

# Machine Learning and Sensor Fusion for Estimating Continuous Energy Expenditure

Nisarg Vyas, Jonathan Farrington, David Andre and John (Ivo) Stivoric

BodyMedia Inc., One Gateway Center, 420 Fort Duquesne Boulevard, Suite 1900, Pittsburgh, PA 15222 USA

{nvyas, jfarrington, dandre, jstivoric}@bodymedia.com

## Abstract

In this paper we provide insight into the BodyMedia FIT® armband system - a wearable multi-sensor technology that achieves the goals of continuous physiological monitoring (especially energy expenditure estimation) and weight management using machine learning and data modeling methods. This system has been commercially available since 2001 and more than half a million users have used the system to track their physiological parameters and to achieve their individual health goals including weight-loss. We describe several challenges that arise in applying machine learning techniques to the health care domain and present various solutions utilized in the armband system. We demonstrate how machine learning and multi-sensor data fusion techniques are critical to the system's success.

## Introduction

In the United States alone, approximately \$2.3 trillion was spent on health care in 2008. It is well recognized that regular and accurate self-monitoring of physiological parameters and energy expenditure (calorie burn) can provide important feedback that increases self-awareness for personal health. Such awareness and tracking are pre-requisites for cost-effective health management, illness reduction, health-conscious decision making and long-term lifestyle changes. There are several technologies available for physical activity, energy expenditure tracking and weight management. Many of these are accurate, but are bulky, expensive and only used in laboratory settings (Holdy 2004). At the other extreme, there are several single-sensor devices in the market (predominantly accelerometer-based) that are cheaper and light-weight at the expense of accuracy (Beighle *et al.* 2001, Crouter *et al.* 2003). Moreover, there is the doubly labeled water technique, a medical procedure that is guaranteed to give accurate measures of energy expenditure (Schoeller *et al.* 1986), but is very expensive and only gives readings for a 10-14 day period, making it impractical for continuous or short term monitoring.

We believe that a physiological monitoring device providing estimates such as energy expenditure should be accurate, provide continuous feedback to the user of the system, be easy-to-use and be able to work during all the activities of the user's daily life (free-living conditions). Moreover, the device should be cost-effective. The presented BodyMedia FIT armband system (BodyMedia 2011) achieves these goals. The effective use of machine-learning methodologies and a combination of basic sensors used in a smart manner can rival medical grade equipment in terms of accuracy. The key reasons that the BodyMedia FIT system is able to provide accurate free-living estimates are:

- i) It uses a data-centric approach and *estimates* (as opposed to *measures*) most of the key physiological parameters using state-of-the-art data modeling and machine learning techniques.
- ii) It uses multiple sensors - providing a sense of the current activity context of the user, and then it provides a context sensitive estimate of the physiological parameters.

This paper will describe some problems of estimating energy expenditure, the BodyMedia FIT armband, the machine learning techniques used in developing the estimation algorithms, the results of several studies on the accuracy of the device, and results indicating the utility of the device in a weight loss scenario.

## Background

Overview of the system: (Fig. 1) shows the armband device (model MF). It is worn on the upper arm. The current commercial version uses five types of sensors: a three-axis accelerometer tracks the movement of the upper arm and body and provides information about body position. A synthetic heat-flux sensor measures the amount of heat being dissipated by the body to the immediate environment. Skin temperature and armband-cover temperature are measured by sensitive thermistors. The armband also measures galvanic skin response (GSR), the conductivity of the wearer's skin, which varies due to sweating and emotional stimuli. The armband contains a transceiver radio and a USB (Universal Serial Bus) port, allowing wireless transmission as well as wired

downloading of data. The armband is made predominantly of natural ABS (Acrylonitrile Butadiene Styrene) and 304 grade stainless steel and attaches to the arm with an elastic Velcro strap. The armband is approximately 55x62x13 mm (2.2x2.4x0.5 inch) and weighs 45.4 grams (1.6 oz), it stores more than 14 days of continuous body data and has enough power for 5-7 days of wear from a rechargeable battery, when worn 23 hours a day. Each sensor is sampled 32 times per second. Other BodyMedia armband monitors are available that record the same sensor information but differ in other features such as using Bluetooth® wireless or increased memory capacity.

The system collects physiological data on a continuous basis from the person wearing the armband. Data is conditioned, analyzed, interpreted and stored within the device. The device's on-board algorithms estimate and provide real-time estimation of key physiological measures of interest such as the energy expenditure, total number of steps, number of minutes of moderate activity and number of minutes of vigorous physical activity. These key measurements can be displayed wirelessly on a BodyMedia FIT display device or a Bluetooth-enabled cell phone using a mobile application (such as the iPhone

application shown in Fig. 1(d)). Additionally, the data can later be transferred electronically (via USB or wirelessly) to a computer or to a BodyMedia web account, where the software re-analyzes the data and makes a definitive high-level analysis of the data with algorithms that are too computationally expensive to run in the device's firmware.

### Introduction to Energy Expenditure Measurement

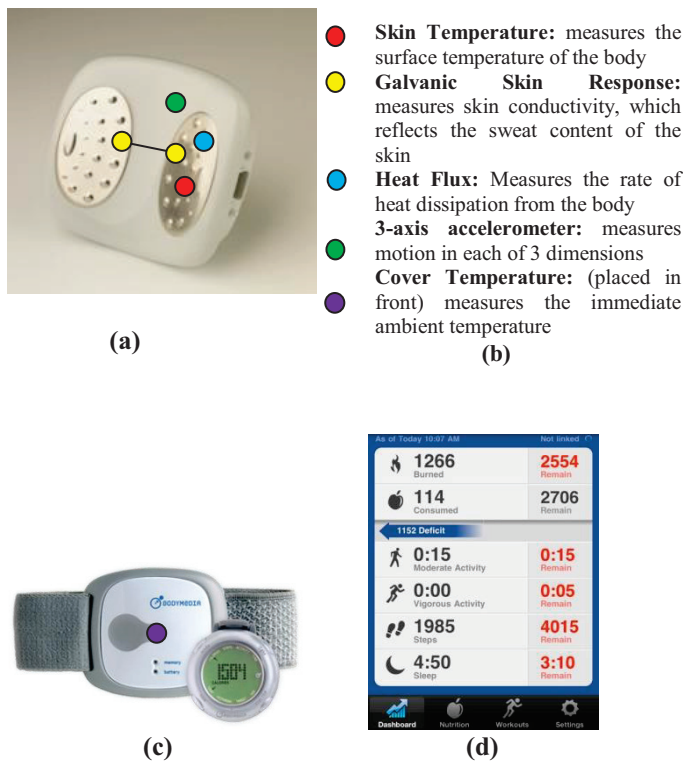
The number of calories a person burns is an important and actionable parameter for many health goals and disease conditions. These include metabolic disorders such as diabetes, weight control (loss, gain, or maintenance), and sports performance. True total energy expenditure (TEE) is very difficult to measure, and nearly all techniques make use of approximations of one kind or another. The following are a few methods commonly used for energy expenditure estimation:

**Indirect Calorimetry:** Metabolic carts measure the oxygen and carbon dioxide a person inhales and exhales and then indirectly compute the calories burned. This technique of measurement is currently very widely accepted in the sports medicine research community. Based on a survey of the literature, devices of this category differ from one another by 5–10% for EE measurements and differ even on repeated measurements of the same activity by around 5–10% (Yates *et al.* 2004, Wells *et al.* 1998). Most metabolic carts are rather large and bulky and are not suited for monitoring outside of the laboratory setting while portable devices are not as accurate. These devices are expensive, costing upward of US\$20,000 for a basic system and US\$40,000 for a portable oxygen analyzer (Holdy 2004, Berntsen *et al.* 2010).

**Doubly Labeled Water (DLW):** The DLW stable isotope method is considered the gold standard for measuring TEE during free living (Schoeller *et al.* 1986). This method is based on the principle that in a loading dose of  $^2\text{H}_2^{18}\text{O}$ , the  $^{18}\text{O}$  is eliminated as  $\text{CO}_2$  and water, while deuterium is eliminated from the body as water. The rate of  $\text{CO}_2$  production, and thus energy expenditure, is calculated from the difference of the two elimination rates. Limitations of the DLW method include a high cost, the need for specialized equipment and expertise to implement the techniques, and the fact that the method can only be used to measure expenditure over a long period of time (e.g. 10–14 days).

**Self-Report Techniques:** Self-report methods include questionnaires, interviews, and activity diaries. There are some advantages to using self-reports or 24-hour recalls, as they are inexpensive and easy to administer. However estimating duration and energy expenditure with these can only provide a rough and inaccurate estimate of activity level.

**Pedometers:** Pedometers, by definition, measure footfalls. The clear advantage of pedometers is their low cost,



**Fig. 1.** BodyMedia System overview (a) BodyMedia armband device and its sensor layout. The figure shows the side of the device worn against the skin. (b) Sensor description (c) Front of the armband (showing the cover temperature sensor) and a display device that provides real-time feedback of physiological parameters to the end-user (d) iPhone application that provides real-time feedback

ranging from \$15 to \$300 (Beighle *et al.* 2001; Freedson *et al.* 2000). In general, pedometers are not accurate when used for activities that do not involve footfalls (e.g. weight lifting, biking, household activities).

**Heart Rate Monitors:** Heart rate is one of the fundamental vital signs and is related to the level of physical exertion. Especially for moderate to strenuous activity, a person's heart rate increases linearly with oxygen consumption (Freedson *et al.* 2000; Welk 2002). However for activities such as rest that do not involve large muscle groups, heart rate does not have a linear relationship to energy expenditure. Heart rate monitoring is quite common and is often used as part of an exercise prescription. Chest-strap heart rate monitors can be a burden to participants because of the constriction required across the chest to maintain good skin contact. Electrode-based heart monitors are difficult to wear, as placement, skin treatment and irritation can be significant issues and detriments to long-term wear.

**Accelerometers:** Accelerometers operate by measuring acceleration along a given axis, using any of a number of technologies, including piezoelectric, micromechanical springs and changes in capacitance. Often, multiple axis measurements are bundled into a single package, allowing two and three axes accelerometers. There is no simple linear relationship between acceleration and energy expenditure, but for certain activities, for example walking, there are well-understood basic principles in physics (work) and physiology (muscle efficiency) that can aid in developing appropriate equations.

## Modeling and Design

**Challenges :** BodyMedia's approach in addressing the estimation of energy expenditure is non-conventional and different from the approaches mentioned above, since it employs machine learning to solve the problem. Machine learning in this context is faced with some significant challenges.

First and foremost is the need for high quality data, which can be very expensive to obtain on a broad enough set of activities and subjects.

The second issue is in the inherent variability present in the data set and in the target user-base. Each armband user is different in terms of their physical characteristics such as age, weight, gender and fitness levels, and all of these characteristics affect the relationships between measured parameters and energy expenditure. Moreover, there is variation in data due to geographical and environmental differences, including different humidity and external temperatures. Additional sources of variation include the different calibration of medical gold-standard equipment. It is required that the final models should be robust enough to tackle the aforementioned variations whether known to be present in the training data or not.

The third challenge is with the fact that the gold-standard data used for building the EE estimation models can only be collected on a limited set of activities in lab settings, whereas the actual use-case of these models is in free-living settings where the users perform a multitude of complex activities. Making EE estimation models in these circumstances violates a fundamental assumption of machine learning that both the training and testing distribution should be the same.

The fourth challenge is in the fact that the models themselves have to satisfy multiple objectives. For example, the model should be accurate for a minute-to-minute real-time feedback for specific activities as well as for weeks-long free-living protocols comprised of a multitude of activities. These different use cases can make model selection difficult.

Another challenge is that algorithms need to continue to work as hardware improves (including miniaturization and simplifications that result in reduced costs).

Finally, many models have a requirement of providing real-time results. In this event, an on-board processor that has limited memory and computational capability computes the algorithm. This influences the underlying features and machine learning methods, in that we prefer methods that are efficient in terms of time and space complexity.

**Modeling Process :** BodyMedia's modeling process can be defined in the following steps:

- i) Data collection
- ii) Data Cleaning
- iii) Feature Generation
- iv) Development of Context Detectors
- v) Development of Regression Models
- vi) Internal and External Validation

**Data Collection:** Any non-trivial machine learning method needs good data. To tackle the challenge of obtaining high quality data, BodyMedia conducts data collection studies at multiple clinical sites spread across the globe. We have worked to enlist many academic researchers as colleagues and advisors, allowing us to obtain data from far more studies than we could fund ourselves. Data collection is designed in a manner so that it provides sufficient samples to capture the variability present in the domain. As specific examples, the data used in the algorithms range from 5-year old children to retirees in their 70s; it represents unhealthy subjects suffering from multiple diseases at one end of the spectrum to highly elite athletes participating in sports events at the other. We capture data from people engaging in many different activities as well, ranging from restful activities such as sleep and lying down to highly vigorous activities such as sprinting, stair-master, rowing and mountaineering. The collected data is either free-living (user-annotated activities) or it could be a part of a

strict protocol in a laboratory or controlled environment. For most lab studies, data from high accuracy gold-standard equipment (such as metabolic carts or metabolic chambers) is also collected for training and testing purposes. It is to be noted that the free-living data consists of many activities and is used for activity classification. The lab data is limited to a certain subset of activities and it is used primarily for building EE estimation models.

**Data Cleaning:** We have developed a rigorous process for cleansing data and preparing it for machine learning. Armband sensor data in each data file is verified; cases of sensor malfunction are detected by comparing the armband sensor data with the sensors' standard distribution. If outliers are found in the sensor values, those data points are discarded. In the cases where gold-standard medical equipment data is also being collected, each data point is carefully aligned using semi-automated procedures. Occasionally, the gold-standard equipment is not properly calibrated, and the equipment has a tendency to overestimate or underestimate. Such cases are identified based on its compliance to standard METs (Metabolic Equivalent, essentially energy cost per unit of mass) ranges for the corresponding activities (Ainsworth *et al.* 2000). Moreover, the data sets are checked for correctness of the activity annotations (in most cases inserted manually by the user or experimenter). The activity annotations are verified for correctness by comparing the sensor values with the standard sensor value distribution for the activity and applying other such heuristics. As an example, most cases of resting activities provide little or minimal changes in motion sensors and GSR sensors. If it is found that the sensors recorded high amount of motion and/or steep rises in GSR, it is very likely that the activities were annotated incorrectly. For cleaning of the activity annotations, the philosophy is to err on the side of caution, as we have found that allowing even small amounts of poorly labeled or misaligned data can have outsized effects on algorithm performance.

**Feature Generation:** The sensors used in the armband are sampled at 32 Hz, whereas the armband records data every minute (this can be adjusted through software). Thus compressed and summarized features are calculated and created from the raw data. Currently more than 50 features of this multi-dimensional raw data stream are gathered as separate channels. For example, the variance of the heat flux is a channel, as is the average of the heat flux values. Some channels are fairly standard such as standard deviation, frequency, peaks and averages. Others are complex proprietary algorithms embedded in the on-board processor of the armband. Then typically these summary features for each minute epoch are stored and the raw data discarded to conserve memory. Typically, we refrain from calculating computationally complex and storage-expensive operations (such as Fourier transforms) in order

to conserve the on-board processor's memory and finish the calculation of the feature in each epoch / duty cycle. Approximations are also used if the actual feature is computationally expensive to create.

The next stage of feature generation is done on the recorded data retrieved from the armband. This is done to find features that aid us in recognizing patterns of activity and for calibrating various measures against one another. For example, relative values of GSR are often more useful than absolute readings. Multiple methods are used to extract these features. Some features are derived using domain knowledge of exercise and physiology. Some are derived using an automated feature generation technique similar to genetic programming where features must pass a few standard statistical tests (such as high correlation to the ground-truth EE in all or some activities). Some features are derived using standard machine learning feature generation techniques such as Principal Components Analysis (PCA) and Independent Component Analysis (ICA). Some are added based on intuition and visual observation. At the end of this phase, a feature space of more than 500 variables is created.

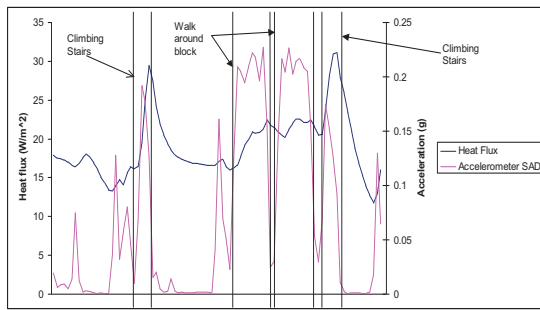
**Development of Context Detectors:** The next stage of the modeling process is to develop a series of classifiers that break down a user's activity into primary components for which good models of energy expenditure can be created. Classifiers are created for the following basic activities: walking, running, stationary biking, resting, weight-lifting, motoring, road biking, sleep and rowing. Many popular computationally inexpensive machine learning methods such as naive Bayes and decision trees are tried out for feature selection and training for the classifiers.

To avoid overfitting, all the feature selection and classification algorithms use k-fold cross-validation. In our experience it is not sufficient to perform k-fold data-based cross-validation (i.e. creating folds at the level of individual data points) to ensure robustness and generalization capabilities, as there are many subject-specific traits present in the data. Instead, the entire data of each subject is assigned to one of the k folds. We refer to this strategy as k-fold by-subject cross-validation. Using by-subject cross-validation results in algorithms that generalize well to unseen subjects. Moreover, to avoid overfitting by the classifiers, it was found necessary to avoid features that include subject-specific traits (such as demographic information).

It should be noted that the use of multiple sensors provides orthogonal sets of features in the feature space, helping to provide more discriminating capability to the classifiers.

Fig. 2 represents signal values of one accelerometer-based feature and a heat-flux based feature during various activities. It is seen that the values for the accelerometer feature for "climbing stairs" and "walking around the





**Fig. 2.** A sample of armband signals during various activities. Multiple sensors allow better activity classification.

block" are very similar, making it complex for the classifiers to infer the activity based on the accelerometer data. With the introduction of heat flux, even a simple classifier can distinguish between the two activities.

Typically, the classifier model is designed as a hierarchical combination of various sub-classifier responses. A base classifier classifies the data into generic activities, and the next level of the classifier model provides a more fine-grained label to the activities. For example, one can think of the base classifier only classifying an activity as "biking", and the second-phase classifier classifying all the "biking" data points into "stationary biking" and "road biking". The classifier model makes use of several sub-classifiers spread across multiple levels of hierarchy.

**Development of Regression Models:** In this phase, several regression models are built that provide energy expenditure estimates. Usually the models are built for a specific activity (or for a set of very similar types of activities). The regressions are then combined according to the probabilities output from the activity classifiers. Many prevalent AI-based regression techniques such as robust-regression and locally-weighted regression are used for fitting the data. Feature selection and training for the regression models is also performed using k-fold by subject cross-validation.

Most of the physiological measures of interest estimated by the armband are dependent on subject-specific traits (e.g. mass). Rather than predicting absolute measures, regressions are tuned to predict relative measures that are subsequently adjusted for the subject. For example, in the case of energy expenditure, the regression models are actually trained on the relatively subject-independent unit METs (Ainsworth *et al.* 2000) instead of absolute units such as kJoules or kcalories. These steps of activity prediction and value estimation overlap one another and are addressed simultaneously. Multiple iterations result in improved algorithms.

**Internal and external validation:** For each algorithm release cycle, certain data sets are kept untouched for the

entire development period, and performance of the model is evaluated on those validation sets. The models are approved and released only if they pass pre-defined criteria on the validation sets as well as the training sets. Just like the training data sets, the validation sets are ensured to provide sufficient data samples for each activity with a broad variety of subjects. Some of the validation sets target particular areas of concern such as a demographic group (children, unhealthy adults and athletes) or specific activities. Some types of data sets are only good to serve as validation sets, for example the doubly labeled water dataset, where there is only one reading of TEE (Total Energy Expenditure) every two weeks. At the alpha and beta stages of the release, results of the models are observed, and minor changes are made to the model if necessary. Many researchers also carry out independent external validation and performance evaluation of the system, providing helpful cues to further improvement (St. Onge *et al.* 2007, Welk *et al.* 2007, Jakicic *et al.* 2007, Malavolti *et al.* 2007).

## Results

BodyMedia armbands have been commercially available since 2001, and currently the fifth generation of the system is in the market. There are more than half a million users of the system spread around the world. To date, BodyMedia has collected more than 10 billion minutes of armband data. The system has recorded more than 170 billion steps and estimated more than 20 billion calories.

**Data Sets:** In the most recent energy expenditure algorithm created at BodyMedia, a data set with roughly 1 million minutes featuring around 800 users was used for training the context detectors. All the minutes were carefully annotated by the users and they were cleaned to make it suitable for modeling. Developing the regressions required a gold-standard data set which had 658 subjects and approximately 40,000 data points.

The datasets had a wide range of demographic variations: age range varied from 5 years to 78 years, weight range varied from 18kg to 152kg (40lbs to 335lbs). The data was collected from more than 50 different studies, conducted at external clinician sites spread across the world and from studies conducted in-house.

**Classification:** Tables 1 & 2 show classification results for the most recent algorithm for major activities. Notice that some of the true-positive rates are seemingly quite low. Further inspection reveals that much of the misclassification happens between similar types of activities. For example, misclassification between motoring and resting occurs often, but from an energy

	True Positive Rate %	True Negative Rate%	Accuracy
Walking	91.5	94.2	94.0
Running	90.6	99.1	98.6
Stationary Biking	60.4	96.0	95.1
Rest	82.6	72.6	75.8
Weight Lifting	34.7	98.4	97.0
Motoring	62.4	94.8	90.0
Road Biking	89.0	99.3	98.9
Sleep	69.6	97.7	94.4
Rowing	83.6	99.8	99.8

**Table 1:** Classification results for prominent activities, evaluated using by-subject cross-validation

expenditure (EE) standpoint the misclassification does not cost much because their EE ranges are very similar.

Table 2 shows the four most frequently predicted classes for each true class (the confusion in classification). Tables 1 & 2 show results evaluated by using by-subject cross-validation. The model also generalizes well to unseen subjects' data, with the overall accuracy of the unseen subjects' dataset just 1% less than the accuracy obtained for by-subject cross-validation.

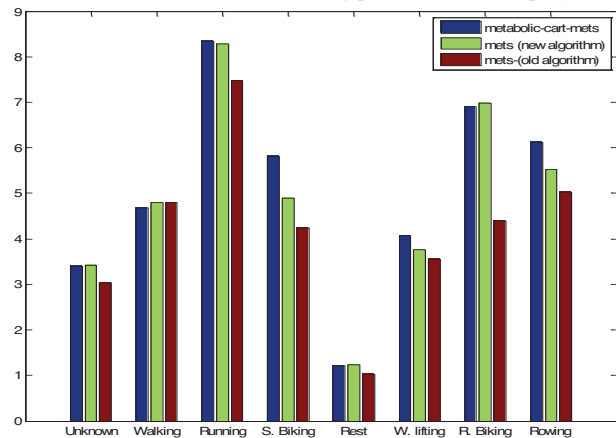
**Regression Models:** Fig. 3 shows results of average METs for a new release candidate versus a recent model already in use per each activity. The METs value can be thought of as the relative activity intensity and energy requirement. It is seen that the release candidate bars are much closer to true average METs, hence showing the improvement in the algorithm.

Typically, the errors on the regression models are measured in Mean Absolute Percentage Error (MAPEs). MAPEs are calculated at each minute, as well as for each entire session of continuous observation. The new release candidate algorithm has 15% session MAPEs over all the lab data. The new release candidate algorithm proved to be providing robust estimates for children too, with the session MAPEs as low as 13.7%. Daily MAPEs for adults are expected to be lower than the lab data suggests. The lab data is comprised predominantly of subjects engaged in exercise, whereas a typical day is made up of mostly sleep (about 30%), restful activities (about 60%), and a small amount of moderate to vigorous activity.

**Doubly Labeled Water Data:** As mentioned earlier, doubly labeled water is the most accurate method to estimate energy expenditure, but provides only one reading per 14 days.

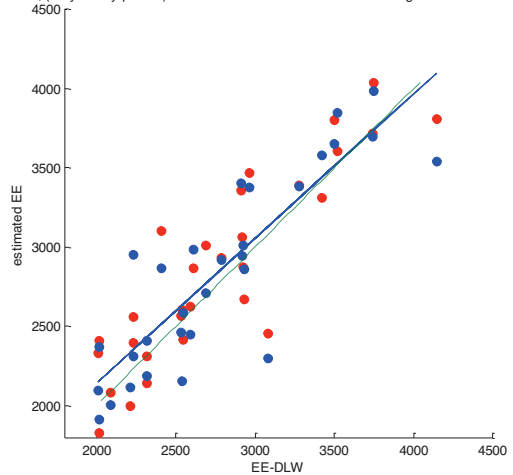
Walking	91.5% walking, 6.2% running
Running	90.6% running, 9.1% walking
Stationary Biking	60.4% stationary biking, 9.9% rest, 9.6% motoring, 9.1% walking
Rest	82.6% rest, 7.9% motoring, 4.9% sleep, 3.1% S. Bike
Weight Lifting	34.7% weight lifting, 28.3% rest, 17.0% stationary biking, 9.9% walking
Motoring	62.4% motoring, 27.6% rest, 4.5% stationary biking, 2.0% sleep
Road Biking	89.0% road biking, 3.1% stationary biking, 2.7% rest, 2.0% motoring
Sleep	69.6% sleep, 29.6% rest
Rowing	83.6% rowing, 13.3% walking, 2.2% stationary biking

**Table 2:** Four most predicted classes per activity. Only the activities that have more than 1% of prediction are displayed.



**Fig. 3.** Average METs per activity - (ground truth computed from metabolic carts, for current algorithm and new release candidate)

(daily EE by person) DLW EE vs. estimated EE corr=0.88 Avg. Error = 8.44 %



**Fig. 4.** DLW results compared with estimated EE for 30 adult subjects (14 days per subject). Blue is armband model Pro3, red is model MF.

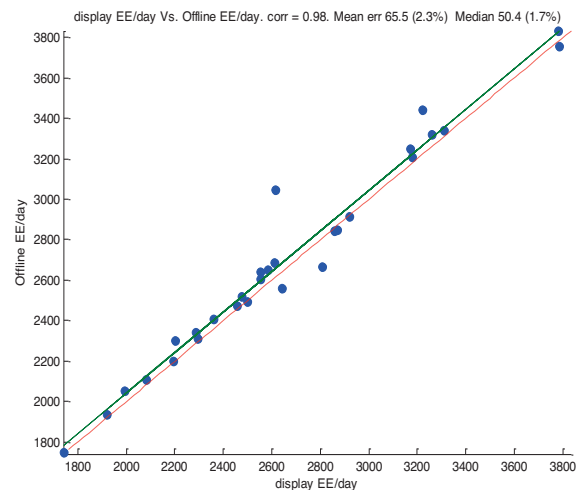
Fig. 4 shows a scatter plot showing that the estimated TEE values match well with the actual TEE calculated from the DLW method. The data was collected on 30 adult individuals and 30 children wearing two versions of armbands (2008-2010 model Pro3 and 2010 model MF), one on each arm for a period of two weeks. The MAPE value is less than 10% for adults, and the correlation between the true and estimated TEE is 0.88 (Johansen *et al.* 2010). A similar study was also conducted for children, and the MAPE value was under 15% (Calabró *et al.* 2011).

An independent study (Bernsten *et al.* 2010) validated the accuracy of the armbands in simulated free-living conditions, where 20 subjects participated in 60 to 120 minutes of realistic daily activity. The estimation error from the armbands was less than 10%. These results demonstrate that the models are generic enough to work for unseen subjects performing free-living activities.

**Accelerometers-only Versus Multi-sensors:** A study was conducted to measure the efficacy of models built based on the current sensor set versus models built only on the accelerometer and motion based signals (but still using Bodymedia's pattern recognition methods). Over 30 subjects participated in various exercise activities. It was found that models that used all the sensors had 8% per-subject error, whereas the models that used only the accelerometers had 12-15% per-subject error. The study also provided comparative evaluation of BodyMedia armband devices with other commercially available energy expenditure estimation devices and it was found that BodyMedia armband system provided the most accurate results (Lee *et al.* 2011) in comparison to other devices. The next best device had 14% per-subject error, in comparison to BodyMedia system's 8% per-subject error.

**Commercial Applications:** The BodyMedia armband is in use in several commercial applications, including the BodyMedia FIT product and the bodybugg® product from 24 Hour Fitness. Not only is the armband in use by many users but it appears to significantly help users in achieving their weight-loss and lifestyle goals. A study performed at the University of South Carolina showed that participants who used armbands in their weight-loss program lost more than twice the weight compared to the subjects who did not use the armbands (Barry *et al.* 2010; Sui *et al.* 2010). A weight-loss study done at the University of Pittsburgh achieved similar results (Pellegrini *et al.* 2011).

Various versions of the armband have been in active use by hundreds of thousands of users over the last nine years. The earlier products were larger, heavier, and more expensive to manufacture and the earliest of these had only a two-axis accelerometer rather than the current three axis model. Additionally, the algorithms have been updated numerous times over the years as more lab data provides



**Fig. 5.** Comparison of Display EE Vs. Offline EE. The correlation between them is 0.98 and the mean difference is 2.3%.

for more accurate and refined algorithms. By having the computations performed on our websites, modifying the algorithms is straightforward. From time to time, when users upload data, firmware updates are pushed out to the armbands.

**Real-time Vs. Offline EE Estimates:** It has been observed that the real-time EE estimates match accurately with the offline EE estimates. The data from the DLW experiment mentioned earlier (30 adult subjects, 14 days wear) found that the mean difference between the Real-time EE (display-EE) and the offline-EE was 2.3% per day (about 66 kcal) with a median difference of 1.7% per day, Fig. 5.

## Conclusions and Future Work

With healthcare costs spiraling each year, people can benefit from an effective, inexpensive, easily wearable and accurate physiological monitoring device: BodyMedia armbands attempt to provide that solution in a non-conventional way using sensor fusion and state-of-the-art machine learning and artificial intelligence techniques.

In this paper, the modeling process for estimating the physiological parameters (especially energy expenditure) has been described. The results presented here demonstrate the capability of the armband sensors and models to provide accurate results for various activities for a large range of users in both lab and free-living settings.

BodyMedia is engaged in continued refinements to the platform and the development of new body monitoring capabilities. These include the integration of new sensors and the ongoing development of data models to extract new physiological features and contextual activities. Some of the other projects that BodyMedia is focusing on include: blood glucose estimation (Vyas *et al.* 2010, Rollins *et al.* 2009), heart rate estimation (Al-Ahmad *et al.*

2004), critical care parameter estimation (Convertino *et al.* 2010), fine-grained sleep detail and estimation of total calories consumed. All of these projects extensively use data-driven methods and sophisticated machine learning techniques.

## Acknowledgements

The authors would like to thank their colleagues at BodyMedia, especially Rachel Jackson, Scott Boehmke and Ray Pelletier for their continuous efforts in data collection study design and management, hardware design, and assistance in modeling respectively. We would like to thank Dr. Astro Teller for the conceptualization and work on early designs of the armbands. We would also like to thank Dr. Greg Welk, Associate Professor at Iowa State University (ISU) for graciously accepting our request to share and present some of the results (both published and unpublished) on studies conducted by him at ISU.

## References

- Ainsworth B. *et al.* (2000), Compendium of physical activities: an update of activity codes and MET intensities. *Medicine & Science in Sports & Exercise* 32(9 Suppl):S498-504
- Al-Ahmad, A., Homer M., Wang P. (2004), Accuracy and Utility of Multi-Sensor Armband ECG Signal Compared With Holter Monitoring. *New Arrhythmia Technologies Retreat*
- Barry V., Shuger S., Beets M., Sui X., Wilcox S., Hand G., McClain A., Blair S. (2010), Wearing the SenseWear Armband and Weight Loss in Sedentary Overweight and Obese Adults. *American College of Sports & Medicine meeting*
- Berntsen S., Hageberg R., Aandstad A., Mowinkel P., Anderssen S., Carlsen K. and Andersen L. (2010), Validity of physical activity monitors in adults participating in free living activities. *British Journal of Sports & Medicine* 44:657-664
- Beighle A., Pangrazi R., Vincent S. (2001), Pedometers, physical activity, and accountability. *Journal Of Physical Education, Recreation and Dance* ;72(9):16-9.
- Bodymedia Inc.(2011) What is Bodymedia Fit. [www.bodymedia.com/Shop/Learn-More/What-is-BodyMedia-FIT](http://www.bodymedia.com/Shop/Learn-More/What-is-BodyMedia-FIT)
- Calabró, M. A., Lee J., St-Maurice P., Welk G. (2011), Validation of pattern-recognition monitors in children using the doubly labeled water method. Submitted to *Medicine & Science in Sports & Exercise*
- Convertino V., Hurst G., Ryan K., Vyas N., Ward K., Rickards C. (2010), Bleeding or Active? Validation of a Machine-Learning Algorithm for Remote Determination of Blood Volume Status, *American College of Sports and Medicine*, 42(5)- 285
- Crouter S., Schneider P., Karbult K., Bassett D. (2003), Validity of 10 electronic pedometers for measuring steps, distance, and energy cost. *Medicine & Science in Sports & Exercise*;35 (8):1455-60
- Freedson P., Miller K. (2000), Objective monitoring of physical activity using motion sensors and heart rate. *Research Quarterly for Exercise & Sport*. 71(2 Suppl):S21-9
- Holdy K. (2004), Monitoring energy metabolism with indirect calorimetry: instruments, interpretation, and clinical application. *Nutrition in Clinical Practice* 19(5):447-54
- Jakicic J., Marcus M., Gallagher K., Randall C., Thomas E., Goss F.L., Robertson R.J. (2004), Evaluation of the SenseWear Pro armband to assess the energy expenditure during exercise. *Medicine & Science in Sports & Exercise*, May 2004; 36(5) 897-904
- Johannsen D.L., Calabro M.A., Stewart J., Franke W., Rood J.C., Welk G.J. (2010), Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Medicine & Science in Sports & Exercise* 42(11):2134-40
- Lee J., Welk G. (2011), Validity of consumer-based physical activity monitors under free-living conditions Submitted to *Journal of Physical Activity and Health*
- Malavolti M., Pietrobelli A., Dugoni M., Poli M., Romagnoli E., DeCristofaro P., Battistini N.C. (2007), A new device for measuring resting energy expenditure (REE) in healthy subjects, *Nutrition, Metabolism & Cardiovascular Diseases*; 17(5) 338-343
- Pellegrini C.A. (2010), The Comparison of a Technology-based system and an In-Person Behavioral Weight Loss Intervention in Overweight and Obese Adults. PhD Dissertation, University of Pittsburgh
- Rollins D., Bhandari N., Kleinedler J., Kotz K., Strohbehn A., Boland L., Murphy M., Andre D., Vyas N., Welk G. and Franke W. (2010), Free-living Inferential Modeling of Blood Glucose Level using only Noninvasive Inputs, *Journal of Process Control*, Volume 20, Issue 1, pp 95-107
- Schoeller D., Ravussin E., Schutz Y., Acheson K., Baertschi P., Jequier E. (1986), Energy expenditure by doubly labeled water: validation in humans and proposed calculations. *American Journal of Physiology*, 250(5 Pt 2):R823-30
- St-Onge M., Mignault D., Allison D., Lhoret-Rabasa R. (2007), Evaluation of a portable device to measure daily energy expenditure in free-living adults, *American Journal of Clinical Nutrition*, 85 (742) 9
- Sui X., Meriwether R., Hand G., Wilcox S., Dowda M., Blair S. (2010), Electronic Feedback in a Diet and Physical Activity-Based Lifestyle Intervention for Weight Loss: Randomized Controlled Trial, *American Heart Association Annual Conference*
- Vyas N., Andre D., Rollins D., Stivoric J., Jackson R. (2009), Development of a Personalized Non-invasive Glucose Monitoring System for Free-living Environments. *Annual Meeting of Diabetes Technology*
- Welk G. (2002), Physical activity assessments for health-research. *Human Kinetics*
- Welk G., McClain J., Eisenmann J., Wickel E. (2007), Field validation of the MTI Actigraph and Bodymedia armband monitor using the IDEEA monitor. *Obesity*, 15 (4) 918-928
- Wells J. and Fuller N. (1998), Precision and accuracy in a metabolic monitor for indirect calorimetry. *European Journal of Clinical Nutrition*. 52(7):536-40
- Yates J., Cullum M., Pittsley J. (2004), Validation of a portable, indirect calorimeter for the measurement of resting metabolic rate. *Medicine & Science in Sports & Exercise*. 36(5):S247.