Author(s): Levikari, Saku, Antti Immonen, Mikko Kuisma, Heikki Peltonen, Mika Silvennoinen, Heikki Kyröläinen & Pertti Silvennoinen

Title: Improving Energy Expenditure Estimation in Wrist-Worn Wearables by Augmenting Heart Rate Data With Heat Flux Measurement

Year: 2021

Version: Published version

Please cite the original version: S. Levikari *et al*., "Improving Energy Expenditure Estimation in Wrist-Worn Wearables by Augmenting Heart Rate Data With Heat Flux Measurement," in *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-8, 2021, Art no. 4003108, doi: 10.1109/TIM.2021.3053070.

Improving Energy Expenditure Estimation in Wrist-Worn Wearables by Augmenting Heart Rate Data With Heat Flux Measurement

Saku Levikari, *Member, IEEE*, Antti Immonen, *Member, IEEE*, Mikko Kuisma, Heikki Peltonen, Mika Silvennoinen, Heikki Kyröläinen, and Pertti Silventoinen, *Member, IEEE*

***Abstract*— Wearable electronics are often used for estimating the energy expenditure of the user based on heart rate measurement. While heart rate is a good predictor of calorie consumption at high intensities, it is less precise at low intensity levels, which translates into inaccurate results when estimating daily net energy expenditure. In this study, heart rate measurement was augmented with heat flux (HF) measurement, a form of direct calorimetry. A physical exercise test on a group of 15 people showed that HF measurement can improve the accuracy of calorie consumption estimates especially during rest and low-intensity activity when used in conjunction with heart rate information and vital background parameters of the user.**

***Index Terms*— Biomedical instrumentations, sensors, temperature and thermal.**

I. Introduction

THE advent of wearable electronics has enabled consumers to measure their vital signs during their everyday life. Devices, such as activity trackers, smart watches, and rings, are gaining ground in the measurement of biometric signals from the user. Typically, these devices contain one or more sensors for measuring signals, such as heart rate, skin temperature, humidity, and movement [1]. Optical heart rate (HR) tracking, or photoplethysmography (PPG), is one of the most common biometric measurements, found in up to 98% of modern smart watches [2]. PPG, among other biometric measurements, gives the wearer the ability to track the level of their physical activity, which may be beneficial for applications such as sports performance monitoring. Nonsports-related applications, such as weight loss [3], sleep tracking [4], and health monitoring [5], have also seen an increase in popularity.

A common application for wearable technology is to estimate the user's energy expenditure (EE), often by utilizing heart rate measurements [6]. However, several studies have reported that the EE estimate provided by commercial wearables during common physical activities is often relatively inaccurate. Scherbina *et al.* [7] evaluated the accuracy of EE and HR estimates on seven different wearables during walking and cycling activities, reporting net EE errors between 25% and 50%. Pope *et al.* [8] tested the EE and HR estimation accuracy of four wearables during exergaming (playing a boxing game on a Nintendo Wii gaming console), with errors in the mean EE ranging from 10% to 40%. In another study, Pope *et al.* [9] reported mean EE errors between 25% and 50% during walking, jogging, and running, and errors of over 50% when the subjects were at rest. While the aforementioned studies [7], [8] reported reasonably good accuracies for the heart rate estimates provided by wearables, the error in the energy expenditure estimates may limit the usability of applications that rely on calorie consumption information, such as sports performance metering. In particular, the high error rates for rest energy expenditure can negatively affect applications for diet control and weight loss because the resting metabolic rate (RMR) typically accounts for 60%–80% of the total daily energy expenditure [10].

Some of the key reasons for the inaccurate EE estimates can be found in the placement and form factor of wearable devices, which impose limitations on the sensor technology. For example, PPG-based heart rate measurements are easily hampered by motion artifacts, especially when the wearable device is worn loosely for comfort. Moreover, the biosignals being measured may not provide enough information to accurately determine the EE in various situations. While heart rate has been observed to be a good indicator of EE at moderate-to-high intensity activities, the error increases during low-intensity activity [11]. This is especially problematic for smart watches and activity trackers intended for monitoring everyday activities, most of which are of low intensity. Therefore, in order to achieve a higher EE estimation accuracy in the setting of everyday activity monitoring, heart rate tracking should be complemented with other biometric measurements.

A. Measuring the Human Metabolic Heat

The human body generates metabolic heat as a sum of basal metabolism and physical activity. Under most conditions, the ambient temperature is lower than the human body

Manuscript received July 31, 2020; revised December 9, 2020; accepted December 30, 2020. Date of publication January 22, 2021; date of current version February 11, 2021. This work was supported by Business Finland (Project Q-Health). The Associate Editor coordinating the review process was Bruno Ando. *(Corresponding authors: Saku Levikari.)*

Saku Levikari, Antti Immonen, Mikko Kuisma, and Pertti Silventoinen are with the Department of Electrical Engineering, Lappeenranta-Lahti University of Technology LUT, 53850 Lappeenranta, Finland (e-mail: saku.levikari@lut.fi; antti.immonen@lut.fi; mikko.kuisma@lut.fi; pertti.silventoinen@lut.fi).

Heikki Peltonen and Heikki Kyröläinen are with the Faculty of Sport and Health Sciences, University of Jyväskylä, 40014 Jyväskylä, Finland (e-mail: heikki.peltonen@jyu.fi; heikki.kyrolainen@jyu.fi).

Mika Silvennoinen is with Doseco Oy, 40500 Jyväskylä, Finland (e-mail: mmjsilvennoinen@gmail.com).

Digital Object Identifier 10.1109/TIM.2021.3053070

temperature, i.e., heat is dissipated to the environment. As the amount of heat generated depends on the level of physical activity, the measurement of heat exchange between a person and the environment can be used to provide information about energy expenditure. Under light physical activity, the primary modes of heat transfer from skin to the environment are convection (air currents), conduction (through objects in contact with the human body), and radiation [12]. As the level of physical activity increases, the role of evaporative heat transfer (perspiration) increases [13].

The aim of this study was to find out whether direct measurement of heat transfer from a human to the environment could be used to improve the accuracy of energy expenditure estimation over heart-rate-based methods. To this end, local heat transfer measurements were performed using a thermoelectric heat flux (HF) sensor, in conjunction with a humidity sensor to consider evaporative heat transfer.

While HF sensors have existed for decades [14], their use in the field of wearable technology has remained limited to only a few commercial products, such as devices with a form factor of an armband [15]. However, wrist-worn wearables, such as activity trackers and smart watches, are the most common type of wearable device. Thus, if HF measurement was to be employed in the form factor of a smart watch, the most likely measurement location would be either the dorsal or medial side of the user's wrist. Therefore, this work focuses on the use of a single, wrist-worn HF sensor. The location of the HF sensor is likely to affect the relationship between the measurement and the user's energy expenditure because excess heat produced at the site of muscular activity will take some time to propagate to the measurement location. Therefore, the suitability of HF measurement for assessing the user's energy expenditure was evaluated for physical activities performed at different intensities: sitting, standing, walking, cycling, and using an arm crank ergometer.

Physiological trials were conducted on 15 persons, and the subjects' EE was estimated by a linear model using combinations of HF, humidity, and heart rate measurements. The results show that HF measurement with heart rate information can improve the accuracy of the EE estimate over a heart-rate-based model, especially during rest or low-intensity activity. The accuracy is increased even further if a measurement of the evaporation rate is included in the model.

This work is an extended version of a previously published conference article, originally presented at the IEEE International Instrumentation and Measurement Technology Conference (I^2MTC 2020) [16].

II. Methods

A. Heat Flux Measurement

In this study, HF measurement was used to provide information about the test subjects' metabolic activity and, thus, energy expenditure. According to Fourier's law [17], heat flows from a higher temperature to a lower. The HF q describes the transfer of thermal energy, i.e., heat flow

$$\vec{q} = -k\nabla T \quad (1)$$

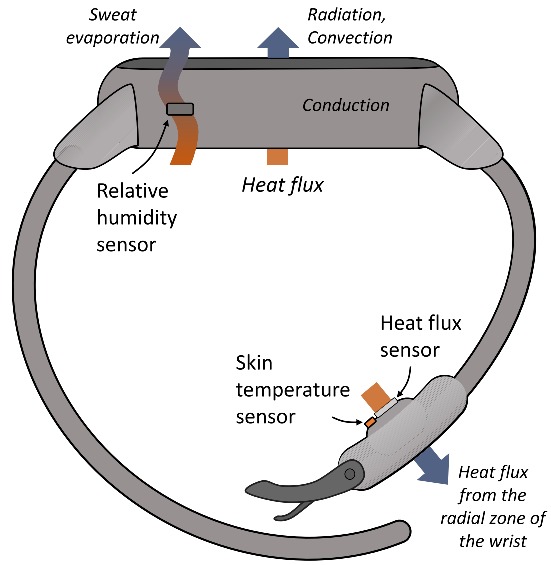

Fig. 1. Schematic of the prototype bracelet equipped with sensors for HF, humidity, and skin temperature measurements. The arrows indicate heat transfer from the user to the environment. When being worn, the HF sensor measures the conduction of bioheat from the radial zone of the user's wrist to the bracelet. From there, the heat is then transferred to the environment by means of convection and radiation. To consider the evaporative heat transfer, the bracelet is equipped with a humidity sensor. The skin temperature sensor was only used for control purposes, and the data were not considered in this study.

which in a 1-D, steady-state case, can be expressed as

$$q = -k\frac{T_{\text{hot}} - T_{\text{cold}}}{h} \quad (2)$$

where k is the thermal conductivity of an object and h is the distance between the hot and cold points. While HF could be estimated according to (2), this would require knowing the values for k and h. Moreover, the thermal mass associated with the object will restrict the rate of change for T_{hot} and T_{cold} during transients in the HF. Instead, HF can be measured, e.g., using a thermoelectric HF sensor, which generates voltage proportional to the heat flow through the sensor structure [18]. Besides few exceptions, HF sensors have not been used in wearable electronics, in part due to the lack of low-cost, mass-produced sensors. However, in recent years, HF sensor designs suitable for mass production have emerged [19], [20], thus facilitating direct HF measurement in wearable electronics. A commercially available HF sensor was used in this study to measure the transfer of metabolic heat to the environment. The purpose of this measurement was to provide information about the metabolic state and activity of a person because metabolism and physical activity produce heat that is transferred to the environment.

B. Experimental Setting

In order to experimentally verify whether a single-site HF measurement could provide useful information for energy expenditure estimation, low-to-medium intensity randomized physical exercise routines were performed in laboratory conditions. During the physical activity, the subjects were equipped with a respiratory calorimeter, which provided the ground-truth data on the subjects' energy expenditure. In addition to the

Fig. 2. Experimental setup: the subject is wearing a custom-made bracelet equipped with HF and temperature sensors on his left hand. A reference EE value was obtained using a spirometry calorimeter; furthermore, the subject's heart rate was monitored using an ECG chest strap.

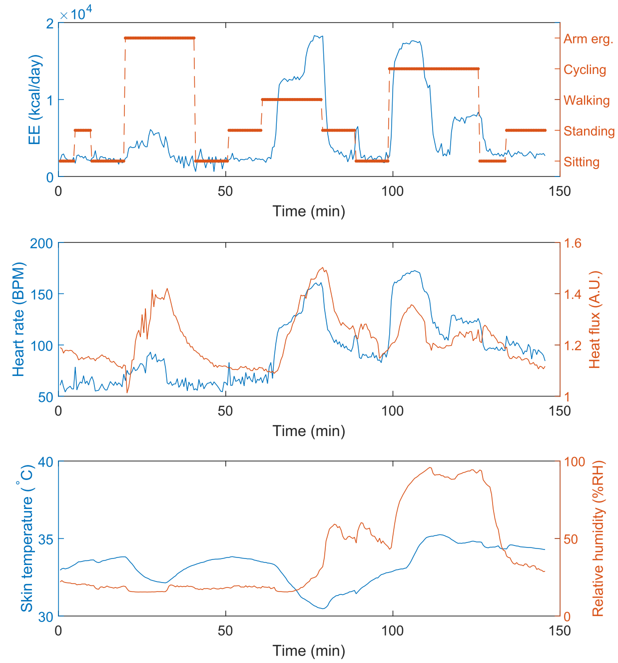

Fig. 3. Example of raw data collected from a subject during a 2.5-h exercise protocol. The subject (male, 29, 165 cm, 72 kg, and overall level of physical activity 6/10) performed a randomized protocol of various activities with different intensities. The ground-truth EE values were obtained using spirometry. HF, humidity, and skin temperature values were measured using a prototype bracelet; the skin temperature values were used for control purposes only and were not included in the further analysis. The heart rate of the subject was measured using an ECG chest strap.

respiratory calorimeter, the subjects were equipped with a custom-made bracelet, with a design similar to modern smart watches (see Fig. 1). The bracelet was wirelessly connected to a separate data acquisition device and equipped with the following sensor hardware:

1) HF sensor (greenTEG gSKIN-XM) [21] on the medial side of the wrist;
2) temperature sensors (Texas Instruments LMT70) on the medial side of the wrist (one for skin temperature and other for heat sink/ambient);
3) humidity sensors (Texas Instruments HDC2080) on the dorsal side of the wrist (one 2–3 mm from skin and other for ambient humidity).

The HF sensor was positioned on the medial side of the user's wrist, above the radial artery. This location was chosen over the dorsal side, which has little soft tissue and blood circulation, resulting in a lower HF density. The sensor was attached to a small heat sink, which would then dissipate the heat into the environment. Instead of equipping the bracelet with an optical heart rate sensor, the subjects' heart rate was recorded using an ECG chest strap, as this method can be considered more accurate and reliable than PPG-based methods found in typical wearables. Fig. 2 shows a test subject wearing the sensor bracelet and a spirometer during an exercise protocol.

C. Exercise Protocol

The experiments for this study were conducted on 15 healthy persons (nine male and six female) between 23 and 45 of age (mean±S.D. 34.7 ± 7.0 years). Each subject's exercise protocol included 2–5 different types of activities from five categories: sitting, standing, treadmill walking, cycling, and arm crank ergometry. The order of the activities was randomized for each person, with the durations of individual activities ranging from 5 to 45 min. For walking activities, the speed and climb angle of the treadmill were randomized. Furthermore, speed and intensity were randomized for each set of cycling and arm ergometry activities. The total duration of the exercise protocols ranged from 96 to 163 min per subject, with a total of 33.5 h of physical activity data recorded.

D. Data Acquisition and Preprocessing

The task of estimating the subject's energy expenditure was approached as a regression problem. To this end, the time series data collected from the subjects were composed into a single data set containing heart rate (HR), HF, and percent relative humidity (%RH) as predictor variables and spirometry data (EE) as the response variable, i.e., ground-truth values.

During the experiments, the predictor (HR/HF/%RH) and response (EE) variables were sampled at different rates. To construct an equally spaced time series data set, preprocessing and resampling were performed on the measurement data. The ECG equipment registered each R-R interval, while the respiratory calorimeter analyzed each exhaled breath. These values were composed into 30 s intervals, from which the mean values of each signal were extracted. On the other hand, the HF and temperature measurements onboard the bracelet were sampled at 20 Hz. Following the convention with the ECG and calorimeter equipment, these values were downsampled into corresponding 30 s intervals using nearest neighbor interpolation.

An example of measurement data collected from a test subject in Fig. 3 shows that heart rate and HF measurements

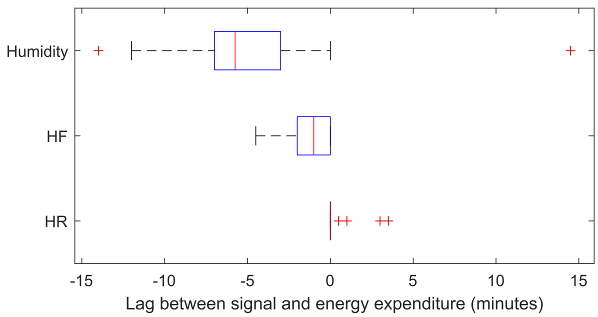

Fig. 4. Experimentally determined lag between the measured biosignals and energy expenditure. Each box contains the mean lag values for each of the 15 subjects. The magnitude of the lag values was determined by applying a cross correlation between the biosignals and energy expenditure. The delay associated with the signals is mainly of physiological origin: heart rate produces a near-instantaneous response to changes in EE, whereas HF responds with a slight delay because of thermal masses. Perspiration often follows physical activity with a longer delay.

react to changes in energy expenditure in different ways. HF has a slightly delayed response, while heart rate changes more immediately. Moreover, HF shows a more pronounced response in particular to walking and arm crank ergometry.

The longer rise and fall times present in HF data possibly result from the thermal mass of the sensor bracelet, as well as heat accumulation under the measurement location. This lag effect caused by the thermal mass was estimated by calculating cross correlation between HF and EE for all subjects and finding the lag value at the maximum of the cross correlation. The results in Fig. 4 show that the maximum cross correlation between HF and EE is reached within 2 min for 75% of the subjects. Furthermore, because of the effect of thermal accumulation seen in Fig. 3, the HF signal may contain useful information about physical activity outside momentary measurement values, as the past heat output affects the current value. To consider these effects, three predictor variables were derived from the HF data:

1) current HF value (at 30 s intervals);
2) mean HF from the past 30 s–2 min;
3) mean HF from the past 2.5 min–15 min.

Similar lag terms were introduced into the humidity measurements, with longer intervals to accommodate for the longer rise and fall times:

1) mean %RH value from the past 0–5 min;
2) mean %RH value from the past 5–10 min;
3) mean %RH value from the past 10–15 min.

While the lag terms provide information about the past physical activity, they do not remove the effect of longer rise and fall times associated with HF and humidity measurements. However, the future values of predictor variables (i.e., by introducing lag into the EE estimate) were not considered to keep the estimated EE values real time in nature.

In addition to the measured signals, the physical parameters of the test subjects were also considered as background variables for EE estimation. These included each person's age, gender, height, weight, and overall level of physical activity on the scale from 1 to 10.

The effect and importance of the biosignals were evaluated using different combinations of the predictor variables. A heart-rate-based linear model was used as the baseline EE estimator, against which different combinations of HF and humidity measurement were compared. In each test scenario, the output from the respiratory calorimeter was used as the ground-truth value for the EE. A summary of variable combinations is shown in Table I.

TABLE I
PREDICTOR VARIABLES USED FOR EE ESTIMATION IN THE TEST SCENARIOS

Variable	HR	HF	HR+HF	HR+RH	HR+HF+RH
$\bar{HR}$ (0–30 s)	x		x	x	x
$\bar{HF}$ (0–30 s)		x	x		x
$\bar{HF}$ (30s–2 min)		x	x		x
$\bar{HF}$ (2–15 min)		x	x		x
$\%\bar{RH}$ (0–5 min)				x	x
$\%\bar{RH}$ (5–10 min)				x	x
$\%\bar{RH}$ (10–15 min)				x	x
Height (cm)	x	x	x	x	x
Body mass (kg)	x	x	x	x	x
Age (y)	x	x	x	x	x
Gender (0=F/1=M)	x	x	x	x	x
Phys. activity (1-10)	x	x	x	x	x

E. Energy Expenditure Estimation

Because of the preliminary nature of this study and the limited amount of data available, the energy expenditure y was modeled as an ordinary least squares (OLS) problem

$$\mathbf{y} = \mathbf{H}\theta + \mathbf{v} \tag{3}$$

with the goal of estimating θ by minimizing the objective function

$$\hat{\theta} = \arg\min_{\theta} \|\mathbf{y} - \mathbf{H}\theta\|^2 \tag{4}$$

where $\mathbf{H}$ is the observation matrix, θ is a vector of model parameters, and $\mathbf{v}$ is an error vector for which $\langle \mathbf{v} \rangle = 0$. For each test subject, the observation matrix $\mathbf{H}$ was constructed by horizontally concatenating the predictor variables $\mathbf{x}_i$ of T time steps in Table I with repeated values of the subject's background parameters and a vector of ones as

$$\mathbf{H} = \begin{pmatrix} x_1^{t=1} & \cdots & x_n^{t=1} & \text{Height} & \cdots & \text{Act.} & 1 \\ x_1^{t=2} & \cdots & x_n^{t=2} & \text{Height} & \cdots & \text{Act.} & 1 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ x_1^{t=T} & \cdots & x_n^{t=T} & \text{Height} & \cdots & \text{Act.} & 1 \end{pmatrix}. \tag{5}$$

For the OLS model in (3), the estimates for the parameters in θ were found by vertically concatenating all observation matrices (5) and minimizing the least squares criterion (4) using the pseudoinverse

$$\hat{\theta} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{y} \tag{6}$$

from which the EE estimates were obtained as $\hat{\mathbf{y}} = \mathbf{H}\hat{\theta}$. To prevent data leakage between adjacent concatenated observation matrices, the time-lagged features in Table I were extracted before the concatenation process. The order of the observation matrices themselves has no effect on the resulting model, as (4) minimizes the total sum of squared errors between the model outputs $\hat{y}$ and all the EE data points y.

Finally, a moving average filter of five time steps (2.5 min) was applied to the estimates, as it was empirically observed to reduce noise in the output signal while maintaining the ability to follow the rising and falling edges within the calorimeter data.

In order to consider the generalizability of the model (3), all testing was performed using leave-one-out cross validation: the model was fitted into the data from all but one subject, and the performance of the model was tested with the left-out subject, repeating the process for all subjects. For evaluating the overall performance of the model using different sets of predictor variables, the coefficient of determination, R^2, was chosen as the performance metric

$$R^2 = 1 - \frac{\mathrm{mean}(y - \hat{y})^2}{\mathrm{mean}(y - \bar{y})^2}. \tag{7}$$

The reason for choosing R^2 over other commonly used metrics, such as the root-mean-square error (RMSE), mean absolute error (MAE), or net EE error, was that the duration and composition of the exercise protocol were randomized for each subject. Because R^2 sets the mean squared error of $\bar{y}$ as the baseline at $R^2 = 0$, the resulting value considers the variance in energy expenditure within each individual exercise protocol. However, for evaluating the model accuracy at individual types of activity, mean absolute percentage error was used

$$\mathrm{MAPE} = \mathrm{mean}\left|\frac{y - \hat{y}}{y}\right| \cdot 100. \tag{8}$$

III. Results

In order to verify whether direct HF measurement has a positive effect on the accuracy of the EE estimate, the OLS model (3) was tested using different combinations of input variables: HR only, HF only, HR and HF combined, HR and %RH combined, and, finally, by combining HR, HF, and %RH. All activity types were used for both fitting and evaluating the OLS model, and all results were obtained using leave-one-out cross validation on a per-subject basis. An example of the output of the OLS model using different predictors is shown in Fig. 5, where the model with HR+HF+%RH inputs gives the best results, except for the last 30 min.

The cross-validation results in Fig. 6 show that the median R^2 values show no significant difference between different sets of predictors, with the exception of using only HF, which yields significantly lower R^2 scores. In terms of median R^2 values, using only heart rate as a predictor yields nearly as good results as heart rate complemented with HF and/or humidity measurements. Instead, the key differences are seen in the distribution of the R^2 scores; complementing the heart rate measurements with either HF or relative humidity measurements yields a notable decrease in variance of the R^2 scores, with the best results obtained by combining HR, HF, and %RH. Adding either HF or %RH measurements also significantly increases the R^2 scores of the lower 50 % of the subjects, for outliers in particular. This, in turn, increases the mean R^2 values by a significant amount.

The overall effect of each predictor on the R^2 score was examined by calculating the permutation importances [22] of the variables. This was done by concatenating the measurements (5) of all the test subjects and randomly partitioning the rows of the resulting observation matrix into training (80%) and testing (20%) sets. First, a baseline score was obtained by fitting the model into the data in the training set and testing the R^2 score on the testing set. Then, the data in the test set (column vectors in (5), with the exception of the last column), were randomly shuffled one predictor at a time, and the decrease in the R^2 score with respect to the baseline was evaluated for each corrupted predictor. Because the range of numerical values was different for each predictor, the data were normalized to zero mean and unit variance before conducting the analysis.

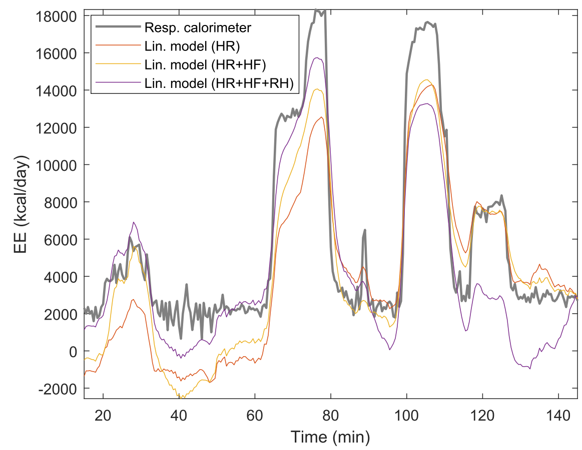

Fig. 5. Example results of EE estimates from linear models using different sets of predictor variables. The gray line indicates the ground-truth EE values, measured using the respiratory calorimeter. The input data and exercise protocol are shown in Fig. 3.

The results from the feature importance analysis in Fig. 7 agree with the cross-validation results in Fig. 6: out of all predictors, heart rate is the single most important feature. However, the current and lagged HF values have a greater impact on R^2 of the outcome than the subject's weight, height, or age. The importance of the HF values is also significantly pronounced over the humidity values, even though the cross-validation results show that both signals are approximately equally important. A likely explanation for this disparity is that the humidity values change more slowly over time than the HF values; thus, the lagged values of %RH are largely redundant, i.e., permuting one of them has little effect on the outcome. Conversely, HF shows a more immediate response for changes in EE, as also suggested by the highest importance of the nonlagged HF value in Fig. 7.

Finally, the performance of the OLS model was evaluated separately over each category of activity (sitting, standing, walking, cycling, and arm ergometry). As the data were partitioned into mutually similar activities, MAPE (8) was chosen as the performance metric. The results in Fig. 8 clearly show that the addition of HF and %RH values to HR yields a notable decrease in error when the subject is sitting, standing, or walking. For the first two categories, appending both HF and humidity measurements results in a reduced error, whereas during walking, adding the %RH data has little effect on the error. During cycling, neither of the added biosignals has a considerable effect on the accuracy, whereas for arm crank ergometry, the addition of HF measurement actually results in a slightly increased error. This decrease in accuracy is

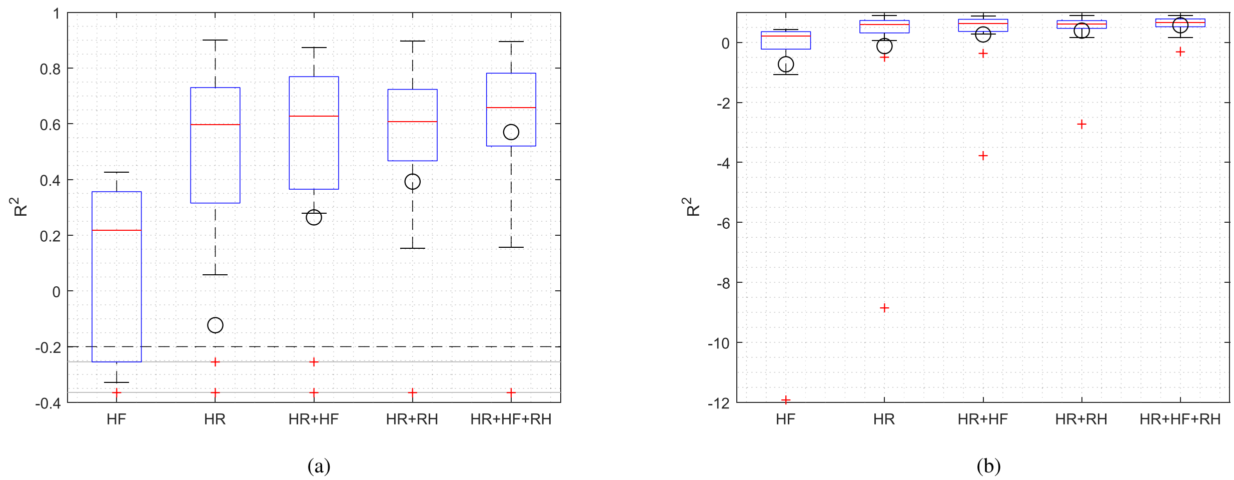

Fig. 6. Cross-validation results for the EE estimation model using different predictor combinations, represented as box plots. Each vertical box represents the distribution of the R^2 scores for the 15 test subjects. The median of each distribution is indicated by a red line, and the blue box around the median represents the interquartile range (IQR), i.e., 50% of the samples. The whiskers outside the blue boxes represent the remaining 25% of the samples each. The samples more than 1.5 IQR away from the blue box are marked as outliers (red "+" sign). The mean of each distribution is indicated by a black circle. (a) R^2 values between -0.2 and 1, zoomed-in view from (b). Values below -0.2 have been compressed between the gray lines. (b) Full range of the R^2 values, showing the impact of the outliers (indicated by red crosses) on the mean R^2 values.

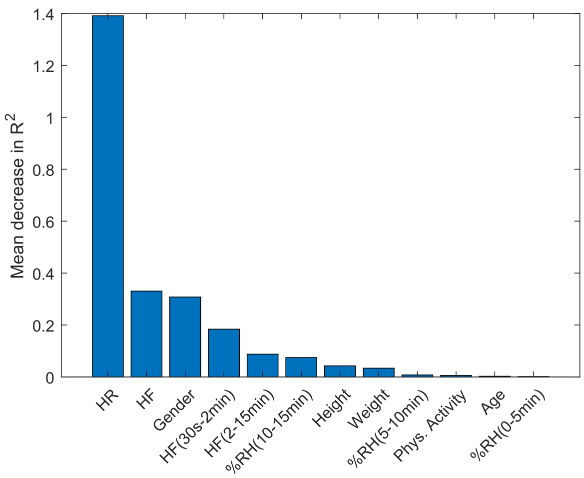

Fig. 7. Visualization of the relative importance of each variable for the EE estimation model. The importances were calculated using the permutation method: the OLS model was fitted into randomly selected 80% of the full data set, and baseline accuracy was obtained from the remaining 20% of data. Each predictor variable in the test set was then randomly shuffled one by one, and the decrease in the R^2 score with the corrupted data was calculated. The higher the decrease in R^2 is, the more important the feature is considered for the model. The results were obtained by repeating the process 1000 times and calculating the mean decrease in R^2 for each predictor.

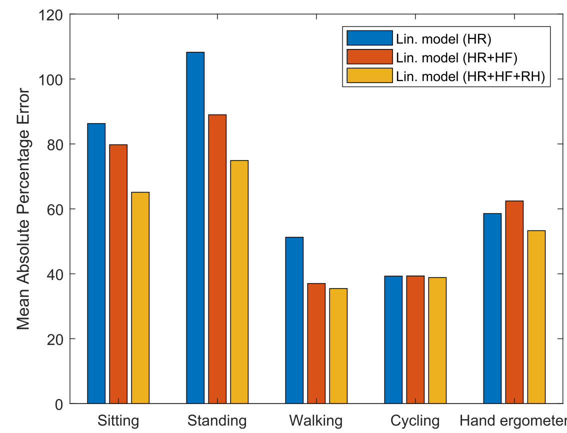

Fig. 8. Effect of added HF and RH measurements on the error of EE estimates at different types of activity. The results were obtained by cross-validating the OLS model across different types of physical activity using various combinations of predictor variables. Adding HF and humidity data to heart rate information yields a notable decrease in error during low-intensity activities, i.e., sitting, standing, and walking.

probably due to the positioning of the HF sensor, which results in motion artifacts and possibly intermittent thermal contact during extreme wrist movement. Overall, the results in Fig. 8 suggest that the inclusion of HR and %RH data has the greatest effect at low energy expenditure levels. To verify this, the mean absolute error of each model was inspected at different levels of EE. The results in Fig. 9 show that the addition of HF and %RH data has indeed a pronounced effect on the MAPE scores below 4000 kcal/day, although they also show a moderate positive effect on levels above 14 000 kcal/day.

The reasons for the effects of added HF and %RH measurements across different levels and types of activity are probably due to several factors. For instance, heart rate variability (HRV) is known to be greater at lower heart rates because of the cycle length dependence [23], decreasing the precision of HR-based EE estimates. However, in this study, the mean HR value is calculated over a period of 30 s, which mitigates the effect of HRV. A more significant factor is the change in the stroke volume of the heart between low and moderate intensity, i.e., a change in heart rate does not fully reflect the total change in the cardiac output [24].

On the other hand, HF measurement can provide a more stable predictor for EE estimation at low intensities, as a result of factors such as more stable skin contact and reduced air currents. As the intensity of physical activity is increased, the correlation between heart rate and energy expenditure increases. In contrast, the HF signal may suffer from changes in skin contact during physical activity, and the magnitude of HF response depends on the type of activity, as shown in Fig. 3. The relationship between the type of physical activity and the resulting HF response can be considered to be governed by the location of muscular activity and the resulting

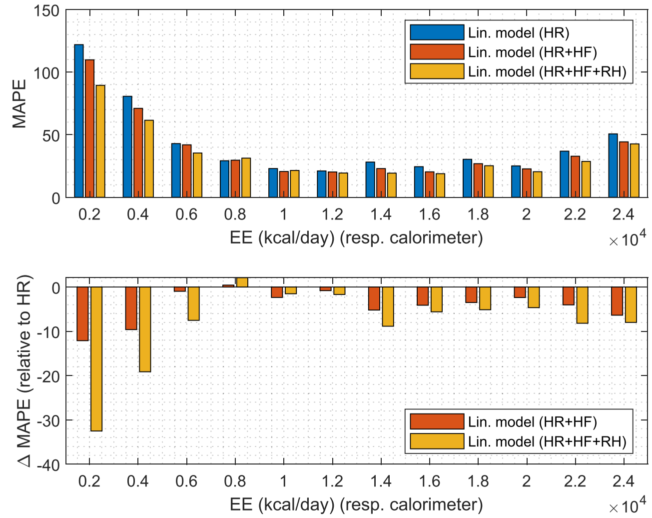

Fig. 9. MAPE for the linear models across all test subjects at different levels of energy expenditure. Top: absolute MAPE values using different combinations of predictors. Bottom: change in the MAPE relative to the baseline (HR-based) model. The inclusion of HF and %RH data results in a significant decrease in the MAPE at EE levels below 4000 kcal/day. The results were obtained by binning the EE estimates according to the true energy expenditure level, as measured using the respiratory calorimeter. Each bin (group of bars) is 2000 kcal/day wide, and the x-axis labels represent the endpoint of each bin (the data for the bin 2000 kcal/day range from 0 to 2000 kcal/day and so on).

increase in blood perfusion, with respect to the location of the HF measurement.

Based on the results in Figs. 6, 8, and 9, the addition of humidity data also has a positive impact on the accuracy of the EE estimates. This is an expected result because the HF sensor alone does not register the amount of heat transfer caused by perspiration. However, the humidity signal experiences a significant lag with respect to the EE values, which has to be considered when processing the measurements.

IV. Discussion

Based on the results, the combination of HF and heart rate data yields more accurate estimates of energy expenditure than using heart rate only. The improvements are mainly seen as increased mean R^2 scores because of the reduced effect of outliers, even though the variance of the scores is also reduced. Both of these effects are more pronounced when humidity values, i.e., the level of perspiration, are included in the model. These findings suggest that HF measurement (including the measurement of evaporative heat) can be applied to construct more robust models for the estimation of energy expenditure. Further inspection revealed that the greatest impact of the addition of HF data is seen during rest and low-intensity activity, i.e., sitting, standing, and walking.

In terms of overall predictive power, heart rate measurement surpasses HF data by a considerable amount, at least when using the current linear model. However, the results indicate that HF measurement could be used to improve the accuracy and reliability of EE models, especially during low-activity conditions and rest. The advances of HF over HR likely stem from the changes within heat rate variability and stroke volume under these conditions, while the variance within HF data is decreased, due to a more stable contact between the HF sensor and skin. These properties could be advantageous in situations where the quality of long-term, low-level EE estimates is crucial, such as applications for weight loss and diet control. The impact of better low-level EE estimates could be significant for persons with a sedentary activity level because the contribution of physical activity to the total energy expenditure is diminished. In particular, HF measurement could be used in determining the resting metabolic rate of a subject more accurately than current methods, which are often based on the person's physiological parameters. In addition to weight control, improved RMR information could be useful in preventing malnutrition among elderly people because the resting metabolic rate changes with age [10].

The authors would like to stress that the results presented are preliminary in nature and obtained using a simple linear model, which may not utilize the added HF measurements to their full extent. Using a nonlinear model for EE estimation would probably result in higher accuracy in terms of R^2, even more so if the model took into account the temporal nature of the measurement data. To this end, the use of models, such as recurrent neural networks, should be investigated. Furthermore, the use of HF measurements should be evaluated over a longer period of time to observe the effect on the daily net energy expenditure.

V. Conclusion

Estimation of low-to-moderate intensity energy expenditure was performed on 15 subjects based on combinations of heart rate, HF, and humidity measurements. The results showed that the best accuracy in terms of R^2 values was achieved by augmenting heart rate measurement with both HF and humidity data. Moreover, the addition of HF appears to be particularly effective at low intensities or when the subject is at rest. The results suggest that HF measurement could be an advantageous feature for wearable devices.

Acknowledgment

The authors would like to thank greenTEG AG, Switzerland, for providing the heat flux sensors used in this study.

References

[1] J. Heikenfeld *et al.*, "Wearable sensors: Modalities, challenges, and prospects," *Lab Chip*, vol. 18, no. 2, pp. 217–248, 2018, doi: 10.1039/C7LC00914C.

[2] C.-C. Chang, C.-T. Wu, B. I. Choi, and T.-J. Fang, "MW-PPG sensor: An on-chip spectrometer approach," *Sensors*, vol. 19, no. 17, p. 3698, Aug. 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/17/3698

[3] S. W. Cheatham, K. R. Stull, M. Fantigrassi, and I. Motel, "The efficacy of wearable activity tracking technology as part of a weight loss program: A systematic review," *J. Sports Med. Phys. Fitness*, vol. 58, no. 4, pp. 534–548, Apr. 2018.

[4] M. De Zambotti, N. Cellini, A. Goldstone, I. M. Colrain, and F. C. Baker, "Wearable sleep technology in clinical and research settings," *Med. Sci. Sports Exerc.*, vol. 51, no. 7, pp. 1538–1557, Jul. 2019.

[5] M. Wu and J. Luo, "Wearable technology applications in healthcare: A literature review," *Online J. Nursing Informat.*, vol. 3, p. 23, Nov. 2019. [Online]. Available: https://www.himss.org/resources/wearable-technology-applications-health%care-literature-review

[6] *An Energy Expenditure Estimation Method Based on Heart Rate Measurement*, Firstbeat Technol., Yväskylä, Finland, Mar. 2012. [Online]. Available: https://assets.firstbeat.com/firstbeat/uploads/2015/10/white_paper_ener%gy_expenditure_estimation.pdf

[7] A. Shcherbina *et al.*, "Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort," *J. Personalized Med.*, vol. 7, no. 2, p. 3, May 2017, doi: 10.3390/jpm7020003.

[8] Z. C. Pope, J. E. Lee, N. Zeng, and Z. Gao, "Validation of four smartwatches in energy expenditure and heart rate assessment during exergaming," *Games for Health J.*, vol. 8, no. 3, pp. 205–212, Jun. 2019, doi: 10.1089/g4h.2018.0087.

[9] Z. C. Pope, N. Zeng, X. Li, W. Liu, and Z. Gao, "Accuracy of commercially available smartwatches in assessing energy expenditure during rest and exercise," *J. for Meas. Phys. Behaviour*, vol. 2, no. 2, pp. 73–81, Jun. 2019. [Online]. Available: https://journals:humankinetics:com/view/journals/jmpb/2/2/article-p73:xml

[10] T. M. Manini, "Energy expenditure and aging," *Ageing Res. Rev.*, vol. 9, no. 1, pp. 1–11, Jan. 2010, doi: 10.1016/j.arr.2009.08.002.

[11] J. Parak, M. Uuskoski, J. Machek, and I. Korhonen, "Estimating heart rate, energy expenditure, and physical performance with a wrist photoplethysmographic device during running," *JMIR mHealth uHealth*, vol. 5, no. 7, p. e97, Jul. 2017. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/28743682

[12] M. S. Ferreira and J. I. Yanagihara, "A transient three-dimensional heat transfer model of the human body," *Int. Commun. Heat Mass Transf.*, vol. 36, no. 7, pp. 718–724, Aug. 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0735193309000773

[13] T. E. Bernard, F. N. Dukes-Dobos, and J. D. Ramsey, "Evaluation and control of hot working environments: Part II—The scientific basis (knowledge base) for the guide," *Int. J. Ind. Ergon.*, vol. 14, no. 1, pp. 129–138, 1994. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0169814194900108

[14] S. Z. Sapozhnikov, V. Y. Mityakov, and A. V. Mityakov, "Gradient heat-flux sensors: Possibilities and prospects of use," *Thermal Eng.*, vol. 53, no. 4, pp. 270–278, Apr. 2006, doi: 10.1134/S0040601506040033.

[15] K. Lyden, T. Swibas, V. Catenacci, R. Guo, N. Szuminsky, and E. L. Melanson, "Estimating energy expenditure using heat flux measured at a single body site," *Med. Sci. Sports Exerc.*, vol. 46, no. 11, pp. 2159–2167, Nov. 2014. [Online]. Available: https://journals.lww.com/acsm-msse/Fulltext/2014/11000/Estimating_Energ%y_Expenditure_Using_Heat_Flux.16.aspx

[16] S. Levikari *et al.*, "Improving energy expenditure estimation in wearables using a heat flux sensor: First observations," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, Dubrovnik, Croatia, May 2020, pp. 1–5.

[17] J. H. Lienhard IV and J. H. V. Lienhard, *Transfer Textbook Textbook*, J. H. Lienhard and J. H. Lienhard, Eds. Cambridge, MA, USA: Phlogiston Press, 2015.

[18] S. Z. Sapozhnikov, V. Y. Mityakov, and A. V. Mityakov, *Heatmetry*, 1st ed. Cham, Switzerland: Springer, 2020.

[19] A. Immonen, S. Levikari, F. Gao, M. Kuisma, and P. Silventoinen, "MEMS heat flux sensor," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, Dubrovnik, Croatia, May 2020, pp. 1–5.

[20] W. Glatz, S. Muntwyler, and C. Hierold, "Optimization and fabrication of thick flexible polymer based micro thermoelectric generator," *Sens. Actuators A, Phys.*, vol. 132, no. 1, pp. 337–345, Nov. 2006.

[21] GreenTEG. (2020). *Heat Flux Sensor Gskin Datasheet*. Accessed: Mar. 9, 2020. [Online]. Available: https://shop.greenteg.com/heat-flux-measurement-heat-flux-sensor-gskin-%xm/

[22] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[23] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers Public Health*, vol. 5, Sep. 2017, doi: 10.3389/fpubh.2017.00258.

[24] A. Guyton and J. Hall, *Textbook of Medical Physiology*, 9th ed. Philadelphia, PA, USA: W. B. Saunders Company, 1996.

Saku Levikari (Member, IEEE) was born in Finland, in 1991. He received the M.Sc. degree from the Lappeenranta University of Technology (LUT), Lappeenranta, Finland, in 2018. He is currently pursuing the Ph.D. degree with the LUT School of Energy Systems, Laboratory of Applied Electronics.

His research interests are data analytics and machine learning, with topics ranging from biosignal measurements to power electronics reliability.

Antti Immonen (Member, IEEE) was born in Finland in 1993. He received the B.Sc. and M.Sc. degrees from LUT University, Lappeenranta, Finland, in 2018 and 2019, respectively. He is currently pursuing the Ph.D. degree with the LUT Laboratory of Applied Electronics.

His main research topics include thermoelectric devices, sensors, and biophysical instrumentation.

Mikko Kuisma was born in Finland, in 1971. He received the M.Sc. degree in electrical engineering and control systems and the D.Sc. degree in electronics from the Lappeenranta University of Technology (LUT), Lappeenranta, Finland, in 1997 and 2004, respectively.

He became an Associate Professor of applied electronics in 2002. Since 1995, he has been working in the fields of electronics, sensors, acoustics, IoT, EMC, and engineering education. His current research interests include measurement electronics and sensors in various applications.

Heikki Peltonen received the Ph.D. degree in biomechanics from the University of Jyväskylä, Finland, with a focus on isometric force-time parameters in monitoring of strength training—with special reference to acute responses to different loading resistances.

He is currently working as a Post-Doctoral Researcher with the Faculty of Sport and Health Sciences, University of Jyväskylä. He is a member of the Advisory Board of the NeuroMuscular Research Center. His research interests include health monitoring, acute neuromuscular responses, and long-term adaptations to different types of exercise training. He has several years of experience in the development process of performance monitors and healthcare solutions.

Mika Silvennoinen received the Ph.D. degree in exercise physiology in 2016.

He worked as a Ph.D. Researcher with the Faculty of Sport and Health Sciences, University of Jyväskylä. He has over 15 years of experience in the field of exercise physiology and has published 28 peer-reviewed research reports in high-level international journals. He has a broad range of skills and understanding of laboratory methods in exercise physiology, biochemistry, and molecular biology.

Heikki Kyröläinen received the Ph.D. degree with the University of Jyväskylä, Jyväskylä, Finland.

He is a Professor with the Faculty of Sport and Health Sciences, University of Jyväskylä. He was found competent for professorships in exercise physiology in 2003, biology of physical activity in 2005, kinesiology in 2006, and biomechanics in 2007. He has authored or coauthored over 180 peer-reviewed international scientific articles and about 300 chapters in books, abstracts, proceedings, and domestic publications. His research interests are wide in the field of biology of physical activity but, however, a major research line has been mechanical efficiency and economy during human locomotion since 1987.

Pertti Silventoinen (Member, IEEE) was born in Simpele, Finland, in 1965. He received the D.Sc. degree from the Lappeenranta University of Technology (LUT), Lappeenranta, Finland, in 2001.

He became a Professor of applied electronics in 2004. His current research interests include power electronics systems in various applications.