inter.noise
HAMBURG 2016

WIND SPEED ESTIMATION AND WIND-INDUCED NOISE REDUCTION USING A 2-CHANNEL SMALL MICROPHONE ARRAY

Shumpei SAKAI[1]; Tetsuro MURAKAMI[2]; Naoto SAKATA[3]; Hirohumi NAKAJIMA[4]; Kazuhiro NAKADAI[5];

[1-2,4] Kogakuin University, Japan

[3] Clarion, Japan

[5] Honda Research Institute, Japan

ABSTRACT

In this paper, 3 types of coherence analysis are performed for a wind-induced noise using a 2-channel small microphone array. The results indicated that there were linear correlations in the low-frequency band when the microphone spacing was very close each other, and the propagation speed of the wind-induced noise is not the speed of sound but the speed of wind. These results could lead to the wind-induced noise reduction by a linear signal processing with the transfer functions considering the speed of wind. Applying the beamforming technique to the wind-induced noise, Signal to Noise Ratio (SNR) was improved about 4-10 dB.

Keywords : Wind-induced noise, the speed of wind, Noise reduction, Microphone array, Beamforming
I-INCE Classification of Subjects Number(s): 22.3

1. INTRODUCTION

Since a wind-induced noise (Wind-N) degrades the sound quality in the outdoor recording session, the Wind-N reduction is inevitable. This is because it is impossible to remove Wind-N completely even if a large windscreen is installed. As Wind-N is highly non-stationary signal in time and spatial domains (1), only nonlinear noise reduction methods such as power spectrum subtraction or "post-wiener-filter" methods have been utilized. However, the nonlinear approach might introduce the nonlinear distortion and fail to restore the phase information. Although the linear processing approach is preferable in terms of the sound quality, there are few reports for Wind-N reduction based on the approach. This paper investigates 3 types of coherence function for Wind-N reduction using the linear beamforming technique.

2. SOME PREVIOUS WORKS AND THE FEATURES OF THIS PAPER

There have been several reports about Wind-N investigation (1, 2) and reduction (3-6). The basic property of Wind-N was clearly summarized in (1). It was reported that the correlation values of Wind-N between 2 microphones could be estimated by $\exp(-3.2x)$ for downwind condition and $\exp(-7x)$ for crosswind condition, where x means the microphone interval in wavelength (2). These estimators were heuristically derived from the field experiments using large aperture microphone arrays. As a Wind-N reduction method, the use of 3 microphones capsuled in acoustical resistive material was proposed (3). Another Wind-N reduction method was implemented by adjusting the optimum gain in the frequency domain using a 2-channel microphone array (4). Power spectrum of the target signal was estimated from the noisy power spectrum using a 3-channel microphone array based on coherence functions (5). A post-filter was introduced to a reduction method for spatially correlated noise (6). All the methods (4-6) are based on the nonlinear processing approach such as spectrum gain

[1] em16008@ns.kogakuin.ac.jp
[2] t.murakami@cc.kogakuin.ac.jp
[3] naoto_sakata@clarion.co.jp
[4] nakajima@cc.kogakuin.ac.jp
[5] nakadai@jp.honda-ri.com

control. Regarding the linear processing for Wind-N, it was shown that the speed of wind could be estimated using a microphone array (7). The features of this paper can be clearly summarized as follows;
1. 3 types of coherence function are analyzed using a microphone array with various spacings (10-100 mm), while a large spacing array was prepared in the other reports (2, 3, 7).
2. The speed of wind is estimated from the phase information using a small microphone array with 4 mm spacing, while a general spacing array with 600 mm was utilized in a report (7).
3. A Wind-N reduction based on the linear processing is performed, while the nonlinear processing was applied in the other methods (3-6).

3. COHERENCE OF WIND-INDUCED NOISE

3.1 Experimental settings

The experiments are conducted to investigate the coherences. A photo in Figure. 1 shows a room that the experiments were conducted. The 2-channel array with omni-directional microphones (AV-LEADER TCM-370) having 4 mm diameter is shown in Figure 2. Figure 3 is the fan (NEOVE FTS30-T12) to generate an artificial wind. Figure 4 and 5 show the IC-recorder (TASCAM DR-05) to record sound signals, and the wind speed meter (CELESTRON CE48021) to measure the speed of wind, respectively. The experimental conditions for the recordings are summarized in Table 1. The 3 major parameters are the microphone spacing, the speed of wind and the wind direction defined in Figure. 6. The recording time for the signal in every condition is 30 seconds with 44.1 kHz sampling rate.

Figure 1 – Experimental room

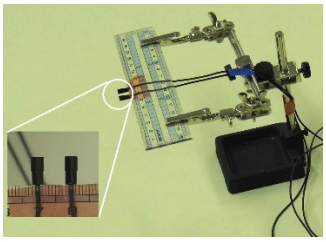

Figure 2 – Microphone array

Figure 3 – Fan

Figure 4 – IC recorder

Figure 5 – Wind speed meter

Table 1 – Recording conditions

Distance (fan to mic.), cm	Wind direction	Mic. spacing : d_m, mm	Wind speed (at mic.) : v_w, m/s	Gain setting (IC recorder)
158	Crosswind or Downwind	10, 20, 30, 50, 100	2.2, 1.7, 1.1	30

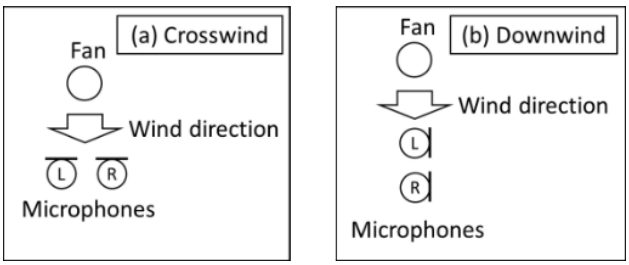

Figure 6 – Definition of the wind direction

3.2 Analysis for coherence functions

3 types of coherence function, "Complex coherence", "Amplitude coherence" and "Power coherence" are analyzed here. The general coherence $\gamma(A_l(\omega), B_l(\omega))$ is defined as

$$\gamma(A_l(\omega), B_l(\omega)) = \frac{\overline{A^*{}_l(\omega) B_l(\omega)}}{\sqrt{\overline{|X_l(\omega)|^2}}\sqrt{\overline{|Y_l(\omega)|^2}}} \tag{1}$$

where $A_l(\omega)$ and $B_l(\omega)$ are the complex spectrum, ω means the angular frequency, l indicates the time frame number, $\overline{A_l(\omega)}$ is the time frame average of A_n ($=\frac{1}{L}\sum_{l=1}^{L} A_n(\omega)$), L is the number of time frames and * is the complex conjugate operator. Complex coherence $\gamma_c(\omega)$ is calculated as

$$\gamma_c(\omega) = \gamma(X_{1,l}(\omega), X_{2,l}(\omega)) \tag{2}$$

where $X_{1,l}(\omega)$ and $X_{2,l}(\omega)$ are the complex spectra of input signal for channel 1 and 2, respectively. $X_{1,l}(\omega)$ is calculated using FFT with an window function $w(n)$ as

$$X_{1,l}(\omega) = \sum_{n=0}^{N-1} x_1(n + Sl)w(n)e^{i\omega n/f_s} \tag{3}$$

where $x_1(n)$ is the input signal for channel 1, n is the discrete time, N is the frame length (= FFT length), S is the frame shift length and f_s is the sampling frequency. $X_{2,l}(\omega)$ is obtained from $x_2(n)$ in the same manner as $X_{1,l}(\omega)$. As a practical setting, N=4096 and S=N/2 are prepared and the hanning window is employed for $w(n)$.

Amplitude coherence $\gamma_a(\omega)$ is defined as

$$\gamma_a(\omega) = \gamma\left(|X_{1,l}(\omega)| - \overline{|X_{1,l}(\omega)|}, |X_{2,l}(\omega)| - \overline{|X_{2,l}(\omega)|}\right) \tag{4}$$

Power coherence $\gamma_p(\omega)$ is defined as

$$\gamma_p(\omega) = \gamma\left(|X_{1,l}(\omega)|^2 - \overline{|X_{1,l}(\omega)|^2}, |X_{2,l}(\omega)|^2 - \overline{|X_{2,l}(\omega)|^2}\right) \tag{5}$$

Complex coherence is often used to evaluate a performance for the linear signal processing, such as correlation-based delay estimation or noise reduction based on beamforming. On the other hand, amplitude and power coherence can evaluate a performance for the nonlinear processing such as noise reduction based on power spectral subtractions.

3.3 Results and discussions

Figure 7 shows 3 types of coherence value for Wind-N. As for complex coherence, the amplitude of the complex coherence $\gamma_c(\omega)$ is displayed in Figure 7. In the frequency band lower than 125 Hz, every coherence value was roughly 0.3-0.8 including complex coherence. This means that Wind-N has a linear correlation over the low frequencies. In the middle-frequency band from 125 to 500 Hz, the complex coherence value was very small, whereas the amplitude and power coherences were around 0.4-0.6. This implies that only amplitude or power spectral subtraction-based methods would be effective in this band. In the frequency band between 1 and 8 kHz, every coherence value became quite high. This was caused by the motor noise emitted from the fan and the fan itself, not by Wind-N at the array. Figure 8 shows the results of analyzing the complex coherence with respect to the microphone spacing. High coherence values (> 0.7) were given in the low-frequency band when the microphone spacing was very small (<10mm). Figure 9 shows the results of analyzing the complex coherence with respect to the wind direction such as a comparison between crosswind and downwind. Higher coherence values were obtained in the low-frequency band when the microphone array was the case of the downwind location depicted in Figure 9. The green and cyan lines show theoretically estimated lines reported in (2), which have good agreement with the measurement data.

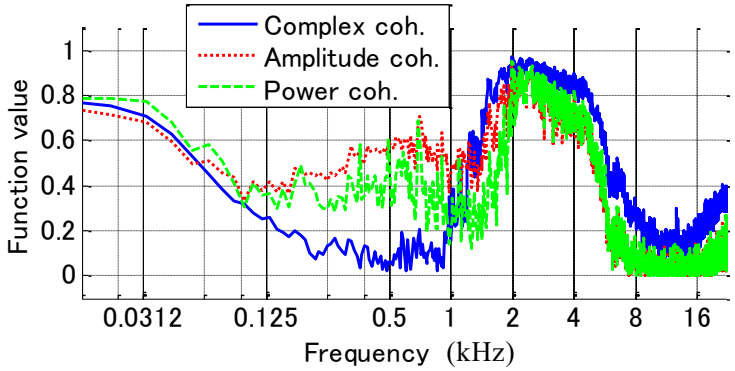

Figure 7 – 3 types of coherence functions

(Downwind, v_w = 2.2 m/s, d_m = 10 mm)

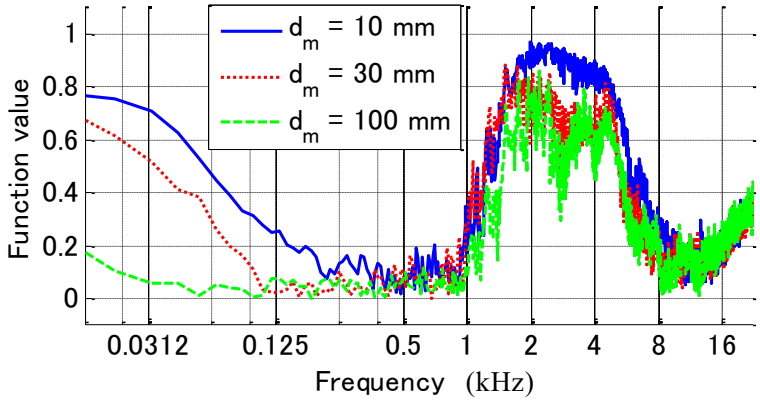

Figure 8 – Complex coherences for the microphone spacing (Downwind, v_w = 2.2 m/s)

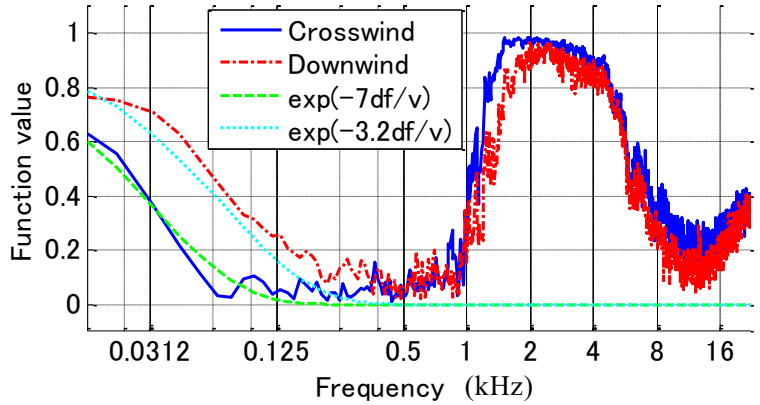

Figure 9 – Comp. coherence vs. wind direction (d_m = 10 mm, v_w = 2.2 m/s)

4. AN ESTIMATION FOR THE SPEED OF WIND BY A SMALL MICROPHONE ARRAY

4.1 Experimental settings

The estimation for the speed of wind using a small microphone array was conducted. Figure 10 is a photo of an anechoic room where this experiment was performed. Figure 11 shows the microphone array whose elements are composed of TCM-370. Three conditions for the microphone spacing (4 mm, 10 mm, 20 mm) were tested. Figure 12 shows the audio interface (M-AUDIO fast track pro) used for the recordings. Two conditions for the wind (Strong and Weak) by a circulator (DAS-KJ191 shown in Figure 13) were prepared. The speeds of wind were 1.8 m/s for strong and 2.6 m/s for weak, which were measured by the wind speed meter shown in Figure 5. A configuration of the equipments and the wind direction to the microphone array is shown in Figure 14. The recordings were conducted for 10 seconds in 44.1 kHz sampling frequency, and for each condition in the speeds of wind.

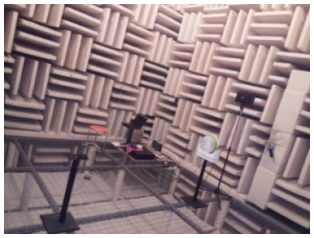

Figure 10 – An anechoic room

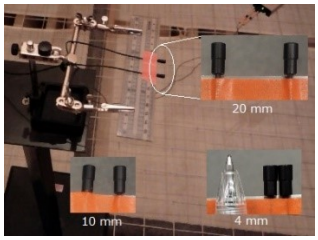

Figure 11 – Mic Array

Figure 12 – Audio Interface

Figure 13 – Circular

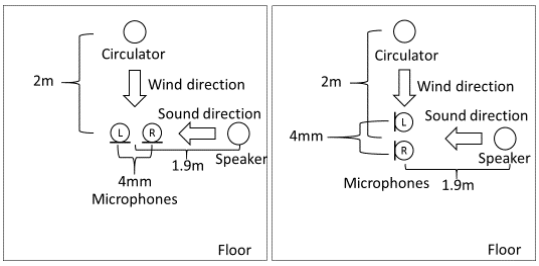

Figure 14 – Configuration of the equipments

4.2 Algorithm for the estimation of the speed of wind

The speed of wind is estimated from the phase information of complex coherence in Eq. (2). An unwrapped phase difference function $\varphi(\omega)$ is calculated as

$$\varphi(\omega) = \mathrm{unwrap}\{\arg\{\gamma_c(\omega)\}\} \tag{6}$$

where arg{ } corresponds to an argument of the complex number and unwrap{ } is the phase unwrapping. The estimated delay $\hat{\tau}$ is obtained by minimizing the cost function $J(\tau)$ as

$$J(\tau) = \sum_{\omega=0}^{2\pi f_{max}} |\varphi(\omega) - (-j\omega\tau)|^2 \tag{7}$$

where f_{max} is the upper limits of the frequency band in Hz and adjusted manually in experiments. This means the liner approximation for $\varphi(\omega)$ by the $(-j\omega\tau)$. For practical reasons, $\hat{\tau}$ is calculated as

$$\hat{\tau} = -\frac{1}{2\pi}\boldsymbol{f}^{+}\boldsymbol{\varphi} \tag{8}$$

where $\boldsymbol{f}$ is the column vector having discrete frequency values from 0 to f_{max}, $\boldsymbol{\varphi}$ is the column vector having phase difference values represented as $\boldsymbol{\varphi} = \varphi(2\pi\boldsymbol{f})$, and $\boldsymbol{A}^{+}$ means the pseudo inverse matrix of $\boldsymbol{A}$. The speed of wind v_w is estimated as $\hat{v}_w = \hat{\tau}/d_m$, where d_m is the microphone spacing.

4.3 Results

Figure 15 shows the amplitude of the complex coherence, $|\gamma_c(\omega)|$ for the strong downwind condition. According to the Figure, the complex coherence with 4 mm microphone spacing was much higher than the others in the low-frequency band. The unwrapped phase differences $\varphi(2\pi f)$ is shown in Figure 16. It can be seen that a linear slope of the phase difference was observed up to 1,400 Hz for the 4 mm microphone spacing. The estimation results for the speed of wind were summarized in Table 2. The relative estimation errors were at most 22%. This accuracy is comparable with one using 600 mm spacing microphone array in (2).

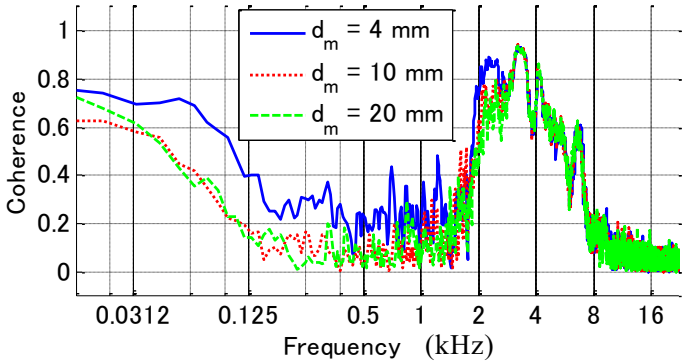

Figure 15 – Amplitude of complex coherence $|\boldsymbol{\gamma_c(\omega)}|$

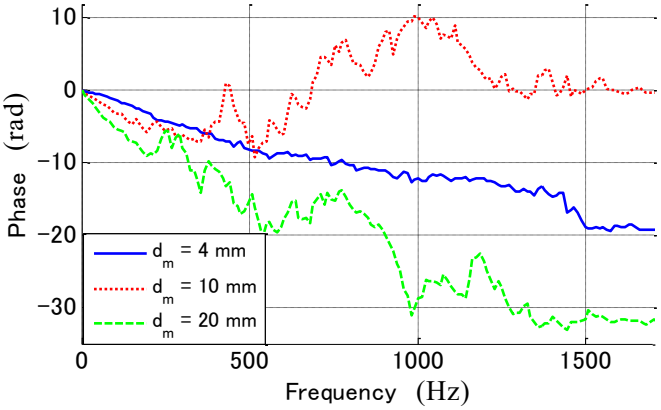

Figure 16 – Unwrapped phase differences

Table 2 – Results for wind speed estimation

Wind speed v_w, m/s	Mic. distance d_m, mm	Upper freq. lim. f_{max}, Hz	Estimated speed $\hat{v}_w$, m/s	Relative error e_r, %
1.8	4	549	1.4	22
1.8	10	96.9	1.8	0
1.8	20	118	2.2	22
2.6	4	1346	2.1	19
2.6	10	194	2.2	15
2.6	20	183	2.6	0

5. WIND-INDUCED NOISE REDUCTION BASED ON LINEAR BEAMFORMING

Wind-N reduction based on 2-channel beamformer (BF) is performed. Coefficient design of BF, experimental setting and its results are described in the following section.

5.1 Coefficients design for 2-channel beamformer

A general BF having convolution-add structure is employed. The output $y(n)$ of the BF is calculated as

$$y(n) = g_1(n) * x_1(n) + g_2(n) * x_2(n) \tag{9}$$

Here, $g_1(n)$ and $g_2(n)$ are the BF coefficients, $x_1(n)$ and $x_2(n)$ are the input signal of each channel, and $*$ indicates the convolution operation. 2 types of the BF coefficients are described below.

5.1.1. Linear subtraction beamformer (LS-BF)

A BF coefficient $g_2(n)$ that minimizes the power of the BF output for Wind-N subject to $g_1(n) = 1$ is designed here. To obtain the optimal coefficient in a practical environment, a recorded input signal is utilized. $g_2(n)$ is calculated as

$$g_2(n) = ifft\left\{-\frac{\overline{X_{1l}(\omega) \cdot X_{2l}^*(\omega)}}{\overline{|X_{2l}(\omega)|^2}}\right\} \tag{10}$$

where $X_{1l}(\omega)$ and $X_{2l}(\omega)$ are input spectrum of the Wind-N signal for channel 1 and 2, respectively and $ifft\{\}$ shows the inverse-fft operation and *means the complex conjugate.

5.1.2. Source separation beamformer (SS-BF)

BF coefficients are designed to separate a target source from a Wind-N regarded as a disturbing source. In the frequency domain, the coefficients can be obtained as a solution of

$$\begin{pmatrix} K_1(\omega) & K_2(\omega) \\ H_1(\omega) & H_2(\omega) \end{pmatrix} \begin{pmatrix} G_1(\omega) \\ G_2(\omega) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{11}$$

where $G_1(\omega)$ and $G_2(\omega)$ are fft of $g_1(n)$ and $g_2(n)$, $K_1(\omega)$ and $K_2(\omega)$ are the transfer functions between the target source and microphone channel 1 and 2, respectively. $H_1(\omega)$ and $H_2(\omega)$ are the transfer functions with respect to the Wind-N. $G_1(\omega)$ and $G_2(\omega)$ are the solution of Eq. (11), which are expressed as

$$G_1(\omega) = \frac{H_2(\omega)}{H_2(\omega)K_1(\omega) - H_1(\omega)K_2(\omega)} \tag{12}$$

$$G_2(\omega) = \frac{-H_1(\omega)}{H_2(\omega)K_1(\omega) - H_1(\omega)K_2(\omega)} \tag{13}$$

These transfer functions are given as

$$K_1(\omega) = 1,\ K_2(\omega) = \exp(-j\omega\tau_t) \tag{14}$$

$$H_1(\omega) = 1,\quad H_2(\omega) = \exp(-j\omega\tau_j) \tag{15}$$

where τ_t and τ_j are arrival delay times of the microphones regarding the target and the disturbance, respectively. τ_t is set according to the speed of sound and the incident angle of the target source. τ_j is calculated from the speed of wind that is parallel to the microphone axis. With Eqs. (14)-(15), the denominator of Eq. (12) and (13) can be rewritten as

$$d(\omega) = \mathrm{e}^{-i\omega t_j} - \mathrm{e}^{-i\omega t_t} \tag{16}$$

Actually, this equation could have some periodic zeros at the multiple frequencies. To avoid divergence of coefficients, a regularity parameter α is introduced and the coefficients are calculated as

$$G_1(\omega) = -\mathrm{e}^{-i\omega t_u}\frac{d(\omega)^*}{|d(\omega)|^2 + \max(|d(\omega)|^2) \times \alpha} \tag{17}$$

$$G_2(\omega) = \frac{d(\omega)^*}{|d(\omega)|^2 + \max(|d(\omega)|^2) \times \alpha} \tag{18}$$

where max() shows the maximum over the frequency ω.

5.2 Experimental settings and evaluation methods

The experimental conditions are the same as in Sec. 5 except that a loudspeaker (Bose 101MM) with an amplifier (YAMAHA MX-1) is used as a target source. The source location is depicted as shown in Figure 14. The source signal is 10 seconds of music extracted from "Popular music" No. 8 in RWC-Database (8). The performances are evaluated by SNR differences between "before" noise reduction and "after" noise reduction. SNR in dB is obtained as

$$\mathrm{SNR} = 10\log_{10}\left(\frac{\sum_\omega P_S(\omega)}{\sum_\omega P_{S+N}(\omega) - \sum_\omega P_S(\omega)}\right) \tag{19}$$

where $P_{S+N}(\omega)$ and $P_S(\omega)$ are the power spectra of the BF output signals with and without Wind-N, respectively.

5.3 Results and discussion

Figure 17 shows a comparison between the power spectra under the strong downwind condition. "Original" and "Signal only" in the figure corresponds to the input signal with and without Wind-N. "Original" was the highest in the comparison in the low-frequency bands (< 500 Hz) and comparable with the others around 1.5 kHz.

"LS-BF" and "SS-BF" show the Wind-N reduction results. Wind-N was reduced by the methods in the low frequency bands (< 250 Hz). The SNR improvements are summarized at Table 3. SS-BF resulted in a better performance than LS-BF and achieved the SNR improvements by 4-10dB.

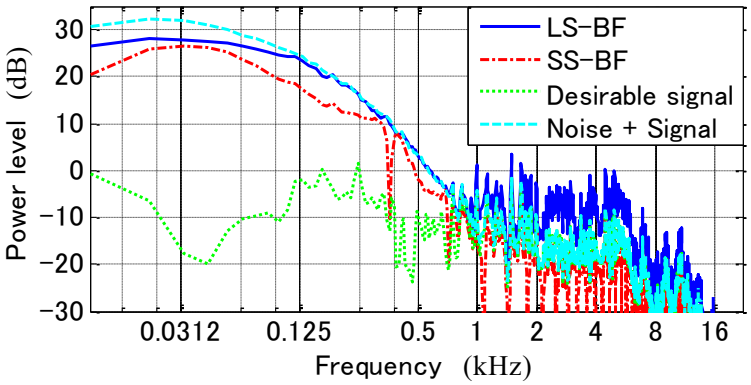

Figure 17 – Comparison of power spectrum

Table 3 – SNR improvement in dB

Method	Crosswind $v_w = 1.8$ m/s	Crosswind $v_w = 2.6$ m/s	Downwind v_w=1.8 m/s	Downwind v_w=2.6 m/s
LS-BF	2.7	3.0	2.4	2.1
SS-BF	10.2	10.2	4.0	5.3

6. CONCLUSIONS

Wind-induced noise coherences using a small microphone array were analyzed. The results showed that linear correlations existed in low frequency band. Apply the linear correlation, wind speeds are estimated within 20% errors and wind-noise reductions based on linear beamforming achieved 4-10dB SNR improvements.

REFERENCES

1. R. Raspet, J. Webster, and K. Dillion, “Framework for wind noise studies,” J. Acoust. Soc. Am., Vol. 119 (2), pp. 834-843, 2006.
2. F. Shields, “Low-frequency wind noise correlation in microphone arrays,” J. Acoust. Soc. Am., Vol. 117 (6), pp. 3489-3496, 2005.
3. D. Herman, “Wind noise rejection apparatus,” U. S. Patent, US 8,391,529 B2, 2013.
4. M. Yoshida and T. Oku, “Wind noise reduction device,” U. S. Patent, US 8,428,275 B2, 2013.
5. Y. Chung, “Rejection of flow noise using a coherence function method,” J. Acoust. Soc. Am., Vol. 62 (2), pp. 388-395, 1977.
6. K. Kumatani, B. Raj, R. Singh, and J. McDonough, “Microphone Array Post-filter based on Spatially-Correlated Noise,” Proc. of Interspeech 2012, 2012.
7. H. Bass, R. Raspet, and J. Messer, “Experimental determination of wind speed and direction using a three microphone array,” J. Acoust. Soc. Am., Vol. 97 (1), pp. 695-696, 1995.
8. Masataka Goto, et. al., “RWC Music Database: Popular, Classical, and Jazz Music Databases,” Proc. of the 3rd Int'l Conf. on Music Information Retrieval (ISMIR 2002), pp.287-288, 2002.