Pseudo-spectrum based methods for estimating the wind speed and direction based on closely spaced microphone signals FREE

Daniele Mirabilii ; Emanuel A. P. Habets

Check for updates

CrossMark

View Online

Export Citation

Articles You May Be Interested In

Wind speed and direction estimation based on the spatial coherence of closely spaced microphones

J. Acoust. Soc. Am. (January 2022)

Generating coherence-constrained multisensor signals using balanced mixing and spectrally smooth filters

J. Acoust. Soc. Am. (March 2021)

Pseudospectrum-based methods for estimating the wind speed and direction based on closely spaced microphone signals

J Acoust Soc Am (March 2023)

Proceedings of Meetings on Acoustics

184th Meeting of the Acoustical Society of America

Chicago, Illinois

8-12 May 2023

Computational Acoustics: Paper 2pCA2

Pseudo-spectrum based methods for estimating the wind speed and direction based on closely spaced microphone signals

Daniele Mirabilii
Department of Signal Processing, WSAudiology, Erlangen, Bavaria, 91058, GERMANY;
danielemirabilii@gmail.com

Emanuel A. P. Habets
International Audio Laboratories Erlangen, Erlangen, Bavaria, 91058, GERMANY;
emanuel.habets@audiolabs-erlangen.de

Acoustic array processing can be employed to measure the wind speed and direction based on microphone signals. Turbulent pressure fluctuations picked up by microphones are referred to as wind noise. According to Taylor's frozen turbulence hypothesis, turbulent eddies retain their shape while advecting at nearly the mean wind speed and in the wind direction. Consequently, wind noise propagates across a microphone array when the inter-microphone distance is smaller than the turbulence wavelength. This property can be exploited to track the orientation of the turbulence advection and hence to characterize the wind flow. We propose beamforming and signal subspace-based methods to estimate the wind speed and direction using a compact microphone array. In particular, the pseudo-spectrum of measured wind noise is computed against candidate pairs of wind speed and direction. The wind speed and direction estimates are then obtained as the maximizers of the pseudo-spectrum. In addition, we extend an existing time difference of arrival-based method originally derived for three microphones to an arbitrary number of microphones. We evaluate the estimation accuracy of the proposed methods separately for the wind speed and direction. Possible applications include highly integrable, portable, and inexpensive anemometers for smart sensors or action cameras.

1. INTRODUCTION

Atmospheric wind measurement is crucial in various fields of application, like wind energy, aviation, and meteorology. Mechanical or ultrasonic anemometers are the industry standard to measure wind speed and direction, referred to as wind velocity in the following. Alternatively, recent studies demonstrated that pressure sensors like microphones and hydrophones could also be employed for wind measurements. In acoustical ocean meteorology, numerical models are commonly identified that describe the relation between underwater acoustics and above-surface phenomena (e.g., rainfall or wind speed) using ambient noise measured with static or mobile hydrophones. For example, specific models[4, 5, 15, 20] were identified to delineate the relationship between the underwater sound pressure level and the above-surface wind speed.

Other studies used microphone signals to estimate the *in-situ* wind velocity. Wind noise in microphones[16, 17] is mainly caused by pressure fluctuations induced by air turbulence. The atmospheric wind is inherently turbulent and represents the primary source of wind noise outdoors. Bass, Raspet, and Messer[2] proposed a method to experimentally determine the wind speed and direction by estimating the time difference of arrival (TDOA) of the turbulent eddies traveling across a planar three-microphone array. Based on Taylor's frozen turbulence hypothesis, the authors assumed that turbulent eddies propagate in the wind direction and at the wind speed, and that their shape is retained across a distance that is large compared to their size. The TDOA between each microphone pair was estimated as the time lag corresponding to the peaks in the cross-correlation computed over 5 minutes. The estimated TDOAs and the microphone locations were then used to determine the wind speed and direction. The microphones were arranged in an equilateral triangle with an inter-microphone distance of 0.61 m at 0.6 m above the ground. The wind speed ground truth was measured with a hot-wire anemometer in the range 2-3 ms^{-1}, while the direction ground truth was simply observed and not measured during the experiment.

Wilson and White[19] used wavelet-based processing to compare the characteristics of outdoor wind noise with propagating acoustic signals. The analysis of data from a 49-microphone array with an inter-microphone distance of 0.152 m showed that wind noise consists of non-planar and weakly correlated microgusts, which are caused by turbulent structures that propagate across the array. By applying the MUltiple SIgnal CLassification (MUSIC) method in the continuous wavelet transform domain, the authors could determine the approximate phase speed and orientation of the microgusts, which were observed to be close to the mean wind speed and direction. However, the MUSIC accuracy in estimating the wind velocity was not extensively evaluated against the ground truth values.

In our early works,[13, 14] we exploited the spatial coherence of wind noise signals measured with a compact microphone array to infer the wind velocity. These works were motivated by the observation made by the present authors[11, 12] that the wind noise spatial coherence agreed well with the Corcos model,[6] which depends on airflow velocity. Fitting the measured coherence to the analytical model thereby provided wind speed and direction estimates. In the first work,[14] the measured spatial coherence was used to estimate the wind speed using a feed-forward neural network. The network was trained using synthetic noise exhibiting a spatial coherence as predicted by the Corcos model. In the second work,[13] we fitted the measured coherence to the Corcos model in the least squares sense with candidate speed and direction pairs. We then selected the pair that yielded the smallest squared residuals with the measured coherence as speed and direction estimates.

In this work, we propose three methods based on pseudo-spectrum maximization to estimate the wind speed and direction using an array of closely spaced microphones. In particular, the spatial pseudo-spectrum is computed based on the wind noise covariance matrix averaged over a finite time interval and evaluated against candidate pairs of speed and direction. The pair corresponding to the pseudo-spectrum maximum is then selected as wind speed and direction estimates. We describe two beamforming methods, namely the Delay and Sum (DS) and the Minimum Power Distortionless Response (MPDR), and the signal-subspace method MUSIC. In addition, we extend the TDOA-based method presented by Bass, Raspet, and Messer[2] to an arbitrary microphone array geometry, i.e., an arbitrary number of sensors and microphone locations. Differently from Wilson and White,[19] we evaluate MUSIC in the Short-Time Fourier Transform (STFT) domain using a microphone array with an inter-microphone distance approximately ten times smaller than the array used in their experiment. Due to the rapid dissipation of turbulent eddies across space, an inter-microphone distance smaller than the turbulence wavelength is required to correctly resolve the wind noise phase speed and orientation. In this case, coherent turbulent structures induce direction-dependent pressure fluctuations across the array. Assuming that such pressure signals propagate in a rectilinear path at nearly the wind speed and in the wind direction, the pseudo-spectrum based methods are suitable to estimate the wind velocity. We employ at least three sensors in a planar configuration to avoid the undetermined problem of a zero phase shift and ambiguity for directions symmetric to the microphone axis. The four methods are evaluated using wind noise collected outdoors

with a 4-MEMS circular microphone array with a radius of 0.01 m. The estimates are compared to the wind speed and direction ground truth measured with an ultrasonic anemometer and synchronized with the audio recordings using a single-board computer.

The remainder of the paper is organized as follows. In Section 2, we introduce the notation and the signal model, and we formulate the problem of estimating the wind velocity using microphone signals. In Section 3, we extend the existing TDOA-based[2] method to an arbitrary microphone array geometry. In Section 4, we propose the three pseudo-spectrum based methods. In Section 5, we evaluate the accuracy of the TDOA- and pseudo-spectrum based methods in estimating the wind speed and direction using a circular microphone array by comparing the estimated values with the ground truth collected with an ultrasonic anemometer.

2.   SIGNAL MODEL AND PROBLEM FORMULATION

Let us assume N closely spaced microphones that are spatially distributed on a plane and exposed to a wind flow characterized by a given speed and direction. For simplicity, we consider wind velocities with no vertical component, i.e., for which the pressure fluctuations propagate without any elevation with respect to the microphone array plane. We model the microphone signals as follows

$$y_i(t) = w_i(t) + n_i(t), \tag{1}$$

where $y_i(t)$ denotes the i-th microphone signal, $w_i(t)$ denotes the turbulent pressure signal at the i-th microphone, referred to as wind noise in the following, $n_i(t)$ denotes an additive stationary noise term, e.g., microphone self-noise, and t denotes the discrete-time sample.

Alternatively, we can write Eq. (1) in the Short-time Fourier Transform (STFT) domain as

$$Y_i(l, k) = W_i(l, k) + N_i(l, k), \tag{2}$$

where l and k denote the time and frequency indices, respectively. We assume mutual statistical independence between the stationary noise terms, i.e., $\mathrm{E}\{N_i N_j\} = 0$ with $i \neq j$, and between the wind noise and the stationary noise terms $\mathrm{E}\{W_i N_j\} = 0$ for all i, j, where $\mathrm{E}\{.\}$ denotes mathematical expectation. Furthermore, we assume homogeneous wind noise and stationary noise contributions at the microphones, i.e., equal wind noise power $\mathrm{E}\{|W_i|^2\} = \phi_w$ and stationary noise power $\mathrm{E}\{|N_i|^2\} = \phi_n$ for all i.

We model the wind noise propagation as a plane wave traveling with a phase speed approximately corresponding to the wind speed u and in the wind direction θ, i.e., with the propagation vector $\mathbf{u} = \nu \cdot \bar{\mathbf{u}}$, where $\bar{\mathbf{u}} \in \mathbb{R}^2$ denotes the unit-norm vector oriented in the wind direction and $\nu = \|\mathbf{u}\| = 1/u$, i.e., the magnitude of $\mathbf{u}$ is inversely proportional to the wind speed. The position vector of the i-th microphone is denoted by $\mathbf{r}_i$, with $i \in \{1, 2, \ldots, N\}$, so that $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$, with $j \in \{1, 2, \ldots, N\}$ and $i \neq j$. The distance between the i-th and the j-th microphone is given by $d_{ij} = \|\mathbf{r}_{ij}\|$. Without loss of generality, we use the first microphone as a reference and use its position as the origin of our coordinate system so that $\mathbf{r}_i = \mathbf{r}_i - \mathbf{r}_1$ for $i \neq 1$, and $d_i = \|\mathbf{r}_i\|$. To account for the turbulence dissipation across space and frequency, we introduce an exponential decay factor.[6,9]

To model the inter-microphone relations more explicitly, we rewrite Eq. (2) as follows

$$Y_i(l, k) = G_i(k)W(l, k) + N_i(l, k), \tag{3}$$

where $G_i(k)$ denotes the relative transfer function modeling the wind noise propagation between the i-th and the reference microphone. Based on the aforementioned model description, we can express the relative transfer function as

$$G_i(k) = \exp\left(-\alpha(\mathbf{u}, \mathbf{r}_i)\,\omega_k\right)\,\exp\left(\iota\omega_k \mathbf{u} \cdot \mathbf{r}_i\right) \tag{4}$$

where $\alpha(\mathbf{u}, \mathbf{r}_i)$ denotes a decay parameter, which depends on the microphone distance and the wind propagation vector, $\omega_k = 2\pi F_s k/K$ denotes the discrete angular frequency, F_s denotes the sampling frequency, K denotes the discrete Fourier Transform (DFT) length, and $\iota = \sqrt{-1}$. The term $\mathbf{u} \cdot \mathbf{r}_i$ corresponds to the TDOA of the wind noise propagating between the i-th and the reference microphone

$$\tau_i = \mathbf{u} \cdot \mathbf{r}_i = \frac{d_i \cos(\theta_i)}{u}, \tag{5}$$

where θ_i denotes the wind direction with respect to the i-th and reference microphone axis, and u denotes the wind speed in ms^{-1}. The term $-\alpha(\mathbf{u}, \mathbf{r}_i)$ relative to the exponential decay delineates the weakly coherent nature of wind noise, and can be modeled according to, e.g., turbulent boundary layer stochastic spatial pressure distribution models.[3,6,7,9] In our previous works,[13,14] the decay parameter was defined as

$$\alpha(\mathbf{u}, \mathbf{r}_i) = \alpha_1 |\mathbf{u} \cdot \mathbf{r}_i| + \alpha_2 |\mathbf{u}_\perp \cdot \mathbf{r}_i|, \tag{6}$$

where $\mathbf{u}_\perp = \nu \cdot \bar{\mathbf{u}}_\perp$ with $\bar{\mathbf{u}}_\perp$ denoting the unit-norm vector perpendicular to the airflow direction. As the scalar products $\mathbf{u} \cdot \mathbf{r}_{ij}$ and $\mathbf{u}_\perp \cdot \mathbf{r}_i$ depict the relative displacement of the i-th and reference microphone in the streamwise (parallel to the airflow) and spanwise (orthogonal to the airflow) directions, respectively, Eq. (6) indicates anisotropic attenuation in magnitude in the streamwise and spanwise directions.

At sufficiently low frequency and for a sufficiently small microphone distance $\exp(-\alpha(\mathbf{u}, \mathbf{r}_i)\omega_k) \approx 1$. In this case, Eq. (3) corresponds to a directional source propagating at a variable speed u in the direction θ, i.e.,

$$G_i(k) \approx \exp(\iota \omega_k \mathbf{u} \cdot \mathbf{r}_i). \tag{7}$$

Finally, we express the signal model in Eq. (3) for all microphone signals in a vector notation

$$\mathbf{y}(l, k) = \mathbf{g}(k) W(l, k) + \mathbf{n}(l, k) \tag{8}$$

where

$$\mathbf{y}(l, k) = [Y_1(l, k), Y_2(l, k), \ldots, Y_N(l, k)]^\mathrm{T}$$
$$\mathbf{g}(k) = [1, G_2(k), G_3(k), \ldots, G_N(k)]^\mathrm{T}$$
$$\mathbf{n}(l, k) = [N_1(l, k), N_2(l, k), \ldots, N_N(l, k)]^\mathrm{T}.$$

The objective of this work is to estimate the wind speed u and direction θ (computed with respect to a reference axis) given the microphone observations $\mathbf{y}(l, k)$ collected with closely spaced microphones.

3. TDOA-BASED METHOD FOR ARBITRARY MICROPHONE GEOMETRIES

The method proposed by Bass, Raspet, and Messer[2] was formulated for a three-microphone array arranged in an equilateral triangle with a given inter-microphone distance of 0.61 m. The estimates were obtained specifically for this array geometry. In this section, the TDOA-based method is generalized for an arbitrary microphone geometry, i.e., for an arbitrary number of microphones N and arbitrary microphone locations $\mathbf{r}_i$, with $i \in \{1, 2, \ldots, N\}$. By employing a so-called TDOA spherical set,[1] i.e., TDOAs between the i-th and a reference microphone, Eq. (5) identifies an overdetermined linear system of N-1 equations in two variables, namely the wind speed and direction. By rearranging the terms of Eq. (5) in a matrix form, we obtain the linear system

$$\mathbf{A}\mathbf{u} = \mathbf{b}. \tag{9}$$

where

$$\mathbf{A} = [\mathbf{r}_2, \mathbf{r}_3, \ldots, \mathbf{r}_N]^\mathrm{T},$$
$$\mathbf{b} = [\tau_2, \tau_3, \ldots, \tau_N]^\mathrm{T}.$$

The least-squares solution of $\mathbf{u}$ is given by the product of the pseudoinverse of $\mathbf{A}$ and $\mathbf{b}$

$$\hat{\mathbf{u}} = \mathbf{A}^\dagger \mathbf{b} = (\mathbf{A}^\mathrm{T} \mathbf{A})^{-1} \mathbf{A}^\mathrm{T} \mathbf{b}. \tag{10}$$

The estimated wind propagation unit vector embedding the wind direction is obtained as

$$\hat{\bar{\mathbf{u}}} = \frac{\hat{\mathbf{u}}}{\|\hat{\mathbf{u}}\|_2}, \tag{11}$$

and the estimated wind speed as

$$\hat{u} = \frac{1}{\|\hat{\mathbf{u}}\|_2}. \tag{12}$$

In practice, only the matrix $\mathbf{A}$ containing the microphone locations in Eq. (9) is known, and the TDOAs vector $\mathbf{b}$ must be estimated from the microphone signals. As in the method proposed by Bass, Raspet, and Messer,[2] the TDOA estimation can be performed via cross-correlation maximization, i.e., by identifying the time lag corresponding to the cross-correlation peak between two microphone signals as

$$\hat{\tau}_i = \operatorname*{argmax}_{\tau} R_i(\tau), \tag{13}$$

where R_i denotes the cross-correlation between the reference and the i-th microphone signal. To overcome the impact of adverse acoustic conditions on the TDOA estimation, the Generalized Cross-correlation (GCC)[8] method is frequently used. The GCC computes the inverse Fourier transform of a weighted cross-power spectrum, where different weighting functions can be employed. One commonly used weighting is the Phase Transform (GCC-PHAT), which disregards the magnitude of the cross-power spectrum, i.e.,

$$R_i(\tau) = \mathcal{F}^{-1}\{\Psi_i(l,k)\}, \tag{14}$$

where $\mathcal{F}^{-1}\{\cdot\}$ denotes the inverse discrete Fourier transform and

$$\Psi_i(l,k) = \frac{Y_i(l,k)Y_1^*(l,k)}{|Y_i(l,k)Y_1^*(l,k)|}. \tag{15}$$

where $(\cdot)^*$ denotes the complex conjugate.

4. PROPOSED PSEUDO-SPECTRUM BASED METHODS

In this section, we propose three methods based on pseudo-spectrum maximization to estimate the wind speed and direction using wind noise signals captured with a compact microphone array. In particular, the spatial pseudo-spectrum is computed from the microphone signals covariance matrix and evaluated at candidate wind speed and direction pairs. The pair that yields the pseudo-spectrum maximum is selected as the estimates of wind speed and direction, i.e.,

$$\left[\hat{u}, \hat{\theta}\right] = \operatorname*{argmax}_{u,\theta} P(k; u, \theta), \tag{16}$$

where $P(k; u, \theta)$ denotes the pseudo-spectrum evaluated within a suitable range of wind velocities. The pseudo-spectrum based methods require the computation of the signal covariance matrix and the definition of the steering vectors. The signal covariance matrix is defined as

$$\mathbf{\Phi_y}(l,k) = \mathrm{E}\{\mathbf{y}(l,k)\mathbf{y}^{\mathrm{H}}(l,k)\}, \tag{17}$$

where $(\cdot)^{\mathrm{H}}$ denotes the Hermitian operator. In practice, the mathematical expectation can be approximated by the sample covariance matrix, i.e., by averaging the instantaneous power-spectrum matrix $\mathbf{y}(l,k)\mathbf{y}^{\mathrm{H}}(l,k)$ over a specific time interval

$$\widehat{\mathbf{\Phi}}_\mathbf{y}(k) = \frac{1}{L}\sum_{l=1}^{L}\mathbf{y}(l,k)\mathbf{y}^{\mathrm{H}}(l,k), \tag{18}$$

where L is the total number of considered time frames. The pseudo-spectrum is evaluated based on candidate steering vectors containing the relative transfer functions between the microphone locations. As described in Section 1 in case of wind noise propagating across a compact microphone array, we can model the relative transfer functions at sufficiently low frequencies with pure delays, i.e.,

$$G_i(k; u, \theta) = \exp\left(\iota\omega_k \tau_i(u, \theta)\right), \tag{19}$$

where $\tau_i(u, \theta)$ is the candidate TDOA that depends on the wind speed and direction. The steering vector is defined based on relative transfer functions as

$$\mathbf{g}(k; u, \theta) = [1, G_2(k; u, \theta), G_3(k; u, \theta), \ldots, G_N(k; u, \theta)]^{\mathrm{T}}. \tag{20}$$

Firstly, we introduce two spatial filtering approaches, the DS and the MPDR beamformers, that have been adapted to accommodate variable propagation velocities. For these approaches, we make no assumption on the microphone

signal covariance structure, and we linearly combine the microphone signals to enhance coherent propagating sources traveling in a specific direction and at a specific speed while minimizing uncorrelated noise sources or directional sources impinging from different directions and traveling at a different speed. This is accomplished by applying complex weights to the microphone signals., i.e., $Z(l, k) = \mathbf{h}^{\mathrm{H}}(k)\mathbf{y}(l, k)$. The pseudo-spectrum is then computed as the beamformer output power. Secondly, we describe the MUSIC algorithm, where we assume that the signal covariance structure can be factorized into a directional source subspace and a noise subspace and exploit the orthogonality of the steering vectors and the noise subspace basis.

A.   DELAY AND SUM BEAMFORMER

The DS beamformer aims to properly delay the microphone signals by a time corresponding to the expected time delay of the wind noise propagating across the microphones from a specific direction at a specific speed. The delayed signals are then summed to obtain the output signal of the beamformer. If the delay is correctly chosen, such that the signals are in phase, their sum yields the maximum output power. This way, the wind noise signals add coherently while the mutually uncorrelated noise is attenuated. The beamformer output is obtained by selecting an appropriate phase shift identified by a candidate pair $[u, \theta]$

$$Z_{\mathrm{DS}}(k; u, \theta) = \frac{1}{N} \sum_{i=1}^{N} \exp\left(-\iota\omega_k\tau_i(k; u, \theta)\right) Y_i(k) = \mathbf{h}^{\mathrm{H}}(k; u, \theta)\mathbf{y}(k), \tag{21}$$

where the dependency on the time frame was omitted for brevity. The complex weights of the DS beamformer are given by

$$\mathbf{h}(k; u, \theta) = \frac{1}{N}\,\mathbf{g}(k; u, \theta). \tag{22}$$

The DS beamformer output power is computed as the expectation of the magnitude squared value of Eq. (21), i.e., $\mathrm{E}\{|Z_{\mathrm{DS}}(k; u, \theta)|^2\} = \mathrm{E}\{|\mathbf{h}^{\mathrm{H}}(k; u, \theta)\mathbf{y}(k)|^2\} = \mathbf{h}^{\mathrm{H}}(k; u, \theta)\widehat{\mathbf{\Phi}}_y(k)\mathbf{h}(k; u, \theta)$, yielding the pseudo-spectrum

$$P_{\mathrm{DS}}(k; u, \theta) = \frac{\mathbf{g}^{\mathrm{H}}(k; u, \theta)\widehat{\mathbf{\Phi}}_y(k)\mathbf{g}(k; u, \theta)}{N^2}. \tag{23}$$

B.   MINIMUM POWER DISTORTIONLESS RESPONSE BEAMFORMER

The MPDR beamformer[18] aims to minimize the total power of the signals received by the microphones subject to the constraint that the beamformer response in the direction of the desired signal is unity, i.e.,

$$\mathbf{h}^{\mathrm{H}}(k; u, \theta)\mathbf{g}(k; u, \theta) = 1. \tag{24}$$

The idea is to minimize the energy of interfering sources and noise while keeping a distortionless response for signals traveling from a specific direction at a specific speed $[u, \theta]$. The optimum solution for the MPDR complex weights is given by[18]

$$\mathbf{h}(k; u, \theta) = \frac{\widehat{\mathbf{\Phi}}_{\mathbf{y}}^{-1}(l, k)\mathbf{g}(k; u, \theta)}{\mathbf{g}^{\mathrm{H}}(k; u, \theta)\widehat{\mathbf{\Phi}}_{\mathbf{y}}^{-1}(l, k)\mathbf{g}(k; u, \theta)}. \tag{25}$$

The beamformer output power is obtained as

$$P_{\mathrm{MPDR}}(k; u, \theta) = \frac{1}{\mathbf{g}^{\mathrm{H}}(k; u, \theta)\widehat{\mathbf{\Phi}}_{\mathbf{y}}^{-1}(l, k)\mathbf{g}(k; u, \theta)}. \tag{26}$$

C.   MULTIPLE SIGNAL CLASSIFICATION

The microphone signal covariance matrix in Eq. (17) can be defined as a superposition of the propagating wind noise covariance matrix $\mathbf{\Phi}_w(l, k)$ and the stationary noise covariance matrix $\mathbf{\Phi}_n(l, k)$, i.e.,

$$\mathbf{\Phi}_{\mathbf{y}}(l, k) = \mathbf{\Phi}_w(l, k) + \mathbf{\Phi}_n(l, k) = \phi_w(l, k)\mathbf{g}(k)\mathbf{g}^{\mathrm{H}}(k) + \phi_n(l, k)\mathbf{I}, \tag{27}$$

where $\mathbf{I}$ denotes an $N \times N$ identity matrix. The covariance matrix $\mathbf{\Phi}_{\mathbf{y}}$ can be factorized into a signal subspace and a noise subspace by computing an eigenvalue decomposition. The decomposition yields a matrix $\mathbf{Q}$ containing the
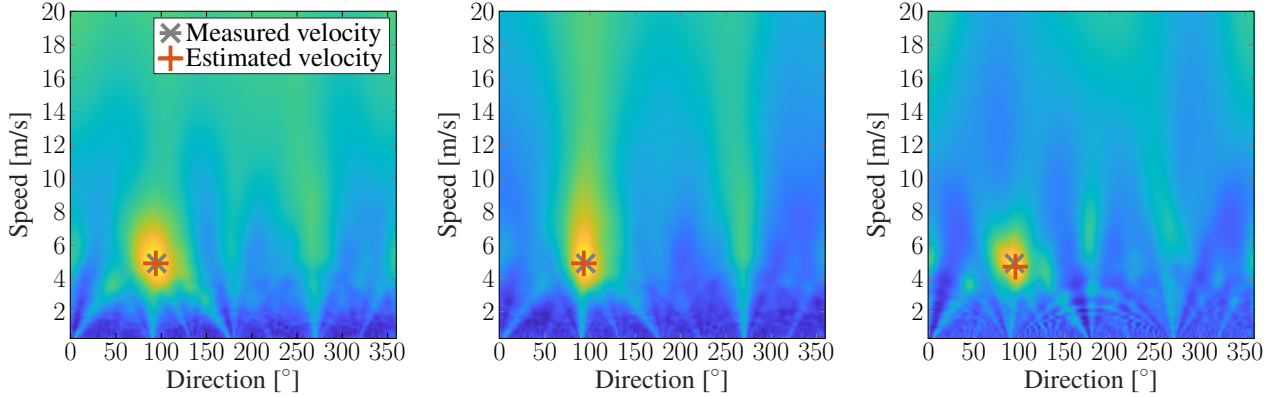
Figure 1: Pseudo-spectra obtained with the DS beamformer (left), the MPDR beamformer (center), and MUSIC (right). The measured wind velocity is depicted with a grey cross (4.9 ms^{-1}, 95°), and the estimated wind velocity with a red plus marker. The pseudo-spectra are normalized, where yellow indicates values close to one and blue values close to zero.

(column-wise) eigenvectors $[\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N]$ and a diagonal matrix $\mathbf{\Lambda}$ containing the eigenvalues (here assumed to be in descending order) associated with the eigenvectors such that

$$\mathbf{\Phi}_\mathbf{y} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\mathrm{H} = \lambda_1 \mathbf{q}_1 \mathbf{q}_1^\mathrm{H} + \phi_n \mathbf{V}\mathbf{V}^\mathrm{H}, \tag{28}$$

where λ_1 is the eigenvalue associated with $\mathbf{q}_1$ and the columns of the $N \times (N-1)$ matrix $\mathbf{V} = [\mathbf{q}_2, \ldots, \mathbf{q}_N]$ span the stationary noise subspace.

MUSIC exploits the orthogonality of the steering vector, which is here related to the wind noise propagation characteristics (speed and direction), and the eigenvectors spanning the noise subspace, i.e., $\mathbf{g}^\mathrm{H}\mathbf{V} = \mathbf{0}$ where $\mathbf{0}$ is a $1 \times (N-1)$ matrix of zeros. Therefore, MUSIC projects the stationary noise subspace onto a grid of possible speeds and directions, and finds the speed and direction at which the projection is the smallest, i.e., $\mathbf{g}^\mathrm{H}\mathbf{V}\mathbf{V}^\mathrm{H}\mathbf{g} \approx 0$. The MUSIC pseudo-spectrum is defined as

$$P_{\mathrm{MUSIC}}(k; u, \theta) = \frac{1}{\mathbf{g}^\mathrm{H}(k; u, \theta)\widehat{\mathbf{V}}(k)\widehat{\mathbf{V}}^\mathrm{H}(k)\mathbf{g}(k; u, \theta)}, \tag{29}$$

where $\widehat{\mathbf{V}}(k)$ contains the eigenvectors related to the stationary noise subspace obtained via eigenvalue decomposition of the estimated covariance matrix in Eq. (18).

Figure 1 shows the DS, MPDR, and MUSIC normalized pseudo-spectra obtained with 30 seconds of microphone signals captured with a 4-MEMS microphone array exposed to a wind flow at 4.9 ms^{-1} and 95° with respect to the true north. The estimated velocity (red plus) corresponds to the pseudo-spectra maxima, and the velocity measured by an ultrasonic anemometer is depicted with a grey cross.

5.   PERFORMANCE EVALUATION

In this section, we evaluate the TDOA-based and the pseudo-spectrum based methods in terms of accuracy in estimating the wind speed horizontal component and direction. First, we describe the method employed to capture wind noise and label the audio recordings with measurements of an ultrasonic anemometer. Then, we introduce the evaluation metrics, the parameters involved in the proposed methods, and practical considerations. Finally, we present outcomes from an outdoor experiment, demonstrating each method's accuracy in estimating wind speed and direction. The data used in this study was collected on the rooftop of the Fraunhofer IIS (Erlangen, Germany) on windy days. Although wind noise was predominant, especially between 0 and 1 kHz, different acoustic sources were active, and generic ambient noise was present.

A circular array consisting of $N = 4$ omnidirectional MEMS microphones with a radius of 1 cm was used to record 3 hours and 40 minutes of wind noise. The wind speed and direction were also measured with an ultrasonic anemometer ATMOS 22 by METER Group[10] and synchronized to the audio recordings with a time resolution of 3
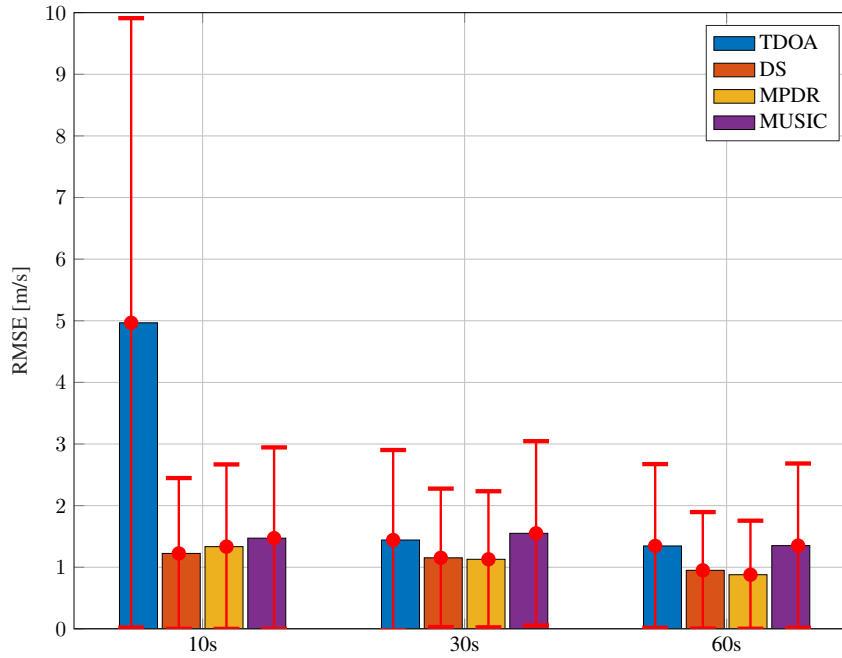
Figure 2: Speed RMSE obtained with the TDOA-based method (in blue), the DS (in orange), the MPDR (in yellow), and MUSIC (in purple) with three time averaging intervals. The red lines depict the error standard deviation for each method.

seconds. The measured speed and direction were used as ground truth values in this study. The wind speed varied within the range $0 \leq u \leq 10$ m s^{-1} and was well described by a Weibull distribution with a mean value of ≈ 2 m s^{-1}. The wind direction was mostly north, east, and south. The anemometer specifications provide an accuracy of $\pm[0.3$ m s^{-1}, $5°]$ and a resolution of $[0.01$ m s^{-1}, $1°]$ for the wind speed horizontal component and direction, respectively. The microphone array was placed 8 cm above the anemometer to avoid shadowing effects between the two instruments. The synchronization and data logging were performed using a single-board computer Raspberry Pi3B+. The wind noise was recorded at a sampling frequency $F_s = 8$ kHz after being prompted by the anemometer measuring a wind speed $u > 0.5$ m s^{-1}. The speed and direction values were stored in the 5th channel of the audio file, and the proposed methods were evaluated offline. The accuracy of the proposed method was evaluated by computing the root mean square error (RMSE) between the wind speed and direction estimates and the ground truth values as

$$\text{RMSE}_u = \sqrt{\frac{1}{Q}\sum_{q=1}^{Q}\left(\hat{u}(q) - u(q)\right)^2} \tag{30}$$

expressed in m s^{-1} for the airflow speed, and

$$\text{RMSE}_\theta = \frac{180}{\pi}\sqrt{\frac{1}{Q}\sum_{q=1}^{Q}\left(\angle\left\{e^{\iota\left(\hat{\theta}(q) - \theta(q)\right)}\right\}\right)^2} \tag{31}$$

expressed in degrees $(\cdot)°$ with respect to the true north axis for the direction, where q denotes the measurement index, $\hat{u}(q), \hat{\theta}(q)$ denote the estimated values, $u(q), \theta(q)$ denote the ground truth values, Q denotes the total number of considered measurements. The variability of the estimates was assessed with the error standard deviation.

As all methods presented in this work require time averaging, three time resolutions of $T = 10, 30, 60$ seconds were evaluated. In particular, these time intervals were used to compute the cross-correlation in Eq. (13) for the TDOA-based method and estimate the microphone signal covariance matrix in Eq. (18) for the pseudo-spectrum
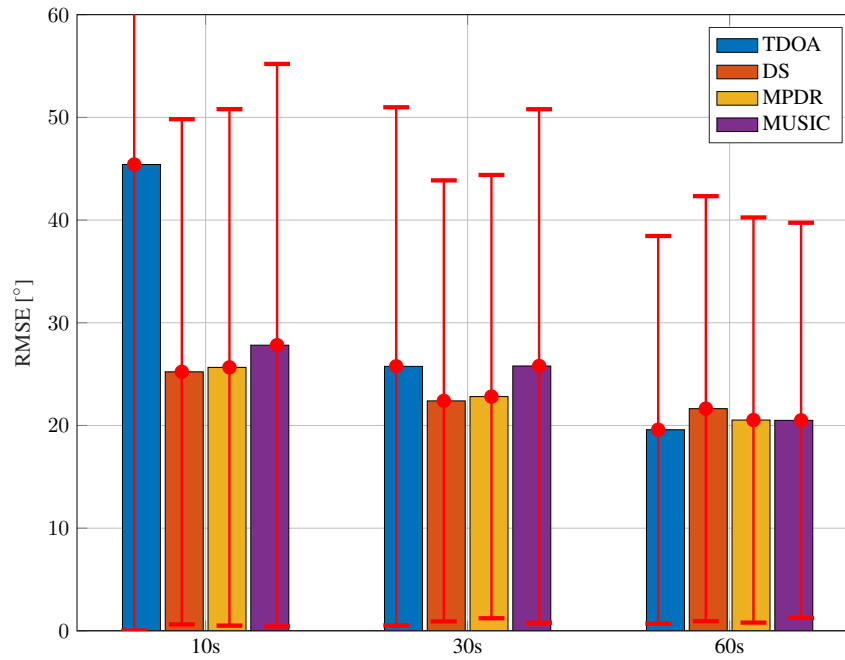
Figure 3: Direction RMSE obtained with the TDOA-based method (in blue), the DS (in orange), the MPDR (in yellow), and MUSIC (in purple) with three time averaging intervals. The red lines depict the error standard deviation for each method

based methods. Therefore, each method yielded one estimate of speed and direction every T seconds. The ground truth values were also averaged over the same time interval. A short T enabled a fast tracking of the wind velocity but increased the error variance, while a longer T decreased the error variance at the cost of reducing the tracking capability of short-term wind gusts. For the TDOA-based method, T seconds of wind noise were extracted, and the cross-correlation between microphone pairs in the spherical set was computed. The time lag corresponding to the cross-correlation peaks was identified and used to solve the linear system in Eq. (9). The PHAT weighting in Eq. (15) degraded the accuracy of the TDOA-based method and was, therefore, not included in this work. For the pseudo-spectrum based method, T seconds of wind noise were extracted and transformed into the STFT domain with a Hann window of $K = 512$ samples and 75% of overlap. Due to the considerations made in Section 1, we restricted the computation of the pseudo-spectrum up to the frequency instance $k_{\max} = 32$ corresponding to approximately 500 Hz. The reason was twofold, as the wind noise is most dominant in this frequency region and, based on our assumption, behaves as propagating noise. Although Eq. (16) yields narrowband estimates of speed and direction, we are interested in broadband estimates that could reflect the mean wind velocity. For this reason, the pseudo-spectrum was averaged across frequency before selecting the maximum as follows

$$P^{\mathrm{M}}(u, \theta) = \frac{1}{\mathcal{K}} \sum_{k=1}^{k_{\max}} P(k; u, \theta), \tag{32}$$

where $\mathcal{K}$ denotes the cardinality of the considered frequency instance set, yielding the new cost function

$$[\hat{u}, \hat{\theta}] = \underset{u, \theta}{\operatorname{argmax}} \, P^{\mathrm{M}}(u, \theta). \tag{33}$$

The optimization of the cost function in Eq. (33) was performed using a grid search across a speed range $u \in \{0.5, 12\}$ m s^{-1} with a resolution of 0.1 m s^{-1}, and a range of azimuthal direction angles $\theta \in \{0, 359\}°$ with a resolution of 1°.

In Fig. 2, we show the results obtained in terms of wind speed RMSE (colored bars) and error standard deviation (red line) with each method and for each time-averaging interval. Overall, the DS and MPDR beamformers yielded the
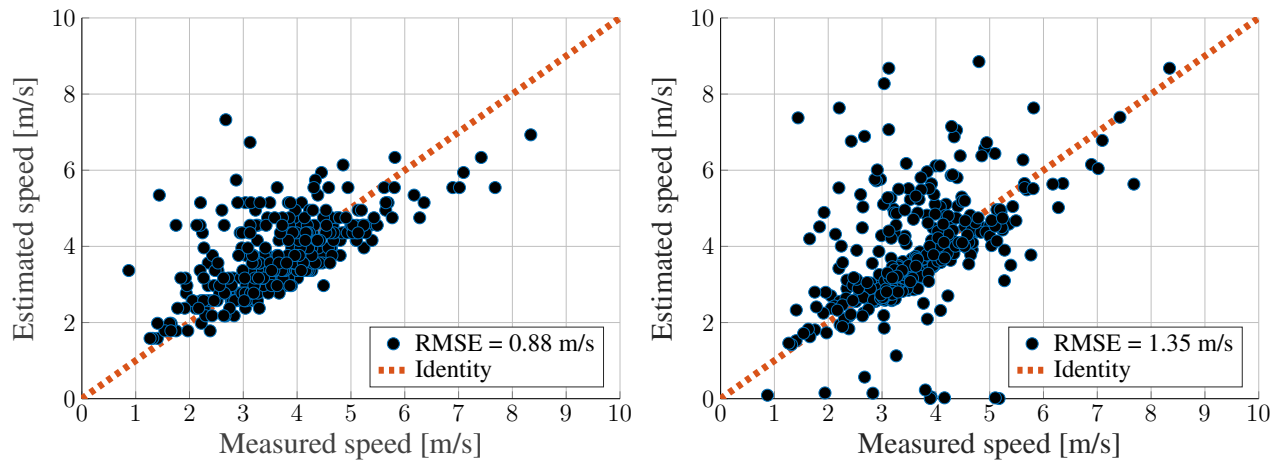
Figure 4: Scatter plots of wind speed estimated using the MPDR method (left) and the TDOA-based method (right) in the y-axis, and the wind speed measured with the anemometer in the x-axis. The time resolution was 60 seconds.
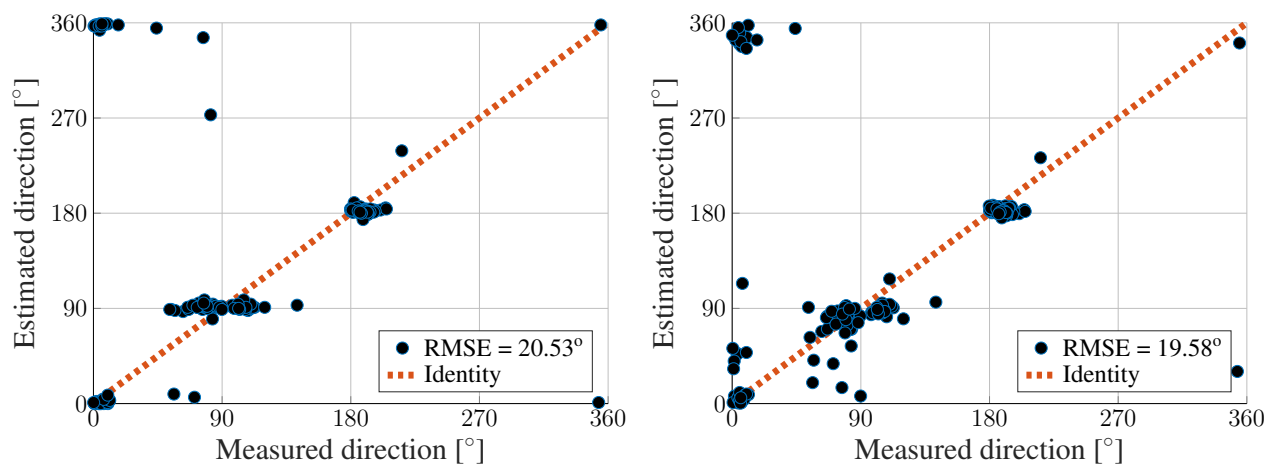
Figure 5: Scatter plots of wind direction estimated using the MPDR (left) and the TDOA-based method (right) in the y-axis, and the wind direction measured with the anemometer in the x-axis. The time resolution was 60 seconds.

best accuracy compared to the TDOA-based and MUSIC methods. The smallest error was obtained with the MPDR and 60 seconds of averaging corresponding to an RMSE of 0.88 m s^{-1} and a standard deviation of 0.87 m s^{-1}. A time-averaging interval of 10 seconds was insufficient to consistently obtain usable correlation peaks in the TDOA method, whereas the latter outperformed MUSIC with a time interval of 30 seconds. In Fig. 3, we show the results obtained in terms of wind direction RMSE (colored bars) and error standard deviation (red line) with each method and for each time-averaging interval. Similar to the speed evaluation, a time-averaging interval of 10 seconds was insufficient. For 30 and 60 seconds, the three methods performed similarly, with the TDOA-based method yielding the smallest RMSE of 19.58° with an error standard deviation of 18.87°.

Figures 4 and 5 show scatter plots of the estimates obtained with the MPDR method (left) and TDOA-based method (right) against the ground truth values of speed and direction for a time resolution of 30 seconds. Points close to the identity line (red) are associated with accurate estimates. In Fig. 4, it can be noticed that the TDOA-based method yielded more outliers, and the speed estimates were sparser compared to the MPDR. However, in Fig. 5, the
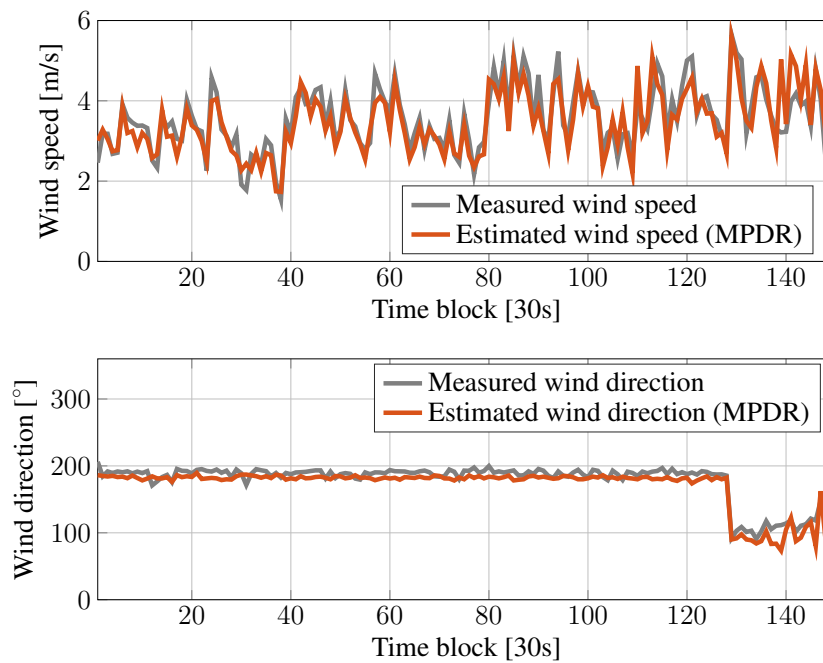
Figure 6: Wind speed (top) and direction (bottom) estimated with the MPDR beamformer (red lines) compared to the ground truth values measured with the anemometer (grey lines). Each data point corresponds to an estimate using 30 seconds of time averaging for a total of ≈ 75 minutes.

TDOA-based method performs similarly to the MPDR in estimating the wind direction. Note that points distributed around the upper left or the bottom right corners of the plots in Fig. 5 are associated to accurate estimates due to the angular wrapping between 360° and 0° despite being far from the identity line. In Fig. 6, we show an example of wind speed and direction (red lines) estimated with the MPDR beamformer over 75 minutes using 30 seconds of time averaging compared to the ground truth values measured with the anemometer (grey lines).

6.   CONCLUSION

We presented pseudo-spectrum based methods to estimate the wind speed and direction using signals captured with a compact microphone array, namely the DS and MPDR beamformers and the MUSIC algorithm. In addition, we reformulated an existing method based on TDOAs between microphone pairs to account for an arbitrary number of sensors and arbitrary sensor locations. The main idea was to exploit turbulent coherent structures advecting with the mean wind flow and generating correlated wind noise. Assuming that wind noise propagates in the wind direction and at nearly the wind speed across the array, the wind characteristics can be inferred using acoustic array processing. The proposed methods aim at evaluating the spatial pseudo-spectrum across candidate speed and direction values. The wind speed and direction estimates are obtained as the pair that maximizes the pseudo-spectrum. We assessed the accuracy of the extended TDOA-based and the pseudo-spectrum based methods with wind noise data collected outdoors using a 4-MEMS microphone array with a radius of 1 cm. The ground truth was measured with an ultrasonic anemometer and synchronized with the audio recordings. The evaluation was performed using three time resolutions of 10, 30, and 60 seconds over which the wind noise cross-correlation and covariance matrix were estimated. Based on the estimation error analysis, the MPDR beamformer yielded the lowest RMSE for the estimated speed. In contrast, the TDOA-based method yielded the lowest RMSE for the estimated direction when averaging across 60 seconds. The TDOA-based method required more than 10 seconds of averaging to yield usable results but was computationally more efficient than the pseudo-spectum-based methods. Future work could investigate using more sensors or specifically-designed pressure sensors for the task at hand, along with a higher sampling frequency.

REFERENCES

[1] P. Annibale, J. Filos, P. A. Naylor, and R. Rabenstein. TDOA-based speed of sound estimation for air temperature and room geometry inference. *IEEE Trans. Acoust., Speech, Signal Process.*, 21(2):234–246, 2012.

[2] 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.*, 97(1):695–696, 1995.

[3] W. K. Blake. *Mechanics of flow-induced sound and vibration, Volume 2: Complex flow-structure interactions*. Academic press, 2017.

[4] 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*, 35(12):2305–2321, 2018.

[5] 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.*, 57(4):2097–2106, 2018.

[6] G. Corcos. The structure of the turbulent pressure field in boundary-layer flows. *Journal of Fluid Mechanics*, 18(3):353–378, 1964.

[7] B. Efimtsov. Characteristics of the field of turbulent wall pressure-fluctuations at large reynolds-numbers. *Soviet Physics Acoustics-USSR*, 28(4):289–292, 1982.

[8] C. Knapp and G. Carter. The generalized correlation method for estimation of time delay. *IEEE Trans. Acoust., Speech, Lang. Process.*, 24(4):320–327, 1976.

[9] R. H. Mellen. On modeling convective turbulence. *J. Acoust. Soc. Am.*, 88(6):2891–2893, 1990.

[10] MeterEnvironment. ATMOS22 Ultrasonic Anemometer, 2019.

[11] D. Mirabilii and E. A. P. Habets. On the difference-to-sum power ratio of speech and wind noise based on the Corcos model. In *Proc. IEEE Intl. Conf. on the Science of Electrical Engineering (ICSEE)*, 2018.

[12] D. Mirabilii and E. A. P. Habets. Simulating multi-channel wind noise based on Corcos model. In *Proc. Intl. Workshop Acoust. Echo Noise Control (IWAENC)*, 2018.

[13] D. Mirabilii and E. A. P. Habets. Wind speed and direction estimation based on the spatial coherence of closely spaced microphones. *J. Acoust. Soc. Am.*, 1:346–357, 2022.

[14] D. Mirabilii, K. Lakshminarayana, W. Mack, and E. A. P. Habets. Data-driven wind speed estimation using multiple microphones. In *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 576–580, 2020.

[15] 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*, 2015, 2015.

[16] R. Raspet, J. Webster, and K. Dillion. Framework for wind noise studies. *J. Acoust. Soc. Am.*, 119(2):834–843, 2006.

[17] G. Van den Berg. Wind-induced noise in a screened microphone. *J. Acoust. Soc. Am.*, 119(2):824–833, 2006.

[18] H. L. Van Trees. *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2002.

[19] D. K. Wilson and M. J. White. Discrimination of wind noise and sound waves by their contrasting spatial and temporal properties. *Acta Acustica United With Acustica*, 96(6):991–1002, 2010.

[20] P. Xiao, K.-d. Yang, and Z.-x. Lei. Sequential filtering for surface wind speed estimation from ambient noise measurement. *China Ocean Engineering*, 31(1):74–78, 2017.