

Desirée Câmara Miraldo

**Open dataset and algorithm based on linear
multiple regression for gait-event estimation with
inertial sensors**

São Bernardo do Campo, SP - Brazil

2019

Desirée Câmara Miraldo

Open dataset and algorithm based on linear multiple regression for gait-event estimation with inertial sensors

Text presented to the Biomedical Engineering Graduate Program of the Federal University of ABC (UFABC), as a requirement to obtain a Master's degree in Biomedical Engineering

Federal University of ABC
Biomedical Engineering Graduate Program

Supervisor: Prof. Dr. Marcos Duarte
Co-supervisor: Prof. Dr. Renato Naville Watanabe

São Bernardo do Campo, SP - Brazil

2019

Desirée Câmara Miraldo

Open dataset and algorithm based on linear multiple regression for gait-event estimation with inertial sensors/ Desirée Câmara Miraldo. – 2019.

[44](#) p. : il.

Supervisor: Prof. Dr. Marcos Duarte

Co-supervisor: Prof. Dr. Renato Naville Watanabe

Dissertation (Master) – Federal University of ABC

Biomedical Engineering Graduate Program, São Bernardo do Campo, SP - Brazil,
2019.

1. Gait dataset. 2. Gait recognition. 3. Open access. 4. Accelerometers. 5. Gyroscopes. I. Marcos Duarte. II. Renato Naville Watanabe. III. Biomedical Engineering Graduate Program. IV. Title

Este exemplar foi revisado e alterado em relação à versão original, de acordo com as observações levantadas pela banca no dia da defesa, sob responsabilidade única do autor e com anuênciā de seus orientadores.

São Bernardo do Campo, SP - Brazil, ____ de _____ de 2019.

Desirée Câmara Miraldo

Autora

Prof. Dr. Marcos Duarte

Orientador

Prof. Dr. Renato Naville Watanabe

Coorientador

Acknowledgements

This work was supported in part by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP/Brazil) under grants 2014/13247-7 and 2015/14810-0 and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES/Brazil) under grants 1673715.

Abstract

Falls and injuries while walking are recurrent problems for people who suffer from foot drop. When this condition is caused by a lesion in the central nervous system, it is possible to restore part of the gait function by applying functional electrical stimulation during specific gait events. This study had the following goals: to create an open dataset of inertial, magnetic, foot-ground contact, and electromyographic signals from wearable sensors during walking at different speeds; and to develop a multiple regression method to estimate gait events based on the data from this open dataset. Employing wearable sensors, we acquired data from 22 healthy adults and one with foot drop walking at self-selected comfortable, fast, and slow speeds, and standing still. All data are publicly available at Figshare (DOI: [10.6084/m9.figshare.7778255](https://doi.org/10.6084/m9.figshare.7778255)). The novel algorithm we proposed is based on linear multiple regression. The open dataset contains 9,661 gait strides for the healthy subjects and 496 for the subject with the foot drop. The proposed algorithm for estimating the toe-off gait event showed a median accuracy across the healthy subjects and gait speeds of 88.8%, and an accuracy of 97.3% for the affected limb of the subject with foot drop. The open dataset we created will enable researchers to test algorithms for gait-event estimation against a common reference. The algorithm presents comparable performance to other existing algorithms concerning healthy subjects and a promising result based on one subject with foot drop. It is potentially adjustable for real-time application.

Keywords: Gait dataset. Gait recognition. Open access. Accelerometers. Gyroscopes.

Resumo

Quedas e lesões durante a caminhada são problemas recorrentes para pessoas que sofrem de pé equino. Quando esta condição é causada por uma lesão no sistema nervoso central, é possível restaurar parte da função da marcha aplicando estimulação elétrica funcional durante eventos específicos da marcha. Este estudo teve os seguintes objetivos: criar uma base de dados aberta de dados inerciais, magnéticos, eletromiográficos e de contato dos pés durante a caminhada em diferentes velocidades; e desenvolver um método de regressão múltipla para estimar os eventos de marcha com base nos dados deste conjunto de dados abertos. Empregando sensores vestíveis, adquirimos dados de 22 adultos saudáveis e um com pé equino caminhando nas velocidades confortável, rápida e lenta, e parados. Todos os dados estão disponíveis publicamente no Figshare (DOI: [10.6084/m9.figshare.7778255](https://doi.org/10.6084/m9.figshare.7778255)). O novo algoritmo que propusemos é baseado em regressão linear múltipla. O conjunto de dados aberto contém 9.661 passadas para os sujeitos saudáveis e 496 para o sujeito com pé equino. O algoritmo proposto para estimar o evento toe-off da marcha mostrou uma acurácia mediana de 88,8% entre os indivíduos saudáveis e as diferentes velocidades, e uma acurácia de 97,3% para o membro afetado do sujeito com pé equino. O conjunto de dados aberto que criamos permitirá que os pesquisadores testem algoritmos para estimativa de evento de marcha em comparação a uma referência comum. O algoritmo apresenta desempenho comparável a outros algoritmos existentes em relação a indivíduos saudáveis e um resultado promissor baseado em um sujeito com pé equino. É potencialmente ajustável a aplicações em tempo-real.

Palavras-chave: Conjunto de dados de marcha. Reconhecimento de marcha. Acesso aberto. Acelerômetros. Giroscópios.

List of Figures

Figure 1 – Gait events, phases, and sub-phases of a typical walking gait cycle.	13
Figure 2 – Placement of the wearable units on the subject’s legs and feet.	17
Figure 3 – Flowchart of signals preprocessing.	19
Figure 4 – Exemplary data of part of one trial for the measured signals as available in the open dataset.	20
Figure 5 – Flowchart of the algorithm for TO-event estimation.	23
Figure 6 – Violin plots (boxplots plus kernel density estimations) across the 22 healthy subjects from the open dataset for the gait variables: stride duration, support duration, stride length, and speed.	28
Figure 7 – Violin plots (boxplots plus kernel density estimations) for the subject with foot drop from the open dataset for the gait variables: stride duration, support duration, stride length, and speed.	29
Figure 8 – Mean ± 1 standard deviation across the 22 healthy subjects from the open dataset for the measured signals, over the left and right gait cycles walking at a comfortable speed.	30
Figure 9 – Mean ± 1 standard deviation for the subject with foot drop from the open dataset, compared to mean ± 1 standard deviation across the 22 healthy subjects for the measured signals, over the left and right gait cycles walking at a comfortable speed.	31
Figure 10 – Mean ± 1 standard deviation for the subject with foot drop from the open dataset, compared to mean ± 1 standard deviation across the 22 healthy subjects for the measured signals, over the left and right gait cycles walking at a slow speed.	32
Figure 11 – Mean ± 1 standard deviation for the subject with foot drop from the open dataset, compared to mean ± 1 standard deviation across the 22 healthy subjects for the measured signals, over the left and right gait cycles walking at a fast speed.	33
Figure 12 – Exemplary ROC curves of estimator for subjects from the open data set for four speed conditions over both legs: comfortable speed, fast speed, slow speed, and all the three speeds combined.	35
Figure 14 – Results of TO-event estimation for subject ‘s00’ from the open dataset, using data from all speeds for training and testing.	36
Figure 13 – Results of TO-event estimation for subject ‘s03’ from the open dataset, using data from all speeds for training and testing.	37

List of Tables

Table 1 – Main characteristics of studies since 2010 about different algorithms for gait-event estimation using wearable inertial sensors.	15
Table 2 – Sensitivity, specificity and accuracy of the TO-event estimation at different gait speeds.	36

List of abbreviations and acronyms

EMG	Electromyographic signal
FSR	Force-sensitive Resistor
HS	Heel Strike
HO	Heel Off
ROC	Receiver operating characteristic
TA	Tibialis Anterior muscle
TS	Toe Strike
TO	Toe Off

Contents

1	INTRODUCTION	12
1.1	Gait	12
1.2	Methods to measure gait	13
1.3	What has already been done	13
1.4	Goals	14
2	METHODS	16
2.1	Subjects	16
2.2	Data acquisition	16
2.3	Task	17
2.4	Preprocessing	17
2.5	Detection of gait events	21
2.6	Data visualization	21
2.7	Algorithm for gait-event estimation	21
3	RESULTS	26
3.1	Open dataset	26
3.2	Gait events estimation	34
4	DISCUSSION	38
5	CONCLUSION	40
	BIBLIOGRAPHY	41

Preamble (in Portuguese)

Este texto de qualificação é uma versão expandida de um manuscrito submetido para publicação em uma revista científica na área de Engenharia Biomédica. A maior parte da forma e conteúdo do presente texto reflete esta escolha.

1 Introduction

Falls and injuries while walking are recurrent problems for people who suffer from foot drop. This condition refers to weakness of the ankle dorsiflexor muscles, of which the primary one is the tibialis anterior (TA) [1, 2], impairing, for example, the ability to raise the foot and toes to prevent them from hitting the ground during the swing phase of walking. Foot drop can be caused by lesions of specific peripheral nerves or central nervous system lesions such as stroke, cerebral palsy, and multiple sclerosis [1–3]. When the lesion affects the central nervous system, the electrical excitability of the associated peripheral nerves is likely preserved [3], so functional electrical stimulation (FES) may be used to restore adequate movement patterns among people who suffer from foot drop. A person with foot drop would wear a portable FES device that would electrically stimulate the tibialis anterior muscle just before the expected swing phase of the affected inferior limb, evoking a flexion of the ankle and foot during walking. In this scenario, it is fundamental to correctly estimate the moment at which to trigger the stimulus to the muscle (in this case, the instant when the foot should leave the ground under normal conditions).

1.1 Gait

Gait is the term used to describe the pattern of movement with alternating load across the limbs during locomotion over a solid substrate, and for humans gait is typically bipedal and resumes to walking and running [4]. The gross movement patterns of a healthy person's gait are cyclic, so the sequence of events that describes the gait is repeated after a certain period [4]. The movement patterns of a walking gait can be described in more detail when divided into events and phases, as illustrated in Figure 1. A complete normal gait cycle (or stride) begins and ends with the same event, usually the initial ground contact of the leading limb, a *heel strike* (HS) [5]. Then the leading limb takes over the body weight by placing the whole foot on the ground, in a sub-phase called *loading response*. The moment when the toes touch the ground is the *toe strike* (TS). Next, in the *mid-stance* sub-phase, the body is moved forward while the opposite limb is in the swing phase. The heel loses ground contact, the *heel off* (HO) event, and the body continues to be propelled forward until the *pre-swing* sub-phase starts. Still due to this propulsion, the *toe off* (TO) event occurs when the toes leave the ground, starting the *swing phase*. During the swing phase the swinging limb is accelerated forward, then it passes the opposite limb (*mid-swing* sub-phase) and is decelerated until the next HS event, which will finish the swing phase. The phase from the HS event until the TO event of the same lower limb is the *support phase* for that lower limb. A stride starting and ending with the left heel strike (LHS) consists of one right step followed by one left step; the inverse occurs for a gait stride starting and ending with the right heel strike (RHS).

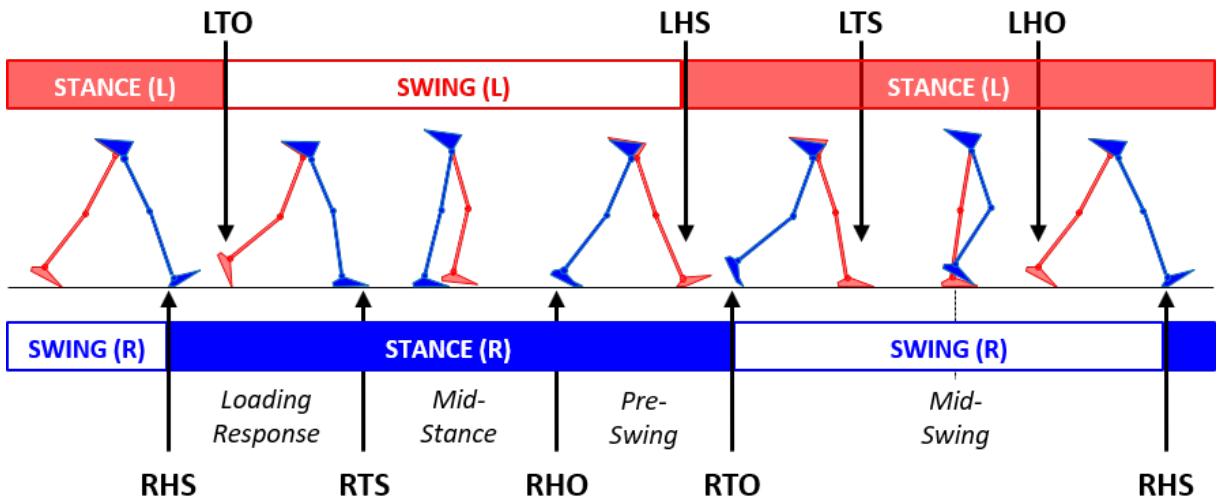


Figure 1 – Gait events, phases, and sub-phases of a typical walking gait cycle. Letters L or R before the event names indicate the lower limb side, left (red) or right (blue), respectively. The horizontal stripes illustrate the stance phase of each lower limb. HS: heel strike; HO: heel off; TO: toe off; TS: toe strike.

1.2 Methods to measure gait

Temporal events that identify different phases in a person’s gait can be consistently defined by the interaction forces between the left and right feet and the ground (see section 1.1). Accordingly, force sensors embedded at the ground (e.g., force platforms) and wearable sensors (e.g., portable foot switches at the sole) have been used as the gold-standard method to detect gait events. A method to estimate gait events, convenient for its portability, low cost, and practicality, has attracted increasing interest. The method is based on inertial and magnetic wearable sensors integrated as a single inertial measurement (IM) unit [6, 7]. A typical portable IM unit that can be applied in gait analysis consists of a microelectromechanical system with gyroscope (angular position rate sensor), an accelerometer (acceleration sensor) and a magnetometer (orientation sensor) with one, two, or three axes at each sensor [6, 8, 9].

1.3 What has already been done

Several algorithms based on some or all signals of an IM unit, or even of multiple IM units, have been proposed in the literature to estimate specific events during normal and impaired gait for potential real-time use in daily living situations. We performed a review of the studies about different methods for gait event identification, where we included the studies which matched one or more of the following search keys in Medline and Google Scholar: gait detection, gait identification, gait events, inertial sensors and inertial measurement unit. Table 1 summarizes the main characteristics of the selected recent studies on this topic since 2010 (see [6] for a review on older articles). Despite the intense development in this field, there is no accepted, robust algorithm for gait-event estimation based on IM units for real-time applications.

A limitation to the development of robust methods for gait-event estimation is that there is no public dataset consisting of raw signals to enable gait analysis (e.g., force data from foot-ground interactions and raw IM-unit signals) that researchers could use to test algorithms. We are aware of two open datasets with some of these characteristics: the MAREA gait database [10] contains data from foot-switch sensors (foot–ground contact data) and from accelerometers, but not gyroscope or magnetic sensor data. The OSHWSP gait dataset [11] contains data from triaxial accelerometers and gyroscopes, but not foot-ground contact data. There is thus a demand for an open dataset of gait-event-related signals. Such a dataset should contain data on different walking speeds, because various speeds are present in daily life activities. Data on the timing of tibialis anterior muscle activation in healthy subjects during walking could also be useful for gait-event estimation and the development of an FES device for people with foot drop. This timing can be measured via surface electromyography. A robust public gait dataset would make it possible to test algorithms against a common dataset, improving the replicability and transparency of such studies [12, 13] and attracting research groups around the world to the problem of gait-event estimation, which has otherwise been inaccessible owing to lack of data.

1.4 Goals

With the proposal to develop a reliable method for gait event identification and that could be used in real-time with wearable sensors, the goals of this study were twofold: to create an open dataset of inertial, foot-ground contact, and electromyographic data during walking at different speeds; and to develop a new method based on multiple regression to estimate gait events using wearable inertial sensors with this open dataset.

Table 1 – Main characteristics of studies since 2010 about different algorithms for gait-event estimation using wearable inertial sensors. a: acceleration; ω : angular rate; B: magnetic field. 1, 2, 3: number of sensor's dimensions; HS: heel strike; TS: toe strike; HO: heel off; TO: toe off; c: comfortable; f: fast; s: slow; ?: not reported. For a review on older studies, see [6].

Ref.	Sensor	Number of subjects		Foot drop	Detected gait event				Algorithm	Real-time	Speeds
		Healthy	Impaired		HS	TS	HO	TO			
[14]	a^3	6	-	-	✓	-	-	✓	Symbolization	?	c s
[15]	ω^1	7	-	-	-	-	-	✓	Rule-based	-	c
[16]	$a^3 \omega^3$	-	1	✓	-	-	-	✓	Rule-based	Adjustable	c
[17]	a^3	6	-	-	-	-	✓	-	Peak detection	✓	c f s
[18]	ω^1	6	-	-	-	-	✓	✓	Hidden Markov models	-	c f s
[19]	ω^1	10	10	-	✓	-	✓	-	Hidden Markov models	Adjustable	c f
[20]	$a^3 \omega^3$	10	32	-	✓	✓	✓	✓	Rule-based	Adjustable	c
[21]	$a^3 \omega^3 B^3$	10	-	-	-	-	-	✓	Decision tree	Adjustable	c f
[22]	ω^1	9	-	-	-	-	-	✓	Hidden Markov models	-	c f s
[23]	ω^1	7	-	-	-	-	-	✓	Rule-based	-	c
[24]	ω^1	10	-	-	✓	-	✓	-	Hidden Markov models	Adjustable	c
[25]	a^3	10	10	-	-	-	-	✓	Rule-based	Adjustable	c
[7]	$a^3 \omega^3 B^3$	10	30	-	✓	-	-	✓	Rule-based	-	c f
[26]	$a^3 \omega^3$	5	-	-	✓	-	-	✓	Rule-based	-	c f s
[27]	$a^3 \omega^3 B^3$	7	1	-	✓	✓	✓	✓	Rule-based	✓	?
[28]	ω^3	16	-	-	✓	-	-	✓	Rule-based	✓	c
[29]	a^3	7	-	-	✓	✓	✓	✓	Rule-based	-	c
[30]	ω^1	10	10	✓	✓	✓	✓	✓	Hidden Markov models	-	c s
[31]	$a^3 \omega^3$	14	5	✓	-	-	-	✓	Rule-based	✓	c f s
[32]	$a^3 \omega^3 B^3$	10	-	-	✓	-	-	✓	Rule-based	-	c
[33]	$a^3 \omega^3$	10	32	-	✓	-	-	✓	Hidden Markov models/SVM	-	c f
[10]	a^3	20	-	-	✓	-	-	✓	Time-frequency analysis	-	c f
[34]	$a^3 \omega^3 B^3$	2	-	✓	✓	-	-	✓	Cycle-extremum/Threshold updating	✓	?
[35]	a^3	20	-	-	✓	-	-	✓	Rule-based	-	c f
[36]	a^3	11	-	-	✓	-	-	✓	Rule-based	-	c f
[8]	$a^3 \omega^3 B^3$	11	15	-	✓	-	-	✓	Peak detection	-	c
[37]	$a^3 \omega^3 B^3$	57	-	-	✓	-	-	✓	Rule-based	✓	c
This work	$a^3 \omega^3 B^3$	22	1	✓	✓	✓	✓	✓	Rule-based	Adjustable	c f s

2 Methods

2.1 Subjects

A convenience sample recruited from students and employees at our University composed by 22 healthy subjects (10 males and 12 females) and one female subject with a foot drop gait abnormality voluntarily participated in this study. The group of healthy subjects averaged (± 1 standard deviation) 28.1 ± 7.4 years of age, 71.1 ± 12.0 kg of body mass, 169.6 ± 10.5 cm of height, and 24.7 ± 2.8 kg/m² of body-mass index. The subject with foot drop was 25.2 years of age and had 50.0 kg of body mass, 161.0 cm of height, and 19.29 kg/m² of body-mass index. The subject's foot drop abnormality was caused by congenital cerebral palsy. Data for each subject are presented in the open dataset (see section 3.1 on how to access it). This study was approved by the local ethics committee of the Federal University of ABC (CAAE: 53063315.7.0000.5594), and all subjects signed a consent form prior to data collection.

2.2 Data acquisition

To measure the inertial variables and electrical activity of the tibialis anterior of both legs and the contact of the heel and toe of both feet with the ground, we employed an integrated solution composed of six wireless wearable units and one portable data logger (Trigno EMG System, Trigno Personal Monitor, Delsys Inc., Natick, USA), as shown in Figure 2. The first unit (Trigno IM with 10 channels, Delsys Inc.), referred to here as an IM+EMG unit, had a triaxial accelerometer (with a sampling period of 6.75 ms/sample per channel), a triaxial gyroscope (6.75 ms/sample per channel), a triaxial magnetometer (13.5 ms/sample per channel), and an electromyographic (EMG) channel (900 μ s/sample per channel). This IM+EMG unit was fixed to the shank, over the tibialis anterior muscle, after skin preparation and sensor placement were performed according to the SENIAM recommendations [38]. A second IM+EMG unit was fixed to the forward flat part of the tibia bone, aligned with its long axis at the same height of the first IM+EMG unit. The third unit (Trigno 4-Channel FSR Adapter, Delsys Inc.) was connected with two force-sensitive resistor (FSR) sensors to measure the heel and toe contacts with the ground, with a sampling period of 6.75 ms/sample per channel. Each FSR sensor (a 1.5-cm diameter circular pad) was fixed underneath the toe and heel with double-sided adhesive tape. The remaining three wearable units were fixed to the other leg and foot in a similar configuration. Two IM+EMG units were attached on each leg for a potential future study on signal reproducibility (these data are also available in the open dataset). The exact axes of orientation of the IM+EMG units depended on the subject's leg shape and how he or she walked; in a standing position with the feet parallel to each other, the Y axis was approximately vertical

and the positive direction pointed downward, the Z axis was approximately horizontal and the positive direction pointed forward, and the X axis direction can be found by the right-hand rule (see Figure 2). The data logger was fixed with a belt to the subject's waist. A software code for the data logger managed the data acquisition, and, after the session, the data were uploaded to a computer in a single file for each trial via the data logger's software (EMGWorks, version 4.3, Delsys Inc.). The data from different channels were acquired at different rates and stored in the file with corresponding timestamps.

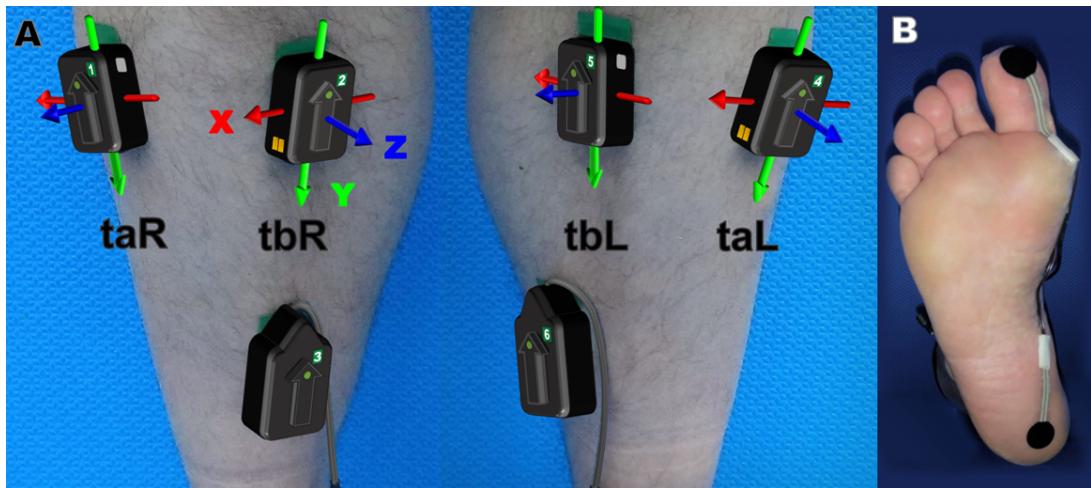


Figure 2 – (A) Placement of the wearable units on the subjects' legs. ‘taR’ and ‘taL’ are the IM+EMG units over the right and left tibialis anterior muscles, ‘tbR’ and ‘tbL’ are the IM+EMG units over the tibia bones, and the other two units are the FSR adapters. The coordinate directions X, Y, and Z of the IM+EMG units are represented using the RGB (red, green, blue) color code, respectively. (B) Placement of FSR sensors on the heel and toe of the right foot.

2.3 Task

After the sensors were attached to the subject and the task was explained, the subject walked barefoot six times at each of three self-paced speeds (comfortable, slow, and fast) on a 40-m long and 2-m wide walkway, without curves, with a flat and rigid surface (trials 1, 3, and 5 were in one direction, and trials 2, 4, and 6 were in the opposite direction). The order of the speeds was randomized among subjects. Each trial lasted from 30 s to 60 s. In addition, one trial was acquired with the subject standing upright and as still as possible for 10 s, for a potential calibration of the sensors. Data collection for each subject was performed in a single session, which lasted 40 min on average.

2.4 Preprocessing

All subsequent steps, including file reading and writing, data processing, analysis, and visualization, were implemented in the Python language using the SciPy library [39]. Figure 3

shows the flowchart for signals preprocessing. There were a few instances of missing data for short periods (less than 100 ms) during the data collection, probably related to wireless transmission, and these missing data appear as zeros in the files. Missing values were identified and the data were reconstructed by linear interpolation. After this reconstruction, the data were filtered with different frequency cutoffs based on the original sampling and data characteristics: accelerometer and gyroscope data were low-pass filtered with a 60-Hz cutoff frequency, data from the magnetometer were low-pass filtered with a 30-Hz cutoff frequency, and EMG data were band-pass filtered between 20 and 450 Hz. These four signals were filtered using a fourth-order zero-phase Butterworth filter. Due to its impulse-response characteristic, the data from the FSR sensors were low-pass filtered with a second-order zero-phase critically damped filter with a 30-Hz cutoff frequency. Subsequently, all data were resampled to a common frequency of 1,000 Hz using a polyphase algorithm [40] (function ‘resample_poly’ from the SciPy library). The amplitude of the FSR data was normalized to the interval 0–1 for each trial. Finally, all data for each trial were saved in an ASCII (text) file and are available in the open dataset. An example of the data from the open dataset is shown in Figure 4.

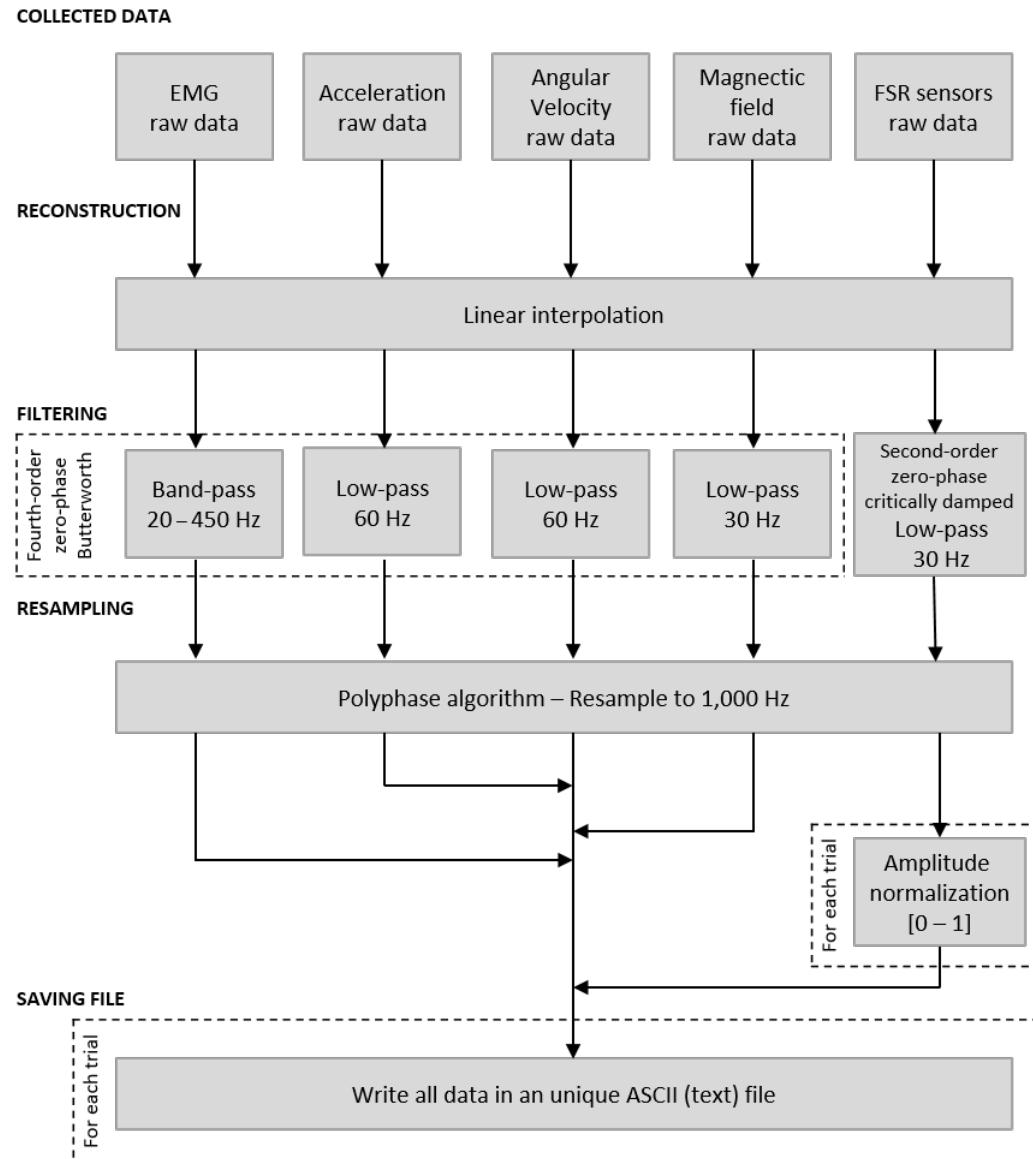


Figure 3 – Flowchart of signals preprocessing. The files written at the last step are available in the open dataset (see section 3.1 on how to access it).

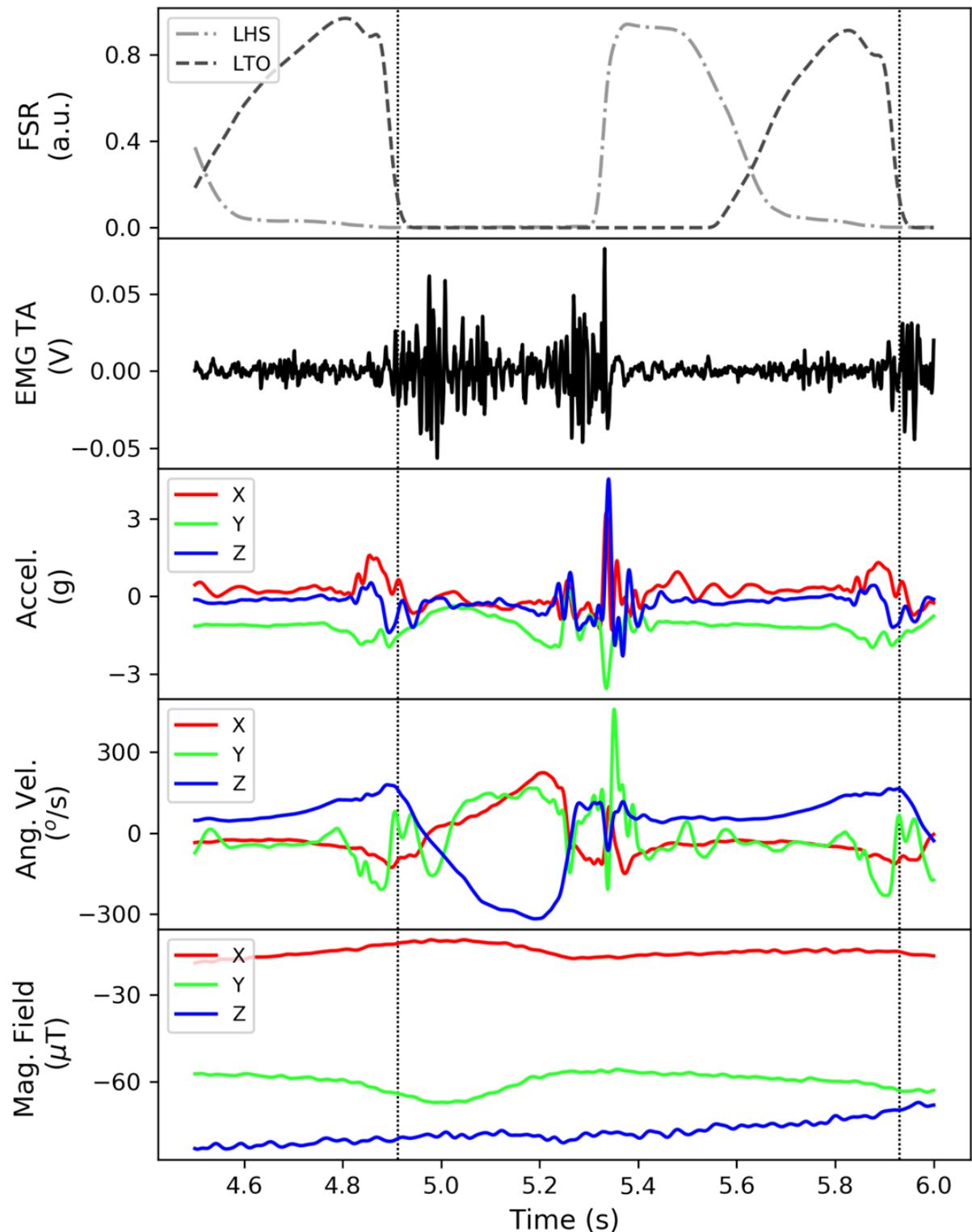


Figure 4 – Exemplary data of part of one trial for the measured signals as available in the open dataset (with minimal processing). FSR: heel and toe contact forces; EMG TA: tibialis anterior electromyographic activity; Accel.: acceleration; Ang. Vel.: angular velocity; Mag. Field: magnetic field. Data are from subject ‘s03’, trial 1, comfortable speed, left side.

2.5 Detection of gait events

Data from the FSR sensors under the heel and toe of the right and left feet were employed to identify the following gait events (see section 1.1): right and left HS, HO, TS, and TO. However, the FSR data often presented fluctuations at the baseline, so before the event detection, this fluctuation was subtracted from the original data, yielding moving-minimum filtered data (using the function ‘move_min’ of the *bottleneck* Python library with a window size of 500 points, 0.5 s, in a dual-pass forward and backward filtering to not introduce any phase lag). To detect gait events from the FSR data, we employed a Python function ‘detect_onset.py’ (available with the dataset), which performs onset detection based on an amplitude-threshold method with a parameter specifying a minimum number of samples above threshold to detect as onset, and other parameters to tune the detection. (After a few trials with visual inspection of the function output, the following parameters were chosen: $threshold = 2$, $n_above = 1$, $n_below = 2$, $threshold2 = 3$, $n_above2 = 1$.) The gait events are stored in a separate file as indices corresponding to the rows in the sensor data files for a given trial, and they are also made available in the open dataset.

2.6 Data visualization

To visualize patterns in the measured signals, we computed average data across subjects. First, an estimation of the EMG amplitude was calculated using a moving RMS filter with a window of 100 points (0.1 s). The other signals were low-pass filtered with a fourth-order zero-phase Butterworth filter and a 10-Hz cutoff frequency. Second, for each trial, we segmented the data in cycles or strides (see section 1.1). The data of each stride were normalized in time from 0% to 100% in steps of 1% and averaged across trials to obtain the mean gait cycle for the given subject/condition. The mean and standard deviation of the gait cycle across subjects were calculated, repeating the same process with the data of all subjects.

2.7 Algorithm for gait-event estimation

To reiterate, we want to estimate temporal gait events from wearable inertial sensor signals, aiming at a potential real-time, day-to-day application. Using system identification [41, 42], we will approach the underlying phenomenon as a system where inertial signals are inputs and a gait event is the output. We will exploit the possibility of using one or more inertial signals to predict a gait event by mathematically representing the system as a parametric empirical model where both its structure (inputs and their connections to the output) as well its parameters are unknown. In addition, as an initial approach to the problem, we will model the system as linear, static, discrete in time, and time-invariant; this will make possible the use of standard and robust statistical methods that are straightforward for embedding into a wearable device. The model calibration, the estimation of its structure and parameters, and initial model

validation are performed in two steps. First, we select features based on the inertial signals and build all possible combinations between inputs and output. Each combination is a candidate for the model structure. The parameters of each combination (the coefficients) are found by multiple linear regression between the features and a mathematical representation of the true gait event (measured with the FSR sensors) using a set of experimental data. Second, of all these linear regression equations, the one that best predicts the gait event is found by testing all these equations with a new set of experimental data with known gait events.

Considering the potential application of this method to the foot drop condition, we will limit the algorithm implementation and test only the estimation of the TO gait event, but this algorithm could be applied to estimation of other gait events (for the flowchart, see Figure 5).

The following features were selected from the data (after the preprocessing described in section 2.4): linear accelerations from the triaxial accelerometer, the magnitude of total linear acceleration, angular velocities from the triaxial gyroscope, and the magnitude of total angular velocity. The magnitudes of total linear acceleration and total angular velocity were computed as the Euclidean norms from their corresponding components (Equation 2.1):

$$\begin{aligned} a_{mag} &= \sqrt{a_X^2 + a_Y^2 + a_Z^2} \\ \omega_{mag} &= \sqrt{\omega_X^2 + \omega_Y^2 + \omega_Z^2} \end{aligned} \quad (2.1)$$

where a_{mag} and ω_{mag} are the magnitudes of total acceleration and total angular velocity, respectively, and a_X , a_Y , a_Z , ω_X , ω_Y and ω_Z are the accelerations and velocities at each direction. Additionally, a constant signal was used as another feature to make possible the adjustment of a linear combination of the features to a non-zero-mean signal [42]. We decided to not use the magnetometer signals because of their dependence on the direction of movement and location in relation to Earth's magnetic field, which would make this signal less robust to everyday applications. In addition, we use only the data from the IM+EMG unit placed over the tibialis anterior muscle of each leg.

Because the frequency content of the inertial signals was below 50 Hz, the data were resampled from 1,000 Hz to 100 Hz to remove redundant information contained in the signals (using the function ‘scipy.signal.decimate’). Furthermore, the data were divided into segments of 5 s (500 samples) to improve the performance of the least-squares method employed in the multiple linear regression [43]. For the analysis, the 5-s duration segments were randomly split into n_{tr} training subsets and n_{ts} test subsets, with twice as many segments for the training than for the testing subset. For each of the 5-s segments, we created a signal \mathbf{Y} representing the instants immediately before the TO event, a 100-ms (100 samples) duration window ending at the instant of the TO event (to be estimated using the TO event detected by the FSR sensor as the true value). The signal \mathbf{Y} was defined as having amplitude 1 inside the 100-ms window and 0 outside. The duration of the rectangular window equal to 100 ms was chosen to accommodate the variability of the onset of tibialis anterior muscle activation close to the TO event.

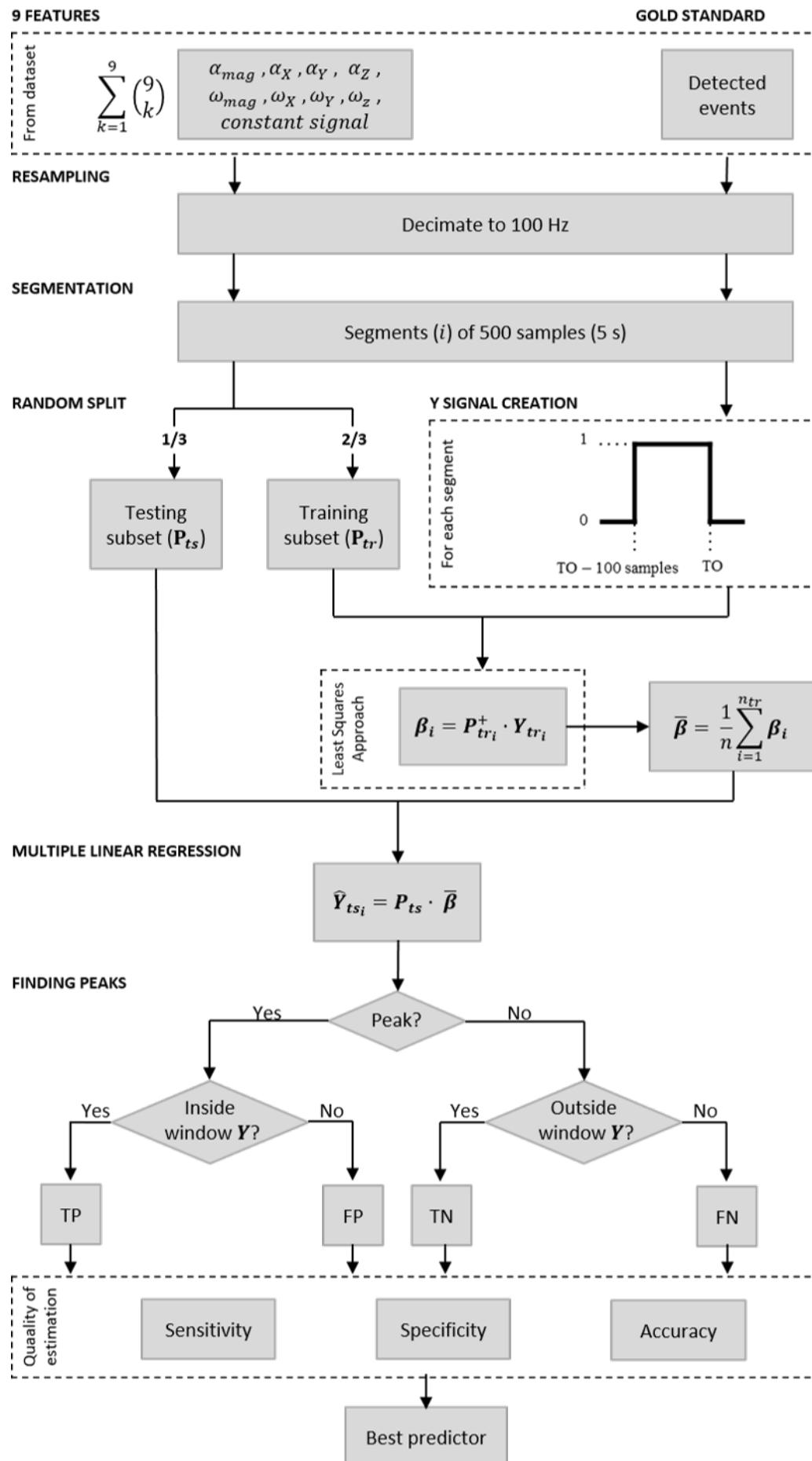


Figure 5 – Flowchart of the algorithm for TO-event estimation. This algorithm could be applied to estimation of other gait events by adapting its Y signal.

For each of the possible $\sum_{k=1}^9 \binom{9}{k} = 2^9 - 1 = 511$ combinations of the features, the following procedure was performed:

1. n_{ts} matrices \mathbf{P}_{ts_i} and n_{tr} matrices \mathbf{P}_{tr_i} were built. Each of these \mathbf{P}_i matrices was built with 500 lines, each line corresponding to each sample of the i^{th} data segment, and with k columns. Each column of the matrix contained the data of one of the features of a given combination.
2. Each \mathbf{P}_{tr_i} matrix and its corresponding signal \mathbf{Y}_{tr_i} were used to obtain the weights β_i for each of the features by using a multiple least squares approach:

$$\beta_i = \mathbf{P}_{tr_i}^+ \cdot \mathbf{Y}_{tr_i} \quad (2.2)$$

where the symbol $^+$ represents the pseudoinverse of the matrix \mathbf{P}_{tr_i} .

3. The final weight vector $\bar{\beta}$ is obtained by computing the mean of the weights obtained for each segment of 5 s:

$$\bar{\beta} = \frac{1}{n_{tr}} \sum_{i=1}^{n_{tr}} \beta_i \quad (2.3)$$

4. The computed weights $\bar{\beta}$ were used in the matrices of features previously separated for test purposes, \mathbf{P}_{ts} , by multiplying each feature by the corresponding weight:

$$\hat{\mathbf{Y}}_{ts_i} = \mathbf{P}_{ts} \cdot \bar{\beta} \quad (2.4)$$

5. A function to find the peaks of signals above a certain threshold (function ‘find_peaks’ available with the dataset) based on comparison of neighbouring values was used to find the peaks of each $\hat{\mathbf{Y}}_{ts_i}$ signal. This procedure was performed considering thresholds between the negative of the maximum value and the maximum value of the signal $\hat{\mathbf{Y}}_{ts_i}$, with steps of 1% of the maximum value.
6. The peaks were considered as the estimated instants of TO. If the peak of $\hat{\mathbf{Y}}_{ts_i}$ was inside the corresponding window of the \mathbf{Y}_{ts_i} signal, it was considered a true positive (TP). Otherwise, it was considered a false positive (FP). Likewise, if no peak of $\hat{\mathbf{Y}}_{ts_i}$ was found outside the window of \mathbf{Y}_{ts_i} it was considered a true negative (TN), and, if no peak of $\hat{\mathbf{Y}}_{ts_i}$ was found inside the window of \mathbf{Y}_{ts_i} it was considered a false negative (FN).

7. The quality of the estimation was assessed by determining the true-positive rate (also known as sensitivity, (2.5)) and the false-positive rate (computed as $1 - Specificity$, (2.6)), where:

$$\text{Sensitivity} = \frac{\sum TP}{\sum TP + \sum FN} \quad (2.5)$$

$$\text{Specificity} = \frac{\sum TN}{\sum TN + \sum FP} \quad (2.6)$$

8. The true-positive and false-positive rates were used to build a receiver operating characteristic (ROC) curve, in which each point of the curve corresponds to a threshold.
9. The accuracy of the estimator was computed for the threshold with the point of the ROC curve nearest to the upper-left point of the graph (point (0,1)), where the accuracy is given by:

$$\text{Accuracy} = \frac{\sum TP + \sum TN}{\sum TP \sum FN + \sum TN + \sum FP} \quad (2.7)$$

The most accurate predictor, (i.e., the multiple linear regression equation with its coefficients, the structure and parameters β of the model), were selected to estimate the TO gait event.

3 Results

3.1 Open dataset

The dataset, consisting of all the data of the 22 healthy subjects plus one subject with foot drop, is available as an open repository accessible on the internet (DOI: [10.6084/m9.figshare.7778255](https://doi.org/10.6084/m9.figshare.7778255)), under the CC0 license (<https://creativecommons.org/publicdomain/zero/1.0/>). The data are stored in ASCII (text) format and can be downloaded separately or as a single compressed file. The dataset has three types of contents: data of the measured signals (data files), data of the gait events (event files), and metadata about the subjects (metadata file).

The data file contains a time column ('Time', in seconds), along with tab-separated columns with data from the four IM+EMG units over the tibialis anterior ('ta') muscle and over the tibia bone ('tb') of the right ('R') and left ('L') legs with triaxial ('x', 'y', 'z', see subsection 2.2 for the axis convention) accelerometers ('ACC', in units of gravitational acceleration), triaxial gyroscopes ('GYR', in °/s), triaxial magnetometers ('MAG', in μ T), EMG of the tibialis anterior ('EMG', in mV), and from the force sensitive resistors ('FSR', in arbitrary units, normalized from 0 to 1) under the heel ('hs') and toe ('to') of both feet, resulting in a total of 43 columns, all sampled at 1,000 Hz. Accordingly, each file has the following header indicating the type of data in each column:

```
Time, EMG_taR, ACCx_taR, ACCy_taR, ACCz_taR, GYRx_taR, GYRy_taR, GYRz_taR,
MAGx_taR, MAGy_taR, MAGz_taR, ACCx_tbR, ACCy_tbR, ACCz_tbR, GYRx_tbR, GYRy_tbR,
GYRz_tbR, MAGx_tbR, MAGy_tbR, MAGz_tbR, EMG_taL, ACCx_taL, ACCy_taL, ACCz_taL,
GYRx_taL, GYRy_taL, GYRz_taL, MAGx_taL, MAGy_taL, MAGz_taL, ACCx_tbL, ACCy_tbL,
ACCz_tbL, GYRx_tbL, GYRy_tbL, GYRz_tbL, MAGx_tbL, MAGy_tbL, MAGz_tbL, FSR_hsR,
FSR_toR, FSR_hsL, FSR_toL
```

These files are named 's<nn><c><t>.txt', where <nn> refers to the number of the subject from '00' to '22' ('00' is the subject with foot drop); <c> refers to the walking speed ('c': comfortable, 's': slow, or 'f': fast); and <t> refers to the trial (from '1' to '6'). For each of the 23 subjects there are three speeds and six trials (18 files), plus one file for the standing still task (named 's<nn>up.txt'), for a total of 437 files).

The event file contains the indices (the line numbers in the corresponding data file) for the following eight gait events (see subsection 1.1) identified in the header: RHS, RHO, LTS, LTO, LHS, LHO, RTS, and RTO. The name of the event file follows the same convention as that of the data file, followed by the letters 'ev' at the end of the file name. There is one event file for each walking trial, for a total of 414 event files.

The metadata file, named 'info.txt', contains the subjects' numbers and information from

their anamneses. Following is the coding for the metadata (the first word identifies the name of the column in the header):

- Subject: number of the subject (from ‘00’ to ‘22’).
- Gender: gender (‘F’ or ‘M’).
- DateBirth: date of the subject’s birth (yyyy/mm/dd)
- Age: subject’s age in years, months, and days.
- Illness: whether the subject has any self-declared illness (‘Yes’ or ‘No’).
- Illness2: type of illness (‘No’ if the subject does not have any illness).
- Mass: mass in kg (measured with a calibrated scale).
- Height: height in cm (measured with a calibrated stadiometer).
- BMI: body mass index in kg/m².
- DateAcquisition: date of the subject’s evaluation (yyyy/mm/dd).

In total, there are 852 files in the dataset, occupying 6.35 GB. The dataset for the healthy subjects contains data for a total of 9,661 gait strides, with stride length varying from 0.93 m to 2.22 m and walking speed varying from 0.63 m/s to 2.46 m/s (see Figure 6). For the subject with the foot drop abnormality, there are data for 496 gait strides, with stride length from 1.03 m to 1.48 m and walking speed from 0.85 m/s to 1.77 m/s (see Figure 7).

Figure 8 shows plots of the ensemble averages over the gait cycle at comfortable speed for the following variables: EMG activity of the tibialis anterior muscle, three-dimensional acceleration, and angular velocity of the left and right legs. Figures from 9 to 11 show the plots of averages for the same variables for the subject with foot drop over the gait cycle at different speeds, comparing to the ensemble averages. Because the subjects walked six times each, with trials 1, 3, and 5 in one direction and the other trials in the reverse direction, we did not compute the ensemble average across trials for the Earth’s magnetic field data.

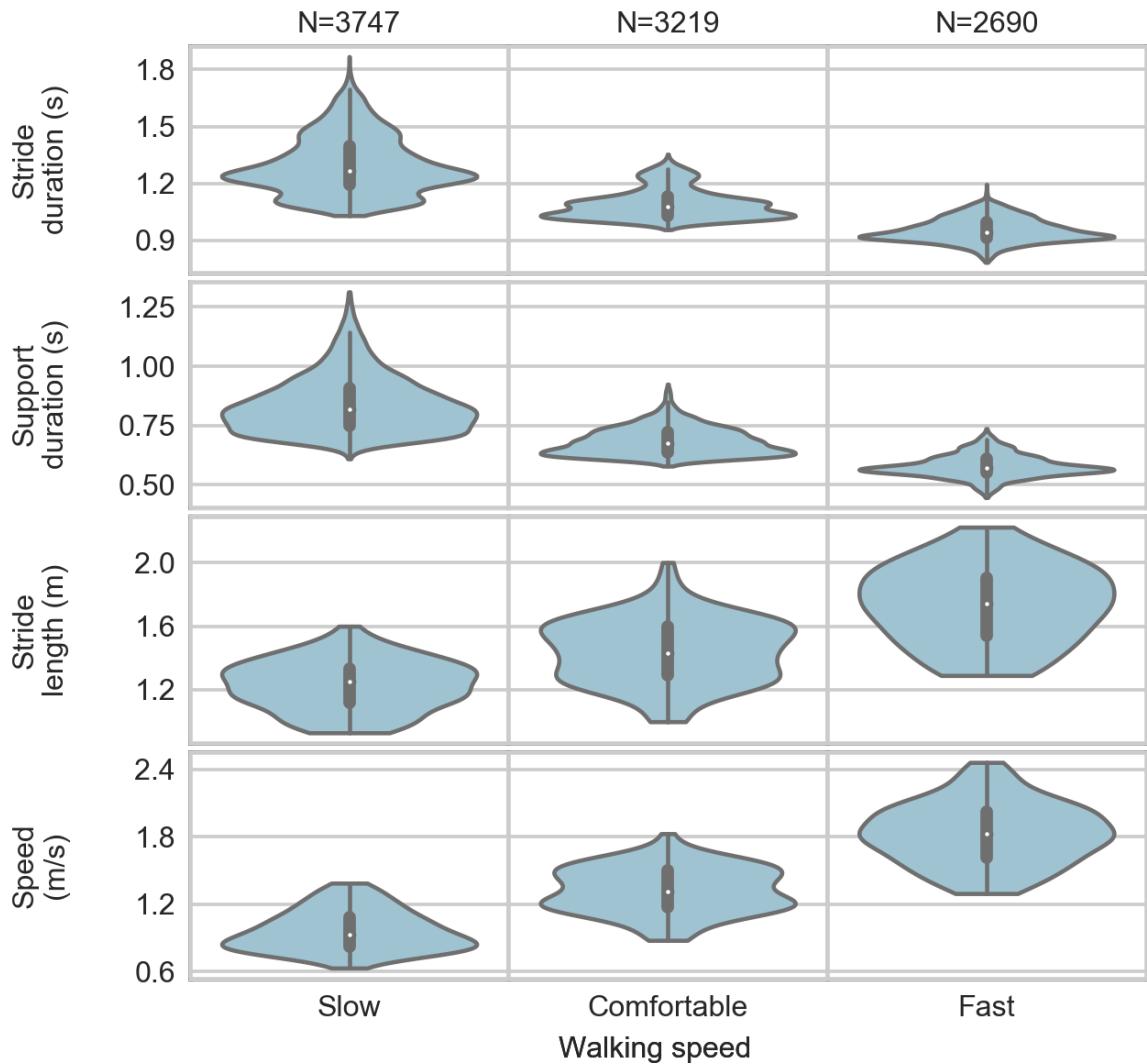


Figure 6 – Violin plots (boxplots plus kernel density estimations) across the 22 healthy subjects from the open dataset for the gait variables: stride duration, support duration, stride length, and speed, calculated using the data from the force-sensitive resistor under the right foot for the different walking speeds. The numbers shown at the top of each column indicate the total number of gait strides available in the dataset at each speed (and used to generate these plots). For each variable, the curve represents an estimation of the data distribution, the vertical black line represents the interval for 95% of the data, the black box represents the interquartile range, and the central dot represents the median value.

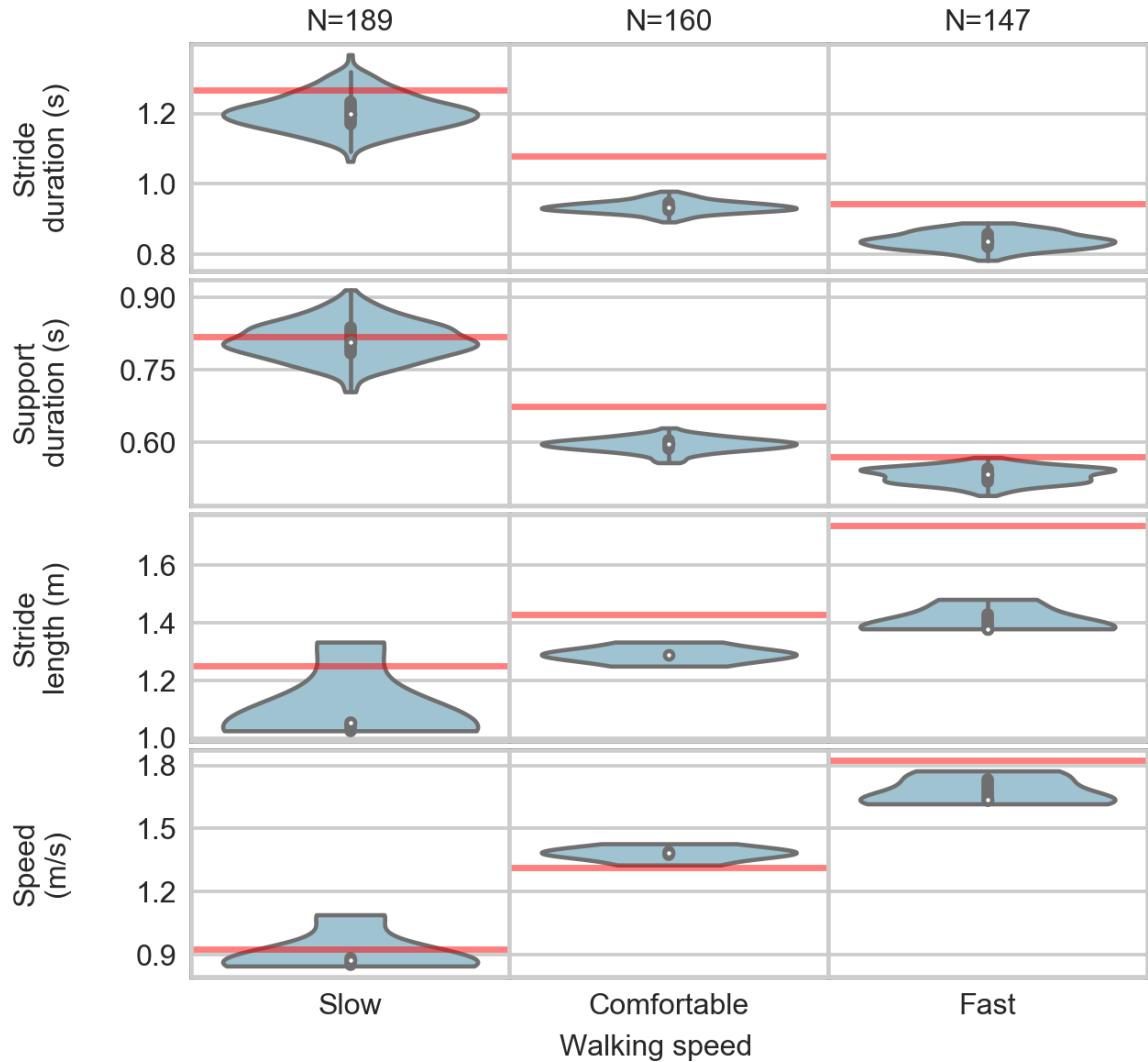


Figure 7 – Violin plots (boxplots plus kernel density estimations) for the subject with foot drop from the open dataset for the gait variables: stride duration, support duration, stride length, and speed, calculated using the data from the force-sensitive resistor under the right foot for the different walking speeds. The numbers shown at the top of each column indicate the total number of gait strides available in the dataset at each speed (and used to generate these plots). For each variable, the curve represents an estimation of the data distribution, the vertical black line represents the interval for 95% of the data, the black box represents the interquartile range, and the central dot represents the median value. The horizontal red line represents the median value across the 22 healthy subjects.

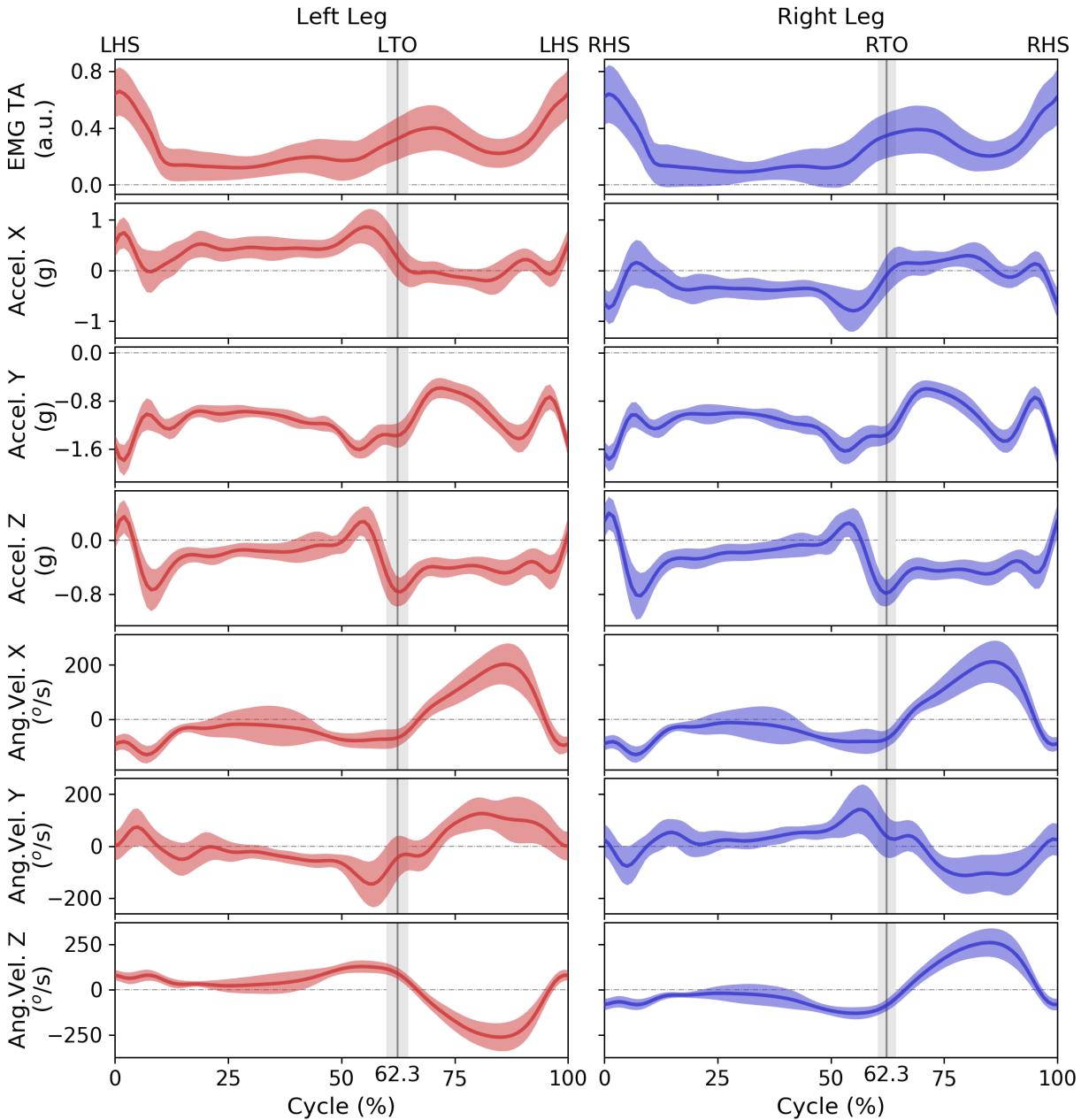


Figure 8 – Mean ± 1 standard deviation across the 22 healthy subjects from the open dataset for the measured signals: tibialis anterior electromyographic activity (EMG TA), three-dimensional acceleration (Accel.), and angular velocity (Ang. Vel.), over the left and right gait cycles walking at a comfortable speed (see the section 2.2 for the axes convention). The mean ± 1 standard deviation termination of the gait support phase, indicated by the LTO or RTO events, are shown by the vertical lines and shaded areas of the plots. The gait events were determined using the data from the force-sensitive resistor under the right and left feet. The abbreviations LHS, RHS, LTO, and RTO denote the gait events: left and right heel strike and left and right toe-off, respectively. These curves are based on a total of 3,222 gait strides.

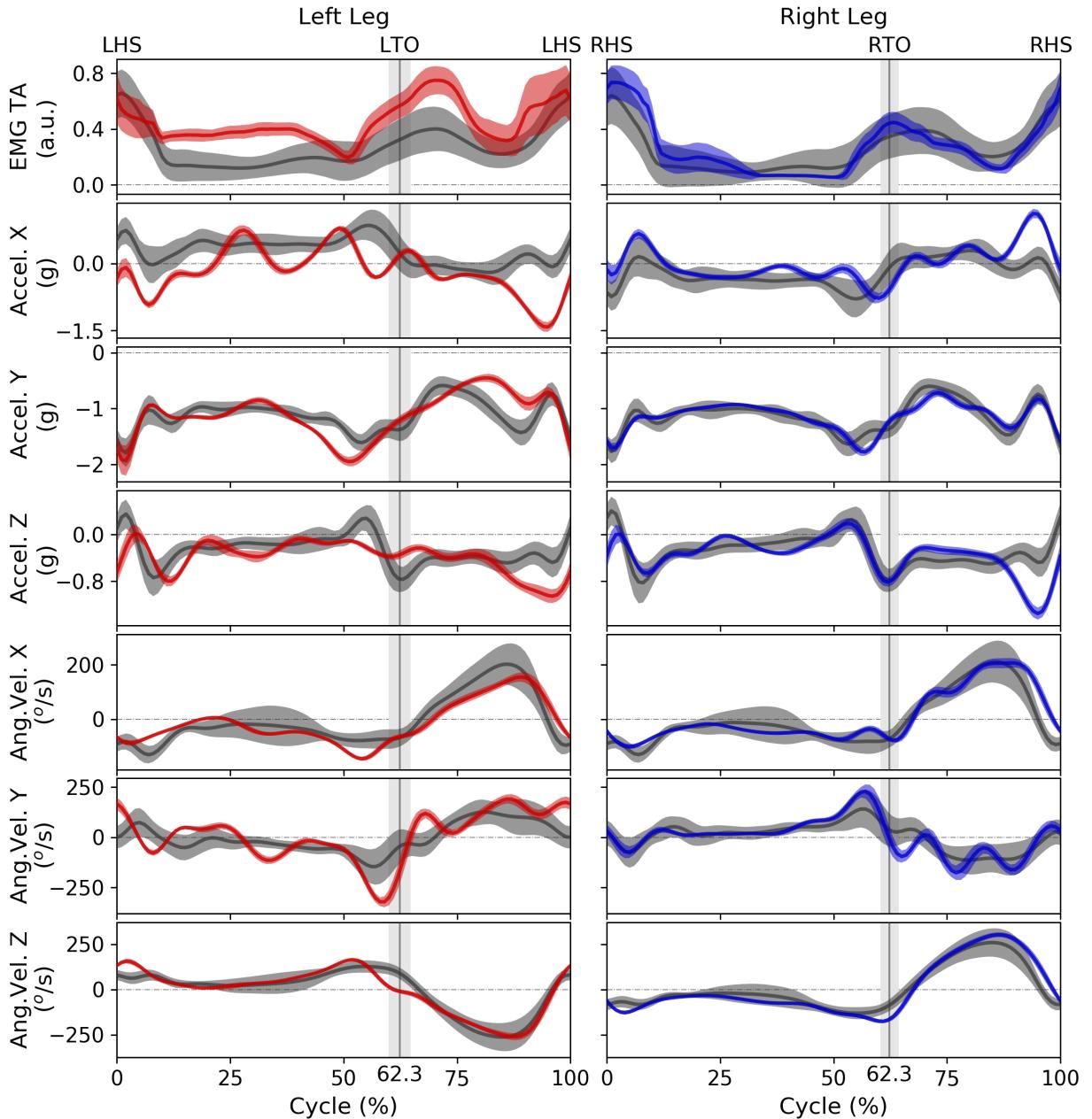


Figure 9 – Mean \pm 1 standard deviation for the subject with foot drop (red and blue) from the open dataset, compared to mean \pm 1 standard deviation across the 22 healthy subjects (gray) for the measured signals: tibialis anterior electromyographic activity (EMG TA), three-dimensional acceleration (Accel.), and angular velocity (Ang. Vel.), over the left and right gait cycles walking at a comfortable speed (see the section 2.2 for the axes convention). The mean \pm 1 standard deviation termination of the gait support phase, indicated by the LTO or RTO events, are shown by the vertical lines and shaded areas of the plots. The gait events were determined using the data from the force-sensitive resistor under the right and left feet. The abbreviations LHS, RHS, LTO, and RTO denote the gait events: left and right heel strike and left and right toe-off, respectively. These curves are based on a total of 160 gait strides.

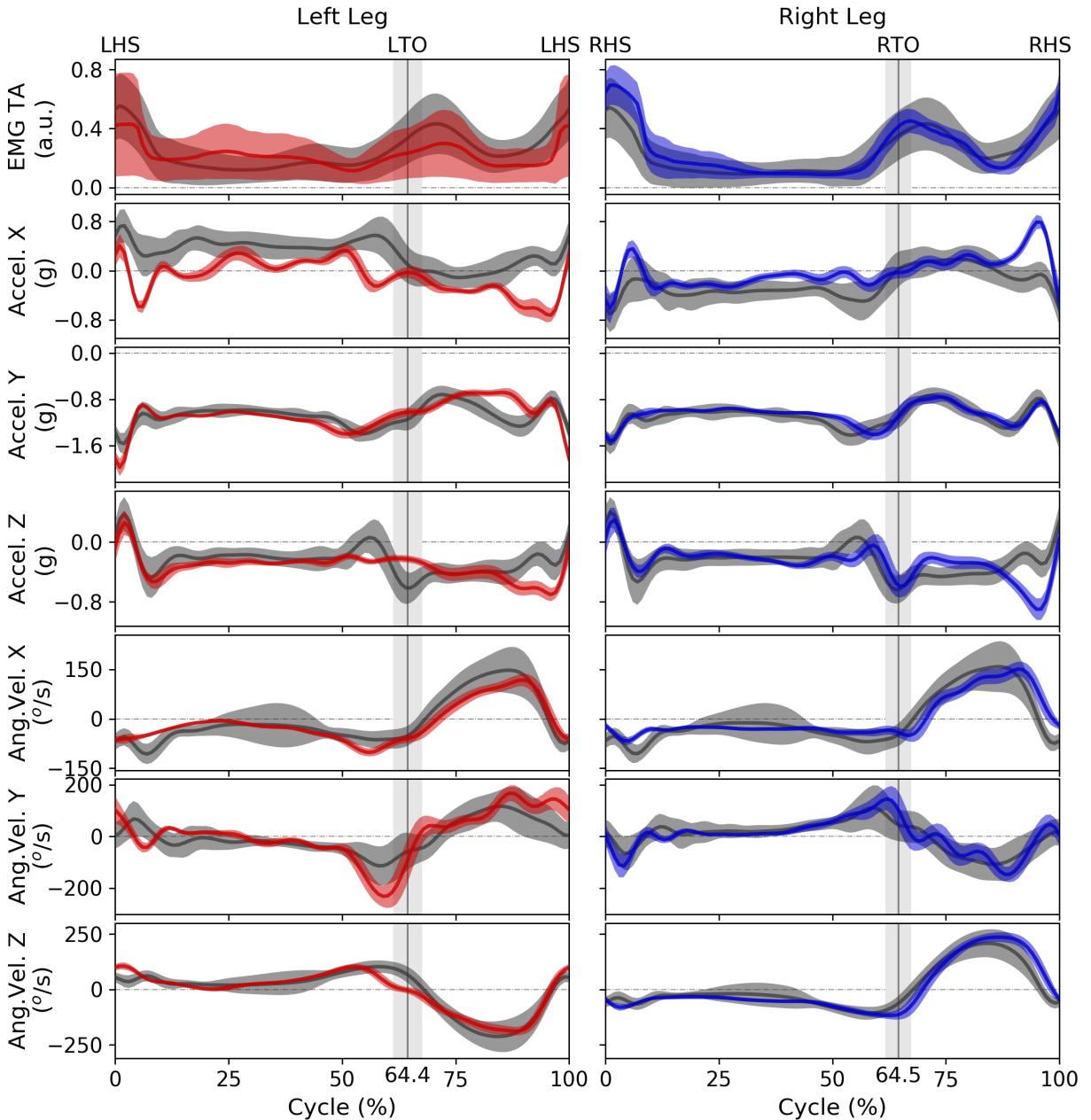


Figure 10 – Mean ± 1 standard deviation for the subject with foot drop (red and blue) from the open dataset, compared to mean ± 1 standard deviation across the 22 healthy subjects (gray) for the measured signals: tibialis anterior electromyographic activity (EMG TA), three-dimensional acceleration (Accel.), and angular velocity (Ang. Vel.), over the left and right gait cycles walking at a slow speed (see the section 2.2 for the axes convention). The mean ± 1 standard deviation termination of the gait support phase, indicated by the LTO or RTO events, are shown by the vertical lines and shaded areas of the plots. The gait events were determined using the data from the force-sensitive resistor under the right and left feet. The abbreviations LHS, RHS, LTO, and RTO denote the gait events: left and right heel strike and left and right toe-off, respectively. These curves are based on a total of 189 gait strides.

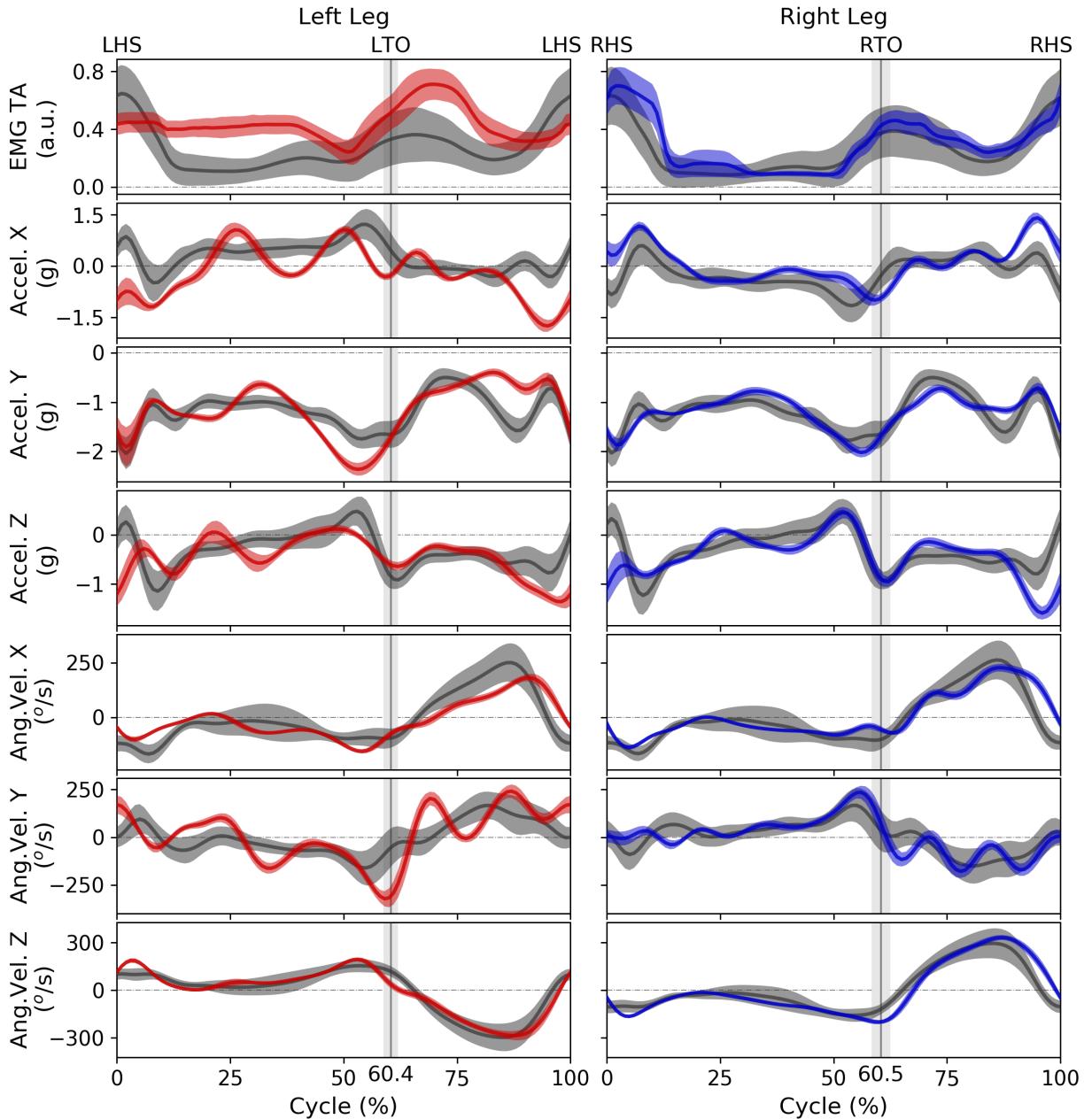
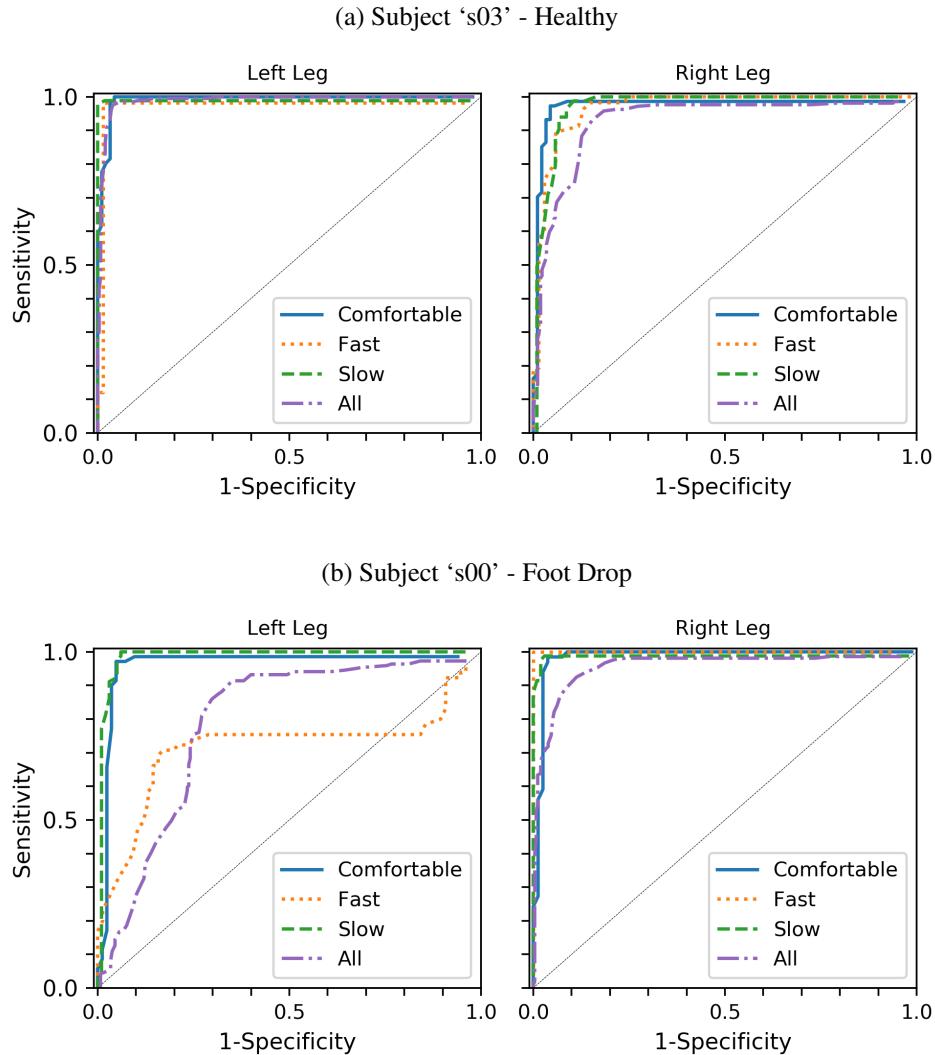


Figure 11 – Mean ± 1 standard deviation for the subject with foot drop (red and blue) from the open dataset, compared to mean ± 1 standard deviation across the 22 healthy subjects (gray) for the measured signals: tibialis anterior electromyographic activity (EMG TA), three-dimensional acceleration (Accel.), and angular velocity (Ang. Vel.), over the left and right gait cycles walking at a fast speed (see the section 2.2 for the axes convention). The mean ± 1 standard deviation termination of the gait support phase, indicated by the LTO or RTO events, are shown by the vertical lines and shaded areas of the plots. The gait events were determined using the data from the force-sensitive resistor under the right and left feet. The abbreviations LHS, RHS, LTO, and RTO denote the gait events: left and right heel strike and left and right toe-off, respectively. These curves are based on a total of 147 gait strides.

3.2 Gait events estimation

The estimation of the TO events was performed for different data groups. Each algorithm execution considered data from one subject, one of his or her legs, and one speed condition, in which the possible speed conditions were: comfortable, fast, slow, and all speeds combined. Because we evaluated data from the two legs of 23 subjects at four speeds, the algorithm was run 184 times, 176 times for healthy subjects and 8 for the subject with foot drop abnormality. For each execution, we computed the set of features whose true-positive rate and false-positive rate corresponded to the largest areas under the ROC curves and calculated the accuracy for this set of features (see Table 2). Examples of ROC curves with the largest areas for the four speed conditions are shown in Figure 12, that contains information about a healthy subject ('s03') and the one with foot drop ('s00'). After the best set of features and the threshold for estimation of the TO event were defined, they were used as parameters for the multiple linear regression and subsequent event estimation.

Figure 12 – Exemplary ROC curves of estimator for subjects from the open data set for four speed conditions over both legs: comfortable (solid), fast (dotted), slow (dashed), and all the three speeds combined (dash-dotted).



We chose to exemplify estimation results by the regression of the data group of all speeds combined, to demonstrate the coverage of the estimator. Although this was not the data group whose ROC curves pointed to the most promising estimator, the TO events were successfully identified by the estimator, as show Figures 13 and 14.

Table 2 – Sensitivity, specificity and accuracy of the TO-event estimation at different gait speeds. For the healthy group, median and [5th, 95th percentiles] across subjects are shown. For the subject with foot drop, the left side is affected.

Speed	Group			
	Healthy (n=22)		Impaired (n=1)	
	Left	Right		
Sensitivity				
Comfortable	0.9680 [0.9259, 1.0000]	0.9714	0.9848	
Fast	0.9813 [0.9649, 1.0000]	0.7077	1.0000	
Slow	0.9521 [0.8906, 1.0000]	0.9747	0.9756	
All	0.8905 [0.7969, 0.9886]	0.8416	0.9249	
Specificity				
Comfortable	0.9543 [0.9198, 1.0000]	0.9518	0.9620	
Fast	0.9677 [0.9524, 1.0000]	0.8289	1.0000	
Slow	0.9521 [0.8906, 1.0000]	0.9495	0.9804	
All	0.8984 [0.8144, 0.9789]	0.7132	0.8872	
Accuracy				
Comfortable	0.9603 [0.8477, 1.0000]	0.9608	0.9724	
Fast	0.9744 [0.8452, 1.0000]	0.7730	1.0000	
Slow	0.9328 [0.5829, 0.9991]	0.9663	0.9783	
All	0.8879 [0.7041, 0.9836]	0.7737	0.9043	

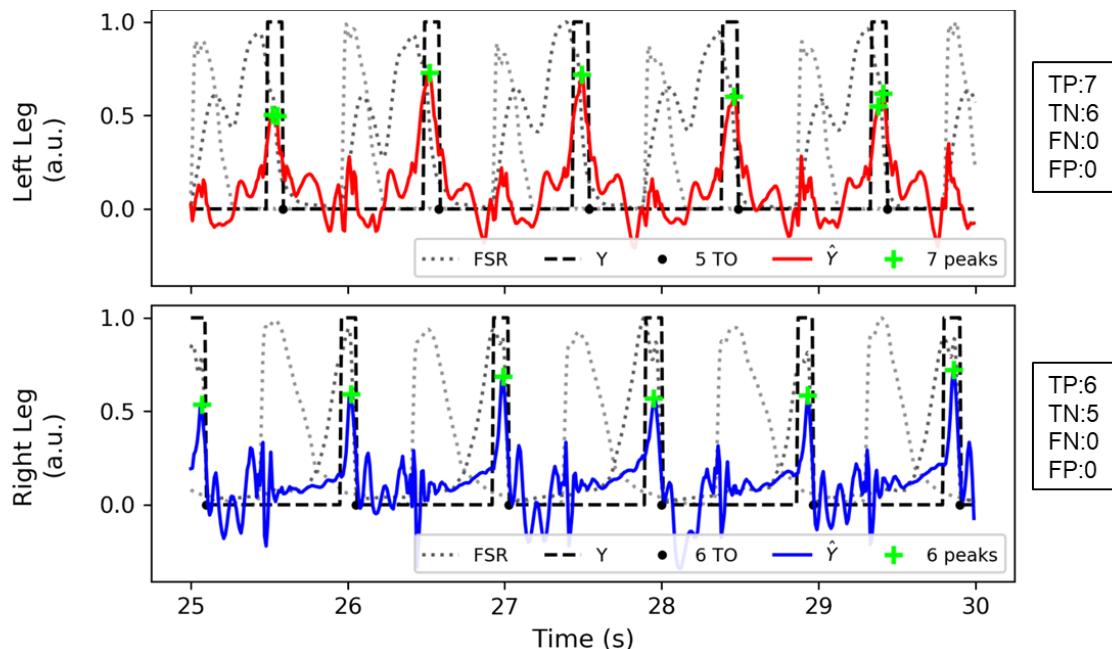


Figure 14 – Results of TO-event estimation for the subject ‘s00’ from the open dataset, using data from all speeds for training and testing. For the left leg, the following features were used in the regression: ‘ACC’ (magnitude), ‘ACCx’, ‘ACCz’, ‘GYR’ (magnitude), ‘GYRx’ and ‘GYRz’ (estimation threshold 0.52). For the right leg, the following features were used: ‘ACCx’, ‘ACCz’, ‘GYRx’, ‘GYRy’, ‘GYRz’ and the constant signal (estimation threshold 0.45). The references are signals from force-sensitive resistor (dotted) and y (dashed, see section 2.7). The crosses matches peaks found on regression signal \hat{y} (solid). The shown segment is from trial 1 at comfortable speed.

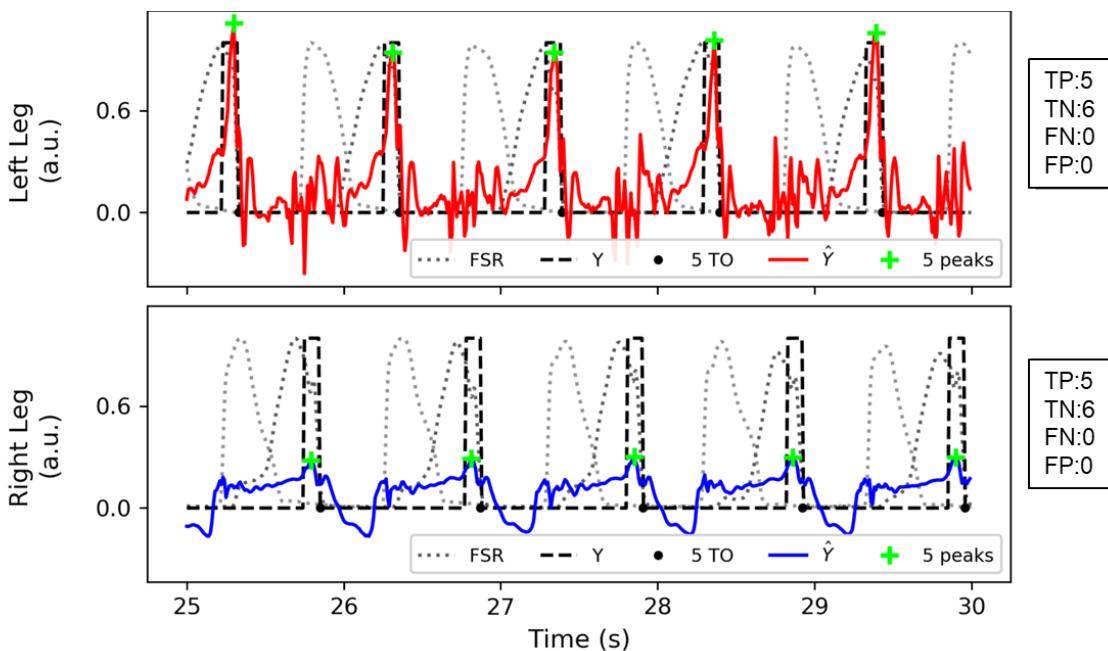


Figure 13 – Results of TO-event estimation for subject ‘s03’ from the open dataset, using data from all speeds for training and testing. For the left leg, the following features were used in the regression: ‘ACC’ (magnitude), ‘ACCx’, ‘ACCz’, ‘GYR’(magnitude), ‘GYRx’, ‘GYRy’, and ‘GYRz’ (estimation threshold 0.46). For the right leg, the following features were used: ‘GYRy’, ‘GYRz’, and the constant signal (estimation threshold 0.61). The references are signals from force-sensitive resistor (dotted) and y (dashed, see section 2.7). The crosses match peaks found on regression signal \hat{y} (solid). The shown segment is from trial 1 at comfortable speed.

4 Discussion

In this study we created an open dataset of inertial, magnetic, electromyographic, and foot-ground contact data from wearable sensors placed on both legs and feet during walking at different speeds and standing still. Data were generated by 22 healthy subjects and one subject with a foot drop gait abnormality. We developed a method based on multiple linear regression for the estimation of gait events using the inertial data from this open dataset.

The open dataset, available at Figshare (DOI: [10.6084/m9.figshare.7778255](https://doi.org/10.6084/m9.figshare.7778255)) under a CC0 license, contains raw data (with minimal processing) sampled at 1,000 Hz totaling 9,661 gait strides of healthy subjects and 496 gait strides of the subject with the foot drop, along with data from the same sensors of each subject standing still. All the healthy subjects exhibited consistent intra- and inter-subject patterns concerning the EMG activity of the tibialis anterior muscle and the three-dimensional acceleration and angular velocity of the left and right legs at the different gait speeds. Also included in the dataset are data files with indices of the actual gait events for each stride and a file with information about the subjects' health characteristics. This open dataset can be used in future studies related to gait-event estimation based on inertial sensors. The dataset will enable researchers to test algorithms for gait-event estimation against a common reference, potentially improving the replicability and transparency of those studies; these are some of the known benefits of open data [12, 13].

The only two other open datasets of inertial gait signals are useful but include only limited signal types, limiting their application. The MAREA gait database [35] comprises triaxial accelerometer data and foot-ground contact of 20 healthy subjects in different gait activities, but no other inertial signals. The OSHWSP gait dataset [11] contains data from triaxial accelerometers and gyroscopes of 12 healthy subjects walking at self-selected speed, but not foot-ground contact data.

The algorithm we developed to estimate gait events based on multiple linear regression produced satisfactory results; the median accuracy for the healthy subjects in the situation where all different speeds were combined, a more real scenario for day-to-day application, was 88.8%. When the present algorithm was applied to estimate the toe-off event for the impaired side of the subject with foot drop abnormality, the worst accuracy of the estimation was 77.3%, at the fast speed. However, the subject reported that, because of her gait abnormality, she is not used to walking at a fast speed, so this condition may not represent a valid situation where the algorithm can be evaluated. If we disregard the fast condition, the accuracy of TO event detection increased to 97.3% for the combination of comfortable and slow speeds. A direct comparison of these results with the literature is problematic, because the use of different gait conditions, algorithms, data analyses, and metrics affect how the performance of gait-event estimation is defined. In the studies most similar to the present one, the accuracy of gait-event estimation varies from

from 66.7% to 90.5% [33], 92.5% [26], 93% [23], 95% [27, 28], and 98% [15]. We conclude that the present algorithm presents comparable performance to other existing algorithms in the literature concerning healthy subjects and the preliminary result based on one subject with foot drop abnormality is promising.

The proposed algorithm presents some limitations. The brute-force search we adopted to determine the best set of features for the model structure (which inertial signal or combination of signals) and the corresponding parameters (regression coefficients) resulted in different sets of features and coefficients for different subjects in the healthy group. Because the overall intra- and inter-subject patterns for the measured variables were consistent (e.g., see the ensemble averages in Figure 8), we were expecting that the best model structure would have the same set of features across subjects. It is not clear why we obtained different sets for different subjects. This characteristic of the algorithm implies that, for use in a real-time application, one would have first to perform a calibration phase with known foot-ground contact forces (basically repeating the steps we conducted in this study) for the intended subject.

5 Conclusion

The open dataset we created contains 9,661 gait strides for the healthy subjects and 496 for the subject with the foot drop. It will enable researchers to test algorithms for gait-event estimation against a common reference.

The algorithm we developed to estimate gait events based on multiple linear regression produced satisfactory results (a median accuracy across the healthy subjects and gait speeds of 88.8%, and an accuracy of 97.3% for the affected limb of the subject with foot drop) and is potentially adjustable for real-time application.

Despite the apparent drawback of the calibration step that is required for real-time application, the presented method has the advantage of computational efficiency at the event estimation step, since it is a simple multiple linear regression. Once the multiple linear regression equation is determined, it can be easily embedded in the wearable device for real-time gait-event estimation.

Bibliography

- 1 STEWART, J. D. Foot drop: where, why and what to do? *Practical Neurology*, BMJ Publishing Group Ltd, v. 8, n. 3, p. 158–169, 2008. ISSN 1474-7758. [12](#)
- 2 WESTHOUT, F. D.; PARÉ, L. S.; LINSKEY, M. E. Central causes of foot drop: Rare and underappreciated differential diagnoses. *Journal of Spinal Cord Medicine*, v. 30, n. 1, p. 62–66, 2007. ISSN 10790268. [12](#)
- 3 LYONS, G. M. et al. A review of portable fes-based neural orthoses for the correction of drop foot. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, v. 10, n. 4, p. 260–279, 2002. ISSN 15344320. [12](#)
- 4 MCN., A. R. Bipedal animals, and their differences from humans. *Journal of Anatomy*, v. 204, n. 5, p. 321–330, 2004. [12](#)
- 5 WHITTLE, M. W. *Gait Analysis: An introduction*. 4th. ed. [S.l.]: Elsevier Ltd., 2007. 255 p. [12](#)
- 6 RUETERBORIES, J. et al. Methods for gait event detection and analysis in ambulatory systems. *Medical Engineering and Physics*, v. 32, n. 6, p. 545–552, 2010. ISSN 13504533. [13](#), [15](#)
- 7 TROJANIELLO, D. et al. Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors: Application to elderly, hemiparetic, parkinsonian and choreic gait. *Journal of NeuroEngineering and Rehabilitation*, v. 11, n. 1, p. 152–164, 2014. ISSN 17430003. [13](#), [15](#)
- 8 CARCREFF, L. et al. What is the best configuration of wearable sensors to measure spatiotemporal gait parameters in children with cerebral palsy? *Sensors (Basel)*, v. 18, n. 2, p. 394, 2018. ISSN 14248220. [13](#), [15](#)
- 9 O'DONOVAN, K. J. et al. An inertial and magnetic sensor based technique for joint angle measurement. *Journal of Biomechanics*, v. 40, n. 12, p. 2604–2611, 2007. ISSN 00219290. [13](#)
- 10 KHANDELWAL, S.; WICKSTRÖM, N. Gait Event Detection in Real-World Environment for Long-Term Applications: Incorporating Domain Knowledge Into Time-Frequency Analysis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, v. 24, n. 12, p. 1363–1372, 2016. ISSN 15344320. [14](#), [15](#)
- 11 LLAMAS, C. et al. Open source platform for collaborative construction of wearable sensor datasets for human motion analysis and an application for gait analysis. *Journal of Biomedical Informatics*, v. 63, p. 249–258, 2016. ISSN 15320464. [14](#), [38](#)
- 12 FERBER, R. et al. Gait biomechanics in the era of data science. *Journal of Biomechanics*, v. 49, n. 16, p. 3759–3761, 2016. ISSN 0021-9290. [14](#), [38](#)
- 13 NOSEK, B. A. et al. Promoting an open research culture. *Science*, American Association for the Advancement of Science, v. 348, n. 6242, p. 1422–1425, 2015. ISSN 0036-8075. [14](#), [38](#)

- 14 SANT'ANNA, A.; WICKSTRÖM, N. A symbol-based approach to gait analysis from acceleration signals: Identification and detection of gait events and a new measure of gait symmetry. *IEEE Transactions on Information Technology in Biomedicine*, v. 14, n. 5, p. 1180–1187, 2010. ISSN 10897771. [15](#)
- 15 CATALFAMO, P.; GHOUSSAYNI, S.; EWINS, D. Gait event detection on level ground and incline walking using a rate gyroscope. *Sensors (Basel)*, v. 10, n. 6, p. 5683–5702, 2010. ISSN 14248220. [15, 39](#)
- 16 KOTIADIS, D.; HERMENS, H. J.; VELTINK, P. H. Inertial gait phase detection for control of a drop foot stimulator. inertial sensing for gait phase detection. *Medical Engineering and Physics*, v. 32, n. 4, p. 287–297, 2010. ISSN 13504533. [15](#)
- 17 PATTERSON, M.; CAULFIELD, B. A novel approach for assessing gait using foot mounted accelerometers. In: *Proceedings of the 5th International ICST Conference on Pervasive Computing Technologies for Healthcare*. Dublin, Ireland: [s.n.], 2011. p. 218–221. ISBN 978-1-936968-15-2. [15](#)
- 18 MANNINI, A.; SABATINI, A. M. Gait phase detection and discrimination between walking–jogging activities using hidden Markov models applied to foot motion data from a gyroscope. *Gait & Posture*, v. 36, n. 4, p. 657–661, 2012. ISSN 09666362. [15](#)
- 19 ABAID, N. et al. Gait Detection in Children with and without Hemiplegia Using Single-Axis Wearable Gyroscopes. *PLoS ONE*, v. 8, n. 9, p. e73152, 2013. ISSN 19326203. [15](#)
- 20 MARIANI, B. et al. Quantitative estimation of foot-flat and stance phase of gait using foot-worn inertial sensors. *Gait and Posture*, v. 37, n. 2, p. 229–234, 2013. ISSN 09666362. [15](#)
- 21 NOVAK, D. et al. Automated detection of gait initiation and termination using wearable sensors. *Medical Engineering and Physics*, v. 35, n. 12, p. 1713–1720, 2013. ISSN 13504533. [15](#)
- 22 MANNINI, A.; GENOVESE, V.; SABATINI, A. M. Online decoding of hidden markov models for gait event detection using foot-mounted gyroscopes. *IEEE Journal of Biomedical and Health Informatics*, v. 18, n. 4, p. 1122–1130, 2014. ISSN 21682194. [15](#)
- 23 FORMENTO, P. C. et al. Gait event detection during stair walking using a rate gyroscope. *Sensors (Basel)*, v. 14, n. 3, p. 5470–5485, 2014. ISSN 14248220. [15, 39](#)
- 24 TABORRI, J. et al. A novel HMM distributed classifier for the detection of gait phases by means of a wearable inertial sensor network. *Sensors (Basel)*, v. 14, n. 9, p. 16212–16234, 2014. ISSN 14248220. [15](#)
- 25 RUETERBORIES, J.; SPAICH, E. G.; ANDERSEN, O. K. Gait event detection for use in FES rehabilitation by radial and tangential foot accelerations. *Medical Engineering and Physics*, v. 36, n. 4, p. 502–508, 2014. ISSN 18734030. [15](#)
- 26 FRACCARO, P. et al. Real-world Gyroscope-based Gait Event Detection and Gait Feature Extraction. *eTELEMED 2014, The Sixth International Conference on eHealth, Telemedicine, and Social Medicine*, Barcelona, Spain, p. 247–252, 2014. [15, 39](#)
- 27 CHEN, B. et al. A new strategy for parameter optimization to improve phase-dependent locomotion mode recognition. *Neurocomputing*, Elsevier, v. 149, n. PB, p. 585–593, 2015. ISSN 18728286. [15, 39](#)

- 28 GOUWANDA, D.; GOPALAI, A. A. A robust real-time gait event detection using wireless gyroscope and its application on normal and altered gaits. *Medical Engineering and Physics*, v. 37, n. 2, p. 219–225, 2015. ISSN 18734030. [15](#), [39](#)
- 29 BOUTAAYAMOU, M. et al. Development and validation of an accelerometer-based method for quantifying gait events. *Medical Engineering and Physics*, v. 37, n. 2, p. 226–232, 2015. ISSN 18734030. [15](#)
- 30 TABORRI, J. et al. Validation of inter-subject training for hidden markov models applied to gait phase detection in children with Cerebral Palsy. *Sensors (Basel)*, v. 15, n. 9, p. 24514–24529, 2015. ISSN 14248220. [15](#)
- 31 MULLER, P.; SEEL, T.; SCHAUER, T. Experimental Evaluation of a Novel Inertial Sensor Based Realtime Gait Phase Detection Algorithm. In: *Proc. of the 5th European Conference on Technically Assisted Rehabilitation - TAR 2015*. Berlin, Germany: [s.n.], 2015. [15](#)
- 32 STORM, F. A.; BUCKLEY, C. J.; MAZZÀ, C. Gait event detection in laboratory and real life settings: Accuracy of ankle and waist sensor based methods. *Gait and Posture*, v. 50, p. 42–26, 2016. ISSN 18792219. [15](#)
- 33 MANNINI, A. et al. A Machine Learning Framework for Gait Classification Using Inertial Sensors: Application to Elderly, Post-Stroke and Huntington's Disease Patients. *Sensors (Basel)*, v. 16, n. 2, p. 134, 2016. ISSN 1424-8220. [15](#), [39](#)
- 34 GAO, Y. et al. A Novel Gait Detection Algorithm Based on Wireless Inertial Sensors. *IFMBE Proceedings*, Springer Singapore, Singapore, v. 62, p. 300–304, 2017. [15](#)
- 35 KHANDELWAL, S.; WICKSTRÖM, N. Evaluation of the performance of accelerometer-based gait event detection algorithms in different real-world scenarios using the MAREA gait database. *Gait and Posture*, v. 51, p. 84–90, 2017. ISSN 18792219. [15](#), [38](#)
- 36 MO, S.; CHOW, D. H. K. Accuracy of three methods in gait event detection during over-ground running. *Gait and Posture*, v. 59, p. 93–98, 2018. ISSN 18792219. [15](#)
- 37 ŠPRAGER, S.; JURIČ, M. B. Robust stride segmentation of inertial signals based on local cyclicity estimation. *Sensors (Switzerland)*, v. 18, n. 4, 2018. ISSN 14248220. [15](#)
- 38 HERMENS, H. J. et al. Development of recommendations for semg sensors and sensor placement procedure. *Journal of Electromyography and Kinesiology*, v. 10, n. 5, p. 361–374, 2000. ISSN 1050-6411. [16](#)
- 39 JONES, E. et al. *{SciPy}*: Open source scientific tools for *{Python}*. Disponível em: <<http://www.scipy.org/>>. [17](#)
- 40 BELLANGER, M. G.; BONNEROT, G.; COUDREUSE, M. Digital filtering by polyphase network: Application to sample-rate alteration and filter banks. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v. 24, n. 2, p. 109–114, 1976. ISSN 00963518. [18](#)
- 41 CAMERON, I. T. et al. *Process Modelling and Model Analysis*. 1st. ed. [S.l.]: Elsevier Science, 2001. 541 p. (Process Systems Engineering). ISBN 9780080514925. [21](#)
- 42 ISERMANN, R.; MUENCHHOFF, M. *Identification of dynamic systems*. Berlin: Springer-Verlag, 2011. 705 p. ISBN 978-3-540-78879-9. [21](#), [22](#)

- 43 BILLINGS, S. A. *Nonlinear system identification*. Chichester, UK: John Wiley & Sons, Ltd., 2013. 574 p. [22](#)