Predicting the Heart Rate Response to Outdoor Running Exercise

Xiaoli Liu
University of Oulu
Oulu, Finland
xiaoli.liu@oulu.fi

Xiang Su
University of Helsinki
Helsinki, Finland
xiang.su@helsinki.fi

Satu Tamminen
University of Oulu
Oulu, Finland
satu.tamminen@oulu.fi

Topi Korhonen
Polar Electro Oy
Kempele, Finland
topi.korhonen@polar.com

Juha Röning
University of Oulu
Oulu, Finland
juha.roning@oulu.fi

Abstract—Heart rate is a good measure for physical exercise as it accurately reflects exercise intensity and is easy to measure. If the heart rate response to a complete exercise session is predicted beforehand, information related to the exercise can be inferred, such as exercise intensity and calorie consumption. While most current heart rate prediction models are developed and tested for the scenarios of indoor running exercise or low running speed exercise, we adopt a nonlinear Ordinary Differential Equation (ODE) model for complete outdoor running exercise sessions to predict the heart rate response and identify the parameters of the model with machine learning algorithms. The proposed model enables us to predict a complete outdoor running exercise session instead of predicting the heart rate for a short duration. Model validation is carried out both on the training and testing sets. Our results show that the proposed model captures very stable prediction performance.

Index Terms—Predictive models, machine learning, parameter estimation, wearable devices

I. INTRODUCTION

Heart rate has been proven to be a good measure for physical exercises, because it is simple, accurately reflects exercise intensity, and well understandable. Heart rate based training is a well-known technique to improve the effectiveness of training and prevent over exercising. If the heart rate responding to a complete session is predicted beforehand, many values can be calculated based on the predicted heart rate, which means information related to the exercise can be inferred, such as exercise intensity and calorie consumption.

Existing approaches to model and predict the heart rate response to running exercises can be divided into two categories: (1) analytic models, which contain closed-form equations that represent a response of the heart rate to strain; and (2) machine learning approaches, which do not encode any prior information but will learn and generalize the response model in the learning process. While approaches in the first category benefit from their analytic closed-form notations, the approaches in the second category are attractive because they take environmental parameters into account such as altitude, slope, or any other relevant information that are not represented in the analytic equations. Most research performing the heart rate prediction in the exercise focuses on indoor exercise and little research has been conducted on outdoor running exercise. Indoor treadmill running usually suffers speed limitation and the model can not be straightforwardly used for outdoor running exercise.

In this paper, we present an adopted nonlinear Ordinary Differential Equation (ODE) model to describe the heart rate response in the outdoor running exercise. The parameters of the model are identified by using the Levenberg-Marquardt algorithm on all the training exercises simultaneously. Model evaluation is performed both on the training and testing sets. Our results show that the performance of our model is quite stable. The difference between model performance on the training and testing sets is small. The Root Mean Squared Error (RMSE) of the proposed model is greatly reduced after 3 minutes.

The rest of this article is organized as follows: Section II presents the background and related work. Section III describes model identification. We perform model evaluation in Section IV and conclude the paper by suggesting future research in Section V.

II. BACKGROUND AND RELATED WORK

Currently, widely accepted wearable devices monitor the subjects during physical exercises and support data analytic technologies to enable the discovery of valuable insights into the behavioral patterns. Wearable devices are utilized to predict the heart rate for different exercise activities. For example, Sumida et al. predict the heart rate dynamics via smart phone sensor data with neural network algorithms [1]. Tapia et al. present a real-time algorithm for the automatic recognition of not only physical activities but also their intensities, using five triaxial wireless accelerometers and a wireless heart rate monitor [2].

In general, heart rate prediction based on physical activity enables advanced smart devices and analytic tools properly controlling and monitoring the strain that a training device imposes on a person during physical exercises. Nowadays, many different kinds of wearable devices monitor the heart rate of people during physical exercise and most devices enable people to be informed whether the heart rate is above or below a given threshold and recommend certain exercises. However, the recommendations are rather straightforward and include only limited insights. Therefore, accurate prediction of heart rate is crucial in models used for training control. Some physical exercises, such as high-speed running, should be performed in a controlled way to maximize efficiency and

to avoid any negative effects on the health caused by an over-strain of the body.

A few studies have been carried out for modeling heart rate response during exercise. As one of the earliest research, Cheng et al. present a nonlinear model for various workload intensities [3] [4]. They introduce a nonlinear model to predict the heart rate behavior of a subject based on the running velocity on a treadmill and ergometer. Paradiso et al. present a slightly modified model to regulate the heart rate using a bicycle ergometer and generalize this model for different sports activities [5]. Baig et al. present a Linear Time Varying (LTV) model to describe the response to cycling, walking and rowing exercises [6]. They estimate the parameters in the LTV model using Kalman Filter. Wang et al. [7] utilize Support Vector Regression (SVR) to study the nonlinear behavior of cardiovascular variables. They generate a nonparametric model that quantitatively describes the observations made. Feedforward Neural Network (FNN) is utilized to estimate the activity based on a given time series of heart rate data [8].

However, current models are mostly designed and tested for lower speed indoor running. Less research is carried out for predicting the heart rate response during outdoor running exercise. With using the real outdoor running exercise data to build the model, our research work is very meaningful and serves as the foundation for developing novel tools for outdoor running exercise.

III. MODEL IDENTIFICATION

We adopt the ODE in [4] to model the heart rate changing during the running exercise, which takes the following form:

$$\dot{x}_1(t) = -a_1 x_1(t) + a_2 x_2(t) + a_2 u^2(t),$$
$$\dot{x}_2(t) = -a_3 x_2(t) + \phi(x_1(t)), \qquad (1)$$
$$\phi(x_1(t)) = \frac{a_4 x_1(t)}{1 + exp(-(x_1(t) - a_5))},$$
$$y(t) = x_1(t),$$

where $u(t)$ is the running speed at time t, a_1, a_2,..., a_5 are positive scalars, and $x_i(0) = 0$ for $i = 1, 2$ are the initial values. The output $y(t)$ describes the heart rate changes at time t related to the resting heart rate, and output $x_2(t)$ describes the slower effects from, such as peripheral metabolism and body temperature increase. The model assumes the increase of the heart rate is proportional to the square of the running speed.

A. Data collection and pre-processing

In this research, the data of a subject with exercise type outdoor running using Polar wearable devices are collected. Running exercise data, including speed, heart rate, distance, ascent, descent, cadence etc. are collected every second. Heart rate data related to resting period (before exercise) and recovery period (after exercise) are not recorded. Data of running speed and corresponding heart rate in the outdoor running exercise are used for performing the analysis. Fifteen representative exercises are selected, where ten exercises are

used as training sets for fitting the model (or estimating the parameters) and five exercises are used for testing sets. Table I describes the statistics of the running speed in ten training sets. Figure 1 presents the speed and corresponding heart rate of three outdoor running samples.

TABLE I
RUNNING SPEED (KM/H) IN THE TRAINING DATA SETS

Exercise	Maximum Speed	Mean Speed	Standard deviation
1	10.5	8.83	0.84
2	10.1	8.91	0.86
3	10.5	9.19	0.80
4	12.4	9.10	1.60
5	11.5	10.09	0.77
6	10.1	9.07	0.68
7	12.5	9.36	1.39
8	11.7	10	0.93
9	10.6	9.29	0.71
10	12.1	9.24	1.41

Fig. 1. Running speed and corresponding heart rate of three samples.

In certain situations, the speed of the subject might dramatically reduce. For example, in the exercise shown in figure 2, the subject seems to stop for some seconds and the speed drops to 0km/hour immediately. This only takes some seconds and the subject continues the running exercise. In such situations, data are imputed with the nearest values to improve the data quality.

B. Parameter estimation

The identified model in [4] can not be used in outdoor running exercise. Figure 2 illustrates the real heart rate response (red curve) in exercise eight and the prediction (green curve) using identified model in [4]. Figure 2 shows that the heart rate response increases too much with the identified model in [4], which suits medium speed exercise. While in the high running exercise, the hear rate response shows a slowly increasing behaviour.
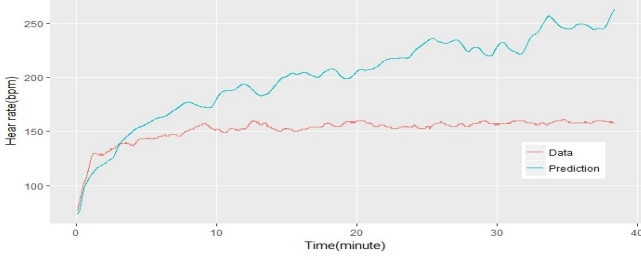
Fig. 2. Real heart rate and predicted heart rate response using model in [4] for exercise eight.

In this section, we identify the model for outdoor running exercise. Let $s(t)$ denote the running speed at time t, normalize the input $s(t)$ and rescale the output $x_1(t)$ by defining

$$u(t) = s(t)/13,$$
$$x_1(t) = x_1(t)/4, \qquad (2)$$

where $s(t)$ is the running speed, $x_1(t) = hr(t) - 70$, $hr(t)$ is the heart rate at time t and 70 bpm denotes resting heart rate. $x_1(t)$ is the scaled change of heart rate from the resting heart rate and $u(t)$ is the normalized running speed.

Define the cost function $F(a)$ as follows:

$$F(a) = \frac{1}{2n} \sum_t (y(t) - \bar{y}(t))^2, \ t \in T$$

where $a = (a_1, a_2, a_3, a_4, a_5)$ are parameters that need to be estimated, $y(t)$ is the solution of equation (1) or prediction at time t with input $u(t)$, $\bar{y}(t)$ is the scaled observation at time t from the resting heart rate, n is the number of observations in the time series, and T is the time range of observations.

The optimal parameters $\hat{a} = (\hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4, \hat{a}_5)$ are defined as the values that minimize the cost function $F(a)$. Levenberg-Marquardt is adopted to find the parameters that minimize $F(a)$. Levenberg-Marquardt algorithm is a standard technique for addressing nonlinear least squares problems. It could be regarded as a combination of the gradient descent algorithm and Gauss-Newton algorithm, as it adaptively varies the parameter updates between them. Please refer to [9] [10] [11] [12] for more details about the algorithm and its variations.

Parameters are identified using ten exercise simultaneously, denote

$$x^i(t) = [x_1^i(t), x_2^i(t)],$$
$$y_i(t) = x_1^i(t),$$

where $i = 1, 2, ..., 10$ and represents exercise i.

Consider the following multiple-input and multiple-output system

$$\dot{x}(t) = f(x(t), a, u(t)),$$
$$y(t) = Ax(t), \qquad (3)$$

where f is a nonlinear function, $x(t) = [x^1(t), ..., x^{10}(t)]' \in R^{20}$, $u(t) = [u_1, ..., u_{10}]' \in R^{10}$, $a = [a_1, a_2, a_3, a_4, a_5]' \in$

R^5, $y(t) = [y_1, ..., y_{10}]' \in R^{10}$ and $x(0) = 0$. The transformation matrix $A \in R^{10 \times 20}$ and the value of element $A(i, j)$ is defined as

$$A(i, j) = \begin{cases} 1, j = 2i - 1, \\ 0, j \neq 2i - 1, \end{cases}$$

for $i = 1, 2, ..., 10$ and $j = 1, 2, ..., 20$. The input $u(t)$ is the normalized running speed and output $y(t)$ is the scaled change of heart rate from the resting heart rate. The observed heart rate data are related to the running period, without observed data in the resting and recovery periods. The average heart rate where starting running speed 0 is used for the resting heart rate. Particularly for this research subject, 70 bpm is identified as the resting heart rate. When solving the ODE system (3), a vector $u(t) \in R^{10}$ will be interpolated instead of the scalar. Employing the Levenberg-Marquardt algorithm to minimize the sum of the squares of the residuals. The estimated parameters are shown in table II.

TABLE II
SIMULTANEOUSLY ESTIMATED PARAMETERS

Estimated parameters				
$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$
1	37.13	2.08×10^{-4}	2.31×10^{-5}	12.81

Combining equations (1) and (2) with the estimated parameters in table II, the identified model is described as:

$$\dot{x}_1(t) = -x_1(t) + 37.13x_2(t) + 0.22s^2(t),$$
$$\dot{x}_2(t) = -2.08 \times 10^{-4}x_2(t) + \phi(x_1(t)), \qquad (4)$$
$$\phi(x_1(t)) = \frac{2.31 \times 10^{-5}x_1(t)}{1 + exp(-(x_1(t) - 12.81))},$$
$$hr(t) = 4x_1(t) + h_r,$$

where $s(t)$ is the running speed km/hour in time t, $hr(t)$ is the heart rate at time t with the input speed $s(t)$, h_r is the resting heart rate, and time is in the unit of minute.

IV. MODEL PERFORMANCE EVALUATION

RMSE of the identified model (4) is calculated on training and testing sets separately. The minimum RMSE on the training sets is 6.05 bpm, the maximum RMSE is 11.76 bpm, and the median RMSE is 9.20 bpm. The minimum RMSE on the testing sets is 7.46 bpm, the maximum RMSE is 10.86 bpm, and the median RMSE is 9.35 bpm. The dynamics of the model is quite stable on the training and testing sets.

Figure 3 presents the prediction results of the identified model (4) of one exercise in the training sets. The red curve indicates the observed data and the green curve indicates the prediction. The prediction is performed on a complete exercise, rather than a short duration of a few seconds or minutes. The length of the predicted exercise is about 40 minutes. The residuals are also plotted at the bottom of Figure 3. For the initial 3 minutes, the residuals are quite large, and after that the residuals vary in a relative stable interval. The RMSE starting from 3 minutes to the exercise ending is 3.77 bpm compared with RMSE 6.59 bpm on the whole exercise session.

Figure 4 describes the prediction of an exercise on the testing set compared to original heart rate observation. Residuals are plotted at the bottom of Figure 4. The prediction is higher than the observed data in the peak and lower in the valley. The prediction RMSE from 3 minutes to the exercise ending is 8.43 bpm and RMSE is reduced 2.07 bpm compared to RMSE on the whole exercise session. In a nutshell, our model can capture the trend of the heart rate, but oscillation amplitude is subjected to the input $u(t)$.
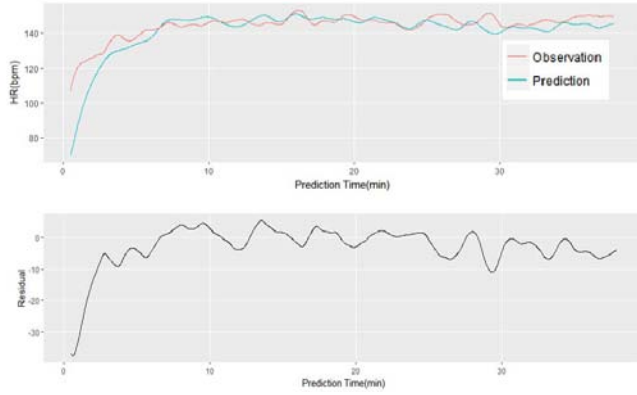

Fig. 3. Prediction and residual of a complete exercise on the training set.
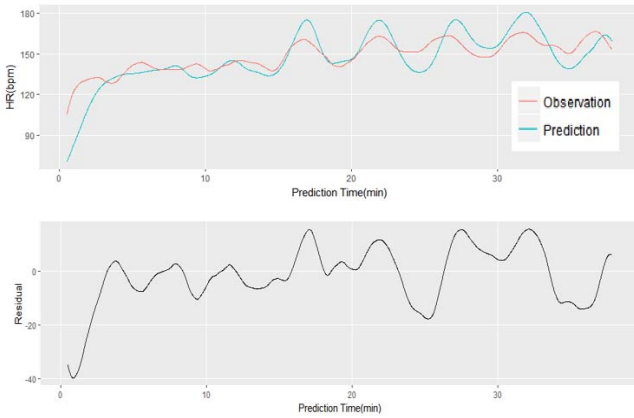

Fig. 4. Prediction and Residuals of a complete exercise on the testing set.

V. CONCLUSION AND FUTURE WORK

This paper proposes a nonlinear ODE model, which is adopted to describe the heart rate response on the outdoor running exercise. Most current research regarding to the heart rate prediction on the exercise focuses on indoor exercise and little research has been conducted on the outdoor running exercise. Indoor treadmill running usually suffers speed limitation and the models that suitable for treadmill running can not be used for outdoor running. Our model is proposed to predict the heart rate response on a complete outdoor running exercise session, instead of predicting the heart rate response for a

few seconds or minutes. We utilize the data of subject with different speed types outdoor running exercise in our analysis. The parameters of the model are identified by employing the Levenberg-Marquardt algorithm using all training exercises simultaneously. Our model achieves good performance as the difference of the model performance on the training and testing sets is not obvious. The performance of model (4) is quite stable and RMSE is greatly reduced from 3 minutes to the exercise ending.

The next steps include: 1) improve the model performance to reach the RMSE with 5 bpm; 2) perform analysis using different algorithms and models; 3) perform scalable experiments with more subjects holding different backgrounds with various speed to develop a general model that can be used as a baseline model; and 4) consider other affecting factors, such as running environment, which is an important factor for outdoor running exercise. The running exercise should be distinguish initially, for example, build a general model can be used for common subjects running on near flat ground and then consider the general model incorporating running exercise with higher ascent and descent slope.

REFERENCES

[1] M. Sumida, T. Mizumoto, and K. Yasumoto, "Estimating heart rate variation during walking with smartphone," Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 245. ACM Press 2013.

[2] E.M. Tapia, S.S. Intille, W. Haskell, K. Larson, J. Wright, A. King, and R. Friedman, "Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor," Proc. of the 11th IEEE Intl. Symp. on Wearable Computers (ISWC 2007), pp. 37–40, 2007.

[3] T.M. Cheng, A.V. Savkin, B.G.Celler, L. Wang, and S.W. Su, "A nonlinear dynamic model for heart rate response to treadmill walking exercise," IEEE International Conference on Engineering in Medicine and Biology Society (EMBS), pp. 2988–2991, 2007.

[4] T. M. Cheng, A. V. Savkin, B. G. Celler, S. W. Su, and L. Wang, "Nonlinear Modeling and Control of Human Heart Rate Response During Exercise With Various Work Load Intensities," IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, vol. 55, no. 11, NOVEMBER 2008.

[5] M. Paradiso, S. Pietrosanti, S. Scalzi, P. Tomei, C. Verrelli, "Experimental heart rate regulation in cycle-ergometer exercises," IEEE Trans. Biomed. Eng. Vol. 60, no. 1, pp. 135–139, 2013.

[6] D.Z. Baig, H. Su, T.M. Cheng, A. v. Savkin, S.W. Su, and B.G Celler, "Modeling of human heart rate response during walking, cycling and rowing," 2010 IEEE International Conference on Engineering in Medicine and Biology (EMBC), pp. 2553–2556, 2010.

[7] L. Wang, S.W. Su, B.G. Celler, G.S.Chan, T. M. Cheng, and A.V. Savkin, "Assessing the human cardiovascular response to moderate exercise: feature extraction by support vector regression," Physiol. Meas. Vol. 30, pp. 227–244, 2009.

[8] M. Yuchi and J. Jo, "Heart rate prediction based on physical activity using feedforwad neural network," International Conference on Convergence and Hybrid Information Technology, ICHIT 2008, pp. 344–350, 2008.

[9] M.I.A. Lourakis, "A brief description of the Levenberg-Marquardt algorithm implemented by levmar," Technical Report, Institute of Computer Science, Foundation for Research and Technology - Hellas, 2005.

[10] H.P. Gavin, "The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems," http://people.duke.edu/ hp-gavin/ce281/lm.pdf, March 2007,

[11] W.H. Press, S.A. Teukosky, W.T. Vetterling, and B.P. Flannery, "Numerical Recipes in C," Cambridge University Press, second edition, 1992.

[12] R. Shrager, A. Jutan, R. Muzic, and O. Till, "leasqr.m Octave-Forge, A collection of packages providing extra functionality for GNU Octave," 1992-2016.