ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΒΙΟΪΑΤΡΙΚΗΣ

UNIVERSITY OF WEST ATTICA
FACULTY OF ENGINEERING
DEPARTMENT OF BIOMEDICAL ENGINEERING

Validation of a system to measure gait cycle parameters by wearable sensors and computational methods

Verros Konstantinos

Student ID: 48016016

Supervisor

Glotsos Dimitris, Professor

Trier, 06/10/2023

The Defense Committee

Supervisor	Member of the examination committee	Member of the examination committee
Prof. Dr. Dimitrios Glotsos	Dr. Stefano Magni	Dr. Rene Peter Bremm
[Digital Signature]	[Digital Signature]	[Digital Signature]

Digitally signed by DIMITRIOS GKLOTSOS
Date: 2023.12.19 16:45:25 +02'00'

STEFANO MAGNI

STEFANO MAGNI
19 déc. 2023

RENE PETER BREMM

RENE PETER BREMM
Nov 11, 2023

DIPLOMATIC THESIS AUTHOR’S STATEMENT

The undersigned **Konstantinos Verros** of **Ioannis**, with student ID number **(480)16016**, student of the Department of **Biomedical Engineering**, Faculty of **Engineering** of the University of West Attica, hereby declare responsibly that:

«I am the author of this thesis, and any assistance I received in its preparation is fully acknowledged and referenced in the thesis. Furthermore, any sources from which I have used data, ideas, or words, whether directly or paraphrased, are cited in their entirety, with complete references to the authors, the publishing house, or the journal, including any sources that may have been used from the Internet. Additionally, I affirm that this work has been written exclusively by me and constitutes intellectual property belonging to both me and the institution.

Violation of the above academic responsibility constitutes substantial grounds for the revocation of my diploma».

Date

06/10/2023

The declarant

Konstantinos Verros

ABSTRACT

Gait abnormalities in Normal Pressure Hydrocephalus (NPH) and Parkinson's Disease (PD) patients share similar characteristics. However, certain quantities differ between the two diseases. In particular, gait asymmetry, due to the difference between the movement of the left and right limbs during walking, is known to be more common in PD. To compare gait between PD and NPH, within the MoveSenseAI (MSAI) study, of which this thesis is part, we use prototype wearable sensors (ActiSense, IEE) to measure time-series data which we then employ to compute gait cycle parameters such as cadence, gait cycle duration, stance phase and swing phase. This system is the foundation of the MSAI study which has the final objective of developing a Machine Learning (ML) algorithm to classify the disease according to the patients' gait. The purpose of this thesis is twofold: the validation of a) the system employed within MSAI, composed of the prototype sensors and a newly developed computational pipeline, and the validation of b) its capability to detect gait asymmetry. For this purpose, 2 different groups of respectively 9 and 12 healthy subjects were recruited, in order to measure their gait and a) compare it to the measurements performed by means of an instrumented treadmill (Gaitway 3D, h/p/cosmos) representing the golden standard in the field and b) analyze gait asymmetry using standardized clinical tests. For objective a) 9 healthy subjects, while equipped with ActiSense, performed 3 walks on Gaitway 3D at the speeds of 2, 3 and 4 km/h, each for a duration of 60 seconds with both systems recording in parallel. For objective b) we measured the gait of 12 healthy subjects considered as controls and of the same subjects performing the walking tests with loads added, alternatively, to one or the other of their lower limbs, to artificially induce asymmetry in their gait in a controlled and reproducible manner. The results from part a) gave us insight into how to calibrate the estimation of the gait cycle parameters by comparing the results of both systems with statistical analysis, ultimately allowing us to validate our system. The results of part b) assessed the potential of the system to detect asymmetry based on symmetry ratios of gait cycle parameters which will subsequently be applied to data from PD and NPH patients within the MSAI study.

Keywords: Wearable Sensors, Gait Analysis, Gait Cycle Parameters, Gait Asymmetry, Statistical Analysis

Acknowledgements:

I would like to express my deepest gratitude to the following individuals and organizations for their invaluable support, guidance, and encouragement throughout the course of my thesis:

My supervisor: **Professor Dimitrios Glotsos** – He went above and beyond by helping with the intricate details, organizing everything seamlessly, and providing valuable input to ensure our work met the university's high standards.

I wish to acknowledge **Professor Klaus Peter Koch,** whose substantial contributions to my academic and professional development have been invaluable. His support has played a pivotal role throughout my journey. Professor Koch graciously provided me with the opportunity to undertake my traineeship at the Trier University of Applied Sciences, initiated our thesis collaboration, and provided essential guidance.

My scientific advisor: **Dr. Stefano Magni –** It is difficult to put into words the impact Dr. Magni had on this journey. He spent countless hours, including his personal time, assisting me with this work and other professional matters, making sure that everything is explained thoroughly and that no question was left unanswered. He contributed significantly to my academic, professional, and personal growth. Your expertise, support, guidance, and unwavering commitment have been nothing short of extraordinary, and for that, I am truly grateful.

My scientific advisor: **Dr. Rene Peter Bremm –** His technical expertise and the guidance he provided significantly enriched the depth and quality of this work. His guidance, patience, and belief in my potential to perform this research have advanced my growth and development in the field. I am deeply thankful for the privilege of working alongside him.

I thank **Dr. Frank Hertel, MD, MA** of the Interventional Neuroscience Group of the Centre Hospitalier du Luxembourg for leading MSAI. **Professor Jorge Goncalves** of the Systems Control Group as well as **Dr. Andreas Husch** of the Imaging and Artificial Intelligence group for hosting my regular visits at the University of Luxembourg.

I would also like to thank **Professor Steffen Müller**, of the Physiotherapy Department of the Trier University of Applied Sciences for granting us access to his lab to validate our system on the h/p/cosmos treadmill and especially for his detailed feedback on the measurement protocol and his support in operating the treadmill.

I equally thank all of the subjects that volunteered to participate in the studies conducted for this thesis. I also thank **IEE Sensing Inc**. and **Luxembourg's National Research Fund** for funding the MSAI study.

Last but not least, I thank my family and friends for their love, encouragement, and support in this life-shaping journey I decided to take.

Thank you all for being part of this academic journey, for your patience, and for believing in the significance of this research.

Table of Contents

Table of Abbreviations

Abbreviation	Definition
NPH	Normal Pressure Hydrocephalus
PD	Parkinson’s Disease
MSAI	MoveSenseAI
ML	Machine Learning
CHL	Centre Hospitalier du Luxembourg
UL	University of Luxembourg
GA	Gait Analysis
GC	Gait Cycle
GCD	Gait Cycle Duration
STP	Stance Phase
SWP	Swing Phase
DSTP	Double Stance Phase
STT	Step Time
CP	Cerebral Palsy
CSF	Cerebrospinal Fluid
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
IQR	Interquartile Range
SD	Standard Deviation

Table of Figures

1. Introduction

Reaching an elderly age comes with a natural result: our body is no longer as strong and versatile, and it is of high importance to prevent the risk of falls wherever we can to maintain and improve quality of life. Brain disorders such as Parkinson's Disease (PD) and Normal Pressure Hydrocephalus (NPH) that affect the motor system in their own ways increase the risk of fall. Although these two diseases share similar characteristics, gait asymmetry, due to the difference between movement of left and right limbs during walking, is more common in PD [1].

I joined MoveSenseAI (MSAI), a research team of the Centre Hospitalier du Luxembourg (CHL) and the University of Luxembourg (UL) for my diploma thesis. The MSAI team and their project, a clinical research study conducted at CHL, are investigating gait abnormalities between PD and NPH patients [2]. To compare their gait, prototype wearable sensors are used to measure data which are then employed to compute gait cycle parameters. This has the ulterior motive of developing a Machine Learning algorithm, with the goal of classifying the disease by the patient's gait, to bring us a step closer to the diagnosis.

My work frame consisted of 1) validating this prototype system (composed by sensors and computational pipeline) against an instrumented treadmill and 2) assessing the potential of the system to detect asymmetry in gait. My contribution to this research comprises defining the protocol used for the studies I was assigned, recruiting healthy subjects to participate, recording their gait along with Dr. Bremm and performing the analysis to support this study. Also, I proposed the weight load test in 2.3. for the asymmetry, described in 2.3.1. – which, to the best of my knowledge has not been employed previously for this particular reason.

An introduction to the quantities and means used to characterize these diseases will be followed by the methods used to validate the system and assess its potential to detect asymmetry. Further on, results of the validation and corrective actions that took place in order to get the most reliable parameters are presented.

1.1. Human Gait Definition

The way we walk has a pattern and it is similarly structured for everyone. Walking can be described for each foot by 4 key events: Heel on, Toe on, Heel off, Toe off. These Heel-Toe events happen in a consecutive order between both feet and form the gait pattern. In order for this to happen, «all voluntary movement, including walking, results from a complicated process involving the brain, spinal cord, peripheral nerves, muscles, bones and joints» [3,4]. This process requires perfect collaboration of the systems mentioned for movement to be produced in our bodies, should one of those be malfunctioning for any reason, gait abnormalities would be introduced to our daily lives inducing asymmetry in gait, temporary in some cases while permanent in others. However, studies have shown that healthy subjects present natural asymmetry [5] and we have also observed the same trend in our study, presented in 4.2.3. Sadeghi H. *et al*., in a review, conclude that gait asymmetry does not appear to be a consequence of abnormality, but rather relates to the contribution of each limb to propulsion and control tasks [6].

Gait Cycle Parameters

For every system or effect we want to measure, it is important that we define the associated quantities first. Thus, to perform Gait Analysis (GA), we start by defining the Gait Cycle (GC), the fundamental of GA before we define the Gait Cycle parameters [7]. For every one of them to be introduced, we provide a parameter for the right foot and a parameter for the left foot. Since we measure time-series data, parameters are expressed in Time (s), Percentage of GC or Steps/min.

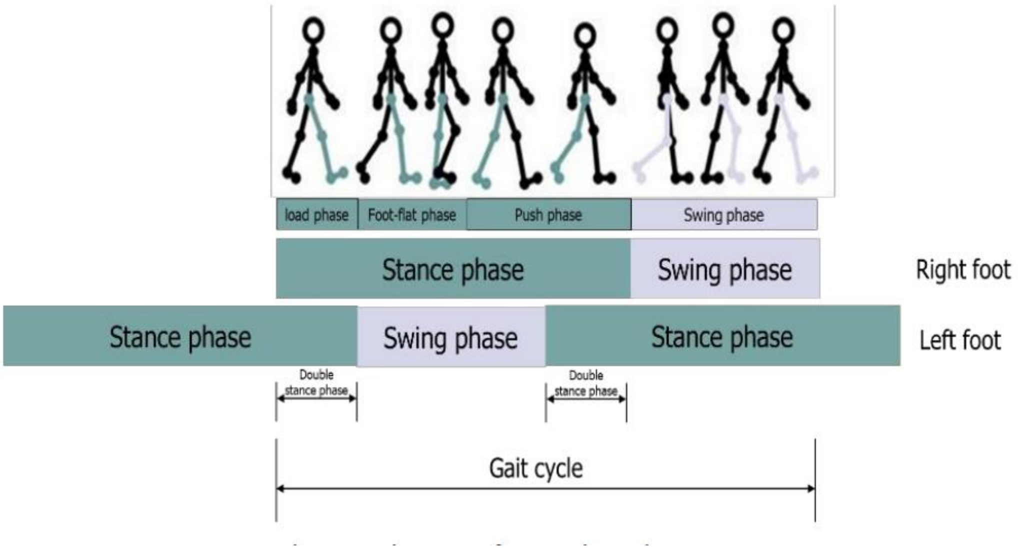


Figure 1. Diagram of one gait cycle [7].

- **Gait Cycle Duration (s) (GCD)** is defined by the very first heel contact (Heel on) of one foot with the ground until the same foot encounters the ground again. Divided into two main phases, Stance Phase (STP) and Swing Phase (SWP).
- **Stance Phase (%) (STP)** is the percentage of the GC beginning from the first heel contact to the moment the toe leaves the ground (Toe off).
- **Swing Phase (%) (SWP)** is the percentage of the GC starting from the moment the toe leaves the ground to the next heel contact.
- **Double Stance Phase (%) (DSTP)** is the amount of time when both feet are in Stance Phase within a GC.
- **Cadence (Steps/min)** is the number of steps per minute.
- **Step Time (s) (STT)** is the amount of time between one foot's heel contact to the other foot's heel contact.

GC can be characterized by many more parameters, but these are the ones we extract, and which will be validated.

1.2. Gait Abnormalities

Any change in our usual gait that is caused without our will can be characterized as abnormal. If for example we twist an ankle, first it will start to get swollen and then the body will cure it over time. This is a natural reflex and a function of our body to avoid pain and inform us that something is wrong. Pain on our ankle in this example, isn't only our damaged tissue recovering but also a security feature that prevents more damage to the same weak point. This of course is an example where recovery is a natural process, and the gait abnormality is temporary and for that period of time our ankle will be in pain, and we will try not to apply as much pressure to reduce load [8].

Other types of Gait Abnormalities which occur due to either a neural or spinal disorder are:

- **Antalgic gait:** is the result of pain that affects the gait as the example described above.
- **Scissors Gait**: this is an effect where knees or thighs hit or cross the same way a scissor does and affects people diagnosed with spastic cerebral palsy (CP) which is a static neurologic condition resulting from brain injury that occurs before cerebral development is complete [9].
- **Spastic Gait (hemiplegic gait)**: also common among people diagnosed with CP but also multiple sclerosis or hemiplegia. Characterized by a "stiff leg" where it needs to be lifted or dragged to move.
- **Steppage Gait (neuropathic gait)**: in this type where the hip is elevated to lift the leg higher than normal the foot may appear floppy when it drops. Usually, the toes point down and scrape the ground during gait. Muscle atrophy or a peroneal nerve injury can cause a steppage gait.
- **Waddling Gait**: causing the upper body movement to exaggerate creates this duck-like walk. A cause for that can be hip dislocation.
- **Crouching Gait**: ankles, knees and hips tend to flex while walking. Also common in cerebral palsy, can appear like the patient is about to bend down.
- **Propulsive Gait (Parkinsonian gait)**: affecting people diagnosed with Parkinson's disease, characterized by a stooping, rigid posture and head and neck bend forwards with a shorter step length to feel more stable and maintain the center of gravity low.

There are also other types such as Ataxic, Shuffling and Lurching gait [8].

Gait abnormalities are more common as we age for many reasons mentioned above.

1.3. Normal Pressure Hydrocephalus

NPH is a brain disorder where excess cerebrospinal fluid (CSF) accumulates in the brain's ventricles. CSF works as a protection layer, like a "cushion" for the brain, although when more of that builds up it cannot exit the scull, resulting in applying pressure to the brain. Especially, this pressure enlarges the ventricles, and they can disrupt or damage nearby brain tissue, leading to difficulty walking, problems thinking and reasoning, and loss of bladder control [10].

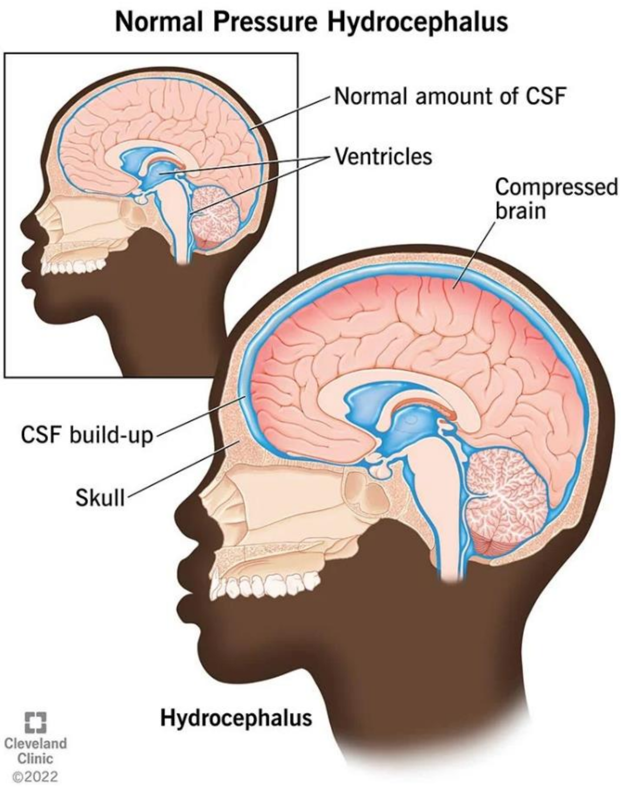


Figure 2. Normal Pressure Hydrocephalus example [8].

Etiology

NPH cases are categorized into Idiopathic NPH (iNPH) where there is no identifiable cause and Symptomatic NPH where trauma or infections of the central nervous system and intracranial hemorrhage are known causes of hydrocephalus. Both share the same characteristics, and the prognosis procedure is the same with the main difference that iNPH is affecting the elderly while symptomatic NPH can affect everyone [10].

Gait Abnormalities

One of the very first symptoms of NPH is gait abnormalities. Characterized by a broad-based gait, with outwardly rotated feet and a diminished step height [10].

Diagnosis/Treatment

Non-invasive imaging applications such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) can help support the diagnosis of iNPH although the tests of CSF drainage increase the prognostic accuracy above 80% and are either,

a) spinal tap test: Lumbar puncture, removing 30-70 mL of CSF or,
b) subarachnoid drainage from Lumbar spine, removing 150-200 mL of CSF daily for a minimum of 2 days up to a week.

These tests have proven to be effective for the prognosis of patients performing the standardized clinical tests offering a great improvement in gait compared to the before and after measurements. The placement of a ventriculoperitoneal shunt implanted with an adjustable valve for iNPH is a standard treatment solution [10].

1.4. Parkinson's Disease

PD is also a brain disorder caused by a loss of brain cells in substantia nigra, the area of the brain where dopamine is produced. This lack of dopamine is what causes malfunction since it's a neurotransmitter that helps us coordinate our movements, the signal transferring is impared and thus the motor system cannot deliver its commands to the muscles properly. Tremor, stiffness and difficulty to balance and coordinate the body are the main characteristics of this disease. Some other symtoms are speech and writing changes[11-13].

Etiology

It is known that the loss of dopamine producing nerve cells is what causes this disease. This is a result of genetic factors such as mutations in genes such as LRRK2 and PINK1 [12], infection, trauma and/or environmental factors (e.g., exposure to many pesticides) [13].

Gait Abnormalities

Gait in PD is characterized by bradykinesia, shorter step length and postural instability.

Figure 3. Parkinson's Disease sketch [19].

Diagnosis/Treatment

Main characteristics of PD are bradykinesia, tremor, and postural instability and the first test is levodopa response on the motor symptoms. Since the etiology remains unknown there is also no treatment. However, there is management of the disease and that's with pharmaco-therapeutic treatment of levodopa. There is of course the wearing off effect of the drugs over the years [11].

1.5. Comparison Normal Pressure Hydrocephalus & Parkinson's Disease

When comparing the groups between them and each one to healthy control subjects there is a variety of results. In one study, Stolze *et al*. (2001) in a comparative analysis of the gait disorder of NPH and PD finds «The symmetry of the step length was not significantly different between the groups, although it was more asymmetric in normal pressure hydrocephalus» [14]. However, it is also known that gait impairments should be symmetric in NPH unless coexisting musculoskeletal disorders cause asymmetry [15]. Moreover, it is found that PD patients tend to be significantly more asymmetric than healthy controls [1]. Considering all the above, symmetry in gait is of significant importance for these diseases and worth examining further.

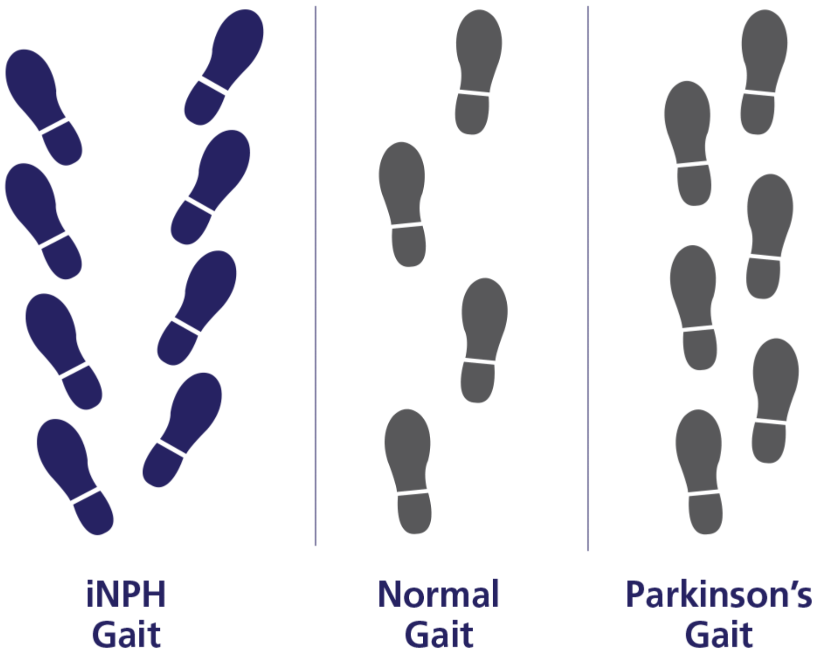


Figure 4. Comparison between Normal Pressure Hydrocephalus & Parkinson's Disease gait [20].

1.6. Diagnostic Equipment

Representing the "Golden Standard", non-wearable systems such as instrumented treadmills and motion tracking systems are the tools providing the best and most accurate gait analysis to date. «The advantage of these systems is that they isolate the study from external factors which could affect the measurements, thus allowing a more controlled analysis of the parameters being studied and obtaining high repeatability and reproducibility levels». They do, however, serve their purpose at the place where they have been installed when it comes to treadmills and unfortunately, they are not cheap enough to distribute them in every facility. Cameras on the other hand, can be portable and placed on other facilities but they usually require additional tools; reflective markers that are placed on the points of interest of the subject or full bodysuits that come preinstalled with those reflective markers to operate [16].

In order to avoid all these expensive and supervision mandatory solutions we are validating a wearable system, where so far only a trained specialist is required. Hopefully in the near future this system will provide a commercially affordable solution for home-care monitoring in a user-friendly application without the need of expertise.

1.7. Standardized Clinical Tests

«To assess gait performance and changes following interventions, clinically and scientifically robust measurement tools are needed» . Flansbjer *et al.* have evaluated the reliability of 6 gait performance tests in individuals in a test reliability study [17].

MSAI uses the "10 Meters Walk" test in order to evaluate our patients before and after shunt/surgery for NPH or medication for PD patients. It allows for a meaningful tool that gives us a fair number of steps and time to evaluate the patient without exhausting their energy. With the 10 Meters Walk we can perform 2 repetitions, the first and the second 10 Meters Walk to have a second measurement and a consistency check, which is used to verify the first walk.

Figure 5. 10 Meters Walk [21].

The "10 Meters Walk" was also used in the study described in this thesis to access the potential of the sensors to detect asymmetry in gait and every time it was recorded a second time on the way back.

1.8. Goal and organization of this thesis

The goals of this thesis are 1) to validate the system which records and extracts the gait metrics against an instrumented treadmill representing the "Golden Standard" and 2) to induce asymmetry in healthy patients gait to test the capability of the system to detect asymmetry. To validate the system, we asked healthy subjects to walk on the instrumented treadmill at different speeds and with wearables and treadmill recording in parallel, then we compare their results with statistical methods. To test the capability to detect asymmetry, another group of healthy subjects was asked to perform a 10-meters-walk as the baseline and perform the 10-meters-walk twice again with weight load on one foot at a time to artificially introduce asymmetry to their gait in a controlled and reproducible manner and compare the results of the 3 walks (each repeated twice). The structure of the thesis is organized by introducing the materials and methods used, presenting, and discussing the results of the methods and the limitations that were taken into consideration with an outlook for future research.

2. Materials & Methods

2.1. System Overview

Subjects are recruited and with the wearable sensors equipped we record their data. Afterwards, we import the raw data on a computational pipeline that extracts Gait Cycle Parameters to further conduct an analysis. Subjects for my validation were healthy subjects while for MSAI, diagnosed iNPH and PD patients.

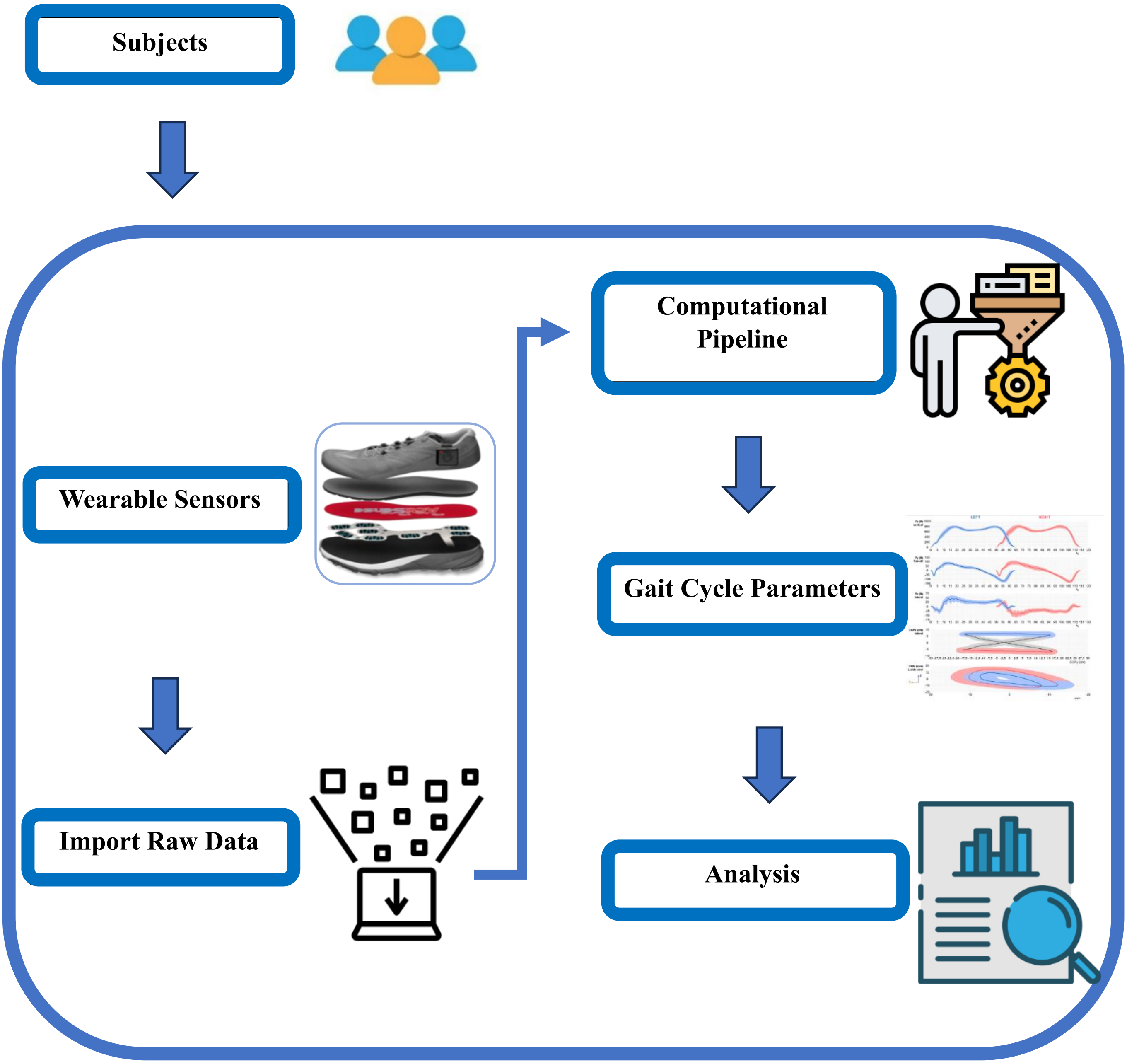


Figure 6. Logic Flow Chart of the System

2.1.1. Wearable Sensors

IEE ActiSense is a wearable, non-invasive analytical system for gait analysis. It consists of eight (8) pressure sensing cells Figure 7 (a), a 3-axis accelerometer, and a 3-axis gyroscope embedded in a housing (b, c), for each shoe for a total of 28 measurement channels recording time-series data at 200 Hz, Figure 8. They are embedded in a pair of shoes and record time-series data, stored internally in the IMU for post processing.

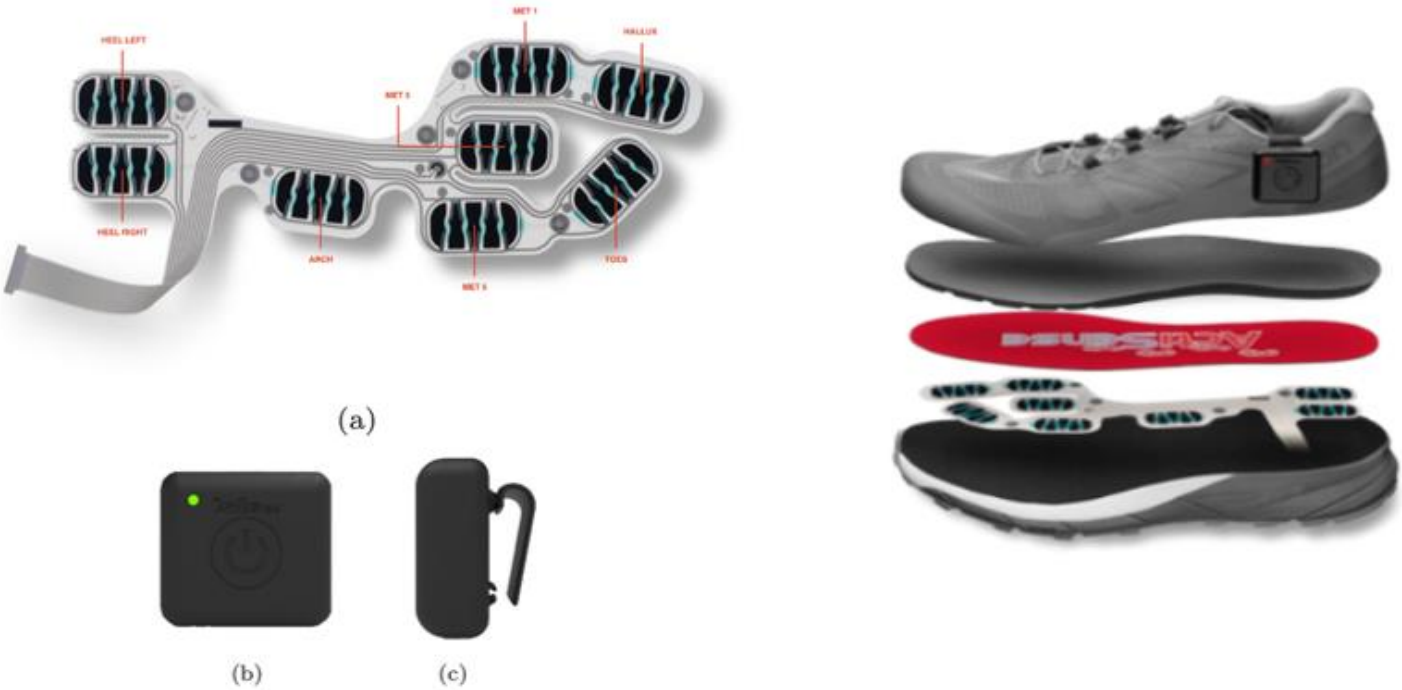


Figure 7. IEE ActiSense, pressure cells (a), IMU front view (b), IMU side view (c), embedded in a shoe (right) [22].

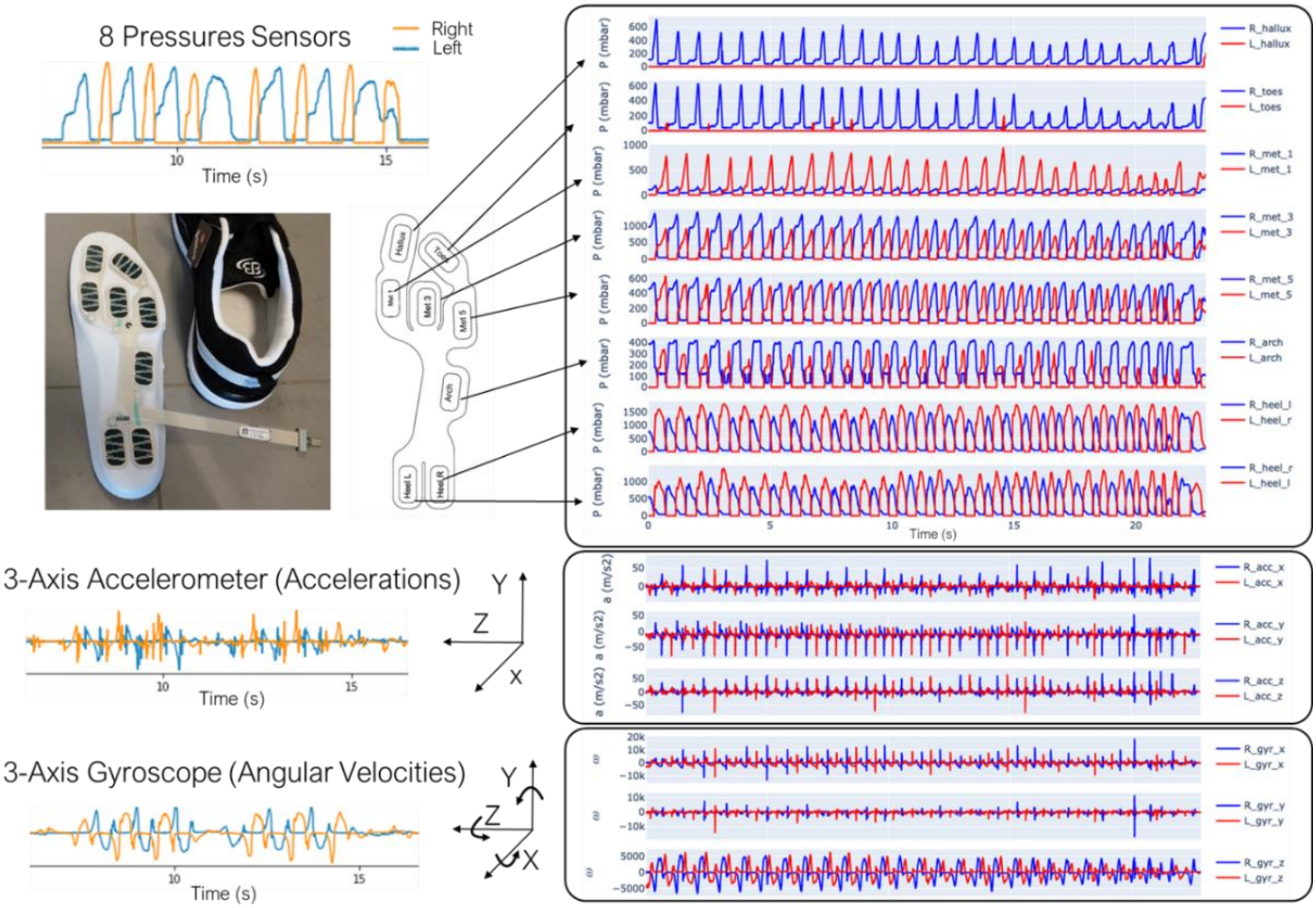


Figure 8. Time-Series Data from ActiSense. Source: MSAI

2.1.2. Computational Pipeline

Adapting methods from a paper [7], Dr. Stefano Magni, scientist of MSAI developed the computational pipeline “EstimateKinematicGaitParameters”, in a Jupyter Notebook form (in Python). Within MSAI, the pipeline is employed to extract Gait Cycle Parameters from the NPH and PD patient data recorded to develop the Machine Learning (ML) algorithm that classifies the disease by the gait.

In short, the structure of the pipeline is the following whether the subject is a healthy subject (as in this thesis) or a diagnosed patient (as in MSAI):

1. Data Preprocessing:

a) Gaussian Filtering is applied on Pressure data, Figure 9:

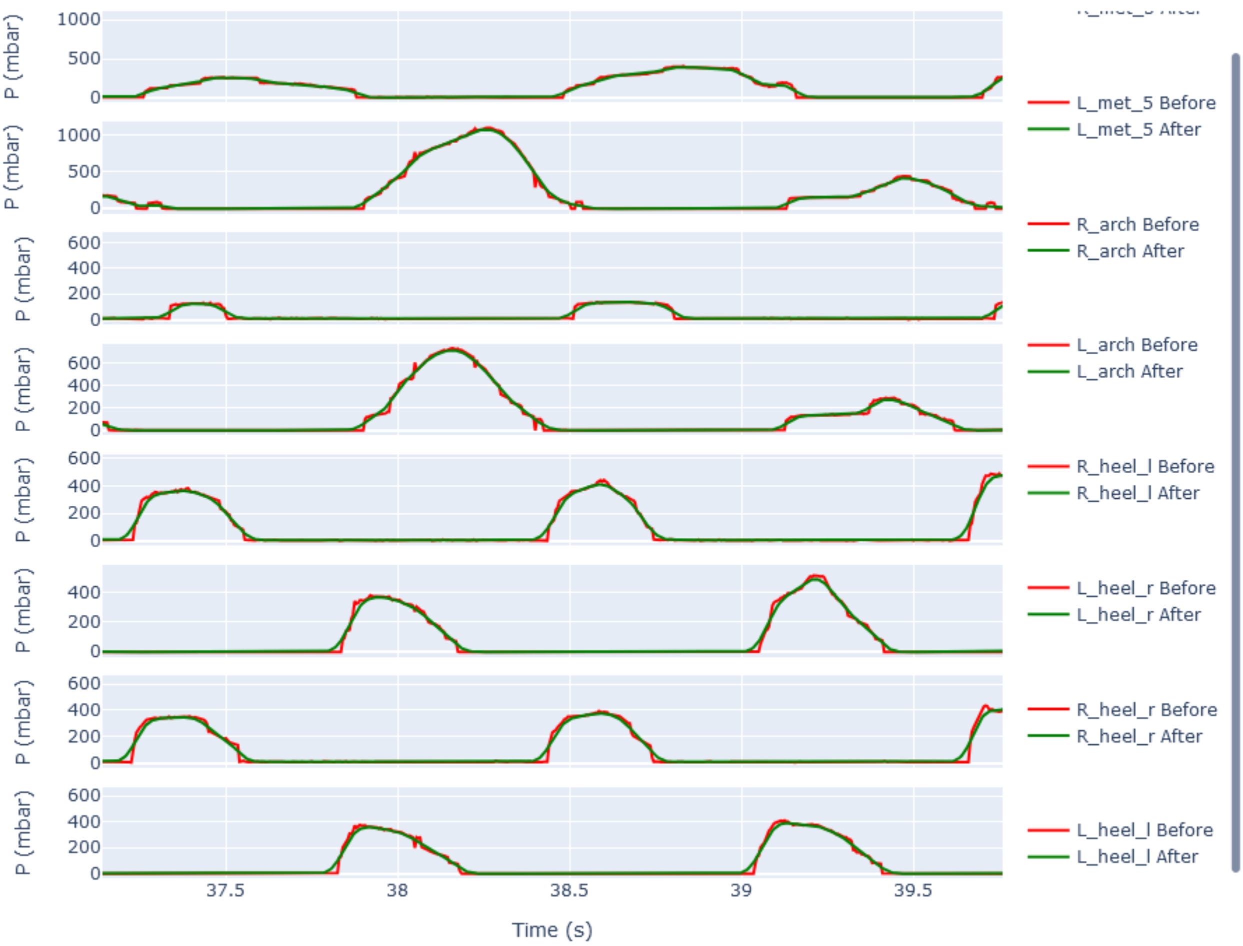


Figure 9. Gaussian Filtering applied on Pressure Data. Source: MSAI

b) Low Pass Filtering is applied on Accelerometer and Gyroscope data, Figure 10:

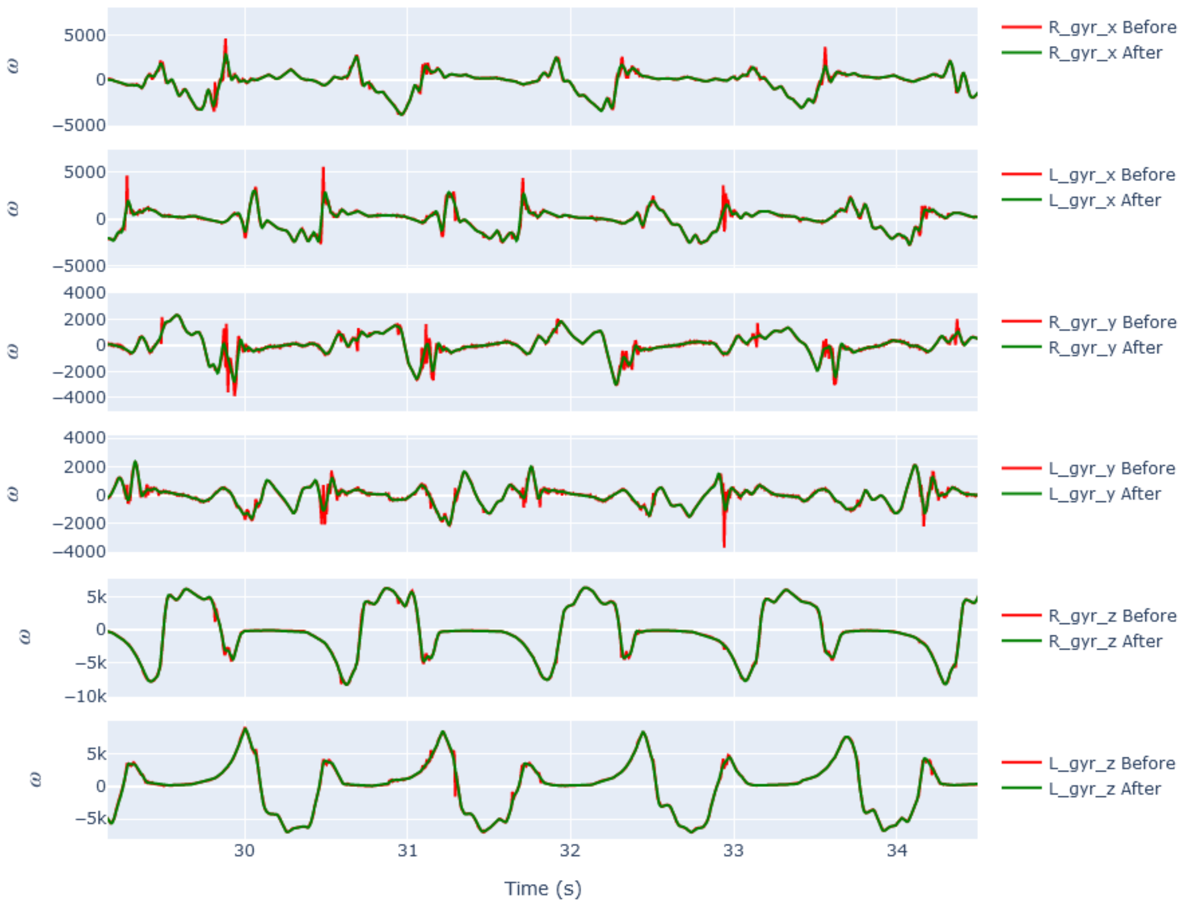


Figure 10. Low Pass Filtering applied on Gyroscope Data. Source: MSAI

2. Identification of "Key Events":

- From Gyroscope: Toe on, Heel off.

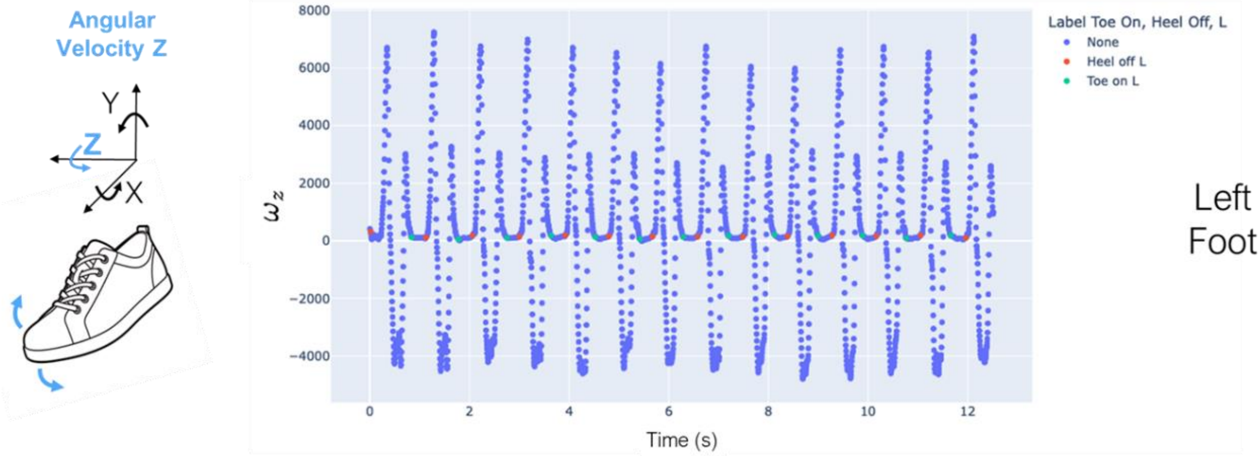


Figure 11. Identification of the Key Events "Toe on", "Heel off". Source: MSAI

- From Pressure: Heel on, Toe off.

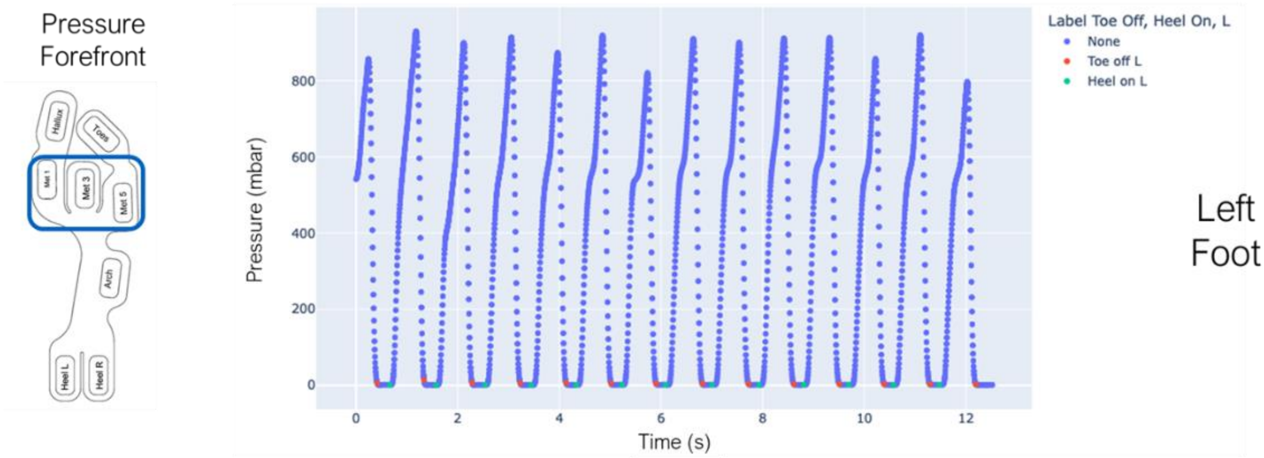


Figure 12. Identification of the Key Events "Heel on", "Toe off". Source: MSAI

3. Computation of Gait Cycle Parameters.

Once gait cycle is identified, the following calculation of the parameters is performed.

$$\text{Gait Cycle Duration (of Current Step)} = \text{Heel Strike of Next Step} - \text{Heel Strike of Current Step}$$

$$\text{Stance Phase} = \frac{\text{ToeOff} - \text{HeelStrike}}{\text{Gait Cycle Duration}} * 100$$

$$\text{Swing Phase} = \frac{\text{HeelStrike} - \text{ToeOff}}{\text{Gait Cycle Duration}} * 100$$

$$\text{Double Stance Phase} = \frac{Time\ \{both\ feet\ in\ Stance\ Phase\}}{\text{Gait Cycle Duration}} * 100$$

$$\text{Cadence} = 60\ \frac{\text{Total number of steps taken}}{\text{Time period (s)}}$$

$$\text{StepTime}_{\text{Right}} = \text{HeelStrike}_{\text{Left}} - \text{HeelStrike}_{\text{Right}}$$

$$\text{StepTime}_{\text{Left}} = \text{HeelStrike}_{\text{Right}} - \text{HeelStrike}_{\text{Left}}$$

4. Outlier Removal

Using the interquartile range (IQR) method where Q1 (lower quartile) and Q3 (third quartile) are calculated by taking the difference of Q3 and Q1 as a measure of variability. Data points outside the IQR are considered outliers and can be removed from the dataset to improve the normality of the distribution [18]. The reason this step is taken into consideration is to avoid technical outliers for the cases where the algorithm misidentifies steps because the signal was not "clear" enough to capture the key events correctly.

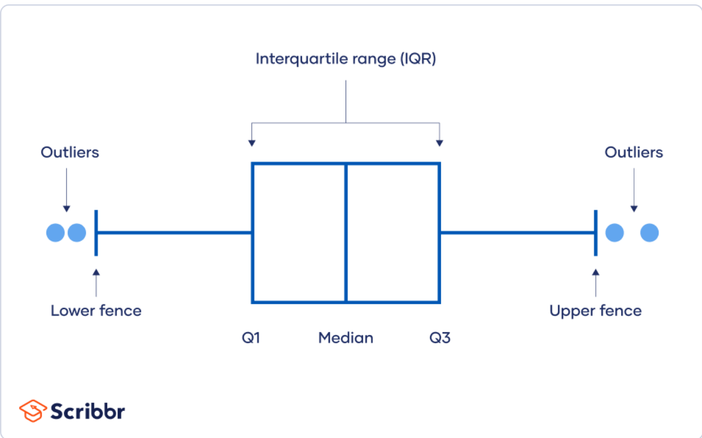


Figure 13. Outlier detection with IQR Method [23].

5. Generates a .csv file with the mean values of the parameters for Left and Right foot which we then use and process further.

	1 VarName1	2 Test	3 MeanGaitCycleDurationRs	4 CVOfGCDR	5 MeanGaitCycleDurationLs	6 CVOfGCDL	7 MeanStancePhaseR	8 CVOfStancePhaseR
1	0	"Walk 10 m,...	1.1888	0.0751	1.2035	0.0169	70.6551	0.0718
2	1	"Walk 10 m,...	1.2002	0.0134	1.1978	0.0437	74.0861	0.0147

Figure 14. Generated .csv file data, processed in MATLAB "GaitCycleFreatures_OutliersRemoved_ID"

6. Further analysis by Machine Learning.

   In order to separate PD and NPH by their gait, the MSAI study uses the GC parameters and train a ML classifier (Support Vector Machine with linear kernel).

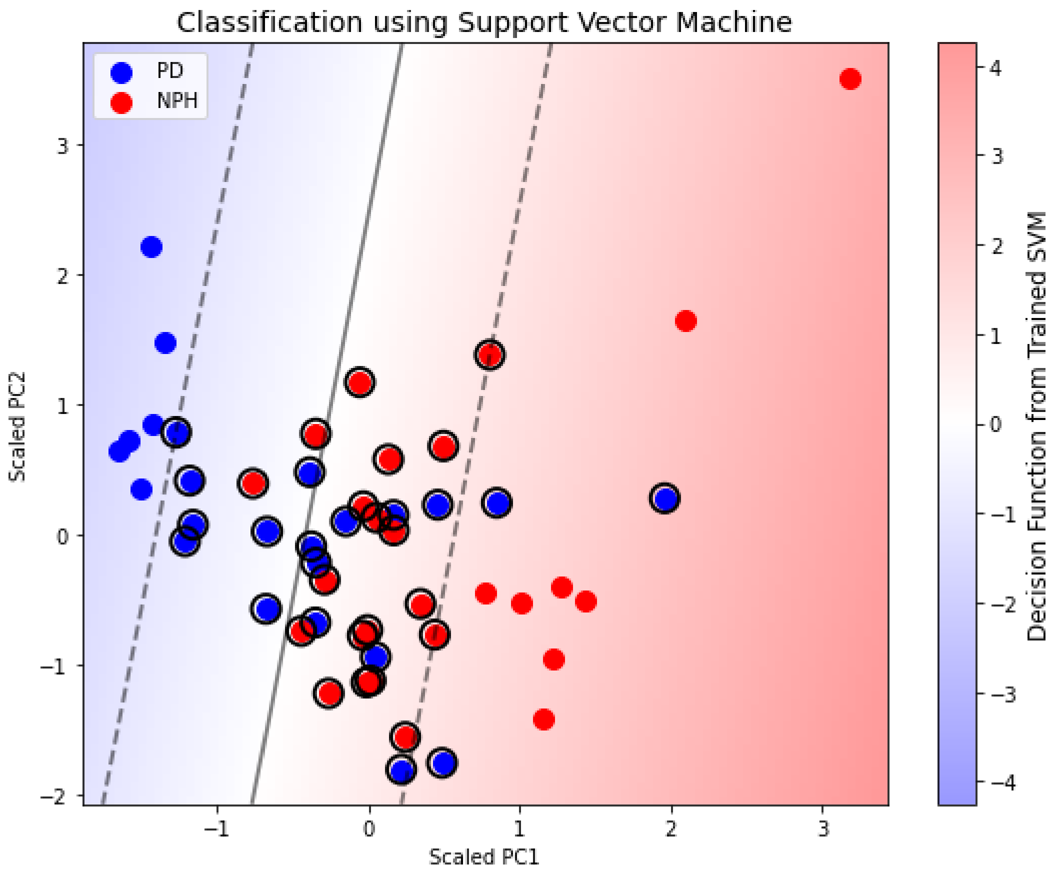


Figure 15. Linear Super Vector Machine used for ML. Source: MSAI

More information on the ML classifier, Support Vector Machine (SVM) is provided in section [8. Supplementary material](#).

2.2. Validation of the System

2.2.1. Objectives

This combination of two prototypes (Wearable Sensors + Computational Pipeline) cannot be considered accurate until we validate and calibrate it if required. For this purpose, we need to compare the gait cycle parameters estimated by our system to a high accuracy and precision system, measuring the same parameters, representing the "Golden standard". We had the opportunity to use an instrument treadmill ("Gaitway 3D", h/p/cosmos) in the facilities of the Trier University of Applied Sciences which records the gait of the subject and generates a report with over 40 gait parameters.

Figure 16. h/p/cosmos Gaitway 3D [24].

2.2.2. Participants

A group of 9 healthy subjects was recruited at Trier University of Applied Sciences. Statistics are presented in Table 1 and are given as Mean ± Standard Deviation, Counts in case of Sex and Range for Shoe Size.

Informed consent was obtained from all healthy volunteers for the two pilot studies. Both studies are part of the MoveSenseAI project. No medical supervision was required.

The first study, referred to as the VitalMove study, which focuses on the validation of the wearable sensor (ActiSense system) on the h/p/cosmos treadmill, was conducted at the

University of Applied Sciences in Trier, Germany, according to the principles of the Declaration of Helsinki with prior approval of the Ethics Committee of the State Chamber of Medicine in Rhineland-Palatinate, Germany

Table 1. Statistics of Healthy Subjects for Validation Study.

Biometrics	Parameters
Sex [m/f]	8/1
Age [years]	28.00 ± 4.97
Height [cm]	181.22 ± 7.33
Weight [kg]	80.04 ± 13.19
Shoe Size [EU]	[39 - 45]

Some key inclusion and exclusion criteria were defined in the proposal of the study. Key **inclusion** criteria for the healthy subjects:

- Age at the time of enrolment 18-85 years.
- No medical condition that affects the participant's gait pattern.
- Participants with shoe sizes 36 to 45.

Key **exclusion** criteria for the healthy subjects:

- Participants are unable to give informed consent.
- Participants are unable to follow instructions.
- Pregnant women, individuals with mental disorders, people with anxieties and mentally deficient persons.

2.2.3. Measurement Protocol

Each healthy subject was given a pseudo code (patient ID) and their data were anonymized. Step by step the protocol is as follows.

Part I

- Participant is welcomed and informed about the study. Given enough time to think about it.
- The purpose of the study is clear and possible questions were answered.
- Consent about data acquisition. The subject is aware and consents to acquiring and using their data for the study.
- Inclusion and exclusion criteria are checked intensively.

Part II

- Biometrics of the subject are being noted (Sex, Age, Shoe size (EU)).
- Height (m) is being measured (without shoes).
- We check the sensors and the synchrony between them and then the subject is equipped with ActiSense.
- There is a "Zero Calibration" step on the treadmill before the subject is allowed to step on.
- The subject is invited on the treadmill and via User Interface (UI) of the Gaitway 3D, their Weight (kg) is being acquired.
- The subject is equipped with an obligatory torso vest safely secured on a rope hanging on top of the treadmill with a safety feature so that in case of fall, the treadmill stops instantaneously and holds the subject before their knees touch the surface of the treadmill to avoid injuries.
- Everyone is asked if they have previous experience with a treadmill and in any case, they are given some time to get familiar with it. If there is no experience (which was never the case) even more time is given.

Part III

Each time before we start, stop, or change the speed, we announce it to the subject. For the speeds of 2, 3, and 4 km/h we give the subject a minute before we record each speed. It was suggested by Prof. Muller to randomize the runs and not have them in consecutive order for all subjects to avoid fatigue of any kind.

- Recording of the 3 runs at different speeds for 60 s each.
- Power off on ActiSense (via application).
- Progressively slow down treadmill to 0 km/h.

Part IV

- Remove Safety Straps.
- Remove ActiSense.
- Extract Sensor data and perform a quick raw data check. In the meanwhile, we allow the subject to see the Report of the treadmill, which is generated on the spot. We hide the monitors of the treadmill during the measurement to avoid giving biofeedback.
- As long as everything is okay with the data from the sensors, we thank the participant and walk him out of the room.

2.3. Gait Asymmetry

2.3.1. Objective

All of the parameters, including gait asymmetry, need to be accurately estimated so that subsequent analysis can be relied upon. This necessity to quantify and test gait asymmetry originally appeared through the gait analysis in MSAI over the past 2 years. Strong asymmetry was noticed in the gait of PD and partially NPH patients (within the MSAI project) while examining their GC parameters and in particular Step Time. While Step Time was considered a reliable parameter, it was found out within the MSAI project that for distinguishing both patient groups, PD vs NPH, in many cases it provided a concerning value for "Squared Weight in the SVM decision function" implemented in the ML algorithm when it comes to relevance of features to differentiate Right vs Left steps. This significant difference over the other parameters could be the key to providing the best results but it also raised questions. Initial thoughts were that the walking aid of some patients or the nurse providing mandatory support to prevent risk of fall was causing this great impact on the Step Time of those patients. Although this could be an external factor (walking aid/support) it could also be part of the disease or the gait of the individual. Thus, to examine this parameter further and eliminate the factors that could "harm" our results, we recruited another group of healthy subjects to perform the "10 Meters Walk" and compare it to the influence of artificially induced asymmetry to their gait. The idea to which we concluded as the best was to add weight load to one of the subjects' lower limbs, heavy enough to have a significant influence on their gait but not so heavy to "anchor" them. With ActiSense equipped, we asked the subjects to perform the standardized clinical test 2x "10 Meters Walk" once to record their "Control Walk". Afterwards, with commercial gym weights with Velcro straps (Figure 17) with a total weight load of 2.6 kg thanks to having two of them attached to one foot at a time to perform two more 2x "10 Meters Walk" to record the "Weight load on Right foot" and "Weight load on Left foot". Asymmetry ratios were computed, dividing each parameter value for the right foot by the corresponding value for the left foot.

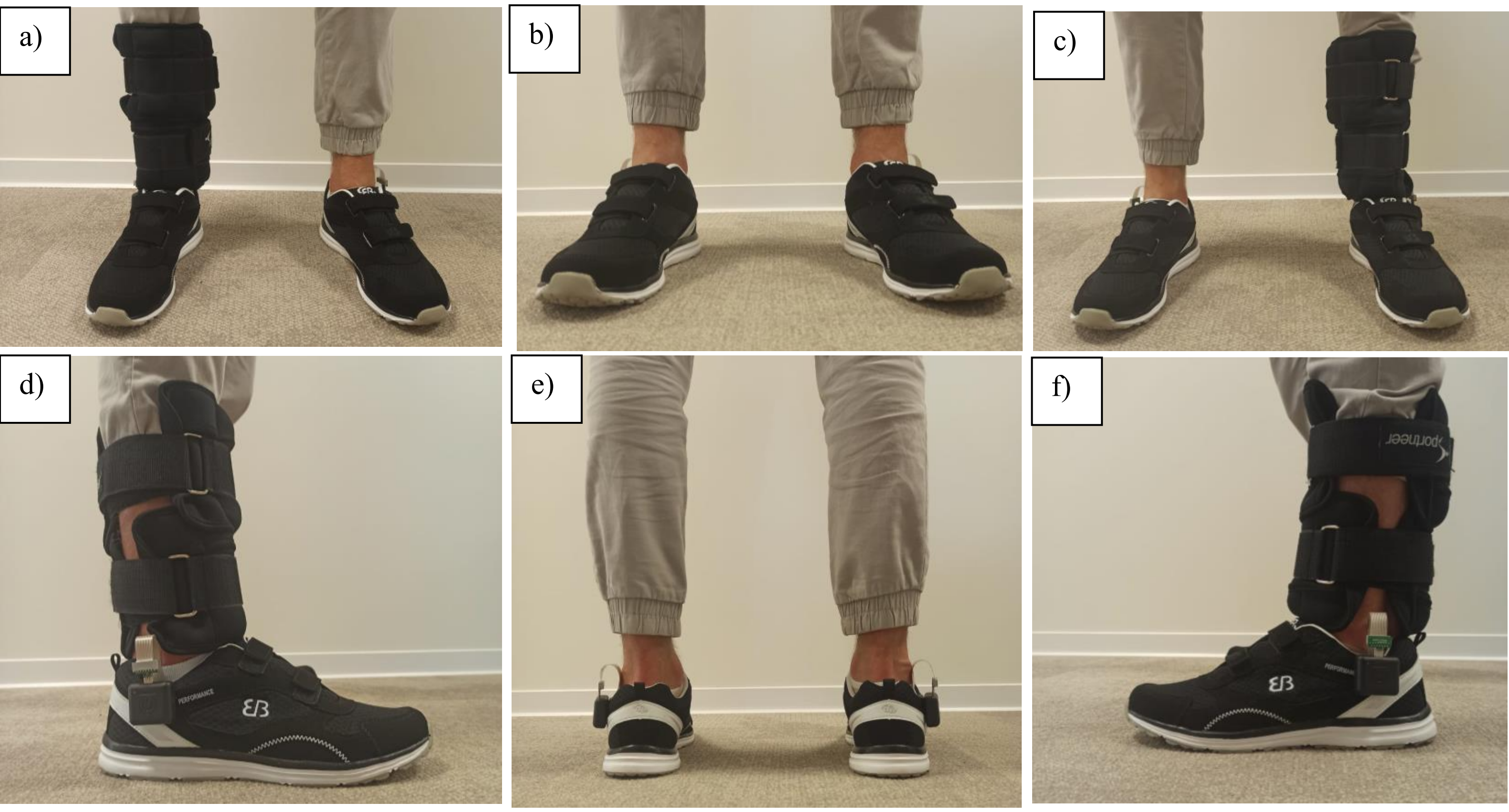


Figure 17. Weight Load with Velcro Straps setup.
a) Weight load on Right foot (front view), b) Control (no weight, front view), c) Weight load on Left foot (front view)
d) Weight load on Right foot (side view), e) Control (no weight, rear view), f) Weight load on Left foot (side view)

2.3.2. Participants

A group of 12 healthy subjects was recruited at the Luxembourg Centre for Systems Biomedicine. Statistics presented in Table 2 are given as Mean ± Standard Deviation, Counts in case of Sex and Range for Shoe Size.

Informed consent was obtained from all healthy volunteers for the two pilot studies. Both studies are part of the MoveSenseAI project. No medical supervision was required.

The second study, referred to as MoveSenseAI Clinical Study (CNER, N. 202101/02 V2.0), which focuses on the analysis of asymmetry using the wearable sensor (ActiSense system), was conducted at the Centre Hospitalier de Luxembourg according to the principles of the Declaration of Helsinki and received a favorable opinion from the Ministry of Health and the National Research Ethics Committee (CNER) Luxembourg.

Table 2. Statistics of Healthy Subjects for Asymmetry Study.

Biometrics	Parameters
Sex [m/f]	6/6
Age [years]	32.66 ± 8.41
Height [cm]	172.08. ± 1.35
Weight [kg]	67.91 ± 15.54
Shoe Size [EU]	[37 - 45]

For this study the inclusion and exclusion criteria were the same as the previous.

Key **inclusion** criteria for the healthy subjects:

- Age at the time of enrolment 18-85 years.
- No medical condition that affects the participant’s gait pattern.
- Participants with shoe sizes 36 to 45.

Key **exclusion** criteria for the healthy subjects:

- Participants are unable to give informed consent.
- Participants are unable to follow instructions.
- Pregnant women, individuals with mental disorders, people with anxieties and mentally deficient persons.

2.3.3. Measurement Protocol

Each subject was given a pseudo code (Patient ID) and their data were anonymized. Step by Step the protocol is as follows.

Part I

- Participant is welcomed and informed about the study. Given enough time to think about it, the purpose of the study is clear and possible questions were answered.
- Consent about data acquisition. The subject is aware and consents to acquiring and using their data for the study.
- Inclusion and exclusion criteria are checked intensively.

Part II

- Biometrics of the subject are noted (Sex, Age, Height (m), Weight (kg), Shoe size (EU)).
- In the meanwhile, we check the sensors and the synchrony between them before the subject is equipped with ActiSense.

  Marked on the floor we had the Starting point (-2 meters) the Recording Start Point (0 meters), the Recording End Point (10 meters) and the End point (12 meters). We added the ± points to let the subject reach walking speed so that the measurement was a pure "10-meter-walk" at regular walking speed.

- Recording of the "Control Walk", 2x "10 Meters Walk".
- Adding weight load on Right foot and recording of the "Weight load on Right foot", 2x "10 Meters Walk".
- Placing weight load on Left foot and recording of the "Weight load on Left foot", 2x "10 Meters Walk".

Part III

- Power off of ActiSense (via application).
- Remove weight load.
- Remove ActiSense.
- Extract Sensor data and perform a quick raw data check.
- As long as everything is okay with the data from the sensors, we thank the participant.

3. Results

3.1. Validation versus Treadmill Results

Data acquired by our system were processed with MATLAB, after running our python pipeline to extract gait cycle parameters. Due to some differences such as, our stance phase being expressed in a percentage of the GC and Gaitway 3D giving that result in time (ms), or our Cadence being measured as Cadence-Right and Cadence-Left while the treadmill giving a Cadence for both feet, some gait cycle parameters had to be transformed in order to be compared. This procedure was automated in a MATLAB script which I developed ex novo.

Presented in this chapter, the results of each parameter extracted from the sensors are compared to those from Gaitway 3D. Each plot represents a comparison of parameters from Wearable with respect to those from Treadmill with a mean and ± Standard deviation for both systems (except for Cadence of ActiSense, where only the mean is available). We repeated the comparison, over the different subjects at different speeds. With the same order for each subject, SubjectXX-Ykm/h-Sensor, SubjectXX-Ykm/h-Treadmill (XX= ID of subject, Y=Treadmill speed value).

The subplots below use formula (1) which displays the percentage error of the sensor for the respective parameter compared with respect to the treadmill values centered at 0.

$$Error\ \% = \frac{Parameter\ Wearable - Parameter\ Treadmill}{\text{Parameter Treadmill}} * 100 \quad (1)$$

3.1.1. Gait Cycle Duration

In Figure 18 we can compare the values of gait cycle duration between wearable and treadmill and see that not only the mean value of our system is very close to the Treadmill value but also that the standard deviation is only lightly larger for our system. We can also see that we follow the treadmill values, and the gait cycle duration values of both systems have a tendency which appears to grow almost linearly with increasing speeds. We also present on Figure 19 the average percentage error between wearable and treadmill for the Gait Cycle Duration parameter for every speed recorded plus the average for all the speeds. As mentioned, treadmill values have been centered at 0, representing the ground truth and for a better visualization of the comparison. Overall, for the gait cycle duration we find a very good agreement between wearable and treadmill.

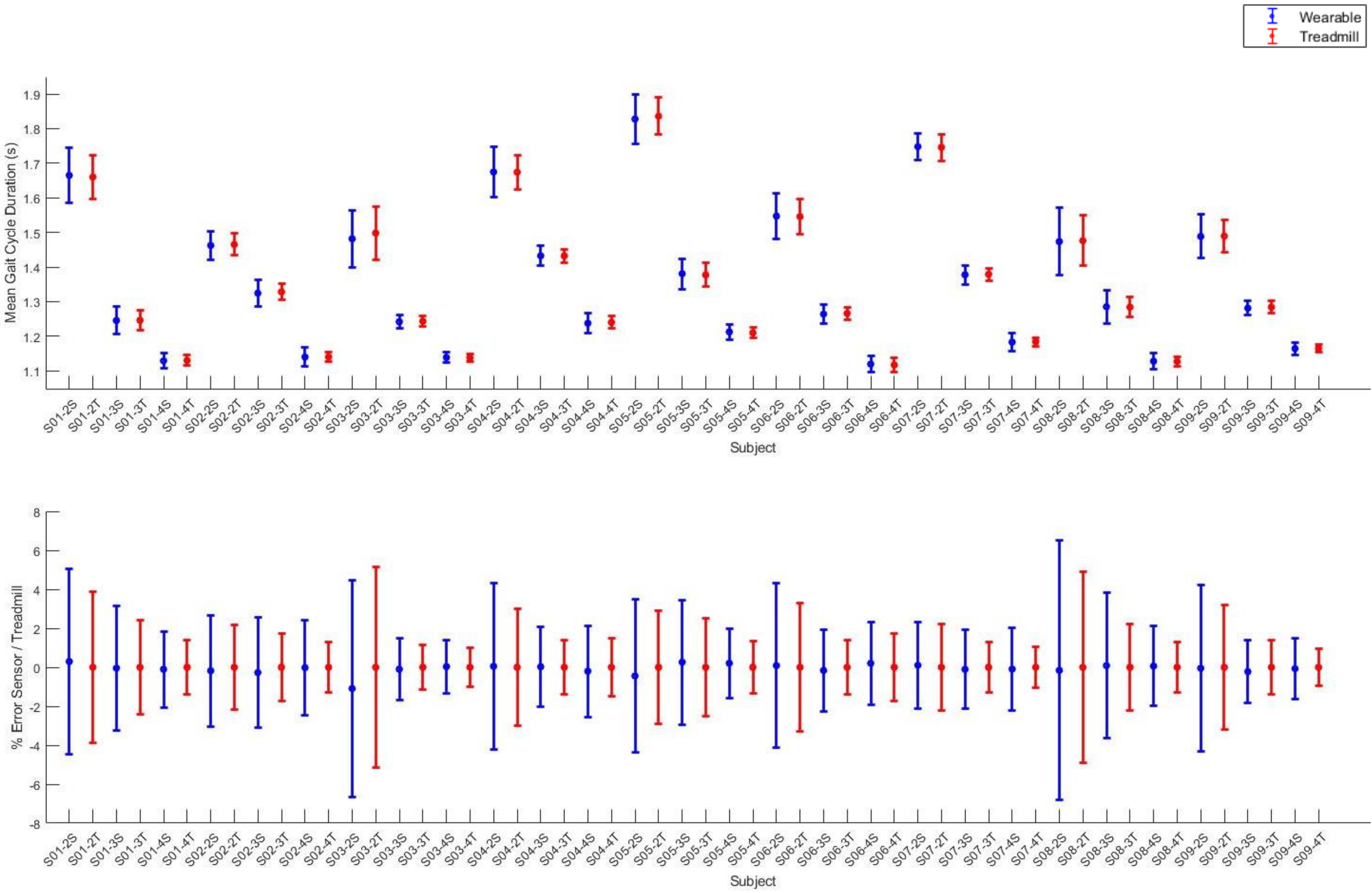


Figure 18. Validation Analysis of Gait Cycle Duration.

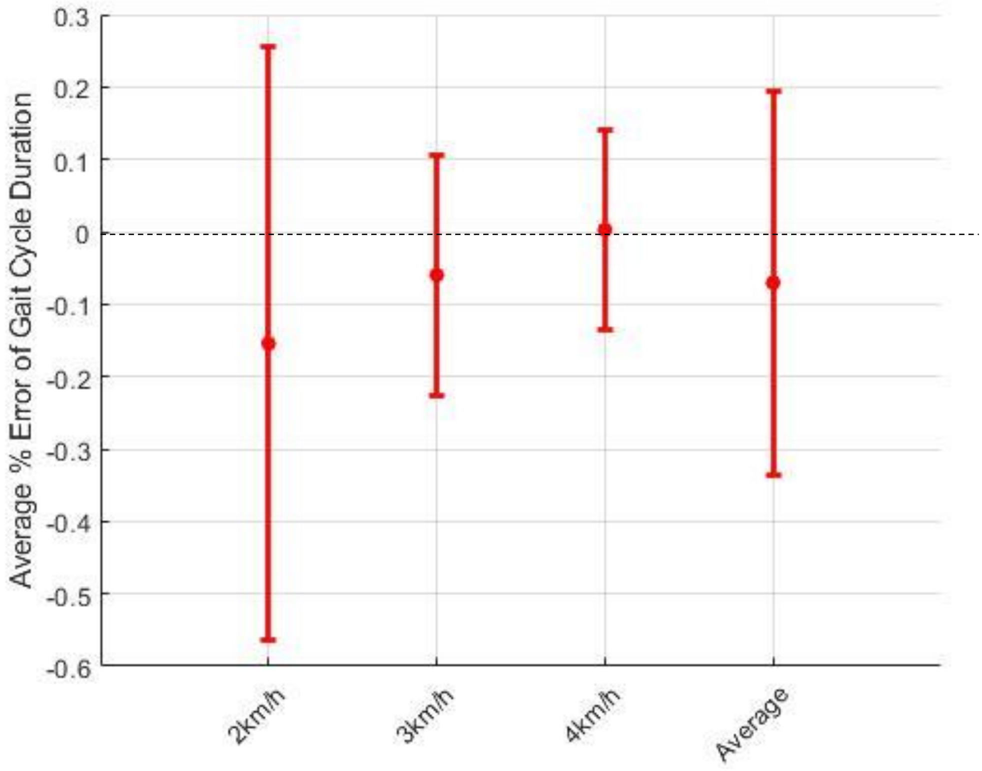


Figure 19. Average % Error of Gait Cycle Duration Analysis.

3.1.2. Cadence

On Cadence, Figure 20, comparing the values of wearable and treadmill, mean value of wearable is within the mean ± Standard Deviation (SD) provided by the treadmill and the error presented in Figure 21 is less than 1% in all cases. Unfortunately, in this case we do not provide an SD for the number of steps. While the treadmill provides an integer and an SD, we give a number of steps/min approximated to a few decimals. If we were to round our number to the closest integer, we would be on the exact same number for Cadence. Overall, for cadence we find a very good agreement between wearable and treadmill.

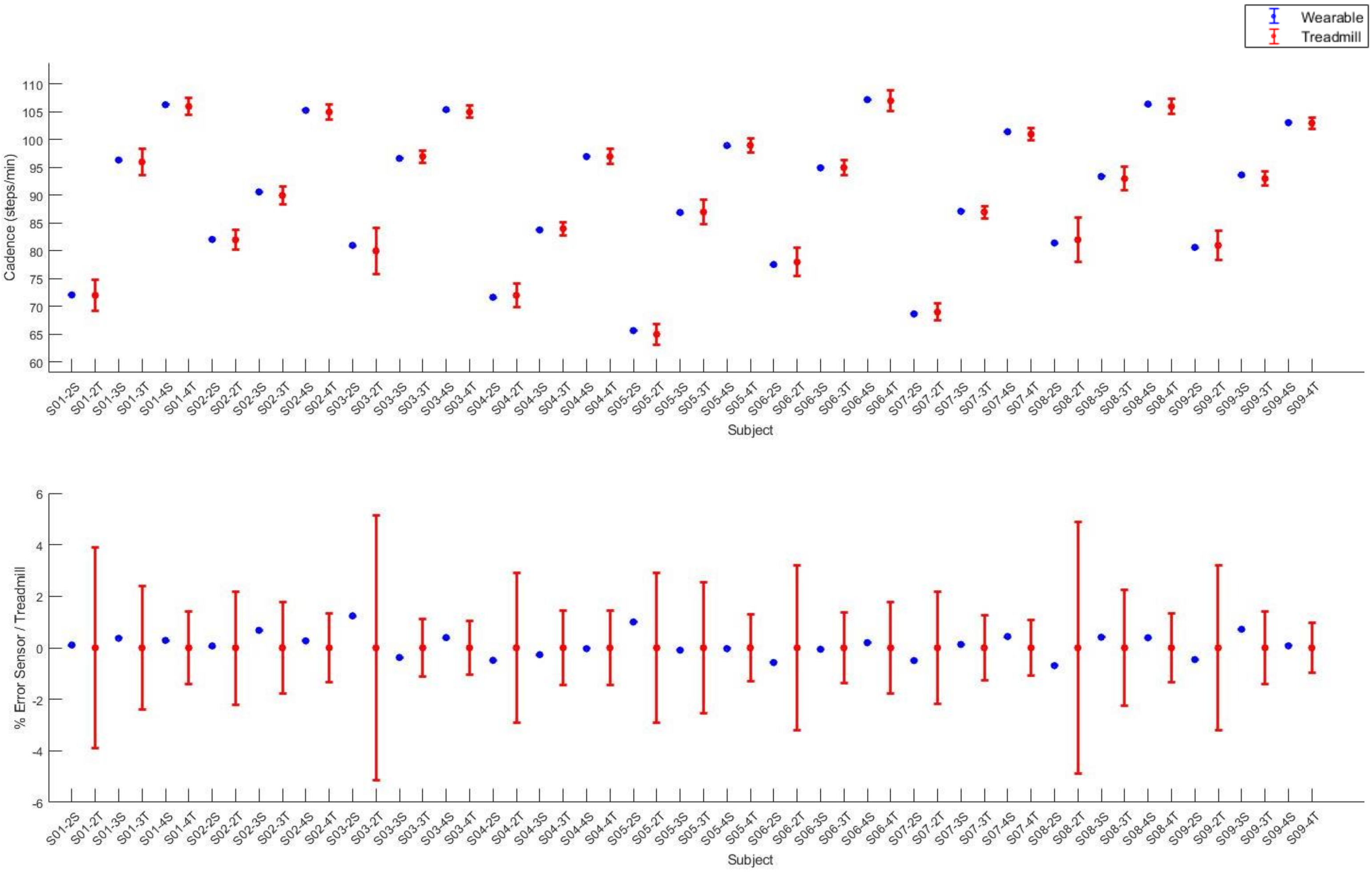


Figure 20. Validation Analysis of Cadence.

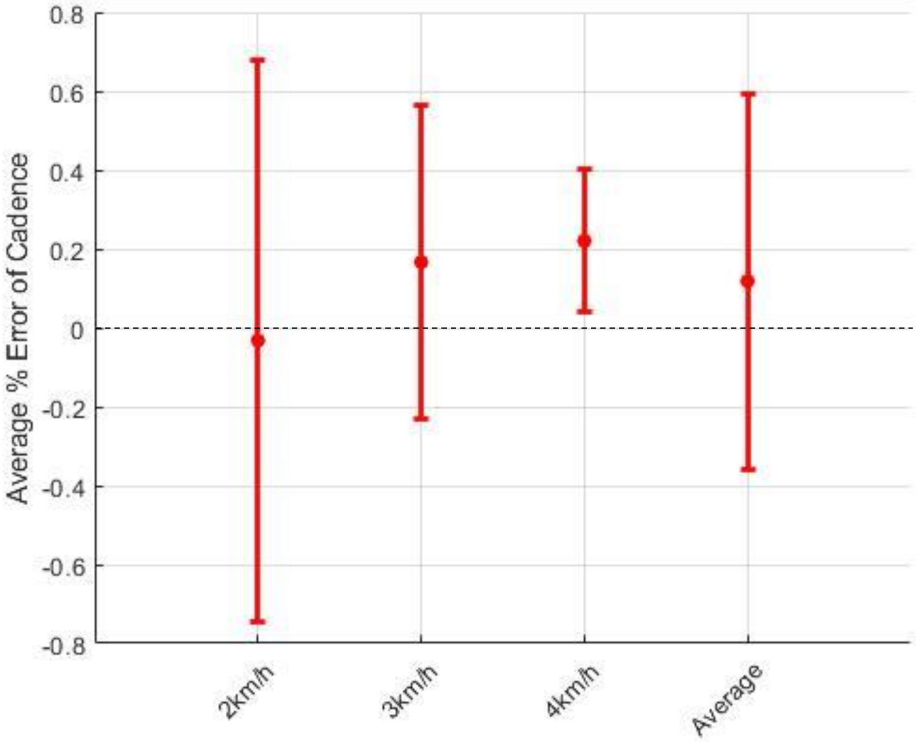


Figure 21. Average % Error of Cadence Analysis.

3.1.3. Stance Phase

From Figure 22 we can see that the Stance Phase is not exactly the same between treadmill and wearable, the wearable values are close to the Treadmill values, but with a systematic shift. It is important to note that our values have a systematic shift, increasing with speed, always in the same direction, of a similar magnitude such that the wearable value overestimates the treadmill value by 11.26% on average, as displayed on Figure 23.

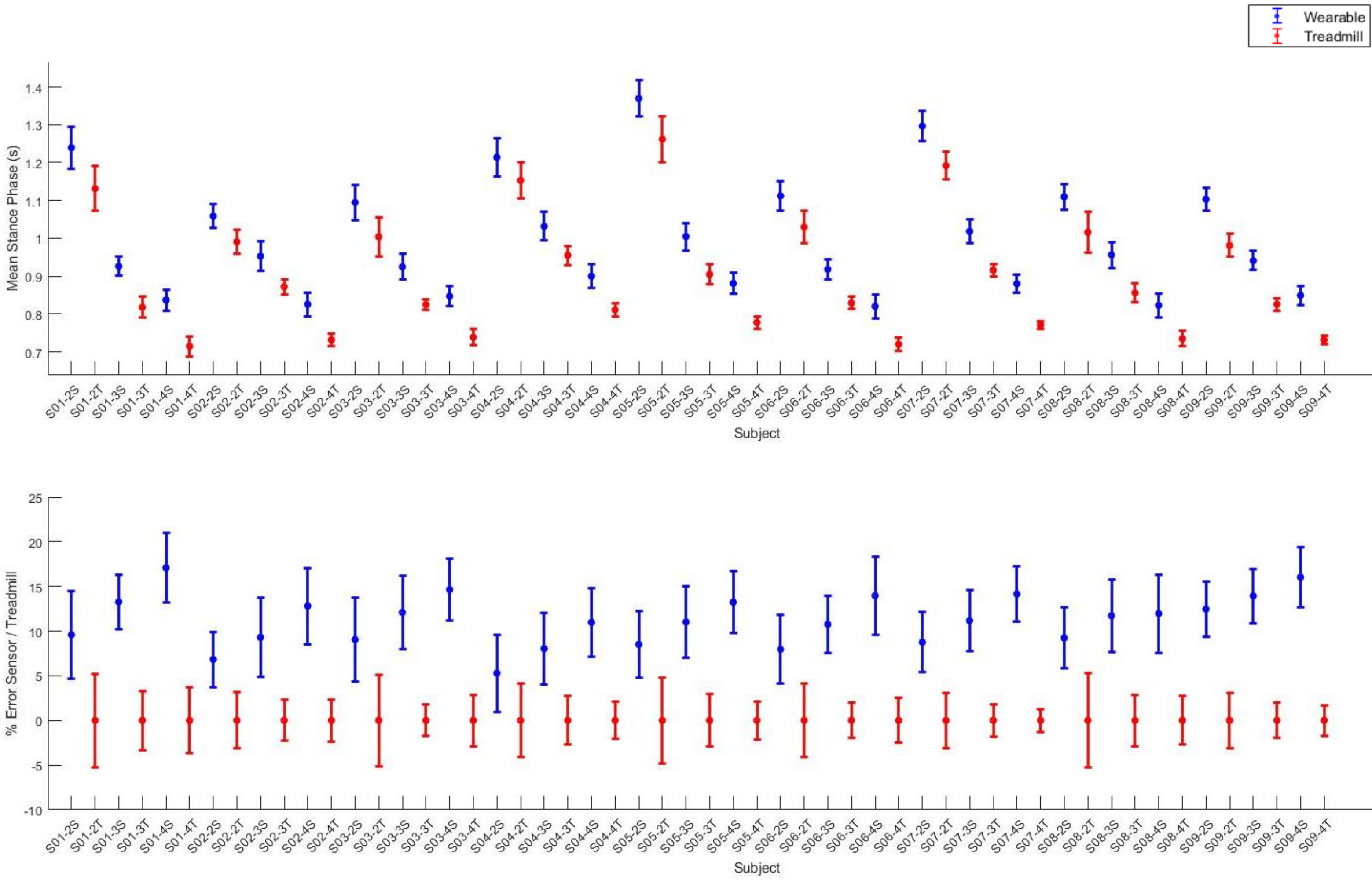


Figure 22. Validation Analysis of Stance Phase.

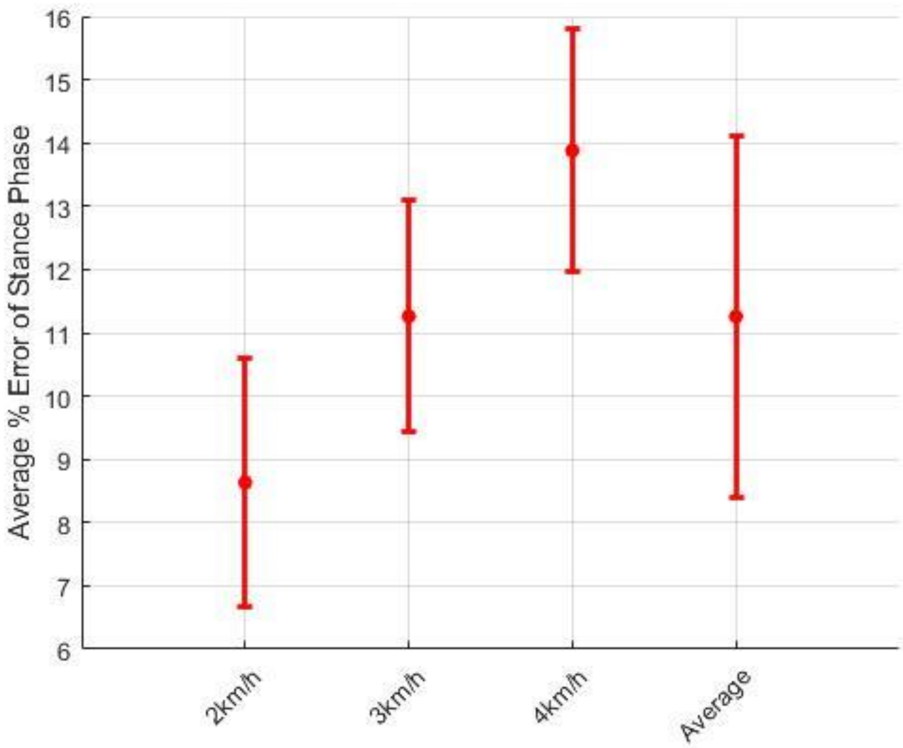


Figure 23. Average % Error of Stance Phase Analysis.

3.1.4. Swing Phase

Swing Phase again has a systematic error as the Stance phase, which is understandable given that they must sum up to 100% of the GC. Again, always on the same side and of a similar magnitude, increasing with speed and underestimating the parameter by around 21.92% on average the value provided by the treadmill (Figure 25).

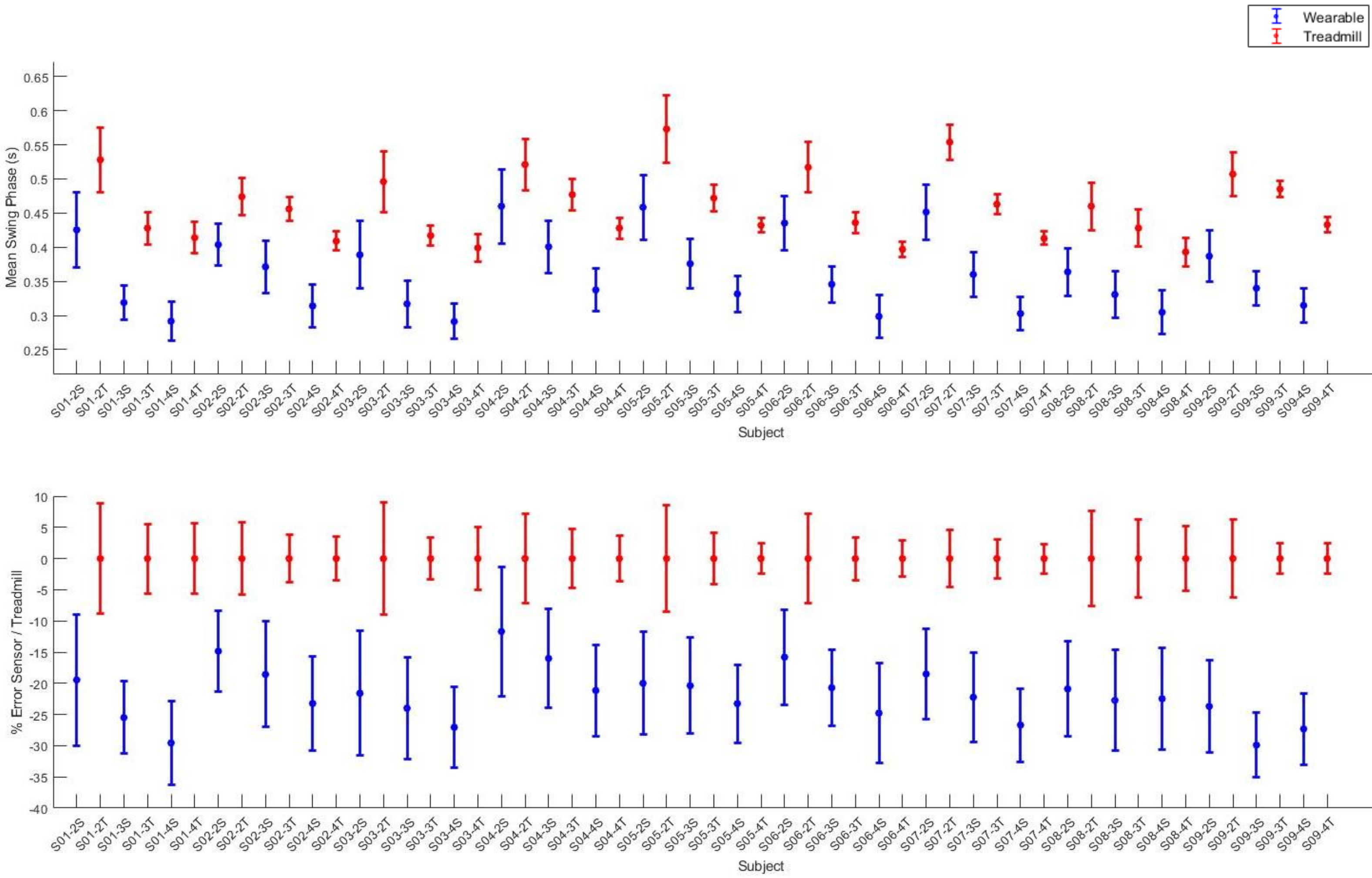


Figure 24. Validation Analysis of Swing Phase.

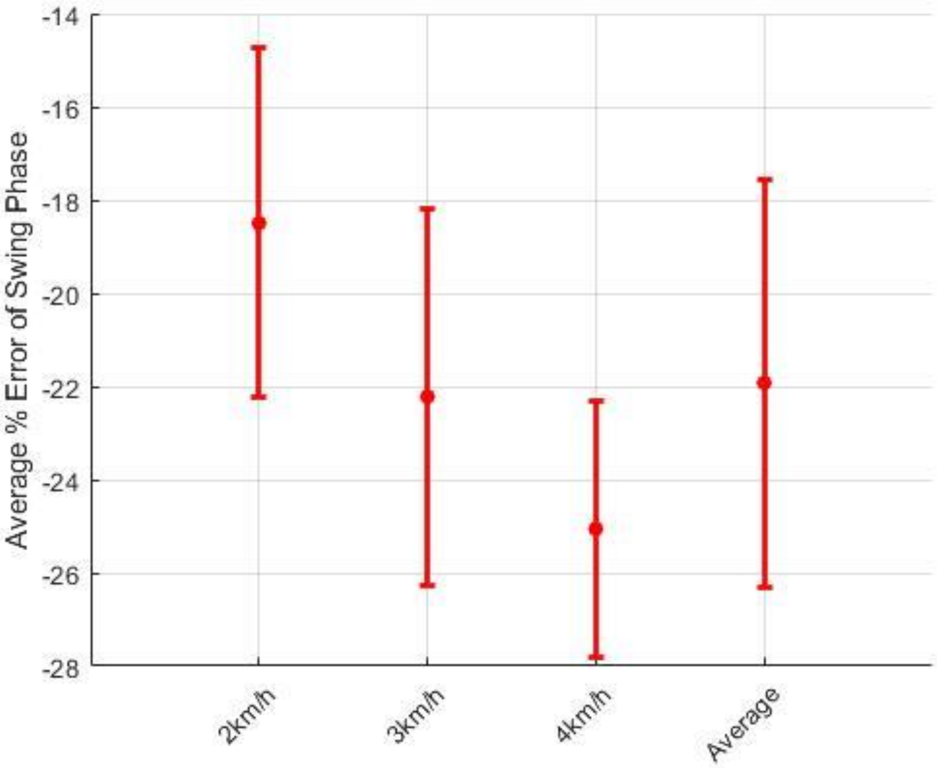


Figure 25. Average % Error of Swing Phase Analysis.

3.1.5. Double Stance Phase

In Figure 26 we can see that the wearable values are very far away from the treadmill ones for the Double Stance Phase estimation. It is the combination of the systematic errors of Stance Phase and Swing Phase mentioned previously, this time overestimating by 194.80% on average, presented in Figure 27. For Stance, Swing and Double Stance phase we could identify and characterize the source of these systematic errors and propose ways to fix it to a certain extent, which is one of the main results of this thesis; this will be presented in chapter 4.1.2.

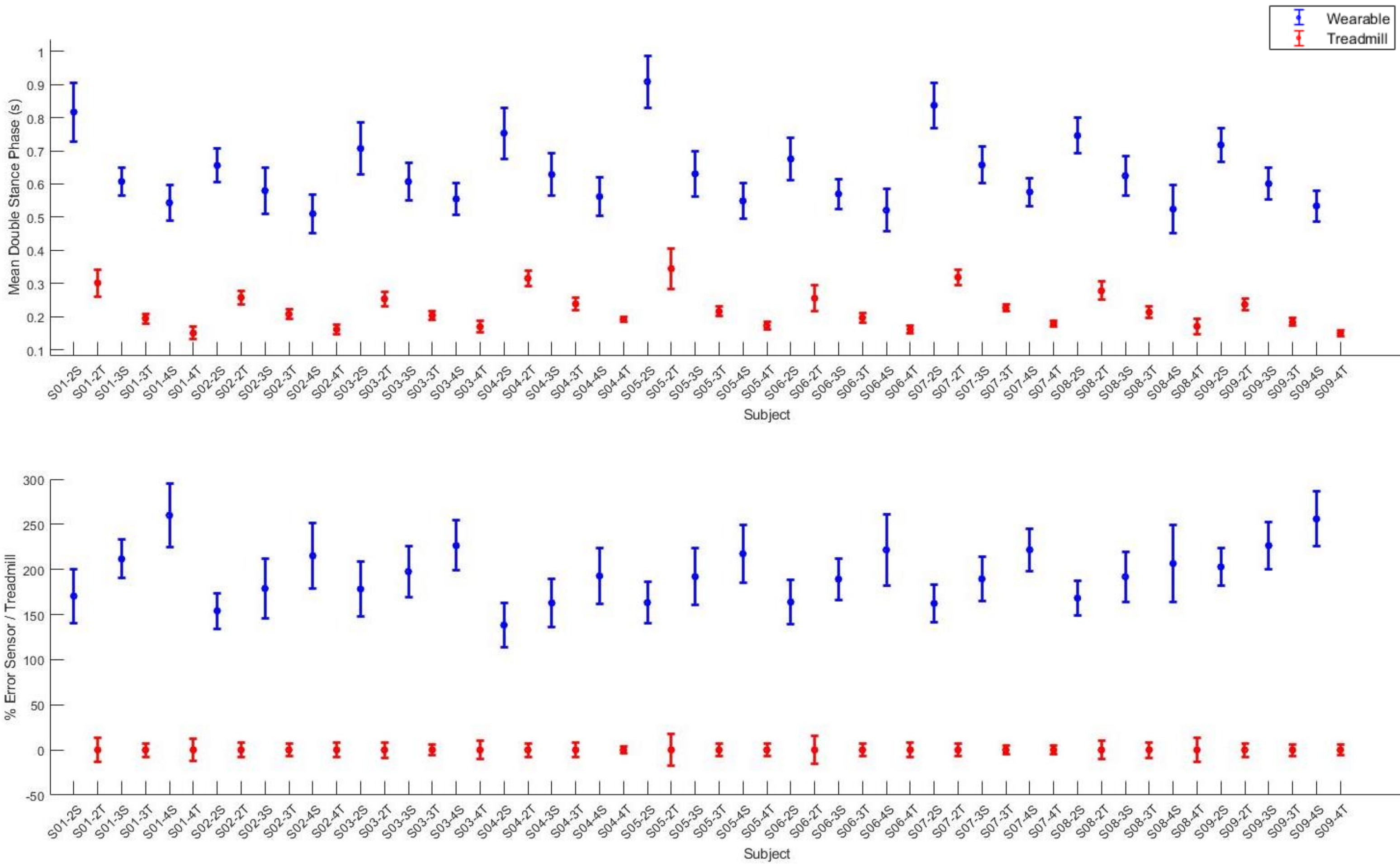


Figure 26. Validation Analysis of Double Stance Phase.

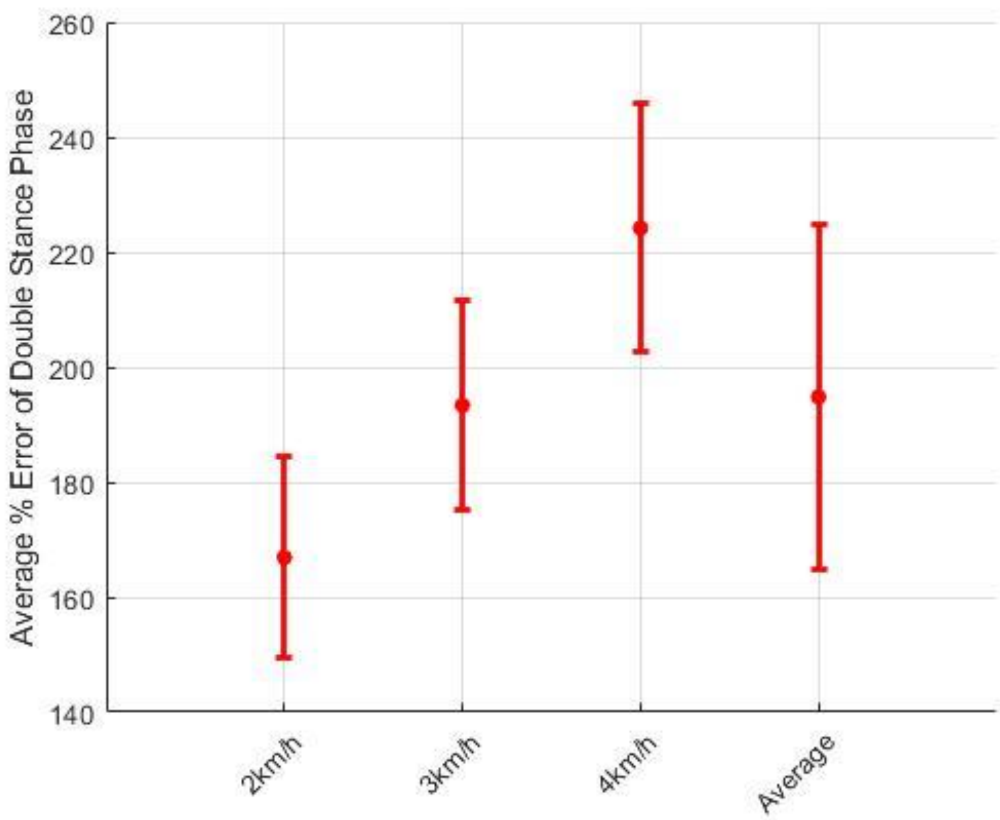


Figure 27. Average % Error of Double Stance Phase Analysis.

3.1.6. Step Time

Step Time, Figure 28, is also a parameter in which we find an agreement between wearable and treadmill. Wearable values are very close to the treadmill value and error in Figure 29 is less than 1%. As mentioned earlier, some strong asymmetry was noticed in the ratio of Step Time for certain subjects, this required further investigation. As we can see in figures 30 and 31 for each subject if the mean of the error for StepTimeL is positive the mean of the error for StepTimeR is negative, and the other way around. In some cases (S01-4, S02-4, S03-4, S07-3) the difference is obvious compared to the rest of the results, and it appears of an equal magnitude and opposite sign for the opposite, corresponding foot.

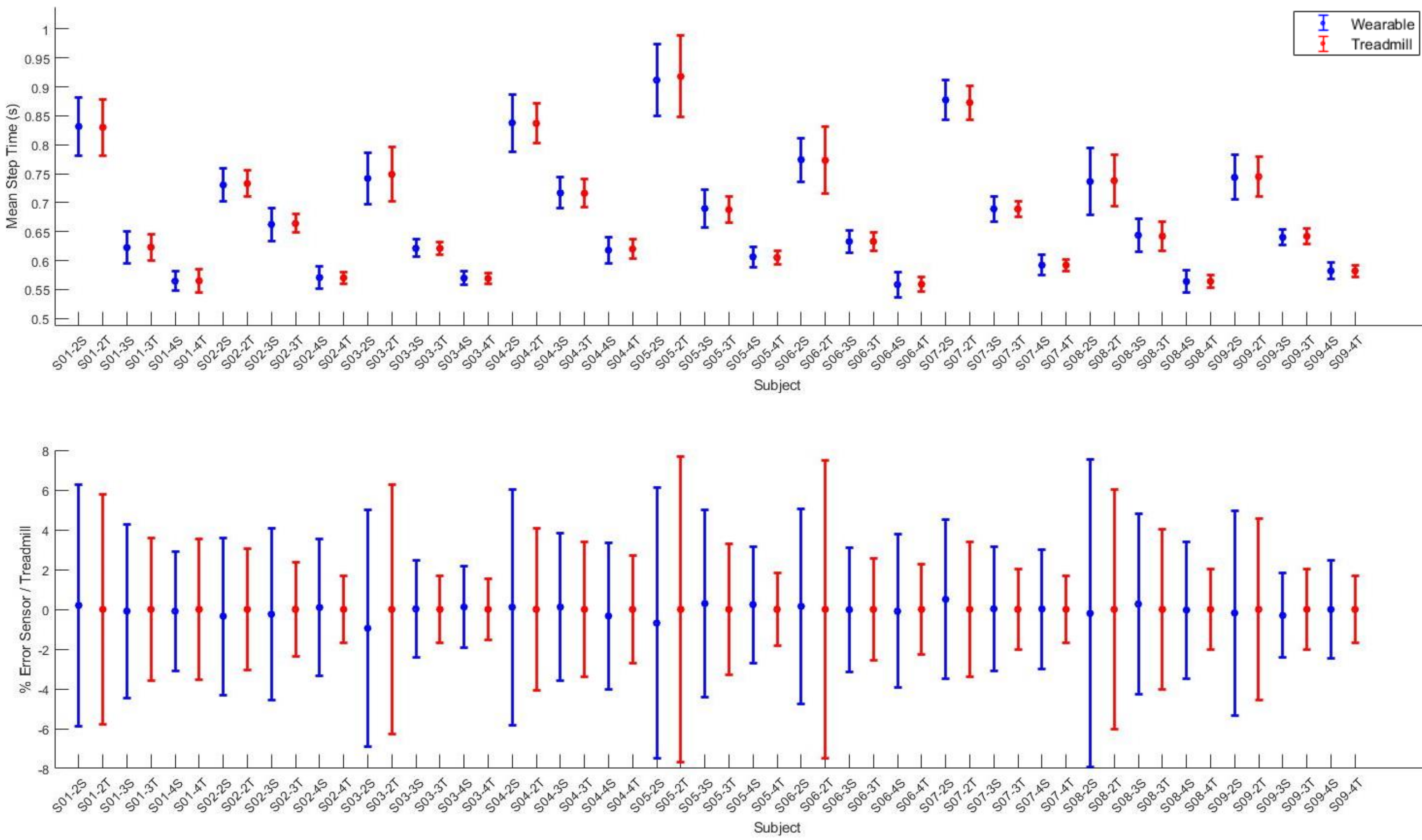


Figure 28. Validation Analysis of Step Time.

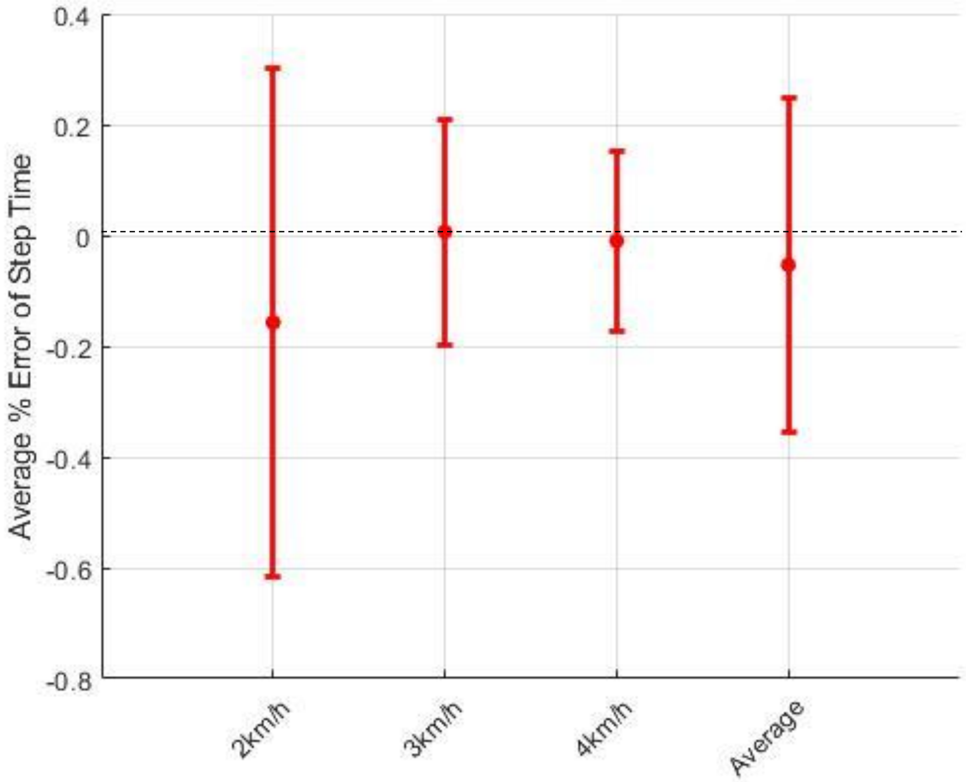


Figure 29. Average % Error of Step Time Analysis.

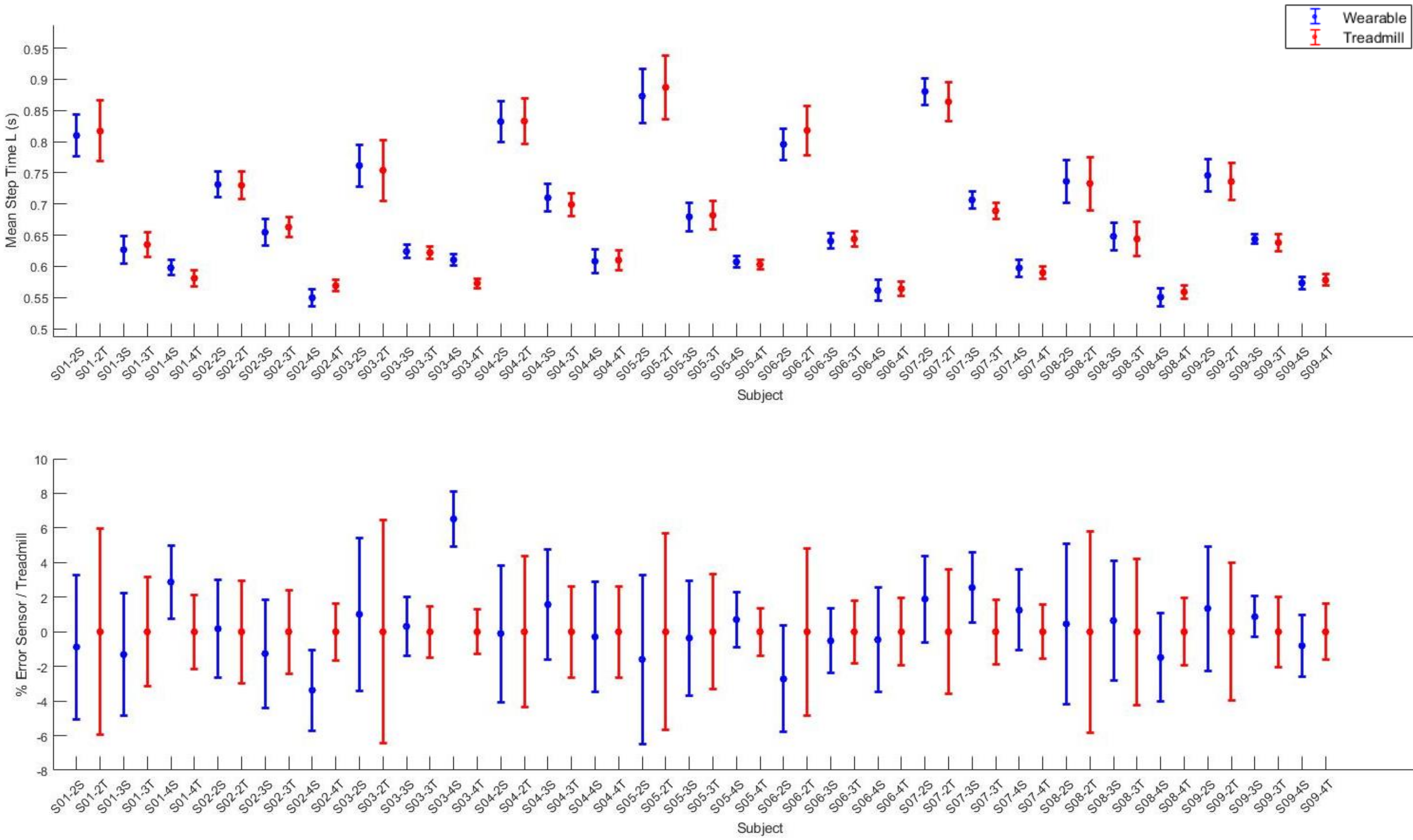


Figure 30. Validation Analysis of Step Time Left.

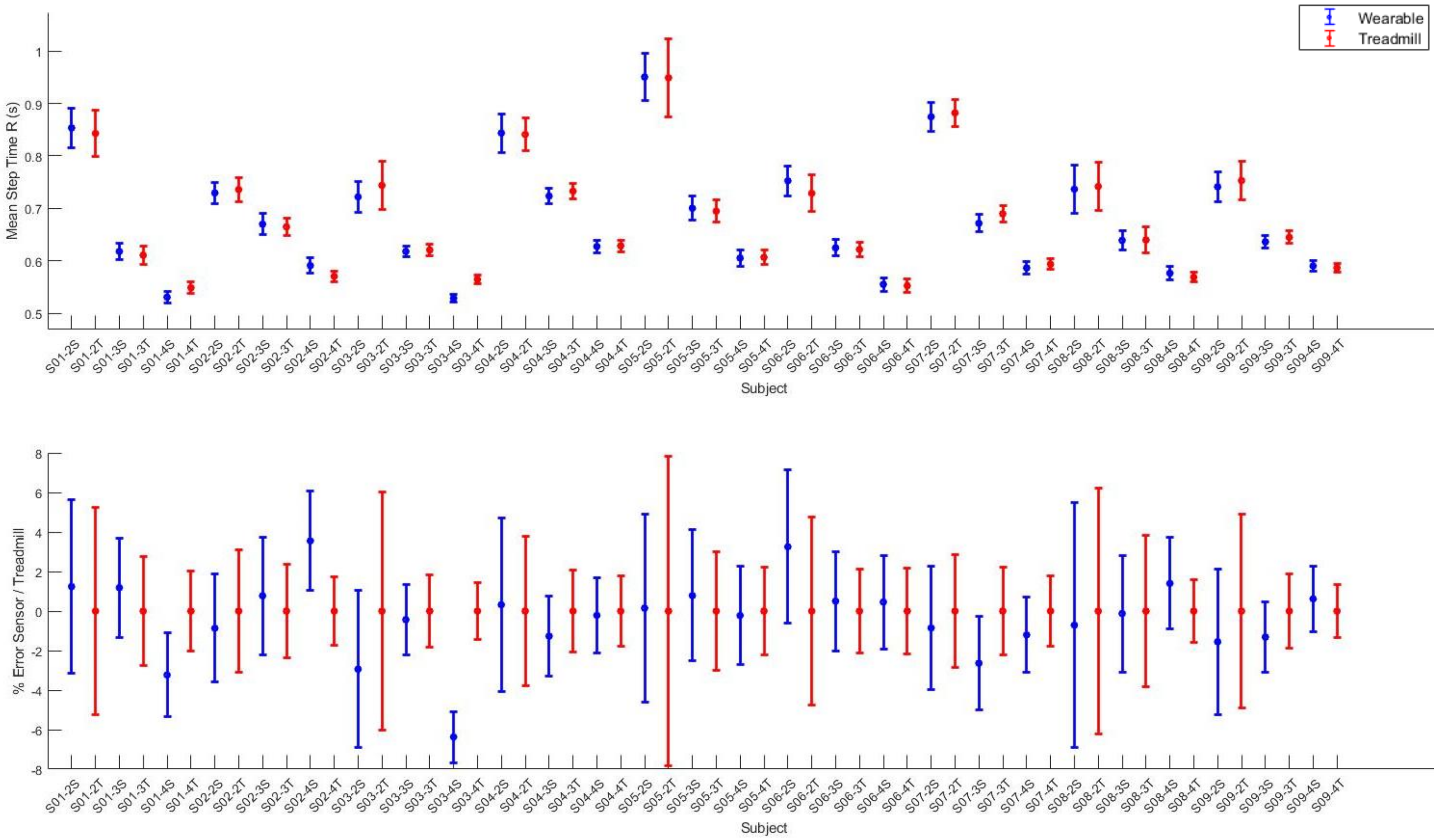


Figure 31. Validation Analysis of Step Time Right.

3.2. Asymmetry Results

For the asymmetry between left and right gait cycle parameters, we represent in Figures 32-37 the symmetry ratios with symmetry ratio of 1st "10 Meters Walk" on x-axis and Symmetry ratio of 2nd "10 Meters Walk" on y-axis, for each gait cycle parameter. We compute the symmetry ratio according to (2) so the result is a positive number, 1 represents completely symmetric gait parameter, greater than 1 means Parameter Right > Parameter Left and smaller than 1 means Parameter Right < Parameter Left.

$$\text{Symmetry Ratio} = \frac{Parameter\ Right}{Parameter\ Left} \qquad (2)$$

3.2.1. Gait Cycle Duration

What we observe in Figure 32 is that, for all the participants' walks, the symmetry ratios are gathered very closely around 1, which should be the case by definition because the gait cycle duration should not display any asymmetry, being defined as the time to perform a full left and a full right step.

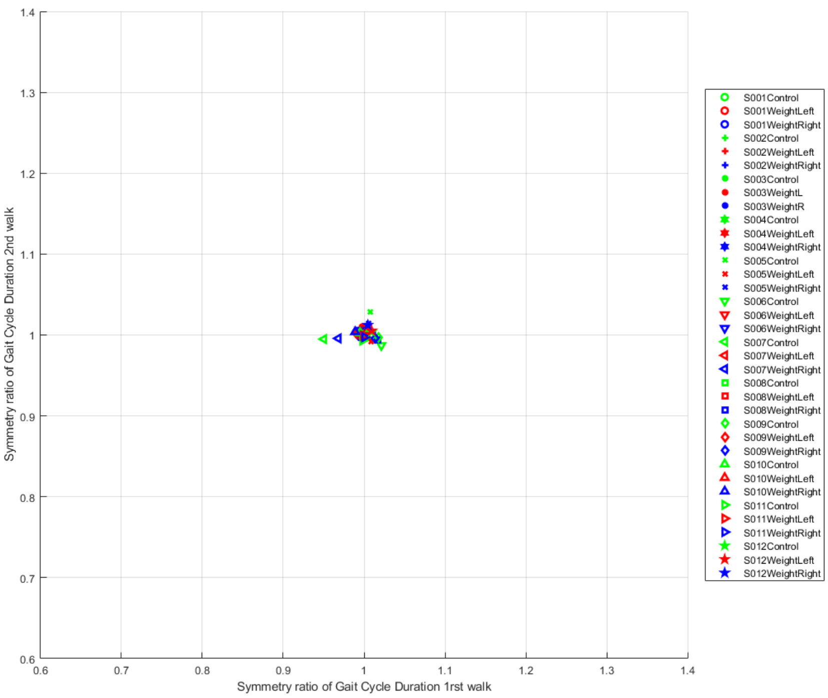


Figure 32. Symmetry Ratio of 1st and 2nd walk for Gait Cycle Duration.

3.2.2. Cadence

In cadence we observe the same result as Gait Cycle duration. This again is expected, according to the definition of cadence, which being the number of steps per minute, should not change considerably depending on from which side we count (for every right step, also a left step is performed, subjects walking do not make two consecutive steps with the same foot).

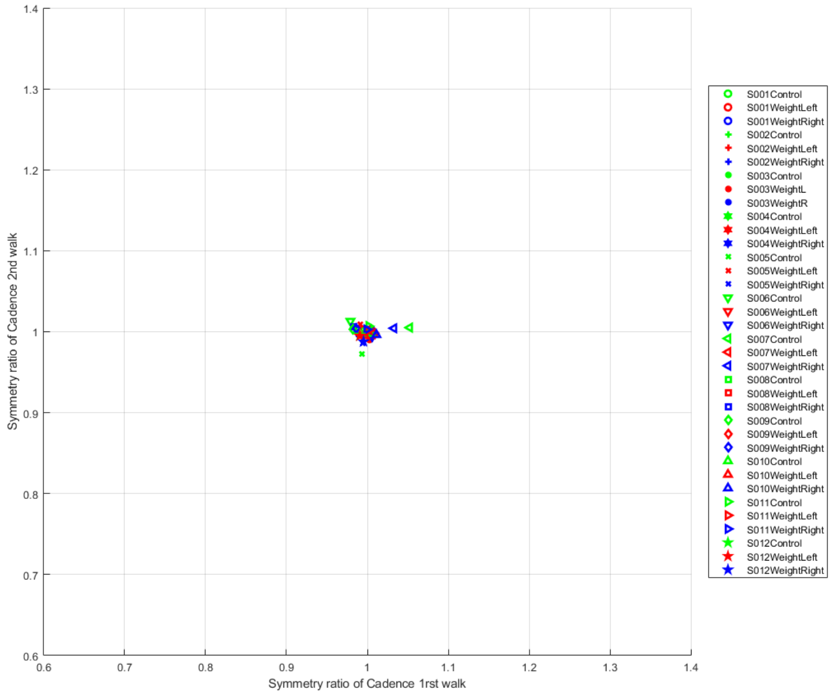


Figure 33. Symmetry Ratio of 1st and 2nd walk for Cadence.

3.2.3. Stance Phase

In Stance Phase, Figure 34, we observe a diagonal spread of the datapoints. Since x and y axes report the values from the two repetitions of the “10 Meters Walk”, we expect that the same subject with the same “conditions” and on the same walk, should provide consistent values and the results should be on the diagonal. Marked with green color, the control walk of the healthy subjects is moderately spread around the value of 1 which indicates that their walk is close to symmetric. In blue and red which are the walks of the subjects with weight loads, we observe a spread further away from the value of 1 when the weight load is added to the lower limb of the subject and in those cases stance phase lasts less for the weight-loaded limb than the control.

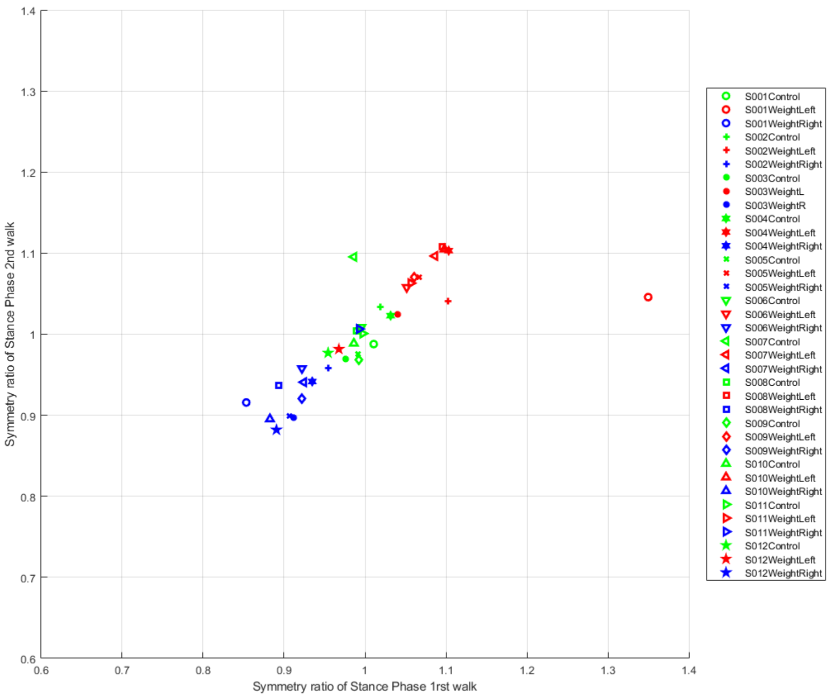


Figure 34. Symmetry Ratio of 1st and 2nd walk for Stance Phase.

3.2.4. Swing Phase

In Swing Phase, Figure 35, we observe a similar result as STP with a bigger spread due to the fact that the values of the SWP are smaller in comparison to the values of the STP and these two quantities sum up a complete duration of 100% of a gait cycle. Once again, the control datapoints with green representing the control walk gathered around 1, compared to the datapoints in red and blue representing the weight load left and right respectively, which are clearly further away from the value of 1 with respect to the datapoints of the control walks.

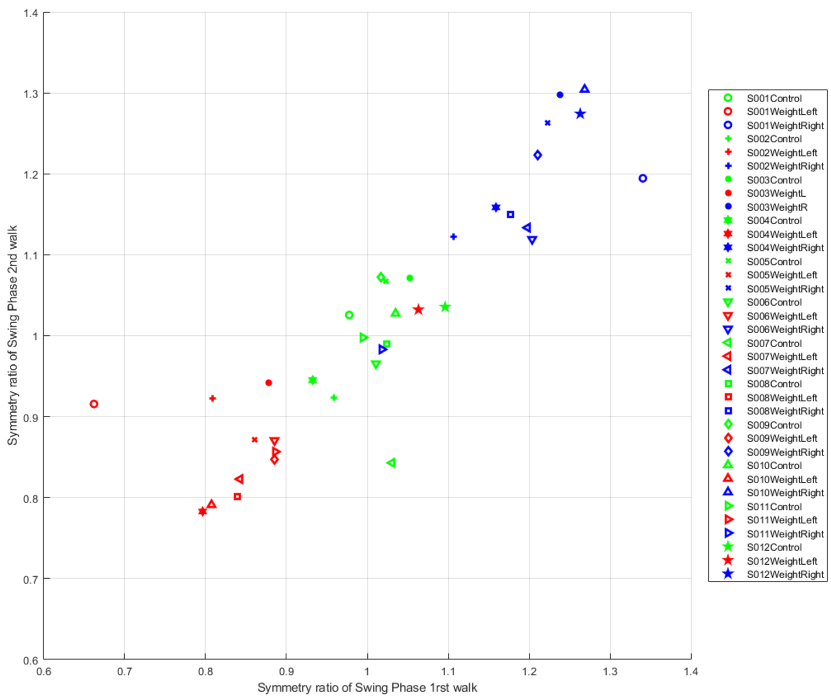


Figure 35. Symmetry Ratio of 1st and 2nd walk for Swing Phase.

3.2.5. Double Stance Phase

Symmetry ratio of Double Stance Phase in Figure 36, gathered around 1 with a few datapoints spread 10-20% far from 1, green for control walks, red and blue for weight left and right respectively. The results from this test do not display a clear separation nor provide a diagnostic value with this presentation.

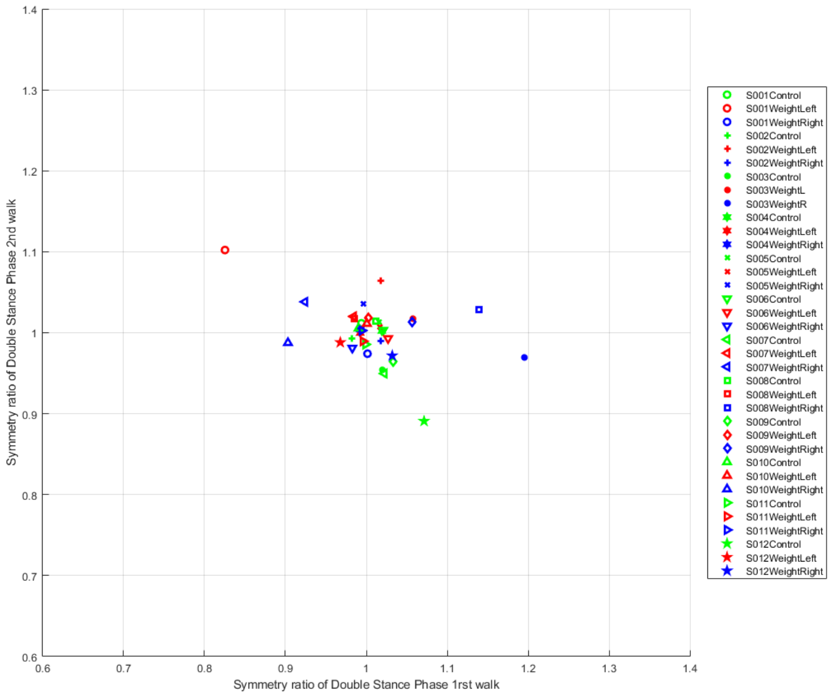


Figure 36. Symmetry Ratio of 1st and 2nd walk for Double Stance Phase.

3.2.6. Step Time

Figure 37, presenting the symmetry ratio of step time, looking closer to 1 we can see the green data points gathered together and red and blue datapoints are gathered further away from the value of 1 as the pattern observed in stance and swing phase. What is obvious here is the outliers greater than 2 and another greater than 7 which do not come from a problem in our procedure and can tell us there is another problem hidden in the system. We decided to investigate these datapoints further, and this led to identifying an issue with the sensors. This investigation is discussed in 4.1.3.

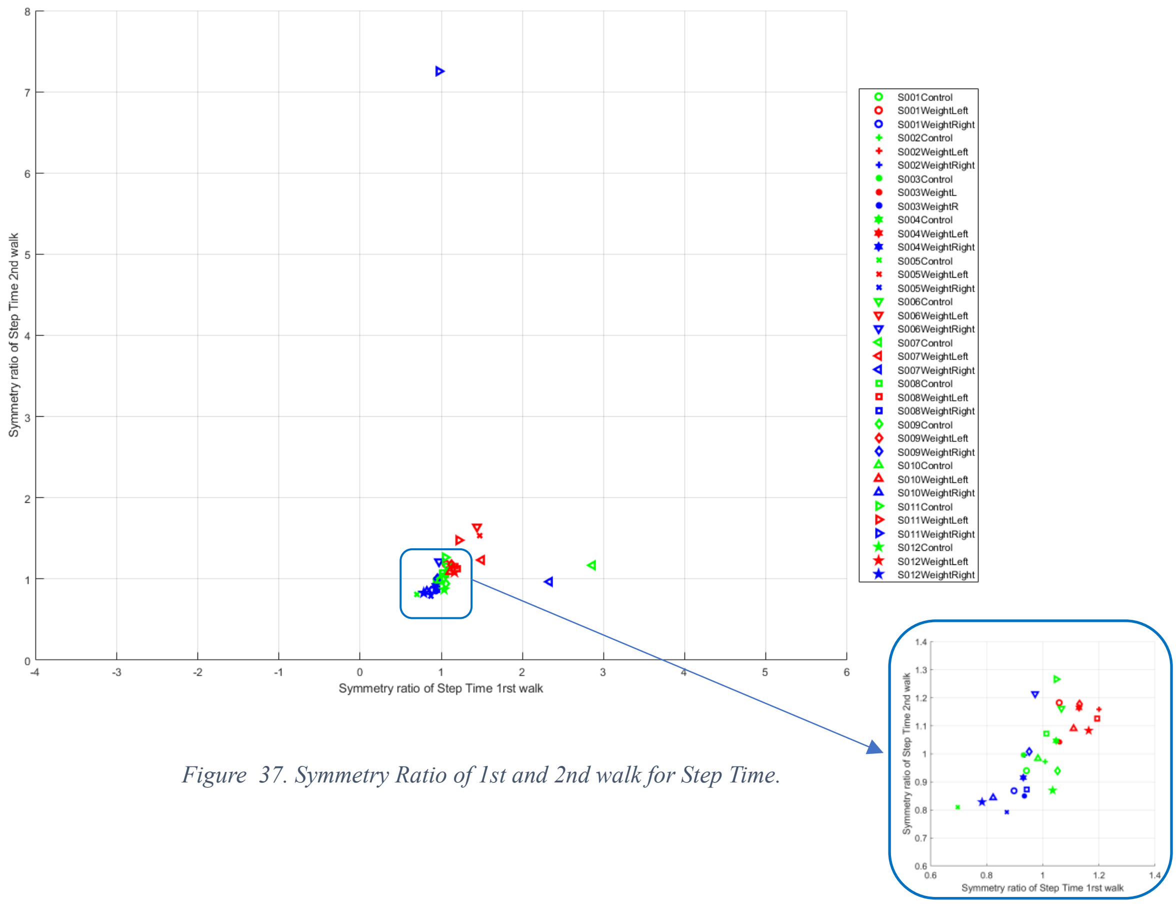


Figure 37. Symmetry Ratio of 1st and 2nd walk for Step Time.

4. Discussion

4.1. Validation

4.1.1. Gait Cycle Duration & Cadence

- **Gait Cycle Duration [3.1.1.]:** Means and standard deviations give us very good comparison results, with mean values being very close between wearable and treadmill and SD provided by the sensor being slightly bigger than the one of the treadmills. Thus, we can conclude that for this parameter the wearables are as accurate and almost as precise, as the treadmill (while the wearables are also much cheaper, and their use in home settings is foreseeable for the future unlikely for treadmills). We can expect a slight difference due to the fact that the treadmill detects the pressure as soon as the sole of the shoe is in contact where sensors detect pressure applied to the insole milliseconds later. Average error is smaller than 0.1% so we can conclude in a solid manner that our system is very accurate on this parameter, and concerning this parameter we can trust it for use within the MSAI study.
- **Cadence [3.1.2.]:** We do not provide an SD for this parameter, and the mean value provided by the treadmill is a rounded number while our system gives an approximate number for left and right foot. Adding these numbers and if we were to round the sum, we would in almost every case be equal to the treadmill. Also, a very accurate parameter we can trust for use within the MSAI study.

4.1.2. Stance, swing, and double stance phase offset

In all of the parameters discussed below a systematic error between the values from the treadmill and those from the wearable is found.

- **Stance Phase [3.1.3.]:** In this parameter we observe a systematic error, always overestimating Stance Phase by 11.26% on average. This error is always of the same direction and magnitude, varying slightly for different speeds (it increases in magnitude linearly with increasing speed).
- **Swing Phase [3.1.4.]:** A similar (and related) systematic error is observed, this time underestimating the Swing Phase by 21.92% on average. Once again always of the same direction and magnitude, varying for different speeds (it increases in magnitude linearly with increasing speed).
- **Double Stance Phase [3.1.5]:** Here the results come from the combination of the systematic errors mentioned above, direction and magnitude are the same in every case, but the error is on average at 194.80%. This problem occurs from a combination of a systematic error of Stance and Swing Phase, but this parameter suffers also from another error which will be further discussed on 4.1.3.

What is causing the systematic error on the Stance Phase, Swing Phase and Double Stance Phase was expected to some extent, being due to a step in the computational pipeline to compute the gait cycle parameters, so part of the scope of the validation which is performed in this thesis was indeed to quantify the magnitude of this effect and investigate potential solutions. The raw data from the sensors are affected by noise (when it comes to patient data with the tremor they might have) and in order to train the algorithm to identify the key events (Heel on, Toe on, Heel off, Toe off), smoothing was necessary. Pressure data are filtered by Gaussian filtering to smooth the signal removing noise. When smoothing a signal, it is expected that you modify the true original value; in this case the curves will lower the maximal amplitude but not only. The rising point, the point at which the pressure curve starts to depart from zero, will not be on the same spot, it will move some fractions of a second earlier (and the converse for the point at which the pressure curve returns to zero, which gets delayed) thus adding a systematic error to our estimation of the stance, swing and double stance phase parameter.

Some early tests were done with the Gaussian Filtering during the process of developing the pipeline and the same filtering method (Gaussian filter) that is being used now was tested for different Filter orders σ, described as sigma=1,2…,11. We can see in Figure 38 how filtering can change the original data and initially the idea was to use a mild filtering option (sigma=5) so that the data wouldn't be influenced greatly, but still would be sufficiently smoothed for the subsequent steps of the computational pipeline to work properly.

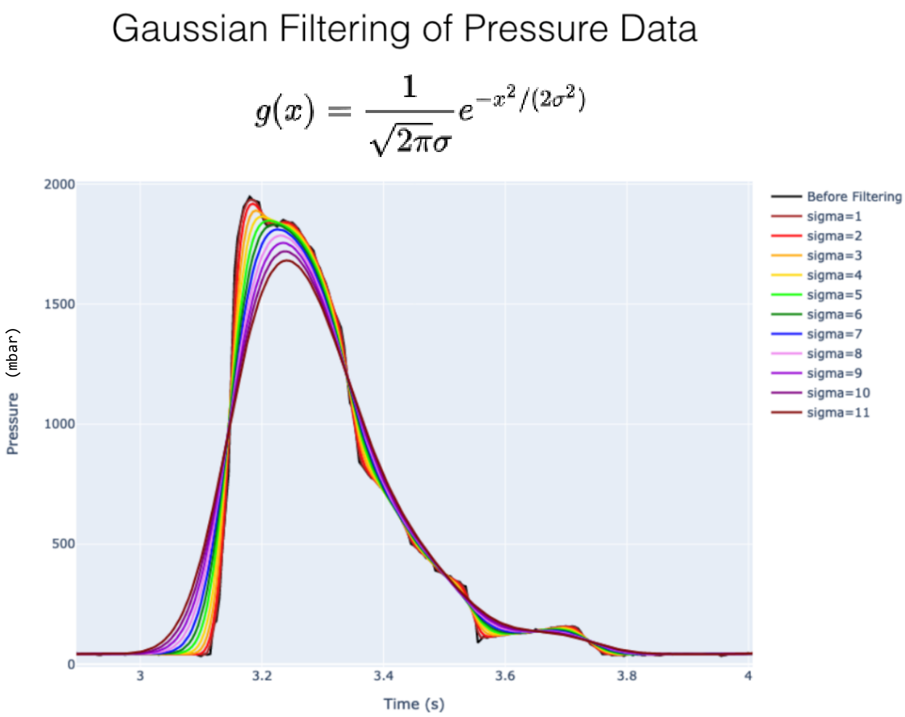


Figure 38. Gaussian Filtering Tests on Pressure Data Example. Source: MSAI.

To verify this effect, I performed a test on one of the subjects walks with different filtering in the pipeline to verify that this is the source of the systematic errors and in Figures 39 and 40 are the results of the Gaussian filtering with different orders.

Gait Cycle Duration, Cadence and Step Time are not affected by this systematic error, but their Standard Deviation is slightly increased with lower filter order (Figure 39).

We can observe on figures 39 and 40 the effects of Gaussian filtering order 7 to order 1 in blue (left to right) and with red we have the treadmill value always from the same report. This verifies the effects of smoothing the data on the calculation of Stance, Swing and Double Stance Phase: we can conclude here that the systematic errors which we have observed in stance, swing and double stance phase are due to the chosen order of the Gaussian filtering. Unfortunately, reducing the order of the Gaussian filtering in the MSAI study might not be an option because this might leave too much noise in the data (which therein come from PD and NPH patients and thus include even more noise due to symptoms such as tremor) for the subsequent steps of the computational pipeline to work properly.

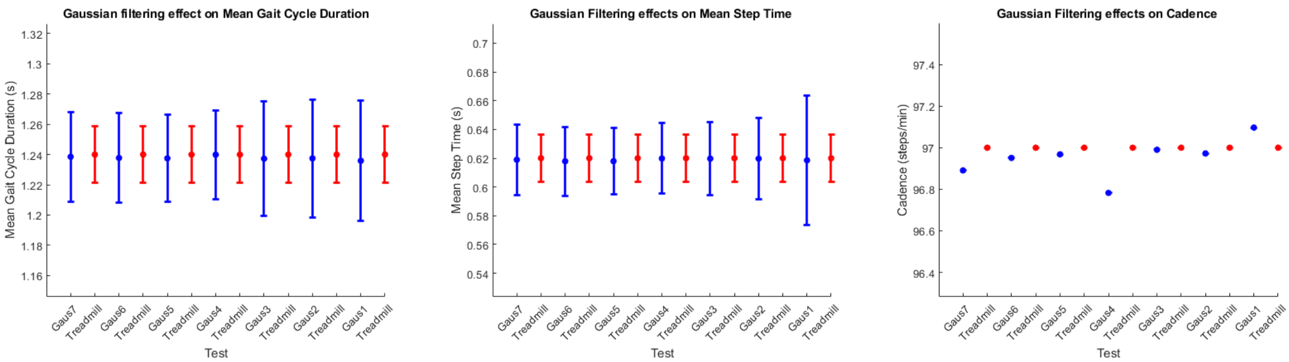


Figure 39. Gaussian Filtering Effect on Gait Cycle Duration (left), Step Time (middle) and Cadence (right).

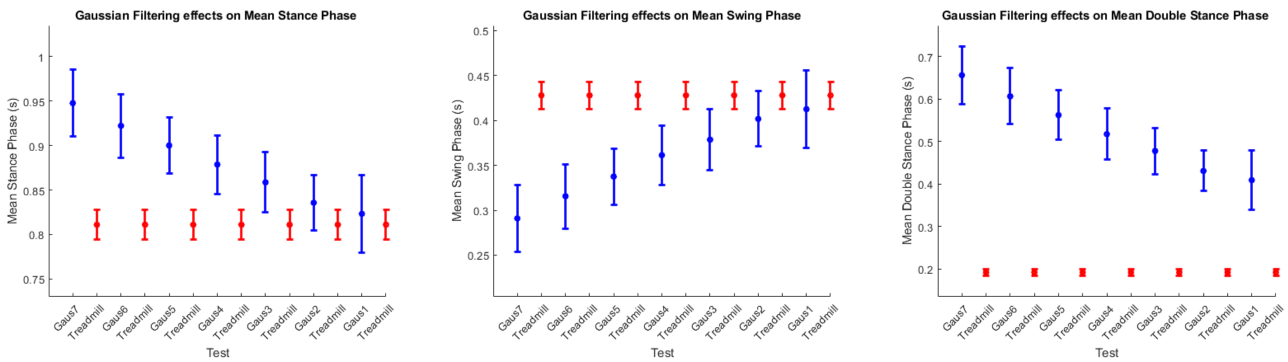


Figure 40. Gaussian Filtering Effect on Stance Phase (left), Swing Phase (middle) and Double Stance Phase (right).

Proposed solution no. 1: A posteriori Correction

Since we know that the filtering which was used for all the healthy subjects that we processed had the same order, and we have computed the average value of the error in Figure 23, the simplest solution is to fix all the Wearable **Stance Phases** by shifting them by the average error (11.26%), the results are presented in Figures 41 & 42.

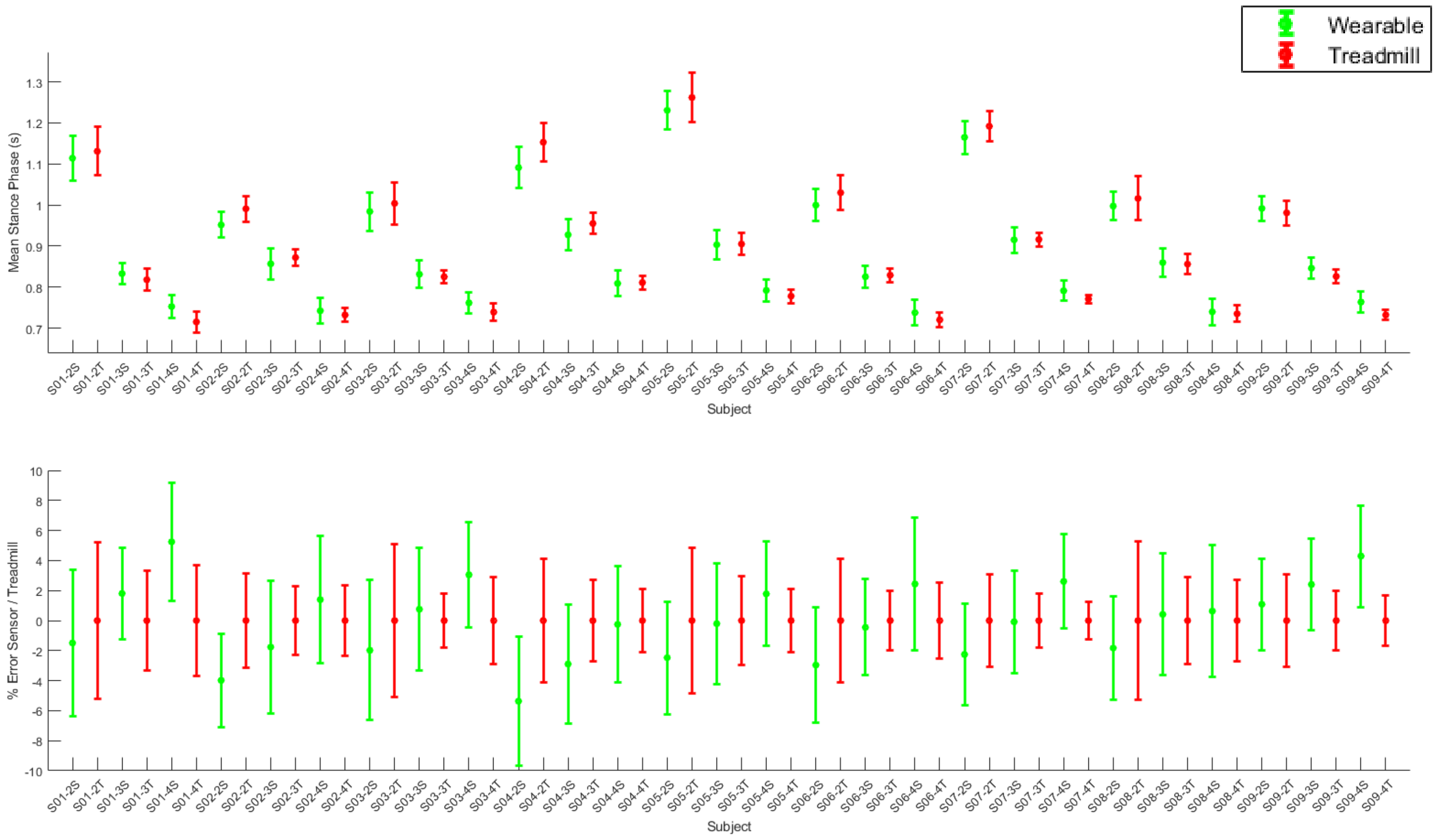


Figure 41. Validation Analysis Corrected for Stance Phase.

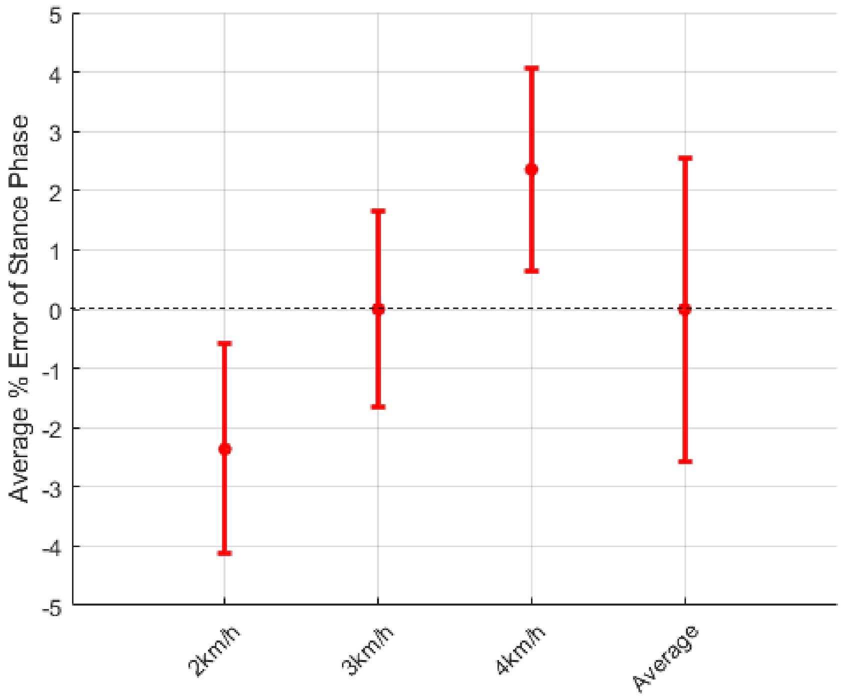


Figure 42. Average % Error for Stance Phase Corrected.

We apply the equivalent correction for **Swing Phases**, according to Figure 25, shifting them by the average error (21.92%). The results are presented in Figures 43 & 44.

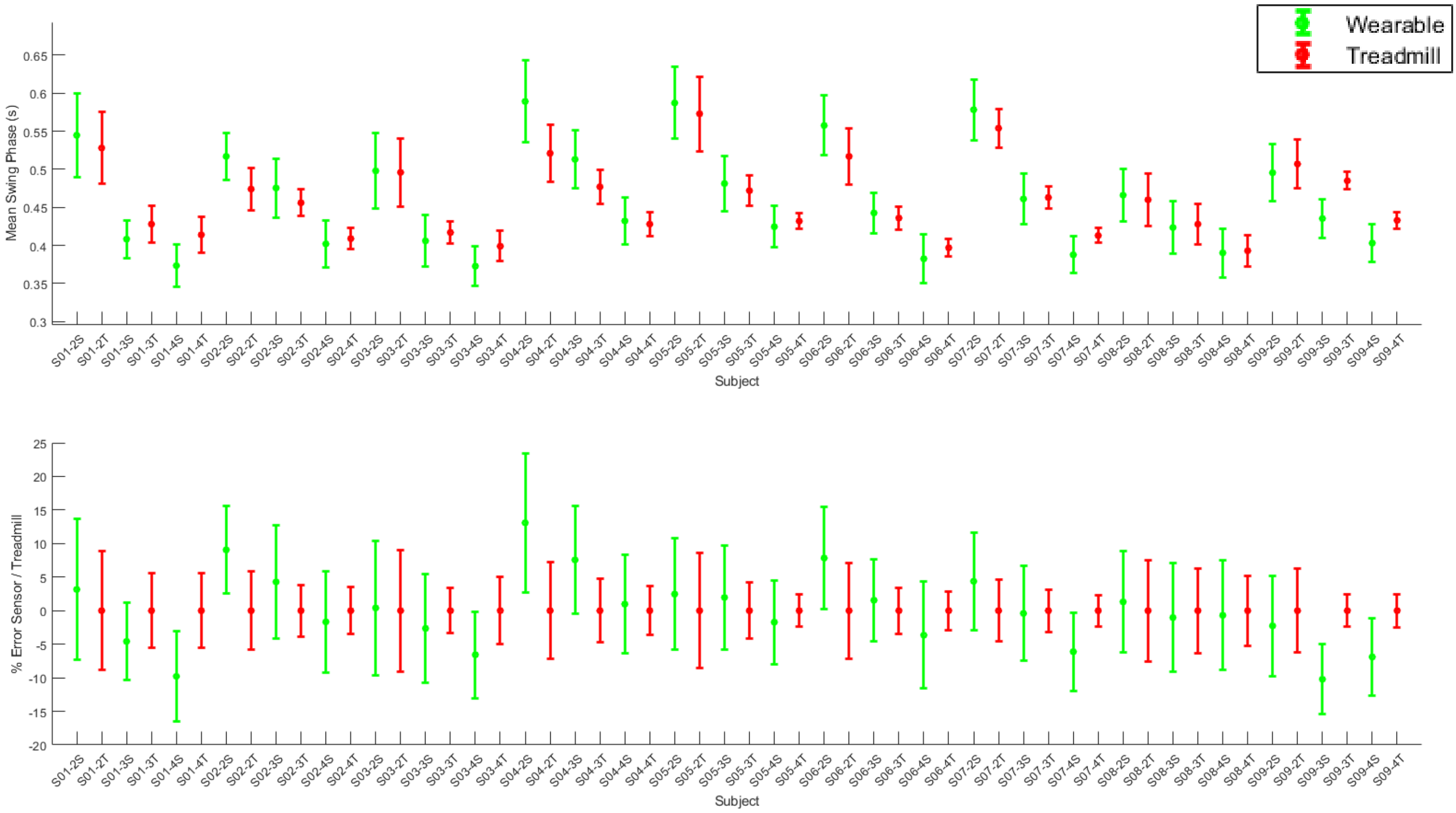


Figure 43. Validation Analysis Corrected for Swing Phase.

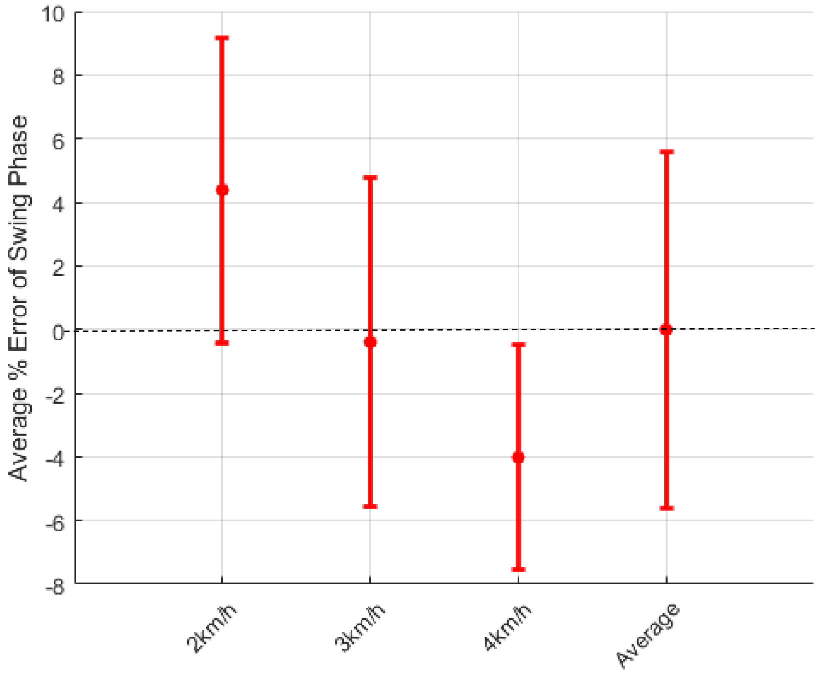


Figure 44. Average % Error for Swing Phase Corrected.

And lastly the correction for **Double Stance Phase**.

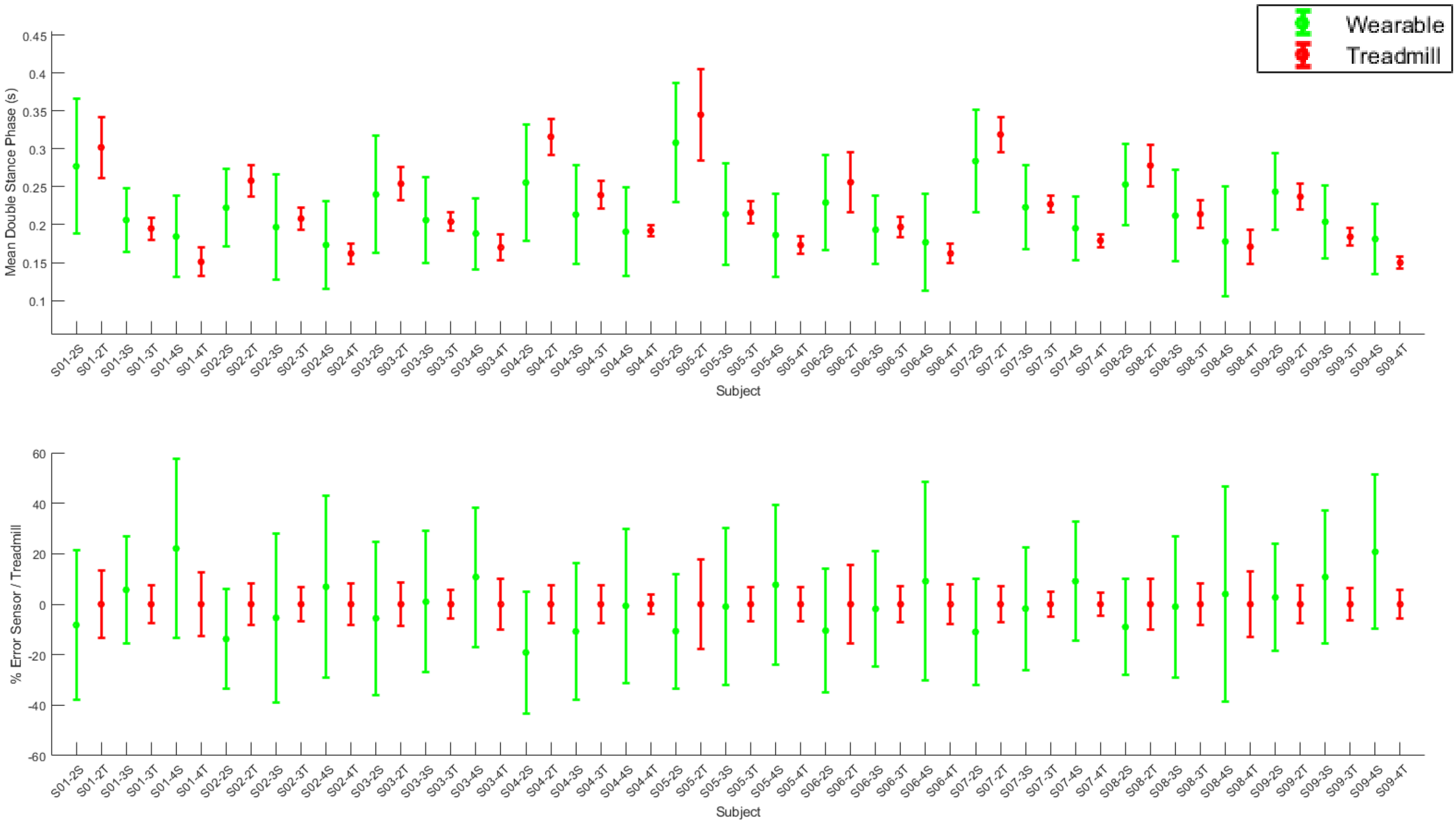


Figure 45. Validation Analysis Corrected for Double Stance Phase.

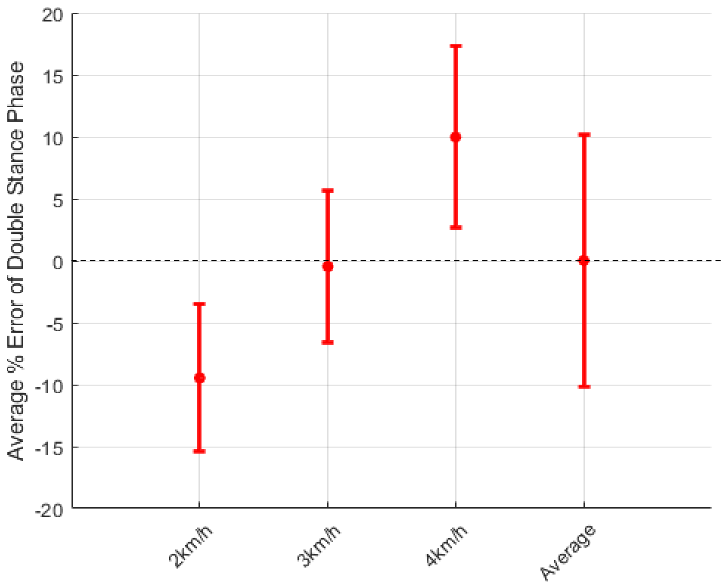


Figure 46. % Error on Double Stance Phase Corrected.

Proposed solution no. 2: Re-Running the notebooks with Gaussian filter with the order of 2 instead of 5

Another approach to solve this problem of the filtering is to reduce the order of the filtering so that we have less smoothing. I re-run all the subjects notebooks with filter order 2 instead of 5 and the results for Stance phase, Swing phase and Double Stance phase are presented bellow in Figures 47-52.

Stance Phase

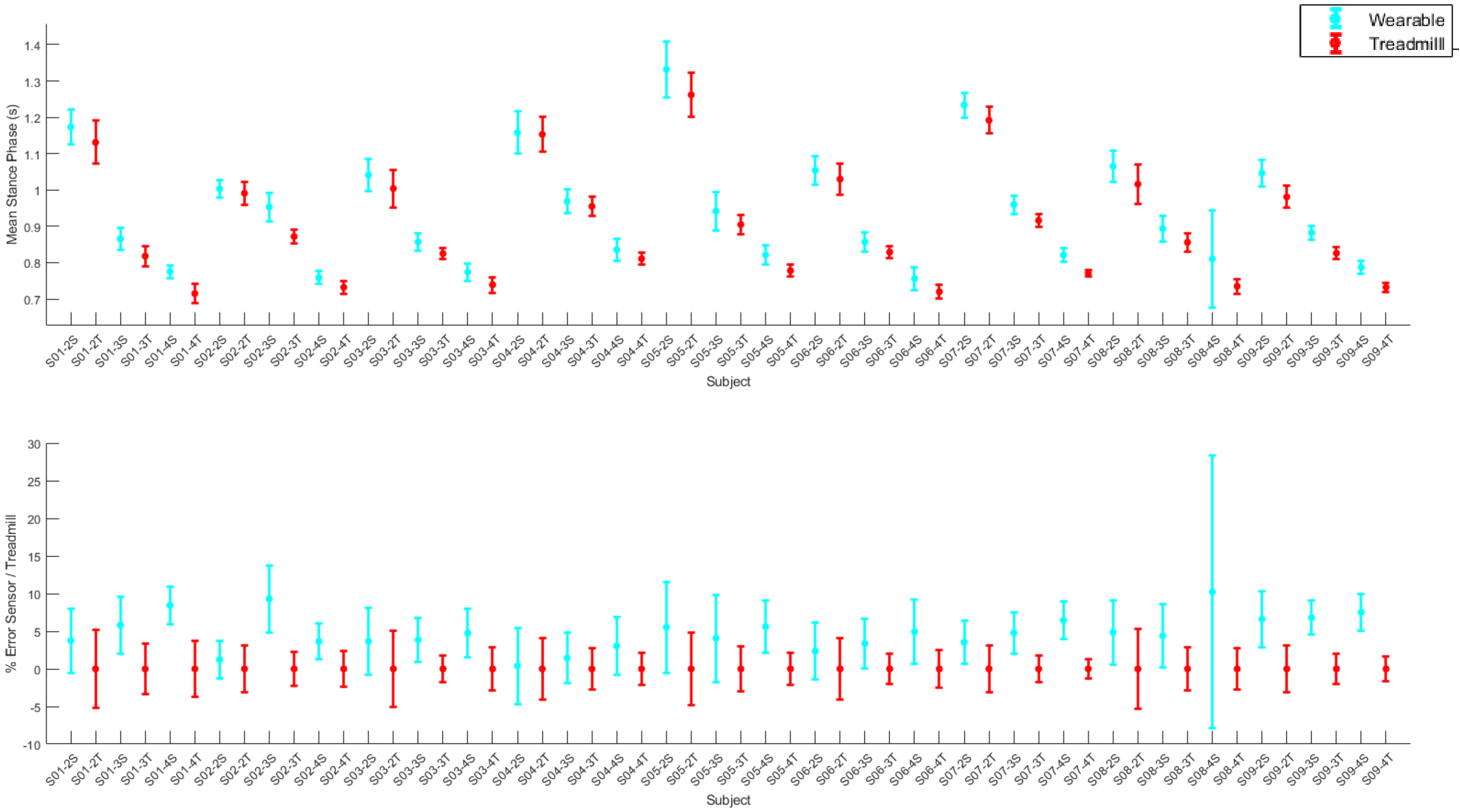


Figure 48. Stance Phase with Gaussian filtering order 2.

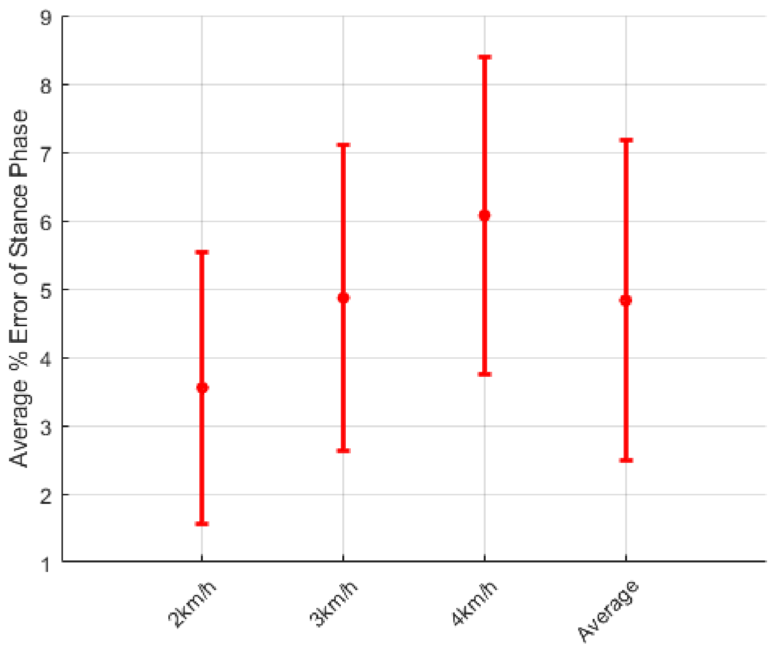


Figure 47. % Error on Stance Phase after Gaussian filtering order 2.

Swing phase

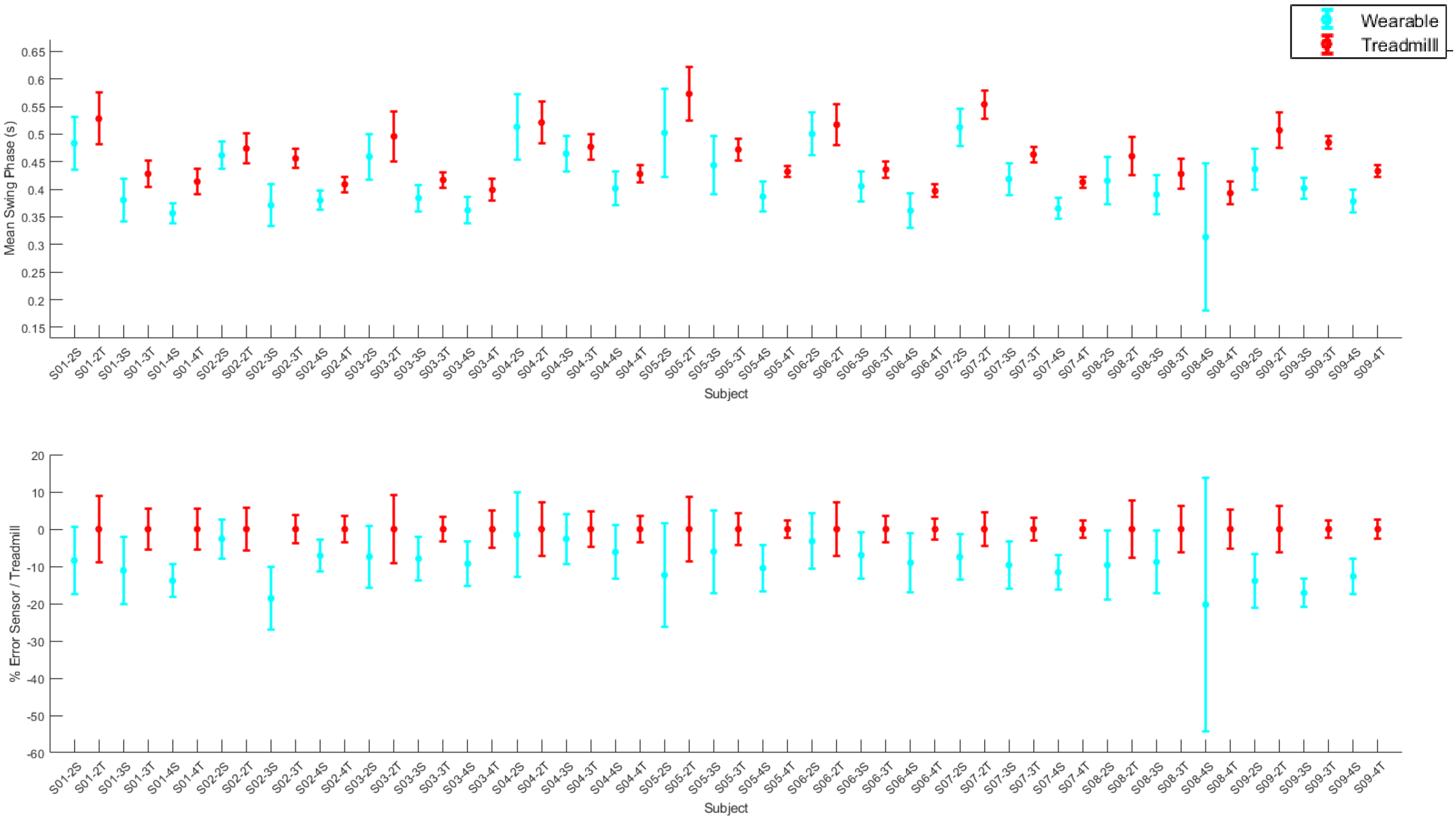


Figure 49. Swing Phase with Gaussian filtering order 2.

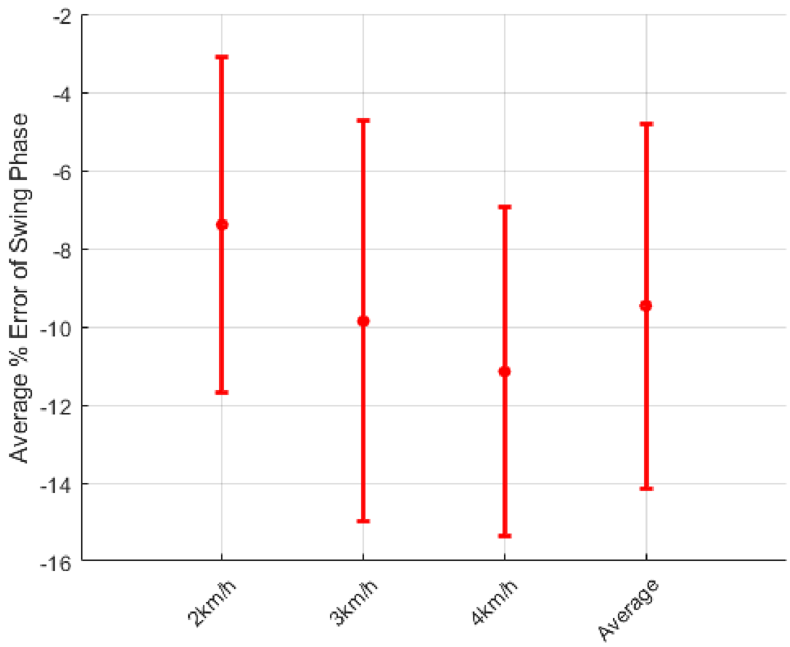


Figure 50. % Error on Swing Phase with Gaussian filtering order 2.

Double Stance Phase

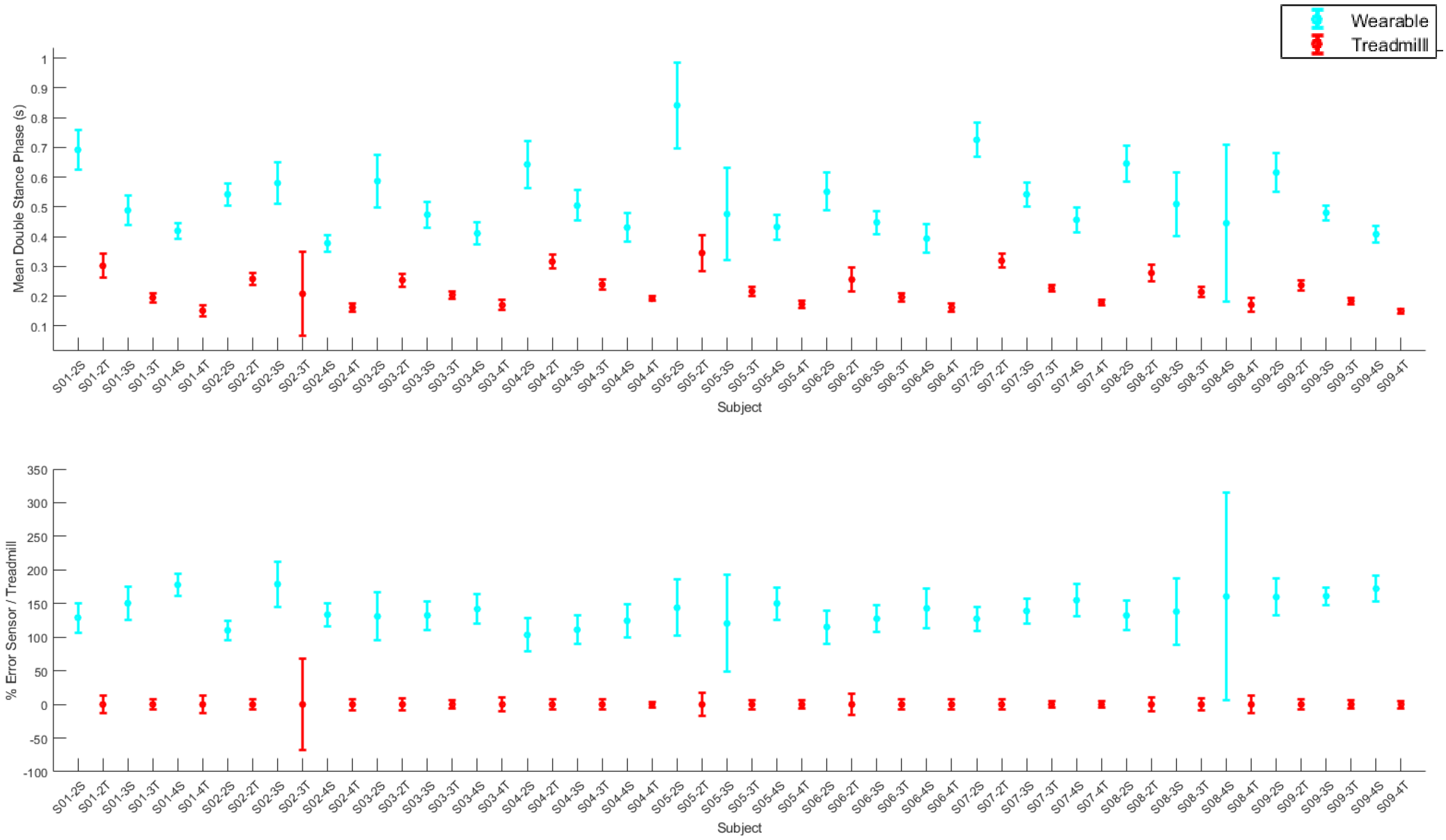


Figure 51. Double Stance Phase Gaussian filtering order 2.

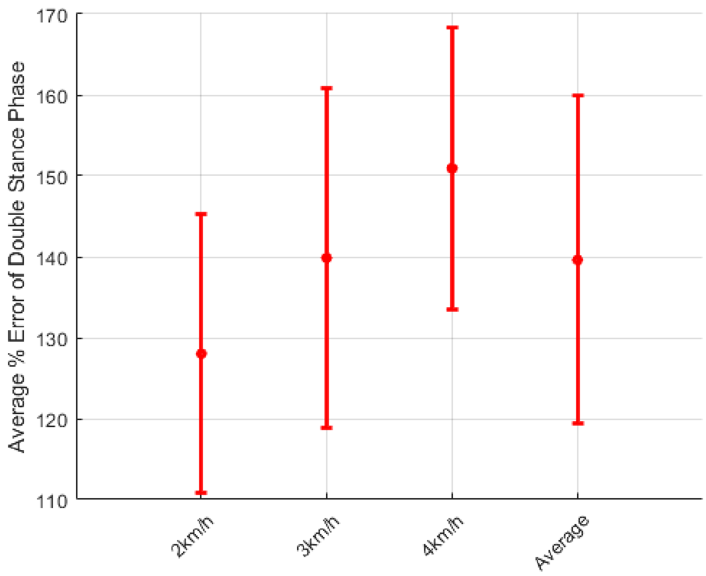


Figure 52. % Error on Double Stance Phase with Gaussian filtering order 2.

This solution supports our conclusion on the origin of the problem. Filtering less does bring the parameter values from the wearables closer to the treadmill values, but it does not solve the problem completely, an offset is, on average, still present (Figures 48, 50, 52). The same trends as before are being observed; Stance and Double Stance phase are overestimated, and Swing phase underestimated. While this solution is practical and fairly easy to implement for all patients it is not the best option for a few reasons:

1. Identifying steps (by the computational pipeline) becomes less accurate due to more noisy signals and valuable information might get lost, considered as an outlier. More specifically, the identification of the key events becomes less and less accurate and more prone to errors the less the data are filtered (this is why the signal was very considerably smoothed in the first place)
2. The algorithm, not being able to correctly identify the steps due to excessive noise, faces the risk of not reaching its end and while signal is relatively less noisy in healthy subjects, signal from patients is much noisier and it is a lot harder to identify the key events with the tremor and other abnormalities they might have.
3. It increases the Standard Deviation of the estimated values.
4. Error estimated is greater than the proposed solution no.1.

For the reasons mentioned above, we can consider this a feasible option but not optimal in our case which brings us to adopt instead for the MSAI study the first proposed solution which is faster to implement, has less risk and brings us closer to the true value according to the treadmill.

4.1.3. Step Time

Mean Step Time [3.1.6.] The agreement between Step Time value measured by the wearable and the one measured by the treadmill seems to be as good as Gait Cycle Duration and Cadence. Although Step Time as well as Double Stance Phase, due to the fact that they require data from both feet in order to be calculated, face the same issue. We investigated it and we found that it is a random error due to the sensors which sometimes lose synchrony (at times which do not show an apparent pattern) as a result, when they try to re-sync one of the IMUs' is "catching up" the other one with the result of missing some time points crucial to the calculation of these parameters. We can notice on figure 53 highlighted on Step Time L and Step Time R respectively that few subjects have the wearable value overestimating one and underestimating the other with respect to the treadmill. This is also a good example of why we considered randomizing the walks, one could say that this only happens at the speed of 4km/h, but it is obvious also at the speed of 3km/h for Subject 07.

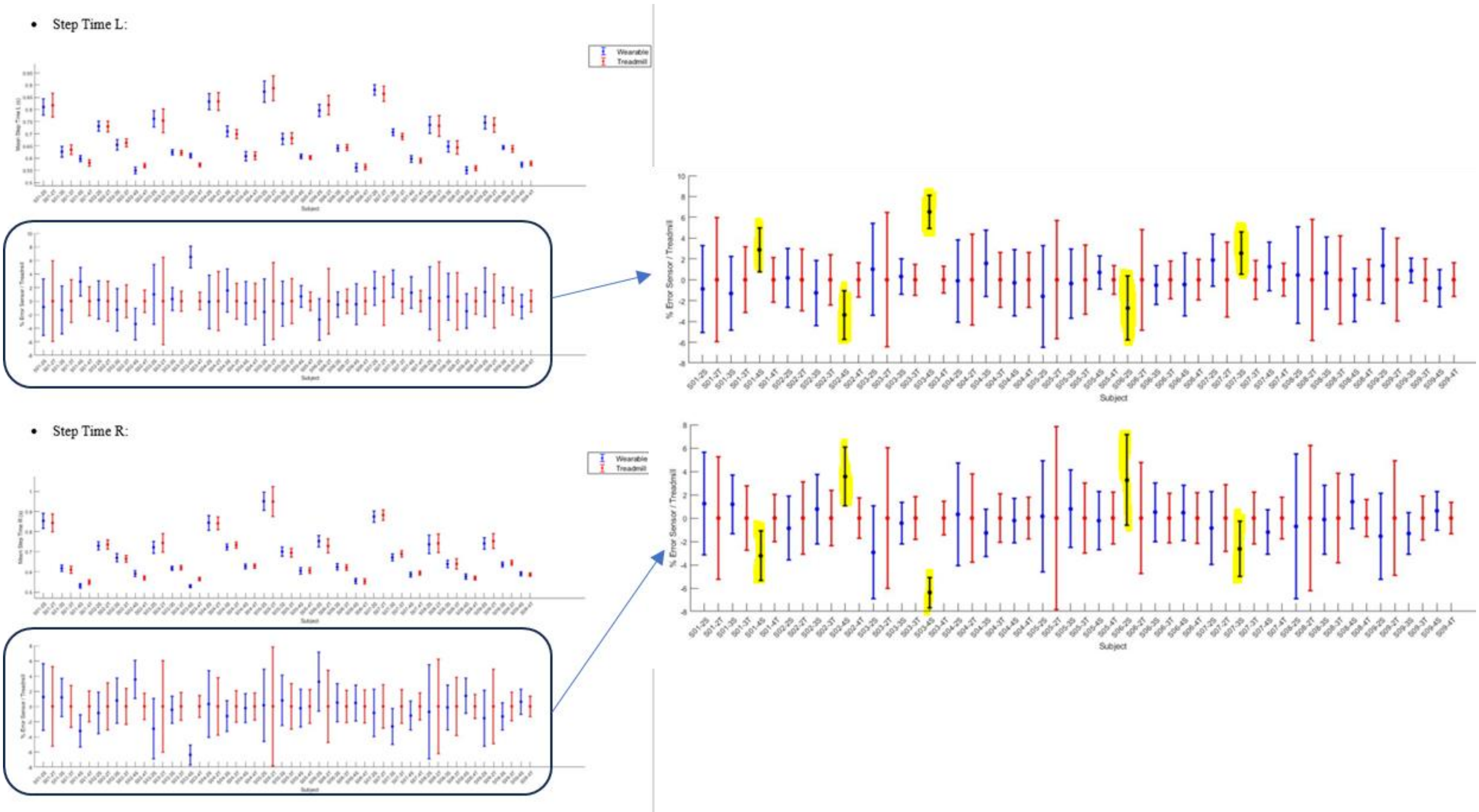


Figure 53. Comparison of Step Time Left and Step Time Right, % Error of Sensors / Treadmill

To help us understand this effect, Figure 54 displays the raw data from one of the subjects performing the asymmetry test with weight load on right foot, for which this asynchronization issue occurs.

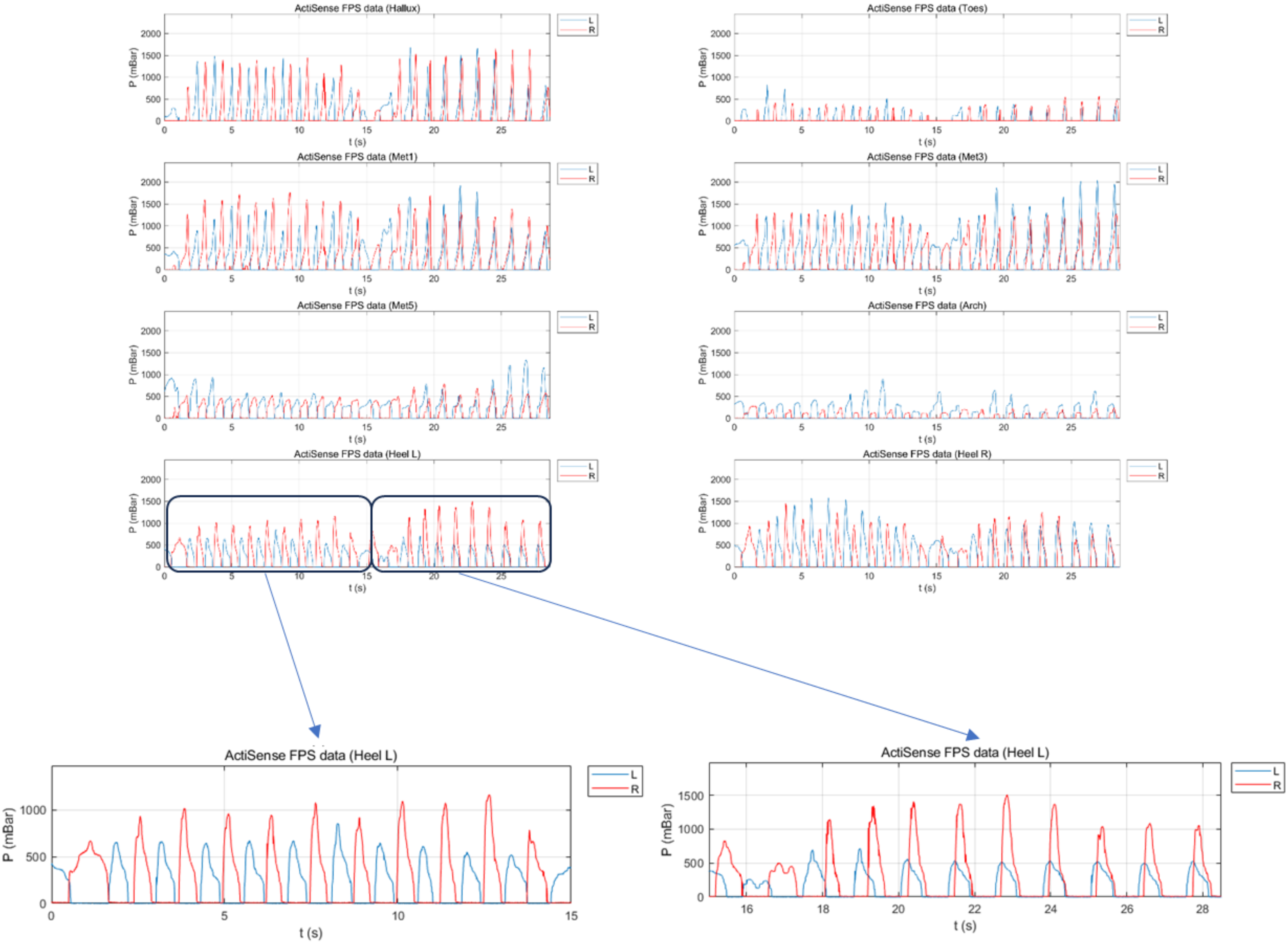


Figure 54. Example of Raw Data and Synchrony Loss.

While at the beginning of the measurement everything goes well, sometime during the walk sensors synchronize again resulting in this asynchrony. This is clearly visible in the panel at the bottom of Figure 54, where in order to generate left and right pressure peaks that occur almost at the same time, the subject would need to be jumping with both feet at the same time, which was not happening during the measurement (visually assessed by the MSAI team).

One might argue that this effect only happens at the end of the recording but below we report also another data set only 8 seconds long (Figure 55) where this effect appears at the 3rd second already.

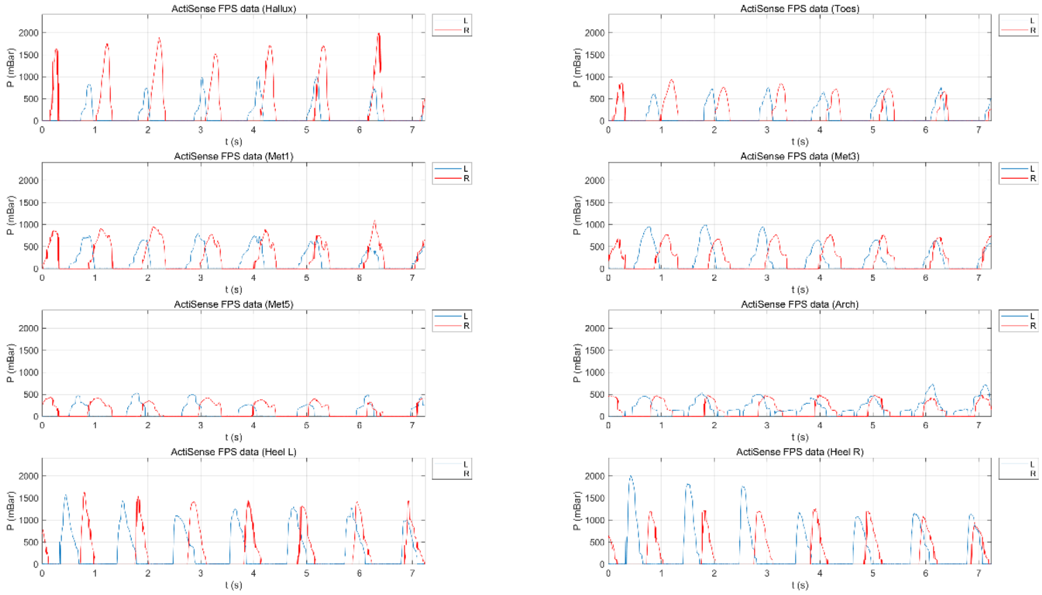
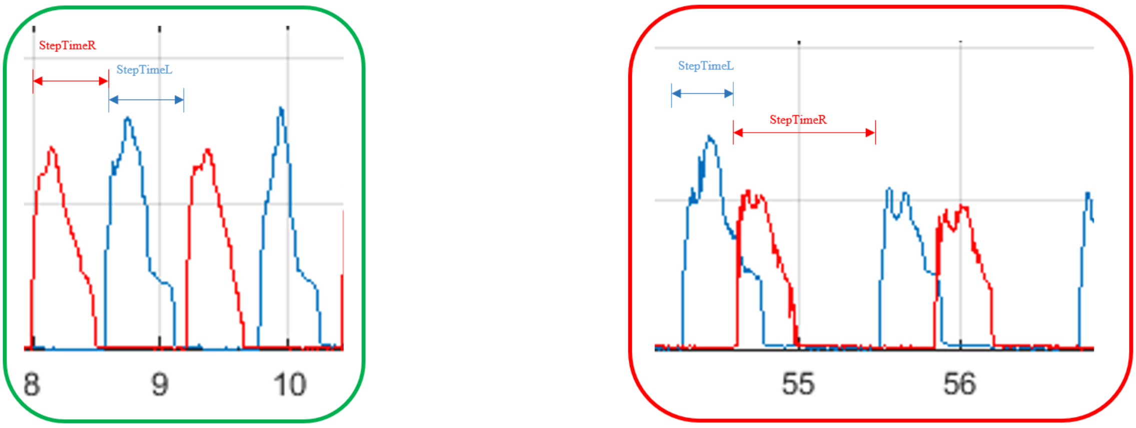


Figure 55. Another Example of Raw Data and Synchrony Loss for a recording duration of 8 seconds.

While this issue isn't affecting the parameters that are calculated for each foot separately (Gait Cycle Duration, Cadence, Stance Phase, Swing Phase) it has a great effect on Step Time and Double Stance Phase for the reason explained below (Figure 56).

On the green Time-series data (1st walk) we see the pressure peaks in consecutive order and StepTimeL and StepTimeR are almost equal 0.60 s and 0.61 s as we can see on the 1st walk with minor differences. However, when the asynchrony effect is present, pressure peaks pile up, presented in the red Time-series data (2nd walk), StepTimeL drops down to 0.43s and StepTimeR extends to 0.79s giving the impression of this huge asymmetry in healthy controls which is an artifact of the sensors' IMU identifying (correctly for the pipeline but not realistically for the actual gait that was performed).



Mean Step Time, L (s)	CV of STT, L
0.6077499999999998	0.05435315065707401
0.43611111111111023	0.18591308179554702

1st walk
2nd walk

Mean Step Time, R (s)	CV of STT, R
0.6175000000000002	0.06264192854199623
0.7986363636363646	0.09489488968562493

Figure 56. Time-Series Data and Mean Step Time L & R example.

4.2. Asymmetry

4.2.1. Gait Cycle Duration & Cadence

By definition these parameters should be giving us a Symmetry Ratio of 1, this is indeed verified on the Asymmetry plots (Figure 32, 33).

- **3.2.1. Gait Cycle Duration:** Either left to right or right to left, the total time for these steps should be identical no matter the asymmetry in the gait since it is defined by the total time from first heel contact to second heel contact of the same foot.
- **3.2.2. Cadence:** It is possible that a step recorded, might be considered an outlier for previous reasons mentioned (filtering, identifying key events, outliers) in the total walk we might measure one extra step (e.g., 30 Left Steps and 29 Right Steps the Symmetry Ratio would be 0.966). However, it will never be the case that we have 30 Left Steps and 20 Right Steps, as that would mean the subject is skipping steps on one foot, which doesn't occur during one walk of a healthy subject.

4.2.2. Stance Phase

3.2.3. Symmetry Ratio of Stance phase gives us a very clear picture, i.e., a plot where data points separate very clearly the "ControlWalk", "WeightLeft" and "WeightRight" groups. The groups were found to be significantly different, tested with a (t-test) p-value threshold of **0.001**.

4.2.3. Swing Phase

3.2.4. Swing Phase was found to be the parameter that allows us to visualize more clearly the results of added weight to the limbs with a clear separation of the data points for "ControlWalk", "WeightLeft" and "WeightRight" groups. This is due to the fact that in this case the nominator and denominator have a greater difference when it comes to symmetry ratio whereas in stance phase the difference is of a smaller magnitude. For example, a SwingPhaseL of 30% and SwingPhaseR of 25% would give us a result of 0.833, where the corresponding StancePhaseL of 70% and StancePhaseR of 65% would give us a result of 0.928. Swing phase symmetry ratio allows us to see a greater difference from 1 with the same weight load, same test on different subjects and this result makes it the most suitable to visualize the effects due to the larger spread. The groups were found to be significantly different, tested with a (t-test) p-value threshold of **0.001**.

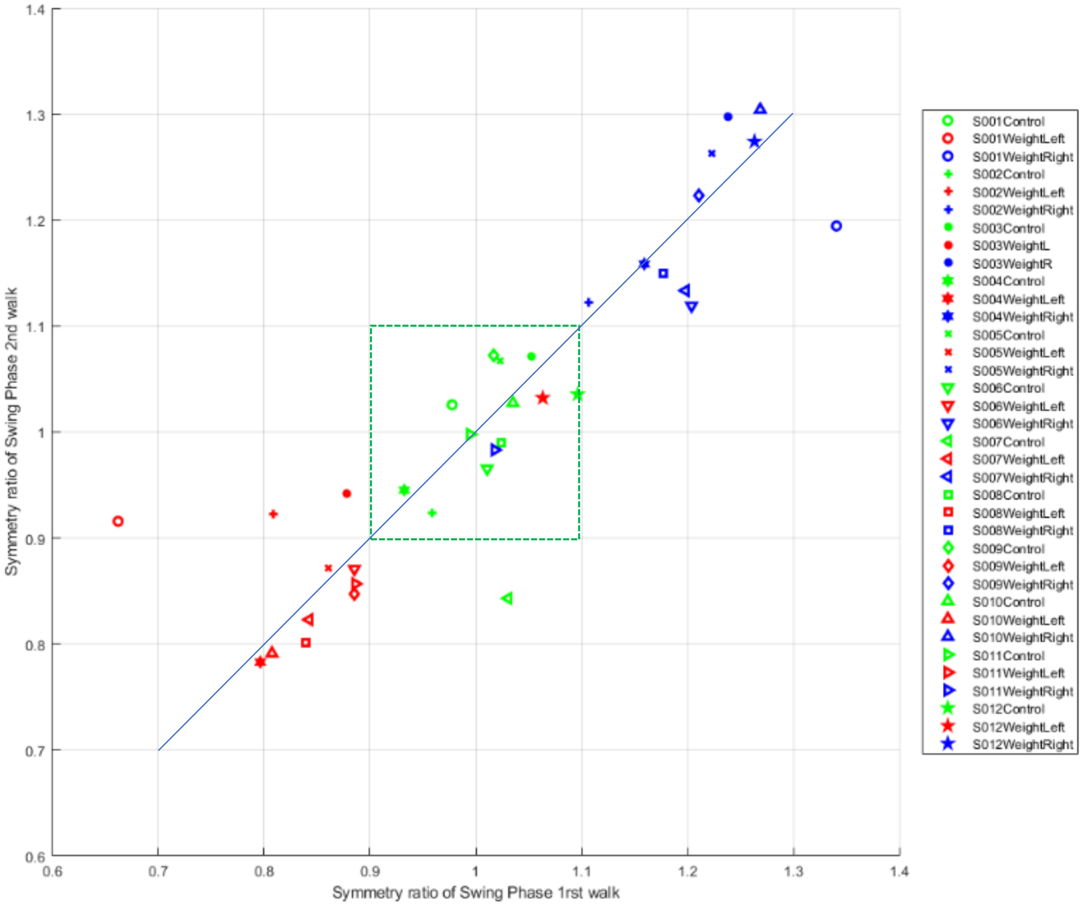


Figure 57. Symmetry Ratio of 1st and 2nd walk for Swing Phase.

In Figure 57, the blue diagonal is a consistency check, x and y axes report the values from the two repetitions of the "10 Meters Walk". We expect that the same subject with the same "conditions" and on the same walk, should provide consistent values and the results should be on the diagonal. Groups are clearly separated, and we can see that all subjects present asymmetry in their gait when load is added to their limbs. We can also see that almost every healthy subject presents some natural asymmetry within the 10% area marked with the green square. This experimental setup with the weight loads is proven to be effective and along with the system this can successfully identify gait asymmetry and distinguish between physiological gait asymmetry and pathological gait asymmetry.

4.2.4. Double Stance Phase & Step Time

- **3.2.5. Double Stance Phase:** This parameter's results do not show any evident pattern when it comes to symmetry ratio, and we cannot draw conclusions since this parameter is affected by the huge offset plus the synchronization issue may or may not be present. We simply do not take into consideration the Symmetry Ratio of Double Stance Phase any further for the analysis within MSAI.
- **3.2.6. Step Time:** This parameter also shows some separation between the walks, but it is also not trustworthy due to the synchrony error, and we can clearly see it on figure 36 with one data point providing a symmetry ratio of more than 7.

5. Limitations

The limitations of this study include the following:

- The validation of the sensors was performed on healthy subjects and not on patients (affected by e.g., PD or NPH) which does not give us the complete picture of the system's performances for subjects whose gait cycle is impaired. Furthermore, more subjects with the same age groups (groups of age 30-40 and 40-50) might have displayed a correlation between age and gait.
- Validation and asymmetry tests were conducted on a rather small group of subjects. Many more subjects e.g., 50-100 per study would allow a better understanding of the system's performances and perhaps would explore some other factors we didn't discover.
- Attempts were made to contact the manufacturing company early enough to address the asynchrony issue, but it wasn't possible to discuss this issue within the time-frame of the thesis.

Despite these limitations that do not affect the results of this study but could improve the outcome we consider satisfactory the results we obtained in the thesis time period and with the means available, and we consider that they do validate the system of sensors plus pipeline, to the extent detailed in the text.

6. Conclusions and outlook

Aiming to develop a Machine Learning classifier within the MSAI study that can distinguish gait of PD and NPH patients by their gait, it was a necessity to validate the equipment used to collect the gait data and the computational pipeline developed to extract gait cycle parameters from such data. One part of this study aimed to validate the system against an instrumented treadmill and for this part we recruited a group of 9 healthy subjects and we recorded in parallel their gait by wearable sensors and the treadmill. We then compared the results of the gait parameters estimated by our system to those provided by the treadmills reports. Some of our system's parameters presented a very good agreement with respect to the treadmill's ones, while others were found to have an offset and were thus calibrated to minimize considerably the average error. The other part of this study aimed to assess the potential of our system to quantify asymmetry in gait. This was achieved by recruiting another group of 12 healthy subjects, recording their gait, and then adding weight load to one of their lower limbs at a time, to artificially induce asymmetry in a controlled and reproducible manner. We then compared the healthy walks to the groups of the artificially asymmetric walks and found that the symmetry ratio of the swing phase is the parameter that allows us to see more clearly the separation of the datapoints. This parameter is indeed the one allowing the best discrimination between physiological (in control subjects) and pathological (in impaired subjects) gait asymmetry. The symmetry ratio of swing phase can now be used to determine if the walk of the subject presents a physiological or a pathological gait asymmetry, according to the magnitude of the distance from the value of 1 representing perfect symmetry. Based on our results, we consider a 10% departure from 1 in the symmetry ratio of swing phase (i.e., a value in the range between 0.9 and 1.1) to be physiological, while more than 10% to be potentially pathological. Out of the initial 6 gait cycle parameters tested within this thesis, 4 were validated (two of which after calibration) and are thus being employed in the subsequent statistical and ML analysis performed within the MSAI study, together with the corresponding symmetry ratios.

7. Bibliography

[1] Fling, B. W., Curtze, C., & Horak, F. B. (2018). Gait Asymmetry in People With Parkinson's Disease Is Linked to Reduced Integrity of Callosal Sensorimotor Regions. Frontiers in neurology, 9, 215. https://doi.org/10.3389/fneur.2018.00215 (Accessed 4 September 2023)

[2, **in preparation**] Magni, S.*, Bremm, R. P.*, Lecossois S., Verros K., He X., Jelke F., Goncalves J., Hertel F. (2023/2024). Comparing gait cycle impairments in patients with Parkinson’s disease and normal pressure hydrocephalus by wearable sensors and machine learning. (* means equal contribution)

[3] Whittle, M. W.,4th edition (2006): “Gait analysis: an introduction”, 240 pages, Published by Butterworth-Heinemann, **ISBN 0-7506-0045-4.** (Accessed 4 September 2023)

[4] Malanga, G., & DeLisa, J. A. (1998). Section One. *Gait Analysis In The Science Of Rehabilitation*, *2*, 1. (Accessed 4 September 2023)

[5] Nasirzadeh, Alireza & Sadeghi, Heydar & Mokhtarinia, Hamid & Rahimi, Abbas. (2017). A Review of Selected Factors Affecting Gait Symmetry. Physical Treatment. 7. 3-12. (Accessed 5 September 2023)

[6] Sadeghi, H., Allard, P., Prince, F., & Labelle, H. (2000). Symmetry and limb dominance in able-bodied gait: a review. Gait & posture, 12(1), 34–45. https://doi.org/10.1016/s0966-6362(00)00070-9 (Accessed 5 September 2023)

[7] Wang, C., Wang, X., Long, Z., Tian, T., Gao, M., Yun, X., ... & Li, J. (2018, December). Estimation of spatial-temporal gait parameters based on the fusion of inertial and film-pressure signals. In 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (pp. 1232-1239). IEEE. (Accessed 5 September 2023)

[8] https://my.clevelandclinic.org/health/symptoms/21092-gait-disorders (Accessed 5 September 2023)

[9] Krigger K. W. (2006). Cerebral palsy: an overview. American family physician, 73(1), 91–100. (Accessed 7 September 2023)

[10] M Das, J., & Biagioni, M. C. (2023). Normal Pressure Hydrocephalus. In StatPearls. StatPearls Publishing. (Accessed 7 September 2023)

[11] Lew M. (2007). Overview of Parkinson's disease. Pharmacotherapy, 27(12 Pt 2), 155S–160S. https://doi.org/10.1592/phco.27.12part2.155S (Accessed 7 September 2023)

[12] Davie C. A. (2008). A review of Parkinson's disease. British medical bulletin, 86, 109–127. https://doi.org/10.1093/bmb/ldn013 (Accessed 7 September 2023)

[13] Walter, J., Bolognin, S., Poovathingal, S. K., Magni, S., Gérard, D., Antony, P. M. A., Nickels, S. L., Salamanca, L., Berger, E., Smits, L. M., Grzyb, K., Perfeito, R., Hoel, F., Qing, X., Ohnmacht, J., Bertacchi, M., Jarazo, J., Ignac, T., Monzel, A. S., Gonzalez-Cano, L., … Schwamborn, J. C. (2021). The Parkinson's-disease-associated mutation LRRK2-G2019S alters dopaminergic differentiation dynamics via NR2F1. Cell reports, 37(3), 109864. https://doi.org/10.1016/j.celrep.2021.109864(3). (Accessed 7 September 2023)

[14] Stolze, H., Kuhtz-Buschbeck, J. P., Drücke, H., Jöhnk, K., Illert, M., & Deuschl, G. (2001). Comparative analysis of the gait disorder of normal pressure hydrocephalus and Parkinson's disease. Journal of neurology, neurosurgery, and psychiatry, 70(3), 289–297. https://doi.org/10.1136/jnnp.70.3.289 (Accessed 8 September 2023)

[15] Williams, M. A., & Malm, J. (2016). Diagnosis and Treatment of Idiopathic Normal Pressure Hydrocephalus. Continuum (Minneapolis, Minn.), 22(2 Dementia), 579–599. https://doi.org/10.1212/CON.0000000000000305 (Accessed 8 September 2023)

[16] Muro-de-la-Herran, A., Garcia-Zapirain, B., & Mendez-Zorrilla, A. (2014). Gait analysis methods: an overview of wearable and non-wearable systems, highlighting clinical applications. Sensors (Basel, Switzerland), 14(2), 3362–3394. https://doi.org/10.3390/s140203362 (Accessed 8 September 2023)

[17] Flansbjer, U. B., Holmbäck, A. M., Downham, D., Patten, C., & Lexell, J. (2005). Reliability of gait performance tests in men and women with hemiparesis after stroke. Journal of rehabilitation medicine, 37(2), 75–82. https://doi.org/10.1080/16501970410017215 (Accessed 8 September 2023)

[18] https://en.wikipedia.org/wiki/Interquartile_rangehttps://en.wikipedia.org/wiki/Interquartile_range (Accessed 8 September 2023)

[19] https://en.wikipedia.org/wiki/Parkinson%27s_disease

[20] https://www.lifenph.com/file/v420399890841905720/general/iNPH_Gait%20(002).gif

[21] *https://podiatrist.borginsole.com/en/bts-g-walk-sensor-software*

[22] https://iee-sensing.com/health-tech/medical/smart-footwear-sensing-solutions/

[23] https://www.scribbr.com/wp-content/uploads/2022/01/interquartile-range.png

[24] https://www.hpcosmos.com/en/products/software-measuring-technology/biomechanics-gaitway-3d

[25] Khoury, N., Attal, F., Amirat, Y., Oukhellou, L., & Mohammed, S. (2019). Data-Driven Based Approach to Aid Parkinson's Disease Diagnosis. Sensors (Basel, Switzerland), 19(2), 242. https://doi.org/10.3390/s19020242

8. Supplementary material

Classification algorithm

Support Vector Machine

Support Vector Machine (SVM) is a well-known machine learning model which attempts to linearly classify training data. If linear classification is not feasible SVM employs one of its hyper-parameters, kernel, which transforms training data to a higher dimensional space and the linear separator (also known as hyperplane) which maximizes the margins between each class. Some of the other hyper-parameters are:

- C, which makes a compromise between the number of the misclassified instances and the margins width of the hyperplane. Higher C value means the number of misclassified samples is low which can lead to overfitting.
- Gamma which defines the influence that one data point has compared to the others. Higher value means greater influence range, also prone to overfitting.

Overall, it can produce good results even with a high-dimensional feature space and a great advantage of SVM is that it provides a unique solution [25].

Tuning parameters within MSAI

Within the MSAI project, SVM was employed to perform binary classification to distinguish PD and NPH patient's data. 11 PD and 12 NPH patients where employed. For most patients, 2 repetitions of the 10m walk where available, while for few patients, 1 or 3 repetitions were available. For each 10m walk, the computational pipeline described above allowed to compute gait cycle parameters, namely gait cycle duration, stance phase, swing phase and cadence, for left and for right foot. In addition, the symmetry ratio for each parameter was computed. Moreover, an alternative version of the symmetry ratio was also computed, where the largest gait cycle parameter value was always kept at the numerator. This allows to have only numbers equal to 1 (symmetry), or larger than 1 (asymmetry), without the information on which side (left or right) has a larger parameter. This makes a total of 16 features, i.e., each 10m walk led to a data point in a 16-dimensional space, in which SVM is run.

When training a machine learning classifier, a common good practice is to split the data in training and test set, so that the classifier (SVM in this case) is trained on a part of the data, and its performances are assessed on another part of the data, not sued during training. In order to reduce the dependency of this process by the actual realization of this random splitting of the data into training and test set, usually a k-fold validation procedure is also employed, where this split and training process is repeated k times, and the performance values are averaged over all models trained. Here we perform a leave-pair-out cross-validation loop, where a part of patients (one PD and one NPH) is left out as test data (and all walks from the same patient are kept in the same data set). This leads to 11*12 = 132 models being trained. Moreover, the training set is further split for each of these models into a further training and validation set, in order to perform leave-one-out cross-validation. This step is employed to fine-tune the hyper-parameter C, and the best performing value of C is used for the corresponding SVM model of the outer, leave-pair-out loop. Model performance is assessed by computing accuracy.

The final classification accuracy on the unseen test data, averaged over the 132 models of the outer-loop, is 0.7 +- 0.28. This is not excellent, but good, and significantly better than a random classification (which would lead to accuracy = 0.5), thus it showcases the potential of our approach to distinguish PD and NPH patients. The large standard deviation of 0.28 indicates that these results change considerably when a different splitting of the patients into the various datasets is performed, which is due to a high patient to patient variability. Thus, a much larger patients' cohort would be necessary to make these results robust and generalizable [2].