Detecting spontaneous brain activations through *f-MRI* scans

Ernesto Bocini

Christian Martin Hennig

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

THE GLOBAL UNDERGRADUATE AWARDS

Introduction

- **Brain activity during sleep:** Our state of consciousness changes significantly during stages of sleep. We investigate on this through blood oxygenation level dependent (BOLD) at rest. This kind of study is possible thanks to modern functional magnetic resonance imaging (fMRI) technologies in addition to ECG.
- **Difficult data to obtain:** MRIs are difficult data to obtain. The machines are very noisy, making it hard for patients to reach a state of deep sleep. However, "MRI scans measure neural activity by detecting the hemodynamic response of structures throughout the brain, thereby providing important information in addition to EEGs." ∼ Dimitri Van De Ville.
- **Project Objective:** Develop and optimize an algorithm that is able to detect spontaneous brain activations at rest.

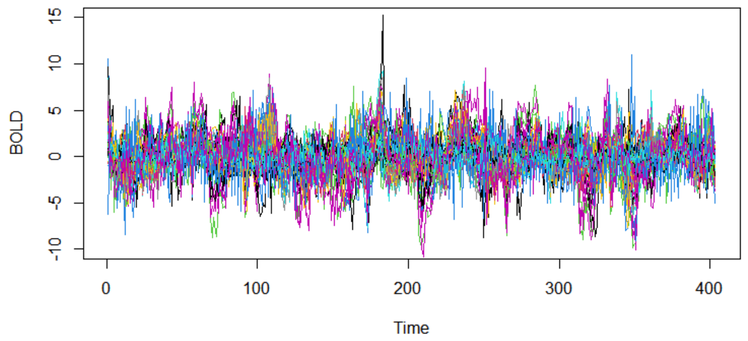

Figure 1. Dynamic activity time series of each brain region for the second subject. 1

Project overview and methods

In order to achieve the Project Objective mentioned earlier, we proposed the following two methods:

1. Analysis using repeated **ADF tests**,
2. Analysis using **outlier detection**.

In addition, given the complexity of the arrays at hand, lots of energies have been spent on dimensionality reduction for multivariate time series and feature engineering.

Analysis using repeated ADF tests.

At first, we tried a straightforward approach based on the stationarity of the time series describing each subject. The Dickey-Fuller (DF) test is the most popular test for unit root ($unit\ roots \Rightarrow non-stationarity$). It uses the following statistic $\hat{\tau}$:

$$\hat{\tau} = \frac{\hat{\phi}-1}{\hat{\sigma}(\sum_{t=2}^{n} X_{t-1}^2)^{-\frac{1}{2}}};\ \hat{\tau} \xrightarrow{d} \frac{\int W\,dW}{\left(\int W^2 dW\right)^{1/2}} \tag{1}$$

where $\hat{\phi}$ and $\hat{\sigma}^2$ are the least squares estimates of ϕ and σ^2. However, under the null hypothesis of non-stationarity, it follows $\hat{\phi} \overset{A}{\sim} N(1,0)$, which is undefined because when the standard deviation is zero, the Gaussian PDF turns into the Dirac delta function. This happens because under H_0, the process X_{t-1} on the right-hand side of our model is non-stationary. Therefore it is highly persistent (non-ergodic) and the autocorrelation decays to zero very slowly. This implies that the usual sample moments do not converge to fixed constants. Instead, Phillips (1987) showed that the sample moments of X_t converge to random function of Brownian motion, that can be seen on the right of the t-statistic above (1).
In this first approach, we are basing the analysis on a series of Augmented Dickey-Fuller tests. The idea is to run an ADF test for all the brain regions (70) of each subject (22) and respectively store the number of rejected tests. In this way, we obtain a statistic that is easily comparable across subjects and that can be modeled as a function of the individuals' traits, an example is shown in Figure 2.

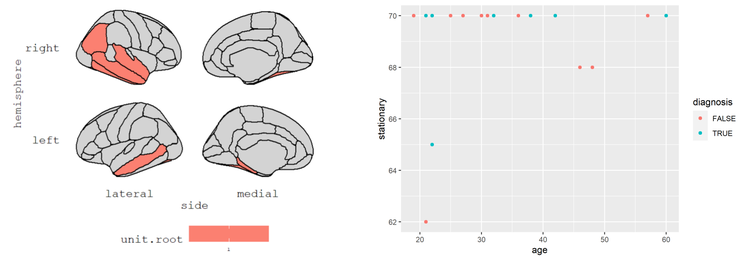

Figure 2. Brain regions for Subject 9 accordingly to the Desikan atlas parcellation into 70 regions.

Analysis using outlier detection procedure

Given the dynamic structure of time series, there are various definitions of outliers: (AO) Additive Outlier, (IO) Innovational Outlier, (TC) Temporary Change, and (LS) Level Shift. See Fig. 3.

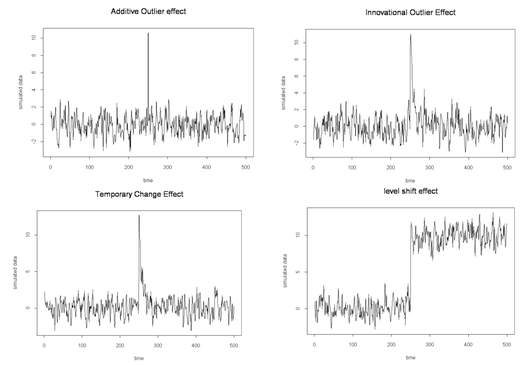

Figure 3. Effects of the four types of outliers.

Outlier detection algorithm:

1. Build a VARIMA model for the series under study, assuming no outliers and let $\hat{a}_t$ be the estimated residuals and $\hat{\Pi}_i$ the estimated coefficients of the autoregressive representation of the time series.
2. Estimate the outlier size by using $\hat{\omega}_{i,h};\ \ i = I, A, L, T$:

$$\hat{\omega}_{I,h} = \hat{a_h}$$

$$\hat{\omega}_{A,h} = -\left(\sum_{i=0}^{n-h} \hat{\Pi}_i' \Sigma^{-1} \hat{\Pi}_i\right)^{-1} \sum_{i=0}^{n-h} \hat{\Pi}_i' \Sigma^{-1} \hat{a}_{h+i} \quad (\Pi_0 = -I)$$

3. Test the null hypothesis $H_0 : \omega = 0$ against the alternative $H_1 : \omega \neq 0$. To do so, two statistics are used :

$$J_{i,h} = \hat{\omega}_{i,h}' \Sigma_{i,h}^{-1} \hat{\omega}_{i,h}; C_{i,h} = \max_{1 \leq j \leq k} \frac{|\hat{\omega}_{j,i,h}|}{\sqrt{\sigma_{j,i,h}}} \tag{2}$$

4. Remove outlier impact.

The procedure is now performed on all the subjects, and the number of outliers is stored. The outliers are interpreted as spontaneous brain activations.

Conclusions

Two main methods have been proposed. The second one gave better results: it turned out that considering spontaneous brain activations as outliers of heavy tailed distributions was a successful idea. In addition, we successfully dealt with the problem of dimensionality reduction for multivariate time series and finally we created an R code that was able to identify and store all the interesting outliers accordingly to the mentioned algorithm.