2022 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT)
March 10-11, 2022. Texas A&M University - EnMed, Houston, TX, USA

2022 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT) | 978-1-6654-9615-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/HI-POCT54491.2022.9744065

A Wearable Exam Stress Dataset for Predicting Grades using Physiological Signals

Md. Rafiul Amin, *Student Member, IEEE*, Dilranjan S. Wickramasuriya, *Member, IEEE*, and Rose T. Faghih*, *Senior Member, IEEE*

***Abstract*— The study of psychological stress in real-world scenarios presents several challenges. Consequently, datasets available to researchers are also scarce. The aim of our study is to acquire such a dataset containing skin conductance measurements and use it to predict human performance. We collected skin conductance and skin temperature data from 10 subjects during three exams using wearable devices. We filter the skin conductance signals to obtain coarse-grained trendlines and then train classifiers to predict high and low grades based on the trendline features. We obtained maximum classification accuracies in the 70–80% range. We also obtained the mean trendlines indicating the general variation of stress levels during the exams. The findings indicate the preliminary viability of using wearable devices to predict performance during real-world stressors. Wearable monitoring presents unique challenges and it is our hope that this publicly-available dataset will aid in addressing some of them.**

I. Introduction

The World Health Organization dubbed stress the "Health Epidemic of the 21st Century" [1]. Stress takes its toll on physical health, affects work productivity, and incurs substantial annual expenses in terms of healthcare costs and costs to industries [1]. Not surprisingly, there has been significant research interest in developing automated technologies that can detect physiological signs of stress for helping people better manage their mental and emotional well-being. While stress is known to be associated with increased cardiovascular disease risk [2] and negative mental health outcomes [3], it is of interest to note that stress does not affect work *productivity* and *performance* purely in a negative manner. According to the Yerkes-Dodson law (also known as the inverted U law) [4], an individual's ability to perform a task at hand is negatively affected primarily when stress levels are either too high or too low. A moderate amount of stress *positively* impacts how well an individual performs. Thus, from a human performance perspective, it is of importance to know how much stress an individual is experiencing and if he/she is at an optimal level of productivity.

Much research has been done in recognizing biomarkers for predicting the mental/emotional state of an individual and a number of experimental studies have been published in this regard. Human emotions can be visualized on a 2D plane with orthogonal axes known as valence and arousal—valence denotes the pleasant/unpleasant nature of an emotion, while arousal captures its corresponding intensity or activation [5]. Experimental studies in the literature, which come under the general umbrella of automated emotion recognition or affective computing, can be broadly categorized into three groups. The first group of studies consists of those where signals are acquired from subjects while they are shown images, movie clips, music videos, etc. corresponding to different valence-arousal ratings or emotional categories (e.g., [6], [7]). Here, the objective is to develop machine learning algorithms that can predict valence, arousal, or specific emotions based on physiological features. A second group of experiments, such as those described in [8], [9], seek to emulate a real-world stress experience (e.g., office work-like tasks, driving) and in so doing attempt to decode an individual's mental state. A third miscellaneous group of studies involving different types of stimuli or tasks is being presented to the subjects. The stimuli may include mathematical tasks, movie clips, fragrances, standard psychological stress tests such as the Trier Social Stress Test, the Cold Pressor Test, etc. (e.g. [10]–[12]). Again, the objective is to identify patterns in biosignals with eventual application to decoding mental states and emotions.

While these studies have certainly helped advance our understanding of emotion and the associated psychophysiology, they have a notable drawback regarding stress. To illustrate, we pose the question as to what stressors out a person in the real world? Real-world stressors include job interviews, performance reviews, exams, office meetings, medical procedures, etc. Unfortunately, unless datasets are collected in these real-world scenarios, attempts to *emulate* their effect in the laboratory or elsewhere will fall short. Consequently, fewer real-world stress datasets are also available for researchers to use. Such datasets, conducted over a period of time on the *same* individuals for longitudinal studies are even more rare. Moreover, real-world scenarios are unconstrained environments. Often, research-grade equipment cannot be used and motion artifact contamination is widespread. These remain some of the primary obstacles for automated emotion decoding outside the research laboratory in everyday life.

Our contribution here is twofold. We first conducted an experiment where physiological data were collected from a group of students during three exams. For the convenience of the students enrolled in the study, we used a wearable watch-

This work was supported in part by by the U.S. National Science Foundation under Grants 1942585 – CAREER: MINDWATCH: Multimodal Intelligent Noninvasive brain state Decoder for Wearable AdapTive Closed-loop arcHitectures and 1755780 – CRII: CPS: Wearable-Machine Interface Architectures.

M. R. Amin, D. S. Wickramasuriya and, *R. T. Faghih are with the University of Houston, Houston, TX, USA (correspondence e-mail: mamin@uh.edu, rfaghih@nyu.edu). Rose T. Faghih served as the senior author.

like device (Empatica E4) to collect skin conductance (SC), skin temperature (SKT) and wrist movement data.We next analyze the SC data for developing a preliminary decoding tool for predicting grade type (performance) based on SC features. Secondly, we make the data publicly available in the hope that other researchers would be able to use it for further analysis and algorithm development (e.g., using other physiological signals). The dataset also highlights the need for robust emotion recognition methods that work well in the presence of highly prevalent motion contamination in unconstrained environments.

II. Methods

A. Study Procedure and Data Acquisition

The study protocol was reviewed and approved by the University of Houston Institutional Review Board (IRB). The study involved 11 students (nine males and two females) enrolled in an undergraduate course on circuit analysis at the University of Houston. Participants were asked to read and sign an informed consent form prior to enrolling in the study. We collected data during three exams held in the semester: mid-term 1, mid-term 2, and the final exam. Both mid-term exams were 1.5 h in duration while the final was 3 h long. Each participant was given an Empatica E4 device approximately 5 min before the start of each exam to be worn on their nondominant hand. The Empatica E4 records several different physiological signals including SC. One of the students was provided additional time per the University of Houston's disability accommodation guidelines. This involved a relocation of the students to a different room after the regular time had ended. Since this involves an additional factor not uniform to the other participants, we discarded the data from this particular subject and only consider the data from the remaining ten. During mid-term 1, we collect SC data and accelerometer data for motion information. We also collected SKT in the subsequent exams. The de-identified data can be accessed from the repository which will be made publicly available soon. The immediate access to the de-identified dataset will be provided by the corresponding author upon request.

B. Skin Conductance

Changes in SC occur due to tiny variations in sweat secretions. Since the sweat glands are innervated by nerve fibers belonging to the sympathetic branch of the nervous system [13], an SC signal provides an index of *sympathetic arousal* [14]. Now different markers of sympathetic arousal are present in SC signals. Additionally, since the the body's *stress response* is closely linked to the activation of the sympathetic nervous system, these markers of arousal also capture information regarding an individual's stress levels. An SC signal consists of a slow-varying tonic component on top of which a faster-varying phasic component is superimposed [15] (Fig. 1). This faster-varying phasic component consists of bi-exponentially-shaped skin conductance responses (SCRs). Each of these SCRs can be modeled as being generated by a burst of neural activity to the sweat glands [16], [17]. The rates at which SCRs occur and their amplitudes are markers of sympathetic arousal (higher SCR rates and amplitudes typically indicate higher levels of arousal) [18], [19]. Meanwhile, the tonic component $y_s(t)$, despite being related to thermoregulation and being dependent on ambient temperature and humidity, also captures some general variations in arousal [20].

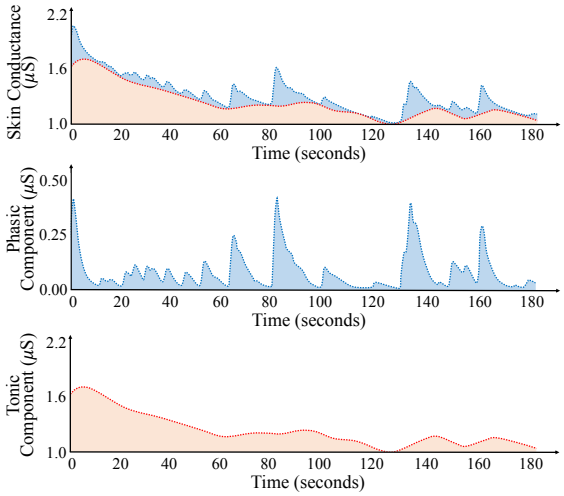

Fig. 1: **Components of an SC Signal.** The first sub-panel depicts an SC signal and the sub-panels below it depict the separated-out phasic and tonic components, respectively [16].

Typically, one of the first steps in SC signal analysis is determining the sequence of neural impulses that are responsible for phasic SCRs. This step is either preceded by or simultaneously performed along with tonic-phasic separation. In a recent work [21], we developed a sparse deconvolution algorithm for tonic-phasic separation and the extraction of the neural impulse train underlying SCRs that outperformed several others in the literature including [22]–[24]. However, none of these methods (including ours) incorporate the ability to handle motion artifacts well and therefore suffer badly in the presence of such contamination. This happens largely because model-based deconvolution approaches only consider Gaussian noise. The presence of non-Gaussian noise thus leads to erroneous tonic-phasic separation and neural impulse detection. Consequently, we were also unable to use our prior state estimation methods (e.g., [18], [25]) that utilize the deconvolved neural impulses for continuously decoding sympathetic arousal. This reflects one of the challenges previously highlighted with motion artifacts in unconstrained wearable device environments.

C. SC Trendline Extraction and Filtering

As noted earlier, the slow-varying tonic component $y_s(t)$ does capture certain changes in sympathetic arousal [20]. Since $y_s(t)$ varies slowly, a heuristic approach for extracting it from an SC signal involves straightforward filtering. For instance, high-pass filtering with a low cut-off frequency is used in [26], [27] for separating out the tonic component (the high-pass filtered component is subtracted from the SC

TABLE I: Lowpass Filter Cut-off Frequencies used to Extract SC Trendlines.

Participant ID	Cut-off Frequency (Hz)		
	Mid-term 1	Mid-term 2	Final
1	0.002	0.002	0.002
2	0.002	0.001	0.002
3	0.0002	0.0002	0.001
4	0.002	0.002	0.002
5	0.001	0.001	0.001
6	0.002	0.002	0.002
7	0.001	0.0002	0.0002
8	0.0002	0.002	0.002
9	0.002	0.002	0.002
10	0.002	0.002	0.002

signal to finally extract the tonic component). Here we use a lowpass FIR filter to obtain a coarse-grained SC trendline that approximates $y_s(t)$. Since the noise levels are different for each SC signal, we selected different cut-off frequencies (Table I) to extract these trendlines after manual inspection. The cut-off frequencies were selected based on a trade-off of capturing the general tonic variation vs. not letting in high-frequency noise and was verified by two different viewers by visual inspection. Later in the discussion section, we show how using a low-pass filter for extracting $y_s(t)$ is almost immune to noise compared to one of the more popular model-based decomposition methods.

D. Feature Extraction and Classification

We denote by $\mathbf{y}$ a vector of all the samples in the SC trendline for each subject for a particular exam. For feature extraction, we first begin by extracting three data segment windows from the beginning, middle, and end of each trendline. We use T_w to denote the duration of a data segment window and T to denote the total signal duration. We choose segments from the beginning, middle, and end of the exam based on an assumption of how stress dynamics may vary during the test. For instance, we assume that participants may be more stressed during the initial and final stages of the exam. Stress experienced early on may be related to a participant's personal sense of being adequately prepared and the type of exam questions seen at first glance [28]. Stress may gradually diminish as the exam progresses and participants begin answering. Concerns regarding the amount of time left, fatigue, and the need to complete answering the required number of questions may govern feelings of stress felt towards the end [29].

Fig. 2 illustrates how data segment windows are extracted. Consider T_w = 5 min as an example. For this value of T_w, we extract the first 5 min of the SC trendline, 5 min from the middle, and the final 5 min of the trendline. Similarly, for T_w = 15 min, we extract the first 15 min, 15 min from the middle and the final 15 min of the trendline. As features, we first compute the trendline mean and variance within each window. This provides $3 \times 2 = 6$ features. Mathematically, this can be expressed as follows. Defining the three trendline

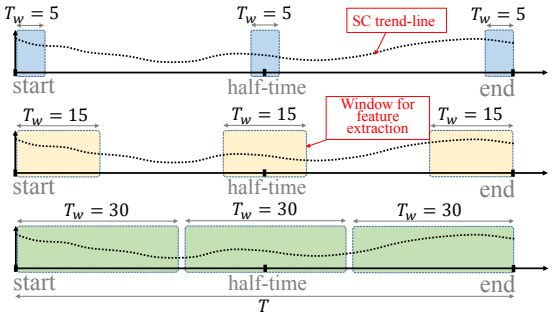

Fig. 2: **Window-based Feature Extraction.** The figure depicts the extraction of different windows of data based on T_w. Here the unit for T_w is minutes.

segments for a given T_w as:

$$\mathbf{y}_{\text{start}} = \mathbf{y}_{1:T_w},$$
$$\mathbf{y}_{\text{mid}} = \mathbf{y}_{(\frac{T}{2}-\frac{T_w}{2}+1):(\frac{T}{2}+\frac{T_w}{2})},$$
$$\text{and } \mathbf{y}_{\text{end}} = \mathbf{y}_{(T-T_w+1):T},$$

where $\mathbf{y}_{i:j}$ denotes the vector comprising of the samples from index i to j, we calculate:

$$\mu_{\text{q}} = \text{mean}(\mathbf{y}_{\text{q}}), \text{ and } \sigma^2_{\text{q}} = \text{var}(\mathbf{y}_{\text{q}})$$

for $q \in \{\text{start}, \text{mid}, \text{end}\}$ (here we have shown the feature calculations in terms of T_w and T represented in an equivalent number of sample points rather than as a duration in minutes). Moreover, since the raw SC trend lines differ between participants, and a high mean or variance value for a particular participant may not necessarily be high for another participant who ordinarily has a higher SC, we calculate the ratio feature

$$\rho = (\mu_{\text{mid}})/(\mu_{\text{start}} + \mu_{\text{end}}), \tag{1}$$

To capture how stress levels may have reduced (or increased) during the middle of the exam compared to what they were at the beginning and what they are at the end. Finally, there is information in how *rapidly* (or slowly) sympathetic arousal increases or decreases in an exam. For instance, if a student is well-prepared, he/she may experience rapidly declining sympathetic activation upon looking through the exam questions for the first time. To capture this *rate* change in trendlines, we calculate a difference signal

$$\mathbf{y}_{\text{diff}} = \mathbf{y}_{2:T_w} - \mathbf{y}_{1:(T_w-1)}, \tag{2}$$

and then obtain its mean and variance:

$$\mu_{\text{diff}} = \text{mean}(\mathbf{y}_{\text{diff}}), \text{ and } \sigma^2_{\text{diff}} = \text{var}(\mathbf{y}_{\text{diff}}). \tag{3}$$

Our final nine-element feature vector is $x = [\mu_{\text{start}}, \mu_{\text{mid}}, \mu_{\text{end}}, \sigma^2_{\text{start}}, \sigma^2_{\text{mid}}, \sigma^2_{\text{end}}, \rho, \mu_{\text{diff}}, \sigma^2_{\text{diff}}]$. We extract these nine features for T_w = 5, 15 and 30 min, and then train separate classifiers for each T_w for each exam. We assign a high or low grade label to each feature vector based on whether or not the participant scored greater than or equal to 80%. Based on the criteria, we have four

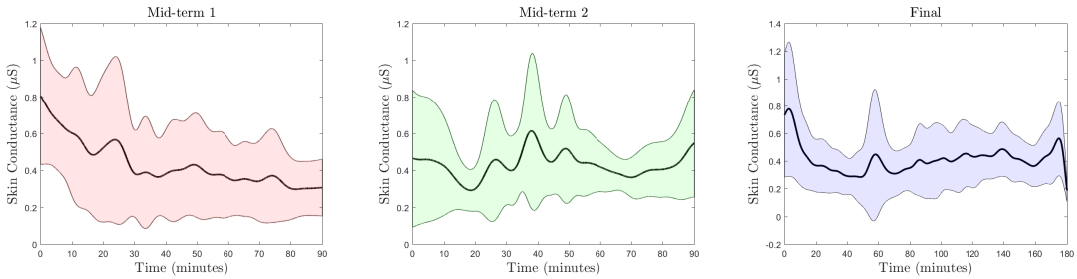

Fig. 3: **Mean SC Trendlines for All Participants.** Each sub-panel depicts the mean SC trendline of the participants in an exam along with its 95% confidence limits.

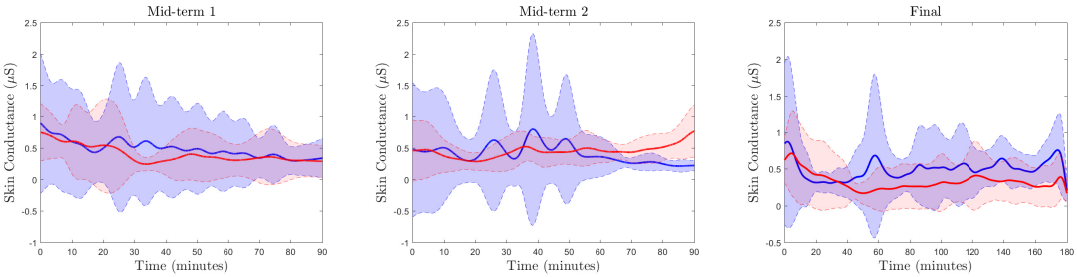

Fig. 4: **Mean SC Trendlines for Two Classes.** Each sub-panel depicts the mean SC trendline of the participants in an exam along with its 95% confidence limits. Blue and red lines correspond to the high and low grade means, respectively.

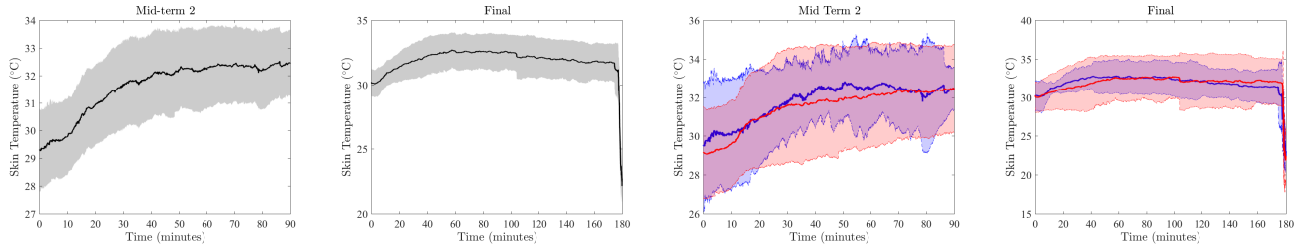

Fig. 5: **Mean SKT Trendlines.** Each sub-panel depicts the mean SKT of the participants in an exam along with its 95% confidence limits. Black, blue, and red lines correspond to the overall, high grade, and low grade means, respectively.

subjects with high grades for mid-terms 1 and 2, and five subjects with a high grade for the final exam. We use all classifiers from MATLAB 2017b classification application with default settings. We evaluate the performance of the classifiers using 10-fold cross-validation. Here, cross-validation protects against overfitting by partitioning the dataset into different folds and then estimating the accuracy within each fold. We avoid classifying a combination of data from all three exams since each exam is a separate stimulus with different levels of difficulty and preparedness.

III. RESULTS

The mean SC trendlines for the three exams are shown in Fig. 3. For mid-term 1, there is a general pattern where the trendline gradually decreases as the exam progresses. A similar pattern is seen for the final exam, although the subsequent decrease after the start is not as prominent. The mean trendline is unusual for mid-term 2 since there are peaks at the beginning, middle, and end with lower valleys in between. Figs. 4 shows the mean SC trendlines for the high and low grade categories, separately.

TABLE II: Grade Classification Accuracies with $T_w = 5$ min (kNN – k-Nearest Neighbors)

Exam	Fine kNN
Midterm 1	60%
Midterm 2	60%
Final	**80%**

TABLE III: Grade Classification Accuracies $T_w = 15$ min (kNN – k-Nearest Neighbors)

Exam	Fine kNN	Ensemble Bagged Trees	Ensemble Subspace kNN
Midterm 1	**70%**	60%	50%
Midterm 2	70%	**80%**	60%
Final	50%	50%	60%

Ambient temperature affects the body's effort to cool itself through sweating. Since we are primarily analyzing a feature related to sweating, we have shown the mean SKT values for mid-term 2 and the final exam in Fig. 5. Motion artifact

TABLE IV: Grade Classification Accuracies $T_w = 30$ min (SVM – Support Vector Machine)

Exam	Cubic SVM
Midterm 1	**70%**
Midterm 2	70%
Final	50%

contamination of the SKT signals is lower and hence low-pass filtering is not required. Mean SKT values appear to gradually rise during the exams. This is rather counterintuitive in light of the pattern observed on the SC trendlines. It may be the case, however, that while *core* body temperature primarily affects the cooling effect through sweating, *skin* temperature may have a more complex relationship with the sweat-based cooling effect. Additionally, the rapid drop in mean SC and SKT at the very end of the final exam is due to some of the participants completing their tests slightly ahead of time and taking off their Empatica E4 devices.

The 10-fold validation accuracies for classification using different machine learning algorithms for different values of T_w are shown in Tables II-IV where cells with '–' symbol denote that the corresponding classification accuracy was less than 50%. The maximum accuracies for each of the three exams are 70%, 80%, and 80%, respectively and are highlighted in blue in each table. The results indicate the feasibility of predicting exam performance from SC signal trendlines which capture sympathetic arousal information. We also used principal component analysis (PCA) to project the features onto a 2D plane for visualization. Fig. 6 shows the PCA plots for T_w = 30, 15 and 5 minutes for the three exams corresponding to the window sizes that yielded the highest classification accuracies.

IV. Discussion

Emotional stress studies encounter a number of challenges in real-world unconstrained environments. Research-grade equipment is often unavailable for signal acquisition and motion artifact contamination is a major challenge. While our group has developed multiple SC decolvolution algorithms [16], [17], [21], [30] and point process sympathetic arousal estimation methods [18], [19], [25], [31]–[33], we were unable to use them on this dataset due to the reasons highlighted earlier. However, with a coarse-grained analysis of wearable data, we were able to demonstrate the preliminary viability of predicting exam performance from SC signals.

As an additional point, we demonstrate the robustness of using a low-pass filter to extract the tonic component compared to a model-based decomposition approach (we use cvxEDA [24] for comparison). We begin by extracting the tonic component from a clean SC signal using both a lowpass filter and cvxEDA. We then introduce Gaussian noise into the signal so that the SNR is 15 dB and reextract the tonic component using low-pass filtering and cvxEDA. The following four scatter plots help us (Fig. 7) evaluate the robustness of both methods in the presence of noise:

- Clean tonic component from cvxEDA vs. clean tonic component from low-pass filter (Fig. 7-(a))
- Clean tonic component from cvxEDA vs. noisy tonic component from cvxEDA (Fig. 7-(b))
- Clean tonic component from cvxEDA vs. noisy tonic component from lowpass filter (Fig. 7-(c))
- Clean tonic component from lowpass filter vs. noisy tonic component from lowpass filter (Fig. 7-(d))

We use a relatively noise-free SC signal from the Driver Stress Data in [34], [35] to perform this comparison. We applied a low-pass filter with a cut-off frequency of 0.001 Hz.

If the method of tonic component extraction is robust to noise, we would expect to see a scatter plot close to the 45° diagonal between the clean and noisy versions. A significant deviation between the clean and noisy versions would indicate a lack of robustness in the presence of noise contamination. Here we add noise to create a 10 dB SNR signal. For each of the scatter plots, we also fit a simple straight line (via regression) and calculate the Pearson correlation coefficient between the variables plotted on either axis.

Consider Fig. 7-(a) which depicts the clean tonic component extracted via cvxEDA plotted against the clean tonic component extracted using low-pass filtering. The Pearson correlation coefficient between the two variables is 0.85 and the correlation via the regression line is also statistically significant ($p < 0.01$). Thus, the simple low-pass filtered trendline contains some of the actual tonic component information extracted using the more sophisticated model-based cvxEDA decomposition method.

The Fig. 7-(b) shows how the tonic component extracted via cvxEDA from the noise-contaminated SC signal loses some of its resemblance to its corresponding clean version (the correlation is also negative). The Fig. 7-(c) and 7-(d), however, depict the robustness of the low-pass filtering approach. The trendline extracted from the noisy SC signal using low-pass filtering remains correlated with a correlation coefficient of 0.85 (same as before) with the clean tonic ground truth and the correlation via the regression line is also statistically significant ($p < 0.01$). Furthermore, it lies very close to the 45° diagonal when plotted against its corresponding clean version showing the robustness to noise.

Different methods have been explored in the literature to suppress the effect of noise and motion artifacts. Among them are multichannel data acquisition (i.e., adding redundancy to the data to eliminate dependency on a single channel) and the use of adaptive filters. Adaptive filters utilize accelerometer readings to derive a noise reference for motion artifact suppression. We plan on developing adaptive filtering methods to improve the quality of wearable SC data in the future.

The small size of the dataset (ten participants) is a limitation of our study. However, we do possess recordings from each participant over three exams. This, therefore, provides the opportunity to perform longitudinal analysis over time. While we obtained reasonable classification accuracy on

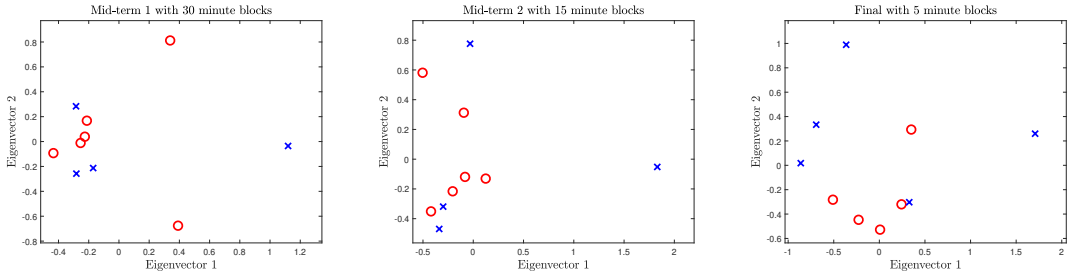

Fig. 6: **SC Feature Distribution.** Each sub-panel depicts the nine-dimensional SC trendline features projected onto a 2D plane using PCA for the three exams. Red circles and blue crosses correspond to low grade and high grade classes, respectively.

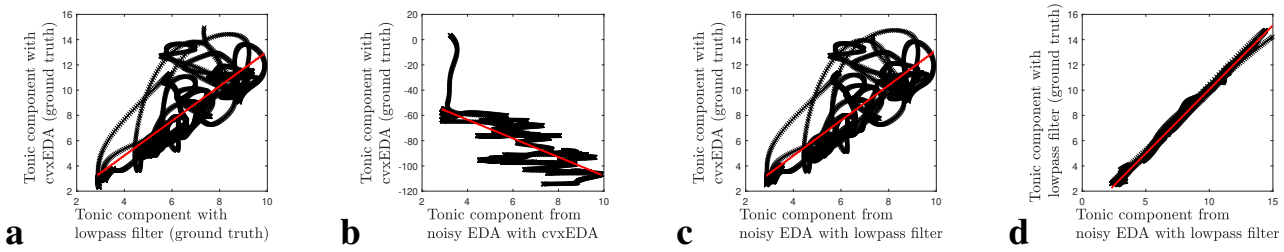

Fig. 7: **Scatter Plots and Regression Line Fits Corresponding to the Extracted Tonic Components from Clean and Noisy Data.** The sub-plots depict the scatter plots (black crosses) and regression lines (red) for: (a) the clean tonic component extracted using cvxEDA vs. the clean tonic component extracted using lowpass filtering; (b) the clean tonic component extracted using cvxEDA vs. the noisy tonic component extracted using cvxEDA; (c) the clean tonic component extracted using cvxEDA vs. the noisy tonic component extracted using lowpass filtering; (d) the clean tonic component extracted using lowpass filtering vs. the noisy tonic component extracted using lowpass filtering.

the current dataset, the inclusion of more data may help improve the accuracy even further. These classification accuracies show a potential link between the information in SC signal and the exam performance, which needs to be further investigated as the future study. The machine learning models attempt to determine the decision surface mapping from physiological signal features to the participants' grades. This is likely a complex mapping and the acquisition of a larger dataset may enable more sophisticated models to determine the mapping even better. We also only utilize SC features for classification. Furthermore, the study has limitations regarding manual filter adjustment via visual inspection depending on individuals. As the future study, researchers should consider an optimal filtering approach based on the motion level (accelerometer data). Finally, the uncertainty to identify the appropriate parts of the signal is another limitation of the study. By making the data publicly available, it is our hope that other researchers could explore additional physiological features and artificial intelligence (AI) techniques as well so that these limitations eradicated.

V. Conclusion

Decoding stress and emotion in the real world is a challenge. Nevertheless, it has important clinical and human performance-related applications for everyday life. In this study, we collected data from a group of students over three exams and attempted to predict their exam grade types based on SC features. Wearable monitoring presents the challenge of artifact contamination in unconstrained real-world settings. We therefore had to rely on coarse-grained measures extracted from the signals for analysis and classification. We extracted SC features from a low-frequency trendline and obtained maximum classification accuracies in the 70-80% range for high vs. low grade prediction. Our results indicate the preliminary viability of using wearable device data for real-world stress-related applications.

Future work would consist of conducting further experiments based on a larger group and/or on a similar-sized group but over a longer duration. We also plan to develop adaptive filters for the suppression of motion artifacts and subsequently use our deconvolution and state-space methods to estimate underlying stress levels in participants. We further plan to develop closed-loop control algorithms based on these state-space models to investigate how stress levels may be adjusted to ensure optimal human performance [36]–[38].

Acknowledgment

The authors wish to thank Dr. David P. Shattuck for his help with the collection of data.

References

[1] G. Fink, "Chapter 1 - stress, definitions, mechanisms, and effects outlined: Lessons from anxiety," in *Stress: Concepts, Cognition, Emotion, and Behavior* (G. Fink, ed.), pp. 3 – 11, San Diego: Academic Press, 2016.

[2] M. Kivimäki and A. Steptoe, "Effects of stress on the development and progression of cardiovascular disease," *Nature Reviews Cardiology*, vol. 15, no. 4, p. 215, 2018.

[3] S. B. Scott, J. E. Graham-Engeland, C. G. Engeland, J. M. Smyth, D. M. Almeida, M. J. Katz, R. B. Lipton, J. A. Mogle, E. Munoz, N. Ram, *et al.*, "The effects of stress on cognitive aging, physiology and emotion (ESCAPE) project," *BMC Psychiatry*, vol. 15, no. 1, pp. 1–14, 2015.

[4] M. Corbett, "From law to folklore: work stress and the yerkes-dodson law," *Journal of Managerial Psychology*, 2015.

[5] J.-R. Zhuang, Y.-J. Guan, H. Nagayoshi, L. Yuge, H.-H. Lee, and E. Tanaka, "Two-dimensional emotion evaluation with multiple physiological signals," in *International Conference on Applied Human Factors and Ergonomics*, pp. 158–168, Springer, 2018.

[6] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 42–55, 2011.

[7] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2011.

[8] S. Koldijk, M. Sappelli, S. Verberne, M. A. Neerincx, and W. Kraaij, "The SWELL knowledge work dataset for stress and user modeling research," in *16th International Conference on Multimodal Interaction*, pp. 291–298, 2014.

[9] D. S. Lee, T. W. Chong, and B. G. Lee, "Stress events detection of driver by wearable glove system," *IEEE Sensors Journal*, vol. 17, no. 1, pp. 194–204, 2017.

[10] J. Birjandtalab, D. Cogan, M. B. Pouyan, and M. Nourani, "A non-EEG biosignals dataset for assessment and visualization of neurological status," in *IEEE International Workshop on Signal Processing Systems*, pp. 110–114, 2016.

[11] B. Winslow, M. B. Carroll, J. W. Martin, G. Surpris, and G. L. Chadderdon, "Identification of resilient individuals and those at risk for performance deficits under stress," *Frontiers in Neuroscience*, vol. 9, p. 328, 2015.

[12] M. Seet, M. R. Amin, N. I. Abbasi, J. Hamano, A. Bezerianos, R. T. Faghih, N. Thakor, and A. Dragomir, "Olfactory-induced positive affect and autonomous response as a function of hedonic and intensity attributes of fragrances," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2020.

[13] P. A. Low, "Chapter 51 - sweating," in *Primer on the Autonomic Nervous System (Third Edition)* (D. Robertson, I. Biaggioni, G. Burnstock, P. A. Low, and J. F. Paton, eds.), pp. 249 – 251, San Diego: Academic Press, third edition ed., 2012.

[14] H. D. Critchley, R. N. Melmed, E. Featherstone, C. J. Mathias, and R. J. Dolan, "Volitional control of autonomic arousal: a functional magnetic resonance study," *Neuroimage*, vol. 16, no. 4, pp. 909–919, 2002.

[15] M. Benedek and C. Kaernbach, "A continuous measure of phasic electrodermal activity," *Journal of neuroscience methods*, vol. 190, no. 1, pp. 80–91, 2010.

[16] M. R. Amin and R. T. Faghih, "Sparse deconvolution of electrodermal activity via continuous-time system identification," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 9, pp. 2585–2595, 2019.

[17] M. R. Amin and R. T. Faghih, "Robust inference of autonomic nervous system activation using skin conductance measurements: A multichannel sparse system identification approach," *IEEE Access*, vol. 7, pp. 173419–173437, 2019.

[18] D. S. Wickramasuriya and R. T. Faghih, "A marked point process filtering approach for tracking sympathetic arousal from skin conductance," *IEEE Access*, vol. 8, pp. 68499–68513, 2020.

[19] D. S. Wickramasuriya and R. T. Faghih, "A mixed filter algorithm for sympathetic arousal tracking from skin conductance and heart rate measurements in Pavlovian fear conditioning," *Plos One*, vol. 15, no. 4, p. e0231659, 2020.

[20] J. J. Braithwaite, D. G. Watson, R. Jones, and M. Rowe, "A guide for analysing electrodermal activity (EDA) & skin conductance responses (SCRs) for psychological experiments," *Psychophysiology*, vol. 49, no. 1, pp. 1017–1034, 2013.

[21] M. R. Amin and R. T. Faghih, "Identification of sympathetic nervous system activation from skin conductance: A sparse decomposition approach with physiological priors," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 5, pp. 1726–1736, 2020.

[22] M. Benedek and C. Kaernbach, "Decomposition of skin conductance data by means of nonnegative deconvolution," *Psychophysiology*, vol. 47, no. 4, pp. 647–658, 2010.

[23] D. R. Bach, G. Flandin, K. J. Friston, and R. J. Dolan, "Modelling event-related skin conductance responses," *International Journal of Psychophysiology*, vol. 75, no. 3, pp. 349–356, 2010.

[24] A. Greco, G. Valenza, A. Lanata, E. P. Scilingo, and L. Citi, "cvxEDA: A convex optimization approach to electrodermal activity processing," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 797–804, 2016.

[25] D. S. Wickramasuriya, M. Amin, and R. T. Faghih, "Skin conductance as a viable alternative for closing the deep brain stimulation loop in neuropsychiatric disorders," *Frontiers in Neuroscience*, vol. 13, p. 780, 2019.

[26] D. R. Bach, J. Daunizeau, N. Kuelzow, K. J. Friston, and R. J. Dolan, "Dynamic causal modeling of spontaneous fluctuations in skin conductance," *Psychophysiology*, vol. 48, no. 2, pp. 252–257, 2011.

[27] S. Subramanian, R. Barbieri, and E. N. Brown, "Point process temporal structure characterizes electrodermal activity," *bioRxiv*, 2020.

[28] J. A. Bosch, H. S. Brand, T. J. Ligtenberg, B. Bermond, J. Hoogstraten, and A. V. N. Amerongen, "Psychological stress as a determinant of protein levels and salivary-induced aggregation of streptococcus gordonii in human whole saliva," *Psychosomatic Medicine*, vol. 58, no. 4, pp. 374–382, 1996.

[29] E. Wressle and K. Samuelsson, "High job demands and lack of time: A future challenge in occupational therapy," *Scandinavian Journal of Occupational Therapy*, vol. 21, no. 6, pp. 421–428, 2014.

[30] M. R. Amin and R. T. Faghih, "Tonic and phasic decomposition of skin conductance data: A generalized-cross-validation-based block coordinate descent approach," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 745–749, 2019.

[31] D. S. Wickramasuriya, C. Qi, and R. T. Faghih, "A state-space approach for detecting stress from electrodermal activity," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3562–3567, 2018.

[32] D. S. Wickramasuriya and R. T. Faghih, "A Bayesian filtering approach for tracking arousal from binary and continuous skin conductance features," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 6, pp. 1749–1760, 2019.

[33] D. Wickramasuriya and R. T. Faghih, "A novel filter for tracking real-world cognitive stress using multi-timescale point process observations," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 599–602, 2019.

[34] J. Healey and R. W. Picard, "Driver stress data," *Retrieved June 26th from MIT Affective Computing Group: http://affect. media. mit. edu*, vol. 124, 2002.

[35] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Transactions on Intelligent Transportation Systems*, vol. 6, no. 2, pp. 156–166, 2005.

[36] H. Fekri Azgomi, D. S. Wickramasuriya, and R. T. Faghih, "State-space modeling and fuzzy feedback control of cognitive stress," in *Annual International Conference of the Engineering in Medicine and Biology Society*, July 2019.

[37] H. Fekri Azgomi and R. T. Faghih, "A wearable brain machine interface architecture for regulation of energy in hypercortisolism," in *IEEE Asilomar Conference on Signals, Systems, and Computers*, pp. 254–258, November 2019.

[38] R. T. Faghih, M. A. Dahleh, and E. N. Brown, "An optimization formulation for characterization of pulsatile cortisol secretion," *Frontiers in Neuroscience*, vol. 9, p. 228, 2015.