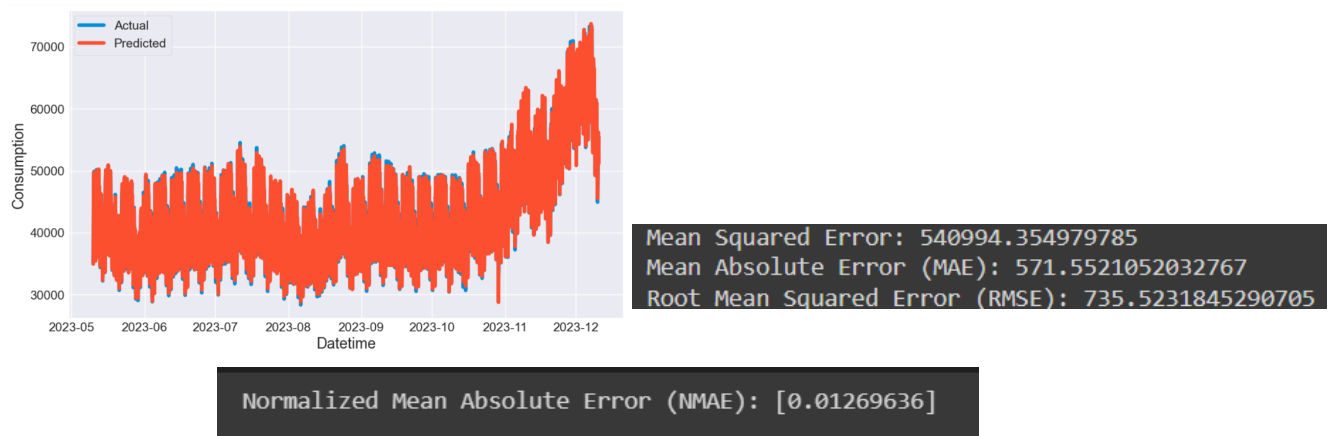


Figure 8: Résultats de performance du modèle LSTM

Bien que le temps d'exécution et d'apprentissage du modèle LSTM soit plus long que celui du RandomForest, il permet d'obtenir de meilleurs résultats dans la prédiction de la consommation d'électricité en France, comme le démontre le graphique des prédictions sur les données tests ainsi que les 3 métriques calculées pour évaluer les modèles (MSE, MAE, RMSE).

Nous avons calculé une métrique en plus (NMAE) pour le meilleur modèle LSTM afin de comprendre sa valeur relativement élevée de MAE qui est de 571,55. Nous constatons que le NMAE est très proche de 0, d'une valeur exacte de 0,0127, ce qui suggère qu'en moyenne, les prédictions sont très proches des valeurs réelles, avec seulement une petite erreur absolue.

7. Déploiement

7.1 Prévion de la production d'énergie éolienne totale

Afin de généraliser à l'ensemble de la France et de calculer la production éolienne totale, nous utilisons la règle de trois qui établit une relation proportionnelle entre la production maximale de chaque parc éolien et la production totale prévue. Cette approche permet d'avoir une vision agrégée de la production éolienne en France en prenant en compte la capacité maximale de chaque parc, offrant ainsi une estimation globale de la contribution éolienne à l'approvisionnement énergétique.

$$\frac{C(\text{La puissance de parc à Saint-Nazaire})}{C(\text{La puissance d'un parc éolien})} = \frac{P(\text{La production de parc à Saint-Nazaire})}{P'(\text{La production d'un parc éolien})}$$

En France, il existe actuellement 3392 parcs éoliens, chacun ayant une production maximale d'énergie éolienne connue. Pour estimer la production totale d'énergie éolienne en France, nous pouvons appliquer la formule mentionnée précédemment à chaque parc éolien et calculer la somme de ces productions individuelles.

En utilisant cette formule pour chaque parc éolien en France, nous pouvons prédire la production totale d'énergie éolienne par heure ou par jour dans toute la France. Il est important de noter que cette estimation est basée sur les productions maximales théoriques de chaque parc et ne prend pas en compte des pannes ou d'autres facteurs imprévus.

7.2 Prévion de l'équation du mix énergétique au 04/12/2023 et évaluation du modèle

Nous avons appliqué la formule de la production éolienne totale au 4 décembre 2023. En raison des restrictions de téléchargement sur les sites météorologiques et dans le but de simplifier le calcul, nous avons choisi des parcs éoliens d'une capacité supérieure à 30 MW comme échantillon pour notre analyse. Il y a ainsi 126 parcs éoliens répondant à ces critères dans toute la France. Nous avons collecté des données météorologiques pour les emplacements de ces 126 parcs éoliens. Par la suite, ces données météorologiques seront appliquées à notre modèle pour prédire la production d'électricité. La valeur prédite, notée P, est basée sur la capacité de production d'électricité du parc éolien de Saint-Nazaire.

Pour estimer la production d'électricité de chaque parc éolien, nous utiliserons la formule présentée à la section 7.1. En agrégeant les productions prédites de tous les parcs éoliens, nous obtiendrons la production totale d'énergie éolienne heure par heure pour l'ensemble de la France au 04/12/2023.

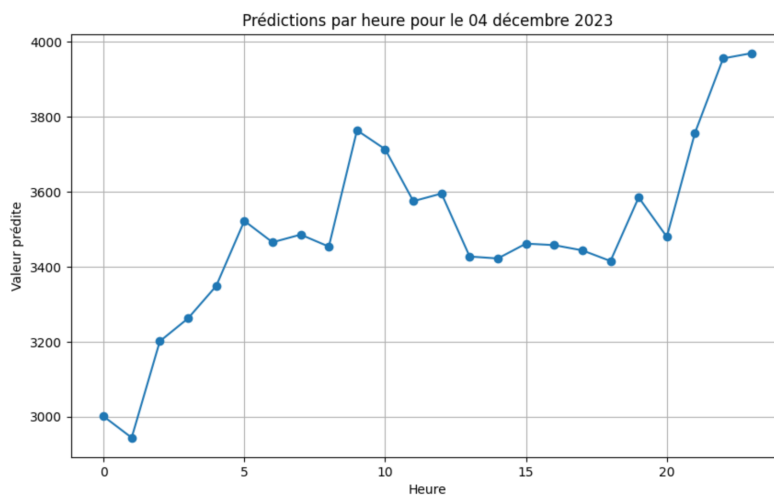


Figure 9 : Production d'électricité prédite heure par heure le 04/12/2023 en France

En parallèle, en se basant sur le modèle de prédiction de la consommation d'électricité heure par heure, nous obtenons ces résultats pour le 4 décembre 2023.

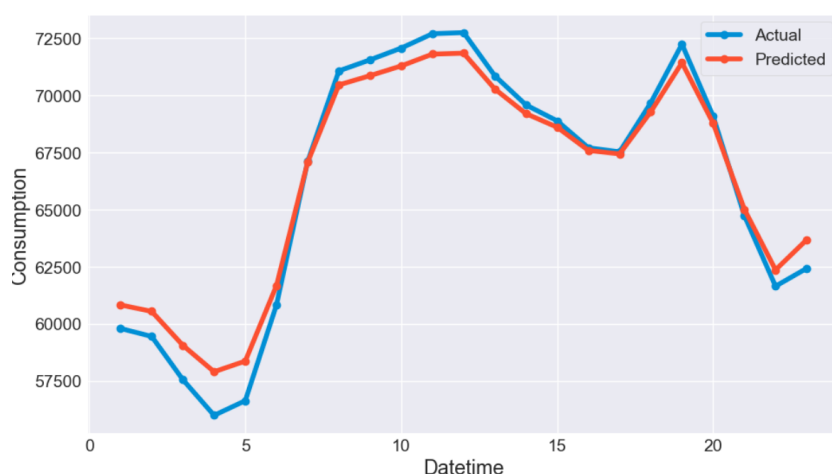


Figure 10: Consommation d'électricité prédite heure par heure le 04/12/2023

Avec ces prévisions, il est aisé d'obtenir le pourcentage de la production éolienne et de calculer la consommation de gaz. Par exemple, à 10 heures, le calcul devient :

Demande d'électricité en France = 63% nucléaire + 11% hydraulique + production éolienne + gaz

- Demande d'électricité = 71 298.33 MW
- Production éolienne = 3 713.47 MW (Pourcentage éolien prédite = 5.21%)
- Production nucléaire = $71\,298.33 \times 63\% = 44\,917.95$ MW
- Production hydrauliques = $71\,298.33 \times 11\% = 7\,842.82$ MW
- Consommation de gaz = $71\,298.33 - 44\,917.95 - 7\,842.82 - 3\,713.47 = 14\,824.09$ MW

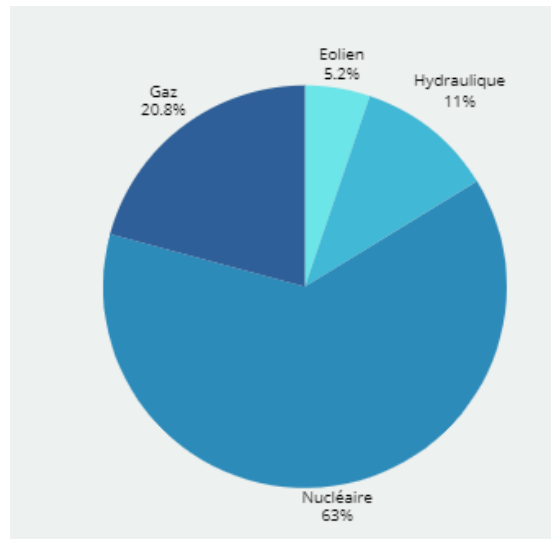


Figure 11: Prédiction de l'équation du mix énergétique le 04/12/2023 à 10h

Ces chiffres montrent qu'une part importante de la production totale d'électricité était générée par des sources non renouvelables telles que le nucléaire et le gaz. En effet, la consommation de gaz pour répondre à la demande électrique était de 14 824.09 MW, ce qui suggère une dépendance significative à cette source non renouvelable pour combler l'écart entre la demande totale et la production provenant des sources renouvelables prises en compte. Il est également crucial de noter que la production éolienne répertoriée ne représente pas l'ensemble de la production éolienne en France, car tous les parcs terrestres d'éoliennes n'ont pas été pris en compte.

8. Limites et pistes d'amélioration

Dans notre étude, nous avons démontré la possibilité de prédire la consommation d'électricité et la production électrique éolienne dans l'équation du mix énergétique. Cependant, nous sommes conscients que la modélisation de cette équation reste empirique, car le pourcentage d'énergie hydraulique produite peut varier en fonction de la disponibilité ou du fonctionnement des barrages hydrauliques. De même, le pourcentage d'énergie nucléaire produit pour répondre à la demande n'est pas constant au cours d'une journée (il y a des temps d'arrêt sur les 24h d'une journée). La production d'électricité à partir d'autres sources vertes telles que le soleil ou la biomasse ont été omises dans notre équation par souci de complexité et de non prédictibilité de ces énergies. Il serait donc important d'ajouter toutes les énergies ainsi que la variabilité de leurs pourcentages (contributions) dans l'amélioration de la modélisation de l'équation du mix énergétique.

De plus, notre modèle de prédiction de la production d'énergie éolienne a été entraîné uniquement sur le parc de St Nazaire, étant donné que les autres parcs éoliens ne mettent pas à disposition leurs données de production horaire. Cela a conduit à l'utilisation d'un calcul basé sur la règle de trois, en se fondant sur les capacités des parcs, pour généraliser le modèle de prédiction aux autres parcs et calculer la production totale en France. Pour l'amélioration des prédictions et du calcul de la production totale, il serait intéressant

d'entraîner le modèle avec les données météorologiques et de productions des autres parcs directement; données qui peuvent être récupérées avec les opérateurs des différents parcs.

Par ailleurs, lors de la généralisation, il a été nécessaire de restreindre la liste des parcs terrestres pris en compte dans le calcul de la production totale d'énergie éolienne. En effet, l'API utilisée pour récupérer les données météorologiques ne permettait pas d'obtenir gratuitement des données pour les 3000 parcs terrestres en France. De plus, nous avons rencontré un problème d'exactitude dans les données météorologiques des parcs éoliens. En effet, les données météorologiques des parcs sont données par des stations qui sont souvent à 50 KM de la localisation exacte du parc, ce qui peut laisser une imprécision sur la météo exacte du parc en question. Pour l'amélioration de la précision des prédictions de production d'énergie éolienne en France à partir des données météorologiques, il faudrait envisager de collaborer avec les fournisseurs de données météorologiques en France pour être en possession de toutes les données météorologiques précises de tous les parcs dans l'étude.

9. Conclusion

Dans cette étude, nous nous sommes proposés de répondre à la problématique portant sur la méthode et la faisabilité de la modélisation et de l'optimisation de l'équation du mix énergétique en avance en France. À la fin de notre étude, nous avons pu déterminer et développer des modèles de machine learning qui nous permettent de prédire, au jour $j+1$, deux éléments de cette équation, à savoir la demande (la consommation) d'électricité et la production totale des énergies éoliennes en France. Toutefois, dans notre étude, nous avons démontré la faisabilité de la prédiction de l'équation du mix énergétique en France. Bien que l'équation soit empirique avec des pourcentages fixes des d'autres sources d'énergie en dehors des éoliens, il est envisageable d'élargir le projet en incluant d'autres énergies et en tenant compte de la flexibilité de leurs pourcentages tout au long de la journée. La collaboration avec des fournisseurs de données météorologiques est recommandée pour améliorer la généralisation et le calcul de la production éolienne totale en France. De plus, l'exploration de l'utilisation d'algorithmes d'optimisation, intégrés au machine learning, pourrait permettre d'optimiser les pourcentages des sources d'énergie de l'équation du mix énergétique en faveur des énergies vertes pour répondre à la demande.

Bibliographie

- [1] <https://prix-elec.com/energie/production/mix-energetique>
- [2] https://www.rte-france.com/actualites/bilan-electrique-2022?fbclid=IwAR158_C84WXimQtYBCu8ughPvk9rUVVkv3WoMI9zvK4zJ5rKA5rN4cjYNig (Bilan Mix 2022)
- [3] https://www.lemonde.fr/les-decodeurs/article/2022/10/06/pourquoi-le-prix-de-l-electricite-d-epend-de-celui-du-gaz-et-autres-questions-sur-les-factures-a-venir_6140985_4355771.html
- [4] <https://www.sciencedirect.com/science/article/pii/S0960148117302550#sec6> (Rapport scientifique sur l'utilisation du Random Forest pour l'estimation de production éolienne)
- [5] <https://www.rte-france.com/eco2mix/telecharger-les-indicateurs> (Données consommations)
- [6] <https://www.visualcrossing.com/weather/weather-data-services> (Données meteorologiques)
- [7] <https://www.services-rte.com/fr/visualisez-les-donnees-publiees-par-rte/production-realisee-par-groupe.html> (Données production électrique parc Saint Nazaire)
- [8] <https://www.georisques.gouv.fr/donnees/bases-de-donnees/eolien-terrestre> (liste des parcs éoliens terrestres)