Validation of time-frequency and ARMA feature extraction methods in classification of mild epileptic signal patterns

V. Sakkalis, M. Zervakis, C. Bigan, T. Cassar, K.P. Camilleri, S.G. Fabri, and Sifis Micheloyannis

Abstract—**Epilepsy is one of the most common brain disorders and may result in brain dysfunction and cognitive disturbances. Epileptic seizures usually begin in childhood without being accommodated by brain damage and many drugs produce no brain dysfunction. In this study cognitive function in mild epilepsy cases is evaluated where children with seizures are compared to controls i.e., children with epileptic seizures, without brain damage and under drug control. Two different cognitive tasks were designed and performed by both the epileptic and healthy children: i) a relatively difficult math task and ii) Fractal observation. Under this context, we propose two frameworks: the first is based on time-frequency analysis using the continuous wavelet transform (WT) and the Compressed Spectral Array (CSA), and the second is based on Auto-Regressive Moving Average (ARMA) modeling to evaluate the EEG signals at rest and during cognitive tasks in both groups. Initially, the analytical capabilities of the proposed feature extraction techniques were assessed in a simulated environment, and finally classification of the actual data was performed. The results suggest that time-frequency analysis methods were able to capture non-stationary activity, whereas ARMA modeling performs better on stationary signals. Classification of the actual data was successful and both approaches reached the same level of accuracy (73.4%). Higher frequency bands (beta and gamma) were apparent on frontal-parietal lobes on both math and fractal tests, while alpha band was diffused across a wider frontal network, only during the math task.**

I. Introduction

EPILEPSY is one of the most common brain disorders and may result in brain dysfunction and cognitive disturbances. Epileptic seizures usually begin in childhood and evaluation as well as treatment of these children is of importance. Most of the cases in childhood are not accommodated by brain damage and many drugs produce no brain dysfunction. Thus we decided to examine children with seizures and compare them with controls to evaluate cognitive function in mild epilepsy cases i.e. children with epileptic seizures, without brain damage and under drug control. Extensive work has been performed on both feature extraction and classification methods of EEG signals. Either time-varying, mean frequency or other oscillatory-based feature vectors are extracted using autoregressive (AR) models [1], Fourier power spectrum or frequency decomposition methods [2] capable of detecting arousal state changes [3]. Traditional spectral analysis techniques with Fourier transform and more specifically the windowed power spectral density estimation function, known as the periodogram [4], forms the most commonly used analytical tool for spectral representation and evaluation of activity on different EEG frequency bands [5][6]. An improved methodology is based on the time-varying spectral analysis that takes into account the non-stationary dynamics of the neuronal processes [7] The Compressed Spectral Array (CSA) and the wavelet transforms (WT) are representative analysis frameworks of this class. The first approach uses a sliding time window, whereas the second one forms the projection of the signal onto several oscillatory kernel-based wavelets matching different frequency bands. WT is typically applied with the relative bandwidth ($\Delta f/f$) held constant, whereas the Fourier approach preserves the absolute band-width (Δf) constant.

In this study we compare the capabilities of the proposed feature extraction measures by initially testing their properties on a simulated environment and then we investigate their performance in real band-limited signals.

Manuscript received June 30, 2006. This work was supported in part by the EC IST project BIOPATTERN, Contract No: 508803.

V. Sakkalis is with the Department of Electronic and Computer Engineering, Technical University of Crete, Chania 73100 and the Institute of Computer Science, Foundation for Research and Technology, Heraklion 71110, Greece (+30-281-0391448; fax: +30-281-0391428; e-mail: sakkalis@ics.forth.gr).

M. Zervakis is with the Department of Electronic and Computer Engineering, Technical University of Crete, Chania 73100, Greece (e-mail: michalis@systems.tuc.gr).

C. Bigan is with the Ecological University of Bucharest, Bucharest, Romania (e-mail: cbigan@yahoo.com).

T. Cassar, K.P. Camilleri and S.G. Fabri are with the Faculty of Engineering, University of Malta, Msida, Malta (e-mail: {trcass; kpcami; sgfabr}@eng.um.edu.mt).

S. Micheloyannis is with the Clinical Neurophysiology Laboratory (L. Widen), Faculty of Medicine, University of Crete, Heraklion 71409, Greece (e-mail: michelogj@med.uoc.gr).

II. Methods

A. Simulated signals

To study the different properties of each of the proposed methods, we consider a simulated dataset consisting of four, 4 seconds EEG signals with a sampling frequency of 512Hz. The activity within each signal is described in Table I. N is equivalent to 1s signal duration. The simulated channels are depicted in Fig. 1. A control set was also considered to have quasi-white noise in each of the four channels.

TABLE I
SIMULATED DATASET

Channel 1	N	= alpha1 activity
	2N	= Gaussian noise (all bands)
	N	= Gamma1 activity
Channel 2	4N	= Alpha1 activity
Channel 3	2N	= Alpha1 activity
	2N	= Gaussian noise (all bands)
Channel 4	4N	= Gaussian noise (all bands)

B. *Real Data Acquisition*

The studied population consisted of twenty mild epileptic subjects and twenty controls. The EEG signals in both groups (controls and mild epileptics) were recorded from 30 cap electrodes (FP1, FP2, F7, F3, FZ, F4, F8, FT7, FC3, FCZ, FC4, FT8, T3, C3, CZ, C4, T4, TP7, CP3, CPZ, CP4, TP8, P3, PZ, P4, PO7, PO8, O1, OZ and O2), according to the 10/20 international system, referred to linked A1+A2

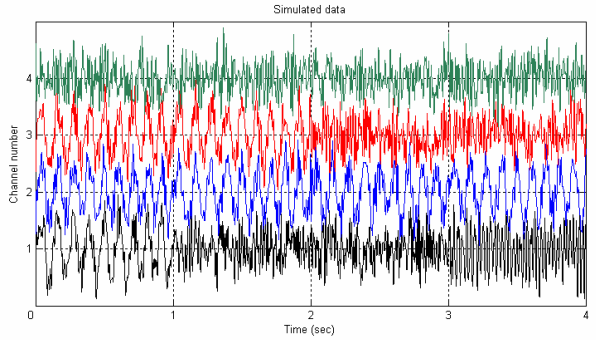

Fig. 1: Simulated dataset resembling EEG activity on different frequency bands.

electrodes. The signals were amplified using a set of Contact Precision Instrument amplifiers, filtered on-line with a band pass between 0.1 and 200 Hz, and digitized at 400 Hz. Off-line, the recorded data were carefully reviewed for technical and biogenic artifacts, so that only artifact free epochs of eight seconds duration are investigated. The procedures used in the study had been previously approved by the University of Crete Institutional Review Board and all subjects signed a consent form after the nature of the procedures involved had been explained to them.

C. *Test Description*

Continuous EEGs were recorded in an electrically shielded, sound and light attenuated room while participants sat in a reclined chair. EEG data were visually inspected for artifacts and epochs of 8 sec were chosen for analysis. We analyzed epochs at rest i.e., while each individual had the eyes fixed on a small point on the computer screen and during two cognitive tasks. The first includes two digits number subtractions or two digits minus one digit, which is thought to be a relatively difficult mathematical task and the second consist of Fractal observation. Stimuli were presented on an LCD screen located in front of the participants. Vertical and horizontal eye movements and blinks were monitored through a bipolar montage from the supraorbital ridge and the lateral canthus.

D. *Compressed Spectral Array (CSA)*

The time frequency visualization method CSA [8] was computed for all channels in both the simulated and real EEG datasets. Compressed spectral array is based on FFT computation of power spectra on successive shifted windows. The spectral result is then plotted on the same axis with a time shift from the bottom to the top.

The FFT used 256 points per 0.5s window, a frequency resolution of 2Hz, and a smoothing factor of 3. A contour plot of CSA is shown in Fig 3, where the power was color coded to a normalized maxima level. The contour plot shows the changes / differences in frequency content and its variation in time.

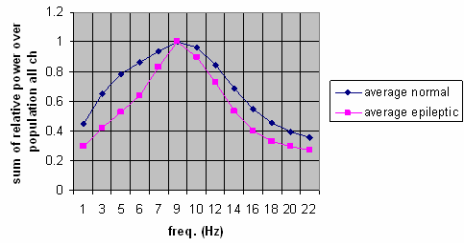

Fig. 2. Average of Spectra on control (upper) and epileptics (lower) subjects

Based on the CSA statistical properties of the 20 normals and 20 epileptics averaged on all channels spectrogram [9] (Fig. 2) and similar to the measures defined in [10], for the purpose of classification the following two features were defined:

$$ind1 = (\Sigma \text{ CSA power delta band} + \Sigma \text{ CSA power in theta}) / (\Sigma \text{ CSA power in Alpha 1}) \quad (1)$$

$$ind2 = (\Sigma \text{ CSA power alpha2 band} + \Sigma \text{ CSA power in beta1}) / (\Sigma \text{ CSA power in Alpha 1}) \quad (2)$$

E. *Statistical feature extraction using the Wavelet Transform (SWT)*

Over the past decade the WT has developed into an important tool for analysis of time series that contain non-stationary power at many different frequencies (such as the EEG signal), as well as a powerful feature extraction method [11][12]. The continuous wavelet transform (CWT) was preferred in this work, so that the time and scale parameters can be considered as continuous variables. In the CWT the notion of scale s is introduced as an alternative to frequency, leading to the so-called time-scale representation domain.

The CWT of a discrete sequence x_n with time spacing δt and N data points (n=0…N-1) is defined as the convolution of x_n with consecutive scaled and translated versions of the wavelet function $\psi_0(\eta)$:

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \psi * \left[(n'-n)\delta t / s \right] \quad (3)$$

$$\psi_0(\eta) = \pi^{-1/4} e^{i\omega_0 \eta} e^{-\eta^2/2} \quad (4)$$

where η and ω_0=6 indicate non-dimensional "time" and frequency parameters, respectively. In our application, $\psi_0(\eta)$ describes the most commonly used wavelet type for spectral analyses, i.e., the *normalized complex Morlet wavelet* given in (4). The wavelet function ψ_0 is a normalized version of ψ that has unit energy at each scale, so that each scale is directly comparable to each other [12]. The power spectrum of the WT is defined by the square of coefficients (1) of the wavelet series as $|W_n(s)|^2$. By adopting the above settings a smooth wavelet power diagram is constructed as in Fig. 4.

TABLE II
FREQUENCY BANDS – SCALE SET MAPPING

Band	Frequency (Hz)	Scale
Theta (θ)	4-8	21, 22, 23, 24
Alpha1 (α1)	8-10	20
Alpha2 (α2)	10-13	18, 19
Beta (β)	13-30	14, 15, 16, 17
Gamma1 (γ1)	30-45	11, 12, 13
Gamma2 (γ2)	45-90	7, 8, 9, 10

As previously noted, there exists a concrete relationship between each scale and an equivalent set of Fourier frequencies, often known as pseudofrequencies. In this study the power spectra is classified in six sequential frequency bands that are coarsely mapped to the scales tabulated in Table II.

The first stage of the feature extraction method is based on capturing the time-averaged power spectrum $\bar{W}_t^2$ for each electrode and scale, which is computed by averaging the power spectrum $|W_n(s)|^2$ over time:

$$\bar{W}_t^2(s) = (1/N) \sum_{n=0}^{N-1} |W_n(s)|^2 \quad (5)$$

Further averaging in scale is performed, in order to map a single feature per frequency band of interest. Thus, the scale-averaged power spectrum $\bar{W}_s^2$ is defined as the weighted sum of the wavelet power spectrum $|W_n(s)|^2$ over scales s_{j_1} to s_{j_2}:

$$\bar{W}_{s,n}^{\ 2} = (\delta j \delta t / C_\delta) \sum_{j=j_1}^{j_2} \left(|W_n(s_j)|^2 / s_j \right) \quad (6)$$

where C_δ is a constant, scale independent factor used for the exact reconstruction of a $\delta(.)$ function from its wavelet transform (for the Morlet wavelet it equals to 0.776) [13].

The aforementioned steps indicate only a significant channel subset, based on task differentiation confidence intervals using Global PS measures. Either synchronization or de-synchronization may be detected compared to the control task. To further refine the features and optimize the classification process, we propose to isolate only those time segments of the EEG signal where notable activity differences occur from the control to the target task. The aim is to further map the EEG signal into a feature vector that best characterizes the EEG pattern of activity for the arithmetic task, in terms of significant temporal and spectral content. As we are interested in ongoing EEG activity within various tasks, the time localization of EEG events is of interest. Notice that we focus on significant (bursty and/or sequential) activations and not on the evolution of brain operation during the task. Thus, we may describe the next step as an attempt to crop up the most significantly different regions from control to mathematics activity out of the bulk initial signal (may be either significant power increase or decrease while performing the requested task compared to the control condition). In fact, this study proposes a way to achieve "crucial" EEG time-segment selection, by testing for significance in the wavelet time domain the "active" task over the control task. The control-task spectrum is defined as the mean time-averaged wavelet power spectrum over all subjects performing the control task.

$$\bar{W}(s) = (1/P) \sum_{p=1}^{P} |W_n^{\ p}(s)|^2 \quad (7)$$

where p is the subject index and $W_n^{\ p}(s)$ may be calculated (3) for each subject. P is the total number of participants. It should be noted that all EEG signals are normalized to zero mean and identity variance. Further rescaling and comparisons are performed using each subject's actual signal variance, in order to include subject-specific information. Significant power increase is calculated using the 95% confidence level at each scale by multiplying the control task spectrum in Eq. 7 by the 95th percentile value for the chi-squared distributed variable χ^2 with two degrees of freedom (DOF) χ^2. This is justified because the wavelet power spectrum uses the Morlet wavelet in a complex product with the signal, so that both the squares of the real and imaginary parts of the result are being χ^2 distributed with one DOF each [13]. In a similar manner, significant power decrease is measured using as upper power limit the 5% confidence level at each scale by multiplying the control task spectrum in Eq. 7 by the 5th percentile value for the chi-squared distributed variable χ^2 with two degrees of freedom (DOF) χ_2^2. Fig. 4 depicts the WT of a simulated signal The significant regions over the time-scale transformed domain that differentiate the two tasks are indicated by the closed contours.

Having derived this significant information, we are now able to form the so-called Significant Power Spectral (Significant PS) features, which are obtained from the signal

energy over those time- and band-localized regions, where apparent significant differentiation is indicated (contours in Fig. 4). For the computation of these features, Eq. (5) is adapted as:

$$\bar{W}_{st}^{\,2}(s) = (1/m)\sum_{m=m_i}^{m_{i+1}} \left|W_m(s)\right|^2, \qquad \mathrm{i} = 1\ldots I \qquad (8)$$

where m is the total number of time points delimited between the boundaries m_i and m_{i+1} of all significant regions I denoted by each contour in Fig. 4 and i is the index of each significant region.

Once the scaled-averaged PS for each of the six EEG bands is calculated for each EEG channel and task, we have a total of 180 (6x30) feature vectors per task (class), representing each participant (subject), which is actually the time-scale averaged PS.

F. Auto-Regressive Moving Average (ARMA)

This analysis is based on parametric modeling of the EEG signals. In particular, an Auto-Regressive Moving Average (ARMA) model [15] was used to model the EEG signals recorded at particular electrodes on the scalp. The ARMA *(m,n)* model for an EEG signal y at time t can be defined as:

$$y_t = -\sum_{j=1}^{m} a_t^{(j)} y_{t-j} + \sum_{k=1}^{n} b_t^{(k)} e_{t-k} + e_t \qquad (9)$$

where $a_t^{(j)}$ and $b_t^{(k)}$ are the AR and MA parameters at time instant t respectively and e is a white noise Gaussian process representing the observation error.

Let θ_t be the vector of ARMA parameters and ψ_t be the regression vector made up of the m past signal values and the n past observation error values:

$$\theta_t = [-a_t^{(1)}, \ldots, -a_t^{(m)}, b_t^{(1)}, \ldots, b_t^{(n)}] \qquad (10)$$

$$\psi_t = [y_{t-1}, \ldots, y_{t-m}, e_{t-1}, \ldots, e_{t-n}] \qquad (11)$$

The ARMA model in (9) can then be re-written as:

$$y_t = \psi_t \theta_t^T + e_t \qquad (12)$$

If random walk is allowed, the update of the parameter vector can be defined as:

$$\theta_{t+1} = \theta_t + \omega_t \qquad (13)$$

where ω_t is a normally distributed white noise process with zero mean and covariance matrix Q. The set of equations (12) and (13) represent the structure of a general state-space formulation where the model parameters θ_t are also referred to as the states of the system. A Kalman smoother [15] was then used to find an optimal estimate of the time-varying model parameter vector θ_t. The advantage of using a smoother is that since data is not being processed in real time, future measurements can be used to find a more accurate estimate of the system parameters at time t.

Once an estimate of the ARMA parameters $a_t^{(j)}$ and $b_t^{(k)}$ is available, the poles and zeros of the system can be found by finding the roots of the denominator and numerator of $H(z)$ respectively, where $H(z)$ is the transfer function of the model obtained by taking the Z-transform of (9):

$$H(z) = \frac{Y(z)}{E(z)} = \frac{1 + \sum_{k=1}^{n} b_k z^{-k}}{1 + \sum_{j=1}^{m} a_j z^{-j}} \qquad (14)$$

These poles and zeros can then be used as features for classifying between epileptics and controls. In particular three sets of feature vectors were considered:

- Frequency and/or Magnitude of the whole set of poles and zeros.
- Frequency and/or Magnitude of poles only.
- Frequency and/or Magnitude of zeros only.

These feature vectors were extracted for signals recorded at each individual channel and using the statistical tool ANOVA (section G), the feature vector from specific channels resulting in the most significant difference between epileptics and controls was used in the classification stage.

G. Feature Selection

This study proposes a statistical method for mining the most significant lobes, resembling the way many clinical neuro-physiological studies evaluate the brain activation patterns. Normality of the log-transformed PS is tested using the D'Agostino-Pearson test [16]. Since normality was met and two classes are being discriminated (test vs. control) t-test or analysis of variance (ANOVA) is the ideal test to use. Greenhouse-Geisser-corrected degrees of freedom are used on data that do not meet the sphericity assumption. For those bands where the significance criterion ($p<0.05$) is fulfilled, follow-up post-hoc tests for each channel are performed to accentuate the best candidates to keep as features, which resemble the most significant channels in terms of activity.

H. Classification

A Linear discriminant classifier (LDC) and a quadratic discriminant classifier (QDC) were used to classify between epileptics and controls. Table III shows the percentage of the correct classification using the leave-one-out scheme for each of the three classes considered (control, math, fractals).

III. RESULTS

A. Testing using the simulated signals

In order to assess the performance of the features obtained by CSA, SWT and ARMA modeling, these methods were applied on the simulated signals shown in Table I.

1) CSA (Fig. 3)

The CSA method was applied, as described in section IID, to the simulated data resulting in the plots shown in Figure 3. These plots show that CSA was able to capture the α1 activity in the first quarter of the signal of Channel 1, although some spectral leakage to nearby bands could be observed. The γ1 activity in this channel was not captured. CSA was able to detect the present α1 activity in both Channels 2 and 3 within the correct time periods, while no activity was detected in Channel 4, which consisted only of random noise.

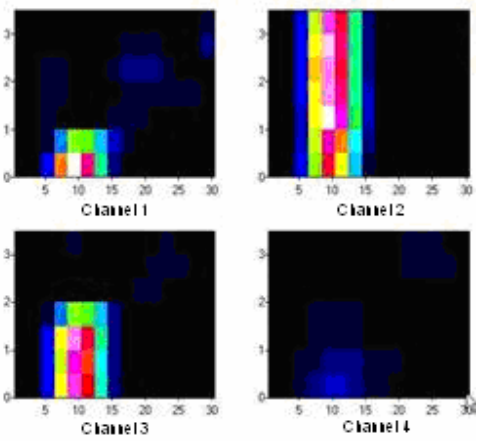

Fig. 3. Activity captured on the simulated dataset using the CSA method (The x-axis represents frequency in Hz and the y-axis represents time in seconds).

2) SWT (Fig. 4)

Similarly, SWT was applied, as described in section IIE, to the simulated data resulting in the time-frequency plots shown in Figure 4. The correct activity was detected for Channels 1, 2 and 3; however, some false random activity was captured in Channel 4.

3) ARMA (Fig. 5)

The ARMA model described in section IIF was applied to each simulated channel by dividing the signals into 0.8s windows overlapped by 50%. The optimum ARMA order was found using AIC [17] in the range of orders between 2 and 20. The estimated ARMA parameters at the end of each window found by the Kalman smoother were converted into the signal's poles and zeros, which were then used to generate the frequency spectrum (see Fig 5a for an example).

Figure 5b shows the frequency content of the signal in each window, extracted from the ARMA modeled frequency spectrum. The simulated activity was detected correctly in most time windows, except for some intervals which were misclassified because the signal within that interval was non-stationary., as can be observed in Channel 3 in the interval between 0.8s and 1.0s.

B. Actual EEG data

These three methods were then applied to the real EEG data described in section IIC giving the classification scores in Table III. The feature vector extracted from CSA was

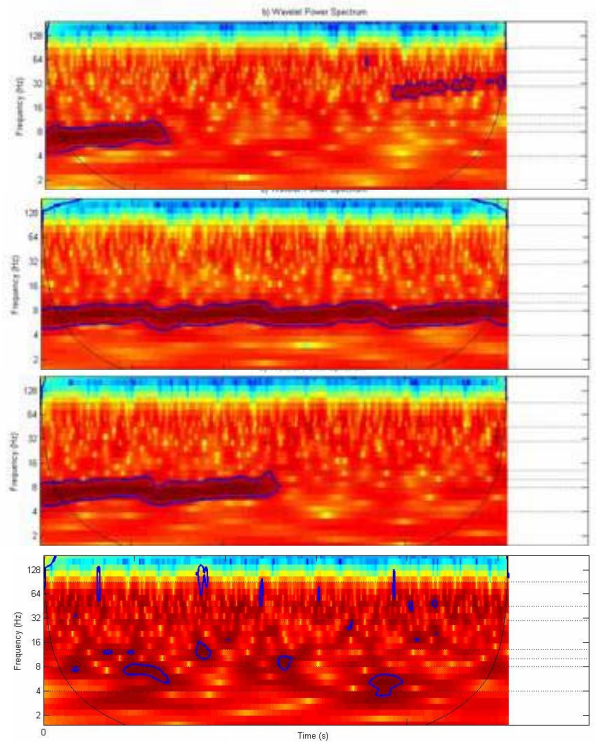

Fig. 4. Activity captured on the simulated dataset using the SWT method. Significant activity is indicated as closed contours. (The y-axis represents frequency in Hz and the x-axis represents time in seconds).

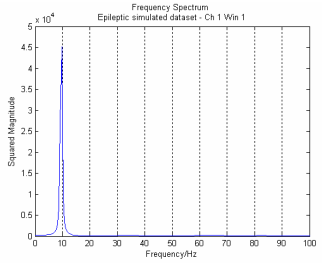

Fig. 5a. Frequency spectrum for simulated Channel 1, Window 1 showing expected alpha activity.

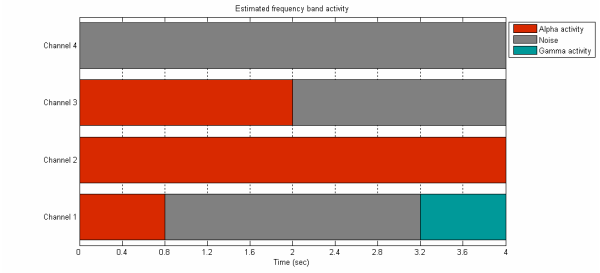

Fig. 5b. Activity captured on the simulated dataset using the ARMA method. Red and blue regions represent alpha and gamma activity, respectively.

ind1+ind2. SWT indicated both significant bands and channels (Table III). For ARMA modeling, the EEG signal at each electrode was modeled by an ARMA *(m,n)* model where *m* and *n* were set to 2 and 3, respectively. This choice of model order was based on the mode of the optimal orders given by AIC when applied across all channels. Since subjects were performing a particular task, a stationary signal was assumed and the model was applied to the whole signal time period. Using ANOVA, it was found that the

TABLE III
CLASSIFICATION RESULTS: SWT SUGGESTS SIGNIFICANT BANDS AND CHANNELS, ARMA SUGGESTS SIGNIFICANT CHANNELS ONLY (COMBINED BANDS) & CSA MAKES NO BAND OR CHANNEL DISTINCTION

	Feature Type	Alpha (α)	Beta (β)	Gamma1 (γ1)	Gamma2 (γ2)	Classification Score
Control (Rest)	SWT (PS)	-	-	FP1	-	67.5 % (QDC)
	ARMA (zeros)			FC3		72.5 % (LDC)
	CSA (ind1+ind2)		Combination of bands All channels were included			72.5 % (LDC)
Math test	SWT (PS)	FP1 FP2 F3 C3 Fz	FP1	FP1	FP1	72.5 % (QDC)
	ARMA (zeros)			F3 FC3 C3 CP3 FCz CPz FC4		73.4 % (LDC)
	CSA (ind1+ind2)		Combination of bands All channels were included			67.5 % (LDC)
Fractal test	SWT (PS)	-	FP2	FP1 FP2	FP1 FP2	65.0 % (LDC)
	ARMA (zeros)			TP7		65.0 % (LDC)
	CSA (ind1+ind2)		Combination of bands All channels were included			65.0 % (LDC)

feature vector made up only of the magnitude of the zeros gave a significant difference between the two subject groups (with $p<0.01$). Table III shows the classification results for the channels with the lowest p-value.

IV. DISCUSSION

Using CSA, in the simulated environment it was shown that the indicator made up of a combination of bands managed to satisfactorily extract the activity present within each channel. The indicator was found to have a higher value for control subjects when compared to epileptics. This was true for both the simulated and real EEG data. For the simulated dataset, the indicator representing the alpha activity was smaller for channel 2 which was made up of 4N alpha activity than for channel 1 having only 1N alpha activity. For the real dataset, different individuals resulted in different indicators which overlap, making the classes not completely separable.

The SWT method was found to have good time-frequency resolution and it gave reasonable classification results which compare well with the other approaches.

The results of the simulated data modeled by the ARMA filter showed that the underlying activity could be well represented by the model poles and zeros. However, the ARMA features are sensitive to the non-stationarity of the signal. Appling the ARMA model to the real EEG data by using the vector of magnitudes of the zeros, the two subject groups were correctly classified with a classification score of up to 73.4%.

The three approaches analyzed in this paper have all provided suitable biomarkers that distinguish between children with mild epilepsy and controls. The classification results obtained from all three methods are satisfactory, considering that no a-priori information was provided. When applied to the real EEG dataset, the added value of SWT approach was the ability to provide both channel and frequency information as compared to the ARMA approach, which provided only channel information and the CSA approach, which provided no band or channel information whatsoever. Although, the classification scores for all three approaches were comparable, the ARMA approach gave the highest overall scores. Better performance may possibly be achieved by adopting other classifiers and by fusing the information from the different approaches.

REFERENCES

[1] E. Haselsteiner, and G. Pfurtscheller, "Using time-dependent neural networks for EEG classification," IEEE Trans. Rehabil. Eng., vol. 8, no. 4, pp. 457–463, 2000.

[2] T. Wang, J. Deng, B. He, "Classifying EEG-based Motor Imagery Tasks by means of Time-frequency Synthesized Spatial Patterns," Clinical Neurophysiology, vol. 115, pp. 2744-2753, 2004.

[3] R.S. Huang, C.J. Kuo, L.L. Tsai, and O.T.C. Chen, "EEG pattern recognition–arousal states detection and classification," Proc. IEEE. Int. Conf. on Neural Networks, vol. 2, pp. 641–646, 1996.

[4] P. Stoica, and R.L. Moses, Introduction to Spectral Analysis, Prentice-Hall, pp. 24-26, 1997.

[5] G. Dumermuth, and L. Molinari, "Spectral analysis of the EEG. Some fundamentals revisited and some open problems," Neuropsychobiology, vol. 17, pp. 85–99, 1987.

[6] P.G. Simos, E. Papanikolaou, E. Sakkalis, and S. Micheloyannis, "Modulation of gamma-band spectral power by cognitive task complexity," Brain Topogr., vol. 14, pp. 191-196, 2002.

[7] A.M. Bianchi, L.T. Mainardi, S. Cerutti: "Time-frequency analysis of biomedical signals," Trans. of the Institute of Measur. and Controls, vol. 22, pp. 321-336, 2000.

[8] S.Blanco, S.Kochen, O.A.Rosso, P.Salgado, Applying Time-Frequency Analysis to Seizure EEG Activity, IEEE Eng. In Med.&Biol.Jan/Feb1997 pp.64-71

[9] W.J. Wiliams, H.P.Zaveri J.C.Sackellares, Time-Frequency Analysis of Electrophysiology Signals in Epilepsy, IEEE Eng. In Med.&Biol.March/April1995 pp.133-143.

[10] J.-P.Lanquart, M.Dumont, P. Linkowski, QRS artifact elimination on full night sleep EEG, Med. Eng.&Phis 28 (2006) pp.156-165.

[11] I. Daubechies, "The Wavelet transform time-frequency localization and signal analysis," IEEE Trans. Inform. Theory, vol. 36, pp.961-1004, 1990.

[12] C.S. Burrus, and R.A. Copinath, and H. Gao. Introduction to wavelets and wavelet transforms: a primer. Englewood Cliffs, NJ: Prentice Hall, 1998.

[13] C. Torrence, and G.P. Compo, "A practical Guide to Wavelet Analysis," Bull. Am. Meteorol. Soc., vol. 79, pp. 61-78, 1998.

[14] R.H. Shumway, D.S. Stoffer, "An Approach to Time Series Smoothing and Forecasting using the EM algorithm," Journal of Time Series Analysis, 3(4), 1982.

[15] M.P. Tarvainen, J.K. Hiltunen, P.O. Rantaaho, P.A. Karjalainen, "Estimation of Nonstationary EEG with Kalman Smoother Approach: An Application to Event-Related Synchronization (ERS)," IEEE Transactions on Biomedical Engineering, 51(3), 2004.

[16] JH. Zar, Biostatistical Analysis, New Jersey, USA: Prentice-Hall, 1999.

[17] M.B. Priestley, "Spectral Analysis and Time Series," Volume 1: Univariate Series, Academic Press, 1981.