Stochastic Models for the EEG Frequencies

[1]A.ABUTALEB, [2]H. ABDELALEEM, [3]K. HEWEDY

[1]School of Engineering, Cairo University, Giza, EGYPT
[1]MIT/Linclon Lab, Lexington, Mass, USA.
[2]School of Medicine, Benyswaif University, Benyswaif, EGYPT
[3]Ministry of Health, Cairo, EGYPT

*Abstract***:** For each of the frequency bands of the EEG, a stochastic model is assumed. The stochastic model assumes that the frequency follows an Ornstein-Uhlenbeck stochastic differential equation. Using the extended Kalman filter and pseudo maximum likelihood, the parameters of the model are first estimated for the simulated data and proved to be accurate. We then applied the same models, to estimate the frequencies, to real data of EEG signals that were obtained from 8 patients suffering from Epilepsy. The estimated frequencies showed that there is no difference between the right lobe and the left lobe signals, except for the beta band, which agrees with physiological findings.

Keywords: EEG, Epilepsy, Stochastic Differential Equation, Ornstein-Uhlenbeck Process, Frequency Estimation, Fourier Transform, Extended Kalman Filter.

1 Introduction:

A typical EEG time series appears to be a summation of waves with contributions from every part of the spectrum, from 0.5 to 150 Hz, appearing with fluctuating phases and variable amplitudes that need to be estimated. Over the years several models have been developed to describe the EEG. For example, [1] used the autoregressive model for EEG signals, [2] used the Welch method to estimate the periodogram for the analysis of the EEG signals, [3] used limit cycles to generate different frequencies of the EEG.

In general, two main classes of time–frequency analysis (TFA) approaches have been proposed for EEG signals. The first is non-parametric spectral estimation methods. Among these, the most common transform is the short-time Fourier transform (STFT), which takes the fast Fourier transform (FFT) of successive and overlapping windows of a signal. However, the STFT has two main disadvantages. Firstly, the time frequency resolution of the STFT cannot simultaneously obtain a high frequency resolution and accurately the timing of any changes in frequency [4 and 5]. Secondly, the spectrum obtained by using the STFT method is characterized by a flatness at the peak, which is unexplained and usually prevents the distinction between spectral peaks or masks low amplitude spectral peaks. Actually this is the subject of the current report. While Fourier transform (FT) has been extensively used to analyze EEG signals, it assumes that the signals are stationary random processes. EEG, however, is a nonstationary process [6 and 7]. Nevertheless, the FT gives insight or preliminary results on the shape and components of the EEG and one could use it as just a transformation of data.

Other time–frequency transforms, including the continuous wavelet transform (CWT), can achieve better time–frequency estimates for nonstationary signals by applying a short window at high frequency and a long window at low frequency [8 and 9]. However, they have degraded frequency resolution for high-frequency contents and degraded time

resolution for low-frequency components, and also displayed spectral smearing due to the finite size of their operator [4]. The choice of the wavelet itself is a subjective matter.

Unlike the non-parametric spectral methods, the power spectral density (PSD) estimates based time-varying autoregressive (TVAR) is a parametric approach [10]. Parametric spectral analysis methods can usually achieve higher frequency resolution than non-parametric approaches provided that an appropriate model should be obtained [8]. The TVAR model is often utilized as an efficient tool to reveal the dynamics of signals due to its simplicity and effectiveness [11 and 12]. Adaptive Kalman filter algorithm and basis function expansion and regression methods have been used to find the TVAR parameters [13 and 10].

In this report we focus on the estimation of the random frequencies of the EEG signal. We first use FT and explain the observed flatness in the peaks when the simulated signal has random frequency. We then propose a stochastic differential equation (SDE) model for the frequency. This model is an Ornstein-Uhlenbeck (OU) model that represents a frequency bouncing around a steady state value in a random fashion. The SDE is driven by a Wiener process. The Ornstein-Uhlenbeck (OU) model is known to be the limit of a TVAR model [14].

To analyze the Epilepsy and seizure and to estimate the EEG frequencies, a group of 8 patients were studied. We obtained the EEG signal at different times pre ictal changes, ictal changes and post ictal changes (pre, event and post) respectively. The EEG spectrum was then separated into its bands through band pass filters. The median frequency in each band was modeled as an OU process. Notice that we observe the sinusoidal wave not the frequency. Two methods were then used to estimate the parameters of the OU model; (1) pseudo maximum likelihood based on Girsanov theory, (2) the extended Kalman filter.

We also assumed that the EEG could be modeled as the summation of five sine waves with different frequency bands (Delta, Theta, Alpha, Lower beta and high beta). Assuming deterministic and constant frequencies, we divided the time series of the EEG into small time windows and we estimated the unknown five amplitudes and five frequencies in each time window. We then performed statistical tests on the estimated parameters, of the left lobe and the right lobe, from the OU models. The results show that there are no differences between the left lobe and the right lobe except for the beta band. This agrees with the physiological findings.

This report is organized as follows: in Section 2, we present the conventional methods to describe the EEG signal. In Section 3, we present the proposed OU model of the frequency of the EEG and how to find an estimate for the model parameters. In Section IV, we present results obtained from real data and conclusions.

2 Problem Formulation:

The EEG could be modeled as a summation of sinusoidal components with random amplitude, random frequency, and random phase. There are, predominantly, five (5) components that belong to the five frequency bands; 0.1-3.5 Hz (*delta*). 4-7.5Hz (*theta*), 8-13 Hz (*alpha*), 14-30 Hz (*betaI and betaII*) and >30 Hz (*gamma*).

The ith frequency component $S_i(t)$ could be modeled as:

$$S_i(t) = a_i Sin(2\pi\ f_i t + \phi_i), \qquad i = 1,...,5 \tag{2. 1}$$

where a_i ith amplitude that could have its own SDE

f_i ith frequency that could have its own SDE

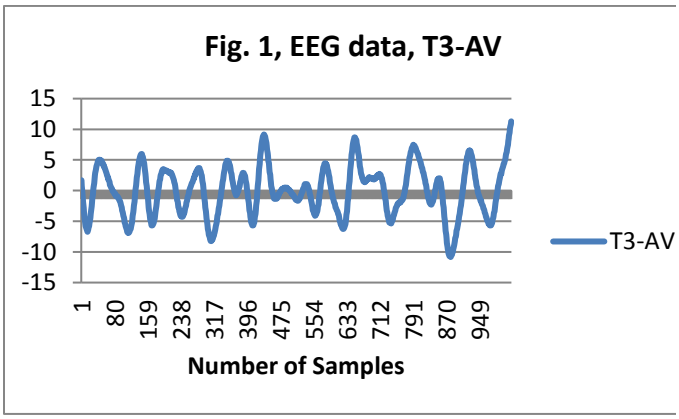
ϕ_i ith phase that could have its own SDE

The EEG observed signal, z(t), could be modeled as:

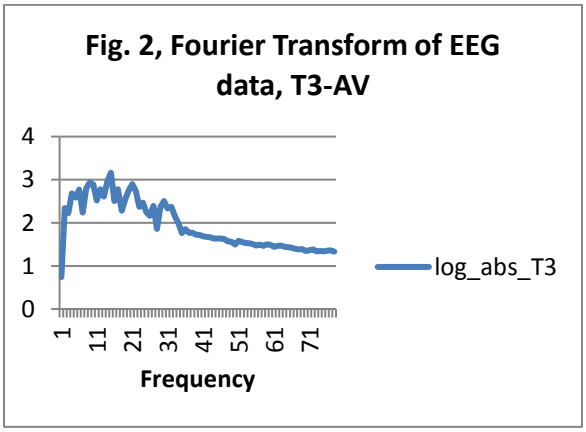
$$z(t) = \sum_{i=1}^{5} S_i(t) = \sum_{i=1}^{5} a_i Sin(2\pi\ f_i t + \phi_i) \quad (2.2)$$

One has to find an estimate for the amplitude, frequency, and phase of each component. This is not an easy task if any or all the components are random or could be represented by SDE.

A typical EEG is shown in Fig. 1 for Channel T3-AV where the sampling interval is 1 mSec and the total observation period is 1 second.



The Fourier transform of this signal is shown in Fig. 2



Notice the presence of around five (5) flattened peaks at frequencies 4 Hz., 8 Hz., 14 Hz., 20 Hz., and 29 Hz. This suggests that channel T3-AV of the EEG could be modeled as the sum of five sinusoids; the frequency of each is random and each could be modeled as an OU process. This is explained in more details in Section 3.

3 The Proposed Approach and Parameter Estimation:

As we notice in Fig. 2, the peaks of the EEG spectrum are flat. This suggests the presence of random frequency or random phase and we must develop a methodology to estimate this randomness.

3.1 Simulation of Single Sinusoid with Random Frequency:

In order to understand this phenomena of flat peaks, we simulated a pure sinusoidal wave with the values: (1) $a_i = 1$, (2) $f_i = 12\,Hz.$, (3) $\phi_i = \pi/3$. The sampling interval is 0.001 seconds, and the length of the data is 1024 points. The Fourier transform (FT) is shown in Fig. 3. Notice the sharp peak of the spectrum located at 12 Hz. If we add random component to the phase while assuming deterministic and constant amplitude and frequency we get the model:

$$S_i(t) = a_i Sin(2\pi\ f_i t + \gamma_i W_i(t)), \qquad i = 1,...,5 \quad (3.1)$$

where $W_i(t)$ Wiener process with variance t.

Equation (3. 1) is simulated with $\gamma_i = 5$. As the value of γ_i is increased the peak of the spectrum flattens and multiple peaks start to emerge. This model was previously investigated and the model parameters were obtained using the likelihood method [15 and 16].

In this report, however, we propose another model; an OU stochastic process model for the frequency f_i. This model represents a frequency f_i that is bouncing around a reference value λ_i. The speed of return to the reference value is controlled by μ_i. The strength of excursion is controlled by σ_i. The Ornstein-Uhlenbeck(OU) model for the frequency is given as:

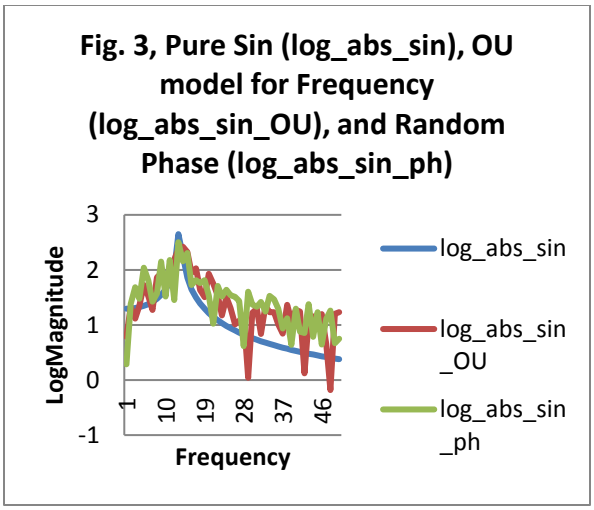
$$df_i(t) = \mu_i[\lambda_i - f_i(t)]dt + \sigma_i dW_i(t), \qquad i = 1,...,5 \tag{3. 2}$$

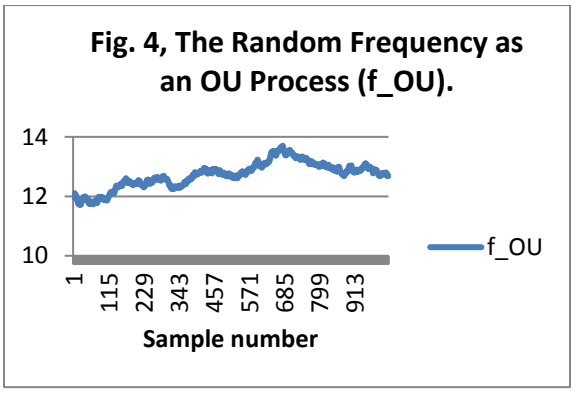
where μ_i controls the return of f_i to the value λ_i

λ_i the reference frequency or the median frequency for the EEG ith band

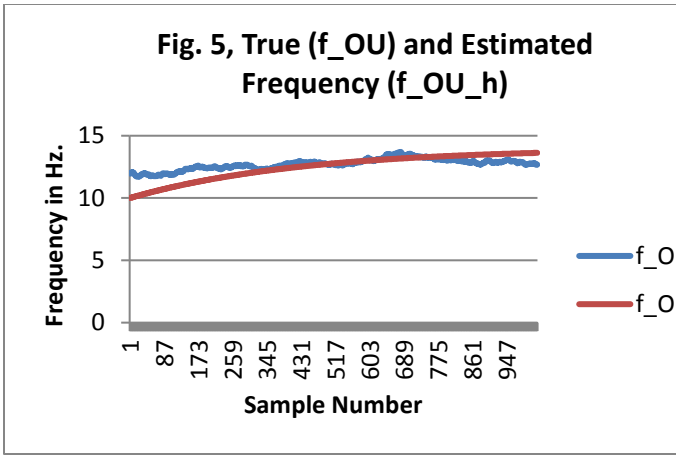
σ_i controls the strength of excursion around λ_i

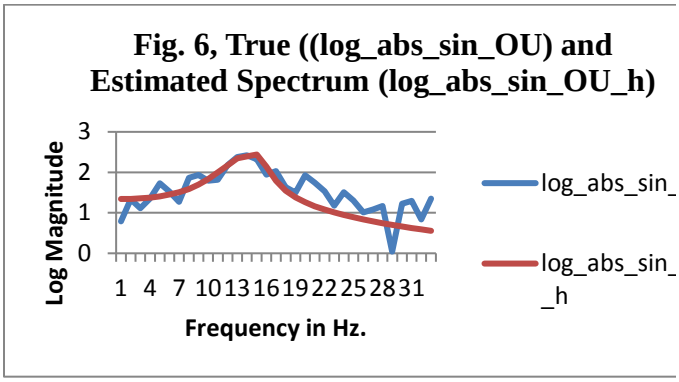
Equation (2. 1) is simulated with the frequency given in equation (3. 2) as shown in Fig. 4. The values are: $\mu_i = 0.1$, $\lambda_i = 12$, and $\sigma_i = 1$. The log magnitude of the Fourier transform of the simulation of a single frequency component is shown in Fig. 3. We also show the simulations of equations (3. 1) and (3. 2). Notice the effect of including randomness; the spectrum flattens and the noise level increases. Compared to the random phase model of equation (3. 1), the flatness of the peak is clearer when we use the OU model for the frequency (equation (3. 2)). That is why we adopt this model in this paper.





The estimates of the different unknowns in equation (3. 2) are obtained through the methods of pseudo maximum likelihood and the extended Kalman filter as will be explained next. The estimated and true frequencies are shown in Fig. 5. The Fourier transform of the estimated single sinusoid with random frequency, along with the true simulated sinusoid with random frequency are shown in Fig. 6. This simulation shows the efficacy of the proposed method.





3.2 The SDE of the Single Sinusoid when the Frequency is an OU Process:

In this subsection we derive the SDE governing the observations of a single sinusoid that has its frequency following an OU process. This is a straightforward application of the Ito lemma. The ith signal is modeled as before as:

$$S_i(t) = a_i Sin(2\pi\ f_i(t)t + \phi_i), \qquad i = 1,...,5 \tag{3. 3}$$

Thus,

$$f_i(t) = \left(\frac{\arcsin \frac{S_i(t)}{a_i} - \phi_i}{2\pi t} \right), \qquad i = 1,...,5 \tag{3. 4}$$

We now derive an SDE for the observed signal $S_i(t)$. Using Ito lemma we get:

$$dS_i(t) = \frac{\partial S_i(t)}{\partial t} dt + \frac{\partial S_i(t)}{\partial f_i(t)} df_i(t) + \frac{1}{2} \frac{\partial^2 S_i(t)}{\partial f_i^{\ 2}(t)} (df_i(t))^2$$

where $\frac{\partial S_i(t)}{\partial t} = 2\pi\ a_i\ f_i(t) \cos(2\pi\ f_i(t)t + \phi_i)$

$$\frac{\partial S_i(t)}{\partial f_i(t)} = 2\pi\ a_i\ t \cos(2\pi\ f_i(t)t + \phi_i)$$

$$\frac{\partial^2 S_i(t)}{\partial f_i^{\ 2}(t)} = -(2\pi\ t)^2\ a_i \sin(2\pi\ f_i(t)t + \phi_i)$$

$$(df_i(t))^2 = \sigma_i^2 dt$$

Collecting terms we get:

$$dS_i(t) = \begin{bmatrix} 2\pi\ a_i\ f_i(t) \cos(2\pi\ f_i(t)t + \phi_i) \\ + 2\pi\ a_i\ t \cos(2\pi\ f_i(t)t + \phi_i) \mu_i (\lambda_i - f_i(t)) \\ - \frac{1}{2} (2\pi\ t)^2\ a_i \sin(2\pi\ f_i(t)t + \phi_i) \sigma_i^2 \end{bmatrix} dt + [2\pi\ a_i\ t \cos(2\pi\ f_i(t)t + \phi_i) \sigma_i] dW_i(t)$$

Substituting "$a_i \cos(2\pi\ f_i(t)t + \phi_i) = \sqrt{a_i^2 - S_i^2(t)}$ ", we get:

$$dS_i(t) = \begin{bmatrix} 2\pi\ f_i(t) \sqrt{a_i^2 - S_i^2(t)} + 2\pi\ \mu_i t \sqrt{a_i^2 - S_i^2(t)} (\lambda_i - f_i(t)) \\ - \frac{1}{2} \sigma_i^2 (2\pi\ t)^2\ S_i(t) \end{bmatrix} dt + \left[2\pi\sigma_i\ t \sqrt{a_i^2 - S_i^2(t)} \right] dW_i(t) \qquad i = 1,...,5 \tag{3. 5}$$

Using equation (3. 4), the observation equation (3. 5) could be represented in the compact format:

$$dS_i(t) = \left[A_i(t,S_i) + \mu_i\lambda_i B_i(t,S_i) + \mu_i C_i(t,S_i) + \sigma_i^2 D_i(t,S_i)\right]dt + \left[\sigma_i\ B_i(t,S_i)\right]dW_i(t)$$

$$= \left[A_i(t,S_i) + \underline{\xi}_i^T(t,S_i)\underline{\theta}\right]dt + \left[\sigma_i\ B_i(t,S_i)\right]dW_i(t) \quad (3.\ 6)$$

where $\underline{\xi}_i(t,S_i) = \left[B_i(t,S_i) \quad C_i(t,S_i) \quad D_i(t,S_i)\right]^T$

$$A_i(t,S_i) = 2\pi\left(\frac{\arcsin\frac{S_i(t)}{a_i} - \phi_i}{2\pi\, t}\right)\sqrt{a_i^2 - S_i^2(t)}$$

$$B_i(t,S_i) = 2\pi\, t\,\sqrt{a_i^2 - S_i^2(t)}$$

$$C_i(t,S_i) = -2\pi\, t\,\sqrt{a_i^2 - S_i^2(t)}\left(\frac{\arcsin\frac{S_i(t)}{a_i} - \phi_i}{2\pi\, t}\right)$$

$$= -\sqrt{a_i^2 - S_i^2(t)}\left(\arcsin\frac{S_i(t)}{a_i} - \phi_i\right)$$

$$D_i(t,S_i) = -\frac{1}{2}(2\pi\, t)^2\ S_i(t)$$

The unknown parameters are the vector $\left[a_i \quad \phi_i \quad \mu_i \quad \mu_i\lambda_i \quad \sigma_i\right]$. If we have only one frequency component, one is able to normalize the amplitude to 1 and the phase to zero. Thus, the reduced vector of unknown parameters becomes $\underline{\theta} = \left[\mu_i \quad \mu_i\lambda_i \quad \sigma_i^2\right]^T$ where T stands for transpose.

Thus, the drift term of the observation equation (eqn. 3. 6) is linear in the vector of unknowns $\underline{\theta} = \left[\mu_i \quad \mu_i\lambda_i \quad \sigma_i^2\right]^T$. This will facilitate the estimation process.

There are several methods to find the estimates of the unknowns [17, 15 and 16]. We next present the estimation methods adopted in this report.

3.3 Pseudo Maximum Likelihood Estimate of the OU Parameters and the Extended Kalman Filter:

For a known value of σ_i , the log likelihood function of the observations at time "t", $L\left(t, S_i(t); \underline{\theta}\right)$, is given as [18 and 15]:

$$L\left(t, S_i(t); \underline{\theta}\right) = \frac{-1}{2}\int_0^t \frac{\left[A_i(s,S_i) + \underline{\xi}_i^T(s,S_i)\underline{\theta}\right]^2}{\left[\sigma_i\ B_i(s,S_i)\right]^2}\,ds + \int_0^t \frac{\left[A_i(s,S_i) + \underline{\xi}_i^T(s,S_i)\underline{\theta}\right]}{\left[\sigma_i\ B_i(s,S_i)\right]^2}\,dS_i(s) \quad (3.\ 7)$$

Since σ_i has to be estimated, equation (3. 7) represents a pseudo log likelihood function. Notice that $L\left(t, S_i(t); \underline{\theta}\right)$ is quadratic function in the unknown vector $\underline{\theta}$. This will facilitate the estimation steps and the calculations of the estimate properties such as bias and variance. In the next subsection we maximize the likelihood function and convert the stochastic integration to Riemann integration. We then present the statistical properties of the estimates. The equations for the extended Kalman filter are presented and a summary of the algorithm is given.

3.3.1 Estimation of the OU parameters:

For an estimated value $\hat{\sigma}_i$, maximizing $L\left(t, S_i(t); \underline{\theta}\right)$ with respect to the unknown vector $\underline{\theta}$ will yield the pseudo maximum likelihood estimates $\hat{\underline{\theta}}$ as follows:

$$\left.\frac{\partial L(t,S_i(t);\underline{\theta})}{\partial \underline{\theta}}\right|_{\underline{\theta}=\hat{\underline{\theta}}} = \underline{0} = -\int_0^t \frac{\underline{\xi}_i(s,S_i)\left[A_i(s,S_i)+\underline{\xi}_i^T(s,S_i)\underline{\theta}\right]}{\left[\sigma_i\ B_i(s,S_i)\right]^2}ds + \int_0^t \frac{\underline{\xi}_i(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2}dS_i(s)$$

$$= -\left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2}ds\right)\hat{\underline{\theta}} - \int_0^t \frac{\underline{\xi}_i(s,S_i)A_i(t,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2}ds + \int_0^t \frac{\underline{\xi}_i(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2}dS_i(s)$$

Rearrange we get:

$$\hat{\underline{\theta}}(t) = \left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2}ds\right)^{-1} \left\{-\int_0^t \frac{\underline{\xi}_i(s,S_i)A_i(t,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2}ds + \int_0^t \frac{\underline{\xi}_i(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2}dS_i(s)\right\} \quad (3.\ 8)$$

This is the pseudo maximum likelihood estimate, $\hat{\underline{\theta}}(t)$, given the observations up till time t. In equation (3. 8) we have a stochastic integration term $\int_0^t \frac{\underline{\xi}_i(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2}dS_i(s)$. We convert this stochastic integration to Riemann integration using Ito lemma [19 and 17]. For example, we know that $\underline{\xi}_i(t,S_i) = \left[B_i(t,S_i) \quad C_i(t,S_i) \quad D_i(t,S_i)\right]^T$ and $B_i(t,S_i) = 2\pi t\sqrt{a_i^2 - S_i^2(t)}$, we need to find

$$\int_0^t \frac{B_i(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2}dS_i(s) = \frac{1}{\sigma_i^2}\int_0^t \frac{1}{B_i(s,S_i)}dS_i(s) = \frac{1}{2\pi\sigma_i^2}\int_0^t \frac{1}{s\sqrt{a_i^2 - S_i^2(s)}}dS_i(s).$$

Define $\frac{\partial g(s,S_i)}{\partial S_i} = \frac{1}{s\sqrt{a_i^2 - S_i^2(s)}}$. Then $g(s,S_i) = \int \frac{dS_i}{s\sqrt{a_i^2 - S_i^2(s)}} = \frac{1}{a_i s}\arcsin\left(\frac{S_i}{a_i}\right)$,

$\frac{\partial g(s,S_i)}{\partial s} = \frac{-1}{s^2\sqrt{a_i^2 - S_i^2(s)}}$, $\frac{\partial^2 g(s,S_i)}{\partial S_i^2} = \frac{S_i(s)\left(a_i^2 - S_i^2(s)\right)^{-3/2}}{s}$. Using Ito lemma for $g(s,S_i)$ we get:

$$g(S_i(T),T) - g(S_i(0),0) = \int_0^T \frac{\partial g}{\partial S_i}dS_i + \frac{1}{2}\int_0^T \frac{\partial^2 g}{\partial S_i^2}\sigma^2(S_i,s)ds + \int_0^T \frac{\partial g}{\partial s}ds$$

It is rearranged as:

$$\int_0^T \frac{\partial g}{\partial S_i}dS_i = g(S_i(T),T) - g(S_i(0),0) - \frac{1}{2}\int_0^T \frac{\partial^2 g}{\partial S_i^2}\sigma^2(S_i,s)ds - \int_0^T \frac{\partial g}{\partial s}ds$$

Substituting we get:

$$\int_0^T \frac{1}{s\sqrt{a_i^2 - S_i^2(s)}}dS_i(s) = \frac{1}{a_i T}\arcsin\left(\frac{S_i(T)}{a_i}\right) - \frac{1}{2}\int_0^T \frac{S_i(s)\left(a_i^2 - S_i^2(s)\right)^{-3/2}}{s}(2\pi)^2 s^2\left(a_i^2 - S_i^2(t)\right)\sigma_i^2 ds + \int_0^T \frac{1}{s^2\sqrt{a_i^2 - S_i^2(s)}}ds$$

$$= \frac{1}{a_i T}\arcsin\left(\frac{S_i(T)}{a_i}\right) - \frac{(2\pi)^2\sigma_i^2}{2}\int_0^T \frac{S_i(s)}{\sqrt{a_i^2 - S_i^2(s)}}s\,ds + \int_0^T \frac{1}{s^2\sqrt{a_i^2 - S_i^2(s)}}ds$$

Since we normalize, the amplitude is set to 1 i.e. $a_i = 1$. The above is integration with respect to time. Numerical methods are used to find this integration and we use the measured

observations $S_i(s)$ for 0<s<T. Similar steps are taken for the other components of equation (3. 8).

3.3.2 Statistical Properties of the Pseudo Estimates $\underline{\hat{\theta}}(t)$:

To find the statistical properties of the estimates $\underline{\hat{\theta}}(t)$, we need another expression for $\underline{\hat{\theta}}(t)$. We do this by substituting the expression of $dS_i(s)$ of equation (3. 6) into equation (3. 8) and we get:

$$\underline{\hat{\theta}} = \left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]^2} ds \right)^{-1} \left\{ \begin{array}{l} -\int_0^t \frac{\underline{\xi}_i(s,S_i)A_i(t,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]^2} ds \\ \int_0^t \frac{\underline{\xi}_i(s,S_i)\left[A_i(s,S_i)+\underline{\xi}_i^T(s,S_i)\underline{\theta}\right]}{\left[\sigma_i\, B_i(s,S_i)\right]^2} ds \\ + \int_0^t \frac{\underline{\xi}_i(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]} dW_i(s) \end{array} \right.$$

which is reduced to:

$$\underline{\hat{\theta}}(t) = \underline{\theta} + \left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]^2} ds \right)^{-1} \left\{ \int_0^t \frac{\underline{\xi}_i(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]} dW_i(s) \right\} \quad (3.\ 9)$$

Assuming that $\left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]^2} ds \right)^{-1}$ could be approximated as a deterministic quantity by replacing $S_i(t)$ with its measured values, an SDE for $\underline{\hat{\theta}}(t)$ is approximated as:

$$d\underline{\hat{\theta}}(t) \approx \left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]^2} ds \right)^{-1} \frac{\underline{\xi}_i(t,S_i)}{\left[\sigma_i\, B_i(t,S_i)\right]} dW_i(t) \quad (3.\ 10)$$

with initial conditions $\underline{\hat{\theta}}(0) = \underline{\theta}$.

Taking expectations of both sides of equation (3. 9) we get:

$$E\left\{\underline{\hat{\theta}}(t)\right\} \approx \underline{\theta} + \left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]^2} ds \right)^{-1} E\left\{ \int_0^t \frac{\underline{\xi}_i(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]} dW_i(s) \right\} \approx \underline{\theta} \quad (3.\ 11)$$

i.e. the estimate is approximately unbiased.

Similarly the variance of the estimate is obtained as:

$$E\left\{\left[\underline{\hat{\theta}}(t)-\underline{\theta}\right]\left[\underline{\hat{\theta}}(t)-\underline{\theta}\right]^T\right\} \approx \left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]^2} ds \right)^{-1} E\left\{ \int_0^t \frac{\underline{\xi}_i(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]} dW_i(s) \int_0^t \frac{\underline{\xi}_i^T(u,S_i)}{\left[\sigma_i\, B_i(u,S_i)\right]} dW_i(u) \right\} \left(\left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\, B_i(s,S_i)\right]^2} ds \right)^{-1} \right)^T$$

$$\approx \left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2} ds \right)^{-1} E\left\{ \int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(u,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2} ds \right\}$$

$$\left(\left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2} ds \right)^{-1} \right)^T$$

As we did before replacing $S_i(t)$ with its measured values, we get the approximate expression for the variance of the estimate as:

$$E\left\{ \left[\underline{\hat{\theta}}(t) - \underline{\theta}\right]\left[\underline{\hat{\theta}}(t) - \underline{\theta}\right]^T \right\} \approx \left(\int_0^t \frac{\underline{\xi}_i(s,S_i)\underline{\xi}_i^T(s,S_i)}{\left[\sigma_i\ B_i(s,S_i)\right]^2} ds \right)^{-1} \quad (3.\ 12)$$

We could have also used equation (3. 10) to find the approximate mean and variance of the estimate $\underline{\hat{\theta}}(t)$.

3.3.3 Summary of the Algorithm:

(1) Use bandpass filters to separate the different EEG bands.
(2) For each band, normalize the amplitude to one and phase to zero.
(3) For each band, use equation (3. 8) to get an estimate for the unknowns $\underline{\theta} = \left[\mu_i \quad \mu_i\lambda_i \quad \sigma_i^2\right]^T$.

3.3.4 A Reduced Form Model and the Extended Kalman Filter for Parameter Estimation [20 and 21]:

In this subsection we present another approach to estimate the parameters of the OU model given that we observe sine waves and we do not observe the frequency. Since the observation equation is nonlinear in the frequency, we use the extended Kalman filter. Extended Kalman filter is used when the states equations and/or the observations equations are nonlinear [21]. Effectively we linearize around some reference set of values that are obtained from the equations themselves (effectively we transform the nonlinear observations of the frequency into linear observations of the frequency). We linearize around the reference frequency $f_{Ri}(t)$ which is the solution of the deterministic equation:

$$df_{Ri}(t) = \mu_i\left[\lambda_i - f_{Ri}(t)\right]dt, \qquad i = 1,...,5 \quad (3.\ 13a)$$

Define $x(t) = \delta f_i(t) = f_i(t) - f_{Ri}(t)$

Thus, $d\delta f_i(t) = df_i(t) - df_{Ri}(t)$

For every frequency component, the change in frequency is relatively small (less than 10 % of the reference). Thus, we develop another OU model for the change in frequency $\delta f_i(t)$ as:

$$d\delta f_i(t) = df_i(t) - df_{Ri}(t)$$

$$= \mu_i\left[\lambda_i - f_i(t)\right]dt + \sigma_i dW_i(t) - \mu_i\left[\lambda_i - f_{Ri}(t)\right]dt$$

$$= -\mu_i\delta f_i(t)dt + \sigma_i dW_i(t) \quad (3.\ 13b)$$

Define $x_i(t) = \delta f_i(t) = f_i(t) - f_{Ri}(t)$, and $u_i(t) = \sigma_i dW_i(t)/dt$

Equation (3. 13b) now has the form:

$$dx_i(t) = F_i(t)x_i(t)dt + \sigma_i dW_i(t)$$

Or $$dx_i(t)/dt = F_i(t)x_i(t) + u_i(t) \quad (3.\ 14)$$

$u_i(t)$ is zero mean Gaussian noise with variance $R_u(t) = \sigma_i^2$, $F_i(t) = -\mu_i$

Similarly for the observation equation which is nonlinear in $f_i(t)$:

$$S_i(t) = a_i Sin(2\pi\ f_i(t)t + \phi_i), \qquad i = 1,...,5 \tag{3. 3}$$

We linearize around the reference frequency $f_{Ri}(t)$.

Define $H_i(t) = a_i \left.\frac{\partial Sin(2\pi\ f_i(t)t + \phi_i)}{\partial f_i(t)}\right|_{f_i(t)=f_{Ri}(t)}$

$= a_i 2\pi\, t\, Cos(2\pi\ f_{Ri}(t)t + \phi_i)$

Using Taylor series expansion around $f_{Ri}(t)$, and keeping only the first order term, we get the approximation:

$$S_i(t) \approx a_i Sin(2\pi\ f_{Ri}(t)t + \phi_i) + H_i(t)\delta f_i(t)$$

$$\approx a_i Sin(2\pi\ f_{Ri}(t)t + \phi_i) + H_i(t)x_i(t)$$

And the approximate observations become:

$$y_i(t) = H_i(t)\delta f_i(t) + v_i(t), \qquad i = 1,...,5$$

Or

$$y_i(t) = H_i(t)x_i(t) + v_i(t), \qquad i = 1,...,5 \tag{3. 15}$$

where $v_i(t) = a_i Sin(2\pi\ f_{Ri}(t)t + \phi_i)$ is approximated as additive white Gaussian noise with zero mean and variance $R_v(t)$. Equation (3. 14) is the state equation and equation (3. 15) is the observation equation. Both are linear. Thus, we have a set of linear equations and, for every $\underline{\theta} = \left[\mu_i \quad \mu_i\lambda_i \quad \sigma_i^2\right]^T$, we use the linear Kalman filter equations as follows:

$$\hat{x}_i(k+1/k) = F_i(k)\hat{x}_i(k/k) \tag{3. 16}$$

$$\Gamma_i(k+1/k) = F_i(k)\Gamma_i(k/k)F_i^T(k) + R_v(k) \tag{3. 17}$$

$$\hat{x}_i(k+1/k+1) = \hat{x}_i(k+1/k) + K_i(k+1)\left[y_i(k+1) - H_i(k+1)\hat{x}_i(k+1/k)\right] \tag{3. 18}$$

$$K_i(k+1) = \Gamma_i(k+1/k)H_i^T(k+1)\begin{bmatrix} H_i(k+1) \\ \Gamma_i(k+1/k)H_i^T(k+1) \\ + R_u(k+1) \end{bmatrix}^{-1} \tag{3. 19}$$

$$\Gamma_i(k+1/k+1) = \left[I - K_i(k+1)H_i(k+1)\right]\Gamma_i(k+1/k) \tag{3. 20}$$

Where $\hat{x}_i(k+1/k)$ is the estimate of the state at time "k+1" , x(k+1), given observations up to time k with variance $\Gamma(k+1/k)$.

$\hat{x}_i(k+1/k+1)$ is the estimate of the state at time "k+1" , x(k+1), given observations up to time "k+1" with variance $\Gamma_i(k+1/k+1)$.

The initial conditions $\hat{x}_i(0/0)$, and $\Gamma_i(0/0)$ could be guessed or estimated. Their effect on the final solution, however, is limited.

3.3.5 Summary of the extended Kalman filter based approach:

(1) For a given set of estimated parameters $\underline{\theta} = \left[\mu_i \quad \mu_i\lambda_i \quad \sigma_i^2\right]^T$ we use the Kalman filter to find an estimate for $x_i(t) = \delta f_i(t)$ that is substituted in the observation equation to get an estimate for the observations $\hat{S}_i(t)$.

(2) These estimates $\hat{S}_i(t)$ are then compared to the true observations $S_i(t)$ and we change the unknown parameters $\underline{\theta} = \begin{bmatrix} \mu_i & \mu_i\lambda_i & \sigma_i^2 \end{bmatrix}^T$, through the gradient method, to reduce the sum of squared differences between the observed and the actual measurements.

(3) The process is repeated several times till we get no change in the estimates of the parameters with given or estimated initial conditions $\hat{x}_i(0)$ and $\Gamma_i(0)$.

4 Results of Real Data:

We have the recorded EEG from 8 patients. It is known that T3 and T4 sharp waves are EEG representation of abnormality. They are the type of waves that are associated with seizures. So we found that the data generated from T3 and T4 are the most suitable data to perform our study. We converted our data signal from time domain to frequency domain by using Fast Fourier Transform (FFT). For the estimation of the OU parameters, band pass filters were used to separate the EEG five bands. In each band, we used the pseudo maximum likelihood method and the extended Kalman filter (EKF) method as explained in Section 3.

We also obtained similar results when we operated in the time domain. We divided the data into small 200 mSec windows for the three segments of data; before event interictal changes, during event interictal changes and after event interictal changes. In each window, assuming constant and deterministic unknowns, we estimated the five frequencies. Typical estimates are given in Fig. 7.

For the statistical tests, we used the average of the estimates. In all cases and using the different methods, there was no significant difference in the estimates of sensor T3 and T4, except for the beta band, which agrees with physiological findings. We next present the results in some details.

Statistical analysis was conducted using SPSS for windows, version 18 (SPSS, Inc., Chicago, IL). For more details about the statistical test please refer to [22]. The current test involved two independent variables. The first one was the (tested group); between subject factor which had two levels (T3 represented left lobes and T4 represented right lobes). The second one was the (measuring periods); within subject factor which had three levels (pre interictal changes, event interictal changes, and post interictal changes). In addition, this test involved one tested dependent variables (frequency of delta, theta, alpha, low beta, and high beta). Prior to final analysis, data were screened for normality assumption, homogeneity of variance, and presence of extreme scores. This exploration was done as a pre-requisite for parametric calculations of the analysis of difference.

Descriptive analysis using histograms with the normal distribution curve showed that the estimated frequency of theta, alpha, and high beta was normally distributed and do not violate the parametric assumption for the measured dependent variable. Additionally, testing for the homogeneity of covariance revealed that there was no significant difference with p values of > 0.05. The box and whiskers plots of the tested variable were done. Normality test of data using Shapiro-Wilk test was used, that reflect the data was normally distributed for the estimated frequency of theta, alpha, and high beta. All these findings allowed the researchers to conduct parametric analysis. So 2×3 mixed design MANOVA was used to compare the tested variables of interest at different tested groups and types of velocity. With the initial alpha level set at 0.05. While Shapiro-Wilk test and descriptive analysis using histograms with the normal distribution curve revealed the data was not normally distributed for frequency of delta and low beta. So non parametric statistical tests in the form of (Freidman tests) was used to compare among different measuring periods for each group and (Wilcoxon Signed Rank tests)

was used as post ad hoc tests if (Freidman tests) is significant. Also, "Mann-Whitney tests" (nonparametric alternative to the independent t test) was used to compare between T3 and T4 in different measuring periods. The alpha level was set at 0.05.

4.1 Short time windows:

Using STFT and for a typical patient, Figure (7) shows the estimated frequencies in the five bands (Delta, Theta, Alpha, Lower beta and high beta) in the three different times (pre inter ictal changes , event inter ictal changes and post inter ictal changes).

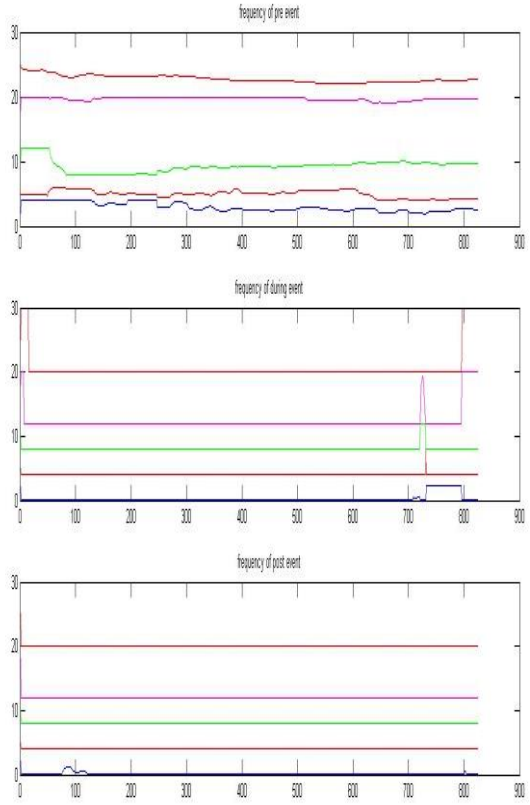


Figure (7): Using STFT, the estimated frequencies for the different five bands [Delta(blue), Theta(red), Alpha(green), Lower beta(purple) and high beta(pink)] at different time conditions (pre , event and post interictal changes).

4.2 Extended Kalman Filter Based Results:

We applied our calculations to the left and the right lobes signal T3 and T4 respectively. We used the equations (3. 16) - (3. 20) to find the estimates of the frequencies. We compared between T3 and T4 results and compared between pre , event and post to know if there is significant difference or not.

Using EKF, Figure (8) shows the result of probability versus frequency from the histogram for frequencies of delta band in three different times (pre inter ictal changes ,event inter ictal changes and post inter ictal changes). The vertical line represents the probability and horizontal line represents frequency in hertz. The first plot shows the result of probability versus frequency of delta before event interictal changes, the second plot shows the result of probability versus frequency of delta during event interictal changes and the last plot shows the result of probability versus frequency of delta after event interictal changes.

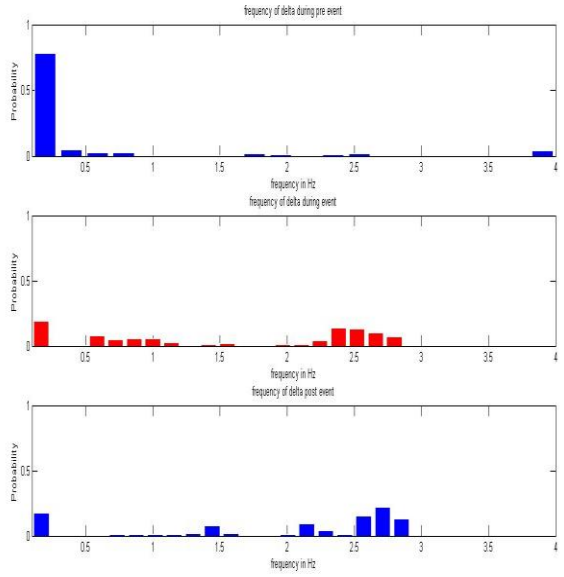


Figure (8):Result of probability versus frequency in Hz from the histogram for the estimated frequency of delta band in three different time (pre interictal changes, event interictal changes and post interictal changes).

Using EKF, Figure (9) shows the result of probability versus frequency from the histogram for frequencies of theta band in three different times (pre interictal changes , event interictal changes and post interictal changes). The vertical line represents the probability and horizontal line represents frequency in hertz. The first plot shows the result of probability versus frequency of theta before event interictal changes, the second plot shows the result of probability versus frequency of theta during event interictal changes and the last plot shows the result of probability versus frequency of theta after event interictal changes.

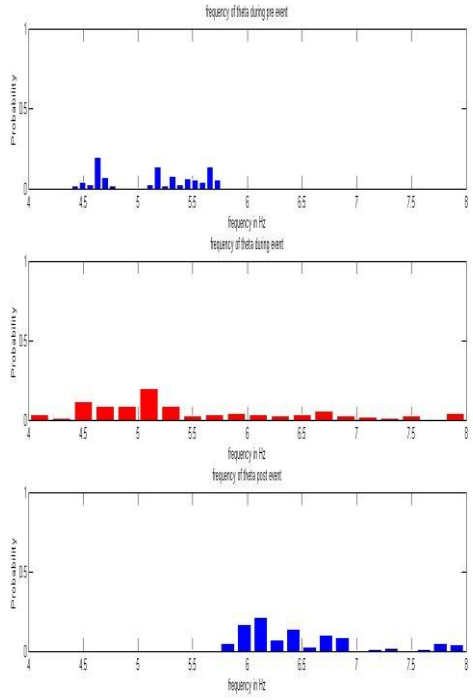


Figure (9): Result of probability versus frequency in Hz from the histogram for the estimated frequency of theta band in three different time (pre interictal changes, event interictal changes and post interictal changes).

Using EKF, Figure (10) shows the result of probability versus frequency from the histogram for frequencies of alpha band in three different times (pre interictal changes, event interictal changes and post interictal changes). The vertical line represents the probability and horizontal line represents frequency in hertz. The first plot shows the result of probability versus frequency of alpha before event interictal changes, the second plot shows the result of probability versus frequency of alpha during event interictal changes and the last plot shows the result of probability versus frequency of alpha after event interictal changes.

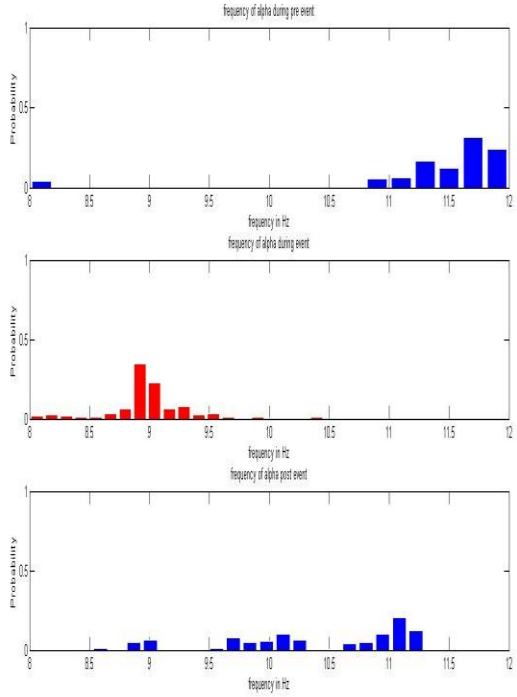


Figure (10):Result of probability versus frequency in Hz from the histogram for the estimated frequency of alpha band in three different time (pre interictal changes, event interictal changes and post interictal changes).

Using EKF, Figure (11) shows the result of probability versus frequency from the histogram for frequencies of lower beta band in three different times (pre interictal changes , event interictal changes and post interictal changes). The vertical line represents the probability and horizontal line represents frequency in hertz. The first plot shows the result of probability versus frequency of lower beta before event interictal changes, the second plot shows the result of probability versus frequency of lower beta during event interictal changes and the last plot shows the result of probability versus frequency of lower beta after event interictal changes.

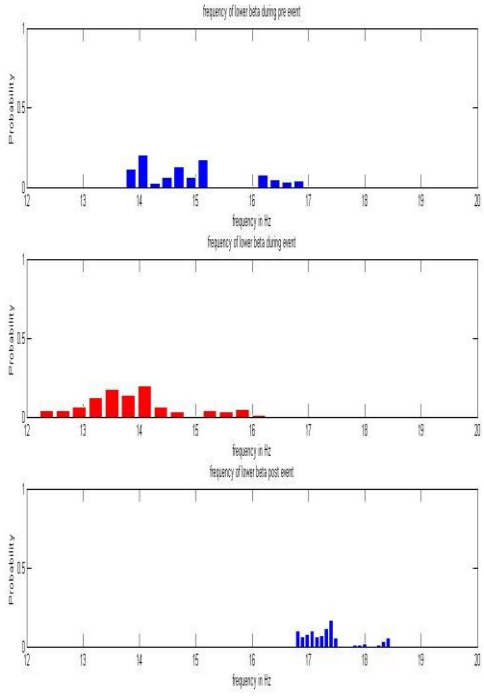


Figure (11):Result of probability versus frequency in Hz from the histogram for the estimated frequency of lower beta band in three different time (pre interictal changes, event interictal changes and post interictal changes).

Using EKF, Figure (12) shows the result of probability versus frequency from the histogram for frequencies of higher beta band in three different times (pre interictal changes , event interictal changes and post interictal changes). The vertical line represents the probability and horizontal line represents frequency in hertz. The first plot shows the result of probability versus frequency of higher beta before event interictal changes, the second plot shows the result of probability versus frequency of higher beta during event interictal changes and the last plot shows the result of probability versus frequency of higher beta after event interictal changes.

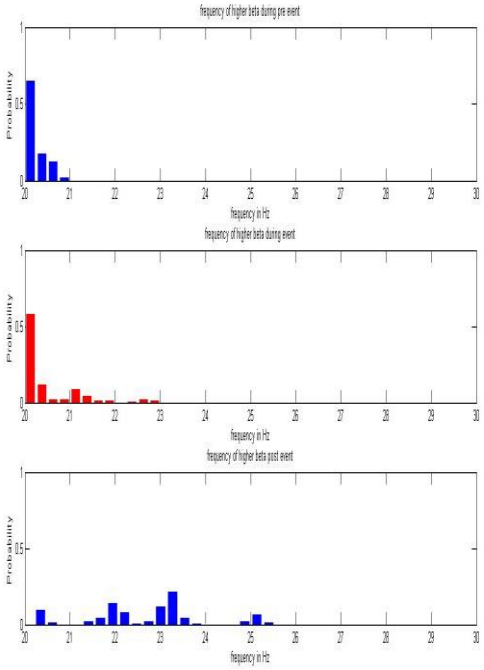


Figure (12):Result of probability versus frequency in Hz from the histogram for the estimated frequency of higher beta band in three different time (pre interictal changes, event interictal changes and post interictal changes).

4.3 OU Model Estimates:

We applied the OU model (equation 3. 2) to the five frequency bands to get the three parameters (mu, lamda and sigma) in each band (delta, theta, alpha, lower beta and higher beta) for T3 and T4 signals and at different time conditions (pre, event and post) and for the 8 patients.

In the OU model of frequency, we compared λ between pre, event and post in T3 and T4 for all bands (delta , theta , alpha , lower beta and higher beta) to know if there is significant difference between interictal changes seizure, before interictal changes and after interictal changes.

4.4.1 Overall effect 2× 3 Mixed Design MANOVA:

Statistical analysis using 2x3 mixed design MANOVA indicated that there were no significant effects of the tested group (the first independent variable) on the all tested dependent variables; frequency of theta, alpha, and high beta (F=0.245, P=0.949). In addition, there were no significant effects of the measuring periods (the second independent variable) on the tested dependent variables (F=1.232, P=0.488). However, the interaction between the two independent variables was not significant, which indicates that the effect of the tested group (first independent variable) on the dependent variables was influenced by the measuring periods (second independent variable) (F=0.572, P=0.789) .

4.4.2The Estimated Frequency of theta (HZ)

As illustrated in figure (10), the mean ± SD values of the estimated frequency of theta in the "Pre", " Event", and "Post " were 6.64 ±1.63, 6.99±1.35, and 6.61±1.28 respectively at the T3. The univariate tests revealed that there was no significant differences in the mean values of the estimated frequency of theta among different measuring periods (F=0.174, P=0.842). As well as, the mean ± SD values of frequency of theta in the "Pre", "Event", and "Post" were 6.14 ±1.55, 7.08 ±1.42, and 6.92 ±1.56 respectively at the T4. The univariate tests revealed that there was no significant differences in the mean values of the estimated frequency of theta among different measuring periods (F=0.937, P=0.415).

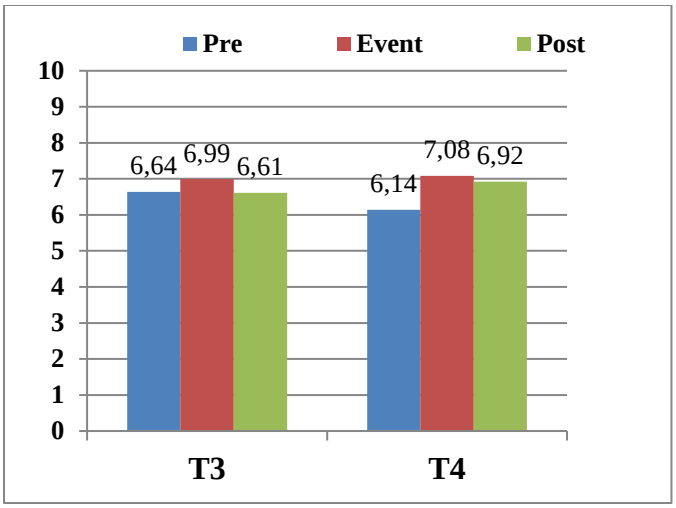


Figure (13):Mean values of the estimated frequency of theta at pre, event, and post in both T3 and T4.

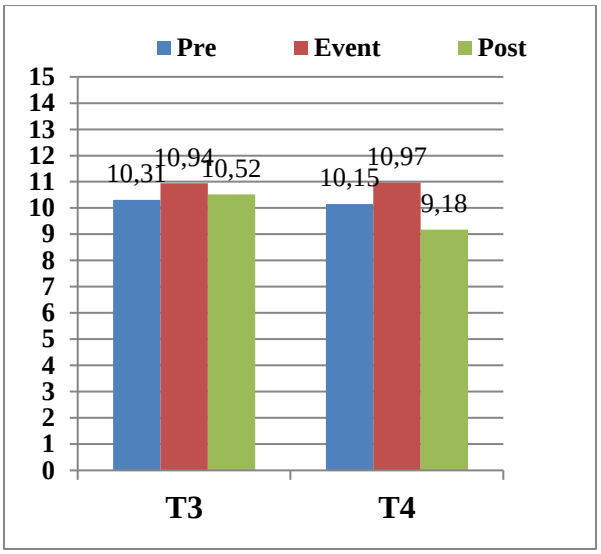


Figure (14):Mean values of frequency of alpha at pre, event, and post in both T3 and T4.

4.4.2 The Estimated Frequency of alpha (HZ)

As illustrated in figure (11), the mean ± SD values of frequency of alpha in the "Pre", " Event", and "Post " were 10.31 ±1.49, 10.94 ±01.27, and 10.52±1.66 respectively at the T3. The univariate tests revealed that there was no significant differences in the mean values of frequency of alpha among different measuring periods (F=0.414, P=0.669).As well as, the mean ± SD values of frequency of alpha in the "Pre", "Event", and "Post" were 10.15 ±1.85, 10.97 ±1.64, and 9.18±1.58 respectively at the T4. The univariate tests revealed that there was no significant differences in the mean values of frequency of alpha among different measuring periods (F=2.808,P=0.094).

4.4.2 The Estimated Frequency of high beta (HZ)

As illustrated in figure (12), the mean ± SD values of frequency of high beta in the "Pre", " Event", and "Post " were 25.5 ±3.64, 21.16 ±2.37, and 21.49 ±2.18 respectively at the T3. The univariate tests revealed that there was no significant differences in the mean values of frequency of high beta among different measuring periods (F=5.164, P=0.02*). So, multiple pairwise comparison tests (Post hoc tests) revealed that there was significant differences at post in compared to pre with (p=0.039*). While, there was no significant differences between (Pre versus Event) and (Event versus Post) with (p= 0.068 and 1.00) respectively. As well as, the mean ± SD values of frequency of high beta in the "Pre", "Event", and "Post" were 23.58 ±3.86, 21.36±1.92, and 24.42 ±4.10 respectively at the T4. The univariate tests revealed that there was no significant differences in the mean values of frequency of high beta among different measuring periods (F=2.147, P=0.154). So, multiple pairwise comparison tests (Post hoc

tests) revealed that there was nosignificant differences between (Pre versus Event), (Pre versus Post), and (Event versus Post) with (p= 0.635, 1.00, and 0.148) respectively.

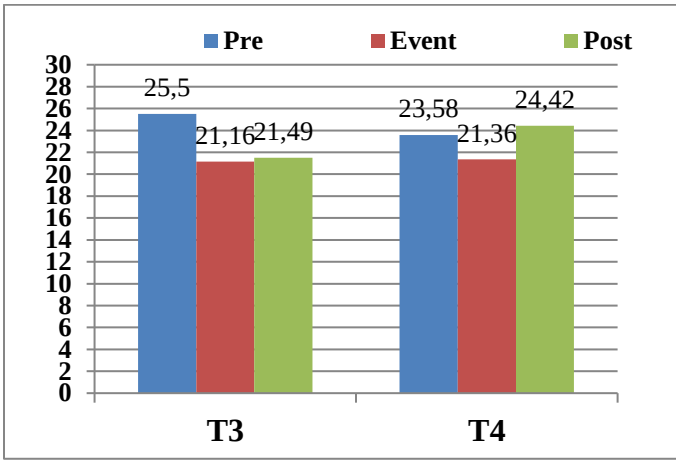


Figure (15): Mean values of frequency of high beta at pre, event, and post in both T3 and T4

4.4.3 The Estimated Frequency of delta (Hz)

As illustrated in figure (13), the mean rank values of frequency of delta in the "Pre", "Event", and "Post" were 1.75, 2.13, and 2.13 respectively at the T3. The Freidman tests revealed that there was no significant differences in the mean rank values of frequency of delta among different measuring periods (F=0.750, P=0.687). As well as, the mean rank values of frequency of delta in the "Pre"," Event", and "Post " were 1.88 , 2, and 2.13 respectively at the T4. The Freidman tests revealed that there was no significant differences in the mean rank values of frequency of delta among different measuring periods (F=0.25, P=0.8)

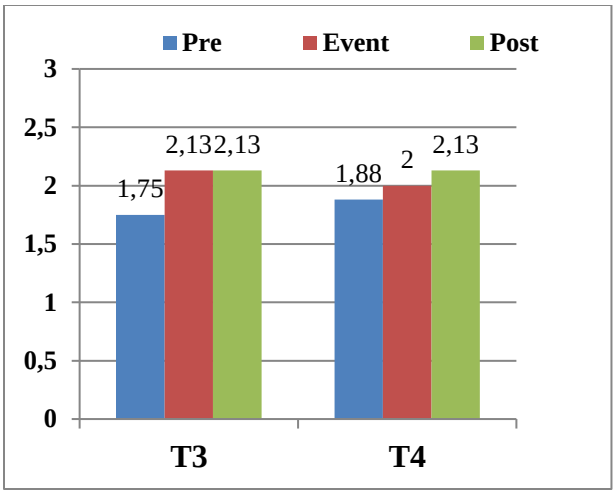


Figure (16): Mean rank values of frequency of delta at pre, event, and post in both T3 and T4.

4.4.4 The Estimated Frequency of low beta (Hz)

As illustrated in figure (14), the mean rank values of frequency of low beta in the "Pre", "Event", and "Post" were 2, 2.13, and 1.88 respectively at the T3. The Freidman tests revealed that there was no significant differences in the mean rank values of frequency of low beta among different measuring periods (F=0.25, P=0.882). As well as, the mean rank values of frequency of low beta in the "Pre"," Event", and "Post " were 1.88, 1.75, and 2.38 respectively at the T4. The Freidman tests revealed that there was no significant differences in the mean rank values of frequency of low beta among different measuring periods (F=1.75, P=0.417).

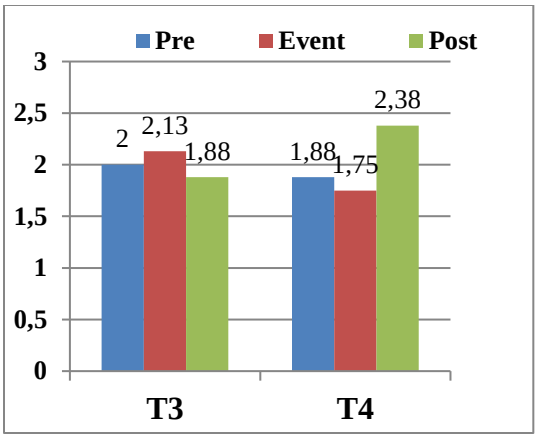


Figure (17): Mean rank values of frequency of low beta at pre, event, and post in both T3 and T4.

5 Summary and Conclusions:

In this report, through simulation, we were able to explain the flatness we observe in the spectrum of the EEG. The randomness of the frequency was shown to be a possible cause. For each EEG frequency band, a model was developed that assumed that the frequency is random and is the output of a stochastic Ornstein-Uhlenbeck (OU) process driven by Wiener process. In this model, the frequency bounces around a fixed value. This model was then used in the observation equation (the measurements of the EEG), and a stochastic differential equation (SDE) was developed for the observed sinusoid (or EEG). The pseudo maximum likelihood method and the extended Kalman filter were then used to find the parameters of the stochastic process that represent the random frequency and the random frequency itself. The estimation procedures were applied to the signals of the T3 and the T4 sensors, of the EEG, and for eight patients and for three intervals of activities. It was shown that there are no statistically significant differences in the estimated parameters between the T3 and the T4 except for the beta band. This result agrees with the literature.

Other models for the EEG could also be considered. A stochastic amplitude with fixed frequency model or a stochastic amplitude with a stochastic frequency model are currently under investigation. The objective is to find accurate estimates of both the amplitude and the frequency for the different EEG bands. It is hypothesized that accurate estimates would result in accurate early diagnosis of diseases.

A much more difficult problem would be the estimation of the amplitudes and frequencies, without band pass filters, when both are stochastic. These issues and others are currently under investigation.

References:

[1] G.Tognola, P.Ravazzani, T.Locatelli, F.Minicucci, F.Grandori, and G.Comi, 1994, "A parametric method for the analysis of temporal and spatial variability in the interictal EEG signal," in Engineering in Medicine and Biology Society, 1994.Engineering Advances: New Opportunities for Biomedical Engineers. Proceedings of the 16th Annual International Conference of the IEEE, pp. 1234–1235, 1994.

[2] M. Naderi and H. MahdaviNasab, "Analysis and classification of EEG signals using spectral analysis and recurrent neural networks," in Biomedical Engineering (ICBME), 17th Iranian Conference of. IEEE, pp. 1–4, 2010.

[3] D. Burke, DYNAMICAL SYSTEM MODELLING OF THE EEG WITH APPLICATION TO DIRECT BRAIN INTERFACING, Ph. D thesis, National University of Ireland, Dublin, 2003.

[4] J. Tary, R. Herrera, M. Van der Baan, "Time-varying autoregressive model for spectral analysis of microseismic experiments and long-period volcanic events", Int. J Geophysics, 2013.

[5] A. Nikoo , A. Kahoo , H. Hassanpour , H. Saadatnia, "Using a time-frequency distribution

to identify buried channels in reflection seismic data", Digital Signal Processing, Vol. 54, Issue C, 2016.

[6] H. Steinberg, T. Gasser, and J.Franke, "Fitting autoregressive models to EEG time series: An empirical comparison of estimates of the order," Acoustics, Speech and Signal Processing, IEEE Transactions on, vol. 33, no. 1, pp. 143–150, 1985.

[7] L.Rankine, N. Stevenson, M.Mesbah, and B.Boashash, "A nonstationary model of newborn EEG," IEEE Transactions on Biomedical Engineering, vol. 54, no. 1, pp. 19–28, 2007.

[8] Z.G. Zhang, Y.S. Hung, S.C. Chan, "Local polynomial modeling of time-varying autoregressive models with application to time–frequency analysis of event related EEG", IEEE Trans. Biomed. Eng. Vol. 58, pp. 557–566, 2011.

[9] H. Bafroui, A. Ohadi, "Application of wavelet energy and Shannon entropy for feature extraction in gearbox fault detection under varying speed conditions", Neurocomputing, Vol. 133, pp. 437–445, 2014.

[10] Yang Li , Mei-Lin Luo , Ke Li, "A multiwavelet-based time-varying model identification approach for time–frequency analysis of EEG signals", Neurocomputing, Vol. 193, pp. 106-114, 2016.

[11] M. Arnold, W.H.R. Miltner, H. Witte, R. Bauer, C. Braun, "Adaptive AR modeling of nonstationary time series by means of Kalman filtering", IEEE Trans. Biomed. Eng. Vol. 45, pp. 553–562, 1998.

[12] Q.H. Huang, J. Yang, Y. Zhou, "Bayesian nonstationary source separation", Neurocomputing, Vol. 71, pp. 1714–1729, 2008.

[13] M. Aboy, O.W. Marquez, J. McNames, R. Hornero, T. Trong, B. Goldstein, "Adaptive modeling and spectral estimation of nonstationary biomedical signals based on Kalman filtering", IEEE Trans. Biomed. Eng., Vol. 52, pp. 1485–1489, 2005.

[14] . Nakajima, and K. Ohashi, "A Cointegrated Commodity Pricing Model", The Journal of Futures Markets, Vol. 32, pp. 995-1032, 2012.

[15] A. Abutaleb, "Instantaneous Frequency Estimation Using Stochastic Calculus and Bootstrapping", EURASIP Journal of Applied Signal Processing, Vol. 12, pp. 1886-1901, 2005A

[16] A. Abutaleb, "INSTANTANEOUS FREQUENCY ESTIMATION WHEN THE AMPLITUDE IS A STOCHASTIC PROCESS USING STOCHASTIC CALCULUS AND BOOTSTRAPPING, Circuits, Systems, and Signal Processing, Vol. 24, #1, pp. 35-52, 2005B.

[17] A. Abutaleb, and M. Papaioannou, Introduction to the Stochastic Calculus and the Malliavin Calculus with Applications to Engineering and Finance, Forthcoming, 2021.

[18] R. Lipster and A. Shiryayev, Statistics of Random Processes, Vol.I and II, Springer, 2007.

[19] S. Govindaraj, "Hypothesis Testing for Diffusion Processes with Continuous Observations: Direct Computation of Large Deviation Results for Error Probabilities", Finance Research Letters, Vol. 2, pp. 234-247, 2005.

[20] A. Ayed, G. Loeper, and F. Abergel, "Forecasting Trends with Asset Prices", arXiv:1504.03934v2 [q-fin.ST], 2015.

[21] P. Maybeck, Stochastic Models Estimation and Control, Academic Press, 1979.

[22] T. MacFarland, Two way analysis of Variance, Springer, 2012.