



UNIVERSITY OF LEEDS

This is a repository copy of *Analysing wireless EEG based functional connectivity measures with respect to change in environmental factors*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/109217/>

Version: Accepted Version

Proceedings Paper:

Biswas, D, Bono, V, Scott-South, M et al. (7 more authors) (2016) *Analysing wireless EEG based functional connectivity measures with respect to change in environmental factors*. In: *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI 2016)*. BHI 2016, 24-27 Feb 2016, Las Vegas, USA. IEEE , pp. 599-602. ISBN 9781509024551

<https://doi.org/10.1109/BHI.2016.7455969>

(c) 2016, IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

Analysing Wireless EEG based Functional Connectivity Measures with Respect to Change in Environmental Factors

Dwaipayan Biswas, Valentina Bono, Michael Scott-South, Shre Chatterjee, Anna Soska, Steve Snow, Catherine Noakes, Janet F. Barlow, Koushik Maharatna and mc schraefel, Member, IEEE

Abstract— In this paper we present a systematic exploration to formulate a predictive model of the human cognitive process with the changing environmental conditions at workplace. We select six different environmental conditions with small change in temperature/ventilation representative of realistic work environment having manual control. EEG data were acquired through 19-channel wireless system from three participants and CO₂, Temperature, Relative humidity were recorded throughout the six conditions. The EEG data was pre-processed using an artifact reduction algorithm and 129 neurophysiological features were extracted from functional connectivity measures using complex network analysis. The environmental data were processed to generate 15 time/frequency domain features. Five best features selected through a ranking algorithm for all the variables across the six conditions were processed to formulate a model (environmental parameters as predictors) using retrospective 10-fold cross-validation in conjunction with multiple linear regression. The model was prospectively evaluated over 10 runs on a test set to predict the EEG variable across the six conditions and parameters corresponding to the run producing least root mean square error were reported. Our exploration shows that the condition having no modulation of the ambient environmental parameters reflects the optimum condition for predicting the EEG features using the examined environmental parameters.

I. INTRODUCTION

The design of modern workplaces (e.g. offices) incorporates the fact that maintaining environmental parameters (e.g. ventilation, air quality, etc.) at predefined levels assures comfort and satisfaction of the occupants which is also directly related to their health and productivity [1]. A number of studies have explored the way occupants perceive poor indoor air quality leading to degraded cognitive function, reduced productivity and overall well-being. However, same indoor environments lead to different subjective responses, primarily due to inter-person differences in perception of environmental factors as well as inherent behavioral factors that influence the decision on comfortability, which is usually beyond the regulation of any standards. It is not possible for HVAC engineers to cater to the varying demands of individuals at workplace.

Analysing the underlying mechanism involved in the perception of changing indoor environmental factors could lead to an objective control or modulation of these factors. This mechanism can be represented by neurophysiological measurement which can access the processes in the

fundamental brain structures [2] responsible for the evolution in perception/emotion dynamics with respect to the changing environment. Electroencephalogram (EEG) is one of the biological signals that are widely used in the field of healthcare for neurofeedback treatment. EEG with its high temporal resolution can detect the immediate responses to external stimuli which affect the perceptual state of the human brain which has been exploited by the research community to understand the effect of changing environmental parameters on human cognition [2], [3]. As the cognitive process (e.g. perception) involves a large-scale network [4] instead of a single brain region [5], a multichannel EEG analysis investigating the interaction among different brain sites could formulate an understanding of the underlying cognitive process. The information exchange between the networks of segregated functional units of the brain which integrate with each other can be described by functional connectivity (FC) measures, which can be quantified by a number of neurobiological features using complex network analysis [6].

In this study we aim to investigate the EEG-based features extracted from brain signals acquired during changing environmental factors (acting as a stimuli) to determine the significant features involved in cognitive processing. These features are used to model the relationship between the cognitive processes (functional units) and the environmental variables. Identification of significant features and the model parameters can help to control the environmental factors and lead towards intelligent workspace design.

For this investigation, we recorded EEG data elicited by six environmental conditions from three participants in a naturally ventilated building. A wireless 19-Channel EEG system was used to collect data from the subjects and relevant processing was done to identify EEG features which were used for modelling the relationship of the cognitive process with each different condition. The acquired EEG data was de-noised using wavelet packet transform-empirical mode decomposition (WPT-EMD) [7] artifact reduction algorithm. The processed EEG data were used to generate FC based measures which were represented with a reduced dimensionality by applying the Brain Connectivity Toolbox (BCT) yielding 129 features. Similarly, we extract 15 time and frequency domain features from the environmental monitoring data acquired through carbon-dioxide (CO₂), temperature (T) and relative humidity (Rh) sensors. We select five significant features that maximizes the variance in the feature space, each from EEG, CO₂, T and Rh data. These features are successfully

*Research supported by EPSRC funded Refresh project, Grant Agreement # EP/K021893/1.

D. Biswas, V. Bono, S. Chatterjee, M. South, A. Soska, S. Snow, K. Maharatna and mc are with the Electronics and Computer Science, University

of Southampton, Hampshire, SO17 1BJ, UK (e-mail: { db9g10, vb2a12, mss2g13, skc105, A.I. Soska, S.Snow, km3, mc }@ecs.soton.ac.uk). C. Noakes is with University of Leeds (email: C.J. Noakes@leeds.ac.uk) and J. Barlow is with the University of Reading (email: j.f.barlow@reading.ac.uk).

used for building a model using multiple linear regression in conjunction with 10 runs of 10-fold cross-validation (CV) to prospectively select the parameters that relate the environmental data to the EEG.

II. BACKGROUND

The research community has for long focused on analyzing the relationship between various environmental and behavioral factors with cognitive processes using dedicated tests (e.g. cognitive executive functions [8]) performed by participants in a controlled environment. These tests are primarily targeted towards evaluating various functionalities of the human brain such as planning, working memory, reasoning etc. However, recent research has brought in its spectrum the analysis of sensor-based physiological signals (e.g. electrocardiogram, EEG) with respect to changing behavioral factors (e.g. physical exercises [9], sleep [10], fatigue [11], etc.) and environmental factors (e.g. temperature [2], CO₂ [3]), thereby nullifying the subjective quotient involved with cognitive test administration/reporting.

A recent study [2] focused on analyzing the effect of changing temperature and emotional status on the brain waves, which to the best of our knowledge appears to be the most recent work in terms of relating cognitive process, temperature and emotional status. However, they mainly looked into the cross-correlation between the EEG channels to trace the movement of the emotional information when the brain carries out its activities under the influence of widely changing temperature values. Moreover, analyzing emotional status has been well researched using brain connectivity measures [12]. Hence, in our study we make a first attempt to analyze the effects of three environmental parameters (CO₂, Rh, T) on the cognitive process by analyzing FC measures (moving beyond the ‘single electrode level analysis’) extracted from EEG signals.

III. DATA ACQUISITION

In this study we chose to measure three common environmental parameters - CO₂, T, Rh and a physiological signal – EEG from three participants who gave their consent, in an independent room within a naturally ventilated building. The room was well lit up, had two glass windows and could hold up to four people. Each individual participant spent approximately half an hour in the room accompanied by the researcher, seated on a standard height chair by a table. The participants were exposed to six different environmental conditions, representative of a realistic workplace scenario (cf. Table I). CO₂, T and Rh data were measured using a HOBO sensors, sampling at 1 sample/sec. EEG signals were recorded throughout the experimental duration using the wireless Enobio system [13] with 19 channels according to the International 10-20 system with a sampling frequency of 500 Hz. It is important to note that throughout the experiments the subjects were instructed to perform no cognitive tasks (e.g. reading, computer work, etc.) as we wished to analyze the baseline cognitive effects of the changing environmental condition. The data in cond1 was primarily targeted towards acquiring the resting stage EEG data used for the artifact removal algorithm used in post-processing the data. The rest of the 5 conditions (cond2 – cond6) each lasting for duration of 5 minutes were representative of the realistic office scenario

having manual control over windows/heater and were targeted towards capturing the perception of the participants towards the changing factors. The small range of the sensor recordings (CO₂: 555 – 924 ppm; T: 22 - 25°C; Rh: 41 – 47%) further highlight that the conditions were representative of a realistic office environment with manual control.

TABLE I. EXPERIMENTAL ENVIRONMENTAL CONDITION

Conditions	Duration	Description
cond1	2 mins	Eyes closed, heating off, windows closed.
cond2	5 mins	Eyes open, heating off, windows closed.
cond3	5 mins	Eyes open, heating on, windows closed.
cond4	5 mins	Eyes open, heating on, windows open.
cond5	5 mins	Eyes open, heating off, windows open.
cond6	5 mins	Eyes open, heating off, windows closed.

IV. METHODS

An overview of the methodology has been illustrated in Figure 1. We first discuss the processing of the EEG signal followed by the environment data.

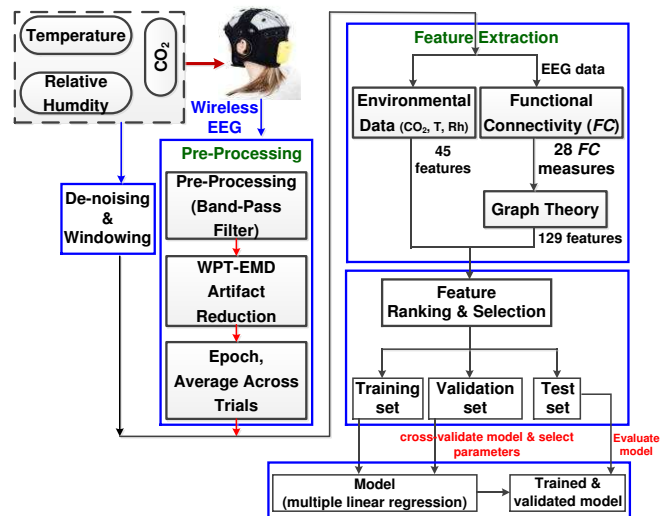


Figure 1. Overview of the methodology.

The EEG recordings were pre-processed with a band pass filter having a cut-off frequency of 0.5 Hz - 42 Hz and further processing involved the following stages – artifact reduction, epoching and feature extraction. The acquired EEG data, contaminated by artifacts (primarily due to eye-blinking and involuntary body-head movements of the participants during the experiments) was de-noised using the artifact reduction algorithm - WPT-EMD [7] prior to epoching. The WPT-EMD uses the resting state EEG, while the subject has their eye closed (cond1). EEG epochs of 8 seconds (deemed sufficient to capture vital information) were extracted from the pre-processed data for all the six conditions, obtaining an ensemble of epochs for each experimental condition. A threshold of 200 μ V was applied on the epochs and those selected within the thresholds were averaged and used for feature extraction.

A. Functional connectivity (FC) - EEG

Hermes Toolbox was used to generate the 29 FC measures; each of these is a matrix with size 19 \times 19 (19 being the number of electrodes). Among the measures, described in Table II, four connectivity measures related to the phase

synchronization (PS) between two signals (i.e. 5-8) have been estimated for each individual band (in Hz) – θ (4-8), α (8-12), β (12-32), γ (32-42) and all bands (6-42), resulting in total of $[(4 \times 5) + 9]$ 29 features.

TABLE II. FUNCTIONAL CONNECTIVITY MEASURES

No.	Measures	Description
1.	CrossCorrelation	linear correlation between two signals as a function of time
2.	Correlation	Pearson’s correlation coefficient (at zero lag)
3.	Coherence	linear correlation between two signals as a function of frequency
4.	Phase Slope Index	estimation of the flow direction of information between two signals as a function of time
5.	Phase Locking Value (PLV)	(PS) inter-trial variability of the phase difference between two signals at time t
6.	Phase-Lag Index (PLI)	(PS) similar to PLV, however rejects phase distributions centered around zero
7.	P Index	(PS) based on Shannon entropy, quantifies the deviation of the distribution of the cyclic relative phase from the uniform distribution
8.	Directionality Phase Indexes (DPI)	(PS) Analysis of the temporal evolution of the phase derivative
9.	Granger Causality	linear parametric method, measures if signal x provides predictive information about signal y
10.	Transfer Entropy	is non-parametric, measures the amount of directed information flow from signals x to y
11.	Partial Directed Coherence	a frequency domain measure of Granger causality, based on modelling time series by multivariate autoregressive (MAR) processes
12.	Direct Transfer Function	similar to PDC, however use a Hermitian transpose instead of a Fourier transform
13.	Mutual Information	measures the amount of information shared between two signals

a. Detailed description of the measures have been provided in [4]

D. Feature extraction: Brain connectivity toolbox (BCT)

BCT uses graph theory analysis on the FC measures except for 1, 3, 4, 11, 12, (cf. Table II) and also the measures Modularity, Radius and Diameter were not computed for the five DPI bands (cf. 8, Table II). Hence this yields a total of 129 features $[24 \text{ FC} \times 6 \text{ GT} - 15]$.

TABLE III. GRAPH THEORETIC MEASURES

No.	Measures	Description
1.	Transitivity	measure of segregation (i.e. how many node’s neighbors are connected among themselves)
2.	Modularity	measure of segregation; it measures how much the network can be divided into subgroups with dense links within-groups and few links between-group
3.	Characteristic path length	measure of integration; measures the average distance between nodes across the entire network
4.	Global efficiency	measure of integration; it is the inverse of the distance between nodes
5.	Radius	measure of shape of network- minimum eccentricity
6.	Diameter	measure of shape of network- maximum eccentricity

b. Detailed description of the measures have been provided in [6]

E. Feature extraction: Environment data

We pre-process the CO₂, T and Rh data using a moving average smoothing filter and divide the data into segments of 8 seconds with an overlap of 1 sample. The following features were extracted from all the windows – time-domain: mean, standard deviation (std), kurtosis, skewness, sample entropy,

signal energy; frequency-domain: maximum and minimum amplitude of the power spectral density using fast Fourier transform (fft), energy content of the fft signal; time-frequency representation: energy content of discrete Wavelet transform (dwt) coefficients for three levels of signal decomposition, using using ‘haar/daubechies’ wavelets.

This concludes the feature extraction process resulting in 129 features extracted from EEG and 45 features in total from the three environmental parameters (CO₂, T, Rh) corresponding to each of the three participants.

V. FEATURE RANKING AND MODELLING

In this work, our target was to formulate a predictive model that relates the EEG data with the environmental data. We rank and select the five best features from either of the feature sets (EEG – 129; CO₂, T, Rh – 15 each), using the low-complexity class-separability measure based on scatter matrices. The algorithm ranks each individual feature for a multiple-class scenario where a high rank represents a small within-class variance and a large between-class distance among the data points in the respective feature space [14].

The selected feature set for each of EEG, CO₂, T, Rh are used within a multiple linear regression model for each of the six environmental conditions. We consider the environment features (CO₂, T, Rh) as the predictors/independent variables and the EEG features as the response/dependent variable. For each of the condition, we randomly separate the 15 data samples (5 features \times 3 participants) into a training set (10 samples) and a test set (5 samples). The training samples are modelled and validated through a 10-fold CV step and the model parameters of the corresponding step which produces the least root mean square error (rmse) are used to prospectively predict the EEG values of the 5 test samples. The process is repeated for 10 runs, selecting a random split of 10 and 5 data samples in each run. This process is repeated for each of the 6 conditions and the model parameters producing the least rmse for the predicted EEG values are selected as the optimal model.

VI. RESULTS

The best ranked features selected for the modelling/validation and prospective testing phase for EEG and environmental data are listed in Table V. The rmse values as result of 10 runs of 10-fold CV using the linear regression model for each of the 6 conditions representing the difference between the original and the predicted EEG values has been listed in Table VI. The parameters for the best model derived which produces the least rmse (cond2) across the six conditions is given in (1).

$$\text{EEG} = 0.6 - 0.652\text{CO}_2 - 0.277\text{T} - 0.029\text{RH} \quad (1)$$

TABLE IV. LIST OF THE BEST FEATURES SELECTED ACROSS THE SIX CONDITIONS FOR THREE PARTICIPANTS COMBINED

Parameters	List of best Features
EEG	Radius[PLI(β)]; Radius[PLI(γ)]; Radius[PLI(α)]; Efficiency[PLV(all bands)]; Efficiency [PLI(γ)]
CO ₂	Mean; Energy_db3_13; Energy_db3_12; Energy_db3_11; Energy_haar_13

T	Mean; std; FFT_power; PSD_max; Energy_haar_I1
Rh	Mean; std; Kurtosis; Energy_db3_I3; Energy_db3_I2

TABLE V. RMSE FOR THE 6 CONDITIONS AS A RESULTS OF 10 RUNS OF 10-FOLD CROSS-VALIDATION OF THE LINEAR REGRESSION MODEL

Conditions	RMSE
cond1	0.484
cond2	0.156
cond3	0.387
cond4	0.272
cond5	0.232
cond6	0.437

Cond2 represents a stable condition when none of the pre-existing (since the start of the experiment) environmental parameters undergo any change. Although we return back to same state in cond6 but only after transiting through states (cond3 – 5) over a duration of 5 minutes each. The accumulative effect of these changing environmental conditions on the EEG data is quantified by the relatively high rmse between the predictor and response variables.

The five best EEG features having a large variance among all the competing classes (i.e. conditions) used for the modelling represents the graph theory based parameters - Radius and Global_efficiency chosen for the FC measures PLI and PLV (phase synchronization and lag) across different frequency bands computed on the EEG data [6]. The Radius represents the shape of the network (minimum distance of a node from all other nodes) whereas the Global_efficiency represents the degree of information integration within the graphical network. The best features selected from the environmental data reflect the importance of the energy components of the dwt coefficients which is a time-frequency representation of the signal, besides highlighting simple time-domain features such as mean, std, reflecting change in environmental parameters over time.

VII. DISCUSSION

In this paper, we describe a systematic exploration using several EEG features and time/frequency domain features extracted from three types of environmental data collected across six different environmental conditions on three subjects, to formulate a linear model. Using this linear relationship (obtained through 10 runs of 10-fold CV) and the best ranked features it is possible to predict the EEG using the CO₂, T and Rh data. Our results show that using mainly the wavelet coefficients and few time domain features on the environmental data we can prospectively predict the EEG data represented by two graph theoretic measures – Radius and Global_efficiency computed on the phase information. The prospective prediction produces the minimal error (rmse) for the second condition out of the six experimental conditions primarily due to the absence of any external environmental modulation.

This exploration is aimed at a ‘proof-of-concept’ first study to relate three environmental parameters to the cognitive processing represented by a network of active functional units of the brain. It is important to note that here we represent the six conditions for a very small duration (5 mins) each having a small difference in the range of the individual values. This simple case-study is representative of the minimal variations

inherent in workplaces having manual control over environmental conditions.

Hence, this study sets the pathway for the following exploration in the near future: 1) analyzing the evolution/change of FC measures computed on each 8 seconds epochs (instead of averaging) along with the change in the feature values computed on the environmental parameters; 2) change in the sequence and duration (longer than 5-minute) of the environmental conditions and seeing how it affects the models and the corresponding results; 3) having the participants perform some cognitive tasks while their EEG is recorded along with changes in environmental parameters. A predictive model using the environmental parameters and the cognitive process (quantified by EEG features) will lay the foundation for intelligent workplace design which could help the occupants to perform at optimal levels without depending on manual control.

REFERENCES

- [1] M. Frontczak and P. Wargocki, “Literature survey on how different factors influence human comfort in indoor environments,” *Building and Environment*, vol. 46, no. 4, pp. 922–937, 2011.
- [2] J.-H. Shin and D.-H. Park, “Analysis for Characteristics of Electroencephalogram (EEG) and Influence of Environmental Factors According to Emotional Changes,” in *Convergence and Hybrid Information Technology*, Springer, 2011, pp. 488–500.
- [3] P. Strøm-Tejsten, et al., “The effect of CO₂ controlled bedroom ventilation on sleep and next-day performance,” in *13th SCANVAC International Conference on Air Distribution in Rooms*.
- [4] G. Niso, et al., “HERMES: towards an integrated toolbox to characterize functional and effective brain connectivity,” *Neuroinformatics*, vol. 11, no. 4, pp. 405–434, 2013.
- [5] I. B. Mauss and M. D. Robinson, “Measures of emotion: A review,” *Cognition and emotion*, vol. 23, no. 2, pp. 209–237, 2009.
- [6] M. Rubinov and O. Sporns, “Complex network measures of brain connectivity: uses and interpretations,” *Neuroimage*, vol. 52, no. 3, pp. 1059–1069, 2010.
- [7] V. Bono, et al., “Artifact reduction in multichannel pervasive EEG using hybrid WPT-ICA and WPT-EMD signal decomposition techniques,” in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, 2014, pp. 5864–5868.
- [8] C. T. Gualtieri and L. G. Johnson, “Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs,” *Archives of Clinical Neuropsychology*, vol. 21, no. 7, pp. 623–643, 2006.
- [9] K. Hötting and B. Röder, “Beneficial effects of physical exercise on neuroplasticity and cognition,” *Neuroscience & Biobehavioral Reviews*, vol. 37, no. 9, pp. 2243–2257, 2013.
- [10] K. Kesper, et al., “ECG signal analysis for the assessment of sleep-disordered breathing and sleep pattern,” *Medical & biological engineering & computing*, vol. 50, no. 2, pp. 135–144, 2012.
- [11] J. B. Coquart, et al., “Influence of successive bouts of fatiguing exercise on perceptual and physiological markers during an incremental exercise test,” *Psychophysiology*, vol. 46, no. 1, pp. 209–216, 2009.
- [12] Y.-Y. Lee and S. Hsieh, “Classifying different emotional states by means of EEG-based functional connectivity patterns,” *PLoS one*, vol. 9, no. 4, p. e95415, 2014.
- [13] “ENOBIO Wearable Wireless EEG Cap.” [Online]. Available: <http://www.neuroelectrics.com/products/enobio/enobio-20/>.
- [14] S. Theodoridis and K. Koutroumbas, “Pattern recognition and neural networks,” in *Machine Learning and Its Applications*, Springer, 2001, pp. 169–195.