Biol. Cybern. 76, 349–356 (1997)

Biological Cybernetics

Simulation of nonstationary EEG

J.P. Kaipio, P.A. Karjalainen

Department of Applied Physics, University of Kuopio, P.O. Box 1627, FIN-70211 Kuopio, Finland

Received: 22 June 1994 / Accepted in revised form: 18 February 1997

Abstract. In this paper we present a systematic method for generating simulations of nonstationary EEG. Such simulations are needed, for example, in the evaluation of tracking algorithms. First a state evolution process is simulated. The states are initially represented as segments of stationary autoregressive processes which are described with the corresponding predictor coefficients and prediction error variances. These parameters are then concatenated to give a piecewise time-invariant parameter evolution. The evolution is projected onto an appropriately selected set of smoothly time-varying functions. This projection is used to generate the final EEG simulation. As an example we use this method to simulate the EEG of a drowsy rat. This EEG can be described as toggling between two states that differ in the degree of synchronization of the activity-inducing neuron clusters.

1 Introduction

In a real-world setting we are hardly ever able to assess the performance of a method by just applying it to observed data (a sample). In many cases a method to be tested is compared with a previos method that is assumed to give the correct answers and thus serves as a reference. In the estimation of properties of stationary processes we have several methods with which we can obtain consistent estimates if the process is ergodic. In this case the time averages can be used to estimate the ensemble averages (Papoulis 1984). We can then make our estimates more accurate by using more data to calculate them. The accurate estimates can subsequently serve as references against which the methods to be tested are compared.

In the case of nonstationary processes we often cannot use time averages. In order to estimate a property of the process at a time instant we have, in the strict sense, just one point of the sample to build our estimate on. To get over this problem the assumption of slow time variation is often adopted. This gives us the possibility of using, foe example, adaptive segmentation which will chop the sample into quasi-stationary (almost stationary) epochs to which we would then apply methods that are intended to be used with samples from ergodic processes (Bodenstein and Praetorius 1977; Ishii et al. 1979). The applicability of this approach depends heavily on the rate of time variation, the criteria on which the segmentation is based and the correctness, or rather the adequacy, of the time series model.

There are cases in which the time variation is so rapid that the process cannot be approximated with a concatenation of stationary epochs. The durations of the quasi-stationary segments would then be too short to enable adequate estimates. This applies to both estimation and simulation of such processes. On the other hand, we cannot allow for a time variation that is totally ‘unpredictable’. This would mean that if we have obtained a sample of the process and aim to estimate its time-varying characteristics, the model may have, loosely speaking, more degrees of freedom than the sample the estimates are based on (Gersch 1991).

If we can make an assumption of *smoothness* of the time variation we can still obtain meaningful estimates and at the same time allow reasonably rapid changes of process characteristics. This follows from the fact that one can have a smooth sequence of parameters describing a process, i.e. a sequence with small second differences, that has large first differences. This means that the parameter rate of change does not have to be small. This applies also to the building of models for time-varying processes and the use of these models in the generation of simulations of these processes.

1.1 Simulation of EEG with autoregressive models

The autoregressive (AR) process model was first applied to the EEG by Zetterberg (1969), Fenwick (1969) and Gersch (1970). Since that time the AR model has been applied to the EEG in at least 160 journal articles (Kaipio 1996). The applicability of AR models to the EEG has usually been verified only experimentally, but recently there has been some interest get in the theoretical justification of this (Steinberg et al. 1985; Wright et al. 1990; Blinowska and Malinowski 1991). AR models have been used in a wide collection of

Correspondence to: J. Kaipio
(Fax: +358-17-162585, e-mail: kaipio@venda.uku.fi)

tasks concerning the EEG, including spectral analysis, classification of brain states, detection of transients, prediction of epileptic seizures and the simulation of the EEG.

The simulation of EEG using an AR model was suggested early in the 1970s (Fenwick et al. 1971; Wennberg and Zetterberg 1971; Zetterberg 1973). Simulations of time-varying AR processes were used by Wennberg and Isaakson (1976) to test the modeling and tracking capabilities of the Kalman filter. The aim in this paper was to assess the applicability of the Kalman filter to the estimation of the time-varying EEG, and the simulations were made to approximate the classical band structure of the EEG. The authors used an approximate factorization of AR model to first- and second-order spectral factors (Zetterberg 1969) which have been shown to correspond to the classical bands of the EEG. To produce time-varying simulations they let the parameters of these factors, i.e. the power, bandwidth and center frequency, be time-varying. They did not, however, discuss how the time variability should be systematically implemented to produce simulations that would exhibit the time-varying characteristics of the EEG. In addition, the classical band structure of the EEG does not always apply, especially when the EEG exhibits at least partly chaotic characteristics. They also discussed the use of simulations to aid in the recognition of certain spectral parameters of the model using visual inspection (Isaksson 1975).

1.2 Motivation and overview of the method

In this paper we suggest a systematic method which can be used to generate simulations of a time-varying EEG. The main motivation of this method is to provide researchers with realistic time-varying EEG simulations. Such simulations are of great importance when the performance of, for example, segmenting and tracking algorithms is to be evaluated objectively. This requires that the true properties of the process are known exactly. Also, it may be the case that one has not enough real EEG data to achieve sufficiently reliable evaluation results. Then the only possibility may be to rely on simulations that are based on the available EEG but are otherwise statistically independent of the observed EEG.

In the first step the existing EEG data are divided, for example, with visual inspection into segments that are then classified. These classes might correspond, for example, to the sleep stages. For each segment the AR model parameters, that is, the prediction coefficients and prediction error variances, are calculated and for each class the mean and the covariance matrix of these parameters are calculated. In this paper it is assumed that the joint distribution of these parameters can be approximated by the normal distribution. The class evolution is then estimated by, for example, Markov models or (in two-class situations) by estimating the probability distributions of segment (epoch) durations for each class.

In the second step the simulation of the EEG is generated. A piecewise constant parameter evolution is formed first for each predictor coefficient and the prediction error variance separately. The durations of constant segments (or state transition times) are drawn from the corresponding estimated distribution. The predictor coefficients and the prediction error variances are drawn from the corresponding joint distributions using the estimated means and covariances. This is done for each segment independently.

In the next step the piecewise constant parameter evolutions are smoothed by projecting them onto set of smoothly time-varying functions. Finally these projections are used as parameters of a time-varying filter with which the simulated EEG is generated.

The proposed method has been used by the authors to evaluate the performance of a root tracking algorithm (Patomäki et al. 1995) and a modified time-varying autoregressive least squares algorithm (Kaipio and Karjalainen 1995).

2 Estimation and simulation of the state evolution and class representers

We assume that the time-varying EEG can be approximated with a progression of states $s_j,\ j = 1, \ldots,$ each of which is a member of a class $\mathscr{C}_k,\ k = 1, \ldots, K$. There are several methods with which one can estimate such a class evolution if the (probabilistic) description is not known a priori. One such a method is to model the class evolution as a Markov model and estimate the parameters of the model. This model can then be used to simulate the class evolution. The state are then drawn from the respective distributions in the order that the class evolution dictates. The estimation of a continuous-time Markov model and the simulation of sleep EEG class evolution (hypnogram) has been discussed in Kemp and Kamphuisen (1986). In addition to the class evolution the model should produce the durations T_j of the states.

We assume that within these states the process can be approximated with an AR(p) process. For the selection of an appropriate order p in different situations, see Vaz et al. (1987). The final selection of p should be the maximum of the adequate orders of all classes. We also assume that within each state the coefficients $\phi_i^j,\ i = 1, \ldots, p$ of the predictor and the prediction error variance γ_j corresponding to the AR(p) model are drawn from the density that can be approximated with a $p + 1$-dimensional normal distribution. Each class has its own distribution. We call the vector $\theta^j = (\phi_1^j, \ldots, \phi_p^j, \gamma_j)'$, where the prime denotes transposition, a representer of the state s_j. We will further assume that the representer of the state s_j is independent of the representers of the states s_ℓ for all $j \neq \ell$ if the class to which s_j belongs is known.

The assumption of approximate normality of representers implies that we only need to estimate the means and covariances of θ^k for each class $\mathscr{C}_k$ to be able to generate representers. If the means and covariances of θ^k are not known, they can be estimated from segmented and classified EEG data by, for example, using any preferred method to calculate the parameters of the AR(p) model and the associated residual error variance for each segment. One such method is the modified covariance method (Marple 1980). This can be used *in lieu* of the predictor coefficients and the prediction error variances. Let there be $n_k > p$ estimates $\widehat{\theta}^k$ for θ^k. Estimates for the mean $\widehat{\mu}_k$ and covariance $\widehat{C}_k$ of $\widehat{\theta}^k$ can then be obtained as

$$\widehat{\mu}_k = n_k^{-1} \sum_{i=1}^{n_k} \widehat{\theta}_i^k \tag{1}$$

$$\widehat{C}_k = (n_k - 1)^{-1} \sum_{i=1}^{n_k} (\widehat{\theta}_i^k - \widehat{\mu}_k)(\widehat{\theta}_i^k - \widehat{\mu}_k)' \tag{2}$$

We can now generate representers from the thus estimated $p + 1$-variate normal distributions

$$f_k(\theta) = (2\pi)^{-p/2} |\widehat{C}_k|^{-p/2} \exp\left(-(\theta - \widehat{\mu}_k)' \widehat{C}_k^{-1} (\theta - \widehat{\mu}_k)\right) \tag{3}$$

where $|\widehat{C}_k|$ denotes the determinant of the $\widehat{C}_k$, as follows: Generate a $p + 1$-vector g of independent normal variables with zero mean and unit variance. Then the variables $\theta = L_k g + \widehat{\mu}_k$ are distributed as $f_k(\theta)$, where $\widehat{C}_k = L_k' L_k$ is the Cholesky decomposition of the $\widehat{C}_k$ (Papoulis 1984).

If the representers have been estimated from relatively short segments, the covariances $\widehat{C}_k$ may include a non-negligible contribution that is due to the small sample estimate properties. In such a case we can reduce the variances of the classes simply by dividing the covariance matrices by a number that is greater than unity. While this method is not the correct way to diminish the small sample contribution, it maintains the eigenvector structure of $\widehat{C}_k$ and only diminishes the eigenvalues. This means that the only result of this operation is that the deviances of the simulated θ^k values from the means $\widehat{\mu}_k$ will be smaller. This will also decrease the possibility of obtaining a temporarily unstable model.

An AR(p) model is stable if and only if the roots $z_i,\ i = 1, \ldots, p$ of the polynomial

$$X(z, \phi) = 1 - \sum_{i=1}^{p} \phi_i z^{-i} \tag{4}$$

fulfill $|z_i| < 1$ for all t. For some classes the roots of $X(z, \phi)$ with $\phi_i = \mu_i,\ i = 1, \ldots, p$ may have almost unit modulus or the covariance may be large, which means that the probability of obtaining an unstable representer is non-negligible. This is why the stability of each representer should be verified by calculating the roots of the polynomial. All roots with modulus greater than or equal to unity should be transferred to lie inside the unit circle and the predictor reassembled. This can be done by factoring the polynomial to (complex) roots and multiplying each unstable root z_i by $|z_i|^{-2}$. This operation will approximately restore the shape of the spectrum in most practical cases (Ljung 1988, subroutine fstab).

3 Generation of the time-varying AR process

We will discuss first two types of direct concatenation that are not feasible for generating time-varying simulations and how this unfeasibility can be overcome. Then we discuss the details of the method that was briefly described at the end of Sect. 1.

3.1 Direct concatenation

The majority of time series methods that are designed to track the characteristics of time-varying processes are based on either direct or indirect minimization of the prediction (residual) error (Ljung 1987). It is thus desirable that the optimal predictor of the simulation behaves in such a way that it makes sense to compare the estimated model with it. For example, if the coefficients $b_k(t)$ of the optimal predictor will decay relatively fast, we would use as the definition of time-varying spectral density (Gersch 1991)

$$P(\omega, t) = \frac{\sigma_e^2(t)}{\left|1 - \sum_{k=1}^{\infty} b_k(t) \exp(-i\omega k T)\right|^2} \tag{5}$$

There are two trivial ways to use concatenation to generate simulations of nonstationary processes which are not directly usable. The first is to generate stationary segments separately and then concatenate them directly. There are two major problems with this kind of simulation. In this case we actually have a nonoverlapping sum of independent processes. Let $x(t)$ be the process with zero mean for all t and t^* be the first instant of a new segment. Then, irrespective of the values at time $t^* - 1$, the correlation of $x(t^*)$ with the entire history of $x(t)$ is zero, the predictor for $x(t^*)$ is zero and the spectrum estimate $P(\omega, t^*) = \sigma_e^2(t^*)$. This is clearly not desirable.

At the neuronal level this kind of sudden change in process characteristics would correspond to a situation where one activity-inducing neuron cluster shuts down abruptly and at the same time another cluster achieves its full effect. This is a totally unrealistic situation.

The second type of concatenation involves the concatenation of the parametric representations of the segments. This concatenation could then be used as a time-varying filter with abrupt filter coefficient changes at the segment borders. As above, this kind of change would correspond to an unrealistic abrupt change in the physical and chemical state of the neuron cluster. In addition, it turns out that the calculation of the optimal predictor is problematic.

For this reason we have chosen to build a smoothly changing parametric representation for the process. We use a time-varying AR(p) model as the representation and generate the simulations by feeding white noise to the corresponding filter. If the coefficients of the model do not change very much during the correlation time of the process, the optimal predictor is approximately equal to the coefficients of the AR(p) at each time. Thus the reference against which the estimates are compared is directly accessible.

3.2 The generation of smooth parameter evolution

We start from a sequence of representers (predictor coefficients and prediction error variances $\phi_j^k,\ j = 1, \ldots, p$ and γ_k, respectively) which are the result of the simulated state evolution as described above. Let the corresponding sequence of the segments durations be $T_k,\ k = 1, \ldots$. We call the concatenate of the segments a *block*, which is of length $T = \sum T_k$. First we build the coefficient matrix Φ^c of piecewise constant parameter evolution

$$\Phi^{\mathrm{c}} = \begin{pmatrix} \left.\begin{matrix} \phi_1^1 & \phi_2^1 & \cdots & \phi_p^1 \\ \vdots & & & \vdots \end{matrix}\right\} T_1 \\ \left.\begin{matrix} \phi_1^1 & \phi_2^1 & \cdots & \phi_p^1 \\ \phi_1^2 & \phi_2^2 & \cdots & \phi_p^2 \\ \vdots & & & \vdots \end{matrix}\right\} T_2 \\ \begin{matrix} \phi_1^2 & \phi_2^2 & \cdots & \phi_p^2 \\ \phi_1^3 & \phi_2^3 & \cdots & \phi_p^3 \\ \vdots & & & \vdots \end{matrix} \end{pmatrix} = [\Phi_1^{\mathrm{c}}\ \Phi_2^{\mathrm{c}}\ \ldots\ \Phi_p^{\mathrm{c}}] \tag{6}$$

Thus Φ^{c} contains in its columns Φ_i^{c} the stepwise constant parameter evolutions of each predictor coefficient. The length of each constant segment is T_k, so that the full row dimension of Φ^{c} is T. The column dimension is the order p of the AR model.

The next step is to smooth each individual parameter evolution Φ_j^{c}. This can be done by assuming that each of them is a linear combination of some fixed set of M linearly independent smooth functions φ_i, $i = 1, \ldots, M$ of duration (length) T. The selection of such functions is discussed later. We call the set φ_i a basis, since it obviously is a basis of an M-dimensional subspace of the T-dimensional space of all possible one-parameter evolutions. We now collect the basis vectors φ_i to matrix $S = [\varphi_1, \ldots, \varphi_M]$ of size $T \times M$. Now, because we assumed that the smoothed parameter evolutions Φ_j are linear combinations of φ_i, they are of the form

$$\Phi_j = Sc_j\ , \tag{7}$$

where c_j is a $M \times 1$ vector. It is well known that if we want the Φ_j vectors to be the best approximations for the Φ_i^{c} vectors in the least squares sense, the vector c_j of coefficients can be calculated from

$$c_j = (S'S)^{-1}S'\Phi_j^{\mathrm{c}} \tag{8}$$

and we have $\Phi_j = S(S'S)^{-1}S'\Phi_j^{\mathrm{c}}$ (Lawson and Hanson 1995). Thus the smoothed parameter vectors are the orthogonal projections of Φ_j^{c} to subspace $\mathscr{S}$ spanned by the columns of S. The projection can be performed simultaneously for all $j = 1, \ldots, p$ giving

$$\Phi = S(S'S)^{-1}S'\Phi^{\mathrm{c}} = SC \tag{9}$$

where $C = [c_1, \ldots, c_p]$ is a $M \times p$ coefficient matrix. The columns of Φ are now the smoothed predictor coefficients Φ_j that are of the form

$$\Phi_j = \Phi_j(t) = Sc_j = \sum_{i=1}^{M} \varphi_i(t)c_{ji}\ , \tag{10}$$

where time t now refers to the row index of the vectors Φ_j and φ_i. To complete the specification we need the correspondingly smoothed evolution of the prediction error variance. As above, concatenate γ_i to give

$$\Gamma^{\mathrm{c}} = (\overbrace{\gamma_1, \ldots, \gamma_1}^{T_1}, \overbrace{\gamma_2, \ldots, \gamma_2}^{T_2}, \gamma_3, \ldots)' \tag{11}$$

Then using the same basis we obtain

$$\Gamma = S(S'S)^{-1}S'\Gamma^{\mathrm{c}} \tag{12}$$

The matrix Φ and the vector Γ now contain all information about the temporal variations of the EEG simulation to be generated.

Next we generate a white noise sequence of zero mean and unit variance $e(t)$, $t = 1, \ldots, T$. The simulated EEG sequence is then obtained by filtering $e(t)$ with the time varying filter defined by Φ and Γ:

$$x(t) = \sum_{j=1}^{p} \Phi_j(t)x(t-j) + \sqrt{\Gamma(t)}e(t)\ . \tag{13}$$

Due to the highly nonlinear mapping of polynomial coefficients to the roots, it would be very burdensome to add the stability condition $|z_k(t)| < 1$ to the projection as a constraint in the least squares problem corresponding to the projection. This constrained least squares problem can be expressed as

$$\min_{\Phi}\{\|\Phi - \Phi^{\mathrm{c}}\|,\ \Phi_j \in \mathscr{S}\ \forall j,\ |z_k(t)| < 1\ \forall k, t\} \tag{14}$$

Since Φ tends to overshoot the stationary representers in parameter space, it is probable that in some cases the roots $z_k(t)$ will be temporarily unstable. This is, however, not a problem if the overshoot and its extent in time are not large since in this case $x(t)$ does not diverge much.

3.3 The choice of basis for $\mathscr{S}$

It is usually necessary to include the constant basis as one of the functions so that the stationary case falls into this setting as a special case whatever the basis. For convenience we will thus assume that $\varphi_1 \equiv 1$. See Kaipio and Karjalainen (1995) for an example where this is not the case.

The basis function selection has been discussed in the literature concerning time-varying AR modeling; for a review see Gersch (1987). The problem of estimating an optimal time-varying AR (TVAR) model can be viewed to be the inverse to the problem of generating simulations. In TVAR modeling the parameter evolution is estimated from an EEG sample and is constrained so that each parameter process is in $\mathscr{S}$.

The basis functions suggested to be used in conjunction with TVAR models include general polynomial, Fourier, Haar, Legendre, spline and prolate spheroidal wave bases. With the exception of the Haar basis, which is a block pulse basis, the suggested bases are able to model smooth changes in the parameter evolution. This ability is dependent on the choice of basis and the parameters concerning it – the dimension of $\mathscr{S}$ in particular.

The problem with these bases is that they are global within the block, so that a small change anywhere in the block induces changes that extend throughout the block. This is obviously not desirable. We wish to maintain the local nature of the basis and simultaneously obtain a smooth parameter evolution.

We have chosen to use a basis consisting of shifted and time-scaled Gaussian functions $\varphi_i(t)$:

$$\varphi_i(t) = \exp\left(-(t - t_i)^2/d_i^2\right), \quad i = 2, \ldots, M \tag{15}$$

Obviously there can be no fixed general choice for M, t_i or d_i that would be optimal in some sense. The choice of these parameters depends heavily on the underlying situation and the demands. In the case of simulation the feasible parameters depend on the number and durations of the

stationary segments, or more generally, the distributions of these.

According to our experience the Gaussian basis fulfills the requirements of locality and smoothness of evolution with relatively simple selection rules of M, t_i or d_i for a wide class of evolutions. The trivial choice for t_i and d_i is such that the half-widths of $\varphi_i(t)$ will exactly cover the length of the block. This leads to $t_i = (i-1)T/(M-1)$, $2Md_i = T$, where T is the length of the block. The selection of M is yet left open. If the block consists of D segments of length L each, then $T = DL$ and we could choose $M = D+1$. This selection leads to projections that are smooth but tend to overshoot the design classes heavily in the parameter space.

We have found that the selection

$$M \approx 6D \tag{16}$$

$$d_i \approx 3T/M \tag{17}$$

$$t_i = (i-1)T(M-1) \tag{18}$$

will produce reasonable projections if the segment length distribution is not very wide.

3.4 Other methods of obtaining a smooth parameter evolution

Naturally it is possible to tailor the bases individually to follow the state evolution and just to smooth the transients of the parameter evolution. One way to implement this indirectly is to use the sigmoidal basis. An example of a sigmoidal type basis function is

$$\varphi_i(t) = (1 + \exp(-c(t - t_i)))^{-1} \tag{19}$$

where t_i are selected to coincide with the segment borders and c is adjusted to give the desired rise times.

It is also easy to build Φ^c and then filter the columns with a (noncausal) zero-phase low-pass filter. Just as with the sigmoidal basis this will, however, lead to a situation where the simulation would be a concatenate of stationary and transition regions which is usually not very realistic.

In addition, it has been shown that the tracking capability of adaptive estimators depends on the type of evolution of the parameters (Benveniste et al. 1990). If the parameters were to be time-invariant for a long period of time, the estimators could perform markedly better with the simulation than with the real EEG data.

In principle it is possible to realize the parameter evolution as a multivariable AR process. For many situtions this could be assumed to be a feasible model for parameter evolution. In time-varying AR estimation (tracking) the use of a Kalman filter would yield such a multidimensional parameter process estimate (Isaksson 1975; Ishii et al. 1980). The probability density function of the parameters is, however, very difficult to control so that the model remains stable at all times. If the parameter processes are forced to the stability region after generation, the smoothness exhibited by, for example, a second-order low-pass type of AR process is difficult to maintain. The testing for stability must be carried out for each time, which would be very burdensome with long simulations.

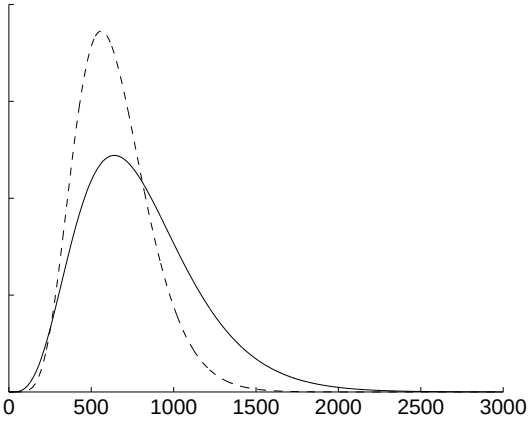

Fig. 1. The maximum likelihood estimates $g_1(T) = g(T; 4, 0.006)$ and $g_2(T) = g(T; 7, 0.013)$ of the segment duration densities for classes 1 (*continuous line*) and 2 (*dashed line*)

4 An example

As an example of the use of the proposed method we simulate the EEG (electrocorticogram) of a drowsy rat. This was chosen for the sake of simplicity since this EEG can be approximated as a process that toggles between states that belong to two classes only. The amount of data on which the statistics we give here are based on, is far too small to enable any true inference. The main point of this example is to illustrate the steps of the proposed method in detail.

4.1 The state evolution

The data were visually classified into two classes. As an aid in the classification the spectrum estimates and root locations of the modified covariance estimates of AR(6) model (Marple 1980) were used.

The exponential distribution induced by the Markov model did not fit the experimental segment duration distribution of either class. We adopted another model for the class evolution. Since there are only two classes, class 1 always changes to class 2 and vice versa. Thus we can model the class evolution simply by the estimated distributions of the segment durations. The gamma distribution

$$g(T; \lambda_1, \lambda_2) = \frac{\lambda_2^{\lambda_1+1}}{G(\lambda_1+1)} T^{\lambda_1} \exp(-\lambda_2 T) \tag{20}$$

seemed to fit the observed histograms of both states, where $G(\cdot)$ is the gamma function. The fitted probability densities $g_1(T) = g(T; \lambda_1^1, \lambda_2^1)$ and $g_2(T) = g(T; \lambda_1^2, \lambda_2^2)$ are shown in Fig. 1.

4.2 The representer estimates

We used the modified covariance method to estimate the AR(6) models for both classes. These coefficients and the corresponding residual variances were used as estimates of the predictor coefficients and prediction error variances. Forty estimates of the representers θ^1 and θ^2 were obtained.

Table 1. The estimated means $\hat{\mu}_1$, $\hat{\mu}_2$ and the covariances $\hat{C}_1$ and $\hat{C}_2$ of classes 1 and 2

	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	γ
$\hat{\mu}_1$	0.5065	−0.0528	0.0619	−0.0828	−0.000	−0.143	0.5686
$\hat{\mu}_2$	0.6722	−0.3583	−0.174	−0.0562	−0.034	−0.281	1.1819
$\hat{C}_1$							
	0.0362	−0.033	0.0182	−0.001	0.0049	−0.005	0.0223
	−0.033	0.0448	−0.018	−0.003	−0.002	0.0031	−0.029
	0.0182	−0.018	0.0283	−0.002	−0.003	−0.000	0.0056
	-0.0016	−0.003	−0.002	0.0139	0.0003	−0.006	−0.005
	0.0049	−0.002	−0.003	0.0003	0.0182	−0.012	0.0026
	−0.005	0.0031	−0.000	−0.006	−0.012	0.0183	0.0082
	0.0223	−0.029	0.0056	−0.005	0.0026	0.0082	0.0444
$\hat{C}_2$							
	0.0177	−0.006	−0.010	0.0127	0.0073	−0.004	−0.011
	−0.006	0.0179	−0.011	0.0111	−0.009	0.0097	−0.007
	−0.010	−0.011	0.0398	−0.034	−0.002	0.0157	0.0083
	0.0127	0.0111	−0.034	0.0408	−0.007	−0.002	−0.017
	0.0073	−0.009	−0.002	−0.007	0.0197	−0.019	0.0066
	−0.004	0.0097	0.0157	−0.002	−0.019	0.0359	−0.001
	−0.011	−0.007	0.0083	−0.017	0.0066	−0.001	0.4413

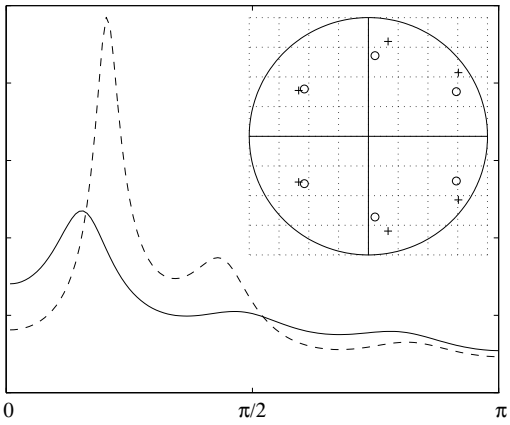

Fig. 2. The spectra of classes 1 (*continuous line*) and 2 (*dashed line*) corresponding to the means μ_1 and μ_2 and the model polynomial root (pole) locations in the complex plane (*open circles*, class 1, *crosses*, class 2)

The marginal densities of each $\phi_i,\ i = 1, \ldots, 6$ were visually examined to support the normal approximation of the predictor coefficients. The marginal densities of the prediction error variances γ_1 and γ_2, however, could not be approximated with the normal distribution, which was to be expected. To use another distribution would be very cumbersome, so we let the joint density be normal but applied hard limits to γ_1 and γ_2 when the representers were drawn from the distribution. As hard limits we used the minimum and maximum of the observed residual variances of both classes.

The means $\widehat{\mu}_1$, $\widehat{\mu}_2$ and the covariances $\widehat{C}_1$ and $\widehat{C}_2$ are given in Table 1. The spectra and root locations corresponding to the means are given in Fig. 2. Note that the frequency corresponding to the second peak in the spectrum of class 2 is approximately 2 times the frequency of the first. This can be taken as a sign of the partially chaotic nature of state 2.

4.3 The generation of simulations

The segment duration sequence was generated by drawing independent random numbers from the distributions g_1 and g_2 in turn. The representer sequence was generated as described in Sect. 2.

The segment durations and representers of a simulation are given in Table 2. The corresponding concatenated predictor coefficient processes Φ_1^c and Φ_2^c and the corresponding smoothed processes Φ_1 and Φ_2 are shown in Fig. 3a. The prediction error variances Γ^c and Γ are shown in Fig. 3b. The corresponding simulation is shown in Fig. 3c and an example of the original EEG in Fig. 3d.

To verify the similarity of the quasi-stationary epochs and the transition region between the original rat EEG data and the simulations, we show in Fig. 4a a segment of original data and, in Fig. 4c, a simulation which is adapted to the original segment as follows: The EEG was divided visually into three parts and the representers for the quasi-stationary parts were estimated. The temporal extent of the representers was broadened to cover the transition region. The concatenated and smoothed parameter evolution is shown in Fig. 4c.

5 Discussion

We have presented a systematic method which can be used to simulate nonstationary EEG. The applicability of the method to, for example, human sleep EEG simulation is obvious. Apart from EEG that can be described by a stochastic class evolution, this method is also applicable to such processes that exhibit a deterministic type state evolution.

An example is the gradual desynchronization of alpha waves as a response to visual stimulation (Markand 1990). In such a case we can take samples of the EEG that are synchronized with the stimulus, estimate representer statistics before the stimulus using one segment $(\mathscr{C}_1; T_1)$ and after the stimulus using several segments $(\mathscr{C}_2, \mathscr{C}_3, \ldots; T_2, T_3, \ldots)$. We can then use the predetermined class evolution (order)

Table 2. The durations and representers of a simulation

Class	$\sum T_i$	T_i	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	γ
1	0	645.5	0.3402	0.0455	−0.02277	−0.08406	0.0358	−0.2347	0.437
2	645	410.2	0.7642	−0.3862	−0.2483	0.06458	−0.01698	−0.3176	2.224
1	1056	804.8	0.25	0.161	−0.1543	0.06727	0.05747	−0.295	0.3[a]
2	1861	532.5	0.9159	−0.3473	−0.53	0.254	0.106	−0.4673	1.159
1	2393	1368	0.4241	0.05525	0.03071	0.01801	0.1109	−0.3348	0.3509
2	3761	555.4	0.774	−0.4261	−0.3595	0.02872	0.1362	−0.5736	1.522

The segment durations T_i drawn from the distributions $g_1(T)$ and $g_2(T)$, and the predictor coefficients ϕ_j, $j = 1, \ldots, 6$ and prediction error variances γ drawn from the normal distribution with means $\hat{\mu}_1$, $\hat{\mu}_2$ and covariances $\hat{C}_1$ and $\hat{C}_2$ are shown. [a] The prediction error has been constrained to lie in the observed region

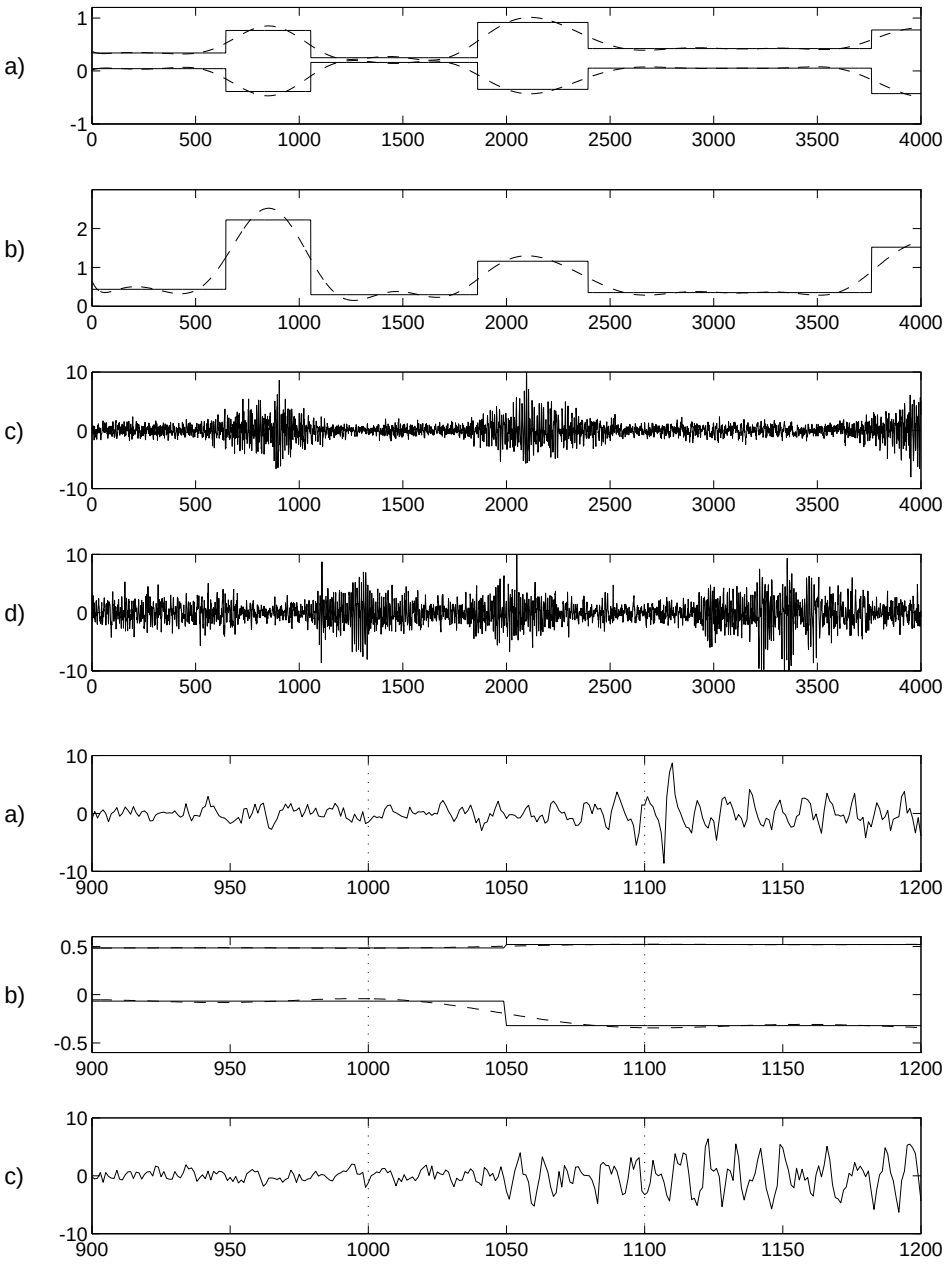

Fig. 3. **a** The concatenated predictor coefficient processes Φ_1^c and Φ_2^c and the corresponding smoothed processes Φ_1 and Φ_2. **b** As in a) but for prediction error variances Γ^c and Γ. **c** The corresponding simulation. **d** An example of original rat EEG

Fig. 4. **a** A segment of original rat EEG data. The *dashed lines* divide the segment to three parts: state 1, transition and state 2. **b** The concatenated parameter evolutions Φ_1^c, Φ_2^c and the corresponding smoothed evolutions. **c** An adapted simulation

$k_1, k_2, \ldots$ and segment durations and vary the representers only. [See Kaipio and Karjalainen (1995), where the performance of a modified TVARLS estimation scheme was evaluated with simulations of this kind.]

Signals such as the electro-oculogram (EOG), electromyogram (EMG) and some event-related potentials can be added to the simulations directly. When doing this, however, it must be noted that the occurrences may correlate some states of the background. An example is the correlation of

EOG with REM sleep. To achieve a realistic situation the occurrence statistics should be estimated and used accordingly.

It is well known that some epochs of the EEG are better described as a mixture of chaotic and stochastic behaviour rather than a regular stochastic process such as an AR(p) model (see, e.g., Pijn et al. 1991). However, for short segments the main difference between chaotic and regular description is that the former exhibits phase correlation whereas the phase of the latter should be independent between any two different frequencies and have uniform distribution between zero and 2π. Short segments with phase coherence can often be approximated with a limiting case of a regular process having a line spectrum with the lines exactly at the multiples of the inverse of the period. Such a process can be approximated also with an AR(p) model. An example is state 2 of the rat EEG and the corresponding simulation of Fig. 4. The visual appearance of the chaotic state 2 does not differ much from the simulation that is regular in construction. See Blinowska (1991) for a discussion of the applicability of the AR(p) model to chaotic EEG.

If this approach is not considered adequate, the prediction error process $\Gamma(t)$ can be smoothly forced to zero and the gap thus obtained filled by a chaotic process that is multiplied with a taper to avoid the abrupt changes in process characterictics discussed in Sect. 3.1. See Freeman (1987) for an example of the simulation of chaotic EEG. Another method of achieving this is to generate a noise process with constant spectrum but with appropriate phase coherence. This can then be fed to the time-varying filter.

References

Benveniste A, Metivier M, Priouret P (1990) Adaptive Algorithms and Stochastic Approximations. Springer-Verlag.

Blinowska K, Malinowski M (1991) Non-linear and linear forecasting of the EEG time series. Biol Cybern 66:159–165

Bodenstein G, Praetorius H (1977) Feature extraction from the electroencephalogram by adaptive segmentation. Proc IEEE 65:642–652

Fenwick P, Mitchie P, Dollimore J, Fenton G (1969) Application of the autoregressive model to EEG analysis. Agressologie 10:553–564

Fenwick P, Mitchie P, Dollimore J, Fenton G (1971) Mathematical simulation of the electroencephalogram using an autoregressive series. Int J Bio-Med Comput 2:281–307

Freeman W (1987) Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. Biol Cybern 56:139–150

Gersch W (1970) Spectral analysis of EEGs by autoregressive spectral decomposition of time series. Math Biosci 7:205–222

Gersch W (1987) Non-stationary multichannel time series analysis. In Handbook of electroencephalography and clinical neurophysiology, vol 1, Methods of Analysis of Brain Electrical and Magnetic Signals Elsevier, Amsterdam, pp 261–296

Gersch W (1991) Smoothness priors. In New directions in time series analysis, part II. Springer, Berlin Heidelberg New York, pp 113–146

Isaksson A (1975) On time variable properties of EEG signals examined by means of a Kalman filter. Technical report 95, Royal Institute of Technology, Stockholm

Ishii N, Iwata A, Suzumura N (1979) Segmentation of non-stationary time series. Int J Systems Sci 10:883–894

Ishii N, Iwata A, Suzumura N (1980) Detection of abrupt change and trend in the time series. Int J Systems Sci 11:557–566

Kaipio J, Karjalainen P (1995) TVAR modeling of event-related synchronization changes. The optimal basis approach. Technical report, Department of Applied Physics, University of Kuopio. To appear also in IEEE Trans Biomed Eng

Kaipio J (1994) Simulation and Estimation of Nonstationary EEG. PhD thesis, University of Kuopio (1996)

Kemp B, Kamphuisen H (1986) Simulation of human hypnograms using a Markov chain model. Sleep 9:405–414

Lawson C, Hanson R (1995) Solving least squares problems. SIAM, Philadelphia, PA

Ljung L (1987) System identification: theory for the user. Prentice-Hall, Englewood Cliffs, NJ

Ljung L (1988) System Identification Toolbox for MATLAB. The Mathworks Inc, Natick, MA

Markand O (1990) Alpha rhythms. J Clin Neurophysiol 7:163–189

Marple L (1980) A new autoregressive spectrum analysis algorithm. IEEE Trans Acoust, Speech Signal Processing 28:441–454

Papoulis A (1984) Probability, Random variables and stochastic processes. McGraw-Hill, New York

Patomäki L, Kaipio J, Karjalainen P (1995) Tracking of nonstationary EEG with the roots of ARMA models. In: IEEE Conference EMBC-95

Pijn J, van Veerven J, Noest A, da Silva FL (1991) Chaos or noise in EEG signals. Dependence on state and brain site. Electroencephalogr Clin Neurophysiol 79:371–381

Steinberg HW, Gasser T, Franke J (1985) Fitting autoregressive models to EEG time series: an empirical comparison of estimates of the order. IEEE Trans Acoust, Speech Signal Processing 33:143–150

Vaz F, Guedes de Oliveira P, Principe J (1987) A study on the best order for autoregressive EEG modelling. Int J Biomed Comput 20:41–50

Wennberg A, Isaksson A (1976) Simulation of nonstationary EEG signals as a means of objective clinical interpretation of EEG. In Quantitative analytical studies in epilepsy, Raven Press, New York, pp 493–509

Wennberg A, Zetterberg L (1971) Application of a computer-based model for EEG analysis. Electroencephalogr Clin Neurophysiol 31:457–468

Wright J, Kydd R, Sergejev A (1990) Autoregression models of EEG. Biol Cybern 62:201–210

Zetterberg L (1969) Estimation of parameters for a linear difference equation with application to EEG analysis. Math Biosci 5:227–275

Zetterberg L (1973) Experience with analysis and simulation of EEG signals with parametric description of spectra. In Kellaway P and Petersen I, editors, Automation of Clinical Electroencephalography, Raven Press, New York, pp 161–201