



17<sup>e</sup> Congrès Français d'Acoustique  
27-30 avril 2025, Paris

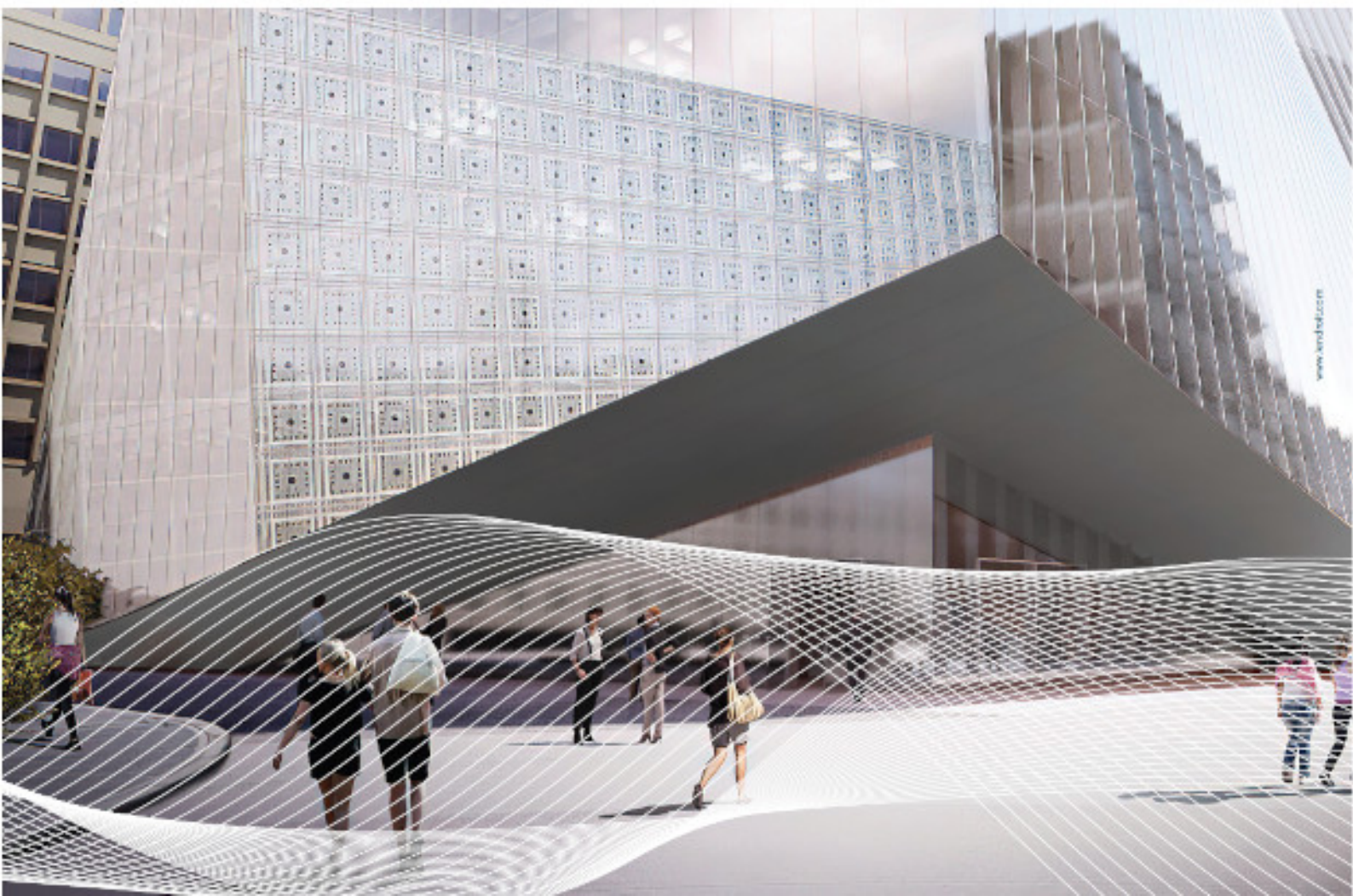
## Estimation du niveau de bruit éolien à l'aide de l'apprentissage profond : apport des données environnementales

A. RKHISS <sup>a</sup>, A. FINEZ <sup>a</sup>, J.-R. GLOAGUEN <sup>a</sup>, J. MAILLARD <sup>b</sup> et G. VASILE <sup>c</sup>

<sup>a</sup> Engie Green, 6 Rue Alexander Fleming, 69007 Lyon, 69007 Lyon, France

<sup>b</sup> CSTB, 24 Rue Joseph Fourier, 38400 Saint-Martin-d'Hères, 38400 Grenoble, France

<sup>c</sup> Grenoble INP-UGA, GIPSA-Lab, 38400 Grenoble, France



La séparation de sources appliquée aux bruits des éoliennes représente un enjeu crucial pour l'analyse acoustique, en raison de la nature complexe et non stationnaire des signaux environnementaux issus de différentes sources du bruit. Ce travail explore l'application des réseaux de neurones récurrents (RNN) à l'extraction du niveau de bruit éolien (particulier) à partir d'un bruit total (ambiant) construit par une synthèse sonore, combinant un bruit particulier simulé et un bruit de fond (résiduel) mesuré. En complément, des données supplémentaires (environnementales), incluant la vitesse du vent mesurées, et la production électrique de l'éolienne, ainsi que des données acoustiques dans la plage de fréquences entre 31,5 Hz et 2 kHz, ont été utilisées pour l'apprentissage et le test de modèle RNN. Cette approche vise à améliorer la performance des réseaux de neurones en leur permettant de mieux identifier les relations entre les facteurs environnementaux et les signaux acoustiques.

## 1 Introduction

L'approvisionnement en énergie a toujours influencé la dynamique géopolitique et la coexistence des peuples. Depuis la fin du XVIII<sup>e</sup> siècle, les combustibles fossiles ont été la principale source d'énergie dans le monde. Cependant, la crise énergétique du début des années 1970, ainsi que la prise de conscience progressive de l'impact sur l'environnement suite à l'augmentation drastique de la température moyenne terrestre, ont progressivement poussé la recherche technologique vers de nouvelles sources d'énergie propre et renouvelable, en premier lieu, les énergies éolienne et photovoltaïque [1]. Cependant, l'impact du bruit généré par les éoliennes sur les riverains est devenu une préoccupation croissante. L'augmentation de la taille des parcs éoliens, rend les questions de bruit de plus en plus sensibles. Ces niveaux de bruit peut affecter la qualité de vie des résidents, notamment en perturbant leur sommeil et leur état de santé général. Freiberg et al.[2] fournit une étude sur les effets potentiels sur la santé du bruit des éoliennes dans les zones résidentielles. En réponse à ces préoccupations, en France on vérifie l'impact sonore d'un parc via le critère d'émergence qui se calcule comme la différence entre le bruit ambiant et le bruit résiduel, à différents période de la journée et pour différents secteurs de vent.

Une stratégie de marche/arrêt des éoliennes est couramment utilisée pour vérifier un plan de bridage employé pour réduire les émissions sonores.

- Un plan de bridage est mis en place afin de limiter l'impact sonore des parcs éoliens et de garantir le respect de la réglementation.
- En phase d'exploitation, ce plan est vérifié par des mesures effectuées lors de cycles de marche/arrêt du parc.
- Cependant, ces mesures présentent plusieurs inconvénients :
  - Leur durée limitée ne reflète pas toujours la variabilité du bruit résiduel.
  - Elles sont coûteuses puisqu'elles nécessitent l'arrêt des éoliennes.
  - Selon les cas, elles peuvent conduire à un sur-bridage ou un sous-bridage du parc.

Pour une surveillance efficace et un plan de bridage dynamique et plus fiable, un outil automatique permettant d'estimer l'émergence sans arrêter les éoliennes serait

bénéfique. Une approche basée sur la factorisation de matrices non négatives (NMF) a déjà été proposée par Gloaguen et al. [3]. malgré des résultats intéressants dans certains cas, l'incertitude de la méthode ne permet pas de la déployer à une échelle industrielle. En conséquence, nous nous orientons vers l'utilisation de réseaux de neurones profonds, compte tenu de leur succès dans le domaine de la séparation des sources [4], de la classification des sons [5] et de la détection d'événements sonores [6]. Une étude précédente basée sur une régression par machines à vecteurs de support (SVR) par Anicic et. al. [7] confirme l'importance d'intégrer des données complémentaires, telles que la vitesse du vent, pour augmenter la précision des modèles de prédiction.

La section (2) donne un aperçu complet de la méthodologie. Commenant par le processus de construction de l'ensemble de données, à partir des mesures du bruit de résiduel et de la synthèse de bruit éolien (2.1). Continuant avec la description des données acoustiques et météorologique (2.2), afin de présenter dans la sous-section (2.3) l'ensemble des prétraitements sur les données. La sous-section (2.4) décrit l'architecture de réseau de neurone et la cellule GRU. Finalement les résultats et discussions sont présentés dans la section (3).

## 2 Méthodologie

### 2.1 Ensemble de données d'apprentissage

L'objectif est d'estimer le niveau de bruit particulier ( $L_{BP}$ ) à partir du bruit ambiant mesuré ( $L_{BA}$ ). Afin d'évaluer l'émergence sonore. Une architecture de réseaux de neurone supervisé est proposée. Pour son entraînement il est essentiel de définir un ensemble de données d'entraînement approprié, avec  $L_{BA}$  et  $L_{BP}$  les niveaux SPL pondéré A du bruit ambiant et particulier respectivement, en bandes de tiers d'octave allant de 31,5 Hz à 2 kHz.  $L_{BA}$  représente l'entrée de modèle et  $L_{BP}$  est la sortie (target). La mesure de  $L_{BP}$  pendant les mesures représente un défi en raison de la présence de bruit résiduel provenant de diverses sources qui ne peuvent pas être isolées. Pour créer des scènes sonores réalistes, une méthode mixte de mesure et de synthèse sonore est adoptée. Cette méthode génère le bruit particulier à partir de données de vent et des spécifications des machines (voir sous-section 2.1.2). Elle lui associe un bruit résiduel mesuré sur un site en développement lors d'un état initial (2.1.1) en cohérence avec des mesures de vent fait à 81 m de hauteur.

### 2.1.1 Bruit résiduel

On a mesuré des niveaux sonores en bandes de tiers d'octave, échantillonné à la secondes du bruit résiduel afin d'avoir une résolution fine et détailler de l'environnement sonore selon ce que peuvent mesurer les sonomètres. Ces mesures sont accompagnées de données de mesure de vitesse de vent échantillonnées sur une période de 10 minutes. La Figure 1, présente un échantillon des niveaux de bruit résiduel mesurés en dB(A).

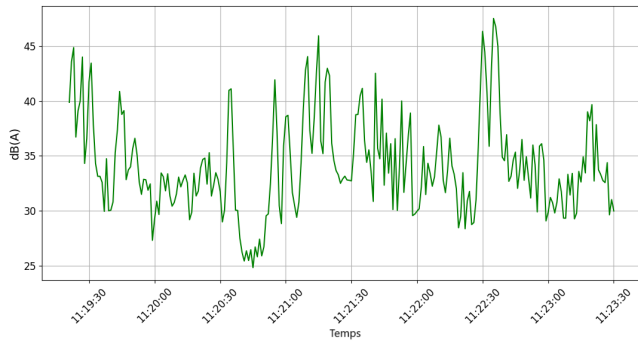


FIGURE 1 – Un échantillon de bruit résiduel mesuré.

### 2.1.2 Bruit particulier

Chaque éolienne est accompagnée d'un document de spécifications de la machine (SM), qui fournit des estimations du niveau de puissance acoustique ( $L_w$ ) par bande de fréquence d'octave, en dB(A), pour différentes valeurs de vitesse de vent, basées sur la norme internationale ISO 61400-11. Une turbine industrielle de 2 MW avec un diamètre de 90 m a été choisie pour cette étude. Ces niveaux sont disponibles pour des vitesses de vent allant de 3 à 25 m/s par pas de 1 m. Ces valeurs peuvent être interpolées à l'aide d'une méthode linéaire, qui aligne les données du document SM avec des enregistrements de la vitesse du vent sur des périodes de 10 minutes, mesurés simultanément aux enregistrements de bruit résiduel. Cela permet d'obtenir  $L_w(F, T)$  en bande d'octave ( $F$ ) pour chaque intervalle de 10 minutes ( $T$ ). Un sur-échantillonnage (up-sampling) est appliqué pour convertir ces dernières résultats ( $L_w(F, T)$ ) en bandes de tiers d'octave ( $f$ ) :  $L_w(f, T)$ . Cette conversion sert simplement à ramener des différentes sources de données sur les mêmes bases fréquentielles. L'étape suivante consiste à modifier la fréquence d'échantillonnage du signal de 10 minutes à une seconde, en utilisant le Zéro-padding, en s'adaptant à la fréquence d'échantillonnage souhaitée avec un rapport de sur-échantillonnage de 1/600, afin d'obtenir  $L_w(f, t)$  en bande de tiers d'octave pour seconde ( $t$ ). La Figure 2 montre le diagramme de l'ensemble des étapes pour obtenir des séries temporelles de puissance acoustique du bruit particulier.

Le niveau de pression acoustique (SPL) du bruit particulier ( $L_{BP}$ ) est calculé pour un scénario simplifié où un parc éolien est composé d'une seule éolienne. Ce niveau est déterminé au point de réception en appliquant un filtre de

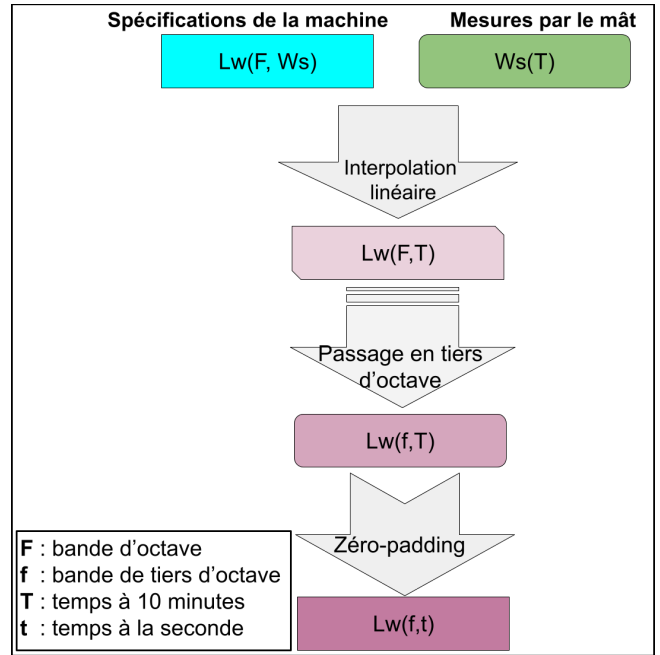


FIGURE 2 – Diagramme des étapes pour construire des séries temporelles avec un échantillonnage de 1 seconde en tiers d'octave de niveau de puissance acoustique du bruit particulier.

propagation acoustique. Ce filtre est basé sur la résolution de l'équation parabolique, modèle de propagation physique prenant en compte différents paramètres tels que l'impédance du sol, le gradient de célérité du son mais aussi la divergence géométrique et l'absorption atmosphérique, et d'un modèle de source éolien étendu [8]. Ces filtres utilisés dans une étude antérieure par Gloaguen et al. [3]. La Figure 3 représente des filtres d'atténuation pour 3 distances distinctes, en se basant sur une hypothèse simplifiée et des conditions de propagation modérément favorables. Ces hypothèses de propagation sont très simplificatrices, car en réalité, les filtres de propagation varient continuellement, notamment en raison des effets de la turbulence atmosphérique.

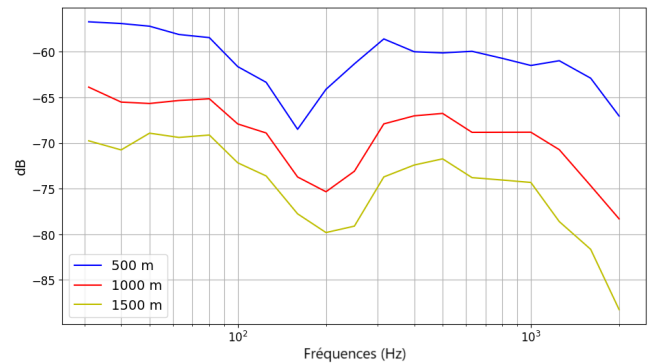


FIGURE 3 – Filtres d'atténuation correspondant à 500 m, 1000 m et 1500 m entre le récepteur (hauteur 1.5 m) et l'éolienne à des conditions de propagation modérément favorables.



### 2.1.3 Bruit ambiant

Enfin, le bruit ambiant  $L_{BA}$  est obtenu en additionnant le bruit particulier synthétisé  $L_{BP}$  et le bruit de résiduel mesuré  $L_{BR}$  :

$$L_{BA} = L_{BR} \oplus L_{BP}, \quad (1)$$

le symbole  $\oplus$  représentant la sommation énergétique. Enfin, nous disposons d'un ensemble de données sous la forme d'une série temporelle de près d'un mois du  $L_{BA}$  et du  $L_{BP}$  pour chaque point de mesure à 500 m, 1000 m et 1500 m.

### 2.2 Vitesse du vent et production électrique

La vitesse du vent est interpolée avec une résolution temporelle d'une seconde et intégrée en tant qu'entrée du modèle RNN. La production électrique de l'éolienne, qui est fonction de ces vitesses de vent, est utilisée conjointement avec les données acoustiques pour l'entraînement du modèle. Cette intégration permet de tenir l'impact de la vitesse du vent sur le bruit généré, et prendre en compte la corrélation entre la production d'énergie et les conditions de vent. La Figure 4 montre un échantillon de la vitesse du vent à la seconde, obtenu à partir de mesures effectuées toutes les 10 minutes, tandis que la Figure 5 représente la production électrique en kW pour chaque vitesse de vent donnée.

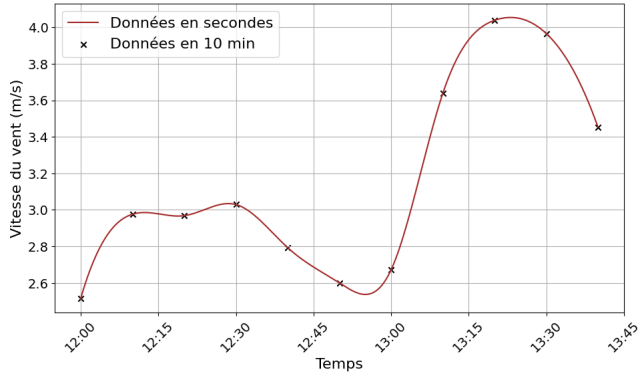


FIGURE 4 – Vitesse du vent interpolée à la seconde pour un échantillon de mesures chaque 10 minutes.

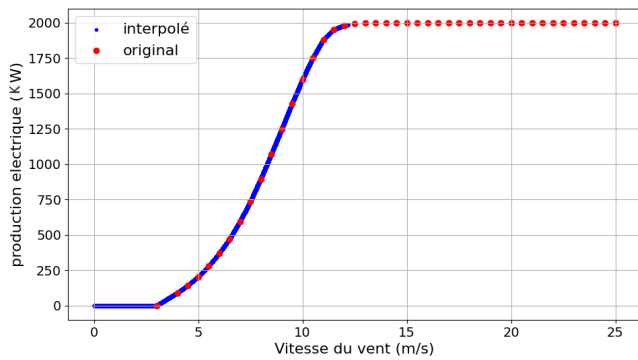


FIGURE 5 – Production électrique en kW de l'éolienne pour chaque vitesse du vent de 0 à 25 m/s.

### 2.3 Prétraitement des données

Les caractéristiques utilisées pour l'apprentissage incluent le bruit ambiant (31 Hz à 2 kHz en dB(A)), la production électrique et la vitesse du vent, dont les échelles varient fortement (0-25 m/s pour le vent, 0-2000 kW pour la production). Le Tableau 1 en présente les propriétés statistiques. Une normalisation est nécessaire pour harmoniser les amplitudes, éviter la dominance de certaines variables et améliorer la convergence, la stabilité et l'efficacité du modèle.

Pour assurer une convergence optimale et éviter l'influence disproportionnée des caractéristiques à grande dynamique, la normalisation Z-score [9] est appliquée. Elle met toutes les caractéristiques à la même échelle, favorisant un apprentissage équilibré et efficace. Les données d'entrée  $Data_i$ , sont standardisées et notées par  $data_{i,std}$  selon l'équation suivante :

$$data_{i,std} = \frac{data_i - \mu}{\sigma}, \quad (2)$$

où  $\mu$  représente la moyenne des données et  $\sigma$  leur écart-type, et  $i$  chaque échantillon à la seconde. Cette normalisation permet d'harmoniser les différentes variables d'entrée en les centrant et réduisant ( $\mu = 0$  et  $\sigma = 1$ ).

Le rapport signal sur bruit ( $SNR$ ) entre le bruit particulier et le bruit résiduel est défini comme suit :

$$SNR = L_{BP} - L_{BR}. \quad (3)$$

Ce rapport segmente les échantillons de test en intervalles de -15 à 15 dB(A). La Figure 6 montre la distribution du  $SNR$  pour l'ensemble de test, représentant 30 % de la base de données. Les échantillons sont répartis équitablement (10 % par distance) selon l'éloignement de l'éolienne. La distribution est centrée autour de 0 dB(A) avec une concentration plus forte près de cette valeur et une légère asymétrie vers les  $SNR$  négatifs. Cette segmentation facilite l'analyse des performances du modèle selon les conditions de  $SNR$ .

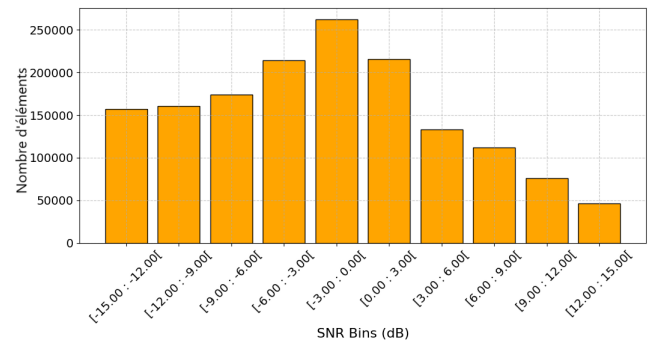


FIGURE 6 – Distribution des données de test selon les  $SNR$ .

### 2.4 Architecture de réseau neuronal

Dans le contexte du traitement des signaux acoustiques, les réseaux neuronaux ont montré des résultats prometteurs pour la séparation des sources, une tâche qui consiste à isoler

TABLEAU 1 – Propriétés statistiques des paramètres d'entrée.

paramètres	Moyen	Écart-type	Maximum	Minimum
acoustiques (dB(A))	16.61	8.88	99.66	-15.41
Vitesse du vent (m/s)	4.88	2.15	12.18	0.00
Production électrique (KW)	306.14	352.37	1985.87	0.00

des composantes sonores individuelles à partir d'un mélange de signaux qui se chevauchent. Des applications telles que la séparation monaurale des sources [4] démontrent les capacités des modèles d'apprentissage profond. Compte tenu de la nature temporelle de l'ensemble de données (le bruit ambiant, vitesse du vent et la production électrique), où chaque ligne représente un pas de temps spécifique  $t$ , et la ligne suivante représente le pas de temps suivant  $t + 1$ , il est intéressant d'adopter des RNNs, tels que Gated Recurrent Unit (GRU) [10].

#### 2.4.1 Cellule GRU

Une cellule GRU comprend deux *portes* : la porte de réinitialisation ( $r_t$ ), qui détermine la quantité de l'état caché précédent à conserver; et la porte de mise à jour ( $u_t$ ), qui contrôle la quantité de l'état caché passé qui doit être reportée au pas de temps suivant.

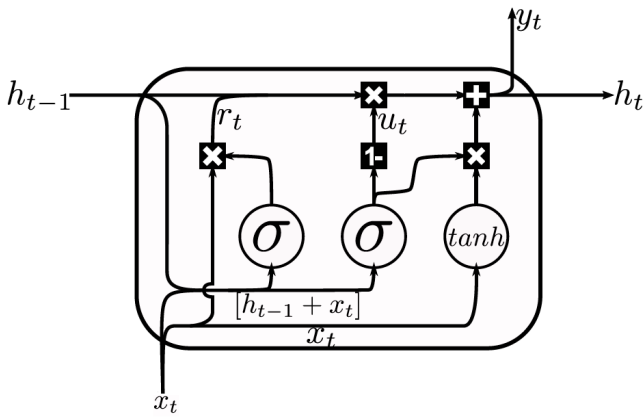


FIGURE 7 – Cellule GRU [11]

La Figure 7 montre une cellule GRU, et l'ensemble des équations, (4), pour un forward-pass dans la cellule :

$$\begin{aligned}
 u_t &= \sigma(W_u x_t + R_u h_{t-1} + b_u) \\
 r_t &= \sigma(W_r x_t + R_r h_{t-1} + b_r) \\
 \tilde{h}_t &= \tanh(W_h x_t + (r_t \odot h_{t-1}) R_h + b_h) \\
 h_t &= (1 - u_t) \odot h_{t-1} + u_t \odot \tilde{h}_t \\
 y_t &= \sigma(W_y h_t + b_y)
 \end{aligned} \quad (4)$$

Les variables  $W$ ,  $R$  et  $b$  représentent les matrices et les vecteurs des paramètres entraînaables (les poids et les biais),  $x_t$  représente l'entrée à l'instant  $t$ . Les opérations utilisées dans les équations du GRU représentent :

- $\sigma(x)$  : La fonction d'activation sigmoïde, définie comme  $\sigma(x) = \frac{1}{1+e^{-x}}$ .
- $\tanh(x)$  : La fonction d'activation tangente hyperbolique, qui produit des valeurs dans l'intervalle  $[-1, 1]$ . Elle est utilisée pour calculer l'état caché candidat  $\tilde{h}_t$ .
- $\odot$  : Le produit élément par élément (produit de Hadamard), utilisé pour contrôler le flux d'information à travers les mécanismes de portes.

#### 2.4.2 Architecture proposée

Dans cette étude, nous proposons une architecture, de type séquence à séquence (seq2seq) illustrée dans la Figure 8, basée sur les cellules GRU, combinant des couches GRU bidirectionnelles et unidirectionnelles afin de prédire le spectrogramme du bruit de l'éolienne à partir d'un spectrogramme contenant le bruit ambiant. L'entrée du modèle est une séquence temporelle de 15 samples, chaque sample contenant 19 caractéristiques spectrales de 31 Hz à 2 KHz (ou avec 2 autres caractéristiques : vitesse du vent et production électrique donc 21 caractéristiques). La première couche du modèle est un GRU bidirectionnel à 32 unités, permettant d'extraire des relations temporelles à la fois passées et futures pour l'apprentissage. Cette couche est suivie d'un GRU unidirectionnel à 64 unités, qui affine la représentation en capturant des dépendances et patterns plus complexes. Ensuite, nous intégrons une couche non linéaire de 32 neurones avec activation ReLU. Enfin, une couche linéaire finale de 19 neurones permet d'estimer le spectrogramme du bruit de la machine sur les 15 pas temporelles en entrée.

#### 2.4.3 Métriques d'évaluation et optimisation

L'erreur absolue moyenne ( $MAE$ ) est choisie comme mesure d'évaluation des performances du modèle RNN pendant l'entraînement. La  $MAE$  mesure l'ampleur moyenne des erreurs entre la valeur prédite du bruit particulier  $\hat{L}_{BP}$  et la valeur réelle  $L_{BP}$  en dB(A) comme suit :

$$MAE = \frac{1}{N} \sum_{i=1}^N |L_{BP_i} - \hat{L}_{BP_i}|. \quad (5)$$

La fonction de perte  $\mathcal{L}$  est définie comme le carré de la différence entre la valeur estimée  $\hat{L}_{BP}$  et le target  $L_{BP}$  :

$$\mathcal{L} = (L_{BP} - \hat{L}_{BP})^2. \quad (6)$$

Ainsi, la fonction de coût pour l'apprentissage des modèles est définie comme la moyenne des pertes sur l'ensemble des

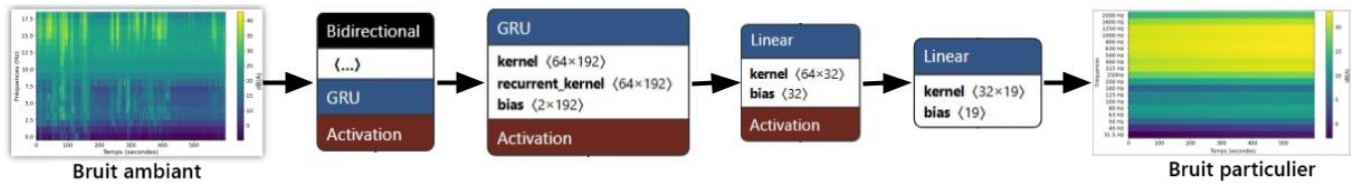


FIGURE 8 – Diagramme de l'architecture réseau utilisée.

$N$  échantillons, où  $i$  désigne un échantillon individuel :

$$MSE = \frac{1}{N} \sum_{i=1}^N (L_{BP_i} - \hat{L}_{BP_i})^2. \quad (7)$$

La dérivée du  $MSE$  par rapport aux prédictions du modèle  $\hat{L}_{BP}$  est linéaire dans le terme d'erreur, ce qui garantit des gradients lisses et cohérents qui facilitent une optimisation efficace selon les algorithmes d'optimisation à base de la descente du gradient. Pour minimiser la fonction de coût, on utilise l'algorithme Adam, qui met à jour les paramètres du modèle  $\theta$  de manière itérative comme suit :

$$\theta = \theta - \eta \cdot \nabla_{\theta} MSE. \quad (8)$$

Ici,  $\nabla_{\theta} MSE$  représente le gradient de la fonction de coût par rapport aux paramètres du modèle  $\theta$ , et  $\eta$  représente le taux d'apprentissage.

### 3 Résultats et discussions

Cette section présente les résultats expérimentaux obtenus sur la base de test, en utilisant deux architectures distinctes. La première architecture repose uniquement sur les données acoustiques (19 caractéristiques), tandis que la seconde inclut, en plus, deux variables supplémentaires : la vitesse du vent et la production électrique, ce qui porte le nombre total de caractéristiques à 21. Les performances pour les 2 modèles d'entrée sont évaluées par le  $MAE$  et  $MSE$  et sont résumées en Tableau 2.

TABLEAU 2 – la fonction de coût  $MSE$  ( $dB(A)^2$ ) et la métrique  $MAE$  ( $dB(A)$ ) pour les deux datasets.

datasets	$MSE$	$MAE$
sans les données environnementales	1.57	0.64
avec les données environnementales	0.78	0.23

Analyser l'impact de l'ajout de ces informations supplémentaires sur la précision du modèle. La Figure 9 illustre la comparaison entre les données réelles et les prédictions avec et sans l'utilisation des données environnementales. On observe que les prédictions prenant en compte les données environnementales (courbe orange) suivent l'évolution de signal réel (courbe bleue), bien que quelques écarts subsistent. En revanche, les prédictions sans données environnementales (courbe rouge) présentent une

plus grande variabilité et s'éloignent davantage des valeurs réelles avec sous estimation.

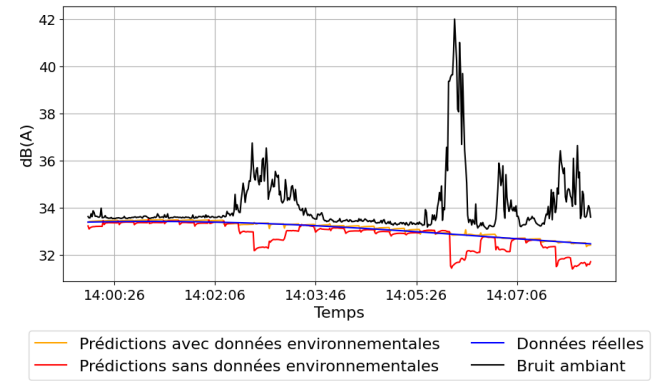


FIGURE 9 – L'évolution temporelle du niveau global de bruit prédit, avec et sans l'intégration des données environnementales, et la valeur cible désirée (données réelles) avec le bruit ambiant.

Le spectre du bruit est estimé à partir de deux ensembles de données et comparé au spectre désiré (label). La Figure 10 présente les spectrogrammes de la différence absolue entre ces spectres,  $|L_{BP} - \hat{L}_{BP}|$ . Avec les données environnementales, cette différence reste inférieure à 0.4  $dB(A)$ , comme l'indiquent les couleurs sombres. En revanche, sans ces données, l'erreur atteint environ 0.6  $dB(A)$ , illustrée par des zones plus claires sur le spectrogramme.

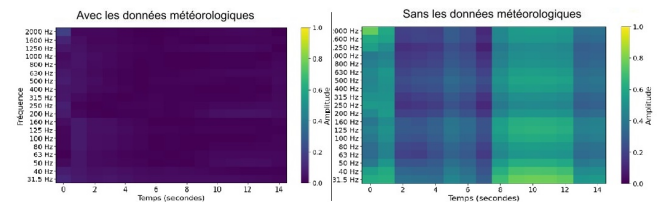


FIGURE 10 – Spectrogrammes de la différence absolue entre le spectrogramme désiré et estimé, représentés pour deux configurations du modèle GRU : avec et sans l'intégration des données environnementales.

L'analyse des erreurs en fonction de la distance démontre l'impact significatif des données environnementales sur les performances du modèle. Comme le montre la Figure 11,  $MAE$  est systématiquement plus faible lorsque les

données environnementales sont intégrées, quelle que soit la distance. À 500 m, l'erreur est réduite de moitié, ce qui confirme une meilleure estimation du spectre sonore. Cet avantage se maintient à 1000 m, où le modèle sans données environnementales présente une erreur nettement plus élevée. À 1500 m, la différence devient encore plus marquée : l'erreur sans données environnementales dépasse le double de celle obtenue avec ces variables, en raison de la faible contribution de  $L_{BP}$ , masquée par  $L_{BR}$ . Cela est également confirmé par la Figure 12, où le MAE est plus élevé pour les SNR négatifs, lorsque  $L_{BP}$  est faible par rapport à  $L_{BR}$  (Equation 3).

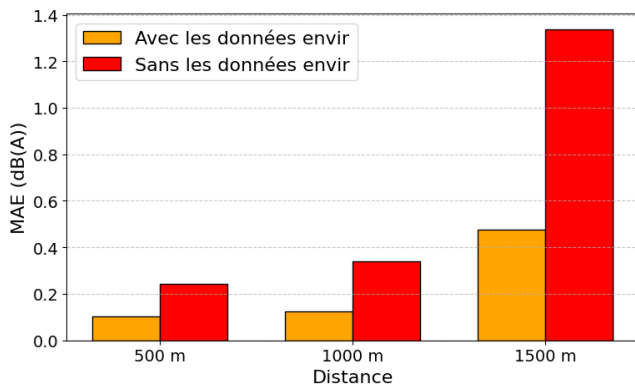


FIGURE 11 – Erreur MAE sur l'ensemble de test pour les données avec et sans variables environnementales, en fonction de trois distances : 500 m, 1000 m et 1500 m.

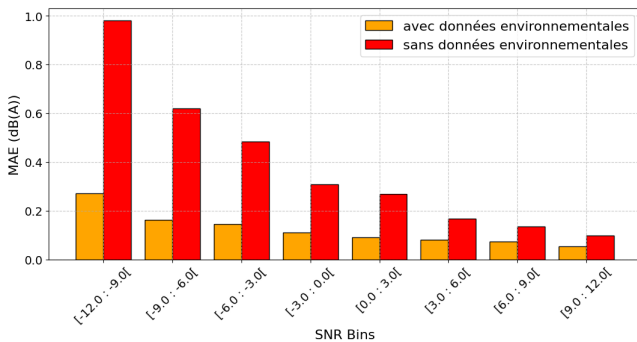


FIGURE 12 – MAE en fonction des intervalles de SNR. La comparaison est effectuée entre les prédictions du modèle avec et sans l'utilisation des données environnementales.

## 4 Conclusion

En conclusion, cette étude met en évidence l'efficacité des réseaux de neurones récurrents, et plus particulièrement des GRU, pour l'estimation du bruit éolien à partir d'une mesure de bruit total, sans avoir à arrêter les éoliennes. L'intégration de données environnementales, telles que la vitesse du vent et la production électrique, permet

d'améliorer considérablement la précision du modèle. L'analyse des résultats par distance et par SNR montre une augmentation progressive de l'erreur avec l'éloignement de la source sonore c'est à dire un bruit particulier faible, donc l'extraction des niveaux est plus difficile dans ce cas. Ces résultats confirment l'importance d'une modélisation prenant en compte les variables environnementales pour une meilleure prévision du bruit éolien. Cependant, une question importante demeure : comment adapter cette approche à des parcs éoliens composés de plusieurs éoliennes ? Faut-il utiliser une production électrique par éolienne ou construire une puissance électrique représentative à partir des éoliennes les plus proches ?

## Références

- [1] Owusu, P. A. Asumadu-Sarkodie, S. A review of renewable energy sources, sustainability issues and climate change mitigation. *Cogent Eng.* 3, 1167990 (2016).
- [2] A. Freiberg, C. Scheffer, and A. Seidler, Health effects of wind turbines on humans in residential settings. *Environmental Research.* **169**, 446–463, (2019).
- [3] Jean-Rémy Gloaguen, David Ecotière, Benoit Gauvreau, Arthur Finez, Arthur Petit, and Colin Bourdat. Automatic estimation of the sound emergence of wind turbine noise with nonnegative matrix factorization. *The Journal of the Acoustical Society of America*, **150**. 3127–3138, (October 2021).
- [4] Po-Sen Huang, Minje Kim. Deep learning for monaural speech separation. *Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1562–1566, (May 2014).
- [5] Chengyun Zhang, Haisong Zhan. Classification of Complicated Urban Forest Acoustic Scenes with Deep Learning Models. *Forests*, **14** :206, (January 2023).
- [6] Arseniy Gorin, and Nickolay Shmyrev. Sound Event Detection System Based on Convolutional Neural Network. *DCASE 2016* (September 2016).
- [7] Anicic, O., Petković, D., Cvetković, S. Evaluation of wind turbine noise by soft computing methodologies : A comparative study. *Renewable and Sustainable Energy Reviews*, **56**, 1122–1128. (2016).
- [8] B. Cotté, “Extended source models for wind turbine noise propagation,” *The Journal of the Acoustical Society of America*. **145**, 1363–1371, (2019).
- [9] J. Han, M. Kamber, and J. Pei, *Data Mining : Concepts and Techniques*, 3rd ed, (Elsevier, 2011).
- [10] Cho K, van Merriënboer B, Gulcehre. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp 1724–1734, (2014).
- [11] Greff, Klaus and Srivastava, and Steunebrink. A comparison of LSTM and GRU networks for learning symbolic sequences, *Lecture Notes in Networks and Systems Intelligent Computing*, p. 771–785, (2023).