A discriminant analysis of the P3b wave with EEG by feature-electrode pairs in schizophrenia diagnosis

Juan I. Arribas[a,b], Luis M. San-José-Revuelta[a,*]

[a] *Department of Electrical Engineering, University of Valladolid, 47011 Spain.*
[b] *Castilla-Leon Neuroscience Institute, University of Salamanca, 37007 Spain.*

Abstract

Schizophrenia is a disease that affects approximately 1% of the population. Its early accurate diagnosis is of vital importance to apply adequate therapy as soon as possible. We present a Statistical Discriminant Diagnosing (SDD) system that discriminates between healthy controls and subjects and that supports diagnosis by a medical professional. The system works with {*feature,electrode*} EEG pairs which are selected based on the statistical significance of the p-values computed over the brain P3b wave. A bank of evoked potential pre-processed and filtered EEG signals is recorded during an auditory odd-ball (AOD) task and serves as input to the SDD system. These EEG signals comprise 20 features and 17 electrodes, both in time (t) and frequency (f) domain. The relevance of the Parieto-Temporal region is shown, allowing us to identify highly discriminant {*feature,electrode*} pairs in the detection of schizophrenia, resulting lower p-values in both Right and Left Hemispheres, as well as in Parieto-Temporal EEG signals. See for instance, the {*PSE,P4*} pair, with p-value=0.00003 for (parametric) t Student and p-

*Corresponding author.
Email address: lsanjose@tel.uva.es (Luis M. San-José-Revuelta)

Preprint submitted to IET Signal Processing *April 27, 2023*

This is the final manuscript (post-print) accepted and published in IET Signal Processing, e12230, vol. 17, Issue 6, pp. 1-20 June 2023. ISSN: 1751-9675, eISSN : 1751-9683. DOI: 10.1049/sil2.12230. (c) John Wiley & Sons Ltd. The Institution of Engineering and Technology.

value=0.00019 for (nonparametric) U Mann-Whitney tests, both under the 15Hz cutoff frequency of a low pass EEG preprocessing filter. The relevance of this pair is in agreement with previously published related results. The proposed SDD system may provide the human expert (psychiatrist) with an objective complimentary information to help in the early diagnosis of schizophrenia.
Keywords: Statistical Discriminant Diagnosing (SDD), Decision Support Systems, EEG, evoked related potential (ERP), schizophrenia, statistical discriminant analysis.

---

1. Introduction

Schizophrenia is a mental disorder that is present in around 1% of global population. In accordance to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) from the American Psychiatric Association, schizophrenia subjects show various symptoms like delusions, hallucinations or disorganized speech [1]. In fact, schizophrenia is one of the most frequent and studied brain disorders due to its fast spreading [2]. Its diagnosis is a very complex task since a large number of factors are involved in its cause and prognosis. One of the current research fields is based on processing electroencephalograms (EEGs) –EEG signals allow to record the electrical brain activity generated in brain– in order to determine significant differences among patients (SZ) and healthy subjects (HC) [3], with the aim of being able to properly discriminate and classify between these two groups of persons [4]. Following this methodology, we will make a comparison of the EEGs recorded from an auditory task to perform a hypothesis test [5]. The P3b sub-component

of the P300 Evoked Related Potential (ERP) brain wave has already been determined in the past as a reliable tool for the diagnosis of schizophrenia by various authors, like, for instance, in [6] and [7]. The P3b event-related electric potential has already been related to working memory and attention deficits and has been successfully used in many recent works in schizophrenia diagnosis, like, for instance, in [8, 9, 10, 11, 12]. In these papers it is proved that schizo patients show an altered response to AOD tasks when the P3b and/or P3a evoked potentials are analyzed.

Non the less, these previous results mainly perform an statistical analysis of the P3b signals and parameters without making a proper discrimination or diagnosis among the various groups of subjects. In addition, the majority of previous works have used time domain (t) EEG features of the P3b wave only for their analysis, and, thus, make no use of any frequency (f) domain features. Only recent works seem to begin to use frequency features. In this work, we also propose to use statistical significance tests that can help to find different responses to P300 wave components (P3b) between healthy (HC) and schizo (SZ) patient groups, so as to determine signs of the disease, and to do it as soon as possible [13]. Schizophrenia is a difficult disease to treat and its early diagnosis and treatment is crucial in its favorable prognosis [14]. In our work, the evoked P3b EEG electromagnetic brain wave has been analyzed and registered using a seventeen electrode cap while an auditory odd-ball (AOD) task is performed. This technique has been widely validated in psychosis and schizophrenia previously.

The statistical significance analysis can be summarized as follows: in short, in a classification task, p-values account for the amount of over-

lap of class probability density functions (posterior pdf's) over the multi-dimensional input (feature) space once an optimized classification hyperplane has been fixed. This way, p-values are defined as the class pdf's overlap over the input space, which are commonly called pdf tails, in such a way that two different classes are said to be more statistically different whenever these overlaps of tails are smaller. This implies that the set of input features in the multi-dimensional input space is able to limit the pdf's overlap to one side and another of the classification hyperplane (classification thresholds).

In general, a hypothesis test is a systematic method of assessing beliefs about reality. This method requires the confrontation of such beliefs with real evidence and decide, in light of this evidence, if such beliefs may be reasonable or should be discarded and considered improbable.

Unfortunately, we have found only a few directly related references in the literature about the statistical significance analysis during an AOD task in P3b wave with EEG in schizophrenia. As related studies, we can cite the recent work by Das and Pachory [15], that uses p-values to validate features (entropies in this case) capability to detect schizo patients, or the paper by Siuly et al. (2020) [16], where a design methodology involving an empirical mode decomposition (EMD) technique is used for the diagnosis of schizo patients from EEG signals. Last paper also checks the significance of the extracted features using the well-known Kruskal-Wallis test. These works show the validity of statistical significance tests to evaluate the discriminant power of selected features.

On the other hand, our study results are properly aligned, and, at least consistent in part, with several relevant previously published papers in the

field, like, for instance, [17] and [18], which use deep learning algorithms to automatically extract significant features at the convolution stage and to classify them. Most significant features are extracted at the max-pooling stage, and signals are classified later on using a fully connected layer.

Also worth it to mention the paper by Roach et al. [19] regarding the use of simultaneous time (t) and frequency (f) domain parameters for proper discrimination, as well as the work by Lundin et al. [20], that studies early and late ERPs in time and frequency domains in a bipolar sample, with and without current psychotic features. In addition, Santos-Mayo et al. [21] proposed an automatic sub-optimal estimation of the most significant electrodes for SZ classification.

Recently, Shalbaf et al. [22] proposed the combination of frontal, central, parietal, and occipital regions applied to a Convolutional neural Network (CNN) for improving classification results, and Barros et al. [23] reviewed several artificial intelligence (AI)-based methods applied to schizophrenia classifications using EEG data. Among its conclusions is the proved capability of these AI systems to identify subjects at high-risk of psychosis conversion and to differentiate schizophrenia from other disorders. Besides, the recent work of Góngora-Alonso et al. [24] includes a literature review of Artificial Intelligence based methods applied to schizophrenia detection, though not based of statistical significance analysis. In this context, it is worth mentioning the recent work by Sairamya et al. [25], where the relaxed local neighbor difference pattern (RLNDiP) technique is proposed and a combination of RLNDiP features from both time domain and time−frequency domain is used. Prominent features describing the effective connectivity are selected

using the Kruskal-Wallis test and fed into an artificial neural network for classification.

Our final goal is to analyze to what degree and extent do the various {*feature,electrode*} pairs are able or not to discriminate schizophrenia using the P3b wave evoked potential signals, at an statistical significant level.

The remaining of the paper is organized as follows: first, Section 2.1 describes the dataset used for numerical simulations and the proposed SDD system; next, definitions concerning electrodes, electrode groupings and features are given in Section 2.2. The main statistical significance tests used in our work are shown in Section 2.3. Numerical simulations and comparative results are described in Section 3, whilst a discussion is included in Section 4. Finally, global conclusions and a brief outline of possible future work lines can be seen on Section 5.

2. Methods

2.1. Dataset acquisition and preprocessing

All EEG signals were registered while participants performed an auditory odd-ball task (AOD). A total of 47 EEG-ERP were recorded over the same number of subjects, from which 16 were SZ diagnosed patients (age = 36.3 $\pm$ 10.5 years) and 31 HC controls (age = 29.9 $\pm$ 9.8 years). Explicit written consent was used for SZ subject inclusion, and volunteer HC subjects received a small remuneration. Exclusion criteria comprised both subjects under 18 years old and recognition of having had drugs in the previous 48h. To properly generate and record both the P3a and P3b waves, EEGs were captured at a sampling rate of 250Hz while subjects went through a 3-stimuli auditory

odd-ball task, comprising a 500Hz target tone ($S3$), a 1000Hz distracter tone ($S2$) and a frequent 2000Hz standard tone ($S1$). Participants heard binaural tone bursts (5 ms rise and fall, 90 dB intensity, 50 ms duration) presented in a uniform random 1000 and 1500 ms stimuli asynchronous onset.

This AOD task consisted in subjects distinguishing among 3 audio stimuli randomly played: a target tone with probability 0.20 evoking the P3b brain wave in subjects; a distracter tone under probability 0.20; and a standard tone with probability 0.60. Participants were requested to stay calm, relaxed and with eyes closed, and were told to press a mouse button using index finger from dominant hand whenever they heard the target tone, discarding non-attended target tones. Previous advise was given to subjects so as to record EEG signals in which most of the power would come from true brain signal and not other noise sources, avoiding this way eye and muscular artifacts as much as possible. The purpose of the closing eyes instruction was to limit eye movement noise due to muscular electrical signals (micro-saccadic). Non the less, if any substantial eye movement artifacts still persist, they will be removed by the Independent Component Analysis (ICA) filtering phase and thus limited impact is expected in the results. EGG signals were pre-processed [26] and low-pass filtered up to 15Hz and 35Hz cutoff frequencies, as shown in the system block diagram, –see Figure 1. Next, a series of differentiating (discriminant) features were computed both in the time (t) and frequency (f) domains. Thus, we have two matrices, one for each filtering scheme, with 340 rows (17 electrodes $\times$ 20 computed features) $\times$ 47 columns, corresponding to the 47 subjects in total. EEGs signals were registered using a BrainVision equipment (Brain Products GmbH; München, Germany), with

a 17 tin electrode cap (Electro-Cap International, Inc.; Eaton, Ohio, USA), in compliance with the 10-20 international standard system [27]. The 17 cap electrode definitions were the following: $F3, F4, F7, F8, Fp1, Fp2, Fz$, $C3, C4, Cz, O1, O2, P3, P4, Pz, T5$ and $T6$. Electrode electrical impedance was kept under 5kΩ. Brain signal potentials were measured relative to the Cz electrode, at a sampling rate of 250Hz, and the EEG recordings were continuous during all AOD task time duration (no pause and no resume).

2.2. Electrodes, electrode groupings and features

The definition of electrodes, electrode groupings and features (characteristics) is next described, followed by a brief description of the main discriminant statistical hypothesis tests used in this work.

2.2.1. Feature definitions

For a proper formal definition of both time (t) and frequency (f) domain features, please see [21].

2.2.2. Electrodes and electrode groupings

A total of 17 electrodes, each one corresponding to a specific position on the head, are defined. These electrodes are grouped as follows (Figure 2):

- **Individual:** Each of the 17 electrodes are individually considered.
- **Total:** grouping formed by the 17 electrodes together.
- **Frontal**: electrodes of the front of the head: *FP1, FP2, F3, F4, F7, F8* and *Fz*.
- **Central:** electrodes of the central part of the head: *C3, C4* and *CZ*.

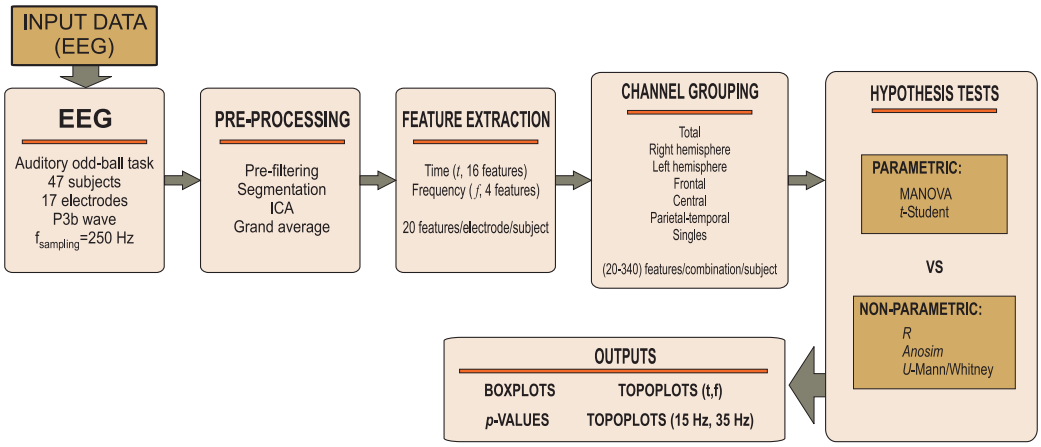


Figure 1: Block diagram of proposed Statistical Discriminant Diagnosing (SDD) system including the process followed during EEG data pre-processing, with either 15Hz or 35Hz cutoff lowpass filters [21]. A total of 47 subjects, 16 SZ and 31 HC, are volunteer enrolled (express written consent), and requested to perform a hearing task (AOD). The measurements carried out while in the auditory task, are recorded using a 17 electrode helmet, distributed over the head. A total of 20 parameters (features) were extracted: 16 in the time (t) domain and 4 in the frequency (f) domain, from which we carry out hypothesis tests (both parametric and non-parametric, both univariate and multivariate) in order to compute statistical significance parameters, including: p-values, p-values topoplots, and p-values boxplots, for single $\{feature, electrode\}$ pairs (univariate) and also for various electrode channel groupings (sets of electrodes, multivariate), both in time (t) and frequency (f) domain.

- **Parieto-Temporal:** electrodes of the parieto-temporal part of the head *P3, P4, PZ, T5* and *T6*.
- **Occipital:** electrodes of the occipital part of the head: *O1* and *O2*.
- **Right Hemisphere:** *FP2, C4, F4, F8, P4, O2* and *T6* are the 7 even electrodes that are located on the right half of the head.

- **Left Hemisphere:** *FP1, F3, F7, C3, P3, O1* and *T5* are the 7 odd electrodes that are located in the left half of the head.

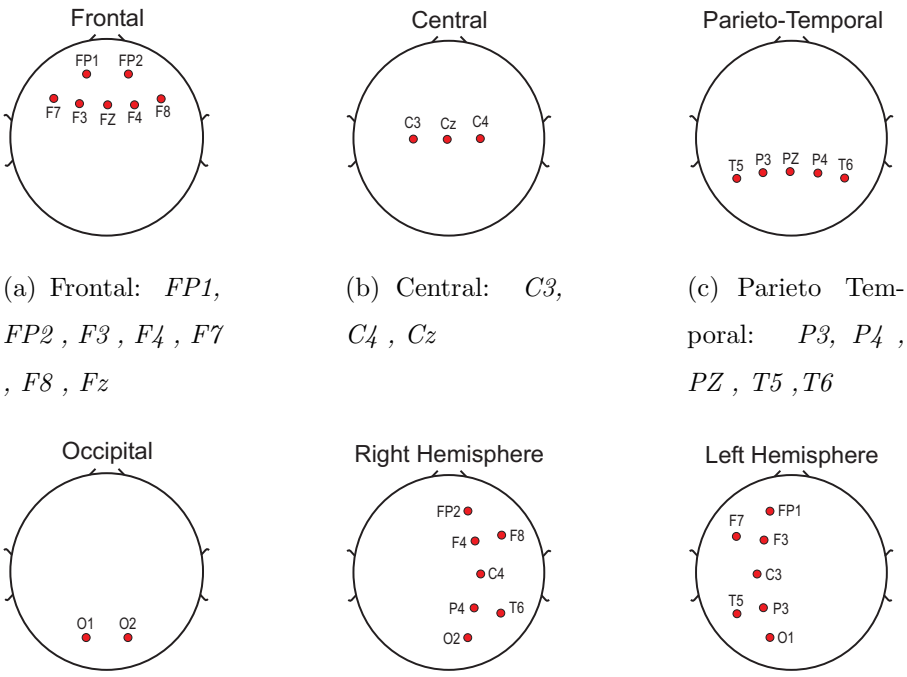
- **Sub-optimal Right Hemisphere:** formed by the top 3 electrodes of the right hemisphere. *FP2, F4* and *P4*.

- **Sub-optimal Left Hemisphere:** formed by the top 3 electrodes of the left hemisphere. *FP1, P3* and *F7*.

2.3. Description of the performed hypothesis tests

Parametric hypothesis tests methods (MANOVA, t-Student) make assumptions about data distribution, while non-parametric tests (R statistic, U-Mann/Whitney, and *Anosim*) are more general, making no assumptions about a specific source data distribution. Whenever needed, we have performed a non-parametric approach. In addition, one can compute univariate (single $\{feature, electrode\}$ pair) or multi-variate (several $\{feature, electrode\}$ pairs) estimations, the later implying several independent variables under analysis.

First, univariate tests for each {*electrode-feature*} pair, individually considered, were performed. The first step of these univariate tests consisted in determining the distribution of samples [28], in order to check the suitability of a parametric hypotheses test. However, due to the non-Gaussian distribution of the variables found –this fact has been proved in several papers, for instance in [29]– it was decided to carry out a non-parametric test [30]. The non-parametric test chosen was the Mann-Whitney test, which is based on comparing the medians of two populations along the independent variable.



(a) Frontal: *FP1, FP2 , F3 , F4 , F7 , F8 , Fz*

(b) Central: *C3, C4 , Cz*

(c) Parieto Temporal: *P3, P4 , PZ , T5 ,T6*

(d) Occipital: *O1, O2*

(e) Right Hemisphere: *FP2, F4, F8, C4, P4, O2, T6*

(f) Left Hemisphere: *FP1, F3, F7, C3, P3, O1, T5*

Figure 2: Electrode loci belonging to the main electrode grouping definitions: a) Frontal, b) Central, c) Parieto-Temporal, d) Occipital, e) Right Hemisphere, and f) Left Hemisphere electrode groupings.

After obtaining these results, MatLab® numerical computation software was used for data visualization in order to obtain meaningful conclusions.

After this phase, a set of electrodes and sub-optimal (discriminant) features were found. Here we also noted that data with discriminant power increased, so performing a multivariate test was needed. Accordingly, we first performed a discriminant analysis to obtain those pairs with higher dis-

criminant power. Then, the MANOVA test was implemented to quantify the discriminatory capacity of each of the pairs. Finally, the statistical analysis software R and the *Anosim* function were used to obtain more reliable results without assuming a normal distribution of the data [31].

2.3.1. Non-parametric Mann-Whitney test

Mann-Whitney non-parametric test makes assumptions about the similarities between equality of the populations' medians. We consider as the null hypothesis (H_0) that the two populations have the same median, and as the alternative hypothesis, that the medians are different [13]:

$$\begin{aligned} H_0 &: m_1 - m_2 = 0 \\ H_A &: m_1 - m_2 \neq 0 \end{aligned} \tag{1}$$

where parameters m_1 and m_2 are the medians of *population-1* and *population-2*, respectively. Our goal will be to reject the null hypothesis so as to be able to say that the two populations have different middle values and, hence, we can say that population medians are different. The test is defined as

$$T = S - \frac{n(n+1)}{2} \tag{2}$$

where n is the number of observations in sample X, and S is the sum of the *assigned hierarchies* to the sample observations of X. The *hierarchies* are assigned as follows: data is sorted from lowest to highest values, and the lowest value is assigned number 1. Remaining values are assigned a higher hierarchy whose value depends on the increase over the minor value.

This statistical test is not distributed according to any known function. This distribution is tabulated and its values are known as *quantiles* of the non-parametric Mann-Whitney statistical test [32].

The null hypothesis (H_0) is rejected if the calculated T value is under W_α, being W_α a critical value of T computed from the quantiles table of the Mann-Whitney statistical test.

2.3.2. Discriminant Analysis (DA)

The main aim of DA is to analyze whether differences between different groups exist, in our case between the SZ and HC groups, regarding the variables considered, and find out the ways these differences are [33]. DA offers a discriminant function of the type

$$D = \alpha X + \beta Y \tag{3}$$

where α and β are the weights of independent variables that make the function reach its highest value in one group, and its lowest one in the others.

Discriminant analysis gives us a discriminant function, which is able to classify a sample into one of the groups, but classification is not our goal. Our main purpose is to see to what extent each of the variables affects that function, and select those variables that have the most discriminatory power, i.e. to obtain a sorted array of the variables based on their discriminant capacity, and then use that information to compute a p-value using the MANOVA multivariate method.

Previous assumptions required to apply this method follow next: (i) data are distributed as a multivariate normal (Gaussian) pdf for each population, and (ii) covariance matrices are equal. These conditions are not verified in all cases of our problem in hand, but this does not constitute a limitation since this discriminant analysis technique is very robust and works well in practice although these restrictions are not met. For this reason, we believe

DA can be applied as a prior step to the MANOVA test execution.

2.3.3. MANOVA test

MANOVA test is a multivariate statistical method used to detect differences between groups of two populations. This method is based on the computation of the variance so as to detect whether populations are equal or not [34]. It consists in a test of multivariate hypotheses, where the following parameters are defined:

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 = ... = \mu_n \\ H_A &: \mu_1 \neq \mu_2 \neq ... \neq \mu_n \end{aligned} \tag{4}$$

with μ_i being the median of the $i-$th population.

There exist several different models of MANOVA, each of them using a different statistic parameter. Next, four types of contrast statistics are defined [35]:

- Trace of Pillai:

$$V = \prod_{i=1}^{s} \frac{\lambda_i}{1+\lambda_i} \tag{5}$$

  where λ_i stands for the eigenvalues of the data matrix.

- Lambda of Wilks:

$$\Lambda = \prod_{i=1}^{s} \frac{1}{1+\lambda_i} \tag{6}$$

  where λ_i denotes the eigenvalues of the data matrix.

- Trace of Hotelling:

$$T = \sum_{i=1}^{s} \lambda_i \tag{7}$$

where λ_i stands for the eigenvalues of the data matrix.

- Roy's largest root:

$$\Theta = max\{\lambda_i\} \tag{8}$$

where $max\{\lambda_i\}$ is the largest eigenvalue of the data matrix.

2.3.4. Anosim test

The *Anosim* function belongs to the *vegan* package of the R statistical software. This package was created for ecological studies of discrimination of natural species, hence it has high discriminatory power even in the absence of normality distribution of the independent variables [36]. In our analysis, the *Anosim* function of this package was used. This function allows us to check whether there are differences between two or more groups of data samples. These differences are quantified by providing a p-value [31] about data distributions. In order to apply the *Anosim* function, groups of samples on which we are interested in determining their dependence, are defined. It is assumed that distances between elements of different groups are larger than those between elements of the same group [37]. To check the significance of data, a permutation-based method in which samples are randomly changed from one group to another, is used. The $\mathcal{R}$ statistical is used in the analysis. This statistical parameter is based on the measured difference between data belonging to different groups (classes), r_B, and data into the same group r_W.

The following formula is applied to both data:

$$\mathcal{R} = 4\frac{r_B - r_W}{N(N-1)} \tag{9}$$

where r_B is the average of all between-groups distances, r_W is the average of all within-group distances and N stands for the total number of samples.

The main advantage of *Anosim* with respect to other parametric techniques, such as MANOVA, is that it is not necessary to make assumptions about data distributions. It is also robust and works well for small data samples, being often used in the analysis of biodiversity, a field of study where much accuracy is needed due to the variety of species found in this area.

Finally, another advantage of this test over other non-parametric methods is that it does not need to compute any matrix inversion, which could pose a problem in those methods that need to compute the determinant for the p-value. In these cases, if matrices are singular, such techniques could be applied (for instance, in those cases where data are linear combinations of each other). Despite the good performance of MANOVA, and thanks to all these advantages, the *Anosim* test can be considered as an improved version of the MANOVA test, and may even be considered as a non-parametric MANOVA version.

3. Results

In this section, univariate (single electrode) results are first shown, followed by multivariate results (electrode groupings), both in visual and numerical form and both at feature and electrode level, including the computation of p-values, p-value topoplots [38], and p-value boxplots, as outcomes.

3.1. Univariate results

Two hypothesis tests are included: the t Student and the Mann-Whitney U tests are here used to evaluate univariate (single variable) results.

3.1.1. t Student's test

First, the t of Student test has been carried out. It is therefore necessary to check whether the condition of data being normally distributed is fulfilled or not [39]. This test was performed with the goodness of fit test of Shapiro-Wilk.

Next, it is necessary to check whether the variances of the two populations (SZ and HC) are identical or different, and depending on this, a type of SPSS t of Student test will be selected. This check is performed using the Levene test, and results are shown in Table A.3 of Appendix A for the 15Hz pre-processing filter and in Table A.4 of Appendix A for the 35Hz case. For the sake of simplicity, only data corresponding to significant pairs are shown in both tables.

Once this test is performed, it was found that of the 720 $\{feature, electrode\}$ pairs of the two filters, only about 100 can be considered to be normally distributed, and, of these, only 13 (8 for the 15Hz filter and 5 for the 35Hz one) have a p-value ≤ 0.05. These pairs are shown in Appendix A Table A.5 for the 15Hz filter and in Appendix A Table A.6 for the 35Hz filter. In these two tables, the p-value of the Student's t test and the p-value of non-parametric Mann-Whitney test (which will be later explained) are compared.

The following conclusions can be drawn after this test is performed:

- Since normally distributed populations are only a very few compared

to the total, it was decided that a non-parametric test should be performed to obtain relevant results, since it does not need to meet any requirement.

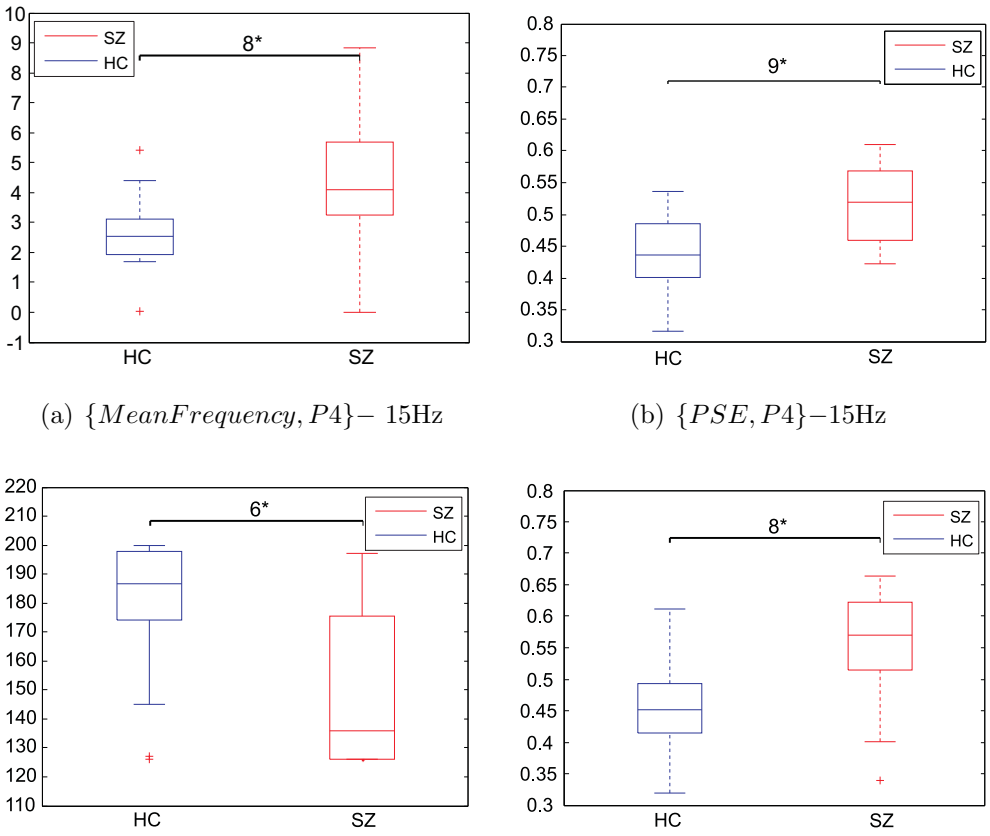
- In the comparison made between significant p-values, it has been observed that p-values of the Student's t test are consistent with those p-values of the Mann-Whitney test, i.e. when some pairs show a p-value ≤ 0.05, the other test leads to similar results.

3.1.2. Mann-Whitney non-parametric U test

After the failed result from the Student's t test, the Mann-Whitney non-parametric U test [40] is performed. Figure 3 shows the four pairs that have proved to be more discriminant since they exhibit the less overlapping boxplots among the two classes. Notice that these boxplots do not show p-values; they show the difference of the measured data for each pair comparing both populations, SZ and HC.

Raw data from the measurement of each pair are related to the p-value of the Mann-Whitney test by a certain number of asterisk symbols (*), which are shown at the top of the boxplots. The relationship between the number of asterisk symbols (*) and the p-values upper bounds is next shown:

$$p \leq 0.0001 \Longrightarrow 10*$$
$$p \leq 0.0005 \Longrightarrow 9*$$
$$p \leq 0.001 \Longrightarrow 8*$$
$$p \leq 0.0025 \Longrightarrow 7*$$
$$p \leq 0.005 \Longrightarrow 6*$$
$$p \leq 0.01 \Longrightarrow 5*$$



(a) $\{MeanFrequency, P4\}-$ 15Hz

(b) $\{PSE, P4\}-15$Hz

(c) $\{P300Lat, F7\}-35$Hz

(d) $\{PSE, P4\}-35$Hz

Figure 3: Comparison of measured data from electrode for what can be considered as the four best { *feature-electrode*} pairs. In these figures, we can highlight how medians of both subject groups, SZ and HC, are well apart one from the other with almost no overlap between the boxes. (a) $\{meanFreqency, P4\}-15$Hz, (b) $\{PSE, P4\}-15$Hz, (c) $\{P300Lat, F7\}-35$Hz, and (d) $\{PSE, P4\}-35$Hz (red: SZ, blue: HC).

$$p \leq 0.02 \Longrightarrow 4*$$
$$p \leq 0.03 \Longrightarrow ***$$
$$p \leq 0.04 \Longrightarrow **$$
$$p \leq 0.05 \Longrightarrow *$$

Next, results reached using the Mann-Whitney U test for electrocardiograms are shown using boxplots, and for features with boxplots and topoplots, the latter being an interpolated graphical option of the EEGLab MatLab® Toolbox [38].

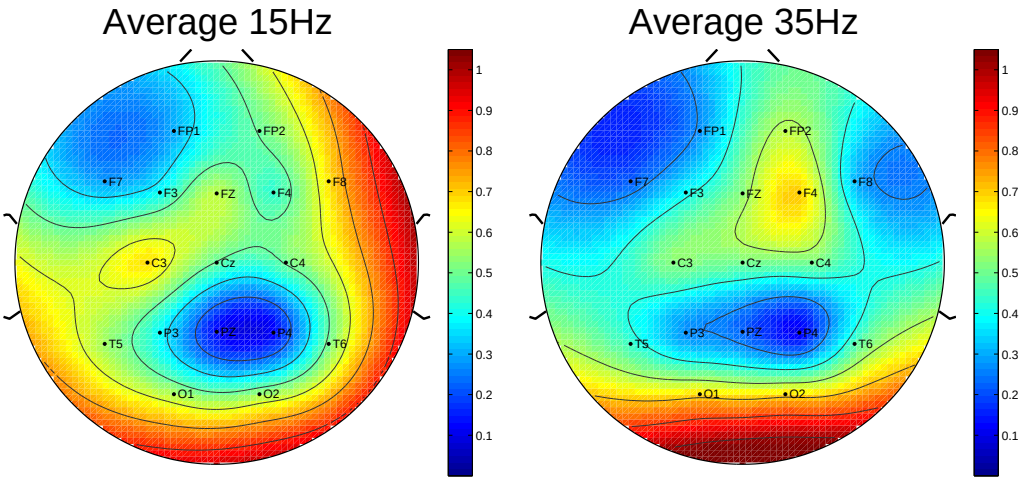
3.1.3. Features

We believe that *topoplot* figure plot type [38] is a fast, simple, intuitive and visual way to observe direct conclusions regarding represented p-values. In Section Appendix B.1, topoplots of frequency (f) features are shown (Figure B.13 and Figure B.14), while in Sections Appendix B.2 and Appendix B.3 topoplots of time-domain (t) features for 15Hz, and 35Hz, Figures B.15 and B.16, respectively, are shown. Analyzing these topoplots, the following conclusions concerning p-values of the Mann-Whitney test can be highlighted:

- For the 15Hz filter, a larger amount of lower p-values (blue tone colors) are observed, in contrast to those obtained with the 35Hz filter.
- For the 15Hz filter, features with more blue area are: *mean frequency*, *ZC*, *P300LatAbs*, and, especially, *PSE*. See Appendix B, Figures B.13 and B.15.
- For the 35Hz filter, features with more blue area are: *mean frequency*, *PSE*, and *variance*. See Appendix B, Figures B.14 and B.16.
- For both filters, the most relevant electrodes found are: *P4*, *Pz*, *Fp1*, and *F7*.
- Considering the interpolated topoplot of the Mann Whitney non-parametric U test that averages feature p-values (both f and t domains) for both

preprocessing filters (either 15Hz or 35Hz) –Figure 4–, it can be seen that the 15Hz case has less red colors (higher p-values) than when the 35Hz filter is used. It is also observed that the blue colored electrodes are those previously described as showing a better performance in most topoplots (*P4*, *Pz*, *Fp1*, and *F7*). Note that for each of the 17 electrodes (channels), we are simply averaging the p-values resulting for each of the 20 features under consideration (4 in f domain and 16 in t domain). The aim is to have an idea of what regions of the head may have a higher discriminant power when discriminating between SZ and HC subjects, considering each feature individually (univariate) and then computing the average of all 20 resulting p-values, using only a single feature at the same time for discriminating SZ and HC subjects.

Next, boxplots for each of the 20 features computed from EEG data are shown in Figure 5. Most of them have a low p-value (upper limits are quite low in all of them). However, we can highlight the behavior of some of them, such as: *PSE*, *mean frequency*, *median frequency*, and *mean*, because all of them offer low p-values for most of the electrodes (the upper limit is not very large); besides, the median p-value is quite low in the four cases. Table 1 summarizes the best p-values found for the features of the 15Hz filter. For the 35Hz filter case, the best mean p-values were obtained by the *PSE*, *Mean Frequency* and *Median Frequency* features, while the best median values corresponded to *PSE*, *ATAR* and *Mean Frequency* features.



(a) Electrodes with best averaged 15Hz topoplot p-values: *P4*, *Pz*, *Fp1*, and *F7*.

(b) Electrodes with best averaged 35Hz topoplot p-values: *P4*, *Pz*, *Fp1*, and *F7*.

Figure 4: Topoplot of p-values from the Mann-Whitney non-parametric U test for the averaged feature p-value of each of the 17 electrodes. a) 15Hz pre-processing filter, b) 35Hz pre-processing filter. Please note that we are simply averaging resulting p-values over all 20 features in each of the 17 electrode loci, but each feature is used alone when discriminating between SZ and HC subjects (univariate).

Table 1: Features with best mean and median p-values in the Mann-Whitney non-parametric U test. Preprocessing with 15Hz filter.

mean p-value		median p-value	
Feature	p-value	Feature	p-value
PSE	0.29426	Mean	0.15723
Mean freq.	0.29471	PSE	0.20064
Mean	0.31649	Mean freq.	0.31234

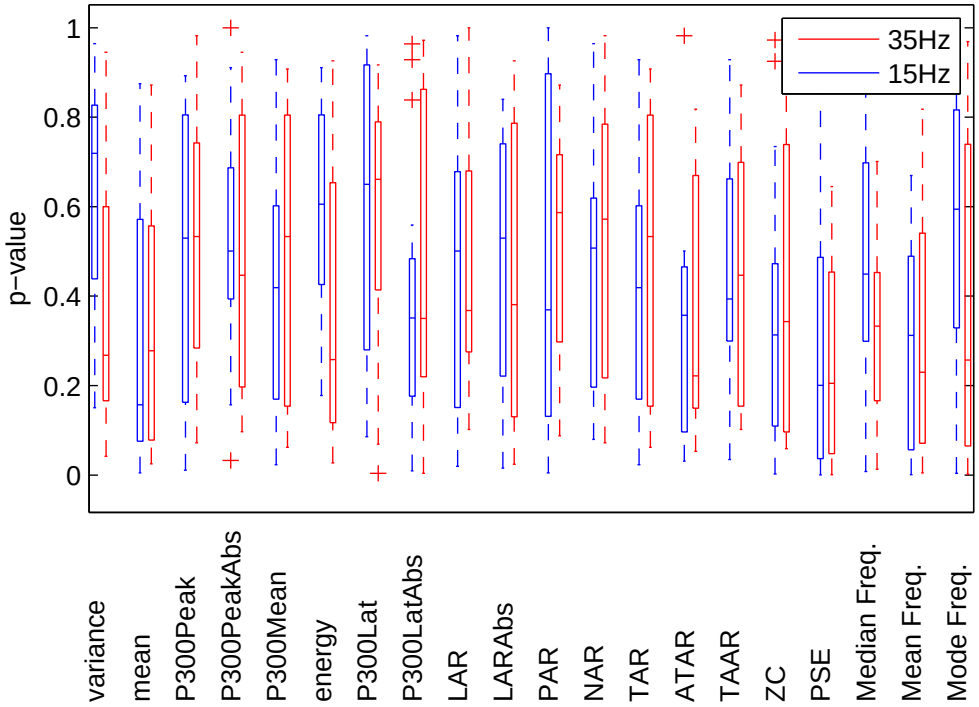


Figure 5: p-values boxplots from the Mann-Whitney non-parametric U test. Boxplots are shown for p-values of each pair, fixing, in each one, the corresponding feature. Every boxplot is constructed using the p-values of the 17 pairs of each feature with every single electrode (channel) (red: 35Hz filter, blue: 15Hz filter).

3.1.4. Electrodes

Figure 6 shows p-value boxplots corresponding to the 17 electrodes. It is worth noting that the best electrodes are P4 and Pz, which obtain small values for most features. We can also highlight that *F3*, *F7*, and *Fp1* electrodes attain low p-values in some pairs.

We can also draw conclusions for the electrodes taking into account those topoplots outlined in previous section, though not so directly. For instance, referring to Figure 4, we can see that, for both heads, the above described electrodes have low p-values, specifically $P4$ and Pz.

Table 2 contains a summary with those electrodes with the best p-values for the 15Hz filter. In case the 35Hz filter is used, both the best mean and median p-values were obtained by the $P4$, $F7$ and Pz features.

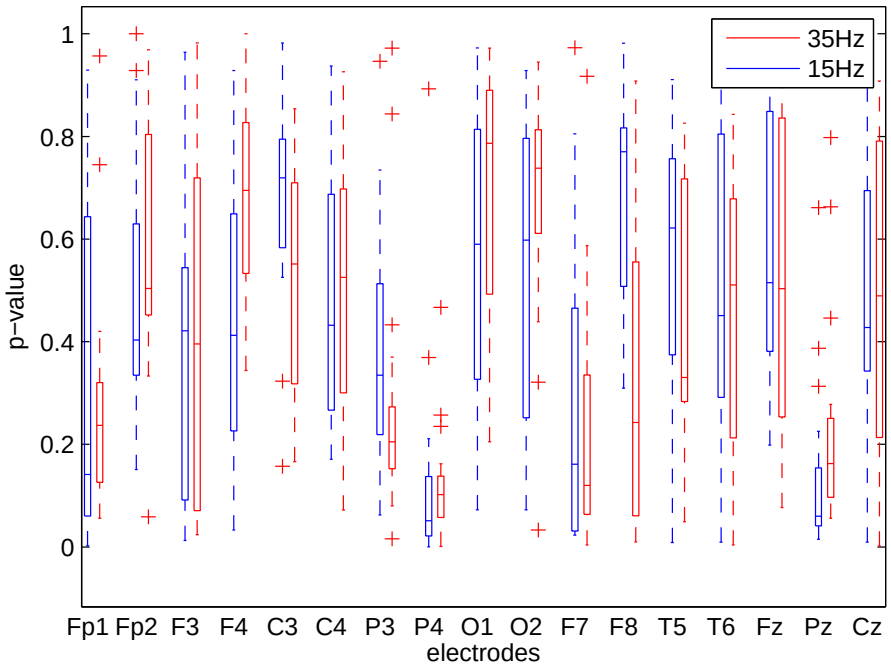


Figure 6: Boxplots of Mann-Whitney U test p-values. Figure shows p-values boxplots for each pair considering the corresponding electrode is fixed (red: 35Hz filter, blue: 15Hz filter).

Table 2: Electrodes with best mean and median p-values in the Mann-Whitney non-parametric U test, for 15Hz preprocessing filter.

mean p-value		median p-value	
Electrode	p-value	Electrode	p-value
Pz	0.12180	P4	0.05131
P4	0.08071	Pz	0.06007
F7	0.26177	F7	0.16121

3.2. Multivariate results

In this section, a multivariate analysis is performed by implementing two different multivariate tests. On the one hand, we have considered the parametric *MANOVA* test, and, on the other hand, the *Anosim* test has been performed. *Anosim* is a version of the non-parametric MANOVA test based on permutations. Both tests were carried out in two ways: first, adding the electrodes/features randomly and obtaining the resulting p-values, and, second, adding the electrodes/features following Fisher's Linear Discriminant Analysis (FLDA) criterion, which is known to be sub-optimal for p-values [41].

First, in Section 3.2.1, this analysis is performed for every feature, next, in Section 3.2.2, for every electrode, and finally, in Section 3.2.3, for different electrode groupings –electrodes belonging to each group can be found in Section 2.2.2, Figure 2.

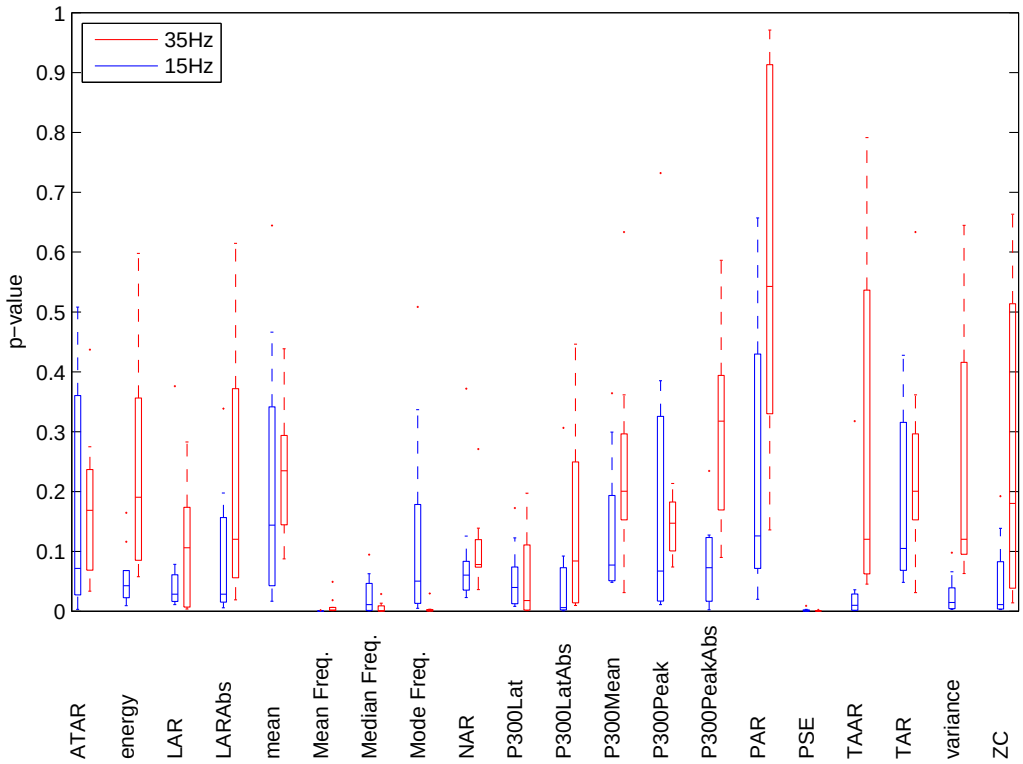


Figure 7: p-values boxplots from MANOVA test for each feature (red: 35Hz filter, blue: 15Hz filter).

3.2.1. Features

Analyzing p-values for each feature, a good performance is observed for the following frequency-domain features (both for the 15 and 35Hz cases): *mean frequency, median* and *PSE*. Besides, features *ATAR, P300LatAbs* and *ZC* also attain low p-values with the 15Hz filtering, while *frequency mode* leads to a low p-value under the 35Hz filtering. Analyzing figures, it is noted that, in general, the 15Hz filter gets lower p-values, and, therefore, better discriminatory capability than the 35Hz filter.

When both MANOVA (Figure 7) and *Anosim* (Figure 8) are compared, it

can be seen that features show a very similar behavior, since features with low p-value using one method, usually have also this tendency with the other. Only small not relevant differences are appreciated. Finally, notice that if an ordering is carried out following the FLDA criterion, lower p-values are reached with respect to a random ordering.

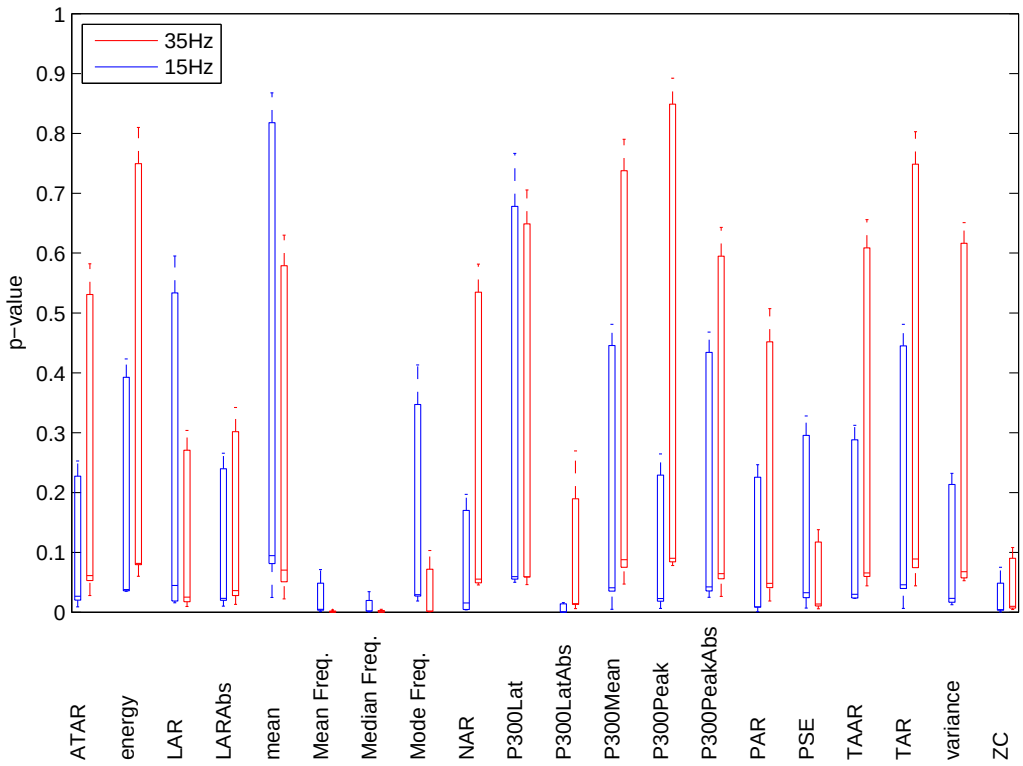


Figure 8: p-values boxplots from *Anosim* test for each feature (red: 35Hz filter, blue: 15Hz filter).

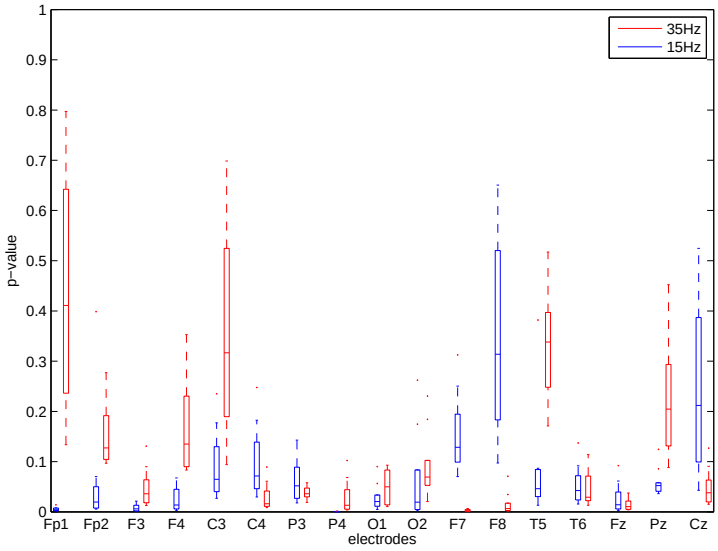


Figure 9: p-values boxplots from MANOVA test for each electrode (red: 35Hz filter, blue: 15Hz filter).

3.2.2. Electrodes

Analyzing these boxplots, electrodes that lead to lower p-values when the 15Hz filtering is used, are: *F4, C3, P4, O1, F7* and *Fz*. On the other hand, electrodes with lowest p-values when the 35Hz filtering is used, are: *C4, O2* and *F7*. It is also noted that 15Hz filtering leads to lower electrode p-values than those achieved with the 35Hz filter.

Making a comparison between MANOVA (Figure 9) and *Anosim* (Figure 10), it can be seen that the *Anosim* method leads to lower p-values than the MANOVA test for most of the electrodes. However, electrodes having low p-values for the MANOVA test, remain low for the *Anosim* function, matching those that have been the most discriminant ones, as found in previous tests. Whenever the FLDA criterion is used to sort electrodes, smaller p-values are

reached than when a random electrodes ordering is used.

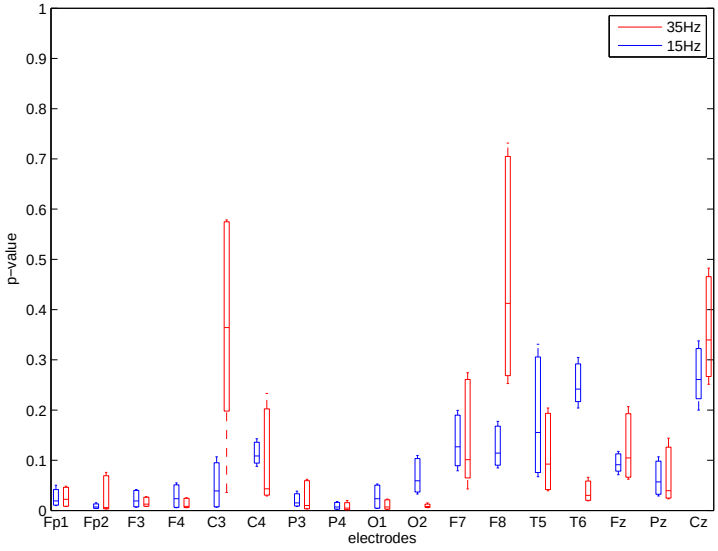


Figure 10: p-values boxplots from *Anosim* test for each electrode (red: 35Hz filter, blue: 15Hz filter).

3.2.3. Groups of electrodes

Analyzing the electrode groupings boxplots, the following conclusions can be drawn: those groupings with the lowest p-values are the Frontal and Parieto-Temporal groups. This performance is reasonable since front group includes *F4*, *F7* and *Fz* electrodes, and, as we have seen in previous sections, these electrodes have high discriminatory power. On the other hand, the Parieto-Temporal grouping includes electrode *P4*, which has very low p-values, as well as being located in the part of the head where the event related potentials of p300 (P3b) wave is believed to be generated, and also in accordance with results recently published in [22], reached after using convolutional Neural Networks (CNN) and Support Vector Machine (SVM)

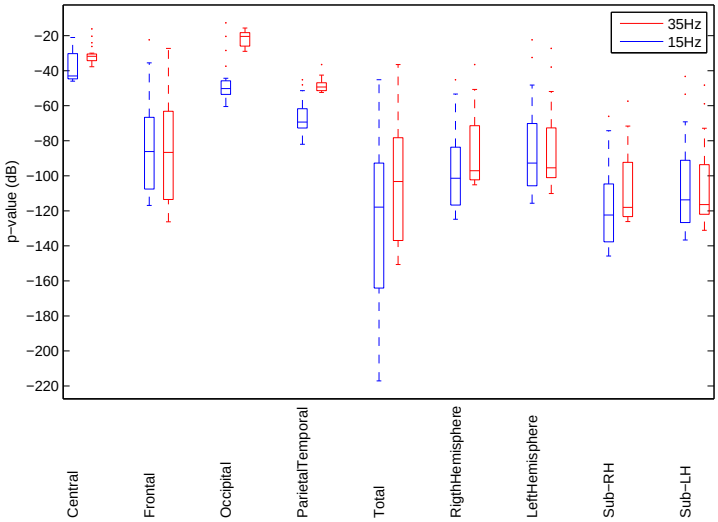


Figure 11: Boxplots of p-values for MANOVA test for different electrode groupings (red: 35Hz filter, blue: 15Hz filter).

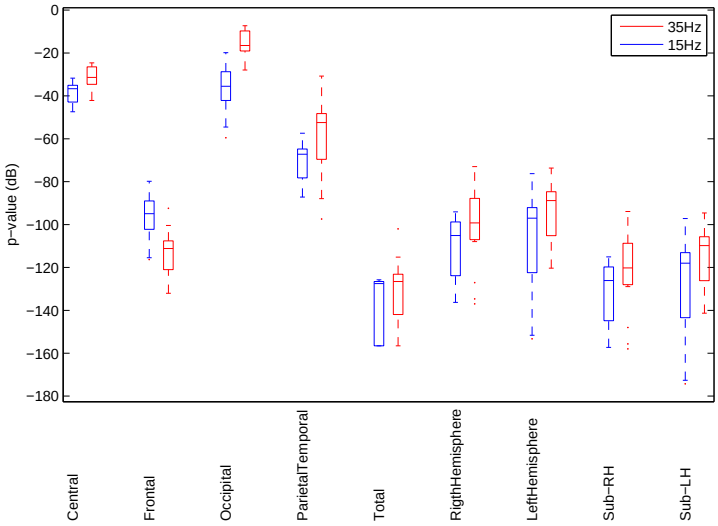


Figure 12: p-values boxplots from *Anosim* test for different electrode groupings (red: 35Hz filter, blue: 15Hz filter).

techniques.

Boxplots show how results improve as the number of electrodes belonging to the grouping is increased. The results of the hemispheres and the total head improve the results of the other groups. If we compare the two hemispheres, it can be seen that the left hemisphere reaches lower p-values, although it is true that the right hemisphere has more low p-values (besides, RH has the lowest average). The 15Hz filter led to lower p-values, accordingly with previous simulations where this specific filtering led to better results, [21]. Regarding the boxplots of sub-hemispheres –whose definition can be seen in Section 2.2.2– it can be seen that they have lower p-values than those obtained for the right and left hemispheres, though not lower than those reached with the Total grouping. This behavior is common for both tests, MANOVA (Figure 11) and *Anosim* (Figure 12). In relation to the organization in which pairs of clusters are caught, we see that lower p-values are reached if FLDA is used instead of the uniform random selection scheme.

4. Discussion

In our work, a univariate analysis of the input samples has been first carried out, followed by a non-parametric hypothesis test (the Mann-Whitney U test). Results from both parametric and non-parametric approaches were conveniently compared. Second, a multivariate analysis was performed; a discriminant analysis was run in order to find the optimal –in terms of better discriminatory capabilities– pairs of features and electrodes. Third, a MANOVA test was performed so as to compute p-values and compare groups

of pairs of features and electrodes. Last, a non-parametric *Anosim* test based on permutations was carried out and compared with the MANOVA test.

Optimal combinations of electrode grouping, filtering, and feature selection based on computed p-values have been provided. Electrode groupings boxplots lead to the following results: Frontal and Parieto-Temporal groupings get lower p-values. This is reasonable since Frontal grouping includes *F4*, *F7*, and *Fz* electrodes, which have, as we have seen in other tests, good discriminatory properties. On the other hand, the Parieto-Temporal grouping includes electrode *P4*, which leads to very low p-values, as well as being located in an area of the skull where the P3b wave is believed to be generated, being in accordance with previous results shown in [21] as well as with those recently reported by Shalbaf et al. in [22].

Boxplots show that results improve as the number of electrodes in the group becomes higher. Results of both right and left hemispheres and the Total grouping are better than those obtained with the other electrode groupings. On the other hand, Left Hemisphere attains lower p-values, although Right Hemisphere gets a higher number of low p-values (and also has the lowest average p-value). As previously proved with other tests, the 15Hz pre-processing filter led to lower p-values. Boxplots of sub-hemispheres have inferior p-values than those corresponding to the left and right hemispheres, but higher than those of the Total grouping. This behavior is general for both MANOVA and *Anosim* tests. Besides, we have seen that lower p-values are reached if FLDA ordering is used instead of the uniform random selection scheme.

These results, regarding electrode groupings multivariate analysis, are

consistent with those observed by Palaniyappan et al. in [17], where schizophrenia is discriminated from bipolar disorders at the single subject level. In fact, [17] concludes that "even though certain features were themselves not significantly different between the groups, they increased separability between the groups when used in combination with other features, emphasizing the strength of the multivariate nature of classifiers". Nevertheless, it must be noted that the inclusion of too many non discriminant features may also reduce the final discriminatory power of the classifier for detecting schizophrenia. Thus, a trade-off is needed when selecting the number of features and the particular $\{feature, electrode\}$ pairs, in the problem of discriminating between healthy controls (HC) and schizophrenia (SZ) patients. In our work, optimized sets of electrode grouping, filtering, and feature selection based on computed p-values have been provided.

This use of statistical significance tests to validate the discriminant capability of features is in accordance with recently published works such as [15], [16], [42], and several of those cited in the review paper by C. Barros et al. [23], all of them describing AI-based schizophrenia detection systems.

5. Conclusion and future work

In this paper, a novel Statistical Discriminant Diagnosing (SDD) system that discriminates between healthy controls and subjects with schizophrenia has been presented. The proposed system works with $\{feature, electrode\}$ EEG pairs which are selected based on the statistical significance of the p-values computed over the brain P3b wave. EEG signals comprise 20 features and 17 electrodes, both in time (t) and frequency (f) domain. Obtained

results have been proved to be in agreement with previous hypothesis regarding the relevance of the Parieto-Temporal region, allowing us to identify highly discriminant $\{feature, electrode\}$ pairs in the detection of schizophrenia, resulting lower p-values in both Right and Left Hemispheres, as well as in Parieto-Temporal EEG signals. For instance, the $\{PSE, P4\}$ pair has been found to be highly discriminant. Consequently, the proposed SDD system may provide the human expert (psychiatrist) with an objective complimentary information to help in the early diagnosis of schizophrenia.

As future work, we plan to apply multi-dimensional (17-D) array processing at electrode skull surface loci with the purpose of identifying the multiple simultaneous electromagnetic brain waves and their originating volume. In addition, we expect to extend the database used regarding the number of patients, seeking higher statistical power tests combined with the use of machine learning in the automatic diagnose of schizophrenia.

Acknowledgments

The authors are grateful to E. Illera, for his valuable help in both programming part of the simulation and discussing results; L. Santos-Mayo for his help in EEG signals pre-processing with EEGLab, a MatLab® toolbox, and Prof. V. Molina and colleagues, dep. of Psychiatry, Medical School, University of Valladolid, for providing the EEG input HC and SZ subject groups data.

Funding statement

This work was supported by Instituto de Salud Carlos III under Fondo de Investigaciones Sanitarias (FIS) grant number FIS-PI11/02203, Spain.

Declaration of competing interest

The authors declare that they have no competing interests. The authors declared that they have no conflict of interest and the paper presents their own work, which does not infringe any third-party rights, especially authorship of any part of the article is an original contribution, not published before and not being under consideration for publication elsewhere. The order of authors listed in the manuscript has been approved by both authors.

Ethical approval

Explicit written consent was used for SZ subject inclusion, and volunteer HC subjects received a small remuneration. The Medical School Ethical Committee, University of Valladolid, gave trials and experiments approval.

Appendix A. *t* Student Tabular values

Table A.3: p-values for W statistic in Levene quality of variance test. Levene test tells us whether we can assume equal variance or not among both classes. This say, if probability associated to Levene statistic parameter is W > 0.05, then equal variance hypothesis can be assumed, vice versa not equal variance counterpart (15Hz filter).

$\{feature, electrode\}$	W (p-value)	Equal variance assumed?
$\{PSE, P4\}$	0.51869	Yes
$\{MeanFrequency, P4\}$	0.00311	No
$\{PSE, F3\}$	0.24185	Yes
$\{P300Peak, FP1\}$	0.52289	Yes
$\{PSE, F7\}$	0.33899	Yes
$\{PSE, FP1\}$	0.20555	Yes
$\{LAR\_Abs, P4\}$	0.06813	Yes
$\{PSE, Pz\}$	0.94265	Yes

Table A.4: p-values for W statistic in Levene quality of variance test. Levene test tells us whether we can assume equal variance or not among both classes. This say, if probability associated to Levene statistic parameter is W > 0.05, then equal variance hypothesis can be assumed, viceversa not equal variance counterpart (35Hz filter).

$\{feature, electrode\}$	W (p-value)	Equal variance assumed?
$\{mean, F3\}$	0.36710	Yes
$\{P300Mean, F7\}$	0.11109	Yes
$\{TAR, F7\}$	0.11109	Yes
$\{MeanFrequency, F3\}$	0.00802	No
$\{MeanFrequency, F8\}$	0.11346	Yes

Table A.5: Comparison between t Student (parametric) and U Mann-Whitney (non-parametric) tests for data fulfilling normality condition. Whenever $p < 0.05$ we can reject the null hypothesis about equal means (15Hz filter).

$\{feature, electrode\}$	t Student p-value	U Mann-Whitney p-value
$\{PSE, P4\}$	0.00003	0.00019
$\{MeanFrequency, P4\}$	0.00708	0.00089
$\{PSE, F3\}$	0.00860	0.01269
$\{P300Peak, FP1\}$	0.01120	0.01118
$\{PSE, F7\}$	0.02419	0.03113
$\{PSE, FP1\}$	0.02590	0.05634
$\{LARAbs, P4\}$	0.03606	0.05930
$\{PSE, Pz\}$	0.04387	0.03887

Table A.6: Comparison between t Student (parametric) and U Mann-Whitney (non-parametric) tests for data fulfilling normality condition. Whenever $p < 0.05$ we can reject the null hypothesis regarding equal means (35Hz filter)

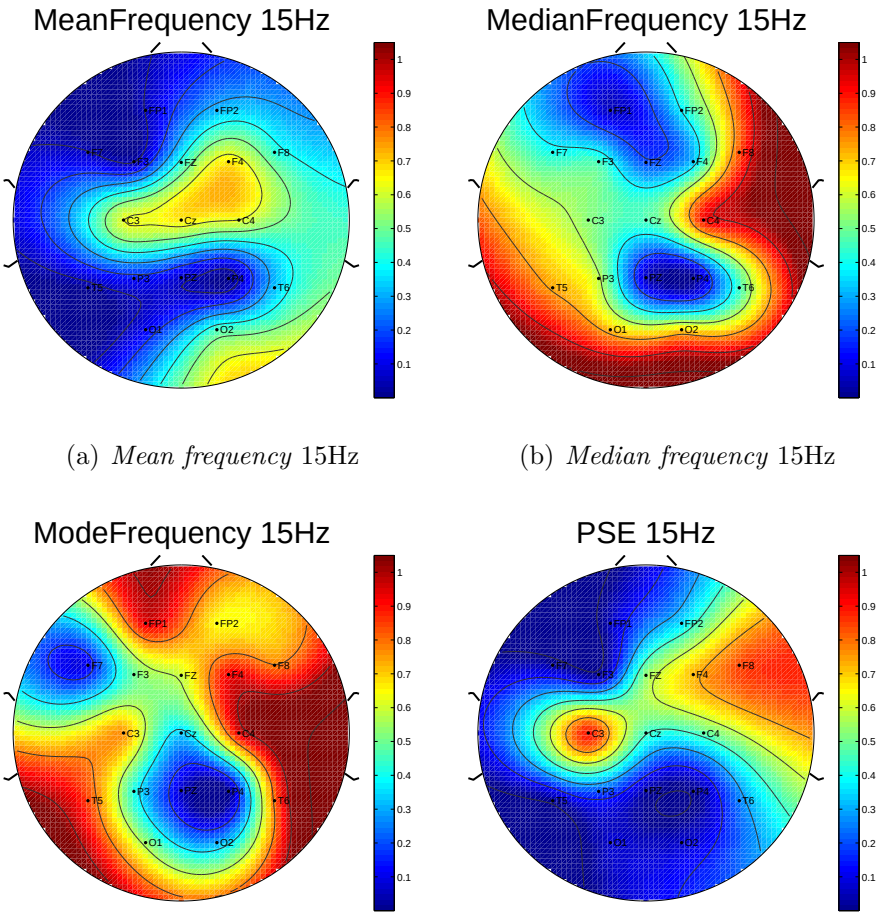
$\{feature, electrode\}$	t Student p-value	U Mann-Whitney p-value
$\{mean, F3\}$	0.04365	0.025
$\{P300Mean, F7\}$	0.03118	0.062
$\{TAR, F7\}$	0.03118	0.062
$\{MeanFrequency, F3\}$	0.04426	0.076
$\{MeanFrequency, F8\}$	0.00947	0.010

Appendix B. Mann-Whitney U test topoplots: univariate

In this appendix, we show the non-parametric Mann-Whitney U p-values topoplots interpolated graphs for each of the 4 frequency (f) and 16 time (t) features extracted from each of the EEG channels (electrodes) considered individually so as to compare the discriminant power in discriminating SZ from HC of each individual feature along head skull in a univariate way. Both 15Hz and 35Hz preprocessing filters have been considered. Thus, lower p-values in interpolated head graphs (blue tones) indicate a higher discriminative power of a particular feature (either frequency (f) or time (t) domains) in distinguishing between SZ and HC subjects.

Appendix B.1. U Mann-Whitney frequency (f) feature topoplots: 15Hz and 35Hz

Figures B.13 and B.14 depict the U Mann-Whitney frequency domain (f) feature topoplots for both 15Hz and 35Hz pre-processing filters, respectively.



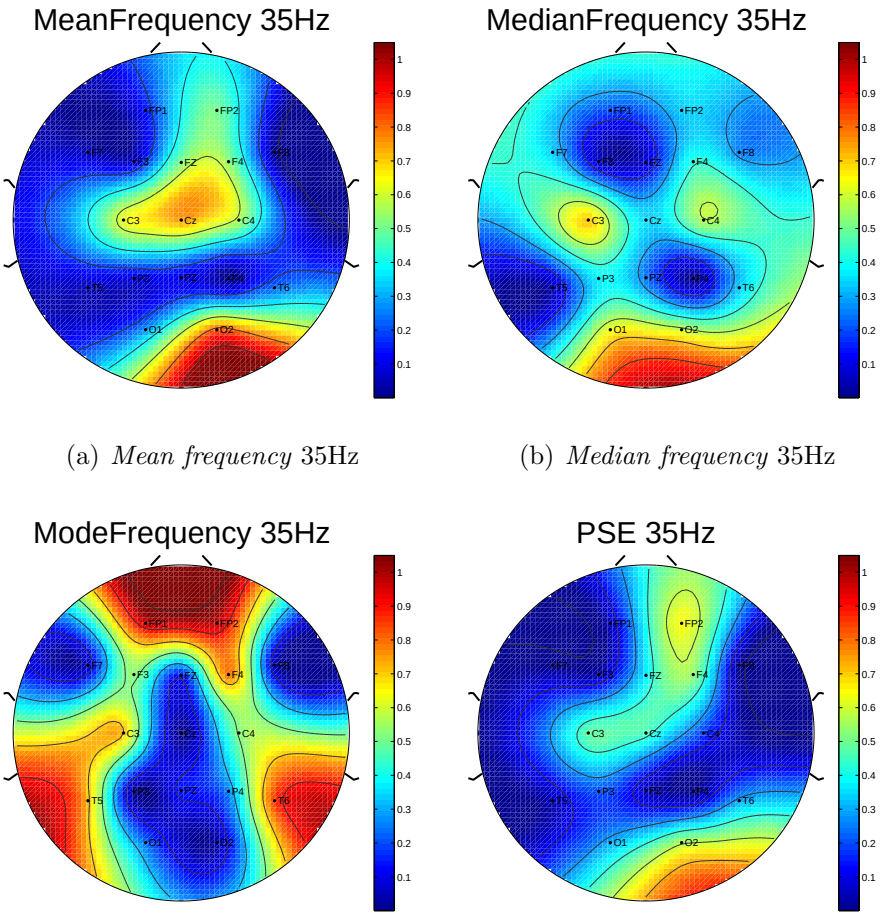
(a) *Mean frequency* 15Hz

(b) *Median frequency* 15Hz

(c) *Mode frequency* 15Hz

(d) *PSE* 15Hz

Figure B.13: U Mann-Whitney frequency (f) feature topoplots for the 15Hz filter. Blue tones indicate lower p-values (highly discriminant) while on the contrary red tones indicate higher (closer to 1) p-values (less discriminant). Topoplots shown are: (a) *Mean frequency* 15Hz, (b) *Median frequency* 15Hz, (c) *Mode frequency* 15Hz, and (d) *PSE* 15Hz. Please note univariate (single feature) results are shown here.



(a) *Mean frequency* 35Hz

(b) *Median frequency* 35Hz

(c) *Mode frequency* 35Hz

(d) *PSE* 35Hz

Figure B.14: U Mann-Whitney test frequency (f) feature topoplots for the 35Hz filter. Blue tones indicate lower p-values (highly discriminant) while on the contrary red tones indicate higher (closer to 1) p-values (less discriminant). Topoplots shown are: (a) *Mean frequency* 35Hz, (b) *Median frequency* 35Hz, (c) *Mode frequency* 35Hz, and (d) *PSE* 35Hz. Please note univariate (single feature) results are shown here.

Appendix B.2. U Mann-Whitney time feature topoplots: 15Hz

Figure B.15 depicts the U Mann-Whitney time domain (t) feature topoplots for the 15Hz pre-processing filter.

Appendix B.3. U Mann-Whitney time feature topoplots: 35Hz

Figure B.16 depicts the U Mann-Whitney time domain (t) feature topoplots for the 35Hz pre-processing filter.

References

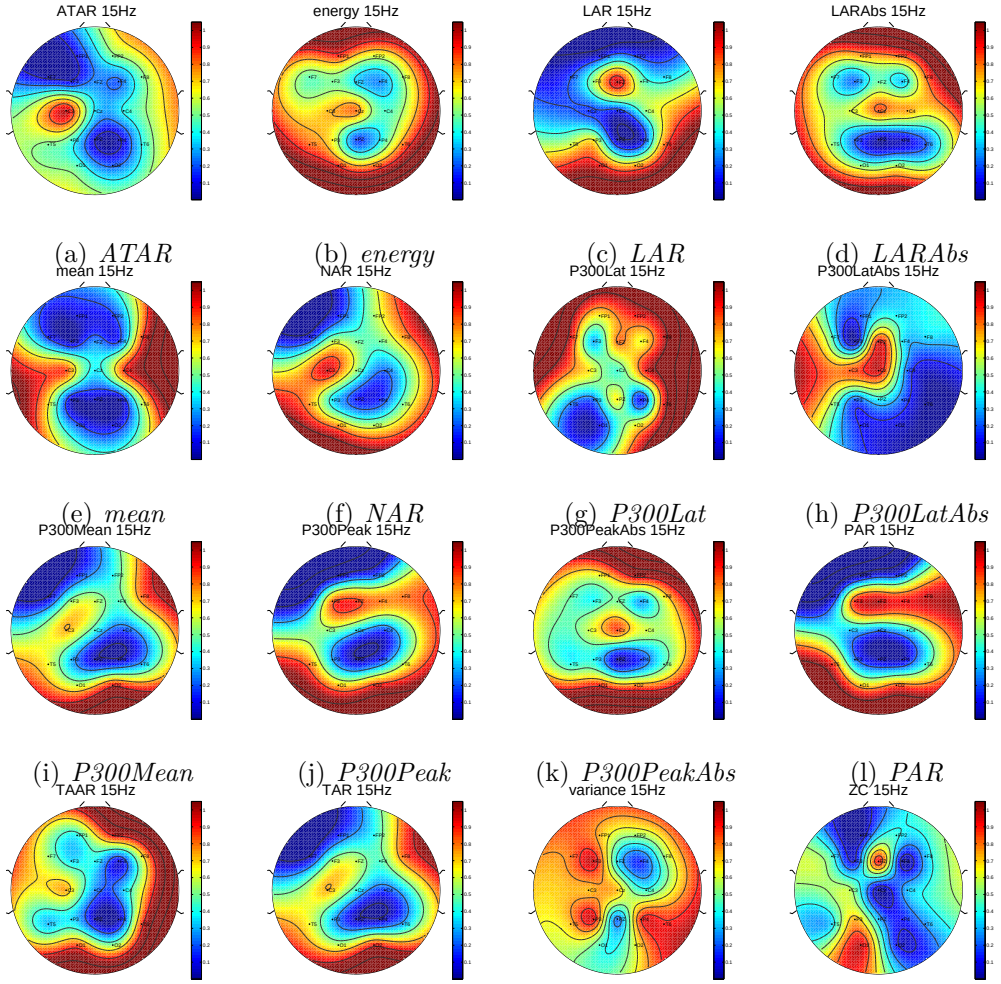
[1] A. P. Association, et al., Diagnostic and Statistical Manual of Mental Disorders DSM-5®, American Psychiatric Association Pubs., 2013.

[2] N. N. Boutros, C. Arfken, S. Galderisi, J. Warrick, G. Pratt, W. Iacono, The status of spectral EEG abnormality as a diagnostic test for schizophrenia, Schizophrenia Research 99 (1) (2008) 225–237.

[3] L. A. Farwell, E. Donchin, Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials, Electroencephalography and Clinical Neurophysiology 70 (6) (1988) 510–523.

[4] S. K. Prabhakar, H. Rajaguru, S. Lee, A framework for schizophrenia EEG signal classification with nature inspired optimization algorithms, IEEE Access 8 (2020) 39875–39897.

[5] M. Nieuwenhuis, N. E. van Haren, H. E. H. Pol, W. Cahn, R. S. Kahn, H. G. Schnack, Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples, NeuroImage 61 (3) (2012) 606 – 612.



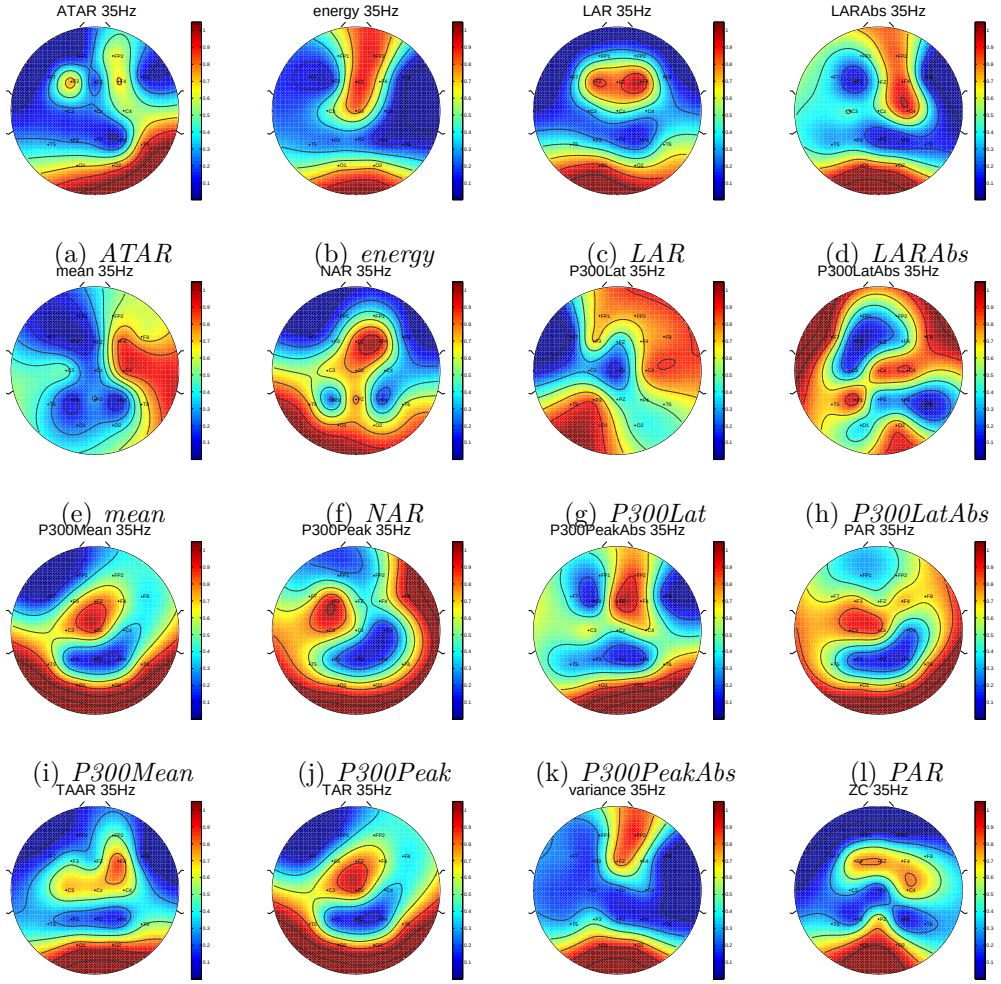
(a) *ATAR* (b) *energy* (c) *LAR* (d) *LARAbs*

(e) *mean* (f) *NAR* (g) *P300Lat* (h) *P300LatAbs*

(i) *P300Mean* (j) *P300Peak* (k) *P300PeakAbs* (l) *PAR*

(m) *TAAR* (n) *TAR* (o) *variance* (p) *ZC*

Figure B.15: U Mann-Whitney test time (t) feature topoplots for the 15Hz filter: (a) *ATAR*, (b) *energy*, (c) *LAR*, (d) *LARAbs*, (e) *mean*, (f) *NAR*, (g) *P300Lat*, (h) *P300LatAbs*, (i) *P300Mean*, (j) *P300Peak*, (k) *P300PeakAbs*, (l) *PAR*, (m) *TAAR*, (n) *TAR*, (o) *variance*, and (p) *ZC*. Blue tones indicate lower p-values (highly discriminant) while on the contrary red tones indicate higher (closer to 1) p-values (less discriminant). Please note univariate (single feature) results are shown here.



(a) *ATAR* (b) *energy* (c) *LAR* (d) *LARAbs*

(e) *mean* (f) *NAR* (g) *P300Lat* (h) *P300LatAbs*

(i) *P300Mean* (j) *P300Peak* (k) *P300PeakAbs* (l) *PAR*

(m) *TAAR* (n) *TAR* (o) *variance* (p) *ZC*

Figure B.16: U Mann-Whitney test time (t) feature topoplots for the 35Hz filte: (a) *ATAR*, (b) *energy*, (c) *LAR*, (d) *LARAbs*, (e) *mean*, (f) *NAR*, (g) *P300Lat*, (h) *P300LatAbs*, (i) *P300Mean*, (j) *P300Peak*, (k) *P300PeakAbs*, (l) *PAR*, (m) *TAAR*, (n) *TAR*, (o) *variance*, and (p) *ZC*. Blue tones indicate lower p-values (highly discriminant) while on the contrary red tones indicate higher (closer to 1) p-values (less discriminant). Please note univariate (single feature) results are shown here.

[6] A. Pfefferbaum, B. G. Wenegrat, J. M. Ford, W. T. Roth, B. S. Kopell, Clinical application of the P3 component of event-related potentials. II. dementia, depression and schizophrenia, Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section 59 (2) (1984) 104–124.

[7] V. Souza, W. Muir, M. Walker, M. Glabus, H. Roxborough, C. Sharp, J. Dunan, D. Blackwood, Auditory P300 event-related potentials and neuropsychological performance in schizophrenia and bipolar affective disorder, Biological Psychiatry 37 (5) (1995) 300–310.

[8] C. Kruiper, B. Fagerlund, M. Ø. Nielsen, S. Düring, M. H. Jensen, B. H. Ebdrup, B. Y. Glenthøj, B. Oranje, Associations between P3a and P3b amplitudes and cognition in antipsychotic-naïve first-episode schizophrenia patients, Psychological medicine 49 (5) (2019) 868–875.

[9] M. Devrim-Üçok, H. Y. Keskin-Ergen, A. Üçok, Novelty P3 and P3b in first-episode schizophrenia and chronic schizophrenia, Progress in Neuro-Psychopharmacology and Biological Psychiatry 30 (8) (2006) 1426–1434.

[10] A. McCathern, Emitted P3a and P3b in chronic schizophrenia and in first-episode schizophrenia, Ph.D. thesis, University of Pittsburgh (2017).

[11] H. K. Hamilton, S. W. Woods, B. J. Roach, K. Llerena, T. H. McGlashan, V. H. Srihari, J. M. Ford, D. H. Mathalon, Auditory and visual oddball stimulus processing deficits in schizophrenia and the psychosis

risk syndrome: forecasting psychosis risk with P300, Schizophrenia bulletin 45 (5) (2019) 1068–1080.

[12] A. G. McCathern, B. A. Coffman, T. K. Murphy, K. L. Ward, S. Haigh, D. Salisbury, Emitted P3a and P3b in chronic schizophrenia and in first-episode schizophrenia-spectrum psychosis, Biological Psychiatry 81 (10) (2017) S80.

[13] E. Niedermeyer, F. L. da Silva, Electroencephalography: basic principles, clinical applications, and related fields, Lippincott Williams and Wilkins, 2005.

[14] J. R. Stevens, L. Bigelow, D. Denney, J. Lipkin, A. H. Livermore, F. Rauscher, R. J. Wyatt, Telemetered EEG-EOG during psychotic behaviors of schizophrenia, Archives of General Psychiatry 36 (3) (1979) 251–262.

[15] K. Das, R. B. Pachori, Schizophrenia detection technique using multivariate iterative filtering and multichannel EEG signals, Biomedical Signal Processing and Control 67 (2021) 102525.

[16] S. Siuly, S. K. Khare, V. Bajaj, H. Wang, Y. Zhang, A computerized method for automatic detection of schizophrenia using EEG signals, IEEE Transactions on Neural Systems and Rehabilitation Engineering 28 (11) (2020) 2390–2400.

[17] L. Palaniyappan, G. Deshpande, P. Lanka, D. Rangaprakash, S. Iwabuchi, S. Francis, P. F. Liddle, Effective connectivity within a triple network brain system discriminates schizophrenia spectrum

disorders from psychotic bipolar disorder at the single-subject level, Schizophrenia Research 214 (2019) 24 – 33.

[18] S. L. Oh, J. Vicnesh, E. J. Ciaccio, R. Yuvaraj, U. R. Acharya, Deep convolutional neural network model for automated diagnosis of schizophrenia using EEG signals, Applied Sciences 9 (14) (2019) 2870.

[19] B. J. Roach, D. H. Mathalon, Event-Related EEG Time-Frequency Analysis: An Overview of Measures and An Analysis of Early Gamma Band Phase Locking in Schizophrenia, Schizophrenia Bulletin 34 (5) (2008) 907–926.

[20] N. B. Lundin, L. A. Bartolomeo, B. F. O'Donnell, W. P. Hetrick, Reduced electroencephalogram responses to standard and target auditory stimuli in bipolar disorder and the impact of psychotic features: Analysis of event-related potentials, spectral power, and inter-trial coherence, Bipolar Disorders 20 (1) (2018) 49–59.

[21] L. Santos-Mayo, L. M. San-Jose-Revuelta, J. I. Arribas, A computer-aided diagnosis system with EEG based on the P3b wave during an auditory odd-ball task in schizophrenia, IEEE Transactions on Biomedical Engineering 64 (2) (2017) 395–407.

[22] A. Shalbaf, S. Bagherzadeh, A. Maghsoudi, Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals, Physical and Engineering Sciences in Medicine 43 (4) (2020) 1229–1239.

[23] C. Barros, C. A. Silva, A. P. Pinheiro, Advanced EEG-based learning approaches to predict schizophrenia: Promises and pitfalls, Artificial Intelligence in Medicine (2021) 102039.

[24] S. Góngora Alonso, G. Marques, D. Agarwal, I. De la Torre Díez, M. Franco-Martín, Comparison of machine learning algorithms in the prediction of hospitalized patients with schizophrenia, Sensors 22 (2022) (7).

[25] N. Sairamya, M. Subathra, S. Thomas George, Automatic identification of schizophrenia using EEG signals based on discrete wavelet transform and rlndip technique with ann, Expert Systems with Applications 192 (2022) 116230.

[26] L. Santos-Mayo, L. M. San-Jose-Revuelta, J. I. Arribas, Diagnóstico automático de la esquizofrenia basado en el estudio de señales EEG de actividad cortical en el cerebro, Master Science Degree. ETSIT, University of Valladolid, Spain, 2015.

[27] N. E. Anderson, R. M. Baldridge, M. S. Stanford, P3a amplitude predicts successful treatment program completion in substance-dependent individuals, Substance Use and Misuse 46 (5) (2011) 669–677.

[28] P. Armitage, G. Berry, J. N. S. Matthews, Statistical methods in medical research, John Wiley and Sons, 2008.

[29] S. K. Prabhakar, H. Rajaguru, S.-W. Lee, A framework for schizophrenia EEG signal classification with nature inspired optimization algorithms, IEEE Access 8 (2020) 39875–39897.

[30] W. W. Daniel, C. L. Cross, Bioestatistics, Wiley, 2013.

[31] R. R. Wilcox, Introduction to robust estimation and hypothesis testing, Academic Press, 2012.

[32] D. B. Owen, Handbook of statistical tables, Addison-Wesley Publishing Company, 1962.

[33] T. Cacoullos, Discriminant analysis and applications, Academic Press, 2014.

[34] R. A. Johnson, D. W. Wichern, Applied multivariate statistical analysis, Prentice hall, 2002.

[35] J. P. Stevens, Applied multivariate statistics for the social sciences, Routledge, 2012.

[36] M. J. Anderson, D. C. Walsh, Permanova, anosim, and the mantel test in the face of heterogeneous dispersions: what null hypothesis are you testing?, Ecological Monographs 83 (4) (2013) 557–574.

[37] G. M. Gaddis, M. L. Gaddis, Introduction to biostatistics: Part 5, statistical inference techniques for hypothesis testing with nonparametric data, Annals of Emergency Medicine 19 (9) (1990) 1054–1059.

[38] A. Delorme, S. Makeig, EEGLab: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis, Journal of Neuroscience Methods 134 (1) (2004) 9–21.

[39] S. Zacks, Parametric statistical inference: Basic theory and modern approaches, Elsevier, 2014.

[40] M. Hollander, D. A. Wolfe, E. Chicken, Nonparametric statistical methods, John Wiley and Sons, 2013.

[41] K. McGarigal, S. A. Cushman, S. Stafford, Multivariate statistics for wildlife and ecology research, Springer Science and Business Media, 2013.

[42] V. Jahmunah, S. L. Oh, V. Rajinikanth, E. J. Ciaccio, K. H. Cheong, N. Arunkumar, U. R. Acharya, Automated detection of schizophrenia using nonlinear signal processing methods, Artificial intelligence in medicine 100 (2019) 101698.