DATA-DRIVEN WIND SPEED ESTIMATION USING MULTIPLE MICROPHONES

Daniele Mirabilii, Kishor Kayyar Lakshminarayana, Wolfgang Mack and Emanuël A.P. Habets

International Audio Laboratories Erlangen*, Am Wolfsmantel 33, 91058 Erlangen, Germany
{daniele.mirabilii,wolfgang.mack,emanuel.habets}@audiolabs-erlangen.de
kishor.kayyar.lakshminarayana@iis.fraunhofer.de

ABSTRACT

A deep neural network (DNN) based approach for estimating the speed of airflows using closely-spaced microphones is proposed. The spatial characteristics of wind noise measured with a small-aperture array are exploited, i.e., the low-frequency spatial coherence of wind noise signals is used as an input feature. The output is an estimate of the wind speed averaged over a specific time interval. The DNN is trained using synthetic wind noise, which overcomes the time-consuming data collection and allows to isolate wind noise from different acoustic sources. The dataset used for testing comprises wind noise measured outdoors with a circular linear array and a ground truth obtained using an ultrasonic anemometer. The obtained model is applied to generated and measured wind noise. The performance of the proposed method is assessed across a wide range of wind speeds and directions, using different time resolutions.

Index Terms— Wind noise, wind speed, multi-channel, Corcos model

1. INTRODUCTION

Resolving an airflow speed is crucial in many applications, like weather stations, monitoring systems for airfield or wind turbines, and smart home sensors. A constrained or unconstrained airstream velocity is measured with an anemometer. State-of-the-art instruments are able to exploit different physical properties to resolve the wind velocity with high accuracy. The most common are mechanical, ultrasonic, and thermal anemometers. Besides being a relatively old technology, the cup anemometer is still the most widely used instrument [1] which is easy to install and relatively cheap. Still, they present significant drawbacks like the finite response due to the start-up torque of the moving parts or the overspeeding phenomenon [2, 3]. Severe weather conditions like ice and dust can further compromise the instruments longevity.

Ultrasonic anemometers [4] present advantages over mechanical instruments, e.g., a faster response to rapid wind speed fluctuations and increased robustness towards icing. These instruments employ ultrasonic transducers. The wind speed is resolved, e.g., by computing the time interval occurring between the emission and the reception of a sound impulse. Ultrasonic anemometers are expensive, due to high-quality transducers and sophisticated signal processing circuitry.

Thermal anemometers exploit the relation between the convective heat transfer of a body and the flow velocity. They present a fine spatial resolution, but are very delicate and therefore not suitable for challenging environments. Generally, commercial anemometers suffer a non-scalable size and low integrability.

In the past decades, two main works focused on the characterization of the airflow using microphone recordings. In [5], the authors exploited a time-domain approach to determine the time difference of arrival (TDOA) of the turbulent eddies by finding the maxima in the cross-correlation functions computed between each pair of sensors in a planar three-microphone array. Taylor's frozen turbulence hypothesis was assumed, which states that the turbulent eddies induced by the airflow preserve their shape across a large distance compared to their size, resulting in a distinctive acoustic coherence. The known geometry of the microphone array allows to resolve the wind speed by simply computing the units of distance per time. Although the results were promising, the wind speed under test was light (7-9 km/h), the microphone spacing was 0.6 meters, and the cross-correlation functions were computed with 5 minutes of noise recordings, which led to a low time resolution.

Alternatively, in [6], the authors exploited a time-domain noise interferometry approach on recorded diffuse-sound fields generated by traffic noise. The acoustic measurements were obtained from an array of 4-5 microphones. The airflow velocity was subsequently obtained subtracting the speed of propagating sound waves (ca. 343 m/s) from the measured effective speed of sound. The estimation errors were in line with the accuracy of a commercial weather station. The latter approach presented two major drawbacks: the distance between the microphones was large (from 15 to 50 m), and as in [5], the average time to compute the cross correlation varied from 6 to 10 minutes. Different approaches exploit underwater acoustics and mobile or fixed hydrophones to determine the above sea-surface wind speed in oceanic environments, as in [7–10].

In this paper, we propose a frequency-domain approach to resolve the airflow velocity based on wind noise measured with a small microphone array. In particular, we exploit the low-frequency spatial correlation of wind noise contributions when measured with closely-spaced microphones. We aim at obtaining a reasonable time resolution and testing the proposed method on a wide range of wind intensities. The benefits of determining the wind speed using a small microphone array are multiple. In contrast to mechanical instruments and ultrasonic transducers, the proposed method does not involve moving parts or transmitters. A minimum requirement of three closely-spaced microphones allows the measurement instrument to be contained in size and costs, and thus highly integrable. Existing general-purpose devices (e.g., smartphones and action cameras) and weather stations can benefit from the abovementioned features. For instance, information on the wind speed could be included in metadata of action camera videos.

The proposed method consists of a deep neural network (DNN) trained with synthetic wind noise data generated using the approach presented in [11]. The training data reflects the spatial characteris-

*A joint institution of the Friedrich-Alexander-University Erlangen-Nürnberg (FAU) and Fraunhofer IIS, Germany.

tics of wind noise measured with closely-spaced microphones. The short-time Fourier transform (STFT) coefficients of the microphone signals are used to compute the pair-wise spatial coherence averaged over a defined time interval, which are subsequently provided as input to the system. The output is an estimate of the horizontal component of the wind velocity.

2. PROBLEM FORMULATION

Let us consider an array of $N \geq 2$ omnidirectional microphones. We denote by $Y_i(l,k)$ the STFT representation of the i-th microphone signal, with $i \in \{1, 2, ..., N\}$, where l and k denote the time and frequency instance index respectively. In the following, we assume the predominance of wind noise in the microphone observations, whereas other acoustic sources and the self noise of the sensors are neglected. The spatial coherence is a normalized function of the similarity between two signals in the frequency domain, which is expressed as follows

$$\gamma_{ij}(m,k) = \frac{\Phi_{ij}(m,k)}{\sqrt{\Phi_{ii}(m,k) \cdot \Phi_{jj}(m,k)}}, \quad (1)$$

where $m \in \{1, 2, 3, ...\}$ denotes the index of a block of L time samples, $\Phi_{ij}(m,k)$ denotes the cross power spectral density (CPSD) and $\Phi_{ii}(m,k)$, $\Phi_{jj}(l,k)$ denote the auto power spectral densities (APSDs) of the i-th and j-th microphones respectively. The latter quantities are estimated using the Welch's periodogram approach [12], i.e.,

$$\widehat{\Phi}_{ij}(m,k) = \frac{1}{L} \sum_{l=L(m-1)+1}^{mL} Y_i(l,k) Y_j^*(l,k), \quad (2)$$

where .* denotes the complex conjugate for the CPSD ($i \neq j$) and the APSD ($i=j$). The number of distinct pair-wise spatial coherence functions is $N(N-1)/2$. In [11], we showed that the spatial coherence of measured wind noise is different from zero when closely-spaced microphones are employed. In particular, we observed that the low-frequency correlation of microphone signals increased by decreasing the microphone distance and/or increasing the wind speed. The measured data agreed well with a fluid-mechanic model identified by Corcos [13]. The Corcos model approximates the pressure-spatial distribution of convective turbulence, expressed as a function of the similarity variables in the horizontal plane, i.e.,

$$\gamma_{ij}(m,k) = \exp\left(\frac{\omega_k d_{ij}[-\alpha(\theta_w) + \iota \cos(\theta_w)]}{U_c}\right), \quad (3)$$

where $\iota = \sqrt{-1}$, $\omega_k = 2\pi k F_s / K$ denotes the discrete angular frequency, K denotes the length of the discrete Fourier transform and F_s denotes the sampling frequency, d_{ij} denotes the relative distance between the i-th and the j-th microphones, θ_w denotes the wind direction of arrival (DOA) in terms of azimuthal angle, $\alpha(\theta_w)$ [11, 14] is a coherence-decay parameter and U_c denotes the convective speed, which is approximately $U_c \approx 0.8\ U$. In addition, a similar model was experimentally derived in [15].

It is noticeable from (3) that the time-variant spatial coherence of a microphone pair is a complex function which depends on the wind speed U and direction θ_w. The correlation between microphone signals is proportional to the wind speed. The wind direction affects the shape of (3) due to the phase term, i.e., oscillations in the real and imaginary part of (3) occur for wind stream which are not orthogonal to the microphone axis, as depicted in Fig. 1.

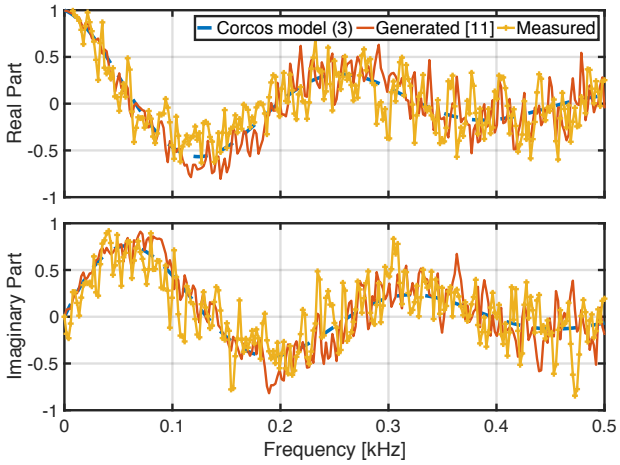

Fig. 1: Real and imaginary part of the spatial coherence approximated by the Corcos model (blue dashed line) compared to the spatial coherence of 5 seconds of wind noise generated by the approach in [11] (red solid line) and 5 seconds of wind noise measured outdoor (yellow marked line). The wind speed was U = 13 km/h and the wind direction $\theta_w = 5°$.

The objective of this work is to obtain an estimate of the wind speed U, denoted by $\hat{U}$, based on the short-term spatial coherence computed from the microphone observations. In principle, $\hat{U}$ could be obtained by solving the non-convex function in (3). Here, we propose to adopt a data-driven approach that could also be used without employing an analytical model. Moreover, by using appropriate training material, this approach could potentially be more robust against ambient noise and non-linear clipping artifacts which are not modelled by (3).

3. PROPOSED SYSTEM

A DNN-based method is proposed which provides an estimate $\widehat{U}$ of the free-field wind speed U, assuming the latter as stationary within a specific time interval, given the input feature representation.

3.1. Input and output feature representation

The model in (3) advocates for the expected value of the spatial coherence as an input feature. In particular, as wind noise is predominant in low frequencies [16], we consider a frequency range $\{0, 560\}$ Hz. We propose to map the low-frequency spatial coherence averaged over a specific time interval to the wind speed. Consequently, the input feature representation is a 3-dimensional tensor of shape $M \times \kappa \times 2$, where $M = N(N-1)/2$ is the number of pair-wise spatial coherences computed with (1) and (2), κ is the number of considered frequency bins, and 2 represents the real and imaginary components. The output is a positive real-valued estimate of the wind speed.

3.2. System Architecture

We propose to use a feed-forward neural network with five layers and rectified linear unit (ReLU) activation [17] after each layer to map the input features to the wind speed. The layers consist of 1500, 1000, 500, 250, and 1 neurons, respectively. To avoid overfitting

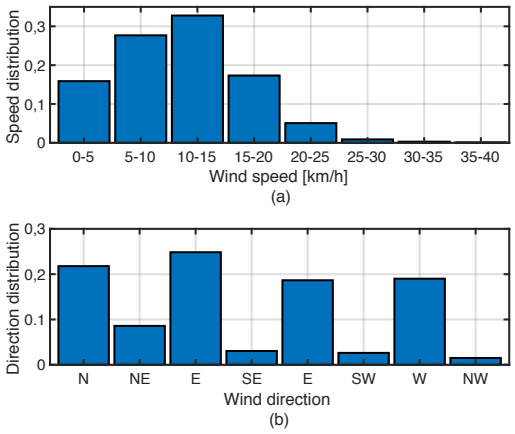

Fig. 2: Distribution of measured testing data across wind speed (a) and direction (b). The direction labels correspond to cardinal coordinates.

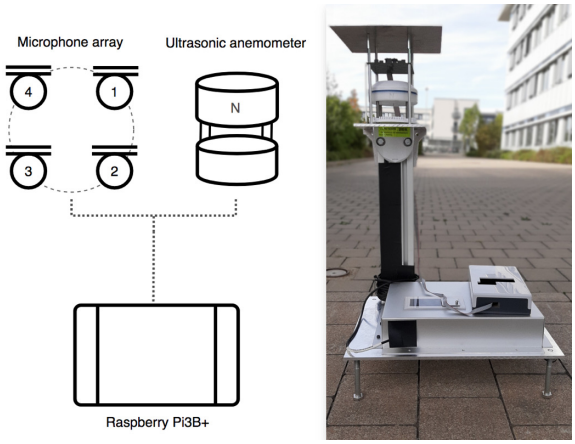

Fig. 3: System used to collect wind noise and label the recordings with the wind speed measured by the ultrasonic anemometer.

during training, dropout was applied to each layer with a probability of 50 percent. The hyperparameters, number of layers and dropout, were selected based on model performance.

4. TRAINING AND TEST DATASETS

The experimental setup consists of a circular linear array (CLA) of $N = 4$ omnidirectional MEMS microphones with a radius of 1 cm. This geometry is used for simulating wind noise data to obtain the training and the validation datasets. The same setup is employed along with an ultrasonic anemometer to collect real data and label the recordings with a ground truth of the wind speed to obtain the test dataset. We consider a sampling frequency F_s = 8 kHz, an STFT with a discrete Fourier transform length of $K = 1024$, and a Hann window of K samples with 75% of overlap.

4.1. Training and validation dataset generation

We generated synthetic multi-channel wind noise using the approach presented in [11]. The advantage of synthetic data is the exclusion of different acoustic sources that are often present in measured data while preserving the temporal, spectral and spatial characteristics of the latter. In addition, it is possible to generate a large number of samples with a uniform distribution over different wind velocities and directions while overcoming the time-consuming data collection outdoors.

The simulated wind noise is characterized by random Gaussian white noise modulated in amplitude by long-term and short-term gains to model the temporal envelope and fine structure. The modulated signals are subsequently low-pass filtered and multiplied with a mixing matrix obtained by decomposing the spatial coherence matrix given by the analytical Corcos model in (3) by means of a Cholesky factorization. The parameters used for generating data were a free-field wind speed of $U \in \{1.5, 31.5\}$ km/h with a step of 0.1 km/h and a wind direction $\theta_{\mathrm{w}} \in \{0, 359\}^{\circ}$ with a step of 1° for every wind speed value. The duration of each sample was 5 seconds. Six different realizations of wind noise samples were generated for every pair $\{U, \theta_{\mathrm{w}}\}$, obtaining in total 648000 samples of 5 seconds each. The employment of Gaussian white noise as the excitation signal allowed multiple realizations of data generated with the same parameters to be independent among each other.

The validation set was generated using the abovementioned approach with a single realization for every pair $\{U, \theta_{\mathrm{w}}\}$, obtaining 108000 samples. Figure 1 depicts the real and imaginary part of the spatial coherence for measured and generated wind noise compared to the Corcos model.

4.2. Test dataset collection and ground truth

The collection and labeling of wind noise measured outdoors were performed by employing an ad-hoc portable system. The main components were a Raspberry Pi 3B+ used as a data logger, the four microphone array and an ultrasonic anemometer ATMOS 22 by METER Group. Figure 3 depicts the employed system. The microphones were placed and aligned at 10 cm above the anemometer and 50 cm above the ground. The audio recordings were synchronized with the wind speed readings of the anemometer, obtaining a reliable ground truth. The oracle wind speed was measured every 3 seconds by the anemometer. The system was placed on the roof of the Fraunhofer IIS (Erlangen, Germany) during windy days. Diverse acoustic sources were present in the recordings, where traffic noise was the most predominant. The measured wind speed had a typical range $U \in \{0, 40\}$ km/h. The distribution of the measured data based on speed and direction is shown in Fig. 2. The total duration of the recordings was approximately 5.5 hours.

5. PERFORMANCE EVALUATION

In this section, we evaluate the results of the proposed method. We trained the model with generated data, as described in Section 4.1. The spatial coherence of the training data was computed by averaging over the entire duration of each training sample, i.e., using $L = 5 \cdot F_s$ in (2). For each of the $M = 6$ pair-wise spatial coherence values, we considered the first $\kappa = 72$ frequency bins. We trained 100 epochs and used the model with the lowest validation loss which was at epoch 20. The obtained model was subsequently applied to (a) a small test dataset of generated data which was not included in the training or validation datasets and to (b) measured data described in Section 4.2, employing different averaging time to evaluate the trade off between the estimation accuracy and the time

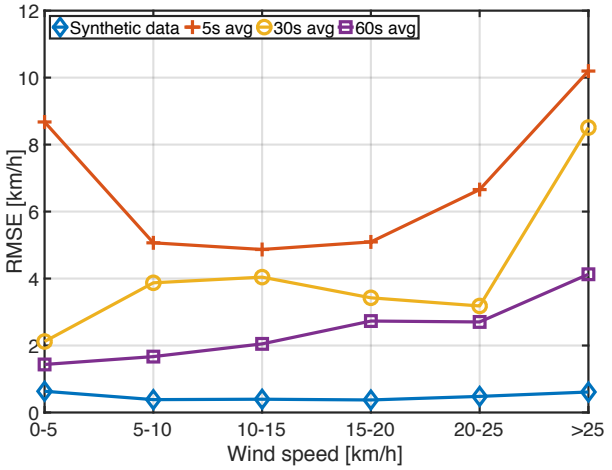

Fig. 4: Performance of the proposed method in terms of RMSE for synthetic data (blue line) averaged over 5 seconds, measured data averaged over 5 seconds (red line), 10 seconds (yellow line) and 60 seconds (purple line) across increasing wind speed.

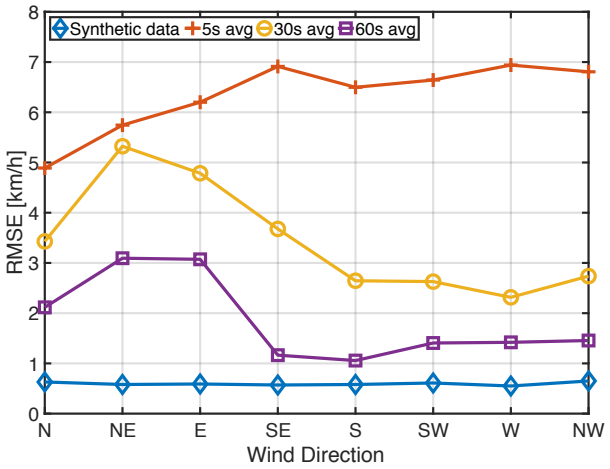

Fig. 5: Performance of the proposed method in terms of RMSE for synthetic data (blue line) averaged over 5 seconds, measured data averaged over 5 seconds (red line), 10 seconds (yellow line) and 60 seconds (purple line) across different wind directions.

resolution. To assess the performance of the proposed approach, we computed the root mean square error (RMSE) between the estimated wind speed $\widehat{U}$ and the ground truth U provided by the anemometer. The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{Q}\sum_{q=1}^{Q}\left(\widehat{U}(q) - U(q)\right)^2} \tag{4}$$

where q denotes the measurement index, and Q denotes the total number of selected measurements. We investigated the performance across different wind speed ranges and wind directions. In this respect, we classified the samples under test first according to the oracle speed and then according to the oracle direction. The speed classes were defined in steps of 5 km/h. The direction classes were defined according to cardinal directions (N, NE, E, SE, S, SW, W, NW). The RMSE was then computed within each class of speed and direction separately.

5.1. Experiment with generated data

We tested the proposed method with a separate dataset of synthetic wind noise to (a) assure that the trained model performed as expected, i.e., obtain a small RMSE with the same type of data used for the training and (b) to define an upper bound for the performance. The data was generated with uniformly distributed wind speed, obtaining in total 2400 samples of 5 seconds each. None of these samples were part of the testing or validation datasets. Each sample was characterized by a fixed pair $\{U, \theta_{\text{w}}\}$. The wind speed was sequentially incremented in steps of 0.1 km/h in the range $\{1.5, 41.5\}$ km/h. For each wind speed, eight directions were randomly selected, one for each class. The spatial coherence was computed by averaging over 5 seconds. As shown by the blue lines in Fig. 4 and Fig. 5, the proposed method resulted in RMSE values close to zero consistently throughout all speed and direction classes.

5.2. Experiment with measured data

We applied the proposed method to outdoor measurements using different time averaging to compute the spatial coherence of the microphone recordings. In particular, we used $L = t \cdot F_s$ with $t = 5, 30, 60$ seconds in (2). The results are shown in Fig. 4 and Fig. 5. The RMSE values corresponding to $t = 5$s (red lines) are considerably higher compared to those obtained with generated data. In this respect, the presence of different acoustic sources and sensor noise could disrupt the estimation accuracy when a short-term averaging is employed. Increasing t resulted in improved performance at the cost of a lower time resolution, where $t = 60$s (purple lines) presented the lowest RMSE overall. The proposed method performed the best in the range $\{0,\ 15\}$ km/h, which corresponds to most of the data in our measured test dataset. Although the performance decreased for wind speed above 25 km/h, it has to be noticed that the samples corresponding to this range represented roughly 1% of the test data. As shown in Fig. 5, the estimation performance of $t = 30, 60$ seconds was approximately uniform for wind directions between south-east and north-west, while we obtained a degradation for wind directions between north and east.

6. CONCLUSION

A DNN-based wind speed estimation approach was proposed, which exploits measurements from a small microphone array. The wind speed can be resolved based on the spatial properties of wind noise measured with closely-spaced microphones, which are well approximated by the Corcos model. Due to the dependency on the free-field airflow speed, the low-frequency spatial coherence of wind noise signals was used as an input of the DNN. The model was trained using synthetic data and tested on wind noise collected outdoors. The estimation performance increased by increasing the samples to estimate the spatial coherence from the measured microphone signals. Further investigations are required to understand the variability in terms of accuracy across different directions and to assess the performance for an *in-situ* wind speed higher than 40 km/h. Future work involves testing the proposed method in a controlled environment, e.g., a wind tunnel, improving the model in terms of robustness towards different acoustic sources, and increasing the time resolution while preserving a reasonable estimation accuracy.

7. REFERENCES

[1] S. Pindado, E. Vega, A. Martínez, E. Meseguer, S. Franchini, and I. P. Sarasola, "Analysis of calibration results from cup and propeller anemometers. influence on wind turbine annual energy production (AEP) calculations," *Wind Energy*, vol. 14, no. 1, pp. 119–132, 2011.

[2] N. E. Busch and L. Kristensen, "Cup anemometer overspeeding," *Journal of Applied Meteorology*, vol. 15, no. 12, pp. 1328–1332, 1976.

[3] D. Westermann, "Overspeeding measurements of cup anemometers compared to a simple numerical model," *Deutsches Windenergie-Institut gemeinnutzige GmbH*, 1996.

[4] S. K. Ammann, "Ultrasonic anemometer," Sept. 6 1994. US Patent 5,343,744.

[5] H. E. Bass, R. Raspet, and J. O. Messer, "Experimental determination of wind speed and direction using a three microphone array," *J. Acoust. Soc. Am.*, vol. 97, no. 1, pp. 695–696, 1995.

[6] O. A. Godin, V. G. Irisov, and M. I. Charnotskii, "Passive acoustic measurements of wind velocity and sound speed in air," *J. Acoust. Soc. Am.*, vol. 135, no. 2, pp. EL68–EL74, 2014.

[7] S. Pensieri, R. Bozzano, J. A. Nystuen, E. N. Anagnostou, M. N. Anagnostou, and R. Bechini, "Underwater acoustic measurements to estimate wind and rainfall in the mediterranean sea," *Advances in Meteorology*, vol. 2015, 2015.

[8] P. Xiao, Z.-x. Lei, *et al.*, "Sequential filtering for surface wind speed estimation from ambient noise measurement," *China Ocean Engineering*, vol. 31, no. 1, pp. 74–78, 2017.

[9] D. Cazau, J. Bonnel, and M. Baumgartner, "Wind speed estimation using acoustic underwater glider in a near-shore marine environment," *IEEE Geosci. Remote Sens. Lett.*, vol. 57, no. 4, pp. 2097–2106, 2018.

[10] P. Cauchy, K. J. Heywood, N. D. Merchant, B. Y. Queste, and P. Testor, "Wind speed measured from underwater gliders using passive acoustics," *Journal of Atmospheric and Oceanic Technology*, vol. 35, no. 12, pp. 2305–2321, 2018.

[11] D. Mirabilii and E. A. P. Habets, "Simulating multi-channel wind noise based on the Corcos model," in *Proc. of the International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2018.

[12] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, 1967.

[13] G. Corcos, "The structure of the turbulent pressure field in boundary-layer flows," *Journal of Fluid Mechanics*, vol. 18, no. 3, pp. 353–378, 1964.

[14] R. H. Mellen, "On modeling convective turbulence," *J. Acoust. Soc. Am.*, vol. 88, no. 6, pp. 2891–2893, 1990.

[15] F. Douglas Shields, "Low-frequency wind noise correlation in microphone arrays," *J. Acoust. Soc. Am.*, vol. 117, no. 6, pp. 3489–3496, 2005.

[16] C. M. Nelke and P. Vary, "Measurement, analysis and simulation of wind noise signals for mobile communication devices," in *Proc. of the International Workshop on Acoustic Signal Enhancement (IWAENC)*, 2014.

[17] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807–814, 2010.