

Ocular Reduction in EEG Signals Based on Adaptive Filtering, Regression and Blind Source Separation

S. ROMERO,^{1,2} M. A. MAÑANAS,^{1,2} and M. J. BARBANOJ³

¹Department of Automatic Control (ESAI), Biomedical Engineering Research Center, Universitat Politecnica de Catalunya (UPC), Barcelona, Spain; ²CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Barcelona, Spain; and ³Drug Research Center (CIM), Research Institute of Sant Pau Hospital, Department of Pharmacology and Therapeutics, Universitat Autònoma de Barcelona (UAB), Barcelona, Spain

(Received 11 April 2008; accepted 20 October 2008; published online 5 November 2008)

Abstract—Quantitative electroencephalographic (EEG) analysis is very useful for diagnosing dysfunctional neural states and for evaluating drug effects on the brain, among others. However, the bidirectional contamination between electro-oculographic (EOG) and cerebral activities can mislead and induce wrong conclusions from EEG recordings. Different methods for ocular reduction have been developed but only few studies have shown an objective evaluation of their performance. For this purpose, the following approaches were evaluated with simulated data: regression analysis, adaptive filtering, and blind source separation (BSS). In the first two, filtered versions were also taken into account by filtering EOG references in order to reduce the cancellation of cerebral high frequency components in EEG data. Performance of these methods was quantitatively evaluated by level of similarity, agreement and errors in spectral variables both between sources and corrected EEG recordings. Topographic distributions showed that errors were located at anterior sites and especially in frontopolar and lateral–frontal regions. In addition, these errors were higher in theta and especially delta band. In general, filtered versions of time-domain regression and of adaptive filtering with RLS algorithm provided a very effective ocular reduction. However, BSS based on second order statistics showed the highest similarity indexes and the lowest errors in spectral variables.

Keywords—Electroencephalography (EEG), Electrooculography (EOG), Ocular artifacts, Regression analysis, Adaptive filtering, Blind source separation (BSS), Independent component analysis (ICA).

INTRODUCTION

Quantitative analysis of electroencephalographic (EEG) signals is a helpful support in the diagnosis of psychiatric and neurological disorders, and in the evaluation of drug effects in functional states of the

brain. EEG changes are often quantified by the calculation of spectral variables in different frequency bands of clinical interest: delta, theta, alpha, and beta. It is known that non-cortical interferences, such as heart, ocular, and muscular activities, contribute to EEG recordings. Procedures to detect and remove these artifacts are very important and necessary because they could lead to wrong results and conclusions.

Ocular artifacts are the most relevant interference because they occur very frequently and their amplitude can be several times larger than brain scalp potentials. As the eyeball moves, the electric field composed by cornea and retina changes and it produces the electrooculographic (EOG) signals. Additionally, some neural activity is recorded by EOG electrodes because they are located near the head. Muscle activity associated with the eyes or near them can also interfere in the EOG signal. Ocular activity propagates across the scalp: vertical ocular projection following the anterior–posterior direction, and horizontal mainly affecting lateral–frontal scalp regions.⁷

Linear regression analysis is the most common approach for reducing eye movement artifacts. This method estimates and removes the EOG component existing in each EEG lead. However, regression procedures do not take into account the bidirectional contamination, i.e., EOG recording is also contaminated by cerebral activity, and consequently this cerebral information could be relevant and can also be cancelled in the EEG recordings after linear subtraction. Gasser *et al.*¹¹ proposed previously to apply regression procedure a low pass filtering of EOG signals in order to reduce the cancellation of high frequency cerebral components from EEG data.

He *et al.*¹² proposed the adaptive filtering by recursive least squares (RLS) algorithm which was applied to simulated data in He *et al.*¹³ in order to compare with time-regression procedure. However, the

Address correspondence to S. Romero, Department of Automatic Control (ESAI), Biomedical Engineering Research Center, Universitat Politecnica de Catalunya (UPC), Barcelona, Spain. Electronic mail: sergio.romero-lafuente@upc.edu

authors did not consider in their study the bidirectional contamination. They recognized that it would deteriorate the method and they pointed out modifications such as smoothing the reference EOG.

In addition, other approaches based on Blind Source Separation (BSS) have been proposed.^{18,22,31} BSS procedures decompose the multichannel EOG and EEG recordings into source components. Sources, which are related with artifacts, are deleted by removing their contributions onto the scalp sensors. The first proposed component-based procedure was Principal Component Analysis (PCA).²⁰ However, it has several drawbacks related to the assumption of orthogonality between neural and ocular activities, and the difficulty to separate eye movement artifacts from cerebral activity when their amplitudes are similar. More recent approaches use blind source extraction based on Independent Component Analysis (ICA). Assumption for ICA-based techniques is that sources must be statistically independent, not just uncorrelated, and this is mostly true with regard to brain and ocular components.

There are different algorithms and principles to estimate the components, generally based on second order and higher order statistics. However, most of the EEG studies with ICA application identify the artifactual source components by visual inspection following a time-consuming subjective criterion. Moreover, quantitative evaluation of the performance of each ocular correction method is difficult because it would be necessary to get available previous knowledge corresponding to the true ocular and cerebral activities, and this is not possible because they cannot be separately measured due to bidirectional contamination. There have been few empirical studies measuring the effectiveness of ocular reduction techniques on simulated data composed by mixtures of cerebral and ocular activities.^{19,32} However, these studies showed opposite conclusions: regression and PCA based algorithms were suggested in Wallstrom *et al.*,³² adaptive filtering was proposed in He *et al.*,¹² and BSS techniques based on second-order statistics were recommended in Kierkels *et al.*¹⁹ and Romero *et al.*²⁵ in spite of adaptive filtering was not evaluated in both. Moreover, in spite of the theoretical advantages of BSS based approaches, a very recent study based on real data and using expert scorers for identifying ocular artifacts concluded that regression procedure could significantly reduce ocular artifacts better than ICA when few EEG channels were available.²⁷

Regarding simulated data, in Wallstrom *et al.*³² coefficients for the linear mixture were determined based partially on normalizing random variates. Due to this random mixture, simulated EEG signals did not correspond to specific EEG leads. In Kierkels *et al.*¹⁹

forced and fast ocular movements were simulated by Boundary element method (BEM). BEM calculated the potential at different positions in an arbitrary shape volume, and with representative tissues conductivity values taken from literature. Furthermore, the location of simulated brain dipoles was chosen randomly based on real data from one EEG channel. This study was based on situations with open eyes. In He *et al.*¹³ an arbitrary gain function for ocular propagation was used to obtain simulated EEG and EOG signals. This function was frequency dependent and constructed based on Gasser *et al.*¹⁰ In Romero *et al.*,²⁵ simulated data was generated from spontaneous EOG and EEG signals with closed eyes and with normal (not fast forced) eye movements, which is the common procedure used in clinical routine. The simplest linear instantaneous mixtures were carried out by factors from regression coefficients and by using EEG recordings considered free of ocular artifacts with a not very restrictive condition. In the present study, simulated EEG and EOG data were obtained by more general linear mixtures whose coefficients were calculated by means of multi-input single-output (MISO) linear models applied to a database composed of real EEG and EOG recordings. A more restrictive condition than used in Romero *et al.*²⁵ was applied in order to consider EEG recordings more free of artifacts. Bidirectional contamination, where EOG recordings are also interfered by some cerebral activity, was also taken into account by linear mixtures in both directions.

In this study, a fully automatic procedure for ocular correction from spontaneous EEG signals based on BSS described in Romero *et al.*²⁵ was used. The objective of the present study was twofold. The first aim was to generate simulated EEG and EOG signals in order to reproduce a real clinical situation in practice. It was achieved by obtaining a proper propagation model of cerebral and ocular activities across the scalp based on the theory of system identification. The second aim was to evaluate the performance of different ocular reduction methods (linear regression, adaptive filtering and BSS) on this simulated data. This evaluation was quantitatively carried out by level of similarity, agreement, and spectral target variables used commonly in clinical EEG studies.

METHODOLOGY

Subjects and Instrumentation

Twenty-four young healthy volunteers of either gender (aged 22.50 ± 3.75 years) were selected for the study from a larger database. Spontaneous EEG and

EOG signals, sampled at 100 Hz, were acquired during 10 non-consecutive 3-min periods following a vigilance-controlled condition with eyes closed. During the vigilance-controlled recordings, the experimenter tried to keep the volunteers alert by acoustic stimulation. Twelve subjects with normal-high eye movements (called ‘ocular group’) were selected following visual inspection of EOG signals. The remaining 12 volunteers (called ‘cerebral group’) were carefully chosen based on the criterion of no apparent eye movements (3 consecutive minutes with EOG amplitudes lower than $25 \mu\text{V}$). Nineteen EEG recordings were obtained by means of scalp electrodes placed according to the international 10/20 system: Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2, referenced to averaged mastoids. Additionally, vertical and horizontal EOG (VEOG and HEOG, respectively) signals were recorded. The VEOG signal was acquired from mid-forehead (2.5 cm above the pupil) to the average of one electrode below the left eye and one electrode below the right eye (2.5 cm below the pupil). The HEOG signal was obtained from the outer canthi. Multichannel EOG and EEG signals were recorded using a band pass filter between 0.3 and 45 Hz.

Simulated Data

Simulated EEG Signals

Neural sources (EEG_s) corresponding to 3-min consecutive epochs were obtained from the 19 EEG channels for each subject belonging to ‘cerebral group’ and selected from the entire large database. The selection criterion was that no samples from VEOG and HEOG signals exceeded $25 \mu\text{V}$. This threshold was lower than $40 \mu\text{V}$, which was used in a previous study,²⁵ in order to better consider a reduced potential ocular activity in EEG data. Neural sources were obtained by high-pass filtering the EEG channels with a cut-off frequency of 0.5 Hz in order to reduce very low frequency components which could be possibly more related to ocular activity. Interference from ocular activity was added to the neural sources. This interference was calculated by estimation of linear models between EOG and EEG recordings following the procedure explained below.

Ten 5-s epochs with high EOG activity were chosen from each one of the 12 volunteers belonging to the ‘ocular group.’ A multi-input single-output (MISO) linear model was estimated for each 5-s epoch and for each lead. Models had two inputs corresponding to VEOG_O and HEOG_O signals (subindex_O referred to signals derived from the ‘ocular group’), and one output that was each one of the 19 EEG_O channels.

Due to bidirectional contamination, it is well known that cerebral activity also affects to EOG signals. This activity was reduced for a better estimation of models in order to obtain the interferences from ocular activity. For this purpose, in each model, EEG_O and EOG_O channels, corresponding to outputs and inputs, respectively, were low-pass filtered with the cut-off frequency corresponding to highest value of the 99% (f_{99}) of the total energy of these VEOG_O and HEOG_O signals. Thus, the remainder 1% of signal energy was not considered as ocular activity (neural activity, electrode noise, power line interference, etc.) in the EEG_O recordings. The 99% cut-off frequencies obtained were 6.34 ± 4.04 and 8.20 ± 5.01 Hz for VEOG_O and HEOG_O , respectively, as mean and standard deviation for all epochs and volunteers. This consideration was supported by some studies which suggested that most of the high frequency range in the EOG signal is of neural origin.¹¹ Besides, this low-pass filtering procedure on EEG_O and EOG_O did not cause apparently shift delays or variations in the ocular waveform.

A linear autoregressive with exogenous input $\text{ARX}(N, M)$ structure, that is described in Eq. (1), was used for the parametric model in each i lead ($i = \text{Fp1}, \text{Fp2}, \dots, \text{O1}, \text{and O2}$), and p epoch ($p = 1, 2, \dots, 10$).

$$\begin{aligned} \text{EEG}_{Oip}(n) &+ \sum_{k=1}^{N-1} a_{ip}(k) \cdot \text{EEG}_{Oip}(n-k) \\ &= \sum_{k=0}^{M-1} b1_{ip}(k) \cdot \text{VEOG}_{Op}(n-k) \\ &+ \sum_{k=0}^{M-1} b2_{ip}(k) \cdot \text{HEOG}_{Op}(n-k) + e(n) \end{aligned} \quad (1)$$

where n represents the sample time; a_{ip} , $b1_{ip}$, and $b2_{ip}$ denote the unknown model parameters; N and M are the orders associated with the output and the inputs, respectively; and $e(n)$ is the unknown error mapping.²¹ The same order M for both inputs was considered because they corresponded to the same kind of activity, ocular, whose propagation is wanted to be modeled; and this propagation takes place in the same medium, through the head, and the same directions: from EOG channels to EEG leads. Akaike’s final prediction error (FPE) was used to select the ARX model orders. Figure 1 shows the FPE as a function of different orders for output and inputs. The orders selected for output and inputs were 4 and 3, respectively, that corresponded to a FPE low enough and it did not decrease sensitively with higher orders. Estimation of model parameters was implemented by least squares estimation method, using QR-factorization for overdetermined linear equations.

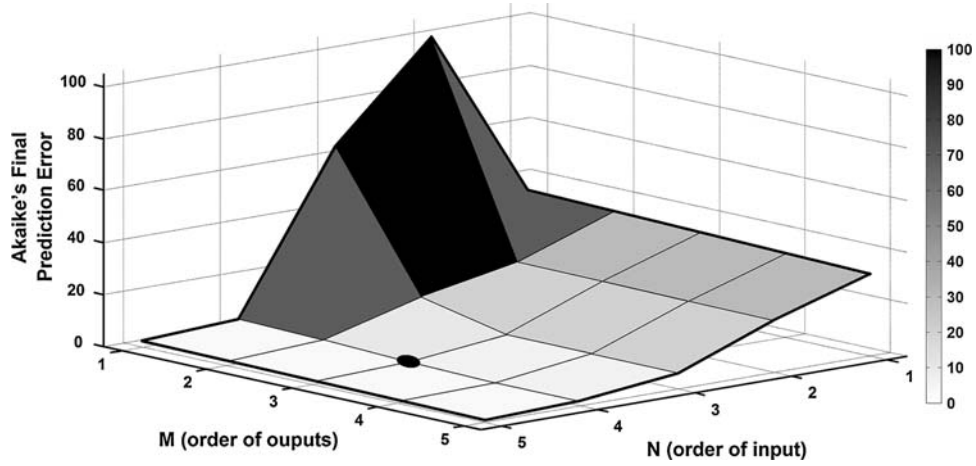


FIGURE 1. Mean Akaike's Final Prediction Error (FPE) as a function of the input and outputs orders of the ARX model representing the ocular propagation. These mean values were calculated by averaging FPE from all EEG leads and all the 12 subjects.

Once the parameters for each of the 10 ARX models were estimated for each EEG lead in each volunteer, two frequency responses (for both vertical and horizontal ocular activity) for every model were evaluated. These 10 frequency responses were averaged in spectral domain for both $VEOG_O$ and $HEOG_O$, for each EEG_O lead and volunteer. That is, they were representative of the contamination of both vertical and horizontal projection of the ocular activity up to each EEG lead in every volunteer. Finally, equivalent 256-order FIR filters $[\bar{b}_{iVEOG}(k) \ \bar{b}_{iHEOG}(k)]$ were estimated from average frequency responses using the frequency sampling method.¹⁷ These FIR filters, that were specific for each channel in each of the 12 volunteers, were applied to EOG_O sources ($VEOG_s$ and $HEOG_s$) in order to simulate EEG recordings (mixed EEG signals: EEG_m) from 12 simulated volunteers.

$$EEG_{im}(n) = EEG_{is}(n) + \sum_{k=0}^{255} \bar{b}_{iVEOG}(k) \cdot VEOG_s(n-k) + \sum_{k=0}^{255} \bar{b}_{iHEOG}(k) \cdot HEOG_s(n-k) \quad (2)$$

for each i channel ($i = Fp1, Fp2, \dots, O1,$ and $O2$).

Figure 2 shows the mean magnitude frequency response of FIR filters corresponding to $VEOG_O$ and $HEOG_O$ for each EEG_O lead. It was obtained by averaging all frequency responses from each subject of the 'ocular group.'

Simulated EOG Signals

Ocular activity sources (EOG_s) corresponding to 3-min consecutive epochs were extracted from the two

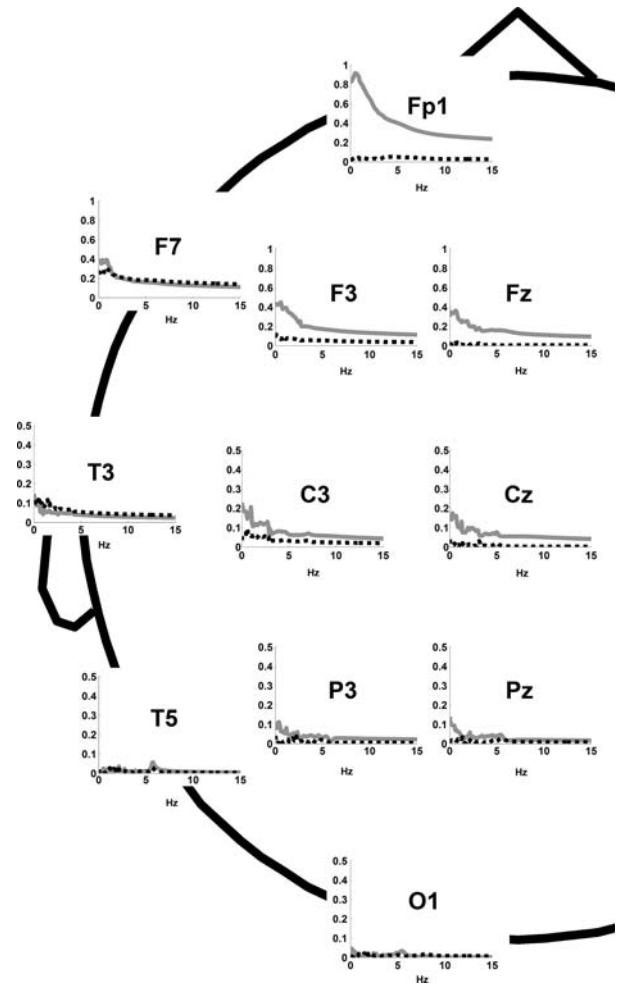


FIGURE 2. Mean magnitude frequency responses corresponding to ocular contamination from vertical (solid grey line) and horizontal (dashed black line) projections to each EEG channel.

EOG channels for each volunteer belonging to ‘ocular group.’ Ocular sources were obtained by low pass filtering these EOG channels with the cut-off frequency corresponding to highest value of the 99% of their total energy. Interference from cerebral activity was added to the ocular sources. Analogously to simulated EEG signals, this interference was calculated by the estimation of linear models between EEG_C and EOG_C recordings (subindex_C referred to signals derived from the ‘cerebral group’) following an analogous procedure that is explained below.

In this case, 10 five-second epochs with no apparent eye movements (absolute EOG_C amplitudes below 15 μV) were selected from each one of the 12 volunteers belonging to the ‘cerebral group.’ Neural contamination of EOG_C channels was obtained from the anterior placed EEG_C electrodes, which are the nearest ones to the eyes. A MISO linear model was identified by an ARX(N , M) structure for each EOG_C channel. It had four inputs, corresponding to the frontopolar EEG_C channels (Fp1 and Fp2) and the lateral-frontal ones (F7 and F8), and one output related to each EOG_C channel. Equation (3) describes as an example the model for VEOG_O channel.

$$\begin{aligned} \text{VEOG}_{C_p}(n) &+ \sum_{k=1}^{N-1} a_{\text{VEOG}_p}(k) \cdot \text{VEOG}_{C_p}(n-k) \\ &= \sum_{k=0}^{M-1} b_{1\text{VEOG}_p}(k) \cdot \text{EEG}_{C_p \text{ Fp1}}(n-k) \\ &\quad + b_{2\text{VEOG}_p}(k) \cdot \text{EEG}_{C_p \text{ Fp2}}(n-k) \\ &\quad + \sum_{k=0}^{M-1} b_{3\text{VEOG}_p}(k) \cdot \text{EEG}_{C_p \text{ F7}}(n-k) \\ &\quad + b_{4\text{VEOG}_p}(k) \cdot \text{EEG}_{C_p \text{ F8}}(n-k) + e(n) \end{aligned} \quad (3)$$

where n , p , a_{VEOG_p} , $b_{1\text{VEOG}_p}$, $b_{2\text{VEOG}_p}$, $b_{3\text{VEOG}_p}$, $b_{4\text{VEOG}_p}$, N , M , and $e(n)$ represent the same as in Eq. (1).

EEG_C signals for the inputs were previously high-pass filtered with a cut-off frequency of 0.5 Hz in order to reduce very low frequency components which could be possibly more related to ocular activity. Besides, the target spectral variables used for the evaluation of ocular filtering performance were calculated in frequencies higher than 0.5 Hz. Thus, this high pass filtering did not affect results and conclusions but it permitted a better estimation of model in order to obtain the cerebral contamination. The selected order for N and M were 4 and 3, respectively, and were chosen by the same criterion as in section ‘‘Simulated EEG Signals.’’ Remainder steps of the procedure were analogous to the ones explained in section ‘‘Simulated EEG Signals.’’ Finally, four 256-order FIR filters were obtained for VEOG_C [$\bar{b}_{\text{Fp1} \rightarrow \text{VEOG}}(k)$, $\bar{b}_{\text{Fp2} \rightarrow \text{VEOG}}(k)$, $\bar{b}_{\text{F7} \rightarrow \text{VEOG}}(k)$, $\bar{b}_{\text{F8} \rightarrow \text{VEOG}}(k)$] and HEOG [$\bar{b}_{\text{Fp1} \rightarrow \text{HEOG}}(k)$, $\bar{b}_{\text{Fp2} \rightarrow \text{HEOG}}(k)$,

$\bar{b}_{\text{F7} \rightarrow \text{HEOG}}(k)$, $\bar{b}_{\text{F8} \rightarrow \text{HEOG}}(k)$] channels from each volunteer belonging to ‘cerebral group.’

These FIR filters, that were specific for each one of the 12 volunteers, were applied to four leads of EEG_C sources (EEG_{Fp1 s}, EEG_{Fp2 s}, EEG_{F7 s}, and EEG_{F8 s}) in order to obtain EOG recordings (mixed EOG signals, VEOG_m and HEOG_m) from 12 simulated subjects:

$$\begin{aligned} \text{VEOG}_m(n) &= \text{VEOG}_s(n) + \sum_{k=0}^{255} \bar{b}_{\text{Fp1} \rightarrow \text{VEOG}}(k) \\ &\quad \cdot \text{EEG}_{\text{Fp1 s}}(n-k) + \sum_{k=0}^{255} \bar{b}_{\text{Fp2} \rightarrow \text{VEOG}}(k) \\ &\quad \cdot \text{EEG}_{\text{Fp2 s}}(n-k) \\ &\quad + \sum_{k=0}^{255} \bar{b}_{\text{F7} \rightarrow \text{VEOG}}(k) \cdot \text{EEG}_{\text{F7 s}}(n-k) \\ &\quad + \sum_{k=0}^{255} \bar{b}_{\text{F8} \rightarrow \text{VEOG}}(k) \cdot \text{EEG}_{\text{F8 s}}(n-k) \end{aligned} \quad (4)$$

$$\begin{aligned} \text{HEOG}_m(n) &= \text{HEOG}_s(n) + \sum_{k=0}^{255} \bar{b}_{\text{Fp1} \rightarrow \text{HEOG}}(k) \\ &\quad \cdot \text{EEG}_{\text{Fp1 s}}(n-k) + \sum_{k=0}^{255} \bar{b}_{\text{Fp2} \rightarrow \text{HEOG}}(k) \\ &\quad \cdot \text{EEG}_{\text{Fp2 s}}(n-k) \\ &\quad + \sum_{k=0}^{255} \bar{b}_{\text{F7} \rightarrow \text{HEOG}}(k) \cdot \text{EEG}_{\text{F7 s}}(n-k) \\ &\quad + \sum_{k=0}^{255} \bar{b}_{\text{F8} \rightarrow \text{HEOG}}(k) \cdot \text{EEG}_{\text{F8 s}}(n-k) \end{aligned} \quad (5)$$

Figure 3 shows the mean frequency response of FIR filters corresponding to the four anterior EEG_C leads for VEOG_C and HEOG_C recordings. It was obtained by averaging all frequency responses from each subject of the ‘cerebral group.’

Ocular Filtering Methods

Linear Regression

Multiple regression analysis is based on subtracting a fraction of EOG channels from contaminated EEG signals. This method assumes that the recorded EEG signals (in this case simulated mixed, EEG_{i m}) are an instantaneous superposition of the true or uncontaminated EEG (EEG_{i s}) and the ocular activity (VEOG_s

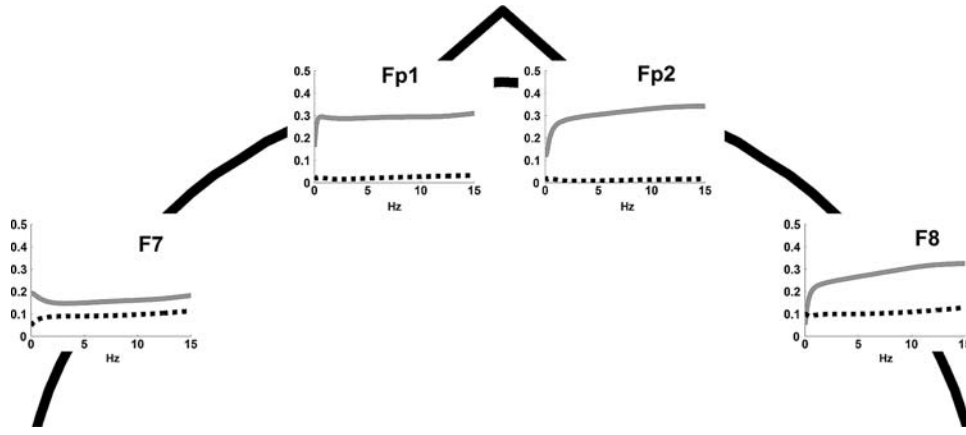


FIGURE 3. Mean magnitude frequency responses corresponding to cerebral activity contamination from frontopolar and lateral-frontal EEG channels to vertical (solid grey line) and horizontal (dashed black line) EOG signals.

and HEOG_s). Then, corrected EEG signals ($\text{EEG}_{i \text{ corr}}$) are calculated by Eq. (6):

$$\text{EEG}_{i \text{ corr}}(n) = \text{EEG}_{i \text{ m}}(n) - \alpha_i \cdot \text{VEOG}_m(n) - \beta_i \cdot \text{HEOG}_m(n) \quad (6)$$

where α_i and β_i denote the propagation factors of the VEOG_m and HEOG_m signals, respectively, up to the $\text{EEG}_{i \text{ m}}$ lead. Equation (4) was applied to each i lead with its corresponding factors α_i and β_i . These factors were calculated using only samples with high VEOG_m or HEOG_m amplitudes in order to improve their estimation.²⁸

This method can easily be implemented but it causes distortion of the corrected EEG signals because it does not take into account the bidirectional contamination between ocular and cerebral activities. In order to improve the performance of the regression procedure, a variant of the method was proposed.¹¹ Cancellation of cerebral information can be reduced by low-pass filtering with a cut-off frequency of 7.5 Hz the VEOG_m and HEOG_m signals before the application of regression subtraction.

Adaptive Filtering by Recursive Least Squares (RLS)

Ocular cancellation by adaptive filtering uses the available references to the interference, in this case, vertical and horizontal EOG channels. Adaptive filters self-adjust a vector of weights $w_i(j)$ according to an algorithm of optimization. These weights model the contamination of the ocular activity to the EEG leads.³⁰ In fact, adaptive filtering is an improvement of linear regression: propagation factors do not need to be neither constant nor frequency independent.²⁹ In terms of modeling, it is assumed that recorded EEG signals are a mixture, not necessarily instantaneous, of cerebral and ocular activities.^{12,13} Corrected EEG signals by RLS method can be calculated by:

$$\begin{aligned} \text{EEG}_{i \text{ corr}}(n) = & \text{EEG}_{i \text{ m}}(n) - \sum_{j=1}^M w_{\text{VEOG}}(j) \\ & \cdot \text{VEOG}_m(n+1-j) - \sum_{j=1}^M w_{\text{HEOG}}(j) \\ & \cdot \text{HEOG}_m(n+1-j) \end{aligned} \quad (7)$$

where w_{VEOG} and w_{HEOG} denote the vector of weights of length M that model the contamination of the VEOG_m and HEOG_m signals, respectively, up to the $\text{EEG}_{i \text{ m}}$ lead.

Although there are several adaptive filtering algorithms, RLS has shown best stability, efficiency and fast convergence. There are two parameters involved in the adaptive filtering method: the number of weights M and the forgetting factor λ . In theory, the value M is determined by characteristics of the EOG-EEG transfer function but the performance of the adaptive filter is not sensitive to this value, as it will be demonstrated in section ‘‘Ocular Artifact Removal.’’ The forgetting factor adjusts the weight of the previous samples to update the filter coefficients, and it depends on the stability of the relationship between the reference inputs (VEOG_m and HEOG_m signals) and the primary input ($\text{EEG}_{i \text{ m}}$ lead). Mathematically, λ is related to a window that indicates the number of previous samples that are used to calculate the current filter coefficients. The size of this window can be estimated by solving $\lambda^N = 0.5$, where N are the number of sample points. A study of the performance of adaptive filtering for several values of M and λ was performed and shown in section ‘‘Ocular Artifact Removal’’ in order to select their proper values.

Component Based Techniques

Component-based approaches decompose multi-channel EOG and EEG data into a mixture of source

ocular and cerebral signals. BSS problem assumes that a set of m recorded EOG and EEG channels are composed by a mixture of n source components, generally with $n \leq m$. Corrected EEG signals can be recovered by a re-mixing process rejecting the ocular sources, i.e., using only the cerebral, or non-ocular, sources. There are several approaches to solve the so-called BSS problem which can be divided in second-order statistics (SOS) techniques and ICA algorithms based on higher-order statistics (HOS).

SOS techniques formulate the hypothesis that sources are only uncorrelated, which is a weak form of statistical independence. Two SOS-based methods were evaluated in this study: PCA and SOBI (Second-Order Blind Identification). PCA transforms multi-channel data set by a rotation, in such a way that components in the new coordinates become uncorrelated and orthogonal. SOBI uses the time structure information provided by the sources to improve the estimation of the model. SOBI decomposition procedure consists on diagonalizing time-lagged covariance matrices.⁴

The difference between SOS- and HOS-based algorithms resides in how the sources are modeled: if sources are assumed mutually independent, HOS are essential to solve the BSS problem. There are several procedures to measure statistical independence, basically based on approaches of non-gaussianity, maximum likelihood estimation and mutual information.¹⁶ Two ICA algorithms were considered in this study: INFOMAX and FastICA. INFOMAX is an information theoretic based algorithm that obtains the independence maximizing the entropy of a neural processor output.³ FastICA is a computationally efficient algorithm that uses a fixed-point iteration scheme that allows faster convergences than gradient descent methods applied in other ICA algorithms. FastICA maximizes non-gaussianity as a measure of statistical independence based on the central limit theorem.¹⁵

Decomposition procedures were performed using the functions included in the ICALAB toolbox v3 for Matlab (<http://www.bsp.brain.riken.jp/ICALAB/>).⁶

Automation of ocular contamination identification is an important step for artifact correction based on decomposition techniques. Classical correction methods have been carried out by visual inspection of source components in order to decide which ones were related to artifacts. In order to overcome this subjectivity, several studies have attempted to find some rules using statistical properties like kurtosis or entropy.^{2,9} In this study, automatic ocular artifact identification was based on frequency and scalp topography aspects of the source components and had been previously described in Romero *et al.*²⁵ The criteria to remove a

source component related to ocular activity were defined by the following rules:

- (1) Relative power in delta band had to be greater than a specific high percentage.
- (2) Projection strength onto the EOG electrodes had to be higher than a threshold.
- (3) Projection strengths onto the EEG electrodes had to follow a gradient decreasing from anterior to posterior brain regions.
- (4) The maximum of the projection strengths on the EEG electrodes had to be higher than a threshold.

Once the components related with ocular artifacts were detected, corrected EEG signals were obtained by reconstruction of the components excluding the ocular related ones.

Evaluation of Ocular Artifact Reduction

Quantitative performance of each ocular correction technique was assessed by using similarity and the most important spectral target variables which are often utilized in clinical EEG studies. In previous studies,^{13,25,32} mean squared error was used to measure similarity between true and corrected EEG data, but it could not be the most appropriate index due to its dependence with the signals scaling. In this article, similarity of waveforms between original cerebral sources and the corrected EEG data was assessed by calculating the Pearson's correlation coefficient between them. Additionally, Bland–Altman analysis is the most direct way to assess agreement between two quantitative measurements.⁵ In this analysis, the difference of the paired two signals (corrected or mixed EEG signals minus neural sources) is plotted against the mean of both signals, and agreement is achieved if the 95% of the data points lie within the ± 1.96 standard deviation (range of agreement) of the mean difference. The more narrow the range of agreement, the more precise the method. Constant biases were checked by calculating the mean of the differences. Finally, regression was performed in order to evaluate proportional biases.

Power spectral density (PSD) functions were calculated for each EEG channel by means of Welch periodogram using a Hanning window of 5-s duration.¹ Then, nine target variables were calculated from PSD function: total power (0.5–35 Hz), and absolute and relative power in four different frequency bands:²⁶ delta (0.5–3.5 Hz), theta (3.5–7.5 Hz), alpha (7.5–13 Hz), and beta (13–35 Hz). Relative spectral powers were computed by dividing the power for a specific band by the total power. These target variables were calculated for further comparison before and after ocular correction. Relative percentage errors for each

target spectral variable were calculated between the values obtained for the original sources and those calculated for corrected EEG channels, following Eq. (8):

$$\begin{aligned} \%ERROR &= 100 \cdot \text{abs} \\ &\left(\frac{\text{spectral variable}_{\text{sources}} - \text{spectral variable}_{\text{corrected}}}{\text{spectral variable}_{\text{sources}}} \right) \end{aligned} \quad (8)$$

where abs denotes the absolute value for further averaging of these percentage errors among simulated subjects.

RESULTS

Simulated Data

Twelve mixtures of 3-min duration were generated by mixing each one of the 12 ocular activity sources from the ‘ocular group’ with each one of the 12 cerebral sources related to the ‘cerebral group.’ Each mixture reproduced a simulated subject whose true ocular and cerebral activities were known. Figure 4 shows, as an example, a time domain 5-s epoch of the simulation procedure. Ocular and cerebral sources of

two subjects are presented in Fig. 4a. Bidirectional contamination between ocular and cerebral activities could be observed in the simulated mixtures in Fig. 4b.

Additionally, extent of eye movement artifacts was evaluated by the topographical distribution pattern of the magnitude frequency responses in Fig. 2: DC gains decreased considerably from anterior to posterior locations for vertical eye movements, and they decreased following a lateral axis for horizontal eye movements. It agreed with the well known physiological information mentioned in section “Introduction.” This similarity as well as the hemispheric symmetry were logical and they provided consistency to the estimated propagation models and hence to the simulated signals.

Ocular Artifact Removal

Visual comparative evaluation of ocular reduction procedures on spontaneous EEG signals can be carried out in Fig. 5. Regression, adaptive and PCA-based correction methods removed some neural activity, which was also recorded in EOG channels, especially on anterior regions (frontopolar leads). Performance of HOS-based techniques seemed to be visually better for posteriorly EEG locations. In fact, these HOS-based methods eliminated quite accurately ocular

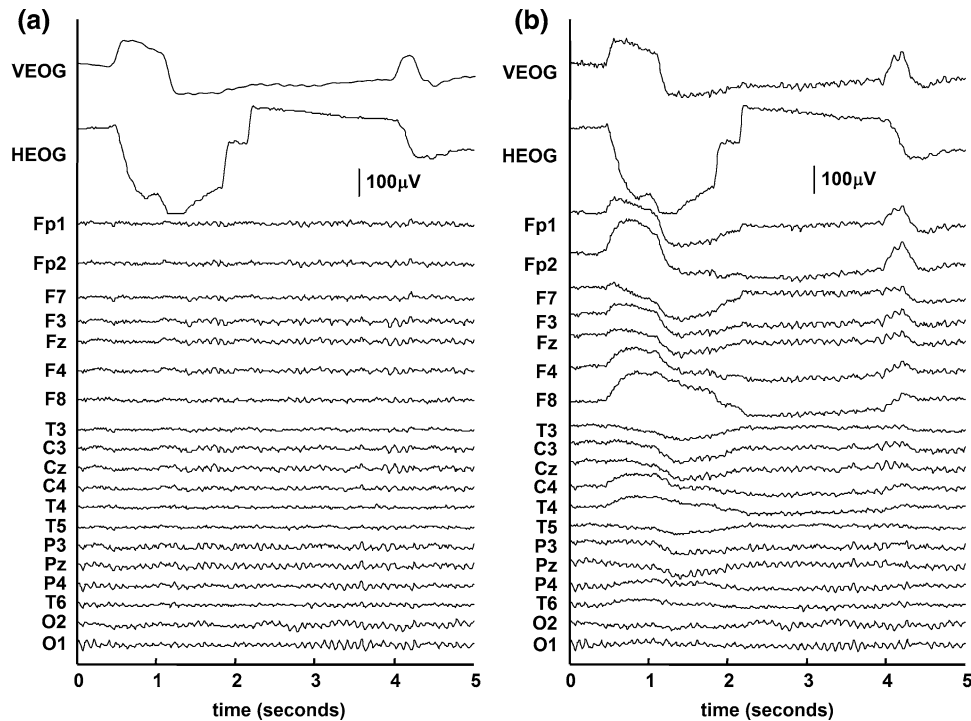


FIGURE 4. (a) Five second segment corresponding to ocular and cerebral activity sources from two subjects belonging to ‘ocular’ and ‘cerebral’ groups, respectively; (b) Five second epoch corresponding to mixed EOG and EEG signals obtained after applying the convolution mixing procedure to the sources in (a).

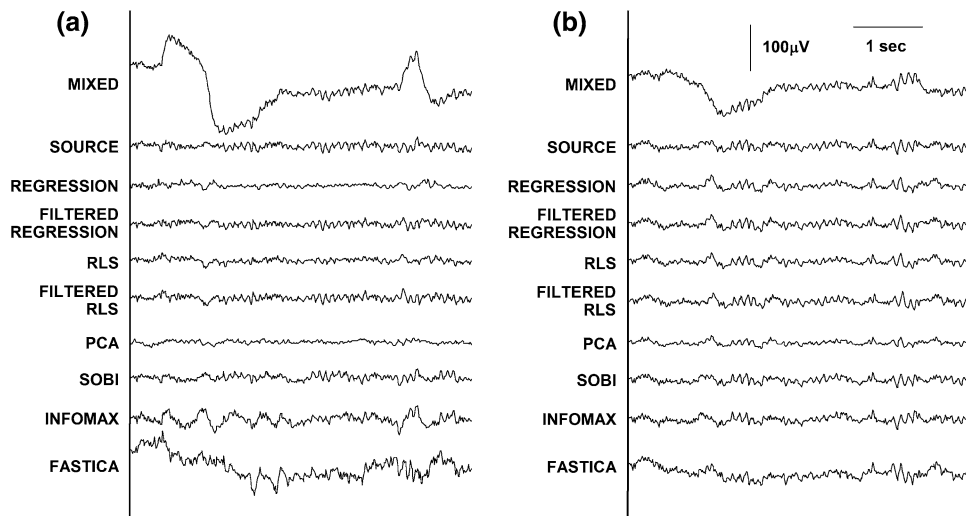


FIGURE 5. Five second time courses corresponding to simulated and corrected EEG signals. Several ocular reduction methods were applied for obtaining corrected EEG signals. Original cerebral activity sources are also displayed. Only two channels are plotted as examples: (a) Fp1 and (b) C3.

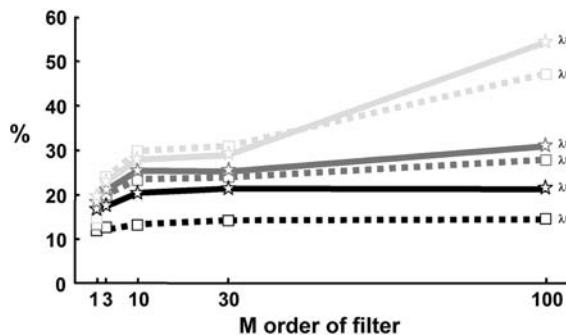


FIGURE 6. Mean average errors calculated for different values of the order of the filter M and of the forgetting factor λ for the adaptive filtering. Solid lines indicate classical RLS algorithm and dashed lines denote filtered version of RLS algorithm.

artifacts. However, corrected EEG signals that were obtained when reconstructing the signals without ocular components, showed modified cerebral activity in frontal sited locations. By visual inspection, SOBI algorithm and filtered versions of regression and adaptive RLS procedures produced more similar corrected EEG signals to the original cerebral sources than other methods in all EEG channels.

Performance of the adaptive filter depends on the order M and on the forgetting factor λ . Both parameters were evaluated by means of Fig. 6. This figure shows the mean percentage errors (average of 19 EEG channels, 12 subjects and for all 9 spectral target variables) calculated for different values of M and λ . Results showed that the lowest errors were obtained with $\lambda = 0.9999$. In addition, errors did not change significantly depending on M with this forgetting factor. By visual inspection of Fig. 6, the filter length of

$M = 1$ and a forgetting factor $\lambda = 0.9999$ were selected because they showed the lowest errors in spectral target variables.

For HOS-based BSS techniques, long data is generally recommended for applying the decomposition procedure.⁸ In this work, all BSS-based procedures were computed with different EEG segment durations from 5 to 180 s in order to evaluate their effect in filtering. Propagation factors in regression-based techniques were calculated considering the whole 3 min available because factors only depended on subject and electrode location at the scalp.¹ Relative errors for all nine spectral target variables were calculated between initial cerebral sources and corrected EEG signals. Results showed that while PCA worked better in shortest segments, errors obtained for the other SOS-based technique (SOBI) were similar in all EEG segment durations longer than 5 s and minimum at 15 s. Finally, in spite of the decreased error in HOS-based procedures by increasing the segment duration, the lowest errors were obtained for SOBI algorithm in any case (around 5%). For further analysis, the epoch duration that provided the lowest error was used for each BSS-based technique: 5 s for PCA, 15 s for SOBI and 180 s for HOS-based algorithms.

Table 1 shows Pearson's correlation values between sources and corrected EEG signals, that were averaged for all subjects and leads corresponding to three brain areas: anterior, central, posterior, and all the brain. Correlation values between sources and non-corrected EEG signals were also depicted. Similarity increased after applying any ocular reduction method and was higher for corrected EEG signals from leads towards posterior brain areas. The highest correlation values

TABLE 1. Pearson's correlation coefficient for EOG correction procedures (mean of all the simulated subjects).

Area	EOG correction algorithms								
	Non-corrected	Regression	Filtered regression	RLS	Filtered RLS	PCA	SOBI	INFOMAX	FastICA
Anterior	0.385	0.740	0.881	0.730	0.871	0.640	0.909	0.782	0.675
Central	0.667	0.968	0.980	0.968	0.982	0.930	0.986	0.954	0.899
Posterior	0.851	0.989	0.992	0.988	0.993	0.976	0.993	0.986	0.965
All EEG channels	0.631	0.891	0.950	0.888	0.945	0.840	0.961	0.902	0.841

TABLE 2. Statistical increases between Pearson's correlation coefficient for EOG correction procedures (paired *t*-tests were used between Method A vs. Method B).

Method A	Method B							
	Non-corrected	FastICA	PCA	RLS	Regression	INFOMAX	Filtered RLS	Filtered Regression
SOBI	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.009	<i>p</i> < 0.013	<i>p</i> < 0.016
Filtered regression	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.024	ns	
Filtered RLS	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.001	<i>p</i> < 0.040		
INFOMAX	<i>p</i> < 0.001	<i>p</i> < 0.009	<i>p</i> < 0.012	ns	ns			
Regression	<i>p</i> < 0.001	<i>p</i> < 0.005	<i>p</i> < 0.006	ns				
RLS	<i>p</i> < 0.001	<i>p</i> < 0.005	<i>p</i> < 0.006					
PCA	<i>p</i> < 0.001	ns						
FastICA	<i>p</i> < 0.001							

Bold depicted *p*-value < 0.01; Italicized *p*-value < 0.05, and *ns* non-significant increases (mean of all EEG channels).

were obtained after applying the SOBI algorithm (0.961 in average for all EEG channels; *p* < 0.016 with respect to the other ocular reduction techniques). Moreover, the highest increase with respect to correlation between EEG sources and non-corrected EEG signals was for the anterior region (from 0.385 to 0.909 after applying SOBI algorithm). These results indicated that filtered EEG signals were very similar to the corresponding sources. Statistical differences (paired *t*-tests) between mean correlation values obtained for each ocular reduction method were shown in Table 2. Results indicated that all techniques increased significantly Pearson's correlation coefficients compared to those obtained from non-corrected EEG signals. No statistical differences were obtained between regression, adaptive filtering and INFOMAX techniques. However, correlation values obtained for filtered versions of regression and adaptive filtering were statistically higher than these three methods. Non-significant differences were obtained between both filtered approaches. Finally, correlation coefficients calculated after applying SOBI algorithm were statistically higher than any other ocular reduction methods.

Figure 7 shows as an example the signal values for both source and SOBI-corrected Fp1 channel in one subject when applying Bland–Altman analysis. Figure 7a indicated high similarity between signals (Pearson's correlation coefficient = 0.899; and slope of

regression line = 0.921). Figure 7b shows the Bland–Altman plot with the differences of source minus SOBI-corrected Fp1 channel. According to this analysis, bias (average of the differences) and precision (95% confidence ranges) were also calculated. Besides, regression was carried out in order to evaluate proportional biases.

Table 3 summarizes the following variables from the Bland–Altman analysis averaged for all subjects and leads corresponding to the anterior brain region: 95% confidence ranges, correlation coefficient and slope of regression line using both source and corrected EEG signals. The bias between source and corrected EEG data was almost zero (lower than 10^{-5}) for all methods so the constant bias was negligible for all methods. Regarding the proportional bias, a higher correlation coefficient and slope indicated a higher proportional bias: the higher values the higher differences. In this case, bias is proportional to the magnitude of the value due to high potentials of ocular contamination. The highest slopes in magnitude are obtained with the non-corrected EEG signals. Afterwards, regression, adaptive RLS, FastICA, and INFOMAX showed lower slopes and with not high correlation coefficients (lower than 0.492). Finally, slopes with values almost zero were obtained for the remainder ocular correction methods (filtered versions of regression and adaptive RLS, PCA and SOBI)

showing that proportional bias was negligible. At least 96% of the data points lied within the 95% confidence range for all methods indicating a high level of agreement. The highest range was obtained with non-corrected EEG signals (± 46.05) and it decreased strongly with all the ocular correction methods and the minimum range was found for SOBI algorithm (± 6.15).

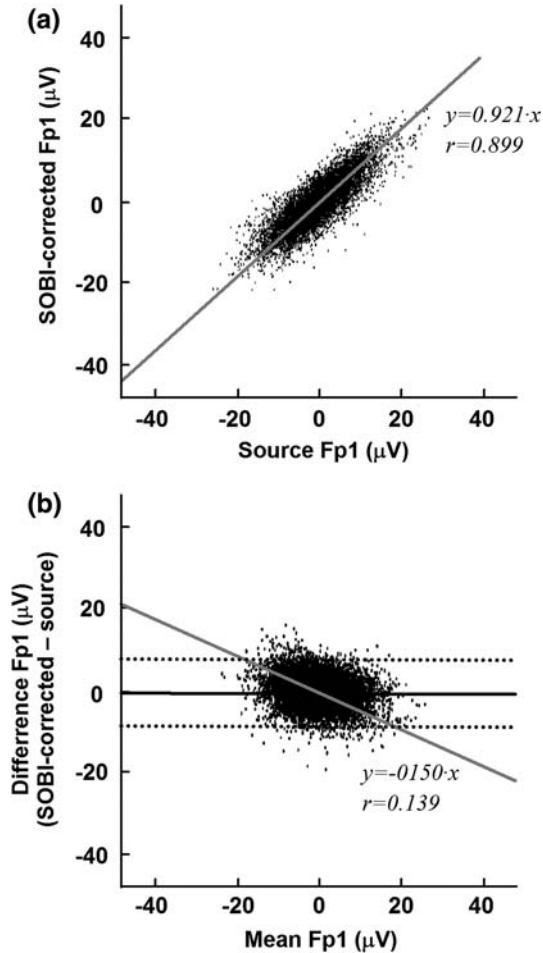


FIGURE 7. Comparison between source and SOBI-corrected Fp1 channels from one simulated subject: (a) Scattergram of source and corrected EEG signal. Gray line is the linear regression of the data; (b) Plot of the differences (Bland–Altman plot) between SOBI-corrected and source Fp1 channel compared with mean values. Black solid line is the average difference, and the dashed lines represent 95% confidence intervals. Gray line is the linear regression of the data.

Results for the other brain areas (central and posterior) were similar when comparing between methods: 95% confidence range, that were lower in general for central and specially posterior areas, decreased in all methods with respect to non-corrected EEG signals and their minimum was found for SOBI algorithm (± 2.419 and ± 1.643 in central and posterior leads, respectively).

Mean percentage errors (average of 19 EEG leads) for each spectral variable obtained with different ocular reduction methods are shown in Table 4. Errors for non-corrected EEG signals are also included and they showed that the highest impact of ocular contamination was in delta band (152 and 45% for absolute and relative delta power, respectively) due to the low frequency components of EOG interferences. All eye movement removal techniques reduced the errors obtained for non-corrected EEG signals in relative powers as well as in absolute delta power (p -values < 0.01 for absolute and relative delta powers). Errors in absolute powers for regression-based and RLS adaptive procedures were high (around 20%) because these methods did not take into account the bidirectional contamination between cerebral and ocular activities, and it also removed neural activity recorded in EOG channels. These errors were reduced with their filtered versions, especially in alpha and beta bands. Errors for HOS-based procedures were high for absolute powers in all bands because independent components extracted and then removed by HOS-based techniques were not only composed of ocular activity but also of some neural one. In average for all variables, the lowest error was obtained with SOBI algorithm (4.25%) and the following higher error was 11.33% for the filtered RLS adaptive approach (p -value < 0.001). This also happened in almost all target spectral variables (see Table 4).

Figure 8 shows the topographic distribution of errors for each ocular reduction technique. Simulated ocular artifacts affected mainly delta EEG band. Results for beta and alpha bands were very similar. Errors in absolute theta and especially in alpha and beta powers revealed that the incidence of ocular propagation in these bands was lower. These topographic results were coherent with the distribution of

TABLE 3. Parameters extracted from Bland–Altman analysis (mean of anterior EEG channels and mean of all the simulated subjects).

Parameter	EOG correction algorithms								
	Non-corrected	Regression	Filtered regression	RLS	Filtered RLS	PCA	SOBI	INFOMAX	FastICA
95% Confidence interval (\pm)	46.054	9.641	6.646	10.061	7.260	12.836	6.152	11.977	17.723
Correlation coefficient	0.759	0.492	0.300	0.407	0.222	0.246	0.132	0.269	0.392
Slope of regression line	1.198	-0.381	-0.136	-0.307	-0.096	-0.084	-0.069	0.221	0.431

TABLE 4. Percentage errors (%) in spectral variables for EOG correction procedures (mean of all EEG channels and mean of all the simulated subjects).

Spectral variables	EOG correction algorithms								
	Non-corrected	Regression	Filtered regression	RLS	Filtered RLS	PCA	SOBI	INFOMAX	FastICA
Total power	47.89	22.54*	11.17**	22.03*	10.56**	20.53*	4.32**	19.52*	43.04
ABS delta	151.78	19.86**	19.78**	18.57**	18.35**	18.83**	8.46**	29.89**	43.10**
ABS theta	3.78	20.39**	18.92**	20.32**	18.54**	22.41**	3.88	21.77**	53.97**
ABS alpha	0.98	24.91**	5.44**	24.77**	5.27**	22.54**	1.89*	13.46**	34.93**
ABS beta	0.87	22.06**	0.82**	22.19**	0.87	23.56**	4.39**	20.59**	67.97**
Mean ABS variables	41.06	21.95	11.23*	21.58	10.72*	21.57	4.59**	21.05	48.60
REL delta	45.24	9.93**	12.21**	10.29**	11.00**	11.47**	5.10**	9.07**	9.03**
REL theta	17.33	8.14*	10.69*	7.37**	10.84*	4.68**	3.17**	5.58**	8.71**
REL alpha	18.45	8.21**	9.47*	8.12**	8.98**	6.17**	3.47**	5.87**	7.75**
REL beta	18.50	14.41	18.64	12.59	17.52	6.73**	3.59**	6.47**	12.76
Mean REL variables	24.88	10.17**	12.75**	9.59**	12.09**	7.26**	3.83**	6.75**	9.56**
Mean all variables	33.87	16.72*	11.90**	16.25*	11.33**	15.21*	4.25**	14.69*	31.25

Statistical differences between errors for ocular correction methods and non-corrected signals were shown (* $p < 0.05$; ** $p < 0.01$).

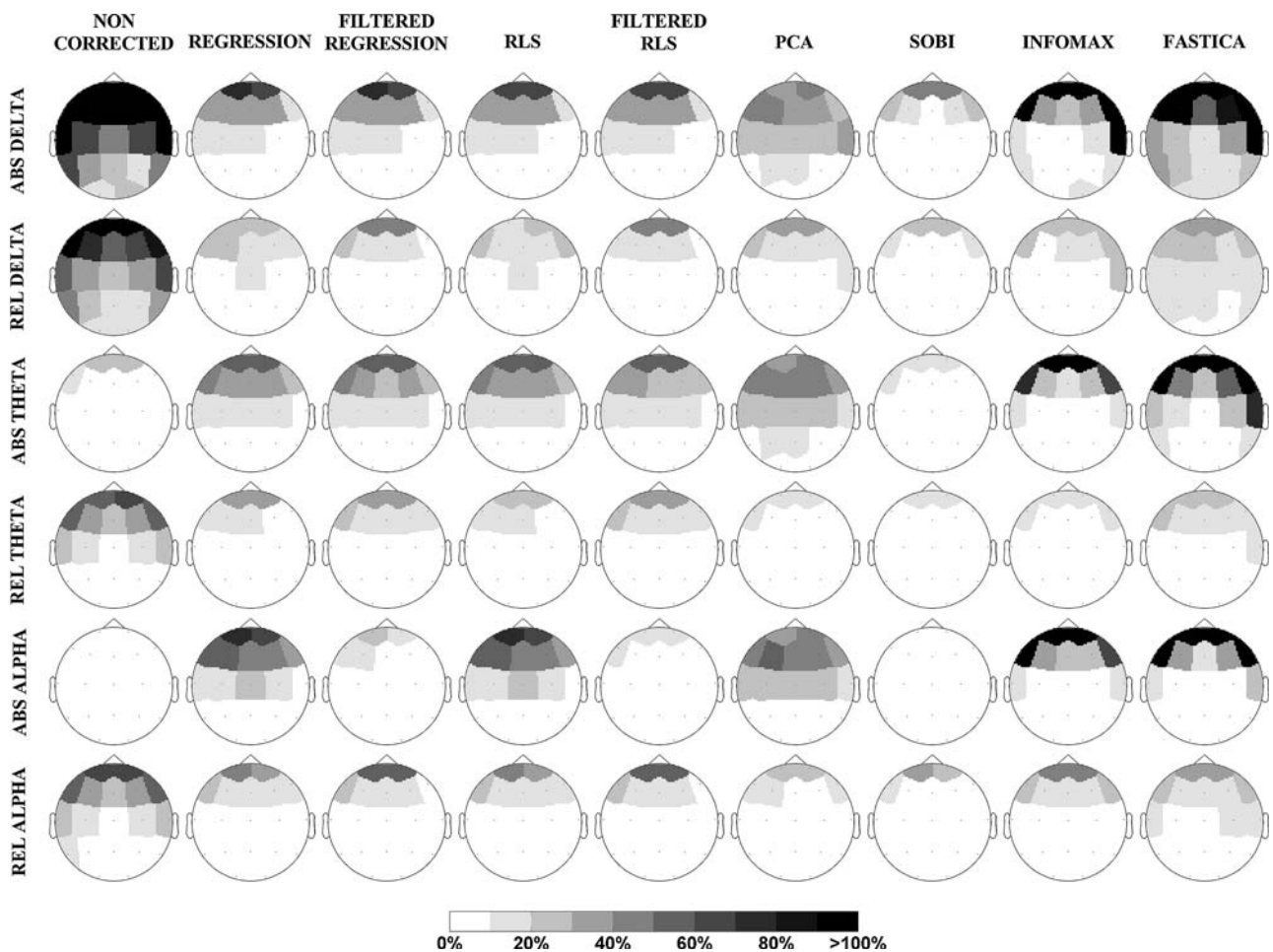


FIGURE 8. Topographic maps of the errors (%) between the cerebral sources and the corrected EEG signals by using different ocular reduction techniques. Errors between sources and non-corrected (mixed) signals are included. Color key is discretized from white to black in 10% stripes.

eye activity contamination across the scalp (see Fig. 2); errors were located at anterior sites and especially in frontopolar and lateral-frontal regions.

SOBI algorithm provided the lowest errors between the original cerebral sources and the corrected EEG signals.

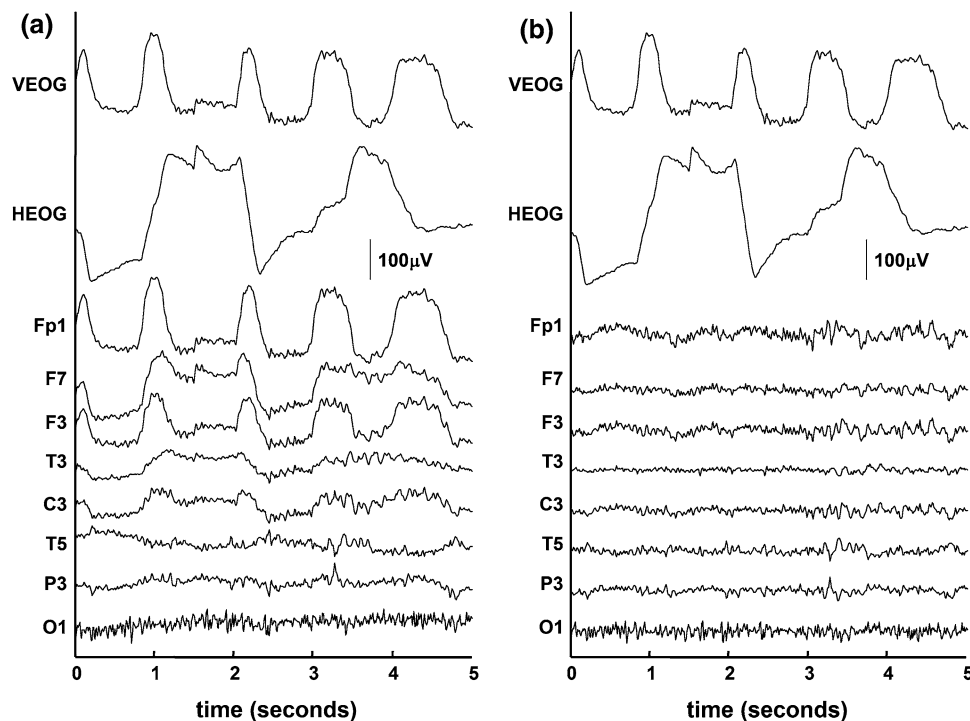


FIGURE 9. (a) Five second epoch of raw EOG and EEG signals containing prominent ocular artifacts. Some EEG channels corresponding to left hemisphere are shown as an example; (b) Raw EOG and corrected EEG signals obtained after applying automatic SOBI-based ocular removal procedure.

Real Data

Figure 9 shows as an example a 5-s epoch corresponding to real EOG and EEG signals displayed before and after applying SOBI-based ocular correction procedure. Propagation of ocular contamination to the different EEG leads and the bidirectional contamination between ocular and cerebral activities can be observed in the raw data (Fig. 9a). SOBI-based decomposition algorithm was performed using epoch durations of 15 s. The effects of ocular correction procedure on different EEG leads are shown in Fig. 9b. By visual inspection, this example demonstrates that SOBI algorithm removed efficiently the ocular artifacts from spontaneous real EEG data.

DISCUSSION AND CONCLUSION

Ocular artifacts in EEG data are a significant trouble in the diagnosis of dysfunctional neural states and in the evaluation of drug effects on the brain. Most used techniques to reduce them from spontaneous EEG signals are based on regression analysis. Adaptive filtering is another approach also proposed for this purpose. Additionally, BSS procedures have been recently applied in order to consider the mutual contamination between cerebral and ocular activities. In this study for

BSS-based methods, an automated procedure using logical rules, which were based on spectral and topographical information, was used in order to identify those components related to eye activity. Quantitative comparison of performance of these different ocular reduction methods required simulated signals where ocular and cerebral activities must be known *a priori*.

In this study, simulated EEG and EOG data were obtained by means of real data recorded according to the common procedure used in clinical routine. For this purpose, linear MISO models corresponding to ocular and cerebral activity contaminations were identified by using an autoregressive structure with exogenous inputs. Although the same model orders were used to cerebral and ocular contaminations through the same medium which is the head, higher and different orders in both contamination could have been selected. However, the identification would not improve sensitively because the FPE was not much smaller (see Fig. 1) and it is known that an unnecessary complex model could produce artifacts. It does not imply that ocular and cerebral propagations are frequency dependent, a topic that has been questioned. Bioelectromagnetism considers that volume conduction is instantaneous for frequencies lower than 1000 Hz; in other words, frequency dependence of ocular propagation is negligible.²³ It is yet to be

definitively demonstrated whether ocular propagation is frequency dependent or independent. In fact, any volume conductor has both resistive and capacitive properties. Nevertheless, the detailed capacitive properties of neocortex at the very low frequency components of most interest in EEG and EOG (0.5–20 Hz) have not been widely studied, but the available evidence suggests that capacitive effects have minimal influence on EEG volume conduction.²⁴ On the other hand, other studies have taken the assumption that ocular propagation exhibits frequency-dependent behavior and apply techniques for artifact cancellation based on filtered EOG signals.^{10,13,29,33}

In this work, for ocular contamination, magnitude frequency responses seemed to correspond to a low pass filtering procedure (see Fig. 2). However, to analyze them in more detail permitted to observe that it was not exactly thus. In this way, 3-dB bandwidths (frequency at which the gain drops 3 dB) of these FIR filters were calculated: 1.71 ± 0.76 Hz for vertical and 3.11 ± 1.94 Hz for horizontal EOG in all channels. These bandwidths corresponded to the 96.67 ± 0.01 and $97.76 \pm 0.01\%$ of the total energy of the vertical and horizontal EOG sources, respectively, for all subjects and channels. Thus, almost whole ocular propagation could be considered practically frequency-independent which was consistent to what would be expected from biophysics knowledge. Regarding cerebral activity contamination, the magnitude frequency responses from pre-frontal and lateral frontal sites to EOG locations (see Fig. 3) corresponded to quasi-all-pass filters in the EEG frequency range and they could be approximated by constant propagation coefficients. Their filter gains showed a higher contamination in the vertical ocular channel than in the horizontal one.

Other studies that have been previously published compared ocular reduction methods in simulated EEG signals. However, the following approaches had not been compared so far: linear regression, adaptive filtering and BSS. Besides, filtered versions of the linear regression, which was already applied in Wallstrom *et al.*,³² as well as of the adaptive filtering, which was suggested but not applied in He *et al.*,¹³ were included in this study. Additionally, performance of these methods was quantitatively evaluated by similarity between sources and corrected signals and by errors in the spectral target variables used commonly in clinical EEG studies.

Time and frequency results indicated that all methods reduced significantly the ocular contamination from EEG data. However, corrected EEG signals from frontal regions showed that a partial neural component was also subtracted by regression, adaptive RLS- and PCA-based methods, possibly due to the mutual contamination between EEG and EOG signals (see Fig. 5).

This cerebral information removal was partly reduced in the alpha and beta bands for regression and RLS techniques by applying a low-pass filter to EOG recordings before subtraction. Other BSS procedures based on SOS or HOS found more appropriate, and not necessarily orthogonal like PCA, source components. Assumptions about source modeling are different in SOS- and HOS-based algorithms: while SOS methods extract uncorrelated sources, HOS-based techniques use statistic information (like higher-order cumulants, entropy or maximum likelihood estimation) in order to obtain sources by minimizing an approximation of mutual information. Ocular sources extracted by HOS-based algorithms contained generally some cerebral activity, especially in anteriorly placed electrodes while it did not happen with SOBI algorithm. Errors between spectral variables for initial cerebral sources and corrected EEG signals calculated with HOS-based algorithms decreased when increasing data length, but they were far from the low errors obtained with some SOS-based methods (see Table 4: 31.25% for FastICA (n.s., when compared to non-corrected errors) and 14.69% for INFOMAX ($p < 0.015$) compared to 4.25% for SOBI ($p < 0.001$), in average). Similar errors and correlation values were obtained for INFOMAX, regression and adaptive filtering methods. Topographic distributions showed that errors were located at anterior sites and especially in frontopolar and lateral-frontal regions. Additionally, filtered version of linear regression and adaptive filtering provided good results in average: 0.950 and 0.945 for Pearson's correlation coefficients, and 11.90 and 11.33% for errors in spectral variables, respectively. However, correlation coefficients were much higher and errors much lower for SOBI algorithm (see Tables 1 and 4: 0.961 and 4.25% in average, respectively). Moreover, agreement level by 95% confidence ranges using Bland–Altman analysis was narrower for SOBI algorithm (see Table 3). This algorithm did not show either constant or proportional bias. Consequently, this method showed much higher similarity between EEG sources and corrected EEG signals.

BSS-based approaches are based in some *a-priori* hypothesis: sources must be uncorrelated for SOS-based algorithms, and statistically independent for HOS-based ones. In the present paper, BSS techniques were used to separate ocular from cerebral activities in order to eliminate the former. The fact that ocular and cerebral activities in a simulated subject came from different real subjects ensured the statistically independence between these activities. Additionally, in all cases, normalized correlations between each ocular (two) and each cerebral (19) sources were calculated. Results showed that correlation values were always lower than 0.023 indicating the fulfillment of the *a-priori* hypothesis of uncorrelatedness for SOS-based

methods. In the case of SOBI, *a-posteriori* correlations between ocular activities and corrected-EEG signals were calculated for each subject. Results showed values lower than 0.062 indicating a poor level of correlation. Analogously, statistical independence was assessed by calculating the normalized cross mutual information¹⁴ between ocular and cerebral sources. Cross mutual information quantifies the amount of information gained about one signal from the measurement of a second one. Normalized mutual information is zero when signals are independent, while it has a maximum value of one if both signals are identical. Results showed values between ocular and cerebral sources lower than 0.004 indicating that they were statistically independent. Moreover, mutual information was also calculated between ocular activities and INFOMAX-based corrected EEG signals obtaining values always lower than 0.030.

Besides, these results can be considered another test of validation of the ocular reduction effectiveness: very low values of correlation and mutual information between ocular sources and corrected EEG signals ensured that there were no relation between them and hence ocular contamination was efficiently removed from EEG data.

Thus, time and frequency results indicated that SOBI algorithm reduced better the ocular artifacts preserving by far more cerebral activity (see Figs. 5 and 8). This algorithm, that is simple, fast to compute and robust, presented the lowest errors in all leads using only short data segments of 15 s. Because HOS-based methods needed longer durations, sources were less stationary which is one of the assumptions of BSS. Besides, SOBI algorithm used the temporal structure provided by EOG and EEG data to achieve the separation. In this case, the use of additional well-defined second-order statistics (as covariances at different time lags) improved the estimation of the BSS model and the separation of ocular and cerebral sources.

Finally, based on artificially generated both spontaneous and corrupted EOG and EEG recordings, we concluded that SOBI was the most effective and efficient techniques for eye movement reduction with respect to similarity between sources and corrected EEG signals and to spectral target variables that are frequently used in a clinical real situation, even when available data length was short.

ACKNOWLEDGMENT

This study was partially supported by CICYT (TEC2008-02754/TEC del *Ministerio de Ciencia e Innovación*) from Spain.

REFERENCES

- ¹Anderer, P., H. V. Semlitsch, B. Saletu, and M. J. Barbanoj. Artifact processing in topographic of electro-encephalographic activity in neuropsychopharmacology. *Psychiat. Res.* 45:79–93, 1992. doi:10.1016/0925-4927(92)90002-L.
- ²Barbati, G., C. Porcaro, F. Zappasodi, P. M. Rossini, and F. Tecchio. Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. *Clin. Neurophysiol.* 115:1220–1232, 2004. doi:10.1016/j.clinph.2003.12.015.
- ³Bell, A. J., and T. J. Sejnowski. An information maximization approach to blind separation and blind deconvolution. *Neural. Comput.* 7:1129–1159, 1995. doi:10.1162/neco.1995.7.6.1129.
- ⁴Belouchrani, A., K. Abed-Meraim, J. F. Cardoso, and E. Moulines. A blind source separation technique using second-order statistics. *IEEE Trans. Signal Process.* 45:434–444, 1997. doi:10.1109/78.554307.
- ⁵Bland, J. M., and D. G. Altman. Measurement error and correlation coefficients. *Br. Med. J.* 313:41–42, 1996.
- ⁶Cichocki, A., S. Amari, K. Siwek, T. Tanaka, *et al.* ICA-LAB Toolboxes for Signal and Image Processing. Available from <http://www.bsp.brain.riken.jp/ICALAB/>. Accessed 10 March 2008.
- ⁷Croft, R. J., and R. J. Barry. Removal of ocular artifact from the EEG: a review. *Neurophysiol. Clin.* 30:5–19, 2000. doi:10.1016/S0987-7053(00)00055-1.
- ⁸Delorme, A., and S. Makeig. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134:9–21, 2004. doi:10.1016/j.jneumeth.2003.10.009.
- ⁹Delorme, A., T. Sejnowski, and S. Makeig. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *Neuroimage* 34:1443–1449, 2007. doi:10.1016/j.neuroimage.2006.11.004.
- ¹⁰Gasser, T., L. Sroka, and J. Möcks. The transfer of EOG activity into the EEG for eyes open and closed. *Electroencephalogr. Clin. Neurophysiol.* 61:181–193, 1985. doi:10.1016/0013-4694(85)91058-2.
- ¹¹Gasser, T., P. Ziegler, and F. Gattaz. The deleterious effect of ocular artifacts on the quantitative EEG, and a remedy. *Eur. Arch. Psy. Clin. N.* 241:241–252, 1992. doi:10.1007/BF02191960.
- ¹²He, P., G. Wilson, and C. Russell. Removal of ocular artifacts from electro-encephalography by adaptive filtering. *Med. Biol. Eng. Comp.* 42:407–412, 2004. doi:10.1007/BF02344717.
- ¹³He, P., G. Wilson, C. Russell, and M. Gerschütz. Removal of ocular artifacts from the EEG: a comparison between time-domain regression and adaptive filtering method using simulated data. *Med. Biol. Eng. Comp.* 45:495–503, 2007. doi:10.1007/s11517-007-0179-9.
- ¹⁴Hoyer, D., B. Pompe, K. H. Chon, H. Hardraht, C. Wicher, and U. Zwiener. Mutual information function assesses autonomic information flow of heart rate dynamics at different time scales. *IEEE Trans. Bio-Med. Eng.* 52:584–592, 2005.
- ¹⁵Hyvärinen, A., and E. Oja. A fast fixed-point algorithm for independent component analysis. *Neural Comput.* 9:1483–1492, 1997. doi:10.1162/neco.1997.9.7.1483.
- ¹⁶Hyvärinen, A., J. Karhunen, and E. Oja. Independent Component Analysis. New York: John Wiley & Sons, p. 481, 2001.

- ¹⁷Jackson, L. B. FIR filter design techniques. In: *Digital Filters and Signal Processing*, 3rd ed. Boston: Kluwer Academic Publishers, 1995, pp. 301–307.
- ¹⁸Jung, T.-P., S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin. Neurophysiol.* 111:1745–1758, 2000.
- ¹⁹Kierkels, J. J. M., G. J. M. van Boxtel, and L. L. M. Vogten. A model-based objective evaluation of eye movement correction in EEG recordings. *IEEE Trans. Bio-Med. Eng.* 53:246–253, 2006.
- ²⁰Lins, O. G., T. W. Picton, P. Berg, and M. Scherg. Ocular artifacts in recording EEGs and event-related potentials II: source dipoles and source components. *Brain Topogr.* 6:65–78, 1993. doi:[10.1007/BF01234128](https://doi.org/10.1007/BF01234128).
- ²¹Ljung, L. *System Identification—Theory for the User*. 2nd ed. Upper Saddle River, NJ: PTR Prentice Hall, p. 609, 1999.
- ²²Makeig, S., A. J. Bell, T.-P. Jung, and T. J. Sejnowski. Independent component analysis of electro-encephalographic data. *Adv. Neural Inf. Process. Syst.* 8:145–151, 1996.
- ²³Malmivuo, J., and R. Plonsey. Volume source and volume conductor. In: *Bioelectromagnetism. Principles and Applications of Bioelectric and Biomagnetic Fields*. New York: Oxford University Press, 1995, pp. 133–147.
- ²⁴Nunez, P. L., and R. Srinivasan. *Electric Fields of the Brain*. 2nd ed. New York: Oxford University Press, 2006.
- ²⁵Romero, S., M. A. Mañanas, and M. J. Barbanj. A comparative study of automatic techniques for ocular artifact reduction in spontaneous EEG signals based on clinical target variables: a simulation case. *Comput. Biol. Med.* 38:348–360, 2008. doi:[10.1016/j.compbiomed.2007.12.001](https://doi.org/10.1016/j.compbiomed.2007.12.001).
- ²⁶Saletu, B., P. Anderer, K. Kinsperger, and J. Grünberger. Topographic brain mapping of EEG in neuropsychopharmacology—Part II. Clinical applications (pharmacological EEG mapping). *Meth. Find. Exp. Clin. Pharmacol.* 9:385–408, 1987.
- ²⁷Schlögl, A., C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller. A fully automated correction method of EOG artifacts in EEG recordings. *Clin. Neurophysiol.* 118:98–104, 2007. doi:[10.1016/j.clinph.2006.09.003](https://doi.org/10.1016/j.clinph.2006.09.003).
- ²⁸Semlitsch, H. W., P. Anderer, P. Schuster, and O. Presslich. A solution for reliable and valid reduction of ocular artifacts applied to the P300 ERP. *Psychophysiology* 23:695–703, 1986. doi:[10.1111/j.1469-8986.1986.tb00696.x](https://doi.org/10.1111/j.1469-8986.1986.tb00696.x).
- ²⁹Sörnmo, L., and P. Laguna. EEG signal processing. In: *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Elsevier Academic Press, 2005, pp. 55–180.
- ³⁰Vasegui, S. V. Adaptive filters. In: *Advanced Digital Signal Processing and Noise Reduction*, 2nd ed. Chichester: John Wiley & Sons, 2000.
- ³¹Vigario, R. N. Extraction of ocular artifacts from EEG using independent component analysis. *Electroencephalogr. Clin. Neurophysiol.* 103:395–404, 1997. doi:[10.1016/S0013-4694\(97\)00042-8](https://doi.org/10.1016/S0013-4694(97)00042-8).
- ³²Wallstrom, G. L., R. E. Kass, A. Miller, J. F. Cohn, and A. F. Nathan. Automatic correction of ocular artifacts in the EEG: a comparison of regression-based and component-based methods. *Int. J. Psychophysiol.* 53:105–119, 2004. doi:[10.1016/j.ijpsycho.2004.03.007](https://doi.org/10.1016/j.ijpsycho.2004.03.007).
- ³³Woestenburg, J. C., M. N. Verbaten, and J. L. Slanger. The removal of eye-movement artifact from the EEG by regression analysis in the frequency domain. *Biol. Psychol.* 16:127–147, 1983. doi:[10.1016/0301-0511\(83\)90059-5](https://doi.org/10.1016/0301-0511(83)90059-5).