Body Weight-Normalized Energy Expenditure Estimation Using Combined Activity and Allometric Scaling Clustering

Marco Altini[1], Julien Penders[2] and Oliver Amft[3]

Abstract— **Wearable sensors have great potential for accurate estimation of Energy Expenditure (EE) in daily life. Advances in wearable technology (miniaturization, lower costs), and machine learning techniques as well as recently developed self-monitoring movements, such as the Quantified Self, are facilitating mass adoption. However, EE estimations are affected by a person's body weight (BW). BW is a confounding variable preventing meaningful individual and group comparisons. In this paper we present a machine learning approach for BW normalization and activities clustering. In our approach to activity-specific EE modeling, we adopt a genetic algorithm-based clustering scheme, not only based on accelerometer (ACC) features, but also on allometric coefficients derived from 19 subjects performing a wide set of lifestyle and gym activities. We show that our approach supports making comparisons between individuals performing the same activities independently of BW, while maintaining accuracy in the EE estimate.**

I. INTRODUCTION AND RELATED WORK

Accurate estimation of Energy Expenditure (EE) in ambulatory settings is a key element in determining the causal relation between physical activity (PA) and health [1]. New technologies seamlessly integrated in everyone's life, able to monitor behavior objectively and non-invasively, can provide unprecedented insights on these links. Currently, accelerometers (ACC) are the most widespread tools used to objectively gather information about PA. The rationale is that body motion measured close to the body center of mass, is linearly related to EE. However, a single ACC worn close to the center of mass cannot be used to detect specific motion activities of lower and upper body, thus does not provide enough detail on the relation of ACC and EE. Recent work showed that activity type can be detected with wearable ACC, enabling more precise, activity-specific EE estimation [2]. A few activity-specific algorithms have been reported [3-7]. Activity-specific algorithms first recognize activities, and then apply an activity-specific model for EE estimation. Splitting the EE estimation problem into sub-problems showed consistent improvements compared to other methods [3]. However, EE estimations are highly dependent not only on the activity performed, but also on a person's body weight.

In order to understand whether differences in aerobic or anaerobic performance in subjects with varying body size

[1]M. Altini is with Holst Centre/imec The Netherlands, Eindhoven, Netherlands, and the ACTLab, Signal Processing Systems, Eindhoven University of Technology, Eindhoven, The Netherlands (e-mail: marco.altini@imec-nl.nl)

[2]J. Penders is with Holst Centre/imec The Netherlands, Eindhoven, Netherlands, (e-mail: julien.penders@imec-nl.nl)

[3]O. Amft is with the ACTLab, Signal Processing Systems, Eindhoven University of Technology, Eindhoven, The Netherlands, (e-mail: amft@tue.nl)

are due to differences in physiology or body weight (BW), an appropriate scaling technique must be used [8,9,10]. Appropriate scaling of EE estimation using BW will allow users $a)$ to compare an individual against standards in EE assessments; $b)$ to compare study groups; and $c)$ to compare longitudinal results of one individual to study e.g., growth, weight loss, or training. BW scaling is typically done using the ratio standard (i.e. $kcal/kg$ [15]). Previous research showed that the ratio standard over-corrects for body weight, resulting in invalid conclusions on the relations between PA and physiology [9]. Alternatives are the use of linear models to take into account the effect of BW, relying on a linear relation that should not be assumed [3,5,6] or allometric models that develop power function ratios. Activity-specific linear models that include BW to predict EE do not allow users to compare between individuals or longitudinally for one individual, since estimated values depend on BW. Allometric scaling is an appropriate mathematical procedure for clarifying the relation between anthropometric measures and physiological variables [9]. However, up to now allometric scaling has not been incorporated into machine learning based EE estimators. To date, a variety of power functions has been investigated for allometric scaling of EE [8-10], with no consensus on an optimal approach. Exponents depend on the level of exertion and extent to which the activity is weight bearing, thus the allometric coefficients are activity-dependent. As a result, in past research allometric scaling was often inconclusive, and researchers struggled to find a global coefficient to suit all activities [13]. In this paper, we present a new machine learning approach to normalize EE estimations by BW using activity-specific allometric coefficients. Our contribution is threefold:

- We determine normalization coefficients for a wide set of sedentary, lifestyle and sport activities, by means of allometric modeling, which can be used for scaling EE estimations for differences in BW.
- We use genetic algorithm-based clustering to group activities and optimize cluster distance for activities and allometric coefficients.
- We compare EE estimates in both BW-dependent and BW-independent forms, showing that the proposed normalization does not affect performance, while enabling us to compare between individuals.

II. METHODS

A. BW scaling of EE estimates

Fig. 1 shows the inaccuracy of current state of the art methods when dealing with BW in the context of EE
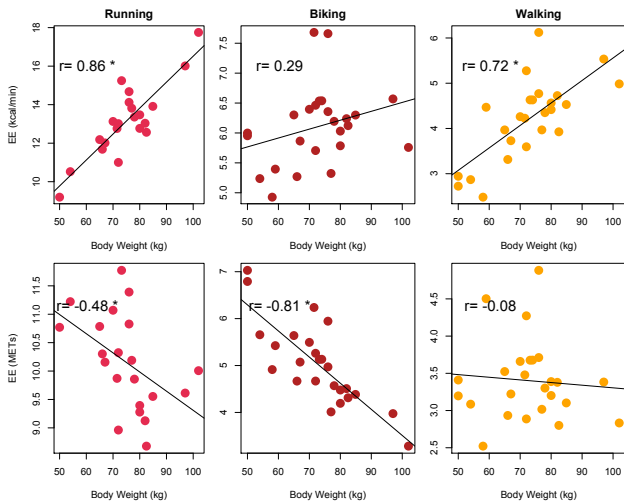
Fig. 1. Relation between BW and EE during running, biking and walking. EE estimates were obtained from an indirect calorimeter device as detailed in Sec. III.B.2. Statistically significant relations are marked with $*$ ($p < 0.05$). Top row: EE in $kcal/min$ and Pearson correlation with BW. Bottom row: EE in METs (using the ratio standard) and Pearson correlation with BW, showing the over-correction effect.

estimation. The top row shows the dependency of EE on BW when no normalization is performed. Correlation (Pearsons coefficient) is significant for weight-bearing activities, such as walking and running. The second row shows the over-correction effect when using the ratio standard (i.e. $METs$ or $kcal/kg$). This is especially evident for non-weight bearing and intense activities such as biking. An alternative way to clarify the relation between anthropometric measures and physiological variables is by means of allometric models. Allometric models are expressed as $y = k * x^{\beta}$, where y is the physiological variable of interest (i.e. EE), x is BW, k a constant, and β is the scaling exponent (if $\beta = -1$, the ratio standard is obtained). The equation can be linearized by applying log-linear regression; $logy = \beta logx + logk$, where the slope β is equal to the allometric exponent. Using this approach, we derived allometric coefficients for the 37 activities that are part of our protocol, and used the coefficients as input features to cluster activities.

B. Automated Activity Clustering

For activity-specific EE estimation, activities are initially clustered. The different activity groups are subsequently used to train an activity classification algorithm. Each activity is mapped to an activity-specific model, used to derive EE [3-7]. Up to now, the problem of clustering activities has often been approached using expert knowledge: researchers manually grouped different activities into clusters. This manual clustering resulted in different activity groupings, e.g. based on motion patterns of the activity performed [3,6], or on EE level of the activity [14]. None of these strategies take into account BW normalization specifically, even though some of them use linear models including BW to predict EE. The latter approach does not allow users to make comparisons (between individuals or longitudinally for one individual),
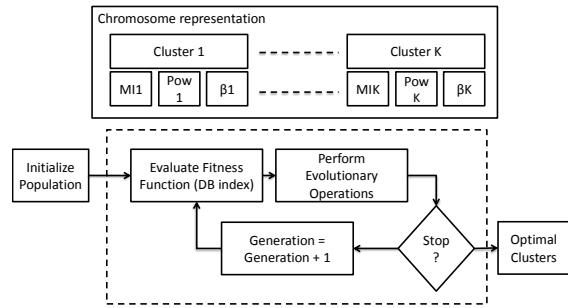

Fig. 2. Block diagram of the chromosome representation and flow diagram of the operations performed by the GA to cluster according to ACC features and BW allometric coefficients.

since BW is part of the linear model. We propose to combine ACC features and allometric coefficients to automatically cluster groups of activities and normalize EE by BW.

C. Combined Activity and BW Clustering

To automate the process of clustering activities, we generated a feature space composed by the following features:

- Allometric coefficients (β).
- Motion Intensity (MI), (sum of the absolute value of the bandpassed (between 1 and 10 Hz) ACC signal).
- Signal Power (Pow) on the vertical axis (sum of the squared FFT components between 1 and 10 Hz).

With this feature space, we searched an optimal solution using Genetic Algorithms (GAs). GAs are adaptive heuristic search algorithms based on the evolutionary ideas of natural selection and genetics. GAs are well suited for multi-objective optimisation problems as the activity clustering for activity classification and BW normalisation. Fig. 2 shows a block diagram of the genetic algorithm. In order to exploit GAs for clustering, we encoded the chromosomes using one gene per cluster center and the Davies-Bouldin (DB) index as fitness function to evaluate each generation. The DB [16] index applies well to the problem of clustering, since it aims to maximize the inter-cluster distance, and at the same time, to minimize the distance between points in a cluster. The optimal number of clusters was five, determined by selecting k that maximized the Calinski-Harabasz index [16].

D. Classification and Energy Expenditure Estimation

Using the activity clusters derived in section II.C, we implemented an activity recognition algorithm, using standard Support Vector Machines (SVMs), 4 seconds non-overlapping windows for segmentation, and the following features: *mean of the absolute bandpassed signal, median, variance, main frequency peak* and *high frequency band signal power*. Following our methodology [3] for activity-specific EE estimation, we developed five multiple-linear-regression models, one for each cluster. Each model is already BW-independent, due to allometric scaling, and does not require anthropometric characteristics to be included. Allometric coefficients were calculated as the averages of the allometric coefficients of each activity belonging to a cluster (see Table 1). Activity-specific ACC features for each

model were selected using linear forward selection, in order to model intra-individual differences in EE. Intra-individual differences are due to the fact that in each cluster there are activities that can be carried out at different intensities, resulting in different EEs.

III. MEASUREMENT SETUP AND DATA COLLECTION

A. Participants Characteristics

Participants were 25 (19 male, 6 female) healthy imec-nl employees (mean age $30.7 \pm 5.6\,years$, mean weight $72.7 \pm 12.7\,kg$, mean height $176.8 \pm 9.4\,cm$, mean BMI $23.1 \pm 2.7\,kg/m^2$). Imec's internal Ethics Committee approved the study, and each participant signed an informed consent form. In this paper, we considered male participants only.

B. Instruments

1) ECG Necklace: The ECG Necklace [11] is a low power wireless ECG platform. The necklace was configured to acquire one lead ECG data at $256Hz$, and ACC data from three axis at $32Hz$. The sensor was placed on the chest with an elastic belt. ECG data were not used for this study.

2) Indirect Calorimeter: Breath-by-breath data were collected using the Cosmed $K4b^2$ indirect calorimeter. The Cosmed $K4b^2$ weights $1.5kg$, battery included, and showed to be a reliable measure of EE [12]. The system was manually calibrated before each experiment according to the manufacturer instructions.

C. Data Collection Protocol

Participants were invited for recordings on two separate days. They reported at the lab at 8.00 a.m., after refraining from drinking (except for water), eating and smoking in the two hours before the experiment. The protocol included a wide range of lifestyle and sport activities. An exhaustive list of the activities can be found in Table 1. Each activity was carried out for about 5 minutes.

TABLE I

ACTIVITIES CLUSTERS DERIVED BY THE GENETIC ALGORITHM

| ID	Number of Activities	β	Activities
1	5	0.99	Cleaning table, washing windows, vacuuming, folding clothes, stacking groceries
2	4	0.70	Running 7, 8, 9 and 10 km/h
3	6	0.05	Biking low, medium and high resistance levels, 60 and 80 rpms
4	10	0.55	Lying, sitting, reading, desk work, writing, PC work, watch TV, standing, cooking, washing dishes
5	12	0.80	Walking 3,4,5,6 km/h, walking carrying weights, walking moving boxes, walking self paced, walking incline (3,5 km/h, 5,10%)
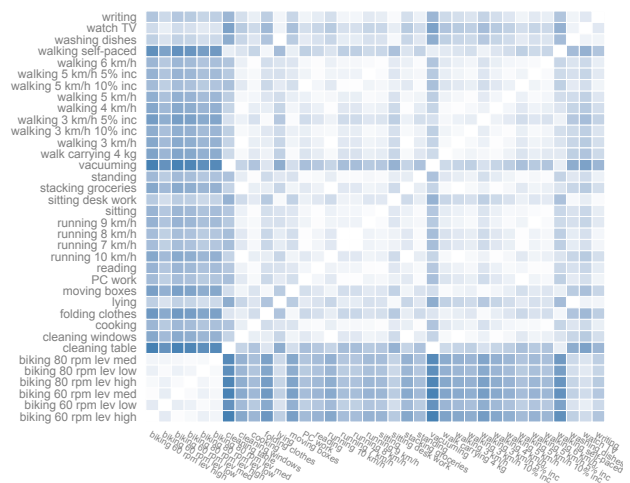
Fig. 3. Heat map of the differences in allometric coefficient (β) between activities. Light colored cells indicate that similar coefficients were found for activities. Dark regions indicate coefficient differences.

D. Performance Measures and Statistics

All analysis were performed independent of the participant (leave-one-out cross validation). Performance of the activity recognition model was evaluated using the percentage of correctly classified instances for each cluster. EE was evaluated according to the Root Mean Square Error (RMSE). A paired t-test was used to compare RMSEs between models.

IV. RESULTS

A. Relation between allometric coefficients and activities

Fig. 3 shows a heat map of the difference between allometric coefficients in different activities. The map illustrates which activities are affected by BW similarly. It is important to note that activities with large EE (e.g. biking and running) can have very different coefficients depending on the extent to which the activity is weight-bearing, and clustering activities based on level of EE, as done in [14], prevents normalization. Nevertheless grouping activities based on allometric coefficients alone might generate clusters that are not practically recognizable with motion sensors. Thus, we used features representative for an activity's motion patterns and allometric coefficients, in the clustering (Sec. II.C).
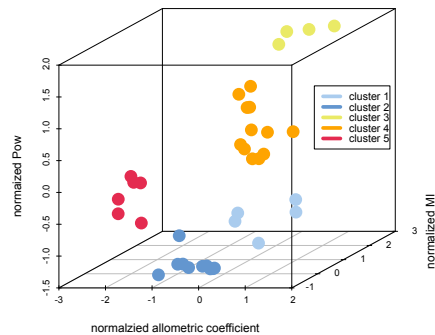

Fig. 4. Features space for the five activities clusters derived by the genetic algorithm. All 37 activities part of our protocol are shown.
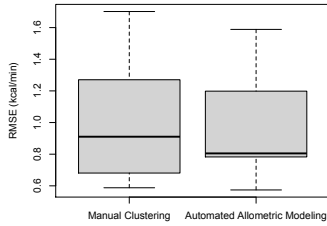
Fig. 5. Boxplot of the RMSE averaged between clusters for manual clustering with no BW normalization as performed in [3] compared to the proposed approach.

B. Activity Clustering

The result of the clustering is shown in Fig. 4 and listed in Table 1. Activities are grouped into clusters optimal for both activity recognition and normalized EE estimation. Accuracy of the activity recognition algorithm is 95.1%.

C. EE Estimation Comparison

Root Mean Square Error (RMSE) for EE was $0.68 kcal/min$ for cluster 1, $1.49 kcal/min$ for cluster 2, $1.46 kcal/min$ for cluster 3, $0.37 kcal/min$ for cluster 4 and $1.23 kcal/min$ for cluster 5. Fig. 5 shows a comparison of the RMSE of the proposed approach compared to the non-normalized and manually clustered approach previously published by our group [3], indicating no loss in performance ($p = 0.66 > \alpha = 0.05$). The RMSE mean and variance across all subjects is slightly reduced. This is an additional benefit of the normalization. Activity-specific multiple linear regression models do not have to deal with differences in body weight anymore. As a result, the linear models can better represent intra-individual variations in EE for a cluster, due to e.g. different level of motion intensity, captured by ACC features, independently of BW.
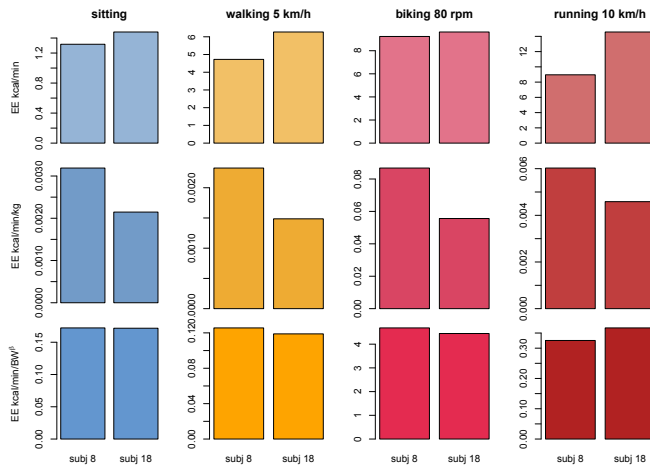

Fig. 6. Differences in EE for subjects with different BW (subj 8, 53.9 Kg, and subj 18, 81.3 kg), for four activities performed at the same intensity. Row 1: No BW scaling, resulting in higher EE estimates depending on BW. Row 2: Over-correction when applying the ratio standard. Row 3: Results when using the proposed cluster based BW normalization.

Fig. 6 gives a representation of the effectiveness of the proposed approach when analyzing subjects of different BW. The new approach reduces differences and provides a way to compare EE levels between heterogeneous subjects, even when performing different activities.

V. CONCLUSIONS

In this paper we proposed a new approach to tackle the problem of BW normalization in EE estimation. We used genetic algorithm-based clustering to develop an automated method to optimally group activities. The clustering was based on allometric coefficients of the clusters, as well as on specific ACC features. So far, we used data from male participants only, which did not allow us to investigate sex differences. Nevertheless, our approach shows promising results to effectively obtain BW-independent EE estimates.

The proposed algorithm allows users to compare EE between different populations, as well as longitudinal results of one individual undergoing e.g. growth, weight loss or training. Such comparisons can provide useful insight into whether differences in aerobic or anaerobic performance in heterogeneous subjects are due to actual difference in physiology or in body weight, as well as provide motivation for subjects participating in physical activity interventions.

REFERENCES

[1] L. Cordain, A. Gotshall, and S. Eaton. Evolutionary aspects of exercise. World Rev Nutr Diet, 81:4960, 1997.

[2] L. Bao and S. Intille. Activity recognition from user-annotated acceleration data. In Pervasive 04, volume 3001, pages 117. 2004.

[3] M. Altini, J. Penders, O. Amft. Energy Expenditure estimation using wearable sensors: a new methodology for activity-specific models. In Proceedings of Wireless Health 2012.

[4] F. Albinali, et al. Using wearable activity type detection to improve physical activity energy expenditure estimation. In UbiComp, pages 311320, 2010.

[5] M. Rumo, O. Amft, G. Troster, and U. Mader. A stepwise validation of a wearable system for estimating energy expenditure in field-based research. Physiological measurement, 32(12):19832001, 2011.

[6] A. G. Bonomi, et al. Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. Journal of Applied Physiology, 107(3):655661, 2009.

[7] E. Tapia. Using machine learning for real-time activity recognition and estimation of energy expenditure. In PhD thesis, MIT, 2008.

[8] D. Rogers et al. (1995) Scaling for the VO2-to-body size relationship among children and adults. J. Appl. Physiol.79:958967.

[9] A. Nevill et al. Scaling, normalizing, and per ratio standards: an allometric modeling approach. J Appl Physiol 79: 10271031, 1995.

[10] I. Zakeri et al. Normalization of energy expenditure data for differences in body mass or composition in children and adolescents. J Nutr 2006; 136: 13716.

[11] J. Penders. A low-power wireless ECG necklace for reliable cardiac activity monitoring on-the-move. In EMBC 11, 2011.

[12] J. E. McLaughlin, et al. Validation of the COSMED K4 b2 portable metabolic system. International J of sports medicine, 22(4):280284, May 2001.

[13] A. M. Prentice et al. Physical activity and obesity: problems in correcting expenditure for body size. Int J Obes Relat Metab Disord 1996; 20: 688691.

[14] L. Atallah, et al.: Sensor Positioning for Activity Recognition Using Wearable Accelerometers. IEEE Transactions on Biomedical Circuits and Systems 5(4), 320329 (2011).

[15] B. E. Ainsworth et al. Compendium of physical activities: an update of activity codes and met intensities. Medicine & Science in Sports & Exercise, 32(9 Suppl):S498504, 2000.

[16] U. Maulik, S. Bandyopadhyay, Performance evaluation of some clustering algorithms and validity indices, IEEE Trans. Pattern Anal. Mach. Intell. 24 (12) (2002) 16501654.