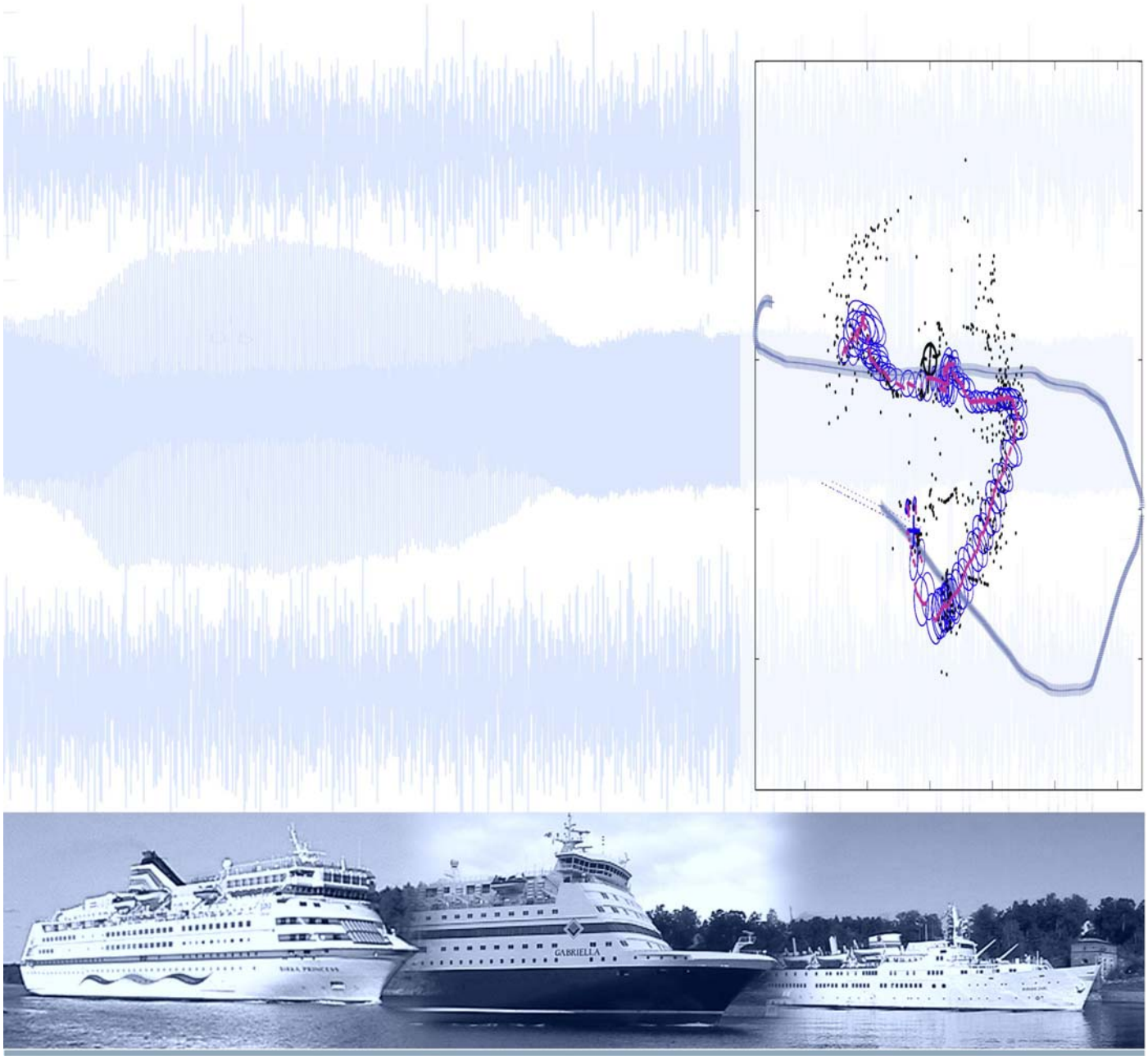
FOI

Passive acoustic and electromagnetic underwater tracking and classification using data fusion

Eva Dalberg, Andris Lauberts,
Ron Lennartsson, Mika Levonen,
Leif Persson



FOI-R--1727 --SE
ISSN 1650-1942

Systems Technology

Command and Control Systems

September 2005

FOI is an assignment-based authority under the Ministry of Defence. The core activities are research, method and technology development, as well as studies for the use of defence and security. The organization employs around 1350 people of whom around 950 are researchers. This makes FOI the largest research institute in Sweden. FOI provides its customers with leading expertise in a large number of fields such as security-policy studies and analyses in defence and security, assessment of different types of threats, systems for control and management of crises, protection against and management of hazardous substances, IT-security and the potential of new sensors.

FOI
Defence Research Agency
Systems Technology
SE-164 90 Stockholm

Phone: +46 8 555 030 00
Fax: +46 8 555 031 00

www.foi.se

Passive acoustic and electromagnetic underwater tracking and classification using data fusion

FOI-R--1727--SE
ISSN 1650-1942

Systems Technology
Technical report

Command and
Control Systems

September 2005

Issuing organization	Report number, ISRN	Report type
FOI – Swedish Defence Research Agency Systems Technology SE-164 90 Stockholm	FOI-R--1727--SE	Technical report
	Research area code 4. C4ISTAR	
	Month year September 2005	**Project no.** E60702
	Sub area code 43 Underwater Surveillance, Target acquistion and Reconnaissance	
	Sub area code 2	

Author/s (editor/s)	Project manager
Eva Dalberg Andris Lauberts Ron Lennartsson Mika Levonen Leif Persson	Eva Dalberg
	Approved by Monica Dahlén
	Sponsoring agency Swedish Armed Forces
	Scientifically and technically responsible Mattias Karlsson

Report title

Passive acoustic and electromagnetic underwater tracking and classification using data fusion

Abstract (not more than 200 words)

An interesting possibility for improved surveillance capabilities in challenging underwater environments and against targets with low acoustic signatures is the use of multisensor systems combined with data- or information fusion. This work describes how data fusion can be used for tracking and classifying targets using passive underwater acoustic and electric field sensors.

Tracking of a synthetic underwater target has been performed on data from a passive acoustic uniform line array (ULA) and a long base-line electrode array for underwater electric field measurements. A Kalman filter was applied on bearing estimates from the acoustic data and estimates of the target position from the electrode array. The results show that the electrode sensors are useful, within the limits of their sensitivity, for adding range estimates to the acoustic bearing information.

Underwater acoustic and electric field signatures have been used to classify surface ships according to their size and propulsion. Two feature extraction methods have been used, an AR (Auto Regressive) model and a non-linear DDE (Delay Differential Equation) model. A Bayesian minimum-error classifier has been implemented on the extracted features. It is shown that data fusion can improve the classification performance in some cases. However, improvement requires that each underlying classifer work properly. In other cases, the performance might be degraded.

Keywords

underwater acoustics, electromagnetics, localisation,bearing estimation, tracking, classification, data fusion

Further bibliographic information	Language English
ISSN 1650-1942	**Pages** 30 p.
	Price acc. to pricelist

Utgivare	Rapportnummer, ISRN	Klassificering
FOI - Totalförsvarets forskningsinstitut Systemteknik 164 90 Stockholm	FOI-R--1727--SE	Teknisk rapport
	Forskningsområde 4. Ledning, informationsteknik och sensorer	
	Månad, år September 2005	**Projektnummer** E60702
	Delområde 43 Undervattenssensorer	
	Delområde 2	

Författare/redaktör	Projektledare
Eva Dalberg Andris Lauberts Ron Lennartsson Mika Levonen Leif Persson	Eva Dalberg
	Godkänd av Monica Dahlén
	Uppdragsgivare/kundbeteckning Försvarsmakten
	Tekniskt och/eller vetenskapligt ansvarig Mattias Karlsson

Rapportens titel (i översättning)

Passiv akustisk och elektromagnetisk undervattensklassificering och -målföljning med hjälp av data fusion

Sammanfattning (högst 200 ord)

En intressant möjlighet för att bedriva undervattensspaning i särskilt svåra miljöer och mot tysta mål är multisensor-system kombinerat med data- och informationsfusion. Denna rapport beskriver hur datafusion kan tillämpas för målföljning och klassificering vid spaning med passiva akustiska och elektriska undervattenssensorer.

Målföljning har genomförts på data insamlade med en hydroakustisk linjeantenn samt ett elektrodsystem av långbas typ för mätningar av elektriska fält under vattnet. Ett Kalmanfilter användes på bäringsberäkningar från akustiska data tillsammans med positionsbestämningar beräknade med hjälp av data från elektrodantennen. Resultatet visar att elektrodsensorerna, inom gränsen för sin känslighet, är användbara för att addera avståndsinformation till den akustiska bäringsinformationen.

Akustiska och elektriska undervattenssignaturer har använts för att klassificera ytfartyg beroende på deras storlek och framdrivning. Två metoder har tillämpats för att ta fram särdrag, en AR (Auto Regressive) modell och en icke-linjär DDE (Delay Differential Equation) modell. Bayes beslutskriterium, som resulterar i minsta antal felbeslut, har använts på de extraherade särdragen. Resultatet visar att datafusion kan förbättra klassificeringsförmågan i några fall, men att utfallet är beroende av att varje underliggande klassificerare fungerar väl. I annat fall kan klassificeringsförmågan försämras.

Nyckelord

hydroakustik, elektromagnetik, lokalisering, bäringsbestämning, målföljning, klassificering, datafusion

Övriga bibliografiska uppgifter	Språk Engelska

ISSN 1650-1942	Antal sidor: 30 s.
Distribution enligt missiv	Pris: Enligt prislista

Contents

1. INTRODUCTION

Traditionally passive underwater surveillance is dominated by acoustic methods. However, in littoral environments the hydrographic and bathymetric conditions can be such that acoustic systems have a limited range of operation. Noise sources such as heavy shipping may further degrade the system performance. If surveillance has to be done under such conditions against targets with very low acoustic signatures the use of other sensor types is a possibility that has to be considered. One way of improving the performance of the acoustic systems is to fuse that data with data from other sensors. One possibility is electrode systems, used for low frequency underwater electric fields, systems that FOI have a long experience with.

The longterm goal of this work is to be able to evaluate in which contexts multisensor techniques could be useful for improved underwater surveillance performance. This could mean improved detection ranges for a specified target, more robust tracking of targets with low signatures in areas with very heavy shipping, or more reliable, automated classification of possible threats. The work presented in this report should be seen as a first step toward that goal. Two methods for underwater surveillance using data fusion has been studied. The first is aimed at tracking of an underwater target. The other studies classification of surface vessels. Both use data fusion on underwater acoustic and electric field data.

1.1. Data fusion

In this context data fusion is a process that combines data and knowledge from different sources with the aim to maximise the value of the data, which can be uncertain, incomplete or contradictory. Frequently the sensors work in complementary ways with raw data in different formats - in our case hydrophones and electromagnetic sensors.

In these cases data fusion is preferably carried out at a higher level of information content. Normally, extracted features, observed by each sensor, are fused into a common unit. Data fusion involves several cooperating and partly overlapping parts. Detection and classification may be done by independent sensors at different times. Association makes another important point for a successful fusion of sensor data. The common information is fused to form a clearer picture of the target position and identity.

Although there is a conceptual line between classification and data fusion their goal is usually the same, to detect and identify a sought target. In addition, data fusion may act on everything from data (features) in a sequence of 'snap-shots' from a sensor to a common set of features from different sensors or separate decisions. Some methods are very general and can be applied to many different cases, whereas other methods involve highly specialised algorithms for a certain application.

1.2. Contents of this report

This work has been primarily aimed at making the whole analysis chain work, from raw data acquired by separate sensor systems to a fused track or classification. In this report we describe the methods we have used and point out the lessons we have learned. Chapter 2 contains the work on tracking. The electromagnetic and acoustic tracking methods and the resulting estimates of the tracks are presented in chapters 2.1 and 2.2, respectively. This is followed by a description of the Kalman filter tracker, and the results of applying it to the data. Chapter 3 contains the work on classification. Finally, we summarise the work and look forward at future work in this field.

2. TRACKING

Tracking has been performed using data from two passive arrays, one acoustic and the other with electrodes. Both arrays were placed on the sea floor, side by side, in a north-south direction. Fig. 1 shows an overview of the

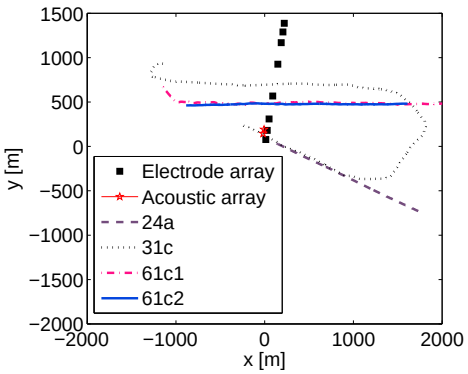


Figure 1: Overview of the sea trial site where the data used in the tracking analysis was recorded. A passive acoustic 31-element array (in red starts) with total aperture 45 m and a passive electrode array with 8 sensors (black squares) was placed on the seafloor. Four tracks have been analysed, tracks 24a, 31c, 61c1 and 61c2.

trial site, with the arrays and the four tracks used in the analysis shown.

2.1. Target localisation using underwater electric fields

Low frequency underwater electric fields from surface vessels or underwater vehicles are caused by electric currents flowing between the hull and the propulsion system of the vessels. The currents have their origin in the use of materials with different electrochemical potential that are in contact with the sea water which acts like an electrolyte. Typical frequencies of these signatures are in the Extreme Low Frequency (ELF) range, usually defined as 0.3-3000 Hz. In this frequency range the propagation conditions in shallow waters are such that the fields can be detected a few kilometres away from the source.

2.1.1. Low frequency underwater electric fields

Low frequency underwater electric fields can be measured using pairs of electrodes. The electric potential is measured between the pairs and the electric field is derived as $E = \Phi/d$, where Φ is the electric potential in V and d is the distance between the electrodes in m. The distance between the electrodes varies between a few meters in systems primarily used for signature measurements, up to a few hundred meters in surveillance systems optimised for long detection ranges.

A target vessel can be modelled as an electric dipole source, with the dipole moment $I \cdot dl$ as its source strength. The propagation can be modelled using a horizontally stratified model of the underwater environment, where the most important parameters are the conductivity and thickness of each layer. Such a model can be more or less refined, but it is often sufficient to use a simple model where the sea water constitutes one layer, one or two sediment layers and finally the bedrock. Using such an environmental model, together with an electric dipole model of the vessel, the electric field caused by a vessel can be calculated at any point in the area surrounding the target. If the environment in an area is known, measurements of the underwater electric field may be used to localise the source of the fields by minimising the difference between the calculated and the measured fields. Such an inversion scheme, originally used to determine the underwater electromagnetic environment using controlled dipole sources at known positions, has been used here to track a source moving through an area where electric field sensors had been deployed.

2.1.2. The Experiment

In this study, data from a linear array with 8 electric field sensors have been used. Each sensor consisted of a pair of carbon fiber electrodes, with 50-200 m distance between the electrode pairs. The array was placed on the sea floor in a strait where the water depth varied between 10-45 m. The area is a typical Swedish archipelago environment with many small islands and a lot of variations in the bathymetric conditions. Details on the experiment can be found in [1].

A synthetic dipole source was towed at a depth of 2 m along different tracks in the

area. The position of the synthetic source was continuously logged with a GPS-system. Continuous sinusoids of a few Hz were transmitted at different dipole moment strengths. In this study, four tracks were used. During track 24a a low dipole strength of around 4 Am was used at 2 Hz. The signal-to-noise ratio (SNR) was high enough in four of the sensors, whereas the other four sensors were not used in the analysis of that track. On the other hand, the depth profile of that track was more uniform than the other tracks used. During tracks 31c, 61c1, and 61c2, the source strength was 28 Am, which enabled the use of all eight sensors. Such a high source strength can be compared with the estimated dipole strengths of medium sized commercial ships or ferries [2] and [3]. During the tracks 61c1 and 61c2 the target vessel was manoeuvred over relatively even depths, whereas the bathymetry along track 31c was more varied.

2.1.3. Localisation method

The amplitude of the electric field at a chosen frequency is calculated through an FFT. In this study we have used 5 s long data segments. These amplitudes are then compared with calculations of the electric field from the model Nlayer 2.0 [4], using the inv_nlayer program [5]. The residual $\mathbf{r}$, between the measured and estimated fields is calculated and minimised. The residual is defined as,

$$\mathbf{r}(\mathbf{x}) = \mathbf{m}(\mathbf{x}, \mathbf{y}) - \mathbf{d} \qquad (1)$$

where $\mathbf{m}$, and $\mathbf{d}$, are the estimated and measured fields respectively.

In the model the parameters $\mathbf{x}$ are unknown, and $\mathbf{y}$ are known. The unknown parameters are the horizontal position, and the heading of the source. They are estimated from the minimisation of the residual. The source was towed at a depth of 2 m, and that information was used as a fixed parameter. The dipole source strength can be estimated using the inversion program, but in this study it was fixed at the strength used during the transmissions. The sensors were assumed to be located on the seabed, at the average depth of the sensors used in the inversion (i.e. all eight sensors for all tracks, except track 24a, where only four sensors were used). The same depth was also used as the waterdepth in the horizontally stratified, 3-layer environment (water+sediment+bedrock). The other parameters in that model environment had typical values estimated from previous experiments in the area [6]-[8].

For each step in the track (i.e. for each 5-second block of data), an estimation of the source position and heading was performed, using 50 or 100 different initial values (50 for track 31c, 100 for tracks 24a, 61c1 and 61c2). The initial values for the heading were chosen randomly from a uniform distribution of all angles. The initial x- and y-coordinates were chosen in a sector with an opening angle of ± 5 degrees around the true bearing of the target, from the acoustic array, reaching out to a maximum radius of 2 km. Fig. 2 explains how this was done. This restriction was used assuming that bearing information would be available from the acoustic array, and that the electric field data would add the range information to the target with the given bearing. This scheme was chosen since bearing information is easily obtained with the use of a horizontal acoustic array, whereas range information is more difficult and less accurately obtained, unless more than one array is available. At the same time, when inverting electric field data for target position, it was found that the range information is slightly more accurate than the bearing information. By restricting the initial positions to a sector around the true bearing, the intention was to increase the efficiency of the method and obtain more reliable range estimates. Although the initial positions were confined to a sector of a circle surrounding the array, the inversion results were allowed to lie within a rectangular area encompassing that sector.

The minimisation was performed using a cost function which is based on the residual between the forward model and the measured fields, the square of Eq. 1. As explained above, 50 or 100 different initial values were used, to avoid reaching a local rather than the global minimum. The smaller the sum of

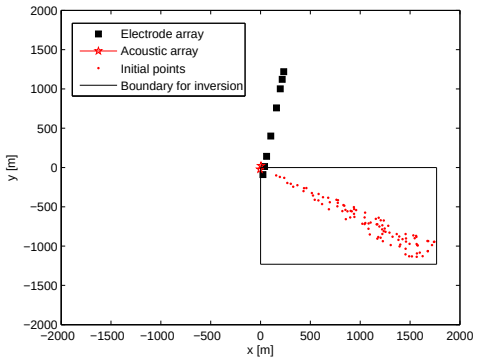


Figure 2: Schematic overview of how the initial points and boundaries for the localisation inversion is performed.

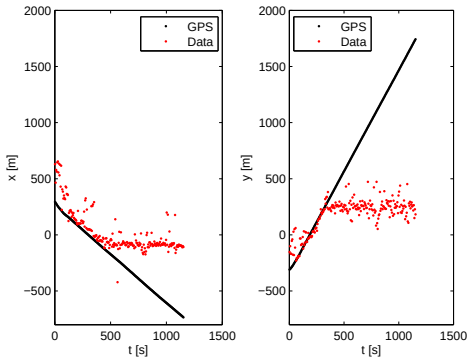


Figure 3: x- (north) and y- (east) coordinates as a function of time along track 24a. The GPS-track is shown in black and the estimated target positions are shown in red.

squared differences is, the better is the match between model and data. In this study, the best solutions have been chosen for use in the higher level data fusion process. Choosing tracks with the smallest residual between model and measured data means that the electric field amplitudes are matched in the best way. It does not mean that the tracks with least deviation between the estimated and true track are chosen. Simulations have shown that in the ideal case, where the environment is known and the noise is insignificant, there is a good correlation between small residual sum of squares and accurate estimates of the target position along a track. However, when using real data the situation is more complex. Even if in general smaller residuals means a better accuracy, choosing the solution with smallest sum of squares difference does not necessarily mean that the most accurate track has been found, since there are noise, spatial variations in the environment and other aspects in data that the model does not take into account. However, in a surveillance situation, this would be a realistic scheme that could be followed.

2.1.4. Results

Track 24a

Track 24a is the track with the lowest electric dipole source strength, which made it necessary to use only 4 of the 8 available sensors, since the others had no signal above the background noise level. It also means that the signal-to-noise ratio (SNR) rapidly gets too low as the range between the target and the array increases. In the localisation algorithm measured amplitudes of the electric field are compared with amplitudes computed by the forward model. As a consequence, the larger the measured amplitude is, the larger is the distance to the target. Measured amplitudes contain both signal and noise, but when the target is far away data will be completely dominated by noise. At this point the measured amplitude will be interpreted as a signal from a target that is at the range where the signal from the target is equal to the noise amplitude. This can be seen in Fig. 3 where the x (north)- and y (east)-coordinates are shown as a function of time, together with the true coordinates recorded by the GPS system. The coordinates are relative to one of the end sensors in the electrode array, see Fig. 1 for an overview of the coordinate system and the tracks. When the source is close to the sensor array, the reconstructed track is close to the true track. Whereas at longer distances, the reconstructed coordinates deviate from the true track at a fixed range from the array.

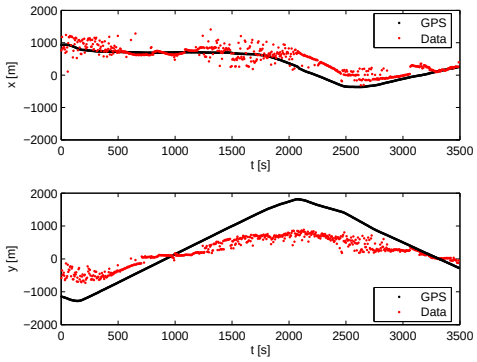


Figure 4: x- (north) and y- (east) coordinates as a function of time along track 31c. The GPS-track is shown in black and the estimated target positions are shown in red.

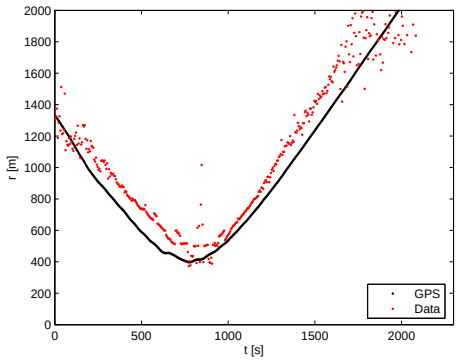


Figure 5: The distance between the sensor array and the target as a function of time along track 61c1. The GPS-track is shown in black and the estimated track in red.

Track 31c

The localisation algorithm was least successful for this track. Fig. 4 shows the reconstructed x- and y-coordinates of the track. The algorithm only manages to reconstruct the general trends of the tracks, but the offsets are large, except for ranges very close to the array when the track follows a sea line in the area (the end of the track). Our interpretation is that this is due to a mismatch between the underwater environment in the area and the environmental model. The model with infinite horizontal layers with equal thickness and properties in the whole area fails to represent the true environment. The first part of the track is also more dependent on the sensors in the northern part of the array, which are at more shallow depths than the sensors in the southern part of the array. The variations in depths between one sensor and the next are also larger, as they are placed on a slope from deeper to more shallow water.

Tracks 61c1 and 61c2

Fig. 5 shows the reconstructed and true range of the target during track 61c1 as a function of time. The track starts on one side of the array, the western side, it pass the array which is directed around 10 degrees from north, and then goes more or less straight to the east. It can be seen that the range is overestimated on both the east and the west sides of the array. By looking at Fig. 6, the error in range can be attributed mainly to the y-coordinate. Errors are smaller for the x-coordinate. The same behaviour is seen in track 61c2 (see figures 7 and 8), which is basically the same track, this time east to west. This means that the east-west component of the source signal is smaller than expected, since the algorithm interprets it as coming from a greater distance than the true position. This indicates that the environment is more absorbing from the source to the sensor array than predicted by the model or that there are other variations in the environment that are important.

2.1.5. Factors influencing the target localisation performance

The performance of the localisation algorithm can only be expected to be successful in cases where the signal to noise ratio (SNR) in the sensors is high enough. As can be seen in Figures 9 to 12, there is a fairly strong correlation between the SNR and the estimation errors.

So, how high does the SNR have to be in order to achieve a certain accuracy in the estimated target position? Since sensors are separated, the SNR varies from one sensor to another. We have looked at the correla-

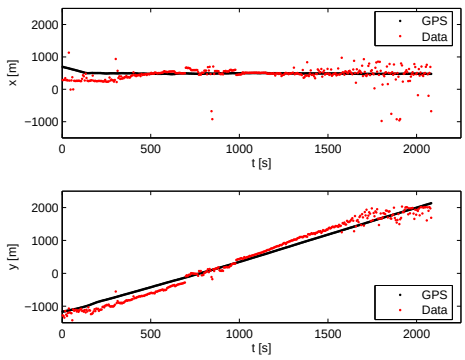


Figure 6: x- (north) and y- (east) coordinates as a function of time along track 61c1. The GPS-track is shown in black and the estimated target positions are shown in red.

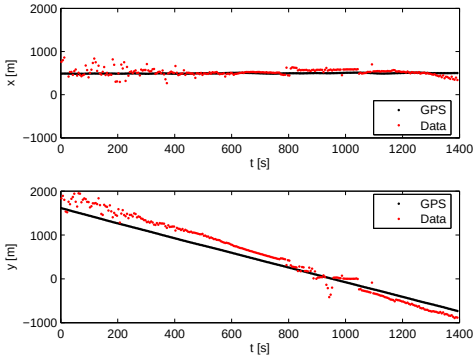


Figure 8: x- (north) and y- (east) coordinates as a function of time along track 61c2. The GPS-track is shown in black and the estimated target positions are shown in red.

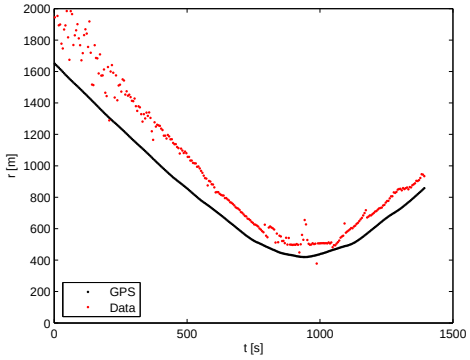


Figure 7: The distance between the sensor array and the target as a function of time along track 61c2. The GPS-track is shown in black and the estimated track in red.

tion between the error in the reconstructed distance to the target (the radius) and

- the SNR in the sensor with highest signal level,
- the mean SNR calculated for all sensors,
- the median SNR, and
- the sum of the SNR level in the sensors used.

The picture is very similar for all cases, except for the median value, which does not have a good correlation with the estimation accuracy. Figure 13 shows the correlation between the mean SNR for the eight sensors and the error in the radius estimation for tracks 31c, 61c1 and 61c2. The behaviour is similar for track 24a. As it can be seen, making requirements on the SNR value can be used to reduce or eliminate points where the reconstruction fails. However, some of the points where the accuracy is good will also be removed by such a scheme. It is possible to set a threshold on the SNR for a required performance. This SNR threshold can be used to calculate the maximum range where a target can be tracked accurately, if the source strength of the target is known.

For a target with a certain source strength, the signal level at a sensor depends on the environment in the area. The bathymetry in the area of the northern part of the array is more rapidly changing than at the southern part of the array. This is probably one of the reasons why the localisation performance is worse during the first part of track 31c, compared with the last part and tracks 24a and 61c1/c2, since sensors 5-8 are more important than sensors 1-4 at that point. Figure 14 shows the bathymetry along the tracks. Apart from the bathymetry, the bottom properties are important. They affect the wave propagation, and as discussed in connection with tracks 61c1 and 61c2, the bottom parame-

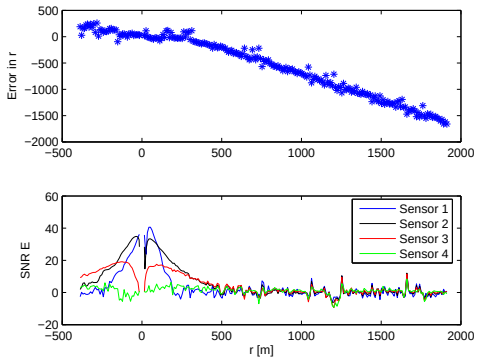


Figure 9: The upper panel shows the error in the estimated distance from the array to the target as a function of the true distance for track 24a. The lower panel shows the SNR in the 4 sensors used in the estimate of the track.

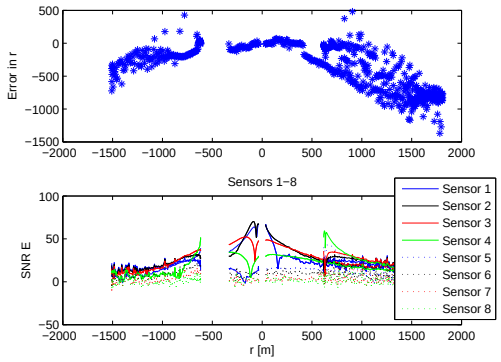


Figure 10: The upper panel shows the error in the estimated distance from the array to the target as a function of the true distance for track 31c. The lower panel shows the SNR in the 8 sensors used in the estimate of the track.

ters used might underestimate the absorption along those tracks.

2.1.6. Suggestions for future improvements

There are several things that can be done in order to improve the overall localisation accuracy, and the range where reliable estimates of the target position can be obtained.

First, the signals can be pre-conditioned by performing reference filtering before applying a localisation algorithm on them. Since most of the low frequency electric field background consists of distant sources, it is a common technique to use signals from a reference system placed some distance away from the array to filter the signals from common noise. This has not been done in this analysis.

Also, the number of sensors used in each step along a track can be varied, depending on the SNR in each sensor at each time. Such a scheme implies that all sensors that have useful information are included, while sensors that only record noise are not allowed to interfere in a negative way.

This analysis is based on data from a linear array. Such an array is effective as a tripwire, i.e. for secure detection of passing platforms. However, in terms of tracking it would be better to have access to data from more than one horizontal axis. Such data is available, and will be used in future studies.

In the current localisation algorithm, data from only one time-window is used simultaneously. If data from several consecutive time windows are used in parallel it is plausible to assume that tracking would become more robust.

Environmental models taking into account 3D-effects of the bathymetry are available. They are more elaborate than the less sophisticated propagation model used here, and are currently not seen as tools for tracking algorithms at the moment. They could however be used to verify the influence of the variability of the underwater environmental parameters on the tracking performance, and might be useful for tracking in the future.

2.1.7. Summary

Localisation of a towed low frequent electromagnetic source has been performed using data from a passive long base line electrode array. The localisation of the target has been performed by minimisation of the difference between the measured electric field ampli-

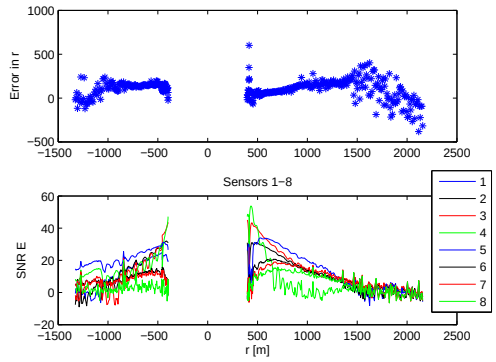


Figure 11: The upper panel shows the error in the estimated distance from the array to the target as a function of the true distance for track 61c1. The lower panel shows the SNR in the 8 sensors used in the estimate of the track.

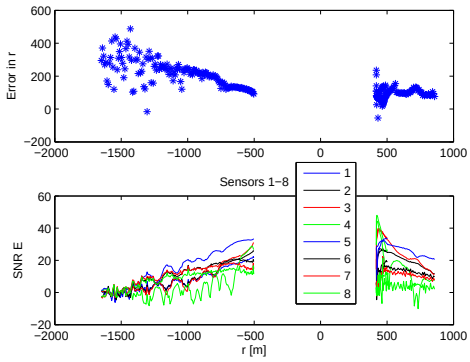


Figure 12: The upper panel shows the error in the estimated distance from the array to the target as a function of the true distance for track 61c2. The lower panel shows the SNR in the 8 sensors used in the estimate of the track.

tudes and calculated amplitudes from a forward model.

A satisfactory tracking performance is achieved at relatively short ranges, where the signal-to-noise ratio is large enough. The numerical model used needs good knowledge of the underwater environment and performs best in areas where the variability of the bathymetry is smallest.

2.2. Acoustic bearing estimation

Localisation is important in underwater array processing, i.e. bearing and range estimation. During the last decades several researchers have developed robust and sensitive high resolution methods for directions of arrival estimation of impinging wave fronts, e.g. [9],[10] and [11]. Here, the main purpose is to use robust acoustic bearing estimations as an input for range estimation by means of inversion of underwater EM-data. The fundamentals of array processing are discussed briefly, and the properties of conventional beamforming and two of the most used high-resolution methods, MUSIC (Multiple Signal Classification) and MV (Minimum Variance), are mentioned.

For more than 50 years, passive and active, underwater sensor systems have employed array processing for localisation of sources. In sonar array processing it is possible to estimate both the temporal (e.g. time-delays) and spatial (e.g. transmission loss) structure of the received wave field. However, the estimation of direction-of-arrival (DOA) of wavefronts, i.e. bearing estimation, is the most fundamental task in sonar array signal processing. To obtain bearing estimates several sensors are used to form sensor arrays in different configurations. This is done to gain the advantage of directivity, resolution and beamwidth compared to single sensor systems. The most basic approach for estimating bearings is based on applying sets of delays and weights on the sensor outputs for scanning each bearing of interest. This method is called conventional beamforming. The resolution for conventional beamforming (CB) is bounded by the wavelength of the received signals, sensor configuration and the aperture of the array. In many underwater situations it is important to resolve sources within the conventional beamwidth. The targets (e.g. ship or a submarine) may have typical characteristics due to the distribution of different sources as machinery and shafts. Sometimes a higher resolution is needed to resolve these sources. During the past decades several high-resolution eigen-

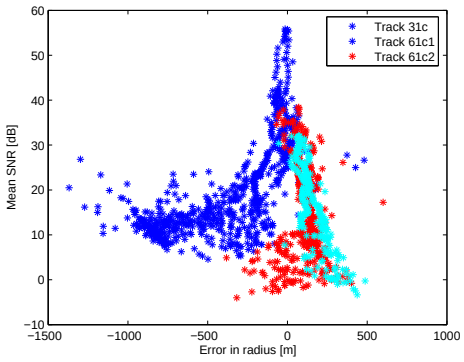


Figure 13: Mean SNR in the sensors used in the estimate of the target position vs the error in the estimated distance to the target.

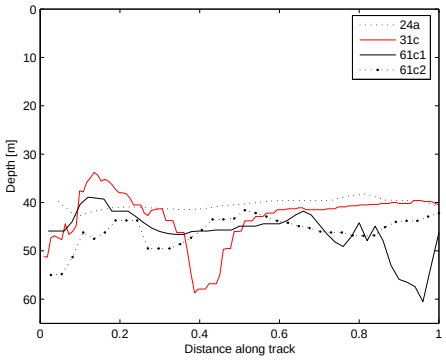


Figure 14: The bathymetry along the tracks, shown as a function of the normalised distance along each track.

structure based DOA estimators have been developed and is found in the literature (e.g. [12]).

The signal model is based on the assumption that a source signal $x(t)$ is present, generating a wave field at a bearing θ relative to a linear array of M sensors. A general model for the received wave field consists of the signals attenuated with α_n and α_m, and then delayed with $\tau_n(\theta)$ and $\tau_m\theta$ at sensor n and m as follows.

$$x_n(t) = \alpha_n x(t - \tau_n(\theta)) + \phi_n(t) \quad (2)$$

$$x_m(t) = \alpha_m x(t - \tau_m(\theta)) + \phi_m(t) \quad (3)$$

The noise terms $\phi_n(t)$ and $\phi_m(t)$ represents the additive ambient noise at each sensor.

The bearing estimation problem is related to finding each delay and performing time-delay estimation. The time-delays are compared to a model, where for each bearing the expected time-delays are calculated. If the shape of the wavefronts is known (mostly assumed planar), it is possible to estimate the time delays at each sensor.

2.2.1. Beamforming

The first method developed and still the most commonly used is conventional beamforming (CB). The advantages with this method are that it is easy to implement, the number of sources is not required as an input, and is fairly robust against coherent signals (for further details about beamforming see i. e. [9] and [13]). Some of the disadvantages are high sidelobes causing lower signal to noise ratio in some directions and the bearing resolution [9] as:

$$\Theta_{3dB} = k\lambda/L \quad (4)$$

where λ equals the maximum wavelength of the signal, L is the aperture of the array and k is a constant, equal to 50.5 degrees for a linear array. For the linear array used in this study, with an aperture of $L = 45$ m and a wavelength of $\lambda \cong 3$ m, the beamwidth is $\Theta_{3db} = 3.4$ degrees. This is the maximum 3 dB level for a conventional bearing resolution. For example, a single ship with a length of 50 m (or a combination of sources and ships within a width of 50 m) at a distance of 1000 m perpendicular to the array cover an angle of only 2.9 degrees, i.e. slightly smaller than the beamwidth resolutions above. In both active and passive sonar it is desirable to obtain higher resolution without increasing the array aperture. Therefore, it is a motif for using other methods with a higher resolution than conventional beamforming.

The received data at the sensor array output is defined as 'snapshots' at time t for all sensors in the vector form $\mathbf{x}(t) = [x_1(t), x_2(t), ...x_M(t)]$. The input for the bearing estimation methods is the sensor output covariance matrix formed as:

$$\mathbf{R}_{xx} = \frac{1}{N}\sum_{t=1}^{N}[\,\mathbf{x}(t)-\hat{\boldsymbol{\mu}}_x\,]\,[\,\mathbf{x}(t)-\hat{\boldsymbol{\mu}}_x\,]^H. \quad (5)$$

where N is the number of snap shots, H is the conjugate transpose and $\hat{\boldsymbol{\mu}}_x$ is the estimated mean (over N snap shots). Beamforming can be defined as steering the main beam of the array in directions of interest by applying weights a_i and delays τ_i to the M channel sensor outputs x_i,

$$A(t,\Theta) = \sum_{i=1}^{M} a_i x_i(t-\tau_i) \quad (6)$$

The weights a_i can for simplicity be set to unity and the time delays are equal to $\tau_i = \frac{d(i-1)}{v}\sin\Theta$, where Θ is the bearing, v the mean propagation velocity, and d is the inter-sensor distance.

If the output $A(t,\Theta)$ in Eq. 6 is transformed to frequency domain the output power of the array can be defined as;

$$P(\Theta) = E\{|A(\omega)|^2\}. \quad (7)$$

The standard definition of conventional beamforming is obtained, e.g. [9]

$$P(\Theta) = \mathbf{C}^{\mathbf{H}}(\Theta)\mathbf{R}_{xx}\mathbf{C}(\Theta). \quad (8)$$

The steering vector

$$\mathbf{C}(\Theta) = [1, e^{j\Phi(1)}, ..., e^{j\Phi(M-1)}]$$

steer the main beam in the Θ directions of interest, for a linear array at sensor m that yields $\Phi(m) = \frac{2\pi d sin(\Theta)}{\lambda}m$, where d is the sensor spacing and λ is the wave length.

2.2.2. High resolution methods

There are many different sensitive high resolution methods. In this study we used two high-resolution methods, namely MUSIC [10] and minimum variance (MV) by [14]. They represents two different categories of methods. The common definition of the term high bearing resolution is that, at least two sources are resolved within the array beam width, e.g. [9] and [10]. We use the standard formulation of high resolution methods formed by the orthogonal decomposition of the covariance matrix $\mathbf{R}_{xx}$ Eq. 5 into a signal subspace $\mathbf{R}_{ss}$ and a noise subspace $\mathbf{R}_{nn}$,

$$\mathbf{R}_{xx} = \mathbf{R}_{ss} + \mathbf{R}_{nn} \quad (9)$$

$$\mathbf{R}_{ss} = \sum_{r=1}^{P}\lambda_r\mathbf{V}_r\mathbf{V}_r^H \quad (10)$$

$$\mathbf{R}_{nn} = \sum_{r=P+1}^{M}\lambda_r\mathbf{V}_r\mathbf{V}_r^H \quad (11)$$

The signal subspace Eq. 10 is spanned by the P 'largest' eigenvalues λ_r and eigenvectors $\mathbf{V}_r$ of $\mathbf{R}_x x$, and it is separated from the noise subspace, Eq. 11, formed by the 'smallest' eigenvalues and vectors. One drawback with the signal and noise subspace methods is that the size of the signal subspace P needs to be estimated or known a priori. This is a difficult task in real situations, e.g. [15]. Here, the Minimum Description Length, MDL, is used for recursive estimation of the size of the subspace, [12]. The estimation of the number of independent sources P is similar to the problem of the model order selection for parametric AR and ARMA models. The key to high performance ARMA [16] is to secure the set of equations to be overdetermined (i.e. $N > 2P$ with N equal the data length). Unfortunately, if the number of sources P is either too large or too small, then the decomposition of the signal subspace in Eq. 10 and noise subspace in Eq. 11 is wrong, the subspace methods can produce biased or false bearings. Therefore, it is important to ensure that the subspaces are correctly sized.

The minimum variance (MV) method is often called the maximum likelihood method or the minimum energy method in the literature. Here, the MV is formulated as a signal subspace method by using the singular value decomposition of the covariance matrix above, Eq. 9 as:

$$P_{MV}(\Theta) = \frac{1}{\mathbf{C}^H(\Theta)\mathbf{R}_{SS}^{-1}\mathbf{C}(\Theta)} \quad (12)$$

The MUSIC estimator [10] can be formulated as in Eq. 13 by using the noise subspace. If

there are signals in some directions Θ, the denominator in Eq. 13 becomes small, making $P_{MUSIC}(\Theta)$ large in signal directions.

$$P_{MUSIC}(\Theta) = \frac{1}{\sum_{P+1}^{M} |\mathbf{C}^H(\Theta)\mathbf{V}_r|^2} \quad (13)$$

The steering vector $C(\Theta)$ in Eq. 12 and Eq. 13 is defined as in Eq. 8. MV uses the noise free signal subspace in Eq. 12 for bearing estimation and is therefore called a signal subspace method. The formulation of MUSIC with the noise subspace in Eq. 13 is a class of estimators often called noise subspace methods. The two estimates are often used in parallel with a beamformer for calibration of the peaks related to source bearings.

2.2.3. Bearing estimation analysis of sea trial data

For the analysis of the sea trial data, we use the aforementioned three bearing estimation methods (CB, MV and MUSIC). The acoustic data was recorded using a 31-element uniform line array (ULA), with sensor spacing 1.5 m, placed on the seafloor at a water depth of around 42 m. Half of that array was placed in a vertical position during tracks 61c1 and 61c2. In this case, only data from the 15 elements remaining in the horizontal part of the array has been used. The acoustic array was placed some 100 m west of the electrode array. An acoustic source was towed after a ship in parallel with the electromagnetic source during tracks 31c and 61c1/c2, while the only acoustic source during track 24a was the ship itself.

The quality of the covariance matrix is crucial in bearing estimation. The condition number is used as a measure of the numerical stability for each bearing estimate. The condition number used here is defined as the ratio of the largest singular value to the smallest. Large condition numbers indicate a badly conditioned covariance matrix. Also, the MDL estimate of the number of sources is used as a quality control measure. Furthermore, in this analysis, planar wave fronts are assumed to propagate over the array of hydrophones.

Table 1: *Summary of RMS, mean and standard deviation of difference between GPS bearing and estimated bearing for track 24a.*

Method	RMS value	Mean	STD
CB	4.05	-2.55	3.15
MV	4.27	-1.04	4.14
MUSIC	5.25	-2.25	4.74

The final bearing track is made by allowing bearing changes for consecutive bearing estimates of no more than 10 degrees. Also, a fifth order smoothing filter is applied to all bearing estimates.

Analysis of track 24a

Track 24a is a stable track with small changes in bearing during the 900 seconds of recordings, with 31 channels. In Figure 15 the three bearing methods are displayed as a function of time. Also the root-mean-square (RMS), mean and standard deviation of the difference between the bearing estimates and the GPS bearings are given in Table 1. The frequency band we use for the estimation is between 125 Hz to 500 Hz. The overall best results are for CB but the bias is smaller for both MV and MUSIC. The condition value of the covariance matrix (Figure 16) and the number of estimated sources (Figure 17) are quite stable and are not indicating any estimation errors. The errors are comparable with the assumed accuracy for this this type of array at the given circumstances.

Analysis of track 31c

Track 31c is a longer recording than track 24a and the bearings to the target shows a larger variation. The frequency band for estimation is here between 125 and 300 Hz. The outputs from the three bearing estimation methods are displayed in Figure 18 as a function of time. For comparison, the root-mean-square (RMS), mean and standard deviation of the difference between the bearing estimates and the GPS bearings are given in Table 2. As for track 24, the errors are smaller for CB compared with MV and MUSIC, but the bias is largest for CB. The number of estimated

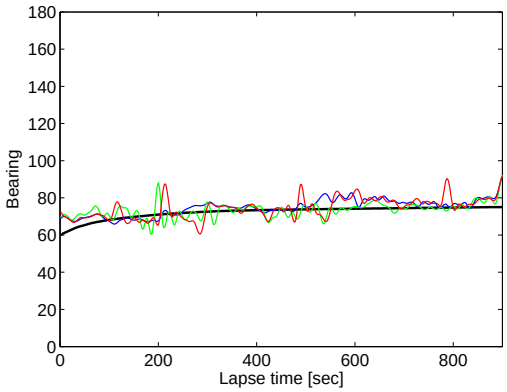


Figure 15: Bearing estimates of track 24a for CB (blue), MV (green) and MUSIC (red) and bearing from GPS (black).

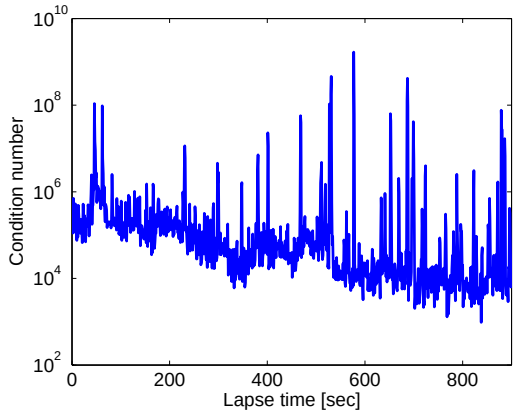


Figure 16: Condition number for the covariance matrix at each second of track 24a.

sources and the condition value are displayed in Figs. 20 and 19. The largest bearing errors are seen at around a lapse time of 2700 sec and in Fig. 20 with more than five sources.

Analysis of track 61c1 and 61c2

Track 61 is divided into two parts (61c1 and 61c2). The frequency band is the same as for track 31c (125 to 300 Hz). The statistics of the bearing results from 61c1 and 61c2 are presented in Table 3 and Table 4, respectively. The RMS error is greater for tracks 61c1 and 61c2, compared to the other tracks (tracks 24a and 31c). However, this is expected since the array is much smaller in this case (only 15 sensors were used). For both parts of track 61 CB and MV show the lowest RMS and standard deviation. The bias is lowest for MUSIC in the first part. The bearing estimates for both tracks are displayed in Figure 21 and 22. The number of estimated sources by the MDL algorithm and the condition number are for both tracks displayed in Figure 23 and 24 and in Figure 25 and 26. We can see a decrease in performance at regions with high condition value, e.g. at lapse time of 700 sec and 1700 sec for part 1. Also the performance of MUSIC is affected by the estimated number of sources, e.g. for parts with a too high estimated number of sources the variance of the estimate is increased.

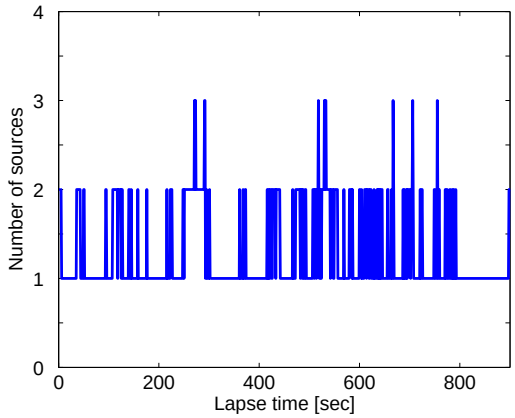


Figure 17: Number of estimated sources by MDL for each second of track 24a.

Table 2: *Summary of RMS, mean and standard deviation of difference between GPS bearing and estimated bearing for track 31c.*

Method	RMS value	Mean	STD
CB	12.07	4.09	11.36
MV	16.32	0.91	16.30
MUSIC	18.05	-0.66	18.04

2.2.4. Conclusions of bearing analysis

A performance comparison of three different types of bearing estimators (CB, MV and MUSIC) is the main issue in this section. The intention is to find the most stable estimator to be used as an input for the tracking using data fusion. Tracks 24 and 31 with the

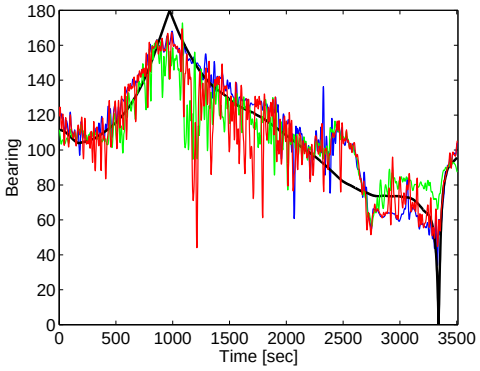


Figure 18: Bearing estimates of track 31c for CB (blue), MV (green) and MUSIC (red) and bearing from GPS (black).

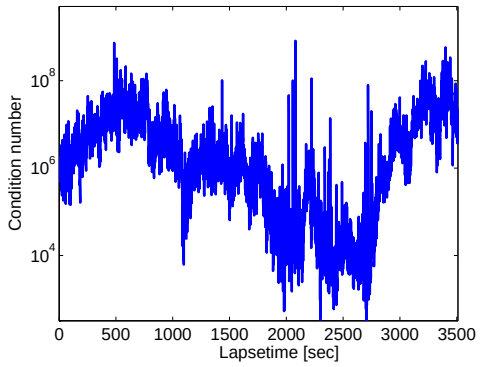


Figure 19: Condition number for the covariance matrix at each second of track 31c.

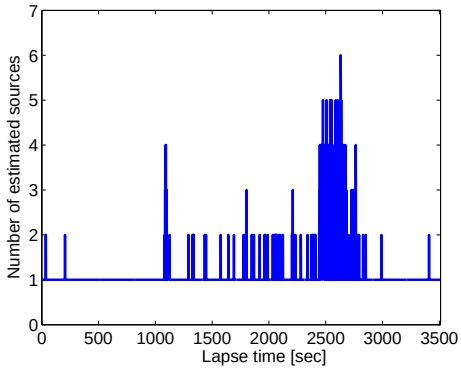


Figure 20: Number of estimated sources by MDL for each second of track 31c.

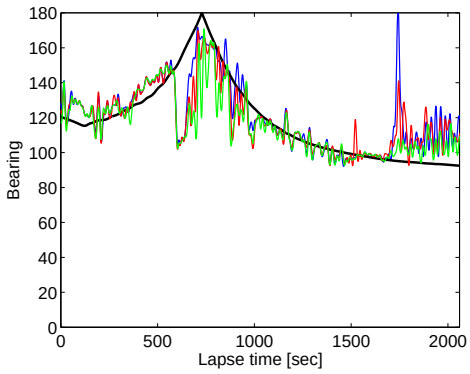


Figure 21: Bearing estimates track 61c1 for CB (blue), MV (green) and MUSIC (red) and bearing from GPS (black).

larger aperture show a smaller variance in the estimate as expected. However, all three estimators have overall comparable statistic performance although partly some differences are data related, supported by the quality control measures MDL and condition value of the spatial covariance matrix. For the further tracking process, the MV bearing estimator has been used.

2.3. Kalman filtering

This study incorporates data fusion in tracking by combining positional information using data from sensor arrays. In the following section the basic equations for a Kalman filter in standard form [17] are given. The Kalman filter provides a general solution to the recursive minimised mean square linear estimation problem. The mean square error will be minimised as long as the target dynamics and the measurement noise are accurately modelled. In addition, the Kalman filter provides a convenient measure of the estimation accuracy through the covariance matrix, and the gain sequence automatically adapts to changing detection histories.

2.3.1. Kalman filter definition

Derivations of the Kalman filter are, for example, presented in [18] and [19], therefore only the resulting equations are given here. Assume that the target dynamic process can

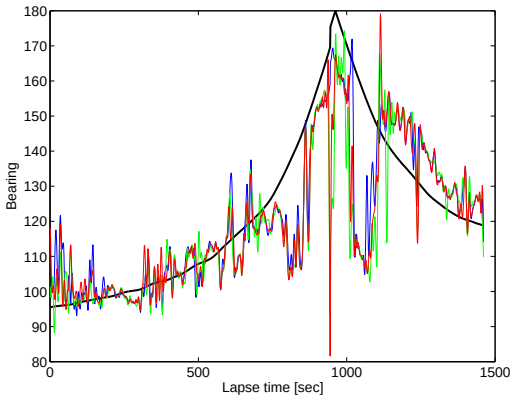


Figure 22: Bearing estimates track 61c2 for CB (blue), MV (green) and MUSIC (red) and bearing from GPS (black).

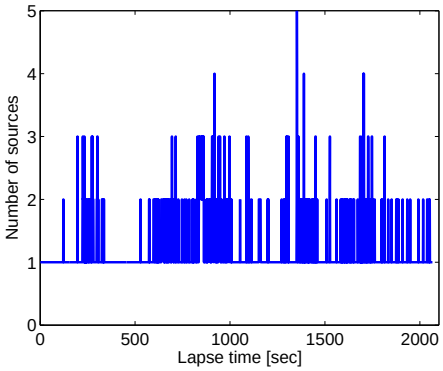


Figure 23: Number of estimated sources by MDL for each second of track 61c1.

be modelled in the discrete-time Markov form

$$\mathbf{x}(k+1) = \mathbf{\Phi}\mathbf{x}(k) + \mathbf{q}(k) + \mathbf{f}(k+1|k) \quad (14)$$

where $\mathbf{x}(k)$ is the n-dimensional target state vector with time index k, $\mathbf{\Phi}$ is the assumed state transition matrix, $\mathbf{q}(k)$ is the zero-mean, white, Gaussian process noise with assumed known covariance $\mathbf{Q}$, and $\mathbf{f}(k+1|k)$ is the assumed known deterministic input. For this application, the discrete-time Markov process, described by the difference equation 14, implies that the statistical representation of the process in the future (scan $k+1$) is completely determined by the present state (scan k). Assume further that the scan steps k are advanced by equal time steps T.

Measurements are in the form of linear combinations of the system state variables, corrupted by uncorrelated noise, modelled as

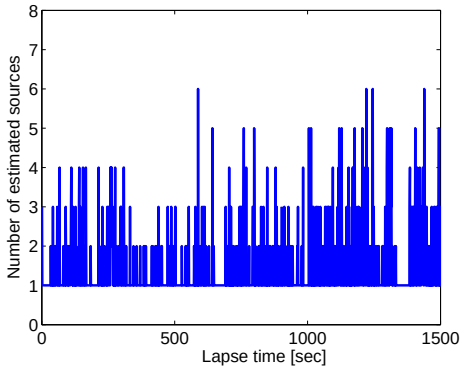


Figure 24: Number of estimated sources by MDL for each second of track 61c2.

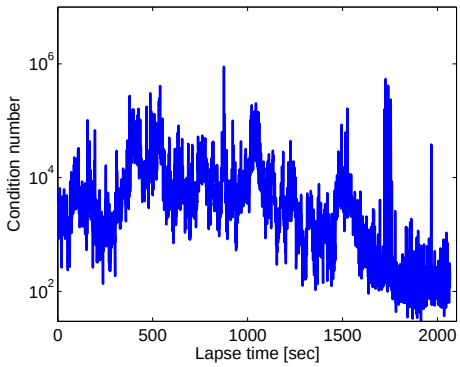


Figure 25: Condition number for the covariance matrix at each second of track 61c1.

$$\mathbf{y}(k) = \mathbf{H}\mathbf{x}(x) + \mathbf{v}(k) \quad (15)$$

where $\mathbf{y}$ is the M-dimensional measurement vector, $\mathbf{H}$ is the $M \times n$ measurement matrix, and $\mathbf{v}$ is zero-mean, white, Gaussian noise with covariance $\mathbf{R}$. Note that in general $\mathbf{Q}$ and $\mathbf{H}$ may also vary with time, though kept constant in our application.

Given the target dynamics and measurement models from Eq. 14 and Eq. 15, the Kalman filter equations become

$$\hat{\mathbf{x}}(k|k) = \hat{\mathbf{x}}(k|k-1) + \cdots$$
$$\mathbf{K}(k)\left[\mathbf{y}(k) - \mathbf{H}\hat{\mathbf{x}}(k|k-1)\right] \quad (16)$$

$$\mathbf{K}(k) = \cdots$$

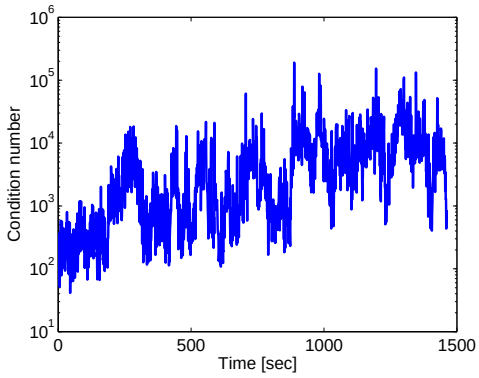


Figure 26: Condition number for the covariance matrix at each second of track 61c2.

Table 3: *Summary of RMS, mean and standard deviation of difference between GPS bearing and estimated bearing for track 61c1.*

Method	RMS value	Mean	STD
CB	14.40	2.44	14.19
MV	15.45	-2.51	15.25
MUSIC	14.46	0.52	14.55

$$\mathbf{P}(k|k-1)\mathbf{H}' \left[\mathbf{H}\mathbf{P}(k|k-1)\mathbf{H}' + \mathbf{R}\right]^{-1} \quad (17)$$

$$\begin{aligned}\mathbf{P}(k|k) = \cdots \\ [\mathbf{I} - \mathbf{K}(k)\mathbf{H}]\mathbf{P}(k|k-1)[\mathbf{I} - \mathbf{K}(k)\mathbf{H}]' + \cdots \\ \mathbf{K}(k)\mathbf{R}\mathbf{K}'(k)\end{aligned} \quad (18)$$

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{\Phi}\hat{\mathbf{x}}(k|k) + \mathbf{f}(k+1|k) \quad (19)$$

$$\mathbf{P}(k+1|k) = \mathbf{\Phi}\mathbf{P}(k|k)\mathbf{\Phi}' + \mathbf{Q} \quad (20)$$

Table 4: *Summary of RMS, mean and standard deviation of difference between GPS bearing and estimated bearing for track 61c2.*

Method	RMS value	Mean	STD
CB	13.32	-3.62	12.82
MV	16.76	-5.77	15.74
MUSIC	15.39	-4.85	14.61

where $\mathbf{H}'$ denotes the transpose of $\mathbf{H}$, etc. and Eq. 17 gives the stabilised form of the covariance update. The state covariance matrix is defined in terms of the zero-mean Gaussian estimation error vector

$$\mathbf{P}(k) = E\{[\mathbf{x}(k) - \hat{\mathbf{x}}]\,[\mathbf{x}(k) - \hat{\mathbf{x}}]'\} \quad (21)$$

2.3.2. Tracker Implementation

The tracker implementation adopts mainly the MATLAB code given in [20]. In the following a 2-D planar system with state vector $\mathbf{x} = [x_1, x_2, v_1, v_2]$ is assumed. The main program, *track_uu*, sets parameters and feeds sensor data (bearings and positions computed earlier) to an administration program, *adm_kalman_filter2*. Here data are translated and co-rotated to the line of sight, to be processed by the inner most wheels, *prep_update* (checks for new data at regular times, flagging update/predict only) and *kalman_update* (the update engine). The output is stored in batch variables and displayed optionally.

Track initiation is done by selecting "rule-of-thumb" parameter values for measurement noise, process noise, state information, state estimate and covariance. Since we expect very slow manoeuvres, the state transition matrix $\mathbf{\Phi}$, the process noise covariance matrix $\mathbf{Q}$, the (unrotated) noise covariance matrices $\mathbf{R}_0$ (one for each sensor type), and the measurement matrix $\mathbf{H}$ have constant, simple forms

$$\mathbf{\Phi} = \begin{pmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (22)$$

$$\mathbf{Q} = q \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (23)$$

$$\mathbf{R}_0 = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \quad (24)$$

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad (25)$$

where T is the time step, q is a scale factor, and the components σ_x^2 and σ_y^2 correspond to measurement $\mathbf{x}$ and $\mathbf{y}$ errors. The state $\mathbf{x}$ is initialised to $\mathbf{x}_0 = [x_{10}, x_{20}, 0, 0]$, taking the coordinates of first accepted detection. The state covariance $\mathbf{P}$ is initialised to a matrix times a scale factor. The matrix $\mathbf{V}$ co-rotates $\mathbf{R}_0$ through a bearing angle b into R

$$\mathbf{R} = \mathbf{V}\mathbf{R}_0\mathbf{V}' \quad (26)$$

$$\mathbf{V} = \begin{pmatrix} \cos b & -\sin b \\ \sin b & \cos b \end{pmatrix} \quad (27)$$

Typically, a new target is first observed by an outpost single sensor. An ellipse, defined by the $\mathbf{R}$ matrix, indicates the observation uncertainty region (highly elongated in the bearing direction having bearing data only). The first position estimate is put at the centre of this ellipse. Subsequent observations from the same sensor add more information, sometimes with the effect to artificially "attract" the target to the sensor. Eventually, other (arrays of) sensors contribute with independent observations improving the tracking. An important point is the association of an observation to the target. Therefore, valid observations are gated within a maximum window defined by a normalised distance,$(\mathbf{y}-\hat{\mathbf{x}})'\mathbf{P}_p^{-1}(\mathbf{y}-\hat{\mathbf{x}}) < maxgate$, where $\mathbf{P}_p$ is a submatrix of the state covariance, containing only the position covariance.

2.3.3. Tracking targets using data from field experiments

The time signals from hydrophone arrays and electromagnetic arrays analysed in this section are the ones mentioned in previous chapters, (tracks 24a, 31c, 61c1, and track 61c2).

Track 24a

Hydrophone (Ak) and electromagnetic (EM) data from track 24a have been analysed resulting in reconstructed coordinates for the target. The acoustic array delivered bearings and the EM array gave positions of a signal emitter towed after a vessel. A GPS on board the vessel provided reference positions showing the "true" track. The acoustic signals were segmented into 1 s windows and the EM signals were segmented into 5 s windows suggesting a 0.5 s time step in the Kalman filter. For each data point we have assumed that the acoustic array gives the target bearing within ±5 degrees about the true direction. Consequently the EM data are mainly used to get the target distance.

The electromagnetic measurement covariance has been scaled with the local mean square residual between model and data, influencing the Kalman tracker. Occasional very big residuals thus make a negligible contribution to the track estimate.

Figure 27 shows a track estimate using data from track 24a. The track estimate starts in a position, in the upper left part of the figure, calculated from the first EM measurement. The initially big error ellipses indicate an uncertain start, followed by some very quick steps past the hydrophone array centre. The crowd of EM points surrounding position (100,-800) lie at the limit of EM sensitivity so all further estimates rely on acoustic bearings only. Not surprisingly this later part of the track estimate deviates a lot from the true track given by GPS readings. The final track estimate is running freely, out of reach of both sensors.

Track 31c

Hydrophone and electromagnetic data from track 31c were analysed in a similar manner

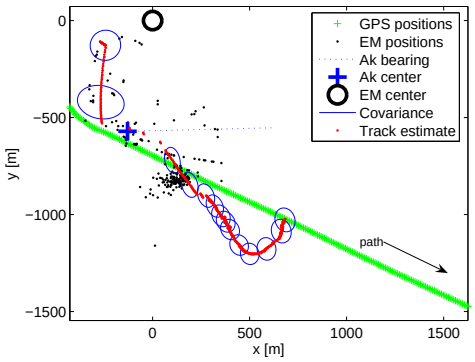


Figure 27: Track 24a. Estimated target track (red dots each time step) using bearings from a hydrophone array (Ak) near coordinates (-200, -500)), and position data from an electromagnetic array (EM) at (0, 0). Black dots denote calculated EM positions, green crosses denote GPS positions. State error ellipses are shown in blue every 100 steps. Relatively high weight is put on acoustic data compared to EM data.

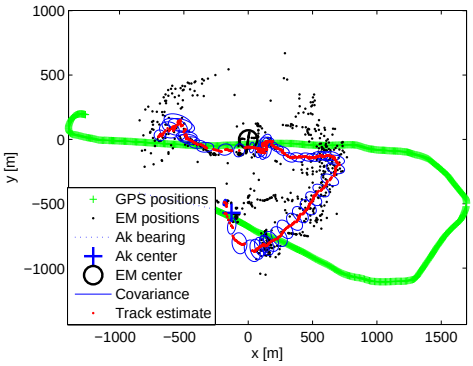


Figure 28: Track 31c. Estimated target track (red dots) using bearings from a hydrophone array (Ak) at coordinates (0, -500)), and position data from an electromagnetic array (EM) at (0, 0). Black dots denote calculated EM positions, green crosses denote GPS positions. State error ellipses are shown in blue every 100 steps. Relatively high weight is put on Ak data compared to EM data.

to that used for track 24a.

Figure 28 shows a track estimate using data from track 31c. Note how the state error ellipses steadily grow after passing the swarm of EM positions just right of the picture centre. A bit surprising is that using EM data only the resulting track turns out to be about the same, indicating the weak influence of the accompanying sonar array. The large discrepancy between EM and GPS positions in this part of the track evidently calls for some radical new interpretation of EM data or extra supporting sensors.

Using the GPS positions a second hydrophone array positioned at (1400,-1100) is simulated, see Figure 29. The simulated acoustic array (Ak2) outputs bearings at the same rate and error characteristics as the real one (Ak1). This time, thanks to the extra information from the second acoustic array, the track estimate follows the "true" track (GPS readings) much better, completely avoiding the central crowd of false EM positions. It should be noted, however, that the second array must be positioned carefully in order to support the bearings from the first array.

Tracks 61c1 and 61c2

Electromagnetic and acoustic signals from tracks 61c1 and 61c2 were analysed in a similar manner to that used for track 31c resulting in reconstructed coordinates for the target. EM data was segmented into 5 second windows using the E-field from all eight antenna channels.

Figure 30 shows a track estimate using data from track 61c1, using positions from the EM array and bearings from the acoustic array. Except from an initial discrepancy between the track estimate and GPS positions the main remaining part of the track estimate follows the GPS positions rather well. The differently sized error ellipses reflect the changing state estimate errors, getting smaller towards the EM centre. The actual position error, however, has a local maximum at this point. This is because the used EM model does not take into account nearfield effects. The acoustic array, on the other hand, has the best performance when the track is broadside to the array. In this case, the track pass the end of the array. This means that the geometry of the track and the arrays are such that the acoustic bearings are of little

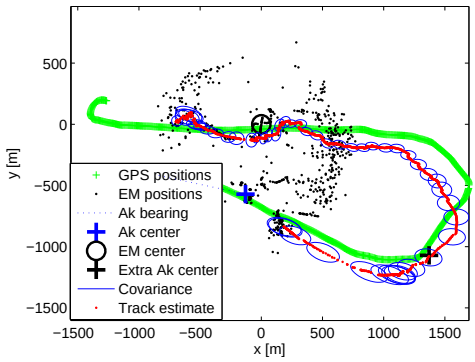


Figure 29: Track 31c. Estimated target track (red dots) using bearings from two hydrophone arrays, Ak1 (real, at -100, -600) and Ak2 (simulated, at 1400, -1100), and position data from an electromagnetic array, EM. See Figure 28 for details.

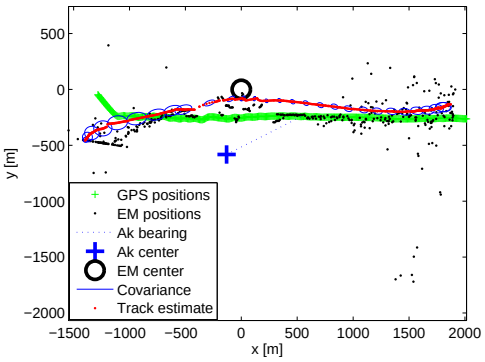


Figure 30: Track 61c1. EM positions fused with acoustic bearings. The track goes from left to right.

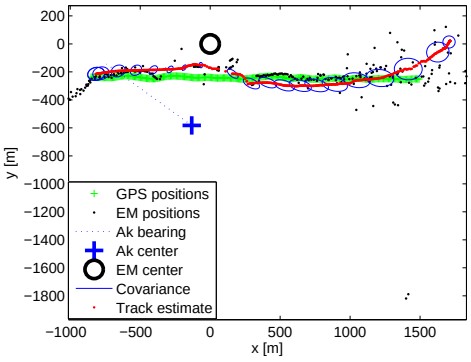


Figure 31: Track 61c2. EM positions fused with acoustic bearings. The track goes from right to left.

help here.

The returning part of the track, running from right to left in Figure 31, uses EM and acoustic data from track 61c2. Except from a uncertain start, the remaining track estimate lies somewhat closer to the GPS readings than in case 61c1. Again there is a maximal deviation between EM and Ak locations, which could be blamed on the particular layout of these sensors.

2.3.4. Summary

Tracking of underwater targets is normally improved using complementary data from hydrophone arrays (bearings) and electromagnetic arrays (positions). The EM sensors have, compared with the acoustic sensors, relatively short detection ranges. In addition to that, target ranges are underestimated in the current analysis, even when the SNR is good. This is probably caused by a mismatch between the actual underwater environment and the environment model used. Filtering of EM data, constrained by sonar bearings, help to some extent. A real improvement can be achieved by adding carefully positioned extra sensors, e.g. another sonar array.

3. CLASSIFICATION

In this Chapter a classification study on real-world extremely low frequency electric (ELFE) signatures and acoustic signatures is presented. The classification problem can be divided into two main steps, feature extraction and decision making. Two feature extraction methods are investigated, one that is based on an AR (Auto Regressive) model and one based on a nonlinear DDE (Delay Differential Equation) model. Utilising the estimated model coefficients as features Bayesian minimum-error-classifiers can be implemented. Separate classifiers based on features extracted from the ELFE and the hydroacoustic sig-

natures are designed. Then the ELFE and underwater acoustic classifiers are fused with Bayes combination rule.

3.1. Feature extraction

The first step in the classification procedure is feature extraction. The features have to represent the raw data in a compact way, maintaining the relevant information that is required in order to perform classification. Two methods are used to extract features from the measured signal.

The first method is based on the parameters of an AR model. An AR model of order p can be written as

$$x(n) = -\sum_{k=1}^{p} a_k x(n-k) + e(n). \qquad (28)$$

where $x(n)$ is the measured signal and $e(n)$ is white noise with variance σ^2. The signal is divided into observation windows of length N. The model is fitted to each observation window of the signal with Yule-Walkers method [21]. The model parameters a_1 to a_p are used-as classification features. The model order p is selected based on the classification performance.

The second method is based on the parameters of a nonlinear DDE model [22] defined by

$$\dot{x} = a_1 x_{\tau_1} + a_2 x_{\tau_2} + a_3 x_{\tau_1} x_{\tau_2} \qquad (29)$$

where $\dot{x} \equiv \dot{x}(t)$ and $x_\tau \equiv x(t-\tau)$. Since we do not have access to the continuous signal $x(t)$ we use the discrete time signal $x(n)$. The signal is divided into observation windows of length N. The model coefficients a_1, a_2, and a_3 in Eq. 29 are estimated for each observation window (with a correlation method) [23],[24], and they comprise the classification feature space. The model has to be tuned to the signals of interest, which is done by finding the delays τ_1 and τ_2 that result in the best possible classification performance.

3.2. Minimum-Error-Rate Classification

The model coefficients from the AR and the DDE model are used as classification features since they reflect the signal properties in a compact way. In the estimation process these features are standardized by normalising each observation window to zero mean and unit variance. The actual classification (i.e. partitioning of the feature space) can be done with any standard classification method e.g. Bayesian, nearest neighbour or neural networks.

A Bayesian approach is chosen to build a minimum-error-rate classifier [25]. For classification purposes S signal classes $\omega_1, ..., \omega_S$ are used. For each class ω_n there are N_n observations using a fixed window length. By applying one of the feature extraction methods discussed above to each class ω_n a set of features $\left\{\mathbf{A}_i^{(\omega_n)}\right\}_{i=1}^{N_n}$ is obtained, where $\mathbf{A} = (a_1, ..., a_p)^t$ is the feature vector. There are many ways to represent a Bayesian minimum-error-rate classifier, here the following formulation is used: assign a feature vector $\mathbf{A}$ to class ω_j if

$$P(\omega_j|\mathbf{A}) > P(\omega_i|\mathbf{A}) \quad \text{for all} \quad i \neq j. \qquad (30)$$

Where $P(\omega_j|\mathbf{A})$ is the a posteriori probability, which can be calculated from Bayes rule:

$$P(\omega_j|\mathbf{A}) = \frac{p(\mathbf{A}|\omega_j)P(\omega_j)}{\sum_{j=1}^{S} p(\mathbf{A}|\omega_j)P(\omega_j)} \qquad (31)$$

where $p(\mathbf{A}|\omega_j)$ is the state conditional probability function for $\mathbf{A}$ conditioned on ω_j and $P(\omega_j)$ is the a priori probability of the true state of nature is ω_j. Further, if the feature distributions are multivariate normal and the a priori probabilities of all classes are equal, then Eq. 31 is reduced to

$$P(\omega_j|\mathbf{A}) = \frac{p(\mathbf{A}|\omega_j)}{\sum_{j=1}^{S} p(\mathbf{A}|\omega_j)} \qquad (32)$$

where

$$p(\mathbf{A}|\omega_j) = \frac{1}{(2\pi)^{p/2}|\mathbf{\Sigma}_j|^{1/2}} \times \cdots$$

$$\exp\left[-\frac{1}{2}(\mathbf{A}-\boldsymbol{\mu}_j)^t \mathbf{\Sigma}_j^{-1}(\mathbf{A}-\boldsymbol{\mu}_j)\right], \qquad (33)$$

where $\boldsymbol{\mu}_j$ is the mean vector and $\boldsymbol{\Sigma}_j$ is the covariance matrix for class ω_j, which often are unknown and therefore have to be estimated from the measured data.

An important objective in this report is to investigate if it is possible to achieve performance gains by data/information fusion. Therefore it is of great interest to fuse the decisions from the classifiers working on either ELFE or hydroacoustic signatures. This can be done in many ways (e.g. Maximum, Mean, Sum, Product or simply by voting). A well established method in information fusion [26], is to use the Bayes combination rule to compute the combined a posteriori probability that a observation belongs to class ω_j given a measurement $\mathbf{Z}$.

$$P(\omega_j|\mathbf{Z}) = \frac{\prod_{k=1}^{K} P(\omega_j|\mathbf{Z}_k)}{\sum_{j=1}^{S}\left(\prod_{k=1}^{K} P(\omega_j|\mathbf{Z}_k)\right)} \quad (34)$$

where $P(\omega_j|\mathbf{Z}_k)$ is the output a posteriori probability from classifier k and K is the total number of classifiers. Now a fused decision can be made with the following decision rule:

$$\text{Decide class } \omega_j \text{ if } \quad P(\omega_j|\mathbf{Z}) > P(\omega_i|\mathbf{Z}) \quad \cdots$$
$$\text{for all } \quad i \neq j. \quad (35)$$

Hence, the class with the maximum combined a posteriori probability is chosen. Bayes combination rule as stated in Eq. 34 assumes that all classifiers are equally reliable.

3.3. Data Analysis

The data set analysed here was recorded at a sea trial conducted in Oxdjupet, October 2000 (for details see [2]). Both electromagnetic sensors and hydrophones were used to record signatures of ships passing the sensor system. All data was recorded with the sampling rate of 2000 Hz. During the recording the electrode data was low pass filtered at 10 Hz and the hydrophone data was low pass filtered at 500 Hz. In Table 5 all passages which include data from both electrodes and the hydrophone are listed. Here we design and evaluate two classifiers, one that tries to determine the number of engines and one that tries to discriminate between small and big ships.

Some ships have two engines and some have four engines (see the fourth column of Table 5), The aim is to design classifiers that can determine the number of engines. One potential problem is that the data set only includes three recordings of ships with two engines and 20 recordings of ships with four engines.

In the third column of Table 5 the weight (or displacement in tons) of the ships are listed. We have divided them into small (up to 21484 tons) and big (above 21484 tons) ships based on their weight. The threshold 21484 tons between small and big ships is chosen to result in data classes of equal size.

Table 5: *The table shows the available data set: ship name (column 1), track number (column 2), displacement in tons (column 3) and the number of engines (column 4).*

Ship Name	Track nr.	Disp.	#
Amorella	92	32483	4
Baltic Star	66, 86	2890	2
Birka Princess	79, 88	21484	4
Gabriella	100	35492	4
Isabella	83, 101	34937	4
Mariella	96	37799	4
Mikhail Sholokov	64	12798	4
Palanga	95	11630	2
Rosella	89	10757	4
Sea Wind	63, 103	15879	4
Silja Europa	62,84,102	59914	4
Silja Festival	73, 94	33818	4
Silja Symphony	75, 99	58377	4
Star Wind	93	13788	4
Transest	65	6040	4

The electrode data is reference filtered and then band pass filtered between 0.4 to 10 Hz and down sampled to 100 Hz. The hydrophone data is low pass filtered at 400 Hz. Next, one segment of length 60 seconds from each passage listed in Table 5 is selected. This segment is selected so that it includes

the closest point of approach. The data segments are windowed, with a 4 second long observation window without overlap. Now the model parameters for the AR model in Eq. 28 and the DDE model in Eq. 29 can be estimated for each observation window. Since the electrode sensor has three axes we get three ELFE time series from each passage. The model parameters are estimated for each of the three axes separately and then averaged to result in one set of parameters for each observation window.

As mentioned earlier, both methods have to be tuned to the classification problem at hand. For the AR method the model order p has to be determined, and for he DDE method the two delays have to be selected. The criteria for selection for both methods is to find the model that results in the best classification performance. In addition for the AR method the model order p should be kept as small as possible.

We implemented one classifier that tries to determine if the ship has two or four engines and one that tries to classify the ship as small or big. This is done with features extracted with both the AR and the DDE model from both the ELFE and from the hydroacoustic signal. For model selection purposes the classification performance is evaluated with a leave one out approach. That is the classifier is trained (i.e. the mean vectors and covariance matrices for all classes are estimated) with all but one observation window, and then the trained classifier is used to classify the removed observation window. This is repeated until all observation windows have been removed and assigned to one of the classes. The best classification result together with the selected model parameters for the AR and DDE model are listed in Table 6 and Table 7, respectively. As can be seen from the tables all classifiers work fairly well (during the model selection) with a correct classification rate between 79% and 98%.

Next, the classification performance is validated by using separate training and testing sets. A randomly selected subset (70%) from all classes are used as a training set (i.e. for estimating the mean and covariance matrices). The remaining 30% of the distributions are then used for testing of the classifier. The training and testing procedure is repeated a hundred times to remove fluctuations in the estimated classification performance. The results of the numerical analysis above is summarised in confusion matrices. A confusion matrix is a table showing the true class of the input features, versus the output of the classification algorithm using the test set of features.

In addition we investigate if the classification performance can be improved by data fusion. The classifiers working on features extracted from either ELFE or hydroacoustic signals are fused with Bayes combination rule Eq. 34, which computes the combined a posteriori probability for each class. The class with the highest combined a posteriori probability is selected.

Table 6: *In the table below the best classification performance (during the model selection process) and the selected model order p for the AR model are listed.*

Classifier	% Cl.	p
2 vs. 4 engines (ELFE)	85%	15
2 vs. 4 engines (Acoustic)	92%	15
Small vs. big (ELFE)	89%	20
Small vs. big (Acoustic)	98%	20

Table 7: *In the table below the best classification performance (during the model selection process) and the selected delays (given in samples) for the DDE model are listed.*

Classifier	% Cl.	τ_1, τ_2
2 vs. 4 engines (ELFE)	82%	$(31, 3)$
2 vs. 4 engines (Acoustic)	92%	$(90, 6)$
Small vs. big (ELFE)	79%	$(58, 4)$
Small vs. big (Acoustic)	84%	$(48, 7)$

3.3.1. Classification of the number of engines with AR features

In this section we summarise the classification results achieved with AR features as input to classifiers that tries to determine the number of engines (2 or 4).

In Table 8 the results from classification of the number of engines using AR features estimated from the electrode data are listed. As can be seen the classifier does not work at all for the case with 2 engines, it only classifies 49.9% correctly. For the four engine case the correct classification rate is 98.8%. This classifier is practically useless since it does not perform better than guessing for one of the classes.

The performance for classification of the number of engines using AR features extracted from the hydrophone data are listed in Table 9. Here the classifier works nearly perfect for the 4 engine case while it has a error of 30% in 2 engines case.

In Table 10 the fused classification result (of the classifiers in Table 8 and Table 9) of classification of the number of engines based on AR features are listed. In this case data fusion does not improve the performance compared to the best stand alone classifier (with AR features extracted from hydro-acoustic data). This is expected since one of the underlying classifiers does not work at all (guessing would give as much information about the true class).

In conclusion, classification of the number of engines only works with AR features extracted from hydrophone data. On ELFE signals the classifier does not work at all, guessing would give a better result. Since one of the classifiers does not work at all fusing the two classifiers only results in a performance degradation (compared to the best stand alone classifier).

Table 8: *Classification of the number of engines based on AR features extracted from the electrode data.*

Output Decision$\Longrightarrow$ True Class $\Downarrow$	2 engines	4 engines
2 engines	49.9%	50.1%
4 engines	1.2%	98.8%

Table 9: *Classification of the number of engines based on AR features extracted from the hydrophone data.*

Output Decision$\Longrightarrow$ True Class $\Downarrow$	2 engines	4 engines
2 engines	70.1%	29.9%
4 engines	1.0%	99.0%

Table 10: *Classification of number of engines based on AR features extracted from both the electrode and hydrophone data.*

Output Decision$\Longrightarrow$ True Class $\Downarrow$	2 engines	4 engines
2 engines	61.9%	38.1%
4 engines	0.0%	100.0%

3.3.2. Classification of ship size with AR features

In this section we summarise the classification results achieved with AR features as input to classifiers that tries to determine the ship size (small or big).

In Table 11 the results from classification of ship size using AR features estimated from the electrode data are listed. As can be seen the classifier works well for both small and big ships, with a average correct classification rate of 87%.

In Table 12 the results from classification of ship size using AR features estimated from the hydrophone data are listed. As can be seen the classifier works very well for both small and big ships, with a average correct classification rate of 97%.

In Table 13 the fused classification result (of the classifiers in Table 11 and Table 12) of classification of the number of engines based on AR features are listed. Here fusion slightly improves the performance compared to the best stand alone classifier (with AR features extracted from hydroacoustic data).

Classification of ship size with AR features works well for both hydroacoustic and ELFE signatures. However, AR features extracted from hydroacoustic data results in 10 % better classification performance than the AR features extracted from electrode data. Fusing the two classifiers results in a performance gain even if it is small (1%).

Table 11: *Classification of ship size based on AR features extracted from the electrode data.*

Output Decision⟹ True Class ⇓	Small	Big
Small	85.6%	14.4%
Big	11.1%	88.9%

Table 12: *Classification of ship size based on AR features extracted from the hydrophone data.*

Output Decision⟹ True Class ⇓	Small	Big
Small	99.0%	1.0%
Big	5.5%	94.5%

Table 13: *Classification of ship size based on AR features extracted from both the electrode and hydrophone data.*

Output Decision⟹ True Class ⇓	Small	Big
Small	98.9%	1.1%
Big	3.6%	96.4%

3.3.3. Classification of number of engines with DDE features

In this section we summarise the classification results achieved with DDE features as input to classifiers that tries to determine the number of engines (2 or 4).

The performance for classification of the number of engines using DDE features extracted from the electrode data are listed in Table 14. The classifier works well with a correct classification rate of 79% for both classes.

The performance for classification of the number of engines using DDE features extracted from the hydrophone data are listed in Table 15. The classifier works well with a average correct classification rate of 84% for both classes.

In Table 16 the fused classification result (of the classifiers in Table 14 and Table 15) of classification of the number of engines based on DDE features are listed. The average correct classification rate is 84% which is the same as for the classifier working only on hydroacoustic data. It seems like both stand alone classifiers have problems with the same events, and therefore we do not get a performance gain by fusing the classifiers.

Classification of the number of engines based on DDE features works quite well for both hydroacoustic and ELFE signatures. However the best performance is achieved with DDE features extracted from hydroacoustic data. Fusing the two classifiers, results in the same average classification rate as the best stand alone classifier.

Table 14: *Classification of number of engines based on DDE features extracted from the electrode data.*

Output Decision⟹ True Class ⇓	2 engines	4 engines
2 engines	79.4%	20.6%
4 engines	20.8%	79.2%

Table 15: *Classification of number of engines based on DDE features extracted from the hydrophone data.*

Output Decision⟹ True Class ⇓	2 engines	4 engines
2 engines	82.4%	17.6%
4 engines	13.7%	86.3%

Table 16: *Classification of number of engines based on DDE features extracted from both the electrode and hydrophone data.*

Output Decision⟹ True Class ⇓	2 engines	4 engines
2 engines	79.6%	20.4%
4 engines	11.6%	88.4%

3.3.4. Classification of ship size with DDE features

In this section we summarise the classification results achieved with DDE features as input to classifiers that tries to determine the ship size (small or big).

In Table 17 the results from classification of ship size using DDE features estimated from the electrode data are listed. As can be seen the classifier works quite well for both small and big ships, with a average classification rate of approximately 79.4% for both.

In Table 18 the results from classification of ship size using DDE features estimated from the hydrophone data are listed. As can be seen the classifier works quite well for both small and big ships, with a average correct classification rate of 84%.

In Table 19 the fused classification result (of the classifiers in Table 17 and Table 18) of classification of ship size based on DDE features are listed. In this case fusion results in a performance gain. The correct classification rate is improved with 10% compared to the ELFE classifier and with 6 % over the hydroacoustic classifier.

Classification of ship size based on DDE features works well both on hydroacoustic and ELFE signatures. Again the hydroacoustic signature gives the best classification result. In this case combining the two classifiers results in a performance gain compared to both stand alone classifiers. This result demonstrates that fusion can improve the overall classification performance on real data.

Table 17: *Classification of ship size based on DDE features extracted from the electrode data.*

Output Decision⟹ True Class ⇓	Small	Big
Small	79.9%	20.1%
Big	21.0%	79.0%

3.3.5. Simulation Results

In this section two computer simulations are presented in order to further illustrate how fusion can improve classification performance.

Table 18: *Classification of ship size based on DDE features extracted from the hydrophone data.*

Output Decision⟹ True Class ⇓	Small	Big
Small	82.5%	17.5%
Big	14.8%	85.2%

Table 19: *Classification of ship size based on DDE features extracted from both the electrode and hydrophone data.*

Output Decision⟹ True Class ⇓	Small	Big
Small	91.1%	8.9%
Big	12.2%	87.8%

In Figure 32 the combined a posteriori probability, $P(\omega_j|\mathbf{Z})$, of two classifiers for the true class ω_j is shown. We have two classifiers with output a posteriori probabilities $P(\omega_j|\mathbf{Z}_1)$ (classifier 1) and $P(\omega_j|\mathbf{Z}_2)$ (classifier 2). On the y-axis the combined a posteriori probability $P(\omega_j|\mathbf{Z})$ is shown and on the x-axis the a posteriori probability $P(\omega_j|\mathbf{Z}_2)$ of classifier 2 is shown. In the figure six graphs are displayed, each corresponding to a given a posteriori probability of classifier 1. The classifier makes the right decision if the combined a posteriori probability is greater than 0.5. As can be seen, if the combined a posteriori probability is greater than the a posteriori probability of classifier one, then the a posteriori probability of classifier two is greater than 0.5. Hence if the second classifier works (i.e has a a posteriori probability greater than 0.5) fusion will improve the classification performance. In conclusion the fused classification performance is better than the stand alone classifier if the added classifier has a posteriori probability greater than 0.5.

In Figure 33 the combined a posteriori probability, $P(\omega_j|\mathbf{Z})$, of up to ten classifiers (each individual classifier with equal a posteriori probabilities P) are shown. On the y-axis the combined a posteriori probability is shown and on the x-axis the number of classifiers is displayed. In the figure five graphs are displayed, each corresponding to various (between 0.5 and 0.9) a posteriori probabilities of the individual classifiers. As can be seen from the plot the combined a posteriori probability is improved as soon as the individual a posteriori probability is greater than 0.5.

Having ten classifiers (independent) in a real application is probably not realistic. However, just having two or three classifiers improves the probability of correct classification.

These simulations have shown that combining classifiers can improve the classification performance if the individual classifier works properly. If the classifiers have equal probability of correct classification the combined probability of correct classification increase with the number of classifiers used.

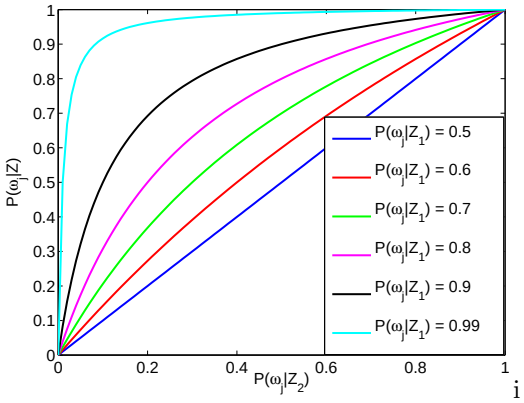


Figure 32: The combined a posteriori probability, $P(\omega_j|\mathbf{Z})$, of two classifiers for the true class ω_j is displayed.

3.4. Summary

Classification of the number of engines works best with DDE features for both ELFE and hydroacoustic signatures. Classifiers with features extracted from hydroacoustic data results in the best stand alone classification performance for both feature types. Fusion does not improve the classification ability in

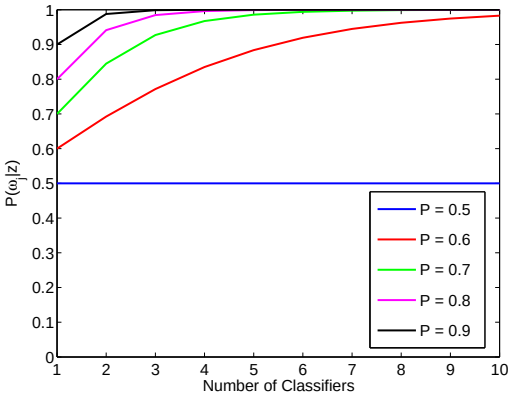


Figure 33: The combined a posteriori probability ,$P(\omega_j|\mathbf{Z})$, of up to ten classifiers (each individual classifier with equal a posteriori probabilities P) are shown

this case. For the AR method the performance is even degraded, however this is expected since the classifier working on ELFE signatures does not work at all. For the DDE method fusion leaves the classification performance unchanged (at 84%) compared to the best stand alone classifier.

Classification of ship size works best with AR features for both ELFE and hydroacoustic signatures. Classifiers with features extracted from hydroacoustic data results in the best stand alone classification performance for both feature types. Here, fusion improves the classification performance for both methods. For the AR method the performance gain achieved by fusion is minor, the correct classification rate is 98% (fused result) compare to 97 % (stand alone result). For the DDE method fusion improves the correct classification rate with 10% compared to the ELFE classifier and with 6 % compared to the hydroacoustic classifier. This result demonstrates that fusion can improve the overall classification performance on real data.

In addition computer simulations have shown that combining classifiers can improve the classification performance if the individual classifier works properly. If the classifiers have equal probability of correct classification the combined probability of correct classification increase with the number of classifiers used.

4. CONCLUSIONS

4.1. Summary of tracking results

Tracking was performed on data from a sea trial performed in 2004, where two passive underwater sensor arrays were used. One of the arrays was an acoustic uniform line array (ULA) placed on the seabed, and the other one was a long base line electrode array for underwater electric field measurements. The sea trial took place in a shallow water archipelago environment, where acoustic and electric field sources were towed from a ship, transmitting signals simulating an underwater target. Data from four such tracks have been analysed. Bearing information from the acoustic array have been fused with target positions derived from the electric field data using an inversion algorithm. The bearing- and positioning information was fused using a Kalman filter to estimate the track of the synthetic target.

Underwater acoustic beamforming techniques have been used a long time, and there is a multitude of algorithms that have been tuned for different applications. Tracking using electric field sensors is a much less mature discipline. It is evident that there is a lot more research that can be done, in order to use electric field sensors in an optimal way.

This work shows that it is possible to use existing methods for localisation of underwater targets based on electric field data together with acoustic bearing information to form a target track. The target tracking is done using an algorithm based on a Kalman filter. This technique can be useful for applications where accurate tracking of a target is needed within a limited area.

4.2. Summary of classification results

Classification was performed on acoustic and electric field data from a sea trial in 2000,

where signatures of 15 surface ships were recorded.

The classification was done using a Bayesian minimum-error-rate classifier. The ships were divided into two different classes, large or small, depending on their weight. They were also divided into two groups depending on the number of engines for the propulsion (either 2 or 4 engines).

The features used in the classification were extracted using two different models, Auto Regressive (AR), and Delay Differential Equation (DDE). The features extracted from the AR model are the 15 or 20 parameters used to represent the signature part of the measured signals. The non-linear DDE model extracts three parameters, which are also used as features for the signatures.

Classification due to weight is most successful using the AR features. The classification performance is improved for both the AR and the DDE method when fusion is used. This is not the case when dividing the ships in classes based on the number of engines. The features extracted by the DDE method performs best in this case, both for acoustic and electric field data. Fusion does not improve the classification performance in this case.

This work demonstrates that fusion can improve the classification performance on real data. However, it is important that each underlying classifier work properly. In other cases, the performance might be degraded.

4.3. Future work

Achieving the best possible performance gain from multisensor data fusion implies that one has complete control over the entire chain from sensor design to fusion algorithm. In this work the focus was kept on passive acoustic and electric field underwater surveillance, which leaves many possible approaches of data fusion unexamined. This opens up future research areas and points out several matters that can be improved on in the future.

In the case of tracking more work is needed on both acoustic bearing estimation and especially on electric field localisation of the target position. Additional methods for acoustic localisation, not only bearing algorithms, should be used. Another possibility would be to use combined matched field precessing. Also, it would be interesting to evaluate the relative benefits of using one acoustic and one electric field array, compared with two acoustic arrays. As the results in this report point out, the positioning of the sensors is very important for the tracking performance.

For classification, more work is needed primarily on extraction of features, both for the acoustic and the electric field signatures. It would be interesting to study features extracted with some other signal processing method, e.g. wavelets or FFT combined with PCA (Principal Component Analysis). A conceptually different method would be to estimate some physical parameter of the source and use that as a feature for automated classification. Here the key is to find measurable physical properties that are unique for each class. In addition if different methods for partitioning of the feature space should be examined, one promising method is SVM (Support Vector Machines). Finally, more data has to be analysed, both with the methods used in this report and with the methods mentioned above.

Since underwater surveillance is done with acoustic methods by default, a different approach would be to evaluate existing acoustic techniques, and use the data fusion algorithms to answer the question of what performance a complementary sensor would have to achieve in order to improve the overall performance of a multisensor system.

REFERENCES

[1] L. Crona et al., *An experiment on passive multi-sensor underwater surveillance*, FOI-R–1336–SE, September 2004.

[2] T. Fristedt, P. Lindqvist, *Utvärdering av sjöförsök med sensorkedja*, FOI-R–0137–SE, 2001.

[3] B. Lundqvist, *Analysis of electric field data recorded on ship-passages in 2000 and 2004*, FOI-Memo 1292, 2005.

[4] L. Abrahamsson, B-L. Andersson, *User's Guide to NLAYER 2.0*, FOA-R–01310-409–SE, 1999.

[5] P. Krylstedt, J. Mattsson, *INV_NLAYSCA*, FOA-R–00-01548-409–SE, 2000.

[6] D. Asraf et al., *Analys av data frn fltfrsk 2001*, FOI-R–0584–SE, 2002.

[7] P. Krylstedt, J. Mattsson, *Environment assessment for underwater electric sensors*, Proc. UDT Europe, Malmö, 2003.

[8] D. Berg, *Matched-field inversion in the time-domain for seabed conductivity*, FOI-R–120–SE, 2004

[9] S. Haykin (ed.), *Array Signal Processing*, Prentic-Hall, New Jersey, 1985.

[10] R. O. Schmidt, *Multiple emitter location and signal parameter estimation*, in Proc. RADC spectral Estimat. Workshop, pp. 243-258, 1979.

[11] T. J. Shah, T. Kailath, *Adaptive beamforming for coherent signals and interference*, IEEE Trans. Acoust., Speech, Signal Processing, ASSP-33, pp. 527-536, 1985.

[12] J. Rissanen, *Modeling by shortest data description*, Automatica, vol 14, pp. 465-471, 1978.

[13] E. Dudgeon, *Fundamentals of digital array processing*, Proc., of the IEEE, vol 65, no 6, pp. 898-904, 1977.

[14] J. Capon, *High-resolution frequency-wavenumber spectrum analysis*, Proc., of IEEE, vol. 57, pp. 1408-1418, 1969.

[15] J.W.C. Robinson, *Spatio-temporal weighting for high resolution direction-of-arrival estimation*, FOA-R-96-00319-2.2–SE, 1996.

[16] J. A. Cadzow, 1980, *High performance spectral estimation - A new ARMA method*, IEEE Trans., Acous., Speech, Signal Proc., vol. ASSP-28, no 5. pp. 524-529., 1980

[17] S. Blackman, R. Popoli,*Design and Analysis of Modern Tracking Systems*, Boston, MA, Artech 1999, pp. 157-160.

[18] A. S. Gelb, *Applied Optimal Estimation*, Cambridge, MA: MIT Press, 1974.

[19] Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, Orlando, FL: Academic Press, 1988.

[20] K. Murphy. *Kalman filter toolbox*, http://www.ai.mit.edu/ murphyk/Software/kalman.html

[21] J.G Proakis, D.G. Manolakis, *Digital signal processing,* Prentice Hall, 1996.

[22] J.B. Kadtke, *Classification of highly noisy signals using dynamical models,* Physics Letters A **203**, 1995.

[23] J.B. Kadtke, M. Kremliovsky, *Estimating dynamical models using generalized correlation functions,* Physics Letters A **260**, 1999.

[24] R.K. Lennartsson, *Classification with dynamical models estimated with higher order statistical moments,* FOA-R-99-01292-313–SE,1999.

[25] R.O. Duda, P.E. Hart,*Pattern classification and scene analysis*, John Wiley & Sons, 1973.

[26] D.L Hall, J. Llinas, *Handbook of multisensor data fusion*, CRC Press LLC, 2001.