Article

Floating Epoch Length Improves the Accuracy of Accelerometry-Based Estimation of Coincident Oxygen Consumption

Henri Vähä-Ypyä [1] , Pauliina Husu [1] , Tommi Vasankari [1,2] and Harri Sievänen [1,*]

1   The UKK Institute for Health Promotion Research, 33500 Tampere, Finland;
     henri.vaha-ypya@ukkinstituutti.fi (H.V.-Y.); pauliina.husu@ukkinstituutti.fi (P.H.);
     tommi.vasankari@ukkinstituutti.fi (T.V.)
2   Faculty of Medicine and Health Technology, Tampere University, 33014 Tampere, Finland
*   Correspondence: harri.sievanen@ukkinstituutti.fi

Abstract: Estimation of oxygen consumption (VO_2) from accelerometer data is typically based on prediction equations developed in laboratory settings using steadily paced and controlled test activities. These equations may not capture the temporary changes in VO_2 occurring in sporadic real-life physical activity. In this study, we introduced a novel floating epoch for accelerometer data analysis and hypothesized that an adaptive epoch length provides a more consistent estimation of VO_2 in irregular activity conditions than a 6 s constant epoch. Two different activity tests were conducted: a progressive constant-speed test (CS) performed on a track and a 6 min back-and-forth walk test including accelerations and decelerations (AC/DC) performed as fast as possible. Twenty-nine adults performed the CS test, and sixty-one performed the AC/DC test. The data were collected using hip-worn accelerometers and a portable metabolic gas analyzer. General linear models were employed to create the prediction models for VO_2 that were cross-validated using both data sets and epoch types as training and validation sets. The prediction equations based on the CS test or AC/DC test and 6 s epoch had excellent performance (R^2 = 89%) for the CS test but poor performance for the AC/DC test (31%). Only the VO_2 prediction equation based on the AC/DC test and the floating epoch had good performance (78%) for both tests. The overall accuracy of VO_2 prediction is compromised with the constant length epoch, whereas the prediction model based on irregular acceleration data analyzed with a floating epoch provided consistent performance for both activities.

Keywords: accelerometer; energy consumption; measurement; movement; physical activity; prediction

1. Introduction

Physical activity (PA) offers significant health benefits and mitigates health risks [1]. It is important to regularly measure population-level PA with valid methods. Such methods should be able to measure the frequency, duration, and intensity of PA, and desirably also the type of activity and its context [2]. These methods can be divided into self-reports and device-based methods. Self-reports provide pertinent information on the specific type or context of PA but are known to overestimate the exercise time while underestimating the active time accumulated during daily routines [3]. The device-based methods can assess PA in a more standardized manner regardless of the individual's current fitness level and body weight, factors that may influence the subjectively perceived and reported intensity of PA [4]. In population studies, accelerometry is a commonly used cost-efficient and feasible method to collect dense data over long periods, allowing a detailed examination of daily behavior [2]. Also, accelerometry shows good compliance among participants [2] and provides relevant information on the dose–response relationship between physical behavior and health outcomes [1].

Accelerometer measurements of physical activity and estimations of its intensity are typically based on equations developed in laboratory settings using steadily paced activities [5]. However, they may offer little insight into how methods will perform in real-life conditions where activity and postural transitions are frequent and occur at irregular intervals [6]. Habitual PA comprises accelerations, turning, and braking, which all increase oxygen consumption (VO_2) compared to constant-speed locomotion [7]. The extra VO_2 depends on the rate of acceleration and deceleration and the number of changes in the direction of movement [8,9]. Changes in the movement speed or direction require applying additional force and energy to change the original momentum and kinetic energy to a new level [10].

The epoch length is the duration of time for which PA data is collected and processed for estimating coincident VO_2. During continuous PA performed constantly at the same intensity, the accuracy of VO_2 estimates from the accelerometer data is independent of the selected epoch length [11]. During intermitted conditions, short epoch lengths (less than 10 s) are effective in identifying PA and stationary phases, whereas longer epoch lengths (over 30 s) tend to smooth the acceleration signal and lead to missed short bursts of PA [5,11]. However, the longer epochs may be more appropriate for evaluating the physiological impacts of PA on health outcomes [11]. In a real-life context, different epoch lengths yield substantially different estimates of the total accumulated PA time, which can be confusing when the individual volumes of daily PA are evaluated [12].

Although short epoch lengths can detect short bursts of PA, they are not able to detect the higher energy expenditure caused by changes in the direction or speed of movement. A higher speed increases both the VO_2 and accelerometer output but changing the body direction by 180°, i.e., performing a U-turn increases the VO_2 while the accelerometer output does not respond similarly [7,13]. Current accelerometric prediction models cannot estimate the increased metabolic requirements caused by temporary or sporadic accelerations, decelerations, and rotations of the body, exceeding the metabolic requirement of constant-speed movements [9,14].

The accelerometer measurements of human movements can be contaminated with noise, which can be mitigated by applying a low-pass filter to the raw data [15]. The choice of the filter and its parameters depends on the noise characteristics and the desired trade-off between noise reduction and preserving meaningful motion data. For example, Wundersitz et al. (2015) found a low-pass filter with a cutoff of 12 Hz appropriate for measuring peak impacts during different types of movements [16]. Correspondingly, Fridolfsson et al. (2018) found that signal frequencies above 10 Hz did not include relevant information relative to movement [17]. Under normal conditions, the peak frequency rarely exceeds 6 Hz [18], and most relevant information ranges up to 4 Hz, corresponding to the step frequency [17].

In this paper, we introduce a novel approach to using a floating epoch length for estimating VO_2 from raw triaxial accelerometer data and compare the performance of developed VO_2 estimation models in walking tests, which were performed either at a constant speed or contained both acceleration and deceleration phases. The floating epoch algorithm estimates the change in the kinetic energy by setting the epoch length according to the coincident stride cycle. This approach can help ensure that all relevant parts of acceleration data are considered for analysis while providing the possibility to estimate the energetic cost of the changes in the kinetic energy. We hypothesize that the continuous adjustment of the epoch length improves the prediction of VO_2 compared to the conventional fixed epoch length approach. Our objective is to develop a more accurate and comprehensive prediction equation for VO_2 compared to existing prediction models.

2. Materials and Methods

2.1. Accelerometry

This study employed the raw acceleration signals from a triaxial accelerometer (Hookie AM20, Traxmeet Ltd., Espoo, Finland) collected during a constant-speed (CS) walking test and a walking test including both acceleration and deceleration phases (AD/DC).

This accelerometer employs the commonly used 13-bit digital triaxial acceleration sensor module (ADXL345; Analog Devices Inc, Norwood, MA, USA). The measurement range of each measurement axis of the accelerometer is ±16 g (g denotes the Earth's gravity, 9.81 m/s^2), and the data is measured at a 100 Hz sampling frequency. The peak acceleration with hip-worn accelerometer can exceed 10 g in foot ground contacts during running [19] and a sampling frequency lower than 100 Hz can lead to missed movement events [20]. The accelerometer was attached to a hip-mounted elastic belt at the level of the iliac crest (Figure 1).

Figure 1. The placement of the accelerometer at the level of the iliac crest on an elastic hip-band.

2.2. Test Protocols

The data employed in the present analysis were collected in our previous studies, where the participants were selected with convenience sampling [21,22]. The CS test was carried out at the premises of Vierumäki Sports Institute near the city of Lahti, Finland, and the participants were recruited from the Lahti region by the staff of the Vierumäki Sports Institute. The AC/DC test was carried out at the premises of the UKK Institute in the city of Tampere, Finland, and the participants were recruited from the Tampere region by the staff of the UKK Institute. Participants were informed about the experimental test protocol, and they gave their written informed consent before performing their test. This study conformed to the code of Ethics of the World Medical Association Declaration of Helsinki and was approved by the Regional Ethics Committee of the Expert Responsibility area of Tampere University Hospital with codes R11115, R12246, and R13040.

Before testing, participants' body height and weight were measured with standard methods. During both the CS and AC/DC tests, VO$_2$ was continuously measured with a portable respiratory gas analyzer (Oxycon Mobile, CareFusion, Yorba Linda, CA, USA), which employs a telemetry system for data collection. A metabolic cart is designed to evaluate the body's energy expenditure and respiratory gas exchange. The device gathers and analyzes information about how the body utilizes oxygen and produces carbon dioxide during different activities. Before each test, the flow meter, oxygen sensor, and carbon dioxide sensor of the metabolic cart were calibrated against known volumes and contents of oxygen and carbon dioxide gases.

In the CS test, 29 adults (15 males and 14 females) performed a pace-conducted non-stop test on a 200 m track. The initial speed was 0.6 m/s, and the speed was increased by 0.4 m/s every 2.5 min. Each stage lasted 2.5 min. The pace was controlled by light sources alongside the track, which were automatically lit up one at a time according to the required pace (Figure 2a). The participant had to keep up with the light sources. All light sources were also visible to the test supervisor. The selected gait type was recorded for each stage, and only completely walked stages were included in the present analysis. As a measure of steady-stage VO$_2$ for a given speed, the mean VO$_2$ of the final minute of the corresponding stage was used. The acceleration was analyzed for the final 2 min of each stage [21].
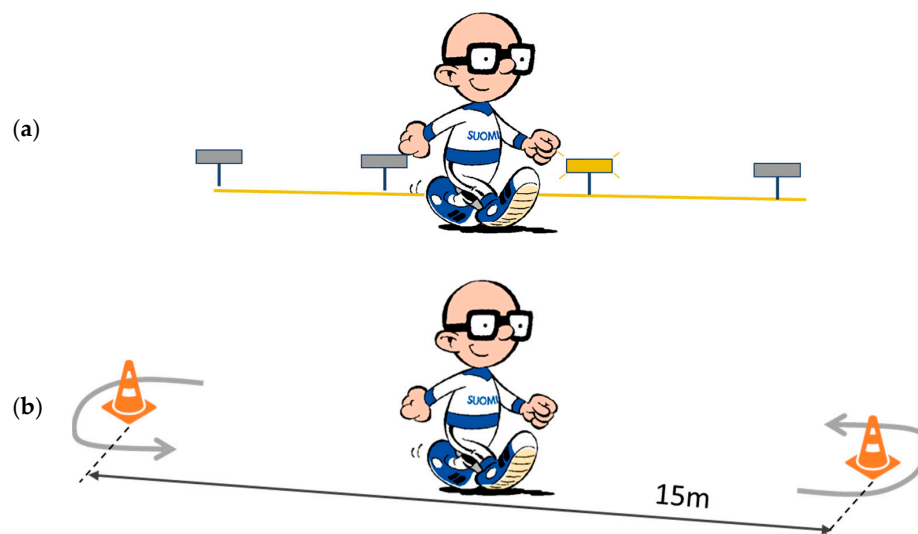
Figure 2. Illustration of the walking tests: (**a**) in the constant speed (CS) test, the participants were asked to follow the pace set by an advancing light source on a 200 m track; (**b**) in the acceleration/deceleration (AC/DC) test, the participants were asked to walk back and forth around the plastic cones on a 15 m track as fast as possible.

In the AC/DC test, 61 adults (35 males and 27 females) walked back and forth on a 15 m track as fast as possible for 6 min (Figure 2b). Like the CS test, the steady-stage VO$_2$ for a given speed was based on the mean VO$_2$ of the final minute of the test. The acceleration was collected and analyzed throughout the test. Before the test, the participants familiarized themselves with the metabolic cart and the 15 m test track. The participants were instructed to walk back and forth around the cones as fast as possible for six minutes. At the end of every minute, the tester informed the passing of walking time and counted the number of laps each participant walked [22].

2.3. Data Analysis

The acceleration data was analyzed with constant 6 s epochs and floating epochs using the MAD algorithm [21]. The acceleration data was first preprocessed with a low-pass filter to remove the noise. Secondly, the data was bandpass filtered to extract the cycle length of movement, which sets the epoch length.

The length of the floating epoch was set by the stride time of the gait. A single stride time starts from the initial contact of one foot and continues to the subsequent contact of the same foot [23]. Thus, a single stride contains two steps, one with each leg. After analyzing the data within one epoch, the analysis window was then shifted forward by the duration of one step to begin analyzing the next epoch. The data analysis with the floating epoch was based on the MAD value of each epoch, the difference in the MAD value between the adjacent epochs, and the inverse of the step time (i.e., step frequency).

Tri-axial acceleration was measured from all three orthogonal measurement axes (*x*, *y*, and *z*) in actual g-units and stored for further analysis. In short, each measurement point (*i*) consisted of samples *x*, *y* and *z*. The resultant acceleration (*r*), which defines the magnitude of the acceleration vector and contains both dynamic and static component of acceleration, was calculated for each (*i*) time point with Equation (1).

$$r[i] = \sqrt{x[i]^2 + y[i]^2 + z[i]^2} \tag{1}$$

The resultant acceleration data was smoothed with a 2nd order low-pass filter with a 0.12 Hz cut-off frequency. The low-pass filtered value (R_{LF}) for the time point (*i*) was calculated with Equation (2), where R_{LF} is the low-pass filtered value of three adjacent

points (i, $i-1$ and $i-2$), a is the filter coefficient ($127/128$ for the 0.12 Hz cutoff frequency) and $r[i]$ is the resultant acceleration for the point (i).

$$R_{LF}[i] = 2aR_{LF}[i-1] - a^2 R_{LF}[i-2] + \left(1 - 2a + a^2\right) r[i] \tag{2}$$

The duration of the single step, defined by the start time point (j) and the number of included samples (N), was determined by an algorithm comprising two digital filters for removing low- and high-frequency noise and for detecting the zero-crossing. First, the resultant acceleration r was filtered with 2nd order Butterworth low-pass filter with a 12 Hz cutoff frequency (R_{LF12}). The R_{LF12} for the time point (i) was calculated with Equation (3), where the filter coefficient value a_1 is $46/512$, a_2 $93/512$, b_1 $125/128$, and b_2 $44/128$. The accelerometer data filtered this way is demonstrated to have an acceptable level of concurrent validity compared to a 3-dimensional motion analysis system [16].

$$R_{LF12}[i] = a_1 r[i] + a_2 r[i-1] + a_1 r[i-2] + b_1 R_{LF12}[i-1] - b_2 R_{LF12}[i-2] \tag{3}$$

Secondly, the low-pass filtered data, R_{LF12}, was bandpass filtered (R_{BP}) with a 2nd order Butterworth filter with bandwidth from 1.0 Hz to 3.0 Hz for the time point (i) with Equation (4), where a_1 is $16,139/2^{22}$, a_2 $32,278/2^{22}$, b_1 $15,565/4096$, b_2 $22,279/4096$, b_3 $14,239/4096$, and b_4 $3429/4096$.

$$R_{BP}[i] = a_1 R_{LF12}[i] - a_2 R_{LF12}[i-2] + a_1 R_{LF12}[i-4] + b_1 R_{BP}[i-1] - b_2 R_{BP}[i-2] + b_3 R_{BP}[i-3] - b_4 R_{BP}[i-4] \tag{4}$$

Finally, the duration of a single step time was determined by the zero-crossing method with hysteresis. The hysteresis was used to prevent artefacts or noise caused by false crossings. It involved two threshold levels: one for the rising edge (positive crossing threshold, 0.05 g) and another for the falling edge (negative crossing threshold, −0.05 g). Counting of the single step time started when the bandpass filtered acceleration value, R_{BP}, went from negative to positive and crossed the rising edge threshold, 0.05 g level. The step time calculation could not stop until the signal value went below the falling edge threshold (−0.05 g) before the next positive zero crossing (0.05 g), i.e., step time, can be detected.

The MAD value of the kth step was calculated with Equation (5), where N is the number of samples in the step and j is the time point of the start of the step.

$$MAD_k = \frac{1}{N} \sum_{i=j}^{j+N-1} |r[i] - R_{LF}[i]| \tag{5}$$

The floating MAD value was calculated for two adjacent steps with Equation (6), where $MAD_{f,k}$ is the floating epoch MAD-value of the kth step, MAD_k and MAD_{k-1} are the MAD values for the kth and preceding step, and N_k and N_{k-1} is the number of samples in the kth and preceding step. Thus, for bipedal gait, each epoch contained successive left and right footsteps or vice versa. As the floating epoch had one step overlap, each step was analyzed twice.

$$MAD_{f,k} = \frac{MAD_{k-1} N_{k-1} + MAD_k N_k}{N_{k-1} + N_k} \tag{6}$$

The change in the MAD values between adjacent epochs ($dMAD$) was calculated with Equation (7), where $MAD_{f,k}$ and $MAD_{f,k-1}$ are the MAD values of the two adjacent floating epochs windows.

$$dMAD_{f,k} = \left| MAD_{f,k} - MAD_{f,k-1} \right| \tag{7}$$

The step frequency, f_s, for the kth epoch was calculated with Equation (8), where 100 Hz is the sampling frequency of the accelerometer.

$$f_{s,k} = 2 \frac{100 \text{ Hz}}{N_{k-1} + N_k} \tag{8}$$

Figure 3a illustrates the calculation of the *MAD* values for single steps during a 6 s period of the CS test. The calculation of the stepwise *MAD* values is based on the use of Equations (1)–(5). Figure 3b illustrates the floating epoch *MAD* values, 6 s epoch *MAD* values, *dMAD* values and stride frequency during a 60 s period of the AC/DC test. The floating epoch *MAD* values can capture the turns during the AC/DC test in contrast to the 6 s epoch. The values were calculated using Equations (6)–(8).
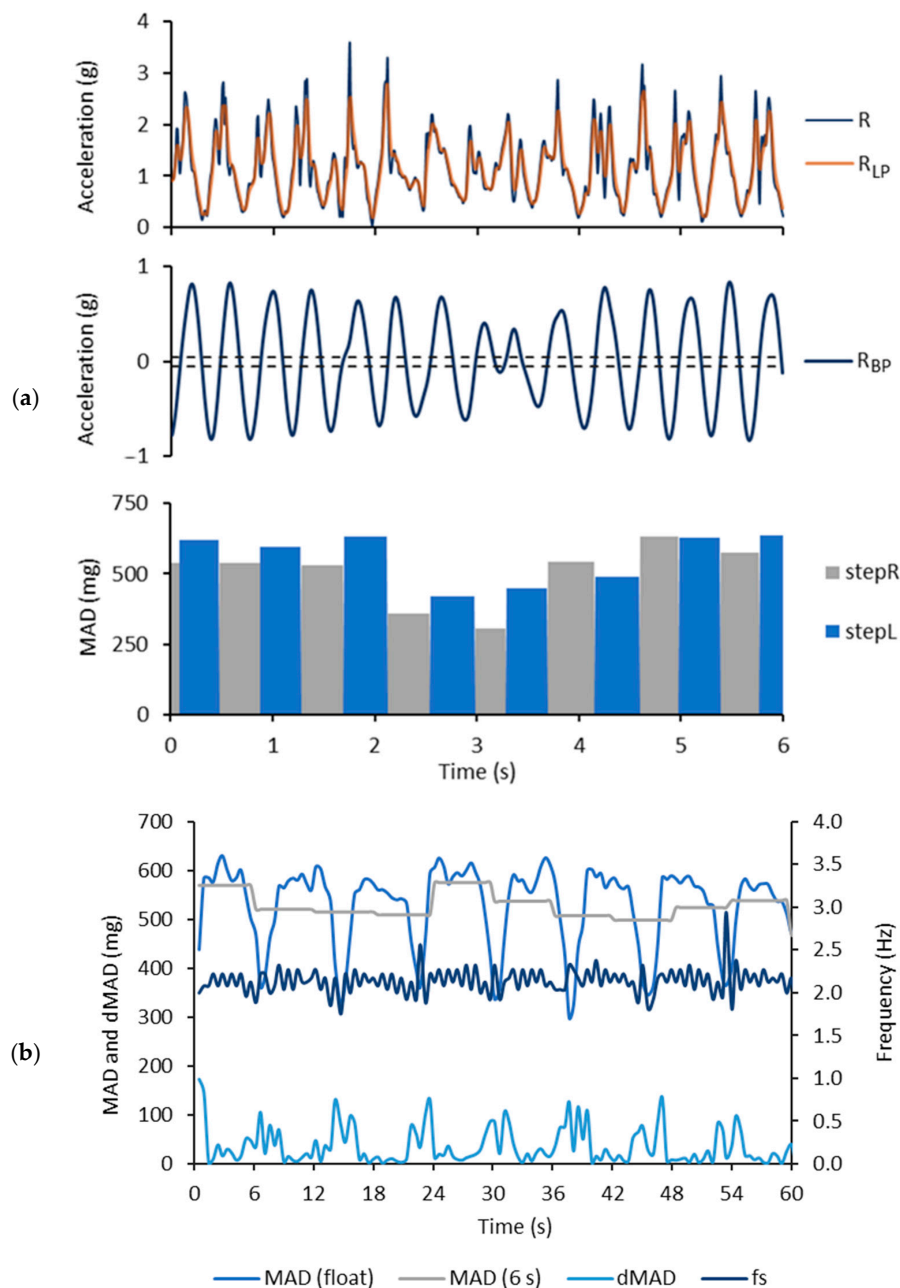

Figure 3. Example of measured and calculated parameter values during the AC/DC test: (**a**) The upper row shows the magnitude of the acceleration (*R*) and low-pass filtered acceleration (R_{LP}) during the 6 s time containing one turn. The middle row shows the band-pass filtered acceleration (R_{BP}). The dashed lines show the rising and falling edge thresholds for detecting the positive and negative crossings. The bottom row shows the mean amplitude deviation (*MAD*) value of each step. (**b**) The curves show the floating and 6 s epoch *MAD* values, change in the *MAD*-values between two adjacent floating epochs (*dMAD*) and step frequency (f_s) during a 60 s time.

2.4. Statistical Analysis

The VO_2 prediction equations based on the constant 6 s epochs were determined from the data of the CS test and AC/DC test and validated with the AC/DC test and CS test, respectively. Similarly, the VO_2 prediction equations based on the floating epochs were determined using both data sets as training and validation sets. Mean *MAD* values for the 6 s and floating epochs, *dMAD*, and f_s were calculated for all stages of the CS test and for the entire AC/DC test.

Generalized linear models were employed to create prediction models for estimating VO_2. Incident VO_2 was the dependent variable, while the coincident *MAD* values based on the 6 s epochs or coincident *MAD* values, *dMAD* value, and f_s based on the floating epoch served as the independent predictor variables. Coefficient of determination (R^2), standard error of estimate (SEE), and mean absolute percent error (MAPE) were used to assess the predictive error of the prediction models. All statistical analyses were conducted using the statistics software (IBM SPSS Statistics for Windows, Version 29.0, Armonk, NY, USA)

3. Results

The participants' characteristics are presented in Table 1. There were more participants in the AC/DC test than in the CS test. The proportion of males was somewhat higher in the AC/DC test while the number of males and females was almost equal in the CS test. The participants of the AC/DC test were on average 15 years older and 12 kg heavier than those of the CS test.

Table 1. Number of participants (N) and their background characteristics, mean (SD, range).

	AC/DC Test	**CS Test**
N	62	29
Male (%)	35 (56.5)	14 (48.3)
Female (%)	27 (43.5)	15 (51.7)
Age (years)	50.1 (12.5, 19.0–68.0)	36.6 (9.7, 20.0–59.0)
Height (cm)	172.9 (7.9, 154.0–191.0)	172.0 (10.0, 156.0–196.0)
Weight (kg)	82.6 (19.7, 52.4–146.8)	70.8 (12.7, 49.8–100.0)
BMI (kg/m^2)	27.5 (5.4, 19.3–41.5)	23.5 (2.3, 19.3–29.6)

The VO_2 values ranged from 8.7 to 30.2 mL/kg/min in the CS test and from 12.4 to 41.1 mL/kg/min in the AC/DC test. The mean values of the 6 s epoch *MAD* ranged from 51 to 660 mg, the floating *MAD* from 56 to 660 mg, *dMAD* from 5 to 27 mg, and step frequency from 0.54 to 1.20 Hz for the CS test. For the AC/DC test, the mean values of the 6 s epoch *MAD* ranged from 280 to 662 mg, the floating *MAD* from 281 to 662 mg, *dMAD* from 20 to 63 mg, and step frequency from 0.89 to 1.37 Hz. The Pearson correlation coefficient between the 6 s epoch *MAD* and the floating *MAD* was 0.9998 for both the CS and AC/DC tests.

Figure 4 illustrates the relationships between the measured VO_2 and calculated parameters. The *MAD* values of the AC/DC test (6 s and floating epoch) tend to be smaller than the values of the CS test for VO_2 values greater than 20 mL/kg/min. The *dMAD* values of the AC/DC test were higher than the values of the CS test.

Table 2 presents the VO_2 prediction equations derived from the 6 s epoch data. Notably, the equation developed using the CS test demonstrated high predictive accuracy when applied to the CS test data but performed poorly when applied to the AC/DC test data. Conversely, the equation generated from the AC/DC test exhibited high predictive accuracy for the CS test data but performed poorly when applied to the AC/DC test data.
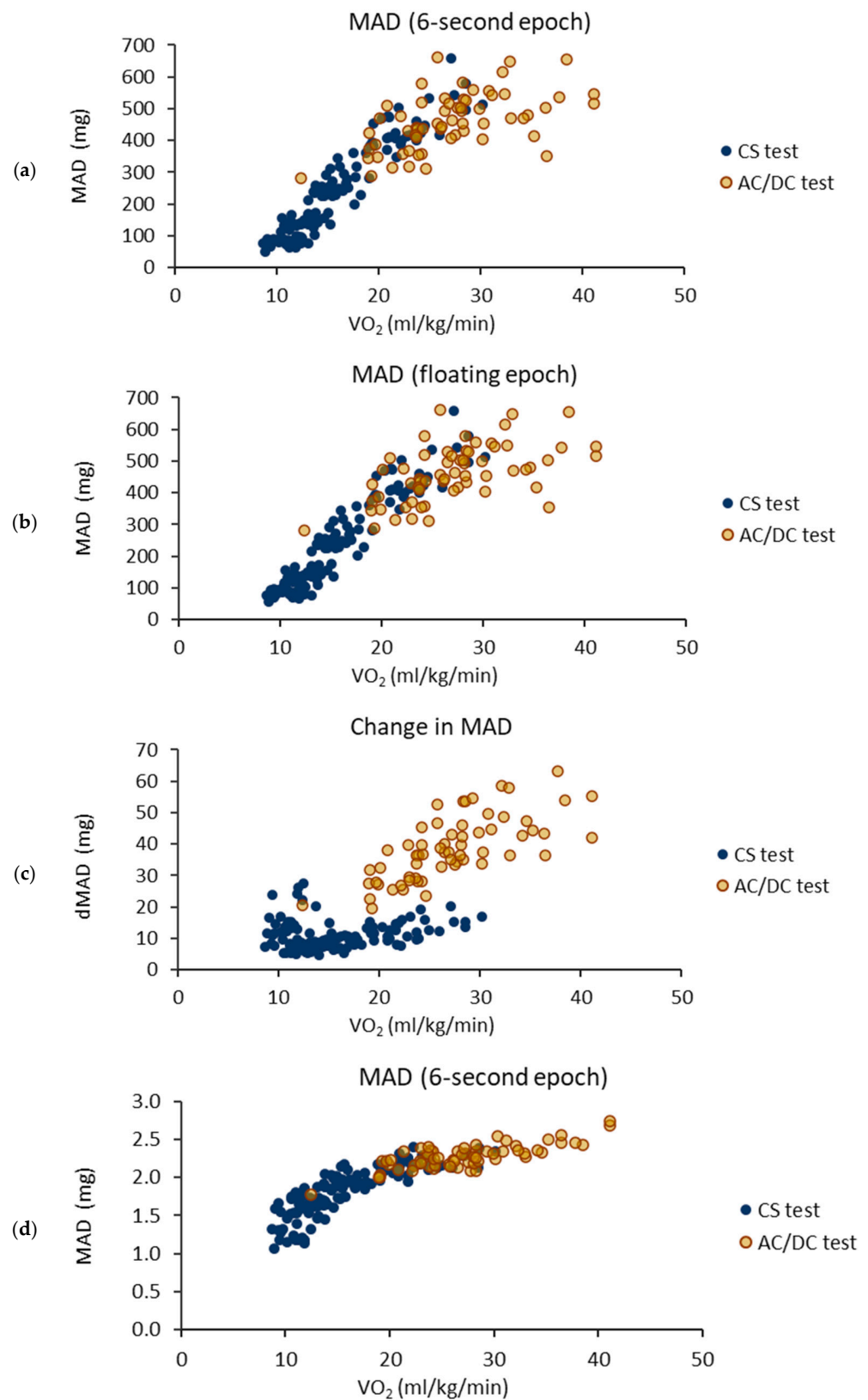
Figure 4. Mean values of measured accelerometer parameters for each participant as a function of coincident oxygen consumption (VO_2) frequency during the acceleration and deceleration (AC/DC) test and constant speed (CS) test. (**a**) Mean amplitude deviation (*MAD*) for the 6 s epochs; (**b**) *MAD* for the floating epochs; (**c**) absolute change in *MAD* between the adjacent epochs (*dMAD*); and (**d**) step frequency.

Table 2. Coefficients and performance of VO$_2$ (mL/kg/min) prediction models based on 6 s epochs.

Training Set		CS Test	AC/DC Test
	Constant (*p*-value)	7.929 (<0.001)	10.379 (0.002)
	MAD (*p*-value)	0.033 (<0.001)	0.036 (<0.001)
	Bias	0	0
	SEE	1.68	4.81
	R^2	88.8%	31.0%
	MAPE	8.7%	13.6%
Validation Set		**AC/DC Test**	**CS Test**
	Bias	4.13	−3.32
	SEE	4.82	1.76
	R^2	31.0%	88.8%
	MAPE	16.9%	23.2%

CS test = constant-speed walking test, AC/DC test = acceleration and deceleration walking test, MAD = mean amplitude deviation, SEE = standard error of estimate, R^2 = coefficient of determination, MAPE = mean absolute percentage error.

Table 3 presents the VO$_2$ prediction equations for the floating epoch. The prediction was based on the *MAD* values, *dMAD* values and the exponent of the f_s. Notably, the equation developed using the CS test demonstrated excellent predictive accuracy for the CS test but performed poorly when applied to the AC/DC test. In contrast, the equation derived from the AC/DC test demonstrated high predictive accuracy for both the AC/DC and CS tests.

Table 3. Coefficients and performance of VO$_2$ (ml/kg/min) prediction models based on the floating epoch.

Training Set		CS Test	AC/DC Test
	Constant (*p*-value)	7.186 (<0.001)	−3.160 (0.295)
	MAD (*p*-value)	0.033 (<0.001)	0.005 (0.585)
	dMAD (*p*-value)	0.068 (0.049)	0.218 (0.008)
	Exp(f_s) (*p*-value)	−0.004 (0.986)	2.004 (<0.001)
	Bias	0.0	0.0
	SEE	1.64	2.9
	R^2	89.3%	74.9%
	MAPE	8.4%	9.1%
Validation Set		**AC/DC Test**	**CS Test**
	Bias	2.29	2.56
	SEE	4.62	2.37
	R^2	36.4%	78.2%
	MAPE	13.1%	19.6%

CS test = constant-speed walking test, AC/DC test = acceleration and deceleration walking test, *MAD* = mean amplitude deviation, *dMAD* = absolute change of MAD value between adjacent epochs, Exp = exponent, f_s = step frequency, SEE = standard error of estimate, R^2 = coefficient of determination, MAPE = mean absolute percentage error.

Figure 5 illustrates the performance of the VO$_2$ prediction models trained with the 6 s epochs and the CS test. The model underestimates over 25 mL/kg/min VO$_2$ values for the AC/DC test. The Bland–Altman difference plot further shows that the limit of agreement is wider for the AC/DC test, indicating a greater variability in the predicted values. In addition, the Bland–Altman plot depicts an apparent trend in the AC/DC test results, as the difference between measured and predicted VO$_2$ values tends to increase with increasing mean VO$_2$ values.
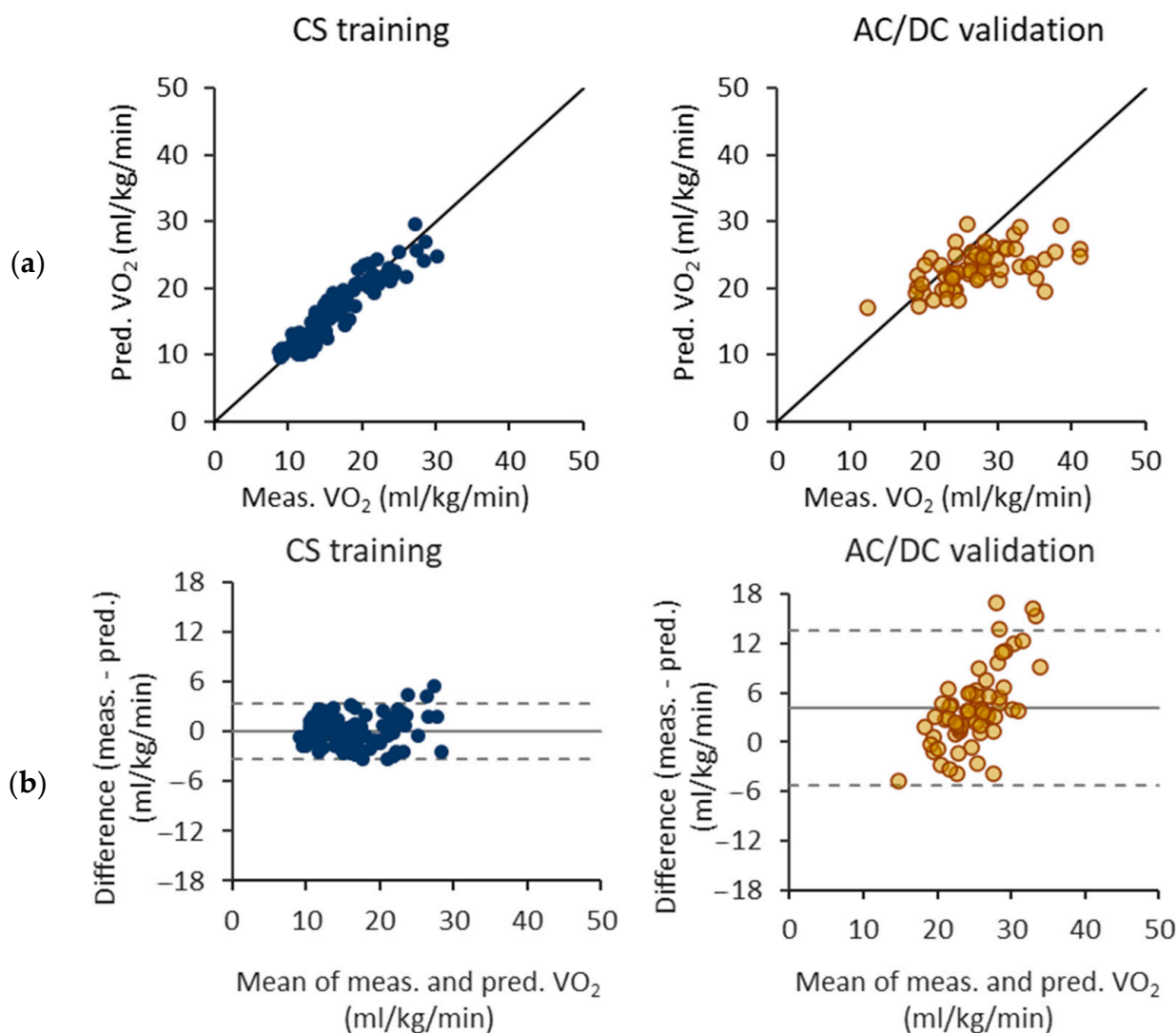
Figure 5. Performance of oxygen consumption (VO_2) prediction equations trained with the 6 s epochs and constant speed (CS) test. (**a**) Measured and predicted VO_2 for the CS test as a training set and the acceleration/deceleration (AC/DC) test as a validation set. (**b**) Respective Bland–Altman difference plots of measured and predicted VO_2 values The solid line represents the mean difference and the dashed lines the 95% limits of agreement.

Figure 6 illustrates the performance of the VO_2 prediction model trained with the 6 s epoch and the AC/DC test. Although the results of the AC/DC test show no bias, there is an evident trend in the Bland–Altman plot. The difference between the measured and predicted VO_2 tends to become larger as the mean VO_2 increases. The CS test has a narrower limit of agreement in the Bland–Altman plot than the AC/DC test, but the model overestimates VO_2 for the CS test.

Figure 7 illustrates the performance of the VO_2 prediction equations trained with the floating epochs and the CS test. Notably, the model consistently underestimates VO_2 values exceeding 25 mL/kg/min for the AC/DC test. The Bland–Altman difference plot further reveals that the AC/DC test exhibits a wider limit of agreement compared to the CS test, signifying a greater variability in predictions. Additionally, this plot demonstrates a clear trend in the AC/DC test results, with the difference between measured and predicted VO_2 values increasing as the mean VO_2 increases.
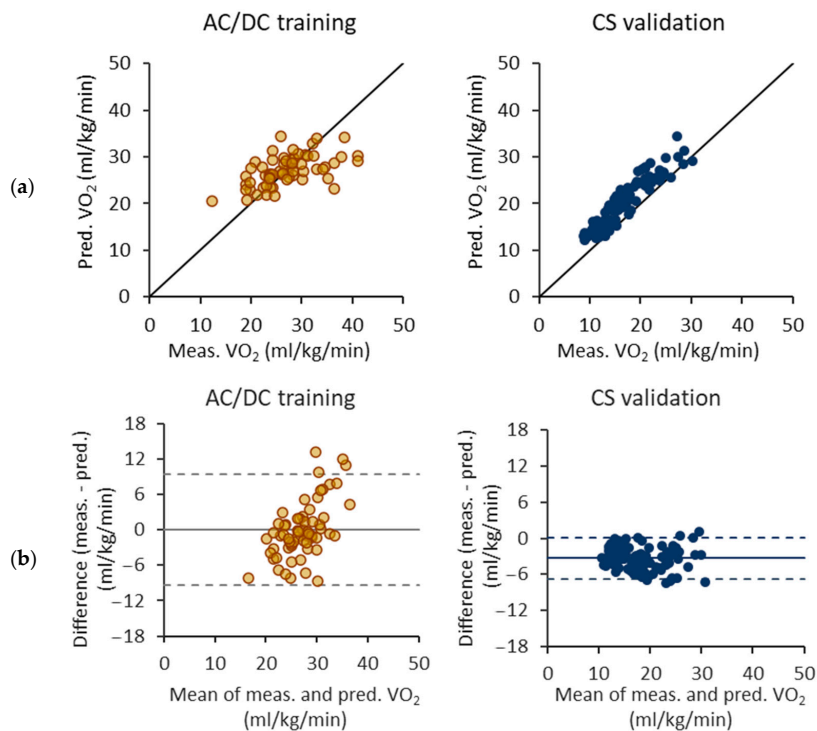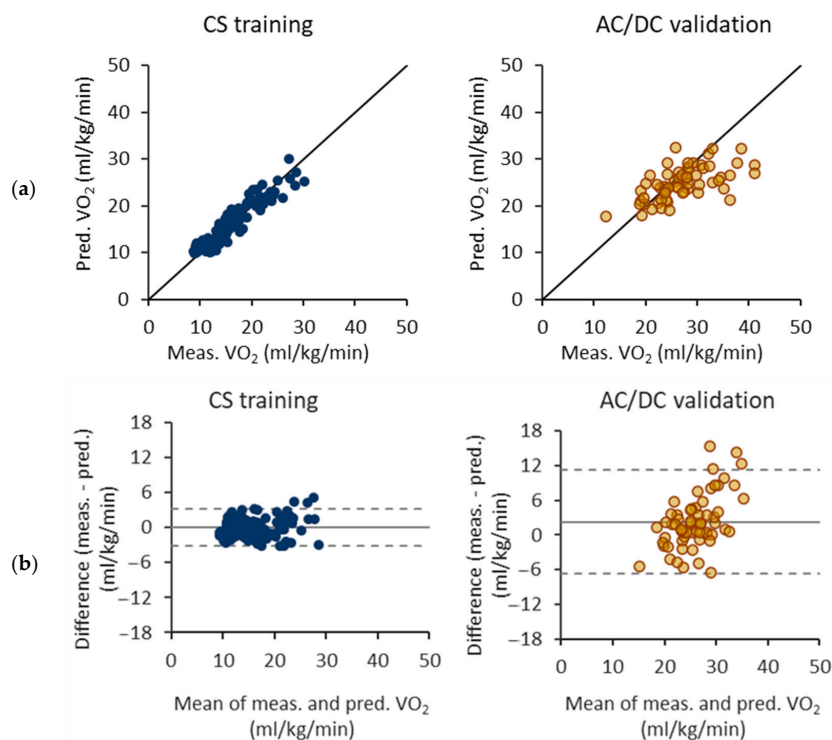
Figure 6. Performance of oxygen consumption (VO$_2$) prediction equations trained with the 6 s epochs and acceleration/deceleration (AC/DC) test. (**a**) Measured and predicted VO$_2$ for AC/DC test as a training set and constant speed (CS) test as a validation set. (**b**) Respective Bland–Altman difference plots of measured and predicted VO$_2$ values. The solid line represents the mean difference and the dashed lines the 95% limits of agreement.

Figure 7. Performance of oxygen consumption (VO$_2$) prediction equations with floating epochs and constant speed (CS) test. (**a**) Measured and predicted VO$_2$ for CS test as a training set and AC/DC test as a validation set. (**b**) Respective Bland–Altman difference plot of measured and predicted VO$_2$. The solid line represents the mean difference and the dashed lines the 95% limits of agreement.

Figure 8 illustrates the performance of the VO$_2$ prediction model trained using the floating epochs and the AC/DC test. Notably, the VO$_2$ values are consistently underestimated for the CS test. The limits of agreement are similar for both the AC/DC test and CS test, suggesting a comparable predictive performance. Unlike the other models, the Bland–Altman difference plot for the AC/DC test displays no visible trend, indicating more stable results over a broad range of VO$_2$.
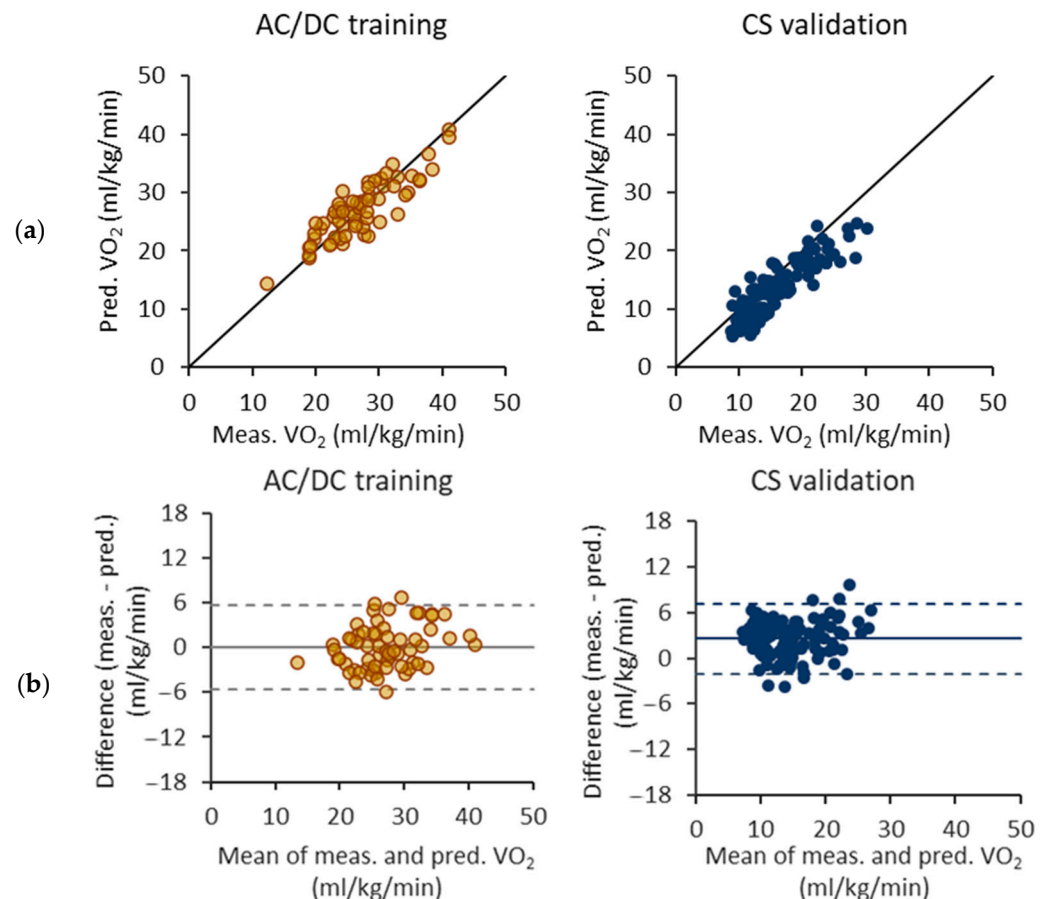

Figure 8. Performance of oxygen consumption (VO$_2$) prediction equations trained with the floating epochs and acceleration/deceleration (AC/DC) test. (**a**) Measured and predicted VO$_2$ for AC/DC test as a training set and constant speed (CS) test as a validation set. (**b**) Respective Bland–Altman difference plots of measured and predicted VO$_2$ values. The solid line represents the mean difference and the dashed lines the 95% limits of agreement.

4. Discussion

The VO$_2$ prediction equation using the floating epoch and developed with the unsteady activity performed adequately with steady activity but not vice versa. The present findings support the notion that VO$_2$ estimation models based on steady activities may contribute to the poor accuracy of VO$_2$ estimations during daily activities [7]. Habitual activities comprise inherently sporadic and irregular patterns compared to steadily paced patterns occurring in walking or running for exercise. Therefore, the assessment of free-living PA could be more reliable with methods developed using irregular activities. Both models trained with the CS test data performed well on the training data (CS test) but poorly on the validation data (AC/DC test). This means that the prediction models based on steady activities are compromised by generalization problems. This kind of model may perform well on the training data but fails to transfer its predictive ability to new, unseen data [24,25]. This lack of generalization is likely to lead to poor performance in real-world measurements [26].

In the present study, the CS and AC/DC tests had different ranges of the measured VO_2 values. The participants of the CS test walked at different speeds, while the participants of the AC/DC test were asked to walk as fast as possible. Thus, the AC/DC test did not contain low VO_2 values typical for light PA and most daily habitual activities [27]. An increased range of walking speeds during the AC/DC test would have likely resulted in a better performance for the floating epoch model. In its present form, the models trained by the AC/DC test yielded narrow limits of agreement for the CS test with both the 6 s and floating epochs. However, both models, the 6 s and the floating epoch, produced slightly biased estimates of VO_2.

The floating epoch length model improved the VO_2 estimation by detecting the variations between steps due to braking and acceleration. Previously, shorter epoch lengths have been found to be more accurate for capturing the actual PA levels during intermittent PA than longer epochs [11]. Like the short epoch length, the floating epoch model can capture short bursts of PA. In addition, when calculating the change in the epoch-wise accelerometer output values, it became possible to estimate the extra VO_2 due to braking and propulsive forces. Noteworthy, the floating epoch length model also had good performance for the steady activity of the CS test, where the change in the accelerometer output between the adjacent epochs was small. The different influence of the floating epoch and fixed epoch lengths on the predicted VO_2 from a standing position to walking is illustrated in Figure 9. The predicted VO_2 reached the highest values during the acceleration phase of walking with the floating epoch model, but during the steady walking pace, both models showed a comparable prediction of VO_2.
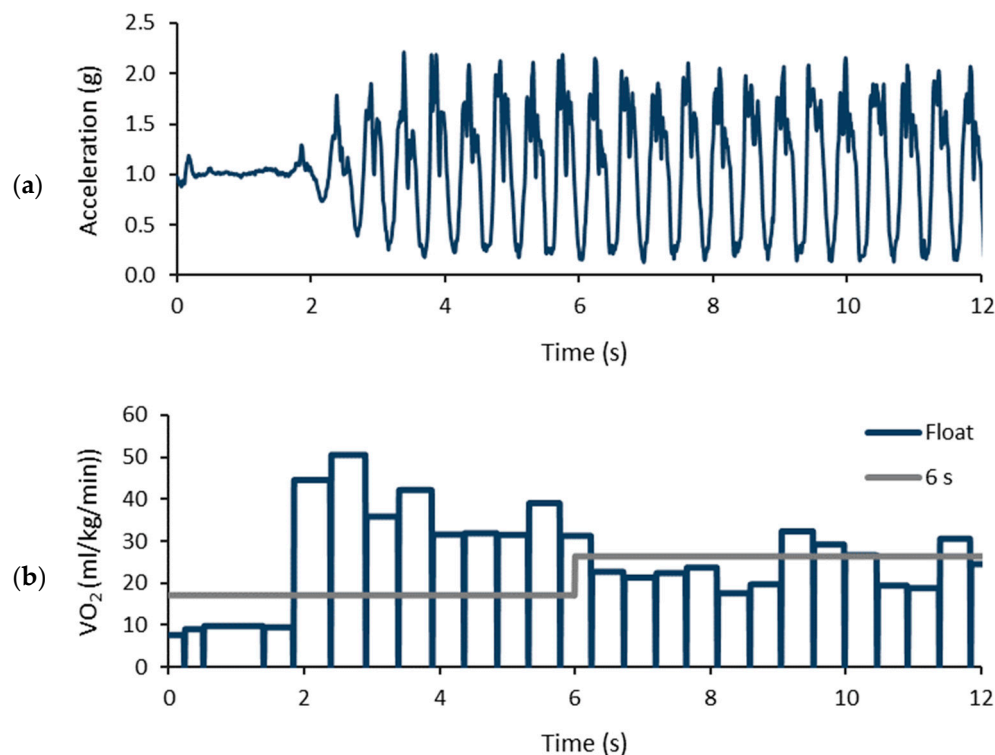

Figure 9. Illustration of a 12 s period containing a change in activity from standing to walking: (**a**) the magnitude of the measured acceleration; and (**b**) predicted oxygen consumptions using the floating epoch and 6 s epochs.

The fixed epoch length can result in the over-smoothing of acceleration data and lead to a loss of potentially important details and features from the data [28]. In this study, the fixed, 6 s epoch length missed the changes in the *MAD* levels during braking and accelerating, which yielded underestimated VO_2 values from 25 mL/kg/min upwards for the AC/DC test (Figures 3–5). The largest difference between the measured and estimated

VO_2 value for the AC/DC test with the 6 s epoch was 17 mL/kg/min, when the CS test was used as the training set, and 13 mL/kg/min, when the AC/DC test was used as the training set. Correspondingly, the largest difference with the floating epoch was 15 mL/kg/min, when the CS test was used as the training set, and 7 mL/kg/min, when the AC/DC test was used as the training set.

Bipedal gait involves continuous fluctuations in energetics, differing from wheeled vehicles, like bicycles [29]. The increases in kinetic energy due to acceleration require extra positive mechanical work and the same applies to deceleration, where negative mechanical work is required to reduce the kinetic energy. The absence of energy fluctuations at the constant speed makes any speed changes energetically very expensive for wheeled vehicles, but on the contrary, for bipedal gait, limited alterations in the speed can be performed without a remarkable increase in metabolic energy consumption [29]. In this study, the *MAD* and VO_2 values were at the same level for the AC/DC and CS tests up to the approximately 20 mL/kg/min level in terms of the VO_2 (Figure 2). For higher intensities, the VO_2 values were higher for the AC/DC test than for the CS test with the same *MAD* values. Therefore, the greatest benefits of the floating epoch model can be achieved with intensities over 20 mL/kg/min corresponding primarily to vigorously performed PA.

The limitation of this study is that the floating epoch model was tested only with walking and the speed range was narrow in the AC/DC test. The floating epoch model should also be tested with other than bipedal physical activities, like swimming, cycling, skiing, and gym training, and in transportation activities, like vehicle driving. The other activities will likely have an influence on the selection of appropriate filter characteristics. The filter type and cutoff frequency determine the level of smoothing applied to the raw accelerometer data and directly impact on the detection of the movement cycle patterns [30,31]. An excessively aggressive filter may result in over-smoothing, obscuring important high-frequency features and leading to missed or distorted cycle events, while an insufficiently smoothing filter can allow noise and fine-scale variations in the raw data to interfere with cycle detection, potentially leading to inaccuracies in movement detection and the classification of movements [32]. Therefore, the judicious choice of the filter is pivotal in achieving a precise and robust cycle detection, striking a balance between noise reduction and the preservation of essential information of coincident movements.

In free-living measurements, the floating epoch model can capture the pulsatile and sporadic activity behaviors. With fixed and longer, over 15 s epochs, the intensity of these kinds of movements becomes underestimated or can even be undetected and classified as stationary time [33,34]. On the other hand, a short interruption in PA can be picked up as a stationary behavior with shorter epoch lengths [34]. Therefore, the floating epoch will contribute to the accumulation of both stationary time and time spent in light, moderate, and vigorous PA. It is likely that the accumulated time of light PA is reallocated to other activities, if the accumulated time is calculated using the floating epochs instead of the fixed epoch length.

When employing floating epochs, the use of an analyzing window with a size that aligns with specific parameters of interest in coincident PA might be a reasonable approach. The time-course of physiological responses to acute PA depends on the various factors, including the type and duration of activity and individual fitness level [35]. The analyzing window size can modulate data smoothing, a long window size smooths data more than a short window size [36]. The longer analyzing window size is likely more suitable for responses with a longer time-course while it removes effectively short bursts of PA in the middle of stationary behavior or short breaks in the middle of a long steady PA period. Similarly, a short window size suits responses with a shorter time-course while it can capture shorts bursts of activity or breaks. A short window size can also improve the type of human activity recognition as the number of windows containing multiple activities can be minimized [37,38]. The optimal size of the analyzing window, i.e., the level of smoothing, depends on the goals of the PA measurement and analysis, and the trade-off between noise reduction and preserving meaningful information. These topics call for further research.

5. Conclusions

The floating epoch model can accurately capture the intensity of the short bursts of PA effectively and may thus enhance the performance of the laboratory-trained prediction models for VO_2 in free-living conditions. This model shows a heightened accuracy in predicting spontaneous and sporadic PA, which is advantageous for research focused on child and adolescent PA or activities associated with occupational tasks. The assessment of PA in population studies may also benefit from this approach. Furthermore, its utility may extend to exercise interventions characterized by short, high-intensity intervals, establishing the floating epoch model as a valuable tool for advancing our understanding and application of PA-related interventions.

References

1. Bull, F.C.; Al-Ansari, S.S.; Biddle, S.; Borodulin, K.; Buman, M.P.; Cardon, G.; Carty, C.; Chaput, J.-P.; Chastin, S.; Chou, R.; et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br. J. Sports Med.* **2020**, *54*, 1451–1462. [CrossRef] [PubMed]
2. Dowd, K.P.; Szeklicki, R.; Minett, M.A.; Murphy, M.H.; Polito, A.; Chigo, E.; van der Ploeg, H.; Ekelund, U.; Maciaszek, J.; Stemplewski, R.; et al. A systematic literature review of reviews on techniques for physical activity measurement in adults: A DEDIPAC study. *Int. J. Behav. Nutr. Phys. Act.* **2018**, *15*, 15. [CrossRef] [PubMed]
3. Sattler, M.C.; Ainsworth, B.E.; Andersen, L.B.; Foster, C.; Hagströmer, M.; Jaunig, J.; Kelly, P.; Kohl, H.W., III; Matthews, C.E.; Oja, P.; et al. Physical activity self-reports: Past or future? *Br. J. Sports Med.* **2021**, *55*, 889–891. [CrossRef] [PubMed]
4. Burchartz, A.; Anedda, B.; Auerswald, T.; Giurgiu, M.; Hill, H.; Ketelhut, S.I.; Kolb, S.; Mall, C.; Manz, K.; Nigg, C.R.; et al. Assessing physical behavior through accelerometry–state of the science, best practices and future directions. *Psychol. Sport Exerc.* **2020**, *49*, 101703. [CrossRef]
5. Fabre, N.; Lhuisset, L.; Bois, J. Accelerometer-Based Physical Activity Assessment During Intermittent Conditions: Effect of Epoch Length on Energy Expenditure Estimate. *Res. Q. Exerc. Sport* **2023**, *94*, 202–209. [CrossRef] [PubMed]
6. Keadle, S.K.; Lyden, K.A.; Strath, S.J.; Staudenmayer, J.W.; Freedson, P.S. A framework to evaluate devices that assess physical behavior. *Exerc. Sport Sci. Rev.* **2019**, *47*, 206–214. [CrossRef] [PubMed]
7. Crossley, S.G.; Mackintosh, K.A.; Wilson, R.P.; Lester, L.J.; Griffiths, I.W.; McNarry, M.A. Energy expenditure associated with walking speed and angle of turn in children. *Eur. J. Appl. Physiol.* **2018**, *118*, 2563–2576. [CrossRef] [PubMed]
8. Stevens, T.G.; De Ruiter, C.J.; Van Maurik, D.; Van Lierop, C.J.; Savelsbergh, G.J.; Beek, P.J. Measured and estimated energy cost of constant and shuttle running in soccer players. *Med. Sci. Sports Exerc.* **2015**, *47*, 1219–1224. [CrossRef]
9. Zamparo, P.; Zadro, I.; Lazzer, S.; Beato, M.; Sepulcri, L. Energetics of shuttle runs: The effects of distance and change of direction. *Int. J. Sports Physiol. Perf.* **2014**, *9*, 1033–1039. [CrossRef]
10. Hatamoto, Y.; Yamada, Y.; Sagayama, H.; Higaki, Y.; Kiyonaga, A.; Tanaka, H. The relationship between running velocity and the energy cost of turning during running. *PLoS ONE* **2014**, *9*, e81850. [CrossRef]

11. Fabre, N.; Lhuisset, L.; Bernal, C.; Bois, J. Effect of epoch length on intensity classification and on accuracy of measurement under controlled conditions on treadmill: Towards a better understanding of accelerometer measurement. *PLoS ONE* **2020**, *15*, e0227740. [CrossRef] [PubMed]

12. Vähä-Ypyä, H.; Sievänen, H.; Husu, P.; Tokola, K.; Vasankari, T. Intensity paradox—Low-fit people are physically most active in terms of their fitness. *Sensors* **2021**, *21*, 2063. [CrossRef] [PubMed]

13. McNarry, M.A.; Wilson, R.P.; Holton, M.D.; Griffiths, I.W.; Mackintosh, K.A. Investigating the relationship between energy expenditure, walking speed and angle of turning in humans. *PLoS ONE* **2017**, *12*, e0182333. [CrossRef] [PubMed]

14. Oxendale, C.L.; Highton, J.; Twist, C. Energy expenditure, metabolic power and high speed activity during linear and multi-directional running. *J. Sci. Med. Sport* **2017**, *20*, 957–961. [CrossRef] [PubMed]

15. Allseits, E.K.; Agrawal, V.; Prasad, A.; Bennett, C.; Kim, K.J. Characterizing the impact of sampling rate and filter design on the morphology of lower limb angular velocities. *IEEE Sens. J.* **2019**, *19*, 4115–4122. [CrossRef]

16. Wundersitz, D.W.; Gastin, P.B.; Robertson, S.; Davey, P.C.; Netto, K.J. Validation of a trunk-mounted accelerometer to measure peak impacts during team sport movements. *Int. J. Sports Med.* **2015**, *36*, 742–746. [CrossRef] [PubMed]

17. Fridolfsson, J.; Börjesson, M.; Arvidsson, D. A biomechanical re-examination of physical activity measurement with accelerometers. *Sensors* **2018**, *18*, 3399. [CrossRef]

18. Beneke, R.; Taylor, M.J. What gives Bolt the edge—AV Hill knew it already! *J. Biomech.* **2010**, *43*, 2241–2243. [CrossRef]

19. Vähä-Ypyä, H.; Vasankari, T.; Husu, P.; Suni, J.; Sievänen, H. A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. *Clin. Physiol. Funct. Imaging* **2015**, *35*, 64–70. [CrossRef]

20. Mason, R.; Pearson, L.T.; Barry, G.; Young, F.; Lennon, O.; Godfrey, A.; Stuart, S. Wearables for running gait analysis: A systematic review. *Sports Med.* **2023**, *53*, 241–268. [CrossRef]

21. Vähä-Ypyä, H.; Vasankari, T.; Husu, P.; Mänttäri, A.; Vuorimaa, T.; Suni, J.; Sievänen, H. Validation of cut-points for evaluating the intensity of physical activity with accelerometry-based mean amplitude deviation (MAD). *PLoS ONE* **2015**, *10*, e0134813. [CrossRef] [PubMed]

22. Mänttäri, A.; Suni, J.; Sievänen, H.; Husu, P.; Vähä-Ypyä, H.; Valkeinen, H.; Tokola, K.; Vasankari, T. Six-minute walk test: A tool for predicting maximal aerobic power (VO$_2$ max) in healthy adults. *Clin. Physiol. Funct. Imaging* **2018**, *38*, 1038–1045. [CrossRef] [PubMed]

23. López-Nava, I.H.; Muñoz-Meléndez, A.; Pérez Sanpablo, A.I.; Alessi Montero, A.; Quiñones-Urióstegui, I.; Núñez Carrera, L. Estimation of temporal gait parameters using Bayesian models on acceleration signals. *Comp. Meth. Biomech. Biomed. Eng.* **2016**, *19*, 396–403. [CrossRef] [PubMed]

24. O'Driscoll, R.; Turicchi, J.; Hopkins, M.; Duarte, C.; Horgan, G.W.; Finlayson, G.; Stubbs, R.J. Comparison of the validity and generalizability of machine learning algorithms for the prediction of energy expenditure: Validation study. *JMIR mHealth uHealth* **2021**, *9*, e23938. [CrossRef] [PubMed]

25. Wang, J.; Lan, C.; Liu, C.; Ouyang, Y.; Qin, T.; Lu, W.; Chen, Y.; Zeng, W.; Yu, P.S. Generalizing to unseen domains: A survey on domain generalization. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 8052–8071. [CrossRef]

26. Montoye, A.H.; Westgate, B.S.; Clevenger, K.A.; Pfeiffer, K.A.; Vondrasek, J.D.; Fonley, M.R.; Bock, J.M.; Kaminsky, L.A. Individual versus Group Calibration of Machine Learning Models for Physical Activity Assessment Using Body-Worn Accelerometers. *Med. Sci. Sports Exerc.* **2021**, *53*, 2691–2701. [CrossRef] [PubMed]

27. Husu, P.; Tokola, K.; Vähä-Ypyä, H.; Sievänen, H.; Suni, J.; Heinonen, O.J.; Heiskanen, J.; Kaikkonen, K.M.; Savonen, K.; Kokko, S.; et al. Physical activity, sedentary behavior, and time in bed among Finnish adults measured 24/7 by triaxial accelerometry. *J. Meas. Phys. Behav.* **2021**, *4*, 163–173. [CrossRef]

28. Brailey, G.; Metcalf, B.; Price, L.; Cumming, S.; Stiles, V. Raw acceleration from wrist-and hip-worn accelerometers corresponds with mechanical loading in children and adolescents. *Sensors* **2023**, *23*, 6943. [CrossRef]

29. Minetti, A.E.; Gaudino, P.; Seminati, E.; Cazzola, D. The cost of transport of human running is not affected, as in walking, by wide acceleration/deceleration cycles. *J. Appl. Physiol.* **2013**, *114*, 498–503. [CrossRef]

30. Charvátová, H.; Procházka, A.; Vyšata, O.; Suárez-Araujo, C.P.; Smith, J.H. Evaluation of accelerometric and cycling cadence data for motion monitoring. *IEEE Access* **2021**, *9*, 129256–129263. [CrossRef]

31. Rindal, O.M.H.; Seeberg, T.M.; Tjønnås, J.; Haugnes, P.; Sandbakk, Ø. Automatic classification of sub-techniques in classical cross-country skiing using a machine learning algorithm on micro-sensor data. *Sensors* **2017**, *18*, 75. [CrossRef] [PubMed]

32. Fridolfsson, J.; Börjesson, M.; Buck, C.; Ekblom, Ö.; Ekblom-Bak, E.; Hunsberger, M.; Lissner, L.; Arvidsson, D. Effects of frequency filtering on intensity and noise in accelerometer-based physical activity measurements. *Sensors* **2019**, *19*, 2186. [CrossRef] [PubMed]

33. Ahmadi, M.N.; Pavey, T.G.; Trost, S.G. Machine learning models for classifying physical activity in free-living preschool children. *Sensors* **2020**, *20*, 4364. [CrossRef] [PubMed]

34. Altenburg, T.M.; Wang, X.; Van Ekris, E.; Andersen, L.B.; Møller, N.C.; Wedderkopp, N.; Chinapaw, M.J. The consequences of using different epoch lengths on the classification of accelerometer based sedentary behaviour and physical activity. *PLoS ONE* **2021**, *16*, e0254721. [CrossRef] [PubMed]

35. Benítez-Flores, S.; Magallanes, C.A.; Alberton, C.L.; Astorino, T.A. Physiological and psychological responses to three distinct exercise training regimens performed in an outdoor setting: Acute and delayed response. *J. Funct. Morphol. Kinesiol.* **2021**, *6*, 44. [CrossRef]

36. Vähä-Ypyä, H.; Sievänen, H.; Husu, P.; Tokola, K.; Mänttäri, A.; Heinonen, O.J.; Heiskanen, J.; Kaikkonen, K.M.; Savonen, K.; Kokko, S.; et al. How adherence to the updated physical activity guidelines should be assessed with accelerometer? *Eur. J. Publ. Health* **2022**, *32* (Suppl. S1), i50–i55. [CrossRef]

37. Noor, M.H.M.; Salcic, Z.; Kevin, I.; Wang, K. Adaptive sliding window segmentation for physical activity recognition using a single tri-axial accelerometer. *Pervasive Mobile Comput.* **2017**, *38*, 41–59. [CrossRef]

38. Ahmadi, M.N.; Chowdhury, A.; Pavey, T.; Trost, S.G. Laboratory-based and free-living algorithms for energy expenditure estimation in preschool children: A free-living evaluation. *PLoS ONE* **2020**, *15*, e0233229. [CrossRef]