



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)

## Data Article

# A semi-simulated EEG/EOG dataset for the comparison of EOG artifact rejection techniques

Manousos A. Klados<sup>a,b,\*</sup>, Panagiotis D. Bamidis<sup>b</sup><sup>a</sup> Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig 04103, Germany<sup>b</sup> Lab of Medical Physics, School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

## ARTICLE INFO

*Article history:*

Received 22 January 2016

Received in revised form

12 June 2016

Accepted 21 June 2016

Available online 29 June 2016

*Keywords:*

EEG

EOG

Artifact Rejection

## ABSTRACT

Artifact rejection techniques are used to recover the brain signals underlying artifactual electroencephalographic (EEG) segments. Although over the last few years many different artifact rejection techniques have been proposed ([\[1\]](http://dx.doi.org/10.1109/JSEN.2011.2115236), [\[2\]](http://dx.doi.org/10.1016/j.clinph.2006.09.003), [\[3\]](http://dx.doi.org/10.3390/e16126553)), none has been established as a gold standard so far, because assessing their performance is difficult and subjective ([\[4\]](http://dx.doi.org/10.1109/ITAB.2009.5394295), [\[5\]](http://dx.doi.org/10.1016/j.bspc.2011.02.001), [\[6\]](http://dx.doi.org/10.1007/978-3-540-89208-3_300)). This limitation is mainly based on the fact that the underlying artifact-free brain signal is unknown, so there is no objective way to measure how close the retrieved signal is to the real one. This article solves the aforementioned problem by presenting a semi-simulated EEG dataset, where artifact-free EEG signals are manually contaminated with ocular artifacts, using a realistic head model. The significant part of this dataset is that it contains the pre-contamination EEG signals, so the brain signals underlying the EOG artifacts are known and thus the performance of every artifact rejection technique can be objectively assessed.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

\* Corresponding author at: Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig 04103, Germany.  
E-mail address: [klados@cbs.mpg.de](mailto:klados@cbs.mpg.de) (M.A. Klados).

## Specifications Table

Subject area	<i>Neuroscience</i>
More specific subject area	<i>EEG Artifact Rejection</i>
Type of data	<i>EEG and EOG signals in MATLAB arrays, links</i>
How data was acquired	<i>Electroencephalography, Electrooculography</i>
Data format	<i>Filtered</i>
Experimental factors	<i>Filtering, and artificially contaminated with EOG signals</i>
Experimental features	<i>This dataset can be used to assess the performance of an artifact rejection technique.</i>
Data source location	<i>Thessaloniki, Greece</i>
Data accessibility	<i>Data is within this article and is downloadable without restriction using the following URL: <a href="https://data.mendeley.com/datasets/wb6yvr725d/1">https://data.mendeley.com/datasets/wb6yvr725d/1</a></i>

## Value of the data

- First EEG dataset artificially contaminated with EOG artifacts using a realistic contamination model.
- Subjective assessment of EOG artifact rejection technique.
- Serves as a reference point for the scientific community in order to compare various algorithms in the same dataset.

## 1. Data

This work presents a semi-simulated EEG dataset, where artifact-free EEG signals are manually contaminated with ocular artifacts following the model proposed by [7]. The significant part of this dataset is that it contains the pre-contamination EEG signals, so the brain signals underlying the EOG artifacts are known and thus the performance of every artifact rejection technique [1–3] can be objectively assessed [4–6]. The main differences of the proposed dataset compared to others (p.e. see [8,9]) is that it is focused only on EOG artifacts, using a realistic model for the contamination of artifact-free EEGs and not a random procedure.

The data are available for downloading without any restriction using this URL: <https://data.mendeley.com/datasets/wb6yvr725d/1>.

## 2. Experimental design, materials and methods

EEG data were obtained from twenty-seven healthy subjects, 14 males (mean age:  $28.2 \pm 7.5$ ) and 13 females (mean age:  $27.1 \pm 5.2$ ), during an eyes-closed session. Nineteen EEG electrodes (FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, Pz) were placed according to the 10–20 International System, with odd indices referenced to the left and even indices to the right mastoid respectively, while the central electrodes (Fz, Cz, Pz) were referenced to the half of the sum of the left and right mastoids. Signals' sampling frequency was 200 Hz and a band pass filtered at 0.5–40 Hz and notch filtered at 50 Hz were applied. From these twenty-seven subjects we have obtained fifty-four datasets in total and each one has 30 s duration. The obtained datasets were carefully inspected in order to ensure that there is no significant contamination by biological or external artifacts.

Moreover, EOG signals were obtained from the same subjects, during an eyes-opened condition, using four electrodes placed above and below of the left eye and another two on the outer canthi of each eye. This process gave rise to two bipolar signals, namely, vertical-EOG (VEOG), which is equal to the upper minus lower EOG electrode recordings and horizontal-EOG (HEOG), which is equal to the left minus right EOG electrode recordings. These EOG signals were band-pass filtered at 0.5–5 Hz [10].

In order to produce this semi-simulated EEG dataset, we used the contamination model proposed by [7] and it follows the next equation:

$$\text{Contaminated\_EEG}_{ij} = \text{Pure\_EEG}_{ij} + a_j \text{VEOG} + b_j \text{HEOG}$$

where *Contaminated\_EEG* are the artificially contaminated EEG signals and *Pure\_EEG* are the signals obtained during the eyes-closed session. The *VEOG* and *HEOG* are the EOG signals that were previously described, while vectors  $a_j$ ,  $b_j$  describe the contamination coefficients for VEOG and HEOG, respectively, initialized according to [10]. Finally the index  $i$  indicates the subject's number, while  $j$  denotes the electrode's number.

The contamination coefficients  $a_j$ ,  $b_j$  were computed for each subject separately. For each subject, EEGs and EOGs were taken from an eyes-opened session. The VEOG signals were used to detect all the blink segments (start – peak – end) and, after confirming that in the same segments the EEGs are also corrupted by blinks, we used linear regression among their amplitudes to compute the  $a_j$ . The same procedure was also applied in HEOGs for the  $b_j$  computation, with the only difference that the input to linear regression was the amplitude of the horizontal plateaus generated by the horizontal eye movements.

## Transparency document. Supplementary material

Transparency data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2016.06.032>.

## References

- [1] N. Mammone, F. La Foresta, F.C. Morabito, Automatic artifact rejection from multichannel scalp EEG by wavelet ICA, *IEEE Sens. J.* 12 (2012) 533–542, <http://dx.doi.org/10.1109/JSEN.2011.2115236>.
- [2] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, G. Pfurtscheller, A fully automated correction method of EOG artifacts in EEG recordings, *Clin. Neurophysiol.* 118 (2007) 98–104, <http://dx.doi.org/10.1016/j.clinph.2006.09.003>.
- [3] N. Mammone, F. Morabito, Enhanced automatic wavelet independent component analysis for electroencephalographic artifact removal, *Entropy* 16 (2014) 6553–6572, <http://dx.doi.org/10.3390/e16126553>.
- [4] M.A. Klados, C.L. Papadelis, P.D. Bamidis, REG-ICA: a new hybrid method for EOG Artifact Rejection, in: Proceedings of 2009 9th International Conference on Information Technology and Applications in Biomedicine, IEEE, 1–4, <http://dx.doi.org/10.1109/ITAB.2009.5394295>, 2009.
- [5] M.A. Klados, C. Papadelis, C. Braun, P.D. Bamidis, REG-ICA: a hybrid methodology combining Blind Source Separation and regression techniques for the rejection of ocular artifacts, *Biomed. Signal Process. Control.* 6 (2011) 291–300, <http://dx.doi.org/10.1016/j.bspc.2011.02.001>.
- [6] M.A. Klados, C. Papadelis, C.D. Lithari, P.D. Bamidis, The removal of ocular artifacts from EEG signals: a comparison of performances for different methods, in: Proceedings of the 4th European Conference of the International Federation for Medical and Biological Engineering, 1259–1263, [http://dx.doi.org/10.1007/978-3-540-89208-3\\_300](http://dx.doi.org/10.1007/978-3-540-89208-3_300), 2009.
- [7] T. Elbert, W. Lutzenberger, B. Rockstroh, N. Birbaumer, Removal of ocular artifacts from the EEG—a biophysical approach to the EOG, *Electroencephalogr. Clin. Neurophysiol.* 60 (1985) 455–463 (accessed 10.04.13), <http://www.ncbi.nlm.nih.gov/pubmed/2580697>.
- [8] X. Yong, M. Fatourechi, R.K. Ward, G.E. Birch, Automatic artefact removal in a self-paced hybrid brain-computer interface system, *J. Neuroeng. Rehabil.* 9 (2012) 50, <http://dx.doi.org/10.1186/1743-0003-9-50>.
- [9] A.K. Abdullah, C.Z. Zhang, A.A.A. Abdullah, S. Lian, Automatic extraction system for common artifacts in EEG signals based on evolutionary stone's BSS algorithm, *Math. Probl. Eng.* 2014 (2014) 1–25, <http://dx.doi.org/10.1155/2014/324750>.
- [10] O. Lins, T. Picton, P. Berg, M. Scherg, Ocular artifacts in EEG and event-related potentials I: scalp topography, *Brain Topogr.* 6 (1993) 51–63, <http://dx.doi.org/10.1007/BF01234127>.