BeBeC-2014-10

PRELIMINARY BENCHMARKING OF MICROPHONE ARRAY METHODS

Gert Herold and Ennes Sarradj
Brandenburg University of Technology Cottbus - Senftenberg
Chair of Technical Acoustics
Siemens-Halske-Ring 14, D-03046 Cottbus, Germany

ABSTRACT

Microphone-array-based methods have become a common tool for the characterization of acoustic sources. One of the main objectives when using these methods is determining the exact magnitude of the noise radiated by each separate sound source, which can be described by the sound pressure level at a certain distance and direction from the source.
In this contribution, simulated as well as measured data of known source configurations are being evaluated using array-based methods like DAMAS, CLEAN, CLEAN-SC, Orthogonal Beamforming, and Covariance Matrix Fitting. The results obtained using these methods and varying several of their parameters will be compared.
It will be shown that, in part, these results differ considerably in terms of calculated position and magnitude of the sound sources. This underlines the need for defining criteria to assess the reliability of microphone array methods under different boundary conditions.

1 INTRODUCTION

Microphone arrays have become a common tool for acoustic measurements in recent years. Processing measured data with dedicated algorithms does not only provide the location of acoustic sources but allows to assess the amount of emitted sound by each source quantitatively.

Stationary acoustic sources can be described by their position, their power in terms of radiated sound (e.g. described by the measured sound pressure level in a distinct distance from the source), their directivity, and their coherence to each other.
These properties have an effect on the radiated sound and therefore on the recorded signals. Depending on the sound propagation model used for the acoustic source reconstruction some of these properties are being assumed to have a certain characteristic. While there are methods that include the directivity (e.g. SODIX [6]) or coherent sources (e.g. DAMAS-C [3] and MACS

[12]) in most cases all sources are assumed to be omni-directional (i.e. only monopole sources) and incoherent.

The methods compared in this paper work in the frequency domain and are based on the cross spectral matrix (CSM). Classical delay-and-sum (DAS) beamforming based methods include the widely established DAMAS algorithm [2], the CLEAN algorithm [5], and the also well-known CLEAN-SC method [9]. Furthermore, orthogonal beamforming (OB) [7] and covariance matrix fitting (CMF) [1, 11], which work independent from DAS beamforming, are featured in this paper as well. These microphone-array based methods will be applied to a reference setup, and it will be shown that these methods yield differing results while using the same original data.

After reviewing the theoretical bases of the aforementioned algorithms in section 2, the setup of the reference case used is described in section 3. Section 4 comprises the results obtained applying the methods to measured and simulated data. The findings are finally summarized in section 5.

2 THEORY

The classical DAS beamforming in the frequency domain consists of shifting the phase of all microphones such that they “focus" on a chosen point in space, i.e. taking into account the retarded time from the focus point to the microphones. Considering a grid of N focus points with possible source positions and using an array consisting of M microphones the beamformer output at the assumed source positions $\boldsymbol{x}_\mathrm{t}$ can be calculated with

$$B(\boldsymbol{x}_\mathrm{t}) = \boldsymbol{h}^\mathrm{H}(\boldsymbol{x}_\mathrm{t})\, \boldsymbol{C}\, \boldsymbol{h}(\boldsymbol{x}_\mathrm{t}) \,, \qquad t = 1 \ldots N \,, \tag{1}$$

where $^\mathrm{H}$ denotes the Hermitian transpose. The entries of the steering vector $\boldsymbol{h}$ are calculated via

$$h_m = \frac{1}{r_{\mathrm{t},m} \sqrt{M \sum_{l=1}^{M} r_{\mathrm{t},l}{}^{-2}}} \mathrm{e}^{-\mathrm{j}k(r_{\mathrm{t},m} - r_{\mathrm{t},0})} \,, \qquad m = 1 \ldots M \,, \tag{2}$$

which corresponds to formulation IV in [8]. Due to the finite measurement time, $\boldsymbol{C} \approx \boldsymbol{p}\boldsymbol{p}^\mathrm{H}$ is always only an approximation of the CSM and is obtained by averaging the windowed and FFT-transformed time signals measured by the microphones. The beamformer output in equation (1) can be interpreted as a source distribution being convoluted with a point spread function (PSF).

Algorithms

The DAMAS method iteratively solves the system

$$\boldsymbol{b} = \boldsymbol{P}\, \boldsymbol{d} \tag{3}$$

using a Gauss-Seidel algorithm. The vector $\boldsymbol{b}$ contains the results at the grid points calculated with DAS beamforming, $\boldsymbol{P}$ represents the matrix of point spread functions from each focus

	DAMAS	CLEAN	CLEAN-SC	OB	CMF
Removed main diagonal of CSM	yes	yes	yes	yes	yes
FFT block size	1024	1024	1024	1024	1024
Steering vector	eq. (2)	eq. (2)	eq. (2)	eq. (2)	–
Max. number of iterations	500	500	500	–	500
Damping factor	–	0.6	0.6	–	–
LassoLARS [10] weight factor	–	–	–	–	10^{-9}
Number of eigenvalues	–	–	–	19	–

Table 1: Fixed parameters for the used methods.

point to all focus points (one PSF per column), and $\boldsymbol{d}$ is the deconvoluted map.

The CLEAN algorithm looks for the highest value in the map obtained with equation (1), copies it (or part of it) into an empty map, and subtracts its convolution with the corresponding PSF from the original map. This process is reiterated using the new maps until no more significant sources are found.

The fast CLEAN-SC method works in a similar way. Instead of using a PSF to determine what portion to subtract, this method identifies parts that are spatially coherent to the highest peak and removes them from the original map.

The OB technique does not require à priori DAS beamforming. It uses an eigenvalue decomposition of the CSM to determine strength and position of incoherent sources. The eigenvalues are used for estimating the source level and, in combination with the corresponding eigenvector, to find the source location via component-wise beamforming.

The CMF method solves the convex optimization problem

$$\text{minimize}_{\{d_n\}_{n=1}^{N}} \|\boldsymbol{C} - \boldsymbol{A}\boldsymbol{D}\boldsymbol{A}^{\mathrm{H}}\|_{\mathrm{F}}^{2}\,, \tag{4}$$

with $\boldsymbol{C}$ being the approximated CSM with removed main diagonal [4]. $\boldsymbol{A}$ denotes the transfer matrix for monopoles from the grid of possible sources to the microphones and $\boldsymbol{D}$ is a diagonal matrix containing the unknown cross spectrum of the sound pressure amplitudes at the grid points.

Each of the methods has one or more parameters that affect the behavior and performance of the algorithm. These parameters have been set to fixed values for the purpose of this analysis. A list of the parameters (not all methods share the same set of parameters) and their respective values is shown in table 1.

3 REFERENCE CASE

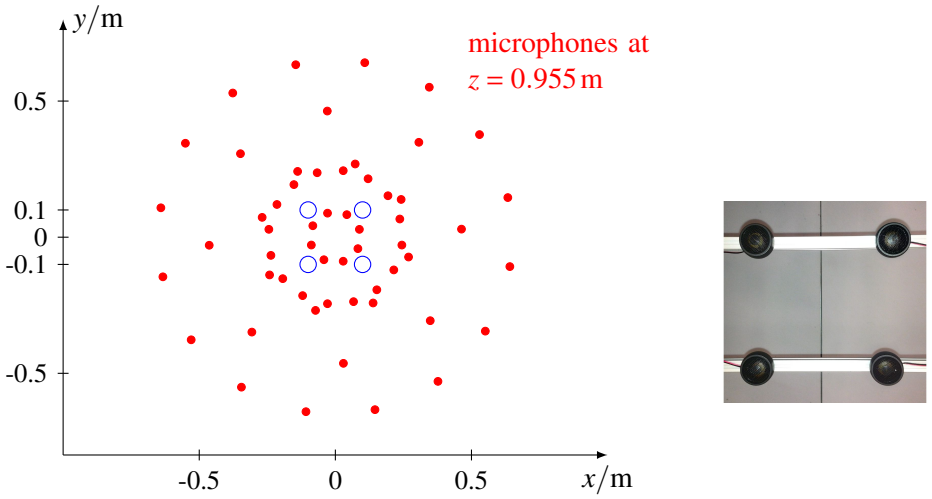

Figure 1: Setup used for the measurements and simulations, view from above. Left: Schematic representation including the array; the blue circles represent the loudspeakers. Right: Tweeters that have been used for the measurements.

The setup used for reference consists of four tweeters, positioned in the corners of a square of 20 cm edge length. Measurements with all speakers emitting incoherent white noise at the same level were made as well as with a sound pressure level difference of 3 dB and 6 dB between each speaker (descending level in clock-wise order). Measurement time was 40 s with a sampling frequency of 51200 Hz.
The array consists of 56 microphones, has a maximum diameter of 130 cm, and is mounted 95.5 cm above the speakers. A schematic representation of the microphone array and the speaker setup is shown in figure 1.

Additionally, simulations representing the experimental setup were executed. The speakers were simulated as monopole point sources emitting a white noise signal, which was generated with a random normal distribution of values.

4 RESULTS

The sound maps calculated for evaluation contain an area of $0.6\,\mathrm{m} \times 0.6\,\mathrm{m}$ with a regular-spaced focus grid with 2 cm intervals. For quantitative analysis, an integration area with a radius of 4 cm around each speaker position is defined, and the computed sound pressure levels at the focus points within these areas are summed.

In figure 2 calculated sound maps based on measured data with all speakers emitting sound at the same level are shown for three exemplary third-octave bands (1000 Hz, 5000 Hz, and 20000 Hz) and the five methods.

Not depending on the method, the integration over the designated areas at 1000 Hz does not yield equal results for all four areas , which is most likely due to varying characteristics of the tweeters at low frequencies.
With all sources at the same level, DAMAS and OB yield approximately the same results in terms of source position and extent. The DAMAS reconstructed level is almost always about 1 dB below the level calculated using CMF. CLEAN and CLEAN-SC detect quite distinct peaks, which is intrinsic to these algorithms. However, the CLEAN method proves to be rather unstable and often omits one or more sources. Quantitatively, the levels detected using CLEAN-SC are in the same range as those calculated by DAMAS, while the CLEAN levels are slightly lower. OB yields levels also slightly higher than DAMAS, however the absolute values of the four integration areas vary more among each other than with any other of the methods except CLEAN. At 1000 Hz the OB algorithm is not any more able to detect the sources at their correct positions.
A comparison of the sound maps at the same frequency bands, but with a 6 dB level interval between the speakers, is shown in figure 3. The displayed sound pressure level range in this figure is extended to 20 dB, which shows the CMF and DAMAS maps at 20 kHz to be noisier than the maps generated using other methods. While all the methods show almost the same characteristic as with equally-loud sources for the loudest source (at the lower right), some more properties are revealed when looking at the lower-leveled speakers. At the 5000 Hz band, CLEAN-SC, OB and CMF reproduce the 6 dB intervals at the integration areas. The DAMAS method seems to overestimate the level difference between the sources, i.e. the low levels are calculated to be lower than anticipated.
This phenomenon is confirmed via a similar analysis using simulated data as shown in figure 4. As is to be expected, using simulated data in most cases leads to clearer maps with distinct peaks at the source positions. At 1000 Hz, however, OB and CLEAN-SC are only able to detect the loudest source and the DAMAS result shows a scattered source region and omits the most quiet source. CLEAN omits several sources at all frequency bands.

Figures 5ff compare the integrated levels at the four areas for third-octave bands from 500 Hz to 20000 Hz. The measurements have been evaluated in figures 5, 7 and 9 for speakers with 0 dB, 3 dB and 6 dB respectively. Figures 6, 8 and 10 feature the corresponding simulated cases. The latter figures also include the sound pressure level at the array center caused by one broadband source at the respective positions. Theoretically, these curves would be straight ascending lines on a double-logarithmic scale. Due to the limited simulation period and the limited block size used for FFT (1024 in this case), however, a slight deviation from a straight line can be seen in the simulated cases.
The tendencies and characteristics of the methods that can be drawn from the diagrams are summarized in the following. It should be kept in mind that these findings are only related to the analyzed case and parameters and cannot be applied at different conditions without further consideration.
The results calculated by the methods become inaccurate at low frequencies. With the exception of CMF, even under computer-simulated conditions the algorithms fail to reconstruct the sources correctly. Of the compared methods, OB is the most sensitive to low frequencies, not being able to correctly detect sources below 2000 Hz. CLEAN-SC can reconstruct the correct levels until around 1250 Hz. Notwithstanding the general unreliability of CLEAN

DAMAS	CLEAN	CLEAN-SC	OB	CMF
2478	9	288	104	2070

Table 2: Average calculation time (without PSF) in ms *for one frequency.*

when detecting multiple sources this frequency seems to be critical here as well. The DAMAS method is only capable of reconstructing the sources and their levels at frequencies above 800 Hz. CMF seems to be able to distinguish all sources down to 500 Hz, the lowest frequency band considered in this contribution.
Additionally, CMF is also the method rendering levels closest to the theoretical values in the simulations, followed by CLEAN-SC and OB. In the measurements with 6 dB intervals (fig. 9), CMF fails to detect the quietest source at frequency bands above 8000 Hz, while CLEAN-SC and OB, albeit with a rather high level, still do.
As mentioned earlier, DAMAS shows the tendency to underestimate the levels of all sources that are not loudest. Simulations as well as measurements show that this effect becomes stronger with higher level differences of the sources.

The calculations were carried out on a 2.2 GHz CPU each. A comparison of the methods in terms of computing time is shown in table 2. The time needed to compute the PSF (which is needed for the DAMAS and the CLEAN algorithm) is excluded from timing. Calculating the PSF is only needed once for a given configuration of microphone array and focus grid.
The fastest method by far is the CLEAN algorithm, followed by OB and CLEAN-SC, which are still reasonably fast. CMF and DAMAS are considerably slower.

5 CONCLUSION

When using microphone arrays for acoustic source characterization a variety of methods can be chosen to obtain position and strength of the sources. To use beamforming and other microphone-array based methods as sophisticated measuring instrument it is necessary for these methods to render an exact and nonambiguous output for a given configuration. The exemplary application of some well-known algorithms on a reference case as presented in this contribution proves that these methods, however, do not yield the same results in every case. It has been shown that the calculated sound pressure levels deviate depending on the used algorithm, examined frequency band, and dynamic range of the source levels.
This analysis represents only a very limited view on the topic, since the possible parameter space is extensive. In order to build up a catalog indicating the reliability or even applicability of a given algorithm on an arbitrary problem, a systematic variation of parameters and the definition of quality criteria are needed and will be the basis for further research.

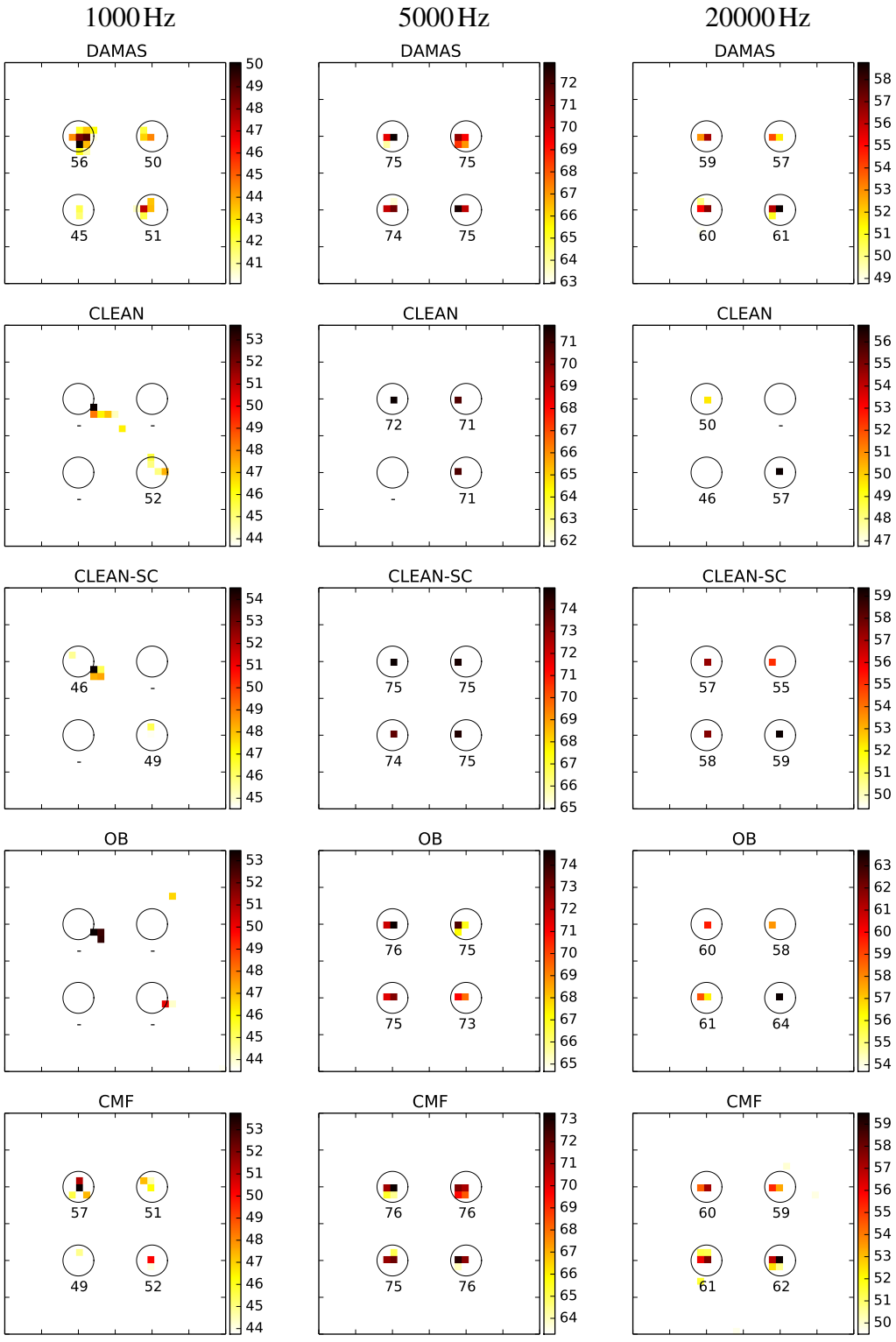

Figure 2: Sound maps evaluated at third-octave bands at 1000 Hz, 5000 Hz, *and* 20000 Hz *using different methods. Measurements with four speakers at the same level. The integrated levels for each sector are indicated under the respective circular areas (all values in* dB*).*

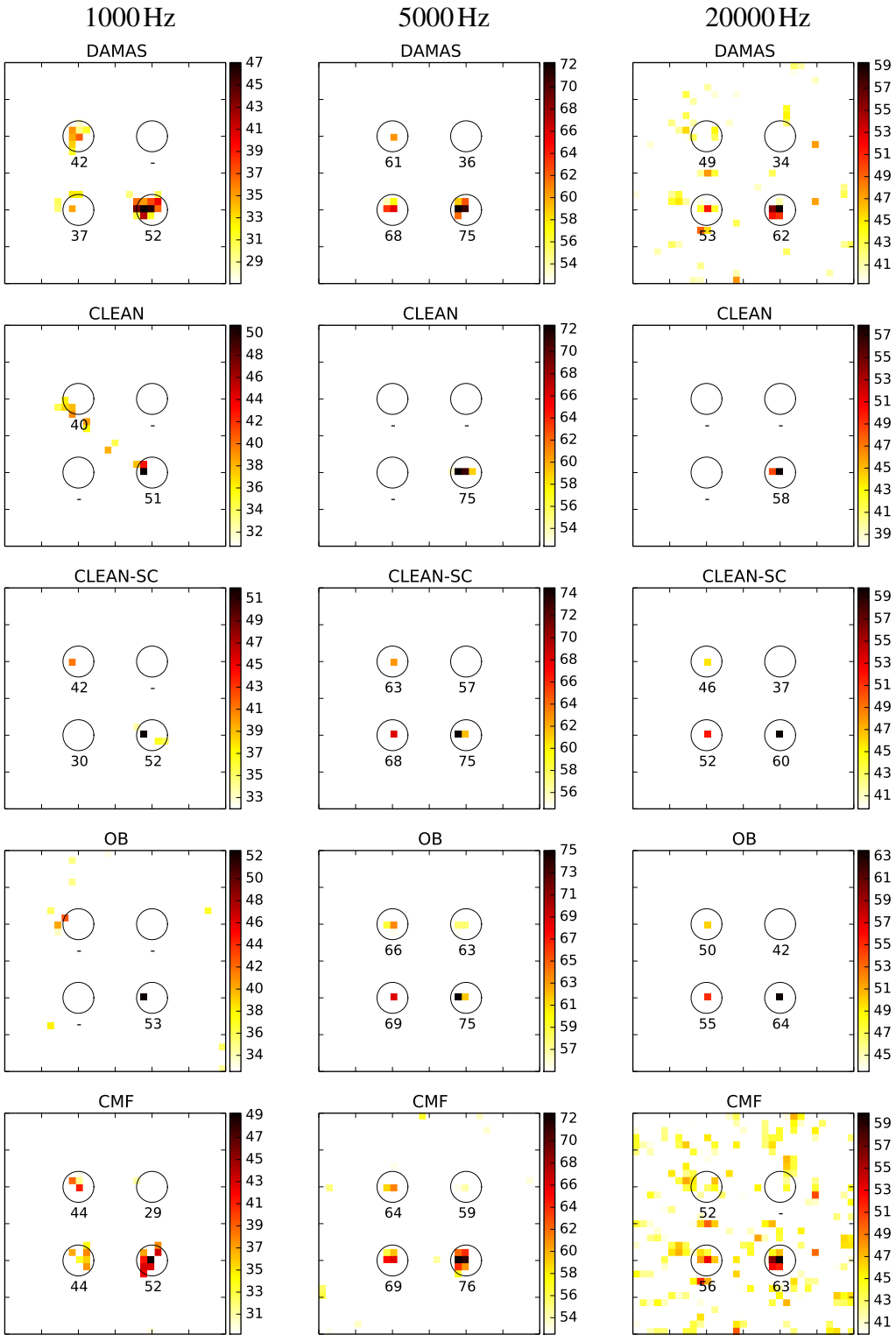

Figure 3: Sound maps evaluated for third-octave bands at 1000 Hz, 5000 Hz, *and* 20000 Hz *using different methods. Four measured sources with differing levels of* 6 dB *(all values in* dB*).*

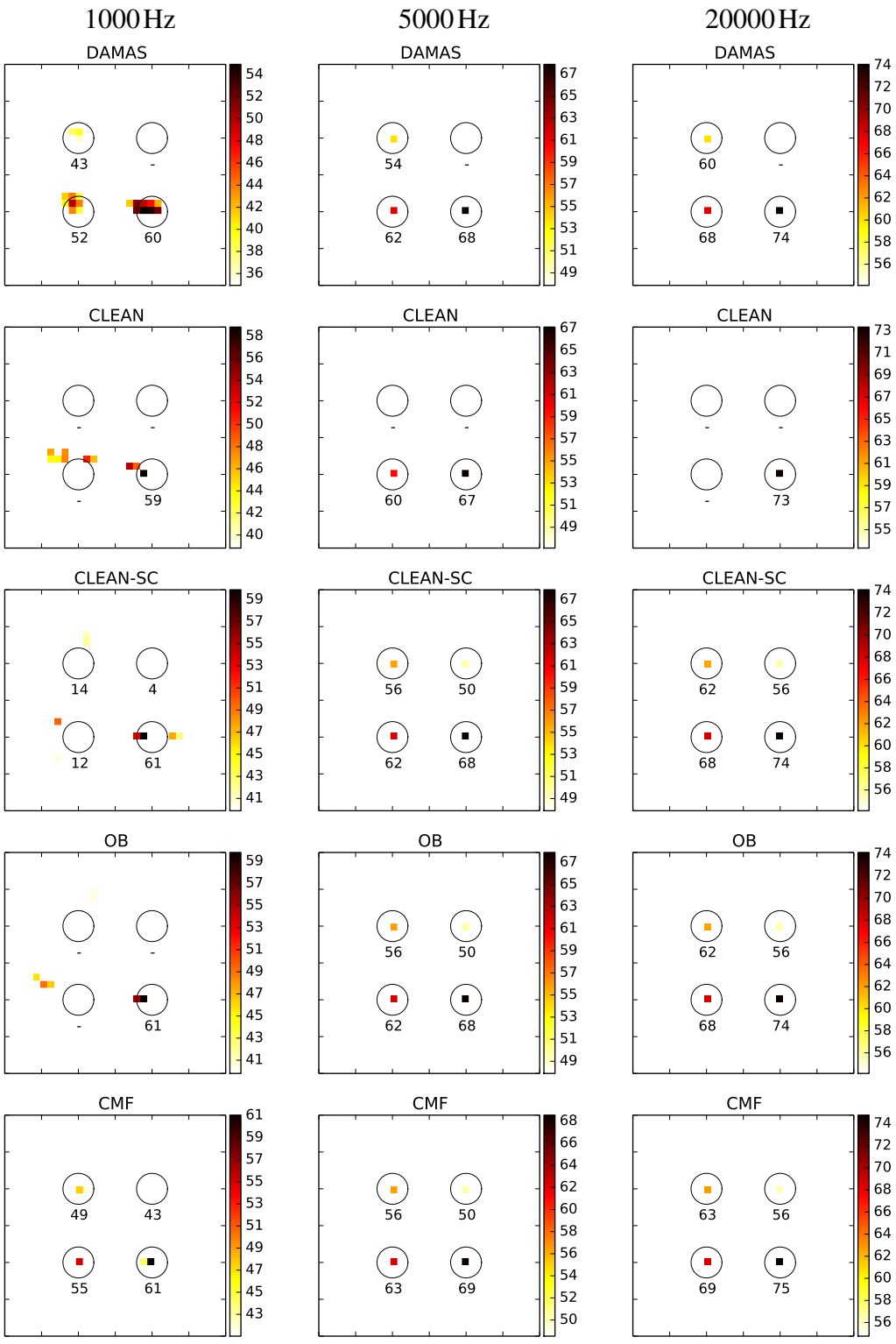

Figure 4: Sound maps evaluated for third-octave bands at 1000 Hz, 5000 Hz, *and* 20000 Hz *using different methods. Four simulated sources with differing levels of* 6 dB *(all values in* dB*).*

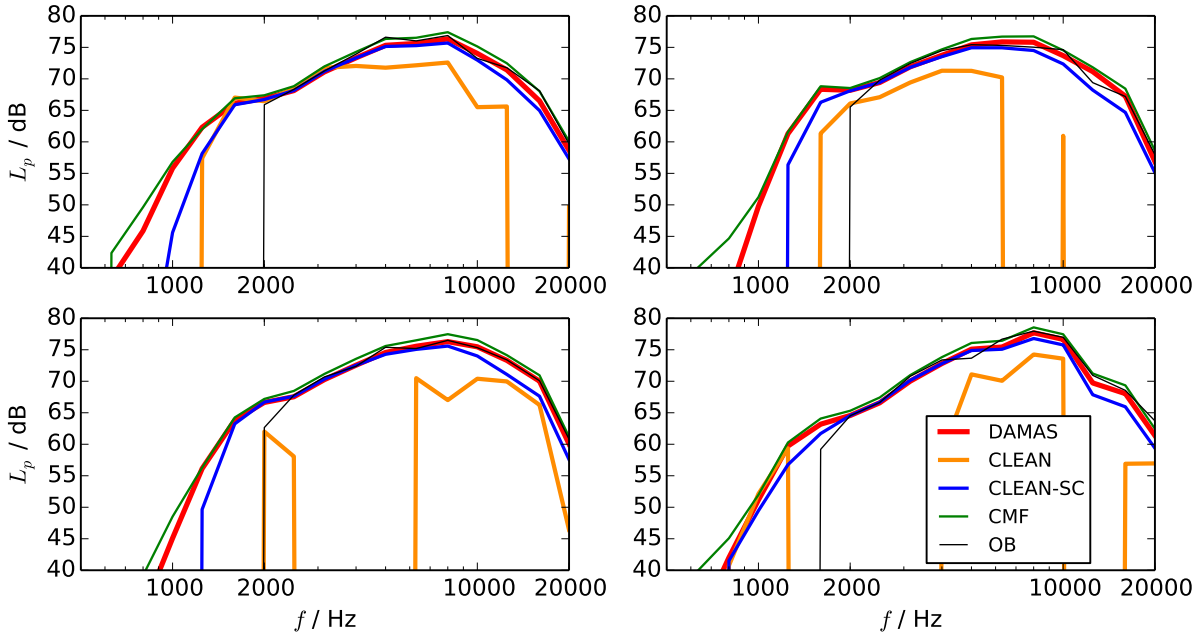

Figure 5: Third-octave band spectra integrated over a circular area ($r = 4\,\mathrm{cm}$) around each loudspeaker. All speakers with approximately the same level. The sub-figures are arranged in the same order as the four integration areas.

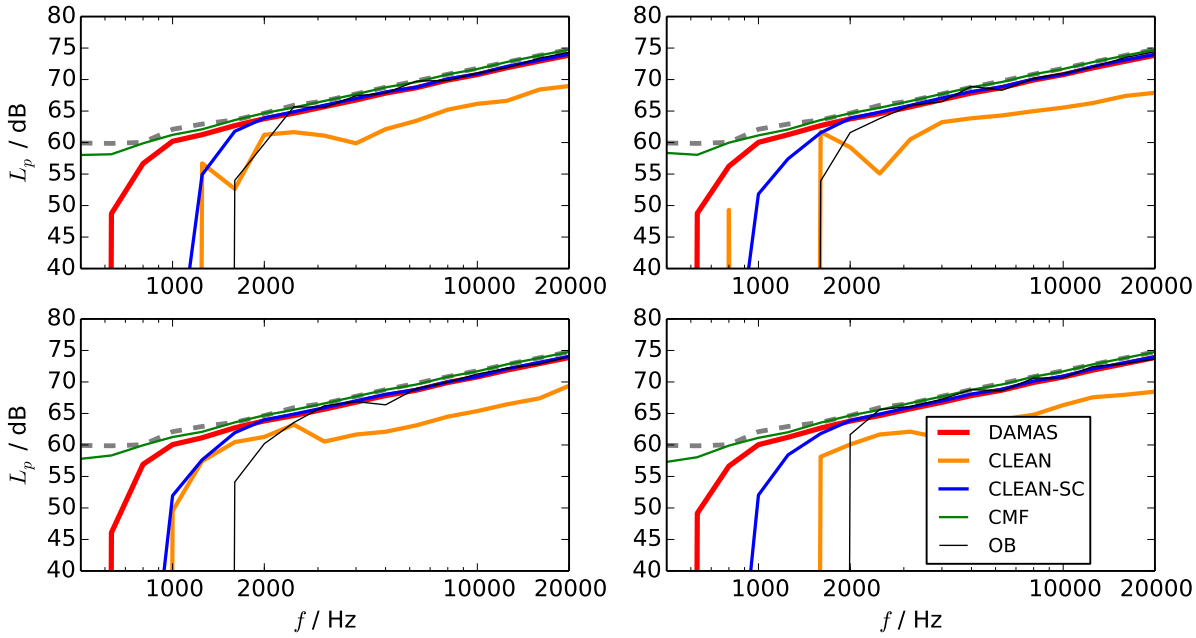

Figure 6: Third-octave band spectra integrated over a circular area ($r = 4\,\mathrm{cm}$) around each simulated source (all with same level). The dotted gray line indicates the actual SPL generated by one simulated source.

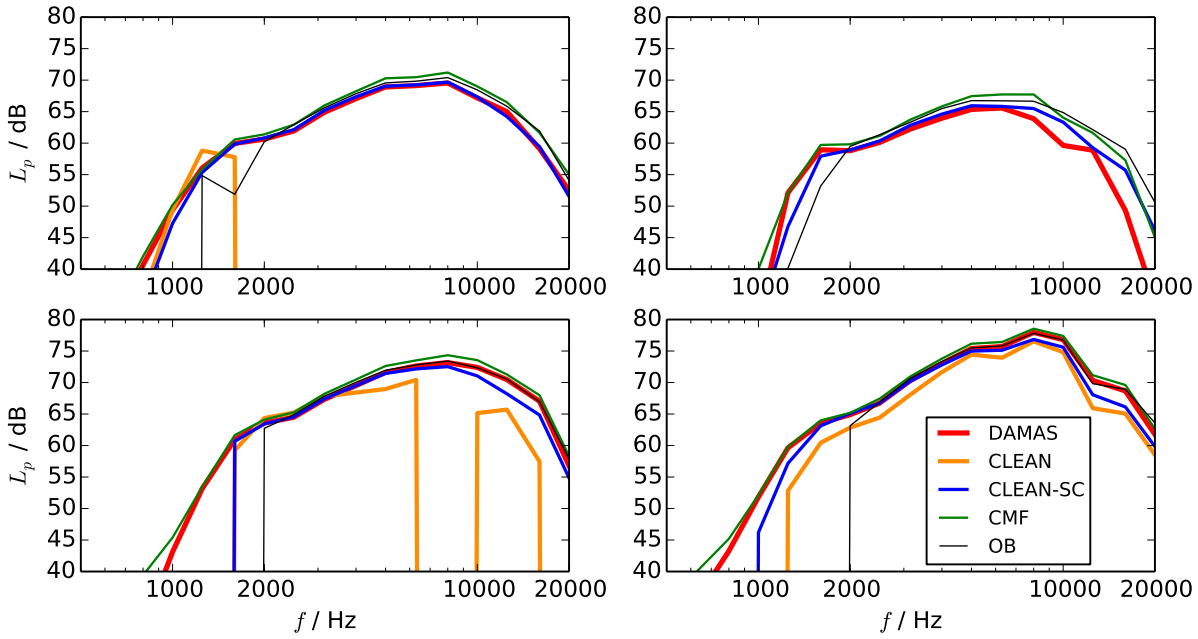

Figure 7: Third-octave band spectra integrated over a circular area ($r = 4\,\mathrm{cm}$) around each loudspeaker. Speakers with 3 dB *level difference.*

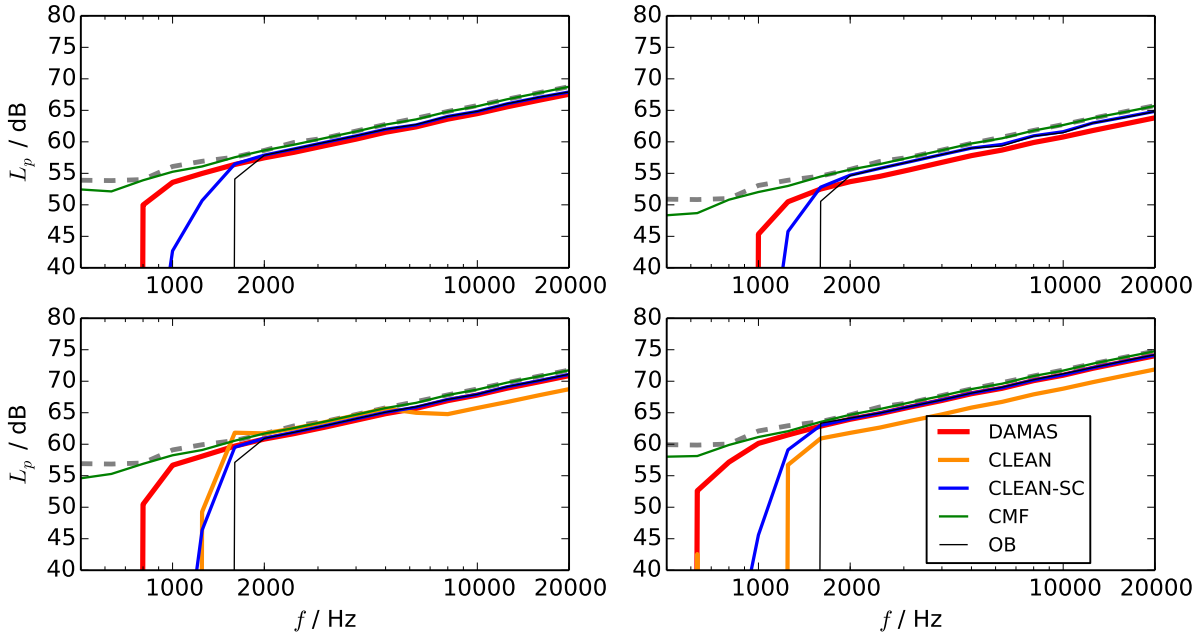

Figure 8: Third-octave band spectra integrated over a circular area ($r = 4\,\mathrm{cm}$) around each source (sources with 3 dB *level difference).*

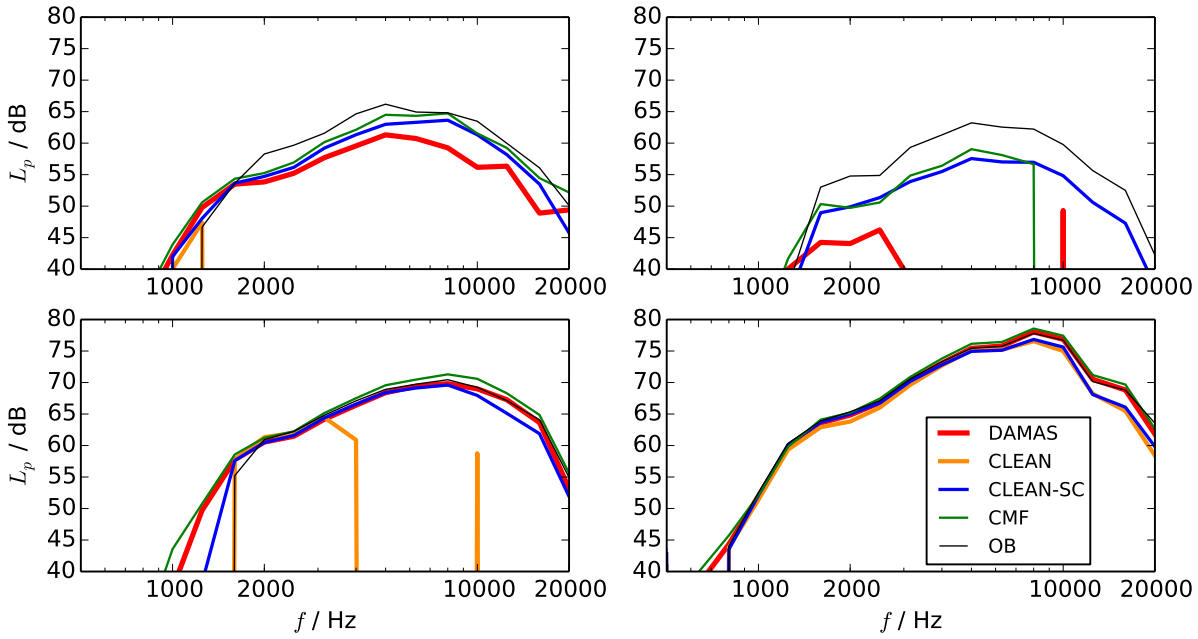

Figure 9: Third-octave band spectra integrated over a circular area ($r = 4\,\mathrm{cm}$) around each loudspeaker. Speakers with 6 dB level difference.

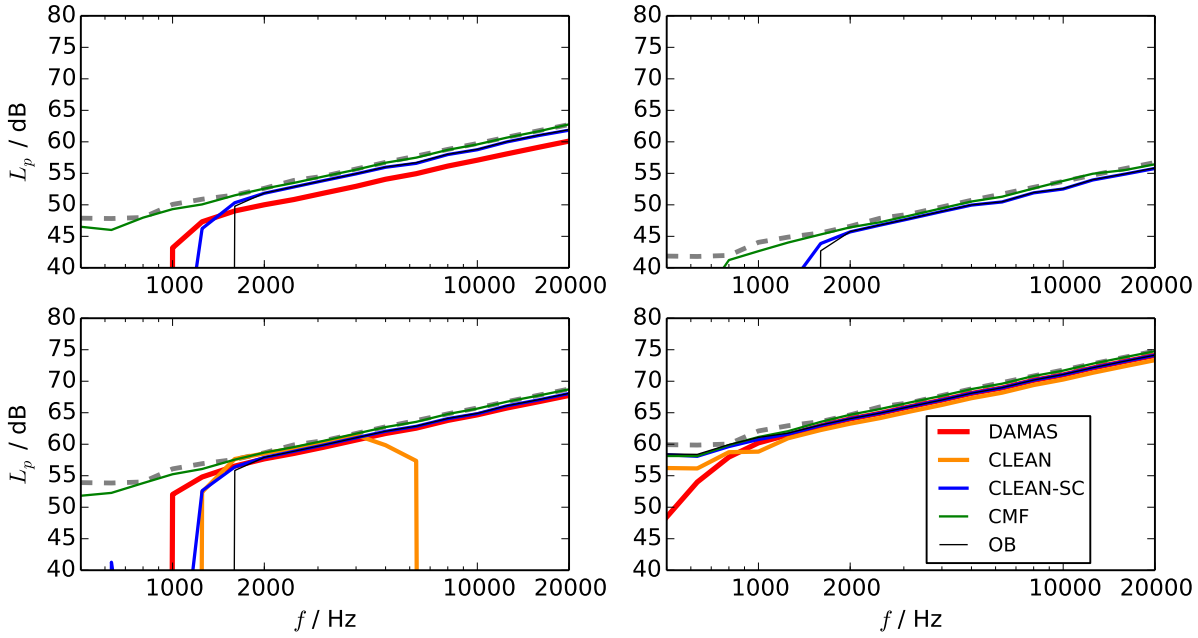

Figure 10: Third-octave band spectra integrated over a circular area ($r = 4\,\mathrm{cm}$) around each source (sources with 6 dB level difference).

References

[1] D. Blacodon and G. Elias. "Level Estimation of Extended Acoustic Sources Using a Parametric Method." *Journal of Aircraft*, 41(6), 1360–1369, 2004. ISSN 0021-8669. doi:10.2514/1.3053.

[2] T. F. Brooks and W. M. Humphreys. "A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays." *Journal of Sound and Vibration*, 294(4-5), 856–879, 2006. ISSN 0022460X. doi:10.1016/j.jsv.2005.12.046.

[3] T. F. Brooks and W. M. Humphreys. "Extension of DAMAS phased array processing for spatial coherence determination (DAMAS-C)." In *Proceedings of the 12th AIAA/CEAS Aeroacoustics Conference*, pages 1–18. 2006.

[4] G. Herold, E. Sarradj, and T. Geyer. "Covariance Matrix Fitting for Aeroacoustic Application." In *AIA-DAGA 2013 Conference on Acoustics*, 11, pages 1926–1928. Merano, 2013.

[5] J. Högbom. "Aperture synthesis with a non-regular distribution of interferometer baselines." *Astronomy and Astrophysics Supplement*, 15(3), 417–426, 1974.

[6] U. Michel and S. Funke. "Noise Source Analysis of an Aeroengine with a New Inverse Method SODIX." In *Proceedings of the 14th AIAA/CEAS Aeroacoustics Conference*. 2008.

[7] E. Sarradj. "A fast signal subspace approach for the determination of absolute levels from phased microphone array measurements." *Journal of Sound and Vibration*, 329(9), 1553–1569, 2010. ISSN 0022460X. doi:10.1016/j.jsv.2009.11.009.

[8] E. Sarradj. "Three-Dimensional Acoustic Source Mapping with Different Beamforming Steering Vector Formulations." *Advances in Acoustics and Vibration*, 2012, 1–12, 2012. ISSN 1687-6261. doi:10.1155/2012/292695.

[9] P. Sijtsma. "CLEAN based on spatial source coherence." *International Journal of Aeroacoustics*, 6(4), 357–374, 2007. ISSN 1475-472X. doi:10.1260/147547207783359459.

[10] R. Tibshirani, I. Johnstone, T. Hastie, and B. Efron. "Least angle regression." *The Annals of Statistics*, 32(2), 407–499, 2004. ISSN 0090-5364. doi:10.1214/009053604000000067.

[11] T. Yardibi, J. Li, P. Stoica, and L. N. Cattafesta. "Sparsity constrained deconvolution approaches for acoustic source mapping." *The Journal of the Acoustical Society of America*, 123(5), 2631–2642, 2008. ISSN 1520-8524. doi:10.1121/1.2896754.

[12] T. Yardibi, J. Li, P. Stoica, N. S. Zawodny, and L. N. Cattafesta. "A covariance fitting approach for correlated acoustic source mapping." *The Journal of the Acoustical Society of America*, 127(5), 2920–2931, 2010. ISSN 1520-8524. doi:10.1121/1.3365260.