Application of the Kalman Filter in Functional Magnetic Resonance Image Data

Valcir J. da C. Farias
Faculty of Statistics- Institute of Exact and Natural Sciences - Federal University of Pará – Brazil.

Marcus P. da C. da Rocha
Faculty of Mathematics- Institute of Exact and Natural Sciences - Federal University of Pará – Brazil.

Heliton R. Tavares (Corresponding author)
Faculty of Statistics- Institute of Exact and Natural Sciences - Federal University of Pará – Brazil.
E-mail: heliton@ufpa.br

Abstract

The Kalman-Bucy filter was applied on the preprocessing of the functional magnetic resonance image-fMRI. Numerical simulations of hemodynamic response added Gaussian noise were performed to evaluate the performance of the filter. After the proceeding was applied in auditory real data. The Kohonen's self-organized map was employed as tools to compare the performance of the Kalman's filter with another type of pre-processing. The results of the application of Kalman-Bucy filter for simulated data and real auditory data showed that it can be used as a tool in the temporal filtering step in fMRI data.

Keywords: Kalman Filter, Temporal Filtering, fMRI, Self-organizing maps.

1. Introduction

There are several techniques used to study the mapping of neural activation in the human brain noninvasively. Among these techniques we can mention positron emission tomography (PET), magnetoencephalography (MEG), Single Photon Emission Computer Tomography (SPECT), electroencephalography (EEG) and the functional magnetic resonance Imaging (fMRI).

From the variety of techniques involved in assessing brain activity through images, the most used technique is the acquisition of functional magnetic resonance images (fMRI) (Huettel et al., 2004). This technique is based on the phenomenon of nuclear resonance. The fMRI analyzes the hemodynamic variation and has the distinction of not submitting to voluntary harmful effects.

In a functional magnetic resonance image experiment, the subject remains within a resonance equipment where he is submitted to various stimuli. The neural activity to stimuli causes an increase in blood flow and oxygenation in blood vessels located in the brain region responsible for the activity in question. The

mechanism of BOLD (Blood Oxygen Level Dependent) contrast is the answer to the application of the experimental protocol consisting of periodic changes between stimulus and non-stimulus.

The temporal evolution of the BOLD effect (Hemodynamic Response Function) is a time series for each voxel of the functional images. The signal acquisition is done over several cycles of stimulation and rest so that the signal of interest is approximately a square wave with constant period.

The BOLD effect gives rise to the change in signal intensity captured which is restricted to a particular area of the brain and this variation is transitory in time with respect to the stimulus. However, these effects are low intensity and corrupted by noise and artifacts. One of the most challenging aspects of fMRI is to find methods to mitigate the noise present in the signal captured, because the technique is susceptible to various noise sources, causing inaccuracies temporal and spatial in the data.

Among the sources that cause these inaccuracies, we can mention the noise due to unintentional movement of the head, or noise derived from the heart beat or noise arising from the respiration or noise caused by other processes, such as differences in acquisition time of images, noise due to imperfections in the hardware and noise caused by interference from other stimuli unrelated to the task of interest.

The time series in an fMRI experiment usually have a low Signal to Noise Ratio (SNR). Thus, it becomes necessary some type of preprocessing for filtering noise in the signal. The pre-processing techniques are used to reduce the influence of most noises mentioned above on the signal. If fMRI data are processed without the application of a technique for pre-processing, the signal to noise ratio (SNR) will be low. Furthermore, the variability that is not associated with an experimental stimulus is reduced with the application of pre-processing. Such variability can reduce or even invalidate the power to detect brain areas considered active or inactive in an experiment. Among preprocessing stages, one can cite: the realignment; spatial filter and time filter.

In temporal filtering, the goal is to attenuate undesired components in the time series, without damaging the signal of interest, that is, it is important to keep after the filtering process the shape of the original signal.

The temporal filtering enhances the quality of fMRI data improving this way the signal to noise ratio. As opposed to the spatial filter which operates on each volume separately, the temporal filtering operates in the time series of the voxels separately, a series of each time. As in fMRI data, many processing techniques are applied directly in the time series, the temporal filtering usually occurs after all stages of the preprocessing.

Several authors applied temporal filters for preprocessing of fMRI data. Kruggel et al. (1999) made comparisons of filtering methods for fMRI datasets, they investigated the correction of baseline and signal restoration. Ngan et al. (2000) applied a time-varying filter based on the theoretical work of Nowak and Baraniuk (1999), such a filter operates under the stationary wavelet transform framework, where they showed that the filter had a good estimate of the true signal. Bannister et al. (1999) applied linear and non-linear filters for fMRI temporal analysis, and their results shown that in general the linear filter is better than the non-linear filter for the signal restoration.

In this paper, the Kalman-Bucy filter (Kalman, 1960; Kalman and Bucy, 1961; Rocha and Leite, 2003; Rocha et al., 2007) will be applied to realize the preprocessing of the fMRI data. It is made an adaptation in this method for filtering periodic signals so as to preserve morphology and amplitude of the original

signal. To make the analysis of the data after the application of the Kalman-Bucy filter, it is used the Kohonen's self-organized map (SOM) (Campelo et al., 2014; Haykin, 2001; Liao et al., 2008). The data will be also analyzed by the SOM without the application of the Kalman-Bucy filter for purpose of comparison and efficiency analysis of the procedure adopted here.

2. Material and Methods

The work of Kalman (Kalman, 1960) are referred in the literature as a significant contribution to the process of retrieving messages contaminated with noise, from the work of Wiener (Wiener, 1949). Essentially, the Kalman filter method is the Wiener solution to the problem of non-stationary processes, by means of transformation of an integral equation at a differential equation most appropriate for numerical calculation. From the description of the simplest Wiener-Kolmogorov problem, the stationary form, whose objective it is to obtain an optimum-filter function, time-invariant, *h(t)*, that operates on the measured signal *z(t)* and minimizes the mean square error between real output $\hat{x}(t)$ (estimated signal) and the desired output *x(t)*.

The cost function of minimizing is given by

$$E\{h(t)\} = E\{[\hat{x}(t) - x(t)]^2\}$$

that results in the normal equations between deviations and observations,

$$E\{[\hat{x}(t) - x(t)]\} \cdot z(t) = 0$$

The basic formulation of filter operation is given by the convolution integral:

$$\hat{x}(t) = \int_{-\infty}^{+\infty} h(\tau)\, z(t-\tau)\, d\tau \tag{1}$$

where *h(t)* is a unknown time-invariant operator (optimum filter) which it must satisfies the Wiener-Hopf equation:

$$\phi_{xz} = \int_{-\infty}^{+\infty} h_w(\tau)\phi_{zz}(t-\tau)d\tau \tag{2}$$

where ϕ_{xz} e ϕ_{zz} are the cross correlation and stochastic autocorrelation functions.

To specify the optimum filter of Wiener-Hopf (WH) it is necessary to solve equation (2), which presents the disadvantage of supposing stationarity of the process and this is not satisfied by the characteristics of the problem in study. It is necessary to make a generalizations of the problem of WH to non-stationarity processes ($h(t,\tau)$) and restrict the operation to a moving window. However, these conditions are not satisfied by integral of convolution. So, equation (2) will be rewritten in the form of a moving average according to Wiener-Kolmogorov theory which established the relation:

$$\underline{\underline{\phi}}_{xz}(t,\sigma) = \int_{t_0}^{t} \underline{\underline{h}}(t,\tau)\underline{\underline{\phi}}_{zz}(\tau,\sigma)d\tau \qquad (t_0 \le \sigma \le t), \tag{3}$$

and the estimated value equation is given by

$$\underline{\hat{x}}(t) = \int_{t_o}^{t} \underline{\underline{h}}(t,\tau)\,\underline{z}(\tau)\,d\tau\ .$$

(4)

Equation (3) is a booton integral and it is usually difficult to solve, but an approach was developed using state variables to describe nonstationary processes (Rocha and Leite, 2003). Kalman and Bucy (1961) have converted the integral in a first-order differential equations system which are more suitable to solve than (3).

2.1 Kalman-Bucy Solution

The problem solution ((3) and (4)) is initiated with an ordinary differential equation of order *N-1* that expresses the relation between the input and output in a system, given by

$$\sum_{n=1}^{N-1} a_n(t)\frac{d^n y(t)}{dt^n} = w(t) \qquad (a_1 = 1)\ .$$

(5)

The transformation of (5) for state variables $x_n(t)$ and $\dot{x}_n(t)$ is done by substituting the derivatives of *y(t)* according to the rule

$$x_1 = y,\ x_2 = \dot{y},\ x_3 = \ddot{y},\ldots,\ x_N = y^{N-1}$$

(6)

$$\dot{x}_1 = x_2,\ \dot{x}_2 = x_3,\ \dot{x}_3 = x_4,\ldots,\ \dot{x}_{N-1} = x_N$$

(7)

$$\dot{x}_N = -(a_1x_1 + a_2x_2 + \cdots + a_Nx_N) + w(t)\ .$$

(8)

The change of variables results in the state dynamic equations. The most general case, (6) and (7), are continuous and time-variant

$$\dot{x}(t) = \underline{\underline{F}}(t)\underline{x}(t) + \underline{\underline{G}}(t)\underline{w}(t) \qquad \text{(system)}$$

(9)

$$\underline{z}(t) = \underline{\underline{H}}(t)\underline{x}(t) + v(t)\,, \qquad \text{(measurement)}$$

(10)

where the vector $\dot{\underline{x}}(t)$ is state variable and the matrices $\underline{\underline{F}}(t),\ \underline{\underline{G}}(t)$ and $\underline{\underline{H}}(t)$ are dependent on *t*; the vector $\underline{w}(t)$ is state generating (signal); and the vector $\underline{z}(t)$ is the selected output from $\underline{\underline{H}}(t)$; $\underline{v}(t)$ is the added white noise to the process.

In Kalman–Bucy solution it is necessary to define the specific priori stochastic properties to the processes $\underline{z}(t), \underline{w}(t)$ and $\underline{v}(t)$ related in the autocorrelation and the stochastic cross correlation. Such a priori

condition is formed by the autocorrelations of white noise series, and by null cross correlations in the window of definition of the governing integral equation, and they are

$$E\{\underline{w}(t)\} = 0; \qquad \underline{\underline{\phi}}_{ww}(t,\tau) = E\{\underline{w}(t)\underline{w}^T(\tau)\} = \underline{\underline{W}}(t)\delta(t-\tau) \tag{11}$$

$$E\{\underline{v}(t)\} = 0; \qquad \underline{\underline{\phi}}_{vv}(t,\tau) = E\{\underline{v}(t)\underline{v}^T(\tau)\} = \underline{\underline{V}}(t)\delta(t-\tau) \tag{12}$$

$$\underline{\underline{\phi}}_{wz}(t,\tau) = E\{\underline{w}(t)\underline{z}^T(\tau)\} = 0 \qquad \underline{\underline{\phi}}_{wv}(t,\tau) = E\{\underline{w}(t)\underline{v}^T(\tau)\} = 0 \tag{13}$$

$$\underline{\underline{\phi}}_{xv}(t,\tau) = E\{\underline{x}(t)\underline{v}^T(\tau)\} = 0 \qquad \underline{\underline{\phi}}_{wx}(t,\tau) = E\{\underline{w}(t)\underline{x}^T(\tau)\} = 0\,, \tag{14}$$

where $\delta(t)$ is the Dirac's delta function that multiplied for $\underline{\underline{W}}(t)$ and $\underline{\underline{V}}(t)$ define diagonal matrices to the autocorrelation as priori conditions. Reorganizing the equations (8)-(11), one obtains the linear and nonlinear differential equations. Solutions in the continuous form are summarized in Table 1.

Table 1 Dynamics equations of the Kalman-Bucy Filter in the continue form.

System	$\dot{x}(t) = \underline{\underline{F}}(t)\underline{x}(t) + \underline{\underline{G}}(t)\underline{w}(t)$
Measurement	$\underline{z}(t) = \underline{\underline{H}}(t)\underline{x}(t) + \underline{v}(t)$
Initial conditions	$E\{x(0)\} = 0,\ E\{(\underline{x}(0) - \underline{\hat{x}}_0)(\underline{x}(0) - \underline{\hat{x}}_0)^T\} = P_0$
State estimate	$\underline{\dot{\hat{x}}}(t) = [\underline{\underline{F}}(t) - \underline{\underline{K}}(t)\underline{\underline{H}}(t)]\underline{\hat{x}}(t) + \underline{\underline{K}}(t)\underline{z}(t)$
Ricatti equation	$\underline{\underline{\dot{P}}}(t) = \underline{\underline{F}}(t)\underline{\underline{P}}(t) + \underline{\underline{P}}(t)\underline{\underline{F}}^T(t) - \underline{\underline{P}}(t)\underline{\underline{H}}^T(t)\underline{\underline{V}}^{-1}(t)\underline{\underline{H}}(t)\underline{\underline{P}}(t) - \underline{\underline{G}}(t)\underline{\underline{W}}(t)\underline{\underline{G}}^T(t)$
Gain matrix	$\underline{\underline{K}}(t) = \underline{\underline{P}}(t)\underline{\underline{H}}^T(t)\underline{\underline{V}}^{-1}(t)$

2.2 Kalman-Bucy Filter and fMRI

To model the hemodynamic response function with paradigm in block, it will be used the model proposed by Friston et al. (2000) with the parameters presented by Hu et al. (2009). Figure 1 show a simulated hemodynamic response function, *x(t)*, with zero mean and variance σ_a^2. To the signal *x(t)* is included an additive noise *v(t)* with zero mean and variance σ_v^2 to obtain the measured signal *z(t)*, that is, *z(t) = x(t) + v(t)*.

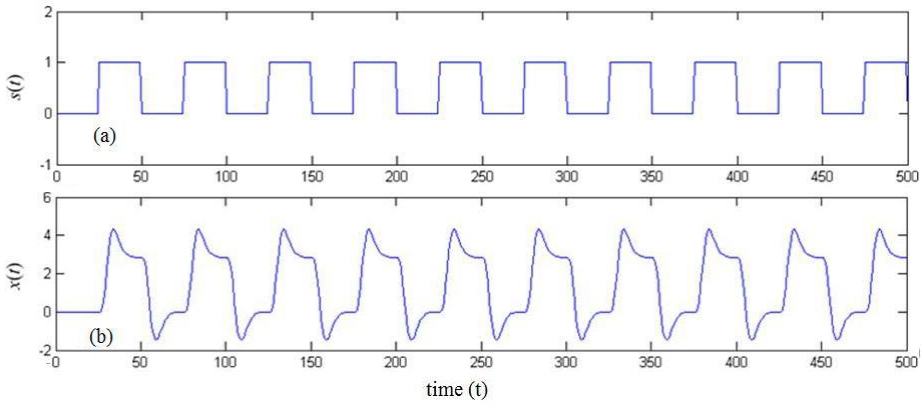


Figure 1 – (a) stimulus function for a block design. (b) Block design response

So, the stochastic properties to model are established, and they are summarized below:

$$E\{\underline{x}(t)\} = 0; \qquad \underline{\underline{\phi}}_{xx}(t_1, t_2) = E\{\underline{w}(t)\underline{w}^T(\tau)\} = \sigma_a^2 \delta(t-\tau) \tag{15}$$

$$E\{\underline{v}(t)\} = 0; \qquad \underline{\underline{\phi}}_{vv}(t_1, t_2) = E\{\underline{v}(t)\underline{v}^T(\tau)\} = \sigma_v^2 \delta(t-\tau) \tag{16}$$

$$E\{\underline{w}(t)\} = 0; \qquad \underline{\underline{\phi}}_{ww}(t_1, t_2) = E\{\underline{w}(t)\underline{w}^T(\tau)\} = \sigma_a^2 \delta(t-\tau) \tag{17}$$

To solve the proposed problem, it is necessary to define the state variables, as below:

$$x_1(t) = x(t) = y(t) \qquad x_2(t) = \dot{x}_1(t) = \dot{x}(t) = w(t)\,. \tag{18}$$

By comparing (18) with (6), (7) and (8), one obtain

$$\dot{x}_1 = 0 + w(t) \qquad y(t) = x_1(t) \tag{19}$$

and

$$\underline{\underline{F}} = 0 \qquad \underline{\underline{G}} = 1 \qquad \underline{\underline{H}} = 1\,. \tag{20}$$

By replacing (19) and (20) in Ricatti equation (Table 1), one obtain

$$\dot{\underline{\underline{P}}}(t) = -\underline{\underline{P}}(t)\underline{\underline{V}}^{-1}\underline{\underline{P}}(t) + \underline{\underline{W}}(t) \tag{21}$$

and the gain matrix is given by

$$\underline{\underline{K}}(t) = \underline{\underline{P}}(t)\underline{\underline{H}}(t)\underline{\underline{V}}^{-1}(t)\,. \tag{22}$$

By rearranging the terms, one obtains a differential equation of the covariance of the error in scalar form

$$\frac{dp(t)}{dt} = -\frac{p^2(t)}{\sigma_v^2} + \sigma_a^2 \tag{23}$$

and the gain in scalar form, given by

$$K(t) = \frac{p(t)}{\sigma_v^2}\,.$$

By solving (23), one obtain

$$p(t) = \sigma_v^2 \sqrt{\gamma}\, \tanh\left(\sqrt{\gamma} t\right)$$

and

$$K(t) = \sqrt{\gamma}\tanh\left(\sqrt{\gamma}t\right) \tag{24}$$

where

$$\gamma = \frac{\sigma_a^2}{\sigma_v^2}.$$

By replacing (24), (19) and (20) in State estimate (Table 1), one obtain the differential equation of the state estimate of the Kalman filter

$$\frac{d\hat{x}(t)}{dt} = -\sqrt{\gamma}\tanh(\sqrt{\gamma}t)\hat{x}(t) + \sqrt{\gamma}\tanh(\sqrt{\gamma}t)z(t). \tag{25}$$

It will be applied to the differential equation (25) which represents the state estimative of the Kalman filter in the scalar form. Based on these theoretical considerations, it was developed a computational algorithm for the implementation of the Kalman-Bucy filter on fMRI data, in which the quality of the filter will be specified by parameter γ. The Kalman-Bucy filter operates in the time series of voxels, *z(t)*, separately, a time series at a time. The application of the Kalman-Bucy filter on time series occurs in windows defined by sequential cycles of rest and stimulation. Figure 2 illustrates the design of the windows for the application of the Kalman-Bucy filter. Each window is formed by a period of activation alternated with a period of rest. Figure 3 shows the application of the Kalman filter with windows in a certain interval ($t_o \le \tau \le T$). The hyperbolic shape of the linear operator *h*(t) is evidenced after the application of the filter in the interval ($t_o \le \tau \le T$). The parameter γ was considered constant in the windows.

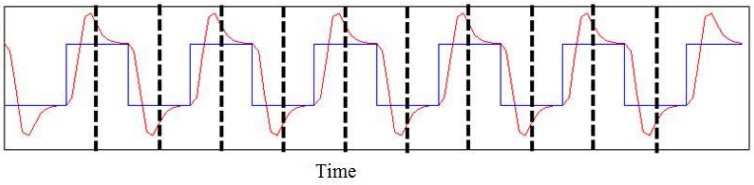


Figure 2 - Windows for application of the filter.

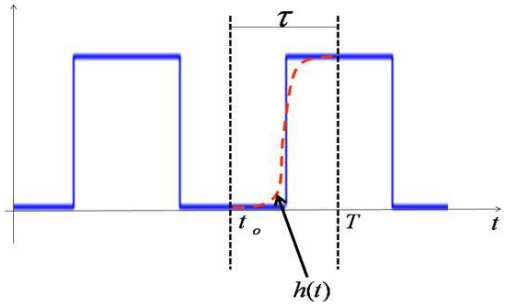


Figure 3 - Illustration of the filter application window ($t_o \le \tau \le T$) and the hyperbolic form of the linear operator

2.3 Self-Organizing Maps

fMRI data were analyzed with Kohonen´s SOM (Kohonen, 2001) using an implementation available in the literature (Fischer and Henning, 1999; Ngan et al., 2002; Peltier et al., 2003; Campelo et al., 2014). Kohonen´s SOM is an artificial neural network where neurons are disposed as a uni- or bi-dimensional grid layout. In a bi-dimensional layout, the geometry is free and can be rectangular, hexagonal, triangular etc. In a SOM, each neuron in a grid is represented by a probability distribution function of the input data. The SOM algorithm responsible for map formation begins initializing the grid neurons weights with random values, which can be obtained from the input data. In the present work, it was used a bidimensional grid of dimension 10 × 10. Each neuron in the grid is connected to every element of the input dataset, i.e., the dimension of weights *m* is the same as the input dataset:

$$m_i = [m_{i1}, m_{i2}, \ldots, m_{in}]^T \in \Re^n$$

where *n* indicates the total amount of points available in the time series generated by the fMRI experiment. After each iteration *t* of the ANN, we selected randomly a vector from the input dataset, given by:

$$x_i = [x_1, x_2, \ldots, x_n]^T \in \Re^n$$

which indicates the time series of a given voxel from the fMRI dataset. Then *x* is compared to all the weights of the grid, with frequent use of the minimum Euclidean distance as similarity criterion for choosing a winner neuron (Fischer and Henning, 1999; Peltier et al., 2003). However, the correlation as a similarity criterion reveals itself better than the conventional Euclidean distance (Liao et al., 2008), the choice of the winning neuron *c* given by:

$$m_c = \arg\max\{corr(x(t), m_i(t))\},$$

with $i = 1, \ldots, M$, where *M* is the total number of neurons in the grid, $m_c(t)$ represents the time series of the winner *c* and *corr(x(t),mi(t))* is the correlation coefficient between *x(t)* and *mi(t)*.

The updating of the weight vector $m(t+1)$ in time *t+1*, with $t = 1, 2, \ldots$ is defined by:

$$m_i(t+1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)] \quad (26)$$

which is applied to every neuron on the grid that is within the topological neighborhood-kernel from the winner neuron. Equation (26) has the goal of approximating the weight vector m_i of neuron *i* towards the input vector, following the degree of interactions h_{ci}. This approach transforms the grid, after training, in a topologically organized characteristic map, in the sense that adjacent neurons tend to have similar weights. A function frequently used to represent the topological neighborhood-kernel h_{ci} is the Gaussian function, which is defined by:

$$h_{ci}(t) = \alpha(t) \exp\left\{-\frac{\|r_c - r_i\|}{2\sigma^2(t)}\right\},$$

where $\alpha(t)$ is the learning rate, which has to gradually decrease along time to avoid that new data gathered after a long training session could compromise the knowledge already sedimented in the ANN; r_c and r_i

determine the discreet position of neurons *c* and *i* in the grid and $\sigma(t)$ defines the topological neighborhood radius, i.e., defines the full-width at half-maximum (FWHM) of the Gaussian kernel. The parameters $\alpha(t)$ and $\sigma(t)$ gradually decrease by t/τ (τ is a time constant) after each iteration *t*, following an exponential decay.

After the learning process of the SOM, the input data with similar patterns appear in groups in neighboring neurons on the map. However, often the amount of groups is unknown or of complex discrimination, making it difficult to distinguish such groups. A proposal for detection and automatic segmentation of the groups in the SOM map is through clustering techniques. The purpose of the clustering techniques is to form similar groups, ie groups that have a high degree of correspondence or similarity.

Several methods for grouping of neurons in the SOM have been proposed, in this paper will apply the hierarchical clustering (HC), for more details see (Liao et al., 2008, Campelo et al., 2014).

There are several mechanisms that can evaluate the quality of the generated map obtained after the learning process. In the present work it was used the quantization error:

$$E_q = \frac{1}{N}\sum \|x - m_c\|^2 \tag{27}$$

The quantization error is defined as the mean square error corresponding to the difference between each characteristic vector *x* and the winner neuron m_c, where *N* is the total number of patterns (Santos et al., 2019).

2.4 Synthetic data

To compare different values of the parameter γ, are generated artificial time series of voxels considered active, i.e. time series containing evidence of BOLD signal. Using an artificial signal, it is possible to compare the original signal, noisy signal and filtered signal, determining in this way, the factor of noise reduction δ, which it will be a metric of performance in this work.

$$\delta = \sqrt{\frac{\sum_k [z(k) - x(k)]^2}{\sum_k [\hat{x}(k) - x(k)]^2}}, \tag{26}$$

where *x(k)* is the original signal without noise, *z(k)* is the signal with noise and the $\hat{x}(k)$ is the filtered signal and *k* is the sample index.

The time series were simulated from block design paradigm, formed of 6 blocks of activation and 6 blocks of rest, where the first block was composed of rest. On the time series it was added a Gaussian noise. From these time series were simulated the 3D data with 64×64 voxels (Figure 4) with 128 samples in time. In Figure 4, the area considered active is composed with 49 voxels, the gray matter was composed of 1349 voxels and the other voxels are of the background of the image. On the time series was added noise with SNR of 5dB, 0dB and -5 dB.

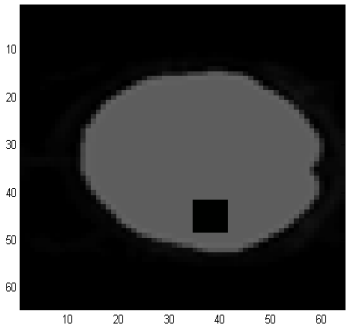


Figure 4- Simulated image: a rectangular area inside the gray matter represents a region considered active.

2.4 Real data

The fMRI experiment used a 1.5 T Siemens scanner (Magnetom Vision, Erlangen, Germany), with the following parameters for EPI (echo-planar imaging) sequences: TE = 60 ms, TR = 4.6 s, FA = 90º, FOV=220 mm, and slice thickness of 6.25 mm. 64 cerebral volumes with 16 slices each were acquired with a matrix dimension of 128x128. During the experimental procedure the subject received auditory stimulation in a blocked design, with 5 stimulation blocks (27.5s each) intercalated with 6 resting blocks (27.5 s each). During the task, the subject listened passively to a complex story with a standard narrative structure. After the test, the subject had to inform to the experimenter its comprehension of the story content.

Acquired images were preprocessed with the software SPM8 (Statistical Parametric Mapping) in order to increase the signal-to-noise ratio (SNR) and to eliminate incident noise associated with the hardware, involuntary movements of the head, cardiac and respiratory rhythms, etc.

3. RESULTS

3.1 Synthetic data

In several tests performed, it will be presented here only the synthetic time series that were contaminated with the noise whose SNR was from -5 dB and the values of 250, 200, 150, 100 and 50 were considered to the parameter γ. Figures 5(a) and 5(b) show the original and contaminated (SNR of -5 dB) time series, respectively. Figures 5(c), 5(d), 5(e), 5(f) and 5(g) show the results of the filtering applying by Kalman-Bucy filter, respectively, with $\gamma = 250$, $\gamma = 200$, $\gamma = 150$, $\gamma = 100$ e $\gamma = 50$. One can observe that smaller the value of the parameter γ, more the signal tends to smoothen. However, the higher the level of smoothing increases the possibility of losing of the desired properties of the signal. Therefore, it is important to choose a suitable value for the parameter γ, such that the filtering remove undesired components of the time series

and keep the desired properties of the signal. That is, the quality of results from the application of the Kalman filter is related to the choice of parameter.

Through factor noise reduction (26) it was selected the parameter γ that best filters the signal. For an SNR of -5 dB it was found a factor of noise reduction δmax = 2.1535 for γ = 20. Figure 6 shows the evolution of the noise reduction factor for SNR =-5dB, and Figure 7(a) shows the comparison between a time series with SNR = -5 dB (blue line) and that series filtered (red line) with γ = 20, which provides the best noise reduction factor (δmax = 2.1535). Based only in Figure 7(a), it can be assumed that there is a loss in signal amplitude significantly after applying the filter, but when comparing the filtered signal with the original signal, it is noticed that the filter has performed well (Figure 7(b)), recovering good part of original form of the signal.

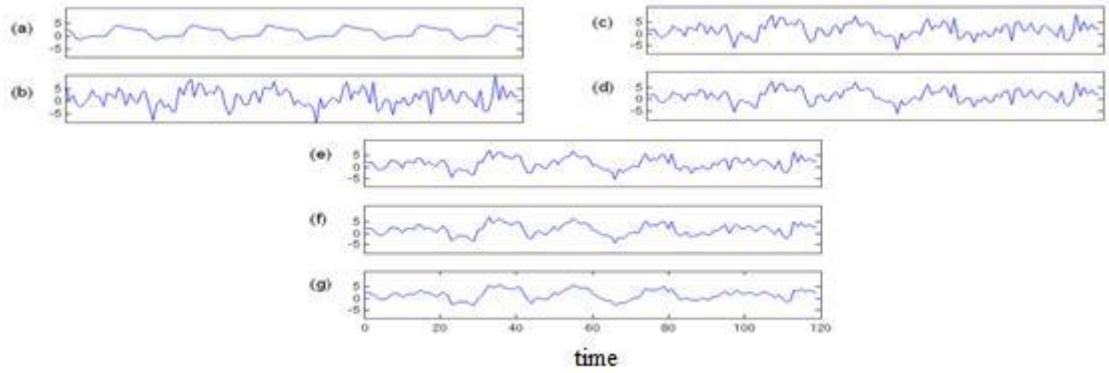


Figure 5. (a) Original signal; (b) Contaminated signal with SNR=-5dB; and filtered signal with: (c) γ=250; (d) γ=200; (e) γ=150; (f) γ=100; (g) γ=50.

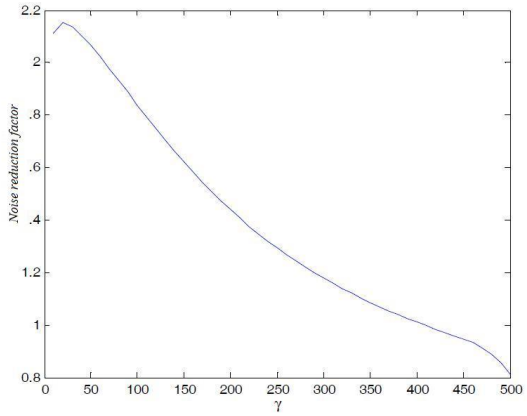


Figure 6. Noise reduction factor for a time serie contaminated with noise of SNR =-5dB.

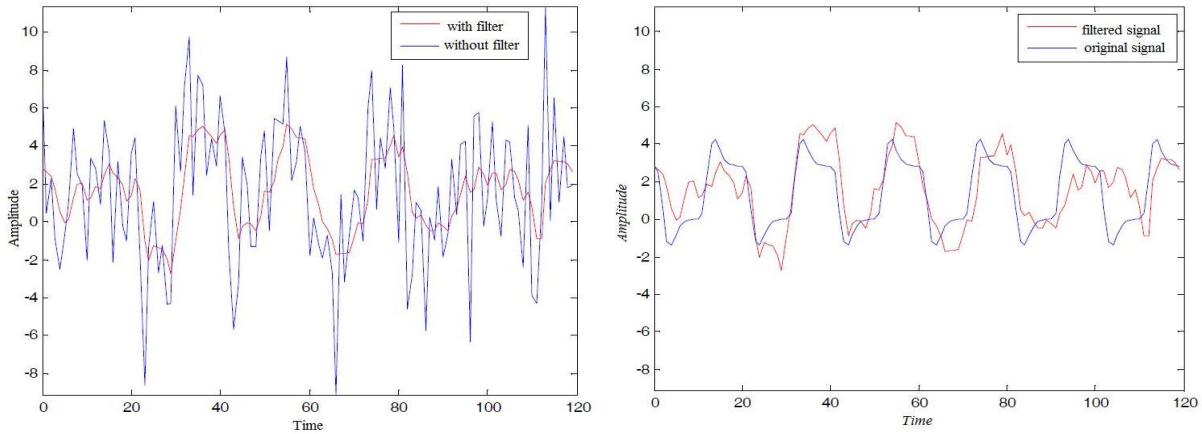
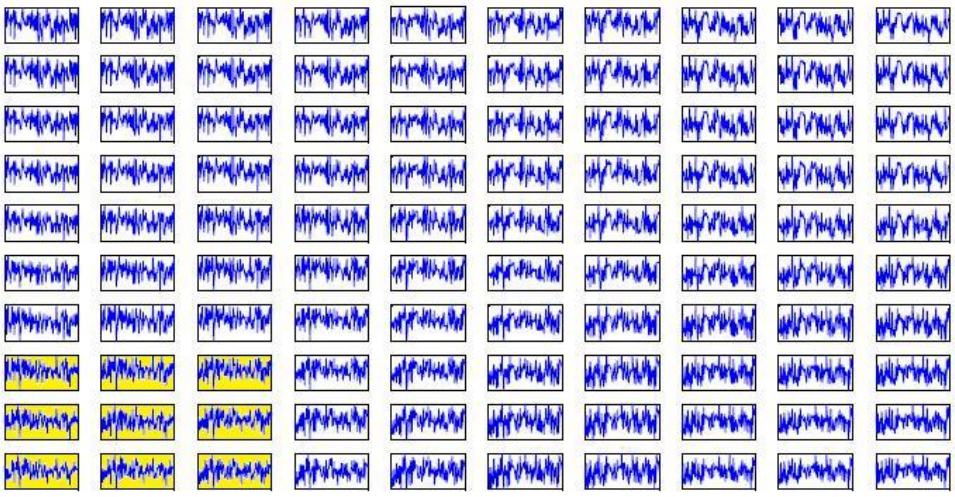


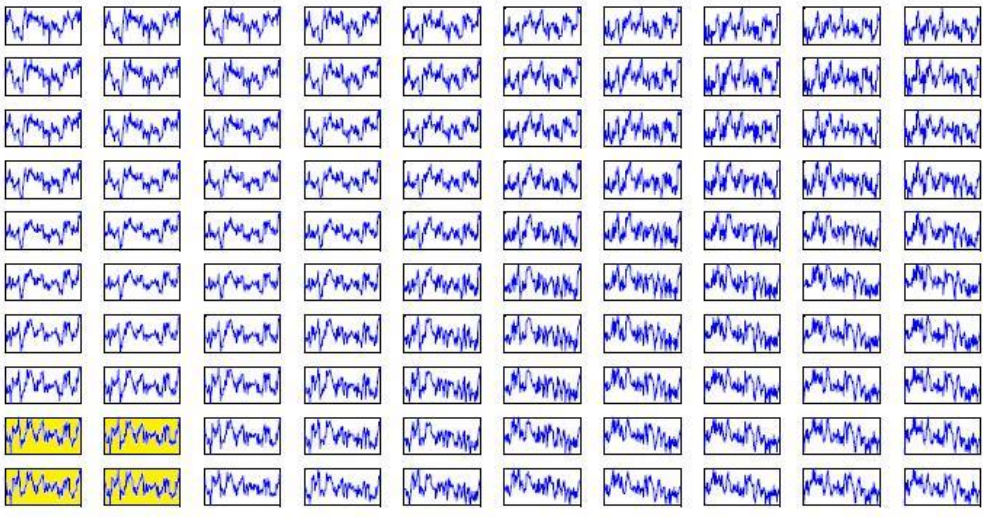
Figure 7. (a) comparison between the time serie contaminated with noise of SNR = -5 and the time serie filtered by the Kalman-Bucy filter, com $\gamma = 20$ e $\delta_{max} = 2{,}1535$. (b) comparison between original signal and the filtered signal.

The self-organizing map (SOM) was applied to verify performance of the Kalman-Bucy filter on the time series of the simulated slice. For the first analysis was considered the time series of the simulated slice without applying the Kalman filter. For this data set, has been started the training stage of the SOM in a total of 150 iterations. Upon completion of this step, the SOM presented a neurons 10x10 grid shown in Figure 8(a). After the training of the SOM it was applied the hierarchical cluster to determinate the group of the neurons that present pattern of the activation. The neurons on the left bottom (highlighted in yellow) represent the activation.

In the second analysis it was applied the Kalman-Bucy filter in the dataset of the slice simulated. Then began the training of the SOM keeping the set of initial weights used in the analysis without the application of the filter (the previous analysis). Figure 8(b) shows the neurons 10x10 grid for this analysis, although to in smaller number, neurons that represent activation signals are sharper than of Figure 8(a), according to visual analysis made at the weights of the neurons grid.



(a)



(b)

Figure 8. (a) Neurons 10x10 grid for data of simulated slice without the application of the Kalman-Bucy filter; (b) Neurons 10x10 grid for data of simulated slice with the application of the Kalman-Bucy filter.

To generate the map of activation (Figure 9) it was calculated the correlation coefficient between all voxels of the areas of the brain and the average weights of the neurons belonging to group candidates. The correlation coefficient between these two signals is used as the threshold to find the voxels with greater possibilities of being actives.

The results were also compared from the quantization error curves (27). The quantization error, among other things also serves as a controller and as a comparator of quality of learning of an artificial neural network. Figure 10 shows the quantization error curves with and without the application of the Kalman filter on time series of the slice simulated. It is observed that the stabilization in both cases occurs approximately from the 70th iteration, with the amplitude of the quantization error of the curve by applying the filter being slightly smaller than the curve without the filter.

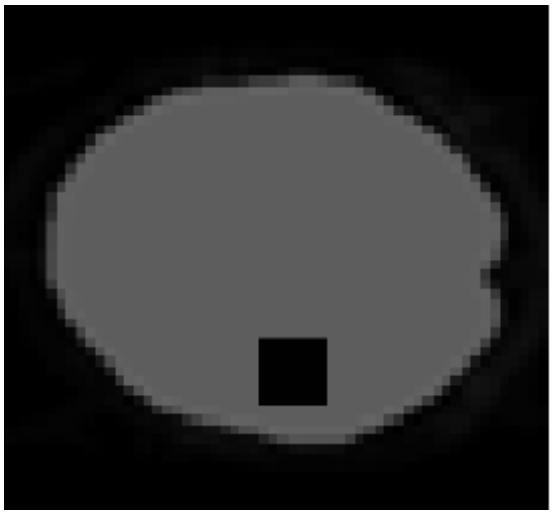


Figure 9. Activation map. The Black area inside the gray matter correspond to active region obtained by SOM.

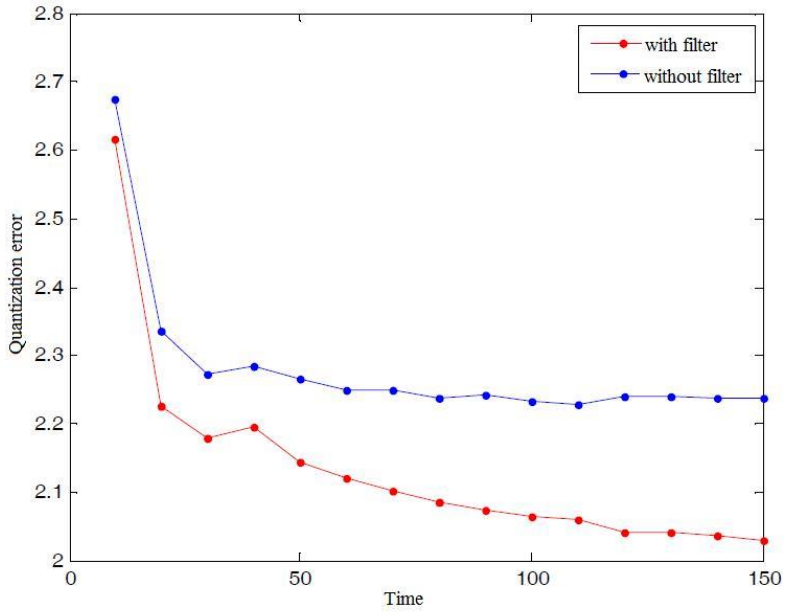


Figure 10. Comparison of the quantization error for the simulated slice with SNR =-5dB: with the application of the Kalman filter (red line), and without the Kalman filter (blue line).

3.2 Real data

In this section it is presented the results of the application of the Kalman filter on experimental fMRI data described at the Section 2.4. Once again, the self-organizing map was applied to assess the real fMRI data without and with the application of the Kalman filter.

Figure 11 (a) and (b) show evolution of the quantization error to data without and with the application of Kalman filter with $\gamma = 80$, respectively. It is observed that the amplitude of quantization error curve of filtered data is lightly lesser than that data without the filtering. After training the SOM and applying the hierarchical clustering on the eighth slice image of fMRI experimental data preprocessed with the Kalman-Bucy filter, the 10 × 10 neuron grid is shown in Figure 12. The group of neurons highlighted in yellow indicates activation pattern.

The map of activation made from the correlation coefficients between the time series of each voxel of the image (eighth slice) and the average of the weights highlighted in Figure 12 is shown in Figure 13. Figure 14 is presented the map of activation for data without the application of the Kalman filter. It is observed that images are similar, however, the auditory areas in Figure 13 are slightly greater than that in Figure 14.

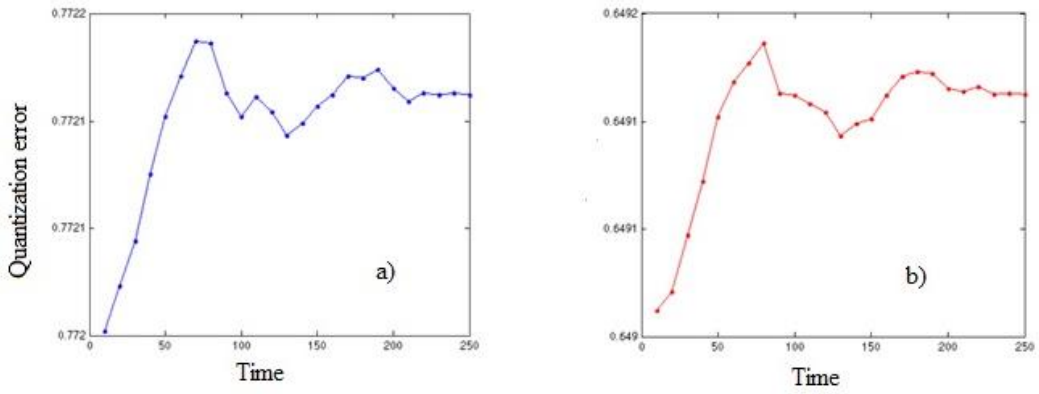


Figure 11. Quantization error: (a) without filter application; (b) with filter application.

The Kalman-Bucy filter left the weight of the neurons of the SOM grid sharper than the signal of the same weight for data without the filter application, It is possible that voxels with low BOLD signal, but that belong to active region have been detected. Explaining, thereby, the small differences between the images.

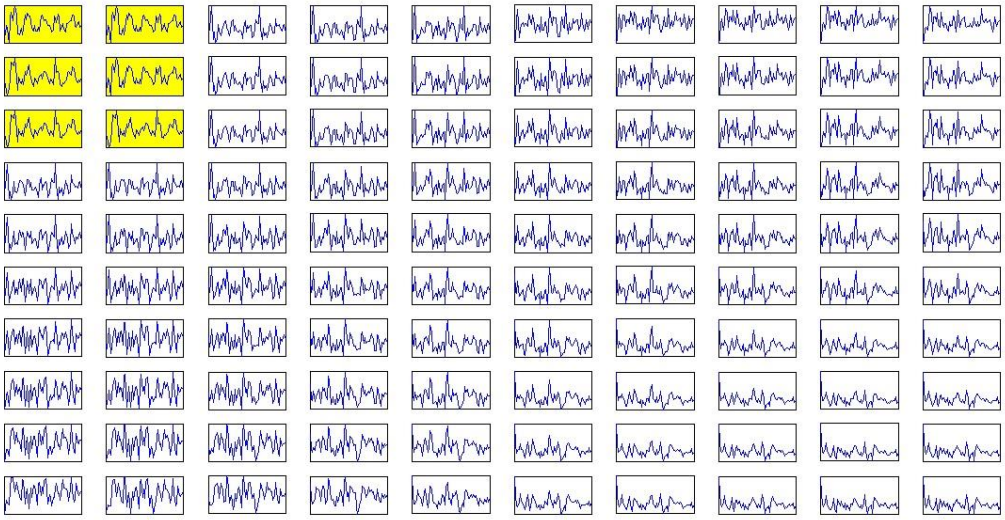


Figure 12. 10x10 neuron grid after executing the SOM algorithm, the highlighted neuron group corresponds to the activity patterns.

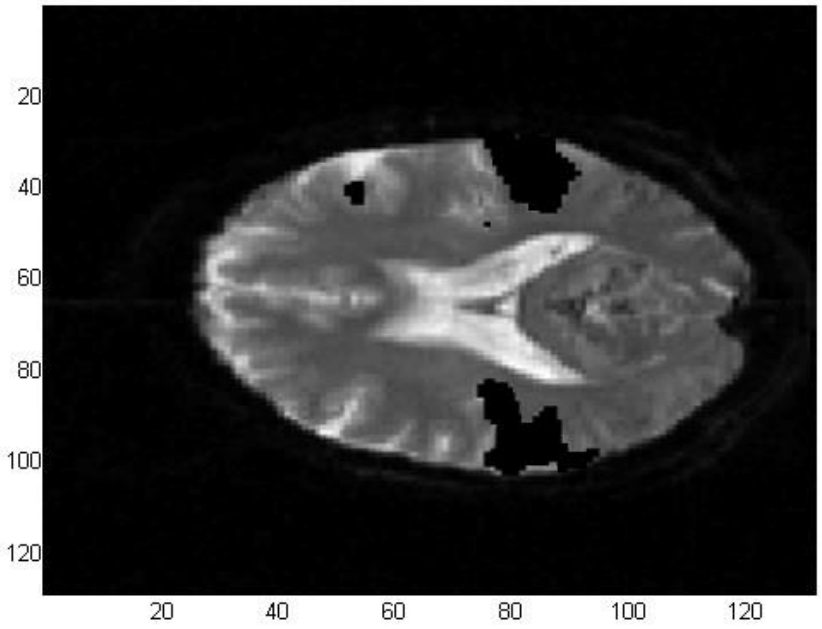


Figure 13. Result of auditory stimulation in the cerebral cortex according to SOM analysis, using data filtered by Kalman with $\gamma = 80$.

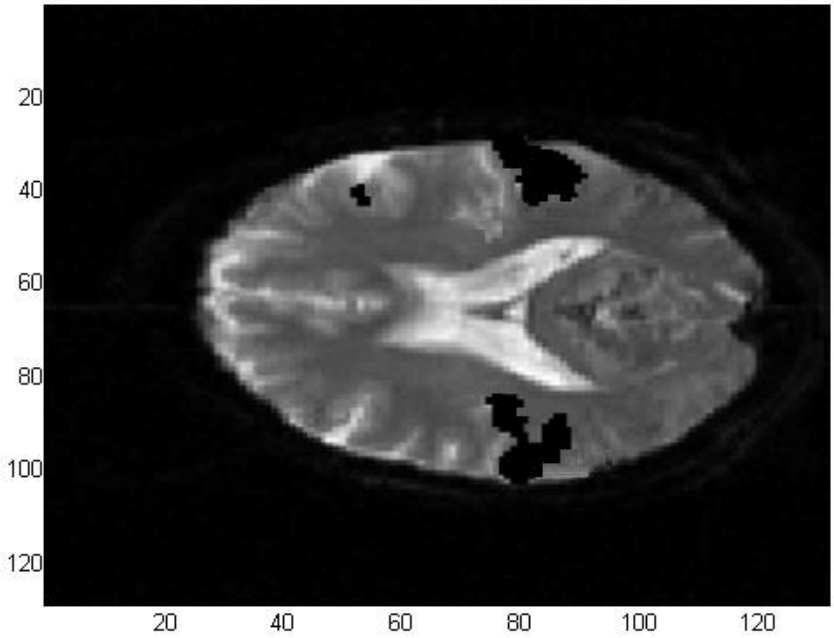


Figure 14. Result of auditory stimulation in the cerebral cortex according to SOM analysis, using data without Kalman filtering.

4. CONCLUSIONS

In this work the Kalman-Bucy approach in the differential form was applied to time series of fMRI data. The experimental results showed that the Kalman filtering can mitigate noise of fMRI data by altering slightly some details and amplitude of events and, however, preserving the characteristic morphology of signal.

In simulated and real data, Kohonen self-organizing map and hierarchical clustering were applied as analyze tools. It was used the quantization error of the SOM to verify the performance of the Kalman filter in the fMRI data preprocessing.

It was observed that the amplitude of the SOM quantization error was minor with the Kalman filter application than without filter application on the data. This result was observed for every SNR values. Regarding the stabilization of the quantization error curve, it was noted that the curves showed similar behaviors. The results showed that Kaman filter can be applied as a tool in the stage of temporal filtering of fMRI data.

5. REFERENCE

Bannister, P.; Smith, S.; Brady, M.; Flitney, D.; Woolrich, M. Evaluating Lowpass Filters for fMRI Temporal Analysis. Department of Clinical Neurology, University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK. www.fmrib.ox.ac.uk/analysi

Campelo, A. D. de S.; FARIAS, V. J. da C.; Tavares, H. R.; ROCHA, M. P. C. Self-organizing maps and entropy applied to data analysis of functional magnetic resonance images. Applied Mathematical Sciences (Ruse), v. 8, pp. 4953-4969, 2014.

Fischer, H., Henning, J., Neural-network based analysis of MR time series. Magnetic Resonance in Medicine 41:124-131, 1999.

Friston, K. J., Mechelli, A., Turner, R., Price, C. J. "Nonlinear responses in fMRI: the balloon model, volterra kernels, and other hemodynamics," NeuroImage, vol. 12, no. 4, pp. 466 - 477, 2000.

Haykin, S. Neural Networks Principles and practice. 2 ed., Bookman, São Paulo, 2001.

Hu, Z.; Zhao, X.; Liu, H.; Shi, P. Nonlinear analysis of the bold signal. EURASIP Journal on Advances in Signal Processing, v. 2009, pp. 1 - 13, 2009.

Huettel, S. A., Song, A. W., e Mccarthy, G., 2004. Functional Magnetic Resonance Imaging. Sinauer Associates.

Kalman, R. E. A new Approach to Linear Filtering and Predction Problems**.** ASME, Series D, Journal of Basic Engineering, 82, pp. 35–45, 1960.

Kalman, R. E., Bucy, R. E. New results in linear filtering and prediction theory**.** ASME, Series D, Journal of Basic Engineering, 83, pp. 95–107, 1961.

Kohonen, T. Self-Organizing Maps. 3 ed.; Springer: Berlim, 2001.

Kruggel, F.; Von Cramon, D. Y.; Descombes, X. Comparison of Filtering Methods for fMRI Datasets. Neuroimage, 10, pp. 530–543, 1999.

Liao, W., Chen, H., Yang, Q., Lei, X. Analysis of fMRI Data Using Improved Self-Organizing Mapping and Spatio-Temporal Metric Hierarchical Clustering. IEEE Transactions on Medical Imaging, vol. 27, pp. 1472-1483, 2008.

Ngan, S. C.; LaConte, S. M.; Hu, X. Temporal Filtering of Event-Related fMRI Data Using Cross-Validation. NeuroImage, 11, pp. 797–804, 2000.

Nowak, R. D., and Baraniuk, R. G. Wavelet-domain filtering for photon imaging systems. IEEE Trans. Image Process., 8, pp. 666–678,1999.

Peltier, S. J., Polk T. A., Noll D.C. Detecting low-frequency functional connectivity in fMRI using a self-organizing map (SOM) algorithm. Hum. Brain Mapp., vol. 20, pp. 220-226, 2003.

Rocha, M. P. C., Leite, L. W. B. Treatment of geophysical data as a non-stationary process. Computational and Applied Mathematics, Vol. 22, N. 2, pp. 149–166, 2003.

Rocha, M. P. C.; Leite, L. W. B.; Santos, M. L.; Farias, V. J. C. Attenuation of multiple in reflection seismec data using Kalman-Bucy filter. Applied Mathematics and Computation, v. 189, p. 805-815, 2007.

Ribeiro, R. C. M., Quadros, T. A., Ausique, J. J. S., Chase, O. A., Campos, P. S. da S., Júnior, P. C. dos S., Almeida, J. F. S. de, & Marques, G. T. (2019). Forecasting incidence of tuberculosis cases in Brazil based on various univariate time-series models. International Journal for Innovation Education and Research, 7(10), 894-909. https://doi.org/10.31686/ijier.vol7.iss10.1841

Wiener, N. Extrapolation, Interpolation, and Smoothing of Stationary Time Series. New York. John Wiley and Sons. 1949.