Machine learning validation of EEG+tACS artefact removal

DOI:

[10.1088/1741-2552/ab58a3](https://doi.org/10.1088/1741-2552/ab58a3)

Document Version

Accepted author manuscript

[Link to publication record in Manchester Research Explorer](#)

Citation for published version (APA):

Kohli, S., & Casson, A. J. (2019). Machine learning validation of EEG+tACS artefact removal. *Journal of Neural Engineering*. <https://doi.org/10.1088/1741-2552/ab58a3>

Published in:

Journal of Neural Engineering

Citing this paper

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.

Machine learning validation of EEG+tACS artefact removal

Siddharth Kohli¹ and Alexander J. Casson^{1,*}

¹School of Electrical and Electronic Engineering, The University of Manchester, Manchester, M13 9PL, UK.

*alex.casson@manchester.ac.uk

ABSTRACT

Motivation: Electroencephalography (EEG) recorded during Transcranial Alternating Current Simulation (tACS) is highly desirable in order to investigate brain dynamics during stimulation, but is corrupted by large amplitude stimulation artefacts. Artefact removal algorithms have been presented previously, but with substantial debates on their performance, utility, and the presence of any residual artefacts. This paper investigates whether machine learning can be used to validate artefact removal algorithms. The postulation is that residual artefacts in the EEG after cleaning would be independent of the experiment performed, making it impossible to differentiate between different parts of an EEG+tACS experiment, or between different behavioural tasks performed. **Methods:** Ten participants undertook two tasks (nBack and backwards digital recall) during simultaneous EEG+tACS, exercising different aspects of working memory. Stimulation during no task and sham conditions were also performed. A previously reported tACS artefact removal algorithm from our group was used to clean the EEG and a Linear Discriminant Analysis was trained on the cleaned EEG to differentiate different parts of the experiment. **Results:** Baseline, baseline during tACS, working memory task without tACS, and working memory task with tACS data segments could be differentiated with accuracies ranging from 65–94%, far exceeding chance levels. EEG from the nBack and backwards digital recall tasks could be separated during stimulation, with an accuracy exceeding 72%. If residual tACS artefacts remained after the EEG cleaning these did not dominate the classification process. **Significance:** This helps in building confidence that true EEG information is present after artefact removal. Our methodology presents a new approach to validating tACS artefact removal approaches.

1 Introduction

Recent years have seen the development of many non-invasive brain stimulation technologies. In particular, transcranial Electrical Stimulation (tES) is a relatively new technique with many applications as a therapeutic and investigative tool [1]. It operates by injecting small amounts of current into the scalp via rubber electrodes that are enclosed in saline soaked sponges. Transcranial Alternating Current Simulation (tACS), where the applied stimulation has a sinusoidal waveform, is of particular interest as it can potentially be used in a closed loop manner, where the frequency and phase of the stimulation are guided based upon concurrently measured brain information [2], [3]. The effects of tES are highly variable between different subjects and studies [4], [5], and it is thought that allowing brain state dependent stimulation paradigms may be a key step towards reducing this variability and understanding its causes [6].

Electroencephalography (EEG) is a widely used brain monitoring technique, with a high temporal resolution which makes it highly suitable for use in closed loop applications. However, tACS guiding via EEG is difficult due to the stimulation artefacts that tACS produces in simultaneous EEG monitoring. These are illustrated in Fig. 1 where it can be seen that the high amplitude, non-linear, tACS artefacts obscure the low amplitude EEG data.

In recent years there have been several investigations into the removal of this artefact in the EEG by applying post-collection signal processing to the raw EEG data [7]–[17]. There have also been investigations into the properties of the tACS artefacts [18]–[22], for example how they spread across the head [19] and how they mix with other bio-signals such as ECG and breathing patterns [18]. A recent review of methodological aspects of EEG+tACS is given in [23]. However, to date these techniques have received only modest use in the research literature, with most studies restricted to only analysing EEG data before and after the tACS stimulation was applied [7]. There remains much debate as to *how good* tACS artefact cleaning methods are, need to be, and the significance of any *post-cleaning* residual artefacts in the EEG trace.

Proving that artefacts have been successfully removed is an intrinsically challenging problem as, in on-person tests, there is no *baseline* available against which the cleaned EEG can be directly compared. In our previous EEG+tACS artefact removal work [7] we introduced the use of a head phantom where pre-recorded EEG was played in to a phantom which then had tACS applied to it, giving a known baseline to compare against for artefact removal validation. This was combined with a number of on-person tests in order to build up the overall confidence that the cleaned EEG was not dominated by residual artefacts

Figure 1. EEG data recorded during tACS stimulation. (a) Stimulation begins at 810 s at which point all recorded data are dominated by a large approximately sinusoidal artefact. (b) Raw collected EEG+tACS data with no band limiting filters applied. Ongoing artefact is not a pure sinusoid at the simulation frequency, but has an approximately 100 μ V ripple present. Images reproduced from [7] under its CC BY 4.0 license.

Figure 2. Top: The nBack working memory task. Stimuli A–G were presented in a randomised sequence and the subject asked to identify whether the current stimuli was the same as the one from n iterations before. Bottom: Shapes used for the nBack task depicting sporting activities. (Images taken from <http://www.clker.com> under a creative commons CC0 public domain license.)

(although some residual artefacts, particularly for high stimulation frequencies, undoubtedly remained).

In this paper we investigate an alternative validation approach for checking EEG+tACS artefact removal, based upon machine learning. In our approach we split the data from an EEG+tACS data collection experiment into different classes: EEG when no tACS was used, and cleaned EEG collected during tACS with artefact removal applied; collected both while the subject was performing a working memory task and while the subject was sitting stationary with no task performed. The underlying assumption is that residual artefacts present after tACS artefact removal would be independent of the task being performed. If residual artefacts dominated the cleaned EEG traces then it would not be possible to use machine learning to discriminate between the different parts of the experiment. We use a previously reported artefact removal approach, Superposition of Moving Averages (SMA) from [7], [12] with a new experiment designed to investigate this validation approach. We believe that it can be used to add further confidence in the performance of an artefact removal technique.

2 Materials and Methods

2.1 Working memory tasks

Subjects were asked to perform two tasks, one after another, exercising similar but different aspects of working memory. Firstly a *Visual nBack* which tested visual working memory span and executive function. Subjects were presented with a sequence of visual stimuli and asked to recall if the current stimuli matched the n^{th} stimuli before the current one, as illustrated in Fig. 2. If the current stimuli matched the stimuli 2 iterations ago this would be a 2-back trial, 3 iterations ago a 3-back trial, and so on. The images used were diagrams of sporting activities, shown in Fig. 2, easily identifiable by participants. A keyboard press was used to indicate whether the stimuli matched or not, and the next image would be presented once a button was pressed.

Secondly, participants did a *Backward digit recall* testing verbal memory span and executive function. A sequence of single digit numbers was presented to the subject, who was then asked to recall them in reverse order. Numbers were displayed on a screen one at a time and each for 0.8 s, followed by a 0.2 s blank screen. The difficulty of the task was varied by changing the length of the number to recall. After all digits were displayed the subject was asked to type them in on a keyboard, in reverse

order of display. Subjects were asked to do as many repetitions of this as possible during the trial block (2 minutes).

Both these tasks utilise the central executive of the multi-modal working memory model but use different forms of working memory capacity. Both tasks were implemented as customised MATLAB scripts, synchronised with our EEG data collection.

2.2 Experimental procedures

2.2.1 Practice session

Before the experimental session with EEG and tACS, subjects were invited to a practice session during which they practised the two working memory tasks at different difficulty levels. The purpose was to familiarise subjects with the tasks and also to establish a baseline difficulty at which each individual's performance peaked (n for the nBack task and the number of digits for the digit recall task). This minimised the influence any learning effects that repetition may have in the experimental session. In the experimental session the difficulty was set for each task and participant as the n or number of digits where their performance was 50% in the practice session.

2.2.2 Experimental sessions

After the practice session, each subject attended 2 experimental sessions. Each session was approximately 1 hour long, including the time to set up and test the EEG and tACS equipment. The protocols for the two sessions were identical with the exception that tACS was applied in one session and the other session was a sham where no tACS was applied (but the tACS electrodes were still placed to ensure the participants were uninformed). Each experimental session was separated from the other by a period of at least 24 hours and a maximum of 1 week.

The experimental session was sub-divided into six *blocks* of EEG+tACS, with each block lasting for 4 minutes. We repeated each working memory task 3 times giving the total of 6 experimental blocks to be performed. All experiments were conducted between 2–5 pm on a weekday to minimise baseline EEG variances and subjects were allowed to take breaks in between blocks to prevent fatigue.

Within one block the parts of the experiment were further sub-divided into the following three stages:

1. *Pre*: A one minute recording of background EEG data before the working memory task was performed. No tACS was applied in this period.
2. *During*: A two minute period during which one of the working memory tasks was performed. EEG data was recorded, and tACS applied (or applied as sham).
3. *Post*: A one minute recording of background EEG data after the working memory task was performed. No tACS was applied in this period.

This division of the experiment is illustrated in Fig. 3. No breaks were present between the *Pre*, *During* and *Post* stages. To prevent fatigue, after each block there was a break of at least 5 minutes, after which the participant was asked when they were ready for the next block.

In-line with our previous tACS artefact removal work [7], deliberately short 2 minute stimulations were used with the intention of avoiding behavioural effects. This is intended to allow the artefact removal problem to be investigated without the confounding factor of potential long term changes in the brain state being present. Here we do not investigate the behavioural result of the impact of tACS on the task performance. Rather the aim is to investigate the artefact removal validation procedure.

All subjects were also asked to do a control session which consisted of a 5 minute tACS stimulation, where subjects were asked to relax, keep their eyes open and minimise movements. This was intended to mirror the *Pre* and *Post* sections of each block, but now with tACS applied. This provided a baseline for the stimulation condition where no working memory task was being performed.

2.3 Participants

For this proof of concept work there were a total of 10 subjects, five male and five female, aged 20–35 years. The study was deliberately not powered to investigate the behavioural result of the impact of tACS on the task performance, as this was not the study aim. Instead 10 subjects was selected as sufficient to investigate the artefact removal validation procedure, and to be in-line with other early stage EEG+tACS artefact removal works (review given in [7] and in the discussion Section here).

All procedures used in this study were reviewed and approved by the University of Manchester Research Ethics Committee, application number 14263. Procedures performed involving human participants were in accordance with the ethical standards of the institutional and informed consent was obtained from all individual participants included in the study.

Control recording
(5 minutes)

B = Baseline
T = Task
T+S = Task + Stimulation
B+S = Baseline + Stimulation

Figure 3. Illustration of the experiment set up and data partitioning, with an example order of blocks. Each block consisted of *Pre*, *During*, and *Post* sections. The four classes for the model 1 and model 2 machine learning was then formed as: The Baseline class (B) from data collected before and after each working memory task. The Task class (T) from when the subject was performing the task without tACS stimulation present (in the nBack task in model 1, and backwards digit recall task in model 2). The Task + Stimulation class (T+S) from EEG data recorded while the subject was performing the task in the stimulation experimental session, and the Baseline + Stimulation (B+S) data came from the control session where no task was performed. Each task was actually repeated 3 times giving six blocks in the stimulation and sham sessions.

2.4 EEG and tACS setup

An 8 channel EEG montage was used for this study with the ground and reference channel placed at Cz. The 8 electrode positions were determined by using the ten–twenty system: Fp1, F3, F4, C3, C4, P4, O1, O2. All electrode impedances were below 10 k Ω and the data was sampled at 500 Hz. EEG was recorded using an Enobio EEG recorder (Neuroelectronics, Spain). Our EEG+tACS artefact removal procedures are channel count independent, and using a low number of electrodes is important for allowing for quicker experimental set ups which may be critical for clinical applications and when working with subjects who are vulnerable [7].

tACS stimulation was delivered using the DC-Stimulator PLUS (Neuroconn, Germany). The tACS electrodes were placed at Fp2 and P3, selected to match montages that target working memory effects in previous tACS research [24], [25]. Stimulation was performed at 1 mA peak-to-peak at 5 Hz for the duration of the working memory tasks (2 minutes) with a ramp up and down time of 1 s each. Again, the choice of 1 mA and 5 Hz stimulation was to mimic previously used choices in experiments looking at the effects of parietal stimulation in working memory [24], [25].

2.5 Data analysis

2.5.1 Pre-processing

All recorded data was analyzed offline in MATLAB (Mathworks, USA). Prior to analysis the EEG data was filtered using 3rd order low pass and high pass Butterworth filters with cutoffs at 45 and 0.5 Hz respectively.

If tACS stimulation was applied, the tACS artefact was removed using the Superposition of Moving Averages (SMA) approach, described in and unmodified from [7]. This gave either raw EEG data (from baseline and sham periods) or *cleaned* raw EEG data (from tACS periods) for use with a machine learning classifier.

2.5.2 Feature matrix generation

For inputting to the machine learning classifier the EEG signal for each channel was split into non-overlapping windows of 5 s. Small window sizes were used to increase the number of epochs available for classification, with 5 s selected to ensure at least 10 cycles of the lowest frequency of interest (4 Hz theta) was present in the window. This duration also ensured that potential non-linear heart beat and breathing artefacts as discussed in [18] would be present in each analysis window.

The EEG data in each 5 s window was downsampled to 256 Hz and decomposed using the Discrete Wavelet Transform and *db1* wavelet basis function. This gave time-frequency estimates of the signal approximately matching the widely used cognitive frequency bands: theta (4–8 Hz), alpha (8–16 Hz), beta (16–32 Hz) and gamma (32–64 Hz). In each band the power and Shannon entropy were calculated, and these *features* used as the inputs to the machine learning. In total 64 features were present for each time window of data: 8 channels, with 2 measures from each of 4 frequency bands. These features were selected as bandpowers are widely used across many EEG applications, and entropy is a measure of order in the brain which would be affected by tACS. For example, some works (e.g. [26]) have suggested that closed loop tACS could be adjusted in terms of entropy slope rather than only amplitude and frequency as commonly considered.

2.5.3 Classification

We created four different machine learning models, fitted and tested with different variants of the feature data available, in order to examine the separability performance in different cases. All models used a Linear Discriminant Analysis (LDA) classifier based on Fisher's discriminant analysis [27], and the associated built-in MATLAB function. This classifier assumes that the data in each class is based on different Gaussian distributions with the same covariance and aims to express each dependant variable (the classes) as a linear combination of the features. It was extended for multiclass models by [28]. In all cases 65% of the data was used to train the classifier and the remaining 35% used for testing the performance of the classifier, with a permutation test (described below) used to cross-validate this performance.

The first model split the data into four classes:

1. Baseline: The *Pre* and *Post* sections of each block, where no task and no tACS was present.
2. Baseline + Stimulation: The control session, where no task was present, but tACS was.
3. nBack task: Data from during the nBack working memory task when no tACS was applied (sham).
4. nBack task + Stimulation: Data from during the nBack working memory task when tACS was applied.

This data split is illustrated in Fig. 3. The procedure was performed separately for each person, with a different LDA trained and tested for each subject.

The second model also split the data into four classes, identical to that above, but this time with the task data from the backwards digital recall task. These first two models are intended to demonstrate the possibility of separating data from during the task from the baseline cases, regardless of whether stimulation was present or not.

The third model split the data into six classes:

1. Baseline: The *Pre* and *Post* sections of each block, where no task and no tACS was present.
2. Baseline + Stimulation: The control session, where no task was present, but tACS was.
3. nBack task: Data from during the nBack working memory task when no tACS was applied (sham).
4. nBack task + Stimulation: Data from during the nBack working memory task when tACS was applied.
5. Backwards digit recall task: Data from during the backwards digit recall working memory task when no tACS was applied (sham).
6. Backwards digit recall task + Stimulation: Data from during the backwards digit recall working memory task when tACS was applied.

In this approach data from both tasks was used to build a single model such that the classifier could be used to differentiate between the two different measures of working memory span (nBack and backwards digit recall) during tACS stimulation.

The three models discussed so far were developed and applied individually to data from each subject. For each analysis, ten different subject dependent classifiers were made—one for each participant using their training and test sections of data. However, in practical use this would necessitate a training session each time a new subject was recruited in order to generate a new model implementation for that person. Our fourth model was a subject independent classifier to investigate the performance of data separation when the same classifier was used for all subjects. This subject independent classifier pooled the training data from all participants to train one classifier. This one classifier was then used on the test data from each participant. It had the benefit of using more data to develop the classifier, which may lead to better performance, at the cost of increased variability in the input data which may lead to poorer performance as the classifier is not optimised for each separate person.

To create the subject independent model, data from the feature matrices of all subjects was normalised by dividing each feature observation with the absolute maximum of said feature for each subject. These normalised features were then concatenated to create a single matrix of normalised features from all subjects. Subsequently, 50% of this data was selected at random to train a classifier which was then used to classify normalised data for all subjects separately. This cross-validation was repeated 1000 times.

2.5.4 Classification validity

In all of the models, since the data sample sizes are not equal (there are different numbers of epochs for each class), the chance level of randomly guessing the class is not one divided by the number of classes present. The actual chance level was calculated by permutation, generating 10,000 random classifier outputs, splitting these to match the proportion of each class of data present, and assessing the chance classification accuracy obtained. These chance levels of performance are reported alongside our results for the actual model classification.

In addition, an assumption for supervised learning using classifiers is that the data is Independent and Identically Distributed (IID) [29]. This indicates that the order of the data is not relevant and thus all epochs are from the same arbitrary distribution [30]. In practice this means that the order in which the data (each epoch) is sorted for input to the classifier should not effect the performance of the classification. This is not automatically satisfied for time series data such as the EEG which evolves over time, where samples collected close together in time may be more similar than those collected further apart in time [31].

To test against this assumption, our data split using 65% of the data to train the classifier and the remaining 35% to test the performance of the classifier, was repeated 1000 times using random selections of data windows to form the training and test data sets. This process meant that the feature data matrix was randomised, mixing the time order of the collected features. The result of this test are included in our results as the error bars present in the classification results graphs. Any change in the performance due to time ordering effects would be captured in these error bars.

3 Results

3.1 Four class classification accuracy

Fig. 4 summaries the fraction of epochs classified correctly by the first three models (4 class nBack, 4 class digital recall, and 6 class), and the chance performance level for each. In all cases, the accuracy of identifying epochs far exceeds that of a chance classifier. For both 4 class models (i.e. regardless of which working memory task was performed) the accuracy was over 80% showing that it was possible to differentiate between cleaned EEG collected from the different parts of the experiment using our LDA classifier.

This is shown in more detail in Fig. 5 which shows the per-class accuracies within these overall performance results. This is shown for both the nBack and backwards digital recall classification models. The recognition accuracy varies depending on the class, with the worst accuracy in recognising the baseline case, and the best in recognising epochs from the baseline+stimulation

Figure 4. Average performance (grey) of each classifier across all 10 subjects. Chance performance level (red). Error bars show the range of performances obtained from different classifiers trained with randomly selected training/test data in a 65/35% ratio for each run.

Figure 5. Fraction of epochs successfully classified using the 4 class model for nBack (red) and backwards digit recall (grey) tasks, averaged for all 10 subjects.

case. With these levels of accuracy it was clearly possible to separate cleaned EEG epochs collected during tACS from those where no tACS was performed.

This could be a potential indicator that residual artefacts dominate the cleaned EEG data. If present, these artefacts may give a distinguishable feature which would make it *easy* to separate baseline from baseline+stim, and task from task+stim. However, this explanation is not supported by the high classification accuracy of both the task+stim and baseline+stim cases. While the accuracy was not perfect, these cases are not being interchanged, suggesting that if residual artefacts were present and dominating the classification decision, these residual artefacts must be different depending on whether or not a working memory task was being performed.

3.2 Six class classification accuracy

The above is further supported by the 6 class classification results. The overall average, across all subjects and classes, is shown in Fig. 4. At 75.1% this accuracy far exceeds that from randomly assigning class labels. The breakdown of per-class accuracy is then shown in Fig. 6. Importantly, the classifier was able to separate the nBack and backwards digit recall tasks, with and without stimulation. During tACS stimulation, epochs from the nBack task were recognised with an 80% accuracy, while epochs from the backwards digital recall task with a 72% accuracy. If residual artefacts remained present after the data cleaning, and these dominated the classification process, then these residual artefacts would also have to vary with which task

Figure 6. Fraction of epochs successfully classified using the 6 class model, averaged for all 10 subjects.

Table 1. Performance and standard deviation of the subject independent classifier for each participant.

Participant	Performance	Standard deviation
1	0.631	0.013
2	0.672	0.016
3	0.722	0.014
4	0.685	0.015
5	0.692	0.011
6	0.683	0.013
7	0.611	0.014
8	0.697	0.011
9	0.670	0.013
10	0.707	0.013
Mean	0.677	0.013
Chance level	0.182	0.015

was being performed. Our methodology cannot rule this out, but the tasks were carefully selected to be similar, yet different, to minimise this possibility. Overall the classification accuracies add considerable weight that EEG data, different between tasks and between task/no task conditions, is indeed being collected after artefact removal.

3.3 Subject independent classifier

As a final validation step, Table 1 gives the performance of the subject independent classifier evaluated on each subject in turn. The performance averaged across all subjects was 67.7%, lower than the subject dependent classifier (75.1% from Fig. 4) but still above chance performance. This reduced performance comes at the benefit of not needing to generate a new set of classifier training data each time a new subject is recruited, which will be important as we move towards closed loop tACS systems. Even with this reduction in performance, the accuracy is still over 3 times the chance level of 18.2%.

4 Discussion

The verification of on-person tACS artefact removal is intrinsically very challenging as there is no ground truth reference available in order to determine whether any differences in the recovered EEG signal are due to residual or higher order artefacts still being present, or whether they are due to tACS having a modulation effect on the brain and so altering the EEG which is collected, or whether it is because recordings with different stimulation settings cannot be performed simultaneously and so there are simply slightly different EEGs present at different times. As a result there is much debate over *how good* the artefact removal process must be in order to start being of practical use.

In our previous work [7] we used a multi-step validation procedure to evaluate the artefact removal performance: performing head phantom tests where the EEG signal to recover was known; inspecting individual on-person tests in the time and frequency domains; comparing the descriptive statistics of raw and cleaned EEG data; and demonstrating the recording of evoked potentials during tACS. Here we propose a new validation step, using a classification based approach to discriminate different

parts of an experiment from EEG data after tACS artefact removal. We were successfully able to separate two different tasks that measure different forms of working memory capacity (nBack and backwards digit recall) during tACS from background EEG.

Our experimental protocol was carefully designed with this objective in mind. A relatively low number of subjects was used as the aim was not to report a behavioural or similar effect of tACS, and the tACS stimulation duration was kept deliberately short in order to avoid inducing changes in the brain. A low number of EEG channels was used with channel count independent artefact removal, and a relatively unsophisticated machine learning classifier (an LDA) was used to avoid overfitting, or over-analysis, of the data. Our tasks were selected to involve similar, but different aspects of working memory, previously shown to be enhanced by tACS stimulation [25], [32]–[34].

Our results suggest that if residual artefacts remained after the data cleaning, and these dominated the classification process, then these residual artefacts would also have to vary with which task (if any) was being performed. We are not aware of either a technological or physiological basis for this. To us the more natural conclusion is that either no residual artefacts remained present after the data cleaning, or that those that did did not dominate the classification process. This is not direct evidence that the cleaned data is *good enough* for use in any arbitrary situation, but does add weight to the overall picture as to the significance of any remaining residual artefacts.

Table 2 summaries the performance validation procedures used in EEG+tACS artefact removal works that we are aware of to date. To our knowledge 10 different methods have been reported, with time and frequency domain visual inspection of on person data being by far the most common validation technique. The use of simulated/synthetic EEG+tACS data, modelling the tACS interference as an additive noise source is the second most common. Most of these models include some second order effects such as timing jitter, however the simple additive tACS model is very debatable [16], [17]. The majority of works use more than one validation approach in order to help build confidence in the artefact removal procedure. We propose our classification approach described here as a new potential validation approach which incorporates an ability to assess the impact of any residual artefacts which may be left in the cleaned EEG data. The current study deliberately used a small number of subjects and a short stimulation time to avoid the stimulation inducing changes in brain state. Future work will focus on longer duration tACS protocols, investigating whether a change in the classification performance is seen for data selected from the start, compared to the end, of stimulation.

5 Conclusions

Machine learning can be used to investigate the performance of tACS artefact removal from simultaneously collected EEG signals. A Linear Discriminant Analysis classifier could separate baseline, baseline during tACS (after artefact removal), working memory task without tACS, and working memory task with tACS (after artefact removal) data segments with accuracies ranging from 65–94%. EEG from the nBack and backwards digital recall tasks could be separated during stimulation, with an accuracy exceeding 72%. This suggests a limited impact of any residual artefacts which may be present after the data cleaning process. If residual tACS artefacts remained after the EEG cleaning these did not dominate the classification process. This new validation method could be used in addition to existing validation processes to build greater confidence in the performance of tACS artefact removal algorithms.

References

- [1] A. Woods, A. Antal, M. Bikson, *et al.*, “A technical guide to tDCS, and related non-invasive brain stimulation tools”, *Clin. Neurophysiol.*, vol. 127, no. 2, pp. 1031–1048, 2016.
- [2] I. Iturrate, M. Pereira, and J. del R. Millan, “Closed-loop electrical neurostimulation: Challenges and opportunities”, *Curr. Opin. Biomed. Eng.*, vol. 8, no. 1, pp. 28–37, 2018.
- [3] F. Mansouri, P. Fettes, L. Schulze, *et al.*, “A real-time phase-locking system for non-invasive brain stimulation”, *Front. Neurosci.*, vol. 12, no. 1, p. 8777, 2018.
- [4] E. Underwood, “Cadaver study challenges brain stimulation methods”, *Science*, vol. 352, no. 6284, pp. 397–397, 2018.
- [5] B. L. Parkin, M. Bhandari, J. C. Glen, *et al.*, “The physiological effects of transcranial electrical stimulation do not apply to parameters commonly used in studies of cognitive neuromodulation”, *Neuropsychologia*, vol. 128, no. 1, pp. 332–339, 2019.
- [6] T. O. Bergmann, “Brain state-dependent brain stimulation”, *Front. Psychol.*, vol. 9, no. 1, p. 2108, 2018.
- [7] S. Kohli and A. J. Casson, “Removal of gross artifacts of transcranial alternating current stimulation in simultaneous EEG monitoring”, *Sensors*, vol. 19, no. 1, e190, 2019.
- [8] F. Schlegelmilch, K. Schellhorn, and P. Stein, “A method for online correction of artifacts in EEG signals during transcranial electrical stimulation”, *Clin. Neurophysiol.*, vol. 124, no. 10, e166–e168, 2013.

Table 2. EEG+tACS artefact removal algorithms and their validation procedure.

Ref.	Artefact removal method	Subjects	tACS settings	Validation procedure
[8]	Linear regression model subtraction	1	10 Hz, 1 mA _{pp} , P3/P4	1. Time domain visual inspection of on person data
[9]	Template average subtraction and notch filter	27	2/6/12/25/40/70 Hz, 250 μ A _{pp} , F3/F4 and Tp9/Tp10	2. Head phantom, no baseline EEG output. 3. Time and frequency domain visual inspection of on person data. 4. Statistical analysis of on person data
[10]	Template average subtraction and PCA	16	10 Hz, 1 mA, Cz/Oz	5. Simulation as additive noise. 6. Identification of ERPs from on-person data
[11]	Dynamic least squares fitting based on a nominal frequency template	Likely 1	6/15 Hz, 0.6/1 mA _{pp} , F3/F4	3. Time and frequency domain visual inspection of on person data
[12]	Template average subtraction	4	5/10/40 Hz, 250 μ A _{pp} , P3/Fp4	3. Time and frequency domain visual inspection of on person data. 7. Quantification of individual alpha frequency change
[13]	Sawtooth waveform and template average subtraction	30	10 Hz, 2 mA _{pp} , Oz/Cz	3. Time and frequency domain visual inspection of on person data. 4. Statistical analysis of on person data. 5. Simulation as additive noise
[14]	Weighted comb filters	Likely 1	11 Hz, amplitude not clear, on arm	1. Time domain visual inspection of on person data. 5. Simulation as additive noise
[15]	Hardware set up and clock synchronisation	Likely 1	10 Hz, 1 mA, positions not clear	3. Time and frequency domain visual inspection of on person data
[7]	Template average subtraction and adaptive filtering	5	5/10/40 Hz, Fp2/P3, 0.25/1 mA _{pp}	3. Time and frequency domain visual inspection of on person data. 4. Statistical analysis of on person data. 6. Identification of ERPs from on-person data. 8. Head phantom, with known signal to recover
[16]	Quadrature regression and subsequent independent vector analysis (q-IVA)	10	Galvanic vestibular stimulation (notes could be applied to tACS)	3. Time and frequency domain visual inspection of on person data. 9. Simulation, including skin impedance effects model. 10. Cleaned <i>during</i> vs <i>post</i> quantitative EEG comparison
[17]	Several. ICA best performing	24	30/40 Hz, 325/675 μ A _{pp} , locations not clear	5. Simulation as additive noise
[35]	Signal-space projection	1	10 Hz, Cz/Oz, 0.5/1 mA _{pp}	3. Time and frequency domain visual inspection of on person data. 6. Identification ERPs from on-person data. 8. Head phantom, with known signal to recover

- [9] U. Voss, R. Holzmann, A. Hobson, *et al.*, “Induction of self awareness in dreams through frontal low current stimulation of gamma activity”, *Nat. Neurosci.*, vol. 17, no. 1, pp. 810–812, 2014.
- [10] R. F. Helfrich, T. R. Schneider, S. Rach, *et al.*, “Entrainment of brain oscillations by transcranial alternating current stimulation”, *Curr. Biol.*, vol. 24, no. 3, pp. 333–339, 2014.
- [11] L. Dubreuil-Vall and G. Ruffini, “Starstim EEG monitoring during tACS stimulation”, Neuroelectrics white paper WP201402, Tech. Rep., 2014.
- [12] S. Kohli and A. J. Casson, “Removal of transcranial a.c. current stimulation artifact from simultaneous EEG recordings by superposition of moving averages”, in *IEEE EMBC*, Milan, Aug. 2015, pp. 3436–3439.
- [13] J. Dowsett and C. S. Herrmann, “Transcranial alternating current stimulation with sawtooth waves: Simultaneous stimulation and EEG recording”, *Front. Hum. Neurosci.*, vol. 10, no. 1, p. 135, 2016.
- [14] Guggenberger, R. & Gharabaghi, A. (2017). ARtACS: Artifact removal for transcranial alternating current stimulation, [Online]. Available: <http://doi.org/10.5281/zenodo.581408>.
- [15] E. Garcia Cossio, S. Wunder, and F. Schlegelmilch, “Overview of hardware artefacts during simultaneous tES and EEG”, in *PEPA ON stimulation*, Cambridge, Jan. 2018.
- [16] S. Lee, M. J. McKeown, Z. J. Wang, *et al.*, “Removal of high-voltage brain stimulation artifacts from simultaneous EEG recordings”, *IEEE Trans. Biomed. Eng.*, vol. 66, no. 1, pp. 50–60, 2019.
- [17] F. Barban, S. Buccelli, D. Mantini, *et al.*, “Removal of tACS artefact: A simulation study for algorithm comparison”, in *IEEE NER*, San Francisco, Mar. 2019, pp. 393–396.
- [18] N. Noury, J. F. Hipp, and M. Siegel, “Physiological processes non-linearly affect electrophysiological recordings during transcranial electric stimulation”, *NeuroImage*, vol. 140, no. 1, pp. 99–109, 2016.
- [19] N. Noury and M. Siegel, “Phase properties of transcranial electrical stimulation artifacts in electrophysiological recordings”, *NeuroImage*, vol. 158, no. 1, pp. 406–416, 2017.
- [20] T. Neuling, P. Ruhnau, N. Weisz, *et al.*, “Faith and oscillations recovered: On analyzing EEG/MEG signals during tACS”, *NeuroImage*, vol. 147, no. 1, pp. 960–963, 2017.
- [21] N. Noury and M. Siegel, “Analyzing EEG and MEG signals recorded during tES, a reply”, *NeuroImage*, vol. 167, no. 1, pp. 53–61, 2018.
- [22] F. H. Kasten, E. Negahbani, F. Frohlich, *et al.*, “Non-linear transfer characteristics of stimulation and recording hardware account for spurious low-frequency artifacts during amplitude modulated transcranial alternating current stimulation (AM-tACS)”, *NeuroImage*, vol. 179, no. 1, pp. 134–143, 2018.
- [23] F. H. Kasten and C. S. Herrmann, “Recovering brain dynamics during concurrent tACS-M/EEG: An overview of analysis approaches and their methodological and interpretational pitfalls”, *Brain Topogr.*, vol. To appear in, no. DOI: 10.1007/s10548-019-00727-7, pp. 1–7, 2019.
- [24] N. Jausovec and K. Jausovec, “Increasing working memory capacity with theta transcranial alternating current stimulation (tACS)”, *Biol. Psychol.*, vol. 96, no. 2, pp. 42–47, 2014.
- [25] N. Jausovec, K. Jausovec, and A. Pahor, “The influence of theta transcranial alternating current stimulation (tACS) on working memory storage and processing functions”, *Acta Psychol.*, vol. 146, no. 2, pp. 1–6, 2014.
- [26] A. Tozzi, J. F. Peters, and M. N. Cankaya, “The informational entropy endowed in cortical oscillations”, *Cogn. Neurodyn.*, vol. 12, no. 5, pp. 501–507, 2018.
- [27] R. A. Fisher, “The use of multiple measurements in taxonomic problems”, *Ann. Hum. Genet.*, vol. 7, no. 2, pp. 179–188, 1936.
- [28] C. R. Rao, “The utilization of multiple measurements in problems of biological classification”, *J. Royal Stat. Soc.*, vol. 10, no. 2, pp. 159–203, 1948.
- [29] A. Zollanvari, “Analytical study of performance of linear discriminant analysis in stochastic settings”, *Pattern Recognit.*, vol. 46, no. 11, pp. 3017–3019, 2013.
- [30] E. Alpaydin, *Introduction to Machine Learning*, 2nd. Boston: MIT Press, 2010.
- [31] J. C. Christensen, J. R. Estep, G. F. Wilson, *et al.*, “The effects of day-to-day variability of physiological data on operator functional state classification”, *NeuroImage*, vol. 59, no. 1, pp. 57–63, 2012.
- [32] O. Meiron and M. Lavidor, “Prefrontal oscillatory stimulation modulates access to cognitive control references in retrospective metacognitive commentary”, *Clin. Neurophysiol.*, vol. 125, no. 1, pp. 77–82, 2014.
- [33] K. E. Hoy, N. Bailey, S. Arnold, *et al.*, “The effect of gamma-tACS on working memory performance in healthy controls”, *Brain Cogn.*, vol. 101, no. 1, pp. 51–56, 2015.
- [34] J. Vosskuhl, R. J. Huster, and C. S. Herrmann, “Increase in short-term memory capacity induced by down-regulating individual theta frequency via transcranial alternating current stimulation”, *Front. Hum. Neurosci.*, vol. 9, no. 1, p. 256, 2015.

- [35] J. Vosskuhl, T. P. Mutanen, T. Neuling, *et al.*, “Signal-space projection suppresses the tACS artifact in EEG recordings”, *bioRxiv*, vol. 823153, no. DOI: 10.1101/823153, pp. 1–30, 2019.

Acknowledgements

This work was funded by the Engineering and Physical Sciences Research Council through their Doctoral Training Partnership with the University of Manchester.

Author contributions statement

S.K. conceived the artefact removal algorithms, performed the on-person testing experiments, and coded the machine learning. A.J.C. helped devise the testing methodology and in the analysis and interpretation of the results. Both authors contributed to writing the manuscript.

Additional information

The authors have no conflicts of interest to report. Data sharing is not possible for this publication as under the ethical approval the authors do not have permission to release the physiological data used outside of the University of Manchester.