BeBeC-2018-S01

ARRAY METHODS: WHICH ONE IS THE BEST?

Ennes Sarradj, Gert Herold and Simon Jekosch
TU Berlin, FG Technische Akustik
Einsteinufer 25, 10587 Berlin, Germany

Abstract

When it comes to process measured array data, the question which processing method to choose arises. With the large number of methods available, it is impossible to clearly decide which method is the best, because all seem to have their advantages and disadvantages. Many different criteria, such as precision of results and speed of computation, can be of interest. The contribution discusses some of these criteria and makes an attempt to define them in a way that they can be estimated from computed results of given test cases. It is also shown that a decision for one method should not be based on the result of one single test case only. Considering this, a Monte-Carlo-method is introduced to compare different methods on the basis of a large number of test cases ($N = 12\,600$). The choice of the test cases is explained and results for multiple criteria and different methods are discussed.

1 INTRODUCTION

If one considers the use of a microphone array for acoustic testing there is a number of things to consider. It has to be decided how to set up the test and which hardware to use. However, since many different signal processing methods are available, thought must be given also to which array data processing method to choose. Among those methods are such that are derived from beamforming, but more interestingly also such that use deconvolution approaches or can be formulated as inverse methods (see [8] for an overview and a slightly different classification). These methods are so numerous and may have very different properties that it is not easy to understand which method(s) would be the best to use in a particular case. However, it is apparent from both theoretical study and from practical application that different methods may behave different and may yield different results. Consequently, the question which one is the best deserves some attention.

Ehrenfried and Koop were among the first who addressed this question to a certain extent by comparing different deconvolution approaches [5] using an example test case. Data from practical array wind tunnel measurements was used for comparison of different methods in

[9] and also in [16]. Comparisons based on synthesized data can be found in [6], in [3] and elsewhere. A benchmarking effort offered multiple test cases with synthesized and measured data sets [1, 12]. These sets were also analyzed with different methods and the results were compared. Moreover, most authors who introduce new methods compare the results for a simple test case with other methods.

If the results from all these comparison approaches were to be used to assess the suitability of a method for a particular application, then one has to rely on the idea that the results depend only on the method and not on the input data. However, as most of the algorithms used for the array processing methods obviously depend on the input data it seems plausible that the data must also be considered somehow in the assessment. Following [7], the present paper concentrates on a possible approach how to include input data properties into the assessment of a method.

2 CRITERIA

The suitability of a method for a particular application can be assessed using various criteria. These could address the ability to characterize acoustic sources with their different properties such as strength, location or directivity or the ease of use, the computational effort or other topics. Which criteria are used and how important they are depends on the case at hand. If only a rough estimate is required, some errors in the determination of source strength can be possibly tolerated better than a long computation time. While it is beyond the scope of this paper to address all possible criteria, some important criteria should be discussed and error metrics for those criteria should be defined in a way that allows for a convenient algorithmic estimation. It shall be noted in passing that the computational cost of a method is often also important, but is not considered here in detail because it depends also on the implementation of the method.

2.1 Definition of error metrics

The ultimate aim of array methods for source characterization is to find estimates for source properties. Consequently, the expected error of these estimates is a measure of the quality of the results from an array method. Usually the error of the estimation is unknown because only the array method results are available.

An alternative option is to use results from cases where the ground truth is available - either using synthesized input data or carefully controlled experiments. This way, the error can be determined just by comparison of estimate and ground truth, where

$$\Delta L_{p,e} = L_{p,estimated} - L_{p,true} \tag{1}$$

is the level difference, an error estimate that can be used for quantities that are given as levels.

A first estimate of interest is the overall level. Ideally, if the scan grid or map region encompasses all sources, the sum of the powers (or mean squares of sound pressures) in the map should give the same value that can be measured by a single microphone at a reference location. The error for the overall level is then given by $\Delta L_{p,e,o}$.

If cases with multiple distinct sources are considered, an estimate of the level for each of the sources can be given. For grid-based methods this is usually achieved by integration over a certain sub-region or sector of the grid. Therefore, this estimator also depends on the size of

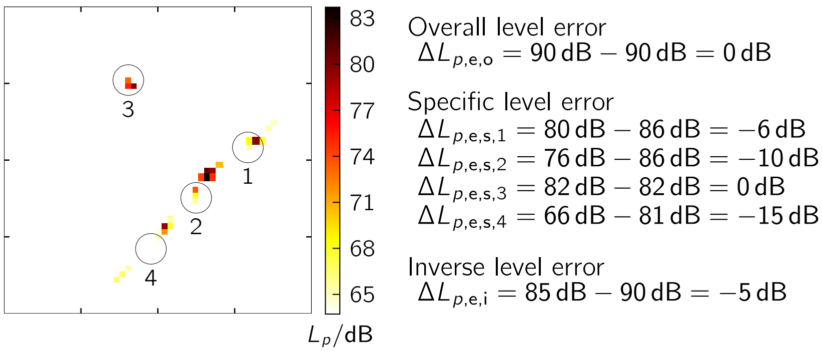


Figure 1: Example of a map with four sources with true source levels of 86 dB, 86 dB, 82 dB and 81 dB. Circles indicate integration regions around the true source positions. The level errors defined in Sec. 2.1 are also shown

this sector. Therefore, the specific level error $\Delta L_{p,e,s}$ depends not only on the source, but also on the integration region used.

Another estimate of interest is the location of the sources. The estimation of the source location is not easy to achieve in a general way. For some methods, like classical delay-and-sum beamforming local maximums in the map can be considered to be the location of sources. Other methods may map sources as "scatter" of some local maximums in close neighborhood. While it may still be possible to assess the location by just "looking at the map", a generally valid estimate for the location which can be computed algorithmically is hard to produce. One computationally expensive option would be the introduction of clustering algorithms that mimic the "look at the map". Another option is the definition of an inverse level error

$$\Delta L_{p,e,i} = L_{p,s,all} - L_{p,o} \tag{2}$$

as the difference of the estimated level of the sum of all source powers and the overall level. If the location of sources is estimated without error, then it can be assumed that no power is in the map outside the integration regions. Thus, all sources together yield the same result as the overall level. While the inverse level error characterizes the quality of the source location estimate, it includes other effects from which the result may suffer: scattering of sources over an area larger than the integration region, "melding" of neighboring sources (insufficient spatial resolution) and spurious sources that are erroneously detected by the method.

Fig. 1 demonstrates the error metrics defined here. A case with four sources is considered. The ground truth is known for all the estimates considered here. Integration of source powers is performed over circular sectors.

2.2 A benchmark test case illustration

In order to illustrate the level error metrics, a test case from a benchmark exercise [12] shall be considered here. The test case consists of four uncorrelated sources emitting white noise and has two subcases. In subcase a all sources have the same power, while in subcase b the sources have different powers (0, -6, -12, -18 dB). Fig. 2 shows the setup chosen for the test case. An

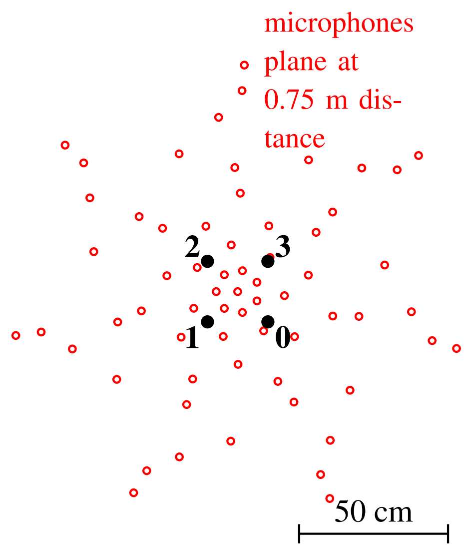
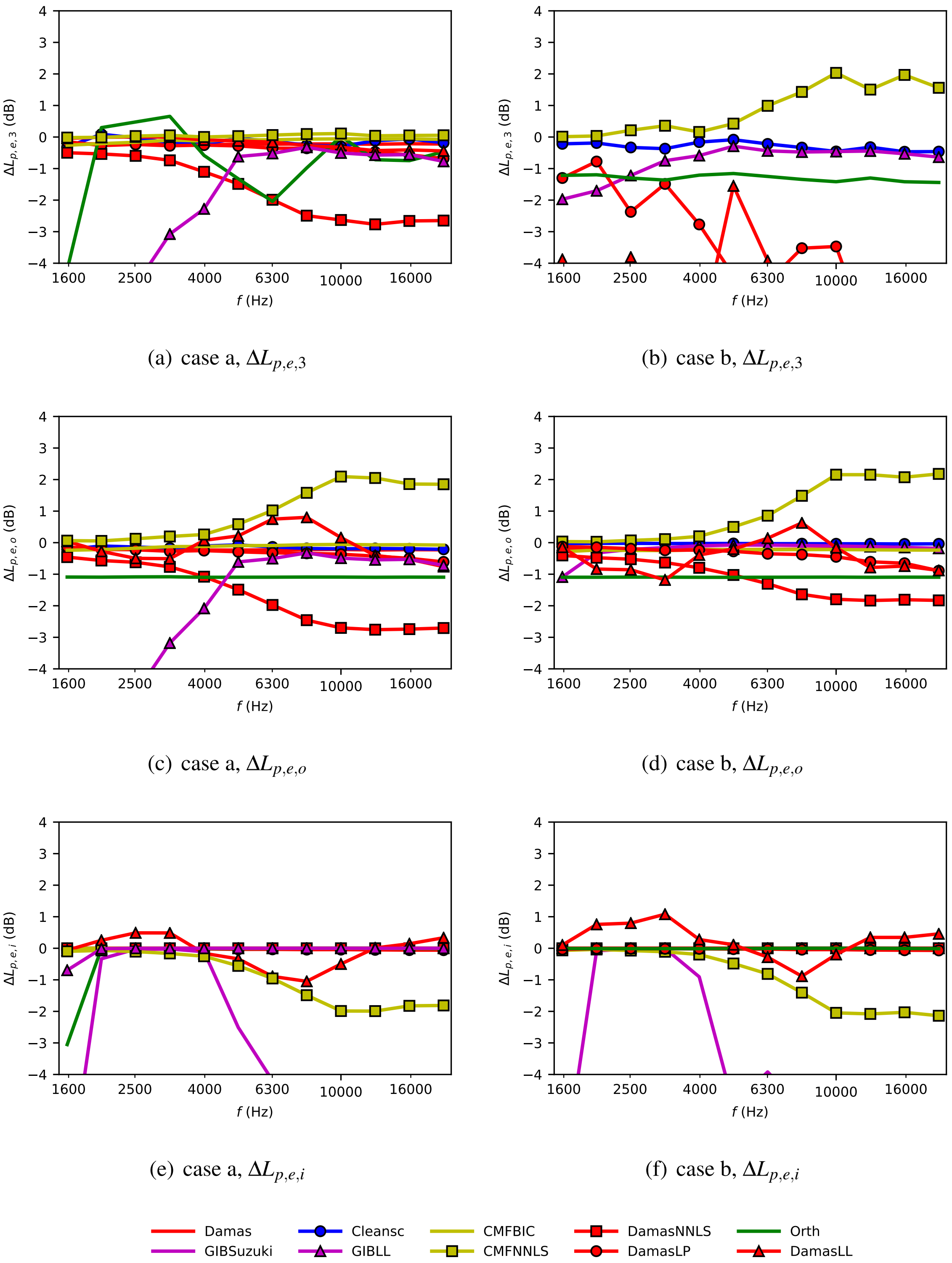


Figure 2: Setup of the test case from [12]. The black dots show the location of the four sources and red dots show the microphone location of a 64 microphone array

array with 64 microphones is used. For the results to be shown, different methods with an open source implementation [11] were used on identical input data:

- DAMAS [2] with the original modified Gauss-Seidel solver as well as with a nonlinear least squares (NNLS) solver, a sparsity enforcing Lasso LARS (LL) solver [15], and linear programming [4]
- Clean SC [13]
- orthogonal deconvolution [10]
- Generalized inverse beamforming [14] with the original solving strategie and with LL solver
- Covariance matrix fitting (CMF) with NNLS and LL solvers

The error metrics computed from the results are shown in Fig. 3. Two important observations can be made. The first is that for nearly all methods the errors are quite low and the results seem to be nearly perfect. This is due to the synthesized input data with only four sources and no additional noise or any uncertainty about microphone position and correct value for the speed of sound. Second, while there some differences between the individual methods are mostly small, the results are somewhat different for the both cases, even when the same method is considered. This means, the benchmark will produce some assessment of the performance of the method for a specific example, but it remains uncertain how a particular method will behave in other cases. Thus, it seems to be difficult to draw a meaningful conclusion from the benchmark in order to decide on a method to apply in a specific case. However, the four sources benchmark can still be used to quickly check whether a method works as intended and also to check the implementation of a method.



(a) case a, $\Delta L_{p,e,3}$

(b) case b, $\Delta L_{p,e,3}$

(c) case a, $\Delta L_{p,e,o}$

(d) case b, $\Delta L_{p,e,o}$

(e) case a, $\Delta L_{p,e,i}$

(f) case b, $\Delta L_{p,e,i}$

Figure 3: Results for both subcases a of the benchmark four point source case.

2.3 Monte Carlo method

Because the quality of the results also depends on the input data, the method assessment preferably needs to be based on input data similar to what can be expected in the respective application. Therefore, an approach is useful where many different cases are considered. A proper assessment is then possible using just the cases that are “similar” to the case of interest.

In order to consider many different cases, the procedure demonstrated in section 2.2 must be repeated many times. The synthesized cases have to consider all the variability that might be of interest. In what follows, test cases are considered that assume a finite number of point sources, each emitting white noise. The sources are assumed to have different levels and are placed within a focus plane in a distance of half an aperture away from an array of 64 microphones. The nominal microphone positions are the same for all test cases. In order to produce different test cases, the number of sources N, the source positions within the source plane and the individual source levels are varied.

Because of the use of synthesized input data, the data perfectly fits the theoretical model. An assessment using this data does not account for the problems that may arise when the data and model deviate somewhat. However, this is always the case for a practical measurement. To include some effects like deviations in the speed of sound and slightly inhomogeneous media properties, the microphone positions used for the synthesis were slightly disturbed. For the analysis, the nominal positions were used. This is equivalent to some uncertainty one may have about the actual microphone positions and therefore about travel times between sources and microphones.

For the results to be shown here, a total of 12600 different cases were analyzed. The array layout and histogram of disturbed source positions are shown in Fig. 4. Instead of a fixed dimension for the aperture, all dimensions in the analysis are given as multiples of the aperture d. Consequently, frequencies f are given in the form of dimensionless Helmholtz numbers $He = fc/d$, where c is the speed of sound. This way, the results can be adapted to any array size and frequency.

The number of sources used for each of the 12600 cases was drawn from a Poisson distribution with $\lambda = 3$. This resulted in cases with sources numbers N between 1 and 13 (4 sources on average, see Fig. 5(a)). This choice was made on the basis of the idea that in most practical scenarios the number of relevant sources is in that range.

The positions of the sources were drawn from a bivariate normal distribution (5(b)). The rationale for using a normal distribution was that in most practical situations the array center is likely to be approximately aligned with the sources of interest.

The source powers were drawn from a Rayleigh distribution of mode $\sigma_R = 5$. This distribution was used because the strength of some naturally occurring phenomena like wind and water waves can be modeled using a Rayleigh distribution. Fig. 6 shows the probability distribution of source powers along with the resulting distribution of minimum and maximum level differences for a particular case. The distribution parameter σ was chosen to get maximum differences that are not too large. This corresponds to many practical applications were the focus is on sources that are not too weak compared to the dominating source.

The data from 12600 test cases was applied as input data to four different methods:

- DAMAS using a modified Gauss-Seidel solver with 500 iterations
- orthogonal deconvolution (OB) using 20 eigenvalues

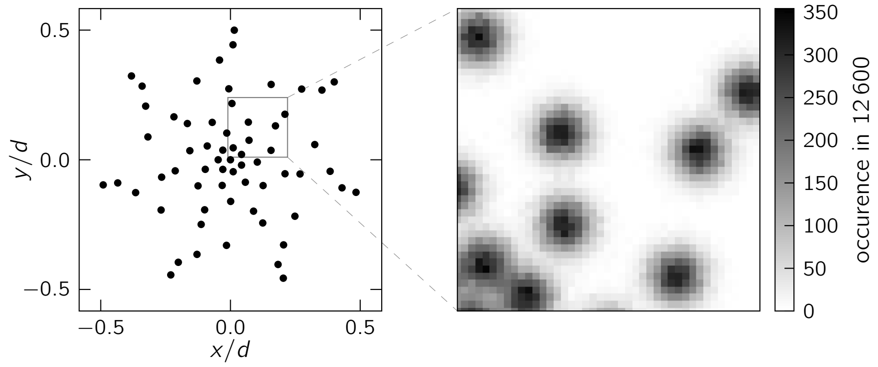
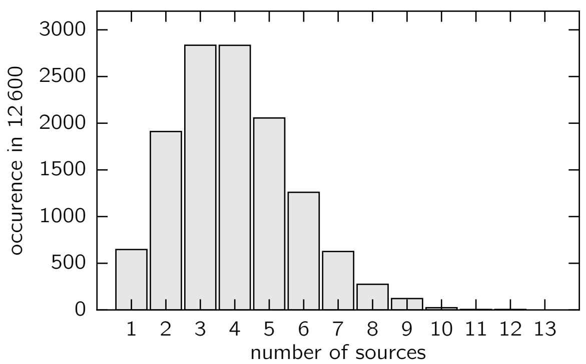
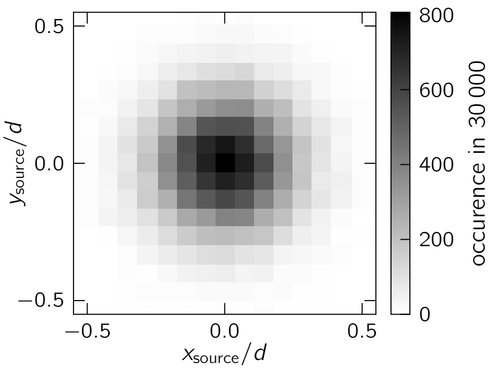


Figure 4: Array layout with 64 microphones in one plane. Nominal positions and representative section with histogram of disturbed microphone positions (12600 cases).



(a) Histogram for number of sources drawn from the Poisson distribution ($\lambda = 3$, 12600 samples).



(b) 2D-histogram for 30000 source positions drawn from the bivariate normal distribution ($\sigma = 0.1688$, values constrained to $x, y \in [-0.5\,d, 0.5\,d]$).

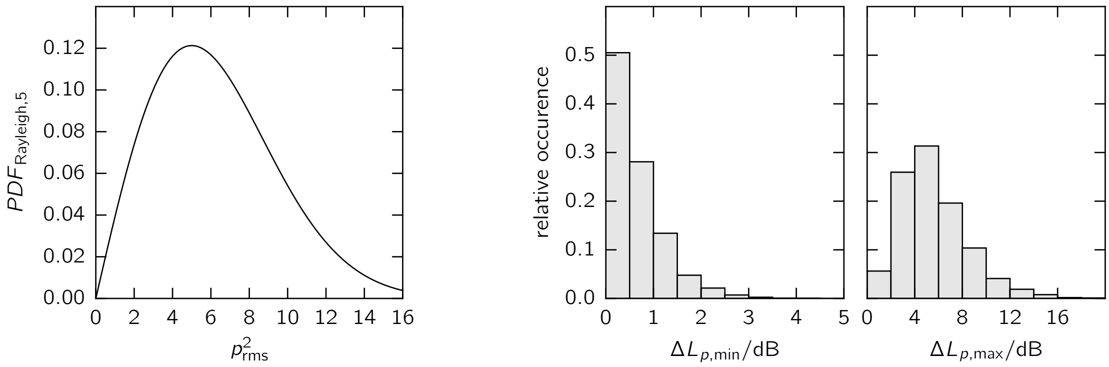
Figure 5: Distributions of source number and source positions

- Clean SC with 500 iterations
- Covariance matrix fitting (CMF) with an LL solver and automatic adjustment of the regularization parameter using the bayesian information criterion [17]

For all cases and all methods, the results were computed for a frequency range of $He = 1 \ldots 16$. Overall, 61100 CPU hours were spent. The methods implemented in Acoular [11] took 69.6%, 0.5%, 1.2%, and 28.7% of this time for DAMAS, OB, Clean SC, and CMF, respectively.

3 RESULTS

The large database of results allows to assess the method performance for many different source scenarios. Because even for the large number of test cases it is very probable that no case exactly



(a) Probability density function for the Rayleigh distribution (mode $\sigma_{\mathrm{R}} = 5$).

(b) Scaled histogram of the minimum and the maximum level differences occurring between 4 source levels randomly drawn from the Rayleigh distribution.

Figure 6: Distribution of source levels

matches the application that is considered and for which the best method has to be found, it makes sense to do a statistical analysis of the results. This analysis can be done for each of the criteria considered (such as the different level errors) and for a number of different parameters of the test case scenarios. Here, only a subset of these is considered. The interested reader is referred to [7], where some more results are discussed.

The following parameters are among those that might be of interest:

- the source rank as defined by a given source being the strongest, second strongest, third strongest and so on up to the weakest source of a given test case
- the level difference of a given source to the strongest source in the test case
- the number of sources within a given test case
- the geometric scattering of the sources in the case

The geometric scattering can be defined as the mean of all source distances in one data set $\langle \Delta r_{\mathrm{all}} \rangle_{\mathrm{geom}}$.

Fig. 7 shows how the specific level error depends on the source rank. Additionally, the frequency dependency is shown. It becomes obvious that the methods exhibit a different behavior with respect to the Helmholtz number. While the fast methods OB and Clean SC perform well at higher *He*, DAMAS and CMF work with less error at low *He*. With the exception of OB, all methods show a slightly increasing error towards very high frequencies, where the true source powers are systematically underestimated. The source rank seems to have a small influence on the error, which is generally higher for the weakest source. This effect can be detected for all methods, but is somewhat stronger for DAMAS and CMF.

The dependency on the level difference to the strongest source shows a more pronounced picture. Fig. 8 gives this dependency for the specific level error. The mean of this error is relatively small for both OB and Clean SC and does not show a strong dependency on the level difference. However, like in Fig. 7 at low *He* the error can be very large for these methods. For

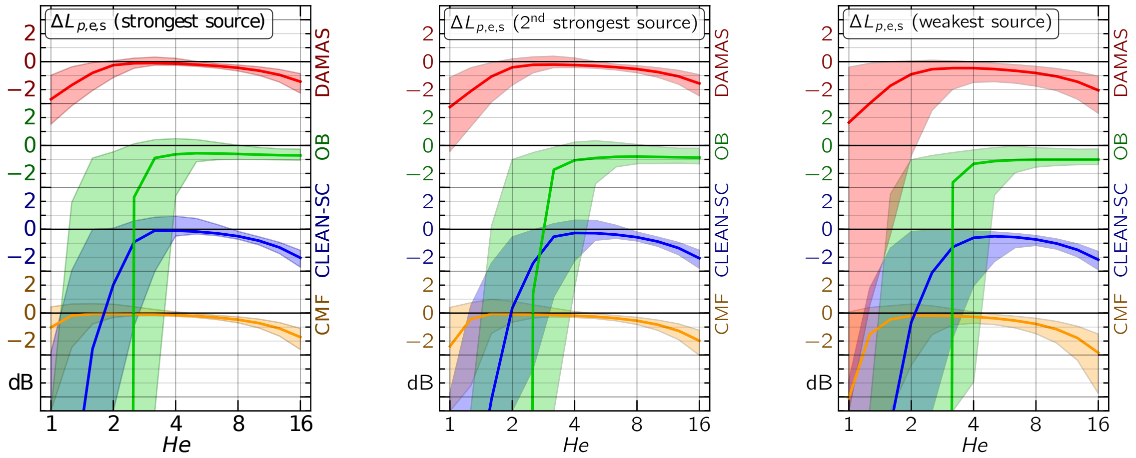


Figure 7: Specific level error $\Delta L_{p,\mathrm{e,s}}$ for the strongest, the second-strongest, and the weakest source in each data set. The colored lines show the median of all level errors for the respective methods, and the filled areas represent the interval between the 16th and 84th percentiles.

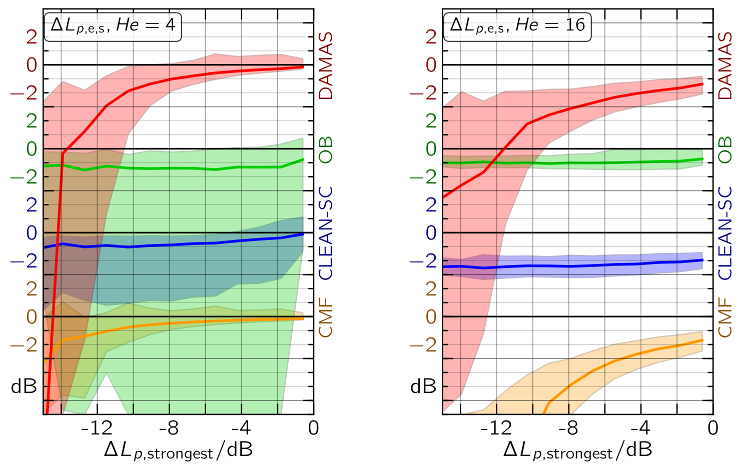


Figure 8: Specific level error depending on the level difference to the strongest source in the data set for He = 4 and He = 16.

DAMAS and CMF a large level difference leads to large errors. This means that these methods do not perform well with minor sources.

The specific level error as a function of the number of sources is shown in Fig. 9. As it can be expected, the error becomes larger with the number of sources for all methods. However, this is not a strong effect except for OB and Clean SC and low frequencies. Consequently, the performance with the number of sources is not helping to find the best method for an certain application case.

Finally, the dependency on geometric scattering of the sources shall be considered at high frequencies. Fig. 10 shows the overall level error as a function of $\langle \Delta r_{\mathrm{all}} \rangle_{\mathrm{geom}}$. For Clean SC and CMF the overall level error increases when the sources are scattered over larger area. This

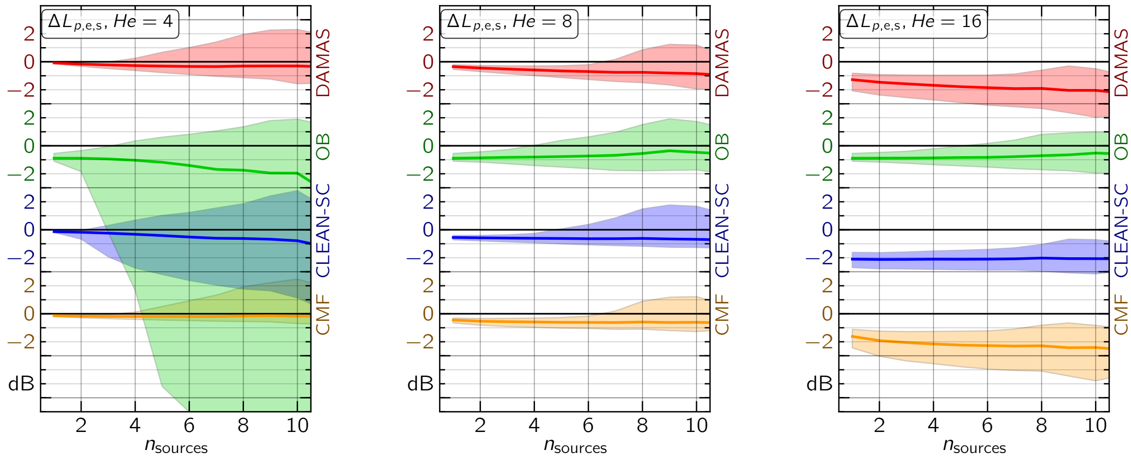


Figure 9: Specific level error for $He = 4$, $He = 8$, $He = 16$ as a function of the number of sources.

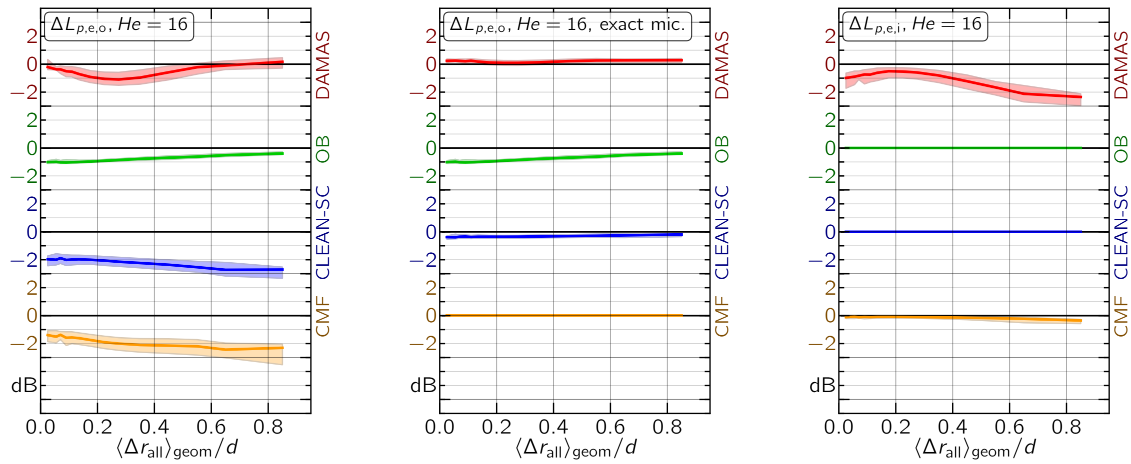


Figure 10: Dependency on the geometric scattering of the sources (geometric mean of all source distances in one data set) at $He = 16$: overall level error, overall level error with undistorted microphone positions, and inverse level error.

effect is much smaller for DAMAS and OB. Interestingly, if the same dependency is considered for undisturbed microphone positions, the overall level error nearly vanishes for all methods. This means that DAMAS and OB can tolerate missing information about the precise location of the microphones better than CMF and Clean SC can. Fig. 10 shows the inverse level error, which is nearly zero for all geometric scattering and all methods except DAMAS. For DAMAS some spurious sources are present in the map if the sources are distributed over a larger area.

4 CONCLUSIONS

These examples of statistical analysis of the results show that the performance of a method depends on the source scenario and that it is not possible to answer the question about the best method in general, but only after the definition of proper error metrics and the use of input data that resembles the application case. The error metrics that are used may vary and the level errors defined are just one possibility.

The procedure for statistical analysis demonstrated here offers some insight into the behavior of the different methods. However, if a real application is considered all the parameters such as number, strength and position of sources are not known a priori. An extension of the statistical procedure could work around this problem by estimating the necessary parameters a posteriori, looking up the best method and then using the results from this method.

Finally, the best method may depend on the frequency of interest. If a larger frequency range is considered, it may even be that a certain application case calls for the use of different methods for different frequencies.

REFERENCES

[1] C. J. Bahr, W. M. Humphreys, D. Ernst, T. Ahlefeldt, C. Spehr, A. Pereira, Q. Leclere, C. Picard, R. Porteous, D. Moreau, et al. "A comparison of microphone phased array methods applied to the study of airframe noise in wind tunnel testing." In *23rd AIAA/CEAS Aeroacoustics Conference, AIAA-paper 2017-3718*. 2017.

[2] T. F. Brooks and W. M. Humphreys. "A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays." AIAA paper 2004-2954, 2004.

[3] Z. Chu and Y. Yang. "Comparison of deconvolution methods for the visualization of acoustic sources based on cross-spectral imaging function beamforming." *Mechanical Systems and Signal Processing*, 48(1-2), 404–422, 2014. ISSN 10961216. doi:10.1016/j.ymssp.2014.03.012. .

[4] R. Dougherty, R. Ramachandran, and G. Raman. "Deconvolution of sources in aeroacoustic images from phased microphone arrays using linear programming." *International Journal of Aeroacoustics*, 12(7-8), 699–718, 2013. ISSN 1475-472X. doi:10.1260/1475-472X.12.7-8.699. .

[5] K. Ehrenfried and L. Koop. "Comparison of Iterative Deconvolution Algorithms for the Mapping of Acoustic Sources." *AIAA Journal*, 45(7), 1584–1595, 2007. doi:10.2514/1.26320.

[6] G. Herold and E. Sarradj. "Preliminary Benchmarking of Microphone Array Methods." In *Proceedings of the 5th Berlin Beamforming Conference*, pages 1–13. Berlin, 2014.

[7] G. Herold and E. Sarradj. "Performance analysis of microphone array methods." *Journal of Sound and Vibration*, 401, 152–168, 2017. ISSN 0022-460X. doi:10.1016/j.jsv.2017.04.030. .

[8] Q. Leclere, A. Pereira, C. Bailly, J. Antoni, and C. Picard. "A Unified Formalism for Acoustic Imaging Techniques : Illustrations in the Frame of a Didactic Numerical Benchmark." In *Proceedings of the 6th Berlin Beamforming Conference*, pages 1–17. 2016.

[9] E. Sarradj. “Aeroacoustic source characterisation using subspace and deconvolution techniques.” In *Noise and Vibration: Emerging Methods*. 2009.

[10] E. Sarradj. “A fast signal subspace approach for the determination of absolute levels from phased microphone array measurements.” *J. Sound Vib.*, 329(9), 1553–1569, 2010. doi: 10.1016/j.jsv.2009.11.009.

[11] E. Sarradj and G. Herold. “A Python framework for microphone array data processing.” *Applied Acoustics*, 116, 50–58, 2017. ISSN 0003682X. doi:10.1016/j.apacoust.2016.09.015. .

[12] E. Sarradj, G. Herold, P. Sijtsma, R. Merino Martinez, T. F. Geyer, C. J. Bahr, R. Porteous, D. Moreau, and C. J. Doolan. “A Microphone Array Method Benchmarking Exercise using Synthesized Input Data.” In *23rd AIAA/CEAS Aeroacoustics Conference*, June. 2017. ISBN 978-1-62410-504-3. doi:10.2514/6.2017-3719. .

[13] P. Sijtsma. “CLEAN Based on Spatial Source Coherence.” In *13th AIAA/CEAS Aeroacoustics Conference, Rome, Italy, May 21-23, 2007*, pages AIAA–2007–3436. 2007.

[14] T. Suzuki. “Generalized Inverse Beam-forming Algorithm Resolving Coherent/Incoherent, Distributed and Multipole Sources.” AIAA-2008-2954, 2008.

[15] 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.

[16] T. Yardibi, N. S. Zawodny, C. Bahr, and F. Liu. “Comparison of microphone array processing techniques for aeroacoustic measurements.” *International Journal of Aeroacoustics*, 9(6), 733–762, 2010. doi:10.1260/1475-472X.9.6.733. .

[17] H. Zou, T. Hastie, and R. Tibshirani. “On the ’degrees of freedom’ of the lasso.” *Annals of Statistics*, 35(5), 2173–2192, 2007. ISSN 00905364. doi:10.1214/009053607000000127.