Abstract—An adaptive autoregressive moving average (ARMA) modelling of nonstationary EEG by means of Kalman smoother is presented. The main advantage of the Kalman smoother approach compared to other adaptive algorithms such as LMS or RLS is that the tracking lag can be avoided. This advantage is clearly presented with simulations. Kalman smoother is also applied to tracking of alpha band characteristics of real EEG during an eyes open/closed test. The observed tracking ability of Kalman smoother, compared to other methods considered, seemed to be better.

Keywords—Nonstationary EEG, Kalman smoother, adaptive algorithms

I. INTRODUCTION

Electroencephalogram (EEG) analysis is a useful tool for studying the functional states of the brain and for diagnosing certain neuropysiological states and disorders. In the analysis of nonstationary EEG the interest is often to estimate the time-varying spectral properties of the signal. A traditional approach to this is the spectrogram method, which is based on Fourier transformation. Disadvantages of this method are the implicit assumption of stationarity within each segment and the rather poor time/frequency resolution. A better approach is to use parametric spectral analysis methods based on e.g. time-varying autoregressive moving average (ARMA) modelling. The time-varying parameter estimation problem can be solved with adaptive algorithms such as least mean square (LMS) or recursive least squares (RLS). These algorithms can be derived from the Kalman filter equations [1], [2].

In this paper we use the Kalman smoother algorithm in tracking of nonstationary properties of EEG. Kalman smoother is compared to LMS and RLS algorithms in tracking of alpha band characteristics of EEG measured during an eyes open/closed test. The Kalman smoother approach is also applied to the detection of alpha waves of EEG. The main advantage of the Kalman smoother algorithm compared to other adaptive algorithms is the fact that the tracking lag can be avoided. This is demonstrated with simulations. Kalman filter has been previously used in EEG analysis in e.g. [3], [4], [5].

II. METHODS

If the signal to be modelled is nonstationary it cannot be modelled as an output of a time-invariant system. It is natural in this case to assume that the system has time-varying parameters.

A. Time-varying linear regression

Here we use a time-varying autoregressive moving average model for the signal. The time-varying ARMA(p,q) model can be written in the form

\[ z(t) = - \sum_{j=1}^{p} a_j(t) z(t-j) + \sum_{k=1}^{q} b_k(t) e(t-k) + e(t) \]  

where \( a_j(t) \) and \( b_k(t) \) are the time-varying ARMA parameters and \( e(t) \) is the driving white noise process. By denoting

\[ \theta_t = (-a_1(t), \ldots, -a_p(t), b_1(t), \ldots, b_q(t))^T \]  

\[ \varphi_t = (z(t-1), \ldots, z(t-p), e(t-1), \ldots, e(t-q))^T \]  

the model can be written in the form

\[ z_t = \varphi_t^T \theta_t + e_t \]  

where \( z_t = z(t) \) and \( e_t = e(t) \). This is clearly a linear observation model, with \( \varphi_t^T \) being the observation matrix and \( e_t \) being the observation error. A typical description for the parameter variation when no a priori information is available, is the random walk model [6]. Thus for the parameters \( \theta_t \) we write a state equation of the form

\[ \theta_{t+1} = \theta_t + w_t \]  

where \( w_t \) is a noise process. Equations (4) and (5) form a specific form of the general state space equations, with the input process \( w_t \). Now the problem is to estimate the time-varying parameters \( \theta_t \), according to the state space model.

B. Kalman filter

The Kalman filtering problem is to find the minimum mean square estimator \( \hat{\theta}_t \) for state \( \theta_t \) given the observations \( z_1, \ldots, z_t \). This has been shown to be equal to the conditional expectation value

\[ \hat{\theta}_t = E \{ \theta_t | z_1, \ldots, z_t \} \]  

We assume here the state and measurement noises \( w_t \) and \( e_t \) to be uncorrelated, zero mean, random processes with covariance matrices \( C_{w_t} = \sigma_w^2 I \) and \( C_{e_t} = \sigma_e^2 I \), so that the individual parameter evolutions are assumed to be independent. The initial state \( \theta_0 \) is assumed to be uncorrelated with \( e_t \) and \( w_t \) with finite variance. The Kalman

0-7803-7211-5/01$10.00©2001 IEEE
<table>
<thead>
<tr>
<th>Report Date</th>
<th>Report Type</th>
<th>Dates Covered (from... to)</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 Oct 2002</td>
<td>N/A</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Title and Subtitle</th>
<th>Contract Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tracking of Nonstationary EEG with Kalman Smoother</td>
<td></td>
</tr>
<tr>
<td>Approach</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Grant Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performer Organization Name(s) and Address(es)</th>
<th>Program Element Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department of Applied Physics University of Kuopio</td>
<td></td>
</tr>
<tr>
<td>Kuopio, Finland</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sponsoring/Monitoring Agency Name(s) and Address(es)</th>
<th>Project Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Army Research, Development &amp; Standardization Group</td>
<td></td>
</tr>
<tr>
<td>(UK) PSC 802 Box 15 FPO AE 09499-1500</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Abstract Notes</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved for public release, distribution unlimited</td>
<td></td>
</tr>
</tbody>
</table>

**Supplementary Notes**
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom. The original document contains color images.

<table>
<thead>
<tr>
<th>Abstract</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Subject Terms</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Report Classification</th>
<th>Subject Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>unclassified</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Classification of Abstract</th>
<th>Subject Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>unclassified</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Pages</th>
<th>Subject Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>
terval.

be calculated from averaged ARMA parameters of that in-
culated for each time instant. However in real applications
it is not usually necessary to calculate PSD estimate for
each time instant, but rather only once for a certain time
ARMA parameters, converges power spectrum can be cal-
adaptive algorithm, used to estimate the time-varying
filter equations can be written in the form
\[
\frac{\sigma^2(t)}{1 + \sum_{j=1}^{q} a_j(t)e^{-i\omega t}}
\]
where \( \sigma^2(t) \) is the prediction error variance. After the
adaptive algorithm, used to estimate the time-varying
ARMA parameters, converges power spectrum can be cal-
culated for each time instant. However in real applications
it is not usually necessary to calculate PSD estimate for
each time instant, but rather only once for a certain time
interval. The average spectrum within certain interval can be
calculated from averaged ARMA parameters of that in-
terval.

\[
\hat{\theta}_{t|t-1} = \hat{\theta}_{t-1}
\]

\[
C_{\hat{\theta}_{t|t-1}} = C_{\hat{\theta}_{t-1}} + C_w_{t-1}
\]

\[
K_t = C_{\hat{\theta}_{t|t-1}} \varphi_t = (I - K_t \varphi_t^T) C_{\hat{\theta}_{t|t-1}}^{-1}
\]

\[
C_{\hat{\theta}_t} = (I - K_t \varphi_t^T) C_{\hat{\theta}_{t|t-1}}
\]

\[
\epsilon_t = z_t - \varphi_t^T \hat{\theta}_{t|t-1}
\]

\[
\hat{\theta}_t = \hat{\theta}_{t|t-1} + K_t \epsilon_t
\]

where \( \hat{\theta}_{t|t-1} \) is the mean square estimator for state \( \theta_t \) given
the observations \( z_1, \ldots, z_{t-1}, \hat{\theta}_t \) is the state estimation
error \( \epsilon_t = \hat{\theta}_t - \hat{\theta}_t \) and \( K_t \) is the Kalman gain matrix.
The adaptation of the filter is primarily affected by \( C_w_{t} \). Note
that the unknown process \( \epsilon_t \) is here estimated with the
prediction error process \( \epsilon_t \) in every step of the iteration.

C. Fixed-interval smoother

The fixed-interval smoothing problem is to determine esti-
mates
\[
\hat{\theta}_{t|T} = E \{ \theta_t | z_1, \ldots, z_T \}
\]
for fixed \( T \) and for all \( t \) in the interval \( 1 \leq t \leq T \). The
solution for this can be written in the form [7]

\[
\hat{\theta}_{t-1|T} = \hat{\theta}_{t-1} + A_{t-1} \left( \hat{\theta}_{t|T} - \hat{\theta}_{t|t-1} \right)
\]

\[
A_{t-1} = C_{\hat{\theta}_{t-1}} C_{\hat{\theta}_{t|t-1}}^{-1}
\]

where \( A_{t-1} \) includes the error covariances stored in the
forward run of Kalman filter. Also the state estimates \( \hat{\theta}_t \) and
\( \hat{\theta}_{t|t-1} \) need to be stored. The smoothed estimates \( \hat{\theta}_{t-1|T} \) are
then obtained by running the stored estimates backwards in
time by taking \( t = T, T-1, \ldots, 2 \). The initialization is
evidently with the filtered estimate.

D. Spectral estimation

Once the time-varying coefficients of the ARMA(\( p,q \))
model (1) are solved the time-varying power spectral den-
sity (PSD) estimation can be obtained in the terms of the
estimated coefficients

\[
P_i(\omega) = \frac{\sigma^2(t)}{1 + \sum_{j=1}^{q} b_j(t) e^{-i\omega t}}
\]

where \( \sigma^2(t) \) is the prediction error variance. After the
adaptive algorithm, used to estimate the time-varying
ARMA parameters, converges power spectrum can be cal-
culated for each time instant. However in real applications
it is not usually necessary to calculate PSD estimate for
each time instant, but rather only once for a certain time
interval. The average spectrum within certain interval can be
calculated from averaged ARMA parameters of that in-
terval.

Fig. 1. AR(2) process estimation with RLS and Kalman smoother
algorithms. The root evolution and the realization are presented in
block (a). Both algorithms were optimized (b). Optimal value for
the state noise covariance coefficient of the Kalman smoother
was \( \sigma^2 = 0.001 \) and the forgetting factor of RLS was \( \lambda = 0.935 \).
The estimates of the modulus and phase angle of the root are
shown in block (c). The true values (black), Kalman smoother
estimates (red) and optimal RLS estimates (blue). The smoother
RLS estimates (green) were calculated by using \( \lambda = 0.98 \).

III. RESULTS

In order to evaluate the performance of the Kalman
smoother algorithm we conduct two simulations, where
Kalman smoother is compared to the popular forgetting
factor RLS algorithm. Finally the Kalman smoother is ap-
plied to time-varying spectrum estimation of real EEG and
for alpha wave detection.

A. Simulations

In the first simulation a time-varying signal was gener-
ated as an AR(2) process. The root evolution and a typi-
cal realization are presented in Fig. 1 (a). The modulus
and phase angle of the root were estimated with Kalman
smoother and RLS algorithms. Parameters controlling the
adaptation were optimized in both algorithms to obtain the
minimum error in AR coefficient estimation. The estimation
errors as a function of adaptation parameters for both
algorithms are presented in Fig. 1 (b). The estimates are
shown in Fig. 1 (c).

Two RLS estimates were calculated to demonstrate the
effect of the forgetting factor on the estimates. RLS esti-
mates with the optimal value for the forgetting factor have
only a small tracking lag but the estimates are far more
unstable compared to the Kalman smoother estimates. By
increasing \( \lambda \) RLS estimates become more stable but the
tracking lag increases at the same time. This simulation
shows clearly the advantages of the Kalman smoother compared to the RLS algorithm. However not much can be said about the performance of the Kalman smoother in tracking of nonstationary EEG based on this simple simulation. Hence we aim to a more realistic simulation of EEG.

In many cases we are interested in tracking of narrow band characteristics of the EEG signal. One such case is the event related desynchronization/synchronization (ERD/ERS) of alpha waves. The occipital EEG recorded while patient having eyes closed shows high intensity in the alpha band (7-13 Hz). With the opening of the eyes this intensity decreases or even vanishes. It can be assumed that EEG exhibits a transition from a stationary state to another. Such a transition was here simulated as an AR(5) process. The roots of the system for both stationary states (obtained from real EEG measurements) and the corresponding power spectrums are presented in Fig. 2 (a), where the strong peak around $0.38\pi$ is due to the synchronization of the alpha waves of EEG.

In order to make the simulation more realistic abrupt transitions of AR coefficients were smoothed as described in [8]. A typical realization of the simulated AR(5) process is presented in Fig. 2 (b). Results of tracking the alpha band characteristics are presented in Fig. 2 (c), where averaged estimates over 100 realizations of the phase angle and the magnitude of the root corresponding to alpha activity are presented. In order to obtain as smooth estimates with RLS as is obtained with Kalman smoother the forgetting factor $\lambda$ must be quite small. However this leads to substantial tracking lag. With larger values of $\lambda$ the tracking lag can be attenuated, but estimates become now more unstable.

**B. ERD/ERS of alpha waves of EEG**

The eyes open/closed test is a typical application of testing the desynchronization/synchronization of alpha waves of EEG. One such transition from desynchronized state to synchronized state is presented in Fig. 3. The performance of the Kalman smoother in tracking of alpha band characteristics is compared to other commonly used adaptive algorithms such as RLS and LMS and also to the traditional spectrogram method. An ARMA(6,2) model was used in all adaptive algorithms. The length of the time-window used in spectrogram was 2 seconds, which is long enough when considering the frequencies of the alpha band (7–13 Hz). The step size of the LMS algorithm was $\mu = 0.0002$ and the forgetting factor of RLS was $\lambda = 0.95$. The state noise covariance coefficient of the Kalman filter was $\sigma^2_w = 0.0003$.
for these concatenated signals are presented in Fig. 4 (c).

The performance of the Kalman smoother was verified by concatenating EEG epochs where alpha waves were detected and those where no detection was made. The PSD estimates, calculated with a traditional FFT based periodogram method, for these concatenated signals are presented in Fig. 4 (c) verifying the absence of alpha rhythm in the lower concatenated signal.

The aim of automatic EEG analysis is often the detection of certain waveforms. Hence the performance of the Kalman smoother on detection of alpha waves of EEG is considered here. Fig. 4 (a) presents a time-varying spectrum for an EEG sample of 15 seconds measured from channel O2.

A gap in alpha rhythm is observed after 9 seconds. The contents of this kind of gaps is considered in Fig. 4.

C. Detection of alpha rhythm of EEG

The aim of automatic EEG analysis is often the detection of alpha waves of EEG. The performance of Kalman smoother in tracking of alpha band characteristics seemed to be most reliable compared to LMS and RLS algorithms. Kalman smoother was also applied to the detection of alpha waves of EEG with success. Also two simulations were conducted showing clearly the main advantages (smooth estimates without tracking lag) of Kalman smoother compared to other adaptive algorithms. The implementation and usability of the Kalman smoother approach are straightforward. The adaptation rate is adjusted simply by setting the state covariance coefficient $\sigma_{w}^{2}$. One problem in modelling the data with adaptive algorithms is the selection of the model order. For time-invariant systems there exist various criteria for selecting the model order [9]. All these criteria are based on the compromise between model fit and model complexity. Also in the time-varying case there exist some criteria for selecting the model order. For example in [10] the use of Akaike’s information criterion (AIC) was justified in the time-varying case under certain conditions. However in the case of tracking alpha rhythm of EEG the ARMA model of order $p = 6$ and $q = 2$ seems to be suitable. The same model order was also used in [11], [12].

REFERENCES