Automatic estimation of the sound emergence of wind turbines using non-negative matrix factorization: a preliminary study

Jean-Rémy Gloaguen, Benoit Gauvreau, David Ecotiere, Arthur Petit, Arthur Finez, Colin Lebourdat

AUTOMATIC ESTIMATION OF THE SOUND EMERGENCE OF WIND TURBINES USING NON-NEGATIVE MATRIX FACTORIZATION: A PRELIMINARY STUDY

Jean-Rémy Gloaguen[1] **David Ecotière**[1] **Benoit Gauvreau**[2]
Arthur Finez[3] **Arthur Petit**[3] **Colin Lebourdat**[3]
[1] UMRAE, Cerema, Univ Gustave Eiffel, Ifsttar, F-67035 Strasbourg, France
[2] UMRAE, Univ Gustave Eiffel, Ifsttar, Cerema, F-44344 Bouguenais, France
[3] ENGIE Green, F-44000 Nantes, France
jean-remy.gloaguen@univ-eiffel.fr

ABSTRACT

The acoustic impact of French wind farms is currently estimated by measuring their sound emergence. These measures require the implementation of *on/off* cycles of the wind farm in order to determine the ambient noise (wind turbines in operation) and the residual noise (stopped wind turbines). These procedures generate very high costs for operators, which strongly limit the duration of emergence measurement periods (2 or 3 weeks). This reduced duration, compared to a full year of different weather conditions, is to the detriment of the representativeness of the estimation of sound emergence. In order to remedy this limitation, we propose to estimate the noise emergence of wind turbines in real time, continuously and without stopping the machines, using a source separation method based on a machine learning technique: Non-negative Matrix Factorization. This technique is tested on a corpus of simulated sound scenes that allows a total control of their composition and especially their emergence. A numerical experiment is conducted to determine, among the various influential parameters of this method, the optimal form that achieves the best estimates of sound emergence over the entire sound corpus. Initial results indicate that this approach generates average estimation errors similar to current methods but depends on the emergence of wind turbine noise. This method makes it possible, under validation by more complex corpora, to estimate the noise emergence of wind farms continuously without having to shut them down which is not the case in the current method.
Keyword: wind turbine noise, sound emergence, *in situ* measurements, non-negative matrix factorization

1. INTRODUCTION

In France, every wind turbine farm has to respect the current regulation which imposes (among several criteria) that the wind turbine sound emergence from the background noise does not exceed 5 dB(A) between 7 a.m. and 10 p.m. and 3 dB(A) between 10 p.m. and 7 a.m., with an ambient sound level superior to 35 dB(A). These criteria aim to preserve the life quality of the inhabitants living near the wind farms. The sound emergence E is defined as the difference between the ambient sound level $L_{A,50,ambient}$ (when the wind turbines work, cycle *on*) and the residual sound level $L_{A,50,residual}$ (when the wind turbines are stopped, cycle *off*):

$$E = L_{A,50,ambient} - L_{A,50,residual}. \tag{1}$$

Since the noise generated by a wind turbine fluctuates slowly over time, the emergence calculation is based on the $L_{A,50}$ statistical indicator, which corresponds to the median sound level based on the 1 second equivalent sound level. This indicator seems adequate in order to limit the impact of the strong noise contributions from residual events, which is not possible with the equivalent sound level. To assure that these regulation levels are respected, a curtailment plan can be implemented in some circumstances (specific meteorological conditions for example) in order to limit the wind turbine and then the generated noise. This plan is usually defined before their installation. Nevertheless, the residual sound level fluctuates over time according to the meteorological conditions (dry weather, rain, wind direction...), the season (winter, summer...) or the surrounding infrastructures (highway, industry...). In consequence, the curtailment plan can sometimes be inadequate: to the detriment of local inhabitants (i.e. sound emergence overcomes the regulation levels) or to the detriment of industrial operator (i.e. the wind farm is too limited in its operation which decreases the electric production). It is possible to adapt this curtailment plan by regularly taking a few measurements. However, this process requires stopping the wind turbine to measure the residual noise level, which again decreases the electrical productivity, but above all, gives only a short temporal representation of the residual noise because these measurement campaigns last a few weeks.

This study aims to develop a tool making it possible to continuously estimate the sound emergence of wind turbines without stopping them. With such a tool, the curtailment plan could be adapted more regularly and more easily. Electricity production would be better optimized and residents would be better protected from louder noise emissions.

The main difficulty here is to correctly extract the wind turbine noise component from the ambient sound component since it is permanently mixed with the residual noise. Gallo et al. [1] proposed a method to overcome this difficulty assessing the wind turbine noise from measurements based on an analytical model. The wind turbine noise is here not considered as a function of the wind speed but as a function of a parameter related to the rotation of all the wind turbines. The residual noise is defined according to the wind speed at ground. To do so, they do not consider, on the used measurements they collected, the anthropic and the animals noises and assumes that the residual noise is mainly due to the wind noise. This assumption can be restrictive because it does not consider the dawn chorus of birds or may not apply depending on the location of the wind farms. This difficulty to isolate a specific sound source in sound mixtures has already been tackled for other noise sources, for example for the estimation of the traffic noise from measurements. In [2], the traffic noise is estimated indirectly by detecting other sound sources. The traffic sound level is them calculated by rejecting the temporal frames where these sound events were detected. In parallel, there is a growing interest in the community dedicated to the sound signal processing for the use of machine learning techniques for environmental sounds [3]. Multiple applications have been investigated in the detection [4], classification [5] and recognition [6] tasks. One possibility offered by the machine learning techniques is to allow the source separation between the different components of a signal [7], [8]. Still in the urban traffic noise, a source separation technique has recently been deployed in order to estimate the urban traffic sound level [9] with the Non-negative Matrix Factorization framework. The source separation approach has the advantage, once the sound of interest is isolated, of making it possible to express the desired indicators such as the sound level. Since this machine learning technique is well-suited for monaural measurements and found multiple applications in the environmental sound field [10], this method is considered to estimate the sound emergence of wind turbines. The purpose of this work is to see if this could be an appropriate choice for wind turbine noise and for such an application.

The first section summarizes the Non-negative Matrix Factorization framework, the second section presents the different used corpora. Then the third and the fourth parts present the numerical experiment carried out and the results.

2. NON-NEGATIVE MATRIX FACTORIZATION

2.1 Principe

Non-negative Matrix Factorization (NMF) [11,12] is a linear approximation method which consists in approximating a non negative matrix $\mathbf{V}_{F\times N}$ in $\mathbb{R}^+$ by the product of two non negative matrices: $\mathbf{W}$, called *dictionary* with $F \times K$ dimensions and $\mathbf{H}$, an activation matrix with K $\times$ N dimensions such as $\mathbf{V} \approx \mathbf{WH}$. In audio domain, $\mathbf{V}$ corresponds to the magnitude of an audio spectrogram and $\mathbf{W}$ is composed of audio spectra. Consequently, F represents the number of frequency bins, N, the number of temporal bins and K, the *rank* of the approximation which corresponds here to the number of spectra in $\mathbf{W}$. The approximation of $\mathbf{V}$ by the matrix product $\mathbf{WH}$ is defined by the minimization of a cost function:

$$\min_{\mathbf{W},\mathbf{H}\geq 0} D_\beta(\mathbf{V}||\mathbf{WH}). \tag{2}$$

$D_\beta(\cdot|\cdot)$ is the β-divergence and is defined as $D_\beta = \sum_{f=1}^{F}\sum_{n=1}^{N} d_\beta\left(\mathbf{V}_{f,n}|\left[\mathbf{WH}\right]_{f,n}\right)$. This class includes 3 particular classes:

- Itakura-Saito divergence (β = 0): $d_\beta(x|y) = \frac{x}{y} - \log\frac{x}{y} - 1$,
- Kullback-Leibler divergence (β = 1): $d_\beta(x|y) = x\log\frac{x}{y} - x + y$,
- Euclidean distance (β = 2): $d_\beta(x|y) = \frac{1}{2}(x-y)^2$.

The minimization problem of NMF (eq. 2) is solved iteratively by updating the matrices $\mathbf{W}$ and $\mathbf{H}$. From the different existing algorithms, the Multiplicative Update [13] is chosen here, since it ensures the convergence of the results and the non-negative results:

$$\mathbf{H}^{(i+1)} \leftarrow \mathbf{H}^{(i)} \otimes \left(\frac{\mathbf{W}^T\left[\left(\mathbf{WH}^{(i)}\right)^{(\beta-2)} \otimes \mathbf{V}\right]}{\mathbf{W}^T\left[\mathbf{WH}^{(i)}\right]^{(\beta-1)}}\right)^{\gamma(\beta)} \tag{3a}$$

$$\mathbf{W}^{(i+1)} \leftarrow \mathbf{W}^{(i)} \otimes \left(\frac{\left[\left(\mathbf{W}^{(i)}\mathbf{H}\right)^{(\beta-2)} \otimes \mathbf{V}\right]\mathbf{H}^T}{\left[\mathbf{W}^{(i)}\mathbf{H}\right]^{(\beta-1)}\mathbf{H}^T}\right)^{\gamma(\beta)}. \tag{3b}$$

where $\gamma(\beta) = \frac{1}{2-\beta}$ for $\beta < 1$, $\gamma(\beta) = 1$ for $\beta \in [1,2]$ and $\gamma(\beta) = \frac{1}{\beta-1}$ for $\beta > 2$. The A $\otimes$ B and A/B operators represent the Hadamard product and ratio. When NMF is used as a dictionary learning method, nor $\mathbf{W}$ nor $\mathbf{H}$ is known, it is then an unsupervised NMF that is carried out. When $\mathbf{W}$ is learned on labeled data and only $\mathbf{H}$ is updated, it is a supervised NMF. Some derived approach have been developed (smoothness [14], source/filter [15]) in order to better adapt it to different cases. Here, it is a variant of NMF that is considered through Thresholded Initialized NMF.

2.2 Thresholded Initialized NMF

TI-NMF has been proposed in [9]. First, an initial dictionary $\mathbf{W_0}$ is learned on a sound database dedicated to the sound source of interest. Then, this matrix is updated, with $\mathbf{H}$, on the sound mixtures. After been updated, the element in the final dictionary $\mathbf{W}'$ can model other sound sources. To only consider the source of interest, a classification step is added to extract the related spectra from $\mathbf{W}'$. To do so, a cosine similarity is computed to estimate the similarity between each element k:

$$D(\mathbf{W_{0,k}}|\mathbf{W'_k}) = \frac{|\mathbf{W_{0,k}} \cdot \mathbf{W'_k}|}{\parallel \mathbf{W_{0,k}} \parallel \parallel \mathbf{W'_k} \parallel}. \tag{4}$$

This indicator gives a value between 0 and 1: 1 means that the k element is exactly the same in $\mathbf{W_0}$ and $\mathbf{W'}$; while 0 means that these elements are completely different. Then by setting a threshold t between 0 and 1, it is possible to classified according to the cosine similarity, which element can be considered as a wind turbine spectra. In the present study, the labeled data are wind turbine spectra (introduced in section 3). Then as output, we can extract the wind turbine component $[\mathbf{WH}]_{WT}$ with $\mathbf{W}_{WT}$ with $F \times J$ dimensions, where $J \leq K$, the number elements considered as wind turbine spectra. In Figure 1, an example of the cosine similarity values sorted in descending order is displayed. All the correspondent elements in $\mathbf{W'}$ that have a similarity with $\mathbf{W_0}$ superior to the threshold (here illustratively chosen at t = 0.6) are considered as wind turbine spectra. In all, 156 elements are retained (J = 156). With their respective vectors in $\mathbf{H}$, the selected elements form the wind turbine component $[\mathbf{WH}]_{WT}$.

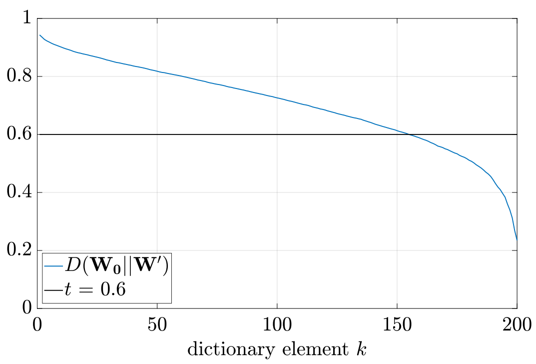

Figure 1. Example of the cosine similarity look (sorted in descend order).

3. ENVIRONMENTAL SOUND CORPUS

To test the performance of NMF on the noise of wind turbines, two different corpus are built. Both are based on data collected during a wind farm measurement campaign in France [16]. In this campaign, *on*/*off* cycles were alternatively performed to measure ambient and residual noise levels at different distances from the wind farm. For this study, measurements close to the machines are used to have labeled data to build the dictionary and to simulate environmental measurements.

3.1 Dictionary corpus

First, TI-NMF (see section 2.2) needs tagged data to learn a dictionary of wind spectra. To obtain the cleanest data possible, only the measurements data corresponding to the IEC 61400-11 standard [17] are taken into account. The microphone being on the ground, the contribution of sound reflection is avoided and we consider that the microphone is close enough to wind turbines (150 m) to neglect the

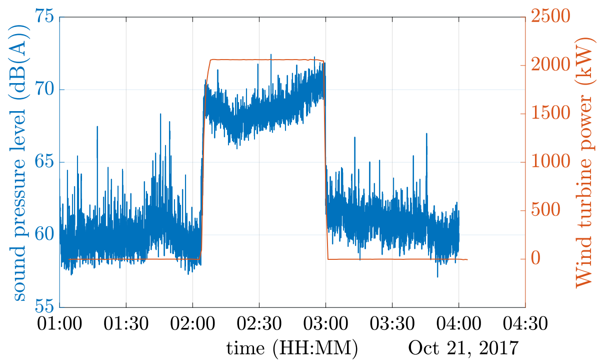

Figure 2. Example of the period extracted for the dictionary in an off/on/off cycle. The residual noise for this time slot is calculated from the samples between 1.30 pm and 2.00 pm and between 3.00 pm and 3.30 pm.

effects of sound propagation. Only the measurements corresponding to stable electrical production and for which the sound emergence of the wind turbine noise is greater than 10 dB (A) are included in the dictionary corpus. To ensure that the cleanest samples are obtained, the residual component (even if less loudly) is filtered out from the measurements by quadratic subtraction. This component is estimated by the equivalent sound level obtained 30 minutes before the *on* cycle and 30 minutes after.

Finally, the dictionary corpus is made up of 14,200 samples of one-second wind turbine spectra in the frequency range $[50 - 5000]$ Hz in 1/3 octave bands (F = 21), which represents approximately 4 hours of cumulative duration.

3.2 Environmental noises corpus

This corpus includes elements allowing to simulate environmental measures. The noise from wind turbines comes from measurements made 300 m from the wind farm, which are the closest measurements made on this wind farm (after those following the IEC standard). Available recordings made at a greater distance do not allow sufficient emergence to guarantee clean samples (i.e. not polluted by residual noise). Again, a selection is made to keep the most emerging wind turbine samples from the residual noise. Those considered samples are filtered of the residual component following the same protocol as for the dictionary corpus. The samples are cut into 10 minutes samples without overlapping. This corpus is finally composed of 277 samples of noise from wind turbines with a duration of 10 minutes of 1 second in third octave bands in the frequency range $[50 - 5000]$ Hz. Finally, 30 samples lasting 10 minutes, composed exclusively of residual noise, are extracted from the measurements at 1250 meters from the wind farm during an *off* cycle. In addition, to bring more diversity in the creation of environmental measures, the corpus is supplemented by isolated residual sound events that can be found in French countryside (tractor, airplane, bird whistles, passing car. . .) [18]. Finally, these 3 corpora (wind turbine noise samples, residual background noise samples and isolated residual sound events) are used to simulate academic measurements on which NMF is applied to estimate the sound emergence of a wind turbine (see section 4.1.

4. SUMMARY OF THE EXPERIMENT

After having introduced the different elements for this study in sections 2 and 3, the measurement corpus, the experimental factors and their modalities involved in TI-NMF are exposed.

4.1 Design of the measurement corpus

To assess the performance of TI-NMF on wind turbine noise, 30 scenes lasting 10 minutes are constructed to compose the measurement corpus. Figure 3 summarizes the different step involved in this process.

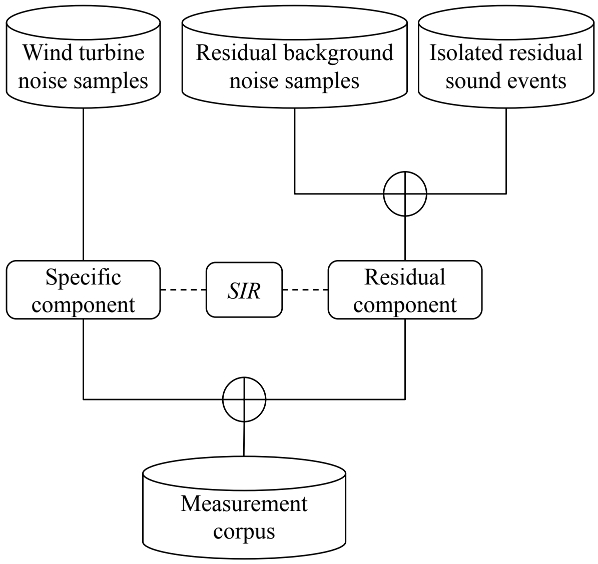

Figure 3. Block diagram of the measurement corpus building.

These scenes are composed of a wind turbine component, taken at random from the 277 samples, a residual background noise and several residual sound events. In order not to complicate the scenes too much and to be representative of what can happen in the countryside, less than 6 isolated sound classes can be present during the duration of every scene. The sound emergence of these isolated events is adjusted according to the background noise empirically by listening to avoid inconsistent behavior. In the rest of the document, the wind turbine noise component will be called the *specific* component while the association of the background noise and additional residual events will be called the *residual* component. The energetic sum of the specific component and the residual component generates the *ambient* component which simulates the measurements recorded in the field when the wind turbines are operating. It corresponds to the spectrogram $\mathbf{V}$ for NMF (see section 2).

In addition, to test different situations, the sound pressure level of the residual component is calibrated according to the sound level of the specific component through the *Signal-to-Interference Ratio* [19]:

$$SIR = L_{A,50,specific} - L_{A,50,residual} \tag{5}$$

with $SIR \in \{-9, -6, -3, 0, 3, 6, 9\}$ dB(A). These values generate an equivalent *Signal-to-Noise-Ratio*, $SNR = L_{A,50,ambient} - L_{A,50,residual}$, which corresponds to the sound emergence E (eq. 1) of wind turbine noise, that worth respectively $SNR \in \{0.5, 1.0, 1.8, 3.0, 4.8, 7.0, 9.5\}$ dB(A).

The advantage of this simulation process is to allow a total control of the content of the measurement because all the components and their contribution to the sound level can be estimated separately, which is never possible with the *in situ* measurements. Thus, the exact sound emergence E of the wind turbine noise can be calculated following the eq. 1 and is compared to that estimated by TI-NMF.

4.2 Design of the dictionary corpus

The complete dictionary corpus is not fully used in TI-NMF as it represents a large matrix which would slow down the computation time and, most of all, it has redundant information in it. In order to reduce it, a K-mean algorithm is applied on the dictionary corpus to obtain initial dictionary $\mathbf{W_0}$ with different size such as $K \in \{25, 50, 100, 200\}$.

4.3 NMF computation

In addition with the different built dictionaries, TI-NMF bring several experimental factors which can take multiple values. The β-divergence parameter is fixed to 3 values, $\beta \in \{0, 1, 2\}$. The threshold t is set from 0.01 to 0.99 with a 0.01 step and 100 iteration are performed. The spectrogram $\mathbf{V}$ and the dictionary $\mathbf{W_0}$ are both expressed with A weighting in 1 second third octave bands. After the 100 updates of $\mathbf{W}'$ and $\mathbf{H}$, the thresholding step makes it possible to extract the wind turbine component $[\mathbf{WH}]_{sp.}$. From this spectrogram, it is then possible to estimate its 1 second A-weighting sound level, $\tilde{L}_{Aeq,1s,specific}$, the residual component, $\tilde{L}_{Aeq,1s,residal}$ and so, the statistical sound level, $\tilde{L}_{A50,residual}$. The estimated sound emergence $\tilde{E}_i$ of a sample i is then obtained: $\tilde{E}_i = L_{A,50,ambient,i} - \tilde{L}_{A,50,residual,i}$.

The retained metric used to assess the performance is the mean absolute error, $MAE = \dfrac{\sum_{i=1}^{N} |\tilde{E}_i - E_i|}{N}$ with N = 30, the number of simulated ambient measurement. The bias, $B = \dfrac{\sum_{i=1}^{N} \tilde{E}_i - E_i}{N}$, is also considered to complete the first metric and to help to better apprehend the results. The entire schema of the experiment is summarized in Figure 4.

4.4 Reference method

To assess the performance of TI-NMF, a reference method is used. It is based on the method currently used to estimate the sound emergence of installed wind turbines. It consists

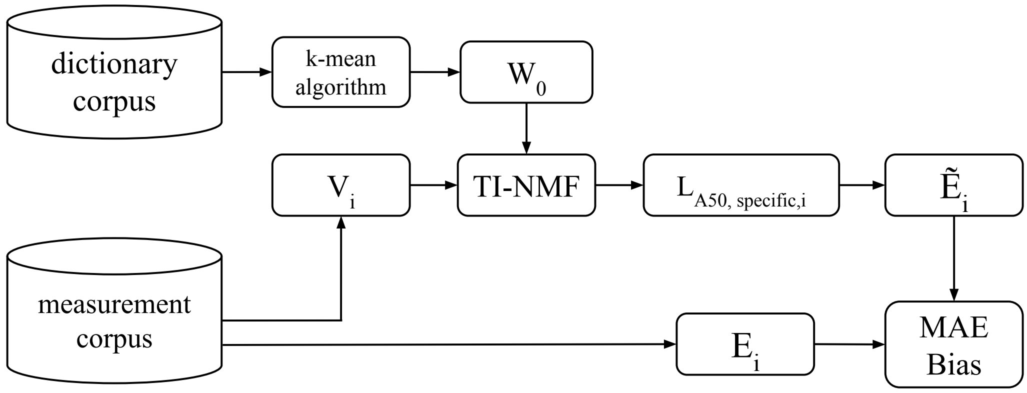

Figure 4. Block diagram of the experience to estimate the sound emergence of wind turbines by TI-NMF.

in evaluating the ambient sound level, $L_{A50,ambient}$ when the wind turbines are operating (cycle *on*) and the residual sound level when the wind turbine is stopped (cycle *off*), $L_{A50,residual,off}$. To do so, for this method, each scene is completed by 10 minutes of residual noise to simulate the cycle *off*. As previously, this component is the combination of the residual background of the corresponding scene during the cycle *on* and some isolated residual sound events, which are different here but with the same distribution (less than 6 different sound classes by scene). To avoid too large variations between the *on* residual and the *off* residual (for each of the 30 scenes), the two components are calibrated to the same equivalent sound level L_{Aeq}. Then with the SIR parameters, the *off* residual is adjusted to keep the same level than the *on* residual. In Figure 5, an example is displayed with the wind turbine component and the residual component for the *on* and *off* cycles. The sound emergence is thus calculated as $E_{ref} = L_{A50,ambient} - L_{A50,residual,off}$.

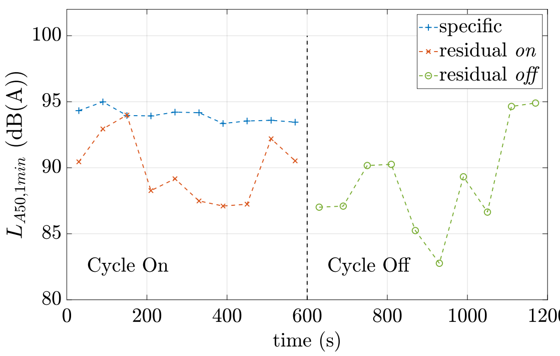

Figure 5. Example of the scene building for the reference method.

5. RESULTS AND DISCUSSIONS

The experience consists in calculating the set of combinations of the parameters. For each set, a MAE error and a bias is obtained. However, it is possible to average the metrics between the 7 values of the SIR to determine the most efficient NMF on the different cases of predominance of the wind turbines. Indeed, in practice, the value of SIR is unknown. The set of NMF parameters that reaches the lowest average MAE is then the most efficient for different wind turbine predominance.

5.1 Global results

In Table 1, the set of parameters that reaches the lowest average MAE errors for each β value are summarized. The error generated by the reference method is also added.

	β	K	t	MAE (dB(A))
reference	-	-	-	3.3 ($\pm$ 2.0)
NMF IS	0	100	0.78	1.7 ($\pm$ 1.5)
	1	**200**	**0.68**	$\mathbf{1.6 (\pm 1.5)}$
	2	200	0.66	1.8 ($\pm$ 1.6)

Table 1. MAE errors and their standard deviation of reference method and TI-NMF for each β value that reached the lowest errors on all the measurement and SIR values. In bold, the set that reached the lowest error.

For each β value, TI-NMF obtains lower estimation error on the emergence than the reference method with lower standard deviation too. Despite different settings in TI-NMF, there is no configuration that is significantly better than the others. The lowest set is obtained with β = 1, K = 200 and for a threshold t = 0.68. This setting is the one that get the best averaged performance for all the SIR values. From these results, it is now possible to extract the errors made by this set for each value of SIR. In Figure 6, the evolution of the MAE errors and the bias are displayed. In addition, the errors generated by the reference method are added.

First, the MAE errors and the bias for the reference method are constant according the SIR values. For both of these metrics, the difference of the estimated emergence $\tilde{E}$ and the exact one E is equivalent to the difference between the residual component of the *off* cycle, $\tilde{L}_{A50,residual}$ and the one of the *on* cycle, $L_{A50,residual}$. This result is independent of the SIR values and reflects the differences in construction between this two residual components. There

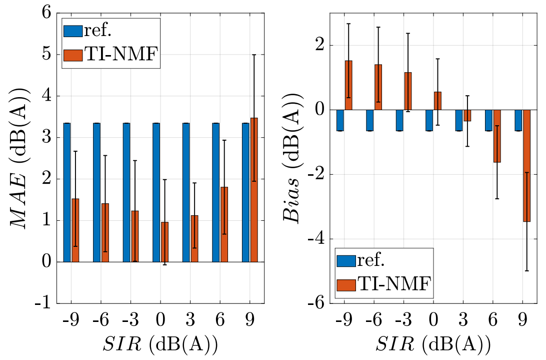

Figure 6. Evolution of the MAE error (on left) and the bias (on right) according to the SIR value with the standard deviation as error bars. For visibility, as constant, the standard deviation of the reference are not displayed.

is an absolute averaged difference of 3 dB(A) and a bias close from zeros. This last metric reveals a balanced distribution of these differences on the 30 scenes: almost half of all scenes have an *off* residual component lower than the *on* residual and vice versa.

The errors generated by TI-NMF are different and evolve according to the predominance of the wind turbine. For $SIR \leq 3$ dB (A), the MAE remains low (<1.5 dB (A)) with a standard deviation less than 1.1 dB (A). Since the predominance of the wind turbine increases, the error also increases to exceed 3 dB (A) at $SIR = 9$ dB (A). The bias allows us to see this evolution: when $SIR \leq 0$ dB (A), the bias is positive, which means that the estimated emergencies are overestimated (less than 2 dB (A)), while when SIR becomes positive, they are underestimated up to -3.5 dB (A) for the highest value of SIR. Finally, TI-NMF offers a satisfactory estimate of the emergence when the wind turbine is less present than the residual component, but its performance decreases as it becomes the predominant sound source. To illustrate this performance, the Figure 7 compares the estimated emergence according to the exact emergence.

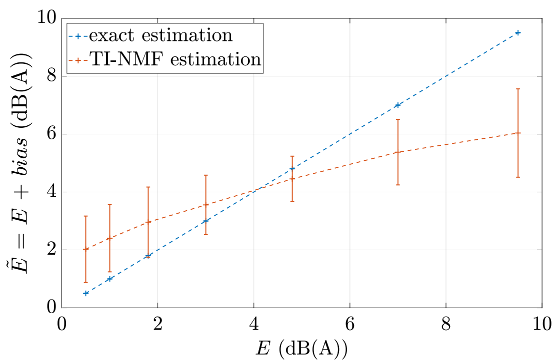

Figure 7. Comparison of the estimated emergence's behavior according to perfect estimations (TI-NMF, β = 1, K = 200, T = 0.68).

The behavior of $\tilde{E}$ is similar to that described above. The current regulations impose not to exceed the residual noise level of 3 dB (A) at night and 5 dB (A) during the day. In the first case, the TI-NMF overestimation is not harmful up to E = 2 dB (A). Afterwards, between an exact emergence of 2 dB (A) and 3 dB (A), while the emergence of the wind turbine still complies with the regulations, the proposed method overestimates the sound emergence and declares the emergence above the regulation limit of 3 dB (A). Consequently, a curtailment plan that is too strong will be planned and will decrease electrical productivity. For an exact emergence between 3 dB (A) and 4 dB (A), the method always overestimates the emergence by 0.5 dB. For the limitation of 5 dB (A), the method underestimates the emergence of sound up to 0.5 dB for an exact emergence between 4 dB (A) and 5 dB (A). Conversely, the risk is now to underestimate the curtailment plan, to produce too much noise and to no longer comply with the regulations. Above, even with the underestimation, in the case where, for whatever reason, the wind turbine is very noisy, despite strong errors MAE, the detection of this case can still be done.

5.2 Estimations of wind turbine and residual sound

To better understand these results, it is necessary to detail the estimate of the specific sound level and the residual level. Their errors MAE and the bias are displayed in the figure 8 for each value of SIR. Here, the metrics are not calculated from the emergences but from the difference between the exact and estimated statistical sound levels of the specific component, L_{A50} and $\tilde{L}_{A50}$.

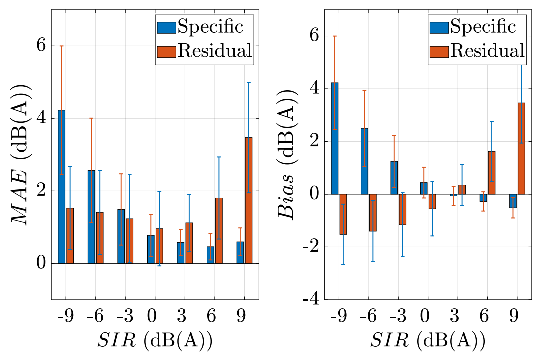

Figure 8. Evolution of the MAE error (on left) and the bias (on right) according to the SIR value with the standard deviation as error bars of the estimation of the wind turbine and the residual components (TI-NMF, β = 1, K = 20, t = 0.68).

The errors in estimating the noise level of wind turbines fluctuate according to the SIR values: when the wind turbine is less predominant than the residual component (SIR < 0 dB(A)), the wind turbine sound level estimation by TI-NMF is less performing (MAE > 1.5 dB(A)) than when it is the predominant sound source (MAE < 0.8 dB(A)). For $SIR \leq 0$ dB(A), the bias is positive which means that the $\tilde{L}_{A50,specific}$ is overestimated compared to the exact value. This overestimation increases strongly as the

SIR tends to decrease. In the opposite, when $SIR > 0$ dB(A), the estimation of the $L_{A50,specific}$ becomes underestimated but less than 1 dB(A). The residual component is estimated from the estimation of the wind turbine component. The evolution of the MAE errors differs as they are low when $SIR \leq 6$ dB(A) (< 2 dB(A)) to increase strongly then. The biases have an opposed behavior to the wind turbine component, with an underestimation of the sound levels for $SIR \leq 0$ dB(A) and then an overestimation. The evolution of the biases for the two components is complementary. For negative SIR values, an overestimation of the wind turbine component by TI-NMF will generate naturally an underestimation of the residual component because too much energy is included in the first one. The MAE errors remain however low since the ambient component is mainly influenced by the residual noise. Consequently, even with strong errors, the residual component (and therefore the emergence $\tilde{E}$) is well estimated. For positive values of SIR, the wind turbine being the main source of noise, the slightest error in its sound level estimation will greatly increase the residual sound level of the component, this behavior is independent of the chosen method. These errors are then transferred in the estimated emergence.

5.3 Influence of the threshold of TI-NMF

In order to apprehend the TI-NMF's behavior and the specific sound level estimation, the evolution of the cosine similarity according to the SIR and that of the MAE errors according to the threshold values are exposed respectively in the Figures 9 and 10.

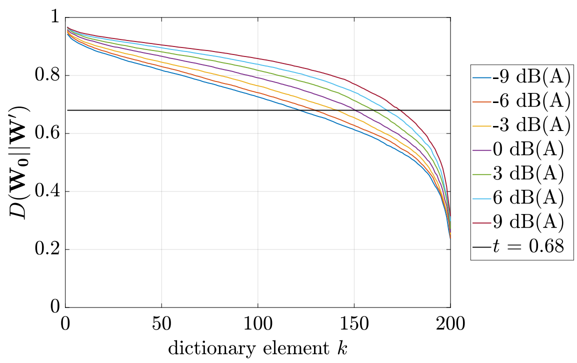

Figure 9. Mean cosine similarity sorted in descending order for each SIR value (TI-NMF, $\beta = 1$, $K = 20$, $t = 0.68$).

Figure 9 allows us to see that, with the update of the initial dictionary $\mathbf{W_0}$ to each scene, the number of the elements that are considered as wind turbine component in $\mathbf{W'}$ varies from scene to scene. The more the wind turbine is predominant, the more the number of elements that describes this sound source is important, while it decreases with the decrease of SIR. This evolution is due to the NMF's behavior defined by the cost function (eq. 2): to minimize it, depending on the predominance of the specific source, a variable number of elements in $\mathbf{W}$ will be dedicated to describe this source. As the residual component becomes predominant, more elements turn into this component which tend to decrease the cosine similarity. They are then less elements to describe the specific source.

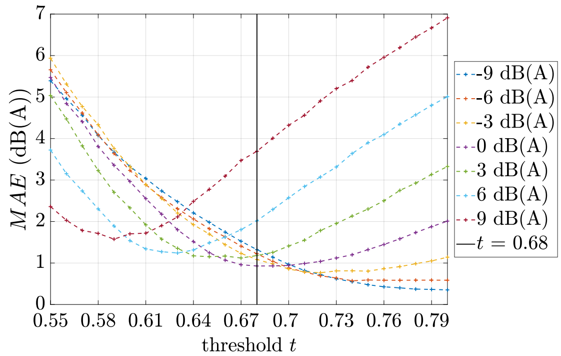

Figure 10. Variation of the MAE errors according to the threshold value (TI-NMF, $\beta = 1$, $K = 20$, $t = 0.68$).

In complement, Figure 10 exposes the sensibility of the MAE errors to the threshold value according to the SIR. Even if the retained set of TI-NMF's parameters reaches the lowest MAE errors on all the SIR values, on each SIR there is an optimal threshold which gets a lowest error. In the case where SIR is negative, the threshold needs to be increased to diminish the error, it will decrease the number of element in $\mathbf{W}_{sp}$. This means that some elements in $\mathbf{W}_{sp}$ are considered as wind turbine spectra because of the similarity of their spectra. In opposite, with positive SIR, it is necessary to decrease the threshold (and to increase the number of elements in $\mathbf{W}_{sp}$. The errors are then due to the fact that some wind turbine spectra in $\mathbf{W'}$ are discarded. However, even if the fixed threshold can miss some wind turbine spectra or can generate false-positive, it has to be reminded that the SIR value is unknown in practice. A fixed threshold stays thus the best compromise to get the best averaged performance.

6. CONCLUSION

Because of the French regulations dedicated to the impact of noise from wind turbines, there is a need to know how to estimate the sound emergence of wind turbines *in situ* without stopping the wind turbine. This will improve electrical productivity and better protect neighboring residents from noise. To do this, Thresholded-Initialized NMF is used to estimate the sound emergence of wind turbines. This machine learning method makes it possible to separate the wind component from the ambient measurements and to deduce its emergence. TI-NMF was applied to simulated ambient measurements where the wind turbine and the residual components are known. This process allows a total control of the noise contribution of the wind turbines. Furthermore, this academic corpus based on *in situ* measurements presents sound diversity by the add of different isolated sound events. This preliminary study estimated a set of parameters of TI-NMF ($\beta = 1$, $K = 200$, $t = 0.68$) which gives satisfactory results on the whole corpus and all the SIR values. The various results exposed reveal a variable behavior of this method according to the predominance of the wind turbine: if TI-NMF obtains small errors for SIR <3 dB (A) (MAE <1.5 dB (A)), as the wind component is more and more present, the error of the sound emergence increases (MAE> 1.8 dB (A) when SIR> 6 dB (A)). The origin of these errors has been identified: when the SIR is negative (or positive), TI-NMF considers two numerous (or too few) elements as wind turbines which will generate an overestimation (or an underestimation) of the sound emergence. The highest errors (SIR> 6 dB (A)) are also due to the fact that the estimation of the residual sound level is more sensitive to the error in estimating the wind turbine noise level. From these promising results, it is possible to consider this approach for further investigation, to compare the method to more complex cases and improve TI-NMF performances. In addition, one of the main objectives will consist in adding sound propagation elements to consider the wind turbine noise at specific distances in order to better simulate *in situ* situations and to test the robustness of the method on different scenarios.

7. REFERENCES

[1] P. Gallo, L. Fredianelli, D. Palazzuoli, G. Licitra, and F. Fidecaro, "A procedure for the assessment of wind turbine noise," *Applied Acoustics*, vol. 114, pp. 213–217, 2016.

[2] J. C. Socoró, F. Alías, and R. M. Alsina-Pagès, "An anomalous noise events detector for dynamic road traffic noise mapping in real-life urban and suburban environments," *Sensors*, vol. 17, no. 10, p. 2323, 2017.

[3] A. Mesaros, T. Heittola, A. Diment, B. Elizalde, A. Shah, E. Vincent, B. Raj, and T. Virtanen, "Dcase 2017 challenge setup: Tasks, datasets and baseline system," in *DCASE 2017-Workshop on Detection and Classification of Acoustic Scenes and Events*, 2017.

[4] J. Li, W. Dai, F. Metze, S. Qu, and S. Das, "A comparison of deep learning methods for environmental sound detection," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 126–130, IEEE, 2017.

[5] J. Salamon and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification," *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 279–283, 2017.

[6] S. Chachada and C.-C. J. Kuo, "Environmental sound recognition: A survey," *APSIPA Transactions on Signal and Information Processing*, vol. 3, 2014.

[7] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, "A consolidated perspective on multimicrophone speech enhancement and source separation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 4, pp. 692–730, 2017.

[8] P. Chandna, M. Miron, J. Janer, and E. Gómez, "Monoaural audio source separation using deep convolutional neural networks," in *International conference on latent variable analysis and signal separation*, pp. 258–266, Springer, 2017.

[9] J.-R. Gloaguen, A. Can, M. Lagrange, and 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.

[10] A. Mesaros, T. Heittola, O. Dikmen, and T. Virtanen, "Sound event detection in real life recordings using coupled matrix factorization of spectral representations and class activity annotations," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 151–155, Apr. 2015.

[11] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, Oct. 1999.

[12] S. Innami and H. Kasai, "Nmf-based environmental sound source separation using time-variant gain features," *Computers & Mathematics with Applications*, vol. 64, no. 5, pp. 1333–1342, 2012.

[13] C. Févotte and J. Idier, "Algorithms for nonnegative matrix factorization with the β-divergence," *Neural Computation*, vol. 23, no. 9, pp. 2421–2456, 2011.

[14] S. Essid and C. Févotte, "Smooth nonnegative matrix factorization for unsupervised audiovisual document structuring," *Multimedia, IEEE Transactions on*, vol. 15, no. 2, pp. 415–425, 2013.

[15] J.-L. Durrieu, G. Richard, B. David, and C. Févotte, "Source/filter for unsupervised mai melody extraction from polyphonic audio signals," *Audio, speech and language processing, IEEE Transactions on*, vol. 18, no. 3, pp. 564–575, 2010.

[16] B. Kayser, B. Gauvreau, Écotière, D., and C. Lebourdat, "A new experimental database for wind turbine noise propagation in an outdoor inhomogenous medium," in *Proc. 17th Long Range Sound Propagation Symposium, Lyon, France*, LMFA, June 2018.

[17] *IEC 61400-11:2012: Wind turbines - Part 11: Acoustic noise measurement techniques*, 2012.

[18] J.-R. Gloaguen, A. Can, M. Lagrange, and J.-F. Petiot, "Creation of a corpus of realistic urban sound scenes with controlled acoustic properties," in *173rd Meeting of the Acoustical Society of America and the 8th Forum Acusticum (Acoustics' 17)*, vol. 30, pp. 4044–4044, ASA, 2017.

[19] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE transactions on audio, speech, and language processing*, vol. 14, no. 4, pp. 1462–1469, 2006.