OPTIMAL METAMODELING TO INTERPRET ACTIVITY-BASED HEALTH SENSOR DATA

Ali Mehmani[*]
Columbia University
New York, NY, 10027
Email: am4515@columbia.edu

Payam Ghassemi[†]
University at Buffalo
Buffalo, NY, 14260
Email: payamgha@buffalo.edu

Souma Chowdhury[‡]
University at Buffalo
Buffalo, NY, 14260
Email: soumacho@buffalo.edu

ABSTRACT

Wearable sensors are revolutionizing the health monitoring and medical diagnostics arena. Algorithms and software platforms that can convert the sensor data streams into useful/actionable knowledge are central to this emerging domain, with machine learning and signal processing tools dominating this space. While serving important ends, these tools are not designed to provide functional relationships between vital signs and measures of physical activity. This paper investigates the application of the metamodeling paradigm to health data to unearth important relationships between vital signs and physical activity. To this end, we leverage neural networks and a recently developed metamodeling framework that automatically selects and trains the metamodel that best represents the data set. A publicly available data set is used that provides the ECG data and the IMU data from three sensors (ankle/arm/chest) for ten volunteers, each performing various activities over one-minute time periods. We consider three activities, namely running, climbing stairs, and the baseline resting activity. For the following three extracted ECG features – heart rate, QRS time, and QR ratio in each heartbeat period – models with median error of <25% are obtained. Fourier amplitude sensitivity testing, facilitated by the metamodels, provides further important insights into the impact of the different physical activity parameters on the ECG features, and the variation across the ten volunteers.

Keywords: Metamodel, Neural Networks, PEMF, health IoT, ECG.

INTRODUCTION
Smart and Connected Health

Internet of things (IoT) is revolutionizing healthcare within and beyond the walls of hospitals and care facilities, e.g., with the promise of unique smart-and-connected health ecosystems in areas such as telemedicine, emergency medicine, sports medicine, smart rehabilitation, and on-field diagnostics. Wearable sensors form an important part of this health IoT ecosystem. In order to produce useful information and actionable knowledge from the data collected by these sensors, machine learning and signal processing tools are typically used. A majority of these tools do not however provide any direct mathematical relationship between different health and activity parameters measured by different sensors. Such relationships could be highly valuable for applications such as: 1) detecting anomalies, 2) fitness assessment, 3) exploring the impact of mobility or different activity (through sensitivity analysis), and 4) compressed storage and communication of data streams (in resource scarce IoT environment). It is important to note that although regression based correlation methods can capture some of these abstractions, the high nonlinearity and stochasticity of dynamic health parameters are often not well capture by smooth parametric regression models.

This gap is analyzing health sensor data can be filled by the metamodeling paradigm, especially by interpolating metamodels that can represent highly nonlinear and noisy relationships, and are typically used in the engineering design arena. In this paper, we adopt this metamodeling paradigm, and go a step further in leveraging a recently developed approach that automatically selects the best metamodel type and composition from multiple choices such as Kriging and RBF models (with various kernel options). An introduction to metamodeling and model evaluation strategies is provided next.

---

[*]Post-Doctoral, Data Science Institute, ASME Professional Member.

[†]Ph.D. Student, Department of Mechanical and Aerospace Engineering.

[‡]Assistant Professor, Department of Mechanical and Aerospace Engineering, ASME Professional Member. Corresponding Author.

METAMODELING OF ACTIVITY-BASED HEALTH DATA

The overarching objective of this research is to investigate the use of metamodels to mathematically map critical features of a persons heart activity onto measures of physical activity. The research framework in this study is illustrated in Figure 1. Key steps in this framework are as follows:

- Extract pertinent ECG features (i.e., heart rate, QRS time, and QR ratio in each heartbeat period) from low sampling rate (50 Hz) ECG data through a combination of specialized filtering and template matching.
- Train global metamodels to define the ECG features as functions of time-based physical activity parameters (e.g., number of steps taken during running/climbing, distance traveled, height climbed, peak arm/leg acceleration during each step, and time since start of activity). In this process the suitability of various metamodels (Kriging, RBFs) will be tested.
- Perform sensitivity analysis to determine and visualize the impact of the physical activity parameters on the heartbeat features.

Metamodeling: Universal Function Approximators

In the era of complex engineering systems, computational models play a central role in systems design and analysis [1, 2]. A popular class of stochastic computational models are *surrogate models*, which are also often known as *metamodels* [3] or *response surfaces* (depending on the research community and the interpretation). Surrogate models are purely mathematical models (i.e., not derived from the system physics) that are used to provide a tractable and inexpensive approximation of the actual system behavior. They are commonly used as an alternative to expensive computational simulations (e.g., CFD [4]) or to the lack of a physical model in the case of experiment-derived data (e.g., creation and testing of new metallic alloys [5]).

Evaluating and Selecting Metamodels

The construction and use of metamodels is almost always plagued with one or both of the following practical concerns: (i) scarcity of data, since high fidelity data is expensive, and (ii) lack of knowledge of the underlying functional form of the system behavior being modeled. As a result, metamodels are generally expected to provide a low fidelity representation of the actual system behavior [6]. Owing to this typical low-fidelity of metamodels [7] and the availability of diverse types of metamodel forms, the selection of a suitable model for a given experimental or simulation data set becomes critical for effective and reliable usage of metamodels in any application.

In the literature, error measures have been used to separately select model type and kernel functions. Popular error measures used for model type selection include [8]: (i) split sample, (ii) cross-validation, (iii) bootstrapping, (iv) Schwarz's Bayesian Information Criterion (BIC) [9], and (v) Akaike's Information Cri-
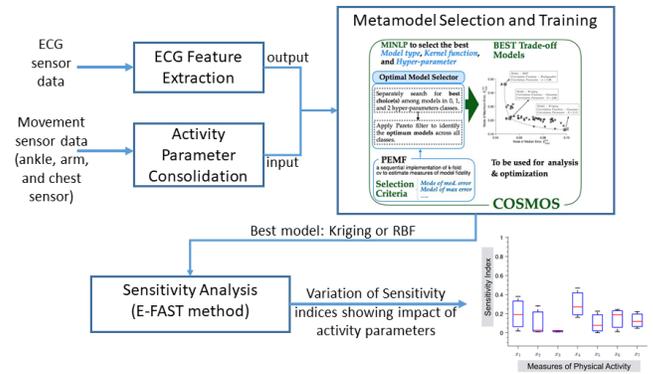

FIGURE 1. Framework and modeling flowchart

terion (AIC) [10, 9]. In addition to the model type and basis (or correlation) function selection, error measures can also be applied in hyper-parameter optimization to select the parameter that minimizes metamodel errors. This hyper-parameter optimization is highly sensitive to the basis functions and the data distribution [11]. Viana et al. [12] applied the cross-validation method to select the best predictor function and weights for different metamodels to construct a hybrid weighted metamodels. Martin and Simpson used Maximum Likelihood Estimation (MLE) and cross-validation methods to find the optimum hyper-parameter value for the Gaussian correlation function in Kriging. The likelihood function in that case defines probability of observing the training data for a particular set of parameters. Gorissen et al. provided the leave-one-out cross-validation and AIC error measures in the SUrrogate MOdeling (SUMO) Toolbox to automatically select the best model type for a given problem.

Activity-based Data: Description and Characterization

In this study, the available public data set Mhealth [13] is used. This data set contains data collected from 10 individuals using wearable sensors with a sampling rate of 50 Hz performing twelve different physical daily activities (e.g., jogging, running, sitting, etc.). Three of these activities: (i) sitting and relaxing, (ii) running, and (iii) climbing stairs are considered in this study. As illustrated in Fig. 2 and defined in Table 1, sensors are located on individual's right lower arm, left ankle, and chest to monitor the acceleration, the rate of turn and the magnetic field orientation of the body, while also collecting ECG (electrocardiogram) signals.

To prepare the training data, different types of ECG features can be extracted from the intervals and amplitudes of ECG waves in different activities. The statistical ECG features of interest extracted in this study include (i) heart rate (HR) per sec., (ii) QRS time in sec., and (iii) QR ratio in each heartbeat period. In this study, we utilize a set of measures of physical activity, extracted from wearable sensor (illustrated in Fig. 2) in each heartbeat period, to define input features. The input features are illustrated in Table. 2. To extract the ECG features, the FTT analysis is used
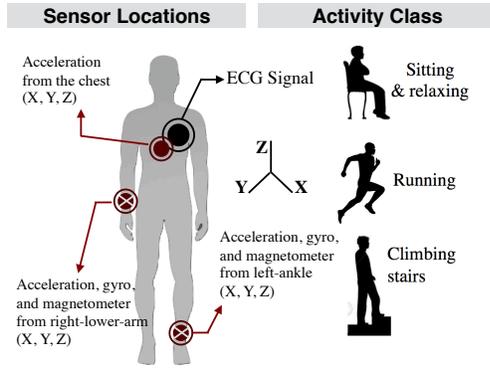
FIGURE 2. Locations to be monitored on the human body during data collection and the selected activities in this study.

TABLE 1. Sensor Description

| Sensor	Sensed Attribute
$S_{x,y,z}^{CH}$	Chest motion in m/s^2 in x, y, and z axises
$S_{x,y,z}^{LA}$	Left ankle motion in m/s^2 in x, y, and z axises
$S_{x,y,z}^{RA}$	Right lower arm motion in m/s^2 in x, y, and z axises
$S_{x,y,z}^{GLA}$	Left ankle angular rate in deg/s in x, y, and z axises
$S_{x,y,z}^{GRA}$	Right lower arm angular rate in deg/s in x, y, and z axises
$S_{x,y,z}^{MLA}$	Magnetic field of left ankle in x, y, and z axises
$S_{x,y,z}^{MRA}$	Magnetic field of right lower arm in x, y, and z axises
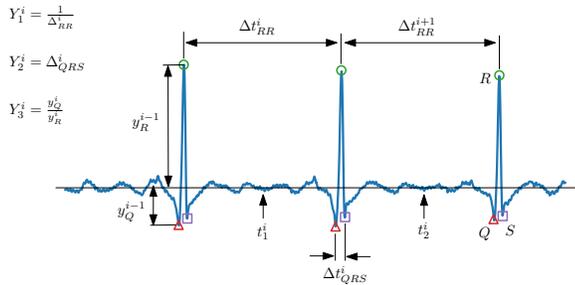
FIGURE 3. Statistical ECG features of interest extracted from the sensor ECG signal

to removing bias variation (low frequency noises) from raw data set. Then, the signal is normalized using the Min-Max scaling, i.e., $y_n = (y - y_{min})/(y_{max} - y_{min})$. As the last step of preprocessing, the normalized signal is subtracted from the average value of signal, it plays important role in the final stage to obtain Q and S locations. Next, the R-peak is estimated using an efficient R-peak detection method (Kathirvel et. al. [14]). Finally, the predicted location of R is tunned by finding the largest value at the local range of R-peak. The features Q and S are then extracted as the smallest negative value before and after the location of the R-peak.

TABLE 2. Description of measures of physical activity in each heart-beat period

Input variable	Description
X_1	Number of steps taken based on chest motion
X_2	Number of steps taken based on left ankle motion
X_3	Horizontal distance moved based on chest motion
X_4	Horizontal distance moved based on left ankle motion
X_5	Maximum acceleration measured of the legs
X_6	Maximum acceleration measured of the arms
X_7	Heartbeat period's center (in [sec])
X_8	height accomplished (only in *Climbing Stains*)

Candidate Metamodels

The effectiveness of the model selection framework (**COSMOS**) is investigated by considering a pool of models comprised of the following three popular model types: (i) Kriging, (ii) and Radial Basis Function (RBF). The different forms of the kernel/basis/correlation functions currently considered in COSMOS are given in Table 5 under Appendix A. It could be said that COSMOS is seeking to be a comprehensive metamodel selection framework both in methodology (3-level selection) and implementation (i.e., in terms of the pool of candidate model-kernels considered). Brief descriptions of the candidate metamodels and the different forms of the kernel/basis/correlation functions currently considered in COSMOS are also provided in A.

In this study, we also investigate the effectiveness of the multiple input-single output (MISO) feed-forward Neural Network model (see Appendix A) in representing the ECG features as functions of time-based physical activity parameters. This model is configured with a single hidden layer with hyperbolic tangent sigmoid transfer function, and a single output layer with the pure linear transfer function. In NN, to avoid over-fitting, the data set is randomly divided into the training set (80% of the available data set) and the testing sets.

TABLE 3. Range of hyper-parameters

Model	Hyper-parameter	Lower bound	Upper bound
RBF	shape parameter, σ	0.1	3
Kriging	correlation parameter, θ	0.1	20

Predictive Estimation of Model Fidelity (PEMF)

In the COSMOS framework, the PEMF method ([15]) is applied to estimate the model selection criteria that represents the (predicted) modal values of the median error for any given candidate metamodel. PEMF has already been successfully employed

to quantify the fidelity of metamodels that participate in a multi-fidelity optimization approach involving adaptive model switching (Mehmani and Chowdhury et al. [16]). In concept, PEMF can be perceived as a novel sequential implementation of *k*-fold cross-validation, with carefully constructed error measures that are in general significantly less sensitive to outliers and the Design of Experiments compared to mean or Root Mean Square error measures. The PEMF method predicts the error by capturing the variation of the metamodel error with an increasing density of training points.

In the PEMF method, for a set of N sample points, intermediate metamodels are constructed at each iteration, t, using M^t heuristic subsets of n^t training points (called intermediate training points). These intermediate metamodels are then tested over the corresponding remaining $N - n^t$ points (called intermediate test points). The median error is then estimated for each of the M^t intermediate metamodels at that iteration, and a parametric probability distribution is fitted to yield the modal value, $E_{med}^{mo,t}$. The use of the modal value of the median error promotes a monotonic variation of error with sample point density, unlike mean or root mean squared error which are highly susceptible to outliers. This approach gives PEMF an important advantage over conventional cross-validation-based error measures, as illustrated by Mehmani et al. ([15]). A similar approach is used to estimate the modal value of the maximum error ($E_{max}^{mo,t}$) at any t^{th} iteration. In the original PEMF method, the probability distribution to be fitted over the median and the maximum errors at each iteration were selected using the chi-square goodness-of-fit criterion ([17]). However, in order to reduce the computational expense of PEMF (that is called to evaluate multiple candidate models within COSMOS), only the lognormal distribution is used. This distribution has been previously observed (from numerical experiments) to be generally effective within PEMF.

In PEMF, once we have the history of the median and the maximum errors at different sample size ($< N$), the variation of the modal values of these error measures with sample density are then represented using the multiplicative ($E = a_0 n^{a_1}$) or the exponential ($E = a_0 e^{a_1 n}$) regression functions. The choice of these regression functions leverage the *monotonically decreasing trend of the modal error values with respect to the training point density*. The root mean squared error metric is used to select the best-fit regression function. These regression functions are then used to predict the modal values and the variance of the median and the maximum errors in the final metamodel, where the final metamodel is trained using all the N sample points.

APPLICATION TO ACTIVITY MONITORING: RESULTS AND DISCUSSION
Metamodeling: Evaluation

The PEMF errors of the best fit metamodels discovered by COSMOS, and their variation across the volunteers, are shown as a box-plot in Fig. 4. It is interesting to note that although the average error for QR ratio and QRS time are smaller (with
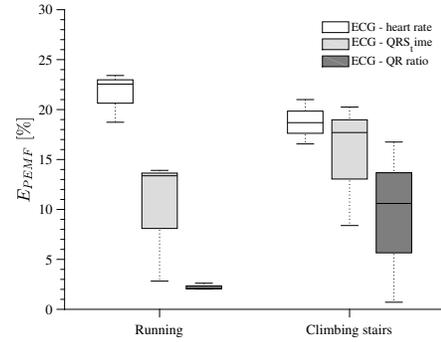

FIGURE 4. Relative prediction error, estimated using PEMF in different metamodels, representing the output ECG features in different activities, selected using COSMOS for 10 individuals .

less variation) in the case of running activity models, compared to the climbing stairs activity models, the ECG heart rate feature exhibits the higher level of error in the former.

Figure 5 illustrates the performance of three NN models, trained to represent the ECG features as functions of time-based physical activity parameters using the sensed attributes in all individuals. The performance reported based on Median Relative Absolute Error (MdRAE) and it is observed that, the error of NN models in predicting QRS heart rate and QRS time in activity 2 are respectively 3% and 1%, while that is 14% in predicting QR ratio. The performance of the NN models across all ten individuals are presented as a box-plot in Fig. 6.

Activity Impact Analysis with Metamodels

In this section, we investigate the sensitivity of extracted features of the ECG to measures of physical activity, X_i, $i = 1, 2, ..., 15$, where the features are predicted using the best-fit metamodels selected by COSMOS. To perform the sensitivity analysis (SA), the Fourier amplitude sensitivity test (FAST) method is used to determine the first-order index that represents the variance of the model output due to each of the input parameters. In Fast, the input variables are transformed into a frequency domain using Fourier transformation. Thus, a multidimensional model is reduced into a model with a single dimension [18]. Assuming a model with n input variables, $X = [X_1, X_2, ..., X_n]$, the output of the model, Y, is expressed as $Y = f(x_1, x_2, ..., x_n)$. In FAST, a search function is defined to allow the input parameter to oscillate periodically in the input space, by assigning a characteristic frequency ω_i, expressed as

$$x_i = G_i(sin \omega_i s), \qquad i = 1, 2, ..., n \qquad (1)$$

Here G_i is a transform function, and $s \in (-\infty, +\infty)$ is a scalar. By applying the properties of Fourier series, $E(Y)$ can be expressed
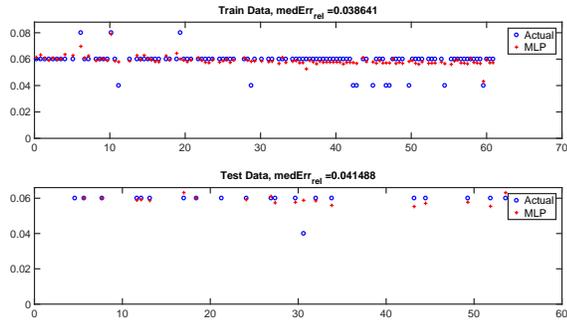
(a) ANN1: output
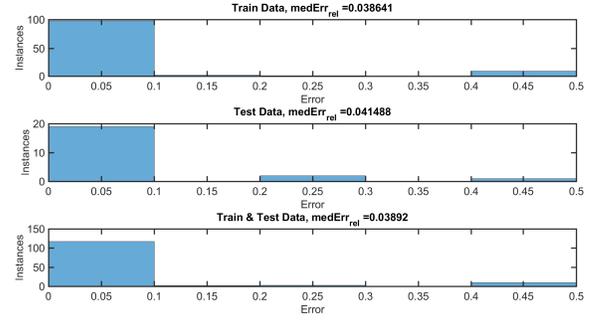
(b) ANN1: relative error
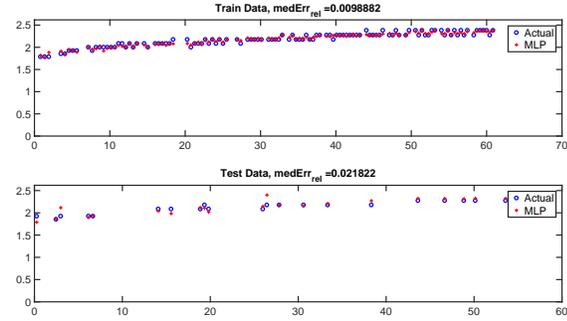
(c) ANN2: output
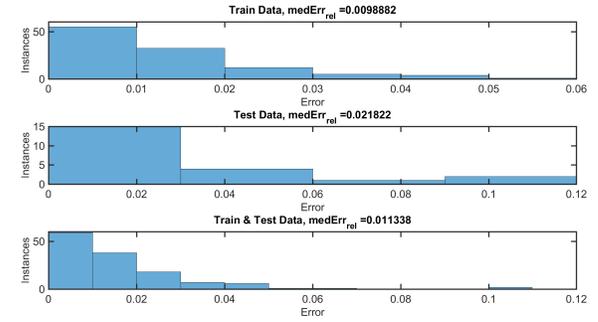
(d) ANN2: relative error
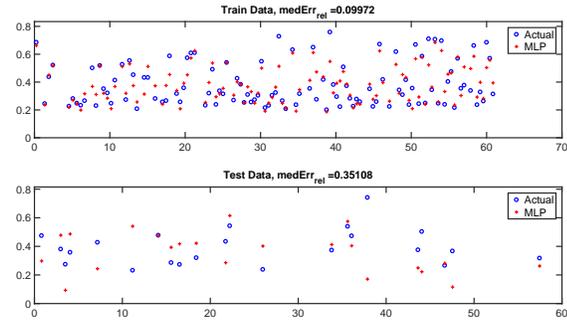
(e) ANN3: output
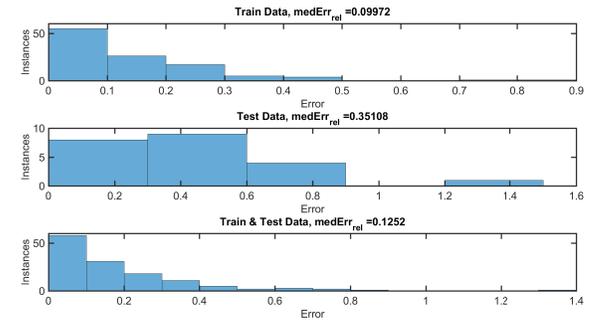
(f) ANN3: relative error

FIGURE 5. The performance of the NN in representing ECG features as functions of time-based physical activity parameters; ANN1, ANN2, and ANN3 are corresponded with the models that are estimating the ECG heart rate per sec., ECG QRS time, and ECG QR ratio, respectively.

as

$$Y = f(s) = A_0 + \sum_{k=1}^{+\infty} \left[A_k \cos(ks) + B_k \sin(ks) \right] \quad (2)$$

where $f(s) = f(x_1(s), x_2(s), \ldots, x_n(s))$, and $i = 1, 2, \ldots, n$; A_0, A_k, and B_k are the Fourier coefficients, defined as

$$A_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(s)\,ds, \text{ and}$$
$$A_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(s) \cos(ks)\,ds, \ B_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(s) \sin(ks)\,ds \quad (3)$$

For practical problems, k must be limited to a reasonable value of the integer N, which indicates the sample size of the input data. The variance of the model output, s_Y^2, can therefore

5
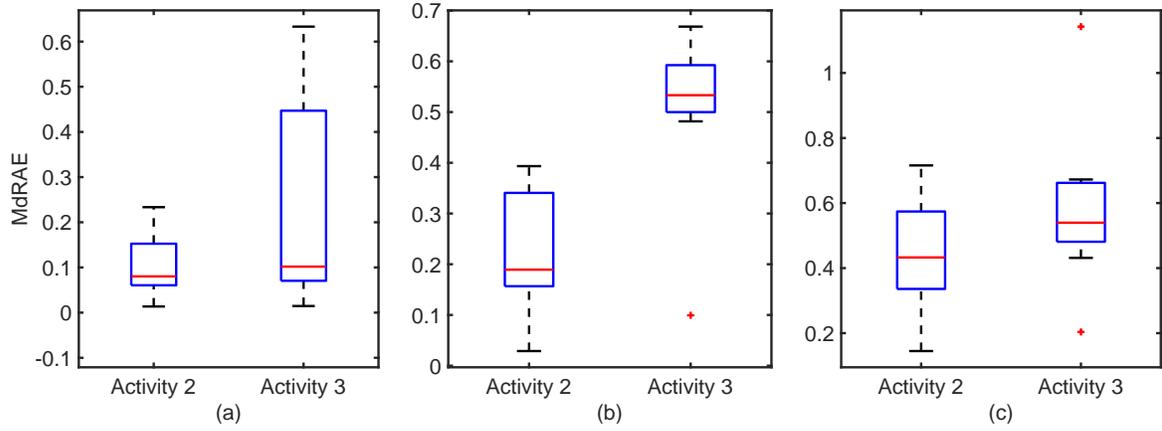
FIGURE 6. Accuracy of the ANN models based on Median Relative Absolute Error (MdRAE) for two activities, running (Activity 2) and climbing stairs (Activity 3), in 10 different individuals: (a) the ECG heart rate per sec, (b) the ECG QRS Time, (c) the ECG QR ratio

be approximated as

$$s_Y^2 = E(Y^2) - [E(Y)]^2 \approx \frac{1}{2\pi} \sum_{k=1}^{(N-1)/2} \left(A_k^2 + B_k^2 \right)$$

where (4)

$$A_k = \frac{1}{\pi} \sum_{j=1}^{N} f(s_j) \cos(s_j k), \; B_k = \frac{1}{\pi} \sum_{j=1}^{N} f(s_j) \sin(s_j k)$$

In the variance-based sensitivity analysis, the first-order sensitivity index of an input parameter, x_i, is defined as the conditional variance of the model output, $s_{E(Y/x_i)}^2$, with respect to the unconditional variance of the model output (s_Y^2). The conditional variance in FAST is approximated by summing up the spectrum values for the basic frequency ω_i and its higher harmonics, as shown below.

$$s_{E(Y/x_i)}^2 \approx \frac{1}{2} \sum_{p=1}^{m} \left(A_{p\omega_i}^2 + B_{p\omega_i}^2 \right) \quad (5)$$

In Eq.(5), $p \in Z$ and $p\omega_i \le (N-1)/2$; and m indicates the order of higher harmonics that are considered [19].

Therefore, the first-order index can be formulated by combining Eq.(5) and Eq.(4), which is expressed as

$$S_i = \frac{s_{E(Y/x_i)}^2}{s_Y^2} \quad (6)$$

Figure 7 illustrates the sensitivity analysis exploring the affect of each measures of physical activity to the selected ECG features, for climbing stairs and running activities.

From the boxplot of sensitivity indices, it can be sen that time has a significant impact on the ECG parameters and also registers higher variation across individuals.

To show the sensitivity of ECG heart rate and QRS time to sensed attributes (e.g., $x_{1,2,...,7}$), when the climbing stairs and running activities are considered, the ECG heart rate and QRS time interval metamodel, selected by COSMOS, is used to generate filled contour plots by varying the input variables (e.g., measures of physical activity) pairwise and keeping the remaining variables at the baseline value. The results for individuals No. 1 and No. 8 are illustrated in Fig. 8 and 9, respectively.

CONCLUSION

This paper provides a preliminary investigation into the use of global approximation models, or metamodels, to interpret health sensor data, more specifically, to represent vital signs parameters (i.e., heartbeat (ECG)) as functions of physical activity parameters. We use a publicly available data set, containing the ECG data and gyro/accelerometer/magnetometer data from three sensors (ankle/arm/chest) for ten volunteers, each performing various activities over one-minute time periods. ECG features, including heart rate, QRS time, and QR ratio in each heartbeat period, are extracted as output features and represented as functions of the measures of physical activity such as number of steps taken during running or climbing and distance traveled. We train both Neural Network models and best-fit metamodels that are chosen for use by the COSMOS framework from among various types of candidate Kriging and RBF models. These metamodels facilitate sensitivity analysis, which is performed using the Fourier amplitude sensitivity test method. The results provide insightful visualization of the differing impacts of the activity/movement parameters on the ECG features as well as their variation across the ten individuals. Future work will focus on the use of metamodels that relate heart behavior to physical activity measures to provide unique insights for (i) rehabilitation planning and progress estimation, (ii) detection of cardiac anomalies, and (iii) the creation of activity-referenced individual baselines to serve as part of their health record.
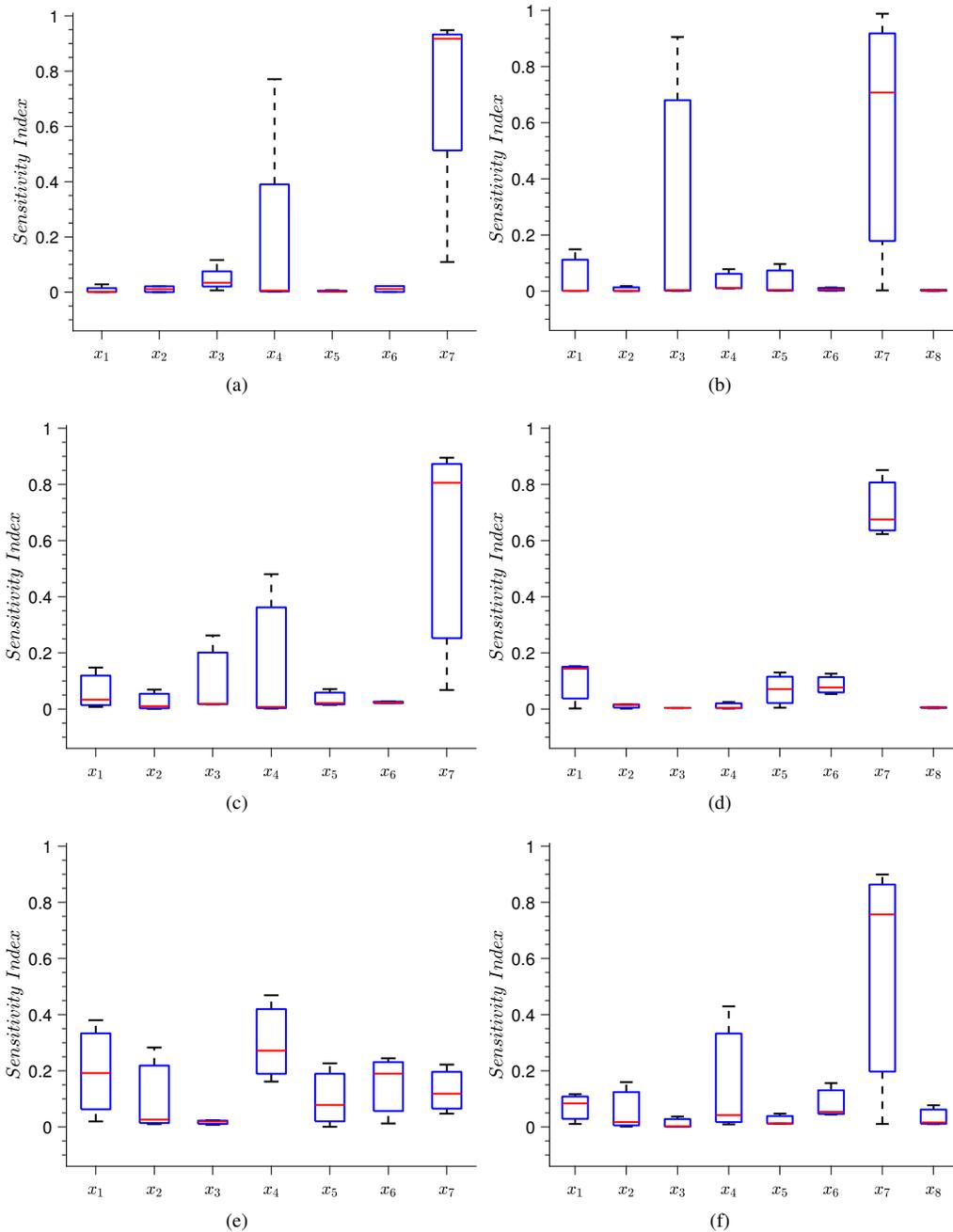
FIGURE 7. SA of the extracted ECG features to all measures of physical activities in 10 different individuals : (a) the ECG heart rate per sec in running activity, (b) the ECG heart rate per sec in climbing stairs activity, (c) the ECG QRS Time in running activity, (d) the ECG QRS Time in climbing stairs activity, (e) the ECG QR ratio in running activity, and (f) the ECG QR ratio in climbing stairs activity

REFERENCES

[1] Bloebaum, C. L., and McGowan, A. R., 2012. "The design of large-scale complex engineered systems: Present challenges and future promise". In 14th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference.

[2] Mesmer, B. L., and Bloebaum, C. L., 2015. "An end-user decision model with information representation for improved performance and robustness in complex system design". *Research in Engineering Design,* **26**(3), pp. 235–251.

[3] Kleijnen, J., 1975. *Statistical techniques in simulation*, Vol. II. New York. New York: Marcel Dekker.

[4] Jeong, S., Murayama, M., and Yamamoto, K., 2005. "Efficient optimization design method using kriging model".
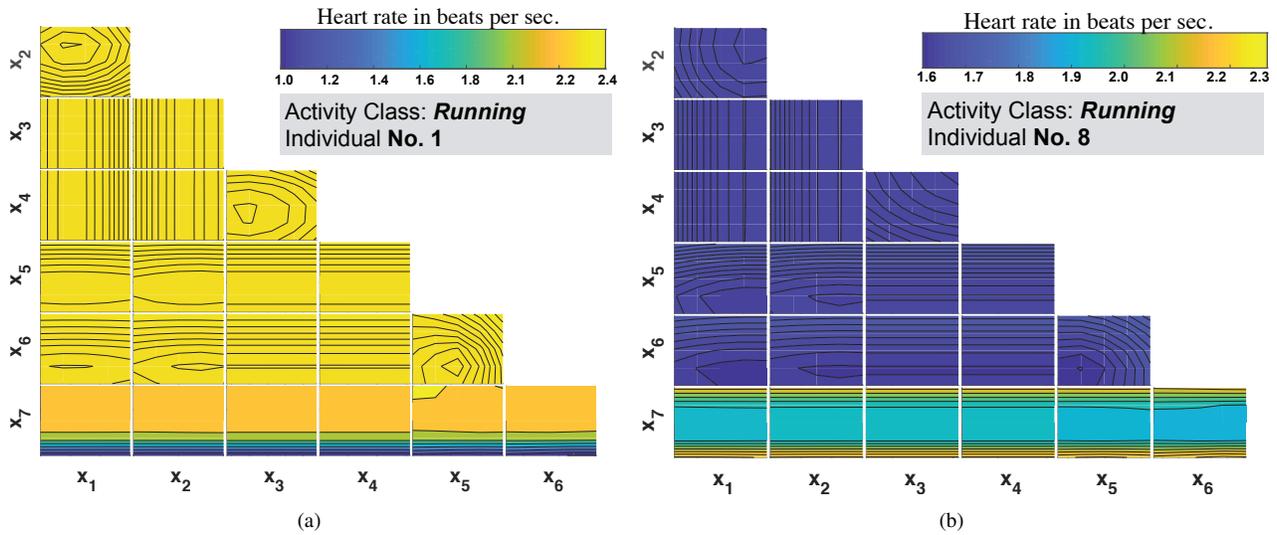
FIGURE 8. The ECG heart rate per sec in (a) individual No. 1 and (b) individual No.8; Each tile shows a contour of the ECG heart rate, versus input variables (measures of physical activities), with the remaining variables held at the baseline value
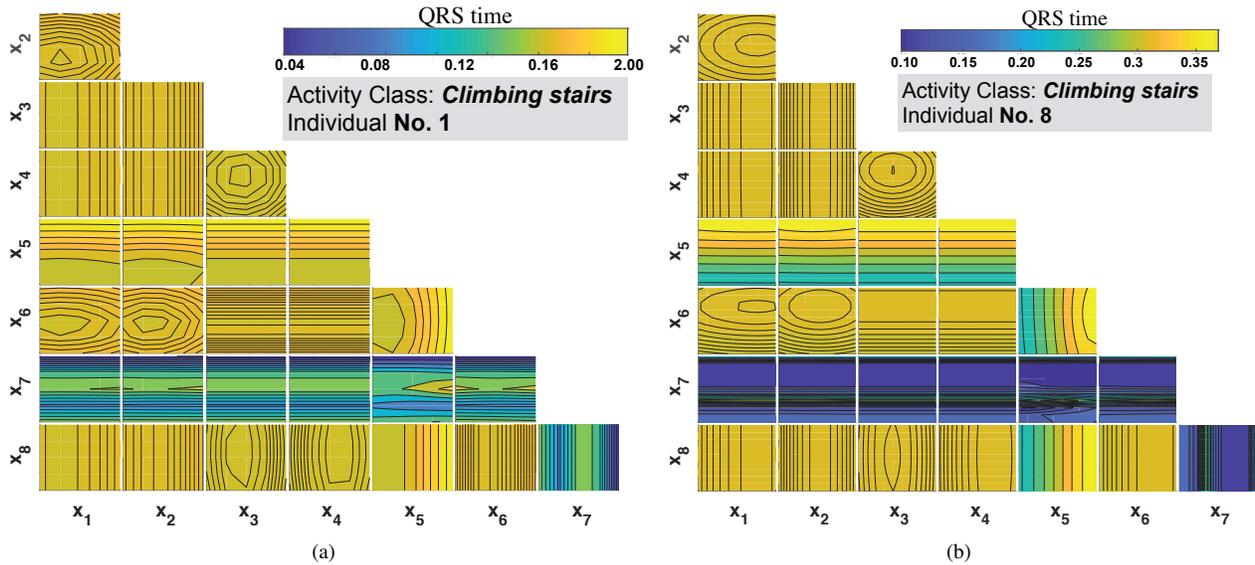

FIGURE 9. The QRS time in (a) individual No. 1 and (b) individual No.8; Each tile shows a contour of the ECG QRS time, versus input variables (e.g., measures of physical activities), with the remaining variables held at the baseline value

Journal of Aircraft, **42**(2), pp. 413–420.

[5] Oktem, H., Erzurumlu, T., and Kurtaran, H., 2005. "Application of response surface methodology in the optimization of cutting conditions for surface roughness". *Journal of Materials Processing Technology,* **170**(1), pp. 11–16.

[6] Jia, G., and Taflanidis, A. A., 2013. "Kriging metamodeling for approximation of high-dimensional wave and surge responses in real-time storm/hurricane risk assessment". *Computer Methods in Applied Mechanics and Engineering,* **261**, pp. 24–38.

[7] Forrester, A., and Keane, A., 2009. "Recent advances in surrogate-based optimization". *Progress in Aerospace Sci-*

ences, **45**(1-3), pp. 50–79.

[8] Queipo, N., Haftka, R., Shyy, W., Goel, T., Vaidyanathan, R., and Tucker, P., 2005. "Surrogate-based analysis and optimization". *Progress in Aerospace Sciences,* **41**(1), pp. 1–28.

[9] Claeskens, G., and Hjort, N. L., 2008. *Model selection and model averaging.* Cambridge Books.

[10] Bozdogan, H., 2000. "Akaike's information criterion and recent developments in information complexity". *Journal of Mathematical Psychology,* **44**, pp. 62–91.

[11] Gorissen, D., Tommasi, L. D., Hendrickx, W., Croon, J., and Dhaene, T., 2008. "Rf circuit block modeling via krig-

ing surrogates". In In Proceedings of the 17th International Conference on Microwaves, Radar and Wireless Communications (MIKON).

[12] Viana, F. A. C., Haftka, R. T., and Steffen, V., 2009. "Multiple surrogates: how cross-validation errors can help us to obtain the best predictor". *Structural and Multidisciplinary Optimization, 39*, pp. 439–457.

[13] Banos, O., Garcia, R., Holgado-Terriza, J. A., Damas, M., Pomares, H., Rojas, I., Saez, A., and Villalonga, C., 2014. "mhealthdroid: a novel framework for agile development of mobile health applications". International Workshop on Ambient Assisted Living, Springer International Publishing, pp. 91–98.

[14] Kathirvel, P., Sabarimalai Manikandan, M., Prasanna, S. R. M., and Soman, K. P., 2011. "An efficient r-peak detection based on new nonlinear transformation and first-order gaussian differentiator". *Cardiovascular Engineering and Technology, 2*(4), pp. 408–425.

[15] Mehmani, A., Chowdhury, S., and Messac, A., 2015. "Predictive quantification of surrogate model fidelity based on modal variations with sample density". *Structural and Multidisciplinary Optimization, 52*(2), pp. 353–373.

[16] Mehmani, A., Chowdhury, S., Tong, W., and Messac, A., 2015. "Adaptive switching of variable-fidelity models in population-based optimization". In *Engineering and Applied Sciences Optimization*, Vol. 38 of *Computational Methods in Applied Sciences*. Springer International Publishing, pp. 175–205.

[17] Haldar, A., and Mahadevan, S., 2000. *Probability, reliability, and statistical methods in engineering design*. Wiley.

[18] Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G., and Schaibly, J. H., 1973. "Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory". *Journal of Chemical Physics, 59*(8), pp. 3873–3878.

[19] Cukier, R. I., 1978. "Nonlinear sensitivity analysis of multiparameter model systems". *Journal of Computational Physics, 26*(1), pp. 1–42.

[20] Hardy, R. L., 1971. "Multiquadric equations of topography and other irregular surfaces". *Journal of Geophysical Research, 76*, pp. 1905–1915.

[21] Mongillo, M., 2011. "Choosing basis functions and shape parameters for radial basis function methods". In SIAM Undergraduate Research Online.

[22] Giunta, A. A., and Watson, L., 1998. "A comparison of approximation modeling techniques: Polynomial versus interpolating models". *AIAA Journal*(AIAA-98-4758).

[23] Cressie, N., 1993. *Statistics for Spatial Data*. Wiley, New York.

[24] Martin, J. D., and Simpson, T. W., 2005. "Use of kriging models to approximate deterministic computer models". *AIAA journal, 43*(4), pp. 853–863.

[25] Lophaven, S. N., Nielsen, H. B., and Sondergaard, J., 2002. Dace - a matlab kriging toolbox, version 2.0. Tech. Rep. IMM-REP-2002-12, Informatics and Mathematical Modelling Report, Technical University of Denmark.

[26] Hornik, K., Stinchcombe, M., and White, H., 1989. "Multilayer feedforward networks are universal approximators". *Neural Networks, 2*(5), pp. 359 – 366.

A Metamodel Candidates
Radial Basis Function (RBF)

The idea of using Radial Basis Functions (RBF) as approximation functions was conceived by [20]. The RBF approximation is a linear combination of the basis functions (Ψ) computed with respect to each sample point, as given by

$$\bar{F}(x) = W^T \Psi = \sum_{i=1}^{n_p} w_i \psi\left(\|x - x^i\|\right) \qquad (7)$$

In Eq. 7, n_p denotes the number of selected sample points; w_i's are the weights estimated using the pseudo inverse method, based on the training data; and ψ is the basis function that is expressed in terms of the Euclidean distance, $r = \|x - x^i\|$, of a point x from a given sample point, x^i. The most effective forms of the basis functions are listed in Table 5 where σ represents the shape parameter of the basis function. The shape parameter in a basis function has a strong impact on the accuracy of the trained RBF. A smaller shape parameter often corresponds to a wider basis function, and the shape parameter, $\sigma = 0$, corresponds to a constant basis function ([21]). The different RBFs considered in this paper are listed in Table 5.

Kriging

Kriging ([22]) is an approach to approximate irregular data. The kriging approximation function consists of two components: (*i*) a global trend function, and (*ii*) a deviation function representing the departure from the trend function. The trend function is generally a polynomial (e.g., constant, linear, or quadratic). The general form of the kriging metamodel is given by ([23]):

$$\bar{F}(x) = \hat{f}(x, \varphi) + Z(x) \qquad (8)$$

where $\bar{F}(x)$ is the unknown function of interest, $Z(x)$ is the realization of a stochastic process with the mean equal to zero, and a nonzero covariance, and $\hat{f}$ is the known approximation function

$$\hat{f}(x, \varphi) = f(x)^T \varphi \qquad (9)$$

where φ is the regression parameter matrix.

The $i, j-th$ element of the covariance matrix, $Z(x)$, is given by

$$COV[Z(x^i), Z(x^j)] = \sigma_z^2 R_{ij} \qquad (10)$$

where R_{ij} is the correlation function between the ith and the jth data points; and σ_z^2 is the process variance, which scales the spatial correlation function. The popular types of correlation functions are listed in Table 5. The correlation function controls the smoothness of the Kriging model estimation, based on the influence of other nearby points on the point of intrest. In Kriging, the regression function coefficients, the process variance, and the correlation function parameters, $\{\varphi, \sigma_z^2, \theta\}$, each can be predefined or estimated using parameter estimation methods such as Maximum Likelihood Estimation (MLE). In this paper, the regression function coefficients and the process variance are estimated using MLE, as given by

$$\varphi = (F^t R^{-1} F)^{-1} F^T R^{-1} Y \qquad (11)$$
$$\sigma_z^2 = \frac{1}{n}(Y - F\widetilde{\varphi})^T R^{-1} (Y - F\widetilde{\varphi})$$

where $Y = [y_1 \; y_2 \; ...]$ represents the vector of the actual output at the training points; R is a correlation matrix; and F is a matrix of $f(x)$ evaluated by Kriging at each training point ([24]). The hyper-parameter, θ, in the correlation function is determined by solving the nonlinear hyper-parameter optimization problem. To estimate the regression function coefficients and the process variance in Kriging, the DACE (design and analysis of computer experiments) package, developed by [25], is used in this paper.

ANN

Artificial Neural Network (ANN) is a set of powerful tools for modeling complex nonlinear systems. Multi-Layer Perception (MLP) [26] is a feed-forward ANN that includes multiple layers with multiple nodes per layer and is also known as a universal metamodel. Hidden Layer Transfer Function (HLTF), Output Layer Transfer Function (OLTF), and Hidden Layer Size (HLS) are three parameters that should be determined in this model.

In this paper, the NN models are trained using scaled conjugate gradient back propagation method three MISO NN models for each output and each activity. The hyper-parameters and obtained results of each model are reported in Table 4.

TABLE 4. Parameters of the trained ANN models

	Output	Activity	HLS	MdRAE
ANN11	Y_1	2	8	0.01
ANN12	Y_1	3	8	0.01
ANN21	Y_2	2	60	0.03
ANN22	Y_2	3	60	0.10
ANN31	Y_3	2	60	0.14
ANN32	Y_3	3	60	0.20

Appendix B: Expressions for Candidate Metamodel-Kernels in COSMOS

TABLE 5.   Basis or Kernel functions and their hyper-parameters in the candidate metamodels

Type of model	Type of basis/correlation/kernel function		Hyper parameter
RBF	Linear:	r	
	Cubic:	r^3	
	Thin plate spline:	$r^2 ln(r)$	shape parameter, σ
	Gaussian:	$e^{(-r^2/2\sigma)}$	
	Multiquadric:	$(r^2 + \sigma^2)^{1/2}$	
Kriging	Linear:	$max(1 - \theta r, 1)$	
	Exponential:	$e^{(-\theta r)}$	
	Gaussian:	$e^{(-\theta r)^2}$	correlation function parameter, θ
	Cubic:	$1 - 0.5\xi + 0.5\xi^2; \ \xi = max(1 - \theta r, 1)$	
	Spherical:	$1 - 3\xi^2 + 2\xi^3; \ \xi = max(1 - \theta r, 1)$	