Detecting Physiological Stress in Pregnant Women using a wireless mHealth sensing system and working towards providing a Just-In-Time Intervention

Wanxin Xu[1], Shuyi Han[1], Andrew Dai[2], Joel Mintz[2], Kristen Rosen[3], Nabil Alshurafa[1,4]

[1]Department of Electrical Engineering and Computer Science, McCormick School of Engineering, Northwestern University, Evanston, IL., [2]Department of Biomedical Engineering, McCormick School of Engineering, Northwestern University, Evanston, IL., [3]Department of Medicinal and Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL., [4]Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL.

Abstract **- Physiological stress experienced during pregnancy has previously been demonstrated as a major factor in the development of negative health risks for both the mother and the fetus. In this paper, we propose a mobile health (mHealth) system using a combination of accelerometer and ECG signals collected using a BioStampRC sensor on the medial chest along with micro-ecological momentary assessments (micro-EMAs) presented on an LG smartwatch and paired with an Android smartphone to develop, train, and test a stress recognition system capable of detecting moments of physiological stress in pregnant women. We trained and tested the performance of 7 classifiers, including a Support Vector Machine (SVM) and a Multilayer Neural Network. Our SVM classifier performed the best, achieving a 92.36% accuracy with an F-score of 0.9234 when using a personalized model and a 72.67% accuracy with an F-score 0f 0.7260 when trained on a generalized model. We hope that this system can be expanded to provide instantaneous, "just-in-time" interventions through the smartwatch/smartphone interface that are capable of reducing physiological stress and improving the quality of life for both the mother and the fetus.**

I. INTRODUCTION

Pre-perinatal "early life" physiological stress (ELS) has significant consequences for mothers and their infants [**1**,**2**,**3**]. Prolonged prenatal stress places pregnant women at increased risk for depression, health risk behaviors, and poor overall health and quality of life. Likewise, fetal exposure to stress places infants at greater risk of neurodevelopmental disruptions [**4**,**5**]. Prior clinical trials have demonstrated that pregnant women, particularly those in low income environments, demonstrate positive health and parenting benefits as a result of receiving prenatal stress reduction interventions [**6**,**7**,**8**].

Previous research has shown that it is possible to develop stress monitoring systems for both in-lab testing and field testing [**9**]. Though the stress classification algorithms developed by Hovsepian et al, have shown to good results overall, they still lack long term usability and wearability for monitoring stress over the length of a gestation cycle. However, advances in flexible epidermal electronic sensing systems, pioneered by Dr. John Rogers, have shown promise in reducing the limitations of existing mobile sensing platforms [**10**]. One such biosensing

technology is the BioStampRC sensor. Long term stress monitoring represents an ideal use of such technology. The sensor is extremely unobtrusive, easy to use, simple to wear, and can adapt to changing physiologic and environmental conditions, all while providing continuous monitoring of quantifiable physiological parameters such as: ECG, accelerometry, and gyrometric data. The wireless sensor will measure the real-time stress response and develop a model for identifying ECG features which correspond with moments of stress. The success and challenges of this model can then be applied towards developing a "just-in-time" intervention during moments of stress.

The short-term goal of this paper is to develop, implement, and test a mobile health (mHealth) system that is able to provide continuous physiological stress monitoring using a combined sensor / smartphone platform that. Once this has been developed and refined, the long-term goal will be to expand the system so that, when triggered, it can provide an appropriate stress reduction intervention in a real world environment. [**11**,**12**]

For our short-term goal, a number of data features calculated from acceleration and electrocardiography (ECG) data collected by the BioStampRC will be compared against the perceived stress levels as measured by a micro-ecological momentary assessment (micro-EMA) provided by the a smartwatch. We hope to validate this model in a small scale study of healthy women, which will serve as the control group for future clinical studies in pregnant women.

II. RELATED WORKS

Stress Assessment

Stress assessment is a widely studied field with numerous papers conducting experiments both in lab and out in the field. The cStress paper presented their own stress model and analyzed its efficacy and accuracy in detecting physiological stress using a number of health signals, with ECG and respiration data via via inductive plethysmography being the two most informative signals for measuring health [**11**]. This provides a good foundation of our choice of ECG as a metric for stress detection. Psychiatric health has also been investigated as a measurement of mental health stress using analysis of GPS, sleep, and acceleration health signals as well as the PHQ-9 [**12**]. The feasibility of continuously collecting data from mobile wireless sensors has also been previously investigated over a 4-week period, concluding that each day has around 11 hours of usable data, an important consideration when moving towards continuous stress monitoring while also attempting to balance memory and battery concerns on whichever sensor is used [**13**].

Using these previous studies looking at stress assessments, this will allow us to further explore both our short-term and long-term objectives with the ultimate long-term potential to combine different sensors, each measuring different data, in a manner which might optimize battery and storage for all the sensors used. For example, if we use GPS data and notice that the individual has not changed locations significantly in the past hour and that it is night time, we may end up relying solely on ECG data and sleep quality data instead of accelerometry in order to save memory and battery. These previous studies give us insight into the applicability of the sensors

we have chosen and the type of data they are collecting for the development and refinement of our own study.

Pregnancy and Stress Detection

A handful of papers published have also delved into the effects of pregnancy and how it affects any collected health data, such as ECG or accelerometry. Osma et al. investigated the viability and potential benefits of implementing a smartphone based sensor system to detect postpartum depression [**14**]. The major takeaway from this paper was that the use of a mobile phone app would be most beneficial for mothers after birth in regions with poor healthcare availability, such as rural areas since it would provide a link between the mother who would otherwise not have good access to a clinician with some sort of objective measurement of stress.

A different study also looked at the implementation of an offline stress detection system using a smartphone based sensor platform and a foetal non-stress test, named non-stress in that it does not place any additional stressors on the fetus during the test [**15**]. This paper was able to detect moments of physiological stress using ECG data from the fetus without having to rely on an internet connection by using an offline model, which again provides significance in how a working, stress detection and intervention model would expand healthcare access to regions with poor access to healthcare.

There is also a currently ongoing investigation into stress detection after birth, looking at whether stress in new mothers is any different than during pregnancy and whether this difference can be picked up by a Microsoft Band 2 smartwatch and the Experience Sampling Method and a Daily Diary as mobile phone apps [**16**].

Of note, however, is that there is currently a space for further investigation into stress during pregnancy specifically looking at stress detection of the pregnant mother and the long-term development of a stress reduction intervention.

Wireless Sensors

The BioStampRC sensor (MC10, Inc., Lexington, MA, USA) has been previously used in many studies due to its ease in conforming to the skin and its ability to measure a wide variety of physiological variables, including raw ECG, gyroscope, and accelerometer data [**17**]. The BioStampRC sensor was previously used to monitor epileptic seizures through accelerometer, gyroscope, and ECG data [**18**]. The BioStampRC has also been used in the automatic detection of spasticity in patients with stroke. A study of 13 people with chronic stroke found that the BioStampRC is able to compute a set of EMG features which, when combined with a machine learning classifier, is able to infer whether spasticity is present during movements of the knee and ankle joints [**19**]. The BioStampRC sensor has also been previously used and validated in its ability to accurately collect data from patients with Huntington's Disease regarding activity recognition at home while remaining unobtrusive and comfortable, and also provide data accurate enough to differentiate between activity data in patients with Huntington's Disease compared to patients without Huntington's Disease [**20**,**21**].

There is room in the field to expand the BioStampRC sensor towards the field of stress detection, as most of the past investigations have dealt with classifying

movement or neurological disorders, such as Huntington's Disease, stroke, and epilepsy [**17**,**18**,**19**,**20**,**21**]. Our goal with this project is to also validate our choice of sensor to see whether the BioStampRC is also capable of providing good detection and classification of physiological stress.

The LG Watch Sport (LG Electronics, Seoul, South Korea) is a type of Android smartwatch. The use of a smartwatch paired with a smartphone to collect EMAs from individuals as an accurate method to obtain self-reports from the wearer has been a hot topic of previous study. EMAs have been previously demonstrated EMAs have been demonstrated as an effective way to collect a subjective self-report about health metrics such as stress, which can be incorporated into the successful development of ecological momentary interventions (EMIs) at the appropriate time [**22**,**23**,**24**,**25**]. Of note, EMAs have been used to detect moments of binge eating and administer instances of cognitive-behavioral therapy (CBT), monitor and detect panic attacks and administer CBT, and even smoking detection so as to attempt to aid individuals to quit smoking [**26**,**27**,**28**,**29**]. The inclusion of an LG Watch Sport smartwatch will allow us to collect EMA information from the subjects during our study, which we can later use to match with the features from the BioStampRC data. The long-term goal is to transform the EMAs so as to also provide an EMI capable of reducing stress, once we detect moments of physiological stress.

III. BACKGROUND

There is some debate as to the appropriate ECG sampling rate required to accurately extract features such as R-R intervals and heart rate variability (HRV) from raw ECG data. Kuusela (2013) recommends a minimum choice of 200 Hz, however Berntson et al. (2007) and other literature in the field recommends a minimum of 500 - 1000 Hz [**30**,**31**,**32**]. Many studies looking at HRV, R-R intervals, and power spectrum density (PSD) features all chose a minimum sampling rate of 1000 Hz for ECG data collection which produced good quality results when calculating the ECG features from the collected raw ECG data [**33**,**34**].

For our study we chose a sampling rate of 500 Hz for our raw ECG data collection. Based on previous literature, we chose to look at the following features: very low frequency power (VLF: $0.0033 \leq f < 0.04$ Hz), low frequency power (LF: $0.04 \leq f < 0.15$ Hz), high frequency power (HF: $0.15 \leq f < 0.40$ Hz), the low frequency power / high frequency power ratio, the total power (VLF power + LF power + HF power), and the R-R intervals in milliseconds, along with the variance, mean, median, max, kurtosis, skew, and the overall heart rate [**11**,**12**,**14**,**33**,**35**].

IV. SYSTEM

The main sensor that will be used for this study is the MC10 BioStampRC sensor. This sensor will be coupled with an LG Watch Sport smartwatch for field studies. The BioStampRC sensor has the capacity to store biometric data on the sensor locally, including ECG, EMC, acceleration, and gyroscope data. In addition, subjective measures of stress will be collected through self-reported mirco-EMAs presented through a mHealth app developed for the LG Watch Sport smartwatch. These micro-EMA's will be provided every hour to the subject and consist of one question that asks whether the subjects are stressed

(yes/no) and if so, asks them to rank their stress on an integer scale from 0 - 6.

Our long-term objective is to expand our system to include an online data processing platform that can deliver a "just-in-time" intervention, however our current study does not allow for online processing due to the fact that the BioStampRC does not allow real-time data processing. Such an intervention might come in the form of a positive message, mediation reminder or a clever meme. Thus the system user can seek to reduce stress, providing better population outcomes to unborn children and their mothers, especially for those of lower socioeconomic status.

V. EXPERIMENTAL SETUP

5 female subjects with a mean age of 25.4 (sd +/- 2.7) used our sensing system. Before each experiment each subject was given thorough instructions on how to use an LG smartwatch, its corresponding charger and how to correctly fill out a micro-EMA using a custom mHealth Android application.

To begin the experiment subjects were asked to fill out a 10 question perceived stress survey (PSS) and input their biometric data into the Biostamp mobile application housed on a Samsung Galaxy tablet. Collected biometric information included the subject's age, height, weight, gender and handedness. Before filling out the PSS, each subject was allowed to read the terms and conditions of the Biostamp and was asked to acknowledge that they had read the terms within Biostamp mobile app. No subjects refused the terms and conditions of the study.

Afterwards, the Biostamp was placed on the medial chest of each subject. To pinpoint the eventual placement of the sensor, the suprasternal notch was identified on each subject and the sensor was placed 2-3 inches inferiorly of the suprasternal notch, depending on the height of the subject. The Biostamp was secured to the skin with tape. Each subject was then asked to wear the LG watch and fill out a pre-study initial micro-EMA on the LG watch before data collection was initialized. ECG data for subject's one and two were collected at 1000 Hz, while ECG data for subjects, three through five were collected with an ECG sampling rate of 500 Hz. The lower sampling rate allowed the subjects to wear the sensor for a maximum of approximately 9 hours. Subjects wore the Biostamp sensor until one of three events occurred: the subject grew uncomfortable with the sensor on their chest, the local memory on was filled, or the data collection period was ended. The total length of experiment time varied from four to nine hours for each subject. After each experiment, data was then uploaded from the Biostamp to the MC10 Cloud Platform and from the LG watch to a computer.

VI. ALGORITHM

All preprocessing, feature extraction, and data analysis was conducted in Python 3.4. Under the PASDAC framework, we used the `scikit-learn` library to import the machine learning classifier functions we used to train and test our model. We also used the `biosspy` library to import the ECG signal processing and smoothing functions we used during preprocessing and features extraction.

Preprocessing

Data was collected from the BioStampRC sensor at a sampling rate of both 1000 Hz

and 500 Hz. This discrepancy is due to our initial choice of 1000 Hz, however, in order to collect longer periods of subject data we changed to a sampling rate of 500 Hz. The data was then smoothed to reduce the effects of noise. Finally, the cleaned data was resampled to 250 Hz, which was used in the data analysis. Downsampling was chosen to reduce the dataset to the sampling rate as recommended by Kuusela [**30**].

The next step was to label the obtained data based on the self-report from micro-EMAs obtained from the LG Watch Sport which will act as our ground-truth labels. In this study, micro-EMAs were provided every hour to the subject. We selected a time window of 5 mins before and after the micro-EMA was prompted and answered, giving us a total of 10 minutes of data corresponding to the micro-EMA response.

Smoothing

For smoothing our data, we used the `biosspy.signals.tools.smoother` function. The input signal is smoothed using an N-point moving average filter. Our implementation uses the convolution of a filter kernel with the input signal to compute the smoothed signal [**36**]. The kernels we used are median, boxcar, and gaussian.

Segmentation

For data segmentation we used a sliding window with a size of 10 seconds and a step size of 1 second. The segments that crossed over discontinuous data were removed and segments that had overlapping labels were sliced in order to enlarge our training dataset.

Feature extraction

For ECG features extraction, we utilized functions from the `biosspy` module `biosppy.signals.ecg`. We used the `engzee_segmenter` function to extract R-peaks from the preprocessed ECG data, which was based off previous literature published by Engelse and Zeelenberg along with Lourenco [**37**,**38**]. The R-R interval was then calculated by finding time difference between R peaks in our given data window. For windows with 3 or more R peaks, we calculated the simple average of R-R intervals within the window as the R-R interval representative of that window. This average value was then associated with the PSD and non-HRV features.

For the PSD features, we computed the average power within our desired frequency range using equations 1-3, listed below. Table 1 lists the frequencies corresponding to the PSD features that we limited our calculations to.

$$P_{\text{bandlimited}} = 2\int_{f_1}^{f_2} S_{xx}(2\pi f)\,df = \frac{1}{\pi}\int_{\omega_1}^{\omega_2} S_{xx}(\omega)d\omega$$

Equation 1. Calculation of power in a specific band frequency from f_1 to f_2. S_{xx} is defined in equation 2 below.

$$S_{xx}(\omega) = \lim_{T\to\infty} \mathbf{E}\left[|\hat{x}_T(\omega)|^2\right].$$

Equation 2. The equation for calculating the power spectral density, from which we can use to find the power within a specific frequency band.

$$\hat{x}_T(\omega) = \frac{1}{\sqrt{T}}\int_0^T x(t)e^{-i\omega t}\,dt.$$

Equation 3. The Fourier Transform of time domain data x(t) to frequency data x(ω) within the time interval 0 to T. This step is performed on all of our time domain preprocessed ECG data, in order to properly calculate the power spectral density and the power within a specific band frequency.

Table 1 below lists all the features we extracted from our preprocessed ECG data.

Table 1: All aggregated ECG features

HRV	Very low frequency power (VLF: $0.0033 \leq f < 0.04$ Hz) Low frequency power (LF: $0.04 \leq f < 0.15$ Hz) High frequency power (HF: $0.15 \leq f < 0.40$ Hz) LF/HF ratio R-R intervals (milliseconds)
Non-HRV	Variance Standard Deviation Absolute Deviation Mean Median Max Heart Rate Kurtosis Skew

Feature normalization

To eliminate the intra-individual factors of the signals, we normalize all the features feature according to equation 4 below.

$$F(i)_{norm} = \frac{F(i) - F_{min}(j)}{F_{max}(j) - F_{min}(j)}$$

Equation 4. Normalization of our features.

Classification

Seven classifiers were used in our experiment which include: Naive Bayes, k-nearest neighbors (KNN), Random Forests (RF), Support Vector Machine (SVM), AdaBoost classifier, Multilayer Neural Network, and Logistic Regression. The performance of the various classifiers we used will be discussed in section V. Results. Both personalized models and generalized models were computed based on the 13 extracted features and the classifier we selected.

Evaluation metrics

Our personalized model was tested on the equally split dataset while the generalized model was validated by using leave-one-subject-out cross-validation.

The metric evaluates the classifier by the total amount of time correctly classified. A confusion matrix was constructed to summarize how many instances of an activity class got confused (misclassified) by the system. Typically, rows correspond to Ground Truth, and columns to Predicted. An example of our output confusion matrix is shown in Table 2 below.

Table 2: Confusion matrix

Truth		Predicted	
		Stressed	Non-stressed
	Stressed	The number of seconds of time that is labeled as being stressed and classified as stressed (TP)	The number of seconds of time that is labeled as being stressed and classified as non-stressed (FN)
	Non-Stressed	The number of seconds of time that is labeled as being non-stressed and classified as stressed (FP)	The number of seconds of time that is labeled as being non-stressed and classified as non-stressed (TN)

The sensitivity and specificity will be calculated using equations 5 and 6 below.

$$sensitivity = \frac{TP}{TP+FN}$$

Equation 5. Sensitivity equation used

$$specificity = \frac{TN}{TN+FP}$$

Equation 6. Specificity equation used

Finally, the overall accuracy was calculated using equation 7 below.

$$accuracy = \frac{TP \times r + TN}{(TP+FN) \times r + (TN+FP)}$$

Equation 7. Accuracy equation used

The ratio r in equation 7 weights TP to TN This biases our classifier accuracy in its ability to detect stress rather than not-stressed moments, since we are more concerned with accurately identifying moments of stress.

To evaluate the performance, a set of standard performance measures were used including the F1 score, area under ROC curve (AUC), and the weighted accuracy.

VII. RESULTS

We constructed a personalized model based on one subject's data and a generalized model based on all the subject data collected.

Personalized model

Subject 1 was chosen to be our dataset and it was equally split into training test and testing set. We built our personalized model based on the training dataset using all 13 features and tested the model on the separated testing set.

Table 3 illustrates the results of each of our classifiers.

Table 3. Results of our 7 classifiers after being tested and trained on subject 1 data.

Classifiers	**F1 score**	**AUC**	**Accuracy**
KNN	0.7770	0.7818	0.7826
RF	0.7650	0.7719	0.7691
SVM	0.9234	0.9227	0.9236
AdaBoost	0.7501	0.7541	0.7521
Naive Bayes	0.8047	0.8043	0.8056
Neural Network	0.9140	0.9133	0.9143
Logistic Regression	0.7753	0.7771	0.7759

As is shown above in Table 3, the SVM with a RBF kernel and penalty parameters C = 1.0, $\gamma = 0.077$ obtained the best result.

Table 4. Results of our 7 classifiers after being tested and trained on subject 1 data using only NON-HRV features.

Classifiers	**F1 score**	**AUC**	**Accuracy**
KNN	0.7659	0.7658	0.7674
RF	0.6009	0.6104	0.6112
SVM	0.8273	0.8268	0.8276

AdaBoost	0.6803	0.6836	0.6837
Naive Bayes	0.7769	0.7820	0.7860
Neural Network	0.6391	0.6544	0.6555
Logistic Regression	0.6737	0.6876	0.6848

Table 5. Results of our 7 classifiers after being tested and trained on subject 1 data using only HRV features.

Classifiers	**F1 score**	**AUC**	**Accuracy**
KNN	0.6504	0.6784	0.6787
RF	0.8081	0.8119	0.8098
SVM	0.8641	0.8643	0.8641
AdaBoost	0.5776	0.6262	0.6265
Naive Bayes	0.5546	0.5805	0.5772
Neural Network	0.8344	0.8349	0.8344
Logistic Regression	0.5891	0.6171	0.6173

Figure 1. Feature importance for Random Forests classifier

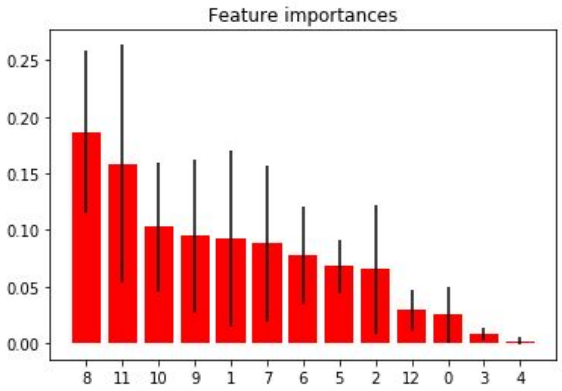
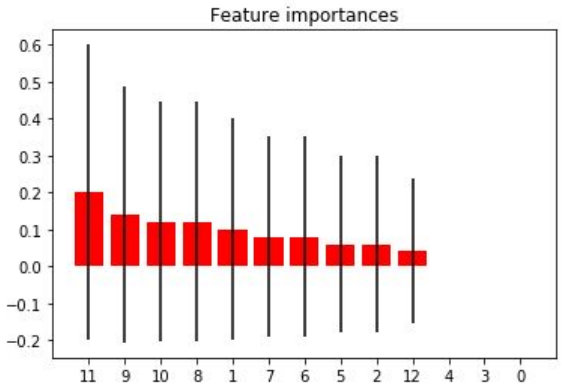


Figure 2. Feature importance for AdaBoost classifier



The labels 0-12 under the red bars represent features which are: mean, median, max, variance, standard deviation, absolute deviation, kurtosis, skew, very low frequency power, low frequency power, high frequency power, LF/HF ratio, and R-R intervals respectively.

In Random Forests and AdaBoost classifier forests of trees can be used to evaluate the importance of features on a classification task. The red bars are the feature importances of the forest, along with their inter-trees variability.

The first plot shown in Figure 1 suggests that for Random Forests classification, the

VLF power and the LF/HF ratio features are the most informative, while the remaining are not. In the AdaBoost classifier however, Figure 2 shows that the LF power and the LF/HF power are the 2 most important features for classification. Among both classifiers, however, the same 4 features (VLF power, LF power, HF power, LF/HF ratio) show up as the top 4 important features albeit in slightly different ranks of importance.

Generalized model

We computed our generalized model using data from Subject 2, Subject 3 and Subject 4 and used Subject 1 data for testing.

Parameter fine-tuning

Table 6. Results of generalized model tested on subject 1 data using SVM (RBF kernel).

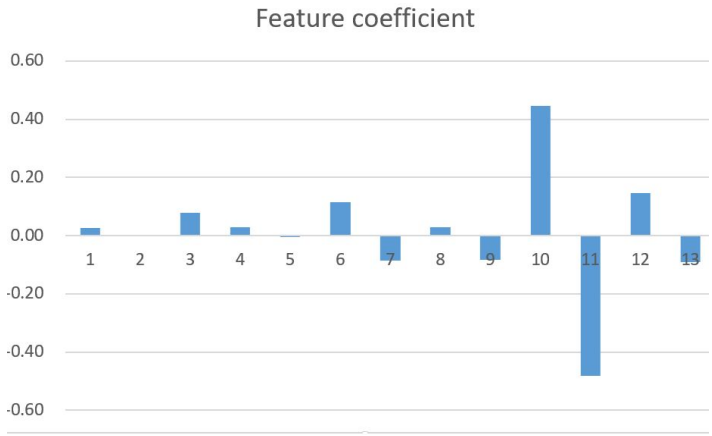
Parameters in SVM	F1 score	accuracy
C = 1, γ = 0.077	0.6655	0.6700
C = 2, γ = 0.01	0.7260	0.7267
C = 2, γ = 0.0055	0.7231	0.7241

Table 6 lists the values of all the performance measures we used for leave-one-subject-out validation on Subject 1, for both categories of features. The table also lists the optimal hyper-parameters: C and γ. For additional reference, Table 7 contains the confusion matrix of the validation test using the optimal hyper-parameters and all features.

Table 7. Confusion matrix (C = 2, γ = 0.01).

	Non Stressed	stressed
Non Stressed	911	261
stressed	386	809

Figure 3. Feature coefficient in Linear SVM



In order to present a ranking of features in terms of contribution to the model performance we employed the Linear SVM classifier, to rank features based on their associated Kernel weight coefficients. As is depicted in Figure 3, the low frequency power and high frequency power contribute most to the constructed model.

VIII. LIMITATIONS / FUTURE STEPS

This work has several limitations and significant potential for future works. Data accuracy could be improved using more convenient data collection methods (e.g. obtaining inter-beat intervals from LG watch instead of BioStampRC which requires the data to be uploaded before any analytics can be performed). Better control of physical activity during the in-field experiment would also result in fewer data segments being removed due to electrode contact issues or motion artifacts.

Additionally, accelerometer data obtained from BioStampRC as well as the LG Sport

Watch could be used in the future to help detect the movements of one subject and analyze the relation between the subject's movement and heart rate. The combination of such data would help to increase the accuracy of the data analysis system, as well as potentially improve differentiation between moments of stress as seen by ECG features, and moments of high physical activity which might result in similar ECG features in the signal.

The quality of the ECG signal collected by the BioStampRC is also heavily reliant on good quality contact between the electrodes and the patient's skin. This is difficult to monitor in an in-field study as most patients will naturally move enough such that the tape is loosened and the sensor moves. The electrode gel applied to the electrodes on the BioStampRC sensor can also run into these same issues as the patient moves around and sweats, which can interfere with the quality of the collected signal.

One major limitation is that our collected dataset was significantly imbalanced. Many subjects only reported all stress instances when responding to the EMA, while others reported all non-stress instances. Two of the subjects reported the same condition throughout the entire length of the 8 hour in-field study. Even when pooled together to make a generalized model, there were not enough instances of non-stress overall in order to properly train and test our classifiers. This could be addressed by increasing the time of data collection in the study. However, it is difficult to ensure a ground truth level of non-stress since even while a subject may be doing relaxing activities such as eating or watching TV, they may be stressed out about upcoming work or deadlines which we cannot truly eliminate. This is one area that we are interested in further exploring for our future studies and improvements.

Another major limitation to translating this work into consistent, reliable at-home monitoring is that the BioStampRC sensor and tablet software interface are still not very user friendly. Starting and stopping data collection is somewhat cumbersome, with the sensor frequently losing Bluetooth connection with the tablet. Additionally, uploading of the data collected took on average 2 - 4 hours and had to be manually started and monitored for issues with bluetooth connection or server timeouts, which significantly impedes the ease of use in an at-home setting. Since the quality of the signal depends on the quality of sensor contact and attachment, placing the BioStampRC sensor on the medical chest may prove difficult for subjects to perform at home without additional help, and is also susceptible to variability in placement location and placement quality.

Further investigation into effective methods of presentation or visualizations of our collected data and present the results to pregnant women in a meaningful way which would help them understand their levels of physiological stress and potentially adjust their lifestyle accordingly. Additionally, we hope to expand the system to include an online data processing platform that can deliver a "just-in-time" intervention. However, this is not feasible using the BioStampRC sensor in its current form. Any data analytics on any sensing modalities (accelerometer, ECG/EMG, gyroscope) that the sensor can collect can only be performed after the data has been manually uploaded online. A good future direction of this work would be the investigation of other ECG-capable sensors which could be integrated with real-time data analytics, in

order to work towards providing a stress reduction EMI for pregnant women.

IX. CONCLUSION

We were able to train an SVM classifier for both personalized and generalized models, resulting in an accuracies of 92.36% and 72.67% with F-scores of 0.9234 and 0.7260, respectively. From our analyzed dataset, the two most important features were very low frequency power and the LF/HF ratio for the personalized model and low frequency power and LF/HF ratio for the generalized model. Future work is needed to expand the scope of this project in order to develop EMIs capable of providing a "just-in-time" stress intervention in a population of pregnant women. Further investigation into a ECG sensor capable of integrating with a real time data analytics program is also necessary, since the BioStampRC sensor in its current form cannot perform real-time data analysis and would not be feasible for developing or providing EMIs capable of stress reduction.

ACKNOWLEDGEMENTS

We would like to thank Professor Nabil Alshurafa, Shibo Zhang, Dzung Nguyen, Zachary King, Lida Zhang, the HABits lab, the Department of Electrical Engineering and Computer Science, and the Feinberg School of Medicine, all at Northwestern University, for their assistance and guidance and in making this project possible. We would also like to thank our research participants as well as members from the Max Näder Lab for Rehabilitation Technologies and Outcomes and the Lab of Dr. Ferdinando Mussa-Ivaldi, both at the Shirley Ryan AbilityLab, for their guidance, assistance, and feedback regarding some aspects of this project.

REFERENCES

1) Talge, N. M., Neal, C., & Glover, V. (2007). Antenatal maternal stress and long‐term effects on child neurodevelopment: how and why?. *Journal of Child Psychology and Psychiatry*, *48*(3‐4), 245-261.
2) Kinsella, M. T., & Monk, C. (2009). Impact of maternal stress, depression & anxiety on fetal neurobehavioral development. *Clinical obstetrics and gynecology*, *52*(3), 425.
3) O'connor, T. G., Heron, J., Golding, J., & Glover, V. (2003). Maternal antenatal anxiety and behavioural/emotional problems in children: a test of a programming hypothesis. *Journal of child Psychology and Psychiatry*, *44*(7), 1025-1036.
4) de Bruijn, A. T., van Bakel, H. J., & van Baar, A. L. (2009). Sex differences in the relation between prenatal maternal emotional complaints and child outcome. *Early human development*, *85*(5), 319-324.
5) Grizenko, N., Fortier, M. E., Zadorozny, C., Thakur, G., Schmitz, N., Duval, R., & Joober, R. (2012). Maternal stress during pregnancy, ADHD symptomatology in children and genotype: gene-environment interaction. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *21*(1), 9.
6) Mendelson, T., Leis, J. A., Perry, D. F., Stuart, E. A., & Tandon, S. D. (2013). Impact of a preventive intervention for perinatal depression on mood regulation, social support,

and coping. *Archives of women's mental health*, *16*(3), 211-218.

7) Tandon, S. D., Leis, J. A., Mendelson, T., Perry, D. F., & Kemp, K. (2014). Six-month outcomes from a randomized controlled trial to prevent perinatal depression in low-income home visiting clients. *Maternal and child health journal*, *18*(4), 873-881.
8) Tandon, S. D., Perry, D. F., Mendelson, T., Kemp, K., & Leis, J. A. (2011). Preventing perinatal depression in low-income home visiting clients: a randomized controlled trial. *Journal of consulting and clinical psychology*, *79*(5), 707.
9) Hovsepian, Karen et al. "Stress: Towards a Gold Standard for Continuous Stress Assessment in the Mobile Environment." *Proceedings of the ... ACM International Conference on Ubiquitous Computing . UbiComp (Conference)*2015 (2015): 493–504. *PMC*. Web. 10 Feb. 2018.
10) Kim, Dae-Hyeong, Nanshu Lu, Fiorenzo G. Omenetto, Yonggang Huang, Todd Coleman, and John Rogers. "Epidermal Electronics." Science 333 (2012): 838-43.
11) Hovsepian, Karen, et al. "cStress: towards a gold standard for continuous stress assessment in the mobile environment." *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. ACM, 2015.
12) Ben-Zeev, Dror, et al. "Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health." *Psychiatric rehabilitation journal* 38.3 (2015): 218.
13) Rahman, Md Mahbubur, et al. "Are we there yet?: Feasibility of continuous stress assessment via wireless physiological sensors." *Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*. ACM, 2014.
14) Osma, Jorge, et al. "Proposal of use of smartphones to evaluate and diagnose depression and anxiety symptoms during pregnancy and after birth." *Biomedical and Health Informatics (BHI), 2014 IEEE-EMBS International Conference on*. IEEE, 2014.
15) Nasim, K. O. Z. O. U. I. E., FOTOHI-GHAZVINI Faranak, and MINAIE-BIDGOLI Behrouz. "Pervasive and Ubiquitous Mobile-health application for pregnancy's mother on android smartphone (Non-Stress Test application)." *Bulletin de la Société Royale des Sciences de Liège* 85 (2016): 409-417.
16) McAvoy, A., Mensching, D., Alshehri, M., & Connelly, K. Detecting Stress in Pregnant Women and New Mothers.
17) Sonawane, A., Manickam, P., & Bhansali, S. (2017). Stability of Enzymatic Biosensors for Wearable Applications. *IEEE reviews in biomedical engineering*, *10*, 174-186.
18) Samson, Kurt. "Wearing the Detectives." Neurology Now, vol. 11, no. 4, 2015, pp. 34–36., doi:10.1097/01.nnn.0000470763.57000.87.
19) Lonini, Luca, et al. "Automatic Detection of Spasticity from Flexible Wearable Sensors."*Proceedings of the 2017 ACM International Joint*

Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers on - UbiComp '17, 2017, doi:10.1145/3123024.3123098.

20) Dinesh, Karthik, et al. "Signal analysis for detecting motor symptoms in Parkinson's and Huntington's disease using multiple body-affixed sensors: A pilot study." *Signal Processing Workshop (WNYISPW), 2016 IEEE Western New York Image and*. IEEE, 2016.
21) Adams, Jamie L., et al. "Multiple Wearable Sensors in Parkinson and Huntington Disease Individuals: A Pilot Study in Clinic and at Home." *Digital Biomarkers* 1.1 (2017): 52-63.
22) Blaauw, Frank J., et al. "Let's get Physiqual–An intuitive and generic method to combine sensor technology with ecological momentary assessments." *Journal of biomedical informatics* 63 (2016): 141-149.
23) Marzano, Lisa, et al. "The application of mHealth to mental health: opportunities and challenges." *The Lancet Psychiatry* 2.10 (2015): 942-948.
24) Heron, Kristin E., and Joshua M. Smyth. "Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments." *British journal of health psychology* 15.1 (2010): 1-39.
25) Riley, William T., et al. "Mobile and wireless technologies in health behavior and the potential for intensively adaptive interventions." *Current opinion in psychology* 5 (2015): 67-71.
26) Norton, Margo, et al. "The use of palmtop computers in the treatment of bulimia nervosa." *European Eating Disorders Review* 11.3 (2003): 231-242.
27) Newman, Michelle G., et al. "Comparison of palmtop-computer-assisted brief cognitive-behavioral treatment to cognitive-behavioral treatment for panic disorder." *Journal of consulting and clinical psychology* 65.1 (1997): 178.
28) Riley, William, et al. "Computer and manual self-help behavioral strategies for smoking reduction: initial feasibility and one-year follow-up." *Nicotine & tobacco research* 4.Suppl_2 (2002): S183-S188.
29) Shiffman, Saul, Arthur A. Stone, and Michael R. Hufford. "Ecological momentary assessment." *Annu. Rev. Clin. Psychol.* 4 (2008): 1-32.
30) Kuusela, T. (2013). "Methodological aspects of heart rate variability analysis," in Heart Rate Variability (HRV) Signal Analysis: Clinical Applications, eds M. V. Kamath, M. A. Watanabe, and A. R. M. Upton (Boca Raton, FL: CRC Press), 10–42.
31) Berntson, G. G., Quigley, K. S., and Lozano, D. (2007). "Cardiovascular psychophysiology," in Handbook of Psychophysiology, eds J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson (New York, NY: Cambridge University Press), 182–210.
32) "Heart rate variability. Standards of measurement physiological interpretation and clinical use",

Circulation, vol. 93, pp. 1043-1065, 1996.
33) Penzel, Thomas, et al. "Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea." *IEEE Transactions on biomedical engineering* 50.10 (2003): 1143-1151.
34) Nussinovitch, Udi, et al. "Reliability of ultra‐short ECG indices for heart rate variability." *Annals of* Noninvasive Electrocardiology 16.2 (2011): 117-122.
35) Shaffer, Fred, Rollin McCraty, and Christopher L. Zerr. "A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability." Frontiers in psychology 5 (2014): 1040.
36) S. W. Smith, “Moving Average Filters - Implementation by Convolution”, http://www.dspguide.com/ch15/1.htm, 1997
37) W. Engelse and C. Zeelenberg, “A single scan algorithm for QRS detection and feature extraction”, IEEE Comp. in Cardiology, vol. 6, pp. 37-42, 1979
38) A. Lourenco, H. Silva, P. Leite, R. Lourenco and A. Fred, “Real Time Electrocardiogram Segmentation for Finger Based ECG Biometrics”, BIOSIGNALS 2012, pp. 49-54, 2012