Detection of temporal changes in operator functional state using statistical process control methods

Jordan Cannon[1] Pavlo A. Krokhmal[1] Yong Chen[1] Robert Murphey[2]

[1]Department of Mechanical and Industrial Engineering, University of Iowa, Iowa City, IA 52242, USA
[2]Air Force Research Lab, Munitions Directorate, Eglin AFB, FL 32542, USA

Abstract: We consider the problem of detecting temporal changes in the functional state of human subjects due to varying levels of cognitive load using real-time psychophysiological data. The proposed approach relies on monitoring several channels of electroencephalogram (EEG) and electrooculogram (EOG) signals using the methods of statistical process control. It is demonstrated that control charting methods are capable of detecting changes in psychophysiological signals that are induced by varying cognitive load with high accuracy and low false alarm rates, and are capable of accommodating subject-specific differences while being robust with respect to differences between different trials performed by the same subject.

Keywords: Statistical process control, control charts, psychophysiological data, electroencephalogram, electrooculogram.

1. Introduction and motivation

Modern systems produce a great amount of information and cues from which human operators must take action. On one hand, these complex systems can place a high demand on an operator's cognitive load, potentially overwhelming them and causing poor performance. On the other hand, some systems utilize extensive automation to accommodate their complexity; this can cause an operator to become complacent and inattentive, which can also lead to deteriorated performance (Wilson & Russell, 2003a; 2003b). An ideal human-machine interface would be one that optimizes the functional state of the operator, preventing overload while not permitting complacency, thus resulting in improved system performance.

The objective of this paper is to advance the methods for monitoring and detection of changes in operator's functional state (OFS), defined as the momentary ability of an operator to meet task demands with their cognitive resources. A high OFS indicates that an operator is vigilant and aware, with ample cognitive resources to achieve satisfactory performance. Low OFS, however, indicates a non-optimal cognitive load, either too much or too little, resulting in sub-par system performance (Wilson & Russell, 1999). In the context of this work, OFS is associated with cognitive load experienced by the operator; with this caveat in mind, we will use both terms interchangeably.

OFS is often measured indirectly, e.g., by using overt performance metrics on tasks. Another indirect measure is the subjective estimate of mental workload, where an operator narrates his/her perceived functional state while performing tasks (Wilson & Russell, 2007). However, indirect measures of OFS are often infeasible in operational settings as performance metrics are difficult to construct for highly-automated complex systems, and subjective workload estimates are often inaccurate and intrusive (Wilson & Russell, 2007; Prinzel et al., 2000; Smith et al., 2001).

OFS can be measured more directly via psychophysiological signals such as electroencephalogram (EEG) and electrooculography (EOG). Numerous studies have demonstrated these signals' ability to respond to changing cognitive load and to measure OFS (see, among others, Wilson & Fisher (1991, 1995); Gevins et al. (1997, 1998); Byrne & Parasuraman (1996)). Moreover, psychophysiological signals are continuously available and can be obtained in a non-intrusive manner, pre-requisite for their use in operational environments.

Reviews on methods for measuring OFS can be found in, e.g., Wilson & Russell (2003a, 2007) and Lotte et al. (2007). Most of these approaches use data mining and pattern recognition techniques to classify mental workload into one of several discrete categories. For instance, given an experiment with easy, medium and hard tasks, and assuming that the tasks induce varying degrees of mental workload on a subject, these methods classify which task is being performed during a given epoch of psychophysiological data. The most common classifiers are artificial neural networks (ANN) and multivariate statistical techniques such as stepwise discriminant analysis (SWDA). ANNs have proved especially effective at classifying OFS as they account for the non-linear and higher order relationships often present in EEG/EOG data; they routinely achieve classification accuracy greater than 80%.

A question that has received much less attention in the literature is that of real-time or on-line detection of *temporal changes* in OFS or, equivalently, in an operator's psychophysiological signals. Such a capability is essential for development of closed-loop adaptive control systems that will aid human operators in time-critical decision making. In this paper we present a new technique for detection of changes in OFS that uses the methods of statistical process control to monitor psychophysiological signals in real-time. We demonstrate that control chart methods are capable of detecting changes in psychophysiological signals that are induced by changes in cognitive load with an accuracy exceeding 80%, and can accommodate subject-specific differences while being robust to between-trial variations.

The paper is organized as follows. The next section describes the dataset used in this study, and Section 3 discusses the processing of the psychophysiological signals and their statistical properties. In Section 4 we outline the control charting methods used to monitor the psychophysiological data, and in Section 5 evaluate the suitability of the various control charting methods to monitoring psychophysiological signals, and analyze their effectiveness in detecting changes in OFS due to varying levels of cognitive load.

2. Data set and data processing

The dataset used in this study originated from experiments conducted at Wright-Patterson Air Force Base in 2007. Data was available for three subjects (A, E, and F), each of whom performed two 14-minute trials (denoted as A01, E02, etc.), which consisted of supervising four unmanned aerial vehicles (UAVs) on a simulated bombing mission. The simulation was designed to present the subjects with tasks of 3 levels of difficulty: low (LL), medium (ML), and high (HL). The LL was the baseline state and encompassed most of each trial. During a trial, the tasks of medium and high difficulty (ML and HL) were presented four times each, in a random order, each lasting approximately 20 seconds. Every trial began in the LL, and each period of ML or HL was followed by LL, thus allowing time for the subject to recover. The experimental design assumed that varying task loads induced corresponding levels of cognitive load on the subject. It is important to emphasize that since the tasks involved monitoring four UAVs executing a bombing mission, the tasks were very visual in nature and were expected to engage the visual processing centers of the brain; see Wilson and Russell (2007) for the complete experimental task design and details.

The data collected for each trial consisted of eight psychophysiological channels of electroencephalogram (EEG), electrooculogram (EOG), and electrocardiogram (ECG) recorded at a sampling frequency of 200 Hz. The EEG channels were recorded from five electrodes: F_7, F_z, P_z, T_5, and O_2, affixed to the subject's scalp according to the 10/20 International electrode system shown in Figure 1. Vertical and horizontal EOG data, termed VEOG and HEOG respectively, were collected for two purposes: primarily, as a measure of cognitive load, and secondly, to eliminate blink artifacts in the EEG signals. Finally, one channel of ECG was collected to measure heart rate. In the present analysis, only the EEG and EOG data were used.

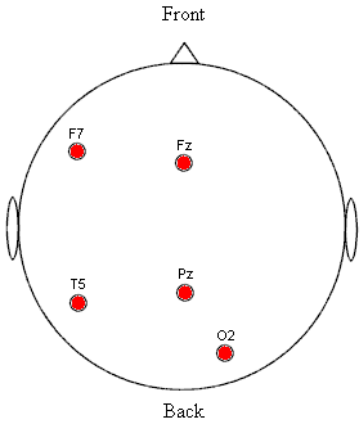

Figure 1. EEG electrode diagram

The data was pre-processed by an online adaptive filter that eliminated blink artifacts, which were especially prominent in the F_7 and F_z electrodes, whose proximity to the eye rendered them most susceptible to contamination (see Figures 1 and 2). The adaptive filter incorporates the VEOG and HEOG signals, s_t^{v} and s_t^{h}, as reference inputs to de-contaminate an EEG signal, say, s_t, for every time moment t. The corresponding artifact-free signal, $\hat{s}_t$, is obtained as

$$\hat{s}_t = s_t - \hat{s}_t^{\mathrm{v}} - \hat{s}_t^{\mathrm{h}}$$

where $\hat{s}_t^{\mathrm{v}}$ and $\hat{s}_t^{\mathrm{h}}$ are the filtered VEOG and HEOG reference signals, respectively:

$$\hat{s}_t^{\mathrm{v}} = \sum_{m=1}^{M} h_m^{\mathrm{v}} s_{t+1-m}^{\mathrm{v}}\,, \quad \hat{s}_t^{\mathrm{h}} = \sum_{m=1}^{M} h_m^{\mathrm{h}} s_{t+1-m}^{\mathrm{h}}$$

and the filter coefficients $h_m^{\mathrm{v}}$and h_m^{h} are obtained and updated by minimizing the weighted squares

$$\min \sum_{i=M}^{n} \lambda^{n-i} \hat{s}_i^2$$

where λ is the "forgetting factor", for each time instant n via a recursive least-squares (RLS) algorithm; see He et al. (2007) for complete details.

Once eye blink artifacts were removed, the psychophysiological signals were transformed to the frequency domain using Fast Fourier Transform (FFT) (see, e.g., Phillips et al., 2007) for every epoch, usually three to five seconds. Thus, a frequency spectrum was created for each epoch, and the powers of particular wavebands, e.g. 9-13 Hz, known as *alpha*, *beta*, etc., were

computed as averages of the frequency's power. These powers have been shown to correlate with changes in cognitive load (see Table 1); for instance, alpha waveband power, 9–13 Hz, increases with relaxation while theta waveband power, 4–8 Hz, generally decreases with relaxation (Smith et al., 2001). From this point forward, the waveband powers of psychophysiological signals recorded at particular EEG/EOG electrodes are referred to as *features*.

Table 1. Waveband frequencies of EEG data and their interpretation

Waveband	Frequency (Hz)	Interpretation
Delta	2-4	Slow wave sleep
Theta	4-8	Arousal
Alpha	9-13	Relaxation
Beta	14-32	Active, alert, working
Gamma	33-43	Cognitive and/or motor function

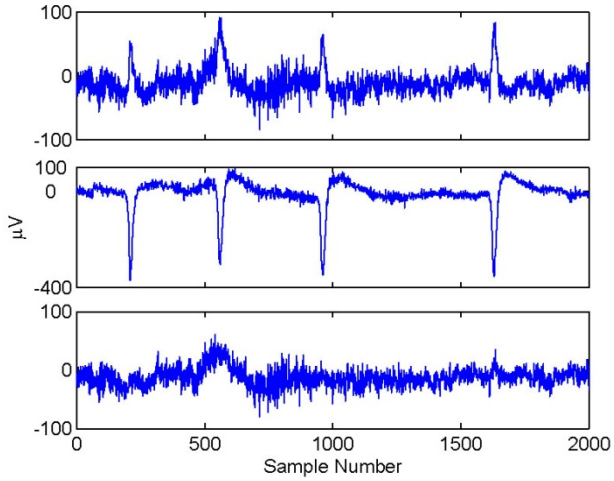

Figure 2. Unfiltered F_7 signal with blink artifacts (top); VEOG signal (middle); filtered F_7 signal (bottom)

3. Characterizing the post-processed data

Empirical analysis was conducted on several psychophysiological features to characterize their properties and behavior with respect to subject and task load. For this analysis, the features were computed in three-second epochs. Although it is common to find one-second epochs in the literature, the corresponding data often exhibit highly volatile behavior; moreover, short epochs induce autocorrelation in the features. And while data aggregated over longer epochs are more stable, excessively long time epochs increase system latency. Three-second epochs were found to provide the best balance between system latency and accuracy.

The features chosen for this analysis are traditional ones used in OFS classification methods, and include the five wavebands listed in Table 1, which have been shown to correlate strongly with cognitive load.

Since all seven electrodes record the same five wavebands for every epoch, it is natural to expect a certain degree of redundancy in the data. In particular, correlation patterns exist for features recorded at the same node (intra-node correlation) as well as for features recorded at

different nodes (inter-node correlation). The reason for such a redundancy is that brain waves behave differently when emitted from different regions of the brain. For instance, alpha power is greatest in parietal regions, e.g. P_z, whereas theta power dominates the frontal region, e.g. F_z. Table 2 illustrates the intra-node correlation of features recorded at VEOG of F01.

In addition, many features exhibited autocorrelation, which was observed especially consistently in beta and gamma features. Figure 3 displays the sample autocorrelation function for a beta feature measured at T_5 of E01. At lag 1, the autocorrelation is the strongest and then exponentially decreases before oscillating around 0, which is characteristic of a first-order, autoregressive time series (AR(1)).

Table 2. Intra-node correlation measured at VEOG of F01

	Delta	**Theta**	**Alpha**	**Beta**	**Gamma**
Delta	1.000	0.922	0.939	0.852	0.582
Theta	0.922	1.000	0.949	0.864	0.655
Alpha	0.939	0.949	1.000	0.914	0.663
Beta	0.852	0.864	0.914	1.000	0.853
Gamma	0.582	0.655	0.663	0.853	1.000

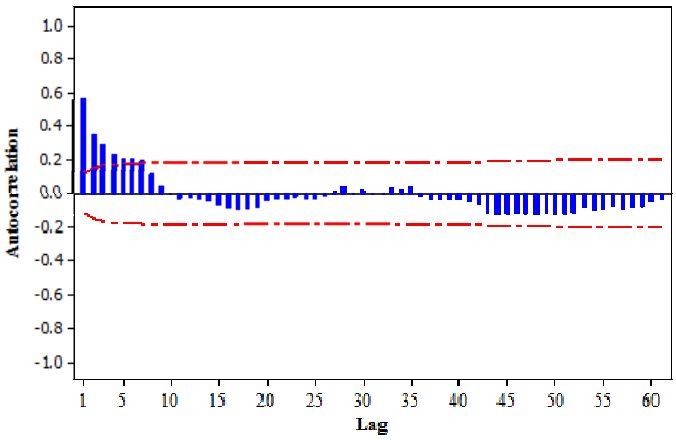

Figure 3. Sample autocorrelation function for T_5 beta of E01
(the dotted lines denote .05 significance level)

Table 3 summarizes feature autocorrelations across subjects. For each waveband, a percentage is displayed representing the proportion of electrodes where the feature exhibited significant autocorrelation. As one can see, theta and alpha were the wavebands with the least autocorrelated features across subjects, and beta and gamma wavebands gave rise to features that were autocorrelated regardless of subject or electrode.

The distributions of theta, alpha, and beta waveband powers computed across all electrodes for every epoch typically exhibited a strong positive skew, and were found to be non-normal at a .05 significance level. In addition, it was observed that none of the features exhibited significant non-stationarity in the mean.

Table 3. Autocorrelations across subjects as proportion of electrodes where the feature exhibited significant autocorrelation

	A01	E01	F01	Average
Delta	0.00%	57.14%	71.43%	42.86%
Theta	14.29%	28.57%	42.86%	28.57%
Alpha	42.86%	28.57%	42.86%	38.10%
Beta	100.00%	100.00%	85.71%	95.24%
Gamma	100.00%	100.00%	100.00%	100.00%

4. Control charts for monitoring psychophysiological data

As mentioned, pattern recognition classifiers are frequently employed to measure OFS in real time. While many of these methods are targeted at classification of OFS into several discrete categories for at a given time epoch, relatively little attention has been given to the problem of dynamically detecting "changes" in OFS. In the present endeavor, the statistical process control methods based on control charts are proposed as an approach for detection of real-time changes in OFS.

In this section, different control charts are presented that will subsequently be evaluated for their ability to detect changes in psychophysiological feature(s) computed in three-second epochs; recall that these changes are induced by changing task load and, correspondingly, cognitive load. For example, a single EEG feature can be controlled by a univariate control chart whose centerline is the feature mean and control limits are calibrated to be breached during HL and ML tasks, while allowing no more than a prescribed number of false alarms. Thus, it is implicitly assumed that the feature mean will change with increasing task load.

The present control chart study used subgroups of size one, i.e., an epoch constituted one data point. It was deemed inappropriate to form subgroups any greater than size one, e.g., averaging a feature over several epochs. This is because psychophysiological data is collected continuously, as opposed to periodic sampling. As such, size one subgroups require specially designed control charts which are discussed in the following sections (Montgomery, 2009).

In order to be amenable to control chart monitoring, the psychophysiological data must comply with a number of assumptions. An often used assumption is that of normality; if this assumption is violated, type I and type II errors will occur at a greater rate than is stipulated by the control limits. In this context, a type II error indicates a failure to detect an out-of-control observation. Some control charts, such as moving average charts, are very robust to violations of normality. Many control charts techniques also assume that the data are independently distributed. A common violation of this assumption comes in the form of autocorrelation, and special control chart methods have been developed in the literature to address this issue. Also, the features are assumed to be stationary, meaning their in-control mean does not change with time (Montgomery, 2009).

The properties of the psychophysiological features under consideration were discussed in Section 3. As demonstrated, some features violated the assumptions of independence and normality, whereas all the features evaluated were stationary.

4.1 Univariate Control Charts

First, we discuss univariate control charts for individual measurements, i.e., for subgroups of size one, including those which accommodate autocorrelated or non-normal data. In the next section, these univariate charts are evaluated as to their ability to detect changes in OFS.

The exponentially-weighted moving average (EWMA) chart and the Shewart individuals chart are commonly combined to form a robust univariate control charting scheme. The EWMA chart is attractive in the context of this study because it is robust to non-normality and can detect small shifts in the mean. On the other hand, the Shewart individuals chart complements the EWMA with its ability to detect large shifts in the mean. The EWMA is susceptible to the "inertia" problem: it is slow to detect changes which occur in the opposite direction from points previously plotted. In contrast, the Shewart chart is impervious to the "inertia" problem and can detect sudden changes, such as outlier points. Outliers can skew the EWMA for several time periods, thus identifying these points via the Shewart chart provides insight in the behavior of the EWMA.

The upper and lower control limits, denoted as UCL and LCL respectively, and the center line, denoted as CL, for the EWMA are given by

$$\begin{matrix} UCL \\ LCL \end{matrix} = \bar{x} \pm L\sigma\sqrt{\frac{\lambda}{(2-\lambda)}[1-(1-\lambda)^{2i}]}, \quad CL = \bar{x}$$

where $\bar{x}$ is the in-control mean of the feature, σ is the sample standard deviation of the feature, λ is a smoothing constant such that $0 \leq \lambda \leq 1$, i is the time epoch, and L is a factor that determines the width of the control limits.

The UCL, LCL, and CL for the Shewart individuals chart have the form

$$\begin{matrix} UCL \\ LCL \end{matrix} = \bar{x} \pm 3\frac{\overline{MR}}{d_2}, \qquad CL = \bar{x}$$

where d_2 is a constant determined from available tables and $\overline{MR}$ is the average moving range. Moving range is calculated from two consecutive epochs as

$$MR = |x_i - x_{i-1}|$$

where x_i represents the feature's magnitude at epoch i.

On the EWMA chart, instead of plotting the actual feature observed, i.e. x_i, a smoothed version, z_i, is plotted for each epoch i, where

$$z_i = \lambda x_i + (1-\lambda)z_{i-1} \tag{1}$$

It is customary to initilize z_0 to the in-control mean of the feature.

The autocorrelated psychophysiological features typically exhibited first-order autoregressive behavior

$$\hat{x}_i = (1-\phi)\mu + \phi x_{i-1} + \epsilon_i \tag{2}$$

where $\hat{x}_i$ is the forecast of the feature made from x_{i-1}, ϕ is a parameter such that $|\phi| < 1$, μ is the expected value of x_i, and ϵ_i are iid error terms. This model is often recognized as the

regression of x_i on x_{i-1} (Montgomery et al., 2008). Time series models similar to (2) can be employed to remove the autocorrelation from affected psychophysiological features, allowing them to be monitored in traditional control charts (Montgomery & Mastrangelo, 1991). To accomplish this, an appropriate time series model is fit to the feature data. This model is then used as a one-step-ahead forecast for the feature, with the residuals (forecast errors) for each epoch plotted on the control chart. Assuming that the feature data was modeled correctly, these residuals will be independent and non-autocorrelated. In addition, residuals from the LL will be different than residuals from the ML or HL data.

Instead of removing autocorrelation, one may simply use the EWMA statistic, as computed in (1), in a moving-centerline exponentially-weighted moving average chart (MCEWMA). This method provides an optimal one-step-ahead forecast for a psychophysiological feature which exhibits the behavior of an integrated moving average time series with first order parameters (IMA(1,1)); this model is defined by

$$\hat{x}_i = x_{i-1} + \epsilon_i - \theta\epsilon_{i-1} \tag{3}$$

where $\theta = 1 - \lambda$. The value of parameter λ is chosen to minimize the sum of squares of the one-step-ahead forecast errors, e_i, defined by

$$e_i = x_i - \hat{x}_i$$

When the feature can be assumed to follow the IMA(1,1) model and λ is selected as above, the forecast in (3) becomes equivalent to the forecast of x_i from z_{i-1}, where z_{i-1} is the EWMA statistic at epoch $i-1$ and the "moving" centerline for the MCEWMA chart. As before, the UCL and LCL were calibrated so as to be breached during the HL and ML periods. The UCL and LCL for the MCEWMA chart for each epoch i are given by

$$UCL_{i+1} = z_i + L\sigma_p(i), \qquad LCL_{i+1} = z_i - L\sigma_p(i)$$

with $\sigma_p(i)$ being the smoothed standard deviation estimated for each epoch as

$$\hat{\sigma}_p^2(i) = \alpha e_i^2 + (1-\alpha)\hat{\sigma}_p^2(i-1)$$

where α is a smoothing constant such that $0 \leq \alpha \leq 1$ and $\hat{\sigma}_p^2(0)$ is estimated by dividing the sum of the squared forecast errors used to derive λ, by n, where n is the number of observations.

4.2 Multivariate Control Charts

Mastrangelo et al. (1996) indicate that multivariate charts monitoring several related variables simultaneously can be superior to univariate charts monitoring those variables separately. Multivariate control charts account for the covariance structure between features to reduce type II errors, while still conforming to an acceptable type I error level. In contrast, type I errors increase when using multiple univariate control charts. Multivariate charts require assumptions similar to those of univariate charts, e.g., multivariate normality, independence, and stationarity. Similarly to the univariate case, a number of multivariate control chart techniques exist in the literature that can accommodate violations of these assumptions. Some of these methods are discussed next and evaluated in the following section.

One of the most popular multivariate control charts is the Hotelling-T^2 chart, a multivariate analogue of the Shewart individuals chart. This chart plots a T_i^2 statistic

(4)

$$T_i^2 = (\mathbf{x}_i - \bar{\mathbf{x}})'\mathbf{S}^{-1}(\mathbf{x}_i - \bar{\mathbf{x}})$$

against a single UCL for every epoch i, where $\mathbf{x}_i \in \mathbb{R}^p$ is a vector of a realization of p features, $\bar{\mathbf{x}} \in \mathbb{R}^p$ is vector of the feature's in-control means, and $\mathbf{S}$ is an estimate of the features' covariance matrix,

$$\mathbf{S} = \frac{1}{2}\frac{\mathbf{V}'\mathbf{V}}{(n-1)}, \quad \text{where} \quad \mathbf{V} = \begin{bmatrix} \mathbf{v}_1' \\ \mathbf{v}_2' \\ \vdots \\ \mathbf{v}_{n-1}' \end{bmatrix} \tag{5}$$

and $\mathbf{v}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$ for $i = 1,2,\dots,n-1$. The UCL for the Hotelling-T^2 chart is determined using the Beta distribution function (β) as

$$UCL = \frac{(n-1)^2}{n}\beta_{\alpha,\frac{p}{2},\frac{n-p-1}{2}}$$

where n is the number of epochs and α is chosen to define an upper quantile in the Beta distribution (Montgomery, 2009).

The multivariate exponentially-weighted moving average control chart (MEWMA) is an extension of the univariate EWMA chart, and is generally more sensitive to small shifts in the mean vector comparing to the Hotelling-T^2 chart. Also, it is robust to violations of normality, which are frequently found in psychophysiological data. The MEWMA chart monitors $\mathbf{z}_i$, a smoothed version of $\mathbf{x}_i$, computed by

$$\mathbf{z}_i = \lambda\mathbf{x}_i + (1-\lambda)\mathbf{z}_{i-1}$$

where $\mathbf{z}_0$ is initialized to zero, assuming the features have zero mean. For each epoch, $\mathbf{z}_i$ is used to compute the ${T_i}^2$ statistic by

$$T_i^2 = \mathbf{z}_i'\,\boldsymbol{\Sigma}_{\mathbf{z}_i}^{-1}\,\mathbf{z}_i, \qquad \text{where} \qquad \boldsymbol{\Sigma}_{\mathbf{z}_i} = \frac{\lambda}{\lambda-1}\left[1-(1-\lambda)^{2i}\right]\mathbf{S} \tag{6}$$

and $\mathbf{S}$ is the estimate of the covariance matrix (5). There is a single UCL for the MEWMA chart which is set to h_4, specified by tables found in the literature (Lowry et al., 1992).

Jackson (1980) has suggested that instead of using the original p features to comprise the ${T_i}^2$ statistic, one may use a subset of their principle components. Principle component analysis (PCA) reduces the dimensionality of multiple features into a subspace of orthogonal components. These components are chosen to capture most of the variability exhibited by the original features. The principle components are computed by estimating the covariance matrix of the features,

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$$

where $\mathbf{x}_i \in \mathbb{R}^p$ is a vector of a realization of the p original features, and it is assumed that the features have been standardized to zero mean and unit variance. By pre- and post-multiplying $\mathbf{S}$ by an orthonormal matrix $\mathbf{U}$, a diagonal matrix $\mathbf{L}$ is obtained whose diagonal elements are equal to the eigenvalues of $\mathbf{S}$

$$\mathbf{U}'\mathbf{S}\mathbf{U} = \mathbf{L},$$

and the columns of $\mathbf{U}$, denoted as $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_p$, are the associated eigenvectors. The principle components are computed from each $\boldsymbol{x}_i$ by

$$\mathbf{p}_i = \mathbf{U}'\mathbf{x}_i$$

where $\mathbf{p}_i \in \mathbb{R}^p$ is the vector of principle components. The *k*-th principle component is then $p_{ik} = \mathbf{u}_k'\mathbf{x}_i$ or, in expanded form, $p_{ik} = u_{1i}x_{1i} + u_{2i}x_{2i} + \cdots + u_{pi}x_{pi}$. In essence, PCA transforms *p* correlated features into *p* uncorrelated principle components for every epoch, such that each component is a linear transformation of the original features.

Dimensionality reduction in PCA is achieved by retaining the subset of the principle components that account for the most variability in the original data, i.e., the components that correspond to largest eigenvalues of $\mathbf{S}$. In this study, m out of p principal components $p_1, p_2, \ldots, p_m$ were selected such that at least 80% of the original data's variability was preserved, where the variability attributed to i-th principal component is estimated as

$$\frac{l_i}{l_1 + l_2 + \cdots + l_p} \tag{7}$$

where l_i is the eigenvalue associated with i-th principal component. These selected principle components are represented by vector $\mathbf{w}_i \in \mathbb{R}^m$, which is computed for every epoch by

$$\mathbf{w}_i = \mathbf{U}_k'\mathbf{x}_i$$

where $\mathbf{U}_k$ is a *p* x *m* matrix, whose columns, $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m$, are the eigenvectors associated with the *m* selected principle components. The vector $\mathbf{w}_i$ is then used to compute the T_i^2 statistic monitored in both the Hotelling-T^2 and MEWMA charts, where T_i^2 is computed in (4) and (6), respectively. In these equations, $\mathbf{S}$ is just the diagonal matrix $\mathbf{L}_k$ of the eigenvalues $l_1, l_2, \ldots, l_m$, and $\mathbf{x}_i$ is substituted with $\mathbf{w}_i$ for every epoch (Mastrangelo et al., 1996).

There were 35 features that could possibly be monitored for each subject (5 wavebands for each of 7 electrodes). However, more than half of these features were autocorrelated, which violated multivariate control chart assumptions. Thus, only the non-autocorrelated subset, typically 10-17 features for each subject, was submitted to PCA analysis.

One drawback to monitoring the subset of principle components is the inability to detect shifts which are orthogonal to the subspace defined by the components. The severity of this problem depends on the process, and in some cases, can be mitigated by monitoring the residuals of predicting $\mathbf{x}_i$ from $\mathbf{w}_i$ (Lowry & Montgomery, 1995; Mastrangelo et al., 1996).

5. Detecting changes in psychophysiological data using control charts

In this section we discuss application of the described above control chart methods to detecting changes in operator functional state due to changes in cognitive load. The following procedure was employed. First, we conducted a preliminary comparative evaluation of the effectiveness of different control chart techniques in detecting changes in OFS induced by varying levels of cognitive loads as described in Section 2 so as to determine the best-performing control charting techniques as well as the features (wavebands) that were best suited for

monitoring with control chart methods. For these purposes, we used only the data from the first trial of each subject (i.e., A01, E01, and F01), see Section 5.1. This was done in order to alleviate potential issues associated with the relatively small data set available for this study; in particular, to avoid potential "overfitting", or "fitting the method to the data".

Then, the best-performing control charting method was further evaluated with respect to its ability to detect OFS changes using data from both trials of each subject (Section 5.2) in the "train-and-test" framework common to data mining literature. Namely, the standard cross-validation method was used, when one trial is used as a "training" data for tuning the parameters so as to achieve the best performance, and the other trial is used for testing the method. Then, the roles of the two trials as testing and training data sets are reversed and the procedure is repeated.

5.1 A comparative study of control chart methods on psychophysiological data

As a part of the preliminary evaluation of the control charts described above, an analysis was conducted to determine how the means of common psychophysiological features changed across task loads. From a control charting perspective, it is desirable for the changes in the mean to be large, and thus easily detectable. For each subject, features from the theta, alpha, and beta wavebands, recorded from all electrodes, were analyzed, and the results are presented in Table 4, where the mean shifts are reported in units of standard deviations of the respective features. The average shift from the LL to the ML across all subjects is .45 standard deviations, and the average shift from the LL to the HL is .57 standard deviations. In a control charting context, such magnitudes of shifts are small and their detection will prove difficult; it can also be expected that changes in OFS that result from the onset of tasks will be detected with substantial delays.

Table 4. Mean shifts in features across task loads

	A			**E**			**F**		
	Theta	Alpha	Beta	Theta	Alpha	Beta	Theta	Alpha	Beta
LL to ML	0.183	0.845	0.064	0.332	0.307	0.506	0.244	0.551	1.021
LL to HL	0.025	1.631	0.050	0.489	1.107	0.294	0.259	0.471	0.825

The parameters for the control chart methods evaluated next were selected based on common recommendations in the literature (e.g., Montgomery, 2009). Particularly, the centerlines for all control charts were computed using the LL task data of the trial, and the control limits were calibrated by following the generally accepted guidelines (e.g., the value of L in EWMA control chart was chosen between 1.65 and 2.24) so as to achieve the highest detection rates subject to constraint on the number of false alarms per trial. The "target" number of false alarms in this study was set at two or fewer, but in some cases we allowed for more false alarms in order to achieve detection rates not close to zero.

Univariate control charts on non-autocorrelated features Here we report the results obtained by using univariate control charts for monitoring non-autocorrelated psychophysiological features. The features monitored were chosen based on research that demonstrated their responsiveness to change in OFS. For instance, Smith et al. (2001) indicate that theta power measured from frontal-midline locations, e.g. F_z, increased with increasing cognitive load, whereas alpha power measured from parietal locations, e.g. P_z, routinely decreased with increasing cognitive load. These and other features responsive to OFS were subsequently analyzed in Gevins et al. (1998) and Wilson & Russell (2003b).

Table 5 summarizes the results from the univariate EWMA-Shewart control chart analysis. The table reports the proportion of the 8 tasks which were detected for each feature and

trial. Subject A was the easiest to monitor, with over 70% of the tasks detected when their results are averaged across all features. Subject F, on the other hand, proved the most difficult to monitor with only half of their tasks detected on average. The features that performed best across subjects were: VEOG theta and P_z alpha. Overall, there was a feature for each subject that could detect at least 75% of the tasks. Subjects A and E each had one or more features which could detect at least 87.5% of the tasks. Finally, the average number of false alarms for the results reported in Table 5 was less than one per trial. Refer to Appendix A for the complete results on false alarms of this and all subsequent control chart analyses. The system "latency", defined as the average number of time epochs elapsed from the onset of a ML or HL task until its detection, given that the detection was made[1], ranged from 2.83 for F_z alpha of A01 to 6.00 for O_2 alpha of F01, with an average of 4.09 epochs, or 12.27 seconds across all subjects and features. The average duration of "out-of-control" ML or HL tasks is approximately 6.67 epochs, or 20 seconds, and the average duration of "in-control" LL tasks is 23.67 epochs, or 71 seconds. Similar latency figures have been observed for other control chart methods discussed below.

Table 5. Detection rates of non-autocorrelated univariate control charts

	VEOG Theta	**F_z Theta**	**F_z Alpha**	**F_7 Theta**	**O_2 Alpha**	**P_z Alpha**	**Average**
A01	0.875	0.500	0.750	0.750	0.625	0.750	**0.708**
E01	1.000	0.000	0.625	0.625	0.500	0.500	**0.542**
F01	0.375	0.500	0.375	0.500	0.500	0.750	**0.500**
Average	**0.750**	**0.333**	**0.583**	**0.625**	**0.542**	**0.667**	

An illustration of EWMA-Shewart chart is shown in Figure B.1 (see Appendix B). In this chart, as well as in other similar charts, the green, yellow, and red vertical lines denote the onset of LL, ML, and HL tasks, correspondingly. The baseline LL task is the starting condition in each trial and occurs everywhere in between the ML and HL conditions.

Univariate control charts on autocorrelated features Next, we examined monitoring of autocorrelated psychophysiological features with univariate control charts. Table 6 summarizes the results of monitoring various beta features in the MCEWMA chart. These features were chosen based on their ability to measure beta power and their responsiveness to changing OFS. More specifically, beta power was found to be strong in the parietal region of the brain, e.g. P_z and T_5, and the F_z feature was weighted heavily in neural network methods (Gevins et al., 1998; Wilson & Russell, 2003b). T_5 beta exhibited the most consistent and accurate performance among subjects, detecting nearly 60% of the tasks on average; coincidentally, this feature also exhibited the most autocorrelation. Overall, there was a beta feature for each subject that could detect 62.5% of the tasks when monitored by the MCEWMA chart. Neither theta nor alpha features ever exhibited strong enough autocorrelation for the MCEWMA chart to be effective, thus only beta features were monitored. On average, this method resulted in over two false alarms per trial. Figure B.2 illustrates an MCEWMA chart monitoring F_7 beta of E01.

[1] If no detections were made during a trial (see, e.g. F_z theta of E01 in Table 5), latency was undefined.

Table 6. Detection rates of MCEWMA control charts

	F_z Beta	P_z Beta	T_5 Beta	Average
A01	0.625	0.625	0.500	**0.583**
E01	0.625	0.625	0.625	**0.625**
F01	0.625	0.625	0.625	**0.625**
Average	**0.625**	**0.625**	**0.583**	

The next results concern control charts which plotted the residuals from time series models. Table 7 summarizes the results of monitoring the time series residuals of various beta features in EWMA-Shewart charts. Notice that these features are identical to those in Table 6, except for the inclusion of HEOG beta; this feature exhibited first-order moving average behavior (MA(1)) which was only conducive to time series modeling, not an MCEWMA chart. The results demonstrate that again, T_5 beta was the feature most conducive to monitoring in this setting, detecting nearly 60% of the tasks when averaged across subjects. If the best performing feature is evaluated for each subject, then 62.5% of the tasks were detected on average. In addition, there were 1.8 false alarms per trial on average. Quantitatively, the performance of control charting residuals was marginally better than the MCEWMA chart, but at the expense of extra effort necessary to fit a time series model. An illustration of an EWMA-Shewart chart monitoring the residuals of an AR(1) model fitted to F_z beta of A01 is given in Figure B.3.

Table 7. Detection rates of time series residuals control charts

	HEOG Beta	F_z Beta	P_z Beta	T_5 Beta	Average
A01	0.375	0.750	0.625	0.625	**0.594**
E01	0.500	0.500	0.500	0.500	**0.500**
F01	0.375	0.500	0.500	0.625	**0.500**
Average	**0.417**	**0.583**	**0.542**	**0.583**	

Multivariate Control Charts Multivariate control charts monitoring several correlated signals were expected to deliver results superior to those obtained by monitoring the same features separately with univariate charts. In our studies, it was observed that monitoring five or less features with multivariate charts led to better performance in terms of detection rates. Consequently, the following analysis was done on two or three features which performed best for each subject in the univariate analysis (see Table 5). These sets of features were monitored in Hotelling-T^2 and MEWMA control charts.

Table 8. Detection rates of multivariate control charts

	Hotelling-T^2 (Select)	MEWMA (Select)	Hotelling-T^2 (PCA)	MEWMA (PCA)	Average
A01	0.625	0.750	0.625	0.750	**0.688**
E01	0.500	0.500	0.500	0.500	**0.500**
F01	0.500	0.625	0.375	0.500	**0.500**
Average	**0.542**	**0.625**	**0.500**	**0.583**	

The results for the Hotelling-T^2 and MEWMA control charts on selected features are contained in the first two columns of Table 8. The MEWMA chart performed better on average than the Hotelling-T^2 chart, detecting 62.5% of the tasks compared to 54.2%, respectively. Finally, the MEWMA chart incurred 2.33 false alarms on average and double that rate for the Hotelling-T^2 chart. These false alarm rates were much worse than their respective univariate analogues.

Next, the subsets of principle components of the non-autocorrelated features were monitored in Hotelling-T^2 and MEWMA charts for each subject. This evaluation was done for a theoretical improvement over monitoring a few subject-specific features. The results from monitoring a subset of the principle components are contained in the last two columns of Table 8. The MEWMA chart performed better on average than the Hotelling-T^2 chart, detecting nearly 60% of the tasks compared to 50%, respectively. The false alarm rates were comparable to monitoring a few subject-specific features. Overall, monitoring a subset of principle components did not yield an improvement over monitoring a few subject-specific features.

Figures B.4 and B.5 illustrate Hotelling-T^2 and MEWMA charts monitoring selected features, and Figures B.6 and B.7 represent the same charts monitoring subsets of 8 features out of, respectively, 17 and 13 non-autocorrelated features of A01 and E01.

The results of the comparative studies demonstrate that control charts can detect real-time change in OFS with varying degrees of accuracy. The most accurate control charting method was to monitor a single subject-specific feature in a univariate EWMA-Shewart chart. Averaged across all subjects, this method resulted in a 75% detection rate with less than 1 false alarm on average; however, this method cannot accommodate autocorrelated features, such as beta and gamma wavebands, and it is unable to monitor multiple features simultaneously. The first issue can be remedied by control charts such as the MCEWMA. But, despite MCEWMA's superiority over other autocorrelated methods, it only detected 62.5% of the changes in OFS on average, with at least 2.33 false alarms per trial on average.

The evaluation of multivariate control charts did not confirm their advantages over univariate control charts when applied to the data at hand. The best multivariate control chart approach was to monitor several subject-specific features in an MEWMA chart. The Hotelling-T^2 chart monitoring the same subject-specific features was less accurate and had twice the false alarms. When averaged across subjects, the best multivariate control charting method detected 62.5% of the changes in OFS, with 2.33 false alarms per trial on average.

Contrary to expectations, monitoring the principle components was not the superior control charting method. In the present analysis, principle components were generated from the subset of non-autocorrelated features. This was done to prevent the principle components from being autocorrelated, and thus, incapable of being monitored in traditional multivariate control charts. This constraint, however, routinely excluded beta and gamma features, which were consistently autocorrelated. A factor pattern analysis conducted on principal components (see Cannon, 2009) has revealed that these higher frequencies were the most prominent features in a subject's brain activity, and thus excluding these features may have been detrimental to the control chart's performance.

The control charts which performed the best shared several common characteristics. First, they smoothed the features, in this case, via an EWMA. It is well known that smoothing time series data can help separate the signal from the noise, reveal important trends, and render non-normal data more normal (Montgomery, 2008). Psychophysiological data is typically quite noisy and volatile; consequently, smoothing filters are common in psychophysiological research, most notably, the simple moving average with overlap and the Hanning window (Freeman et al., 1999; Wilson & Russell, 2007). A second characteristic common to the best performing control charts

was that they monitored subject-specific features. Each subject responded differently to varying task load, thus no feature or subset of features were optimal with respect to all subjects.

5.2. Online detection of changes in OFS using control charts

According to the results reported in the preceding subsection, the EWMA control chart monitoring certain subject-specific features produced the highest detection rates with the lowest number of false alarms, and thus can be considered as the most promising candidate for real-time detection of changes in OFS. In this section we conduct a more detailed investigation of this method, particularly its amenability to be "trained" in order to achieve improved results on a per-subject basis. In contrast to the preceding comparative studies, where all control charting methods were evaluated on a single trial for each subject, and, moreover, for each given control chart technique the same parameters were used across all subjects, in this section we allow for subject-specific "training" of the control chart using the data of one trial, and then testing the obtained control chart on the data of other trial. Cross-validation is then performed by switching the training and testing trials.

In particular, the training phase consisted of consolidating the data corresponding to no-load (LL) periods in one trial and computing the initial control limits and EWMA statistic (the centerline) for testing of the subject's other trial. The rationale behind this is that the centerline is the normal state, and the control limits are calibrated to achieve low false alarm rates and high detection rates for out-of-control conditions associated with the states of elevated task load. Specifically, the values of L were taken at 1.65, 1.96, and 2.24, corresponding to significance levels of .1, .05, and .025, and λ was selected as either .1, .2, or .3. The "best" combination of parameters was determined as such that delivered the highest detection rates with no more than 2 false alarms per trial. Then, during the testing stage, the data of the testing trial was monitored using control charts with these initialized parameters.

For each training/testing dataset and combination of parameters, the proportion of correctly detected tasks and the number of false alarms were determined. With respect to this, several comments are in order. First, we only consider the lower control limit since the phsychophysiological features being tested are known to decrease with task load. Secondly, when the control limit is breached, we consider consecutive out-of-control points as one event rather than individual events. For example, if 5 consecutive points signal out-of-control during a task load, we consider this as single detection, and, similarly, if several consecutive points go out-of-control as false alarms, it is considered as one false alarm. Additionally, we ignore "false alarms" that may occur at the very end of a baseline LL task period (i.e., immediately before onset of a ML or HL task): there are indications that, due to the cyclic nature of the trials, the subjects may have anticipated the approaching start of an elevated-difficulty task, which initiated changes in their psychophysiological signals. Such effects have been observed in our previous studies (Cannon et al., 2010).

An illustration of EWMA-Shewart chart tested on trial A02 is presented in Figure 4, where the vertical lines denote onsets of LL, ML, and HL tasks, and the results of cross-validation testing are presented in Table 9. First, it was determined that the psychophysiological features most suitable for monitoring with a EWMA control charts are different for subjects A and E, and F. Namely, for subjects A and E the best results are obtained using the theta of VEOG (recall that the tasks performed by the subjects were heavily visual), whereas for subject F the best results are obtained using P_z alpha. Moreover, with optimal values of parameters EWMA control charts demonstrate very good detection rates, exceeding 80% for subjects A and E with no more than 2 false alarms per trial, whereas for subject F the average detection rate is markedly worse, at about 25%.

These findings indicate that subject F (or, rather, his/her psychophysiological data) can be considered "different" in some sense from those of subjects A and E. Such a conclusion is well in

accord with the previous studies (Yu, 2009, Cannon, 2009, Cannon et al., 2010), where subject F's data proved to be the most challenging to various machine learning methods.

The last column of Table 9 reports the average number of epochs elapsed before a ML or HL task is detected, given that the this task is correctly detected. Note that average duration of a HL or ML task is 6.67 epochs, or 20 seconds.

It can be seen that the EWMA control charting model is fairly robust with respect to between-trial differences; by this we mean that the same combination of parameters proved to be optimal in both trials. An exception here is subject A, in which case the corresponding optimal pairs of L and λ do not coincide but are very close as well.

Table 9. Cross-validation testing results on EWMA control chart.

Train / Test	Feature	λ	*L*	Detection rate	False alarms	Latency (in epochs)
A01 / A02	VEOG theta	0.2	1.65	0.88	2	3.30
A02 / A01	VEOG theta	0.2	1.96	1.00	2	3.50
E01 / E02	VEOG theta	0.2	1.65	0.75	0	5.50
E02 / E01	VEOG theta	0.2	1.65	1.00	1	3.38
F01 / F02	P_z Alpha	0.3	1.96	0.25	0	5.00
F02 / F01	P_z Alpha	0.3	1.96	0.25	1	5.00

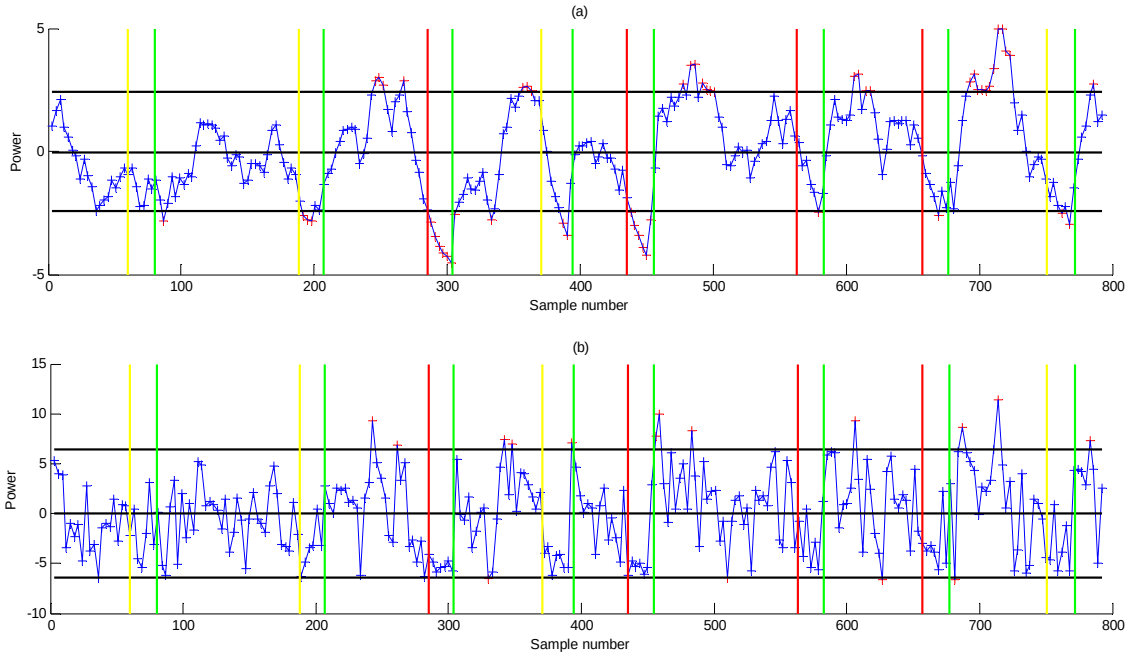

Figure 4. EWMA control chart (a) and Shewart individuals chart (b) monitoring VEOG theta during testing trial A02 (out-of-control events are identified with breaching of the lower control limit)

6. Conclusions

In this paper we considered the problem of detecting temporal changes in the functional state of human subjects due to varying levels of cognitive load using real-time psychophysiological data. We have proposed a new approach that relied on monitoring several channels of electroencephalogram (EEG) and electrooculogram (EOG) using the methods of statistical process control. It was demonstrated that control charting methods are capable of detecting changes in psychophysiological signals that are induced by varying cognitive load with high accuracy and relatively few false alarms, and also can be tailored to a specific subject. Moreover, control charts proved computationally inexpensive and readily interpretable as compared to some pattern recognition methods.

Several control chart methodologies did not perform well in our study; a possible explanation for this is that psychophysiological data is not as well suited for monitoring with control charts as are the types of stochastic processes more traditional to statistical process control applications. In general, however, our study indicates that the control chart methods represent a viable candidate for monitoring real-time psychophysiological data, but further research is needed to develop control charts that effectively exploit the particular properties of psychophysiological signals.

7. Acknowledgements

J. Cannon acknowledges support from the National Science Foundation's Graduate Research Fellowship; P. Krokhmal and R. Murphey acknowledge support from the Air Force Office of Scientific Research.

The authors would also like to thank the Human Effectiveness Directorate of the Air Force Research Lab for providing the dataset used in this study.

References

Byrne, E. A., Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology, 42*(3), 249-268.

Cannon, J. (2009). *Statistical Analysis and Algorithms for Online Change Detection in Real-Time Psychophysiological Data*, M.S. Thesis, Department of Mechanical and Industrial Engineering, University of Iowa, Iowa City, IA.

Cannon, J., Krokhmal, P. A. , Lenth, R. V., Murphey, R. (2010). An algorithm for online detection of temporal changes in operator cognitive state using real-time psychophysiological data. *Biomedical Signal Processing and Control*, to appear.

Fournier, L. R., Wilson, G. F., Swain, C. R. (1999). Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: manipulations of task difficulty and training. *International Journal of Psychophysiology, 31*(2), 129-145.

Freeman, F. G., Mikulka, P. J., Prinzel, L. J., Scerbo, M. W. (1999). Evaluation of an adaptive automation system using three EEG indices with a visual tracking task. *Biological Psychology, 50*(1), 61-76.

Gevins, A., & Smith, M. (2003). Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science, 4*, 18.

Gevins, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., Du, R., et al. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Human Factors, 40*(1), 79-91.

Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex, 7*(4), 374-385.

Hankins, T. C., & Wilson, G. F. (1998). A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. *Aviation Space and Environmental Medicine, 69*(4), 360-367.

He, P., Wilson, G., Russell, C., & Gerschutz, M. (2007). Removal of ocular artifacts from the EEG: a comparison between time-domain regression method and adaptive filtering method using simulated data. *Medical & Biological Engineering & Computing, 45*(5), 495-503.

Inouye, T., Shinosaki, K., Iyama, A., Matsumoto, Y., Toi, S., & Ishihara, T. (1994). Potential Flow of Frontal Midline Theta-Activity during a Mental Task in the Human Electroencephalogram. *Neuroscience Letters, 169*(1-2), 145-148.

Ishii, R., Shinosaki, K., Ukai, S., Inouye, T., Ishihara, T., Yoshimine, T., et al. (1999). Medial prefrontal cortex generates frontal midline theta rhythm. *Neuroreport, 10*(4), 675-679.

Jackson, J. E. (1980). Principal Components and Factor-Analysis .1. Principal Components. *Journal of Quality Technology, 12*(4), 201-213.

Jung, T. P., Makeig, S., Stensmo, M., & Sejnowski, T. J. (1997). Estimating alertness from the EEG power spectrum. *IEEE Transactions on Biomedical Engineering, 44*(1), 60-69.

Klimesch, W., Schmike, H., & Pfurtscheller, G. (1993). Alpha frequency, cognitive load, and memory performance. *Brain Topography, 5*, 10.

Lotte, F., Congedo, M., Lecuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering, 4*, R1–R13.

Lowry, C. A., & Montgomery, D. C. (1995). A Review of Multivariate Control Charts. *IIE Transactions, 27*, 10.

Lowry, C. A., Woodall, W. H., Champ, C. W., & Rigdon, S. E. (1992). A Multivariate Exponentially Weighted Moving Average Control Chart. *Technometrics, 34*(1), 46-53.

Mastrangelo, C. M., Runger, G. C., & Montgomery, D. C. (1996). Statistical process monitoring with principal components. *Quality and Reliability Engineering International, 12*(3), 203-210.

Montgomery, D. (2009). *Statistical Quality Control* (6th edition ed.): John Wiley & Sons, Inc.

Montgomery, D. C., Jennings, C. L., & Kulachi, M. (2008). *Introduction to Time Series Analysis and Forecasting*. Hoboken: John Wiley & Sons.

Montgomery, D. C., & Mastrangelo, C. M. (1991). Some Statistical Process-Control Methods for Autocorrelated Data. *Journal of Quality Technology, 23*(3), 179-193.

Phillips, C., Parr, J., & Riskin, E. (2008). *Signals, Systems, and Transforms* (4 ed.). Upper Saddle River: Prentice Hall.

Prinzel, L. J., Freeman, F. C., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2000). A closed-loop system for examining psychophysiological measures for adaptive task allocation. *International Journal of Aviation Psychology, 10*(4), 393-410.

Scranton, R., Runger, G. C., Montgomery, D. C., & Keats, B. J. (1996). Efficient Shift Detection Using Multivariate Exponentially-Weighted Moving Average Control Charts and Principle Components. *Quality and Reliability Engineering International, 12*, 6.

Smith, M. E., Gevins, A., Brown, H., Karnik, A., & Du, R. (2001). Monitoring task loading with multivariate EEG measures during complex forms of human-computer interaction. *Human Factors, 43*(3), 366-380.

Wilson, G. F., & Fisher, F. (1991). The Use of Cardiac and Eye Blink Measures to Determine Flight Segment in F4-Crews. *Aviation Space and Environmental Medicine, 62*(10), 959-962.

Wilson, G. F., & Fisher, F. (1995). Cognitive Task Classification Based upon Topographic EEG Data. *Biological Psychology, 40*(1-2), 239-250.

Wilson, G. F., & Russell, C. (1999). *Operator Functional State Classification Using Neural Networks with Combined Physiological and Performance Features*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting.

Wilson, G. F., & Russell, C. A. (2003a). Operator functional state classification using multiple psychophysiological features in an air traffic control task. *Human Factors, 45*(3), 381-389.

Wilson, G. F., & Russell, C. A. (2003b). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Human Factors, 45*(4), 635-643.

Wilson, G. F., & Russell, C. A. (2007). Performance enhancement in an uninhabited air vehicle task using psychophysiologically determined adaptive aiding. *Human Factors, 49*(6), 1005-1018.

Ye, N., Borror, C., & Zhang Y. (2002). EWMA Techniques for Computer Intrusion Detection through Anomalous Changes in Event Intensity. *Quality and Reliability Engineering International, 18*, 443-451.

Yu, Z. (2009). *Optimization Techniques in Data Mining with Applications to Biomedical and Psychophysiological Data Sets.* M.S. Thesis, Department of Mechanical and Industrial Engineering, University of Iowa, Iowa City, IA.

Appendix A
False alarm data for control charts

Table A.1 False alarms of EWMA-Shewart control charts

	VEOG Theta	F_z Theta	F_z Alpha	F_7 Theta	0_2 Alpha	P_z Alpha	Average
A01	1	2	1	2	0	1	**1.17**
E01	0	0	1	1	0	0	**0.33**
F01	1	0	1	0	1	0	**0.50**
Average	**0.67**	**0.67**	**1.00**	**1.00**	**0.33**	**0.33**	

Table A.2 False alarms of time series residuals control charts

	HEOG Beta	F_z Beta	P_z Beta	T_5 Beta	Average
A01	1	1	2	1	**1.25**
E01	2	1	1	1	**1.25**
F01	4	4	3	1	**3.00**
Average	**2.33**	**2.00**	**2.00**	**1.00**	

Table A.3 False alarms of MCEMWA control charts

	F_z Beta	P_z Beta	T_5 Beta	Average
A01	3	3	3	**3.000**
E01	2	2	1	**1.667**
F01	2	4	2	**2.667**
Average	**2.333**	**3.000**	**2.000**	

Table A.4 False alarms of multivariate control charts

	Hotelling-T^2 (Select)	MEWMA (Select)	Hotelling-T^2 (PCA)	MEWMA (PCA)	Average
A01	4	2	4	3	**3.25**
E01	4	1	2	2	**2.25**
F01	6	4	4	2	**4.00**
Average	**4.66**	**2.33**	**3.33**	**2.33**	

Appendix B
Graphical illustrations of control charts for psychophysiological data

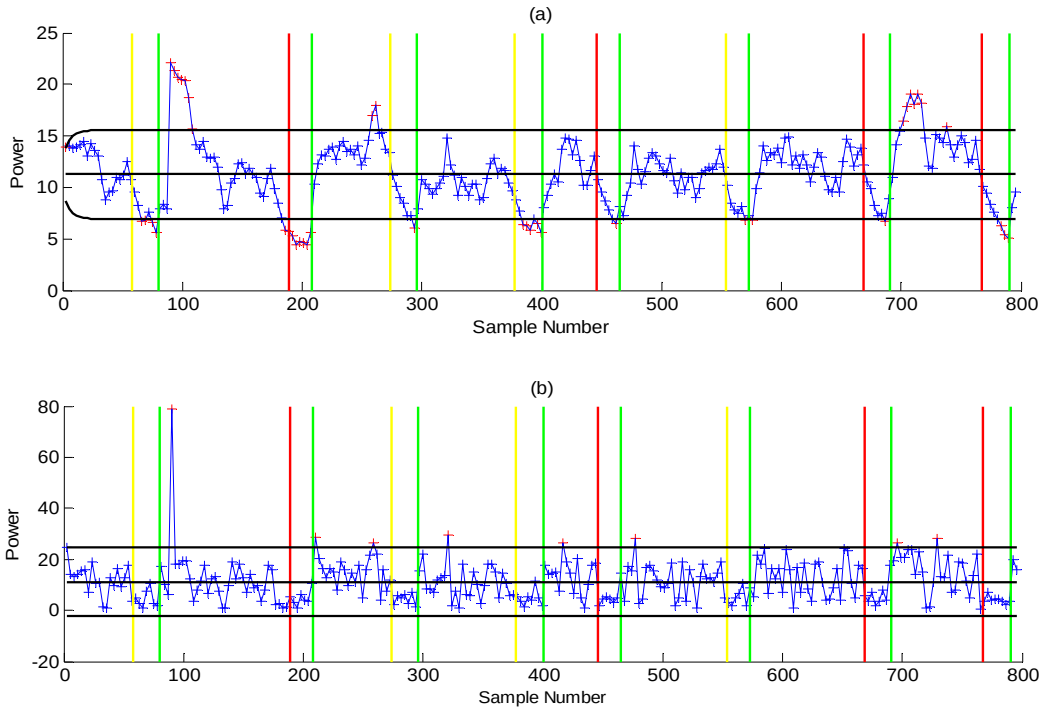

Figure B.1 EWMA control chart (a) and Shewart individuals chart (b) for monitoring VEOG theta of E01

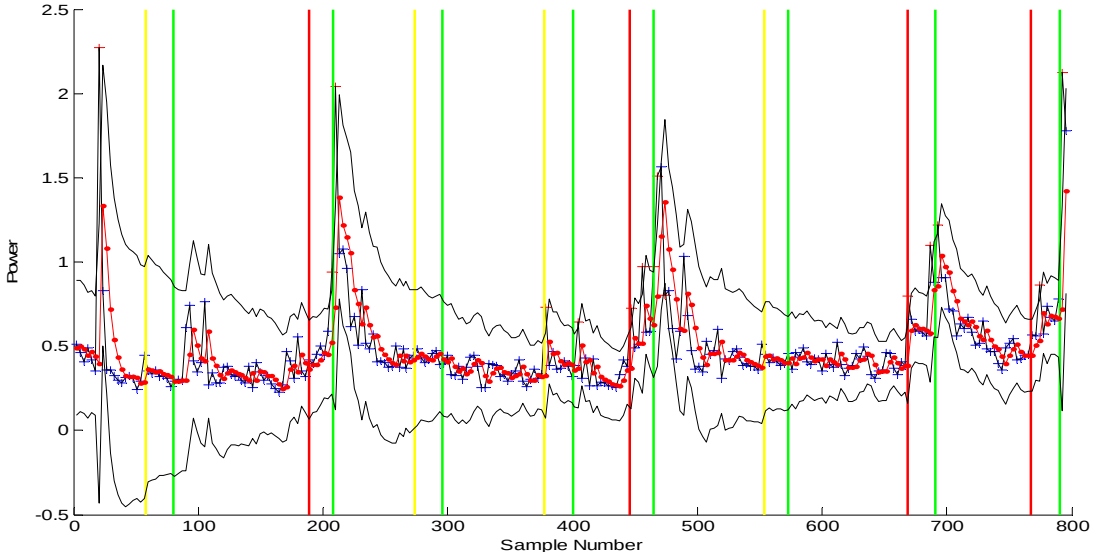

Figure B.2 MCEWMA chart monitoring F_7 beta of E01

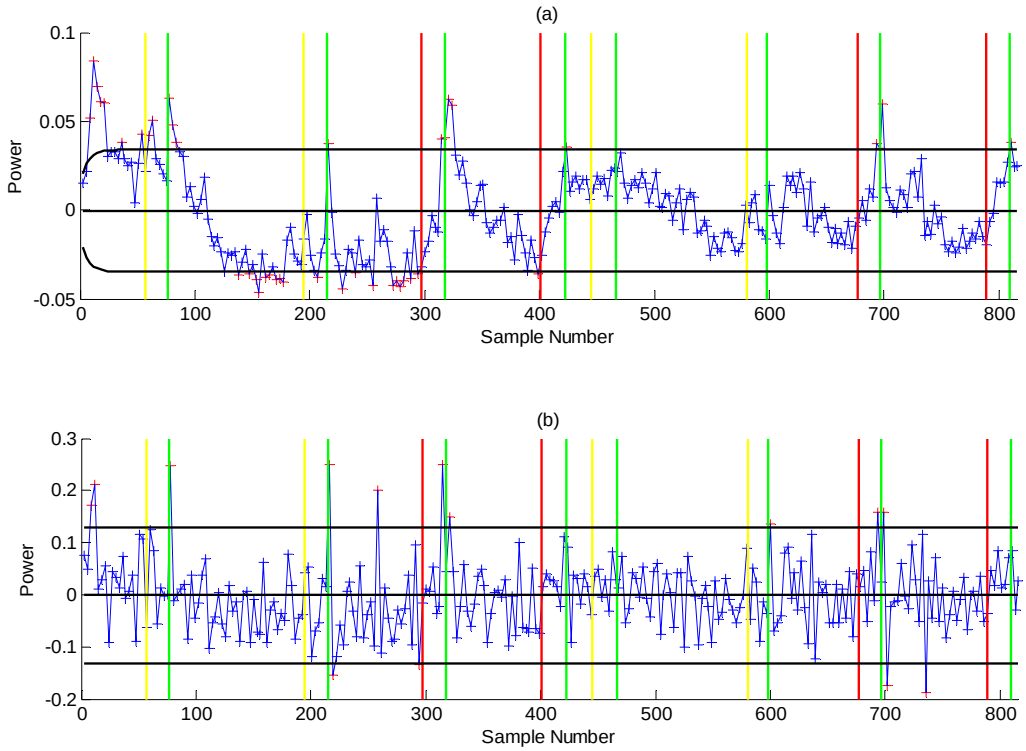

Figure B.3 EWMA chart (a) and Shewart individuals chart (b) for monitoring of time series residuals from F_z beta of A01

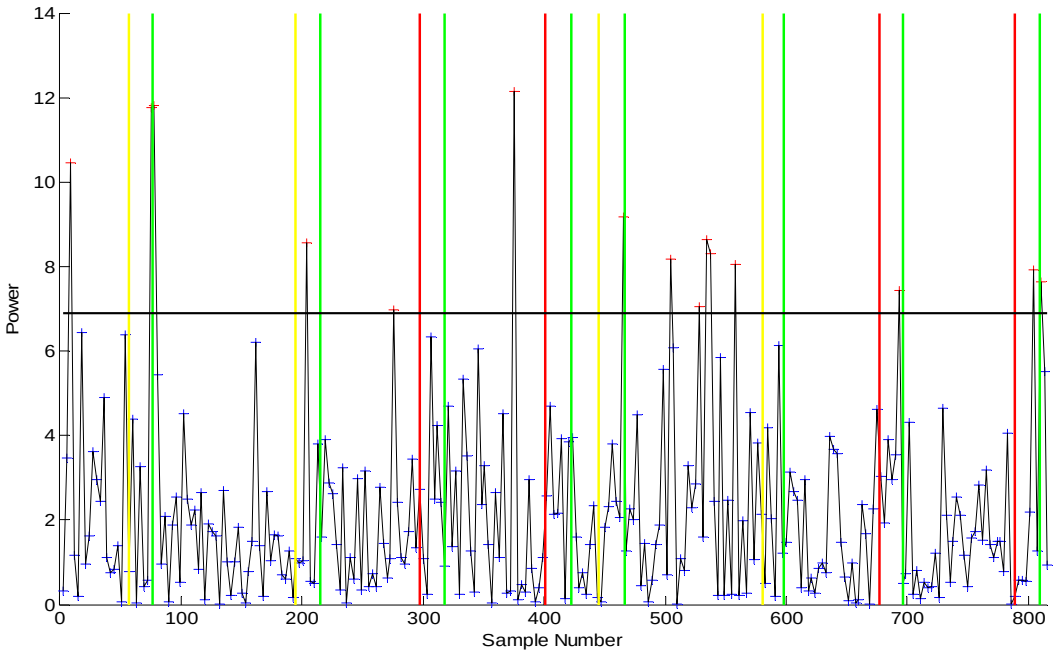

Figure B.4 Hotelling-T^2 chart monitoring select alpha and theta features of A01

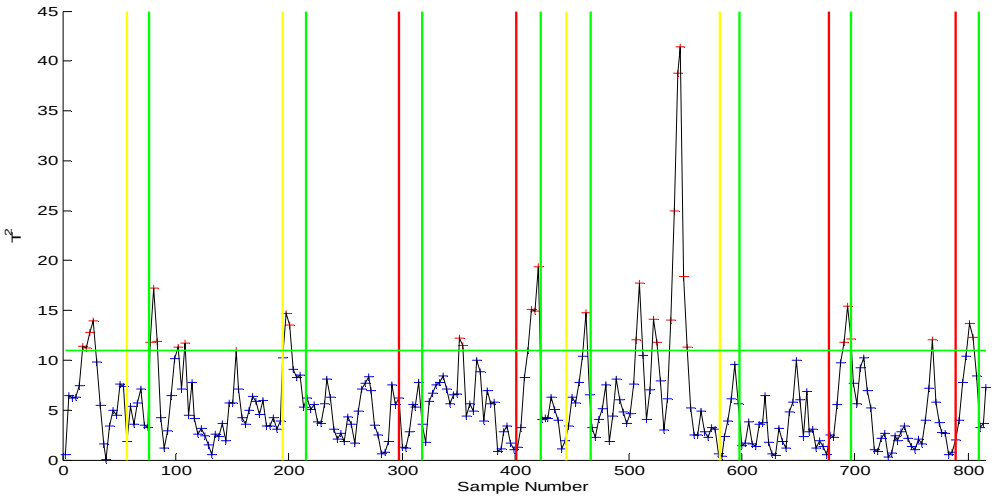

Figure B.5 MEWMA chart monitoring select alpha and theta features of E01

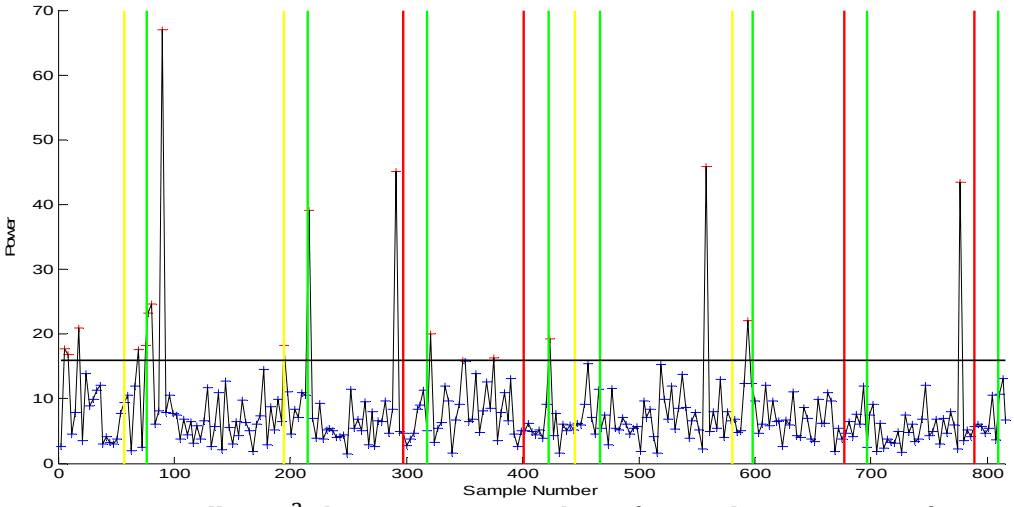

Figure B.6 Hotelling-T^2 chart monitoring a subset of principle components of A01

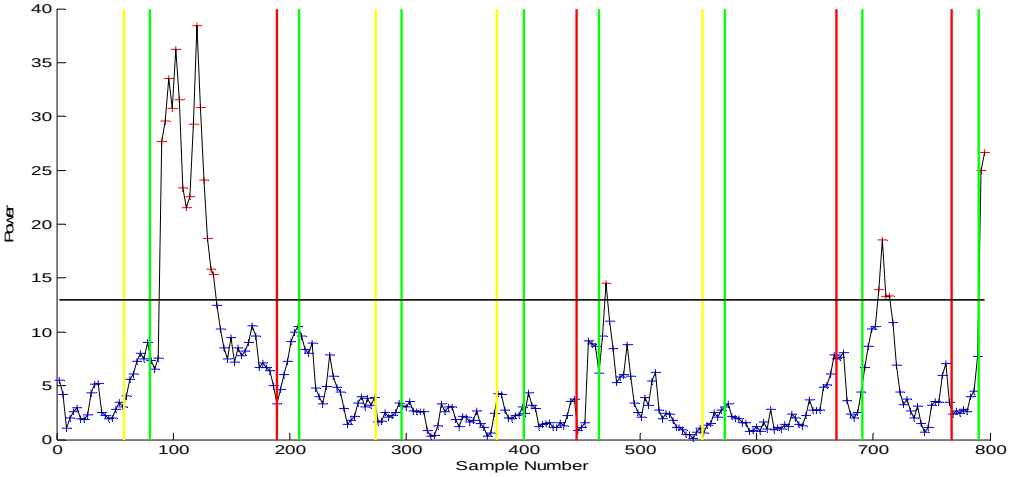

Figure B.7 MEWMA chart monitoring a subset of principle components of E01