PREDICTIVE MODELS IN DIAGNOSIS OF ALZHEIMER'S DISEASE FROM EEG

Lucie Tylova[a,*], Jaromir Kukal[a], Oldrich Vysata[b]

[a] *Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Department of Software Engineering in Economy, Trojanova 13, 120 00 Praha 2, Czech Republic*

[b] *Department of Computing and Control Engineering, Faculty of Chemical Engineering, Institute of Chemical Technology in Prague*

* corresponding author: `tylovluc@fjfi.cvut.cz`

Abstract. The fluctuation of an EEG signal is a useful symptom of EEG quasi-stationarity. Linear predictive models of three types and their prediction error are studied via traditional and robust measures. The resulting EEG characteristics are applied to the diagnosis of Alzheimer's disease. Our aim is to decide among: forward, backward, and predictive models, EEG channels, and also robust and non-robust variability measures, and then to find statistically significant measures for use in the diagnosis of Alzheimer's disease from EEG.

Keywords: Alzheimer's disease, EEG, linear predictive model, quasi-stationarity, robust statistics, multiple testing, FDR.

1. Introduction

Dementia is a set of clinical symptoms, e.g., memory loss and communicative difficulties. Two main categories are cortical and subcortical dementias. The most important cortical dementia, which accounts for about 50 % of the cases, is Alzheimer's disease. In patients with Alzheimer's disease, brain cells die quickly and the chemistry and structure of the brain is changed. EEG is used to study and detect abnormalities caused by dementia.

Biological rest is an endogenously dynamic process. Transient EEG events identify and quantify brain electric microstates as time epochs with a quasi-stable field topography [1]. We can assume better predictability inside microstates, lower predictability during changes between microstates. Higher fluctuations of EEG predictability may be connected with higher frequency of microstate changes.

Falk et al. [2] analysed an EEG signal via its envelope. They first constructed the modulation spectrum and found the region of significant spectral peaks (SP). This technique achieves accuracy 81.3 % with sensitivity 85.7 % and specificity 72.7 %. After the Hilbert transform, they also calculated the percentage modulation energy (PME) with better accuracy 90.6 %, sensitivity 90.5 %, and specificity 90.9 %.

Another approach was used by Ahmadlou et al. [3]. The first step in their approach was based on wavelet decomposition. The resulting patterns were processed by the Visibility Graph Algorithm (VGA). The power spectrum of the VGA structures was used for feature extraction. Two types of classifier (RBFNN, PCA-RBFNN) were used for the final decision. The accuracy was 96.50 % with sensitivity 100 % and specificity 87.03 % in the case of RBFNN. PCA-RBFNN increased the accuracy to 97.75 % with sensitivity 100 % and specificity 91.08 %.

2. Models

The main hypothesis of this work is that predictability of the brain activity differs between groups of patients with Alzheimer's disease (AD) and normal controls (CN). The activity of the human brain is measured via multichannel EEG, which produces a time series. On the basis of the quasi-stationarity of the EEG signal, the time series were decomposed into non-overlapping segments of constant length. Each segment of a given EEG channel and of each patient produced a short time series, the properties of which were studied via linear autoregressive models of three types.

2.1. Predictive model

Let m be the length of a segment. Let n be the model size as a number of parameters. Let $x_1, \ldots, x_m$ be an EEG [4] data segment. The linear predictive model has the form

$$x_k = \sum_{i=1}^{n} a_i x_{k-i} + e_k,, \tag{1}$$

for $k = n+1, \ldots, m$, where e_k is the model error in the k-th measurement and a_i is the model parameter for $i = 1, \ldots, n$. Formula (1) represents the traditional AR (autoregressive) model [5].

2.2. Back-predictive model

The predictive AR model (1) can also be used in the opposite time direction. The resulting model is

$$x_k = \sum_{i=1}^{n} a_i x_{k+i} + e_k, \tag{2}$$

Ch	Predictive			Back-predictive			Symmetric		
	STD	MAD$_1$	MAD$_2$	STD	MAD$_1$	MAD$_2$	STD	MAD$_1$	MAD$_2$
1	0.1027	0.0402	0.0314	0.0711	0.0338	0.0270	0.0503	0.0131	0.0074
2	0.0065	0.0016	$6.52 \cdot 10^{-4}$	0.0031	0.0012	$5.06 \cdot 10^{-4}$	0.0015	$2.16 \cdot 10^{-4}$	$6.00 \cdot 10^{-5}$
3	0.0121	0.0038	0.0014	0.0141	0.0035	0.0013	0.0172	0.0010	$2.24 \cdot 10^{-4}$
4	0.1408	0.0612	0.0337	0.1470	0.0540	0.0308	0.0635	0.0081	0.0029
5	0.2551	0.1906	0.1277	0.2573	0.1690	0.1141	0.2063	0.0867	0.0441
6	0.0643	0.0476	0.0275	0.0647	0.0391	0.0223	0.0417	0.0165	0.0064
7	0.0279	0.0192	0.0103	0.0232	0.0166	0.0086	0.0288	0.0214	0.0091
8	0.0917	0.1619	0.1290	0.0947	0.1474	0.1152	0.0572	0.0908	0.0664
9	0.1780	0.2093	0.1512	0.1815	0.1797	0.1308	0.1862	0.0832	0.0504
10	0.6823	0.8572	0.8429	0.7739	0.8309	0.8136	0.6093	0.6226	0.5553
11	0.2358	0.1763	0.1203	0.2218	0.1540	0.1046	0.1234	0.0527	0.0255
12	0.0910	0.0598	0.0467	0.0924	0.0545	0.0446	0.0359	0.0285	0.0216
13	0.1183	0.2376	0.1806	0.0953	0.2120	0.1607	0.0997	0.1113	0.0602
14	0.1027	0.1964	0.1744	0.1114	0.1779	0.1595	0.0706	0.0827	0.0558
15	0.2297	0.2925	0.2539	0.2363	0.2521	0.2174	0.1673	0.1517	0.0985
16	0.4478	0.5942	0.5282	0.4009	0.5395	0.4806	0.3636	0.3136	0.2170
17	0.0680	0.1197	0.1094	0.0437	0.1070	0.0965	0.0288	0.0304	0.0175
18	0.0418	0.0634	0.0595	0.0545	0.0694	0.0654	0.0296	0.0299	0.0219
19	0.2875	0.3889	0.3288	0.2483	0.3431	0.2868	0.1506	0.1568	0.1025

TABLE 1. Traditional fluctuation measures.

where e_k is again the model error, but for $k = 1, \ldots, m - n$.

2.3. SYMMETRIC MODEL

The third AR model is symmetric, and thus with lower prediction error for smooth signals. Supposing n is even, the adequate model is

$$x_k = \sum_{i=1}^{n/2} a_i x_{k-i} + \sum_{i=1}^{n/2} a_{n/2+i} x_{k+i} + e_k, \qquad (3)$$

where e_k is the model error for $k = n/2 + 1, \ldots, m - n/2$.

2.4. MODEL ERROR

The three AR models above are easily comparable, because they produce an overdetermined system of $M = m - n$ linear equations for n unknown variables $a_1, \ldots, a_n$. The unknown parameters $a_1, \ldots, a_n$ were estimated by the method of least squares (LSQ) [6] and the residues $r_1, \ldots, r_M$ are determined as the difference between an observed value and a predicted value. The estimate of the prediction error inside the given segment is

$$s_e = \sqrt{\frac{\sum_{i=1}^{M} r_i^2}{M - n}}. \qquad (4)$$

3. FLUCTUATION OF THE MODEL ERROR

Three basic characteristics were used to characterize EEG fluctuations: the standard deviation (STD), the mean of the absolute differences from the mean value

(MAD$_1$), and the mean of the absolute differences from the median value (MAD$_2$). However, these characteristics are excessively sensitive to outlier values. We preferred robust measures of EEG fluctuations: the median of the absolute differences from the median (MAD$_3$), the interquartile range (IQR), and the first quartile of the absolute mutual differences (MED).

Let N be the number of EEG signal segments. Let $\boldsymbol{s} = (s_1, s_2, \ldots, s_N)$ be the vector of errors (4) in all segments. Let Q$_1$, Q$_2$, Q$_3$, E be the first, the second, and the third quartile and mean value functions. The fluctuation criteria are defined as

$$\mathrm{STD} = (\mathrm{E}(\boldsymbol{s} - \mathrm{E}(\boldsymbol{s}))^2)^{1/2}, \qquad (5)$$
$$\mathrm{MAD}_1 = \mathrm{E}(|\boldsymbol{s} - \mathrm{E}(\boldsymbol{s})|), \qquad (6)$$
$$\mathrm{MAD}_2 = \mathrm{E}(|\boldsymbol{s} - \mathrm{Q}_2(\boldsymbol{s})|), \qquad (7)$$
$$\mathrm{MAD}_3 = \mathrm{Q}_2(|\boldsymbol{s} - \mathrm{Q}_2(\boldsymbol{s})|), \qquad (8)$$
$$\mathrm{IQR} = \mathrm{Q}_3(\boldsymbol{s}) - \mathrm{Q}_1(\boldsymbol{s}), \qquad (9)$$
$$\mathrm{MED} = \mathrm{Q}_1(|s_i - s_j|). \qquad (10)$$

We obtained the STD, MAD$_1$, MAD$_2$, MAD$_3$, IQR, and MED values of the model fluctuations of each channel for all AD and CN patients. The null hypothesis H$_0$: $\mu_{\mathrm{AD}} = \mu_{\mathrm{CN}}$ was tested via a two-sample t-test [7] against the alternative H$_A$: $\mu_{\mathrm{AD}} \neq \mu_{\mathrm{CN}}$. Here, $\mu_{\mathrm{AD}} = \mathrm{E} \ln \mathit{fluctuation}$ (5–10) for AD group and $\mu_{\mathrm{CN}} = \mathrm{E} \ln \mathit{fluctuation}$ (5–10) for CN group.

4. EXPERIMENTAL PART

Groups of 26 AD and 139 CN patients were used for testing. We used the international 10–20 electrode system with constant sampling frequency 200 Hz. A predictive model (1), a back-predictive model (2), and a

Ch	Predictive			Back-predictive			Symmetric		
	STD	MAD$_1$	MAD$_2$	STD	MAD$_1$	MAD$_2$	STD	MAD$_1$	MAD$_2$
1	0.0029	0.0265	0.0018	0.0035	0.0236	0.0021	$2.6 \cdot 10^{-4}$	0.0025	$2.3 \cdot 10^{-4}$
2	$6.9 \cdot 10^{-6}$	$6.2 \cdot 10^{-5}$	$\mathbf{4.8 \cdot 10^{-6}}$	$1.0 \cdot 10^{-5}$	$5.1 \cdot 10^{-5}$	$\mathbf{4.8 \cdot 10^{-6}}$	$\mathbf{3.9 \cdot 10^{-7}}$	$\mathbf{3.0 \cdot 10^{-6}}$	$\mathbf{5.1 \cdot 10^{-7}}$
3	$\mathbf{3.5 \cdot 10^{-6}}$	$6.5 \cdot 10^{-5}$	$\mathbf{3.5 \cdot 10^{-6}}$	$\mathbf{1.7 \cdot 10^{-6}}$	$4.2 \cdot 10^{-5}$	$\mathbf{3.7 \cdot 10^{-6}}$	$\mathbf{1.8 \cdot 10^{-7}}$	$\mathbf{1.6 \cdot 10^{-6}}$	$\mathbf{3.0 \cdot 10^{-7}}$
4	$2.4 \cdot 10^{-4}$	0.0019	$2.2 \cdot 10^{-4}$	$6.6 \cdot 10^{-4}$	0.0026	$3.4 \cdot 10^{-4}$	$\mathbf{4.8 \cdot 10^{-6}}$	$4.6 \cdot 10^{-5}$	$9.1 \cdot 10^{-6}$
5	0.0017	0.0124	0.0022	0.0038	0.0170	0.0025	$1.4 \cdot 10^{-4}$	0.0015	$2.1 \cdot 10^{-4}$
6	$2.9 \cdot 10^{-4}$	0.0047	$4.1 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$	0.0039	$3.2 \cdot 10^{-4}$	$2.4 \cdot 10^{-5}$	$2.4 \cdot 10^{-4}$	$3.2 \cdot 10^{-5}$
7	$2.4 \cdot 10^{-4}$	0.0025	$1.9 \cdot 10^{-4}$	$1.7 \cdot 10^{-4}$	0.0019	$1.7 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$	0.0014	$2.6 \cdot 10^{-4}$
8	0.0478	0.0787	0.0390	0.0495	0.0679	0.0406	0.0174	0.0384	0.0204
9	0.0159	0.0490	0.0127	0.0130	0.0387	0.0123	0.0013	0.0066	0.0023
10	0.8614	0.6785	0.7914	0.8522	0.6462	0.7958	0.3281	0.2948	0.3613
11	0.0038	0.0227	0.0034	0.0021	0.0151	0.0031	$1.8 \cdot 10^{-4}$	0.0015	$2.4 \cdot 10^{-4}$
12	0.0054	0.0182	0.0082	0.0066	0.0201	0.0121	0.0051	0.0100	0.0083
13	0.0177	0.0722	0.0212	0.0201	0.0730	0.0219	$7.0 \cdot 10^{-4}$	0.0085	$6.9 \cdot 10^{-4}$
14	0.0713	0.1341	0.0873	0.0676	0.1056	0.0885	0.0040	0.0180	0.0053
15	0.0877	0.1351	0.0791	0.0581	0.0994	0.0631	0.0028	0.0139	0.0050
16	0.2547	0.2882	0.2583	0.2338	0.2464	0.2512	0.0131	0.0368	0.0195
17	0.0307	0.0740	0.0359	0.0277	0.0676	0.0338	$4.3 \cdot 10^{-4}$	0.0038	$4.7 \cdot 10^{-4}$
18	0.0511	0.1400	0.0317	0.0451	0.1313	0.0375	0.0032	0.0257	0.0033
19	0.0491	0.1666	0.0492	0.0448	0.1625	0.0439	0.0017	0.0253	0.0021

TABLE 2. Robust fluctuation measures.

	Predictive	Back-predictive	Symmetric
MAD$_3$	3	3	2, 3, 4
IQR			2, 3
MED	2, 3	2, 3	2, 3

TABLE 3. Significant channels.

symmetric model (3) were identified, and the model errors (4) and their fluctuations were studied for segment length $m = 150$ and model size $n = 50$. The number of EEG segments varied patient-by-patient and satistisfied the inequality $352 \leq N \leq 762$.

The significance level for the testing was $\alpha = 0.001$. The hypotheses of mean equity were tested on 19 EEG channels, three predictive models, and six fluctuation characteristics. This is a kind of multiple testing, with 342 potentially dependent tests. The standard False Discovery Rate (FDR) methodology [8] was used to eliminate the acceptance of a false hypothesis.

The corrected critical value was determined as $\alpha_{\mathrm{FDR}} = 4.8347 \cdot 10^{-6}$. The t-test results (p_{value}) for traditional measures are included in Tab. 1. Results for robust measures are collected in Tab. 2. Bold font was used for a p_{value} below the critical probability α_{FDR}. The null hypothesis was rejected only in channels 2, 3, 4, which correspond to the frontal domain of the human brain. Only three robust fluctuation characteristics are significant: $\ln \mathrm{MAD}_3$, $\ln \mathrm{IQR}$, and $\ln \mathrm{MED}$.

The second channel is significant only for $\ln \mathrm{MED}$ or symmetrical prediction. The third channel is significant only for $\ln \mathrm{MED}$, $\ln \mathrm{MAD}_3$ or symmetric prediction. The fourth channel is significant only for $\ln \mathrm{MAD}_3$ together with symmetrical prediction. Tab. 4 summarizes the results.

The best $p_{\mathrm{value}} = 1.8885 \cdot 10^{-7}$ was obtained on the third channel for the symmetric model and the $\ln \mathrm{MAD}_3$ criterion. Figure 1 shows its Receiver operating characteristic (ROC) curve [9]. The area under the curve (AUC) is 0.77, which evaluates the model as good. The boxplot in Fig. 2 displays the differences between AD and CN patients.

5. DISCUSSION

While the autoregressive model is linear and requires a stationary signal, the higher fluctuation of the model error in Alzheimer's subjects may reflect a different structure of brain microstates than in healthy subjects. It may reflect alterations in the brain anatomical cortical connectivity in resting-state networks.

In contrast to applying robust methods and filters, the autoregressive linear model offers a simple and traditional solution that provides results with a sufficient level of significance. These results could also be influenced by the small group of testing data.

FIGURE 1. ROC for ln MAD_3
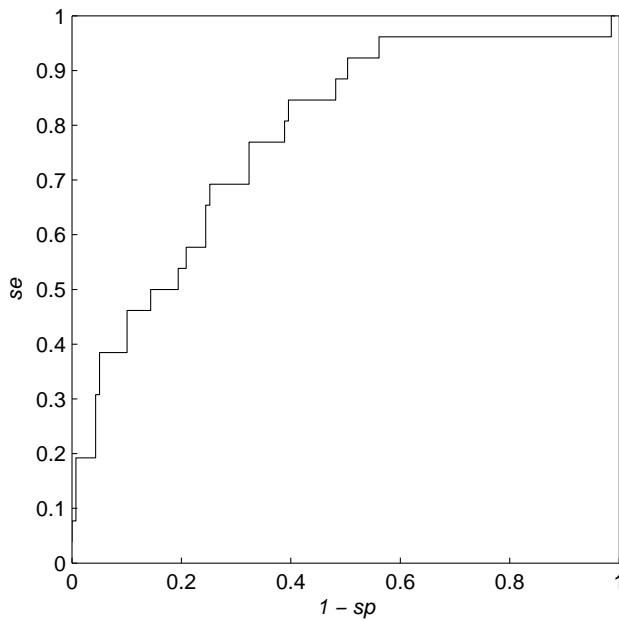
FIGURE 2. Boxplot diagram for ln MAD_3

6. CONCLUSION

Using a symmetric predictive model of the EEG signal and MAD_3, IQR, and MED robust measures of predictive error fluctuations, we recognize significant differences between AD and CN groups in the case of frontal electrodes, which are represented by the second, the third, and the fourth channel of EEG.
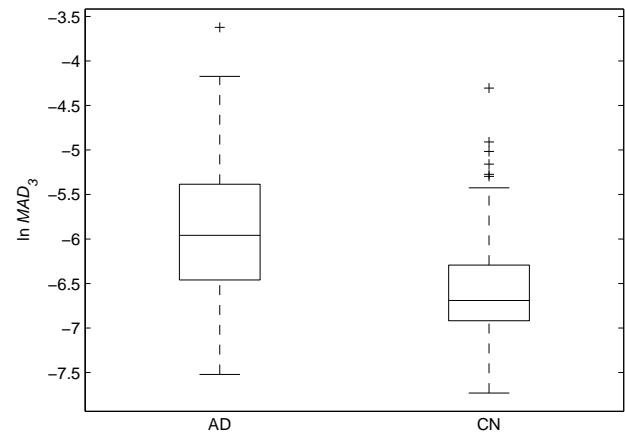
This result is directly applicable to the diagnosis of Alzheimer's disease. The accuracy of our method is 63.64 % with sensitivity 84.62 % and specificity 59.71 %. However, methods based on the modulation spectrum and the Hilbert transform [2] or on the Visibility Graph Algorithm [3] are better in accuracy, sensitivity, and specificity.

ACKNOWLEDGEMENTS

REFERENCES

[1] Musso, F., Brinkmeyer, J., Mobascher, A., Warbrick, T., Winterer, G., Spontaneous brain activity and EEG microstates. A novel EEG/fMRI analysis approach to explore resting-state networks, *Neuroimage*, Vol.52, No.4, 2010, pp. 1149-1161.

[2] Falk, T. H., et al., EEG amplitude modulation analysis for semiautomated diagnosis of Alzheimer's disease, *EURASIP Journal on Advances in Signal Processing*, 2012 2012:192.

[3] Ahmadlou, M., Adeli, H., Adeli, A., New diagnostic EEG markers of the Alzheimer's disease using visibility graph, *EURASIP Journal on Advances in Signal Processing*, 2010 2010:177.

[4] Niedermeyer, E., Lopes da Silva, F., *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* , Lippincott Williams & Wilkins, 2005.

[5] Priestley, M. B., *Non-linear and Non-stationary Time Series Analysis*, Academic Press, 1988.

[6] Björck, A., *Numerical Methods for Least Squares Problems*, SIAM, 1996.

[7] Meloun, M., Militky, J., *The statistical analysis of experimental data*, Academia, 2004.

[8] Benjamini, Y., Hochberg, Y., Controlling the false discovery rate: A practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society*, Vol.57, No.1, 1995, pp. 289-300.

[9] Fawcett, T., An Introduction to ROC Analysis, *Pattern Recognition Letters*, Vol.27, No.8, 2006, pp. 861-874.