

Alma Mater Studiorum Università di Bologna  
Archivio istituzionale della ricerca

Timing estimation for gait in water from inertial sensor measurements: Analysis of the performance of 17 algorithms

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

*Published Version:*

Pacini Panebianco, G., Bisi, M.C., Stagni, R., Fantozzi, S. (2020). Timing estimation for gait in water from inertial sensor measurements: Analysis of the performance of 17 algorithms. *COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE*, 197(December 2020), 1-8 [10.1016/j.cmpb.2020.105703].

*Availability:*

This version is available at: <https://hdl.handle.net/11585/776574> since: 2020-10-29

*Published:*

DOI: <http://doi.org/10.1016/j.cmpb.2020.105703>

*Terms of use:*

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).  
When citing, please refer to the published version.

(Article begins on next page)

# Computer Methods and Programs in Biomedicine

## Timing estimation for gait in water from inertial sensor measurements: analysis of the performance of 17 algorithms

--Manuscript Draft--

<b>Manuscript Number:</b>	CMPB_2020_49R1
<b>Article Type:</b>	Full Length Article
<b>Section/Category:</b>	Biomedical signal and image processing methods
<b>Keywords:</b>	gait timing; IMUs; wearable inertial sensors; gait events; event detection; walking; temporal parameters; water; algorithm
<b>Corresponding Author:</b>	Giulia Pacini Panebianco University of Bologna Bologna, Bologna Italy
<b>First Author:</b>	Giulia Pacini Panebianco
<b>Order of Authors:</b>	Giulia Pacini Panebianco Maria Cristina Bisi Rita Stagni Silvia Fantozzi
<b>Manuscript Region of Origin:</b>	Europe
<b>Abstract:</b>	<p>Background and Objectives: Walking in water is used for rehabilitation in different pathological conditions. For the characterization of gait alterations related to pathology, gait timing assessment is of primary importance. With the widespread use of inertial sensors, several algorithms have been proposed for gait timing estimation (i.e. gait events and temporal parameters) out of the water, while an assessment of their performance for walking in water is still missing. The purpose of the present study was to assess the performance in the temporal segmentation for gait in water of 17 algorithms proposed in the literature. Methods: Ten healthy volunteers mounting 5 tri-axial inertial sensors (trunk, shanks and feet) walked on dry land and in water. Seventeen different algorithms were implemented and classified based on: 1) sensor position, 2) target variable, and 3) computational approach. Gait events identified from synchronized video recordings were assumed as reference. Temporal parameters were calculated from gait events. Algorithm performance was analysed in terms of sensitivity, positive predictive value, accuracy, and repeatability. Results: For walking in water, all Trunk-based algorithms provided a sensitivity lower than 81% and a positive predictive value lower than 94%, as well as acceleration-based algorithms, independently from sensor location, with the exception of two Shank-based ones. Drop in algorithm sensitivity and positive predictive value was associated to significant differences in the stride pattern of the specific analysed variables during walking in water as compared to walking on dry land, as shown by the intraclass correlation coefficient. When using Shank- or Foot-based algorithms, gait events resulted delayed, but the delay was compensated in the estimate of Stride and Step time; a general underestimation of Stance- and overestimation of Swing-time was observed, with minor exceptions. Conclusion: Sensor position, target variable and computational approach determined different error distributions for different gait events and temporal parameters for walking in water. This work supports an evidence-based selection of the most appropriate algorithm for gait timing estimation for walking in water as related to the specific application, and provides relevant information for the design of new algorithms for the specific motor task.</p>
<b>Suggested Reviewers:</b>	Francesco Di Nardo f.dinardo@univpm.it Expert in gait analysis with inertial sensors on healthy adults and elderlies and pathological people.  Martina Mancini mancinim@ohsu.edu

	<p>Using a variety of body worn sensors to characterize movement (i.e EMG, inertial sensors) and brain activity, this line of research will help in determine how the central nervous system integrates sensory information in healthy individuals and those with movement disorders</p>
	<p>Encarna Mico Amigo          Maria.Mico-Amigo@newcastle.ac.uk          Development of easily-deployable and low-cost medical technology to enhance the quality of public health</p>
	<p>Philippe Terrier          ph.terrier@gmail.com          Sensor combination for enhancing physical activity monitoring in real-life conditions</p>
	<p>Silvia Del Din          silvia.del-din@ncl.ac.uk          Strong research focus in human movement analysis, wearable technology and digital health</p>
	<p>Elise K Vonstad          elise.k.vonstad@ntnu.no          PhD in the field of biomechanics, assessing quality of movement in sports</p>
<p><b>Opposed Reviewers:</b></p>	
<p><b>Response to Reviewers:</b></p>	

13<sup>rd</sup> July 2020

Dear Filippo Molinari  
Editor in Chief

*Computer Methods and Programs in Biomedicine*

**Ref: CMPB\_2020\_49**

**Title: “Timing estimation for gait in water from inertial sensor measurements: analysis of the performance of 17 algorithms”**

We received communication of requested major revision for the above-mentioned manuscript and are now resubmitting the revised version of the manuscript.

We would like to thank the reviewers for their comments that certainly helped to improve the manuscript.

We are providing a point by point response to the reviewers’ comments and the revised version of the manuscript, where modified sections are highlighted.

We hope that this revised version of the manuscript can be suitable for publication on *Computer Methods and Programs in Biomedicine*.

We declare that this material has not been and will not be submitted for publication elsewhere except as an abstract and that there are no competing interests to declare.

Each author has been fully involved with the work, has read and concurs with the content in the final manuscript.

Best regards,

Giulia Pacini Panebianco  
Maria Cristina Bisi  
Rita Stagni  
Silvia Fantozzi

## POINT-BY-POINT RESPONSES TO REVIEWERS' COMMENTS:

First, we would like to thank the reviewers for their comments that helped improving the manuscript.

We modified the manuscript according to reviewers' comments and hope that it will be now suitable for publication in *Computer Methods and Programs in Biomedicine*.

In the following lines, we reported the point-by-point answers to reviewers' comments, referring to changes to the manuscript that were highlighted in yellow in the revised version.

**Reviewer #1:** Objective, quantitative gait analysis is a wide field in which inertial measurement units (IMUs) present an intriguing opportunity for low-cost, easy to use analysis of gait patterns in the lab and in ecologic settings. One of the main challenges of using IMUs is extracting valid and reliable information of the temporal and spatial features of a gait cycle, as acceleration and gyroscopic data is not as readily available to analyze as data captured by marker based 3D motion capture systems. Several algorithms have been evaluated since researchers begun using IMUs for gait analysis, and the effects sensor placement and several features of analysis approach has on gait cycle information has been well documented in above-ground and treadmill walking. Walking in water, however, poses a new set of challenges as the forces acting on the musculoskeletal system is very different from walking on dry land. This creates kinematic and kinetic patterns that are wildly different from walking on dry land. Because of this, it is not known how well the aforementioned algorithms for detecting gait cycle events perform on walking in water compared to walking on land.

The current paper aims to evaluate 17 gait event detection algorithms, recently evaluated for overground walking (Panebianco et al 2018), while in walking in water. Overall the paper is well written, and I am confident that the authors took care in designing an experimental setup that could provide high-quality data that can answer the question at hand. Please see my comments below that I recommend the authors address to improve the paper.

a) My main feedback to the authors of this paper is that they should consider giving a more thorough justification for evaluating algorithms designed for dry land walking on gait cycles from walking in water. These are, as the authors also point out, mechanically very different situations because of the effect being in water has on both kinematic and kinetic features of a gait cycle. Therefore, I am a bit concerned that directly applying these algorithms is unfair, as they are not designed for this situation in the first place. However, as it is not known how badly these algorithms will perform in a water walking situation, the paper still adds value to the body of knowledge as it provides insight into how and where some of these algorithms do fail.

The authors thank the reviewer for the sensible comment and agree regarding the fact that, as general methodological approach, algorithms should not be evaluated in conditions different from the ones they were specifically designed for. On the other hand, to authors' knowledge, no novel algorithm was specifically designed and proposed for the automatic identification of gait events for walking in the water and a specific accurate characterization of the motion pattern for walking in water is still missing, not supporting the design of new algorithms yet. Still considering the

functional and mechanical differences between walking on dry land and in water, some of the basic characteristics, considered for the design of available algorithms, are common to both walking conditions, such as body forward progression and lower limb alternate pendulation and ground contact. In addition, algorithms for over ground walking are designed considering the gait pattern of healthy subjects as reference, but are used for the analysis of pathologic gait, generalizing their application out of the original design context.

Therefore, authors proposed the evaluation of the performance of available algorithms for the analysis of walking in water, aiming to investigate their ability to identify features still common to both locomotor conditions. The authors implemented a systematic analysis of the currently available algorithms for the estimation of gait timing from IMU measurements when applied to the water environment, relating the performance (pros and cons) to the intrinsic characteristics of the algorithms (i.e. sensor position and signal characteristics), in order to support the selection of the most performing approach among the available ones, and/or provide criteria for the design of new ones, if necessary. The concurrent evaluation of the repeatability of the gait pattern and its similarity to the corresponding one out of water is meant to support this analysis.

To better clarify this approach, the introduction (page 4, lines 115-133) and the first part of the discussion section (Page 14 lines 377-384) were revised and integrated accordingly.

**Introduction.** Page 4, Lines 115-133.

*In recent years, the use of inertial wearable sensors (IMUs) has become extensive in the analysis of WDL, and a large number of algorithms for gait segmentation [10,22] were proposed for the identification of GEs for healthy subjects. A recent comprehensive analysis [10] showed that specific performance is significantly affected by sensor placement, analysed variable, and computational approach even in ideal conditions. To authors' knowledge, no novel algorithm was specifically designed and proposed for the automatic identification of GEs for WW and a specific accurate characterization of the motion pattern for WW is still missing, not supporting the design of new algorithms yet. On the other hand, still considering the functional and mechanical differences between WDL and WW [1], some of the basic characteristics, considered for the design of the available algorithms, are common to both walking conditions, such as body forward progression and lower limb alternate pendulation and ground contact. Therefore, the evaluation of the performance of these available algorithms for the analysis of WW is deemed suitable to investigate their ability to identify features still common to both locomotor conditions, providing evidence to support the selection of the most performing approach among the available ones, and/or provide criteria for the design of new ones, if necessary.*

*The aim of the present work was to analyse the performance during WW of 17 different algorithms designed for GE estimation in WDL, taking into account the biomechanical changes induced in the gait pattern by the different environmental conditions, i.e. water. To achieve this goal, the algorithms previously identified through a systematic review [10] underwent a systematic analysis, relating the performance (pros and cons) to the intrinsic characteristics of the algorithms (i.e. sensor position and signal characteristics).*

**Discussion.** Page 14, Lines 377-384.

*The purpose of the present study was to assess the performance of 17 gait segmentation algorithms originally implemented for WDL [10] when applied to WW. The implementation of these algorithms is based on the identification of specific features that can be identified in the gait pattern of healthy subjects during WDL. The mechanical characteristics of the water environment during WW can alter these patterns [1], similarly to pathological and other perturbed conditions, for which WDL algorithms are still applied [45-48]. Since no algorithm was proposed for WW temporal segmentation yet, this preliminary work was design to understand and quantify how the alterations of the gait pattern associated to the presence of the water affect the performance of available algorithms.*

b) The authors report (p 6, line 10) employing a cutoff at a sensitivity of 81% for including algorithms in further analysis. When reading the reference for this decision I'm missing a rationale in the current paper, why is the sensitivity used as cut off, and not positive predictive value?

The threshold for sensitivity was originally chosen from literature for the simple interpretation it provided regarding the performance of the algorithm with respect to the selected video-based reference (number of GEs correctly identified by algorithm/number of GEs identified from video). On the other hand, the positive predictive value (PPV - number of GEs correctly identified by algorithm/number of all GEs identified by algorithm) can effectively integrate this information, providing a better and more thorough analysis of algorithm performance, as correctly suggested by the reviewer. Therefore, PPV was also calculated and the corresponding threshold set at 94% as reported by Trojanello et al. (Trojanello et al., 2014). Thresholding was necessary prior to detailed analysis of algorithm performance, because algorithms unable to detect a sufficient number of GEs are considered unsuitable for applications, and further investigation of error characteristics is not relevant.

- D. Trojaniello, A. Cereatti, U. Della Croce, Accuracy, sensitivity and robustness of five different methods for the estimation of gait temporal parameters using a single inertial sensor mounted on the lower trunk, *GaitPosture*. 40 (2014) 487–492.

Methods and Results sections were integrated accordingly (Methods: Page 7, Lines 201-204; Page 8, Lines 214-215; Results: Page 9, Lines 251-257).

**Methods.** Page 7, Lines 201-204

*For each algorithm considering the video as reference, the sensitivity and the positive predictive value (PPV) in GE identification during WW was calculated [22] respectively as:*

$$\text{Sensitivity} = 100 \times \frac{\text{Number of GEs correctly identified by algorithm}}{\text{Number of all GEs as identified from video}} \quad (2)$$

$$\text{PPV} = 100 \times \frac{\text{Number of GEs correctly identified by algorithm}}{\text{Number of all GEs identified by algorithm}} \quad (3)$$

Page 8, Lines 214-215

*'Only the algorithms reporting a minimum sensitivity of 81% and a minimum PPV of 94% [22] were considered for further analysis.'*

**Results.** Page 9, Lines 251-257.

*'After the sensitivity and PPV analysis, only algorithms exploiting a peak identification approach for shank positioned, and only angular velocity based for foot positioned sensor underwent further error analysis.'*

*Algorithms that passed the 81% sensitivity and 94% PPV criterion showed an ICC for WW above or equal to 0.70, with an ICC of WDL stride cycle applied to WW ranging from 0.31 to 0.61. On the other hand, algorithms that did not pass the 81% sensitivity and 94% PPV showed an ICC for WW below 0.60, with an ICC of WDL stride cycle applied to WW ranging from 0.10 to 0.34, with the only exception of Trojanello et al.[11], showing an ICC for WW equal to 0.79 but an ICC of WDL stride cycle applied to WW of only 0.11.'*

c) In my opinion, it can be discussed whether it is necessary to include all 17 algorithms in this analysis. Some of the algorithms found in the 2018 paper does not perform at a sufficient level even in dry land walking, so I'm wondering if these lesser-performing algorithms should be excluded in analysis of an mechanical situation (i.e walking in water) that is more challenging for these algorithms compared to dry land walking.

The doubt raised by the reviewer was also faced by the authors while designing the present study, therefore we totally understand and appreciate the comment. The reviewer is right pointing out that some algorithms exhibited significantly worse performance for WDL based on the results of the previous study, and that the water environment is expected to modify the mechanical characteristics of the gait pattern. On the other hand, the fundamental characteristics of bipedal locomotion associated to the different features targeted by the different algorithms are still present in gait in the two environments; what can change is the repeatability and reliability of such features, and the effectiveness of the algorithms in identifying them in the specific conditions. Given the extremely scarce characterization of WW, the authors did not consider the available evidence sufficient to speculate on the generalization of WDL analysis results of algorithm performance to WW conditions. Due to lacking evidence for WW, paradoxical behaviours could not be *a priori* excluded. Moreover, the inclusion of all the different approaches in the analysis offers the chance to provide further lacking characterization of WW motion pattern for specific analysed variables. Therefore, the authors decided to analyse the repeatability of the patterns of the variables targeted by the different algorithms within WW and for WW with respect to WDL calculating the corresponding ICC, in order to: i) assess the actual existence of a repeatable pattern in the targeted variable in WW and its similarity to the WDL pattern; ii) identify an objective criterion to support the choice of an algorithm and, at least partially, explain the reason why it is expected to provide a certain performance in the identification of GEs (i.e. are there features to identify in the specific WDL condition? Are these features the same as in WW?).

To better clarify this choice, the Introduction (Page 4, Lines 119-133) was modified.

**Introduction**, Page 4, Lines 119-133.

*'... On the other hand, still considering the functional and mechanical differences between WDL and WW [1], some of the basic characteristics, considered for the design of the available algorithms, are common to both walking conditions, such as body forward progression and lower limb alternate pendulation and ground contact. Therefore, the evaluation of the performance of these available algorithms for the analysis of WW is deemed suitable to investigate their ability to identify features still common to both locomotor conditions, providing evidence to support the selection of the most performing approach among the available ones, and/or provide criteria for the design of new ones, if necessary.'*

*The aim of the present work was to analyse the performance during WW of 17 different algorithms designed for GE estimation in WDL, taking into account the biomechanical changes induced in the gait pattern by the different environmental conditions, i.e. water. To achieve this goal, the algorithms previously identified through a systematic review [19] underwent a systematic analysis, relating the performance (pros and cons) to the intrinsic characteristics of the algorithms (i.e. sensor position and signal characteristics).'*

d) Authors report using an ICC measure of repeatability. What ICC algorithm is used, and why?

The ICC used in the present study for the assessment of repeatability is ICC(3,1), which is one of the six forms of ICC proposed by Shrout and Fleiss (Shrout, and Fleiss, 1979). The same approach, ICC(3,1), was previously exploited in other motion analysis studies: to assess the repeatability of specific parameters calculated by two different instruments during walking on a treadmill (De Castro M.P. et al., 2014), to calculate the similarity of knee kinematic patterns during gait of healthy people (Ateshian G.A. et al., 2020), and to compare the pattern of arm cycling movement between healthy and tetraplegic people (Flueck J.L. et al., 2017). Therefore, ICC(3,1) was considered the most suitable to compare the pattern of the different variables over the gait cycle intra-WW and for WW vs WDL.

- P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, Psychol Bull. 86 (1979) 420–428.
- M.P. de Castro, M. Meucci, D.P. Soares, P. Fonseca, M. Borgonovo-Santos, F. Sousa, L. Machado, J.P. Vilas-Boas, Accuracy and Repeatability of the Gait Analysis by the WalkinSense System, Biomed Res Int. 2014 (2014).
- G.A. Ateshian, K.M. Myers, J.M.R.S. Tavares, Computer Methods, Imaging and Visualization in Biomechanics and Biomedical Engineering: Selected Papers from the 16th International Symposium CMBBE and 4th Conference on Imaging and Visualization, August 14-16, 2019, New York City, USA, Springer Nature, 2020.
- J.L. Flueck, Experimental Protocol of a Three-minute, All-out Arm Crank Exercise Test in Spinal-cord Injured and Able-bodied Individuals, J Vis Exp. (2017).

The type of ICC used for the analysis was described in more detail in the Methods section (Page 7, Lines 208-209).

**Methods**, Page 7, Lines 208-209.

*ICC(3,1) of the mean WDL stride cycle over the WW sequence was calculated, to analyse the similarity of the pattern during WW with respect to the one during WDL, assumed as reference for the algorithm design.*

e) Table 2 and 3 - these are not very clear, and should be revised to increase the value they add to the paper. Consider splitting them up?

Table 2 and 3 were revised, improving their format to support readability, still maintaining related information together, to allow direct comparison. The new versions of Table 2 and Table 3 are reported in the revised manuscript.

f) On page 13, the section regarding results from Computation Approach: Consider putting these results into a figure as well, showing the differences in accuracy and repeatability? This might make these results more easily readable.

Thank you, we carefully considered the suggestion. The section regarding results for computational approach draws from the numerical results reported in Table 3, which shows the statistical differences between the two computational approaches and the corresponding error distributions. Introducing an additional figure with the same content is likely to result redundant, failing to effectively convey quantitative information. These pieces of information are very difficult to be summarized in a single Figure, consisting in 6 different parameters (i.e. FC, FO, stride, stance and swing time) for 2 different IMU positions (i.e. shank and foot). Thus, we would need at least 12 figures, or figure sections, to convey the same information schematically depicted in Table 3. Moreover, boxplots of the errors for each algorithm are already graphically depicted in Figure 1.

Nevertheless, Table 3 was revised, improving format to support readability, still maintaining related information together, to allow direct comparison. The new version of Table 3 is reported in the revised manuscript.

g) Page 15, line 16: The abbreviation "WDP" should be "WDL"? I can't seem to find a reference to WDP anywhere else.

We apologise, it was a typo and was corrected.

**Reviewer #2:** This paper propose to study the performance in the temporal segmentation for gait in water, with 17 state-of-the-art algorithms initially implemented for gait out of the water. In addition to the performance analysis itself depending on the algorithm used, conclusions try to be based on the position of the sensor, the “target variable” (type of sensor used: accelerometer or gyroscope) and the "computational approach” (what kind of feature extraction is performed: peak identification, zero crossing, raw...).

a) I personally feel that the aim of the research is interesting but much more experimental and writing work is needed. The contribution of the article, which in theory is to clarify if the algorithms of gait out of water work into water, is not reached due to lack of study in the state of the art and a clarification of the experiments carried out. First of all, I miss a state-of-the-art study section. It is not clear if previous researches have been performed to study this situation.

The authors appreciate the comment, which highlights the unfortunately extremely limited amount of research available in this specific field. The actual quantitative characterization of WW is still very scarce, largely due to the lack of effective approaches suitable to the operative characteristics of this specific environment, and the number of research studies proposing and/or analysing methods to assess WW is even more limited. The purpose of the proposed manuscript was actually to serve as a preliminary contribution addressing the specific problem. To properly address the specific background and clarify the limited available information, the introduction (Page 4, Lines 111-133) and first part of the discussion section (Page 14, Lines 386-390) were revised and integrated to report relevant background and compare available previous results.

**Introduction.** Page 4, Lines 111-133

*A recent review [13] reports for WW slower speed, lower cadence [1,14–20] and shorter step length [1,20,21] as compared to WDL, with similar duration of stance phase relative to gait cycle (62% for WDL and 60% for WW when the subjects walked at their preferred speeds) [20]. Still the characterization of WW is limited by the lack of a standardised approach for the analysis.*

*In recent years, the use of inertial wearable sensors (IMUs) has become extensive in the analysis of WDL, and a large number of algorithms for gait segmentation [10,22] were proposed for the identification of GEs for healthy subjects. A recent comprehensive analysis [10] showed that specific performance is significantly affected by sensor placement, analysed variable, and computational approach even in ideal conditions. To authors’ knowledge, no novel algorithm was specifically designed and proposed for the automatic identification of GEs for WW and a specific accurate characterization of the motion pattern for WW is still missing, not supporting the design of new algorithms yet. On the other hand, still considering the functional and mechanical differences between WDL and WW [1], some of the basic characteristics, considered for the design of the available algorithms, are common to both walking conditions, such as body forward progression and lower limb alternate pendulation and ground contact. Therefore, the evaluation of the performance of these available algorithms for the analysis of WW is deemed suitable to investigate their ability to identify features still common to both locomotor conditions, providing evidence to*

*support the selection of the most performing approach among the available ones, and/or provide criteria for the design of new ones, if necessary.*

*The aim of the present work was to analyse the performance during WW of 17 different algorithms designed for GE estimation in WDL, taking into account the biomechanical changes induced in the gait pattern by the different environmental conditions, i.e. water. To achieve this goal, the algorithms previously identified through a systematic review [10] underwent a systematic analysis, relating the performance (pros and cons) to the intrinsic characteristics of the algorithms (i.e. sensor position and signal characteristics).*

**Discussion.** Page 14, Lines 386-390.

*In accordance with previous studies using video analysis [16,19], IMUs [5] and force platform [13], in the present study mean speed resulted decreased of 50% in WW as compared to WDL, and cadence was lower (0.36 s<sup>-1</sup> for WW vs for 0.86 s<sup>-1</sup>WDL), similarly to what reported by Heywood et al. [13], and stance relative duration slightly decreased during WW compared to WDL (Mean±Standard Deviation of 58.5% ± 2.9% for WW and of 62.2% ± 1.8% for WDL).*

b) Also, it is not explained why the selected methods have been taken, and this methods are not explained in sufficient detail for the reader.

Unfortunately, no algorithm is currently available for the assessment of GEs from IMU measurements during WW, while several algorithms have been proposed for the purpose for WDL. These were identified by means of a systematic review and their performance analysed in a previously published work by these authors (Pacini et al., 2018). A specific accurate characterization of the motion pattern for WW is still missing, not supporting the design of new algorithms yet. On the other hand, still considering the functional and mechanical differences between walking on dry land and in water, some of the basic characteristics, considered for the design of available algorithms, are common to both walking conditions, such as body forward progression and lower limb alternate pendulation and ground contact. In addition, algorithms for over ground walking are designed considering the gait pattern of healthy subjects as reference, but are used for the analysis of pathologic gait, generalizing their application out of the original design context. Therefore, the evaluation of the performance of available algorithms for the analysis of walking in water was proposed by the authors aiming to investigate their ability to identify features still common to both locomotor conditions. The authors implemented a systematic analysis of the currently available algorithms for the estimation of gait timing from IMU measurements when applied to the water environment, relating the performance (pros and cons) to the intrinsic characteristics of the algorithms (i.e. sensor position and signal characteristics), in order to support the selection of the most performing approach among the available ones, and/or provide criteria for the design of new ones, if necessary. The concurrent evaluation of the repeatability of the gait pattern and its similarity to the corresponding one out of water is meant to support this analysis.

- G. Pacini Panebianco, M.C. Bisi, R. Stagni, S. Fantozzi, Analysis of the performance of 17 algorithms from a systematic review: Influence of sensor position, analysed variable and computational approach in gait timing estimation from IMU measurements, *Gait & Posture*. 66 (2018) 76–82.

To better clarify this approach, the introduction (Page 4, Lines 115-133) and the first part of the discussion section (Page 14, Lines 378-384) were revised and integrated accordingly.

Regarding the characteristics of the analysed algorithms, more details about the implementation characteristics of these algorithms were added in the paragraph ‘Analysed algorithms’ of the Methods section (Page 5, Lines 135-163).

**Introduction.** Page 4, Lines 115-133.

*In recent years, the use of inertial wearable sensors (IMUs) has become extensive in the analysis of WDL, and a large number of algorithms for gait segmentation [10,22] were proposed for the identification of GEs for healthy subjects. A recent comprehensive analysis [10] showed that specific performance is significantly affected by sensor placement, analysed variable, and computational approach even in ideal conditions. To authors’ knowledge, no novel algorithm was specifically designed and proposed for the automatic identification of GEs for WW and a specific accurate characterization of the motion pattern for WW is still missing, not supporting the design of new algorithms yet. On the other hand, still considering the functional and mechanical differences between WDL and WW [1], some of the basic characteristics, considered for the design of the available algorithms, are common to both walking conditions, such as body forward progression and lower limb alternate pendulation and ground contact. Therefore, the evaluation of the performance of these available algorithms for the analysis of WW is deemed suitable to investigate their ability to identify features still common to both locomotor conditions, providing evidence to support the selection of the most performing approach among the available ones, and/or provide criteria for the design of new ones, if necessary.*

*The aim of the present work was to analyse the performance during WW of 17 different algorithms designed for GE estimation in WDL, taking into account the biomechanical changes induced in the gait pattern by the different environmental conditions, i.e. water. To achieve this goal, the algorithms previously identified through a systematic review [10] underwent a systematic analysis, relating the performance (pros and cons) to the intrinsic characteristics of the algorithms (i.e. sensor position and signal characteristics).*

**Methods.** Page 5, Lines 135-163

*Analysed algorithms*

*17 algorithms for GE and gait temporal parameters (GTP) estimation, previously selected and analysed by Pacini et al. [20] for WDL, were here analysed for WW. Following the same approach [20], the selected algorithms were classified based on:*

- i) IMU position (i.e. trunk, shanks, feet);*
- ii) Target variable (i.e. acceleration, angular velocity);*
- iii) Computational approach: ‘peak identification’ and ‘zero crossing’, on raw or filtered target variable i.e. finite impulse response (FIR), infinite impulse response (IIR) and wavelet transform (WT) filtering. ‘Peak identification’ aims to identify specific peaks on the target variable, corresponding to specific temporal events: local maxima or minima of the vertical, antero-posterior or 3D components for acceleration-based algorithms; local minima of the medio-lateral component for angular velocity-based algorithms. ‘Zero crossing’ aims to identify the instants of sign change in the target variable, corresponding to specific temporal events: in the antero-posterior component*

for acceleration-based algorithms; in the medio-lateral component for angular velocity-based algorithms.

The selected algorithms have the following features:

- 6 were trunk-based, analysing antero-posterior [23–26], vertical [27] or 3D [28] components of acceleration. Among these algorithms, 3 adopted ‘peak identification’, of which Buganè et al. with IIR [23], Lee et al. with FIR [24] and McCamley et al. with WT filtering [27], and 3 used ‘zero crossing’ approach, of which Shin et al. with raw signal [28], González et al. with FIR filtering [25] and Zijlstra et al. with IIR filtering [26];

- 7 were shank-based, of which 3 analysing antero-posterior [11] or 3D [29,30] components of acceleration with ‘peak identification’, Trojanello et al. with raw signal [11], Lee et al. with IIR [29] and Khandelwal et al. with WT filtering [30], and 4 analysing the medio-lateral component of angular velocity with ‘peak identification’, Salarian et al. [31] and Greene et al. [32] with raw signal, Catalfamo et al. [33] with IIR and Aminian et al. [34] with WT filtering;

- 4 were foot-based, of which Jasiewicz et al. analysing anterior-posterior acceleration with ‘peak identification’ of raw signal [35], and 3 analysing medio-lateral angular velocity, 2 with ‘peak identification’, Ferrari et al. [36] adopting raw signal, Sabatini et al. [37] with IIR filtering and Mariani et al. [38] with ‘zero crossing’ of IIR filtered signal.

#### **Discussion.** Page 14, Lines 378-384

The implementation of these algorithms is based on the identification of specific features that can be identified in the gait pattern of healthy subjects during WDL. The mechanical characteristics of the water environment during WW can alter these patterns [1], similarly to pathological and other perturbed conditions, for which WDL algorithms are still applied [45-48]. Since no algorithm was proposed for WW temporal segmentation yet, this preliminary work was design to understand and quantify how the alterations of the gait pattern associated to the presence of the water affect the performance of available algorithms.

c) The used dataset is too small to extract reliable conclusions. Are there datasets with more subjects? In any case, affirmative or negative, this should be said in the state of the art.

Unfortunately, the research in the specific field is still very limited and no dataset (independently on the number of subjects) is currently available for the characterization of WW, and studies that addressed WW using different assessment methods still include sample of few subjects (Barela A. et al., 2006, Chevutschi A. et al., 2007, Fantozzi S. et al., 2015), due to the complexity of the experimental approach in the specific condition. Therefore, the proposed study is a preliminary step to prime further research in the field. The number of subjects is not large, but the target population is very focused (i.e. young healthy and physically active subjects with no musculoskeletal or neurological disorder) and its characterization with this preliminary sample provides the basic information necessary to dimension samples for future analysis.

- A.M.F. Barela, S.F. Stolf, M. Duarte, Biomechanical characteristics of adults walking in shallow water and on land, *Journal of Electromyography and Kinesiology*. 16 (2006) 250–256.
- A. Chevutshi, G. Lensele, D. Vaast, A. Thevenon, An electromyographic study of human gait both in water and on dry ground, *J Physiol Anthropol*. 26 (2007) 467–473.
- S. Fantozzi, A. Giovanardi, D. Borra, G. Gatta, Gait Kinematic Analysis in Water Using Wearable Inertial Magnetic Sensors, *PLoS ONE*. 10 (9) (2015)

We agree with the reviewer that this aspect must be properly addressed in the manuscript, and a limitation paragraph was added to the discussion section Page 16, Lines 420-422.

**Discussion.** Page 16, Lines 420-422.

*Being a preliminary study and given the population chosen for the analysis, only 10 healthy adults subjects were considered. However, a larger sample size will be acquired in future investigations to build a dataset for WW similar to the ones available for WDL.*

d) In my humble opinion, tables in general are a bit confusing. For example, Table 1 should be simplified by drawing the plots in separate figures, out of the table. Tables have a lot of numbers per cell too.

We thank the reviewer for the comment. The reviewer correctly points out that the information reported in the tables is very dense, due to the attempt to make as much quantitative information available for applicative use to other researchers. The authors tried to organize the information as schematically and accessibly as possible, still keeping it complete and comprehensive: Tables and Figures were planned to show the error distribution associated to each parameter and algorithm (see boxplots in Figure 1) and as related to the specific implementation characteristic (see Table 2 and 3). In Table 1, information about sensitivity, PPV and ICC values are meant to exhaustively support the interpretation of the suitability of algorithms for gait segmentation in the water. To facilitate the interpretation of the numerical values, plots for each target variable and condition was provided to underline any difference in the gait pattern between the two environments. We understand that Table 1 is dense in terms of information, and might require a little more time for critical evaluation, but removing part of the Table, splitting it up or drawing the plots in different Figures out of the Table, would result in a huge number of Figures, more difficult to navigate, losing the global picture of suitable/unsuitable algorithms for gait segmentation in the water.

To improve the readability of Table 1, we simplified the ‘Algorithm information’ and ‘CV’ columns by removing redundant information. Table 2 and 3 were revised, improving their format to support readability, still maintaining related information together, to allow direct comparison. The new versions of Table 1, Table 2 and Table 3 are reported in the revised manuscript.

d) Although there is a section of nomenclatures, it is advisable that the first time they are used in the text their meaning is indicated.

Agreed. The manuscript was checked thoroughly, and verifying that acronyms were properly defined the first time they were used in the text.

e) Discussion starts like an state-of-the-art section, when this is not the right place for it. In addition, discussion is incomplete, which also leads to a conclusion section which cannot clarify which computational approach is better.

Agreed. The initial part of the discussion section was rewritten, following reviewer suggestion, and integrated with the critical discussion of previous results. Moreover, a specific section discussing computational approach results was added.

**Discussion.** Page 14, Lines 377-395.

*The purpose of the present study was to assess the performance of 17 gait segmentation algorithms originally implemented for WDL [10] when applied to WW. The implementation of these algorithms is based on the identification of specific features that can be identified in the gait pattern of healthy subjects during WDL. The mechanical characteristics of the water environment during WW can alter these patterns [1], similarly to pathological and other perturbed conditions, for which WDL algorithms are still applied [45-48]. Since no algorithm was proposed for WW temporal segmentation yet, this preliminary work was design to understand and quantify how the alterations of the gait pattern associated to the presence of the water affect the performance of available algorithms.*

*In accordance with previous studies using video analysis[16,19], IMUs [5] and force platform [13], in the present study mean speed resulted decreased of 50% in WW as compared to WDL, and cadence was lower (0.36 s<sup>-1</sup> for WW vs for 0.86 s<sup>-1</sup>WDL), similarly to what reported by Heywood et al. [13], and stance relative duration slightly decreased during WW compared to WDL (Mean±Standard Deviation of 58.5% ± 2.9% for WW and of 62.2% ± 1.8% for WDL). As expected, the present work highlighted a significant reduction in the sensitivity and PPV of the analysed algorithms for WW.*

*Based on the values of the ICC, sensitivity and PPV still resulted over the 81% and 94% thresholds, respectively, when the gait pattern of the target variable was sufficiently repeatable (ICC for WW above or equal to 0.70), although differing from the WDL pattern (ICC of WDL over WW from 0.31 to 0.61).*

Page 15, Lines 413-419

*For computational approach, no final conclusion can be drawn, given that for Shank-based algorithms only peak identification approach was applied to both acceleration and angular velocity, while both peak identification and zero crossing resulted applied to only angular velocity for Foot-based ones. Therefore, the analysis can hardly evaluate the performance of the computational approach independently from sensor location. Considering Shank- and Foot-based algorithms separately, comparable performance were found between raw signal and different signal filtering (i.e. IIR and WT) in the evaluation of Stride and Step time, while Stance and Swing time resulted under- or over-estimated, depending on the specific signal processing.*

**Reviewer #3:** The study presents a comparative evaluation of performances achieved during walking in water (WW) by 17 different IMU-based algorithms, originally proposed for assessing gait events and temporal parameters during ground walking. Results were provided in terms of sensor position, target variable, and computational approach. Although no novel methodology was proposed, the effort of adapting and testing 17 different algorithms to "walking-in-water" condition is absolutely remarkable and could be potentially very useful for quantitative gait-parameter assessment in rehabilitation programs. To my knowledge, indeed, no specific algorithm for gait segmentation during walking in water is present in literature. Thus, I believe that this manuscript deserves the publication in Computer Methods and Programs in Biomedicine. However, I have some minor issues, which should be handled.

a) Please add the line number to facilitate possible further revisions.

Agreed. The line numbers were added in the manuscript.

b) Please assign a number to the formulas throughout the manuscript.

Agreed. Formulas were numbered.

c) As reported by authors, walking in water represents one of the fundamental motor tasks performed during aquatic therapy. However, readers interested in gait analysis or in rehabilitation (who are familiar with the values of gait events and parameters during standard ground walking) do not necessarily know the values of this parameter in WW and their variations from the standard values. Thus, the Introduction would take advantage from reporting and discussing numerical values of temporal parameters and their deviations from the values adopted during standard walking (ex: Foot off is typically located around 60% of standard gait cycle. Is this true also for WW?).

Thank you for the suggestion. We agreed that providing information about values of temporal parameters during WW would improve the Introduction. The following paragraph was added accordingly:

**Introduction.** Page 4, Lines 111-114.

*A recent review [13] reports for WW slower speed, lower cadence [1,14-20] and shorter step length [1,20,21] as compared to WDL, with similar duration of stance phase relative to gait cycle (62% for WDL and 60% for WW when the subjects walked at their preferred speeds) [20]. Still the characterization of WW is limited by the lack of a standardised approach for the analysis.*

d) Although the 17 algorithms were describes in a previous study from the same group of researchers [10], the readability of the present manuscript and the comprehension of the results could take advantage by a brief description of the main points the algorithms are based on.

Agreed. More details about the implementation characteristics of the selected algorithms were added in the paragraph 'Analysed algorithms' of the Methods section.

**Methods.** Page 5, Lines 135-163.

### *Analysed algorithms*

*17 algorithms for GE and gait temporal parameters (GTP) estimation, previously selected and analysed by Pacini et al. [20] for WDL, were here analysed for WW. Following the same approach [20], the selected algorithms were classified based on:*

*i) IMU position (i.e. trunk, shanks, feet);*

*ii) Target variable (i.e. acceleration, angular velocity);*

*iii) Computational approach: ‘peak identification’ and ‘zero crossing’, on raw or filtered target variable i.e. finite impulse response (FIR), infinite impulse response (IIR) and wavelet transform (WT) filtering. ‘Peak identification’ aims to identify specific peaks on the target variable, corresponding to specific temporal events: local maxima or minima of the vertical, antero-posterior or 3D components for acceleration-based algorithms; local minima of the medio-lateral component for angular velocity-based algorithms. ‘Zero crossing’ aims to identify the instants of sign change in the target variable, corresponding to specific temporal events: in the antero-posterior component for acceleration-based algorithms; in the medio-lateral component for angular velocity-based algorithms.*

*The selected algorithms have the following features:*

*- 6 were trunk-based, analysing antero-posterior [23–26], vertical [27] or 3D [28] components of acceleration. Among these algorithms, 3 adopted ‘peak identification’, of which Buganè et al. with IIR [23], Lee et al. with FIR [24] and McCamley et al. with WT filtering [27], and 3 used ‘zero crossing’ approach, of which Shin et al. with raw signal [28], González et al. with FIR filtering [25] and Zijlstra et al. with IIR filtering [26];*

*- 7 were shank-based, of which 3 analysing antero-posterior [11] or 3D [29,30] components of acceleration with ‘peak identification’, Trojanello et al. with raw signal [11], Lee et al. with IIR [29] and Khandelwal et al. with WT filtering [30], and 4 analysing the medio-lateral component of angular velocity with ‘peak identification’, Salarian et al. [31] and Greene et al. [32] with raw signal, Catalfamo et al. [33] with IIR and Aminian et al. [34] with WT filtering;*

*- 4 were foot-based, of which Jasiewicz et al. analysing anterior-posterior acceleration with ‘peak identification’ of raw signal [35], and 3 analysing medio-lateral angular velocity, 2 with ‘peak identification’, Ferrari et al. [36] adopting raw signal, Sabatini et al. [37] with IIR filtering and Mariani et al. [38] with ‘zero crossing’ of IIR filtered signal.*

e) Pg 5. In my opinion, Eref should be computed in absolute value.

Agreed. The equation was modified accordingly in the Method section (Page 7, Line 196) as:

$$E_{ref} = \max (\text{abs}(GE_{GoPro} - GE_{GRF})) \quad (1)$$

f) Pg 5. Since FC and FO have been visually identified from the video recordings, I wonder if authors checked out if the number of events estimated by the video approach are the same provided by GRF one (human eye can be wrong).

Thank you for the comment. The authors verified that each foot contact and foot off that was visually identified from the video recordings corresponded with the same event provided by the force platform applying a 20 N threshold to the vertical component of the ground reaction force (Zeni J.A. et al., 2008).

- J.A. Zeni, J.G. Richards, J.S. Higginson, Two simple methods for determining gait events during treadmill and overground walking using kinematic data, *Gait Posture*. 27 (2008) 710–714.

We agreed to better clarifying that the number of events estimated using the video approach are the same provided by ground reaction force.

**Methods.** Page 7, Lines 193-195

*'After verifying that the number of events estimated by the video recordings were the same provided by GRF, the measurement error of video-based assessment was estimated considering the GRF as reference:..'*

g) Pg. 6, First formula. This is essentially an assessment of sensitivity, known also as recall [=true positives/(true positives + false negatives)]. Did authors compute also the false positives and thus the precision [=true positives/(true positives + false positives)]? In my opinion, this would complete the information.

Agreed. As suggested, we added the information about precision, setting the threshold for at 94% as reported by Trojanello et al. (Trojanello et al., 2014). The method section was modified accordingly: the cut-off for positive predictive value was introduced with the relative mathematical equation and it was used together with the sensitivity as a criterion to discriminate suitable algorithms for gait events detection during walking in water. Moreover, numerical values of PPV were added in Table 1, and Results section was integrated reporting the calculated values of PPV (Methods: Page 7, Lines 201-204; Page 8, Lines 214-215; Results: Page 9, Lines 251-257).

- D. Trojaniello, A. Cereatti, U. Della Croce, Accuracy, sensitivity and robustness of five different methods for the estimation of gait temporal parameters using a single inertial sensor mounted on the lower trunk, *Gait Posture*. 40 (2014) 487–492.

**Methods.** Page 7, Lines 201-204

*For each algorithm considering the video as reference, the sensitivity and the positive predictive value (PPV) in GE identification during WW was calculated [22] respectively as:*

$$\text{Sensitivity} = 100 \times \frac{\text{Number of GEs correctly identified by algorithm}}{\text{Number of all GEs as identified from video}} \quad (2)$$

$$\text{PPV} = 100 \times \frac{\text{Number of GEs correctly identified by algorithm}}{\text{Number of all GEs identified by algorithm}} \quad (3)$$

**Methods.** Page 8, Lines 214-215

*'Only the algorithms reporting a minimum sensitivity of 81% and a minimum PPV of 94% [22] were considered for further analysis.'*

**Results.** Page 9, Lines 251-257.

*'After the sensitivity and PPV analysis, only algorithms exploiting a peak identification approach for shank positioned, and only angular velocity based for foot positioned sensor underwent further error analysis.'*

*Algorithms that passed the 81% sensitivity and 94% PPV criterion showed an ICC for WW above or equal to 0.70, with an ICC of WDL stride cycle applied to WW ranging from 0.31 to 0.61. On the other hand, algorithms that did not pass the 81% sensitivity and 94% PPV showed an ICC for WW below 0.60, with an ICC of WDL stride cycle applied to WW ranging from 0.10 to 0.34, with the only exception of Trojanello et al.[11], showing an ICC for WW equal to 0.79 but an ICC of WDL stride cycle applied to WW of only 0.11.'*

h) Pg 7: Authors stated: "Maximum measurement error of video- versus GRF-reference resulted 0.05 s for both FC and FO." More information is needed. How many subjects were considered to compute E<sub>ref</sub>? Besides maximum, also mean (or median) value could be useful. Moreover, are the errors always positive? Negative? Maybe a table reporting all E<sub>ref</sub> value or at least average values over each subject (it depends on the number of subjects/strides) could help.

Thank you for the question. The assessment of maximum measurement error was performed on 10 subjects, considering 20 FC and 20 FO per subject. According to the previous request of this Reviewer, the E<sub>ref</sub> was calculated as E<sub>ref</sub>=max(abs(GE<sub>GoPro</sub>- GE<sub>GRF</sub>)). The value of 0.05s represents maximum measurement error of video- versus GRF-reference, thus it is not worth to analyse the positive or negative bias and, as the Reviewer underlined earlier, it can be computed also in absolute value. Differently, errors between sensors and video recordings should be computed with the original sign, to underline any anticipation/delay in the identification of the gait events with different approach and properly assess the performance of the algorithms in relation to the implementation characteristics.

In order to better specify the number of foot contact and foot off considered for the analysis, the Methods section was modified accordingly:

**Methods,** Page 7, Lines 191-193.

*'For WDL, 20 FC<sub>GRF</sub> and 20 FO<sub>GRF</sub> were automatically identified applying a 20 N threshold to the vertical component of the GRF recordings [40], and FC<sub>GoPro</sub> and FO<sub>GoPro</sub> were visually identified from the video recordings for each participant and trial.'*

i) Pg 10: "IMU Position" paragraph:

Authors found significant differences among Shank- and Foot-based algorithms. To compare mean (or median) values, parametric (or non-parametric) test such as t-test or Anova (kruskal-wallis, Wilcoxon) were typically implemented. It is not clear to me (maybe it is only my fault) what statistical test has been used to evaluate this significance. Please discuss.

As the reviewer pointed out, parametric or non-parametric tests can be applied to compare data distribution. However, linear-mixed models are being used ever more frequently in the analysis of experimental data (Luke S.G. et al., 2017). For example, they were recently applied to evaluate the effect of the speed during gait in water (Fantozzi S. et al., 2015) or to analyse differences in gait biomechanical variables by two different prosthetic knee types (Segal A.D. et al., 2006). In this study, a linear mixed model was used to examine the effect of different implementation characteristic (i.e. positioning, target variable and computational approach) on the identification of gait events and gait temporal parameters. For example, considering the positioning, linear mixed models were used to identify the effects of the different location of sensors (i.e. trunk, shank and foot) on each outcome variable. The models were specified as follows:

$$y_{ti} = \beta_0 + \beta_1 \cdot x_{ti} + u_{0i} + \varepsilon_{0ti}$$

where  $\beta_0$  represent the fixed intercepts,  $u_{0i}$  represent the random effects associated to the  $i^{\text{th}}$  subject, the subscripts  $ti$  indicate values of variables measured in the  $t^{\text{th}}$  step clustered in the  $i^{\text{th}}$  subject,  $y$  represents the outcome variable,  $x$  represents the independent variable, and  $\varepsilon_0$  is the random error component.

The linear mixed model was fit using the lme4 package in R software (R-Core Team 2017, Austria, version 3.4.3) (Bates D. et al., 2018), with the significance level set at 0.05.

- S.G. Luke, Evaluating significance in linear mixed-effects models in R, Behav Res. 49 (2017) 1494–1502.
- S. Fantozzi, A. Giovanardi, D. Borra, G. Gatta, Gait Kinematic Analysis in Water Using Wearable Inertial Magnetic Sensors, PLoS ONE. 10 (9) (2015).
- Segal AD, Orendurff MS, Klute GK, et al. Kinematic and kinetic comparisons of transfemoral amputee gait using C-Leg and Mauch SNS prosthetic knees. J Rehabil Res Dev. 2006;43(7):857-870.
- D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models using lme4, ArXiv:1406.5823 [Stat]. (2014).

We agree that more details can be useful, thus, we inserted in Methods the following information about the function used in the linear mixed model:

**Methods**, Page 8, Lines 221-229.

*'The models were specified as follows:*

$$y_{ti} = \beta_0 + \beta_1 \cdot x_{ti} + u_{0i} + \varepsilon_{0ti} \quad (7)$$

where  $\beta_0$  represent the fixed intercepts,  $u_{0i}$  represent the random effects associated to the  $i^{\text{th}}$  subject, the subscripts  $ti$  indicate values of variables measured in the  $t^{\text{th}}$  step clustered in the  $i^{\text{th}}$  subject,  $y$  represents the outcome variable,  $x$  represents the independent variable, and  $\varepsilon_0$  is the random error component. The linear mixed model was fit using the *lme4* package in R software (R-Core Team 2017, Austria, version 3.4.3) [42], with the significance level set at 0.05.'

l) What did authors mean with "but with different error distribution"? Please, explain with more details.

With the statement 'Shank-based algorithms resulted equally accurate and repeatable in FC estimation (Med/Dmed: 0.25/0.15 s for both Shank and Foot), but with different error distribution' the authors intended that Shank- and Foot-based algorithms showed the same interpercentile difference around the median value of the error, i.e. 0.15 s. However, as reported in Table 2, values of 25<sup>th</sup>/75<sup>th</sup> percentiles resulted of 0.20/0.35s for Shank-based algorithms and of 0.15/0.30s for Foot-based algorithms, leading to the same interpercentile difference, but with different distribution of the errors.

We agreed to better clarify the description of the results as follows:

Page 11, Lines 284-289.

*For GE, significant differences were found between Shank- and Foot-based algorithms: Shank-based algorithms resulted equally accurate and repeatable in FC estimation (Med/Dmed: 0.25/0.15 s for both Shank and Foot), but with different values of 25th and 75th percentiles (0.20s and 0.35s for Shank-based algorithms and of 0.15s and 0.30s for Foot-based algorithms). Moreover, Shank-based algorithms were equally accurate and more repeatable in FO estimation (Med/Dmed: 0.20/0.20 s and 0.20/0.30 s, for Shank and Foot, respectively) than Foot-based ones.*

m) The reference reported for ICC definition is Shrout and Fleiss (1979), which is an old study. In this study, authors reported that "there are numerous versions of the intraclass correlation coefficient" and I agree with them. The computation of ICC has been used for the performance evaluation of the segmentation algorithm of an altered gait pattern (WW), thus results could be sensitive to the kind of adopted ICC. Thus, my question is what version of ICC has been implemented in the present manuscript? Please discuss.

The ICC used in the present study for the assessment of repeatability is ICC(3,1), which is one of the six forms of ICC proposed by Shrout and Fleiss (Shrout, and Fleiss, 1979). The same approach, ICC(3,1), was previously exploited in other motion analysis studies, to assess the repeatability of specific parameters calculated by two different instruments during walking on a treadmill (De Castro M.P. et al., 2014), to calculate the similarity of knee kinematic patterns during gait of healthy people (Ateshian G.A. et al., 2020), to compare the pattern of arm cycling movement between healthy and tetraplegic people (Flueck J.L. et al., 2017). Therefore, ICC(3,1) was considered the most suitable to compare the pattern of the different variables over the gait cycle intra-WW and for WW vs WDL.

- P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, *Psychol Bull.* 86 (1979) 420–428.
- M.P. de Castro, M. Meucci, D.P. Soares, P. Fonseca, M. Borgonovo-Santos, F. Sousa, L. Machado, J.P. Vilas-Boas, Accuracy and Repeatability of the Gait Analysis by the WalkinSense System, *Biomed Res Int.* 2014 (2014).
- G.A. Ateshian, K.M. Myers, J.M.R.S. Tavares, *Computer Methods, Imaging and Visualization in Biomechanics and Biomedical Engineering: Selected Papers from the 16th International Symposium CMBBE and 4th Conference on Imaging and Visualization, August 14-16, 2019, New York City, USA, Springer Nature, 2020.*
- J.L. Flueck, Experimental Protocol of a Three-minute, All-out Arm Crank Exercise Test in Spinal-cord Injured and Able-bodied Individuals, *J Vis Exp.* (2017).

The type of ICC used for the analysis was described in more detail in the Methods section (Page 7, Lines 208-209).

**Methods.** Page 7, Lines 208-209

*ICC(3,1) of the mean WDL stride cycle over the WW sequence was calculated, to analyse the similarity of the pattern during WW with respect to the one during WDL, assumed as reference for the algorithm design.*

n) Many techniques were reported to provide IMU-based assessment of gait parameters during standard ground walking. Not so many were reported for WW. However, some data on spatial-temporal parameters during WW could be found [Heywood et al. *Gait Posture*, 2016]. Although this is a comparative study (among 17 algorithms), comparing the performances of these algorithms also with state-of-the-art WW data could really improve the discussion of the study.

Thank you for the suggestion. As the reviewer pointed out, relevant results are reported in the study performed by Heywood et al. (Heywood S. et al., 2016) about spatiotemporal and kinematic parameters, force and muscle activation. This information was added to both introduction and discussion sections. Information concerning spatiotemporal parameters was added to the Introduction and the relevant results critically discussed with respect to those summarised by Heywood et al. (Heywood S. et al., 2016).

- S. Heywood, J. McClelland, P. Geigle, A. Rahmann, R. Clark, Spatiotemporal, kinematic, force and muscle activation outcomes during gait and functional exercise in water compared to on land: A systematic review, *Gait Posture.* 48 (2016) 120–130.

The following paragraphs have been added;

**Introduction.** Page 4, Lines 111-114.

*A recent review [13] reports for WW slower speed, lower cadence [1,14–20] and shorter step length [1,20,21] as compared to WDL, with similar duration of stance phase relative to gait cycle (62% for WDL and 60% for WW when the subjects walked at their preferred speeds) [20]. Still the characterization of WW is limited by the lack of a standardised approach for the analysis.*

**Discussion.** Page 14, Lines 386-390.

*In accordance with previous studies using video analysis[16,19], IMUs [5] and force platform [13], in the present study mean speed resulted decreased of 50% in WW as compared to WDL, and cadence was lower (0.36 s-1 for WW vs for 0.86 s-1WDL), similarly to what reported by Heywood et al. [13], and stance relative duration slightly decreased during WW compared to WDL (Mean±Standard Deviation of 58.5% ± 2.9% for WW and of 62.2% ± 1.8% for WDL).*

o) As I already reported, the results of this study could be very useful. However, they could be much more useful if author would take a position indicating which algorithm/algorithms (among the 17 ones) should be used in WW.

We appreciate suggestion to be more ‘courageous’ in providing a definitive indication to the reader. On the other hand, the present work was meant to provide a comprehensive picture of the performance of analysed algorithm for WW, as well as to propose a systematic approach for their analysis. The possible indication of ‘the algorithm that should be used’ largely depends on the research requirements and constraints of the final user. Moreover, other algorithms could be proposed and compared using the same approach for the analysis, integrating the present one. Finally, the extension of the evaluation dataset could further improve the results. Therefore, we do not think that a definitive indication regarding the algorithm to be preferred is ours to be given, leaving to the user the choice, based on the provided evidence based on its research needs (e.g. if having specific constraints for sensors placement, or measured variable, or target parameter).

We agree to adding a final paragraph describing these considerations, in order to help the reader in understanding the conclusion of the work.

**Conclusions.** Page 16, Lines 433-438.

*All analysed factors resulted to affect gait segmentation during WW. No proposed algorithm can be generally preferred over the others, but the reported results can support researchers in the choice of the most suitable algorithm/algorithms based on experimental condition (e.g. number/type/placement of sensors) and research question (e.g. calculation of specific parameter). Finally, these results can support the design of novel segmentation algorithms, better addressing alterations induced in the gait pattern by the water environment.*

p) Despite the numerous algorithms tested, only 10 subjects were recruited for the study. This may affect the statistical analysis. Please include this in the limitations of the study.

Thank you for the suggestion. We agreed in including this aspect in the Discussion, as a limitation of the study as reported here below.

**Discussion.** Page 16, Lines 420-422.

*Being a preliminary study and given the population chosen for the analysis, only 10 healthy adults subjects were considered. However, a larger sample size will be acquired in future investigations to build a dataset for WW similar to the ones available for WDL.*

## **Highlights**

- Timing estimation with inertial sensors was assessed for gait in water environment
- Performance of 17 algorithms was analysed based on implementation characteristics
- No trunk-based algorithm is suitable for gait timing estimation in water
- Shank/foot-based algorithm performance depends on analysed target variable
- Target variable ICC supports the assessment of algorithm performance in altered gait

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

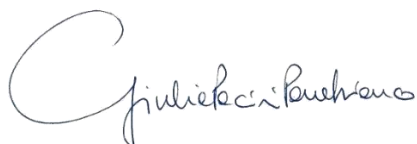
We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

We further confirm that any aspect of the work covered in this manuscript that has involved either experimental animals or human patients has been conducted with the ethical approval of all relevant bodies and that such approvals are acknowledged within the manuscript.

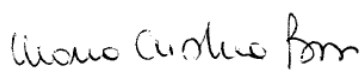
We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and direct communications with the office). She is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address (giulia.pacini2@unibo.it) which is accessible by the Corresponding Author and which has been configured to accept email from the Journal.

Signed by all authors as follows:

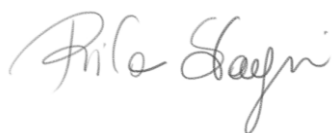
Giulia Pacini Panebianco, 08/01/2020



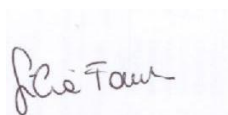
Maria Cristina Bisi 08/01/2020



Rita Stagni 08/01/2020



Silvia Fantozzi 08/01/2020





## 39 Abstract

40 *Background and Objectives:* Walking in water is used for rehabilitation in different pathological  
41 conditions. For the characterization of gait alterations related to pathology, gait timing assessment is of  
42 primary importance. With the widespread use of inertial sensors, several algorithms have been proposed for  
43 gait timing estimation (i.e. gait events and temporal parameters) out of the water, while an assessment of  
44 their performance for walking in water is still missing. The purpose of the present study was to assess the  
45 performance in the temporal segmentation for gait in water of 17 algorithms proposed in the literature.  
46 *Methods:* Ten healthy volunteers mounting 5 tri-axial inertial sensors (trunk, shanks and feet) walked on dry  
47 land and in water. Seventeen different algorithms were implemented and classified based on: 1) sensor  
48 position, 2) target variable, and 3) computational approach. Gait events identified from synchronized video  
49 recordings were assumed as reference. Temporal parameters were calculated from gait events. Algorithm  
50 performance was analysed in terms of sensitivity, positive predictive value, accuracy, and repeatability.  
51 *Results:* For walking in water, all Trunk-based algorithms provided a sensitivity lower than 81% and a  
52 positive predictive value lower than 94%, as well as acceleration-based algorithms, independently from  
53 sensor location, with the exception of two Shank-based ones. Drop in algorithm sensitivity and positive  
54 predictive value was associated to significant differences in the stride pattern of the specific analysed  
55 variables during walking in water as compared to walking on dry land, as shown by the intraclass correlation  
56 coefficient. When using Shank- or Foot-based algorithms, gait events resulted delayed, but the delay was  
57 compensated in the estimate of Stride and Step time; a general underestimation of Stance- and overestimation  
58 of Swing-time was observed, with minor exceptions. *Conclusion:* Sensor position, target variable and  
59 computational approach determined different error distributions for different gait events and temporal  
60 parameters for walking in water. This work supports an evidence-based selection of the most appropriate  
61 algorithm for gait timing estimation for walking in water as related to the specific application, and provides  
62 relevant information for the design of new algorithms for the specific motor task.

63  
64 *Index Terms*— algorithm, event detection, gait events, gait timing, IMUs, temporal parameters, walking,  
65 water, wearable inertial sensors

66

67

68

69

70

71

72 **Nomenclature**

73 WW: Walking in water

74 WDL: Walking on dry land

75 GE: Gait Event

76 GTP: Gait Temporal Parameter

77 FC: Foot Contact

78 FO: Foot Off

79 IMU: Inertial Measurement Unit

80 FIR: Finite Impulse Response

81 IIR: Infinite Impulse Response

82 WT: Wavelet Transform

83 GRF: Ground Reaction Force

84 E: Error

85 **PPV: positive predictive values**

86 Med: Median

87 Dmed: Dispersion around median value

88 ICC: Intraclass correlation coefficient

89 CV: Coefficient of Variation

90

91

## 92 Introduction

93 The water environment plays a relevant role in rehabilitation programs [1]. During aquatic therapy, the  
94 buoyancy, the drag force, and the pressure exerted by water reduce the gravitational load on joints, resist  
95 motion, and increase proprioception [2–5]. In recent years, Walking in Water (WW) has become extensively  
96 used for people with specific gait deviations [1] and represents one of the fundamental motor tasks performed  
97 during aquatic therapy [5]. It is recommended for developing and maintaining cardiorespiratory and muscular  
98 fitness [6], can be adopted to speed up recovery from minor orthopaedic injuries [7], offers clear advantages  
99 over the land-based equivalent for populations with high risk of fall such as older adults and neurological  
100 patients [8], and can be practiced by individuals without swimming skills [9].

101 Assessment of motor performance during WW, based on quantitative motion analysis, would support the  
102 understanding of water-induced biomechanical modifications and the design and/or monitoring of WW based  
103 rehabilitation.

104 In quantitative motion analysis, the assessment of temporal parameters is of primary importance, implying  
105 the correct identification of gait events (GEs), i.e. Foot Contact (FC) and Foot Off (FO) [10,11]. Various  
106 technologies (i.e. force platforms, instrumented mats, footswitches) can be and were used for identifying GE  
107 when Walking on Dry Land (WDL), but no validated approach is available for gait timing identification for  
108 WW, as highlighted by Matsumoto et al., 2008 [12]. Several studies exploited a camera-based approaches  
109 for WW [1,5,6,12], although its drawbacks: i) the limited field of view allowing to analyse only one/two  
110 consecutive steps; ii) the time-consuming set-up and post-processing [1,6].

111 A recent review [13] reports for WW slower speed, lower cadence [1,14–20] and shorter step length  
112 [1,20,21] as compared to WDL, with similar duration of stance phase relative to gait cycle (62% for WDL  
113 and 60% for WW when the subjects walked at their preferred speeds) [20]. Still the characterization of WW  
114 is limited by the lack of a standardised approach for the analysis.

115 In recent years, the use of inertial wearable sensors (IMUs) has become extensive in the analysis of WDL,  
116 and a large number of algorithms for gait segmentation [10,22] were proposed for the identification of GEs  
117 for healthy subjects. A recent comprehensive analysis [10] showed that specific performance is significantly  
118 affected by sensor placement, analysed variable, and computational approach even in ideal conditions. To  
119 authors' knowledge, no novel algorithm was specifically designed and proposed for the automatic  
120 identification of GEs for WW and a specific accurate characterization of the motion pattern for WW is still  
121 missing, not supporting the design of new algorithms yet. On the other hand, still considering the functional  
122 and mechanical differences between WDL and WW [1], some of the basic characteristics, considered for the  
123 design of the available algorithms, are common to both walking conditions, such as body forward progression  
124 and lower limb alternate pendulation and ground contact. Therefore, the evaluation of the performance of

these available algorithms for the analysis of WW is deemed suitable to investigate their ability to identify features still common to both locomotor conditions, providing evidence to support the selection of the most performing approach among the available ones, and/or provide criteria for the design of new ones, if necessary.

The aim of the present work was to analyse the performance during WW of 17 different algorithms designed for GE estimation in WDL, taking into account the biomechanical changes induced in the gait pattern by the different environmental conditions, i.e. water. To achieve this goal, the algorithms previously identified through a systematic review [10] underwent a systematic analysis, relating the performance (pros and cons) to the intrinsic characteristics of the algorithms (i.e. sensor position and signal characteristics).

## Methods

### *Analysed algorithms*

17 algorithms for GE and gait temporal parameters (GTP) estimation, previously selected and analysed by Pacini et al. [20] for WDL, were here analysed for WW. Following the same approach [20], the selected algorithms were classified based on:

- i) IMU position (i.e. trunk, shanks, feet);
- ii) Target variable (i.e. acceleration, angular velocity);
- iii) Computational approach: ‘peak identification’ and ‘zero crossing’, on raw or filtered target variable i.e. finite impulse response (FIR), infinite impulse response (IIR) and wavelet transform (WT) filtering. ‘Peak identification’ aims to identify specific peaks on the target variable, corresponding to specific temporal events: local maxima or minima of the vertical, antero-posterior or 3D components for acceleration-based algorithms; local minima of the medio-lateral component for angular velocity-based algorithms. ‘Zero crossing’ aims to identify the instants of sign change in the target variable, corresponding to specific temporal events: in the antero-posterior component for acceleration-based algorithms; in the medio-lateral component for angular velocity-based algorithms.

The selected algorithms have the following features:

- 6 were trunk-based, analysing antero-posterior [23–26], vertical [27] or 3D [28] components of acceleration. Among these algorithms, 3 adopted ‘peak identification’, of which Buganè et al. with IIR [23], Lee et al. with FIR [24] and McCamley et al. with WT filtering [27], and 3 used ‘zero crossing’ approach, of which Shin et al. with raw signal [28], González et al. with FIR filtering [25] and Zijlstra et al. with IIR filtering [26];
- 7 were shank-based, of which 3 analysing antero-posterior [11] or 3D [29,30] components of acceleration with ‘peak identification’, Trojanello et al. with raw signal [11], Lee et al. with IIR [29]

157 and Khandelwal et al. with WT filtering [30], and 4 analysing the medio-lateral component of angular  
158 velocity with ‘peak identification’, Salarian et al. [31] and Greene et al. [32] with raw signal, Catalfamo  
159 et al. [33] with IIR and Aminian et al. [34] with WT filtering;  
160 - 4 were foot-based, of which Jasiewicz et al. analysing anterior-posterior acceleration with ‘peak  
161 identification’ of raw signal [35], and 3 analysing medio-lateral angular velocity, 2 with ‘peak  
162 identification’, Ferrari et al. [36] adopting raw signal, Sabatini et al. [37] with IIR filtering and Mariani  
163 et al. [38] with ‘zero crossing’ of IIR filtered signal.  
164

## 165 *Experimental analysis*

### 166 *Participants.*

167 Ten young adult healthy participants (5 females, 5 males; age  $26.2 \pm 3.3$  years; height  $1.71 \pm 0.07$  m; weight  
168  $65.4 \pm 8.6$  Kg) were recruited in the study. All participants were physically active and self-reported no  
169 musculoskeletal or neurological disorder. The Bioethics Committee of the University of Bologna approved  
170 the study on 13/07/2018 with protocol number 99412, and informed consent was signed by all participants.  
171

### 172 *Data acquisition.*

173 Each participant walked 5 times back and forth along a 10 m straight pathway at self-selected speed in 2  
174 conditions: i) WDL; ii) WW at 1.2m depth with water temperature of  $28^{\circ}\text{C}$  wearing water shoes and keeping  
175 the arms on the water surface. No device (e.g. metronome or timer) was used to control the walking speed,  
176 not to interfere with the natural walking pattern [39]. Before the WW analysed trial, participants performed  
177 an acclimatisation trial.  
178

179 Five tri-axial IMUs (Cometa, Italy, sf=285 Hz; accelerometer: sensitivity 156,3 mV/g, range  $\pm 8g$ ;  
180 gyroscope: sensitivity 1,3 mV/g, range  $\pm 1000^{\circ}/\text{s}$ ) were attached to the trunk (at L5 level), shanks (about five  
181 centimetres above lateral malleolus), and feet (on the dorsal aspect of each shoe). The walking tasks were  
182 also filmed using a video camera (Hero4, GoPro, USA, sf=240Hz, 848x480 pixels resolution) for GE  
183 detection reference. IMU led flashing was video-recorded and used for time-synchronization of IMU and  
184 video recording.

185 During WDL trails, ground reaction forces (GRF) were also recorded using two force platforms (Kistler,  
186 Winterthur, Switzerland, sf=1000Hz) mounted half-way along the pathway. A trigger signal was generated  
187 by IMU system at the beginning of each trial for synchronization with the force platforms.

### 188 *Data analysis.*

189 Average gait speed was calculated from reference video during WDL and WW as the ratio between straight  
190 walked distance and time.

191 For WDL, 20  $FC_{GRF}$  and 20  $FO_{GRF}$  were automatically identified applying a 20 N threshold to the vertical  
192 component of the GRF recordings [40], and  $FC_{GoPro}$  and  $FO_{GoPro}$  were visually identified from the video  
193 recordings for each participant and trial. After verifying that the number of events estimated by the video  
194 recordings were the same provided by GRF, the measurement error of video-based assessment was estimated  
195 considering the GRF as reference:

196

$$E_{ref} = \max(\text{abs}(GE_{GoPro} - GE_{GRF})) \quad (1)$$

197

198 For WW,  $FC_{GoPro}$  and  $FO_{GoPro}$  were visually identified as for WDL, and  $FC_{IMU}$  and  $FO_{IMU}$  were estimated  
199 from IMU measurements using the selected 17 algorithms [10], implemented in MATLAB (MathWorks  
200 2017a, USA). GTPs were calculated from GE.

201 For each algorithm considering the video as reference, the sensitivity and the positive predictive value  
202 (PPV) in GE identification during WW was calculated [22] respectively as:

203

$$Sensitivity = 100 \times \frac{\text{Number of GEs correctly identified by algorithm}}{\text{Number of all GEs as identified from video}} \quad (2)$$

$$PPV = 100 \times \frac{\text{Number of GEs correctly identified by algorithm}}{\text{Number of all GEs identified by algorithm}} \quad (3)$$

204

205 For each target variable of the 17 selected algorithms, the intraclass correlation coefficient (ICC) (3,1) of  
206 the mean stride cycle over the whole sequence of each trial for WW and WDL was calculated according to  
207 Shrout et al. [41] to analyse the repeatability of the pattern over the trial in each condition.

208 ICC(3,1) of the mean WDL stride cycle over the WW sequence was calculated, to analyse the similarity of  
209 the pattern during WW with respect to the one during WDL, assumed as reference for the algorithm design.

210 For each algorithm and each condition, coefficient of variation (CV) of the analysed target signal was  
211 calculated as:

212

$$\frac{\text{Difference between 75th and 25th percentile over the stride cycle}}{(\text{Abs}(\text{median of the median value over the stride cycle}))} \quad (4)$$

213

214 Only the algorithms reporting a minimum sensitivity of 81% and a minimum PPV of 94% [22] were  
 215 considered for further analysis.

216 For each algorithm, the error was calculated for GE and GTP, considering the video as reference, as:

217

$$E_{GE} = GE_{IMU} - GE_{GoPro} \quad (5)$$

$$E_{GTP} = GTP_{IMU} - GTP_{GoPro} \quad (6)$$

218

219 *Statistical analysis.*

220 For each parameter (FC, FO, Stride Time, Step Time, Stance Time, Swing Time), a linear mixed model [42]  
 221 was applied to test the dependency of error values on each implementation criterion. The models were  
 222 specified as follows:

223

$$y_{ti} = \beta_0 + \beta_1 \cdot x_{ti} + u_{0i} + \varepsilon_{0ti} \quad (7)$$

224

225 where  $\beta_0$  represent the fixed intercepts,  $u_{0i}$  represent the random effects associated to the  $i^{th}$  subject, the  
 226 subscripts  $ti$  indicate values of variables measured in the  $t^{th}$  step clustered in the  $i^{th}$  subject,  $y$  represents the  
 227 outcome variable,  $x$  represents the independent variable, and  $\varepsilon_0$  is the random error component. The linear  
 228 mixed model was fit using the lme4 package in R software (R-Core Team 2017, Austria, version 3.4.3) [42],  
 229 with the significance level set at 0.05. First, the statistical analysis was performed to investigate the influence  
 230 of IMU position and target variable, alone. Then, the influence of analysed variable and computational  
 231 approach were investigated separately for each IMU position.

232 Median value (Med) of the error was calculated to characterize accuracy, and the Dispersion around Med  
 233 (Dmed, 75th percentile – 25th percentile values of the error) to characterize repeatability.

234

235

236

237

238

239

240

241 **Results**

242 Gait speed normalised according to Hof [43] resulted  $0,33\pm 0,06$  for WDL and  $0,17\pm 0,08$  for WW.

243 Maximum measurement error of video- versus GRF-reference resulted 0.05 s for both FC and FO.

244

245 During WW, 32 FCs and FOs were identified and analysed for each participant, for a total of 320 FCs and  
246 FOs.

247 No algorithm exploiting a sensor on the trunk passed the 81% sensitivity and 94% PPV, as well as no  
248 acceleration-based algorithm, independently from sensor placement, with the exception of two, both  
249 exploiting a sensor positioned on the shank and a peak identification approach: i) Khandelwal et al. [44] on  
250 WT signal; ii) Lee et al. [29] on a IIR transformed signal.

251 After the sensitivity and PPV analysis, only algorithms exploiting a peak identification approach for shank  
252 positioned, and only angular velocity based for foot positioned sensor underwent further error analysis.

253 Algorithms that passed the 81% sensitivity and 94% PPV criterion showed an ICC for WW above or equal  
254 to 0.70, with an ICC of WDL stride cycle applied to WW ranging from 0.31 to 0.61. On the other hand,  
255 algorithms that did not pass the 81% sensitivity and 94% PPV showed an ICC for WW below 0.60, with an  
256 ICC of WDL stride cycle applied to WW ranging from 0.10 to 0.34, with the only exception of Trojanello et  
257 al. [11], showing an ICC for WW equal to 0.79 but an ICC of WDL stride cycle applied to WW of only 0.11.

258 No trend was observed for the CV of the variable analysed by each algorithm.

259 Sensitivity, PPV and ICC analysis results are summarised in Table 1.

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

**Table 1:** Sensitivity of algorithms in FC and FO identification; ICC, CV with Minimum, Med (in bold), Maximum, and plots with Med, 25th and 75th percentile of the different target signals in relation to the implementation criteria defined on the normalized gait cycle and referred to WDL, WW, WDL and WW (thicker lines for WDL plots).

Algorithm	Sensitivity (%)	PPV (%)	WDL			WW			WDL and WW	
	FC FO	FC FO	ICC	CV	Plot	ICC	CV	Plot	ICC	Plot
Bugané et al., 2012	< 81	< 94	0.79	1.3 2.7 11.3		0.33	3.3 4.2 6.5		0.24	
Lee et al., 2009			0.79	1.3 2.7 11.1		0.33	3.3 4.2 6.5		0.23	
McCamley et al., 2012			0.90	1.8 2.9 7.9		0.37	164.9 229.1 362.9		0.10	
González et al., 2010			0.82	1.5 2.5 9.0		0.34	3.1 4.0 6.3		0.23	
Shin et al., 2011			0.76	732.0 1063.6 3042.3		0.19	4.5 6.7 13.7		0.15	
Zijlstra et al., 2003			0.93	2.9 5.3 8.3		0.44	2.2 3.0 4.3		0.34	
Lee et al., 2010			98 99	98 100	0.91	0.8 2.2 6.8		0.70	30.6 155.0 490.0	
Trojaniello et al., 2014	< 81	< 94	0.75	1.7 5.4 35.3		0.79	4.1 6.1 13.8		0.11	
Khandelwal et al., 2014	97 99	97 98	0.91	2.1 14.7 31.2		0.76	4.5 18.0 42.0		0.56	
Catalfamo et al., 2010	98 100	98 98	0.97	0.3 1.1 3.5		0.86	0.5 1.5 6.0		0.61	
Greene et al., 2010	98 100	98 100	0.96	0.3 1.1 3.5		0.86	0.5 1.5 6.0		0.61	
Salarian et al., 2004	100 100	99 98	0.96	0.3 1.1 3.5		0.86	0.5 1.5 6.0		0.61	
Aminian et al., 2002	98 100	94 96	0.96	0.3 1.1 3.5		0.86	0.5 1.4 5.9		0.61	
Jasiewicz et al., 2006	< 81	< 94	0.76	0.5 1.6 15.1		0.56	5.7 15.3 38.8		0.19	
Sabatini et al., 2005	99 100	98 99	0.95	1.2 8.7 36.4		0.82	1.5 4.6 12.3		0.56	
Ferrari et al., 2016	93 95	94 95	0.94	1.3 9.1 41.4		0.82	1.6 4.8 13.6		0.54	
Mariani et al., 2013	99 100	99 100	0.94	1.3 8.9 38.6		0.82	1.4 4.4 11.7		0.55	

279 In general, FC and FO estimates resulted delayed, with this delay being compensated in Stride and Step  
 280 time estimates, while Stance Time was under- and Swing Time over-estimated.

281 In more detail, considering the different implementation characteristics:

### 283 *IMU Position*

284 For GE, significant differences were found between Shank- and Foot-based algorithms: Shank-based  
 285 algorithms resulted equally accurate and repeatable in FC estimation (Med/Dmed: 0.25/0.15 s for both Shank  
 286 and Foot), but with different values of 25<sup>th</sup> and 75<sup>th</sup> percentiles (0.20s and 0.35s for Shank-based algorithms  
 287 and of 0.15s and 0.30s for Foot-based algorithms). Moreover, Shank- based algorithms were equally accurate  
 288 and more repeatable in FO estimation (Med/Dmed: 0.20/0.20 s and 0.20/0.30 s, for Shank and Foot,  
 289 respectively) than Foot-based ones.

290 For GTP, no significant difference was found for Stride and Step time estimates, while Shank-based  
 291 algorithms resulted equally accurate but more repeatable in Stance Time (Med/Dmed: -0.05/0.25 s and -  
 292 0.05/0.35 s, for Shank and Foot, respectively) and less accurate but more repeatable in Swing Time estimation  
 293 (Med/Dmed: 0.05/0.25 s and 0.00/0.35 s, for Shank and Foot, respectively) than Foot-based ones.

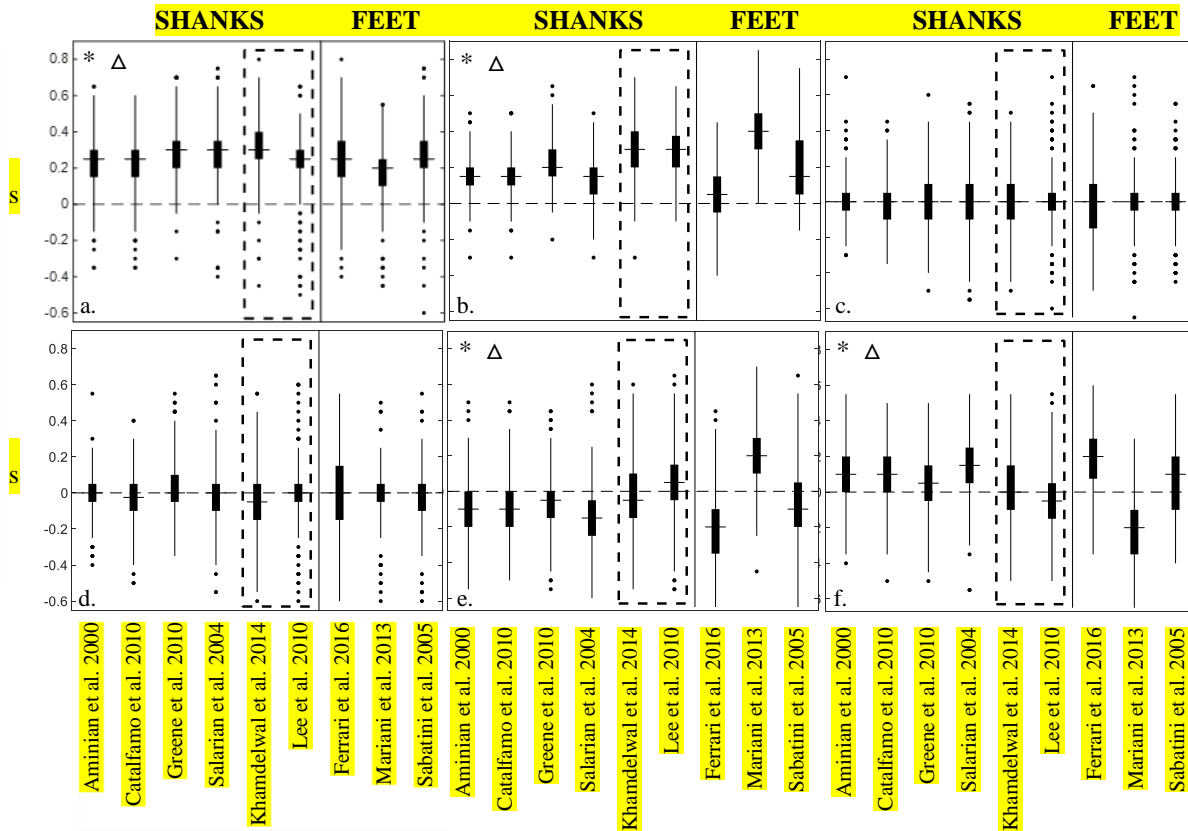
### 295 *Target variable*

296 For GE, Acceleration based algorithms resulted equally accurate but more repeatable for FC (Med/Dmed:  
 297 0.25/0.15 s and 0.25/0.20 s, for Acceleration and Angular velocity, respectively), and less accurate and  
 298 equally repeatable in FO estimation (Med/Dmed: 0.30/0.20 s and 0.15/0.20 s, for Acceleration and Angular  
 299 velocity, respectively) than Angular velocity-based ones.

300 For GTP, no significant difference was found in Stride time and Step time estimation, while Acceleration-  
 301 based algorithms resulted more accurate and equally repeatable in Stance time estimation (Med/Dmed:  
 302 0.00/0.25 s and -0.10/0.25 s, for Acceleration and Angular velocity, respectively), and Swing time estimation  
 303 (Med/Dmed: -0.00/0.25 s and 0.10/0.25 s, for Acceleration and Angular velocity, respectively) than Angular  
 304 velocity-based ones.

306 Error characteristics for FC (a), FO (b), Stride- (c), Step- (d), Stance- (e) and Swing time (f) as related to  
 307 IMU position and target variable are schematically depicted in Figure 1, while error distribution is reported  
 308 in Table 2.

309  
 310  
 311  
 312  
 313  
 314



**Figure 1:** Box plot (minimum, 25th percentile, Med, 75th percentile, maximum values) for FC (a), FO (b), Stride- (c), Step- (d), Stance- (e), and Swing time (f) estimation errors as related to IMU position and target variable. Acceleration-based algorithms are framed in dashes (-). Significant differences were highlighted with asterisk (\*) for IMU position and with triangle (Δ) for target variable.

315  
316  
317  
318  
319  
320  
321

**Table 2:** Results of statistical analysis for IMU positioning and target variable for WW: 25th quartile, Med (in bold) and 75th quartile of estimation error for FC, FO, stride time, step time, stance time and swing time (\* p<0.05).

Parameter	Estimation of errors: IMU position (s)						Level of significance	Estimation of errors: target variable (s)						Level of significance	
	Shanks			Feet				Shanks – Feet			Acceleration				Angular velocity
FC	0.20	<b>0.25</b>	0.35	0.15	<b>0.25</b>	0.30	*	0.20	<b>0.25</b>	0.35	0.15	<b>0.25</b>	0.35	*	
FO	0.10	<b>0.20</b>	0.30	0.05	<b>0.20</b>	0.35	*	0.20	<b>0.30</b>	0.40	0.10	<b>0.15</b>	0.30	*	
Stride Time	-0.10	<b>0.00</b>	0.05	-0.10	<b>0.00</b>	0.05		-0.10	<b>0.00</b>	0.05	-0.10	<b>0.00</b>	0.05		
Step Time	-0.10	<b>0.00</b>	0.05	-0.10	<b>0.00</b>	0.10		-0.10	<b>0.00</b>	0.05	-0.10	<b>0.00</b>	0.10		
Stance Time	-0.20	<b>-0.05</b>	0.05	-0.20	<b>-0.05</b>	0.15	*	-0.10	<b>0.00</b>	0.15	-0.20	<b>-0.10</b>	0.05	*	
Swing Time	-0.05	<b>0.05</b>	0.20	-0.15	<b>0.00</b>	0.20	*	-0.15	<b>0.00</b>	0.10	-0.05	<b>0.10</b>	0.20	*	

322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334

335 *Computational approach*

336 Shank-based algorithms exploited only ‘peak identification’ approach. Raw and WT filtering performed  
337 equally (Med/DL: 0.30/0.15 s), but less accurate and repeatable than IIR (Med/DL: 0.25/0.10 s) in FC  
338 estimation, while IIR and WT performed equally (Med/DL: 0.20/0.20 s), but equally accurate and less  
339 repeatable than Raw (Med/DL: 0.20/0.15 s) in FO estimation. Considering GTP, no significant difference  
340 was found among Raw, IIR and WT for Stride and Step time estimation in terms of accuracy, with decreasing  
341 (Med/DL: 0.00/0.20 s, 0.00/0.15 s, and 0.00/0.10 s, for Raw, IIR and WT, respectively) and equal  
342 repeatability (Med/DL: 0.00/0.15 s, for all) for Stride and Step time estimation, respectively; on the other  
343 hand, Raw, WT and IIR showed and increasing accuracy, respectively, in Stance and Swing time estimation,  
344 but Raw and WT resulted more repeatable than IIR (Med/DL: -0.10/0.20 s, 0.00/0.25 s, and -0.05/0.20 s, for  
345 Raw, IIR and WT, respectively for Stance Time; Med/DL: 0.10/0.20 s, 0.00/0.25 s, and 0.05/0.20 s, for Raw,  
346 IIR and WT, respectively for Swing Time).

347  
348 Foot-based algorithms exploited both ‘peak identification’ (both Raw and IIR for filtering) and ‘zero  
349 crossing’ (only IIR for filtering) approaches, as applied to angular velocity. ‘Zero crossing’ with IIR resulted  
350 more accurate in FC, but less accurate in FO estimate than ‘peak identification’, where IIR resulted equally  
351 accurate in FC estimate and less accurate in FO estimate than Raw, with comparable repeatability.

352 No significant difference was found in Stride and Step time estimation for both computational approach  
353 and filtering technique. On the other hand, ‘zero crossing’ with IIR: i) delayed and resulted more repeatable,  
354 while ‘peak identification’ anticipated Stance time; ii) anticipated and resulted less accurate, while ‘peak  
355 identification’ delayed’ Swing time. For both GTP IIR resulted more accurate than Raw.

356  
357 Error values and statistical analysis for computational approach are reported in Table 3.

358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373

**Table 3:** Results of statistical analysis for computational approach during WW: 25th quartile, Med (in bold) and 75th quartile of estimation error for FC, FO, stride time, step time, stance time and swing time (\* p<0.05).

IMU position	Parameter	Level of significance 'peak identification' vs 'zero crossing'	Estimation of errors: Filtering (s)									
			Filtering within 'peak identification'		Level of significance for filtering within 'peak identification'		Filtering within 'zero crossing'	Level of significance for filtering within 'zero crossing'				
Shank	FC	-	Raw	0.20	<b>0.30</b>	0.35	IIR - Raw *	-				
			IIR	0.20	<b>0.25</b>	0.30	IIR - WT *					
			WT	0.20	<b>0.30</b>	0.35	Raw - WT					
	FO		Raw	0.10	<b>0.20</b>	0.25	IIR - Raw *					
			IIR	0.10	<b>0.20</b>	0.30	IIR - WT					
			WT	0.10	<b>0.20</b>	0.30	Raw - WT *					
	Stride Time		Raw	-0.10	<b>0.00</b>	0.10	IIR - Raw					
			IIR	-0.10	<b>0.00</b>	0.05	IIR - WT					
			WT	-0.05	<b>0.00</b>	0.05	Raw - WT					
	Step Time		Raw	-0.05	<b>0.00</b>	0.10	IIR - Raw *					
			IIR	-0.10	<b>0.00</b>	0.05	IIR - WT					
			WT	-0.10	<b>0.00</b>	0.05	Raw - WT *					
	Stance Time		Raw	-0.20	<b>-0.10</b>	0.00	IIR - Raw *					
			IIR	-0.15	<b>0.00</b>	0.10	IIR - WT *					
			WT	-0.15	<b>-0.05</b>	0.05	Raw - WT *					
	Swing Time		Raw	0.00	<b>0.10</b>	0.20	IIR - Raw *					
			IIR	-0.10	<b>0.00</b>	0.15	IIR - WT *					
			WT	-0.05	<b>0.05</b>	0.15	Raw - WT *					
Foot	FC	*	Raw	0.15	<b>0.25</b>	0.35	IIR - Raw	-				
			IIR	0.20	<b>0.25</b>	0.35	IIR - Raw		IIR	0.10	<b>0.20</b>	0.25
	FO		Raw	-0.05	<b>0.05</b>	0.15	IIR - Raw *		IIR	0.30	<b>0.40</b>	0.50
			IIR	0.05	<b>0.15</b>	0.35	IIR - Raw *		IIR	-0.05	<b>0.00</b>	0.05
	Stride Time		Raw	-0.15	<b>0.00</b>	0.10	IIR - Raw		IIR	-0.05	<b>0.00</b>	0.05
			IIR	-0.05	<b>0.00</b>	0.05	IIR - Raw		IIR	-0.05	<b>0.00</b>	0.05
	Step Time		Raw	-0.15	<b>0.00</b>	0.15	IIR - Raw		IIR	-0.05	<b>0.00</b>	0.05
			IIR	-0.10	<b>0.00</b>	0.05	IIR - Raw		IIR	-0.05	<b>0.00</b>	0.05
	Stance Time		Raw	-0.35	<b>-0.20</b>	-0.10	IIR - Raw *		IIR	0.10	<b>0.20</b>	0.30
			IIR	-0.20	<b>-0.10</b>	0.05	IIR - Raw *		IIR	-0.35	<b>-0.20</b>	-0.10
	Swing Time		Raw	0.10	<b>0.20</b>	0.30	IIR - Raw *		IIR	-0.35	<b>-0.20</b>	-0.10
			IIR	-0.10	<b>0.10</b>	0.20	IIR - Raw *		IIR	-0.35	<b>-0.20</b>	-0.10

374  
375

## 376 Discussion

377 The purpose of the present study was to assess the performance of 17 gait segmentation algorithms  
 378 originally implemented for WDL [10] when applied to WW. The implementation of these algorithms is based  
 379 on the identification of specific features that can be identified in the gait pattern of healthy subjects during  
 380 WDL. The mechanical characteristics of the water environment during WW can alter these patterns [1],  
 381 similarly to pathological and other perturbed conditions, for which WDL algorithms are still applied [45–  
 382 48]. Since no algorithm was proposed for WW temporal segmentation yet, this preliminary work was design  
 383 to understand and quantify how the alterations of the gait pattern associated to the presence of the water affect  
 384 the performance of available algorithms.

385

386 In accordance with previous studies using video analysis [16,19], IMUs [5] and force platform [13], in the  
 387 present study mean speed resulted decreased of 50% in WW as compared to WDL, cadence was lower (0.36

388  $s^{-1}$  for WW vs for  $0.86 s^{-1}$ WDL), similarly to what reported by Heywood et al.[13], and stance relative  
389 duration slightly decreased during WW compared to WDL (Mean±Standard Deviation of  $58.5\% \pm 2.9\%$  for  
390 WW and of  $62.2\% \pm 1.8\%$  for WDL).

391 As expected, the present work highlighted a significant reduction in the sensitivity and PPV of the analysed  
392 algorithms for WW.

393 Based on the values of the ICC, sensitivity and PPV still resulted over the 81% and 94% thresholds,  
394 respectively, when the gait pattern of the target variable was sufficiently repeatable (ICC for WW above or  
395 equal to 0.70), although differing from the WDL pattern (ICC of WDL over WW from 0.31 to 0.61).

396 On the other hand, sensitivity and PPV were below threshold when WW pattern was not repeatable enough  
397 (ICC for WW below 0.60), and when it differed too much from the reference WDL one, like for Trojanello  
398 et al. [11], showing a repeatable WW pattern (ICC for WW 0.79) but very different from WDL one (ICC of  
399 WDL over WW 0.11). According to the reported results, ICC analysis can provide an effective method for  
400 the objective preliminary evaluation of performance of a specific algorithm to the segmentation of an altered  
401 gait pattern.

402 In particular, no Trunk-based algorithm passed the sensitivity criterion, while Shank- and Foot-based ones  
403 provided better performance, with Shank-based ones performing slightly better in terms of repeatability. The  
404 failure of Trunk-based algorithms is associated to the aforementioned drop in WW gait pattern repeatability,  
405 but also to the disappearing of the pendulum pattern characterizing the reference WDL pattern [26], while  
406 for the shanks and feet, ICC values of WDL pattern over WW result higher due to the still occurring alternate  
407 swing of the lower limbs, associated to bi-pedal progression. Nevertheless, due to the alteration of the pattern  
408 during WW, Shank- and Foot-based algorithms did no longer perform as differently as for WDL [10].

409 Considering the influence of target variable, acceleration never passed the sensitivity and PPV criterion,  
410 with the only exception of 2 Shank-based algorithms [29,44]. Nevertheless, these two resulted slightly more  
411 repeatable in FC-, less accurate in FO identification, more accurate in Stance- and Swing-time estimates than  
412 Angular-velocity based ones.

413 For computational approach, no final conclusion can be drawn, given that for Shank-based algorithms only  
414 peak identification approach was applied to both acceleration and angular velocity, while both peak  
415 identification and zero crossing resulted applied to only angular velocity for Foot-based ones. Therefore, the  
416 analysis can hardly evaluate the performance of the computational approach independently from sensor  
417 location. Considering Shank- and Foot-based algorithms separately, comparable performance were found  
418 between raw signal and different signal filtering (i.e. IIR and WT) in the evaluation of Stride and Step time,  
419 while Stance and Swing time resulted under- or over-estimated, depending on the specific signal processing.

420 Being a preliminary study and given the population chosen for the analysis, only 10 healthy adults subjects  
421 were considered. However, a larger sample size will be acquired in future investigations to build a dataset  
422 for WW similar to the ones available for WDL.

423 The limited accuracy of the video-based reference for GE identification for WW can be considered a second  
424 limitation of the present study. On the other hand, no other reference measure is available in the water  
425 environment, and the relevant minimal detectable difference properly taken into account in the statistical  
426 analysis, supporting the reliability of the results.

## 427

### 428 **Conclusion**

429 According to the results of the present research work: i) no available Trunk-based algorithm is suitable to  
430 gait segmentation for WW, due to the disappearance of pendulum mechanics; ii) angular velocity based  
431 algorithms with sensor located on the shank and feet result more reliable in terms of sensitivity and PPV than  
432 acceleration based ones, but not in terms of accuracy and repeatability; iii) no final conclusion can be drawn  
433 regarding the computational approach, independently from sensor location and target variable. All analysed  
434 factors resulted to affect gait segmentation during WW. No proposed algorithm can be generally preferred  
435 over the others, but the reported results can support researchers in the choice of the most suitable  
436 algorithm/algorithms based on experimental condition (e.g. number/type/placement of sensors) and research  
437 question (e.g. calculation of specific parameter). Finally, these results can support the design of novel  
438 segmentation algorithms, better addressing alterations induced in the gait pattern by the water environment.

452 **References**

- 453 [1] A.M.F. Barela, S.F. Stolf, M. Duarte, Biomechanical characteristics of adults walking in shallow water  
454 and on land, *J. Electromyogr. Kinesiol.* 16 (2006) 250–256.
- 455 [2] D.A. Skelton, S.M. Dinan, Exercise for falls management: Rationale for an exercise programme aimed  
456 at reducing postural instability, *Physiother. Theory Pract.* 15 (1999) 105–120.  
457 <https://doi.org/10.1080/095939899307801>.
- 458 [3] E. Watanabe, N. Takeshima, A. Okada, K. Inomata, Comparison of water- and land-based exercise in  
459 the reduction of state anxiety among older adults, *Percept. Mot. Skills.* 91 (2000) 97–104.  
460 <https://doi.org/10.2466/pms.2000.91.1.97>.
- 461 [4] B.E. Becker, Aquatic therapy: scientific foundations and clinical rehabilitation applications, *PM R.* 1  
462 (2009) 859–872. <https://doi.org/10.1016/j.pmrj.2009.05.017>.
- 463 [5] S. Fantozzi, A. Giovanardi, D. Borra, G. Gatta, Gait Kinematic Analysis in Water Using Wearable  
464 Inertial Magnetic Sensors, *PloS One.* 10 (2015) e0138105.  
465 <https://doi.org/10.1371/journal.pone.0138105>.
- 466 [6] K. Masumoto, T. Shono, S. Takasugi, N. Hotta, K. Fujishima, Y. Iwamoto, Age-related differences in  
467 muscle activity, stride frequency and heart rate response during walking in water, *J. Electromyogr.*  
468 *Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.* 17 (2007) 596–604.  
469 <https://doi.org/10.1016/j.jelekin.2006.06.006>.
- 470 [7] K. Kaneda, H. Wakabayashi, D. Sato, T. Nomura, Lower extremity muscle activity during different  
471 types and speeds of underwater movement, *J. Physiol. Anthropol.* 26 (2007) 197–200.
- 472 [8] A.C. Severin, B.J. Burkett, M.R. McKean, M.G.L. Sayers, Biomechanical aspects of aquatic therapy:  
473 A literature review on application and methodological challenges - Kaenz, *J. Fit. Res.* 5 (2016) 49–62.
- 474 [9] S.P. Brown, L.F. Chitwood, K.R. Beason, D.R. McLemore, Deep Water Running Physiologic  
475 Responses: Gender Differences at Treadmill-Matched Walking/Running Cadences, *J. Strength Cond.*  
476 *Res.* 11 (1997) 107.
- 477 [10] G. Pacini Panebianco, M.C. Bisi, R. Stagni, S. Fantozzi, Analysis of the performance of 17 algorithms  
478 from a systematic review: Influence of sensor position, analysed variable and computational approach  
479 in gait timing estimation from IMU measurements, *Gait Posture.* 66 (2018) 76–82.  
480 <https://doi.org/10.1016/j.gaitpost.2018.08.025>.
- 481 [11] D. Trojaniello, A. Cereatti, E. Pelosin, L. Avanzino, A. Mirelman, J.M. Hausdorff, U. Della Croce,  
482 Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-  
483 mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait, *J.*  
484 *NeuroEngineering Rehabil.* 11 (2014) 152. <https://doi.org/10.1186/1743-0003-11-152>.

- 485 [12] K. Masumoto, T. Shono, N. Hotta, K. Fujishima, Muscle activation, cardiorespiratory response, and  
486 rating of perceived exertion in older subjects while walking in water and on dry land, *J. Electromyogr.*  
487 *Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.* 18 (2008) 581–590.  
488 <https://doi.org/10.1016/j.jelekin.2006.12.009>.
- 489 [13] S. Heywood, J. McClelland, P. Geigle, A. Rahmann, R. Clark, Spatiotemporal, kinematic, force and  
490 muscle activation outcomes during gait and functional exercise in water compared to on land: A  
491 systematic review, *Gait Posture.* 48 (2016) 120–130. <https://doi.org/10.1016/j.gaitpost.2016.04.033>.
- 492 [14] L.C. Carneiro, S.M. Michaelsen, H. Roesler, A. Hauptenthal, M. Hubert, E. Mallmann, Vertical reaction  
493 forces and kinematics of backward walking underwater., *Gait Posture.* 35 (2012) 225–230.  
494 <https://doi.org/10.1016/j.gaitpost.2011.09.011>.
- 495 [15] T. Miyoshi, T. Shirota, S.-I. Yamamoto, K. Nakazawa, M. Akai, Functional roles of lower-limb joint  
496 moments while walking in water, *Clin. Biomech.* 20 (2005) 194–201.  
497 <https://doi.org/10.1016/j.clinbiomech.2004.10.006>.
- 498 [16] T. Miyoshi, T. Shirota, S. Yamamoto, K. Nakazawa, M. Akai, Effect of the walking speed to the lower  
499 limb joint angular displacements, joint moments and ground reaction forces during walking in water,  
500 *Disabil. Rehabil.* 26 (2004) 724–732. <https://doi.org/10.1080/09638280410001704313>.
- 501 [17] M.I.V. Orselli, M. Duarte, Joint forces and torques when walking in shallow water, *J. Biomech.* 44  
502 (2011) 1170–1175. <https://doi.org/10.1016/j.jbiomech.2011.01.017>.
- 503 [18] A. Chevutschi, G. Lensel, D. Vaast, A. Thevenon, An electromyographic study of human gait both in  
504 water and on dry ground, *J. Physiol. Anthropol.* 26 (2007) 467–473.
- 505 [19] A. Chevutschi, M. Alberty, G. Lensel, V. Pardessus, A. Thevenon, Comparison of maximal and  
506 spontaneous speeds during walking on dry land and water, *Gait Posture.* 29 (2009) 403–407.  
507 <https://doi.org/10.1016/j.gaitpost.2008.10.059>.
- 508 [20] A.M.F. Barela, M. Duarte, Biomechanical characteristics of elderly individuals walking on land and in  
509 water, *J. Electromyogr. Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.* 18 (2008) 446–454.  
510 <https://doi.org/10.1016/j.jelekin.2006.10.008>.
- 511 [21] A.M. Degani, A. Danna-dos-Santos, The Effect of Water Walking on the Lower Limb Motion of Older  
512 Adults, *Int. J. Aquat. Res. Educ.* 1 (2007). <https://doi.org/10.25035/ijare.01.03.03>.
- 513 [22] D. Trojaniello, A. Cereatti, U. Della Croce, Accuracy, sensitivity and robustness of five different  
514 methods for the estimation of gait temporal parameters using a single inertial sensor mounted on the  
515 lower trunk, *Gait Posture.* 40 (2014) 487–492. <https://doi.org/10.1016/j.gaitpost.2014.07.007>.
- 516 [23] F. Bugané, M.G. Benedetti, G. Casadio, S. Attala, F. Biagi, M. Manca, A. Leardini, Estimation of  
517 spatial-temporal gait parameters in level walking based on a single accelerometer: Validation on normal

- 518 subjects by standard gait analysis, *Comput. Methods Programs Biomed.* 108 (2012) 129–137.  
519 <https://doi.org/10.1016/j.cmpb.2012.02.003>.
- 520 [24] Hyo-Ki Lee, Sung-Jae Hwang, Sung-Pil Cho, Dong-Ryul Lee, Sung-Hyun You, Kyoung-Joung Lee,  
521 Yong-Ho Kim, Ho-Seon Choi, Novel algorithm for the hemiplegic gait evaluation using a single 3-axis  
522 accelerometer, in: 2009 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., IEEE, Minneapolis, MN, 2009:  
523 pp. 3964–3966. <https://doi.org/10.1109/IEMBS.2009.5333650>.
- 524 [25] R.C. González, A.M. López, J. Rodríguez-Uría, D. Álvarez, J.C. Alvarez, Real-time gait event detection  
525 for normal subjects from lower trunk accelerations, *Gait Posture.* 31 (2010) 322–325.  
526 <https://doi.org/10.1016/j.gaitpost.2009.11.014>.
- 527 [26] W. Zijlstra, Assessment of spatio-temporal parameters during unconstrained walking, *Eur. J. Appl.*  
528 *Physiol.* 92 (2004) 39–44. <https://doi.org/10.1007/s00421-004-1041-5>.
- 529 [27] J. McCamley, M. Donati, E. Grimpampi, C. Mazzà, An enhanced estimate of initial contact and final  
530 contact instants of time using lower trunk inertial sensor data, *Gait Posture.* 36 (2012) 316–318.  
531 <https://doi.org/10.1016/j.gaitpost.2012.02.019>.
- 532 [28] S.H. Shin, C.G. Park, Adaptive step length estimation algorithm using optimal parameters and  
533 movement status awareness, *Med. Eng. Phys.* 33 (2011) 1064–1071.  
534 <https://doi.org/10.1016/j.medengphy.2011.04.009>.
- 535 [29] J.-A. Lee, S.-H. Cho, Y.-J. Lee, H.-K. Yang, J.-W. Lee, Portable activity monitoring system for  
536 temporal parameters of gait cycles, *J. Med. Syst.* 34 (2010) 959–966. [https://doi.org/10.1007/s10916-](https://doi.org/10.1007/s10916-009-9311-8)  
537 [009-9311-8](https://doi.org/10.1007/s10916-009-9311-8).
- 538 [30] S. Khandelwal, N. Wickström, Identification of Gait Events Using Expert Knowledge and Continuous  
539 Wavelet Transform Analysis, in: *Proc. Int. Jt. Conf. Biomed. Eng. Syst. Technol. - Vol. 4,*  
540 *SCITEPRESS - Science and Technology Publications, Lda, Portugal, 2014: pp. 197–204.*  
541 <https://doi.org/10.5220/0004799801970204>.
- 542 [31] A. Salarian, H. Russmann, F.J.G. Vingerhoets, C. Dehollain, Y. Blanc, P.R. Burkhard, K. Aminian,  
543 Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring, *IEEE*  
544 *Trans. Biomed. Eng.* 51 (2004) 1434–1443. <https://doi.org/10.1109/TBME.2004.827933>.
- 545 [32] B.R. Greene, D. McGrath, R. O'Neill, K.J. O'Donovan, A. Burns, B. Caulfield, An adaptive gyroscope-  
546 based algorithm for temporal gait analysis, *Med. Biol. Eng. Comput.* 48 (2010) 1251–1260.  
547 <https://doi.org/10.1007/s11517-010-0692-0>.
- 548 [33] P. Catalfamo, S. Ghousayni, D. Ewins, Gait Event Detection on Level Ground and Incline Walking  
549 Using a Rate Gyroscope, *Sensors.* 10 (2010) 5683–5702. <https://doi.org/10.3390/s100605683>.

- 550 [34] K. Aminian, B. Najafi, C. Büla, P.-F. Leyvraz, Ph. Robert, Spatio-temporal parameters of gait measured  
551 by an ambulatory system using miniature gyroscopes, *J. Biomech.* 35 (2002) 689–699.  
552 [https://doi.org/10.1016/S0021-9290\(02\)00008-8](https://doi.org/10.1016/S0021-9290(02)00008-8).
- 553 [35] J.M. Jasiewicz, J.H.J. Allum, J.W. Middleton, A. Barriskill, P. Condie, B. Purcell, R.C.T. Li, Gait event  
554 detection using linear accelerometers or angular velocity transducers in able-bodied and spinal-cord  
555 injured individuals, *Gait Posture.* 24 (2006) 502–509. <https://doi.org/10.1016/j.gaitpost.2005.12.017>.
- 556 [36] A. Ferrari, P. Ginis, M. Hardegger, F. Casamassima, L. Rocchi, L. Chiari, A Mobile Kalman-Filter  
557 Based Solution for the Real-Time Estimation of Spatio-Temporal Gait Parameters, *IEEE Trans. Neural*  
558 *Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.* 24 (2016) 764–773.  
559 <https://doi.org/10.1109/TNSRE.2015.2457511>.
- 560 [37] A.M. Sabatini, C. Martelloni, S. Scapellato, F. Cavallo, Assessment of Walking Features From Foot  
561 Inertial Sensing, *IEEE Trans. Biomed. Eng.* 52 (2005) 486–494.  
562 <https://doi.org/10.1109/TBME.2004.840727>.
- 563 [38] B. Mariani, S. Rochat, C.J. Büla, K. Aminian, Heel and Toe Clearance Estimation for Gait Analysis  
564 Using Wireless Inertial Sensors, *IEEE Trans. Biomed. Eng.* 59 (2012) 3162–3168.  
565 <https://doi.org/10.1109/TBME.2012.2216263>.
- 566 [39] M. Schwartz, A. Rozumalski, J. Trost, The effect of walking speed on the gait of typically developing  
567 children. - PubMed - NCBI, *Eff. Walk. Speed Gait Typically Dev. Child.* 41 (2008) 1639–50.  
568 <https://doi.org/10.1016/j.jbiomech.2008.03.015>.
- 569 [40] J.A. Zeni, J.G. Richards, J.S. Higginson, Two simple methods for determining gait events during  
570 treadmill and overground walking using kinematic data, *Gait Posture.* 27 (2008) 710–714.  
571 <https://doi.org/10.1016/j.gaitpost.2007.07.007>.
- 572 [41] P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, *Psychol. Bull.* 86  
573 (1979) 420–428. <https://doi.org/10.1037//0033-2909.86.2.420>.
- 574 [42] D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models using lme4,  
575 *ArXiv14065823 Stat.* (2014). <http://arxiv.org/abs/1406.5823> (accessed July 2, 2018).
- 576 [43] A.L. Hof, Scaling gait data to body size, *Gait Posture.* 4 (1996) 222–223. [https://doi.org/10.1016/0966-6362\(95\)01057-2](https://doi.org/10.1016/0966-6362(95)01057-2).
- 578 [44] S. Khandelwal, N. Wickström, Identification of Gait Events using Expert Knowledge and Continuous  
579 Wavelet Transform Analysis:, *SciTePress.* (2014) 197–204.
- 580 [45] L. Brognara, P. Palumbo, B. Grimm, L. Palmerini, Assessing Gait in Parkinson’s Disease Using  
581 Wearable Motion Sensors: A Systematic Review, *Diseases.* 7 (2019) 18.  
582 <https://doi.org/10.3390/diseases7010018>.

- 583 [46] Y. Moon, R.S. McGinnis, K. Seagers, R.W. Motl, N. Sheth, J.A.W. Jr, R. Ghaffari, J.J. Sosnoff,  
584 Monitoring gait in multiple sclerosis with novel wearable motion sensors, PLOS ONE. 12 (2017)  
585 e0171346. <https://doi.org/10.1371/journal.pone.0171346>.
- 586 [47] C. Tunca, N. Pehlivan, N. Ak, B. Arnrich, G. Salur, C. Ersoy, Inertial Sensor-Based Robust Gait  
587 Analysis in Non-Hospital Settings for Neurological Disorders, Sensors. 17 (2017).  
588 <https://doi.org/10.3390/s17040825>.
- 589 [48] F.A. Storm, C.J. Buckley, C. Mazzà, Gait event detection in laboratory and real life settings: Accuracy  
590 of ankle and waist sensor based methods, Gait Posture. 50 (2016) 42–46.  
591 <https://doi.org/10.1016/j.gaitpost.2016.08.012>.  
592



**Abstract**

*Background and Objectives:* Walking in water is used for rehabilitation in different pathological conditions. For the characterization of gait alterations related to pathology, gait timing assessment is of primary importance. With the widespread use of inertial sensors, several algorithms have been proposed for gait timing estimation (i.e. gait events and temporal parameters) out of the water, while an assessment of their performance for walking in water is still missing. The purpose of the present study was to assess the performance in the temporal segmentation for gait in water of 17 algorithms proposed in the literature. *Methods:* Ten healthy volunteers mounting 5 tri-axial inertial sensors (trunk, shanks and feet) walked on dry land and in water. Seventeen different algorithms were implemented and classified based on: 1) sensor position, 2) target variable, and 3) computational approach. Gait events identified from synchronized video recordings were assumed as reference. Temporal parameters were calculated from gait events. Algorithm performance was analysed in terms of sensitivity, positive predictive value, accuracy, and repeatability. *Results:* For walking in water, all Trunk-based algorithms provided a sensitivity lower than 81% and a positive predictive value lower than 94%, as well as acceleration-based algorithms, independently from sensor location, with the exception of two Shank-based ones. Drop in algorithm sensitivity and positive predictive value was associated to significant differences in the stride pattern of the specific analysed variables during walking in water as compared to walking on dry land, as shown by the intraclass correlation coefficient. When using Shank- or Foot-based algorithms, gait events resulted delayed, but the delay was compensated in the estimate of Stride and Step time; a general underestimation of Stance- and overestimation of Swing-time was observed, with minor exceptions. *Conclusion:* Sensor position, target variable and computational approach determined different error distributions for different gait events and temporal parameters for walking in water. This work supports an evidence-based selection of the most appropriate algorithm for gait timing estimation for walking in water as related to the specific application, and provides relevant information for the design of new algorithms for the specific motor task.

*Index Terms*— algorithm, event detection, gait events, gait timing, IMUs, temporal parameters, walking, water, wearable inertial sensors

66  
67  
68  
69  
70  
71

72 **Nomenclature**

73 WW: Walking in water

74 WDL: Walking on dry land

75 GE: Gait Event

76 GTP: Gait Temporal Parameter

77 FC: Foot Contact

78 FO: Foot Off

79 IMU: Inertial Measurement Unit

80 FIR: Finite Impulse Response

81 IIR: Infinite Impulse Response

82 WT: Wavelet Transform

83 GRF: Ground Reaction Force

84 E: Error

85 PPV: positive predictive values

86 Med: Median

87 Dmed: Dispersion around median value

88 ICC: Intraclass correlation coefficient

89 CV: Coefficient of Variation

90

91

## 92 **Introduction**

93 The water environment plays a relevant role in rehabilitation programs [1]. During aquatic therapy, the  
94 buoyancy, the drag force, and the pressure exerted by water reduce the gravitational load on joints, resist  
95 motion, and increase proprioception [2–5]. In recent years, Walking in Water (WW) has become extensively  
96 used for people with specific gait deviations [1] and represents one of the fundamental motor tasks performed  
97 during aquatic therapy [5]. It is recommended for developing and maintaining cardiorespiratory and muscular  
98 fitness [6], can be adopted to speed up recovery from minor orthopaedic injuries [7], offers clear advantages  
99 over the land-based equivalent for populations with high risk of fall such as older adults and neurological  
100 patients [8], and can be practiced by individuals without swimming skills [9].

101 Assessment of motor performance during WW, based on quantitative motion analysis, would support the  
102 understanding of water-induced biomechanical modifications and the design and/or monitoring of WW based  
103 rehabilitation.

104 In quantitative motion analysis, the assessment of temporal parameters is of primary importance, implying  
105 the correct identification of gait events (GEs), i.e. Foot Contact (FC) and Foot Off (FO) [10,11]. Various  
106 technologies (i.e. force platforms, instrumented mats, footswitches) can be and were used for identifying GE  
107 when Walking on Dry Land (WDL), but no validated approach is available for gait timing identification for  
108 WW, as highlighted by Matsumoto et al., 2008 [12]. Several studies exploited a camera-based approaches  
109 for WW [1,5,6,12], although its drawbacks: i) the limited field of view allowing to analyse only one/two  
110 consecutive steps; ii) the time-consuming set-up and post-processing [1,6].

111 A recent review [13] reports for WW slower speed, lower cadence [1,14–20] and shorter step length  
112 [1,20,21] as compared to WDL, with similar duration of stance phase relative to gait cycle (62% for WDL  
113 and 60% for WW when the subjects walked at their preferred speeds) [20]. Still the characterization of WW  
114 is limited by the lack of a standardised approach for the analysis.

115 In recent years, the use of inertial wearable sensors (IMUs) has become extensive in the analysis of WDL,  
116 and a large number of algorithms for gait segmentation [10,22] were proposed for the identification of GEs  
117 for healthy subjects. A recent comprehensive analysis [10] showed that specific performance is significantly  
118 affected by sensor placement, analysed variable, and computational approach even in ideal conditions. To  
119 authors' knowledge, no novel algorithm was specifically designed and proposed for the automatic  
120 identification of GEs for WW and a specific accurate characterization of the motion pattern for WW is still  
121 missing, not supporting the design of new algorithms yet. On the other hand, still considering the functional  
122 and mechanical differences between WDL and WW [1], some of the basic characteristics, considered for the  
123 design of the available algorithms, are common to both walking conditions, such as body forward progression  
124 and lower limb alternate pendulation and ground contact. Therefore, the evaluation of the performance of

125 these available algorithms for the analysis of WW is deemed suitable to investigate their ability to identify  
 126 features still common to both locomotor conditions, providing evidence to support the selection of the most  
 127 performing approach among the available ones, and/or provide criteria for the design of new ones, if  
 128 necessary.

129 The aim of the present work was to analyse the performance during WW of 17 different algorithms  
 130 designed for GE estimation in WDL, taking into account the biomechanical changes induced in the gait  
 131 pattern by the different environmental conditions, i.e. water. To achieve this goal, the algorithms previously  
 132 identified through a systematic review [10] underwent a systematic analysis, relating the performance (pros  
 133 and cons) to the intrinsic characteristics of the algorithms (i.e. sensor position and signal characteristics).

## 134 **Methods**

### 135 *Analysed algorithms*

136 17 algorithms for GE and gait temporal parameters (GTP) estimation, previously selected and analysed by  
 137 Pacini et al. [20] for WDL, were here analysed for WW. Following the same approach [20], the selected  
 138 algorithms were classified based on:

- 139 i) IMU position (i.e. trunk, shanks, feet);
- 140 ii) Target variable (i.e. acceleration, angular velocity);
- 141 iii) Computational approach: ‘peak identification’ and ‘zero crossing’, on raw or filtered target variable  
 142 i.e. finite impulse response (FIR), infinite impulse response (IIR) and wavelet transform (WT) filtering.  
 143 ‘Peak identification’ aims to identify specific peaks on the target variable, corresponding to specific  
 144 temporal events: local maxima or minima of the vertical, antero-posterior or 3D components for  
 145 acceleration-based algorithms; local minima of the medio-lateral component for angular velocity-based  
 146 algorithms. ‘Zero crossing’ aims to identify the instants of sign change in the target variable,  
 147 corresponding to specific temporal events: in the antero-posterior component for acceleration-based  
 148 algorithms; in the medio-lateral component for angular velocity-based algorithms.

149 The selected algorithms have the following features:

- 150 - 6 were trunk-based, analysing antero-posterior [23–26], vertical [27] or 3D [28] components of  
 151 acceleration. Among these algorithms, 3 adopted ‘peak identification’, of which Buganè et al. with IIR  
 152 [23], Lee et al. with FIR [24] and McCamley et al. with WT filtering [27], and 3 used ‘zero crossing’  
 153 approach, of which Shin et al. with raw signal [28], González et al. with FIR filtering [25] and Zijlstra et  
 154 al. with IIR filtering [26];
- 155 - 7 were shank-based, of which 3 analysing antero-posterior [11] or 3D [29,30] components of  
 156 acceleration with ‘peak identification’, Trojanello et al. with raw signal [11], Lee et al. with IIR [29]

157 and Khandelwal et al. with WT filtering [30], and 4 analysing the medio-lateral component of angular  
158 velocity with ‘peak identification’, Salarian et al. [31] and Greene et al. [32] with raw signal, Catalfamo  
159 et al. [33] with IIR and Aminian et al. [34] with WT filtering;

160 - 4 were foot-based, of which Jasiewicz et al. analysing anterior-posterior acceleration with ‘peak  
161 identification’ of raw signal [35], and 3 analysing medio-lateral angular velocity, 2 with ‘peak  
162 identification’, Ferrari et al. [36] adopting raw signal, Sabatini et al. [37] with IIR filtering and Mariani  
163 et al. [38] with ‘zero crossing’ of IIR filtered signal.

## 165 *Experimental analysis*

### 166 *Participants.*

167 Ten young adult healthy participants (5 females, 5 males; age  $26.2 \pm 3.3$  years; height  $1.71 \pm 0.07$  m; weight  
168  $65.4 \pm 8.6$  Kg) were recruited in the study. All participants were physically active and self-reported no  
169 musculoskeletal or neurological disorder. The Bioethics Committee of the University of Bologna approved  
170 the study on 13/07/2018 with protocol number 99412, and informed consent was signed by all participants.

### 172 *Data acquisition.*

173 Each participant walked 5 times back and forth along a 10 m straight pathway at self-selected speed in 2  
174 conditions: i) WDL; ii) WW at 1.2m depth with water temperature of  $28^{\circ}\text{C}$  wearing water shoes and keeping  
175 the arms on the water surface. No device (e.g. metronome or timer) was used to control the walking speed,  
176 not to interfere with the natural walking pattern [39]. Before the WW analysed trial, participants performed  
177 an acclimatisation trial.

178  
179 Five tri-axial IMUs (Cometa, Italy, sf=285 Hz; accelerometer: sensitivity 156,3 mV/g, range  $\pm 8g$ ;  
180 gyroscope: sensitivity 1,3 mV/g, range  $\pm 1000^{\circ}/\text{s}$ ) were attached to the trunk (at L5 level), shanks (about five  
181 centimetres above lateral malleolus), and feet (on the dorsal aspect of each shoe). The walking tasks were  
182 also filmed using a video camera (Hero4, GoPro, USA, sf=240Hz, 848x480 pixels resolution) for GE  
183 detection reference. IMU led flashing was video-recorded and used for time-synchronization of IMU and  
184 video recording.

185 During WDL trails, ground reaction forces (GRF) were also recorded using two force platforms (Kistler,  
186 Winterthur, Switzerland, sf=1000Hz) mounted half-way along the pathway. A trigger signal was generated  
187 by IMU system at the beginning of each trial for synchronization with the force platforms.

### 188 *Data analysis.*

189 Average gait speed was calculated from reference video during WDL and WW as the ratio between straight  
190 walked distance and time.

191 For WDL, 20  $FC_{GRF}$  and 20  $FO_{GRF}$  were automatically identified applying a 20 N threshold to the vertical  
192 component of the GRF recordings [40], and  $FC_{GoPro}$  and  $FO_{GoPro}$  were visually identified from the video  
193 recordings for each participant and trial. After verifying that the number of events estimated by the video  
194 recordings were the same provided by GRF, the measurement error of video-based assessment was estimated  
195 considering the GRF as reference:

$$196 \quad E_{ref} = \max(\text{abs}(GE_{GoPro} - GE_{GRF})) \quad (1)$$

197  
198 For WW,  $FC_{GoPro}$  and  $FO_{GoPro}$  were visually identified as for WDL, and  $FC_{IMU}$  and  $FO_{IMU}$  were estimated  
199 from IMU measurements using the selected 17 algorithms [10], implemented in MATLAB (MathWorks  
200 2017a, USA). GTPs were calculated from GE.

201 For each algorithm considering the video as reference, the sensitivity and the positive predictive value  
202 (PPV) in GE identification during WW was calculated [22] respectively as:

$$203 \quad \textit{Sensitivity} = 100 \times \frac{\text{Number of GEs correctly identified by algorithm}}{\text{Number of all GEs as identified from video}} \quad (2)$$

$$204 \quad \textit{PPV} = 100 \times \frac{\text{Number of GEs correctly identified by algorithm}}{\text{Number of all GEs identified by algorithm}} \quad (3)$$

205 For each target variable of the 17 selected algorithms, the intraclass correlation coefficient (ICC) (3,1) of  
206 the mean stride cycle over the whole sequence of each trial for WW and WDL was calculated according to  
207 Shrout et al. [41] to analyse the repeatability of the pattern over the trial in each condition.

208 ICC(3,1) of the mean WDL stride cycle over the WW sequence was calculated, to analyse the similarity of  
209 the pattern during WW with respect to the one during WDL, assumed as reference for the algorithm design.

210 For each algorithm and each condition, coefficient of variation (CV) of the analysed target signal was  
211 calculated as:

$$212 \quad \frac{\text{Difference between 75th and 25th}}{\text{percentile over the stride cycle}} \quad (4)$$

$$213 \quad \frac{\text{Difference between 75th and 25th}}{\text{percentile over the stride cycle}}}{(\text{Abs}(\text{median of the median value over the stride cycle}))}$$

214 Only the algorithms reporting a minimum sensitivity of 81% and a minimum PPV of 94% [22] were  
 215 considered for further analysis.

216 For each algorithm, the error was calculated for GE and GTP, considering the video as reference, as:

217

$$E_{GE} = GE_{IMU} - GE_{GoPro} \quad (5)$$

$$E_{GTP} = GTP_{IMU} - GTP_{GoPro} \quad (6)$$

218

219 *Statistical analysis.*

220 For each parameter (FC, FO, Stride Time, Step Time, Stance Time, Swing Time), a linear mixed model [42]  
 221 was applied to test the dependency of error values on each implementation criterion. The models were  
 222 specified as follows:

223

$$y_{ti} = \beta_0 + \beta_1 \cdot x_{ti} + u_{0i} + \varepsilon_{0ti} \quad (7)$$

224

225 where  $\beta_0$  represent the fixed intercepts,  $u_{0i}$  represent the random effects associated to the  $i^{\text{th}}$  subject, the  
 226 subscripts  $ti$  indicate values of variables measured in the  $t^{\text{th}}$  step clustered in the  $i^{\text{th}}$  subject,  $y$  represents the  
 227 outcome variable,  $x$  represents the independent variable, and  $\varepsilon_0$  is the random error component. The linear  
 228 mixed model was fit using the lme4 package in R software (R-Core Team 2017, Austria, version 3.4.3) [42],  
 229 with the significance level set at 0.05. First, the statistical analysis was performed to investigate the influence  
 230 of IMU position and target variable, alone. Then, the influence of analysed variable and computational  
 231 approach were investigated separately for each IMU position.

232 Median value (Med) of the error was calculated to characterize accuracy, and the Dispersion around Med  
 233 (Dmed, 75th percentile – 25th percentile values of the error) to characterize repeatability.

234

235

236

237

238

239

240

241 **Results**

242 Gait speed normalised according to Hof [43] resulted  $0,33\pm 0,06$  for WDL and  $0,17\pm 0,08$  for WW.

243 Maximum measurement error of video- versus GRF-reference resulted 0.05 s for both FC and FO.

244

245 During WW, 32 FCs and FOs were identified and analysed for each participant, for a total of 320 FCs and  
246 FOs.

247 No algorithm exploiting a sensor on the trunk passed the 81% sensitivity and 94% PPV, as well as no  
248 acceleration-based algorithm, independently from sensor placement, with the exception of two, both  
249 exploiting a sensor positioned on the shank and a peak identification approach: i) Khandelwal et al. [44] on  
250 WT signal; ii) Lee et al. [29] on a IIR transformed signal.

251 After the sensitivity and PPV analysis, only algorithms exploiting a peak identification approach for shank  
252 positioned, and only angular velocity based for foot positioned sensor underwent further error analysis.

253 Algorithms that passed the 81% sensitivity and 94% PPV criterion showed an ICC for WW above or equal  
254 to 0.70, with an ICC of WDL stride cycle applied to WW ranging from 0.31 to 0.61. On the other hand,  
255 algorithms that did not pass the 81% sensitivity and 94% PPV showed an ICC for WW below 0.60, with an  
256 ICC of WDL stride cycle applied to WW ranging from 0.10 to 0.34, with the only exception of Trojanello et  
257 al. [11], showing an ICC for WW equal to 0.79 but an ICC of WDL stride cycle applied to WW of only 0.11.

258 No trend was observed for the CV of the variable analysed by each algorithm.

259 Sensitivity, PPV and ICC analysis results are summarised in Table 1.

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

**Table 1:** Sensitivity of algorithms in FC and FO identification; ICC, CV with Minimum, Med (in bold), Maximum, and plots with Med, 25th and 75th percentile of the different target signals in relation to the implementation criteria defined on the normalized gait cycle and referred to WDL, WW, WDL and WW (thicker lines for WDL plots).

Algorithm	Sensitivity (%)	PPV (%)	WDL			WW			WDL and WW	
	FC FO	FC FO	ICC	CV	Plot	ICC	CV	Plot	ICC	Plot
Bugané et al., 2012	< 81	< 94	0.79	1.3 <b>2.7</b> 11.3		0.33	3.3 <b>4.2</b> 6.5		0.24	
Lee et al., 2009			0.79	1.3 <b>2.7</b> 11.1		0.33	3.3 <b>4.2</b> 6.5		0.23	
McCamley et al., 2012			0.90	1.8 <b>2.9</b> 7.9		0.37	164.9 <b>229.1</b> 362.9		0.10	
González et al., 2010			0.82	1.5 <b>2.5</b> 9.0		0.34	3.1 <b>4.0</b> 6.3		0.23	
Shin et al., 2011			0.76	732.0 <b>1063.6</b> 3042.3		0.19	4.5 <b>6.7</b> 13.7		0.15	
Zijlstra et al., 2003			0.93	2.9 <b>5.3</b> 8.3		0.44	2.2 <b>3.0</b> 4.3		0.34	
Lee et al., 2010	98 99	98 100	0.91	0.8 <b>2.2</b> 6.8		0.70	30.6 <b>155.0</b> 490.0		0.31	
Trojaniello et al., 2014	< 81	< 94	0.75	1.7 <b>5.4</b> 35.3		0.79	4.1 <b>6.1</b> 13.8		0.11	
Khandelwal et al., 2014	97 99	97 98	0.91	2.1 <b>14.7</b> 31.2		0.76	4.5 <b>18.0</b> 42.0		0.56	
Catalfamo et al., 2010	98 100	98 98	0.97	0.3 <b>1.1</b> 3.5		0.86	0.5 <b>1.5</b> 6.0		0.61	
Greene et al., 2010	98 100	98 100	0.96	0.3 <b>1.1</b> 3.5		0.86	0.5 <b>1.5</b> 6.0		0.61	
Salarian et al., 2004	100 100	99 98	0.96	0.3 <b>1.1</b> 3.5		0.86	0.5 <b>1.5</b> 6.0		0.61	
Aminian et al., 2002	98 100	94 96	0.96	0.3 <b>1.1</b> 3.5		0.86	0.5 <b>1.4</b> 5.9		0.61	
Jasiewicz et al., 2006	< 81	< 94	0.76	0.5 <b>1.6</b> 15.1		0.56	5.7 <b>15.3</b> 38.8		0.19	
Sabatini et al., 2005	99 100	98 99	0.95	1.2 <b>8.7</b> 36.4		0.82	1.5 <b>4.6</b> 12.3		0.56	
Ferrari et al., 2016	93 95	94 95	0.94	1.3 <b>9.1</b> 41.4		0.82	1.6 <b>4.8</b> 13.6		0.54	
Mariani et al., 2013	99 100	99 100	0.94	1.3 <b>8.9</b> 38.6		0.82	1.4 <b>4.4</b> 11.7		0.55	

279 In general, FC and FO estimates resulted delayed, with this delay being compensated in Stride and Step  
280 time estimates, while Stance Time was under- and Swing Time over-estimated.

281 In more detail, considering the different implementation characteristics:

### 282 283 *IMU Position*

284 For GE, significant differences were found between Shank- and Foot-based algorithms: Shank-based  
285 algorithms resulted equally accurate and repeatable in FC estimation (Med/Dmed: 0.25/0.15 s for both Shank  
286 and Foot), but with different values of 25<sup>th</sup> and 75<sup>th</sup> percentiles (0.20s and 0.35s for Shank-based algorithms  
287 and of 0.15s and 0.30s for Foot-based algorithms). Moreover, Shank- based algorithms were equally accurate  
288 and more repeatable in FO estimation (Med/Dmed: 0.20/0.20 s and 0.20/0.30 s, for Shank and Foot,  
289 respectively) than Foot-based ones.

290 For GTP, no significant difference was found for Stride and Step time estimates, while Shank-based  
291 algorithms resulted equally accurate but more repeatable in Stance Time (Med/Dmed: -0.05/0.25 s and -  
292 0.05/0.35 s, for Shank and Foot, respectively) and less accurate but more repeatable in Swing Time estimation  
293 (Med/Dmed: 0.05/0.25 s and 0.00/0.35 s, for Shank and Foot, respectively) than Foot-based ones.

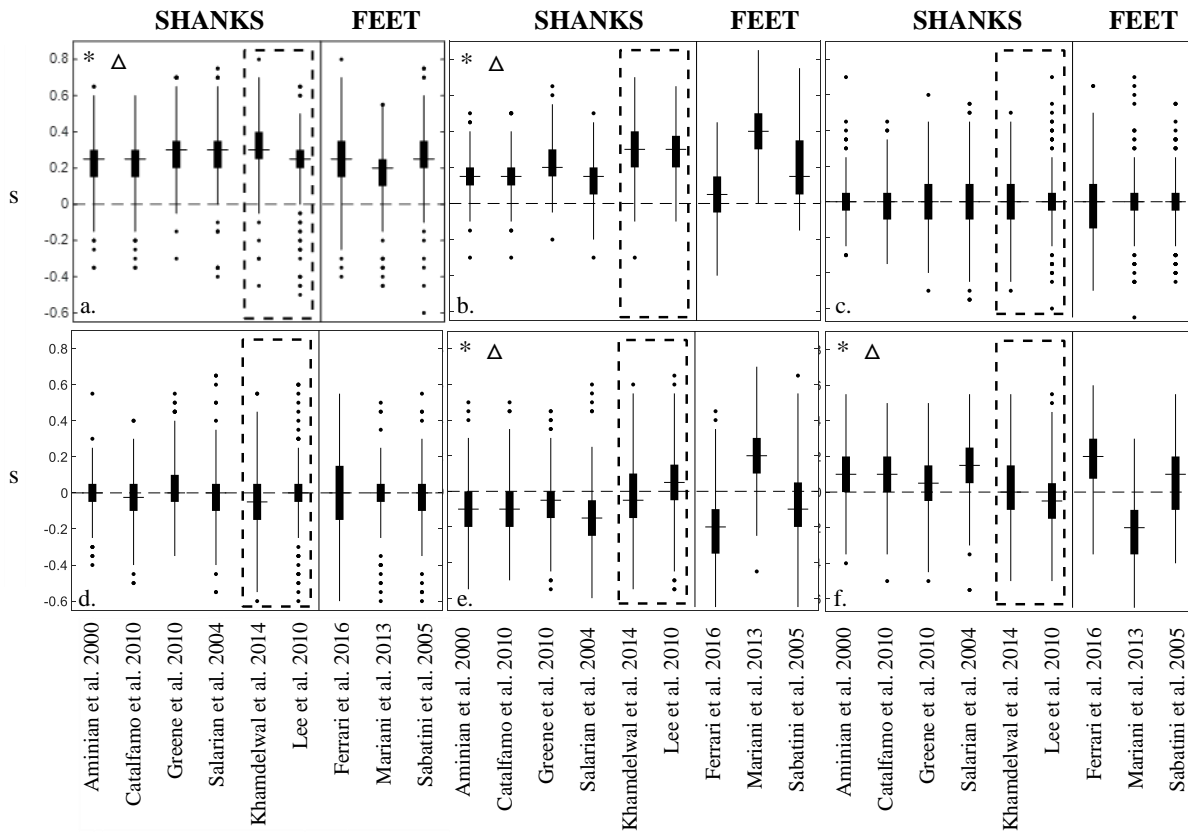
### 294 295 *Target variable*

296 For GE, Acceleration based algorithms resulted equally accurate but more repeatable for FC (Med/Dmed:  
297 0.25/0.15 s and 0.25/0.20 s, for Acceleration and Angular velocity, respectively), and less accurate and  
298 equally repeatable in FO estimation (Med/Dmed: 0.30/0.20 s and 0.15/0.20 s, for Acceleration and Angular  
299 velocity, respectively) than Angular velocity-based ones.

300 For GTP, no significant difference was found in Stride time and Step time estimation, while Acceleration-  
301 based algorithms resulted more accurate and equally repeatable in Stance time estimation (Med/Dmed:  
302 0.00/0.25 s and -0.10/0.25 s, for Acceleration and Angular velocity, respectively), and Swing time estimation  
303 (Med/Dmed: -0.00/0.25 s and 0.10/0.25 s, for Acceleration and Angular velocity, respectively) than Angular  
304 velocity-based ones.

305  
306 Error characteristics for FC (a), FO (b), Stride- (c), Step- (d), Stance- (e) and Swing time (f) as related to  
307 IMU position and target variable are schematically depicted in Figure 1, while error distribution is reported  
308 in Table 2.

309  
310  
311  
312  
313  
314



**Figure 1:** Box plot (minimum, 25th percentile, Med, 75th percentile, maximum values) for FC (a), FO (b), Stride- (c), Step- (d), Stance- (e), and Swing time (f) estimation errors as related to IMU position and target variable. Acceleration-based algorithms are framed in dashes (-). Significant differences were highlighted with asterisk (\*) for IMU position and with triangle (Δ) for target variable.

315  
316  
317  
318  
319  
320  
321

**Table 2:** Results of statistical analysis for IMU positioning and target variable for WW: 25th quartile, Med (in bold) and 75th quartile of estimation error for FC, FO, stride time, step time, stance time and swing time (\* p<0.05).

Parameter	Estimation of errors: IMU position (s)						Level of significance	Estimation of errors: target variable (s)						Level of significance
	Shanks			Feet				Acceleration			Angular velocity			
FC	0.20	<b>0.25</b>	0.35	0.15	<b>0.25</b>	0.30	*	0.20	<b>0.25</b>	0.35	0.15	<b>0.25</b>	0.35	*
FO	0.10	<b>0.20</b>	0.30	0.05	<b>0.20</b>	0.35	*	0.20	<b>0.30</b>	0.40	0.10	<b>0.15</b>	0.30	*
Stride Time	-0.10	<b>0.00</b>	0.05	-0.10	<b>0.00</b>	0.05		-0.10	<b>0.00</b>	0.05	-0.10	<b>0.00</b>	0.05	
Step Time	-0.10	<b>0.00</b>	0.05	-0.10	<b>0.00</b>	0.10		-0.10	<b>0.00</b>	0.05	-0.10	<b>0.00</b>	0.10	
Stance Time	-0.20	<b>-0.05</b>	0.05	-0.20	<b>-0.05</b>	0.15	*	-0.10	<b>0.00</b>	0.15	-0.20	<b>-0.10</b>	0.05	*
Swing Time	-0.05	<b>0.05</b>	0.20	-0.15	<b>0.00</b>	0.20	*	-0.15	<b>0.00</b>	0.10	-0.05	<b>0.10</b>	0.20	*

322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334

335 *Computational approach*

336 Shank-based algorithms exploited only ‘peak identification’ approach. Raw and WT filtering performed  
337 equally (Med/DL: 0.30/0.15 s), but less accurate and repeatable than IIR (Med/DL: 0.25/0.10 s) in FC  
338 estimation, while IIR and WT performed equally (Med/DL: 0.20/0.20 s), but equally accurate and less  
339 repeatable than Raw (Med/DL: 0.20/0.15 s) in FO estimation. Considering GTP, no significant difference  
340 was found among Raw, IIR and WT for Stride and Step time estimation in terms of accuracy, with decreasing  
341 (Med/DL: 0.00/0.20 s, 0.00/0.15 s, and 0.00/0.10 s, for Raw, IIR and WT, respectively) and equal  
342 repeatability (Med/DL: 0.00/0.15 s, for all) for Stride and Step time estimation, respectively; on the other  
343 hand, Raw, WT and IIR showed and increasing accuracy, respectively, in Stance and Swing time estimation,  
344 but Raw and WT resulted more repeatable than IIR (Med/DL: -0.10/0.20 s, 0.00/0.25 s, and -0.05/0.20 s, for  
345 Raw, IIR and WT, respectively for Stance Time; Med/DL: 0.10/0.20 s, 0.00/0.25 s, and 0.05/0.20 s, for Raw,  
346 IIR and WT, respectively for Swing Time).

347  
348 Foot-based algorithms exploited both ‘peak identification’ (both Raw and IIR for filtering) and ‘zero  
349 crossing’ (only IIR for filtering) approaches, as applied to angular velocity. ‘Zero crossing’ with IIR resulted  
350 more accurate in FC, but less accurate in FO estimate than ‘peak identification’, where IIR resulted equally  
351 accurate in FC estimate and less accurate in FO estimate than Raw, with comparable repeatability.

352 No significant difference was found in Stride and Step time estimation for both computational approach  
353 and filtering technique. On the other hand, ‘zero crossing’ with IIR: i) delayed and resulted more repeatable,  
354 while ‘peak identification’ anticipated Stance time; ii) anticipated and resulted less accurate, while ‘peak  
355 identification’ delayed’ Swing time. For both GTP IIR resulted more accurate than Raw.

356  
357 Error values and statistical analysis for computational approach are reported in Table 3.

358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373

**Table 3:** Results of statistical analysis for computational approach during WW: 25th quartile, Med (in bold) and 75th quartile of estimation error for FC, FO, stride time, step time, stance time and swing time (\* p<0.05).

IMU position	Parameter	Level of significance 'peak identification' vs 'zero crossing'	Estimation of errors: Filtering (s)						
			Level of significance		Filtering within 'zero crossing'	Level of significance for filtering within 'zero crossing'			
			Filtering within 'peak identification'	Level of significance for filtering within 'peak identification'				Level of significance for filtering within 'zero crossing'	
Shank	FC	-	Raw	0.20	<b>0.30</b>	0.35	IIR - Raw *	-	
			IIR	0.20	<b>0.25</b>	0.30	IIR - WT *		
			WT	0.20	<b>0.30</b>	0.35	Raw - WT		
	FO		Raw	0.10	<b>0.20</b>	0.25	IIR - Raw *		
			IIR:	0.10	<b>0.20</b>	0.30	IIR - WT		
			WT	0.10	<b>0.20</b>	0.30	Raw - WT *		
	Stride Time		Raw	-0.10	<b>0.00</b>	0.10	IIR - Raw		
			IIR	-0.10	<b>0.00</b>	0.05	IIR - WT		
			WT	-0.05	<b>0.00</b>	0.05	Raw - WT		
	Step Time		Raw	-0.05	<b>0.00</b>	0.10	IIR - Raw *		
			IIR	-0.10	<b>0.00</b>	0.05	IIR - WT		
			WT	-0.10	<b>0.00</b>	0.05	Raw - WT *		
	Stance Time		Raw	-0.20	<b>-0.10</b>	0.00	IIR - Raw *		
			IIR	-0.15	<b>0.00</b>	0.10	IIR - WT *		
			WT	-0.15	<b>-0.05</b>	0.05	Raw - WT *		
	Swing Time		Raw	0.00	<b>0.10</b>	0.20	IIR - Raw *		
			IIR	-0.10	<b>0.00</b>	0.15	IIR - WT *		
			WT	-0.05	<b>0.05</b>	0.15	Raw - WT *		
Foot	FC	*	Raw	0.15	<b>0.25</b>	0.35	IIR - Raw	-	
			IIR	0.20	<b>0.25</b>	0.35	IIR - Raw		IIR 0.10 <b>0.20</b> 0.25
	FO		Raw	-0.05	<b>0.05</b>	0.15	IIR - Raw *		IIR 0.30 <b>0.40</b> 0.50
			IIR	0.05	<b>0.15</b>	0.35	IIR - Raw *		
	Stride Time		Raw	-0.15	<b>0.00</b>	0.10	IIR - Raw		IIR -0.05 <b>0.00</b> 0.05
			IIR	-0.05	<b>0.00</b>	0.05	IIR - Raw		
	Step Time		Raw	-0.15	<b>0.00</b>	0.15	IIR - Raw		IIR -0.05 <b>0.00</b> 0.05
			IIR	-0.10	<b>0.00</b>	0.05	IIR - Raw		
	Stance Time		Raw	-0.35	<b>-0.20</b>	-0.10	IIR - Raw *		IIR 0.10 <b>0.20</b> 0.30
			IIR	-0.20	<b>-0.10</b>	0.05	IIR - Raw *		
	Swing Time		Raw	0.10	<b>0.20</b>	0.30	IIR - Raw *		IIR -0.35 <b>-0.20</b> -0.10
			IIR	-0.10	<b>0.10</b>	0.20	IIR - Raw *		

374  
375

## 376 Discussion

377 The purpose of the present study was to assess the performance of 17 gait segmentation algorithms  
 378 originally implemented for WDL [10] when applied to WW. The implementation of these algorithms is based  
 379 on the identification of specific features that can be identified in the gait pattern of healthy subjects during  
 380 WDL. The mechanical characteristics of the water environment during WW can alter these patterns [1],  
 381 similarly to pathological and other perturbed conditions, for which WDL algorithms are still applied [45–  
 382 48]. Since no algorithm was proposed for WW temporal segmentation yet, this preliminary work was design  
 383 to understand and quantify how the alterations of the gait pattern associated to the presence of the water affect  
 384 the performance of available algorithms.

385

386 In accordance with previous studies using video analysis [16,19], IMUs [5] and force platform [13], in the  
 387 present study mean speed resulted decreased of 50% in WW as compared to WDL, cadence was lower (0.36

388  $s^{-1}$  for WW vs for  $0.86 s^{-1}$ WDL), similarly to what reported by Heywood et al.[13], and stance relative  
389 duration slightly decreased during WW compared to WDL (Mean±Standard Deviation of  $58.5\% \pm 2.9\%$  for  
390 WW and of  $62.2\% \pm 1.8\%$  for WDL).

391 As expected, the present work highlighted a significant reduction in the sensitivity and PPV of the analysed  
392 algorithms for WW.

393 Based on the values of the ICC, sensitivity and PPV still resulted over the 81% and 94% thresholds,  
394 respectively, when the gait pattern of the target variable was sufficiently repeatable (ICC for WW above or  
395 equal to 0.70), although differing from the WDL pattern (ICC of WDL over WW from 0.31 to 0.61).

396 On the other hand, sensitivity and PPV were below threshold when WW pattern was not repeatable enough  
397 (ICC for WW below 0.60), and when it differed too much from the reference WDL one, like for Trojanello  
398 et al. [11], showing a repeatable WW pattern (ICC for WW 0.79) but very different from WDL one (ICC of  
399 WDL over WW 0.11). According to the reported results, ICC analysis can provide an effective method for  
400 the objective preliminary evaluation of performance of a specific algorithm to the segmentation of an altered  
401 gait pattern.

402 In particular, no Trunk-based algorithm passed the sensitivity criterion, while Shank- and Foot-based ones  
403 provided better performance, with Shank-based ones performing slightly better in terms of repeatability. The  
404 failure of Trunk-based algorithms is associated to the aforementioned drop in WW gait pattern repeatability,  
405 but also to the disappearing of the pendulum pattern characterizing the reference WDL pattern [26], while  
406 for the shanks and feet, ICC values of WDL pattern over WW result higher due to the still occurring alternate  
407 swing of the lower limbs, associated to bi-pedal progression. Nevertheless, due to the alteration of the pattern  
408 during WW, Shank- and Foot-based algorithms did no longer perform as differently as for WDL [10].

409 Considering the influence of target variable, acceleration never passed the sensitivity and PPV criterion,  
410 with the only exception of 2 Shank-based algorithms [29,44]. Nevertheless, these two resulted slightly more  
411 repeatable in FC-, less accurate in FO identification, more accurate in Stance- and Swing-time estimates than  
412 Angular-velocity based ones.

413 For computational approach, no final conclusion can be drawn, given that for Shank-based algorithms only  
414 peak identification approach was applied to both acceleration and angular velocity, while both peak  
415 identification and zero crossing resulted applied to only angular velocity for Foot-based ones. Therefore, the  
416 analysis can hardly evaluate the performance of the computational approach independently from sensor  
417 location. Considering Shank- and Foot-based algorithms separately, comparable performance were found  
418 between raw signal and different signal filtering (i.e. IIR and WT) in the evaluation of Stride and Step time,  
419 while Stance and Swing time resulted under- or over-estimated, depending on the specific signal processing.

420 Being a preliminary study and given the population chosen for the analysis, only 10 healthy adults subjects  
421 were considered. However, a larger sample size will be acquired in future investigations to build a dataset  
422 for WW similar to the ones available for WDL.

423 The limited accuracy of the video-based reference for GE identification for WW can be considered a second  
424 limitation of the present study. On the other hand, no other reference measure is available in the water  
425 environment, and the relevant minimal detectable difference properly taken into account in the statistical  
426 analysis, supporting the reliability of the results.

## 427 **Conclusion**

429 According to the results of the present research work: i) no available Trunk-based algorithm is suitable to  
430 gait segmentation for WW, due to the disappearance of pendulum mechanics; ii) angular velocity based  
431 algorithms with sensor located on the shank and feet result more reliable in terms of sensitivity and PPV than  
432 acceleration based ones, but not in terms of accuracy and repeatability; iii) no final conclusion can be drawn  
433 regarding the computational approach, independently from sensor location and target variable. All analysed  
434 factors resulted to affect gait segmentation during WW. No proposed algorithm can be generally preferred  
435 over the others, but the reported results can support researchers in the choice of the most suitable  
436 algorithm/algorithms based on experimental condition (e.g. number/type/placement of sensors) and research  
437 question (e.g. calculation of specific parameter). Finally, these results can support the design of novel  
438 segmentation algorithms, better addressing alterations induced in the gait pattern by the water environment.

452 **References**

- 453 [1] A.M.F. Barela, S.F. Stolf, M. Duarte, Biomechanical characteristics of adults walking in shallow water  
454 and on land, *J. Electromyogr. Kinesiol.* 16 (2006) 250–256.
- 455 [2] D.A. Skelton, S.M. Dinan, Exercise for falls management: Rationale for an exercise programme aimed  
456 at reducing postural instability, *Physiother. Theory Pract.* 15 (1999) 105–120.  
457 <https://doi.org/10.1080/095939899307801>.
- 458 [3] E. Watanabe, N. Takeshima, A. Okada, K. Inomata, Comparison of water- and land-based exercise in  
459 the reduction of state anxiety among older adults, *Percept. Mot. Skills.* 91 (2000) 97–104.  
460 <https://doi.org/10.2466/pms.2000.91.1.97>.
- 461 [4] B.E. Becker, Aquatic therapy: scientific foundations and clinical rehabilitation applications, *PM R.* 1  
462 (2009) 859–872. <https://doi.org/10.1016/j.pmrj.2009.05.017>.
- 463 [5] S. Fantozzi, A. Giovanardi, D. Borra, G. Gatta, Gait Kinematic Analysis in Water Using Wearable  
464 Inertial Magnetic Sensors, *PloS One.* 10 (2015) e0138105.  
465 <https://doi.org/10.1371/journal.pone.0138105>.
- 466 [6] K. Masumoto, T. Shono, S. Takasugi, N. Hotta, K. Fujishima, Y. Iwamoto, Age-related differences in  
467 muscle activity, stride frequency and heart rate response during walking in water, *J. Electromyogr.*  
468 *Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.* 17 (2007) 596–604.  
469 <https://doi.org/10.1016/j.jelekin.2006.06.006>.
- 470 [7] K. Kaneda, H. Wakabayashi, D. Sato, T. Nomura, Lower extremity muscle activity during different  
471 types and speeds of underwater movement, *J. Physiol. Anthropol.* 26 (2007) 197–200.
- 472 [8] A.C. Severin, B.J. Burkett, M.R. McKean, M.G.L. Sayers, Biomechanical aspects of aquatic therapy:  
473 A literature review on application and methodological challenges - Kaenz, *J. Fit. Res.* 5 (2016) 49–62.
- 474 [9] S.P. Brown, L.F. Chitwood, K.R. Beason, D.R. McLemore, Deep Water Running Physiologic  
475 Responses: Gender Differences at Treadmill-Matched Walking/Running Cadences, *J. Strength Cond.*  
476 *Res.* 11 (1997) 107.
- 477 [10] G. Pacini Panebianco, M.C. Bisi, R. Stagni, S. Fantozzi, Analysis of the performance of 17 algorithms  
478 from a systematic review: Influence of sensor position, analysed variable and computational approach  
479 in gait timing estimation from IMU measurements, *Gait Posture.* 66 (2018) 76–82.  
480 <https://doi.org/10.1016/j.gaitpost.2018.08.025>.
- 481 [11] D. Trojaniello, A. Cereatti, E. Pelosin, L. Avanzino, A. Mirelman, J.M. Hausdorff, U. Della Croce,  
482 Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-  
483 mounted magneto-inertial sensors: application to elderly, hemiparetic, parkinsonian and choreic gait, *J.*  
484 *NeuroEngineering Rehabil.* 11 (2014) 152. <https://doi.org/10.1186/1743-0003-11-152>.

- 485 [12] K. Masumoto, T. Shono, N. Hotta, K. Fujishima, Muscle activation, cardiorespiratory response, and  
486 rating of perceived exertion in older subjects while walking in water and on dry land, *J. Electromyogr.*  
487 *Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.* 18 (2008) 581–590.  
488 <https://doi.org/10.1016/j.jelekin.2006.12.009>.
- 489 [13] S. Heywood, J. McClelland, P. Geigle, A. Rahmann, R. Clark, Spatiotemporal, kinematic, force and  
490 muscle activation outcomes during gait and functional exercise in water compared to on land: A  
491 systematic review, *Gait Posture.* 48 (2016) 120–130. <https://doi.org/10.1016/j.gaitpost.2016.04.033>.
- 492 [14] L.C. Carneiro, S.M. Michaelsen, H. Roesler, A. Hauptenthal, M. Hubert, E. Mallmann, Vertical reaction  
493 forces and kinematics of backward walking underwater., *Gait Posture.* 35 (2012) 225–230.  
494 <https://doi.org/10.1016/j.gaitpost.2011.09.011>.
- 495 [15] T. Miyoshi, T. Shirota, S.-I. Yamamoto, K. Nakazawa, M. Akai, Functional roles of lower-limb joint  
496 moments while walking in water, *Clin. Biomech.* 20 (2005) 194–201.  
497 <https://doi.org/10.1016/j.clinbiomech.2004.10.006>.
- 498 [16] T. Miyoshi, T. Shirota, S. Yamamoto, K. Nakazawa, M. Akai, Effect of the walking speed to the lower  
499 limb joint angular displacements, joint moments and ground reaction forces during walking in water,  
500 *Disabil. Rehabil.* 26 (2004) 724–732. <https://doi.org/10.1080/09638280410001704313>.
- 501 [17] M.I.V. Orselli, M. Duarte, Joint forces and torques when walking in shallow water, *J. Biomech.* 44  
502 (2011) 1170–1175. <https://doi.org/10.1016/j.jbiomech.2011.01.017>.
- 503 [18] A. Chevutschi, G. Lensele, D. Vaast, A. Thevenon, An electromyographic study of human gait both in  
504 water and on dry ground, *J. Physiol. Anthropol.* 26 (2007) 467–473.
- 505 [19] A. Chevutschi, M. Albery, G. Lensele, V. Pardessus, A. Thevenon, Comparison of maximal and  
506 spontaneous speeds during walking on dry land and water, *Gait Posture.* 29 (2009) 403–407.  
507 <https://doi.org/10.1016/j.gaitpost.2008.10.059>.
- 508 [20] A.M.F. Barela, M. Duarte, Biomechanical characteristics of elderly individuals walking on land and in  
509 water, *J. Electromyogr. Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.* 18 (2008) 446–454.  
510 <https://doi.org/10.1016/j.jelekin.2006.10.008>.
- 511 [21] A.M. Degani, A. Danna-dos-Santos, The Effect of Water Walking on the Lower Limb Motion of Older  
512 Adults, *Int. J. Aquat. Res. Educ.* 1 (2007). <https://doi.org/10.25035/ijare.01.03.03>.
- 513 [22] D. Trojaniello, A. Cereatti, U. Della Croce, Accuracy, sensitivity and robustness of five different  
514 methods for the estimation of gait temporal parameters using a single inertial sensor mounted on the  
515 lower trunk, *Gait Posture.* 40 (2014) 487–492. <https://doi.org/10.1016/j.gaitpost.2014.07.007>.
- 516 [23] F. Bugané, M.G. Benedetti, G. Casadio, S. Attala, F. Biagi, M. Manca, A. Leardini, Estimation of  
517 spatial-temporal gait parameters in level walking based on a single accelerometer: Validation on normal

- 518 subjects by standard gait analysis, *Comput. Methods Programs Biomed.* 108 (2012) 129–137.  
519 <https://doi.org/10.1016/j.cmpb.2012.02.003>.
- 520 [24] Hyo-Ki Lee, Sung-Jae Hwang, Sung-Pil Cho, Dong-Ryul Lee, Sung-Hyun You, Kyoung-Joung Lee,  
521 Yong-Ho Kim, Ho-Seon Choi, Novel algorithm for the hemiplegic gait evaluation using a single 3-axis  
522 accelerometer, in: 2009 Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., IEEE, Minneapolis, MN, 2009:  
523 pp. 3964–3966. <https://doi.org/10.1109/IEMBS.2009.5333650>.
- 524 [25] R.C. González, A.M. López, J. Rodríguez-Uría, D. Álvarez, J.C. Alvarez, Real-time gait event detection  
525 for normal subjects from lower trunk accelerations, *Gait Posture.* 31 (2010) 322–325.  
526 <https://doi.org/10.1016/j.gaitpost.2009.11.014>.
- 527 [26] W. Zijlstra, Assessment of spatio-temporal parameters during unconstrained walking, *Eur. J. Appl.*  
528 *Physiol.* 92 (2004) 39–44. <https://doi.org/10.1007/s00421-004-1041-5>.
- 529 [27] J. McCamley, M. Donati, E. Grimpampi, C. Mazzà, An enhanced estimate of initial contact and final  
530 contact instants of time using lower trunk inertial sensor data, *Gait Posture.* 36 (2012) 316–318.  
531 <https://doi.org/10.1016/j.gaitpost.2012.02.019>.
- 532 [28] S.H. Shin, C.G. Park, Adaptive step length estimation algorithm using optimal parameters and  
533 movement status awareness, *Med. Eng. Phys.* 33 (2011) 1064–1071.  
534 <https://doi.org/10.1016/j.medengphy.2011.04.009>.
- 535 [29] J.-A. Lee, S.-H. Cho, Y.-J. Lee, H.-K. Yang, J.-W. Lee, Portable activity monitoring system for  
536 temporal parameters of gait cycles, *J. Med. Syst.* 34 (2010) 959–966. [https://doi.org/10.1007/s10916-](https://doi.org/10.1007/s10916-009-9311-8)  
537 [009-9311-8](https://doi.org/10.1007/s10916-009-9311-8).
- 538 [30] S. Khandelwal, N. Wickström, Identification of Gait Events Using Expert Knowledge and Continuous  
539 Wavelet Transform Analysis, in: *Proc. Int. Jt. Conf. Biomed. Eng. Syst. Technol. - Vol. 4,*  
540 *SCITEPRESS - Science and Technology Publications, Lda, Portugal, 2014: pp. 197–204.*  
541 <https://doi.org/10.5220/0004799801970204>.
- 542 [31] A. Salarian, H. Russmann, F.J.G. Vingerhoets, C. Dehollain, Y. Blanc, P.R. Burkhard, K. Aminian,  
543 Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring, *IEEE*  
544 *Trans. Biomed. Eng.* 51 (2004) 1434–1443. <https://doi.org/10.1109/TBME.2004.827933>.
- 545 [32] B.R. Greene, D. McGrath, R. O'Neill, K.J. O'Donovan, A. Burns, B. Caulfield, An adaptive gyroscope-  
546 based algorithm for temporal gait analysis, *Med. Biol. Eng. Comput.* 48 (2010) 1251–1260.  
547 <https://doi.org/10.1007/s11517-010-0692-0>.
- 548 [33] P. Catalfamo, S. Ghousayni, D. Ewins, Gait Event Detection on Level Ground and Incline Walking  
549 Using a Rate Gyroscope, *Sensors.* 10 (2010) 5683–5702. <https://doi.org/10.3390/s100605683>.

- 550 [34] K. Aminian, B. Najafi, C. Büla, P.-F. Leyvraz, Ph. Robert, Spatio-temporal parameters of gait measured  
551 by an ambulatory system using miniature gyroscopes, *J. Biomech.* 35 (2002) 689–699.  
552 [https://doi.org/10.1016/S0021-9290\(02\)00008-8](https://doi.org/10.1016/S0021-9290(02)00008-8).
- 553 [35] J.M. Jasiewicz, J.H.J. Allum, J.W. Middleton, A. Barriskill, P. Condie, B. Purcell, R.C.T. Li, Gait event  
554 detection using linear accelerometers or angular velocity transducers in able-bodied and spinal-cord  
555 injured individuals, *Gait Posture.* 24 (2006) 502–509. <https://doi.org/10.1016/j.gaitpost.2005.12.017>.
- 556 [36] A. Ferrari, P. Ginis, M. Hardegger, F. Casamassima, L. Rocchi, L. Chiari, A Mobile Kalman-Filter  
557 Based Solution for the Real-Time Estimation of Spatio-Temporal Gait Parameters, *IEEE Trans. Neural*  
558 *Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.* 24 (2016) 764–773.  
559 <https://doi.org/10.1109/TNSRE.2015.2457511>.
- 560 [37] A.M. Sabatini, C. Martelloni, S. Scapellato, F. Cavallo, Assessment of Walking Features From Foot  
561 Inertial Sensing, *IEEE Trans. Biomed. Eng.* 52 (2005) 486–494.  
562 <https://doi.org/10.1109/TBME.2004.840727>.
- 563 [38] B. Mariani, S. Rochat, C.J. Büla, K. Aminian, Heel and Toe Clearance Estimation for Gait Analysis  
564 Using Wireless Inertial Sensors, *IEEE Trans. Biomed. Eng.* 59 (2012) 3162–3168.  
565 <https://doi.org/10.1109/TBME.2012.2216263>.
- 566 [39] M. Schwartz, A. Rozumalski, J. Trost, The effect of walking speed on the gait of typically developing  
567 children. - PubMed - NCBI, *Eff. Walk. Speed Gait Typically Dev. Child.* 41 (2008) 1639–50.  
568 <https://doi.org/10.1016/j.jbiomech.2008.03.015>.
- 569 [40] J.A. Zeni, J.G. Richards, J.S. Higginson, Two simple methods for determining gait events during  
570 treadmill and overground walking using kinematic data, *Gait Posture.* 27 (2008) 710–714.  
571 <https://doi.org/10.1016/j.gaitpost.2007.07.007>.
- 572 [41] P.E. Shrout, J.L. Fleiss, Intraclass correlations: uses in assessing rater reliability, *Psychol. Bull.* 86  
573 (1979) 420–428. <https://doi.org/10.1037//0033-2909.86.2.420>.
- 574 [42] D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models using lme4,  
575 *ArXiv14065823 Stat.* (2014). <http://arxiv.org/abs/1406.5823> (accessed July 2, 2018).
- 576 [43] A.L. Hof, Scaling gait data to body size, *Gait Posture.* 4 (1996) 222–223. [https://doi.org/10.1016/0966-6362\(95\)01057-2](https://doi.org/10.1016/0966-6362(95)01057-2).
- 578 [44] S. Khandelwal, N. Wickström, Identification of Gait Events using Expert Knowledge and Continuous  
579 Wavelet Transform Analysis:, *SciTePress.* (2014) 197–204.
- 580 [45] L. Brognara, P. Palumbo, B. Grimm, L. Palmerini, Assessing Gait in Parkinson’s Disease Using  
581 Wearable Motion Sensors: A Systematic Review, *Diseases.* 7 (2019) 18.  
582 <https://doi.org/10.3390/diseases7010018>.

- 583 [46] Y. Moon, R.S. McGinnis, K. Seagers, R.W. Motl, N. Sheth, J.A.W. Jr, R. Ghaffari, J.J. Sosnoff,  
584 Monitoring gait in multiple sclerosis with novel wearable motion sensors, PLOS ONE. 12 (2017)  
585 e0171346. <https://doi.org/10.1371/journal.pone.0171346>.
- 586 [47] C. Tunca, N. Pehlivan, N. Ak, B. Arnrich, G. Salur, C. Ersoy, Inertial Sensor-Based Robust Gait  
587 Analysis in Non-Hospital Settings for Neurological Disorders, Sensors. 17 (2017).  
588 <https://doi.org/10.3390/s17040825>.
- 589 [48] F.A. Storm, C.J. Buckley, C. Mazzà, Gait event detection in laboratory and real life settings: Accuracy  
590 of ankle and waist sensor based methods, Gait Posture. 50 (2016) 42–46.  
591 <https://doi.org/10.1016/j.gaitpost.2016.08.012>.
- 592