



Industrial sound source localization using microphone arrays under difficult meteorological conditions

Dick BOTTELDOOREN¹, Timothy VAN RENTERGHEM², Frits VAN DER EERDEN³, Peter WESSELS⁴, Tom BASTEN⁵, Bert DE COENSEL⁶

^{1,2,6} Ghent University, Belgium

^{3,4,5} TNO, The Netherlands

⁶ ASAsense, Belgium

ABSTRACT

Source localization using microphone arrays has now become common practice for many applications. Their application for long term monitoring of industrial noise nevertheless remains uncommon. Hence, in 2016 we deployed four microphone arrays near a large industrial area for a period of 12 months and explored the challenges involved. The difficult meteorological conditions leading to strongly varying attenuations, 3D refraction, and scattering turned out to be one of these major challenges. Changes in predicted source locations and estimated source power are quite often the result of abrupt changes in local weather conditions. Notwithstanding the careful weather monitoring, assumptions and hypotheses on the sources' temporal behavior and spectra combined with a statistical locator were still needed to obtain satisfactory results. Validation of the proposed approaches on a known source with stable noise emission over time is presented in the paper.

Keywords: Microphone Array, Industrial Noise,
I-INCE Classification of Subjects Number(s): 72, 12

1. INTRODUCTION

When noise complaints arise near large industrial areas such as harbors, it is often quite difficult to pinpoint the main source of noise at the time of the complaint. Often distances between the source and receiver are rather large and therefore noise levels at the receiver strongly depend on instantaneous meteorological conditions. Microphone arrays have now become common tools for finding the direction of a source both in indoor and outdoor conditions (1). However, the use of such arrays for semi-permanent monitoring of industrial noise remains uncommon.

In military context however, very low frequency arrays have been deployed around the world to detect high intensity blasts (2) since the days of the cold war. Civil applications include the detection of fireworks, volcanic activity, etc. In the majority of these applications the sound source produces a short impulse-like signal. Arrival times of this impulse at different sensor locations is the core information used in the algorithms.

Within the industrial context more continuous sounds are expected, hence algorithms that were developed originally for small arrays – where small is defined by the source being in the far field of the array – and to operate in a homogeneous and time invariant propagation medium, were extended to cope with the outdoor propagation conditions and the wider distribution of microphones. The vast majority of algorithms developed for such small arrays uses cross correlation between the signals

¹ dick.botteldooren@ugent.be

² timothy.vanrenterghem@ugent.be

³ frits.vandereeden@tno.nl

⁴ peter.wessels@tno.nl

⁵ tom.basten@tno.nl

⁶ bert.decoensel@asasense.com

received at different sensors as a first step. For linear time invariant propagation, the impulse response and frequency response provide exactly the same information on the propagation, and thus allow to estimate the location of the source with the same precision. In this case a broadband sound source would be as easily located as a short impulse. However, atmospheric turbulence makes the propagation environment time variant. Several researchers have tackled the problem of coherence loss that was identified as a limiting factor on outdoor use of extended sounds sensor networks which is adding another factor to the Cramer–Rao lower bound (CRLB) (3). Often a distinction is made between close sensors forming a sub-array and sensors that are further apart. In (4) it is shown that bearing estimation from the sub-arrays combined with time delay estimation between sub-arrays performs as good as a fully coupled estimator once coherence loss between sub-arrays occurs. Explicitly accounting for loss of coherence between more distant parts of the array considered as a whole, can lead to improved performance of maximum likelihood estimators (5). Others have attempted to make signal subspace techniques more robust against inaccuracies of the estimation of the spatial covariance matrix (6).

In addition to inducing coherence loss, meteorological conditions also directly affect signal to noise ratio at the microphone arrays. Wind increases noise in the array microphones, in particular at low frequencies. In addition to wind noise induced in the microphone, also turbulence carried by the wind is of major importance. Meteorological conditions also influence signal levels as upward refraction could significantly reduce noise levels at the distances from the sources that are of interest.

This manuscript investigates the effect of meteorological conditions on the localization of industrial sound sources in a harbor area of 8 by 6 kilometer. Around this area, four microphone arrays, each consisting of 40 non-equally spaced microphones have been deployed for 12 months. In view of the above mentioned effects of coherence loss, signals from the four arrays are processed individually and combined afterwards. The technological challenges involved have been discussed in (7), the estimation of propagation in (8), and the fusion of data from different arrays in (9). Here we elaborate on the effect of meteorological conditions on performance of the system and introduce uncertainty estimates as a way to account for them during data fusion.

2. Measurement setup

2.1 Arrays and additional microphones

To cover the study area, 4 microphone arrays consisting of 40 microphones each and 6 additional microphones are placed around the industrial area. Technological details can be found in (9). Figure 1 shows the location of these sensors on a sketch of the area layout. Tick marks on the axis are shown every 2 kilometer and give an impression of the size of the area. Data from the arrays are collected for about 30 seconds every 10 minutes during one year. The 10 single microphones monitor the environment using 1/8 second 1/3 octave band spectra continuously.

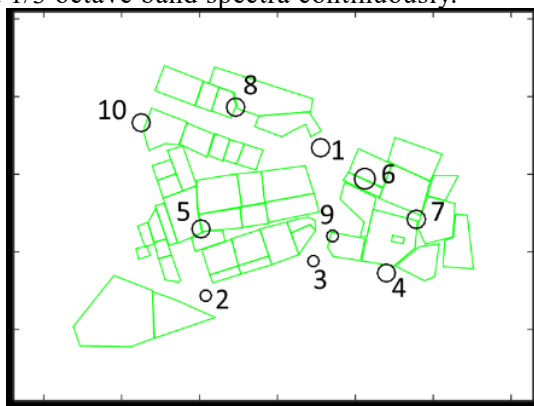


Figure 1 – Schematic layout of the industrial area with the arrays (numbers 1 till 4) and additional microphones (numbers 5 till 10) indicated.

2.2 Weather stations and meteorological data

Four weather stations are deployed in the area to carefully monitor local wind and temperature as well as gradients in wind and temperature. The masts provide data each minute on temperature, humidity, pressure, rain, wind speed and direction. The wind direction and speed are measured at a height of 10 meters. To determine the time varying sound propagation between each source area and

each receiver position the wind, temperature and humidity needs to be known, preferably at several heights. Meteorological data from HiRLAM (High Resolution Limited Area Model) is used which provides data at several heights (10 levels with a maximum of 700 meters are used). Data assimilation is then applied to combine the HiRLAM data with the measurements to obtain data for periods of 10 minutes.

3. Detection of a reference source under varying meteorological conditions

3.1 Calculated contribution of the reference source to overall level at the arrays

One of the sound sources in the industrial area is well known and operates continuously. It constantly emits a broad band noise between 100 and 630 Hz but consists of a large number of smaller incoherent sound sources. In Figure 1 it is shown as the small rectangle close to microphone 7. To illustrate the effect of weather conditions on the localization performance of the array, a specific summer night where a mild wind turns from east to west, is investigated. The assumption that the source under study emits a constant sound power is easily validated by analyzing L_{A50} at microphone 7. This measurement remains within 5 dB of the theoretical estimate during the whole night for all 1/3-octave bands between 100 and 630 Hz. Figure 2 compares the calculated levels at 250 Hz with noise level measurements near the four arrays starting at 21:00 UTC or 23:00 local time. For reference the microphone wind and sensor noise is also shown (see Section 4). According to (3) the signal to noise levels observed would cause the CRLB to be determined by signal to noise ratio rather than to coherence loss, except for array 4. In the latter, the distance between the reference source and the array is less than 1 km which would in turn limit the effect of coherence loss.

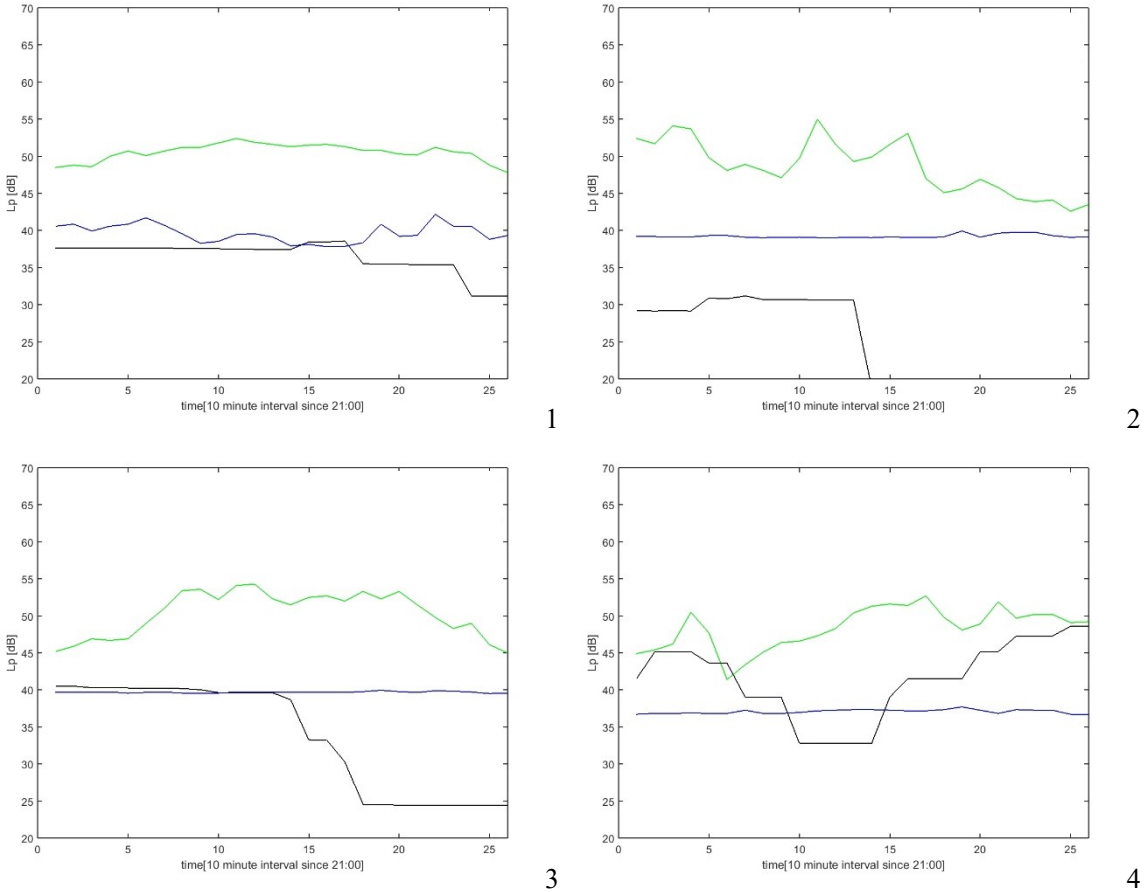


Figure 2 – Comparison of sound levels (L_{50}) at a microphone close to each of the four arrays (green line) with predicted contribution of the known source (black line) and array microphone wind noise and sensor noise (blue line); 250Hz 1/3 octave band; the time axes represents 10 minute observation intervals.

3.2 Atmospheric conditions

The night period that is observed starts at 21:00 UTC or 23:00 local time on a summer day in August. This is well after sunset and thus a temperature inversion starts building up. Figure 3 shows the temperature and wind profiles at the start of the night under investigation and at the end measured at the meteorological tower co-located with Array 4 and extended to greater heights using the HiRLAM model close to the reference source. With the observed temperature inversion and relatively low wind speeds a stable atmosphere and thus low degrees of turbulence are expected.

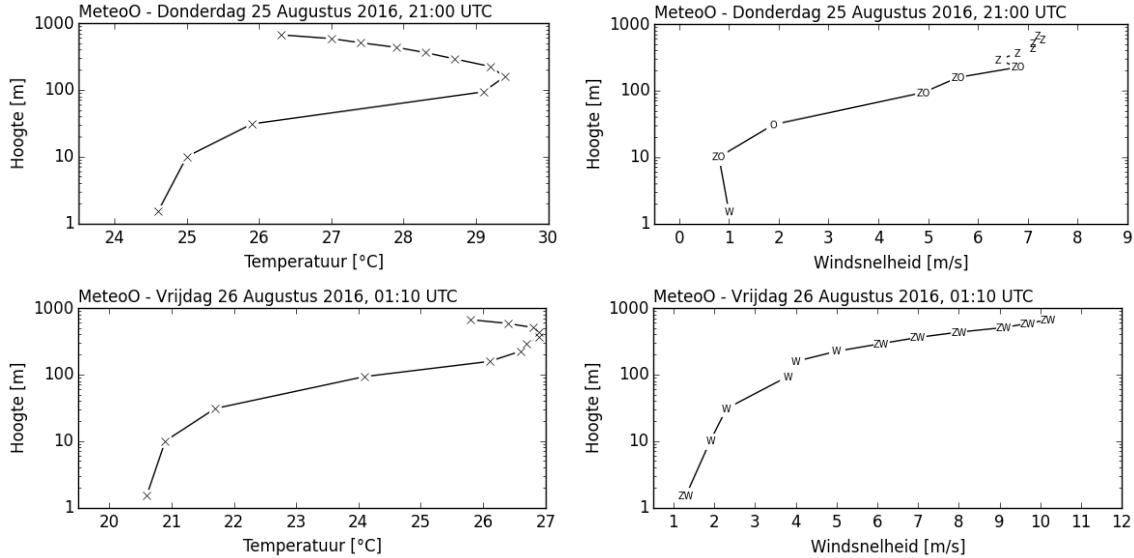
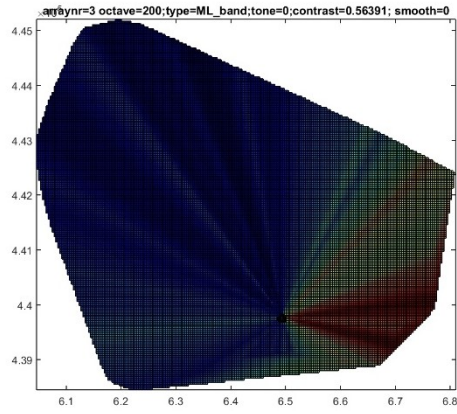
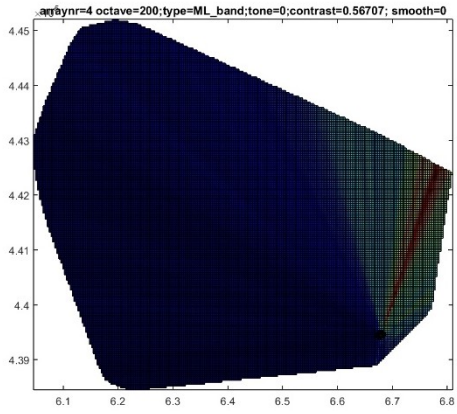


Figure 3 – Temperature as a function of height (left) and wind speed as a function of height (right) at the beginning (top) and the end (bottom) of the focus night.

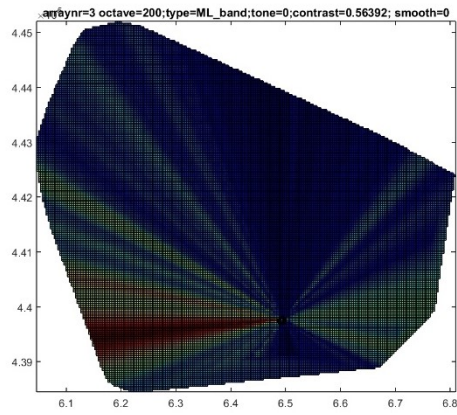
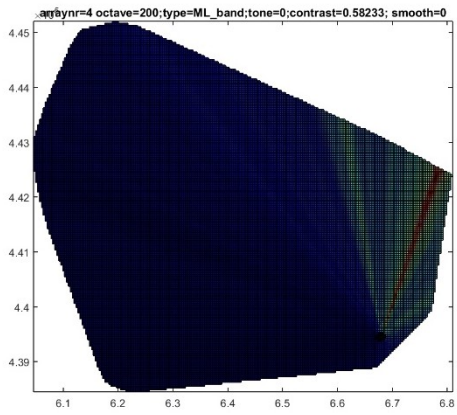
3.3 Beam forming from array 3 and array 4 during changing wind direction

Several beam forming techniques are used and combined in the overall measurement system. To illustrate the effect of meteorological conditions, the results of a maximum likelihood beam former using 36 microphones (ignoring the more closely spaced microphones in the middle of the array) are shown. This beam former is applied 4 times on a 6 second interval and results are combined using a multiplication. The latter is inspired by a desire to remove occasional sounds such as trucks, trains, and ships passing close to the arrays.

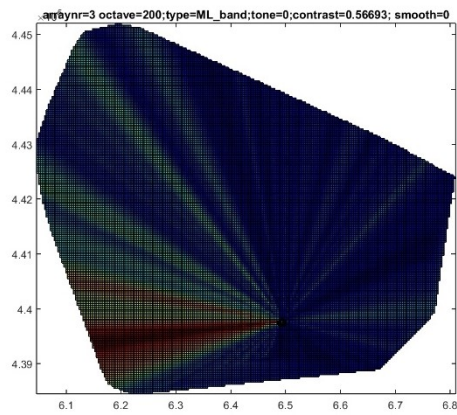
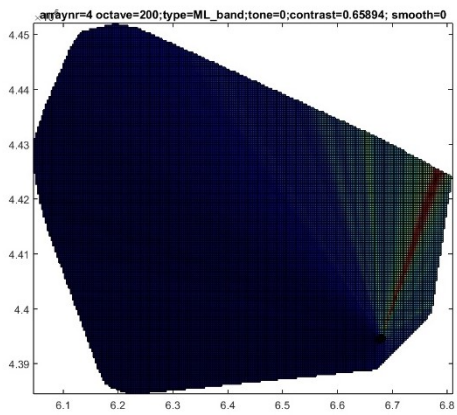
Arrays 3 and 4 are located closest to the reference source, hence the effect of meteorological conditions and other sources on beam forming is investigated for these arrays. The selected night is particularly challenging with a wind direction that causes upward refraction between the reference source and Array 4 and also Array 3 during a long portion of the night. Figure 4 shows the intensity of the beams in different directions at different time steps (10 minute intervals) corresponding to Figure 2. Array 4 (left) points at the reference source consistently. At $t=9$ and $t=11$ other sources are detected (the events could clearly be identified on a spectrogram), but in general even when the expected contribution of the reference source is more than 10 dB below the overall sound level and 5 dB below the sensor noise, detection is still very good. Array 3 (right) on the contrary only vaguely points in the direction of the reference source at $t=1$ but quite quickly points towards the west rather than the east. Until $t=6$ a faint beam can still be detected in the direction of the reference source. From Figure 2 it becomes obvious that this is mainly due to the presence of another important source that increases the overall noise level at the array by up to 10 dB at 200 Hz. Jumping to $t=26$, where the overall noise level has dropped, the faint beam in the direction of the reference source is again visible. This is remarkable as the calculated expected contribution of the reference source to the sound level is now 20 dB below the overall noise level and 15 dB below the sensor noise. This is close to the limit one can expect with a 36 microphone array. The calculated expected contribution during this period with upward refraction may have been underestimated, for example due to stronger turbulence scattering.



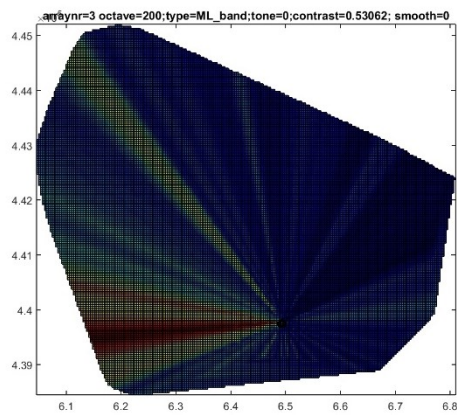
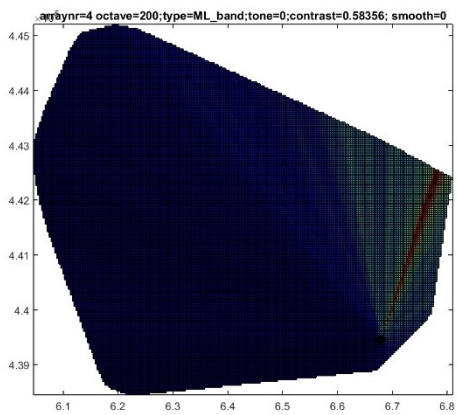
t=1



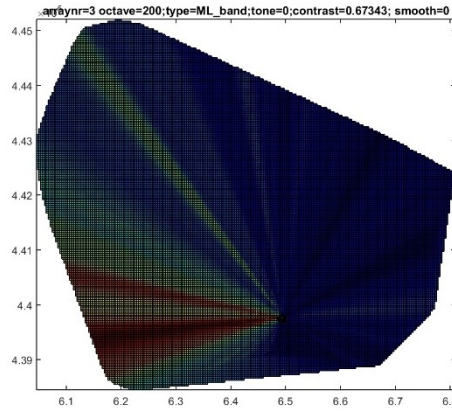
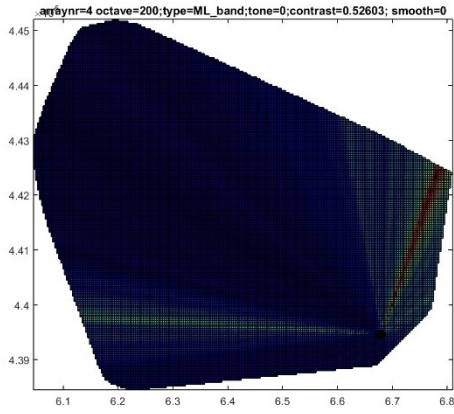
t=2



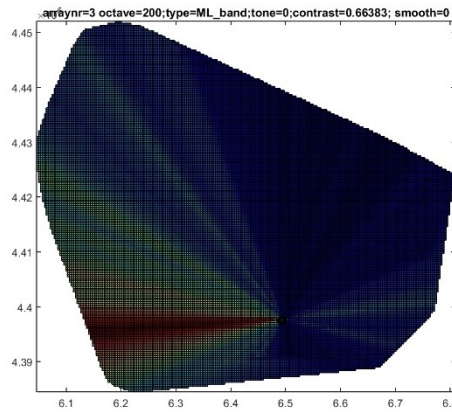
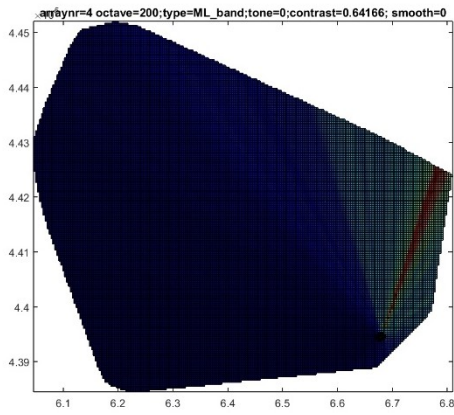
t=3



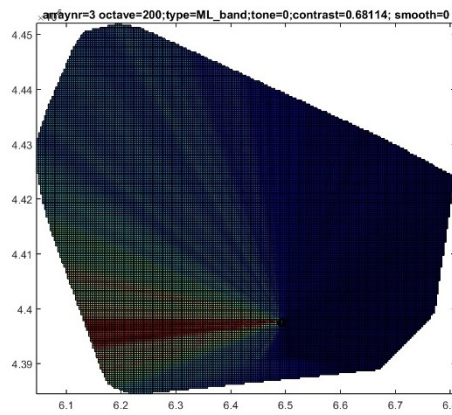
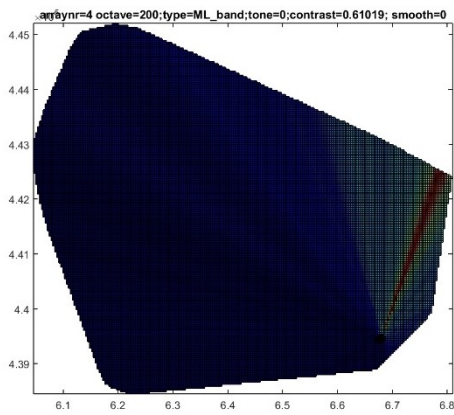
t=4



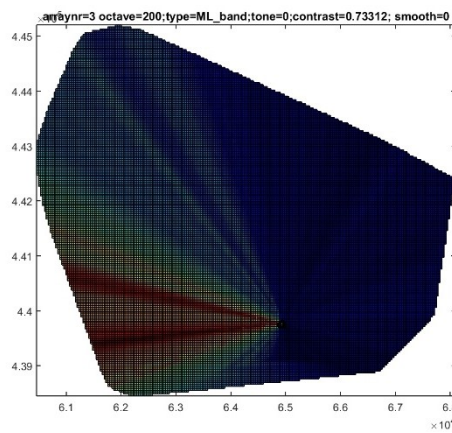
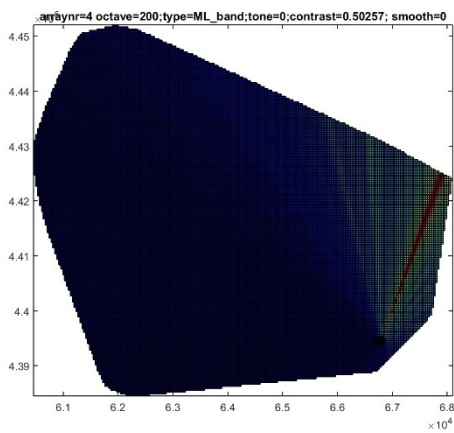
t=5



t=6



t=7



t=8

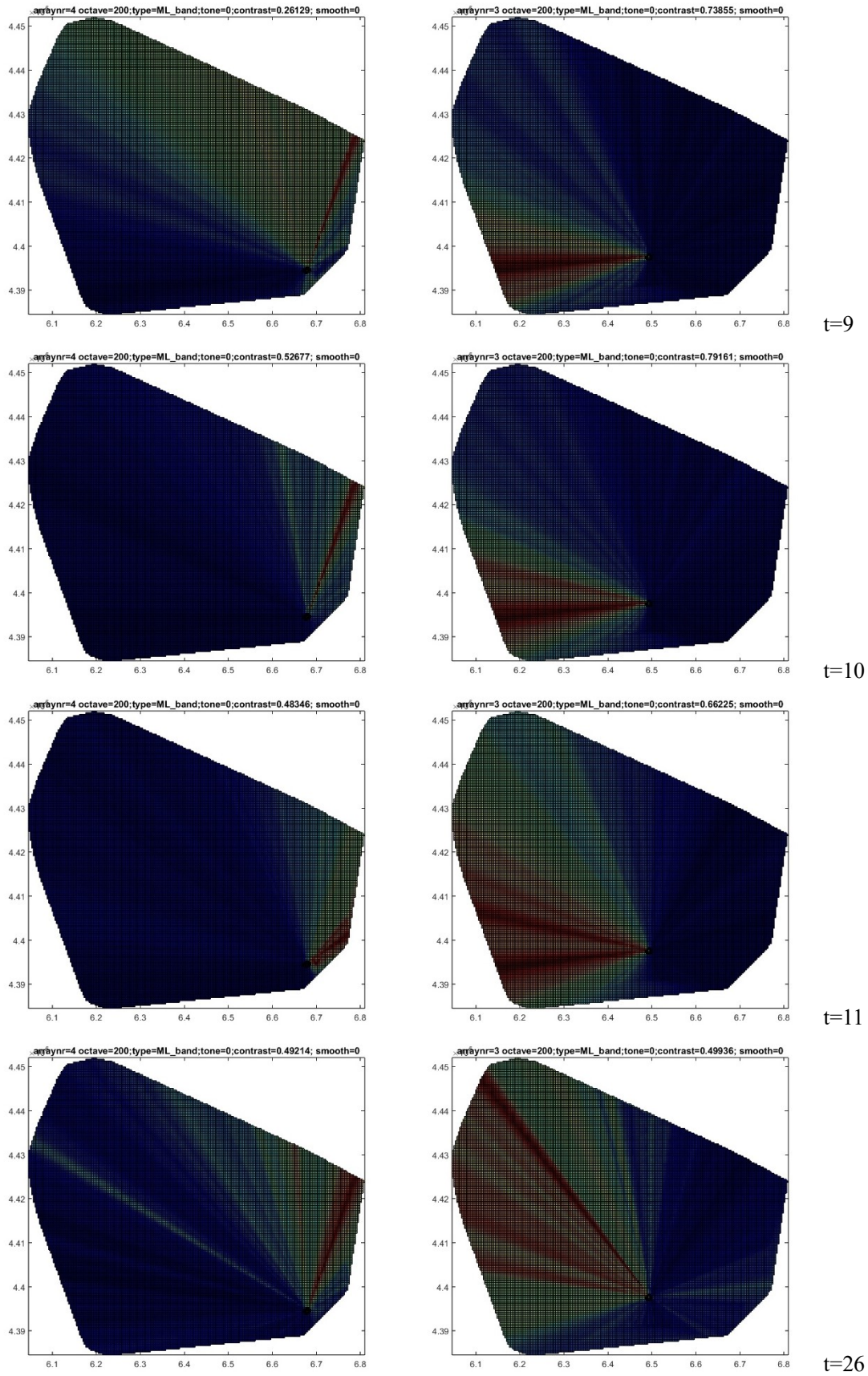


Figure 4 – Intensity plots of beams from array 4 (left) and array 3 (right) at different time steps during the night.

4. Keeping track of accuracy

To estimate sound powers emitted by various sources on the area of study, the information from the four arrays surrounding the industrial area is combined (9). Due to meteorological conditions and simply distance to the array, not all sound sources in the industrial area will be equally well sensed by all arrays. Therefore an estimate of the precision of the information obtained from each array is included in the data fusion process. For this, an ordered weighted average (OWA, 10) of the estimates of the source power obtained from each array is used.

$$L_W(r, f, t) = \sum_{i=1}^4 \left[\max \left(1, \sum_{j=1}^i \mu_{(j)} \right) - \max \left(1, \sum_{j=1}^{i-1} \mu_{(j)} \right) \right] L_{W(i)}(r, f, t) \quad (1)$$

where r , f , and t refer to the place, frequency band (1/3 octave), and time (10 minute step) respectively. $\mu_{(j)}$ is the accuracy of the estimate of the power by array (j). The brackets indicate that the estimates are ordered according to decreasing μ 's.

Several sources of imprecision can be identified. As shown above, the signal to noise ratio of the contributions of a source to the array microphone signal is an important source of imprecision. To estimate the effect of imprecision caused by low signal to noise ratio at the microphones, the signal and noise have to be estimated. The sensor's inherent noise level as a function of frequency has been characterized in quiet anechoic environment and was tabulated. The additional wind induced noise has been estimated on site by assuming that wind noise is dominant at the quietest periods of a month. The signal is the contribution of a possible source at a given location to the level observed by the array microphones. This results in an estimate of the accuracy based on SN given by:

$$\mu_{SN}(r, f, t) = \begin{cases} 0 & SN < SN_{min} \\ \frac{SN - SN_{min}}{SN - SN_{min} + \sigma_{SN}} & SN \geq SN_{min} \end{cases} \quad (2)$$

with the SN-threshold, SN_{min} , estimated at -15 dB and the transition region determined by σ_{SN} . Equation (2) is applied for each array separately. The signal to noise ratio SN_k is estimated for each array k as:

$$SN_k(r, f, t) = \sum_i P(L_{Wi}(f), r, t) (L_{Wi}(f) - A_k(r, f, t) - N_k(f, t)) \quad (3)$$

where P is the probability of finding the power L_{Wi} at location r at time t . As this probability is hard to estimate a priori, it is fixed to a single value of 120 dB for each frequency and each location. Nevertheless post hoc better estimates could be made. A_k is the attenuation between source location r and the k^{th} array, which is calculated using the PE approach.

The second source of imprecision is due to the calculation of the attenuation between each potential sound source location and the arrays and microphones around the industrial area. The attenuation appears in the formula that calculates the source power based on the observation at the array. The standard error on the propagation is calculated on the basis of uncertainty in input of the parabolic equation model that is used to predict the attenuation between source and microphone arrays given the instantaneous meteorological data. Imprecision of the choice of a set of 2D PE models for approximating 3D propagation is not accounted for.

To get the attenuation (for each source-receiver combination), a representative sound speed profile is determined based on the meteorological data, and this profile is fitted to one of 27 profiles for which the propagation has been calculated beforehand. Three different types of ground absorption were used: hard (for sand and water), soft (like grass), and very soft (like heathland or forest). The sound propagation for paths with a combination of ground types is determined by weighting each of the three different propagation results with the corresponding distance. In general, a source height of 15 meters has been used for the industrial area.

To estimate the uncertainty of the propagation due to a quickly varying meteorology, also the two neighboring sound speed profiles were used for the sound propagation. Additionally, a change in source height is used. Based on these four propagation results, the standard deviation, SD , is calculated, with a maximum of 10 dB. For the propagation over industrial terrain a 3 dB increase is used to

account for uncertainty due to scattering. The standard error on the propagation model is transformed to an accuracy in the interval [0,1] using:

$$\mu_A(r, f, t) = e^{-SN^2/\sigma_A^2} \quad (4)$$

where σ_A determines the transition to zero.

Finally, the effect of instantaneous changes in propagation time due to turbulence at various time scales causes deviations in the estimated angle both by instantaneous changes in the angle of arrival of the sound at the array and loss of coherence between the signals detected at the different microphones of the array. This source of imprecision is currently not taken into account.

The accuracy of the estimates used in Equation (1) is simply the product of both sources of both accuracies: $\mu = \mu_{SN} \cdot \mu_A$.

Once the sound power level at each location, $L_W(r, f, t)$, has been determined, the accuracy of this estimate itself, μ_{LW} , can be determined. It is obtained as the product of the linear average of the three largest $\mu_{(j)}$, a term accounting for the mismatch between power estimates originating from each array, μ_{MM} and a calculation of μ_{SN} (Equations (2) and (3)) now based on the real sound power L_W rather than on the probability P of finding this sound power. μ_{MM} is calculated using a formula similar to (4) with a weighted mismatch between the different estimates of sound power and σ_{MM} as arguments.

5. Sound power levels

Using the approach described above, the sound power estimate obtained from the four arrays is combined. Figure 5a shows the estimated sound power level in a bounding box surrounding the reference source. The horizontal lines represent the expected sound power level. It can be seen that during the first 10 time steps the calculated sound power level fluctuates around the expected level, in particular for the 125 and 250 Hz octave bands. The deviations are mainly due to the imprecision in the calculated attenuation. Due to the changing meteorological conditions, at time step 11 the source is no longer detected. This is mainly due to the drop in the contribution of the reference source at array 4 (Figure 2). At time step 15 array 4 again starts to detect the source, but at that time array 2 and 3 no longer receive a signal that is sufficiently high to detect the source.

To avoid the problem of poor detection of the source during periods of adverse meteorological conditions, exponential averaging is introduced on each potential source location in the area, taking into account the expected accuracy of the source estimation at each time step. Figure 5b shows that this results in a much higher probability of finding the stable reference source and hence more precise determination of its sound power level. The deviations from the expected source power now have a diversity of reasons that can hardly be identified precisely. Apart from the accuracy of the estimation of the sound attenuation, the aggregation over longer time intervals conserves the location of occasional sources over several time steps. Hence sound power is attributed to them also after they have stopped working. Thus the stable sources receive a smaller fraction of the total sound power emitted in the area. This immediately also highlights a serious drawback of combining source detection over various time steps: occasional sources remain virtually present over a period of time that is too long compared to the actual time they are present.

6. Discussion and conclusions

Meteorological conditions have a number of effects on the accuracy of detecting sound sources and determining their sound power in a large industrial area based on microphone arrays. Based on theoretical considerations it was expected that coherence loss due to turbulence would play an important role. However, it turned out that upward refraction and the resulting loss of signal amplitude at the array position was the main cause of imprecision in detecting sources. When sources are detected, estimating their sound power correctly requires extremely accurate propagation models that account for the changing refraction and account for turbulence scattering in a timely manner. Although a PE-model and accurate effective sound speed profiles are used in this project, the imprecision of estimated sound power for a reference source remains in the order of 10 dB during the example night with varying wind direction that is shown in this paper. Keeping track of imprecisions in the estimated propagation towards each array improves accuracy but it cannot fully compensate for uncertainty in the sound power level estimation due to meteorological conditions. It is suggested that more meaningful results could be obtained by maintaining sound source positions detected during one of the observations at 10 minute intervals where the meteorological conditions are favorable for a longer

interval of time. Thus detection during different time intervals could be aggregated to a more continuous source power map at least for stable sources. The impact on occasional sources nevertheless needs to be investigated further.

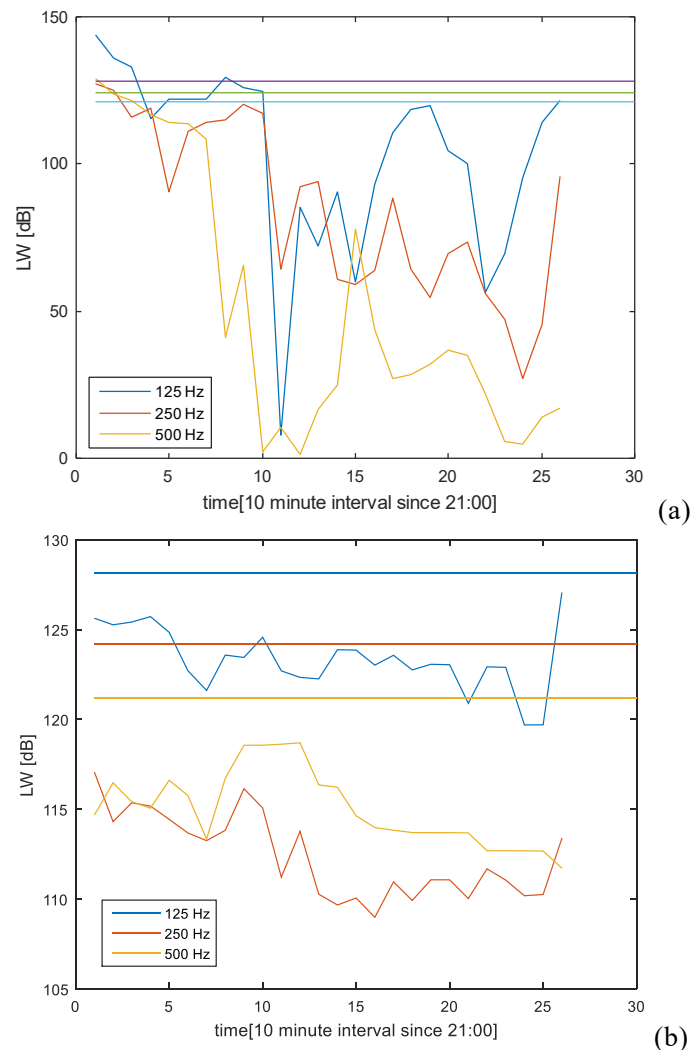


Figure 5 – Estimated sound power level inside a bounding box extending 200m beyond the reference source in three octave bands as a function of time: (a) as detected by the arrays at every time step; (b) integrated over time steps. The horizontal lines represent the expected sound power levels.

ACKNOWLEDGEMENTS

The development of the Maasvlakte sound measurement network ("Geluidmeetnet Maasvlakte") was commissioned by DCMR in the framework of the BRG (Bestaand Rotterdam Gebied), and received financial support from the City of Rotterdam, the Province of South Holland, the Port of Rotterdam and the Government of the Netherlands.

REFERENCES

1. Benesty, J., Chen, J., & Huang, Y. (2008). *Microphone array signal processing* (Vol. 1). Springer Science & Business Media.
2. Brown, D. J., Katz, C. N., Le Bras, R., Flanagan, M. P., Wang, J., & Gault, A. K. (2002). Infrasonic signal detection and source location at the Prototype International Data Centre. *Pure and applied geophysics*, 159(5), 1081-1125.
3. Collier SL, Wilson DK. Performance bounds for passive sensor arrays operating in a turbulent medium: Plane-wave analysis. *The Journal of the Acoustical Society of America*. 2003;113(5):2704-18.
4. Kozick, R. J., & Sadler, B. M. (2004). Source localization with distributed sensor arrays and partial

- spatial coherence. *Signal Processing, IEEE Transactions on*, 52(3), 601-616.
5. Vorobyov, S. A., Gershman, A. B., & Wong, K. M. (2005). Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays. *Signal Processing, IEEE Transactions on*, 53(1), 34-43.
 6. Di Claudio, E. D., & Parisi, R. (2001). WAVES: Weighted average of signal subspaces for robust wideband direction finding. *Signal Processing, IEEE Transactions on*, 49(10), 2179-2191.
 7. De Coensel B, Botteldooren D, Van Renterghem T, Dekoninck L, Spruytte V, Makovec A, Wessels P, van der Eerden F, Basten T. Robust microphone array beamforming for long-term monitoring of industrial areas. In 22nd International Congress on Acoustics (ICA 2016) 2016.
 8. Van Der Eerden F, Segers A, Basten T, De Coensel B, Botteldooren D, Van Renterghem T, Dekoninck L, Spruytte V, Makovec A. Time varying sound propagation for a large industrial area. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings 2016 Aug 21 (Vol. 253, No. 4, pp. 4590-4597). Institute of Noise Control Engineering.
 9. Botteldooren D, Van Renterghem T, De Coensel B, Dekoninck L, Spruytte V, Makovec A, van der Eerden F, Wessels P, Basten T. Fusion of multiple microphone array data for localizing sound sources in an industrial area. In 45th International Congress and Exposition on Noise Control Engineering (Inter-Noise 2016) 2016 (pp. 3127-3134).
 10. Yager, R. R. (1988). On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on systems, Man, and Cybernetics*, 18(1), 183-190.