

## **EFFECT OF SOME POWER SPECTRAL DENSITY ESTIMATION METHODS ON AUTOMATIC SLEEP STAGE SCORING USING ARTIFICIAL NEURAL NETWORKS**

Cuneyt Yucelbas. *Dept. Of Electrical-Electronics Engineering, Selcuk University, Konya, TURKEY.*

Seral Ozsen. *Dept. Of Electrical-Electronics Engineering, Selcuk University, Konya, TURKEY.*

Salih Gunes. *Dept. Of Electrical-Electronics Engineering, Selcuk University, Konya, TURKEY.*

Sebnem Yosunkaya. *Faculty Of Medicine, Necmettin Erbakan University., Konya, TURKEY.*

### **ABSTRACT**

Sleep staging has an important role in diagnosing sleep disorders. It is usually done by a sleep expert through examining sleep Electroencephalogram (EEG), Electrooculogram (EOG), Electromyogram (EMG) signals of the patients and determining the stages of sleep in different time sections named as epochs. Manual sleep staging is preferred among the sleep experts but because it is rather tiring and time consuming task, automatic sleep stage scoring systems get popularity. In this study, we obtained EEG, EMG and EOG signals of four healthy people at sleep laboratory of Meram Medicine Faculty of Necmettin Erbakan University to use them in sleep staging and extracted 20 different features by using some power spectral density estimation methods which are: Fast Fourier Transform (FFT), Welch and Autoregressive (AR). We evaluated the effects of these methods on sleep staging through using ANN classifier. Comparison between these methods was done on each individual whose data were utilized separately from others. According to the results, the maximum test classification accuracy was reported as 79.72% by using of FFT method for subject1. Also, mean of test classification accuracies for all of subjects were obtained as 74.14%, 71,58 and 70.34% with use of FFT, Welch and AR, respectively.

### **KEYWORDS**

Artificial neural networks, automatic sleep stage, EEG, PSD.

## 1. INTRODUCTION

Sleep related problems have begun to affect our daily life more seriously in nowadays than as in the past (Ozsen, 2012). Sleep stage scoring has an important part in determination of sleep related problems. It can be defined as determining sleep stages during the sleep with the use of Polysomnographic (PSG) recordings. These recordings contain electroencephalogram (EEG), electromyogram (EMG), electrooculogram (EOG), electrocardiogram (ECG) and some other signals used to detect chest, body and leg movements. Approximately for 45 years, this scoring process has been done manually analysing the above PSG recordings using some criteria represented by Rechtschaffen and Kales (1969) rules. Although there exist some procedures and definite rules in manual sleep stage scoring, differences between signal interpretations are not uncommon. Also, sleep stage scoring is also a very time consuming task and it can limit the number of subjects that sent to the sleep laboratory for evaluation.

One enters a number of stages during his sleep and the sequence as well as duration of them determines the sleep pattern of the subject. The stages are named as Awake (W), Rapid Eye Movement (REM) and Non-rapid Eye Movement (Non-REM). Non-REM sleep stage can be further divided into 3 phases as Non-REM phase 1, Non-REM phase 2 and Non-REM phase 3 (AASM, 1999). Commonly, Awake stage is seen in the beginning of the sleep and it can be described as a transition stage from the full alertness to the half-sleepy situation and generally it is characterized by alpha rhythms which are signals with frequencies between 8-13 Hz. Non-REM stages includes mixed frequency activity but the frequency of the sleep generally lowers by the increasing deepness of the sleep. In Non-REM 2 stage, sleep spindles and K complexes are searched in the EEG activity. Lastly REM stage is characterized principally rapid eye movements as its name (AASM, 1999).

Some studies were realized with in the fields of signal processing and machine learning in literature, several attempts to develop automated sleep stages have been carried out increasing number of studies for last 20 years. In (Holzmann, et al., 1999), developed an expert system by using ganglionic lattices for infants and researchers obtained an overall performance of 96.4% agreement with the expert on validation data without artefacts and 84.9% agreement on validation data with artefacts. Also in (Oropesa, et al., 1999), To partition EEG signal into 7 particular frequency bands is used Discrete Wavelet Transform for sleep staging. In same study, they obtained an agreement level of 76.6% by computing the energy of these bands and then presented to ANN, researchers reached an agreement level of 76.6%. In other study, used the segmentation and clustering strategies in their application which also involves active participation of sleep operator and obtained 80.6% classification accuracy (Agarwal and Gotman, 2001). In (Piffero, et al., 2004), To reach better results by using fuzzy logic in sleep scoring area is realized a fuzzy based scoring system. Whereas their discernment was good, the results were not so. In their study, Estrada et. al., Utilizing from the EEG signals tried three different feature extraction schemes which were Relative Spectral Band Energy, Harmonic Parameters and Itakura Distance (Estrada, et al., 2004). Also in the studies (Acharya, et al., 2005; Pereda, et al., 1998; Jiayi, et al., 2007; Van Quyen, et al., 2003; He, et al., 2005; Shen, et al., 2003), that were used linear measures, non-linear analysis of EEG signal was also carried out. In some of these studies, linear and quadratic classifiers, k nearest neighbours, parzen kernels and neural networks five classifiers were compared and reported which classifier is more efficient in sleep scoring (Becq, et al., 2005). According to the results obtained understood that the best classifier was Neural Network with a classification accuracy

## EFFECT OF SOME POWER SPECTRAL DENSITY ESTIMATION METHODS ON AUTOMATIC SLEEP STAGE SCORING USING ARTIFICIAL NEURAL NETWORKS

of 72%. Also Hidden Markov Models were also applied to this classification area in (Flexer, et al., 2005). They used two datasets from two labs and reached a classification accuracy of 54% and 42,02% for that labs respectively. wavelet analysis and ANN was applied in their study and obtained a classification rate of 95,55% (Sinha, 2008). Researchers in (Šušmáková and Krakovská, 2008) analyzed 73 characteristics measures in sleep staging process with discriminant analysis by Fisher's quadratic classifier.

In this study, 20 different features were extracted from the power spectral density (PSD) estimations of EEG, EOG and EMG data of four subjects. In obtaining PSD estimations, three well known methods were used and compared in automatic sleep stage scoring. These are, Fast Fourier Transform (FFT), Welch and Autoregressive (AR) methods. After applying these three PSD estimation methods to EEG, EOG and EMG data of four subjects' full night sleep, ANN classifier was trained and tested for optimum parameters that give highest test classification accuracy in each subject. The maximum test accuracies for all subjects were obtained by FFT method. These were 79.72%, 75.62%, 74.91% and 66.32% for subject-1, subject-2, subject-3 and subject-4, respectively.

## 2. MATERIAL AND METHODS

The data to be used in the experimental procedure was recorded at sleep laboratory of Meram Medicine Faculty of Necmettin Erbakan University with the VIASY PSG device. EEG and EOG signals were sampled at a rate of 128 Hz while the sampling rate of EMG was 256 Hz. Roughly 8 hour recordings were separated to 30 sec. long epochs and each epoch was scored as Awake (W), Non-REM1 (N\_1), Non-REM2 (N\_2), Non-REM3 (N\_3) and REM (R) by a sleep expert.

### 2.1 Used Dataset and Preprocessing

The data taken from four voluntary subjects whose stage information is tabulated in Table 1 was used in the following process steps as shown in Figure 1.

Table 1. The distribution of the epochs to the stages in the used data

	Awake	Non-REM1	Non-REM2	Non-REM3	REM	Total
<b>Subject-1</b>	57	68	656	23	162	966
<b>Subject-2</b>	83	26	457	77	162	805
<b>Subject-3</b>	64	31	552	91	157	895
<b>Subject-4</b>	138	54	617	6	160	975
<b>Total</b>	342	179	2282	197	641	3641

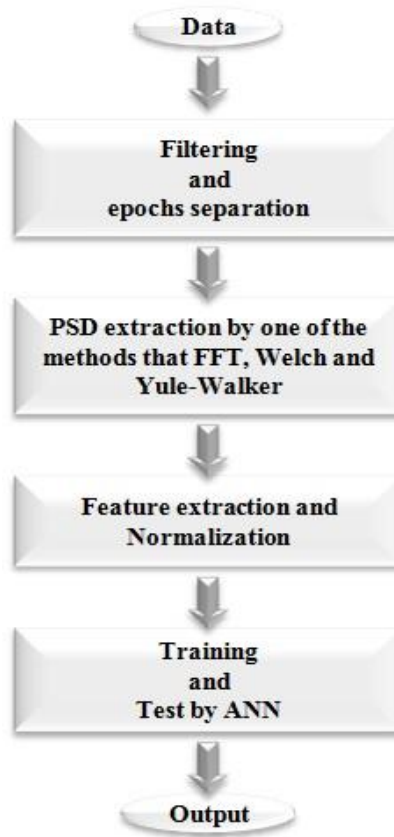


Figure 1. The performed process steps

Data pre-processing: The EEG, EMG, left and right EOG signals of four subjects were processed by filtering, amplification and data cleaning procedures. 6<sup>th</sup> order butterworth bandpass filter was used to filter EEG and EOG signals in 0.3-35Hz band. These cut-off frequencies were 10-70Hz for EMG signals. The filtering frequencies were shown in Table 2. After the processes of filtering, amplification and data cleaning were conducted for all signals of subjects, the epochs of used signals were prepared by separating them into 30 sec. long parts. As a result of this separation, the epochs for EEG, EMG and EOG signals were obtained as given in Table 1. A scoring process for these epochs was realized by a sleep expert and the result of that scoring was taken as the real stages of related epochs.

Table 2. The used filter values

Signal Type	Lower cut-off frequency (Hz)	Upper cut-off frequency (Hz)
EEG	0.3	35
EMG	10	70
Left EOG	0.3	35
Right EOG	0.3	35

## EFFECT OF SOME POWER SPECTRAL DENSITY ESTIMATION METHODS ON AUTOMATIC SLEEP STAGE SCORING USING ARTIFICIAL NEURAL NETWORKS

PSD extraction: Power spectral densities of EEG, EMG, left and right EOG signals were extracted by using FFT, Welch and AR methods in which yule-walker parameter estimation algorithm was used.

Feature extraction: 20 features were extracted for all epochs according to the rules in R&K scoring (Rechtschaffen and Kales, 1969) and they were normalized between [0-1].

ANN classifier: ANN structure is based on a network of artificial neurons connected with each other via weights. By inspiring from the natural neural system, it tries to find best output values for any presented input with this structure. In ANN, an input layer, some number of hidden layers and an output layer form the network of neurons. There exist connections between these layers and it is the values of these connection weights which is learned with the training process. The adaptation process in ANN can be regarded as two-steps. In the first step, the researcher should decide how many hidden layers and how many neurons (named as nodes) should lie in the structure. This is related with the architecture of the ANN. The other step to be determined is to decide which learning algorithms should be used while training ANN to find optimum weight values by back propagation neural network. This aspect must be dealt with carefully because there are many training procedures in literature. Gradient descent training algorithm is one of them which is also a widely used one. In our study, we preferred that algorithm, too. However there are some parameters that should be determined related with this algorithm, the most important ones are learning rate ( $lr$ ) and momentum constant ( $mc$ ) parameter. Learning rate determines the rate of changing in weights in each step and the amount of it is important in this regard. Low learning rate can result in very slow training process while higher learning rates can cause optimum points to be escaped.  $mc$  is the other parameter which is also determines the amount of weight changes in each iteration with  $lr$  but it's value changes according to the gradient of error. Thus, this parameter arranges  $lr$  to increase or decrease with regard to the amount of error change. Thus, it supplies the algorithms to escape from the local minima points.

Training and test of sub-systems: In this study, training and test operations were done separately and the network simulation was performed only one times for each subject but random generator was fixed to a state before the simulations. Thus, training-test data partitioning was done in the scope of each individual. 70% of each stage was used as training data while the remaining 30% of each stage was used for test in each subject. Training of the classifier was realized for the selected features for which optimum ANN architectures were investigated. In ANN, one hidden layer with varying node number was trained with gradient descent algorithm. To find the optimum parameters of ANN was applied following steps:

1. While momentum constant, learning rate and iteration number of ANN were kept constant, node number of hidden layer was increased from 2 to 52 and the optimum hidden layer node number giving the highest classification accuracy was found.
2. While momentum constant, learning rate and node number of hidden layer of ANN were kept constant, iteration number was increased from 100 to 10000 with steps 10 and the iteration number giving the highest classification accuracy was stored.
3. While momentum constant, node number of hidden layer and iteration number of ANN were kept constant, learning rate was increased from 0.1 to 2.0 and the rate with highest classification accuracy was noted.
4. While learning rate, node number of hidden layer and iteration number of ANN were kept constant, momentum constant was increased from 0.1 to 1.0 and the optimum value for it was found by searching highest classification accuracy.

This procedure for finding optimum ANN was repeated for all PSD methods and subjects. In this study, a verbose schematic structure of ANN was not given as in (Sinha, et al., 2003) due to the fact that optimum values of ANN were determined by applying different parameters for each method and subject. Also, an example of experimental design and ANN schematic are shown in the same study by Sinha, et al. (2003).

## 2.2 Experimented PSD Methods

In this study, power spectral density estimation methods which are FFT, welch and AR on automatic sleep stage scoring were used.

Fast Fourier Transform (FFT) is an algorithm to compute the Discrete Fourier Transform (DFT) and its inverse. The DFT is obtained by decomposing a sequence of values into components of different frequencies. This operation is useful in many fields but computing it directly from the definition is often too slow to be practical although an FFT is a way to compute the same result more quickly. An FFT computes the DFT and produces exactly the same result as evaluating the DFT definition directly; the only difference is that an FFT is much faster than DFT (Wikipedia, 2013a).

Welch's method is another widely used PSD estimation method. The method is based on the concept of using periodogram spectrum estimates, which are the result of converting a signal from the time domain to the frequency domain. Welch's method is an improvement on the standard periodogram spectrum estimating method and on Bartlett's method, in that it reduces noise in the estimated power spectra in exchange for reducing the frequency resolution. Due to the noise caused by imperfect and finite data, the noise reduction from Welch's method is often desired (Wikipedia, 2012). In our study, rectangular window with 128 points and 50% overlap were used for this method.

AR method is another experimented PSD method which is a parametric modern spectrum estimate method other than the above ones. In statistics and signal processing, AR model is a type of random process which is often used to model and predict various types of natural phenomena. AR is one of a group of linear prediction formulas that attempt to predict an output of a system based on the previous outputs (Wikipedia, 2013b). Several algorithms were proposed to calculate the parameters of AR such as Yule-Walker, Burg, Covariance and Modified Covariance. In this study we utilized Yule-Walker scheme because it gives better results for long data sequences than other methods. The most important point in AR is determination of model order. Some criteria can be utilized for this as well as trial and error approach which is also not uncommon. We, too, preferred to select model order by experimentation and comparison with PSD graphics of other methods. Model degree was selected as 15, 11 and 9 for EEG, EMG and EOG signals, respectively.

## 2.3 Feature Extraction

After obtaining PSD estimates of EEG, EMG and EOG signals of four subjects, 20 features were extracted from the time and frequency domain EEG, EMG and EOG signals. These features were chosen from the features mentioned in (Ozsen, 2012), as the following:

1. Relative Powers of frequencies in alpha band (EEG)
2. Relative Powers of frequencies in theta band (EEG)
3. Power of theta band/power of alpha band (EEG)

EFFECT OF SOME POWER SPECTRAL DENSITY ESTIMATION METHODS ON AUTOMATIC SLEEP STAGE SCORING USING ARTIFICIAL NEURAL NETWORKS

4. Power of alpha band in related epoch/power of alpha band in previous epoch (EEG)
5. Relative Powers of frequencies in 12-14 Hz (for spindle detection) (EEG)
6. Mean value of the EEG signal in time domain
7. Skewness of the EEG signal in time domain which is calculated with the Equation 1 as the following:

$$x_{ske} = \frac{\sum_{n=1}^N (x(n) - x_m)^3}{(N-1)x_{std}^3} \quad (1)$$

where N is the length of the signal x,  $x_m$  is the mean and  $x_{std}$  is the standard deviation of x.

8. Kurtosis of the EEG signal in time domain which is calculated with Equation 2 as the following :

$$x_{krt} = \frac{\sum_{n=1}^N (x(n) - x_m)^4}{(N-1)x_{std}^4} \quad (2)$$

9. Sum of Powers of frequencies in 0.5-2 Hz (EEG)
10. Sum of Powers of frequencies in delta band (0-4 Hz) (EEG)
11. Sum of Powers of frequencies in 2-6 Hz (EEG)
12. The energy of EMG signal
13. Sum of the Powers in frequency spectrum of EMG signal
14. Sum of the powers of frequencies in 0.5-2 Hz (left eye EOG)
15. Energy of the left eye EOG signal
16. Sum of the powers of frequencies in 0.5-2 Hz (right eye EOG)
17. Energy of the right eye EOG signal
18. Mean value of the signal produced by summing left and right EOG signal
19. Energy of the signal produced by summing left and right EOG signal
20. Standard deviation of the signal produced by summing left and right EOG signal

## 2.4 Testing and Performance Evolution

When a system uses ANN as its classifier three types of criteria can be used to measure the classification performance of that system. These are the mean of squared errors, classification error and classification accuracy. Mean of squared errors is calculated as in Equation 3 and it gives a measure in what degree the outputs of ANN are similar to the real outputs.

$$mse_{error} = \frac{\sum (y_1 - y_2)^2}{n} \quad (3)$$

Here  $y_1$  is the real output while  $y_2$  is the ANN's output and  $n$  is the total number of samples. The  $mse$  error can be read in percentage form by multiplying the result by 100. For example  $mse$  error of 0.03 can also be read as a 3%  $mse$  error. The other criterion is classification error which is calculated by the following equation:

$$err(\%) = \frac{100 \times N_{err}}{N_{test}} \quad (4)$$

Here  $err(\%)$  is the test error in percent,  $N_{err}$  is the number of incorrectly classified epochs and  $N_{test}$  is the number of test data. In calculation of this error, outputs of ANN are rounded and if the output of the ANN is not equal to the real output the  $N_{err}$  is increased. Classification accuracy is calculated as in Equation 5.

$$\text{Classification accuracy}(\%) = 100 - err(\%) \quad (5)$$

We use Equation 4 and Equation 5 to measure the classification performance of that system.

### 3. EXPERIMENTAL RESULTS

As explained in Section 2, three methods of PSD estimation was applied to EEG, EMG and EOG data of four subjects and the effects of them were compared by using ANN which accepts 20 features as inputs extracted from time and frequency domain signals. In Figure 2, PSD estimates of one epoch EEG signal were given which are obtained by all three PSD estimation schemes.

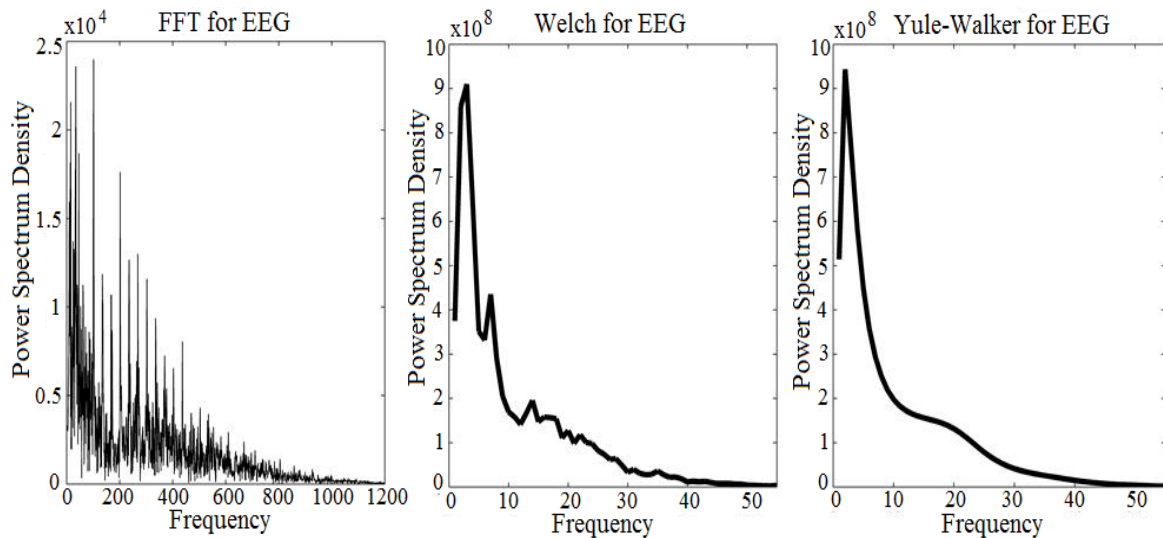


Figure 2. FFT, Welch and AR PSD estimates of EEG signal of one epoch.

As seen from the Figure 2, although Welch and AR methods give more rough estimates of signals, FFT gives many spectral peaks which are hard to comment. But, because of these peaks, FFT gives more accurate PSD estimate of signals as it will be seen in the performance results. Welch and AR methods have the advantage of giving an insight about which frequency regions dominated but because they obtain a rough estimate, some frequency



EFFECT OF SOME POWER SPECTRAL DENSITY ESTIMATION METHODS ON AUTOMATIC  
SLEEP STAGE SCORING USING ARTIFICIAL NEURAL NETWORKS

content in real estimates are underestimated. The classification results for all four subjects and there PSD estimation methods were given in Table 3. As shown in Table 3, the highest test classification accuracies were obtained by FFT method for most of subjects (except from subject-3). The reason for this can be the underestimation characteristics of Welch and AR (Yule-Walker) methods as explained in the above paragraph. The results are also presented in Figure 3 graphically.

Table 3. Obtained classification results for all subjects

SUBJECTS		Subject-1			Subject-2			Subject-3			Subject-4		
METHODS		FFT	WL	AR	FFT	WL	AR	FFT	WL	AR	FFT	WL	AR
<b>RESULTS</b>	<b>TrA %</b>	79,25	70,22	72,14	72,3	71,22	71,04	78,36	77,08	74,03	67,84	66,22	66,81
	<b>TA %</b>	<b>79,72</b>	<b>70,79</b>	<b>72,51</b>	<b>75,62</b>	<b>73,55</b>	<b>73,14</b>	<b>74,91</b>	<b>75,64</b>	<b>69,74</b>	<b>66,32</b>	<b>66,32</b>	<b>65,98</b>
	<b>HNN</b>	30	50	30	7	25	48	31	25	45	15	24	33
	<b>ITER</b>	400	100	100	100	100	100	400	400	100	100	100	100
	<b>LR</b>	0,5	2	1,3	0,7	0,8	0,8	1,4	1,5	0,5	0,8	0,8	1
	<b>MC</b>	0,4	0,1	0,2	0,6	0,80	0,8	0,1	0,4	0,8	0,8	0,1	0,8

TrA: Training accuracy - TA: Test accuracy - HNN: Hidden Node Number - ITER: Iteration number - LR: Learning rate - MC: Momentum constant - FFT: Fast Fourier Transform - WL: Welch - AR: Yule-Walker AR

Whereas there isn't too much difference between PSD estimating methods, even an improvement of 1% in test performance is very important for this problem. In AR method on the other hand, model order selection procedure can be adapted for more efficient results. For welch scheme, the used window type and size can be changed to find highest performance. Thus, this comparison done in our study between these three methods was done based on only a rough estimate of performances. But it is evident that, FFT is an efficient PSD estimation method that can be used confidently in sleep staging process to give frequency content of EEG, EMG and EOG signals.

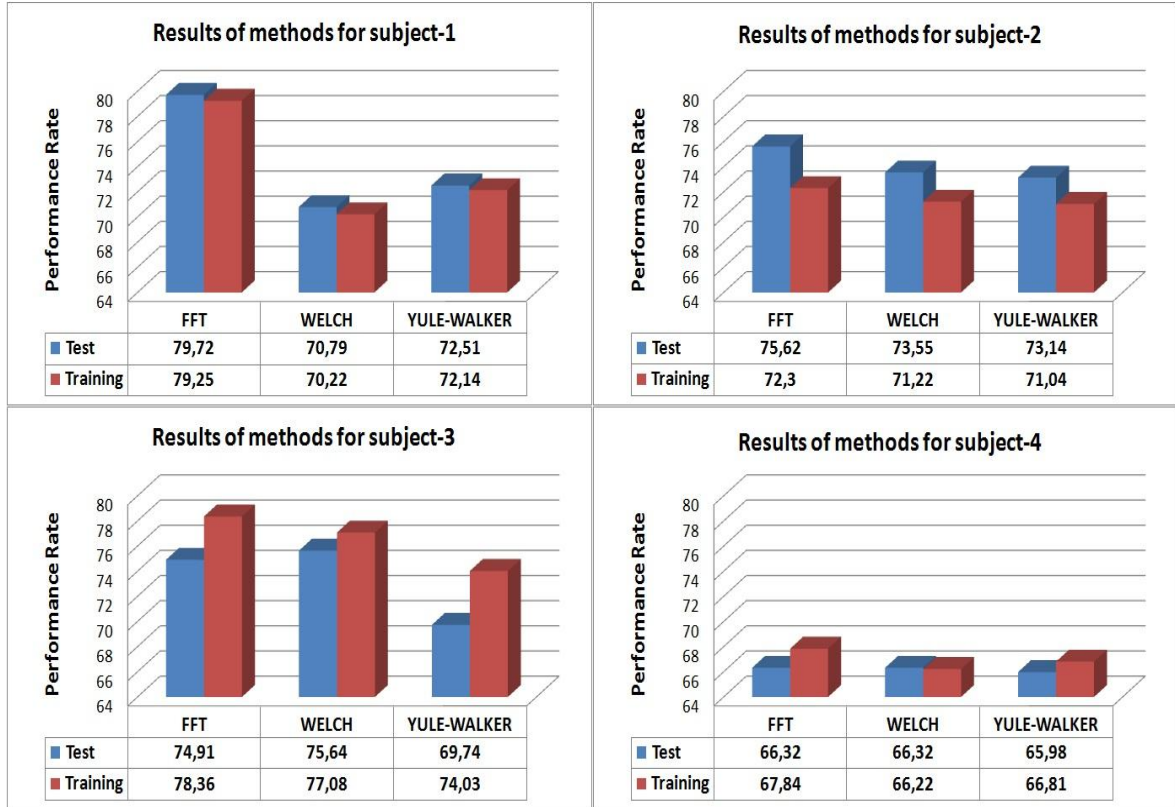


Figure 3. The effects of three methods on training and test classification accuracies for four subjects

As shown in Table 4, the general mean of test classification accuracies for all subjects, 74.14%, 71.58% and 70.34% in FFT, Welch and AR methods were obtained respectively. The results are also presented in Figure 4 and Figure 5 graphically.

Table 4. Obtained test classification results for all subjects

SUBJECTS		Subject-1	Subject-2	Subject-3	Subject-4	MEAN OF TAs (%)
		TA %	TA %	TA %	TA %	
METHODS	FFT	79,72	75,62	74,91	66,32	74,14
	WELCH	70,79	73,55	75,64	66,32	71,58
	YULE WALKER	72,51	73,14	69,74	65,98	70,34

TA: Classification accuracy of test datas (Test Accuracy)

EFFECT OF SOME POWER SPECTRAL DENSITY ESTIMATION METHODS ON AUTOMATIC SLEEP STAGE SCORING USING ARTIFICIAL NEURAL NETWORKS

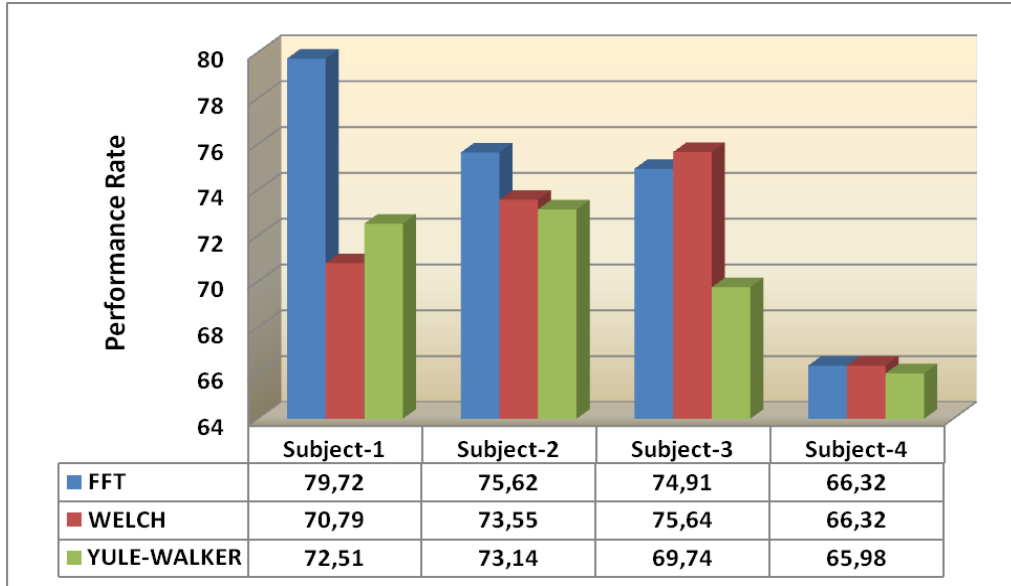


Figure 4. The effects of three methods on test classification accuracies for four subjects

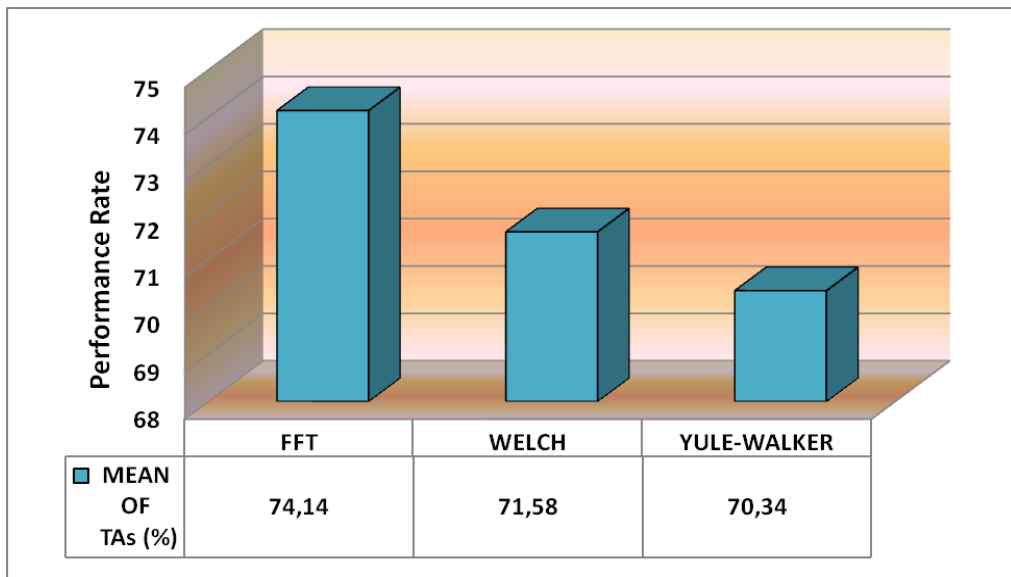


Figure 5. The effects of three methods on general mean of test classification accuracies for four subjects

As seen from the Figure 5, the general mean of test classification accuracies for four subjects, FFT method was obtained as the most efficient method. Welch and Yule-Walker methods followed this method.

## 4. CONCLUSION

In this study, EEG, EMG and EOG signals of four subjects were obtained to use in sleep staging. The effects of three PSD estimation methods were scrutinized on sleep stage classification by ANN. 20 Features extracted from time and frequency domain EEG, EMG and EOG signals were presented to ANN and optimum ANN parameters were searched for each subject's data. According to the results, the maximum test classification accuracies were recorded as 79.72% with FFT, 75.62% with FFT, 75.64% with Welch and 66.32% with FFT and Welch for subject-1, 2, 3 and 4 respectively. The mean values for PSD methods among four subjects are not different from these results. Again FFT stands out as better performed method in our application. As stated in the results section, this is a rough comparison. Better parameter arrangements would give different results but under the conclusions of this study, we can say that FFT can be used as an efficient PSD estimation method for sleep stage classification applications. Also, the results obtained in this study can be considered good according to the outcomes of similar studies in the literature.

We suggest that same system and methods of feature extracting can be applied the different types of data and signal, too.

## ACKNOWLEDGEMENT

This study is supported by the Scientific Research Projects of Selcuk University.

## REFERENCES

- AASM - American Academy of Sleep Medicine Task Force, 1999. *Sleep-Related Breathing Disorders in Adults: Recommendations for Syndrome Definition and Measurement Techniques in Clinical Research*. USA: SLEEP, Vol. 22, No. 5, pp 667-689.
- Acharya, R. et al, 2005. Non-Linear analysis of EEG signals at various sleep stages. *Computer Methods and Programs in Biomedicine*, 80, pp 37-45.
- Agarwal, R. and Gotman, J., 2001. Computer-Assisted Sleep Staging. *IEEE Transactions on Biomedical Engineering*, 48, pp 1412-1423.
- Becq, G. et al, 2005. Comparison Between Five Classifiers for Automatic Scoring of Human Sleep Recordings. *Studies in Computational Intelligence*, 4, pp 113-127.
- Estrada, E. et al, 2004. EEG Feature Extraction for Classification of Sleep Stages. in *Proc. 26<sup>th</sup> annual conference of the IEEE EMB*, San Francisco, USA.
- Flexer, A. et al, 2005. A reliable probabilistic sleep stager based on a single EEG signal. *Artificial Intelligence in Medicine*, 33, pp 199-207.
- He, W.X. et al, 2005. Nonlinear feature extraction of sleeping EEG signals. in *Proc. 27<sup>th</sup> annual conference of IEEE Engineering in Medicine and Biology*, Shangai, China.
- Holzmann, C.A. et al, 1999. Expert-system classification of sleep/waking states in infants. *Medical & Biological Engineering & Computing*, 37, pp 466-476.
- Jiayi, G. et al, 2007. Sample entropy analysis of aleep EEG under different states. in *Proc. IEEE/ICME International Conference on Complex Medical Engineering*, China.

EFFECT OF SOME POWER SPECTRAL DENSITY ESTIMATION METHODS ON AUTOMATIC SLEEP STAGE SCORING USING ARTIFICIAL NEURAL NETWORKS

- Oropesa, E. et al, 1999. Sleep stage classification using wavelet transform and neural network. *International Computer Science Institute*, TR-99-008.
- Ozsen, S., 2012. Classification of Sleep Stages Using Class-Dependent Sequential Feature Selection and Artificial Neural Network. *Neural Computing & Applications (ISI)* , DOI: 10.1007/s00521-012-1065-4.
- Pereda, E. et al, 1998. Non-Linear behavior of human EEG: fractal exponent versus correlation dimension in awake and sleep stages. *Neuroscience Letters*, 250, pp 91-94.
- Piñero, P. et al, 2004. Sleep stage classification using fuzzy sets and machine learning techniques. *Neurocomputing*, 58-60, pp 1137-1143.
- Rechtschaffen, A. and Kales, A., 1969. A Manual of Standardized Terminology. *Techniques and Scoring System for Sleep Stages of Human Subjects* , US Government Printing Office, Washington.
- Shen, Y. et al, 2003. Dimensional complexity and spectral properties of the human sleep EEG. *Clinical Neurophysiology*, 114, pp 199-209.
- Sinha, R.K. et al, 2003. A power spectrum based backpropagation artificial neural network model for classification of sleep-wake stages in rats. *OJHAS*, Vol. 2, No. 1, pp 1-8.
- Sinha, R.K., 2008. Artificial Neural Network and Wavelet Based Automated Detection of Sleep Spindles, REM Sleep and Wake States. *Journal of Medical Systems*, 32, pp 291 – 299.
- Šušmáková, K. and Krakovská, A., 2008. Discrimination Ability of Individual Measures Used in Sleep Stages Classification. *Artificial Intelligence in Medicine*, 44, pp 261-277.
- Wikipedia, 2012. *Welch's method*. [online] Available at: <[http://en.wikipedia.org/wiki/Welch's\\_method](http://en.wikipedia.org/wiki/Welch's_method)> [Accessed 26 January 2013].
- Wikipedia, 2013a. *Fast Fourier transform*. [online] Available at: <[http://en.wikipedia.org/wiki/Fast\\_Fourier\\_transform](http://en.wikipedia.org/wiki/Fast_Fourier_transform)> [Accessed 26 January 2013]
- Wikipedia, 2013b. *Autoregressive model*. [online] Available at: <[http://en.wikipedia.org/wiki/Autoregressive\\_model](http://en.wikipedia.org/wiki/Autoregressive_model)> [Accessed 26 January 2013].
- Van Quyen, M.L. et al, 2003. Exploring the nonlinear dynamics of the brain. *Journal of Physiology*, 97, pp 629-639.