Dynamic Data Driven Approach for Modeling Human Error

Wan-Lin Hu[1], Janette J. Meyer[2], Zhaosen Wang[3], Tahira Reid[1], Douglas E. Adams[2], Sunil Prabnakar[3], and Alok R. Chaturvedi[4]

[1] School of Mechanical Engineering
Purdue University, West Lafayette, Indiana, U.S.A
hu188@purdue.edu
[2] Laboratory of Systems Integrity and Reliability (LASIR)
Vanderbilt University, Nashville, Tennessee, U.S.A.
janette.j.meyer@vanderbilt.edu
[3] Department of Computer Sciences
Purdue University, West Lafayette, Indiana, U.S.A.
[4] Krannert School of Management
Purdue University, West Lafayette, Indiana, U.S.A.

Abstract
Mitigating human errors is a priority in the design of complex systems, especially through the use of body area networks. This paper describes early developments of a dynamic data driven platform to predict operator error and trigger appropriate intervention before the error happens. Using a two-stage process, data was collected using several sensors (e.g., electroencephalography, pupil dilation measures, and skin conductance) during an established protocol - the Stroop test. The experimental design began with a relaxation period, 40 questions (congruent, then incongruent) without a timer, a rest period followed by another two rounds of questions, but under increased time pressure. Measures such as workload and engagement showed responses consistent with what is expected for Stroop tests. Dynamic system analysis methods were then used to analyze the raw data using principal components analysis and the least squares complex exponential method. The results show that the algorithms have the potential to capture mental states in a mathematical fashion, thus enabling the possibility of prediction.

Keywords: bio-sensors, dynamic data-driven application systems (DDDAS), least squares complex exponential (LSCE), error detection

1 Introduction

Human factor studies have shown that nearly 80% of the root causes contributing to major accidents affecting safety, the environment, and/or economics can be attributed to human

operator error [18]. Causes for the human operator error are many. In a report for the TNO Institute for Perception, Raaijmakers [19] discussed the impact of mental and physical stress on human performance in decision-making situations, specifically for time critical decisions. Raaijmakers argued, " ... effects of various stressors (fatigue, sleep loss, time pressure, anxiety, and cognitive strain) are shown to be task-dependent. At lower levels, fatigue and sleep loss seem to be the most important stressors; at higher levels, the largest effects are expected to be time pressure and cognitive strain." (p. 1).

This paper describes preliminary developments of a dynamic data driven platform to predict operator error and trigger appropriate intervention before the error happens. Taking advantage of the advances in commercial grade psychophysiological sensors such as electroencephalography (EEG), electrocardiogram (ECG), galvanic skin response (GSR), and eye-tracking systems, a human operator's mental and physical states before, during, and after each cognitive activity is captured in an android device which doubles as the server for a wireless body area network (WBAN). Thus the operator becomes an integral part of the system through a brain computer interface (BCI). The WBAN server assimilates and fuses data from multiple sensors and streams it to a stress-intensity simulation model, which, in turn may trigger appropriate intervention.

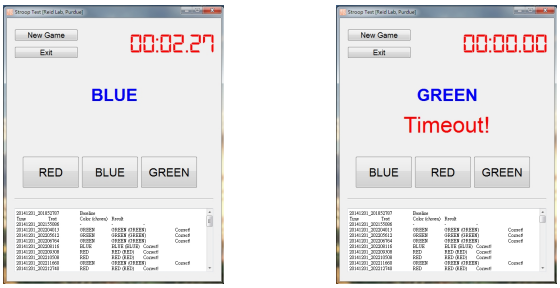
The rest of the paper is organized as follows: background literature is presented in Section 2, Section 3 describes the methodology, Section 4 presents the preliminary results, Section 5 discusses approaches to raw data analysis; Section 6 presents the results and analysis; and Section 7 concludes the paper.

2 Background Literature

By virtue of being human, it is understood that people will make errors [21]. The field of aviation has been the most prominent in developing methods for evaluating and preventing human error through processes and protocols [13, 17]. In recent years, a number of interventions have been developed to help mitigate human errors. In the automotive industry, new technologies have been introduced that automatically slow vehicles down prior to rear-ending a car or potentially hitting a pedestrian [6]. Such methods work through technologies embedded in physical systems.

Psychophysiological measurements provide a means for gathering data directly from people. These data are physiological measurements that correspond to psychological states. EEG is a common sensor that has been used in studies to predict future and present cognitive performance [20] and to develop a driver drowsiness monitoring system [12]. In [9] a body-mounted system is developed to prioritize incoming data streams based on the user's workload, which is measured using EEG and ECG signals.

Recent advances in wireless communication and wearable devices have led to the development of wireless body-area networked systems. Fletcher et al. developed a wearable sensor platform for integrating data from skin conductance and heart rate to monitor emotion and behavior for clinical use. Cognitive Behavioral Therapy could be given automatically based on bio-signals [11]. Choi et al. established a wearable mental stress monitoring system based on heart rate, respiration, electromyography (EMG), and electrodermal activity (EDA) [5]. Kumar et al. further introduced a mobile health system structure to recognize behaviors and deliver timely interventions [15]. Although several studies have proposed wearable biosensor systems to track stress or behavior, the method for preventing errors based on these measures is not well-established.



(a) (b)

Figure 1: Screen shots of the computer-based Stroop test. (a) A congruent question with a 3-second timer. (b) An incongruent question. The participant has ran out of time.

3 Methodology

3.1 Subjects

The study protocol was approved by the Institutional Review Board on human investigation. A total of ten adults volunteered to join the study after passing the prescreening and providing a written consent form. The first five participants completed initial characterization tests to help refine the experimental procedure. Only data from the second five participants (two women and three men) who joined the formal study will be reported here.

3.2 Computer-Based Stroop Test

In order to record synchronized response times and biosignals, a computer-based Stroop test developed on Qt 5.3 (Qt Project Housing, Norway) was used in the study (Figure 1). To make the question colors easy to distinguish, only primary colors (red, blue, and green) were included, which is consistent with established Stroop test protocols [7]. Moreover, participants used arrow keys (left, down, and right) to answer questions. The setup was intuitive, so participants could answer as quickly as possible when they figured out the answer. The result of each question showed on the screen for 0.5 seconds. If the participant did not answer in time, a "timeout" warning was displayed and the answer was recorded as incorrect for that question.

3.3 Equipment

The main approach of this research is to develop a body area network to capture signals from bio-sensors. In order to obtain holistic psychophysiological responses of participants during decision making (and various kinds of cognitive activities), three types of sensors were used in the proposed experiment. A B-Alert X10 EEG headset system (Advanced Brain Monitoring, Inc., USA), a Shimmer3 GSR+ Unit (Shimmer, USA), and a Tobii X60 eye tracker (Tobii Technology, USA) were utilized to collect brain waves, galvanic skin response, and pupil dilation, respectively. All data collection was performed and synchronized through iMotions Attention Tool (iMotions, Inc., USA) on a Windows 7 operating system.

B-Alert X10 is a 9-channel EEG system, which records brain waves from mid-line and lateral sites on the scalp including Fz, Cz, POz, F3, C3, P3, F4, C4, P4 (Figure 2). In addition to raw waves measured from the above nine sites using a 256 Hz sampling rate, B-Alert X10

Figure 2: Locations and labels of data channels acquired using the B-Alert system [1].

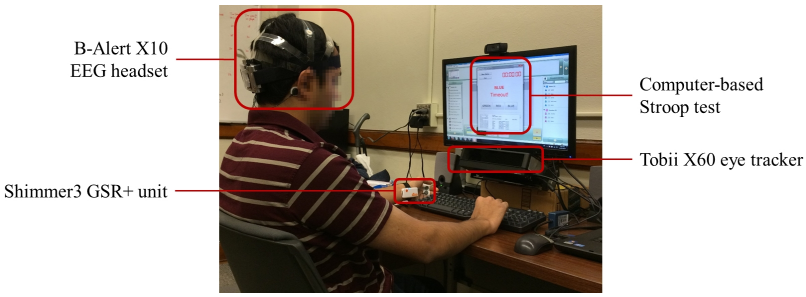


Figure 3: A participant wears the GSR sensor on the left (nondominant) wrist and the EEG head set, and uses the keyboard to do the computer-based Stroop test. The eye-tracking system is mounted under the screen monitor for tracking the participants pupils.

generates four metrics (engagement, workload, distraction, and drowsy) in epochs of one second for representing cognitive states,

Participants' skin resistances between 2 electrodes attached to the index and the middle fingers of the nondominant hand were monitored by the Shimmer 3 GSR+ Unit using a 52 Hz sampling frequency. Skin resistance fluctuates in response to external or internal stimuli [10]. The GSR was included in the body area network to detect the mental stress response.

The proposed experiment is a computer-based test, so a Tobii X60 eye tracker was mounted on the screen to track how participants pupil dilation responded to stimuli on the screen monitor. Figure 3 shows the experimental setup of the EEG headset, the GSR sensor, and the eye-tracking system.

The ultimate goal of this work is to build a cost-efficient, compact, and portable error early warning system, and NeuroSky MindWave Mobile (NeuroSky, Inc., USA) will be a possible solution. NeuroSky is a non-invasive EEG that connects the user to iOS and Android platforms, and transfers all signal information through Bluetooth as opposed to radio. It can detect brainwaves ranged from 0.1 Hz to 30 Hz, and divide them into delta, theta, low/ high alpha, low/ high beta, and low/high gamma parts. An app was developed to acquire the data from Neurosky . All data are sampled once per second. The attention and meditation level are calculated based on the eight wave bands. Attention indicates the user level of focus and meditation indicates the user level of relaxation.

3.4 Procedure

Participants were briefed with the experimental procedure and signed the consent form at the beginning of the experiment. After setting up the EEG and GSR sensors, participants were asked to follow instructions to perform a series of tasks for establishing the EEG baseline. The personalized baseline was used to generate and normalize cognitive metrics: workload, engagement, distraction, and drowsy measures. Then participants were asked to follow a dot on the screen to calibrate the eye tracker. The preparation process took approximate 50 minutes.

The experimental procedure is shown in Figure 4. After finishing calibrations, participants started to do the experiment with a 3-minute baselining session. During this time, they rested and listened to Bach's Harpsichord Concerto No. 5 in F Minor BWV 1056, which meets the

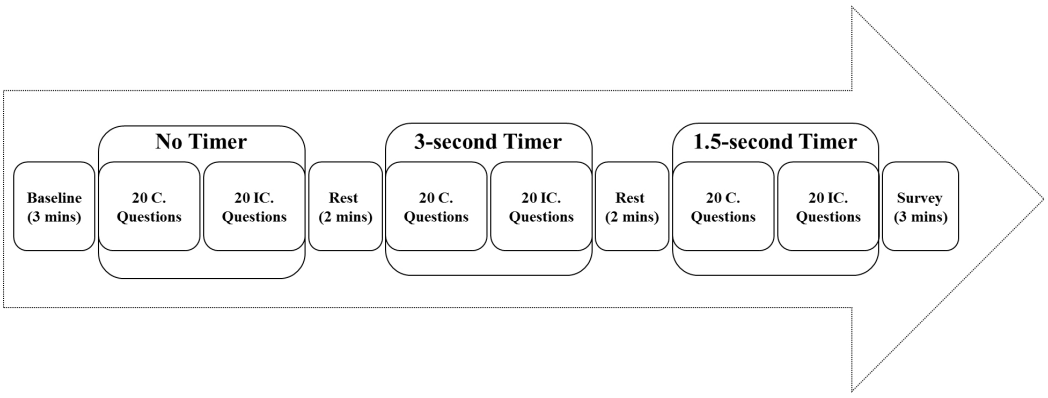


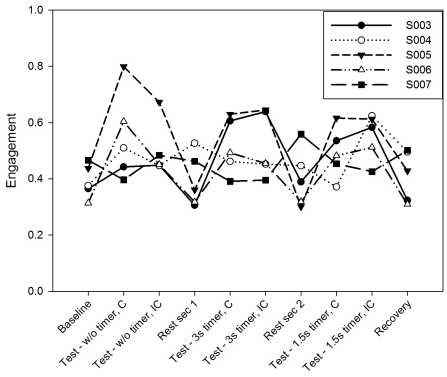
Figure 4: The proposed experimental procedure. C: congruent; IC: incongruent.

criterion for being of a relaxing nature [4]. The biosignals recorded from this session showed that the participants were in a relaxed state. Following the baseline session, participants answered 20 congruent questions without a timer. Within these 20 questions, text and color were the same and participants were able to answer questions at their own pace. Next, they answered 20 incongruent questions without a timer. Participants were then given a 2-minute rest session. Next, they answered 20 congruent and 20 incongruent questions with a 3-second timer. This was followed by another 2-minute rest session, and one more set of 20 congruent and 20 incongruent questions, but with a 1.5-second timer. The Stroop effect interferes with the reaction time of a task, thus it was expected that participants would make more errors when answering incongruent questions with time pressure. Lastly, participants answered a short survey about their previous experience with the Stroop test and their perception of time pressure during the test.

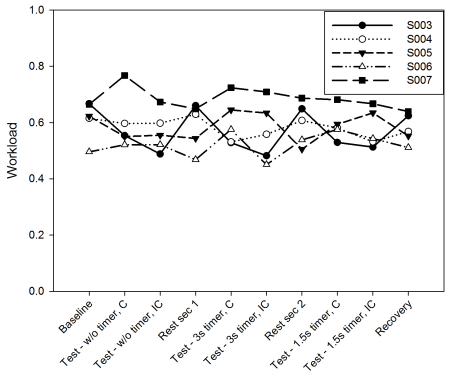
4 Preliminary Results

To determine an overview of each cognitive state, the average values of each metric over the time spent on each session were calculated. Therefore, there were 10 data points of each metric for each participant. Figure 5a and Figure 5b show the level of engagement and workload of all participants through the test. Although there were differences between participants, they tended to show lower engagement during rest sessions including the baseline session, and had higher engagement when getting familiar with the test (i.e., the first 20 congruent questions) and under considerable time pressure (i.e., the last 20 incongruent questions with 1.5-second timers). S003 had a relatively low engagement level in the first test session. This participant had prior experience with doing the Stroop test, which may have influenced this result. The workload level did not change as much as engagement, but it was noticed that native English speakers (S004, S005, S007) had higher average workload when doing tests. S007 showed a somewhat different trend from other participants in both engagement and workload. Based on the survey result, she/he did not have enough sleep in the previous night (about 5 hours) and did not feel stressed when the timer had started.

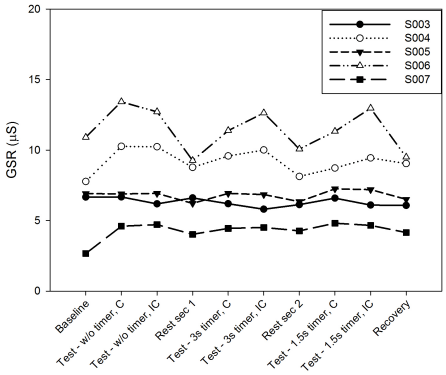
Figure 5c shows the GSR of all participants. The range of skin resistance is individual dependent, so every participant has her/his own baseline. The unit of this figure is μS; a higher



(a) The overview of the engagement level of all participants through the test. Engagement is a scale-normalized metric.



(b) The overview of the workload level of all participants through the test. Workload is a scale-normalized metric.



(c) The overview of the mental stress level of all participants through the test. The range or skin resistance varies with people.

Figure 5

conductance indicates a higher arousal. In the experiment, the arousal level was associated with the stress level. Thus, the curves were fitted for the difficulty level, and were similar to the change in engagement.

To investigate how biosignals change with the outcome of each question, signals measured from the EEG headset, the GSR sensor, the eye tracker, and the time spent on each question were plotted and questions answered incorrectly were identified. For example, Figure 6 is the plot of S005. She/he showed very low engagement/very high distraction in some questions in the last session, where many errors occurred (9/20). Further experiments and analyses are needed for identifying the causality. In addition, when she/he noticed an error was made, or adjusted her/his state from rest to answer questions, the stress level increased. One of the most common reasons for incorrect responses was that the participant ran out of time. Participants tended to answer in a very short period for the question that followed an incorrectly answered

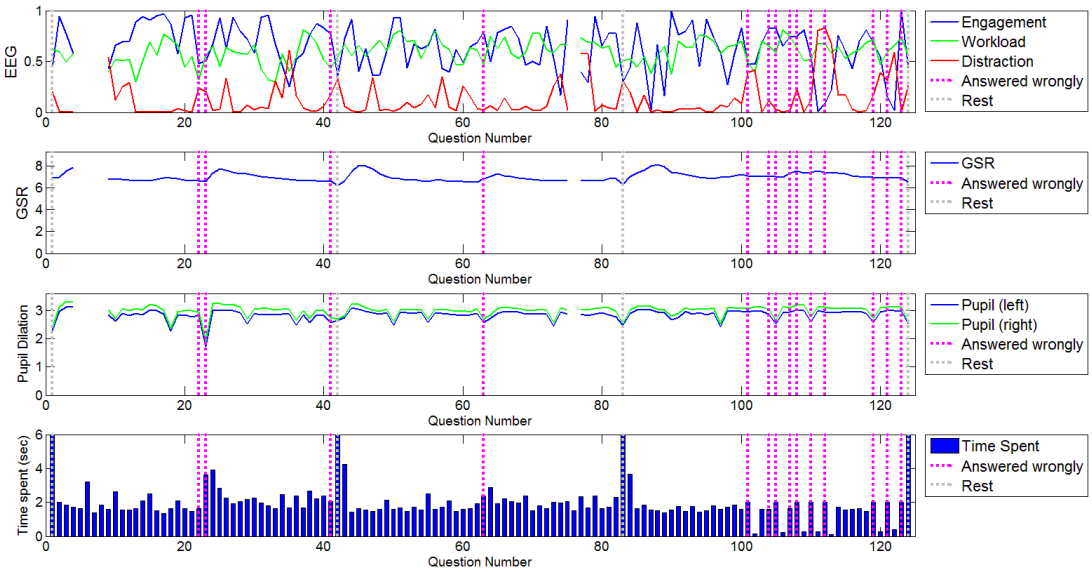


Figure 6: EEG, GSR, and pupil dilation measured from one participant (S005) while doing the test. The last raw shows the time spent on each question and rest sessions.

one. In the next section, the nine raw brain waves will be analyzed.

5 Approach to Raw Data Analysis

In addition to analyzing the GSR, eye movement and EEG metrics, two methods traditionally applied to time series data of dynamic systems were applied to the EEG data sets. First principal component analysis (PCA) was applied to identify differences across data sets acquired under different testing conditions (i.e. timer lengths). Independent component analysis (ICA) is often used in EEG analysis instead of PCA because of its ability to detect and remove artifacts in the data such as eye-blinks [8]. ICA has also been shown able to correlate brain activity to behavioral phenomena including workload [14] and drowsiness [16]. PCA is chosen here in order to investigate its ability to distinguish relatively long time histories acquired under different test conditions. Further, training data is not necessary to perform the PCA analysis presented here.

The second method is the least squares complex exponential (LSCE) parameter estimation technique. The purpose of applying this technique to the EEG data is to investigate the ability of the LSCE algorithm to identify a model for the data and then to use that model to make a one-step-ahead prediction of the data. Both the PCA and the LSCE analysis are described in the following sections.

5.1 Principal Component Analysis

Principal component analysis (PCA) uses singular value decomposition to identify principal values and vectors of a data set. The magnitude of the principal values indicate the contribution of the corresponding vector in describing the data. In this work, PCA was applied to the set of EEG time histories acquired from the bAlert system. Data from the nine data channels

(F_3,F_Z,F_4,C_3,C_Z,C_4,P_4,PO_Z,P_4) shown in Figure 2 were compiled into a single matrix, Q, as follows:

$$Q = \begin{bmatrix} F_3(t_1) & F_Z(t_1) & F_4(t_1) & C_3(t_1) & C_Z(t_1) & C_4(t_1) & P_3(t_1) & PO_Z(t_1) & P_4(t_1) \\ F_3(t_2) & F_Z(t_2) & F_4(t_2) & C_3(t_2) & C_Z(t_2) & C_4(t_2) & P_3(t_2) & PO_Z(t_2) & P_4(t_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ F_3(t_N) & F_Z(t_N) & F_4(t_N) & C_3(t_N) & C_Z(t_N) & C_4(t_N) & P_3(t_N) & PO_Z(t_N) & P_4(t_N) \end{bmatrix} \quad (1)$$

where N is the number of data points in the selected data set. Singular value decomposition was then applied to matrix Q to identify the left (v_L) and right (v_R) singular (principal) vectors and the singular (principal) values (SV):

$$[v_L, SV, v_R] = \text{svd}(Q). \quad (2)$$

The singular values were used to identify which vectors contribute most to the overall measured response. The corresponding vectors were then studied to identify the data channels which were most affected by the changing conditions of the test. This study focused on the effect of the presence of a timer while subjects answered questions, as described in Section 3. Therefore, the data from a chosen subject was partitioned into three sets. The first set contained the data from the first forty questions, in which no timer was present. The second set contained the data from the second forty questions, in which a timer of 3 seconds was imposed on the subject. Finally, the third set contained data from the last forty questions, in which a timer of 1.5 seconds was imposed. The hypothesis explored with this analysis is that changes in the singular values and vectors from the different data sets correspond to changes in the brain function caused by the change in test conditions.

5.2 Least Squares Complex Exponential Method

In dynamic systems analysis, the least squares complex exponential (LSCE) method [2] is used to identify modal parameters from time histories acquired from multiple input-output pairs. A parallel exists between dynamic data acquired using a single excitation source and multiple sensors to measure a system's response and the EEG data acquired here. In both cases, the system is excited by a source. For traditional dynamic systems, the source is typically an applied force to the system. In the experiment presented here, a subject's brain is excited by the presentation of a question. In traditional dynamics applications, the system's response will be affected by its physical properties, boundary conditions, and environmental factors. Similarly, the response of the subject's brain will be affected by his or her current mental state (levels of fatigue, engagement, etc.), the difficulty of the task, and environmental conditions. With these parallels in mind, the LSCE method was applied to the EEG data.

The purpose of the LSCE method is to identify the parameters, α_k, such that

$$\sum_{k=0}^{m} \alpha_k h_n(t_{i+k}) = 0 \quad (3)$$

where m is the model order, $h_n(t)$ is an impulse response function, n indicates the data channel of interest, and i is an arbitrary starting index within the time vector, t. Note that the index i must be selected such that $i + m$ is within the bounds of the data. Because an impulse response function could not be directly determined from the data acquired from the EEG sensors (impulse response function calculations require the input force to be measured), an operational impulse response function was used instead. As detailed in [3], the operational impulse response function of a time history, $x(t)$, is the positive lags of the autocorrelation function, R_{xx}. The

operational impulse response function will be denoted as $\hat{h}(t)$, with the $\hat{}$ distinguishing it from the standard impulse response function. After substituting in the operational impulse response function, normalizing the α_k terms by α_m, and rearranging, Equation 3 becomes

$$\alpha_0\hat{h_n}(t_{i+0}) + \alpha_1\hat{h_n}(t_{i+1}) + \cdots + \alpha_{m-1}\hat{h_n}(t_{i+m-1}) = -\hat{h_n}(t_{i+m}) \tag{4}$$

Equation 4 has $m-1$ unknowns, and, therefore, more equations are necessary in order to determine the values for the α_k's. The additional equations come from the N different data channels as well as altering the starting index i within the same data stream:

$$\begin{bmatrix} \hat{h_1}(t_0) & \hat{h_1}(t_1) & \dots & \hat{h_1}(t_{m-1}) \\ \hat{h_1}(t_1) & \hat{h_1}(t_2) & \dots & \hat{h_1}(t_m) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{h_1}(t_{m-1}) & \hat{h_1}(t_m) & \dots & \hat{h_1}(t_{2m-2}) \\ \hat{h_2}(t_0) & \hat{h_2}(t_1) & \dots & \hat{h_2}(t_{m-1}) \\ \hat{h_2}(t_1) & \hat{h_2}(t_2) & \dots & \hat{h_2}(t_m) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{h_2}(t_{m-1}) & \hat{h_2}(t_m) & \dots & \hat{h_2}(t_{2m-2}) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{h_N}(t_0) & \hat{h_N}(t_1) & \dots & \hat{h_N}(t_{m-1}) \\ \hat{h_N}(t_1) & \hat{h_N}(t_2) & \dots & \hat{h_N}(t_m) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{h_N}(t_{m-1}) & \hat{h_N}(t_m) & \dots & \hat{h_N}(t_{2m-2}) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{m-1} \end{bmatrix} = \begin{bmatrix} \hat{h_1}(t_m) \\ \hat{h_1}(t_{m+1}) \\ \vdots \\ \hat{h_1}(t_{2m+1}) \\ \hat{h_2}(t_m) \\ \hat{h_2}(t_{m+1}) \\ \vdots \\ \hat{h_2}(t_{2m+1}) \\ \vdots \\ \hat{h_N}(t_m) \\ \hat{h_N}(t_{m+1}) \\ \vdots \\ \hat{h_N}(t_{2m+1}) \end{bmatrix}. \tag{5}$$

After solving Equation 5, the α_k and $\hat{h_n}$ values were substituted into Equation 3 and the mean and standard deviation of the model error were calculated. The results of these analyses will be presented in Section 6.

6 Results

6.1 Principal Component Analysis

To explore the potential of using PCA to assess brain function, data from one subject (S005) was analyzed in depth. Future work will expand this approach to include data from multiple subjects.

Figure 7 shows the singular values for the data from the three test sections (i.e., no timer, 3 second timer, 1.5 second timer) acquired from the chosen subject. In general, the trends for all three data sets is similar. For all three data sets, the first singular value is significantly larger than the other eight indicating that the first component contributes most strongly to the overall response. The singular values of components two through five form a cluster of components that contribute less than the first, but are distinct from components six through nine. The dropoff in magnitude of the singular values after component five indicates that the contributions of components six through nine are not as significant as those of the first five. Therefore, only the first five components will be used in subsequent analyses.

Figure 8 shows the amplitudes of the first right singular vectors for the three data sets. Singular vectors determined by singular value decomposition form an orthogonal set of vectors and are analogous to operational mode shapes in traditional vibrations analysis. The singular vectors indicate the relative amplitude of response of each of the electrodes for the selected component. Changes in the amplitude, or "shape", of the vector indicates a change in the measured response for that data set. In Figure 8, the three vectors have distinct shapes. A

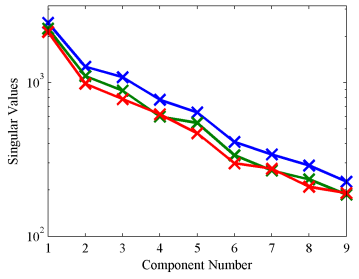


Figure 7: Singular values for the data acquired with no timer (blue), 3 second timer (green) and 1.5 second timer (red).

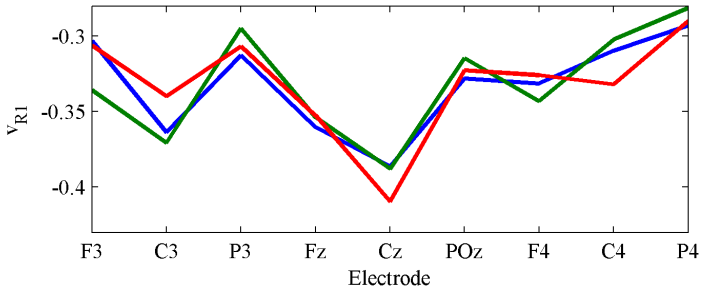


Figure 8: First right singular vectors for the data acquired with no timer (blue), 3 second timer (green) and 1.5 second timer (red).

possible explanation for these differences is that the addition of a timer affected the subject's brain function due to perceived stress caused by the time pressure.

Figure 9 shows the right singular vectors for components two through five for the three data sets. Component two (v_{R2}) is fairly consistent across data sets, but the other three components exhibit distinct differences. In general, the trends exhibited by the vectors associated with the other three components (v_{R3}, v_{R4}, v_{R5}) progress from the un-timed data set (blue), to the data set with a 3 second time (green), to the data set with the 1.5 second timer. For example, the plot of the v_{R4} vectors shows that the shape flattens out near the C_3 node as the time-to-answer is decreased. Conversely, the shape peaks more near the $C4$ node as the time-to-answer decreases. Although these trends are for a single subject, they show the potential for this type of analysis to distinguish mental states. Future work will be needed to establish statistically significant correlations between the trends observed here and the human state.

6.2 Least Squares Complex Exponential Analysis

A long term goal of this work is develop predictive capabilities that include determining when a person is susceptible to poor decision making. Model development is a key component of that effort. To determine the effectiveness of the model presented in Equation 3, Equation 5 was solved using data from the nine electrode channels and a model order of 16. Data was first partitioned by question, and a model was determined for the data from each question individually. This partitioning allowed for comparisons to be made across testing conditions, since the time-to-answer was decreased for the three sets of questions. To determine how well the model describes the data, the distributions of the mean and standard deviation of the percent error for each question were determined and are shown in Figure 10. These results show that the model is most accurate when the time pressure is greatest. Figure 10A shows that for 26 of the 40 questions presented with the shortest timer, the model error is less than 12%. Eighty-five percent of all of the data was modeled with less than 25% error. This model represents one-step-ahead predictive capability. Ongoing work is being conducted to improve both its accuracy and predictive capability.

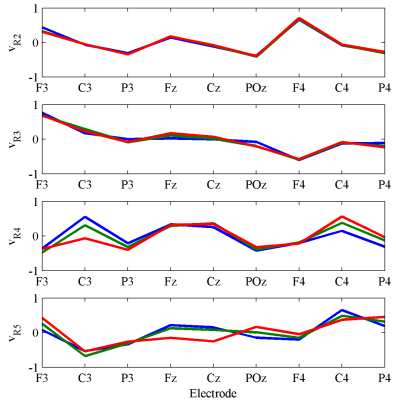


Figure 9: Second through fourth right singular vectors for the data acquired with no timer (blue), 3 second timer (green) and 1.5 second timer (red).

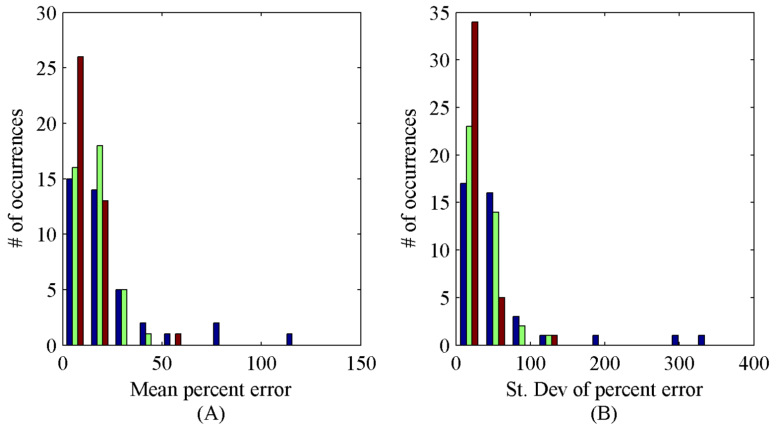


Figure 10: Distributions of (A) mean percent error and (B) standard deviation of percent error for the LSCE model calculated using data acquired with no timer (blue), 3 second timer (green) and 1.5 second timer (red).

7 Conclusions

In this paper, a dynamic data driven approach to predict operator error taking advantage of the advances in commercial grade neuro-physiological sensors is described. Some elements of the overall approach presented in this paper are: (1) a human system DDDAS is implemented as a classical dynamic data driven application system. The human operator is wired with commercial grade B-alert, Shimmer, and eye-tracking sensors. Data from these sensors are assimilated in an android body area network and streamed into the stress-intensity simulation model. The simulation model assesses the stress and intensity of stress in the operator and generates one period prediction, (2) an experimental method to collect data on human subject, and (3) analysis and prediction approach based on principal component analysis with singular value decomposition. The preliminary result of this study is promising. It was shown that under stressful conditions, the model can, with reasonable certainty, predict the occurrence of an error one period in advance. This is significant and with improvement of prediction and sensor fusion algorithms, predictions of errors will be made earlier.

Special thanks to AFOSR for their support of this research (contract number FA9550-14-1-0058).

References

[1] B-alert x10 eeg headset system. `http://www.advancedbrainmonitoring.com/xseries/x10/`. Accessed: 01-26-2015.

[2] DL Brown, RJ Allemang, Ray Zimmerman, and M Mergeay. Parameter estimation techniques for modal analysis. Technical report, SAE Technical paper, 1979.

[3] S Chauhan, R Martell, RJ Allemang, and DL Brown. Unified matrix polynomial approach for operational modal analysis. In *Proceedings of the 25th IMAC, Orlando (FL), USA*, 2007.

[4] Linda Chlan. Effectiveness of a music therapy intervention on relaxation and anxiety for patients receiving ventilatory assistance. *Heart and Lung: Journal of Acute and Critical Care*, 27(3):169–176, 1998.

[5] Jongyoon Choi, Beena Ahmed, and Ricardo Gutierrez-Osuna. Development and evaluation of an ambulatory stress monitor based on wearable sensors. *IEEE transactions on information technology in biomedicine*, 16(2):279–286, March 2012.

[6] Erik Coelingh, Andreas Eidehall, and Mattias Bengtsson. Collision warning with full auto brake and pedestrian detection-a practical example of automatic emergency braking. In *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*, pages 155–160. IEEE, 2010.

[7] P E Comalli, S Wapner, and H Werner. Interference effects of Stroop color-word test in childhood, adulthood, and aging. *The Journal of genetic psychology*, 100(1):47–53, 1962.

[8] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[9] Michael C Dorneich, Patricia May Ververs, Santosh Mathan, Stephen Whitlow, and Caroline C Hayes. Considering etiquette in the design of an adaptive system. *Journal of Cognitive Engineering and Decision Making*, page 1555343412441001, 2012.

[10] Paul Ekman, Robert W. Levenson, and Wallace V. Friesen. Autonomic Nervous System Activity Distinguishes among Emotions. *Science*, 221(4616):1208–1210, 1983.

[11] Richard Ribón Fletcher, Sharon Tam, Olufemi Omojola, Richard Redemske, and Joyce Kwan. Wearable sensor platform and mobile application for use in cognitive behavioral therapy for drug addiction and PTSD. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011:1802–1805, January 2011.

[12] Nikita Gurudath and H. Bryan Riley. Drowsy driving detection by {EEG} analysis using wavelet transform and k-means clustering. *Procedia Computer Science*, 34(0):400–409, 2014. The 9th International Conference on Future Networks and Communications (FNC'14)/The 11th International Conference on Mobile Systems and Pervasive Computing (MobiSPC'14)/Affiliated Workshops.

[13] Robert L Helmreich. Managing human error in aviation. *Scientific American*, 276(5):62–67, 1997.

[14] Christian A Kothe and Scott Makeig. Estimation of task workload from eeg data: new and current tools and perspectives. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 6547–6551. IEEE, 2011.

[15] Santosh Kumar, Wendy Nilsen, Misha Pavel, and Mani Srivastava. Mobile Health: Revolutionizing Healthcare Through Transdisciplinary Research. *Computer*, 46(1):28–35, 2013.

[16] Chin-Teng Lin, Ruei-Cheng Wu, Sheng-Fu Liang, Wen-Hung Chao, Yu-Jie Chen, and Tzyy-Ping Jung. Eeg-based drowsiness estimation for safety driving using independent component analysis. *Circuits and Systems I: Regular Papers, IEEE Transactions on*, 52(12):2726–2738, 2005.

[17] Daniel E Maurino, J Reasonson, N Johnstonton, and Rob B Lee. *Beyond aviation human factors: Safety in high technology systems*. 1995.

[18] Ian Nimmo. It's time to consider human factors in alarm management. *Chemical engineering progress*, 98(11):30–38, 2002.

[19] Jeroen GW Raaijmakers. Decision making under mental and physical stress. *TNO Institute for Human Factors*, 1990.

[20] Maja Stikic, Robin R Johnson, Daniel J Levendowski, Djordje P Popovic, Richard E Olmstead, and Chris Berka. Eeg-derived estimators of present and future cognitive performance. *Frontiers in human neuroscience*, 5, 2011.

[21] Douglas A Wiegmann and Scott A Shappell. Human error analysis of commercial aviation accidents: Application of the human factors analysis and classification system (hfacs). *Aviation, space, and environmental medicine*, 72(11):1006–1016, 2001.