Analysis of hydroacoustic time series by state space modelling

Andreas Galka[1]

[1]Bundeswehr Technical Centre for Ships and Naval Weapons, Maritime Technology and Research (WTD 71)
Klausdorfer Weg 2-24, 24148 Kiel, Germany
E-Mail: andreasgalka@bundeswehr.org

Abstract: *Hydroacoustic time series, recorded in the ocean, may contain signal components originating from natural or artificial sources. The aim of the present paper is to detect, separate and characterise such signal components, with particular emphasis on investigating hydroacoustic signatures of ships. We approach this aim by employing a class of parametric models from time series analysis, known as state space models. The models are fitted to the data by Kalman filtering and maximisation of the likelihood, using suitable algorithms for numerical optimisation. The application of Kalman filtering to such situations corresponds to solving an inverse problem. We demonstrate that by state space modelling parametric estimates of the power spectra of the signal components can be obtained, which have attractive properties, as compared to classical non-parametric methods, furthermore filtering and noise reduction can be accomplished. We also discuss how the framework of state space modelling can be generalised to deal with phenomena which typically arise in the investigation of ship signatures, such as combs of lines and Doppler effects.*

Keywords: *time series analysis, state space modelling, Kalman filtering*

1. INTRODUCTION

Underwater sound can be recorded by suitable sensors, such as hydrophones; the resulting recordings represent *hydroacoustic time series*. By analysing such time series, information regarding the sources of the underlying underwater sound can be obtained. Sources may be artificial, such as ships or explosions, or natural, such as marine mammals or earthquakes. An important part of the analysis of hydroacoustic time series is given by the task of separating mixtures of different sources, or of extracting individual components from such mixtures. If no prior knowledge on the properties both of the sources and of the mixing process is available, such task falls into the realm of *Blind Signal Separation*.

Time series analysis forms a well developed branch of statistical data analysis [1, 2]. Within science and engineering, time series are routinely recorded and analysed in many fields; notable examples are climate research, meteorology, neuroscience and astrophysics, but also econometrics and financial mathematics. Acoustics, in particular underwater acoustics, represents another example. However, in our opinion, the full power of the available mathematical framework for time series analysis has not yet been mustered for the task of analysing hydroacoustic time series.

As an example we mention the possibility to choose representations for a given time series either in frequency domain, or in time domain. While frequency domain representations are well established in underwater acoustics, this is not the case for time domain representations, such as autoregressive modelling. The present paper intends to illustrate the potential which time domain representations have to offer for the analysis of hydroacoustic time series.

The structure of this paper is as follows. In Sections 2 and 3 the mathematical background of parametric predictive modelling by state space models is reviewed. In Section 4 the practical fitting of these models to actual data sets is discussed. In Section 5 parametric estimation of power spectra is discussed, and an example for real data is presented. In Sections 6 and 7 real-data examples are presented for modelling combs of lines and Doppler effects. Finally, in the last Section the results of the paper are briefly summarised.

2. REPRESENTATIONS OF TIME SERIES

We begin by assuming that a time series is given, denoted by $\mathbf{y}_t$, where $t = 0, \ldots, (N-1)$ denotes the time points at which the recordings were taken; time is measured in units of $1/f_S$, where f_S denotes the sampling frequency. N denotes the length of the time series, and n denotes the dimension of the data vectors $\mathbf{y}_t$.

A typical non-parametric frequency-domain representation of $\mathbf{y}_t$ would be given by an (inverse) discrete Fourier transform, according to

$$\mathbf{y}_t = \frac{1}{N} \sum_{f=0}^{N-1} \tilde{\mathbf{y}}_f \exp\left(-2\pi j \frac{ft}{N}\right) \tag{1}$$

where the series of complex numbers $\tilde{\mathbf{y}}_f$, with frequency variable $f = 0, \ldots, (N-1)$, is obtained from the time-domain data by discrete Fourier transform. By Eq. (1) the time-domain data $\mathbf{y}_t$ is expanded into complex exponentials with complex weights, such that the number of expansion coefficients $\tilde{\mathbf{y}}_f$ is equal to the number N of time-domain data points $\mathbf{y}_t$. From a statistical point of view, such choice of representation represents an obvious case of *overfitting*.

As an alternative, we will now discuss the typical setting for *predictive parametric modelling* in time domain; in this setting the time series data is represented according to

$$\mathbf{y}_t = \boldsymbol{\phi}(\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p}; \boldsymbol{\theta}) + \boldsymbol{\nu}_t \tag{2}$$

where $\boldsymbol{\phi}(\ldots)$ denotes a prediction function, aiming at modelling the data vector at time t by a deterministic function of a set of previous data vectors, up to p steps into the past, while $\boldsymbol{\nu}_t$ denotes a series of residual prediction errors, also known as *innovations*. The prediction function $\boldsymbol{\phi}(\ldots)$ depends on a set of internal parameters, collected in a parameter vector $\boldsymbol{\theta}$. In order to avoid overfitting, the dimension of $\boldsymbol{\theta}$ is kept much lower than the number of data points N; if this condition is fulfilled, we may speak of *parsimonious modelling*.

3. LINEAR PREDICTIVE PARAMETRIC MODELLING

While in general the prediction function $\boldsymbol{\phi}(\ldots)$ may be nonlinear w.r.t. $\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots, \mathbf{y}_{t-p}$, in this paper we constrain ourselves to linear prediction functions. As a further constraint, we shall only work with scalar data, i.e., $n = 1$. Scalar hydroacoustic time series may result from employing a single hydrophone, or from applying a beamformer to the data recorded by an array of hydrophones.

A well established class of linear predictive parametric models is known as *autoregressive moving-average* (ARMA) models, given by

$$y_t = \sum_{k=1}^{p} a_k y_{t-k} + \eta_t + \sum_{k=1}^{q} b_k \eta_{t-k} \tag{3}$$

where η_t denotes a noise term, representing the prediction errors and assumed to follow a gaussian distribution with zero mean and variance σ_η^2, and a_k, $k = 1, \ldots, p$, and b_k, $k = 1, \ldots, q$, denote two sets of model parameters. On the right-hand-side of Eq. (3) the first sum represents the deterministic prediction function, while the remaining terms represent the stochastic part of the model. In principle, the weighted sum over past values of η_t, known as the *moving-average* term, can also be interpreted as part of the prediction function.

The integer numbers p and q represent the autoregressive and moving-average model orders, respectively, and the model is denoted as ARMA(p, q) model. We often choose $p = 2$, $q = 1$; by this choice, the model is suitable for describing one stochastic oscillation. Sometimes also higher values may be chosen for the model orders; below we will briefly discuss an example for such choice. Let the frequency of a stochastic oscillation be denoted by φ, then the autoregressive parameters of the corresponding ARMA(2,1) model are given by

$$a_1 = 2\rho \cos\varphi \qquad ; \qquad a_2 = -\rho^2 \tag{4}$$

where $0 < \rho \leq 1$ is a damping parameter; the moving-average parameter b_1 has only little effect on the peak frequency of the model. As a consequence of these relationships, the model parameters a_1 and a_2 can be replaced by φ and ρ.

As a generalisation of ARMA models, we will now discuss *state space models* [2]. A linear state space (LSS) model consists of two equations, according to

$$\mathbf{x}_t = \mathsf{A}\,\mathbf{x}_{t-1} + \boldsymbol{\eta}_t \tag{5}$$

$$y_t = \mathsf{C}\,\mathbf{x}_t + \epsilon_t \tag{6}$$

where $\boldsymbol{\eta}_t$ and ϵ_t denotes two noise terms, representing *process noise* and *observation noise*, respectively. Both noise terms are assumed to follow gaussian distributions with zero means and covariances $\text{cov}(\boldsymbol{\eta}_t) = \mathsf{Q}$ and $\text{cov}(\epsilon_t) = \mathsf{R}$. For scalar data, R will consist of a single variance parameter. The time-dependent vector $\mathbf{x}_t$ denotes the *state vector* which represents the dynamical state of the system under consideration at time t. The dimension of $\mathbf{x}_t$ shall be denoted by m. A and C denote two parameter matrices, to be called *state transition matrix* and *observation matrix*, respectively. The parameter vector $\boldsymbol{\theta}$ then comprises all elements of the four parameter matrices A, C, Q and R. Usually the model parameters are assumed to be constant, however, in special situations, such as the presence of Doppler effects (see Sec. 7), some parameters may be modelled as varying with time.

Unlike with ARMA models, in LSS models we have separate terms for process noise and observation noise; this distinction opens up the possibility to suppress observation noise.

We also mention that ARMA models can be rewritten as LSS models; in particular, it is possible to merge several mutually independent ARMA models into a single LSS model. The resulting model class is suitable for approaching the Blind Signal Separation problem [3].

4. ESTIMATION OF STATES AND PARAMETERS

In order to model a given time series by a LSS model, the model given by Eqs. (5) and (6) needs to be inverted, i.e., the sequence of state vectors $\mathbf{x}_t$ has to be estimated from the data y_t; this task can be accomplished by a *Kalman filter*. Furthermore, the parameter matrices A, C, Q and R have to be estimated; this task is approached by a numerical maximum-likelihood approach, where the (logarithmic) likelihood is computed by the Kalman filter according to

$$\mathcal{L}(\mathsf{A}, \mathsf{C}, \mathsf{Q}, \mathsf{R}) = -\frac{1}{2}\sum_{t=1}^{N}\left(\log \sigma_t^2 + \frac{\nu_t^2}{\sigma_t^2}\right) - \frac{N}{2}\log(2\pi) \tag{7}$$

Here, ν_t denotes the sequence of the innovations, and σ_t^2 denotes the corresponding variances of the innovations; both are computed by the Kalman filter.

We perform maximisation of the (logarithmic) likelihood by a combination of a Newton-Raphson algorithm (Broyden-Fletcher-Gordon-Shanno), a simplex algorithm (Nelder-Mead) and a variant of the expectation maximisation algorithm. The estimation of the parameters of ARMA models that are not embedded within LSS models, can also be performed by a numerical maximisation approach.

5. PARAMETRIC ESTIMATORS OF THE POWER SPECTRUM

Predictive parametric models describe the temporal correlation structure of the given data, therefore they contain information that can be employed for estimating the power spectrum of the data, thereby providing *parametric* estimators.

The parametric estimator resulting from the ARMA model of Eq. (3) is given by

$$P_f = \left| \frac{1 + \sum_{k=1}^{q} b_k \exp(-2\pi k j f)}{1 - \sum_{k=1}^{p} a_k \exp(-2\pi k j f)} \right|^2 \sigma_\eta^2 \tag{8}$$

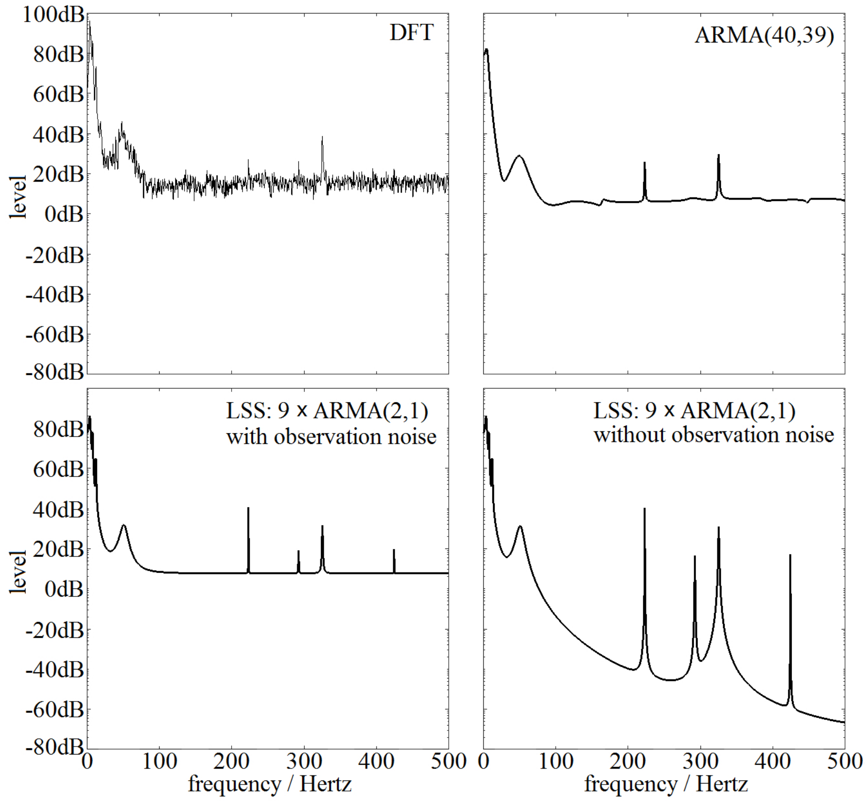

Figure 1: Non-parametric estimate (upper left panel) and three parametric estimates (remaining three panels) of the power spectrum of a hydroacoustic time series segment of 30 seconds length

The parametric estimator resulting from the state space model of Eqs. (5) and (6) is given by

$$P_f = \left| \mathsf{C}\big(\exp(2\pi j f)\mathsf{I}_m - \mathsf{A}\big)^{-1}\mathsf{Q}^{1/2} + \mathsf{R}^{1/2} \right|^2 \tag{9}$$

where I_m denotes the m-dimensional unity matrix, and $\mathsf{Q}^{1/2}$ denotes a matrix square root of the covariance matrix Q, i.e., an upper triangular matrix fulfilling $\mathsf{Q} = \left[\mathsf{Q}^{1/2}\right]^{\mathsf{T}}\mathsf{Q}^{1/2}$; for the case of scalar data, $\mathsf{R}^{1/2}$ denotes an ordinary square root. By omitting the term $\mathsf{R}^{1/2}$ from Eq. (9), we can perform suppression of observation noise.

As an example for the estimation of power spectra for a real hydroacoustic time series, we analyse a time series which has been recorded by a recording buoy, deployed at a fixed position. We choose a segment of 30 seconds length with effective sampling rate $f_s = 2$ kHz, corresponding to a length of $N = 60000$. We estimate the power spectrum of the chosen segment by discrete Fourier transform (DFT), by an ARMA(40,39) model, using Eq. (8), and by an LSS model, consisting of 9 ARMA(2,1) models, using Eq. (9); for the LSS model, we either keep or omit the term corresponding to observation noise. The results are shown in Fig. 1.

Within the chosen time series segment, several sharp lines with peak frequencies between 200 Hz and 500 Hz are present, resulting from shipping; some of these lines are visible in the estimated power spectra. In Fig. 1 it can be seen that the DFT estimate is fairly noisy, while the ARMA and LSS estimates are much smoother; this result illustrates the effect of avoiding overfitting. Due to the noisy estimate, the number of lines detected by the DFT estimate is not well defined; at least three are visible. The ARMA estimate detects two lines, despite its high model order of $p = 40$, $q = 39$; the LSS estimate detects four lines, corresponding to 4 of the 9

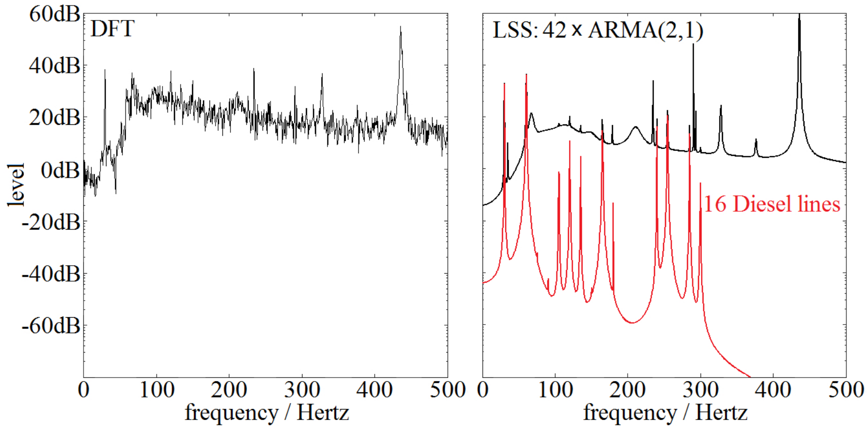

Figure 2: Non-parametric estimate (left panel) and parametric estimate (right panel; black curve) of the power spectrum of a hydroacoustic time series segment of 20 seconds length; in the right panel also the power spectrum of those components representing the comb of lines of the Diesel engine is shown (red curve).

ARMA(2,1) models. The remaining 5 ARMA(2,1) models describe contributions to the power spectrum at frequencies below 100 Hz. The lines are embedded within broad background power representing the observation noise; in the lower right panel of Fig. 1 this observation noise has been suppressed, resulting in a much clearer picture of the four lines that were detected by the LSS estimate.

In addition to providing a parametric estimate of the power spectrum, the LSS model also yields a time-domain decomposition of the given time series into source components, thereby offering a solution to the Blind Signal Separation problem [3].

6. COMBS OF LINES AS PARTS OF SHIP SIGNATURES

The hydroacoustic signature of ships usually consists of lines which originate from the technical systems within the ship, or from oscillating structures, like the propeller. Further characterisation of the ship and its state, beyond the values of the peak frequencies of the lines, becomes possible by detecting relationships between the lines, e.g., in the guise of several lines forming a comb of lines.

As an example we show in Fig. 2 the result of analysing a time series recorded by a deployed buoy in close distance to a moving ship (research vessel ELISABETH MANN BORGESE); effective sampling rate is again $f_s = 2$ kHz. Parametric modelling is performed by a LSS model, consisting of 42 ARMA(2,1) models. In the figure, it can be seen that both the non-parametric and the parametric estimate of the power spectrum show numerous lines at frequencies up to 500 Hz. The peak frequencies of at least 16 of these lines are multiples of a value of approx. 14.96 Hz, thereby forming a comb of lines, which can be attributed to the Diesel engine of the ship.

The presence of a comb of lines considerably simplifies the parameter estimation step, since instead of 16 individual frequencies only the fundamental frequency has to be estimated; furthermore the estimate will have higher precision, if it is based on 16 lines. The same approach can be applied to cases where a line is split up into a series of equidistant lines, as it may occur for lines resulting from propeller singing; in this case there will typically be no fundamental frequency, but a frequency increment, which becomes part of the vector of model parameters.

Since the LSS model is built from independent components, modelled by ARMA(2,1) models, it is possible to plot a parametric estimate of the power spectrum only for a subset of components. In the right panel of Fig. 2 the estimate of the power spectrum only of the 16 Diesel engine lines is shown (red curve); it can be seen that some higher harmonics within the comb of lines are weaker than the overall spectral background of the time series, while others are completely absent.

7. MODELLING OF DOPPLER EFFECTS

We now discuss an effect occurring in the analysis of ship signatures which requires additional effort, with respect to modelling, namely the Doppler effect. For a ship moving at a straight path with constant velocity, the apparently time-varying frequency at the location of a resting observer is given by

$$f(t) = f_0\Big(1 + \frac{v}{c_S} \cos \text{ atan2}\big(d_{CPA}, v(t - t_{CPA})\big)\Big)^{-1} \tag{10}$$

where f_0 denotes the true frequency, v denotes the velocity of the ship, d_{CPA} denotes the distance between the ship and the observer at the point of closest approach (CPA), and t_{CPA} denotes the time point of CPA; c_S denotes the velocity of hydroacoustic sound. In Eq. (10), atan2(.,.) denotes the 4-quadrant inverse tangent function.

If within the time interval covered by a time series that is to be modelled by a state space model, frequencies vary with time due to Doppler effects, also the autoregressive parameters within the state transition matrix A will vary with time, via Eq. (4). The Kalman filter can deal with time-varying model parameters, as long as these are known at each time point. Consequently, instead of estimating the autoregressive parameters a_1, a_2, for each ARMA(2,1) model, we compute these parameters as functions of time, using Eq. (10); however, this is possible only if estimates for the *Doppler parameters* f_0, v, d_{CPA} and t_{CPA} are available. We have succeeded in estimating these parameters within the maximum-likelihood framework mentioned above.

As an example, we analyse a time series of 420 seconds length, recorded by a floating buoy in moderate distance to a moving ship; effective sampling rate was $f_s = 4$ kHz. In the left panel of Fig. 3 a non-parametric estimate of the spectrogram of the time series, restricted to the frequency interval between 492 Hz and 508 Hz, is shown; a ship line is visible, which displays a clear Doppler effect. Parametric modelling of this time series is performed by a LSS model, consisting of 9 ARMA(2,1) models and one ARMA(8,7) model; the purpose of the ARMA(8,7) model is to describe the broad spectral background. The strong line around 500 Hz, visible in the spectrogram, is described by one of the ARMA(2,1) models, while the other ARMA(2,1) models describe further ship lines. The frequencies of all ARMA(2,1) models are modelled as time-varying, according to Eq. (10), while the frequencies of the ARMA(8,7) model are kept constant. We obtain the following maximum-likelihood estimates for the Doppler parameters:

$$f_0 = 499.193 \text{ Hz} \qquad v = 8.895 \text{ m/s}$$
$$d_{\text{CPA}} = 270.24 \text{ m} \qquad t_{\text{CPA}} = 211.31 \text{ s}$$

The corresponding optimal curve describing the time-varying frequency, according to Eq. (10), is shown in the right panel of Fig. 3 (red curve).

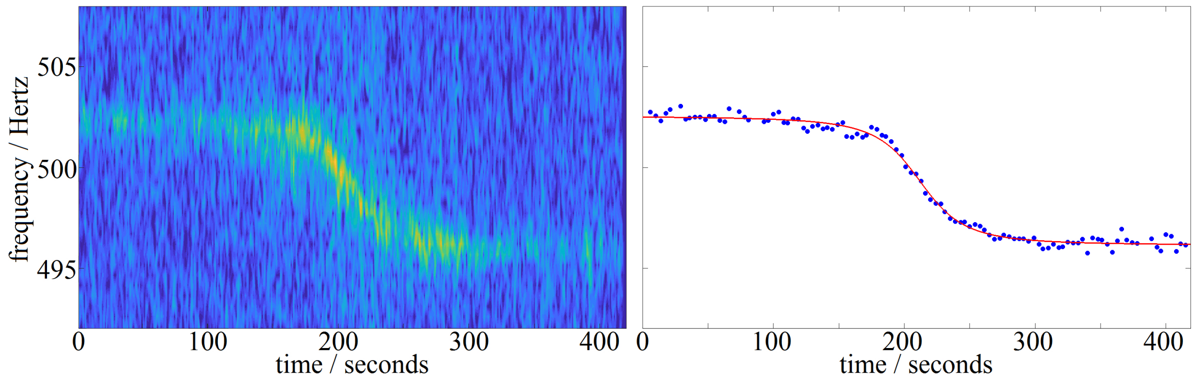

Figure 3: Non-parametric estimate of the spectrogram (left panel) and time-varying peak frequency of the corresponding ARMA(2,1) model within a LSS model (right panel), for a hydroacoustic time series of 420 seconds length; in the right panel, constant estimates from a sequence of short overlapping windows (blue dots) and the Doppler curve according to Eq. (10), using maximum-likelihood estimates for the Doppler parameters (red curve), are shown.

8. CONCLUSION

The purpose of the present paper is to demonstrate the potential power of parametric predictive modelling in time domain for the analysis of hydroacoustic time series. We have shown examples for obtaining parametric estimates of the power spectrum, for modelling combs of lines, and for modelling time-varying frequencies, e.g. in Doppler effects, all within the framework of linear state space modelling. There exists also a generalisation for modelling time-varying amplitudes, but for lack of space we have not discussed it. The state space approach offers attractive benefits, such as the possibility of obtaining noise-reduced estimates of power spectra, which is due to avoidance of overfitting in the modelling step, furthermore observation noise can be explicitly suppressed. Another benefit is given by a solution of the Blind Signal Separation problem, such that source components in time domain can be reconstructed, and by the possibility to perform time-domain filtering; due to the choice of a time-domain approach, this can be done also for components which are not well defined in frequency space.

As a disadvantage of the parametric predictive modelling approach, we mention its considerably higher computational time consumption, as compared to typical non-parametric approaches, such as Discrete Fourier transform.

REFERENCES

[1] G.E.P. Box & G.M. Jenkins: *Time Series Analysis, Forecasting and Control* (Holden-Day, 1976).

[2] J. Durbin & S.J. Koopman: *Time Series Analysis by State Space Methods* (Oxford University Press, 2012).

[3] A. Galka *et al.*: “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)**.