HAL open science

Energy Expenditure Estimation From Respiratory Magnetometer Plethysmography: A Comparison Study

Aya Houssein, Jacques Prioux, Steven Gastinger, Brice Martin, Fenfen Zhou, Di Ge

▶ **To cite this version:**

Aya Houssein, Jacques Prioux, Steven Gastinger, Brice Martin, Fenfen Zhou, et al.. Energy Expenditure Estimation From Respiratory Magnetometer Plethysmography: A Comparison Study. IEEE Journal of Biomedical and Health Informatics, 2023, 27 (5), pp.2345-2352. 10.1109/JBHI.2023.3252173 . hal-04113965

HAL Id: hal-04113965

https://hal.science/hal-04113965

Submitted on 12 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3252173

Energy expenditure estimation from respiratory magnetometer plethysmography: A comparison study

A. Houssein, J.Prioux, S.Gastinger, B. Martin, F.Zhou, and D.Ge

***Abstract*— Physical activity (PA) quantification by estimating energy expenditure (EE) is essential to health. Reference methods for EE estimation often involve expensive and cumbersome systems to wear. To address these problems, light-weighted and cost-effective portable devices are developed. Respiratory magnetometer plethysmography (RMP) is among such devices, based on the measurements of thoraco-abdominal distances. The aim of this study was to conduct a comparative study on EE estimation with low to high PA intensity with portable devices including the RMP. Fifteen healthy subjects aged 23.84 ± 4.36 years were equipped with an accelerometer, a heart rate (HR) monitor, a RMP device and a gas exchange system, while performing 9 sedentary and physical activities: sitting, standing, lying, walking at 4 and 6 km/h, running at 9 and 12 km/h, biking at 90 and 110 W. An artificial neural network (ANN) as well as a support vector regression algorithm were developed using features derived from each sensor separately and jointly. We compared also three validation approaches for the ANN model: leave one out subject, 10 fold cross-validation, and subject-specific. Results showed that 1. for portable devices the RMP provided better EE estimation compared to accelerometer and HR monitor alone; 2. combining the RMP and HR data further improved the EE estimation performances; and 3. the RMP device was also reliable in EE estimation for various PA intensities.**

***Index Terms*— Energy expenditure, machine learning, physical activity, respiratory magnetometer plethysmography, accelerometer, Heart rate monitor.**

I. Introduction

REGULAR physical activity (PA) is essential to maintain and improve health conditions. It plays an important role in the prevention and treatment of cardiovascular diseases, diabetes, obesity, and several types of cancer [1], [2]. Therefore, the quantification of PA is of major interest in epidemiological and physiological research to investigate the relationships and effects of PA to health conditions. Quantification of PA consists of identifying the type of PA and estimating the related energy expenditure (EE).

Several tools exist to measure or estimate EE. Reference methods such as direct calorimetry, indirect calorimetry and doubly-labeled water provide the most accurate measurements of EE. However, these methods are limited due to the specific equipment and associated human expertise, not to mention the high costs [3]. To overcome these problems, researchers have attempted to develop portable devices that are light-weighted and less expensive. Motion sensors (accelerometers) are frequently used in the literature to quantify both PA and EE [4], [5]. However, the use of a single accelerometer appears to be insufficient for activities involving both upper and lower body movements [6], leading to the idea of using multiple accelerometers placed at different body positions [4], at the cost of increasing both the system complexity and the physical constraints in wearing them. Another problem related to accelerometers is that the workload of activities remains difficult to quantify. Physiological variables, such as heart rate (HR), are also frequently used for EE estimation [7]. Unfortunately, HR is also influenced by factors unrelated to EE such as stress, emotion, temperature, etc. [8], [9] and its use requires individual calibration.

Durnin et al [10] and Ford et al [11] have shown that minute ventilation ($\dot{V}_E$) can be effectively used as an index of EE estimation. Recent advances in technologies to estimate $\dot{V}_E$ from wearable sensors have encouraged researchers to use this physiological variable to estimate EE [12]. Among these technologies, the respiratory magnetometer plethysmography (RMP) system has been validated as a reliable method [12]. The study of Gastinger et al. [13] was the first to use the RMP system in estimating EE. They used sample-wise linear regression models to connect the respiration traces from the RMP to the measured EE values, and individual calibration was required for each subject and for each activity, a tedious and impractical process. In addition, the authors evaluated the performances of the RMP system over a limited range of activities with low to moderate intensity, including only sedentary and walking activities. To our knowledge, no study in the literature ever compared the RMP with other devices for the EE estimation performances. The main objective of our work was to develop a new approach to EE estimation using non-linear machine learning models with features extracted from the RMP signals directly related to respiration and EE, and to test its performances with the most frequent PA

A. Houssein, Ecole normale supérieure de Rennes, France (email: aya.houssein@hotmail.com)
J.Prioux, Ecole normale supérieure de Rennes, France (email: jacques.prioux@ens-rennes.fr)
S.Gastinger, laboratoire Mouvement Sport Santé, France (email: gastingersteven@yahoo.fr)
B.Martin, laboratoire Mouvement Sport Santé, France (email: brice.mar34@"gmail.com)
F.Zhou, Ecole normale supérieure de Rennes, France (email: fenfen.zhou@ens-rennes.fr)
D. Ge, Laboratoire traitement du signal et de l'image (LTSI), France (email: di.ge@univ-rennes1.fr)

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3252173

activities of various intensities including sedentary, walking, running and biking. We also compared the performances of the RMP to other widely-used sensors such as HR monitor and accelerometer, and evaluated the effects of model choice and validation approaches.

II. Materials and Methods

A. Subjects

This study was approved by the committee for the protection of persons of ile de France 1 (Ref: CPPIDF1-2019-ND25 cat2). Fifteen healthy subjects (10 men and 5 women) aged 23.84 ± 4.36 years voluntarily participated in this study (height $175.42 \pm 6.77\ cm$, weight $66.63 \pm 9.83\ kg$, body mass index $21.56 \pm 2.19\ kg/m^2$). All subjects were healthy and did not present any pulmonary or heart diseases. After a detailed description of the protocol, subjects provided their informed consent.

B. Experimental protocol

Subjects participated in two sessions of activities conducted on two different days. The order of these sessions was randomized. The first session consists of both PA and sedentary activities. In the second session, a maximal graded test was performed. However, we only used the dataset of the first session for the current study. PA were selected to cover the most frequently performed PA in daily life with a wide range of intensities. During the experiments, subjects were equipped with an HR monitor, an accelerometer, a RMP system and a gaz exchange measurement system for reference EE measurements. During the first session, each subject performed three successive sedentary activities: sitting, standing, lying, with a duration of 6 min separated by a 1 min rest. Then, each subject performed three PA with two different intensities: walking at 4 and 6 km/h, running at 9 and 12 km/h, biking at 90 and 110 W. The duration of each PA was 6 min, with a 3 min rest in between. The rest period was chosen to ensure sufficient recovery and to allow the EE level to return to its value at rest. Both walking and running were performed on a treadmill, while biking on a cycle ergometer. The order of activities was chosen with increasing intensity. All these activities were performed in laboratory under standardized conditions.

C. Materials

During experiments, subjects were equipped with a RMP device, a HR monitor (Polar), a triaxial accelerometer (ActiGraph) and a gas exchange system (Ultima CardioO2) to measure respiratory frequency (RF) and EE in each cycle, and to visualize them in real time. Table I summarizes the main characteristics of the sensors used in our study.

TABLE I: Main characteristics of sensors.

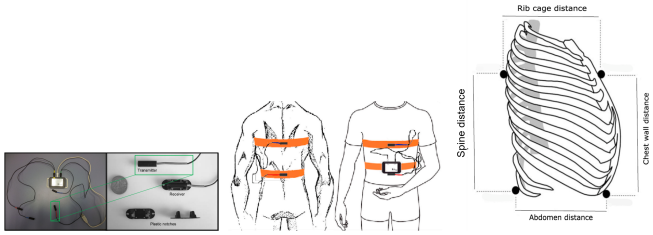
Sensor	Position	Measurements	Sampling frequency	Software
RMP system	Anterior and posterior positions of the sternum and nombril	Antero-posterior distances of the rib cage and abdomen, longitudinal distances of the rib cage and spine	15 Hz	Apios
Accelerometer (ActiGraph wGT3x+)	Hip	Acceleration	30 Hz	ActiLife
Heart rate monitor (Polar RS800CX)	Chest	Heart rate	1 Hz	Polar ProTrainer
Gas exchange system (Ultima CardioO2)	Facial mask	Respiratory frequency Energy expenditure	cycle by cycle	Cardiorespiratory diagnostic

1) Respiratory gas exchange system: Indirect calorimetry is the reference method for measurement of EE. Gas exchange measurements were recorded using an indirect calorimetry, the Ultima CardiO2 (Medical Graphics Corp, St. Paul, MN, USA). It requires the use of a face mask equipped with a Directconnect™ flow sensor (Medical Graphics Corp, St. Paul, MN, USA) [22]. The Ultima CardioO2 was calibrated according to manufacturers' instruction before each test [22]. Data were acquired on a breath-by-breath basis, provinding continous analysis and display of respiratory vaiables such as tidal volume (V_T), RF, $\dot{V}_E$, and EE. The recorded data were saved in the internal database of the Ultima CardioO2 and then transfered to a PC for analysis.

2) RMP system: The RMP system was that used by Dumond et al. [14] (Nomics s.a, Liège Science Park, Belgium). It is a lightweight (80 g) and small ($85\ mm \times 55\ mm \times 16\ mm$) device composed of two pairs of electromagnetic coils connected to a recorder box. Each pair consists of a transmitter and a receiver. The transmitter generates a magnetic field through a circuit with a resonant frequency. The receiver detects the resonant frequency of the magnetic field from the transmitter and converts it into a signal. The intensity of the magnetic field collected provides a measure of the distance between the transmitter and the receiver. The RMP system consists of two transmitters and two receivers. The first transmitter is placed in anterior position at the level of the sternum. The first receiver is placed in a posterior position on the spine, opposite to the first transmitter. The second receiver is placed in the anterior position on the midline of the abdomen just above the umbilicus. The second transmitter is placed in a posterior position on the spine opposite to the second receiver (Figure 1). RMP measures four distances: the anteroposterior displacements of the rib cage and abdomen as well as the axial displacements of the chest wall and the spine. The data were recorded with a sampling rate of 15 Hz. The device records signals and stores them in an internal memory. The memory capacity is 16 MB, equivalent to 100 h of recording. The RMP system is powered by a lithium battery that can be rechargeable via USB. When the battery is fully charged, the autonomy of this system is sufficient to ensure several nights of recording (minimum 60 h in recording mode). The APIOS software is used to transfer a recording from the device to the PC.

3) Accelerometer: The ActiGraph accelerometer is among the most widely used activity monitors for PA quantification [15]. It is a triaxial accelerometer (ActiGraph wGT3X+) with

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3252173



(a) (b) (c)

Fig. 1: a) RMP system. b) Position of RMP system. c) Distances measured by RMP system.

a dimension of 46 mm×33 mm×15 mm and a weight about 19 g. It measures acceleration along three axes: vertical, anteroposterior and mediolateral. The data are recorded with a sampling frequency ranging from 30 to 100 Hz. Before the start of experiments, the accelerometer was initialized using ActiLife software. The sampling frequency can be selected. In our study, we selected a frequency of 30 Hz. The data were stored in a non-volatile memory and then downloaded via a USB cable to a PC using the ActiLife software. The ActiGraph uses a rechargeable lithium battery, with an autonomy depending on the selected sampling frequency, between 11.5 to 22.5 days. The memory capacity is between 12.5 and 42.5 days. The accelerometer was placed at the hip in a pocket with Velcro closure.

4) Heart Rate Monitor: HR was recorded continuously using the Polar RS800CX HR monitor, composed of a Polar Wear-Link W.I.N.D transmitter and a smartwatch. The transmitter consists of an attached transmitter and a chest belt placed around the chest. It collects and transmits HR data to the smartwatch. The smartwatch records HR data with a sampling rate of 1 Hz. It can measure HR for approximately 31 h, either in beats per minute, or in % of maximum or reserve HR. The ProTrainer software allows to visualize and transfer the data to a PC via USB IrDA.

D. Data processing

Signals collected from different sensors were preprocessed to eliminate artifacts unrelated to respiration activities. Figure 2 illustrates the different steps used to process these data.

1) Data processing: Signals from the RMP system are often affected by several types of noises caused by movement and artefacts. Indeed, the belts of RMP system typically moved during intense PA on treadmill, inducing artefacts. Therefore, RMP signals were first low-pass filtered with the cut-off frequency (fc) chosen according to the PA intensity estimated by the measured HR:
fc= 0.4 Hz if HR $<$ 90 bpm
fc= 0.8 Hz if 90 $<$ HR$<$ 170 bpm
fc= 1.1 Hz if HR$>$ 170 bpm
Similarly for the accelerometer, signals were filtered by a 4-th order Butterworth bandpass filter with a pass-band of $0.25 - 2.5$ Hz to reject noise and artefacts outside of the PA-related frequency band.

2) Features extraction: To increase the efficiency of EE estimation model during training, filtered data from different sensors were converted into an appropriate space representation (the features). For this purpose, we divided data using a sliding window of 30 s without overlapping. In total, we obtained 154 temporal segments for sedentary and biking activities, and 165 segments for walking and running activities. Features of RMP, accelerometer and Polar HR signals were then computed from each segment.

RMP features: Two parameters directly related to $\dot{V}_E$ were calculated from RMP signals. The first was the interdecile range of the RMP signals to reflect the variations in thoracoabdominal distances related to tidal volumes. This range is the difference between the 90th and 10th percentiles of the signal in each temporal segment. The second parameter was the RF estimated from each temporal segment. The product of the V_T and RF corresponds to $\dot{V}_E$, a proven index for EE estimation.

To estimate RF from the RMP system, we have adopted a temporal approach. Indeed, the sampling frequency of the RMP signals was not high enough to obtain an accurate estimation of RF with a frequency approach (e.g. Fourier transform, periodogram). The proposed temporal approach first detected the respiratory cycles, by applying a peak (local maxima and minima) detection algorithm with an adaptive thresholding (Figure 3). Then, a quality index was calculated for each RMP signal to identify the one with the highest quality in terms of regularity or periodicity of the respiration cycles. The quality index was defined as the sum of the variation coefficients of peaks. This variation coefficient was simply the quotient of the standard deviation over the mean of cycle durations (consecutive min or max differences). Thus, a lower index indicates a better periodicity and therefore more regular respiratory cycles.

$$CV_{max} = \frac{std(T)}{\bar{T}}$$

$$\bar{T} = \frac{\sum_{i=1}^{N-1} T_i}{N}$$

Where CV_{max} is the variation coefficient of the maximum peaks, T the vector of the cycle durations (between two consecutive maximum peaks), $\bar{T}$ the average cycle durations, std the standard deviation and N the number of maximum peaks.

The RF estimate was simply the inverse of the averaged cycle durations from the RMP segment with the best quality index.

Accelerometer features: We computed the activity count for each axis from the filtered accelerometer data. This count is the sum of the acceleration values of the considered axis in a temporal segment. We calculated also the count of the magnitude vector which is the square root of the sum of the squares of the acceleration counts of the three axes. These features give an indication of the intensity level of PA.

HR features: The sampling rate was low which did not allow the calculation of complex features. Therefore, we only computed the average HR for each temporal segment.

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3252173

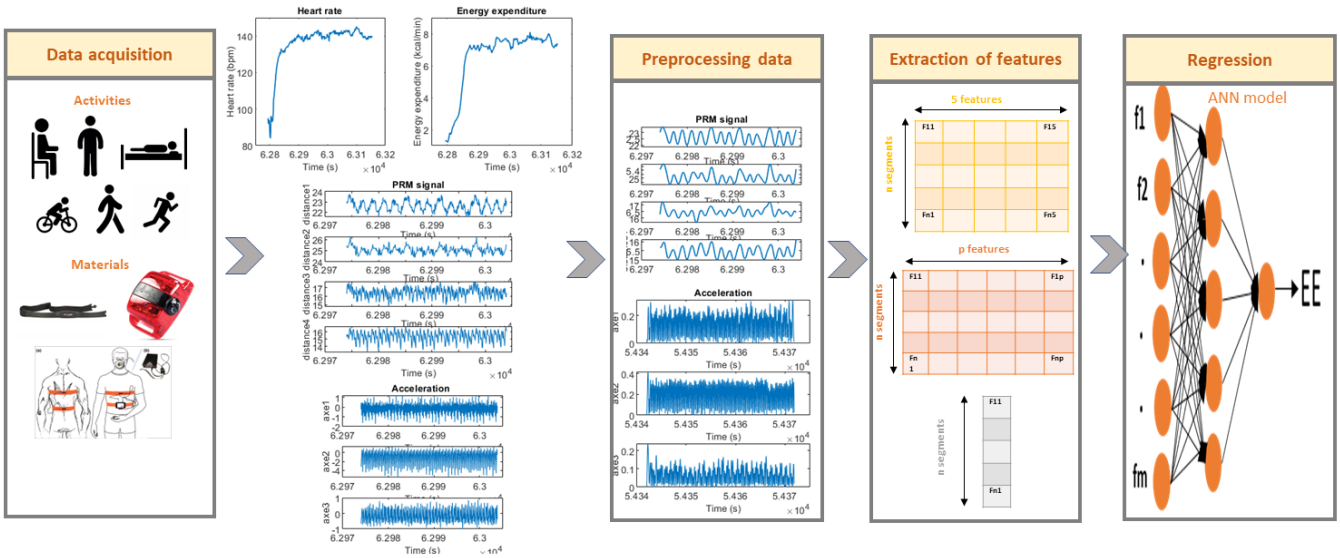


Fig. 2: Differents steps for data processing.
Data were collected from subjects wearing the accelerometer, HR monitor, RMP system and a facial mask during different activities. Signals were extracted, preprocessed for noise reduction, and divided into segments. Then, features were extracted and used as inputs of regression models for EE estimation.

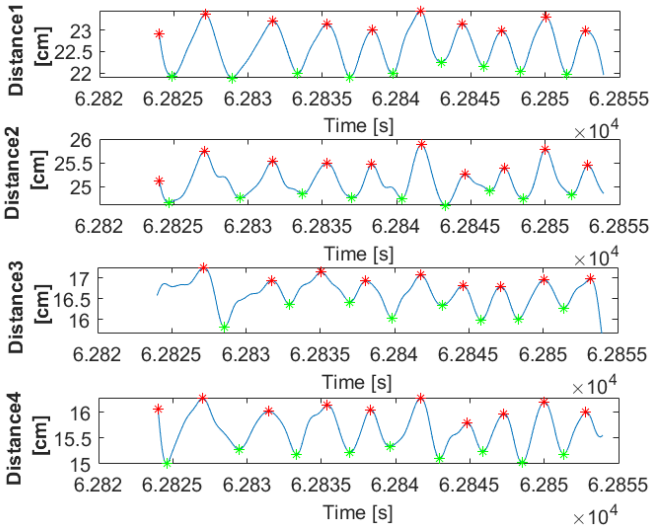


Fig. 3: Example of detection of respiratory cycles from RMP signals.

3) Regression Models: A nonlinear regression model, artificial neural network (ANN), was developed in the MATLAB environment using the Deep learning Toolbox for each sensor separately and then by combining both RMP and HR system, to establish the relationships between EE and the corresponding features. The ANN model implemented in our study was a multilayer perceptron with three layers, an input layer, an output layer and a hidden layer. The input and output neurons were connected by weighted connections. The number of neurons in the input layer was equal to the number of features. The output layer was composed of a single neuron that corresponds to the predicted EE. The hidden layer was composed of 5 neurons. Various architectures with different number of neurons in hidden layer was applied to select the optimal network (Figure 4). A hyperbolic tangent sigmoid transfer function "transig" was used for the hidden layers and a linear transfer function "purelin" for the output layers. The training was done using Levenberg Marquardt optimization function "trainlm" that consists in adjusting the weights and bias of the neural network in order to minimize the mean square error between the value of estimated EE and the measured EE. For all models, a leave-one-out subject validation (LOOS) was used. With this validation, a model was trained with data from all subjects except one that was used to test the model.

To show the effect of the regression model on the estimation performance using RMP system, we compared the ANN regression model to the support vector regression (SVR) model. The SVR model was a nonlinear model in which the data was projected into a larger representation space using a kernel function to find a linear function that estimates EE with minimum error. The kernel function used in our study was a polynomial function of type 2. It was developed and trained by "fitrsvm" MATLAB function using Statistics and Machine learning Toolbox.

In order to evaluate the importance of the individualization of the models for EE estimation using RMP system, we compared LOOS to two other validation approaches: 1) a subject specific approach (SS) in which each ANN model was trained using only data from one specific subject, and the test was then performed by 5 fold cross validation for the

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3252173

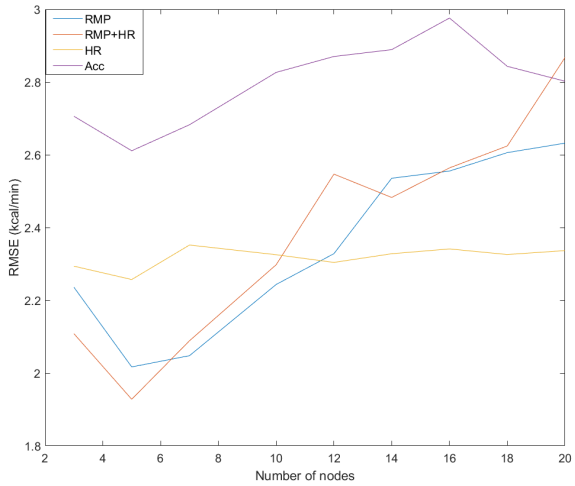


Fig. 4: Root mean square error (RMSE) as a function of the number of nodes in hidden layer.
RMP: respiratory magnetometer plethysmography. HR: heart rate monitor. Acc: Accelerometer.

subject, and 2) a 10 fold cross validation (CV) approach in which the dataset including all subjects was randomly divided into 10 subsets of equal sizes. At each iteration the model was trained with 9 subsets and the remaining was used to test the model.

4) Statistical analysis: The evaluation of the performance of EE estimation was performed by comparing the estimated EE to its reference value from the gas exchange system. Graphical representations of estimated EE versus reference values were plotted with their associated regression line. The most common parameters used for model validation were calculated:
- The root mean square error (RMSE) defined by:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(EE_i - EE_{magi})^2}$$

-The coefficient of determination (R^2) defined by:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(EE_i - EE_{magi})^2}{\sum_{i=1}^{N}(EE_i - \overline{EE})^2}$$

, where N is the number of temporal segments, EE_{magi} the estimated EE, EE_i the reference EE of the i-th segment, and $\overline{EE}$ the average reference EE.
To quantify the significances of differences between two or more regression models, we used a multiple comparison test ANOVA, and a $p < 0.05$ was considered to be significant.

III. Results

A. Performance of RF estimation

We evaluated the performance of RF by first comparing the estimated RF to the reference RF (Table II). Our results showed that they were not significantly different for all sedentary and PA. Our results also showed that error increases with the intensity. Indeed, the noise generated by the movement increases with the intensity, leading to an increase in the estimation error.

TABLE II: Differences between measured and estimated respiratory frequency (RF) for all PA.

Activities	RF reference (cycles/min)	RF estimated (cycles/min)	Estimation error of RF (cycles/min)
Sitting	13.59 ± 4.77	13.57 ± 4.74	0.74 ± 0.65
Standing	13.56 ± 5.36	13.55 ± 5.22	0.81 ± 0.88
Lying	12.96 ± 4.44	12.95 ± 4.43	0.83 ± 0.80
Walking 4 km/h	17.89 ± 4.79	17.69 ± 4.75	1.30 ± 1.14
Walking 6 km/h	19.66 ± 5.71	19.63 ± 5.59	1.26 ± 1.51
Running 9 km/h	27.57 ± 7.22	27.84 ± 7.00	1.60 ± 1.71
Running 12 km/h	32.39 ± 8.51	32.72 ± 7.99	1.69 ± 1.77
Biking 90 W	24.81 ± 6.00	24.98 ± 5.92	1.23 ± 1.36
Biking 110 W	26.21 ± 7.54	26.30 ± 7.40	1.34 ± 1.49

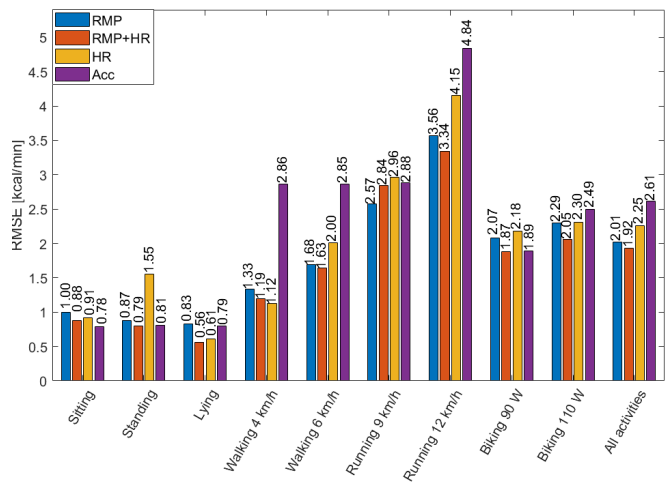


Fig. 5: RMSE of EE estimation for different sensors.

B. Comparison of sensors

Figure 5 illustrated the RMSE of the estimated EE for each activity using the different sensors (RMP, accelerometer, HR monitor) and their combinations (RMP+HR). Our results showed that RMP provides better overall performance than HR monitor and accelerometer. The overall RMSE was lower with RMP (2.01 kcal/min) compared to HR monitor (2.25 kcal/min) and accelerometer (2.61 kcal/min). By combining RMP and HR, the performance of EE estimation was further improved. For all activities, RMSE with RMP and HR ($0.56 - 3.34$ kcal/min) were lower than those with RMP alone ($0.83 - 3.56$ kcal/min) or those with HR alone ($0.61 - 4.15$ kcal/min). The same results were observed for accelerometer alone ($0.78 - 4.84$ kcal/min). In addition, the estimation error increased with activity intensity. RMSE were the lowest for lying (0.56 kcal/min) and the highest for running 12 km/h activity (3.34

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3252173

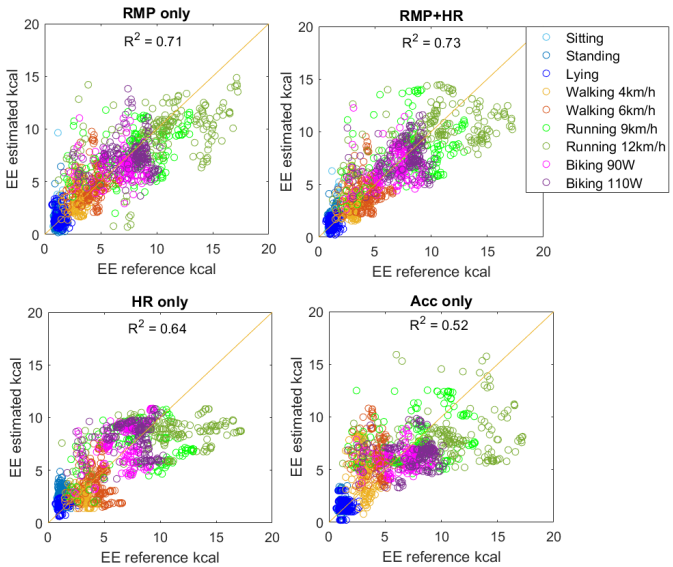


Fig. 6: Estimated EE as a function of reference EE using different sensors.

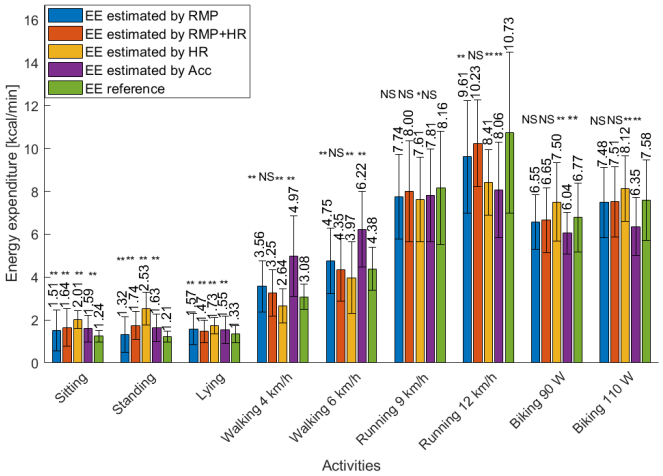


Fig. 7: Mean and standard deviation of reference and estimated EE for all sensors.
NS: non significant different p>0.05 , *: significant different p<0.05, **: high significant different p<0.01

kcal/min). To evaluate the agreement between the measured and estimated EE with either the RMP, HR, accelerometer alone, or the combination of RMP+HR, we have plotted the estimated EE against those measured and then calculated the coefficient of determination (Figure 6). Our results showed that a higher coefficient of determination was achieved by combining the RMP and HR system ($R^2 = 0.73$) compared to using only the RMP system ($R^2 = 0.71$), HR monitor ($R^2 = 0.64$), or accelerometer ($R^2 = 0.52$). The mean and standard deviation of the reference and estimated EE from all sensors are illustrated in Figure 7. The estimated EE using a combination of RMP and HR system were not significantly different from the reference for all PA, whereas using the RMP system alone, they became significantly different for most activities except for biking and running at 9 km/h. For sedentary activities and for all sensors, the estimated EE were significantly different from the reference.

C. Comparison of regression models

To study the model effects on EE estimation, we implemented another frequently used nonlinear models: SVR using the same RMP features and reference values for both training and testing. Our results showed that RMSE of the SVR model

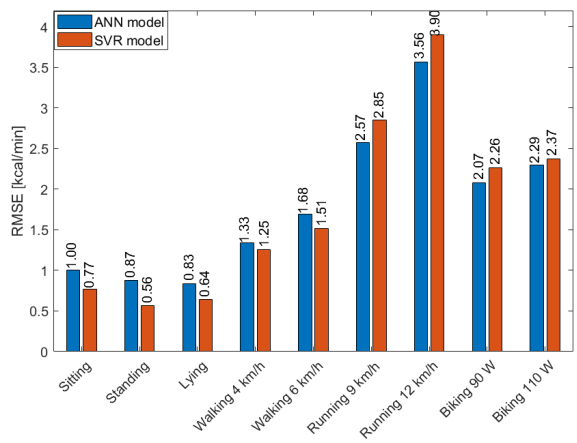


Fig. 8: RMSE for ANN and SVR model.

were lower than those of the ANN for sedentary and walking activities, while the opposite was observed for running and cycling activities (Figure 8). One machine learning model can better fit for certain types of activities, but not for all of them. Though the underlying mecanism is not explored in this study, this may pave the way for more complex hybrid regression methods coupled with an activity type recognition/classification module. Regarding the coefficient of determination, Figure 9 showed that the ANN model ($R^2 = 0.71$) was better than the SVR model ($R^2 = 0.69$). However, the difference was not significant.

D. Comparison of validation approaches

We have compared 3 validation approaches: LOOS, SS and CV. Figure 10 showed that RMSE decreased significantly using SS ($0.49-1.94$ kcal/min) as compared to LOOS ($0.83-3.56$ kcal/min). This result was observed for different physical and sedentary activities. Furthermore, RMSE obtained with CV ($0.47-2.77$ kcal/min) were lower than those with LOOS ($0.83-3.56$ kcal/min). We can conclude that estimation

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3252173

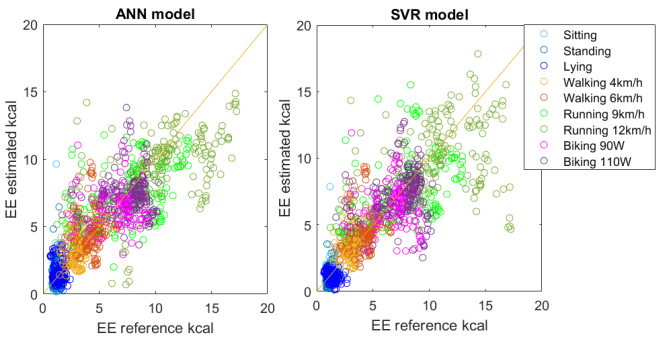


Fig. 9: EE estimated as a function of EE reference for ANN and SVR models.

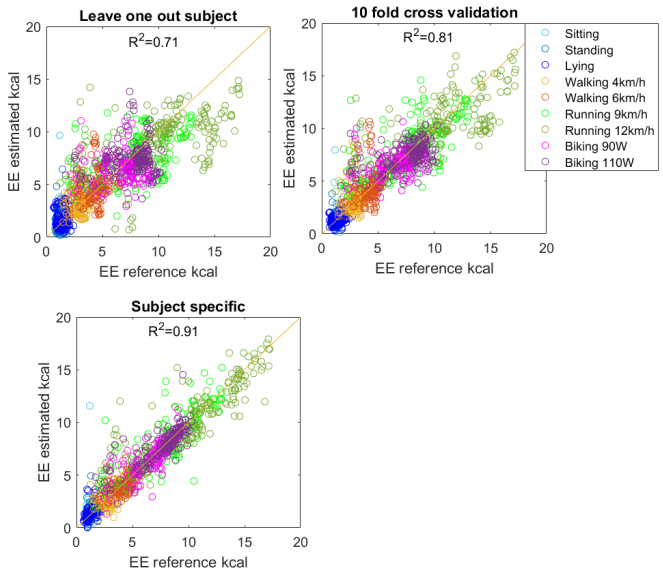


Fig. 11: Estimated EE by RMP as a function of reference EE for different validation approaches.

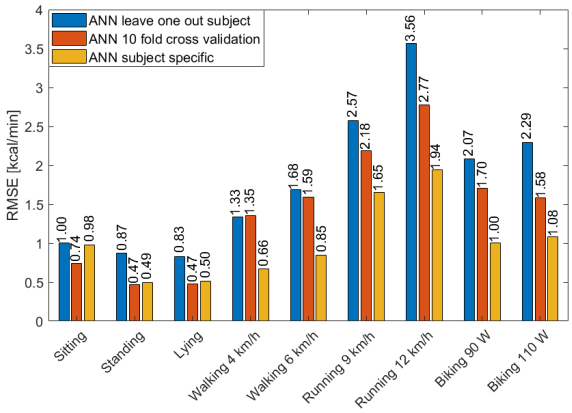


Fig. 10: RMSE for different validation approches of ANN model.

errors were mainly due to the inter-subject variations. Figure 11 showed estimated EE using the RMP as a function of measured EE for each activity and for the three approaches. The coefficient of determination of SS approach ($R^2 = 0.91$) was significantly higher than that of LOOS ($R^2 = 0.71$) and CV approach ($R^2 = 0.81$).

IV. Discussion

The main objective of our study was 1) to evaluate a new approach to EE estimation using RMP, 2) to compare its performances with the most frequently used sensors in the literature such as HR monitor and accelerometer, and 3) to study the effects of model choices. Our dataset included both sedentary activities and PA from low to high intensities of daily life. Better results of EE estimation are obtained with the RMP system with a lower RMSE and a higher coefficient of determination than those with the accelerometer and HR monitor.

Indeed, few studies in literature have attempted to evaluate EE based on respiratory variables [13], [16], [17]. Liu et al. [16] used a portable multi-sensor device to measure acceleration and $\dot{V}_E$ in 50 subjects performing 13 types of activities at different intensities. Their results showed that combining these two sensors allowed to estimate EE with a RMSE of 0.42 METs, an error 22.2% smaller than using a single accelerometer placed at the hip. More recently, Lu et al. [17] used a combination of HR, $\dot{V}_E$, and acceleration derived from a single wearable sensor. Their results showed better performances when combining these three measurements ($R^2 = 0.92$) than those using the HR alone ($R^2 = 0.75$) or combined with the acceleration ($R^2 = 0.86$). Gastinger et al. [13] validated the RMP system but only for postural and walking activities. In addition, these authors used individual calibration, based on a linear equation developed for each subject and for each activity. In our study, the RMP system was evaluated for a much higher number of activities with different intensity levels, including walking, running and cycling. We also used a novel algorithm by applying an artificial neural network to cycle-based respiration features strongly related to EE . Finally, we used an independent subject validation to avoid the complex calibration process.

Our results showed that combining the RMP and HR monitor improved the EE estimation for all PA studied: the estimated EE were not significantly different from their reference values. The HR provided additional information about the PA

This article has been accepted for publication in IEEE Journal of Biomedical and Health Informatics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2023.3252173

intensity directly related to EE variations, but its contribution was only observed for non-sedentary activities. This could be explained by the indistinguishably low variations in both thoraco-abdominal distances and HR values for sedentary activities. Other studies concurred that HR was not a good indicator of EE for sedentary to low-intensity PA [18]. Indeed, these authors showed that HR did not provide any additional information to their model for low-intensity activities, since they are usually performed with normal HR. In order to evaluate the effect of model choice on EE estimation, we compared the estimation performances with two frequently used machine learning algorithms, i.e., ANN and SVR. Our results showed a comparable performances for both models. Kate et al. [21] found similar results in comparing different machine learning algorithms that the EE estimation performances are not strongly affected by model choices.

In our study, we also analyzed the effects of the validation methods. This has been done by comparing three different validation approaches, namely the LOOS, the SS and the CV approach. Our results showed a significant performance increase with the SS ($R^2 = 0.91$) and CV ($R^2 = 0.81$) when compared to the LOOS ($R^2 = 0.71$). The development and validation of the model using data from the same subject can lead to adjustment of the machine learning models to the individual characteristics. On the other hand, when models are validated on an independent subject (as in the case of LOOS), the unavailability of subject-specific data and the inter-subject variations are the main factors affecting the estimation performances. Indeed, two individuals performing the same activity may yield significantly different EE values, as physical conditions and metabolism vary from one person to another [19], [20]. Nevertheless, LOOS has a particular importance in the interoperability assessment of models among a new population.

Our proposed method has several insufficiencies to be noted. Firstly, the performances for sedentary activities could be improvement by assigning a static EE value for these activities using the compendium of PA, or by applying other estimation approaches such as identifying the PA intensity levels and developing a regression model for each level. In either way, it suggests a recognition of the sedentary activity type with a classification tool. Secondly, the dataset included only 15 healthy and young subjects due to constraints in the recruitment. The total volume, however, in terms of samples is sufficient statistically (1430 samples). Thirdly, the study was conducted under controlled laboratory conditions. A larger dataset containing more subjects to cover a larger population characteristics in the future is essential for the model development in order to obtain more accurate results for EE estimation. Finally, it would be of interest to also evaluate the RMP system on patients with respiratory and cardiovascular diseases and to test more complex deep learning models.

References

[1] R.E. Rhodes, I. Janssen, S.D. Bredin, D.E. Warburton, and A. Bauman, "Physical activity: Health impact, prevalence, correlates and interventions," in *Psychol Health.*, vol. 32, no. 8, pp. 942–975, 2017.

[2] D.R. Warburton, and S.D. Bredin, "Health benefits of physical activity: A systematic review of current systematic reviews," in *Curr Opin Cardiol.*, vol. 32, no. 5, pp. 541–556, 2017.

[3] A.P. Hills, N. Mokhtar, and N.M. Byrne, "Assessment of Physical Activity and Energy Expenditure: An Overview of Objective Measures," in *Front Nutr.*, vol. 16, pp. 1–5, 2014.

[4] M. Altini, J. Penders, R. Vullers, and O. Amft, "Estimating energy expenditure using body-worn accelerometers: A comparison of methods, sensors number and positioning," in *IEEE J Biomed Health Inform.*, vol. 19, no. 1, pp. 219–226, 2015.

[5] R. O'Driscoll, J. Turicchi, M. Hopkins, G.W. Horgan, G. Finlayson, and J.R. Stubbs, "Improving energy expenditure estimates from wearable devices: A machine learning approach," in *J Sports Sci.*, vol. 38, no. 13, pp. 1496–1505, 2020.

[6] P. Freedson, D. Pober, and K.F. Janz, "Calibration of accelerometer output for children," in *Med Sci Sports Exerc.*, vol. 37, no. 11, pp. S523–30, 2005.

[7] S. Brage, N. Brage, P.W. Franks, U. Ekelund, and N.J. Wareham, "Reliability and validity of the combined heart rate and movement sensor actiheart," in *Eur J Clin Nutr.*, vol. 59, no. 4, pp. 561–570, 2005.

[8] S. Gastinger, G. Nicolas, A. Sorel, H. Sefati, and J. Prioux, "Energy expenditure estimate by heart-rate monitor and a portable electromagnétic coils system," in *Int J Sport Nutr Exerc Metab.*, vol. 22, no. 2, pp. 117–30, 2012.

[9] S. Gastinger, A. Donnelly, R. Dumond, and J. Prioux, "A review of the evidence for the use of ventilation as a surrogate measure of energy expenditure," in *IJPEN J Parenter Enteral Nutr.*, vol. 38, no. 8, pp. 926–38, 2014.

[10] J.V. Durnin and R.G. Edwards, "Pulmonary ventilation as an index of energy expenditure," in *Q J Exp Physiol Cogn Med Sci.*, vol. 40, no. 4, pp. 370–7, 1955.

[11] A.B. Ford, and H.K. Hellerstein, "Estimation of energy expenditure from pulmonary ventilation," in *J Appl Physiol.*, vol. 14, pp. 891–893, 1959.

[12] A. Houssein, D. Ge, S. Gastinger, R. Dumond, and J Prioux, "Estimation of respiratory variables from thoracoabdominal breathing distance: A review of different techniques and calibration methods," in *Physiol Meas.*, vol. 40, no. 3, pp. 03TR01, 2020.

[13] S. Gastinger, H. Sefati, G. Nicolas, A. Sorel, A. Gratas-Delamarche, and J. Prioux, "A new method to estimate energy expenditure from abdominal and rib cage distances," in *Eur J Appl Physiol.*, vol. 111, no. 11, pp. 2823–35, 2011.

[14] R. Dumond, S. Gastinger, H. Abdul Rahman, A. Le faucheur, P. Quinton, H. Kang, and J. Prioux, "Estimation of respiratory volume from thoracoabdominal breathing distances: comparison of two models of machine learning," in *Eur J Appl Physiol.*, vol. 117, no. 8, pp. 1533–1555, 2017.

[15] S.E. Crouter, J.R. Churilla, and D.R. Bassett, "Estimating energy expenditure using accelerometers," in *Eur J Appl Physiol.*, vol. 98, no. 6, pp. 601–12, 2006.

[16] S. Liu, X. Gao, D. John, J.W. Staudenmayer, P.S. Freedson, "Multisensor Data Fusion for Physical Activity Assessment," in *IEEE Trans Biomed Eng.*, vol. 59, no. 3, pp. 687–696, 2012.

[17] K. Lu, L. Yang, F. Seoane, F. Abtahi, M. Forsman, and K. Lindecrantz, "Fusion of Heart Rate, Respiration and Motion Measurements from a Wearable Sensor System to Enhance Energy Expenditure Estimation," in *Sensors.*, vol. 18, no. 9, pp. 3092, 2018.

[18] B. Cvetkovic, R. Milic, and M. Lustrek, "Estimating Energy Expenditure With Multiple Models Using Different Wearable Sensors," in *IEEE J Biomed Health Inform.*, vol. 20, no. 4, pp. 1081–7, 2015.

[19] J.A. Álvarez-García, B. Cvetković, and M. Luštrek, "A Survey on Energy Expenditure Estimation Using Wearable Devices," in *ACM Computing Surveys.*, vol. 53, no. 5, 2020.

[20] M.N. Ahmadi, A. Chowdhury, T. Pavey, and S.G. Trost, "Laboratory-based and free-living algorithms for energy expenditure estimation in preschool children: A free-living evaluation," in *PLoS One.*, vol. 15, no. 5, pp. e0233229, 2020.

[21] R.J. Kate, A.M. Swartz, W.A. Welch, and S.J. Strath, "Comparative evaluation of features and techniques for identifying activity type and estimating energy cost from accelerometer data," in *Physiol Meas.*, vol. 37, no. 3, pp. 360–379, 2016.

[22] Alcantara JMA, Sanchez-Delgado G, Amaro-Gahete FJ, Galgani JE, Ruiz JR, "Impact of the Method Used to Select Gas Exchange Data for Estimating the Resting Metabolic Rate, as Supplied by Breath-by-Breath Metabolic Carts," in *Nutrients.*, vol. 12, no. 2, 487, 2020.