Computers in Biology and Medicine 37 (2007) 183 – 194

Computers in Biology and Medicine

www.intl.elsevierhealth.com/journals/cobm

Selection of optimal AR spectral estimation method for EEG signals using Cramer–Rao bound

Abdulhamit Subasi*

Department of Electrical and Electronics Engineering, Kahramanmaras Sutcu Imam University, 46601 Kahramanmaraş, Turkey

Received 7 June 2005; received in revised form 27 September 2005; accepted 21 December 2005

Abstract

Electroencephalography is an essential clinical tool for the evaluation and treatment of neurophysiologic disorders related to epilepsy. Careful analyses of the electroencephalograph (EEG) records can provide valuable insight and improved understanding of the mechanisms causing epileptic disorders. The detection of epileptiform discharges in the EEG is an important element in the diagnosis of epilepsy. In this study, EEG signals recorded from 30 subjects were processed using autoregressive (AR) method and EEG power spectra were obtained. The parameters of autoregressive method were estimated by different methods such as Yule-Walker, covariance, modified covariance, Burg, least squares, and maximum likelihood estimation (MLE). EEG spectra were then used to analyze and characterize epileptiform discharges in the form of 3-Hz spike and wave complexes in patients with absence seizures. The variations in the shape of the EEG power spectra were examined in order to obtain medical information. These power spectra were then used to compare the applied methods in terms of their frequency resolution and determination of epileptic seizure. The Cramer–Rao bounds (CRB) were derived for the estimated AR parameters of the EEG signals and the performance evaluation of the estimation methods was performed using the CRB values. Finally, the optimal AR spectral estimation method for the EEG signals was selected according to the computed CRB values. According to the computed CRB values, the performance characteristics of the MLE AR method was found extremely valuable in EEG signal analysis.

Keywords: Electroencephalograph; Epileptic seizure; AR spectral estimation method; Power spectral density; Cramer–Rao bound

1. Introduction

About 1% of the people in the world suffer from epilepsy and 30% of epileptics are not helped by medication [1]. Research is needed for better understanding of the mechanisms causing epileptic disorders. Careful analyses of the electroencephalograph (EEG) records can provide valuable insight into this widespread brain disorder. The detection of epileptiform discharges occurring in the EEG between seizures is an important component in the diagnosis of epilepsy. In this work, autoregressive (AR) methods were used to analyze epileptiform discharges in recorded brain waves (EEG) from patient with absence seizures (petit mal). Absence seizure is one of the main types of generalized seizures and the underlying pathophysiology is not completely understood. Neurologists make an absence seizure epileptic diagnosis primarily through visual identification of a 3-Hz spike and wave complex [1–3].

An EEG contains a wide range of frequency components. However, the range of clinical and physiological interests is between 0.5 and 30 Hz. This range is divided into a number of frequency bands as follows [4]:

Delta (0.5–4 Hz): Delta rhythms are slow brain activities preponderant only in deep sleep stages of normal adults. Otherwise, they suggest disease.

Theta (4–8 Hz): This EEG frequency band exists in normal infants and children as well as during drowsiness and sleep in adults. Only a small amount of theta rhythms appears in the normal waking adult. Presence of high theta activity in awake adults suggests pathological conditions.

Alpha (8–13 Hz): Alpha rhythms exist in normal adults during relaxed and mentally inactive awakeness. The amplitude is mostly less than 50 μV and appears most prominent in the

* Tel.: +90 344 219 1253.
E-mail address: asubasi@ksu.edu.tr.

0010-4825/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2005.12.001

occipital area. Alpha rhythms are blocked by opening the eyes (visual attention) and other mental efforts such as thinking.

Beta (13–30 Hz): Beta activity is mostly marked in fronto-central region with less amplitude than alpha rhythms. It is enhanced by expectancy states and tension.

Since there is no single criterion evaluated by the experts, visual analysis of EEG signals in time domain may be insufficient. Therefore, some automation and computer techniques have been used for this aim. Since the early days of automatic EEG processing, representations based on a Fourier transform have been most commonly applied. This approach is based on earlier observations that the EEG spectrum contains some characteristic waveforms. A number of spectral estimation method have recently been developed and compared to the more standard fast Fourier transform (FFT) method have been studied in the literature [5–9]. AR spectra can be computed by different algorithms such as the Burg method and Yule-Walker method [5–12].

A number of spectral estimation techniques have been developed recently for EEG signal processing. The AR method is the most frequently used among model-based (parametric) methods, since the estimation of the parameters in the AR signal models is a well-established topic and the estimates are found by solving linear equations of the system. The parameters of the AR method can be estimated by using different estimation methods such as Yule-Walker, covariance, modified covariance, Burg, least squares, and maximum likelihood estimation (MLE) [5–12].

In this study, EEG signals were obtained from 30 subjects, 5 with epilepsy and 25 controls. The rest of them had been healthy subjects, were examined by taking into consideration of their power spectral densities (PSDs). The PSDs of the EEG signals were obtained by different parametric methods. The AR parameters were estimated by Yule-Walker, covariance, modified covariance, Burg, least squares, and MLE methods. We provided detailed analysis of the EEG signals; hence spectral distributions of these signals were visualized. These parametric estimation methods were compared in terms of their frequency resolution and the effects in epileptic seizure detection. The Cramer–Rao bounds (CRBs) were derived for the estimated AR parameters and the performance evaluation was performed by using CRB values. According to the computed CRB values, the optimal AR spectral estimation method was selected for the EEG signals.

2. Materials and methods

2.1. EEG data acquisition and representation

Scalp EEG signals are synchronous discharges from cerebral neurons detected by electrodes attached to the scalp. Epileptic seizure is an abnormality in EEG recordings and characterized by brief and episodic neuronal synchronous discharges with dramatically increased amplitude. This anomalous synchrony may occur in the brain locally (partial seizures) which is seen only in a few channels of the EEG signal, or involving the whole brain (generalized seizures) which is seen in every channel of the EEG signal. Four channels of EEG (F7-C3, F8-C4, T5-O1 and T6-O2) recorded from a healthy subject is shown in Fig. 1 and a patient with absence seizure epileptic discharge is shown in Fig. 2.

Currently, analysis of the recorded EEG data is performed primarily by neurologists through visual inspection. Most studies on the characteristics of the 3-Hz spike and slow wave complex have been based on simple visual inspection of data recorded for different channels. EEG signals for both healthy and unhealthy cases were recorded from subjects under relaxation, with their eyes closed. The recording conditions followed Guideline 7 of the American EEG Society and electrodes were placed according to the International 10–20 system. The signals were digitized and transferred to the PC using 12-bit AD converter, storage-sampling rate at 200 Hz.

Two neurologists with experience in the clinical analysis of EEG signals separately inspected every recording included in this study to score epileptic and normal signals. Each event was filed on the computer memory and linked to the tracing with its start and duration. These were then revised by the two experts jointly to solve disagreements and set up the training set for the program, consenting to the choice of threshold for the epileptic seizure detection. The agreement between the two experts was evaluated as the rate between the numbers of epileptic seizures detected by both experts. When revising this unified event set, the human experts, by mutual consent, marked each state as epileptic or normal. They also reviewed each recording entirely for epileptic seizures that had been overlooked by all during the first pass and marked them as definite or possible. Nevertheless, a preliminary analysis was carried out solely on events in the whole set, as each stage in these sets had a definite start and duration.

2.2. Cramer–Rao bound

Since the parameter estimates which are obtained by the estimators having lower variance will be close to the actual values, the parameter estimation method having the lowest variance should be selected for parameter estimation. Cramer–Rao bound can be defined as selection of the estimation method having the lowest variance. Since all information is in material form in the observed data and the underlying probability density function (PDF) for the data, the estimator accuracy depends directly on the PDF. In determination of CRB, PDF of the observed data is defined as the function of the unknown parameter and is referred to as likelihood function: $p(x;\theta)$, where θ denotes the vector of unknown parameters ($\theta=[\theta_1\ \ \theta_2\ \ \cdots\ \ \theta_p]^{\mathrm{T}}$). Then the log-likelihood function is determined. To obtain the CRB, the well-known formula which states that the elements of the Fisher information matrix is used,

$$[I(\theta)]_{ij} = -E\left[\frac{\partial^2 \ln p(x;\theta)}{\partial\theta_i \partial\theta_j}\right], \quad i = 1, 2, \ldots, p, \quad j = 1, 2, \ldots, p, \tag{1}$$

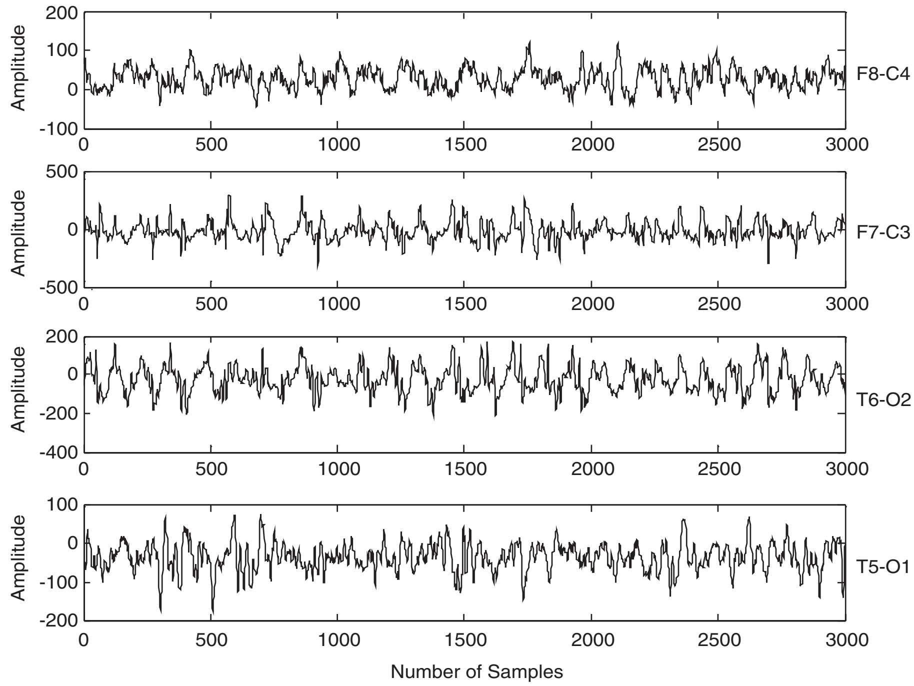

Fig. 1. EEG signal taken from a healthy subject.

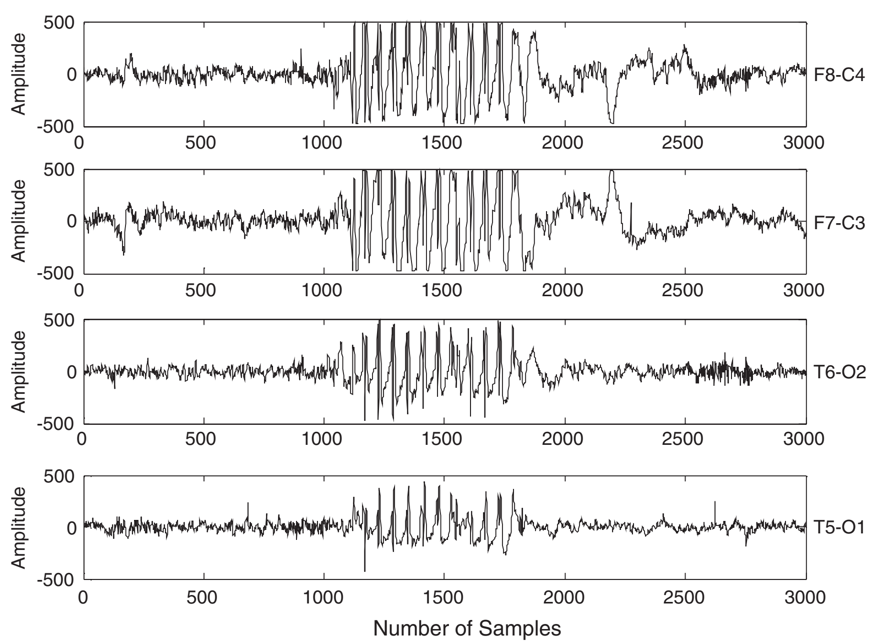

Fig. 2. EEG signal taken from an unhealthy subject (epileptic patient).

where $I(\theta)$ is Fisher information matrix with the dimension of $p \times p$.

The CRB is the inverse of the Fisher information matrix,

$$\mathrm{var}(\hat{\theta}_i) \geqslant [I^{-1}(\theta)]_{ii}. \tag{2}$$

Thus, to evaluate the Fisher information matrix, the derivatives of the log-likelihood function are computed with respect to the various parameters of interest and their expected values are taken [10,13–15].

2.3. AR method for spectral analysis

The model-based (parametric) methods are based on modeling the data sequence $x(n)$ as the output of a linear system characterized by a rational structure. In the model-based methods, the spectrum estimation procedure consists of two steps. The parameters of the model-based method are estimated from a given data sequence $x(n)$, $0 \le n \le N-1$. Then, the PSD estimate is computed from these estimates. Since the estimation of AR parameters can be done easily by solving linear equations, the AR method is the most frequently used parametric method. In the AR method, data can be modeled as output of a causal, all-pole, discrete filter whose input is white noise. The AR method of order p is expressed as the following equation:

$$x(n) = -\sum_{k=1}^{p} a(k)x(n-k) + w(n), \tag{3}$$

where $a(k)$ are the AR coefficients and $w(n)$ is white noise of variance equal to σ^2. The AR(p) model can be characterized by the AR parameters $\{a[1], a[2], \ldots, a[p], \sigma^2\}$. The PSD is

$$P_{\mathrm{AR}}(f) = \frac{\sigma^2}{|A(f)|^2}, \tag{4}$$

where $A(f) = 1 + a_1 \mathrm{e}^{-\mathrm{j}2\pi f} + \cdots + a_p \mathrm{e}^{-\mathrm{j}2\pi f p}$.

In order to obtain stable and most suitable AR method, some factors must be taken into consideration such as selection of the model order, the length of the signal which will be modeled, and the level of stationary of the data [5–16].

Because of the good performance and the computational efficiency of the AR spectral estimation methods, a lot of estimation methods are widely used in practice. The AR spectral estimation methods are based on estimation of either the AR parameters or the reflection coefficients. Apart from the MLE which is based on maximizing the likelihood function, all the model based estimation techniques estimate the parameters by minimizing an estimate of the prediction error power.

2.3.1. Yule-Walker method

In the Yule-Walker method, the AR parameters are estimated by minimizing an estimate of prediction error power,

$$\hat{\rho} = \frac{1}{N} \sum_{n=-\infty}^{\infty} \left| x(n) + \sum_{k=1}^{p} a(k)x(n-k) \right|^2. \tag{5}$$

The samples of the $x(n)$ process which are not observed (i.e., those not in the range $0 \le n \le N-1$) are set equal to zero in Eq. (5). The estimated prediction error power is minimized by differentiating Eq. (5) with respect to the real and imaginary parts of the $a(k)$'s. This may be done by using the complex gradient to yield

$$\frac{1}{N} \sum_{n=-\infty}^{\infty} \left(x(n) + \sum_{k=1}^{p} a(k)x(n-k) \right) x^*(n-l) = 0, \quad l = 1, 2, \ldots, p. \tag{6}$$

In matrix form this set of equations becomes

$$\begin{bmatrix} \hat{r}(1) \\ \vdots \\ \hat{r}(p) \end{bmatrix} + \begin{bmatrix} \hat{r}(0) & \cdots & \hat{r}(-p+1) \\ \vdots & \ddots & \vdots \\ \hat{r}(p-1) & \cdots & \hat{r}(0) \end{bmatrix} \begin{bmatrix} \hat{a}(1) \\ \vdots \\ \hat{a}(p) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

or

$$\hat{r}_p + \hat{R}_p \hat{a} = 0, \tag{7}$$

where

$$\hat{r}(k) = \begin{cases} \dfrac{1}{N} \displaystyle\sum_{n=0}^{N-1-k} x^*(n) \times x(n+k), & k = 0, 1, \ldots, p, \\ \hat{r}^*(-k), & k = (-p+1), (-p+2), \ldots, -1. \end{cases}$$

From Eq. (7) the AR parameter estimates are found as

$$\hat{a} = -\hat{R}_p^{-1} \hat{r}_p. \tag{8}$$

The estimate of the white noise variance σ^2 is found as $\hat{\rho}_{\min}$, which is given by

$$\hat{\sigma}^2 = \hat{\rho}_{\min} = \frac{1}{N} \sum_{n=-\infty}^{\infty} \left| x(n) + \sum_{k=1}^{p} \hat{a}(k)x(n-k) \right|^2. \tag{9}$$

The final result is found by using Eq. (6),

$$\hat{\sigma}^2 = \hat{r}(0) + \sum_{k=1}^{p} \hat{a}(k)\hat{r}(-k). \tag{10}$$

From the estimates of the AR parameters, PSD estimation is formed as [16–18]

$$\hat{P}_{\mathrm{YW}}(f) = \frac{\hat{\sigma}^2}{\left| 1 + \sum_{k=1}^{p} \hat{a}(k) \mathrm{e}^{-\mathrm{j}2\pi f k} \right|^2}. \tag{11}$$

2.3.2. Covariance method

For complex data, a similar estimator may be found by minimizing the estimate of the prediction error power,

$$\hat{\rho} = \frac{1}{N-p} \sum_{n=p}^{N-1} \left| x(n) + \sum_{k=1}^{p} a(k)x(n-k) \right|^2. \tag{12}$$

The only difference between the covariance method and the autocorrelation method is the range of summation in the prediction error power estimate. In the covariance method all the data points needed to be computed from observed $\hat{\rho}$. It is not necessary to take the some part of the data equal to zero. The minimization of Eq. (12) may be effected by applying the complex gradient to yield the AR parameter estimates as the solution of the equations,

$$\begin{bmatrix} c(1,0) \\ \vdots \\ c(p,0) \end{bmatrix} + \begin{bmatrix} c(1,1) & \cdots & c(1,p) \\ \vdots & \ddots & \vdots \\ c(p,1) & \cdots & c(p,p) \end{bmatrix} \begin{bmatrix} \hat{a}(1) \\ \vdots \\ \hat{a}(p) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

or

$$c_p + C_p \hat{a} = 0, \tag{13}$$

where

$$c(\mathrm{j}, k) = \frac{1}{N-p} \sum_{n=p}^{N-1} x^*(n-\mathrm{j})x(n-k).$$

From Eq. (13) the AR parameter estimates are found as

$$\hat{a} = -C_p^{-1} c_p. \tag{14}$$

The white noise variance is estimated as

$$\hat{\sigma}^2 = \hat{\rho}_{\min} = c(0,0) + \sum_{k=1}^{p} \hat{a}(k) c(0,k). \tag{15}$$

From the estimates of the AR parameters, PSD estimation is formed as [10,16–18]

$$\hat{P}_{\mathrm{COV}}(f) = \frac{\hat{\sigma}^2}{\left|1 + \sum_{k=1}^{p} \hat{a}(k) \mathrm{e}^{-\mathrm{j}2\pi f k}\right|^2}. \tag{16}$$

2.3.3. Modified covariance method

For an AR(p) process the optimal forward predictor is

$$\hat{x}(n) = -\sum_{k=1}^{p} a(k) x(n-k), \tag{17}$$

while the optimal backward predictor is

$$\hat{x}(n) = -\sum_{k=1}^{p} a^*(k) x(n+k), \tag{18}$$

where the $a(k)$'s are the AR parameters. In each case the minimum prediction error power is just the white noise variance σ^2. The modified covariance method estimates the AR parameters by minimizing the average of the estimated forward and backward prediction error powers,

$$\hat{\rho} = \tfrac{1}{2}(\hat{\rho}^f + \hat{\rho}^b), \tag{19}$$

where

$$\hat{\rho}^f = \frac{1}{N-p} \sum_{n=p}^{N-1} \left| x(n) + \sum_{k=1}^{p} a(k) x(n-k) \right|^2,$$

$$\hat{\rho}^b = \frac{1}{N-p} \sum_{n=0}^{N-1-p} \left| x(n) + \sum_{k=1}^{p} a^*(k) x(n+k) \right|^2.$$

As in the case of covariance method, the summations are more than the prediction errors that involve observed data samples. Minimization of Eq. (19) can be done by applying the complex gradient to yield

$$\begin{aligned} \frac{\partial \hat{\rho}}{\partial a(l)} = \frac{1}{N-p} &\left[\sum_{n=p}^{N-1} \left(x(n) + \sum_{k=1}^{p} a(k) x(n-k) \right) x^*(n-l) \right. \\ &\left. + \sum_{n=0}^{N-1-p} \left(x^*(n) + \sum_{k=1}^{p} a(k) x^*(n+k) \right) x(n+l) \right] \\ = 0, \quad & l = 1, 2, \ldots, p. \end{aligned} \tag{20}$$

After some simplification, the equation becomes

$$\begin{aligned} \sum_{k=1}^{p} \hat{a}(k) &\left(\sum_{n=p}^{N-1} \times x(n-k) x^*(n-l) \right. \\ &\left. + \sum_{n=0}^{N-1-p} x^*(n+k) x(n+l) \right) \\ &= -\left(\sum_{n=p}^{N-1} x(n) x^*(n-l) + \sum_{n=0}^{N-1-p} x^*(n) x(n+l) \right), \\ l = 1, 2, \ldots, p \end{aligned} \tag{21}$$

or in matrix form,

$$\begin{bmatrix} c(1,0) \\ \vdots \\ c(p,0) \end{bmatrix} + \begin{bmatrix} c(1,1) & \cdots & c(1,p) \\ \vdots & \ddots & \vdots \\ c(p,1) & \cdots & c(p,p) \end{bmatrix} \begin{bmatrix} \hat{a}(1) \\ \vdots \\ \hat{a}(p) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad c_p + C_p \hat{a} = 0, \tag{22}$$

where

$$c(\mathrm{j},k) = \frac{1}{2(N-p)} \left(\sum_{n=p}^{N-1} x^*(n-\mathrm{j}) x(n-k) + \sum_{n=0}^{N-1-p} x(n+\mathrm{j}) x^*(n+k) \right).$$

From Eq. (22) the AR parameter estimates are found as

$$\hat{a} = -C_p^{-1} c_p. \tag{23}$$

The estimate of the white noise variance is

$$\begin{aligned} \hat{\sigma}^2 = \hat{\rho}_{\min} = \frac{1}{2(N-p)} &\left[\sum_{n=p}^{N-1} \left(x(n) + \sum_{k=1}^{p} \hat{a}(k) x(n-k) \right) \right. \\ &\left. \times x^*(n) + \sum_{n=0}^{N-1-p} \left(x^*(n) + \sum_{k=1}^{p} \hat{a}(k) x^*(n+k) \right) x(n) \right], \end{aligned}$$

where Eq. (20) has been used, and finally,

$$\hat{\sigma}^2 = c(0,0) + \sum_{k=1}^{p} \hat{a}(k) c(0,k). \tag{24}$$

It is observed that the modified covariance method is identical to the covariance except for the definition of $c(\mathrm{j},k)$, the autocorrelation estimator. From the estimates of the AR parameters, PSD estimation is formed as [10,16–18]

$$\hat{P}_{\mathrm{MCOV}}(f) = \frac{\hat{\sigma}^2}{\left|1 + \sum_{k=1}^{p} \hat{a}(k) \mathrm{e}^{-\mathrm{j}2\pi f k}\right|^2}. \tag{25}$$

2.3.4. Burg method

The Burg method is based on minimization of the forward and backward prediction errors and estimation of the reflection coefficient. The forward and backward prediction errors for a pth-order model are defined as

$$\hat{e}_{f,p}(n) = x(n) + \sum_{i=1}^{p} \hat{a}_{p,i} x(n-i), \quad n = p+1, \ldots, N, \tag{26}$$

$$\hat{e}_{b,p}(n) = x(n-p) + \sum_{i=1}^{p} \hat{a}^{*}_{p,i} x(n-p+i), \quad n = p+1, \ldots, N. \tag{27}$$

The AR parameters are related to the reflection coefficient $\hat{k}_p$ can be denoted as

$$\hat{a}_{p,i} = \begin{cases} \hat{a}_{p-1,i} + \hat{k}_p \hat{a}^{*}_{p-1,p-i}, & i = 1, \ldots, p-1, \\ \hat{k}_p, & i = p. \end{cases} \tag{28}$$

The Burg method considers the recursive-in-order estimation of $\hat{k}_p$ given that the AR coefficients for order $p-1$ have been computed. The reflection coefficient estimate is given by

$$\hat{k}_p = \frac{-2\sum_{n=p+1}^{N} \hat{e}_{f,p-1}(n)\hat{e}^{*}_{b,p-1}(n-1)}{\sum_{n=p+1}^{N} [|\hat{e}_{f,p-1}(n)|^2 + |\hat{e}_{b,p-1}(n-1)|^2]}. \tag{29}$$

The prediction errors satisfy the following recursive-in-order expressions:

$$\hat{e}_{f,p}(n) = \hat{e}_{f,p-1}(n) + \hat{k}_p \hat{e}_{b,p-1}(n-1), \tag{30}$$

$$\hat{e}_{b,p}(n) = \hat{e}_{b,p-1}(n-1) + \hat{k}^{*}_p \hat{e}_{f,p-1}(n) \tag{31}$$

and these expressions are used to develop a recursive-in-order algorithm for estimating the AR coefficients. From the estimates of the AR parameters, PSD estimation is formed as [10,16–18]

$$\hat{P}_{\mathrm{BURG}}(f) = \frac{\hat{e}_p}{\left|1 + \sum_{k=1}^{p} \hat{a}_p(k) \mathrm{e}^{-\mathrm{j}2\pi f k}\right|^2}, \tag{32}$$

where $\hat{e}_p = \hat{e}_{f,p} + \hat{e}_{b,p}$ is the total least squares error.

2.3.5. Least squares method

Linear prediction of the AR method is to predict the unobserved data sample $x(n)$ based on the observed data samples $\{x(n-1), x(n-2), \ldots, x(n-p)\}$,

$$\hat{x}(n) = -\sum_{k=1}^{p} \alpha_k x(n-k), \tag{33}$$

the prediction coefficients $\{\alpha_1, \alpha_2, \ldots, \alpha_p\}$ are chosen to minimize the power of the prediction error $e(n)$:

$$\rho = E\{|e(n)|^2\} = E\{|x(n) - \hat{x}(n)|^2\}. \tag{34}$$

For minimizing ρ the orthogonality principle is used,

$$r(k) = -\sum_{l=1}^{p} \alpha_l r(k-l), \quad k = 1, 2, \ldots, p, \tag{35}$$

$$\rho_{\min} = r(0) + \sum_{k=1}^{p} \alpha_k r(-k), \tag{36}$$

where $\alpha_k = a[k]$ for $k = 1, 2, \ldots, p$ and $\rho_{\min} = \sigma^2$.

Given a finite set of data samples $\{x(n)\}_{n=1}^{N}$ minimum of $E\{|e(n)|^2\}$ is calculated with respect to α_k $(k = 1, 2, \ldots, p)$.

$$\begin{aligned} f(\alpha) = E\{|e(n)|^2\} &= \sum_{n=N_1}^{N_2} |e(n)|^2 \\ &= \sum_{n=N_1}^{N_2} \left| x(n) + \sum_{k=1}^{p} \alpha[k] x(n-k) \right|^2, \quad k = 1, 2, \ldots, p \\ &= \left\| \begin{bmatrix} x(N_1) \\ x(N_1+1) \\ \vdots \\ x(N_2) \end{bmatrix} + \begin{bmatrix} x(N_1-1) & \cdots & \cdots & x(N_1-p) \\ x(N_1) & \cdots & \cdots & x(N_1+1-p) \\ \vdots & & & \vdots \\ x(N_2-1) & \cdots & \cdots & x(N_2-p) \end{bmatrix} \alpha \right\|^2 \\ &= \|x + X\alpha\|^2. \end{aligned} \tag{37}$$

The vector α that minimizes $f(\alpha)$ is given by

$$\hat{\alpha} = -(X^{*}X)^{-1}(X^{*}x), \tag{38}$$

As seen from Eq. (37) the definitions of X and x depend on the choice of (N_1, N_2). The most common choices N_1 and N_2 are: (i) $N_1 = 1$, $N_2 = N + p$ and this choice yields the Yule-Walker method; (ii) $N_1 = p + 1$, $N_2 = N$ and this choice of (N_1, N_2) is yields the covariance method.

By substitution, autocorrelation function estimates $\{\hat{r}(k)\}_{k=0}^{p}$ and $\hat{\alpha}$ in Eq. (36) $\hat{\rho}_{\min}$ are obtained,

$$\hat{\rho}_{\min} = \hat{r}(0) + \sum_{k=1}^{p} \hat{\alpha} \hat{r}(-k). \tag{39}$$

From the estimates of the AR parameters, PSD estimation is formed as [10,16–18]

$$\hat{P}_{LS}(f) = \frac{\hat{\rho}_{\min}}{\left|1 + \sum_{k=1}^{p} \hat{a}_p(k) \mathrm{e}^{-\mathrm{j}2\pi f k}\right|^2}. \tag{40}$$

2.3.6. Maximum likelihood estimation method

If the MLE of a parameter exists under regular conditions, it is consistent, asymptotically unbiased, efficient, and normally distributed. Likelihood function of $\{x \sim N(0, C(\theta))\}$ Gaussian random process is expressed as

$$p(x;\theta) = \frac{1}{(2\pi)^{N/2} \det^{1/2}(C(\theta))} \exp\left[-\frac{1}{2} x^{\mathrm{T}} C^{-1}(\theta) x\right]. \tag{41}$$

The logarithm of Eq. (41) equals to log-likelihood function,

$$\ln p(x;\theta) = -\frac{N}{2}\ln 2\pi - \frac{N}{2}\int_{-1/2}^{1/2}\left[\ln P(f) + \frac{I(f)}{P(f)}\right]\mathrm{d}f, \tag{42}$$

where $I(f)$ is periodogram of the data,

$$I(f) = \frac{1}{N}\left|\sum_{n=0}^{N-1} x(n)\exp(-\mathrm{j}2\pi f n)\right|^2.$$

The MLE of θ is obtained by calculating the maximum of Eq. (42). After required calculations and derivations, the estimated autocorrelation function is obtained as the following:

$$\hat{r}(k) = \begin{cases} \dfrac{1}{N}\displaystyle\sum_{n=0}^{N-1-|k|} x(n)x(n+|k|), & |k| \leqslant N-1, \\ 0, & |k| \geqslant N. \end{cases} \tag{43}$$

The set of equations to be solved for the MLE of AR parameters,

$$\sum_{l=1}^{p} \hat{a}(l)\hat{r}(k-l) = -\hat{r}(k), \quad k = 1, 2, \ldots, p,$$

or in matrix form

$$\begin{bmatrix} \hat{r}(0) & \hat{r}(1) & \cdots & \hat{r}(p-1) \\ \hat{r}(1) & \hat{r}(0) & \cdots & \hat{r}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{r}(p-1) & \hat{r}(p-2) & \cdots & \hat{r}(0) \end{bmatrix} \begin{bmatrix} \hat{a}(1) \\ \hat{a}(2) \\ \vdots \\ \hat{a}(p) \end{bmatrix} = -\begin{bmatrix} \hat{r}(1) \\ \hat{r}(2) \\ \vdots \\ \hat{r}(p) \end{bmatrix}. \tag{44}$$

Eq. (44) is equal to the estimated Yule-Walker equations and the MLE of AR parameters are calculated from this equation. Then the MLE of σ^2 is found,

$$\hat{\sigma}^2 = \hat{r}(0) + \sum_{k=1}^{p} \hat{a}(k)\hat{r}(k). \tag{45}$$

These estimated parameters are used to compute the AR PSD as [8,10,16–18]

$$\hat{P}_{\mathrm{MLE}}(f) = \frac{\hat{\sigma}^2}{\left|1 + \sum_{k=1}^{p}\hat{a}(k)\mathrm{e}^{-\mathrm{j}2\pi f k}\right|^2}. \tag{46}$$

2.3.7. Selection of AR model orders

One of the most important aspects of the model-based methods is the selection of the model order. Much work has been done by various researchers on this problem and many experimental results have been given in the literature [16–18]. One of the better known criteria for selecting the model order has been proposed by Akaike [19], called the Akaike information criterion (AIC), and is based on selecting the order that minimizes Eq. (47) for the AR method:

$$\mathrm{AIC}(p) = \ln\hat{\sigma}^2 + 2p/N, \tag{47}$$

where $\hat{\sigma}^2$ is the estimated variance of the linear prediction error. Note that the term $\hat{\sigma}^2$ decreases and therefore $\ln\hat{\sigma}^2$ also decreases as the order of the AR method is increased. However, in Eq. (47) $2p/N$ increases with an increase in p. In this situation, a minimum value is obtained for some p in Eq. (47) [10,19] and in this study, model order of the AR method was taken as 9 by using Eq. (47).

3. Results and discussion

Diagnosing epilepsy is a difficult task requiring observation of the patient, an EEG, and gathering of additional clinical information. In this work, we have proposed different model-based AR methods to compute PSDs of EEG signals. The EEG signals are usually interpreted by examining their spectral content. Diagnosis and disease monitoring are assessed by analysis of spectral shape and parameters. During evaluation of epilepsy problems, frequency content and bandwidth parameters can be used for the detection of an epileptic seizure. EEG power spectra describe the distribution of power with frequency. Therefore, it is important to determine the suitability of the available spectral estimation methods for the EEG signals. EEG power spectra were obtained by using different AR methods. Then EEG spectra were used to analyze and characterize epileptiform discharges in the form of 3-Hz spike and wave complexes in patients with absence seizures. PSDs of EEG signals for healthy and unhealthy (epileptic patient) subjects are presented in Figs. 3 and 4, respectively.

In Fig. 3, power spectrums of an EEG signal taken from a healthy person are given. If these spectrums are examined visually, delta activity, alpha activity, and beta activity can be seen easily. These results are true because it is a normal EEG signal. Fig. 4 shows power spectrum of an EEG signal taken from unhealthy person. If these frequency spectrums are examined, it is seen that there are peaks at low frequency range. Since the signal is taken from an epileptic patient, the results fit with the typical characteristics of epilepsy.

In AR parametric methods, model for the signal generation can be constructed with a number of parameters that can be estimated from the observed data. From the model and the estimated parameters, PSD can be computed. AR PSD estimation methods may model spectra with narrow peaks by placing zeroes of the A-polynomial close to the unit circle. The estimation of parameters in the AR signal models is a well-established topic; the estimates are found by solving linear equations of the system. Since the estimated parameters differ according to the estimation methods, the estimated PSDs become different. The estimated EEG spectra were then used to compare the applied AR spectral estimation methods in terms of their frequency resolution and the effects in determination of epileptic seizure. However, from Fig. 3, it is seen that the PSD estimations of healthy subject obtained by different AR spectral estimation methods produce similar spectral characteristics with nearly identical peak frequencies. From Fig. 4, it is apparent that the PSD estimations of unhealthy subject (epileptic patient) obtained by different AR spectral estimation methods differ slightly from each other. If we compare the PSDs of

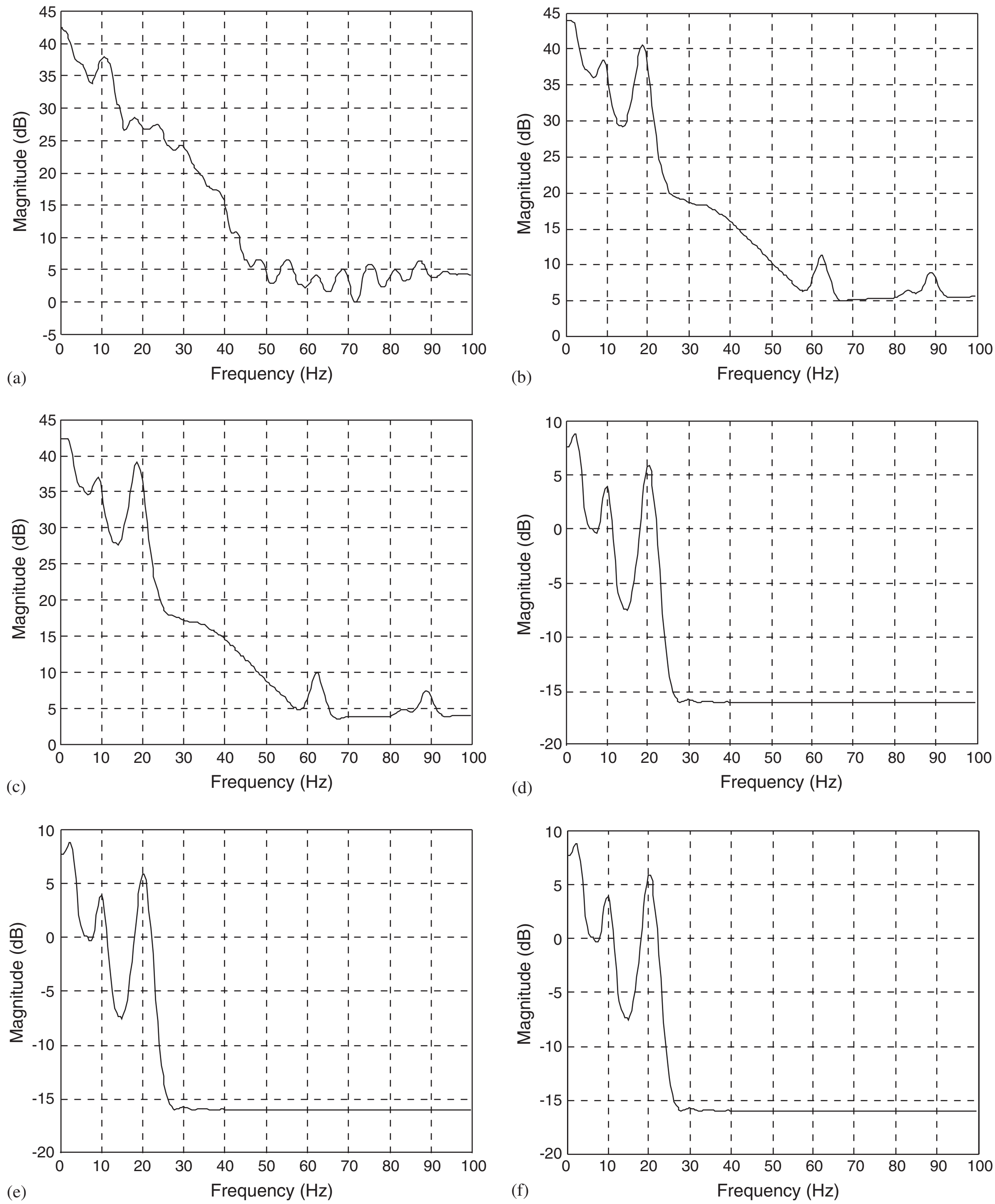

Fig. 3. PSDs of EEG signal taken from healthy subjects: (a) Yule-Walker AR; (b) covariance AR; (c) modified covariance AR; (d) Burg AR; (e) least squares AR; (f) MLE AR methods.

healthy and unhealthy subjects, it can be seen easily that, the PSDs of healthy subjects contains delta, alpha and beta activity; but PSDs of unhealthy subject's shows typical characteristics of epileptic seizure (low frequency component).

The AR spectral estimation methods were compared with the use of statistical tools such as correlation coefficients (r). The correlation coefficients between the AR spectral estimation methods were calculated with a statistical package (SPSS version 10.0). EEG PSD values were used for the calculation of the correlation coefficients. The correlation coefficient is limited with the range $[-1, 1]$. When $r = 1$ there is a perfect positive linear correlation between the two methods' PSD values, which means that they vary by the same amount. When $r = -1$ there is a perfectly linear negative correlation between the two methods'

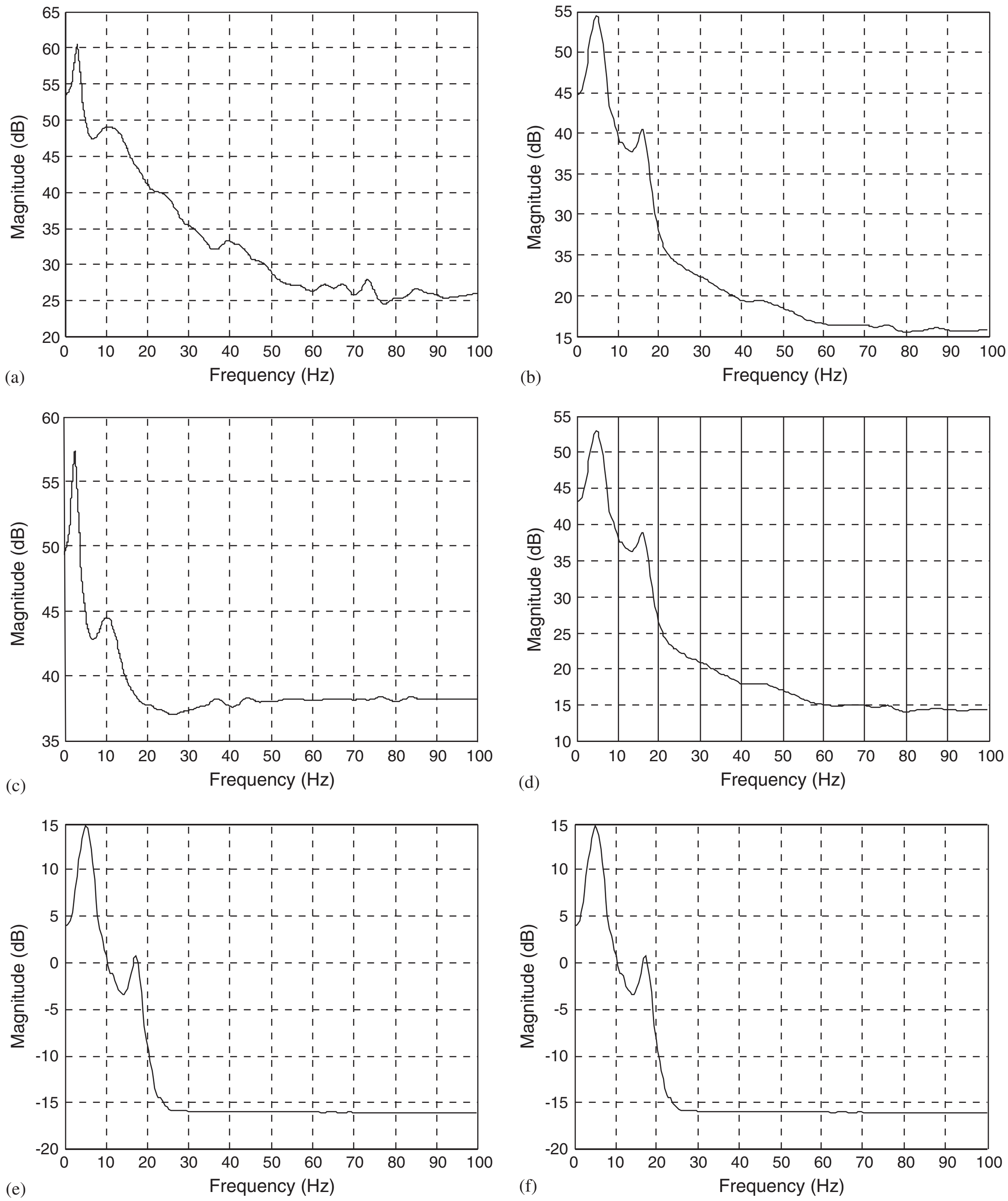

Fig. 4. PSDs of EEG signal taken from unhealthy subjects (epileptic patients): (a) Yule-Walker AR; (b) covariance AR; (c) modified covariance AR; (d) Burg AR; (e) least squares AR; (f) MLE AR methods.

PSD values, that means they vary in opposite ways (when one of the method's PSD values increase, the other method's PSD values decrease by the same amount). When $r = 0$ there is no correlation between the two methods' PSD values (the values are called uncorrelated). Intermediate values describe partial correlations. The correlation coefficients between AR parametric methods were calculated from EEG PSD values of healthy and unhealthy (petit mal) subjects. The calculated r values for healthy subject are varying in the range $[0.991, 0.994]$ and show that there are perfect positive linear correlations among the PSD values. The calculated r values for unhealthy subject are varying in the range $[0.992, 0.996]$ and indicate that there are perfect positive linear correlations among the PSD values. According to these correlation values, there is a perfect positive linear correlation between PSD values of AR parametric methods. However, from Figs. 3 and 4, it is apparent that the AR

Table 1
Mean of the estimated AR parameter values of EEG signals obtained by different AR spectral estimation methods for healthy subjects

Parameters	Estimated values					
	Yule-Walker AR	Covariance AR	Modified covariance AR	Burg AR	Least squares AR	MLE AR
a(1)	−2.0170	−2.0263	−2.0371	−2.0462	−2.0511	−2.0630
a(2)	1.4755	1.4785	1.4866	1.4951	1.4974	1.5542
a(3)	−0.4148	−0.4281	−0.4481	−0.4497	−0.4521	−0.4589
a(4)	−0.2744	−0.2787	−0.2943	−0.2984	−0.2995	−0.3144
a(5)	0.7928	0.7986	0.8067	0.8134	0.8185	0.8257
a(6)	−0.6861	−0.6971	−0.6992	−0.7132	−0.7212	−0.7415
a(7)	−0.4485	−0.4565	−0.4594	−0.4642	−0.4687	−0.4851
a(8)	1.6007	1.6301	1.6379	1.6472	1.6497	1.6584
a(9)	−1.6372	−1.6452	−1.6479	−1.6577	−1.6592	−1.6664
σ^2	3.1719	2.8251	2.4235	2.1562	1.9467	1.6768

Table 2
Mean of the estimated AR parameter values of EEG signals obtained by different AR spectral estimation methods for unhealthy subjects (epileptic patients)

Parameters	Estimated values					
	Yule-Walker AR	Covariance AR	Modified covariance AR	Burg AR	Least squares AR	MLE AR
a(1)	−1.5863	−1.5957	−1.5989	−1.6063	−1.6345	−1.6586
a(2)	0.5187	0.6182	0.6197	0.6245	0.6432	0.6518
a(3)	−0.0640	−0.1624	−0.1740	−0.1887	−0.2642	−0.3064
a(4)	−0.0551	−0.1255	−0.1585	−0.1651	−0.1777	−0.1850
a(5)	0.1825	0.1982	0.2587	0.3225	0.4781	0.5005
a(6)	−0.0104	−0.0241	−0.0604	−0.0711	−0.0842	−0.1201
a(7)	0.3461	0.3587	0.3619	0.3986	0.4619	0.4981
a(8)	−0.2676	−0.2768	−0.2997	−0.3254	−0.3676	−0.3967
a(9)	−0.1223	−0.1327	−0.1541	−0.1652	−0.1723	−0.1914
σ^2	4.1723	4.0024	3.8571	3.5112	2.9115	2.5261

methods' EEG PSDs are similar to each other. In this situation, it is difficult to compare the AR spectral estimation methods' peak frequencies and power levels of the EEG PSDs for healthy subjects and unhealthy subjects (epileptic patients).

Estimation accuracy is often measured with the help of CRB, which is computed for the variance of parameter estimates. The performance comparisons of the AR spectral estimation methods were performed by using CRB and the optimal AR spectral estimation method was selected for the EEG signals. CRBs for the AR parameters were derived with the use of the Fisher information matrix [10].

The estimated AR parameter values of the EEG signals, which were obtained by different AR spectral estimation methods, for healthy subjects and unhealthy subjects (epileptic patients) are given in Tables 1 and 2, respectively. These parameters are the mean estimated parameters of the whole data set. The variance of the estimated AR parameters of the EEG signals was computed by using the derived CRB expressions. The CRBs of the estimated AR parameters of the EEG signals, which were obtained by different AR spectral estimation methods, for healthy subject and unhealthy subject (epileptic patient) are given in Tables 3 and 4, respectively. These CRBs of the estimated parameters are also the mean of the whole estimated data set. From Table 3, it is seen that the CRB values of all AR parameters, which are obtained by the MLE AR spectral estimation method are the lowest for healthy subject's EEG signals. From Table 4, it is seen that the CRB values of all AR parameters, which are obtained by the MLE AR spectral estimation method are the lowest for unhealthy subject's EEG signals.

4. Conclusion

Careful analyses of the EEG records can provide valuable insight and improved understanding of the mechanisms causing epileptic disorders. The detection of epileptiform discharges in the EEG is an important component in the diagnosis of epilepsy. It is important to determine the optimal spectral estimation method for the EEG signals, since clinically useful information can be extracted from EEG power spectrum. Spectral analysis of the EEG signals was performed using AR methods. The parameters of the AR methods were estimated by estimation methods such as Yule-Walker, covariance, modified covariance, Burg, least squares, and MLE. Interpretation and performance of these estimation methods were compared in terms of their frequency resolution and the effects in epileptic

Table 3
Mean of the CRBs of the estimated AR parameter values of EEG signals obtained by different AR spectral estimation methods for healthy subjects

Parameters	CRB values					
	Yule-Walker AR	Covariance AR	Modified covariance AR	Burg AR	Least squares AR	MLE AR
a(1)	2.9689×10^{-6}	2.6737×10^{-6}	2.3362×10^{-6}	2.2954×10^{-6}	2.0673×10^{-6}	1.7235×10^{-6}
a(2)	1.4947×10^{-5}	1.4877×10^{-5}	1.4856×10^{-5}	1.4835×10^{-5}	1.4820×10^{-5}	1.4805×10^{-5}
a(3)	1.5436×10^{-5}	1.5419×10^{-5}	1.5379×10^{-5}	1.5354×10^{-5}	1.5333×10^{-5}	1.5312×10^{-5}
a(4)	1.5545×10^{-5}	1.5535×10^{-5}	1.5524×10^{-5}	1.5517×10^{-5}	1.5506×10^{-5}	1.5490×10^{-5}
a(5)	1.6522×10^{-5}	1.6506×10^{-5}	1.6487×10^{-5}	1.6465×10^{-5}	1.6447×10^{-5}	1.6428×10^{-5}
a(6)	1.6564×10^{-5}	1.6545×10^{-5}	1.6531×10^{-5}	1.6524×10^{-5}	1.6510×10^{-5}	1.6487×10^{-5}
a(7)	1.6651×10^{-5}	1.6645×10^{-5}	1.6634×10^{-5}	1.6627×10^{-5}	1.6616×10^{-5}	1.6603×10^{-5}
a(8)	1.6613×10^{-5}	1.6587×10^{-5}	1.6573×10^{-5}	1.6562×10^{-5}	1.6534×10^{-5}	1.6489×10^{-5}
a(9)	1.6647×10^{-5}	1.6636×10^{-5}	1.6628×10^{-5}	1.6620×10^{-5}	1.6612×10^{-5}	1.6602×10^{-5}
σ^2	3.1456×10^{-4}	2.7540×10^{-4}	2.1123×10^{-4}	1.5091×10^{-4}	1.3756×10^{-4}	1.1522×10^{-4}

Table 4
Mean of the CRBs of the estimated AR parameter values of EEG signals obtained by different AR spectral estimation methods for unhealthy subjects (epileptic patients)

Parameters	CRB values					
	Yule-Walker AR	Covariance AR	Modified covariance AR	Burg AR	Least squares AR	MLE AR
a(1)	-1.3650×10^{-4}	-1.3965×10^{-4}	-1.4082×10^{-4}	-1.4092×10^{-4}	-1.4124×10^{-4}	-1.4706×10^{-4}
a(2)	-2.7982×10^{-4}	-3.0180×10^{-4}	-3.1094×10^{-4}	-3.1356×10^{-4}	-3.3969×10^{-4}	-3.5044×10^{-4}
a(3)	-4.0424×10^{-5}	-4.4635×10^{-5}	-5.1180×10^{-5}	-2.0791×10^{-4}	-2.2310×10^{-4}	-1.6782×10^{-4}
a(4)	1.2103×10^{-5}	1.2044×10^{-5}	1.2033×10^{-5}	1.1905×10^{-5}	1.1852×10^{-5}	1.1719×10^{-5}
a(5)	1.4051×10^{-5}	1.3915×10^{-5}	1.3442×10^{-5}	1.2326×10^{-5}	1.0801×10^{-5}	6.5621×10^{-6}
a(6)	-1.6942×10^{-5}	-1.7370×10^{-5}	-1.7550×10^{-5}	-1.7835×10^{-5}	-1.8177×10^{-5}	-1.9849×10^{-5}
a(7)	-2.5427×10^{-6}	-1.3717×10^{-5}	-1.4668×10^{-5}	-1.6940×10^{-5}	-1.7543×10^{-5}	-1.9108×10^{-5}
a(8)	1.2316×10^{-5}	1.2301×10^{-5}	1.2231×10^{-5}	1.2169×10^{-5}	1.2087×10^{-5}	1.1784×10^{-5}
a(9)	2.0440×10^{-5}	7.9820×10^{-6}	7.5932×10^{-6}	7.5282×10^{-6}	7.2051×10^{-6}	6.5032×10^{-6}
σ^2	5.1842×10^{-4}	4.6024×10^{-4}	3.4645×10^{-4}	2.6478×10^{-4}	2.0079×10^{-4}	1.6692×10^{-4}

seizure detection. The variance of the estimated AR parameters of the EEG signals was computed using the derived CRB expressions. According to the computed CRB values, the performances of the AR spectral estimation methods were evaluated and the MLE AR method was selected as the optimal AR spectral estimation method for the EEG signals. In conclusion, it should be emphasized that the performance characteristics of the MLE AR method had been found extremely valuable for the use in epileptic seizure detection.

References

[1] H. Adeli, Z. Zhou, N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, J. Neurosci. Meth. 123 (2003) 69–87.

[2] A. Subasi, E. Erçelebi, Classification of EEG signals using neural network and logistic regression, Comput. Meth. Prog. Biomed. 78 (2005) 87–99.

[3] A. Subasi, E. Erçelebi, A. Alkan, E. Koklukaya, Comparison of subspace-based methods with AR parametric methods in epileptic seizure detection, Comput. Biol. Med. (2005) (In press).

[4] A. Cohen, Biomedical signals: origin and dynamic characteristics; frequency-domain analysis, in: J.D. Bronzino (Ed.), The Biomedical Engineering Handbook, second ed., CRC Press LLC, Boca Raton, FL, 2000.

[5] J. Pardey, S. Roberts, L. Tarassenko, A review of parametric modelling techniques for EEG analysis, Med. Eng. Phys. 18 (1) (1996) 2–11.

[6] J. Muthuswamy, N.V. Thakor, Spectral analysis methods for neurological signals, J. Neurosci. Meth. 83 (1998) 1–14.

[7] M. Akin, M.K. Kiymik, Application of periodogram and AR spectral analysis to EEG signals, J. Med. Sys. 24 (4) (2000) 247–256.

[8] I. Guler, M.K. Kiymik, M. Akin, A. Alkan, AR spectral analysis of EEG signals by using maximum likelihood estimation, Comput. Biol. Med. 31 (2001) 441–450.

[9] M.K. Kiymik, A. Subasi, H.R. Ozcalik, Neural networks with periodogram and autoregressive spectral analysis methods in detection of epileptic seizure, J. Med. Sys. 28 (6) (2004) 511–522.

[10] E.D. Ubeyli, I. Guler, Selection of optimal AR spectral estimation method for internal carotid arterial Doppler signals using Cramer–Rao bound, Comput. Electr. Eng. 30 (2004) 491–508.

[11] A. Isaksson, A. Wennberg, L.H. Zetterberg, Computer analysis of EEG signals with 4 parametric models, Proc. IEEE 69 (4) (1981) 451–461.

[12] S-Y. Tseng, R-C. Chen, F-C. Chong, T-S. Kuo, Evaluation of parametric methods in EEG signal analysis, Med. Eng. Phys. 17 (1995) 71–78.

[13] B. Friedlander, B. Porat, The exact Cramer–Rao bound for Gaussian autoregressive processes, IEEE Trans. Aerospace Electron. Syst. AES-25 (1) (1989) 3–7.

[14] S.M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice-Hall, New Jersey, 1993.

[15] S. Sando, A. Mitra, On the Cramer–Rao bound for model-based spectral analysis, IEEE Signal Process. Lett. 9 (2) (2002) 68–71.

[16] S.M. Kay, S.L. Marple, Spectrum analysis—A modern perspective, Proc. IEEE 69 (11) (1981) 1380–1419.
[17] J.G. Proakis, D.G. Manolakis, Digital Signal Processing, Principles, Algorithms, and Applications, Prentice-Hall, New Jersey, 1996.
[18] P. Stoica, R. Moses, Introduction to Spectral Analysis, Prentice-Hall, New Jersey, 1997.
[19] H. Akaike, A new look at the statistical model identification, IEEE Trans. Autom. Control AC-19 (1974) 716–723.

Abdulhamit Subasi graduated from Hacettepe University in 1990. He took his M.Sc. degree from Middle East Technical University in 1993, and his Ph.D. degree from Sakarya University in 2001, all in Electronics Engineering. He has been working as an Assistant Professor at Kahramanmaras Sutcu Imam University from 2001. Now, he is a visiting scholar at Georgia Institute of Technology. His areas of interest are application of neural networks, biomedical signal processing, computer networks and security.