Article

Discrimination of Severity of Alzheimer's Disease with Multiscale Entropy Analysis of EEG Dynamics

Chang Francis Hsu [1][ID]**, Hsuan-Hao Chao** [1][ID]**, Albert C. Yang** [2,3]**, Chih-Wei Yeh** [1]**, Long Hsu** [1] **and Sien Chi** [4,*][ID]

[1]   Department of Electrophyics, National Chiao Tung University, Hsinchu 30010, Taiwan; francis-920@hotmail.com (C.F.H.); edsfjk@hotmail.com (H.-H.C.); stevesteve227@gmail.com (C.-W.Y.); long@nctu.edu.tw (L.H.)

[2]   Institute of Brain Science/Digital Medicine Center, National Yang-Ming University, Taipei 11221, Taiwan; accyang@gmail.com

[3]   Division of Interdisciplinary Medicine and Biotechnology, Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, MA 02215, USA

[4]   Department of Photonics, National Chiao Tung University, Hsinchu 30010, Taiwan

[*]   Correspondence: schi@mail.nctu.edu.tw; Tel.: +886-3-5731824

check for updates

Abstract: Multiscale entropy (MSE) was used to analyze electroencephalography (EEG) signals to differentiate patients with Alzheimer's disease (AD) from healthy subjects. It was found that the MSE values of the EEG signals from the healthy subjects are higher than those of the AD ones at small time scale factors in the MSE algorithm, while lower than those of the AD patients at large time scale factors. Based on the finding, we applied the linear discriminant analysis (LDA) to optimize the differentiating performance by comparing the resulting weighted sum of the MSE values under some specific time scales of each subject. The EEG data from 15 healthy subjects, 69 patients with mild AD, and 15 patients with moderate to severe AD were recorded. As a result, the weighted sum values are significantly higher for the healthy than the patients with moderate to severe AD groups. The optimal testing accuracy under five specific scales is 100% based on the EEG signals acquired from the T4 electrode. The resulting weighted sum value for the mild AD group is in the middle of those for the healthy and the moderate to severe AD groups. Therefore, the MSE-based weighted sum value can potentially be an index of severity of Alzheimer's disease.

Keywords: electroencephalography; Alzheimer's disease; multiscale entropy; linear discriminant analysis

1. Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder as well as the most prevalent form of age-related dementia in the modern society [1]. As of 2018, there were an estimated 50 million people with AD worldwide. This number will increase to around 82 million in 2030, and 152 million in 2050 [2]. Increasing focus on detecting or differentiating AD using sophisticated analysis of electroencephalography (EEG) signals has emerged in recent years. Optimization of EEG analysis is critical for developing wearable devices to screen the AD patients in low cost and non-invasive manners. Selection of crucial EEG channels can further help in designing new wearable devices and saving computational resource.

Multiscale entropy (MSE) and MSE-based analyses of EEG signals from AD patients have been explored in many studies [3–13]. MSE measures the complexity of a physiologic time series under different time scales, and reflects the degree of healthiness of a biological system through its output

physiologic signals [14,15]. Many studies observed that the MSE values of EEG signals from healthy subjects are higher than those from the AD patients at small scales, while lower than those from the AD patients at large scales [3–6]. Besides, the slope of MSE vs. scale plot at large scales was found higher for the AD patients than the healthy subjects [6].

Recently, machine learning technique has been further applied to the EEG analysis to improve the classification performance of the AD patients with different severities, including healthy subjects [7,8]. Each time scale of MSE may serve as a feature in the machine learning model. Tzimourta et al. adopted 38 features for machine learning for each EEG channel, which included MSE features as well as other spectral and temporal features extracted from the EEG signals. In total, 24 EEG records were collected from healthy, mild AD, and moderate AD groups. Five binary and one ternary classification problems were addressed [8]. Fan et al. adopted 380 MSE features extracted from 19 EEG channels. The EEG signals from each channel contributed a sequence of 20 different MSE values computed at scales 1–20, separately. In total, 123 EEG records were collected from healthy, very mild AD, mild AD, and moderate to severe AD groups. Six binary classification problems were addressed [8].

Instead of binary or ternary classification, we propose a new AD severity index to measure the degree of AD severity as additional information for clinical doctor by using linear discriminant analysis (LDA). The corresponding machine learning model is so simple that only 2–5 MSE features are required to be extracted from a particular EEG channel. This benefits the extraction of critical time scales as the MSE features in the model.

In this study, we collected the EEG data recorded from 15 healthy subjects, 69 patients with mild AD, and 15 patients with moderate to severe AD. Our study consisted of three parts. In the first part, we used leave-one-out cross validation (LOOCV) method to test the performance of the LDA in differentiating patients with moderate to severe AD from the healthy subjects. In the second part, we obtained the AD severity index by training with the healthy and moderate to severe AD groups. In the third part, we obtained the AD severity index by training with the healthy and mild AD groups.

2. Materials and Methods

2.1. Participants

The study group comprised 84 patients with AD (49 women, 35 men) and with a mean age of 77.0 years (SD: 8.0) who were recruited from the Dementia Clinic at the Neurological Institute, Taipei Veterans General Hospital in Taiwan. The diagnosis for AD was based on the criteria of the National Institute of Neurological and Communicative Disorders and the Stroke/Alzheimer's Disease and Related Disorders Association (McKhann et al., 1984). All patients had received neurological examinations, laboratory tests, EEG monitoring, and neuroimaging evaluation during the diagnostic process. Our study was approved by the Institutional Review Board of Taipei Veterans General Hospital to conduct retrospective analysis of the patients' clinical and EEG data. We excluded patients who had other conditions that caused secondary dementia, such as vascular dementia, Parkinson's disease, hypothyroidism, vitamin B12 deficiency, syphilis, and prior history of major psychiatric illness (e.g., major depression, bipolar disorder, or schizophrenia).

The severity of dementia was assessed by the Clinical Dementia Rating (CDR) scale (Morris, 1993). Patients were categorized as having AD that was mild (CDR = 1; N = 69) or moderate to severe (CDR $\geq$ 2; N = 15). The control group included 15 healthy participants (9 women, 6 men) with a mean age of 69.9 years (SD: 9.5). All healthy subjects were cognitively normal and symptom-free, and they were free of neurological disease or psychiatric illness. For convenience, the healthy control, mild AD, and moderate to severe AD groups are called for short as HC, AD1, and AD2, respectively, in the following context.

2.2. EEG Recordings and Preprocessing

All participants had received routine EEG recording (Nicolet EEG, Natus Medical, Incorporated, San Carlos, CA, USA) in the EEG examination room at the Neurological Institute of Taipei Veterans General Hospital. The EEG recording protocol began with a 5-min habituation to the examining environment, followed by three sessions of 10 s with the eyes closed and then open, and a session of photo stimulus. Recordings were in accord with the international 10–20 system with linked ear reference, 256 Hz sampling rate, high pass filter of 0.05 Hz, low pass filter of 70 Hz, notch filter of 60 Hz, and impedance below 3 kΩ. We recorded 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2). The signals from these 19 electrodes were referenced to linked earlobe electrodes. Vigilance was monitored by the EEG technician, who alerted patients when signs of drowsiness appeared in the tracings. Vertical eyeball movement was detected from electrodes placed above and below the right eye, with the horizontal analog detected from electrodes placed at the left outer canthus. The EEG signals were exported in European Data Format and were processed using MATLAB software (Mathworks Inc., Sherborn, MA, USA).

For each subject, 19 electrodes were used to record the EEG signals of the subject. Each electrode recorded three epochs with each of 10 s for a total of $256 \times 10 = 2560$ data points. The EEG signals length of 10 s is consistent with many other studies with length within 5–12 s [4–9]. Empirical mode decomposition (EMD) method was used for detrending.

2.3. Multiscale Entropy (MSE) Algorithm

MSE is a widely used complexity measure for physiologic signals. MSE measures the irregularity of a time series under different time scales. MSE was proposed as a measure of the complexity of a physiologic time series. It aims to reflect the degree of healthiness of a biological system through its output physiologic signals [14–16]. The algorithm of the MSE analysis method of a time series consists of three steps [5,6]. The first step of MSE method is to divide a time series $\{x_i\}$ into N/τ consecutive and non-overlapping windows $w_j^{(\tau)}$ with each of equal length τ. The second step is to calculate the average value of the τ elements in each window according to

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i, \quad 1 \leq j \leq N/\tau . \tag{1}$$

Thus, all the average values of the N/τ windows form a new coarse-grained time series $\{y_j^{(\tau)}\}$.

The third step is to calculate the MSE value of the original series $\{x_i\}$ as the sample entropy of the coarse-grained time series $\{y_j^{(\tau)}\}$ as follows. Consider the m-length vectors $u_m(p) = \{y_p^{(\tau)}, y_{p+1}^{(\tau)}, \ldots, y_{p+m-1}^{(\tau)}\}$, where $1 \leq p \leq N/\tau - m + 1$. Another vector $u_m(q \neq p)$ is considered similar to $u_m(p)$ if the following equation is satisfied

$$max\{|\, y_{p+k}^{(\tau)} - y_{q+k}^{(\tau)} \,| : 0 \leq k \leq m-1\} \leq r \tag{2}$$

The parameter m is the sequence length, and the parameter r is the similarity criterion. Then, set $n_p^m(r)$ as the number of vectors $u_m(q \neq p)$, which are similar to the template vector $u_m(p)$. The MSE value of the original time series $\{x_i\}$ is given by

$$MSE(\tau) = ln \frac{\sum_p^{N-m} n_p^m}{\sum_p^{N-m} n_p^{m+1}} \tag{3}$$

The above equation can be considered as the irregularity of the coarse-grained time series $\{y_j^{(\tau)}\}$, which is equal to the negative of the natural logarithm of the conditional probability that sequences close to each other for m consecutive data points will also be close to each other when one more point is added to each sequence [15]. The best settings of the parameters $(m, r)\, are\, (2, 0.15)$ in analyzing the

EEG signals from the AD and healthy subjects [4]. Thus, the same settings of $m = 2$ and $r = 0.15$ were used in the following analysis.

2.4. Feature Extraction in Linear Discriminant Analysis (LDA)

The scale factor τ was limited between 1 and 20 in this study. For each subject, the overall MSE value of the EEG signals from an EEG electrode at each τ were obtained as the average of the three MSE values derived from three epochs from the electrode, separately. As a result, there were 20 overall MSE values for τ from 1 to 20 per each EEG electrode for each subject.

For the analysis of each electrode, the number of selected MSE scales S was from 1 to 5 as input features of LDA. We adopted LDA to derive the weighting factors of MSE values under each scale such that the weighted sum values for AD and healthy groups were well separated. In each of the $S = 1$ to $S = 5$ cases, the combinations of scales among the 20 scales were exhaustively examined. The signals from the 19 electrodes were examined one by one.

LDA is a technique used to reduce a high dimensional feature set to a lower dimensional feature set, such that the groups can be more easily separated in the lower dimensional space. LDA has an analytical solution via matrices calculation without optimization process, which makes LDA efficient [17].

2.5. Performance Matrix

We adopted some indices that are often used in binary classification:

$$Specificity = \frac{TN}{TN + FP} \, , \tag{4}$$

$$Recall(Sensitivity) = \frac{TP}{TP + FN} \, , \tag{5}$$

$$Precision = \frac{TP}{TP + FP} \, , \tag{6}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \, , \tag{7}$$

$$F1 = \frac{2 \cdot Recall \cdot Precision}{Recall + Precision} \, . \tag{8}$$

where TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively.

2.6. Analysis Procedure

Our study consists of three parts as follows:

1. Leave-one-out cross validation (LOOCV) method was used to test the performance of the LDA in differentiating 15 HC and 15 AD2 subjects.
2. The expected AD severity indices were obtained by training 15 HC and 15 AD2 subjects using LDA. Then, the models obtained were applied to all the HC, AD1, and AD2 groups to compare their weighted sum values.
3. The 69 AD1 subjects were divided into training and validation sets with 54 and 15 subjects, respectively, to obtain the AD severity indices. Then, the models obtained were applied to the 15 HC, 15 AD1, and 15 AD2 subjects to compare their weighted sum values.

3. Results

3.1. LOOCV Performance in Differentiating the HC from AD2 Subjects

Table 1 illustrates the optimal F1 scores of LOOCV in differentiating subjects into the HC and AD2 groups of each electrode using 1–5 MSE scales (labeled as 1S–5S in the table), respectively. The highest F1 score obtained from one MSE scale corresponded to electrodes T5 and P3, while all the highest scores obtained using 2–5 MSE scales corresponded to electrode T4.

Table 1. Optimal F1 scores of LOOCV in differentiating subjects into the HC and AD2 groups of each electrode obtained using 1–5 MSE scales (labeled as 1S–5S), respectively. Bold highlights the best performances.

Case \ Electrode	1S	2S	3S	4S	5S
Fp1	0.710	0.774	0.903	0.903	0.933
Fp2	0.733	0.759	0.815	0.857	0.857
F7	0.710	0.828	0.933	0.938	0.938
F3	0.710	0.800	0.889	0.889	0.889
Fz	0.710	0.815	0.903	0.903	0.903
F4	0.714	0.800	0.828	0.897	0.897
F8	0.750	0.897	0.933	0.966	0.966
T3	0.733	0.786	0.867	0.867	0.857
C3	0.741	0.800	0.867	0.875	0.897
Cz	0.750	0.774	0.867	0.839	0.857
F4	0.714	0.800	0.828	0.897	0.897
F8	0.750	0.897	0.933	0.966	0.966
T3	0.733	0.786	0.867	0.867	0.857
C3	0.741	0.800	0.867	0.875	0.897
Cz	0.750	0.774	0.867	0.839	0.857
C4	0.800	0.867	0.867	0.897	0.933
T4	0.786	**0.938**	**0.966**	**0.968**	**1.000**
T5	**0.828**	0.903	0.903	0.903	0.933
P3	**0.828**	0.933	0.966	0.966	0.966
Pz	0.786	0.897	0.903	0.903	0.903
P4	0.759	0.903	0.933	0.933	0.968
T6	0.621	0.750	0.750	0.774	0.774
O1	0.800	0.897	0.897	0.897	0.897
O2	0.667	0.690	0.667	0.667	0.645

3.2. AD Severity Index Obtained by Training HC and AD2 Groups

Table 2 illustrates the optimal training F1 scores in differentiating the HC and AD2 groups among the 19 electrodes for each of the 1S–5S cases, respectively. The corresponding MSE scales, specificities, recalls, precisions, and accuracies are also illustrated. Table 3 further shows the weighted sum models for the 2S–5S cases, respectively. Note that the differentiating performance is the same as the performance of LOOCV listed in Table 1.

Table 2. Optimal F1 score and corresponding accuracy, precision, recall, specificity, selected electrode, and scales for each of the 1S–5S cases, respectively.

Case	Electrode	Scales	F1 Score	Accuracy	Recall	Precision	Specificity
1S	T5; P3	{12}; {19}	0.828	0.833	0.813	0.867	0.800
2S	T4	{6, 15}	0.938	0.933	0.882	1.000	0.867
3S	T4	{2, 3, 14}	0.966	0.967	1.000	0.933	1.000
4S	T4	{6, 9, 12, 17}	0.968	0.967	0.938	1.000	0.933
5S	T4	{2, 8, 11, 15, 16}	1.000	1.000	1.000	1.000	1.000

Table 3. The weighted sum models for the 2S–5S cases.

Case	Weighted Sum Model
2S	$0.82 \cdot \text{MSE}\,(\tau = 6) - 0.58 \cdot \text{MSE}\,(\tau = 15) - 0.56$
3S	$0.61 \cdot \text{MSE}\,(\tau = 2) + 0.79 \cdot \text{MSE}\,(\tau = 3) - 0.10 \cdot \text{MSE}\,(\tau = 14) - 0.29$
4S	$0.43 \cdot \text{MSE}\,(\tau = 6) + 0.54 \cdot \text{MSE}\,(\tau = 9) - 0.67 \cdot \text{MSE}\,(\tau = 12) + 0.25 \cdot \text{MSE}\,(\tau = 17) - 1.13$
5S	$0.22 \cdot \text{MSE}\,(\tau = 2) + 0.67 \cdot \text{MSE}\,(\tau = 8) - 0.30 \cdot \text{MSE}\,(\tau = 11) - 0.55 \cdot \text{MSE}\,(\tau = 15) + 0.32 \cdot \text{MSE}\,(\tau = 16) - 0.69$

Figure 1 illustrates the optimal training result of the 2S case listed in Table 3. Figure 1a illustrates the MSE ($\tau = 6$) vs. the MSE ($\tau = 15$) of the 30 EEG signals recorded from the T4 electrode. The 30 signals are from 15 HC and 15 AD2 subjects. Figure 1b illustrates the weighted sum values of the 30 signals. The red dashed line is the differentiating boundary with threshold value 0; the data points above the red line with positive values are differentiated into the healthy group and the data points below the red line with negative values are differentiated into the AD group. In other words, the absolute value of the weighted sum value of each data point shown in Figure 1b is the distance between the corresponding point and the boundary line shown in Figure 1a. Note that the differentiating performance is the same as the performance of LOOCV listed in Table 1.
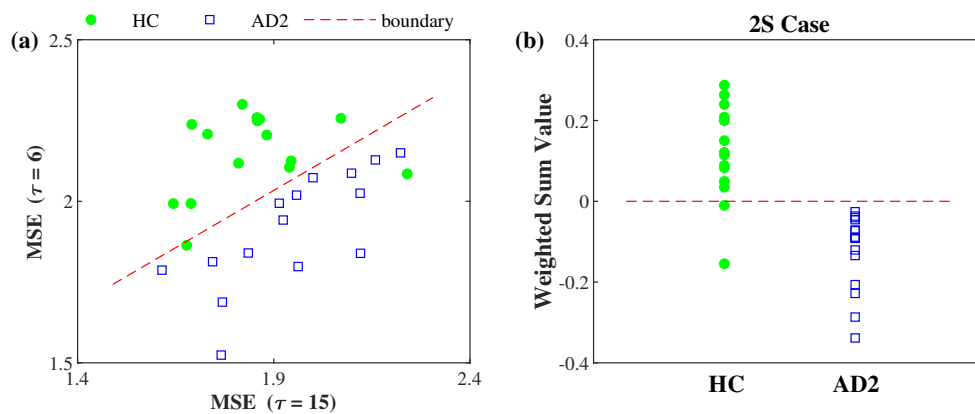

Figure 1. (**a**) The MSE ($\tau = 6$) vs. the MSE ($\tau = 15$) of the 30 EEG signals recorded from the T4 electrode. The 15 circle and 15 square symbols are from 15 healthy subjects and 15 patients with moderate to severe AD, respectively. The red dashed line is the differentiating boundary. (**b**) The weighted sum values of the 30 signals. The red dashed line is the differentiating boundary with threshold value 0.

Figure 2 illustrates the optimal training results of the 3S–5S cases listed in Table 3. Note that the differentiating performances are the same as the performances of LOOCV method listed in Table 1.
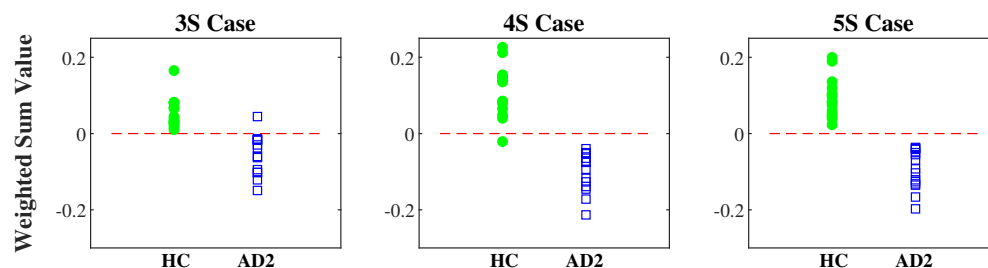

Figure 2. The weighted sum values of the 30 EEG signals recorded from the T4 electrode for the 3S–5S cases. The 15 circle and 15 square symbols are from 15 HC and 15 AD2 subjects, respectively.

Table 4 further compares the means $\pm$ SDs of the weighted sum values among the HC, AD1, and AD2 groups for each of the 2S–5S cases. For each case, the weighted sum value of the healthy group is significantly higher than that of the mild AD group, while that of the mild AD group is significantly higher than that of the moderate to severe AD group (p-value < 0.05).

Table 4. Comparison of MSE-based weighted sum values among the HC, AD1, and AD2 groups. (The values are labeled as mean $\pm$ SD.)

Group / Case	HC N = 15	AD1 N = 69	AD2 N = 15
2S	0.13 ± 0.12	0.05 ± 0.14	-0.13 ± 010
3S	0.05 ± 0.04	0.00 ± 0.05	-0.05 ± 0.05
4S	0.12 ± 0.08	0.06 ± 0.11	-0.12 ± 0.08
5S	0.09 ± 0.05	0.03 ± 0.12	-0.09 ± 0.05

The weighted sum models listed in Table 3 were too complicated for us to understand the underlying meaning. We thus sought a simplified one. The absolute values of the two weighting coefficients shown in the 2S case in Table 3 are not too far away to each other, thus we simplified the weighted sum model to

$$MSE(\tau = 6) - MSE(\tau = 15). \tag{9}$$

Figure 3 shows the differentiating result with the same data as in Figure 1 using the simplified weighted sum model. The corresponding F1 score is 0.933, which is similar to that of the 2S case listed in Table 2. The means $\pm$ SDs of the weighted sum values of the HC, AD1, and AD2 groups are 0.30 ± 0.17, 0.21 ± 0.21, and -0.03 ± 0.13, respectively. The order of the weighted sum values is consistent with that shown in Table 4.
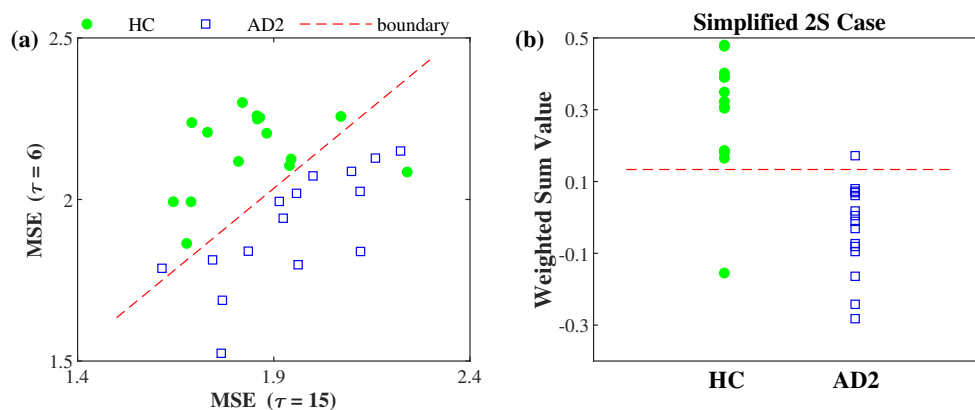

Figure 3. The differentiating result with the same data as in Figure 1 using simplified weighted sum model as MSE ($\tau = 6$) $-$ MSE ($\tau = 15$). The F1 score is 0.933. (**a**) The MSE ($\tau = 6$) vs. the MSE ($\tau = 15$) of the 30 EEG signals recorded from the T4 electrode. The 15 circle and 15 square symbols are from 15 healthy subjects and 15 patients with moderate to severe AD, respectively. The red dashed line is the differentiating boundary. (**b**) The weighted sum values of the 30 signals. The red dashed line is the differentiating boundary with threshold value 0.134.

3.3. AD Severity Index Obtained by Training HC and AD1 Groups

Table 5 illustrates the optimal training F1 scores in differentiating the HC and AD1 groups among the 19 electrodes for each of the 3S–5S cases, respectively. The training set was composed of 15 HC and 54 AD1 subjects, where the 54 AD1 subjects were randomly selected from the 69 total patients in the AD1 group, and 15 AD1 subjects remained for validation. Table 6 further shows the weighted sum models for the 2S–5S cases, respectively.

Table 5. Optimal F1 score and corresponding accuracy, precision, recall, specificity, selected electrode, and scales for each of the 1S–5S cases.

Case	Electrode	Scales	F1 Score	Accuracy	Recall	Precision	Specificity
3S	F8	{2, 3, 12}	0.854	0.797	0.976	0.759	0.933
4S	F7	{2, 5, 6, 12}	0.857	0.797	0.955	0.778	0.867
5S	F7	{1, 2, 4, 15, 17}	0.857	0.797	0.955	0.778	0.867

Table 6. The weighted sum models for the 3S–5S cases.

Case	Weighted Sum Model
3S	$-0.63 \cdot \text{MSE}\,(\tau = 2) + 0.76 \cdot \text{MSE}\,(\tau = 3) - 0.16 \cdot \text{MSE}\,(\tau = 12) - 0.06$
4S	$-0.31 \cdot \text{MSE}\,(\tau = 2) + 0.86 \cdot \text{MSE}\,(\tau = 5) - 0.36 \cdot \text{MSE}\,(\tau = 6) - 0.21 \cdot \text{MSE}\,(\tau = 12 - 0.06$
5S	$-0.10 \cdot \text{MSE}\,(\tau = 1) - 0.47 \cdot \text{MSE}\,(\tau = 2) + 0.78 \cdot \text{MSE}\,(\tau = 4) - 0.40 \cdot \text{MSE}\,(\tau = 15) - 0.05 \cdot \text{MSE}\,(\tau = 17)$

Table 7 compares the means $\pm$ SDs of the weighted sum values among the HC, AD1, and AD2 groups for each of the 3S–5S cases. The validation set was composed of 15 HC, 15 AD1, and 15 AD2 subjects. For each case, the weighted sum value of the healthy group is significantly higher than that of the mild AD group, while that of the mild AD group is significantly higher than that of the moderate to severe AD group (p-value < 0.05).

Table 7. Comparison of MSE-based weighted sum values among the HC, AD1, and AD2 groups. (The values are labeled as mean $\pm$ SD.)

Group Case	HC N = 15	AD1 N = 15	AD2 N = 15
3S	0.02 ± 0.02	-0.03 ± 0.04	-0.07 ± 0.04
4S	0.28 ± 0.04	0.24 ± 0.06	-0.19 ± 0.03
5S	0.03 ± 0.08	-0.05 ± 0.10	-0.13 ± 0.07

Figure 4 illustrates the weighted sum values for the validation set of the 3S–5S cases listed in Table 5.
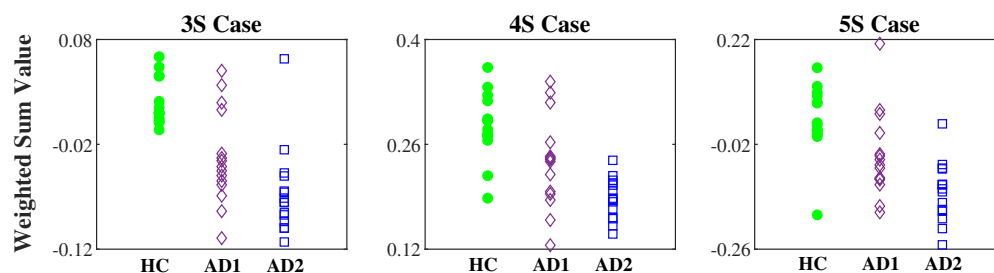
Figure 4. The weighted sum values of the 45 EEG signals recorded from the optimal electrode for the 3S–5S cases. The 15 circle, 15 diamond, and 15 square symbols are from 15 HC, 15 AD1, and 15 AD2 subjects, respectively.

4. Discussion

The results above show that using only a few MSE features as the input of LDA provides a high classification performance in differentiating the healthy subjects from the AD patients. For example, with only two MSE scales at 6 and 15, the F1 score is 0.933 in differentiating the HC and AD2 groups using the simple weighted sum model "$MSE(\tau = 6) - MSE(\tau = 15)$" to reflect the degree of health, as listed in Equation (9). The model indicates that the weighted sum values, the AD severity indices, of the EEG signals from the healthy subjects are higher than those from the AD ones. This implies that the MSE values of the healthy are higher than those of the AD patients at small scales, while lower

than those of the AD patients at large scales. The tendency of the results is consistent with those in the previous studies [3–6].

As to the F1 scores and the related differentiating performances, Table 1 shows that they do not significantly increase from two to five MSE scales for the 19 EEG electrodes, in general. For example, the F1 scores derived from the EEG signals from electrode T4 are 0.938, 0.966, 0.968, and 1.000 under 2 to 5 MSE scales, respectively. Table 4 shows that the significance levels are similar in differentiating the HC and AD1 groups with 2 to 5 scales. Similarly, Figure 4 illustrates that the performances in differentiating the HC, AD1, and AD2 groups do not improved from three to five scales. Indeed, the differentiating performances are high enough using 2 to 3 MSE scales and may not be positively correlated with the number of scales selected [18]. This implies that a simple model is sufficient to reflect the AD severity. Thus, the simple machine learning models in this study provide a basis for designing new simple algorithms to measure the health degree of EEG signals in the future.

MSE was first proposed as a measure of complexity, which reflects the ability of a biological system to adapt and function in an ever-changing environment. Thus, a signal from a diseased system is expected to exhibit a lower complexity than that from a healthy system. Considering heart rate time series as an example, diseased systems suffering congestive heart failure (CHF) exhibit relatively regular behavior. On the contrary, diseased atrial fibrillation systems exhibit highly erratic fluctuations with statistical properties resembling uncorrelated noise. As a result, under large scales τ around 20, the MSE values of the RR interval series from a healthy group are higher than those of the two diseased groups of CHF and AF. However, under between two and five scales, the MSE values for CHF, healthy, and AF subjects are small, medium, and large, respectively [14,15,19]. For comparison, the complexity of the healthy remain higher than those of the CHF and AF at every scale for another analysis called entropy of entropy (EoE) [20]. These particular features imply that MSE is a measure of disorder under relatively small scales and a measure of complexity under relatively large scales. Therefore, the EEG signals from AD patients exhibit low disorder and high complexity. Further explorations are needed to clarify the meaning of complexity of EEG signals.

Three limitations of our study exist. First, although HC and AD2 groups were well separated, as shown in the Results Section, the accuracy in differentiating all the HC, AD1, and AD2 groups still needs to be improved. Second, the EEG signals were only recoded with length of 10 s; longer EEG segment might provide more information and benefit classification. Third, the EEG signals are combinations of cortical electric activities and cannot clearly localize the underlying pathology. Future study of AD classification with EEG source localization may help to identify the specific brain regions of interest.

5. Conclusions

Using LDA, we propose a MSE-based weighted sum value as an index of severity of Alzheimer's disease with only 2–5 MSE features extracted from a single EEG channel. In the case of 19-electrode EEG signals from 15 AD2 and 15 HC, we adopted MSE and LDA to differentiate the AD subjects from the healthy ones. For a single MSE scale, the best differentiating performance corresponds to a F1 score of 0.828 for both electrodes T5 and P3. For 2–5 MSE scales, the best differentiating performances correspond to the F1 scores of 0.938, 0.966, 0.968, and 1.000 for electrode T4. Further, the MSE-based weighted sum values of the HC, AD1, and AD2 groups are large, medium, and small, respectively.

References

1. Kulkarni, N; Bairagi, V. *EEG-Based Diagnosis of Alzheimer Disease: A Review and Novel Approaches for Feature Extraction and Classification Techniques*; Academic Press: Cambridge, MA, USA, 2018.
2. Patterson, C. *World Alzheimer Report 2018: The State of the Art of Dementia Research: New Frontiers*; Alzheimer's Disease International: London, UK, 2018.
3. Mizuno, T.; Takahashi, T.; Cho, R.Y.; Kikuchi, M.; Murata, T.; Takahashi, K.; Wada, Y. Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy. *Clin. Neurophysiol.* **2010**, *121*, 1438–1446. [CrossRef] [PubMed]
4. Yang, A.C.; Wang, S.-J.; Lai, K.-L.; Tsai, C.-F.; Yang, C.-H.; Hwang, J.-P.; Lo, M.-T.; Huang, N.E.; Peng, C.-K.; Fuh, J.-L. Cognitive and neuropsychiatric correlates of EEG dynamic complexity in patients with Alzheimer's disease. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2013**, *47*, 52–61. [CrossRef] [PubMed]
5. Azami, H.; Abásolo, D.; Simons, S.; Escudero, J. Univariate and multivariate generalized multiscale entropy to characterise EEG signals in Alzheimer's disease. *Entropy* **2017**, *19*, 31. [CrossRef]
6. Escudero, J.; Abásolo, D.; Hornero, R.; Espino, P.; López, M. Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy. *Physiol. Meas.* **2006**, *27*, 1091. [CrossRef] [PubMed]
7. Tzimourta, K.D.; Giannakeas, N.; Tzallas, A.T.; Astrakas, L.G.; Afrantou, T.; Ioannidis, P.; Grigoriadis, N.; Angelidis, P.; Tsalikakis, D.G.; Tsipouras, M. G.EEG window length evaluation for the detection of Alzheimer's disease over different brain regions. *Brain Sci.* **2019**, *9*, 81. [CrossRef] [PubMed]
8. Fan, M.; Yang, A.C.; Fuh, J.L.; Chou, C. A.Topological pattern recognition of severe Alzheimer's disease via regularized supervised learning of EEG complexity. *Front. Neurosci.* **2018**, *12*, 1–10. [CrossRef] [PubMed]
9. Song, Z.; Deng, B.; Wang, J.; Wang, R. Biomarkers for Alzheimer's Disease Defined by a Novel Brain Functional Network Measure. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 41–49. [CrossRef] [PubMed]
10. Tang, B. Comment on "analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy". *Physiol. Meas.* **2007**, *28*, 9–11. [CrossRef] [PubMed]
11. Tsai, P.H.; Chang, S.C.; Liu, F.C.; Tsao, J.; Wang, Y.H.; Lo, M.T. A novel application of multiscale entropy in electroencephalography to predict the efficacy of acetylcholinesterase inhibitor in Alzheimer's disease. *Comput. Math. Methods Med.* **2015**, *2015*. [CrossRef] [PubMed]
12. Azami, H.; Fernández, A.; Escudero, J. Refined multiscale fuzzy entropy based on standard deviation for biomedical signal analysis. *Med. Biol. Eng. Comput.* **2017**, *55*, 2037–2052. [CrossRef] [PubMed]
13. McBride, J.; Zhao, X.; Munro, N.; Jicha, G.; Smith, C.; Jiang, Y. EEG multiscale entropy dynamics in mild cognitive impairment and early Alzheimer's disease. In Proceedings of the 2014 Biomedical Sciences and Engineering Conference, Oak Ridge, TN, USA, 6–8 May 2014; pp. 1–4. [CrossRef]
14. Costa, M.; Goldberger, A.L.; Peng, C.-K. Multiscale entropy analysis of complex physiologic time series. *Phys. Rev. Lett.* **2002**, *89*, 68102. [CrossRef] [PubMed]
15. Costa, M.; Goldberger, A.L.; Peng, C.-K. Multiscale entropy analysis of biological signals. *Phys. Rev. E* **2005**, *71*, 021906. [CrossRef] [PubMed]
16. Richman, J.; Moorman, J. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol.-Heart Circ. Physiol.* **2000**, *278*, H2039–H2049. [CrossRef] [PubMed]
17. Fisher, R.A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.* **1936**, *7*, 179-188. [CrossRef]
18. Chao, H.-H.; Yeh, C.-W.; Hsu, C.F.; Hsu, L.; Chi, S. Multiscale Entropy Analysis with Low-Dimensional Exhaustive Search for Detecting Heart Failure. *Appl. Sci.* **2019**, *9*, 3496. [CrossRef]
19. Hsu, C.F.; Lin, P.Y.; Chao, H.-H.; Hsu, L; Chi, S. Average Entropy: Measurement of disorder for cardiac RR interval signals. *Physica A* **2019**, *529*, 121533. [CrossRef]
20. Hsu, C.F.; Wei, S.-Y.; Huang, H.-P.; Hsu, L.; Chi, S.; Peng, C.K. Entropy of Entropy: Measurement of Dynamical Complexity for Biological Systems. *Entropy* **2017**, *19*, 550. [CrossRef]