



HAL
open science

Estimation par la Factorisation en Matrices Non-négatives de l'émergence sonore du bruit éolien sur des mesures réalisées in situ

Jean-Rémy Gloaguen, David Ecotièrre, Benoit Gauvreau, Arthur Finez, Arthur
Petit, Colin Le Bourdat

► To cite this version:

Jean-Rémy Gloaguen, David Ecotièrre, Benoit Gauvreau, Arthur Finez, Arthur Petit, et al.. Estimation par la Factorisation en Matrices Non-négatives de l'émergence sonore du bruit éolien sur des mesures réalisées in situ. 16ème Congrès Français d'Acoustique, CFA2022, Société Française d'Acoustique; Laboratoire de Mécanique et d'Acoustique, Apr 2022, Marseille, France. hal-03847869

HAL Id: hal-03847869

<https://hal.science/hal-03847869>

Submitted on 10 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



16^{ème} Congrès Français d'Acoustique
11-15 Avril 2022, Marseille

Estimation par la Factorisation en Matrices Non-négatives de l'émergence sonore du bruit éolien sur des mesures réalisées in situ

J-R. Gloaguen ^a, D. Ecotièrre ^b, B. Gauvreau ^c, A. Finez ^a, A. Petit ^a, C. Le Bourdat ^a

^a ENGIE Green, 15 rue Nina Simone, 44000, Nantes

^b UMRAE, CEREMA, Univ. Gustave Eiffel, 11 Rue Jean Mentelin, 67200 Strasbourg

^c UMRAE, Univ. Gustave Eiffel, CEREMA, Campus de Nantes, Route de Bouaye, 44340 Bouguenais



L'impact acoustique des parcs éoliens en France est estimé *in situ* par la réalisation de mesures en cycle de fonctionnement marche/arrêt qui permet de déterminer le bruit ambiant (éoliennes en marche) et le bruit résiduel (éoliennes à l'arrêt) pour ensuite déduire son émergence sonore. Cette procédure engendre des pertes d'exploitation des parcs très importantes ce qui nécessite de limiter la durée des périodes de mesures de l'émergence à 1 ou 2 semaines. Cette durée réduite de mesure se fait au détriment de sa représentativité au regard d'une année complète de conditions météorologiques. Récemment, une méthode de séparation de sources sonores, la Factorisation en Matrices Non-négatives (NMF) a montré de l'intérêt dans sa capacité à estimer le bruit éolien parmi des scènes sonores synthétisés de bruit ambiant et à déduire l'émergence sonore. Cette approche ouvre alors la voie vers un monitoring continu des parcs sans arrêter les machines. Nous proposons ici une première application de cette méthode sur l'ensemble des points de mesures d'une campagne de réception d'un parc éolien afin de comparer les estimations de l'émergence par la NMF par rapport à celles déterminées par le processus normatif. Les résultats obtenus révèlent des performances variables de l'estimation de l'émergence par la NMF selon les points de mesures. Certaines estimations se révèlent être bien concordantes avec celles obtenues selon le processus normatif. En revanche d'autres points présentent des résultats plus dispersés, qui pourrait relever de l'influence de l'environnement sonore du point de mesure et de la vitesse de vent sur les performances de l'outil.

1 Introduction

En France, pour toute éolienne *on-shore* installée, la réglementation en vigueur impose, parmi plusieurs critères (niveau de bruit maximal admissible en bordure de périmètre, présence de tonalité marquée), que le bruit émis par un parc éolien n'émerge pas de plus de 5 dB(A) par rapport au bruit résiduel sur le site le jour (de 7 h à 22 h) et de 3 dB(A) la nuit (de 22 h à 7 h) pour un bruit ambiant supérieur à 35 dB(A). Ces critères ont pour objectif de s'assurer que l'implantation d'un parc éolien ne vient pas détériorer la qualité de vie des habitants vivant aux alentours par le bruit qui en sera émis. Cette émergence se définit par la différence du niveau ambiant (niveau sonore lorsque l'éolienne fonctionne) avec le niveau résiduel (lorsque l'éolienne est à l'arrêt) : $E = L_{A,50,ambiant} - L_{A,50,résiduel}$.

Pour certifier que cette émergence n'est pas dépassée, un plan de bridage peut être établi afin de limiter, pour certaines vitesses et direction de vents la vitesse de rotation des pâles des éoliennes et donc le bruit généré. Ce plan de bridage est défini avant l'installation du parc éolien à partir de mesures de bruit résiduel faites au préalable. Cependant, le bruit résiduel est fluctuant au cours du temps selon les conditions climatiques (temps sec, pluie, direction du vent...), les périodes annuelles (été, hiver...) ou les infrastructures environnantes (construction de route, d'habitation...). En conséquence, le plan de bridage peut ne plus être adapté au cours de son fonctionnement. Cette inadéquation engendre 2 cas distincts :

- Le niveau sonore de l'éolienne est trop important et dépasse les limites d'émergences réglementaires. Dans ce cas, il faut limiter le fonctionnement de l'éolienne afin de respecter la réglementation.
- Le bruit résiduel étant plus élevé, le plan de bridage est trop important. Ce dernier cas signifie alors que la production électrique de l'éolienne est trop limitée et pourrait donc être plus importante.

Il est possible de modifier ce plan de bridage en faisant régulièrement de nouvelles mesures de bruit résiduel. Mais ces campagnes de mesures nécessitent de réaliser des cycles marche/arrêt des parcs qui engendrent de nouvelles pertes en production pour leurs exploitants. La durée d'une telle

campagne de mesures n'excède donc pas 1 à 2 semaines, ce qui est peu au regard des variations sonores qui peuvent intervenir sur une année complète et ne permet donc pas de représenter correctement l'évolution du bruit résiduel.

L'objectif de cette étude est de développer un outil d'évaluation automatique de l'émergence sonore du bruit des éoliennes, en temps réel, et sans avoir à mettre à l'arrêt les aérogénérateurs, à partir de mesures acoustiques. Avec un tel outil, il deviendrait possible, à partir de l'émergence sonore estimée, d'adapter dynamiquement le plan de bridage des éoliennes plus facilement et plus régulièrement sans arrêter les parcs éoliens.

Pour ce faire, la méthode de la factorisation en matrices non-négatives (abrégée NMF selon le terme anglais *Non-negative Matrix Factorization*) a été retenue [1] en tant que méthode de séparation de sources sonores, via un apprentissage supervisé. Elle présente notamment l'avantage d'être utilisable pour des mesures réalisées par un seul microphone. Développée il y a 20 ans, elle possède également une théorie stable et une riche littérature sur son fonctionnement avec, en outre, de nombreuses applications pour des sons environnementaux [2, 3].

Une première application pour le bruit éolien a récemment été proposée sur des corpus de bruit ambiant simulées [4]. Ces scènes mélangent une composante de bruit éolien et bruit résiduel exprimées en bande de tiers d'octave avec un pas temporel de 1 seconde, qui est un format compatible avec les mesures réalisés sur le terrain. L'intérêt de ce corpus est de permettre un contrôle total sur la composition des scènes, de l'origine du bruit éolien et résiduel ainsi que de connaître la valeur exacte de l'émergence, ce qui n'est pas possible en pratique avec des mesures *in situ*. Les résultats ont permis de définir un paramétrage optimal de la NMF avec un apprentissage supervisé.

Cet article prolonge ce travail, mais en se basant sur une NMF à l'apprentissage semi-supervisée [5] et en présentant les résultats obtenus par cette méthode sur un jeu de données mesuré lors d'une réception acoustique sur un parc éolien. Dans la partie 2, la NMF semi-supervisée est décrite, puis en partie 3, les jeux de données requis pour cette étude sont présentés. Enfin, dans la quatrième partie, les résultats

obtenus sont résumés et discutés permettant de voir l'intérêt de la méthode mais aussi ces limites.

2 Factorisation Non-négative semi-supervisée

La Factorisation en Matrices Non-négatives (ou NMF) est une méthode d'approximation linéaire qui vise à estimer un signal exprimé sous forme d'une matrice $V_{F \times N}$ par le produit de 2 matrices $W_{F \times K}$ et $H_{K \times N}$ tel que

$$V \approx \tilde{V} = WH. \quad (1)$$

Dans le domaine de l'audio, V s'exprime sous la forme d'un spectrogramme en amplitude. W est appelé *dictionnaire* et est composé de K spectres audio et H est la *matrice d'activation* qui traduit l'évolution en amplitude dans le temps de chacun de ces spectres (voir Figure 1). Seule des données *non-négative* sont considérés ici, ce qui ne permet de considérer que des combinaisons additives des composantes.

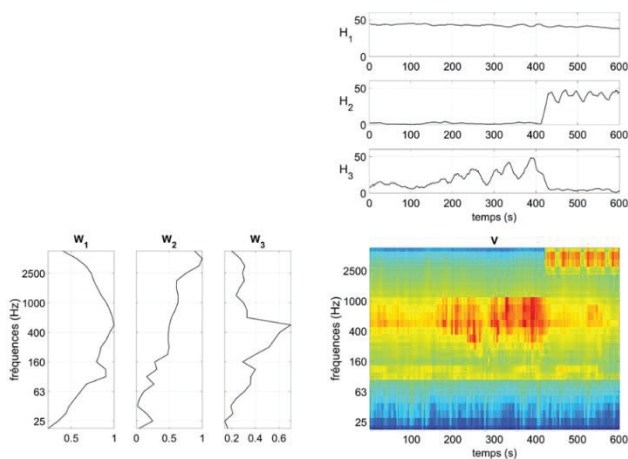


Figure 1 : illustration du principe de la NMF avec $K = 3$.

L'objectif de la méthode est alors de minimiser une fonction coût, basée sur une β -divergence, afin de réaliser la reconstruction du spectrogramme V :

$$\min(D_\beta(V|\tilde{V})) = \min\left(\sum_{F,N} d_\beta(V_{fn}|\tilde{V}_{fn})\right) \quad (2)$$

Parmi les différents opérateurs possibles, la distance euclidienne est celle retenue dans cette étude ($\beta = 2$). Un processus est mis en place venant mettre à jour les matrices W et/ou H itérativement grâce à des algorithmes définis dans la littérature et qui dépendent du choix de la valeur β [6]. Dans cette étude, le bruit éolien est la source sonore commune dans les mesures. Le dictionnaire est donc composé de spectres de bruit éolien fixe. Seule H est l'inconnue à déterminer et donc la seule matrice à être mise à jour.

Il peut être réducteur de ne considérer qu'un dictionnaire éolien dans la NMF, car cela serait les seules informations

disponibles pour approximer le spectrogramme V , ce qui augmente le risque de confusion et d'erreur. Ainsi, plutôt que de définir un second dictionnaire de bruit résiduel fixe difficile à définir, la NMF semi-supervisée (NMF SS) est considérée [7]. Elle consiste à ajouter au dictionnaire W et à la matrice d'activation H d'une seconde partie, respectivement $Y_{F \times J}$ et $Z_{J \times N}$. Ces derniers sont alors définis directement sur chaque spectrogramme V testé, permettant à la NMF de s'adapter plus librement. Ces éléments, de dimension J , sont libres de s'adapter à chaque spectrogramme V testé. La dimension de J est choisie telle que $K \geq J$ afin de maintenir l'utilisation de W dans l'approximation. Si J est trop grand, la minimisation de la fonction coût (équation 2) risque d'être essentiellement réalisée par le produit $[YZ]$ au détriment de la partie supervisée $[WH]$. L'équation 1 devient alors :

$$V \approx \tilde{V} = \sqrt{(WH)^2 + (YZ)^2}. \quad (3)$$

Le choix d'une somme énergétique entre les composantes WH et YZ provient de l'unité des données exprimé en Pa. La forme des algorithmes devient alors :

$$H_{i+1} \leftarrow H_i \frac{W^T [V \otimes [WH]_i]}{W^T [V \otimes \tilde{V}_i]}, \quad (4)$$

$$Z_{i+1} \leftarrow Z_i \frac{Y_i^T [V \otimes [YZ]_i]}{Y_i^T [V \otimes \tilde{V}_i]}, \quad (5)$$

$$Y_{i+1} \leftarrow Y_i \frac{[V \otimes [YZ]_i] Z_i^T}{[V \otimes \tilde{V}_i] Z_i^T}. \quad (6)$$

En parallèle du choix de l'apprentissage, il est possible de contraindre le fonctionnement de la NMF afin d'orienter les mises à jour des matrices vers une forme souhaitée. Ici, une contrainte de régularité temporelle sur la matrice H est choisie. Cette contrainte a pour objectif de limiter les soudaines variations au sein des activations et ainsi d'assurer une forme plus continue et régulière de celles-ci. Elle semble donc adaptée au bruit éolien puisque celui-ci varie lentement dans le temps. Cette contrainte est considérée dans la fonction coût à travers l'ajout d'un terme de régularisation :

$$\min(D_\beta(V|\tilde{V}) + \alpha_{sm} \sum (h_n - h_{n-1})^2). \quad (7)$$

Celui-ci est pondéré par le coefficient α_{sm} . Plus cette valeur est élevée plus la contrainte sur H devient importante. La mise à jour de la matrice H (équation 4) est alors modifiée afin d'intégrer ce terme de pénalité [8].

3 Données expérimentales

3.1 Données d'apprentissage du dictionnaire

Dans la NMF SS, parmi les 4 matrices permettant l'approximation du spectrogramme V , seule la matrice W est une donnée fixée. W contient alors K spectres éoliens permettant de décrire l'allure spectrale de la source d'intérêt. Pour cette étude, ce sont les spectres de puissance théoriques à chaque vitesse de vent (de 3 m/s jusqu'à 15 m/s en bandes

de tiers d'octave de 20 Hz à 10 kHz) calculés par les constructeurs qui sont utilisés. Dans [5], la valeur optimale de K a été estimée à 5 éléments qui est repris dans cette étude. Pour réduire le nombre de spectres disponibles à cette valeur, un algorithme K-means est utilisé. Dans le cas de la NMF SS, il a semblé intéressant d'étendre le nombre de bandes de fréquences afin de mieux dissocier celles où se concentrent l'énergie du bruit éolien (entre 100 Hz et 1000 Hz) et celles du bruit résiduel qui inclût de nombreuses sources sonores pouvant dépasser la bande fréquentielle de 1000 Hz.

Enfin, puisque le bruit éolien se propage à de longues distances, le dictionnaire éolien est filtré via des filtres de propagation afin de mieux faire correspondre l'allure des spectres dans W au bruit éolien propagé dans les mesures. Ces filtres sont définis avec un paramétrage médian afin d'obtenir des conditions favorables à la propagation acoustique du son et représentatives des campagnes. L'ensemble des détails pour la construction de ces filtres est disponible dans [4]. Pour chaque point de mesures, la distance de propagation est alors définie par la distance d entre celui-ci et l'éolienne la plus proche, arrondi à 100 m près. L'ensemble de ces étapes sont résumées en Figure 2. Enfin l'ensemble des données est ensuite calibré tel que $\|W_k\| = 1$.

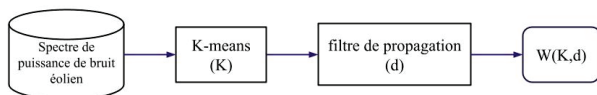


Figure 2 : étapes pour la construction du dictionnaire.

3.2 Mesures de bruit éolien

Le parc considéré pour l'étude est un parc appartement à la société ENGIE Green composé de 6 éoliennes. Une campagne de mesures dans le cadre d'une réception acoustique y a été réalisée pendant 1 mois où une succession de créneaux arrêt/marche de 2 h a été réalisée. La Figure 3 résume les données météorologiques, de puissance et de niveaux sonores mesurés en 1 point de réception de plusieurs heures de cette campagne.

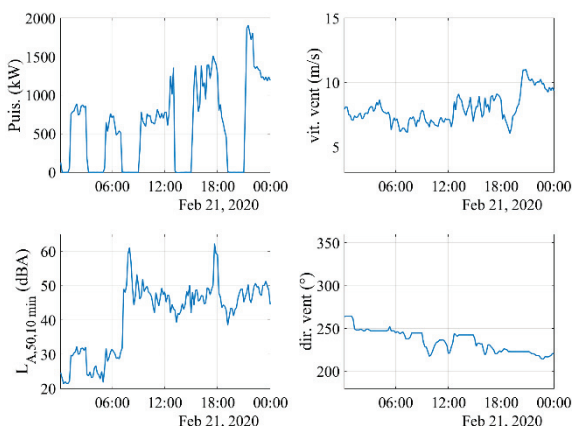


Figure 3 : extrait sur 24h de la puissance active, vitesse et direction de vent moyennes et des niveaux médians mesurés à un point de mesures.

8 points de mesures situés aux abords du parc chez des riverains sont considérés ici dans l'étude. Ces mesures respectent le protocole de mesures du projet de norme PrNF-31-114. À partir des mesures faites lors des périodes de fonctionnement du parc ; les mesures acoustiques sont découpées en échantillons de 10 minutes. Chaque échantillon est décrit avec un pas temporel de 1 seconde ($N = 600$), en bandes de tiers d'octave allant de 20 Hz à 10 kHz ($F = 28$) constituant alors le spectrogramme V sur lequel est appliqué la méthode NMF. Chaque échantillon est associé à des données météorologiques mesurées sur le site, permettant d'y associer une vitesse et une direction de vent.

4 Résultats et discussions

4.1 Résumé des calculs

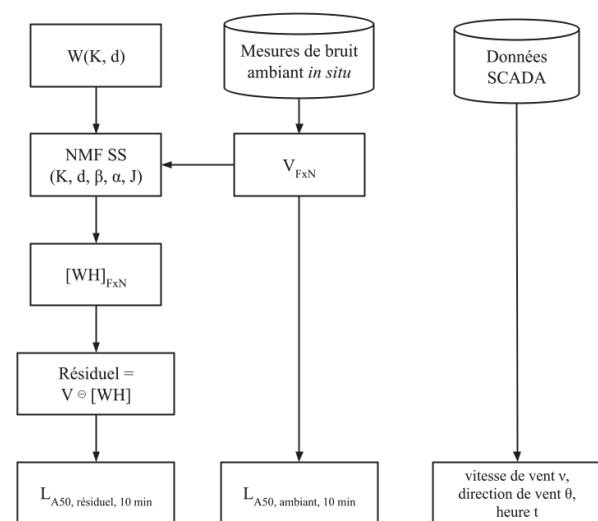


Figure 4 : schéma en blocs du procédé des calculs.

L'objectif est donc d'appliquer la NMF SS sur l'ensemble des périodes de bruit ambiant de la campagne de mesures sur les 8 points de réception retenu. Le paramétrage de la NMF SS est défini à partir d'une étude menée au préalable sur le corpus de scènes simulées dans [5]. L'étude paramétrique a défini des résultats optimaux dans le cas d'un NMF SS pour $\beta = 2$, un nombre réduit de spectres pour les dictionnaires W et J , respectivement $K = 5$ et $J = 5$. La pondération du terme de la régularisation temporelle est définie à $\alpha_{sm} = 0.5$. Enfin, 100 itérations sont réalisées pour converger suffisamment à une solution viable. Les données acoustiques sont d'abord filtrées par un filtre de pondération A puis convertis en Pa pour être compatible avec le formalisme de la NMF. Ce paramétrage de la NMF SS est ensuite appliqué sur des scènes de 10 minutes de bruit ambiant. Puisque c'est la composante éolienne qui est l'élément commun à l'ensemble des mesures testées, c'est le produit $[WH]$ qui est utilisé pour déterminer dans un premier temps la composante éolienne. Puis, celle-ci est soustraite au bruit ambiant pour en estimer la composante du bruit résiduel et son niveau sonore équivalent 1 seconde $\tilde{L}_{A,eq,résiduel,1s}$. De ces 600 valeurs, le niveau médian du bruit résiduel $\tilde{L}_{A,50,résiduel,10\ min}$ de l'échantillon est déduit et est associé aux valeurs météorologiques obtenues via les SCADA des

machines (vitesse de vent v , direction de vent θ et heure t). La Figure 4 résume l'ensemble des étapes de calculs.

4.2 Evolution temporelle

Dans un premier temps, l'utilisation de la NMF SS permet d'observer conjointement les évolutions estimées du bruit éolien, résiduel et ambiant 10 minutes, qui n'est pas possible en pratique avec les méthodes actuelles. La Figure 5 résume pour illustrer cela l'évolution temporelle de ces niveaux sonores médians 10 minutes sur 24 h.

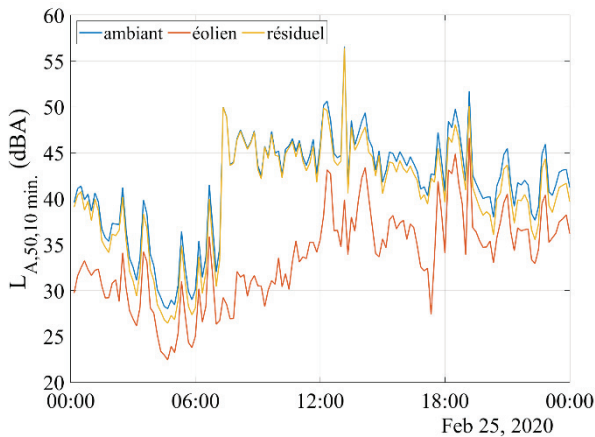


Figure 5 : évolution des niveaux ambiant et des niveaux résiduels et éoliens estimés par la NMF.

Cette visualisation permet de voir conjointement comment évolue le niveau de bruit éolien et résiduel au cours du temps. La période entre 6h00 et 12h00, dans cet exemple, démontre d'une faible émergence du parc en raison notamment des activités humaines dans cette période. Certaines évolutions du bruit ambiant, par leur brièveté, peuvent légitimement être associées à des variations du bruit résiduel et sont correctement identifiées par la NMF. On constate tout de même des périodes où les niveaux éoliens estimés semblent erronés (entre 17h30 et 19h30 par exemple) au regard de son évolution temporelle. Ces passages peuvent avoir plusieurs causes, notamment la confusion dans [WH] qui intègre du bruit résiduel, comme le passage d'un avion, dont la signature sonore partage des similarités avec le bruit éolien. Ils subsistent également des problèmes de convergence où des niveaux sonores éoliens sont parfois supérieurs au niveau de bruit ambiant, ce qui n'est évidemment pas possible en pratique, ce qui donne des valeurs aberrantes du bruit résiduel.

Les émergences sur une échelle de 10 minutes peuvent être alors calculées permettant là encore de suivre temporellement son évolution au cours d'une journée (Figure 6). Ces informations pourraient être utiles en pratique afin notamment de brider dynamiquement le parc éolien selon les émergences constatées.

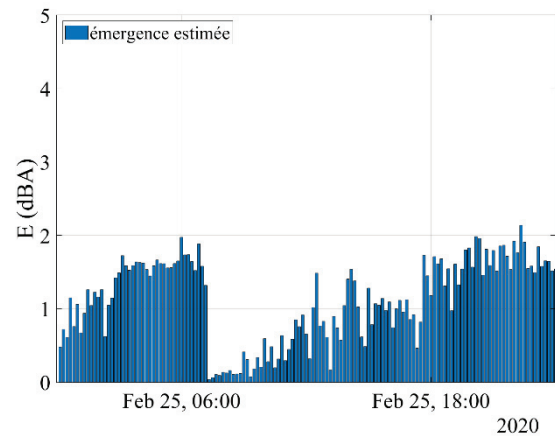


Figure 6 : évolution des émergences 10 minutes estimées par la NMF.

4.3 Comparaison des émergences

Il est possible d'exprimer ces niveaux sonores en suivant le protocole prNF 31-114. Pour cela, l'ensemble des niveaux sonores médians ambiants et ceux estimés du résiduel 10 minutes par la NMF SS sont classés par classe de vent entier V_i ($v \in [V_i - 0.5; V_i + 0.5[$ m/s), par secteur de vent (ici, deux secteurs de vents θ_i , $[135^\circ ; 315^\circ[$ et $[315^\circ ; 135^\circ[$) ainsi que par période de la journée T_i (Jour, $[7h ; 22h[$ et Nuit, $[22h ; 7h[$, voir partie 1). L'apport de la NMF SS est qu'il devient possible de déterminer un niveau médian éolien pour ces classes, comme en Figure 7.

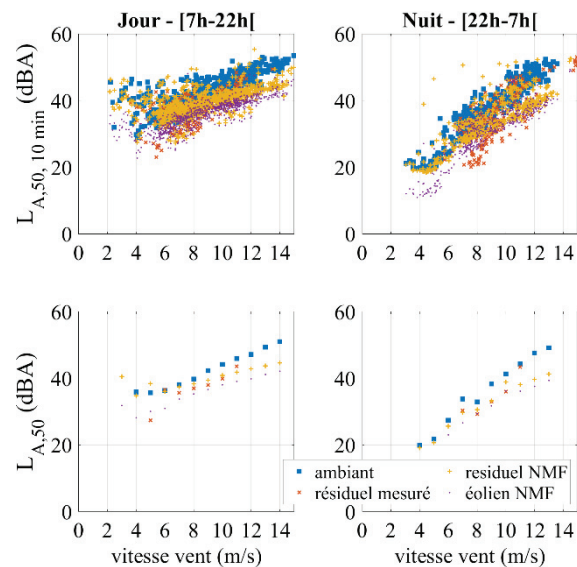


Figure 7 : distribution des niveaux sonores $L_{A,50}$ 10 minutes et médians sur un point de mesure pour au secteur de vent $[135^\circ ; 315^\circ[$.

Les niveaux médians ambiants, éoliens et résiduels par classe de vitesse de vent V , secteur de vent θ et période T sont alors estimés, respectivement $L_{A,50,ambiant}(V_i, \theta_i, T_i)$, $\tilde{L}_{A,50,éolien}(V_i, \theta_i, T_i)$, et $\tilde{L}_{A,50,résiduel}(V_i, \theta_i, T_i)$. Il faut néanmoins 10 valeurs dans cette classe spécifique pour définir une valeur médiane.

On constate que l'évolution du niveau médian du bruit éolien tend à augmenter avec la vitesse du vent. Or en pratique, à partir d'une vitesse de vent élevé, les éoliennes atteignent leur fonctionnement nominal où la vitesse de rotation est constante afin de maintenir la production à son maximum. L'émission sonore de l'éolienne est donc supposément stable, ce qui n'est pas le cas ici. La raison est que le bruit résiduel se charge en basses fréquences à mesure que la vitesse du vent augmente, ce qui tend à générer de la confusion spectrale dans la NMF où du bruit résiduel est associé à un bruit éolien.

À partir de ces niveaux médians, l'émergence du bruit éolien est finalement estimée :

$$E_{\text{NMF}}(V_i, \theta_i, T_i) = L_{A,50,\text{ambient}}(V_i, \theta_i, T_i) - \tilde{L}_{A,50,\text{résiduel}}(V_i, \theta_i, T_i), \quad (8)$$

Cette émergence est alors comparée à l'émergence définie selon le protocole normatif, $E_{\text{norm.}}$, c'est-à-dire en suivant la même procédure, mais en se basant sur les bruits résiduels mesurés lors des arrêts des machines et non durant leur fonctionnement,

$$E_{\text{norm.}}(V_i, \theta_i, T_i) = L_{A,50,\text{ambient}}(V_i, \theta_i, T_i) - L_{A,50,\text{résiduel}}(V_i, \theta_i, T_i), \quad (9)$$

avec $L_{A,50,\text{résiduel}}$ le niveau sonore médian du bruit résiduel mesuré lors des arrêts des machines. Les niveaux sonores 10 minutes de ce résiduel et les niveaux médians sont résumés en Figure 7. La Figure 8 compare les valeurs d'émergences en distinguant la période T et sur le seul secteur de vent $[135^\circ ; 315^\circ[$ car étant le secteur de vent ayant été le plus fréquent et permettant alors d'avoir suffisamment d'échantillons pour être analysé ici. On ajoute également une courbe référence $E_{\text{norm.}} = E_{\text{NMF}}$, qui correspondrait à une concordance parfaite entre les deux méthodes. Cette concordance n'est pas en pratique possible car les niveaux de bruits résiduels sur lesquels se basent ces deux approches ne sont pas les mêmes. Mais elle offre un point de repère pour situer la méthode NMF par rapport au processus normatif. La zone grisée autour de cette courbe de référence correspond à la *Mean Absolute Error* de 1.9 dBA obtenue par la NMF SS dans [5] sur un corpus de scènes de bruit ambiant simulées. Elle permet de vérifier si les résultats obtenus ici restent cohérents avec ceux de la première étude.

Une majorité des points représentés est bien présente dans la zone grisée, ce qui signifie que ces points présentent un écart par rapport à $E_{\text{norm.}}$ inférieure à 1,9 dBA, ce qui conforte les résultats obtenus dans [5]. On constate tout de même plusieurs points en dehors de cette zone que cela soit le jour ou la nuit. Selon la période, ces points ne sont pas les mêmes. Par exemple, les émergences estimées au point 5 en journée sont systématiquement en dehors de cette zone d'erreur. La nuit, elles présentent une meilleure concordance avec $E_{\text{norm.}}$. À l'inverse, les estimations au point 8 présentent un comportement opposé. Sur ce parc, il ne semble donc pas y avoir de dépendance spatiale spécifique qui justifierait nécessairement que les estimations de l'émergence par la NMF sont mieux similaires à $E_{\text{norm.}}$ ou non. On peut observer toutefois que la concordance est meilleure la nuit

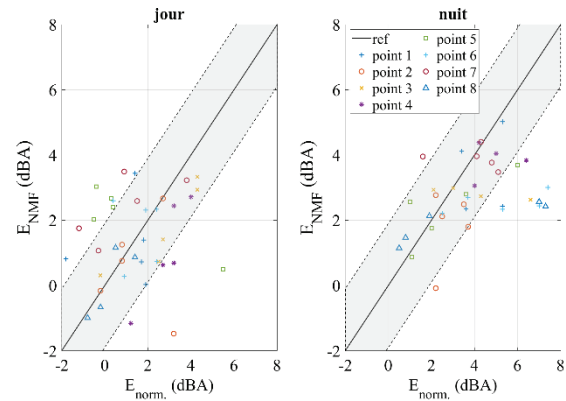


Figure 8 : comparaison des émergences estimées par la NMF et par la procédure normative pour le secteur de direction de vent $[135^\circ ; 315^\circ[$.

probablement parce que les sources de bruit résiduels y sont plus faibles, ce qui permet de mieux retrouver le bruit éolien dans le bruit ambiant. C'est un résultat intéressant car ce sont ces périodes là où les dépassements d'émergences sont le plus susceptibles d'être observés en raison de la diminution des niveaux de bruit résiduel. En journée, les résultats sont plus disparates avec des estimations d'émergences plus ou moins élevées par rapport à $E_{\text{norm.}}$. Si les valeurs où E_{NMF} sous-estime $E_{\text{norm.}}$ ne sont pas encore expliquées, celles qui dépassent le seuil d'erreur de 1.9 dBA se justifient par les hautes vitesses de vent où on a constaté que le bruit éolien continue à croître alors qu'il doit atteindre un niveau stable. Cette surestimation entraîne alors également celle des émergences. Il semble donc que l'utilisation de la NMF SS dans les hautes vitesses (> 10 m/s) en journée doit être traitée avec attention. Toutefois, il faut compléter cette conclusion en raison du nombre réduit d'échantillons mesurés à ces vitesses lors des périodes d'arrêts qui n'a donné qu'un faible nombre d'échantillons à observer.

5 Conclusion

La factorisation en matrices non-négatives semi-supervisée (NMF SS) a été appliquée sur un ensemble de mesures de bruit ambiant réalisé autour d'un parc éolien afin d'en estimer son émergence sonore. Cette application présente l'intérêt de pouvoir mieux suivre les évolutions des émergences sonores et donc ouvre la voie à un bridage dynamique du parc permettant d'optimiser la production électrique tout en vérifiant plus régulièrement la conformité du parc.

La comparaison des émergences estimées par la NMF et celle par la méthode normative présentent des résultats cohérents au regard des performances de la NMF SS observées sur les corpus de scènes simulées dans la précédente étude. Il subsiste tout de même certaines estimations qui divergent fortement. Sur celles-ci, leur position ou encore la période de journée n'ont pas présenté de lien apparent. Cela indiquerait que ces disparités ne sont pas liées à une dépendance spatio-temporelle du point de mesures mais bien au fonctionnement de la NMF avec notamment de la confusion entre les sources sonores dues à des similarités spectrales.

Cette étude doit maintenant être complétée en l'appliquant sur un plus grand nombre de parcs afin d'en étoffer la représentation statistique. En parallèle, il reste des pistes à explorer. La composition du dictionnaire est notamment un élément clé, puisqu'elle définit l'allure spectrale de la source d'intérêt. Plutôt que d'utiliser des niveaux de puissances sonores modélisés par les turbiniers, il peut être envisagé de mesurer les spectres d'émission sonore des parcs directement *in situ* en suivant le protocole de mesures IEC 61400-11. Cette méthode permettrait de définir un dictionnaire propre au parc éolien et pourrait être remis à jour au cours de la durée de vie du parc. Enfin, la formation des filtres peut également être améliorée en affinant les modèles de propagation aux sites éoliens via des modèles plus complets (équation parabolique) ou plus couramment utilisés dans les bureaux d'études comme Harmonoise.

6 Bibliographie

- [1] D. D. Lee et H. S. Seung, «Learning the parts of objects by non-negative matrix factorization,» *Nature*, vol. 401, n° 16755, pp. 788-791, 1999.
- [2] T. Heitola, A. Mesaros, V. T. et E. A., «Sound event detection in multisource environments using source separation,» chez *International Workshop on Machine Listening in Multisource Environments, CHiME*, Florence, Italie, 2011.
- [3] J.-R. Gloaguen, A. Can, M. Lagrange et J.-F. Petiot, «Road traffic sound level estimation from realistic urban sound mixtures by Non-negative Matrix Factorization,» *Applied Acoustics*, vol. 143, pp. 229-238, 2019.
- [4] J.-R. Gloaguen, D. Ecotière, B. Gauvreau, A. Finez, A. Petit et C. Le Bourdat, «Automatic estimation of the sound emergence of wind turbine noise with nonnegative matrix factorization,» *The Journal of the Acoustical Society of America*, vol. 150, n° 14, pp. 3127-3138, 2021.
- [5] J.-R. Gloaguen, D. Ecotière, B. Gauvreau, A. Finez, A. Petit et C. Le Boudart, «Estimation of the sound emergence of wind turbines by semi-supervised learning technique,» chez *INCEEurope*, Dublin, 2021.
- [6] C. Févotte et J. Idier, «Algorithms for nonnegative matrix factorization with the β -divergence,» *Neural Computation*, vol. 23, n° 19, pp. 2421-2456, 2011.
- [7] D. Kitamura, H. Saruwatari, K. Yagi, K. Shikano, Y. Takahashi et K. Kondo, «Music Signal Separation Based on Supervised Nonnegative Matrix Factorization with Orthogonality and Maximum-Divergence Penalties,» *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, vol. E97.A, n° 15, pp. 113-118, 2014.
- [8] C. Févotte, E. Vincent et A. Ozerov, *Single-channel audio source separation with NMF: divergences, constraints and algorithms*, Springer, 2018, pp. 1-24.