11th Edition of the

International Conferences on
Wind Turbine Noise

Copenhagen, Denmark – 10th to 13th June 2025

Impact of environmental data on wind turbine noise level estimation

A. RKHISS [1], **A. FINEZ, J-R. GLOAGUEN.**
Engie Green, Lyon, 69007, France

G. VASIELE
GIPSA-Lab, Grenoble-INP UGA, Grenoble, 38000, France

J. MAILLARD
CSTB, Grenoble, 38000, France

Summary

Wind energy is one of the most widely used renewable energy sources in the world and has grown rapidly in recent years. However, wind turbines generate noise that is often perceived as a disturbance by nearby residents. So, developing tools to assist wind farm developers and regulatory authorities is essential. This study focus on the impact of environmental data on wind turbine noise (WTN) level estimation using recurrent neural networks (RNNs). We compare the performance of an architecture which is based on long-short term memory cells (LSTM). LSTM model trained using only acoustic features in the frequency range of 31.5 Hz to 2 kHz with those incorporating additional environmental features, such as wind speed and wind turbine power accorded to each wind speed value. The results highlight the influence of these factors on noise characterization and demonstrate the extent to which environmental data enhances WTN level estimation.

1. Introduction

The noise generated by wind turbines has raised concerns among residents living near wind farms, as it can negatively impact sleep quality and overall well-being [1]. With the increasing size of modern wind turbines, noise issues have become more significant, prompting several countries, including France, to implement strict regulatory standards to control wind turbine noise (WTN) emissions. These regulations define emergence as the difference between the total noise level during wind turbine operation and the background noise level when the turbines are inactive. Compliance requires that emergence does not exceed 5 dB(A) during the day and 3 dB(A) at night if the total noise is exceeding 35 dB(A).

An on/off strategy for wind turbines is commonly used to verify a curtailment plan employed to reduce noise emissions. This plan is implemented to limit the noise impact of wind farms and ensure regulatory conformity. During the operational phase, it is verified through measurements taken during start/stop cycles of the wind farm. However, these measurements have several drawbacks. Their limited duration does not always reflect the variability of residual noise. They are also costly, as they require shutting down the

[1] abdelazyz.rkhiss@grenoble-inp.fr

turbines. Additionally, depending on the situation, they may lead to excessive or insufficient curtailment of the wind farm.

Several methods have been explored to address this issue. For instance, Gloaguen et al. [2] proposed a non-negative matrix factorization (NMF)-based approach to estimate WTN levels. Despite promising results in some cases, the uncertainty of the method prevents its deployment on an industrial scale. Consequently, we are moving towards the use of deep neural networks, given their success in the field of source separation [3], sound classification [4] and sound event detection [5]. Anicic et al. [6] applied Support Vector Regression (SVR) to predict WTN levels from acoustic and wind speed data, confirmed the importance of integrating environmental data, such as wind speed, to increase the accuracy of prediction models.

The method presented in this paper uses Recurrent Neural Networks (RNNs) for WTN level estimation from total noise monitoring. RNNs proved their effectiveness in extracting features from acoustic scenes[7], and handling long-term dependencies in sound patterns. These models are well-suited for capturing the temporal dynamics and sequential nature of acoustic data, making them good candidates for extracting WTN level from total noise.

Section 2 provides a comprehensive overview of the methodology, beginning with the dataset construction process, from background noise measurements and WTN synthesis, and environmental data description. Section 3 then outlines the data preprocessing steps, provides an explanation of the LSTM cell, and introduces the proposed architecture along with the evaluation metrics and optimization algorithm. The experimental results are discussed in the final section, Section 4.

2. Dataset

To train the supervised LSTM model, it is crucial to have an appropriate training dataset. The dataset includes the time series of the total noise level, L_{TN} as input, while the overall sound pressure level of wind turbine noise, L_{WTN}, serves as the output (label). Both time series consist of one-second equivalent noise levels, measured in third-octave bands ranging from 31.5 Hz to 2 kHz. The levels are expressed in A-weighted decibels (dB(A)).

Measuring L_{WTN} during recordings presents a challenge due to the presence of background noise from various sources that cannot be isolated. To create realistic soundscapes, a hybrid approach combining measurement and sound synthesis is adopted. This method used a background noise measured at a development site during an initial state (2.1), ensuring consistency with wind measurements taken at a height of 81 meters. This background noise combines with a generated WTN based on wind speed measurements and machine specifications (see Subsection 2.2).

2.1 Background noise measurements

Background noise levels were measured in third-octave bands, sampled at a one-second resolution to provide a detailed representation of the sound environment, in line with the capabilities of sound level meters. These measurements are accompanied by wind speed data recorded over 10-minute intervals. Fig. 1 presents a sample of the measured Background noise levels in dB(A).

2.2 Wind turbine noise synthesizing

The estimation of sound power levels (L_W) for wind turbines relies on the Machine Specifications (MS) document, which provides octave-band values in dB(A) for various wind speeds, typically following the ISO 61400-11 standard procedure. As a choice for this study, a 2 MW industrial turbine with a 90 m diameter is used, with L_W values specified for wind speeds ranging from 3 to 25 m/s. To obtain time-varying L_W values, the recorded 10-minute wind speed measurements are used to linearly interpolate the corresponding

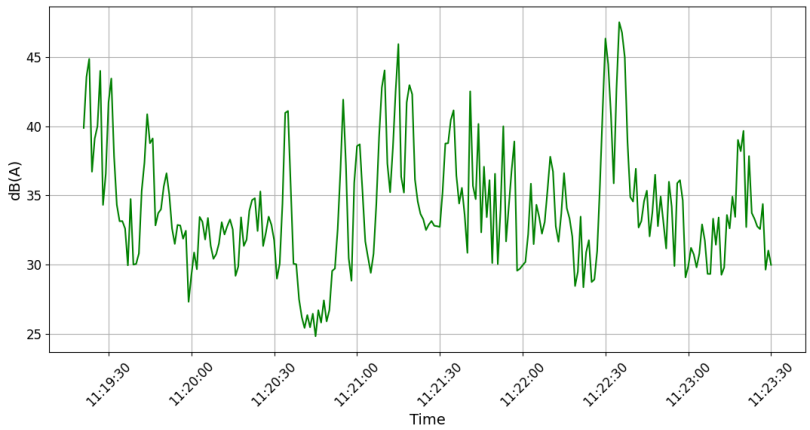

Figure 1 The overall SPL of the measured background noise sample (L_{BN}).

sound power levels from the MS. This process allows for obtaining $L_w(F,T)$ in octave bands (F) for each 10-minute interval (T). An up-sampling procedure is then applied to convert these results ($L_w(F,T)$) into third-octave bands (f), yielding $L_w(f,T)$. This conversion simply serves to bring different data sources onto the same frequency basis. The next step involves adjusting the sampling frequency of the signal from 10-minute intervals to one second by applying zero-padding. This ensures compatibility with the desired sampling frequency using an up-sampling ratio of 1/600, ultimately obtaining $L_w(f,t)$ in third-octave bands per second (t). Fig. 2 illustrates the complete process for generating time series of acoustic power for the WTN component.

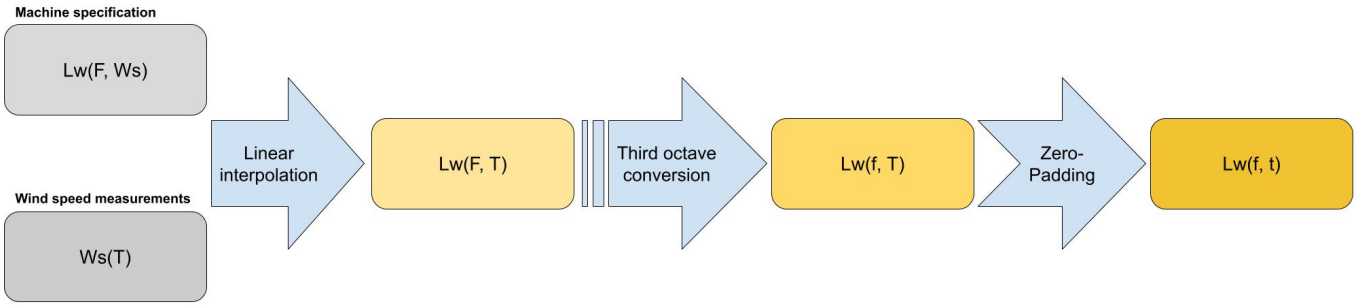

Figure 2 Flowchart illustrating the step-by-step process for obtaining time series of acoustic power levels. The process starts with machine specifications and wind speed measurements, followed by linear interpolation to estimate $L_w(F,T)$. A third-octave conversion, and zero-padding is applied to resample the data to a third octave bands and one-second resolution, resulting in $L_w(f,t)$.

The sound pressure level of the WTN (L_{WTN}) is computed for a simplified scenario in which a wind farm consists of a single wind turbine. This level is determined at the receiver point by applying an acoustic propagation filter. The filter is based on solving the parabolic equation, a physical propagation model that accounts for various parameters such as ground impedance, sound speed gradient, geometric divergence, atmospheric absorption, and an extended wind turbine source model [8]. These filters were previously used in a study by Gloaguen et al. [2]. Fig. 3 illustrates attenuation filters for three different distances, based on a simplified assumption and moderately favorable propagation conditions. However, these propagation assumptions are highly simplified, as in reality, propagation filters continuously vary due to factors such as atmospheric turbulence.

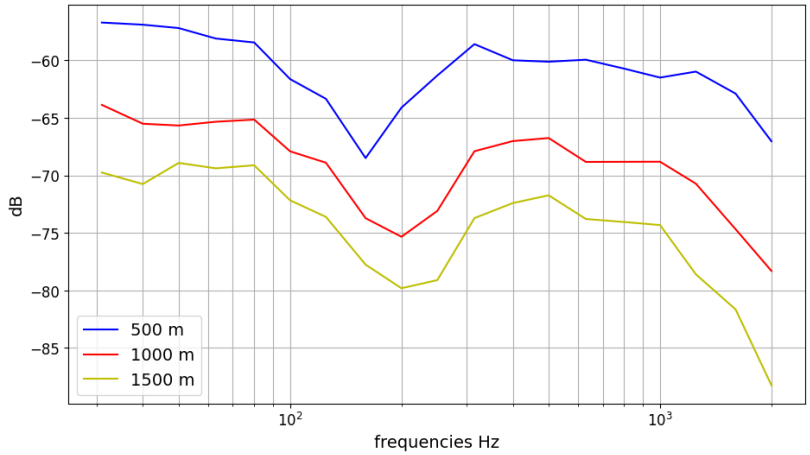

Figure 3 Attenuation filters corresponding to distances of 500 m, 1000 m, and 1500 m between the receiver (at a height of 1.5 m) and the wind turbine under moderately favorable propagation conditions.

2.3 Total noise

The total noise SPL (L_{TN}) is obtained by summing the synthesized L_{WTN} with the background noise L_{BN}, as follows:

$$L_{TN} = L_{WTN} \oplus L_{BN}, \tag{1}$$

where $\oplus$ represents the energetic summation of decibels. Finally, a dataset is produced, consisting of time series of L_{TN} and L_{WTN} for each location at 3 distances from the turbine.

2.4 Environmental data

For simplification purposes, we have defined the electrical production data and wind speed as environmental variables. The wind speed is interpolated with a temporal resolution of one second and integrated as an input into the LSTM model. The wind turbine's electrical production, which depends on wind speed, is used alongside acoustic data for model training. This integration accounts for the impact of wind speed on generated noise and considers the correlation between energy production and wind conditions. Fig. 4 presents a sample of wind speed recorded per second, derived from measurements taken every 10 minutes. It also displays the corresponding electrical production in kW over time, illustrating the correlation between wind speed variations and power generation.

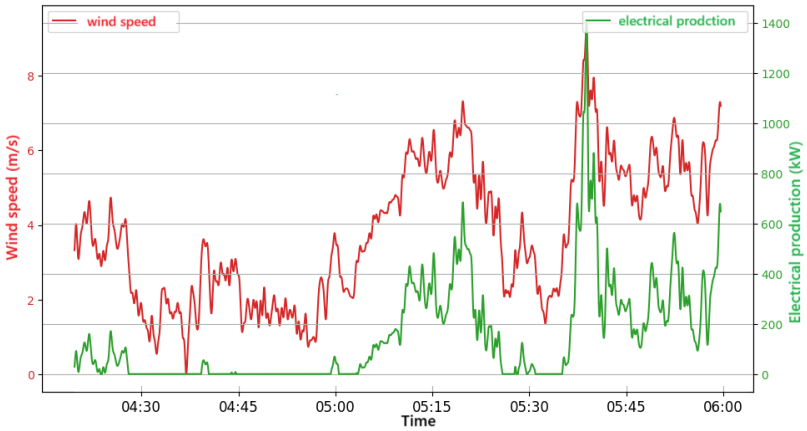

Figure 4 Time series of wind speed (in m/s, red) and electrical production (in kW, green) over a period of approximately 1.5 hours. The wind speed data is interpolated at a resolution of one second, while the electrical production corresponds to the turbine's response to varying wind speed.

3. Methodology

3.1 Data preprocessing

The features used for training include total noise (31 Hz to 2 kHz in dB(A)), electrical production, and wind speed, which have significantly different scales (0-25 m/s for wind speed and 0-2000 kW for production). Normalization is required to harmonize the amplitudes, prevent certain variables from dominating, and enhance the model's convergence, and stability.

To ensure optimal convergence and prevent the disproportionate influence of features with large dynamic ranges, Z-score normalization [9] is applied. This technique scales all features to the same range, promoting balanced and efficient learning.

The input data, denoted as $X_{i,k}$, are standardized using `Z-score` normalization to produce $X_{i,k}^{\text{std}}$:

$$X_{i,k}^{\text{std}} = \frac{X_{i,k} - \mu_k}{\sigma_k}, \tag{2}$$

where:

- $\mu_k = \frac{1}{N}\sum_{i=1}^{N} X_{i,k}$ is the mean of feature k,
- $\sigma_k = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_{i,k} - \mu_k)^2}$ is its standard deviation,
- N is the total number of samples.

Fig. 5 illustrates the impact of Z-score normalization on the wind speed feature (m/s) for our LSTM model. The original wind speed distribution (left plot) shows characteristic right-skewness (mean = 4.89 m/s). Normalization successfully transforms the data to zero mean and unit variance (right plot).

The persistence of skewness post-normalization suggests the feature retains non-Gaussian characteristics, which is inconsequential for RNNs—as they do not assume input normality—but ensures stable gradient updates during training by mitigating scale disparities across features (wind speed, electrical production, acoustic data), without altering the physically meaningful distribution shape of wind speeds. Values extending beyond $Z = +3\sigma$ (corresponding to wind speeds $>\sim$ 10 m/s in the original scale) represent high-wind events. While these could be considered statistical outliers, that the LSTM's activation functions and sequential processing can inherently handle such deviations. This preprocessing step is essential to harmonize feature scales without distorting temporal patterns.

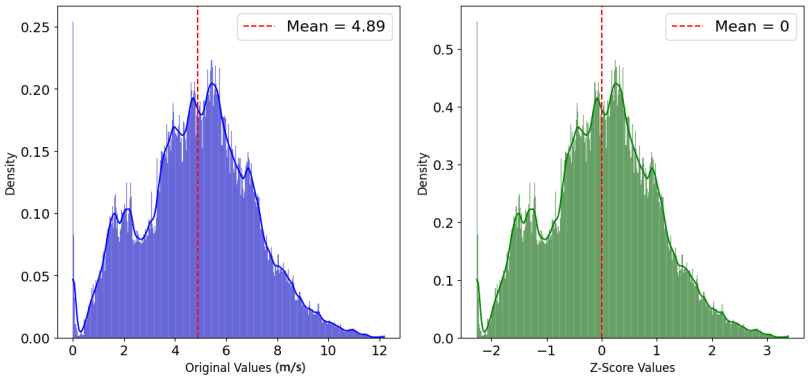

Figure 5 Density distributions of original wind speed values (in m/s) and their z-score normalized ($\mu = 0$, $\sigma = 1$).

3.2 Neural networks

Given the temporal nature of the dataset—comprising total noise, wind speed, and electrical production—where each row represents a specific time step t and the next row corresponds to the following time step $t + 1$, it is beneficial to use Recurrent Neural Networks (RNNs), such as LSTM [10].

3.2.1 LSTM cell

A standard LSTM cell includes three *gates*: the forget gate f_t which determines how much of the previous data to forget; the input gate i_t which evaluates the information to be written into the cell memory; and the output gate o_t which decides how to calculate the output from the current information, calculated from input data x_t and previous hidden state h_{t-1} by a sigmoid function, see Fig. 6.

$$\begin{aligned} i_t &= \sigma(W_i x_t + R_i h_{t-1} + b_i) \\ f_t &= \sigma(W_f x_t + R_f h_{t-1} + b_f) \\ o_t &= \sigma(W_o x_t + R_o h_{t-1} + b_o). \end{aligned} \tag{3}$$

Here, the W, R, and b variables represent the matrices and vectors of trainable parameters. The LSTM unit is defined by

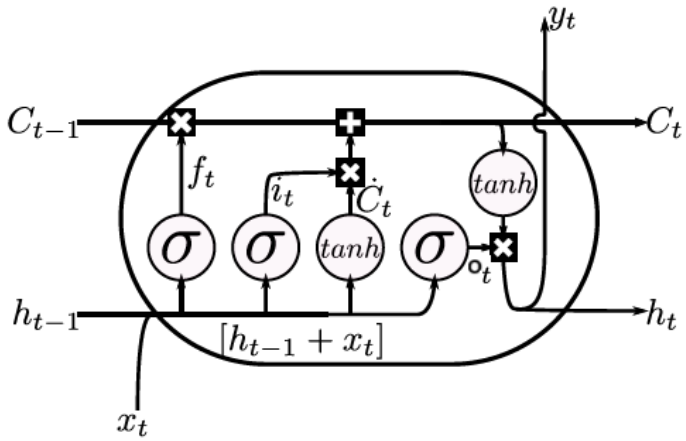

Figure 6 Illustration of a LSTM cell. The diagram represents the flow of information within the cell.

$$\begin{aligned} \dot{C}_t &= \tanh(W_c x_t + R_c h_{t-1} + b_c) \\ C_t &= f_t \odot C_{t-1} + i_t \odot \dot{C}_t \\ h_t &= o_t \odot \tanh(C_t) \\ y_t &= \sigma(W_y h_t + b_y). \end{aligned} \tag{4}$$

In words, the candidate cell state $\dot{C}_t$ is calculated using the input data x_t and the previous hidden state h_{t-1}. The cell memory or current cell state C_t is calculated using the forget gate f_t, the previous cell state C_{t-1}, the input gate i_t and the candidate cell state $\dot{C}_t$. The Hadamard product $\odot$ is simply the element-wise product of the involved matrices. The output y_t is calculated by applying the corresponding weights (W_y and b_y) to the hidden state h_t.

3.2.2 proposed architecture

In this study, we propose a regression architecture, illustrated in Fig. 7, based on a stack of LSTM layers to predict the global sound pressure level (OASPL) of WTN from the total acoustic spectrogram. The input to the model is a time-frequency representation of the total noise, computed over third-octave bands from 31 Hz to 2 kHz. Each time step contains 19 spectral features, forming a sequence that encodes the acoustic context.

The model begins with two sequential unidirectional LSTM layers, each designed to learn temporal dependencies in the acoustic data. These layers extract relevant sequential features that capture the dynamics of the wind turbine's acoustic behavior. Following the LSTM stack, a non-linear dense layer with ReLU activation is introduced to model complex relationships within the extracted temporal features.

To prevent overfitting and enhance generalization, a dropout layer with a rate of 20% is added. Finally, a custom non-linear output layer with a scaled tanh activation function is employed: $f(x) = 60 \cdot \tanh(x)$. This function constrains the output within a plausible acoustic range, centered on 0. It enables predictions up to 60 dB(A), and allows negative values approaching - 60 dB(A), representing situations with no WTN contribution, such as when the turbine is shut down.

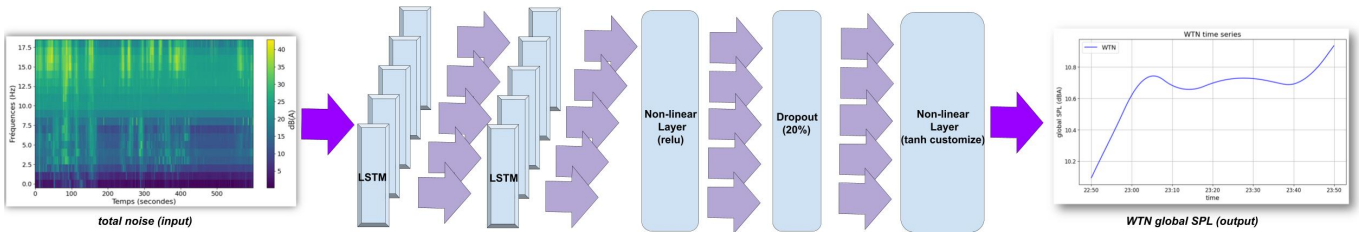

Figure 7 RNN architecture used for the estimation of the OASPL of WTN from total noise spectrograms.

3.3 Evaluation Metrics and Optimization

The Mean Absolute Error (MAE) is selected as the evaluation metric for assessing the performance of the RNN model during training. The MAE quantifies the average magnitude of the errors between the predicted particular noise level $\hat{L}_{WTN}$ and the actual value L_{WTN} in dB(A), and is defined as follows:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left| L_{WTN_i} - \hat{L}_{WTN_i} \right|. \tag{5}$$

To improve robustness against outliers (highly variable noise conditions), the Huber loss function is adopted as the objective function for training, which combines Mean Squared Error (MSE) for small residuals and Mean Absolute Error (MAE) for larger residuals:

$$L_\delta(L_{WTN}, \hat{L}_{WTN}) = \begin{cases} \frac{1}{2}(L_{WTN} - \hat{L}_{WTN})^2, & \text{for } |L_{WTN} - \hat{L}_{WTN}| \le \delta \\ \delta(|L_{WTN} - \hat{L}_{WTN}| - \frac{1}{2}\delta), & \text{otherwise} \end{cases} \tag{6}$$

where δ is a threshold controlling the transition between quadratic and linear behavior.

The derivative of the Huber loss with respect to the model predictions $\hat{L}_{WTN}$ is quadratic for small residuals and linear for larger ones. This ensures smooth and stable gradients in the presence of clean data, while limiting the influence of outliers through a reduced gradient response for large errors. Such behavior facilitates efficient and robust optimization using gradient-based algorithms. To minimize the loss function, we use the Adam optimizer, which updates the model parameters θ iteratively as follows:

$$\theta = \theta - \eta \cdot \nabla_\theta \mathcal{L}_\delta. \tag{7}$$

Here, $\nabla_\theta \mathcal{L}_\delta$ denotes the gradient of the Huber loss with respect to the model parameters θ, and η is the learning rate.

4. Results and Discussion

This section presents the experimental results obtained on the test set using two distinct architectures. The first architecture relies solely on acoustic data (19 features), while the second includes two additional variables—wind speed and power generation—bringing the total number of input features to 21.

The prediction performance of both input configurations is evaluated on the test set using the MAE, as presented in Table 1. The results highlight the impact of including environmental variables (wind speed and power generation) on the model's accuracy.

Table 1 Comparison of prediction performance (MAE) and loss function (Huber) on the test set with and without environmental data.

Test Set Configuration	Loss function dB(A)	MAE dB(A)
Without environmental data	1.86	0.72
With environmental data	0.84	0.26

The results present a significant improvement in prediction accuracy when environmental variables, wind speed and power generation, are included in the model. Both MAE and loss function show substantial reductions, from an MAE of 0.72 dB(A) to 0.26 dB(A), and from a loss of 1.86 dB(A) to 0.84 dB(A). This indicates that the addition of environmental data enhances the model's ability.

Fig. 8 compares the predicted WTN results obtained using two input configurations: with (left) and without (right) environmental data. The real WTN levels (L_{WTN}) and the predicted ones ($\hat{L}_{\mathrm{WTN}}$) are plotted alongside the real and predicted background noise levels (L_{BN} and $\hat{L}_{\mathrm{BN}}$), with the total noise (L_{TN}) in red.

The left panel shows that when wind speed and power generation are incorporated into the model, the predicted WTN signal closely aligns with the actual measurements. The predicted $\hat{L}_{\mathrm{WTN}}$ almost perfectly overlaps with the real L_{WTN}, confirming the model's ability to accurately capture WTN dynamics. This also enables precise estimation of background noise ($\hat{L}_{\mathrm{BN}}$) through decomposition from the total noise.

In contrast, the right panel shows the model performance using only acoustic features. Here, slight deviations between predicted and real WTN are visible, particularly during fluctuations in total noise. These discrepancies propagate to the estimated background noise, as seen in the divergence between real and predicted L_{BN}.

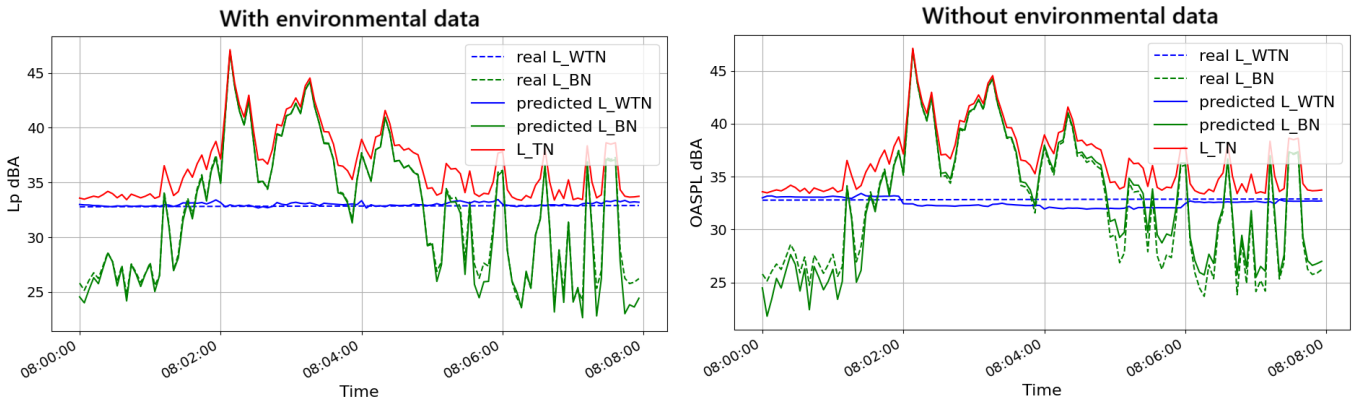

Figure 8 Comparison of predicted wind turbine noise with and without environmental data. Left: model using acoustic and environmental features. Right: model using only acoustic features.

To better visualize the impact of including environmental data, we compare the ground truth WTN OASPL with the predicted WTN using two model configurations. As shown in Fig. 9, the blue curve represents the ground truth, while the red and black curves correspond to the model predictions with and without environmental data, respectively.

The model incorporating environmental inputs (red curve) demonstrates a strong ability to follow the underlying trend of the ground truth. It closely captures both the gradual rise and the eventual plateau of the WTN signal.

In contrast, the model without environmental data (black curve) exhibits higher variance and abrupt fluctuations that deviate significantly from the actual trend. This erratic behavior—especially noticeable during the mid and late intervals—highlights the model's limited capacity to infer WTN characteristics based only on acoustic features in this test set sample, with an observed error trend of approximately ±0.5 dB(A).

This comparison reinforces the earlier observation that environmental data leads to more stable and reliable predictions of WTN.

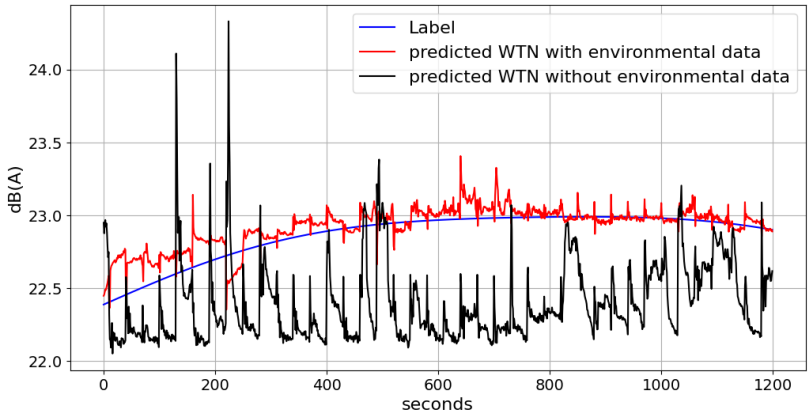

Figure 9 Comparison of WTN OASPL prediction using models with (red) and without (black) environmental data. The blue curve shows the ground truth label.

5. Conclusion

In this work, we proposed and evaluated an LSTM-based sequence-to-sequence (seq2seq) model for estimating wind turbine noise (WTN) levels from acoustic measurements, with particular attention to the role of environmental data. Our findings demonstrate the capability of recurrent neural networks (RNNs) to learn meaningful patterns from acoustic signals, yielding robust performance even under challenging conditions with high WTN levels. While both model configurations showed promising results, the inclusion of environmental variables led to significantly more accurate predictions. This performance gap highlights the importance of contextual information in capturing the temporal and spectral dynamics of WTN signals. To improve robustness, a customized output activation function was adopted along with the Huber loss, which effectively handles outliers and stabilizes training.

For future work, we aim to develop more realistic and diverse simulated WTN datasets to better train and evaluate the models. This includes incorporating additional environmental variables, generating complex and varied acoustic scenes using advanced wind turbine noise models, and leveraging more sophisticated deep learning architectures such as attention mechanisms and convolutional neural networks (CNNs). These enhancements are expected to enable a deeper investigation and comparison of existing deep learning techniques applied to WTN level estimation from acoustic input data.

References

[1] A. Freiberg, C. Schefter, M. Girbig, V. C. Murta, and A. Seidler, "Health effects of wind turbines on humans in residential settings: Results of a scoping review," *Environmental Research*, vol. 169, pp. 446–463, Feb. 2019.

[2] J.-R. Gloaguen, D. Ecotière, B. Gauvreau, A. Finez, A. Petit, and C. 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, pp. 3127–3138, Oct. 2021.

[3] P.-S. Huang, M. Kim, M. Hasegawa-Johnson, and P. Smaragdis, "Deep learning for monaural speech separation," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1562–1566, May 2014.

[4] C. Zhang, H. Zhan, Z. Hao, and X. Gao, "Classification of Complicated Urban Forest Acoustic Scenes with Deep Learning Models," *Forests*, vol. 14, p. 206, Jan. 2023.

[5] A. Gorin, N. Makhazhanov, and N. Shmyrev, *DCASE 2016 Sound Event Detection System Based on Convolutional Neural Network*. Sept. 2016.

[6] O. Anicic, D. Petković, and S. Cvetkovic, "Evaluation of wind turbine noise by soft computing methodologies: A comparative study," *Renewable and Sustainable Energy Reviews*, vol. 56, pp. 1122–1128, Apr. 2016.

[7] S. Mun, S. Shon, W. KIM, D. Han, and H. ko, "A Novel Discriminative Feature Extraction for Acoustic Scene Classification Using RNN Based Source Separation," *IEICE Transactions on Information and Systems*, vol. E100.D, pp. 3041–3044, Dec. 2017.

[8] B. Cotté, "Extended source models for wind turbine noise propagation," *The Journal of the Acoustical Society of America*, vol. 145, pp. 1363–1371, Mar. 2019.

[9] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. Elsevier, 3rd ed., June 2011.

[10] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.