Analysis of hydroacoustic time series by parametric predictive modelling

Andreas Galka

Analysis of hydroacoustic time series by parametric predictive modelling

Andreas Galka
Department GF620, Bundeswehr Technical Center for Ships and Naval Weapons, Maritime Technology and Research (WTD71), Eckernförde, Schleswig-Holstein, 24340, GERMANY; andreas.galka@gmail.com

In this paper we study hydroacoustic time series by parametric predictive modelling, using linear state space models. The aim is to detect, separate and characterise sound components emitted by complex sources such as ships, marine mammals, explosions, etc. The models are fitted to the data by Kalman filtering and maximisation of the likelihood, using various algorithms for numerical optimisation. We demonstrate that by state space modelling parametric estimates of the power spectrum can be obtained which have attractive properties, as compared to classical spectral methods. We present several examples for the application of state space modelling to actual data, including ship sounds and whale vocalisations. Special emphasis is given to the modelling of combs of ship lines, stemming e.g., from Diesel engines, or from propeller singing.

1.   INTRODUCTION

Hydroacoustic time series may contain mixtures of signal components of various origin, such as technical or natural sources. As a prominent example for technical sources we mention ships, while examples for natural sources include biological sources (e.g., marine mammals), but also earthquakes and effects related to weather. For various purposes, it may be desirable to separate such mixtures and to reconstruct and investigate the source components.

As examples for such purposes we mention the analysis of hydroacoustic signatures of ships, possibly with respect to reconnaissance, furthermore acoustic underwater communication, detection of transient events, such as earthquakes or explosions, but also research related to biological sources.

If several simultaneously recorded time series are available, e.g. obtained by using an array of hydrophones, the task of separation of components can be approached by beamforming techniques or by Blind Signal Separation (BSS) algorithms. However, frequently only a single hydrophone is employed, such that only a single scalar time series is available; in such cases, the task of separating several components represents an underdetermined inverse problem.

In this paper we propose to solve this inverse problem within a particular framework for parametric predictive modelling of time series, known as *state space modelling*.[1,2] This framework has been shown to provide solutions to a variety of time series analysis tasks arising in diverse fields such as statistics, physics and engineering, but also econometrics.

A standard element in the analysis of time series is the estimation of the power spectrum (*power spectral density*); usually this task is performed by the Discrete Fourier Transform (DFT), representing a classical non-parametric approach. Through parametric predictive modelling, an alternative approach to the estimation of power spectra becomes available, providing *parametric* estimates of the power spectrum. Compared to their non-parametric counterparts, these offer the advantage of much smoother appearance, i.e., better noise suppression. In this paper we will present examples for parametric estimates of the power spectrum and of the spectrogram of actual hydroacoustic time series.

The structure of this paper is as follows. In Section 2 we review the mathematical background of parametric predictive modelling by state space models. In Section 3 we discuss the practical fitting of these models to actual data sets. In Section 4 results for two actual data sets are presented; the paper closes with a section containing discussion and conclusions.

2.   PARAMETRIC PREDICTIVE MODELS

We denote the time series data to be modelled by $y(t)$, where $t = 1, \ldots, N$ denotes time in units of the sampling time; N denotes the length of the time series. We assume that the mean of the time series has been subtracted, such that afterwards the mean is zero. Two classes of parametric predictive models will be discussed now: Autoregressive moving-average models and state space models.

A.   AUTOREGRESSIVE MOVING-AVERAGE (ARMA) MODELS

A parametric predictive model aims at predicting the value of the recorded quantity at time t, i.e., $y(t)$, by using a set of previous values of the recorded quantity, i.e., $y(t-1), y(t-2), \ldots, y(t-p)$:

$$y(t) = f\big(y(t-1), y(t-2), \ldots, y(t-p)\big) + \eta(t) \tag{1}$$

In this model, $f(.)$ represents some (in general nonlinear) deterministic function, while $\eta(t)$ represents the prediction errors. We assume that $\eta(t)$ follows a Gaussian probability density function, with a mean of zero

and a variance of σ_η^2. In case of a linear function, Eq. (1) becomes

$$y(t) = \sum_{\tau=1}^{p} a_\tau y(t - \tau) + \eta(t) \tag{2}$$

The model of Eq. (2) is known as an *autoregressive* (AR) model; the integer parameter p denotes the model order. The a_τ, $\tau = 1, \ldots, p$, denote a set of AR model parameters, to be estimated from the data. Commonly, the prediction error term $\eta(t)$ is interpreted as an unknown noise input into the model, representing the unpredictable (and therefore stochastic) component of the data. It is then also possible to formulate a linear parametric predictive model entirely in terms of previous values of this noise input:

$$y(t) = \sum_{\tau=1}^{q} b_\tau \eta(t - \tau) + \eta(t) \tag{3}$$

The model of Eq. (3) is known as a *moving-average* (MA) model; the integer parameter q denotes the model order. The b_τ, $\tau = 1, \ldots, q$, denote a set of MA model parameters, to be estimated from the data. A generalised class of models may be defined by merging AR and MA models as follows:

$$y(t) = \sum_{\tau=1}^{p} a_\tau y(t - \tau) + \sum_{\tau=1}^{q} b_\tau \eta(t - \tau) + \eta(t) \tag{4}$$

The model of Eq. (4) is known as a *autoregressive moving-average* model with integer model orders p and q; it is abbreviated as ARMA(p, q). We will usually, without loss of generality, assume that $p = q + 1$.

We have particular interest in the ARMA$(2, 1)$ model:

$$y(t) = a_1 y(t - 1) + a_2 y(t - 2) + \eta(t) + b_1 \eta(t - 1) \tag{5}$$

This model is suitable for modelling exactly one stochastic oscillatory component.

B.   STATE SPACE MODELS

The class of (linear) state space models is defined by a pair of equations:[1,2]

$$\mathbf{x}(t) = \mathsf{A}\mathbf{x}(t - 1) + \boldsymbol{\eta}(t) \tag{6}$$

$$y(t) = \mathbf{c}\mathbf{x}(t) + \epsilon(t) \tag{7}$$

Here $\mathbf{x}(t)$ denotes a series of *state vectors*, which are a priori unknown and have to be estimated, usually by employing a *Kalman filter*; the dimension of $\mathbf{x}(t)$ shall be denoted by m. Eq. (6) represents the *dynamical equation* of the state space model; formally it corresponds to a vector autoregressive model of first order. In this case, the input noise term $\boldsymbol{\eta}(t)$ is given by a vector of dimension m. Again we assume that $\boldsymbol{\eta}(t)$ follows a Gaussian probability density function, with a mean vector of zero and a covariance matrix of Q. Eq. (7) represents the *observation equation* of the state space model; it contains a separate noise term, denoted by $\epsilon(t)$, which represents observation noise. We assume that also $\epsilon(t)$ follows a Gaussian probability density function, with a mean of zero and a variance of σ_ϵ^2. A and $\mathbf{c}$ denote the *state transition matrix* and the *observation matrix*, respectively; in the present case of scalar data, the observation matrix is actually a row vector, therefore a vector notation is used.

In contrast to the case of ARMA(p, q) models, for state space models the data prediction errors do not directly correspond to the input noise term $\boldsymbol{\eta}(t)$, and also not to the observation noise term $\epsilon(t)$.

C.   INDEPENDENT COMPONENTS STATE SPACE MODELS

The popularity of the state space modelling framework partly results from the considerable freedom of the choice of the parameter matrices which it offers to the data analyst. As an example, we may choose a block-diagonal structure for the matrices A and Q:

$$
A = \begin{pmatrix} A^{(1)} & 0 & \dots & 0 \\ 0 & A^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A^{(K)} \end{pmatrix} \qquad Q = \begin{pmatrix} Q^{(1)} & 0 & \dots & 0 \\ 0 & Q^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & Q^{(K)} \end{pmatrix} \tag{8}
$$

where the $A^{(k)}$ and $Q^{(k)}$, $k = 1, \dots, K$, denote two sets of square matrices, defined such that the dimension of each $A^{(k)}$ is equal to the dimension of the corresponding $Q^{(k)}$. Then the resulting state space model is suitable for modelling K mutually independent components, since all model parameters that could potentially generate correlations between these components, either instantaneous via Q, or delayed via A, are zero. Finally, also the observation vector has a partitioned structure:

$$
\mathbf{c} = \begin{pmatrix} \mathbf{c}^{(1)} & \mathbf{c}^{(2)} & \dots & \mathbf{c}^{(K)} \end{pmatrix} \tag{9}
$$

Via this observation vector, the data is modelled as a weighted sum (i.e., a mixing) of these K independent components.

D.   STATE SPACE REPRESENTATION OF ARMA MODELS

A given ARMA(p, q) model can be expressed as state space model, if the parameter matrices of the state space model are chosen as follows:

$$
A = \begin{pmatrix} a_1 & 1 & 0 & \dots & 0 \\ a_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{p-1} & 0 & 0 & \dots & 1 \\ a_p & 0 & 0 & \dots & 0 \end{pmatrix} \qquad\qquad Q = \begin{pmatrix} 1 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{pmatrix} \begin{pmatrix} 1 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{pmatrix}^{\mathsf{T}} \sigma_\eta^2 \tag{10}
$$

$$
\mathbf{c} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \end{pmatrix} \qquad\qquad\qquad \sigma_\epsilon^2 = 0 \tag{11}
$$

For the case of an ARMA$(2, 1)$ model this gives us

$$
A = \begin{pmatrix} a_1 & 1 \\ a_2 & 0 \end{pmatrix} \qquad\qquad\qquad Q = \begin{pmatrix} 1 & b_1 \\ b_1 & b_1^2 \end{pmatrix} \sigma_\eta^2 \tag{12}
$$

$$
\mathbf{c} = \begin{pmatrix} 1 & 0 \end{pmatrix} \qquad\qquad\qquad\qquad \sigma_\epsilon^2 = 0 \tag{13}
$$

However, it is advisable to normalise σ_η^2 to a value of 1.0 and instead to shift the corresponding parameter to the observation vector $\mathbf{c}$, which would then read $\mathbf{c} = \begin{pmatrix} c & 0 \end{pmatrix}$, with a new observation parameter c.

We remark that instead of the parametrisation of A given by Eq. (12) (known as *left companion shape*) also a parametrisation via the eigenvalues of A can be employed; these eigenvalues would either form a complex conjugated pair $\rho \exp(\pm j\phi)$ (where the phase ϕ corresponds to the frequency of the stochastic oscillatory component), or a pair of real numbers.

E.   PARAMETRIC POWER SPECTRA

By studying the z-domain transfer function, it can easily be shown that a given ARMA(p, q) model provides a parametric estimator of the power spectrum, given by

$$p_y(f) = \left| \frac{1 + \sum_{\tau=1}^{q} b_\tau \exp(-2\pi\tau jf)}{1 - \sum_{\tau=1}^{p} a_\tau \exp(-2\pi\tau jf)} \right|^2 p_\eta(f) \tag{14}$$

where $p_\eta(f)$ denotes the power spectrum of the input noise term $\eta(t)$ (usually assumed to be white) and $j = \sqrt{-1}$. For the case of an ARMA$(2, 1)$ model this gives us

$$p_y(f) = \left| \frac{1 + b_1 \exp(-2\pi jf)}{1 - a_1 \exp(-2\pi jf) - a_2 \exp(-4\pi jf)} \right|^2 p_\eta(f) \tag{15}$$

For state space models according to Eqs. (6), (7) the parametric estimator of the power spectrum is given by

$$p_y(f) = \left| \mathbf{c} \left(\exp(2\pi jf)\mathsf{I}_m - \mathsf{A} \right)^{-1} \mathsf{Q}^{1/2} \right|^2 \tag{16}$$

where the absolute-square is to be understood as a complex scalar product (i.e., requiring a conjugate transpose), I_m denotes the m-dimensional unity matrix, and $\mathsf{Q}^{1/2}$ denotes a matrix square root, provided by a Cholesky decomposition.[1] Since in case of state space models the input noise term $\boldsymbol{\eta}(t)$ cannot be estimated, the corresponding power spectral density is assumed to be white, and consequently it has been omitted from the estimator of the power spectrum.

3.   ESTIMATION OF MODEL PARAMETERS

We shall now assume that we intend to fit an independent components state space model to the given time series data, consisting of a set of K ARMA$(2, 1)$ models, i.e., the state transition matrix and the input noise covariance matrix are given by block-diagonal matrices according to Eq. (8), where the individual blocks $\mathsf{A}^{(k)}$ and $\mathsf{Q}^{(k)}$ have the structure given by Eq. (12).

Then this state space model contains several groups of parameters, namely:

- autoregressive parameters $a_\tau^{(k)}$, $\tau = 1, 2$, or alternatively $\phi^{(k)}, \rho^{(k)}$; $k = 1, \ldots, K$
- moving-average parameters $b_1^{(k)}$, $k = 1, \ldots, K$
- observation parameters $c^{(k)}$, $k = 1, \ldots, K$
- variance of observation noise σ_ϵ^2 (which in the general case does not vanish)

Also the initial value of the series of state vectors, $\mathbf{x}(0)$, and the corresponding covariance matrix may be counted to the set of model parameters to be estimated for the given data.

The estimation of model parameters is performed by a maximum-likelihood approach. For this purpose, a Kalman filter estimates the series of state vectors for given data and a given trial set of model parameters; as a by-product, data prediction errors (also known as *innovations*) $\nu(t)$ and the corresponding variances $\sigma_\nu^2(t)$ are computed. Then the logarithmic likelihood can be computed by

$$\log L = -\frac{1}{2} \sum_{t=1}^{N} \left(\log \sigma_\nu^2(t) + \frac{\nu^2(t)}{\sigma_\nu^2(t)} \right) \tag{17}$$

where a constant term has been omitted.

Maximisation of the logarithmic likelihood, as given by Eq. (17), with respect to the model parameters, has to be performed by algorithms for numerical optimisation. The optimisation scheme which we employ consists of iterative application of the following algorithms:

- expectation maximisation (EM) algorithm[2–4]
- Broyden-Fletcher-Gordon-Shanno (BFGS) quasi-Newton algorithm[5]
- Nelder-Mead simplex algorithm[5]

Starting from an initial set of model parameters, obtained by pure autoregressive modelling, a few iterations of the EM algorithm achieve fast improvement, until its rate of improvement begins to slow down; at that point the BFGS and simplex algorithms take over and achieve further improvement until convergence to a (local) maximum.

Since the specific structure of the matrices $\mathsf{A}^{(k)}$, $\mathsf{Q}^{(k)}$ and $\mathbf{c}^{(k)}$ according to Eqs. (12),(13) has to be preserved, we are dealing with a case of constrained optimisation. The introduction of constraints into the EM algorithm for state space modelling has recently been discussed by Holmes[3] and by Galka *et al.*[4]

4.   APPLICATION TO ACTUAL DATA

A.   SHIPPING SOUND

We present results for hydroacoustic recordings obtained in August 2019 at a research cruise of RV ELIS-ABETH MANN BORGESE in the Baltic Sea, close to the port of Warnemünde, Germany. The recording equipment consisted of an AMAR buoy (JASCO Applied Sciences), deployed at a fixed position; the sampling rate is 64 kHz, but for actual data analysis we subsample the data to 2 kHz. The ship passed the buoy at straight course with a speed of 5.6 knots at a minimum distance of approx. 80 metres.

i.   Analysis of a short time series segment

A time series segment of 20 sec. length, around the time point of minimum distance (closest point of approach, CPA), is chosen for further analysis. An independent components state space model consisting of $K = 34$ ARMA$(2,1)$ models is fitted to the data. The value of K is chosen by gradually increasing the number of ARMA$(2,1)$ models, until the innovations appear approximately white.

The estimated state components resulting from the optimised model are shown in Fig. 1, as estimated by a Kalman filter and a subsequent Rauch-Tung-Striebel (RTS) smoother;[1] all components have complex conjugated pairs of eigenvalues, such that they may be labelled by their frequencies (shown on the left side of the figure); however, 4 components which have frequencies larger than 500 Hz, have been omitted from the figure.

It can be seen that some components have almost deterministic sinusoidal shape, while others appear more stochastic. A set of components form a line comb with frequencies

$$f_\kappa = \kappa \times 14.96 \text{ Hz}, \quad \kappa = 1, \dots, 12$$

In the figure, the frequencies of these 12 components are emphasised by red colour. The origin of this line comb is the Diesel engine of the ship, with 15 Hz being a typical ignition rate of individual cylinders. State space modelling allows us to identify and extract this line comb component which forms an important part of the ship's hydroacoustic signature.

Another prominent part of the ship's hydroacoustic signature is given by the component at 435 Hz (bottommost component in Fig. 1); its probable origin is *propeller singing*. This component will be analysed more closely below in Sec. 4.A.iii.
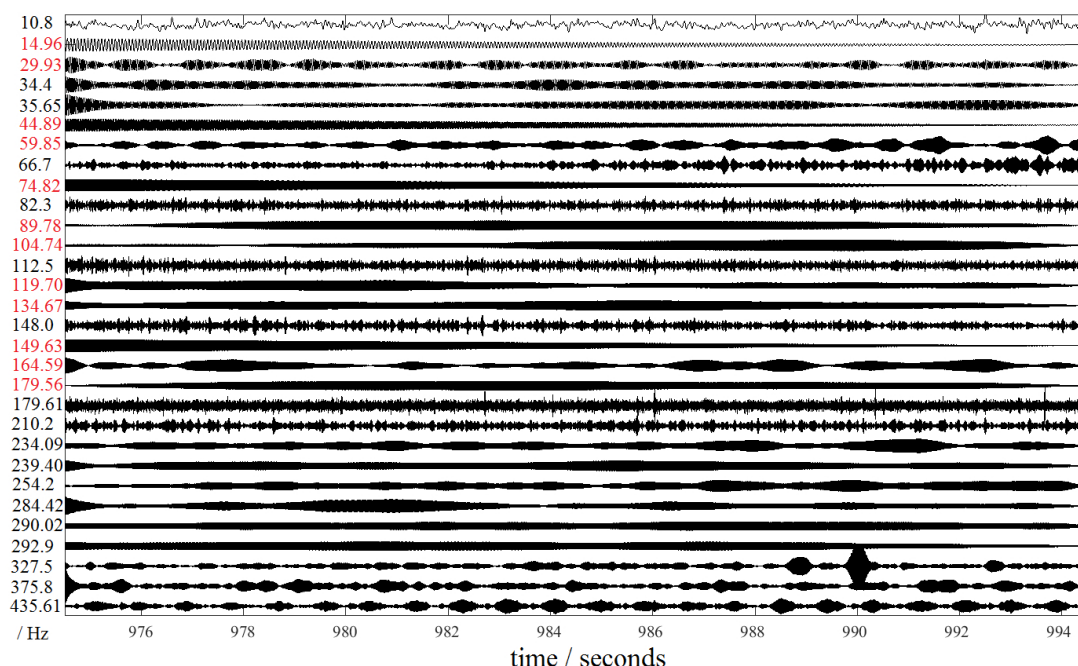
Figure 1: Estimated state components for a time series segment with ship lines. Each component is labelled by its frequency (left vertical axis); components belonging to the line comb of the Diesel engine are emphasised by frequency values in red colour. All components are normalised to unit variance.

The results of analysing the power spectrum of the chosen time series segment with ship lines are shown in Fig. 2. In the upper left panel of the figure the non-parametric estimates of the power spectra of the raw data (blue) and of the innovations after state space modelling (purple) are shown. It can be seen that the power spectrum of the raw data contains various lines, the strongest of which is the propeller-singing line at 435 Hz; on the other hand, the power spectrum of the innovations appears essentially white, indicating successfull predictive modelling. It can also be seen that non-parametric estimates of power spectra are inherently noisy, which is a typical consequence of *overfitting*.

In the upper right panel of the figure the parametric estimate of the power spectrum of the data according to Eq. (16) is shown (blue). By comparison with the non-parametric estimate shown in the left upper panel, it can be seen that the parametric estimate is much smoother.

By using Eq. (15) it is also possible to display the contribution of individual components to a parametric power spectrum; this is illustrated in the two lower panels of Fig. 2. In the lower left panel of the figure the contributions of the 12 components belonging to the Diesel engine line comb are shown (red), in addition to the complete power spectrum (blue). It can be seen that some components are very weak, such that they can be detected only as part of a line comb.

In the lower right panel of the figure the contributions of the remaining 22 components are shown (red), in addition to the complete power spectrum (blue).

ii. Moving-window analysis of a long time series segment

If longer time series segments are to be analysed, one has to expect gradual changes of the power spectrum, due to non-stationarity. For such situations the power spectrum is commonly estimated on a moving window, providing the *spectrogram*. We suggest to compute spectrograms by parametric predictive modelling. In order to demonstrate this approach, we choose a time series segment of 30 minutes length from the same
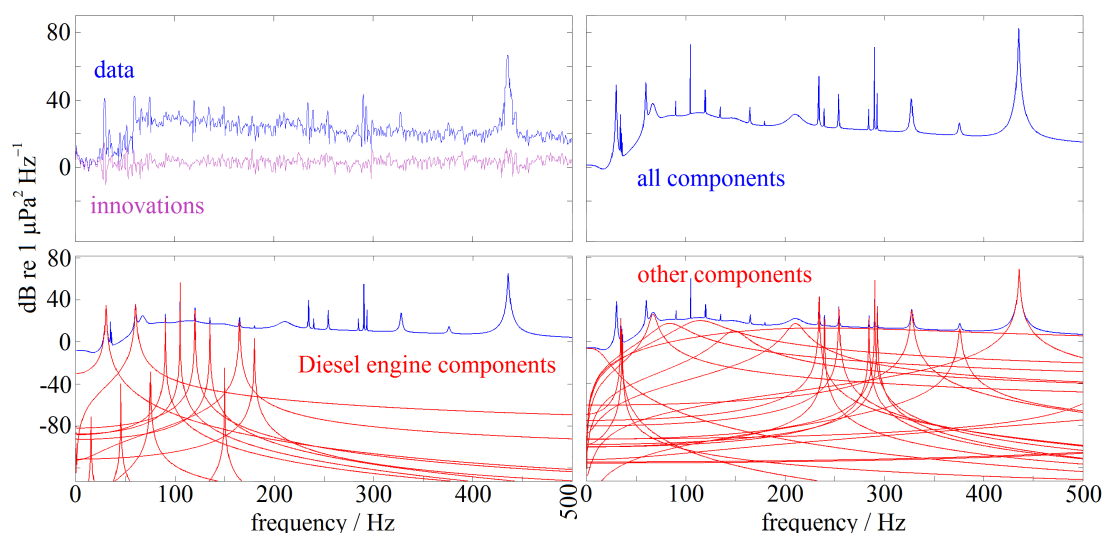
Figure 2: Analysis of the power spectrum of a time series segment with ship lines by non-parametric (upper left panel) and parametric (remaining 3 panels) methods; for details see text.

data set as before, such that the closest point of approach (CPA) falls approx. into the middle of the segment.

In Fig. 3 the spectrogram of the chosen segment is shown, based on non-parametric estimation of the power spectrum of a moving window. A number of ship lines is visible, such as the prominent line at 435 Hz, but also some lines of the Diesel engine line comb, furthermore broad-band components such as the noise between 50 and 100 Hz which onsets after the CPA.

In order to compute the spectrogram of the chosen time series segment by parametric predictive modelling, we divide the segment into 600 windows of 6 sec. length, such that consecutive windows have 3 sec. overlap. An independent components state space model consisting of $K = 16$ ARMA$(2, 1)$ models is fitted to each window. In order to speed up the maximum-likelihood step, for each window the final model of the previous window is employed as initial model. Finally, for all windows the parametric estimate of the power spectra according to Eq. (16) are computed and graphically presented as a spectrogram; the result is shown in Fig. 4.

By comparison of Figs. 3 and 4 it can be seen that again the parametric estimate of the spectrogram is much smoother than the non-parametric estimate; as a consequence, ship lines stand out clearly against the broadband background. However, only lines which have successfully been picked up as independent components during state space modelling will be represented in the spectrogram.

iii.   Analysis of the line at 435 Hz

Finally we report results on an additional analysis of the propeller-singing line at 435 Hz, which in the given data forms the strongest component of the ship's hydroacoustic signature. In Fig. 5 (upper panel) a close-up of the non-parametric power spectrum of the time series with ship lines is shown, limited to the frequency interval between 425 Hz and 445 Hz: it can be seen that the line at 435 Hz has itself a substructure, consisting of a comb of equidistant lines. This comb has presumably its origin in the rotation of the propeller which causes amplitude modulation.

We choose a time series segment of 10 sec. length around the time point of the CPA, subsample it to 8 kHz and bandpass-filter it, such that all components outside the frequency interval between 420 Hz and 450 Hz are removed, then we fit an independent components state space model consisting of $K = 13$ ARMA$(2, 1)$ models to this segment. From these 13 components, 11 components are intended for the actual
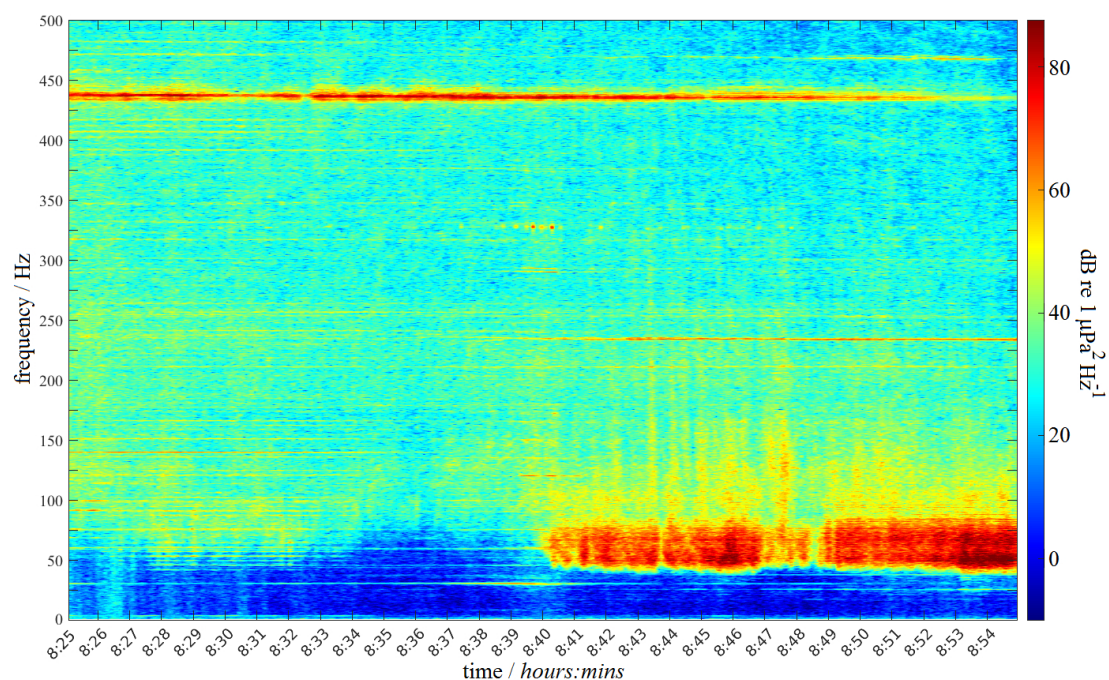
Figure 3: Spectrogram of a time series segment with ship lines, based on non-parametric estimation of the power spectrum of a moving window.
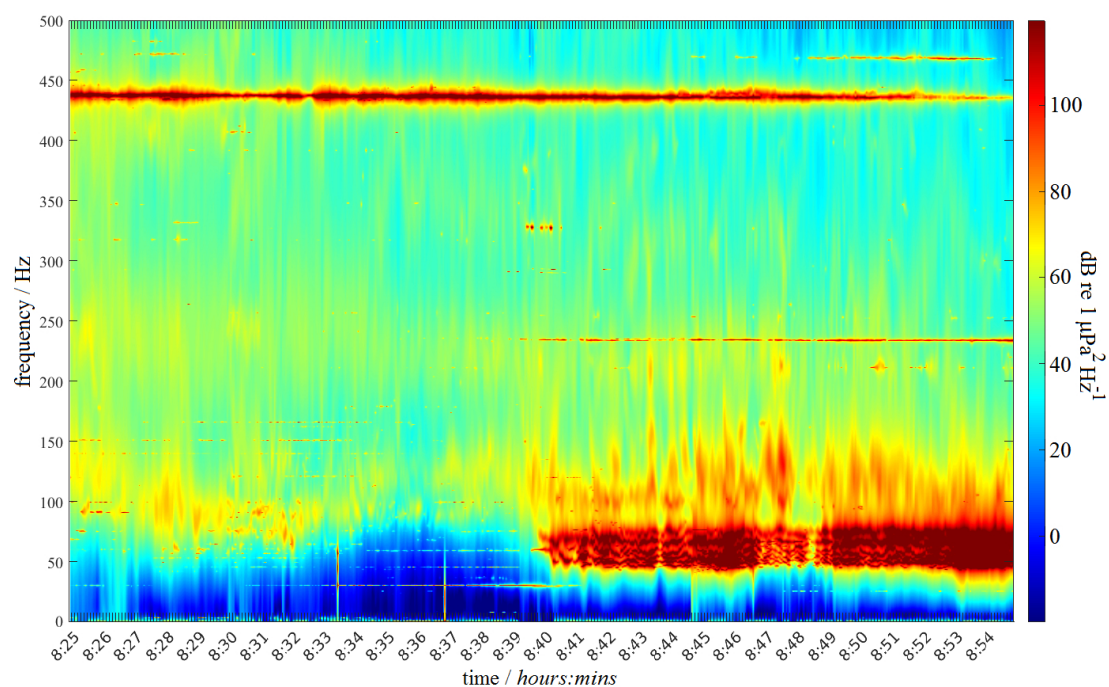

Figure 4: Spectrogram of a time series segment with ship lines, based on parametric estimation of the power spectrum of a moving window.
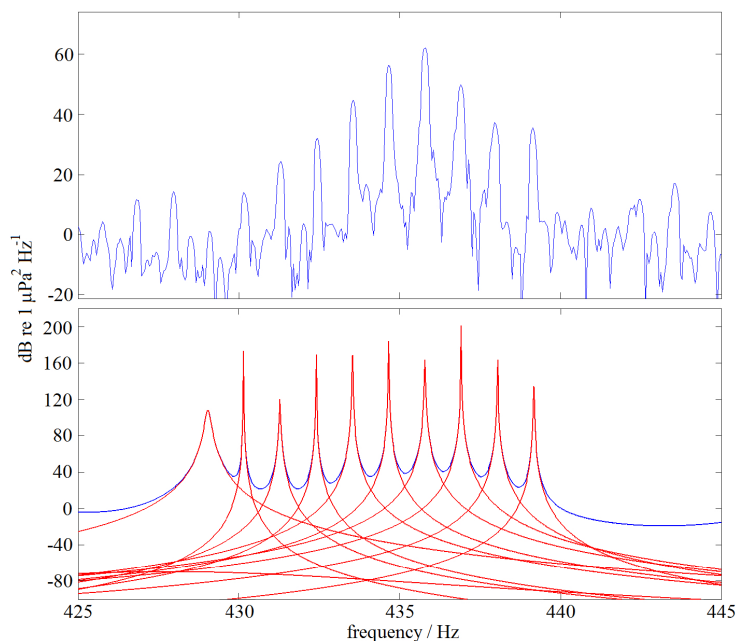
Figure 5: Analysis of the power spectrum of a time series segment with ship lines by non-parametric (upper panel) and parametric (lower panels) methods; the time series segment was bandpass-filtered, such that all components outside the frequency interval between 420 Hz and 450 Hz are removed; for details see text.

lines, while the remaining two components are intended for the background.

In practical model fitting, we choose to optimise frequencies $\phi^{(k)}$ and moduli $\rho^{(k)}$, instead of the autoregressive parameters $a_1^{(k)}$ and $a_2^{(k)}$, $k = 1, \ldots, 11$. Since it is known that the lines within the comb are equidistant, it suffices to optimise one of the frequencies – chosen as the frequency of the strongest line – and the frequency difference between neighbouring lines (i.e., the line spacing) instead of all individual frequencies; by this approach the number of parameters to be optimised is reduced by 9.

The maximum-likelihood estimate of the line spacing is found to be 1.126 Hz, corresponding to a propeller rotation rate of 67.56 rotations per minute, which is a realistic value. This example illustrates that by state space modelling more information than just line frequencies can be extracted from a ship's hydroacoustic signature.

In the lower panel of Fig. 5 the corresponding parametric estimate of the power spectrum (blue) and the contributions of the individual lines (red) are shown. It can be seen that the comb of lines is well modelled, although height and shape of the individual lines do somewhat differ from the non-parametric power spectrum. Closer inspection reveals that one of the ARMA$(2, 1)$ models has failed to "catch" its line, such that only 10 lines are visible.

B.   MARINE MAMMALS

As a second example for the application of state space modelling to actual data we choose a hydroacoustic recording which was obtained in the Atlantic Ocean, in the presence of a blue whale (*Balaenoptera musculus*). The data set has been published by Forst *et al.*;[6] no further details on recording equipment and precise position and time of the recording are available. The length of the available time series segment is 206 sec., the sampling rate is, after subsampling, 220.5 Hz. During the recording low-pass filtering was employed, such that signal components beyond approx. 90 Hz have been suppressed.

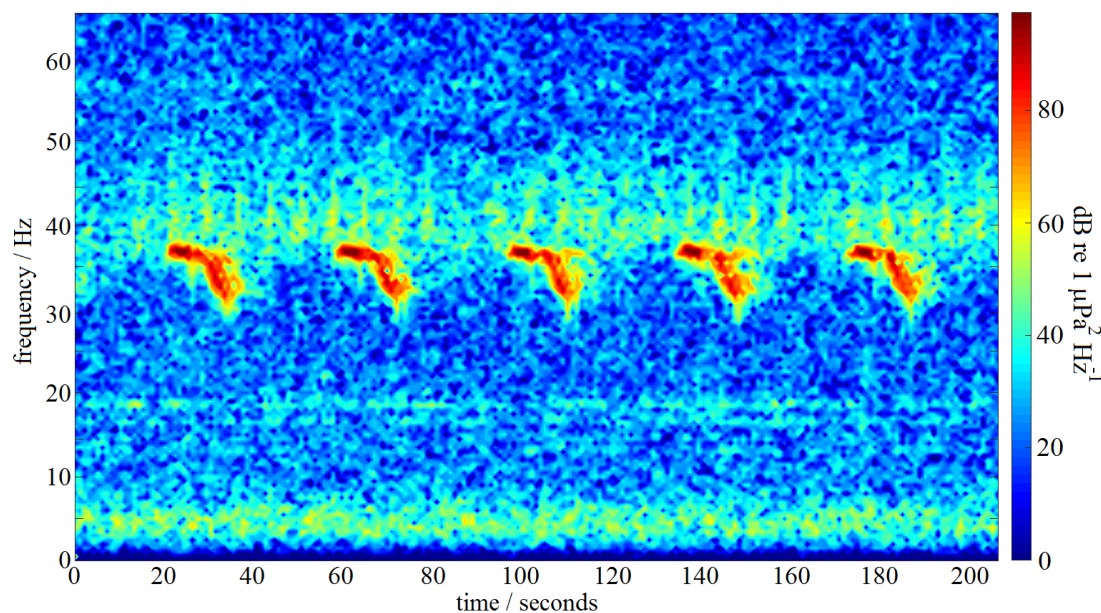In Fig. 6 the spectrogram of the time series is shown, based on non-parametric estimation of the power

Figure 6: Spectrogram of a time series with vocalisations of a whale, based on non-parametric estimation of the power spectrum of a moving window.

spectrum of a moving window. It can be seen that the whale has produced 5 vocalisations within the frequency interval between 30 Hz and 40 Hz. Also some weak lines around 20 Hz are visible, presumably from shipping.

An independent components state space model consisting of $K = 18$ ARMA$(2, 1)$ models is fitted to the first 56 sec. of the time series; after model fitting, the optimised model is applied to the complete time series of 206 sec. length. Since the time series is sufficiently stationary, the Kalman filter performs well out-of-sample. This approach is chosen in order to reduce the computational time expense of the optimisation step.

The estimated state components resulting from the optimised model are shown in Fig. 7, as estimated by a Kalman filter and a subsequent RTS smoother;[1] again, all components have complex conjugated pairs of eigenvalues, such that they may be labelled by their frequencies (shown on the left side of the figure).

It can be seen that three of the components represent well the vocalisations of the whale, with frequencies 32.80 Hz, 33.70 Hz and 37.29 Hz, while the remaining 15 components either represent ship lines or background noise. We emphasise that this successful separation of the whale components has been achieved automatically by the model fitting procedure, without the need for any prior information or user interaction.

The results of analysing the power spectrum of the time series with whale vocalisations are shown in Fig. 8. In the upper left panel of the figure the non-parametric estimates of the raw data (blue) and of the innovations after state space modelling (purple) are shown. Again, it can be seen that in the power spectrum of the raw data the vocalisations of the whale are represented by a strong peak between 30 Hz and 40 Hz; furthermore there are a few lines around 20 Hz. The power spectrum of the innovations appears essentially white, for frequencies below 90 Hz, while for higher frequencies the effect of low-pass filtering is visible.

In the upper right panel of the figure the parametric estimate of the power spectrum of the data according to Eq. (16) is shown (blue). Again the parametric estimate is much smoother than the non-parametric estimate. In the lower left panel of the figure the contributions of the three components belonging to the vocalisations of the whale are shown (red), in addition to the complete power spectrum (blue). In the lower right panel of the figure the contributions of the remaining 15 components are shown (red), in addition to
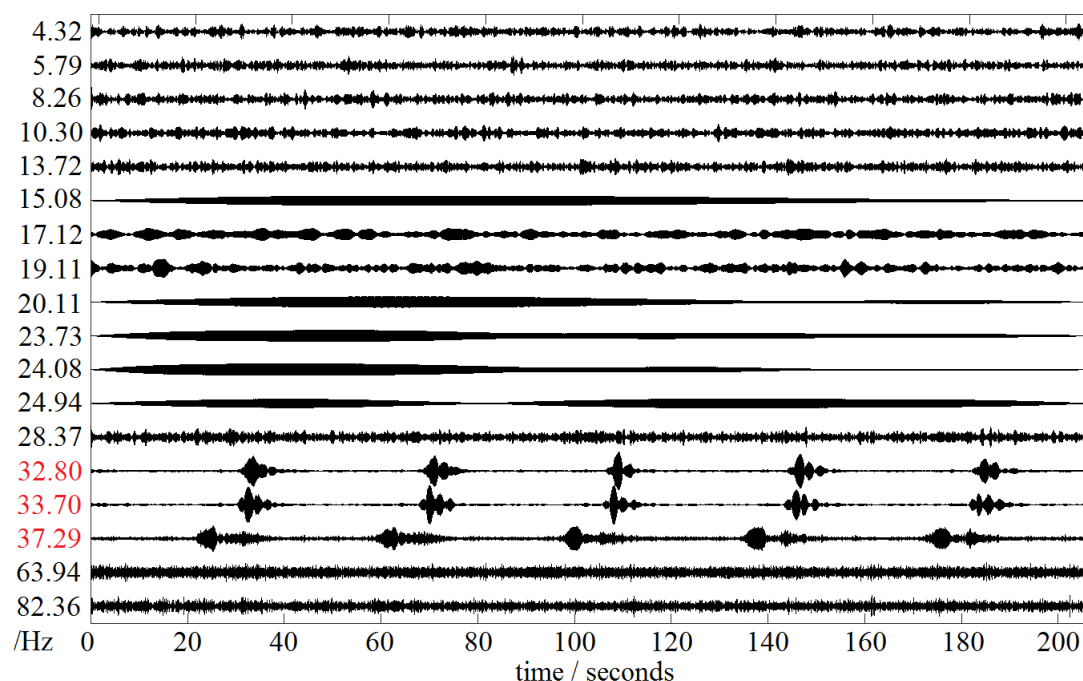
Figure 7: Estimated state components for a time series with whale vocalisations. Each component is labelled by its frequency; components belonging to the vocalisations of the whale are emphasised by frequency values in red colour. All components are normalised to unit variance.

the complete power spectrum (blue); as mentioned before, they either represent narrow-band ship lines or broad-band background noise. Some of the narrow-band components are very weak and may in fact be spurious.

Due to the shortness of this time series, we refrain from computing a spectrogram based on parametric modelling.

5. CONCLUSION

In this paper it has been shown that signal components may be extracted from hydroacoustic time series by independent components state space modelling, and that the power spectra of both the original time series and its individual components can be estimated from the parameters of the resulting models. The smooth appearance of these power spectra is a result of *parsimonious modelling*, in contrast to the overfitting which is inherent in the Discrete Fourier transform.

The task of reconstructing several unknown components from a single mixture represents an example of an underdetermined inverse problem; as demonstrated, such task can be accomplished by the combination of state space modelling, Kalman filtering and maximum-likelihood estimation. In this paper, all components were modelled by $ARMA(2,1)$ models; different models could have been employed, such as $AR(1)$, $ARMA(p, p-1)$ with $p > 2$, or even nonlinear models. However, it has been shown in earlier work that the Kalman filter is very flexible with respect to misspecified models, such that even distinctly nonlinear components could be extracted by using $ARMA(2,1)$ models.[7]

A weak point of the proposed method is the high computational time expense required for performing the required numerical optimisation. Currently this kind of analysis can only be carried out retrospectively, while real-time application is not yet feasible; with our current implementation, modelling of 6 sec. of time
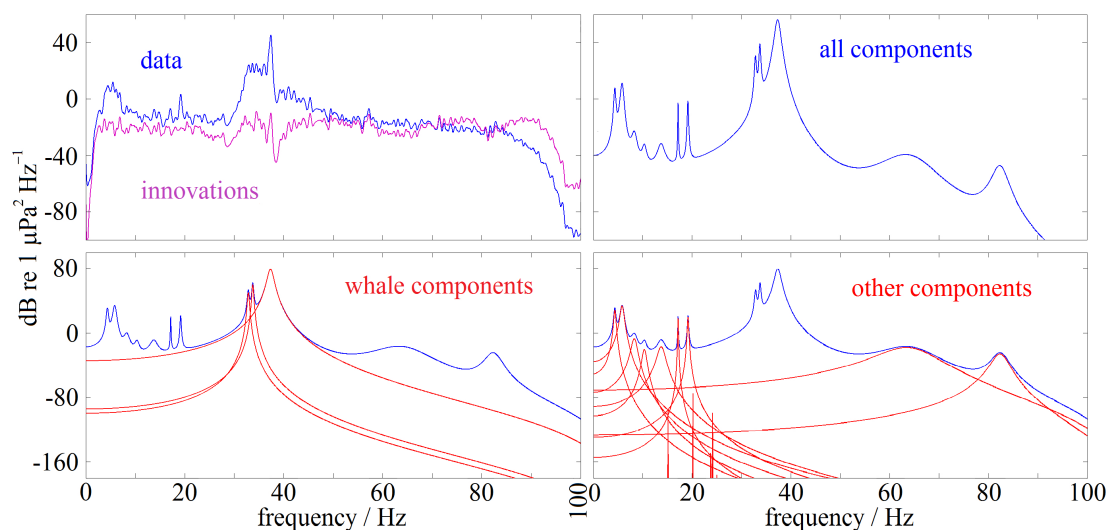
Figure 8: Analysis of the power spectrum of a time series segment with whale vocalisations by non-parametric (upper left panel) and parametric (remaining 3 panels) methods; for details see text.

series data consumes 1-2 hours of computation time. However, if only the task of computing parametric estimates of the power spectrum is retained, while the task of estimation of independent components is dropped, real-time estimation would approximately become possible; to this end full state space modelling would have to be replaced by modelling by pure ARMA modelling.

ACKNOWLEDGMENTS

REFERENCES

[1] T. Kailath, A. Sayed & B. Hassibi, *Linear estimation*, Prentice Hall (2000).

[2] R. H. Shumway & D. S. Stoffer, *Time Series Analysis and its Applications*, 4th edition, Springer (2017).

[3] E. E. Holmes, "Derivation of an EM algorithm for constrained and unconstrained multivariate autoregressive state-space (MARSS) models", *arXiv*, 1302.3919 (2013).

[4] A. Galka, S. Moontaha & M. Siniatchkin, "Constrained expectation maximisation algorithm for estimating ARMA models in state space representation", *EURASIP Journal on Advances in Signal Processing*, **23** (2020).

[5] J. Nocedal & S. Wright, *Numerical Optimization*, 2nd edition, Springer (2006).

[6] C. Forst, L. Ginzkey & P. Wille, *Schall im Meer* (audio compact disc), Forschungsanstalt der Bundeswehr für Wasserschall und Geophysik (FWG), Kiel, Germany (2000).

[7] A. Galka, K. F. K. Wong, U. Stephani, T. Ozaki & M. Siniatchkin (2013), "Blind signal separation of mixtures of chaotic processes: a comparison between Independent Component Analysis and State Space Modeling", *International Journal of Bifurcation and Chaos* **23**, 1350165 (2013).