J Rehabil Med 2022; 54: jrm00277

JRM Journal of Rehabilitation Medicine

ORIGINAL REPORT

ACCURATE ESTIMATION OF UPPER LIMB ORTHOSIS WEAR TIME USING MINIATURE TEMPERATURE LOGGERS

Claudia J. W. HAARMAN, MSc[1,2], Edsko E. G. HEKMAN, MSc[1], Johan S. RIETMAN, MD, PhD[1,3] and Herman van der KOOIJ, PhD[1]

From the [1]Department of Biomechanical Engineering, University of Twente, Enschede, [2]Hankamp Rehab, Enschede and [3]Roessingh Research and Development, Enschede, The Netherlands

Objective: **To propose and validate a new method for estimating upper limb orthosis wear time using miniature temperature loggers attached to locations on the upper body.**

Design: **Observational study.**

Subjects: **Fifteen healthy participants.**

Methods: **Four temperature loggers were attached to the arm and chest with straps. Participants were asked to remove and re-attach the straps at specified time-points. The labelled temperature data obtained were used to train a decision tree classification algorithm to estimate wear time. The final performance (mean error and 95% confidence interval) of the trained classifier and the wear time estimation were assessed with a hold-out data-set.**

Results: **The trained algorithm can correctly classify unseen temperature data with a mean classification error between 1.1% and 3.1% for the arm, and between 1.8% and 4.0% for the chest, depending on the sampling time of the temperature logger. This resulted in mean wear time errors between 0.5% and 8.3% for the arm, and 0.13% and 13.0% for the chest.**

Conclusion: **The proposed method based on a classifier can accurately estimate upper limb orthosis wear time. This method could enable healthcare professionals to gain insight into the wear time of any upper limb orthosis.**

Key words: orthotic device; temperature; treatment adherence and compliance; data accuracy; time; ambulatory monitoring.

Accepted Jan 28 2022; Epub ahead of print Feb 11, 2022

J Rehabil Med 2022; 54: jrm00277

DOI: 10.2340/jrm.v54.043

Correspondence address: Claudia Haarman, Department of Biomechanical Engineering, Faculty of Engineering Technology, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands. E-mail: c.j.w.haarman@utwente.nl

LAY ABSTRACT

Upper limb orthoses are wearable devices that support the impaired shoulder, arm or hand. As orthoses can be effective only if they are worn, information about wear time will help physicians and therapists to evaluate the effectiveness of their prescribed treatment. Currently, physicians and therapists rely mainly on subjective data from patients, such as diaries or questionnaires, which may be biased and inaccurate. This study developed a new, objective method to estimate wear time, based on temperature readings from miniature sensors that can be attached easily to any upper limb orthosis. The results show that the wear time of upper limb orthoses can be assessed accurately using this method.

Upper limb orthoses are frequently prescribed to support patients with impairment of the shoulder, arm and hands. However, an orthosis can be effective only if it is used by patients. Little is known about treatment adherence with wearing upper limb orthoses. Physicians or therapists assume that the patient adheres to their prescribed treatment, but the actual wear time may deviate strongly from the prescribed time. The rate of non-use may be as high as one-third of all devices provided (1). To evaluate the efficacy of a treatment that involves an orthosis, a reliable estimation of the wear time is required.

Subjective methods, such as diaries or questionnaires, are low-cost and easy to implement, but are vulnerable to reporting bias (e.g. social desirability or recall errors) (2, 3). These subjective methods lack accuracy, as most patients tend to overestimate the degree of actual wear time (4). Objective methods, such as measuring force (5–7), acceleration (8) or temperature (2, 4, 7, 9–11), have been proposed to overcome these problems.

Of these proposed methods, temperature seems to be most suitable for measuring orthosis wear time, because of the small size of the sensors, low cost, ease of implementation in existing orthoses, and ability to take measurements over a prolonged period (up to a few months) without human intervention. Currently, no algorithms have been developed to estimate upper limb orthosis wear time using temperature sensors. Previous efforts to estimate wear time have focused mainly on orthopaedic footwear (2, 9) and spinal correction braces (4, 7). In these applications, orthoses are typically worn for long, consecutive periods (from a few hours up to 23 h/day). The device is thus donned and doffed only once or twice per day. In contrast, upper limb orthoses are generally donned and doffed more frequently; for example when performing rehabilitation exercises at home (e.g. 3 times a day for 20 min), or when performing specific daily activities. Algorithms developed previously rely on absolute temperature thresholds (4, 7), or peak detection algorithms to discriminate between on and off states (9). Reported accuracies in these studies range from 86% to 99%, but these results may not be valid if the sensors are applied to other body parts, are donned and doffed more frequently, or if the sensors are not worn directly on the skin.

The aim of this study is to propose and validate a new method to estimate orthotic device wear time using temperature sensors attached to locations on the upper body. The method should accurately estimate wear time during frequent donning and doffing, and while not wearing the sensors directly on the skin. A decision tree classification algorithm was trained using labelled temperature data obtained from healthy participants, and its performance assessed using unseen test data. Instead of using only a single temperature sensor, the study investigated whether a dual sensor configuration (1 sensor directed away from the body and the other directed towards the body) improved the performance of the wear time estimation algorithm. A further aim was to investigate the effect of sampling time on the performance of the algorithm.

METHODS

Temperature sensor data loggers

Thermochron® iButtons® (Maxim Integrated, San Jose, CA, USA) are miniature data loggers that measure and store temperature (Fig. 1A). The sensors are 17 mm in diameter and 6 mm high, and thus are a suitable size for integration into an orthotic device. The DS1922L Thermochron® can store up to 8,192 values with a resolution of 0.5°C. Its sampling time can be programmed from 1 s up to 273 h. With a sampling time of 1 min, the device can store up to approximately 5.5 days of consecutive temperature readings. When the sampling time of the sensor is increased, longer measurements can be performed before data has to be retrieved. Table I shows the maximum periods of unsupervised data collection for different sampling times. Under normal operating conditions (1-min logging interval, and a temperature of 30°C), the DS1922L battery lasts for at least 1 year.

The temperature loggers are secured to the body with 2 elastic straps. One strap is positioned around the chest and the other around the forearm (Fig. 1B).

Table I. Maximum duration of one measurement when different temperature logger sampling times are programmed

Sampling time	Maximum measurement duration
1 min	5.5 days
5 min	28.4 days (~4 weeks)
10 min	56.9 days (~8 weeks)
15 min	85.3 days (~12 weeks)

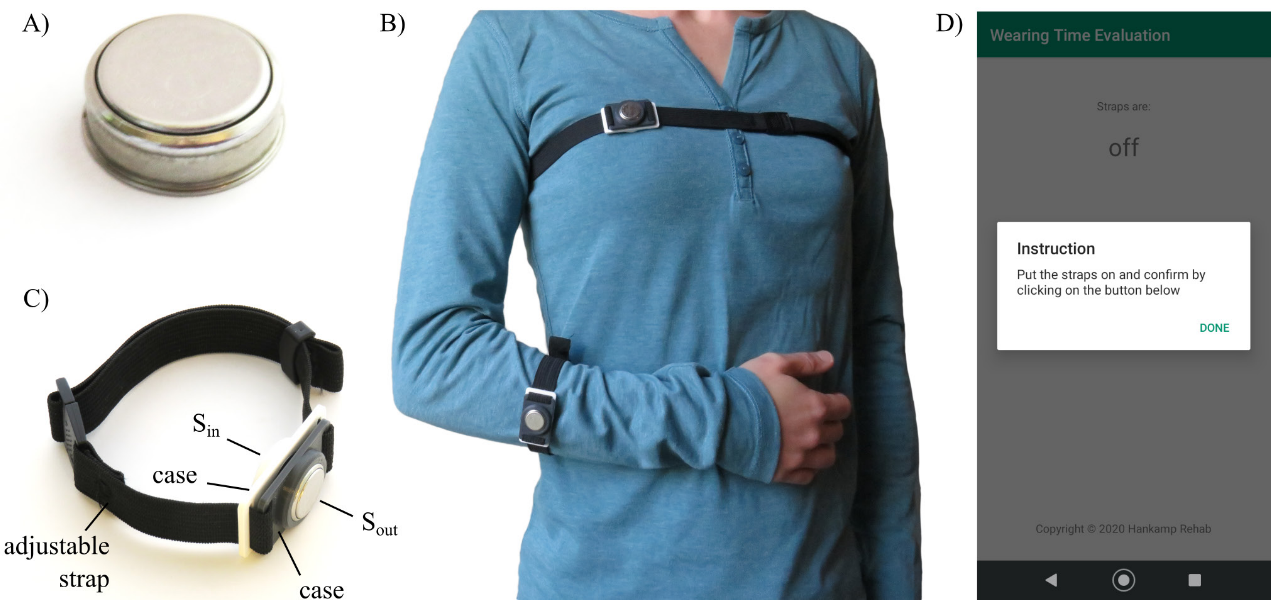

Fig. 1. (A) DS1922L Thermochron® iButton®. (B) Close-up of the adjustable strap with 2 sensors mounted in a 3D-printed case; 1 sensor facing towards the body (S_{in}), and the other sensor facing away from the body (S_{out}). (C) Placement of the straps on the forearm and chest. (D) Screenshot of smartphone app.

Temperature sensors are attached to each strap by means of a custom 3D-printed case (Fig. 1C). One sensor is positioned facing towards the body (S_{in}), such that it touches the participant's clothing. The other sensor is positioned facing away from the body (S_{out}). In order to provide a comfortable interface, the length of the straps is adjusted for each participant.

Smartphone app

To simulate donning and doffing of an orthosis, the participant was asked to remove and re-attach the temperature sensors at specified time-points. For this, a custom Android smartphone application was developed, which cues the participant and registers whether the action has been completed. Based on the timestamps logged by the smartphone app, a label was assigned to each temperature data-point that contains its true state. The app notified participants with an audio signal when it was time to don or doff the straps. Fig. 1D shows a screenshot of the instruction a participant receives after a notification. To record sufficient transitions between use and non-use states (to represent frequent donning and doffing of the orthosis), the intervals between 2 notifications were randomly programmed between 15 and 60 min. The smartphone time was synchronized with the sensor time prior to each measurement.

Measurement protocol

From each participant approximately 24 h of measurement data were collected, consisting of 8 h of active sensor donning and doffing and 16 h of non-use. Both straps (with sensors attached to them) were worn on top of the participant's clothing. During the period in which they actively don and doff the sensors (8 h), participants were instructed to carry out their normal daily routines. Prior to the measurements, participants were instructed to pay attention to the correct sensor orientation (the same sensor should always face towards the body). They were also instructed to immediately confirm their action in the smartphone app after sensor donning or doffing, to minimize the difference between the time of the actual change and the time logged by the smartphone. Participants were allowed to doff the sensors temporarily (e.g. during a clothes change). Unexpected doffing periods that lasted more than 30 s had to be noted. After each measurement session, temperature data from the sensors and timestamp data from the smartphone app were transferred to a computer for further processing and analysis.

Participants

A total of 15 healthy subjects (7 males and 8 females), with a median age of 35 years (range 24–67 years) participated in the study. Inclusion criteria included the ability to follow simple instructions. The study was approved by the ethics committee of the University of Twente (reference number 2020.39). Written informed consent was obtained from all subjects before the study.

Data processing

Each temperature sample was assigned a label 0 ("off") or 1 ("on") according to the timestamps logged by the smartphone app when the sensors were donned and doffed. Seventy percent of the measurement data was randomly assigned to a training set and 30% to a test set. Stratification was applied to maintain an equal distribution of the classes within these 2 sets.

In this study, the data were captured with a sampling time of 1 min. The obtained data-sets were down-sampled by a factor, n, leaving only every nth sample in the data-set. Thus, multiple data-sets with different sampling times were obtained from the original measurement data to allow for an in-depth evaluation of the algorithm performance at different sampling times. The chosen down-sampling factors (n) are: 1, 5, 10 and 15, corresponding to sampling times (t_s) of: 1, 5, 10 and 15 min.

Depending on the configuration used, different features can be calculated and the selected features were used during data processing. The study investigated whether a dual sensor configuration can better estimate wear time compared with a single sensor configuration. In a single sensor configuration, temperature readings from only 1 sensor, directed towards the body (T_{in}), are available. A derived feature is the temperature difference of this sensor from its previous reading ($dT_{in}dt$). In a dual sensor configuration, temperature readings from a sensor directed away from the body (T_{out}) are also available. Thus, the temperature difference between the 2 sensors (ΔT) can also be computed. Table II summarizes the features that were calculated from the measurement data. Table III lists the selected features that were used for data processing in the single and dual sensor configuration.

Binary classification

In total, 16 different data-sets were constructed based on the collected measurement data (Fig. 2). A decision tree classifier was trained to map temperature data (input) to corresponding device states (output) based on example input-output pairs. This can be considered a binary classification problem, as there are only 2 classes to be discriminated for each sample in the data-set: use ("on") or non-use ("off"). A well-trained

Table II. Features that were extracted from the measurement data

Feature	Abbreviation	Equation
Sensor temperature	T_{in} T_{out}	
Temperature difference between current and previous data point	$dT_{in}dt$	$T_{in}(i) - T_{in}(i-1)$
Temperature between 2 sensors	ΔT	$T_{in}(i) - T_{out}(i)$

Table III. Selected features for the single and dual sensor configuration

Sensor configuration	Selected features
Single	$T_{in} + dT_{in}dt$
Dual	$T_{in} + dT_{in}dt + \Delta T$

model correctly predicts these classes ("on" and "off") for each data point obtained during a certain measurement period. Any inconsistency between the predicted state (predicted by the trained model) and the true state (recorded by the smartphone app) will decrease the accuracy with which the class can be correctly determined.

Classifier training and performance evaluation

The classification error rate (e_{class}), or misclassification rate, is defined as the number of incorrect predictions, divided by the total number of predictions. Fig. 3 shows an overview of the classifier training and performance evaluation procedure. Stratified K-fold cross-validation was used to evaluate and optimize the performance (error rate) of the classification algorithm. For this, the training data was split into k subsets (folds). In each fold the class ratios were maintained. For every iteration a different fold was held out as a validation set. The mean classification error rate was then determined over all iterations and the model parameters resulting in the best mean performance were selected

For each data-set (different sampling times and body location), and for the 2 feature subsets (representing a single or dual sensor configuration), the classification error rates and the 95% confidence intervals (95% CI) of the trained model were determined using the held-out test set, containing 30% of the original data. The 95% CIs were calculated using the following equation (1):

$$CI = 1.96 * \sqrt{\frac{e_{class} * (1 - e_{class})}{n}} \quad (1)$$

where e_{class} is the classification error rate.

Wear time estimation (interval data-set)

The cumulative wear time (T_{wear}) is the total time that the orthotic device is worn during a certain measurement period, and is calculated by multiplying the total number of data points labelled "on" by the sampling time. The estimated cumulative wear time (T_{est}) is calculated by multiplying the total number of data points with the classification label "on" by the sampling time.

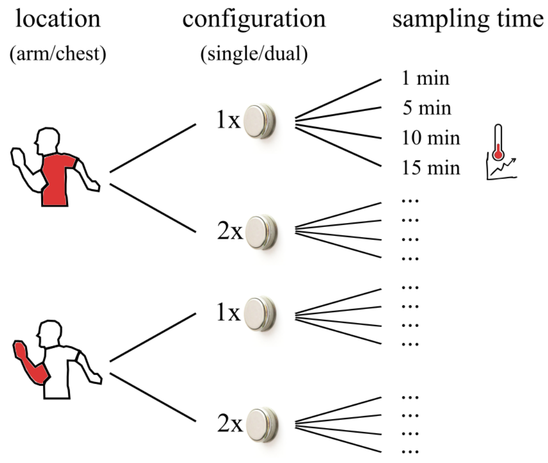

Fig. 2. Construction of different 16 data-sets based on collected measurement data. For all participants 16 data-sets are created, differentiating between location (arm or chest), sensor configuration (single or dual) and sampling time (1, 5, 10 and 15 min).

The wear time error is then defined as the ratio between the estimated and true wear times. The wear time error can be positive (overestimation of the wear time) or negative (underestimation of the wear time). For each data-set mean the wear time errors including their 95% CIs were estimated by bootstrapping the test set (sample size = 80%, iterations = 1000). Bootstrapping is a sampling method to estimate a quantity of a population (12). For this, the test set is randomly sampled many times with replacement.

Each bootstrap sample represents 1 realization of a test set. The resulting mean wear time errors were then calculated, and the 95% CIs were obtained that bound the estimated skill of the trained model. Differences between mean wear time errors of the single and dual sensor configuration are statistically significant ($p < 0.05$) if the degree of overlap of the 95% CIs is $< 50\%$ of the mean margin of error (MOE). The MOE is defined as half the CI width (13).

Wear time estimation (continuous use data-set)

The performance of the trained model to estimate the wear time was also evaluated when sensors were not frequently donned and doffed, but worn for a single long, consecutive period per day. For this purpose, 5 additional data-sets were obtained from healthy participants who only donned and doffed the sensors once (at the beginning and end of an 8-h period of use) during a 24-h measurement session. Temperature and timestamp data from these measurements were processed and provided as test sets to the trained classifier. The

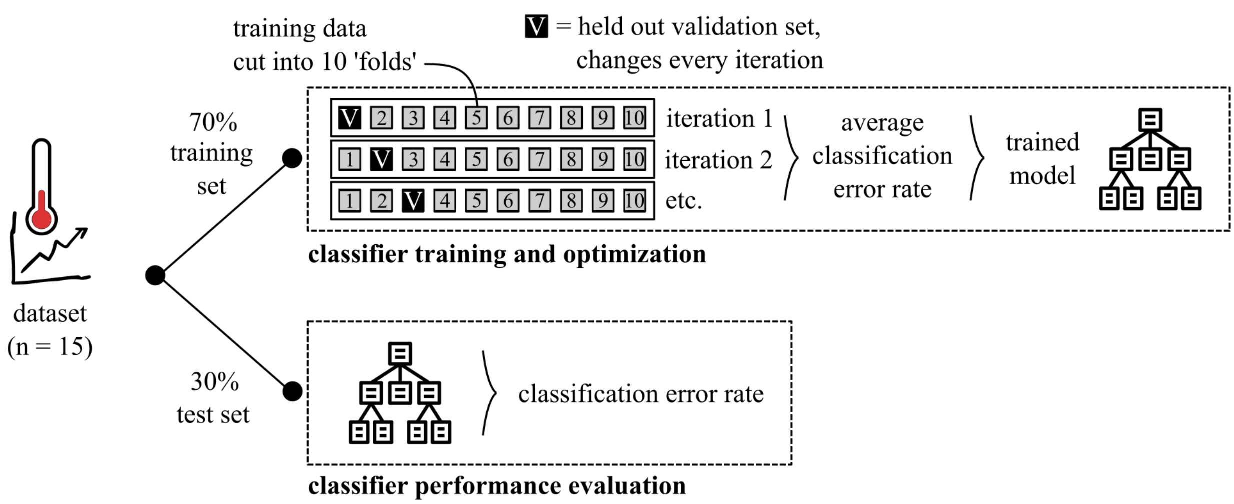

Fig. 3. Graphical overview the classifier training and optimization, and performance evaluation procedure. For each of the 16 data-sets (e.g. single sensor mounted to the chest, and a 5-min sampling time) the temperature data of all 15 participants were used to construct a data-set. Seventy percent of the data were used to train the classifier and the other 30% were used to evaluate the classifier performance of the trained model.

wear time error was calculated in a similar manner as for the original test set.

RESULTS

Data-sets

In total, 22.427 data samples were collected from 15 participants. Unbalanced data-sets (unequal distribution of classes "on" and "off") were obtained due to the nature of the data collection (8 h of repeated donning and doffing, during a 24-h measurement period). Two participants reported intervals with inconsistent sensor orientation during their measurement. In 1 occasion the participant noticed that the donning action was not confirmed in the app. The corresponding samples were removed from their data-set to maintain internal consistency, leaving 22,224 data points for classification. Fig. 4 shows a plot of the 8-h measurement time-period with periods of repeated donning and doffing for a typical participant. This figure shows data from both temperature sensors T_{in} (*black circle*) and T_{out} (*red cross*), which were mounted on the forearm, for increasing sampling times (top to bottom). The *grey areas* indicate the intervals that the sensors were marked "*on*" in the smartphone app. As can be seen from the data, after donning, the temperature increases slowly until the temperature of the contact surface is reached, while after doffing, the temperature slowly returns to the ambient temperature. It can also be seen from the graphs that the temperature T_{in} was higher than the temperature T_{out} throughout the entire measurement period, which lasted over 8 h.

Classifier performance

The classifier performance was evaluated by cross-validating the training set and determining the mean of the calculated classifier errors. In Fig. 5 the classification errors (%) are presented for the arm (left) and chest (right), and for the single (o) and dual (x) sensor configuration and different sampling times. Error bars indicate the 95% CIs of the classification error. Overall, the mean classification error increased for increasing sampling times. Classification errors of the chest sensor(s) were, in all cases, higher than the errors of the arm sensor(s). Fig. 5 shows that the lowest classification errors (1.4% and 1.1%) were seen when the sensor(s) were mounted to the arm and sampled every 1 min. The highest classification errors (4.0% and 3.0%) were made when the sensor(s) were mounted to the chest and sampled every 15 min. For all data-sets, the classification errors in the dual sensor configuration were lower than in the single sensor configuration. In addition, the classifier performance on unseen data (test set containing 30% of held-out data) was also reported (x) in Fig. 5.

Wear time estimation (interval data-set)

Besides the classifier performance, the wear time estimation performance was also obtained by comparing the estimated wear times of the test set with the true wear times. In Fig. 6 (*top*) the mean wear time errors (%) and their 95% CIs are shown for each data-set.

For the arm, the mean wear time errors were –2.3%, –0.5%, –0.1% and 4.9%, respectively, for a 1, 5, 10 and 15 min sampling time. For the chest, these mean errors were 1.5%, –0.4%, –1.9% and –11.5% for a 1, 5, 10

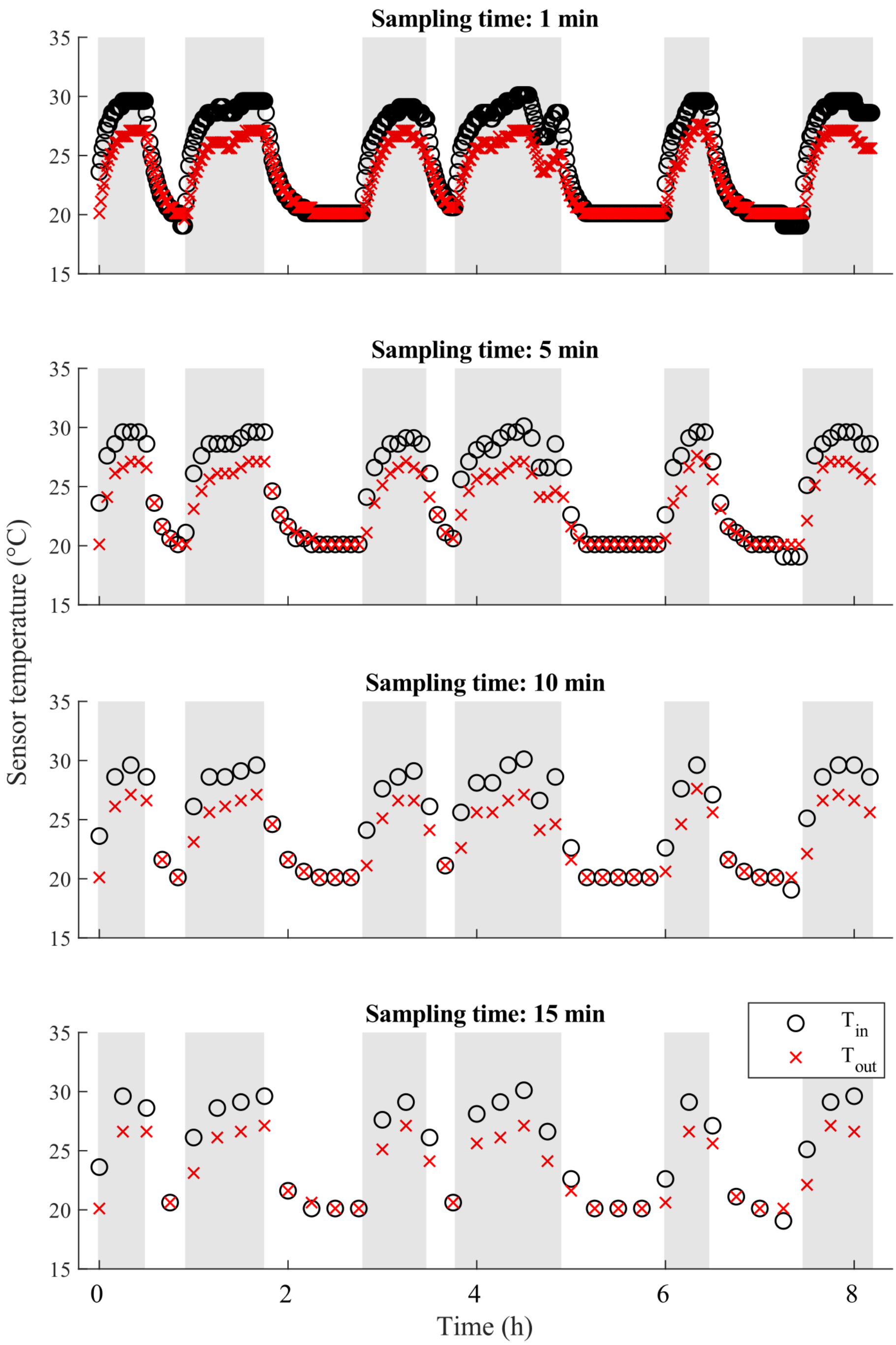

Fig. 4. Typical temperature data from the inward (*black dots*) and outward (*red crosses*) directed sensors on the arm, for different sampling intervals. The top figure shows the original temperature data, the other figures show the down-sampled measurement data (5, 10 and 15 min). The *grey areas* indicate the intervals that the sensors were marked "on" in the smartphone app.

and 15 min sampling time, respectively. The 95% CIs increased with increasing sampling times, meaning that the range of plausible values for the true wear time error increased. The mean wear time errors between the single (o) and dual (x) sensor configuration were not significantly different, as their 95% CIs overlap sufficiently (13).

Wear time estimation (continuous use data-set)

To evaluate the performance of the trained classification model for alternative use cases, an alternative test set was presented to the trained model. The test set consists of data from 5 participants who wore the sensors for 8 h consecutively, and then doffed the sensors. In total, 24 h of data were collected from each participant. For this test set the mean wear time errors were obtained after bootstrapping, based on the trained classification algorithm. Fig. 6 (*bottom*) shows these wear time errors. It can be seen that the algorithm in all cases overestimates the actual wear times (positive error). Also, the mean errors were larger than for the 15–60 min use intervals test set.

The mean wear time errors were 4.3%, 3.8%, 8.5% and 15.3%, respectively, for 1, 5, 10 and 15 min sampling times. For the chest, these mean errors were 5.4%, 5.7%, 5.2% and 11.6%, respectively, for 1, 5, 10 and 15 min sampling times. The 95% CIs, however, were smaller than for the other test set, which indicates a smaller range of plausible values of the true error. The differences between the mean wear time errors between the single (o) and dual (x) sensor configuration were statistically significant in 4 out of 8 data-sets (see Fig. 6, *bottom*).

DISCUSSION

Obtaining information about upper limb orthotic device wear time can improve treatment outcomes. This study proposed a new method, based on a trained decision tree classification algorithm, to estimate the wear time, using miniature temperature loggers attached to the arm and chest, and evaluated its performance. Data were captured during periods of frequent donning and doffing of the sensors, and while wearing the sensors on top of the clothing, mounted on the arm and chest. The study showed that the trained algorithm can correctly classify unseen temperature data with a mean classification error of between 1.1% and 3.1% for the arm, and between 1.8% and 4.0% for the chest. This results in mean wear time errors of between 0.5% and 8.3% for the arm, and between 0.13% and 13.0% for the chest. The wide availability and low cost (approximately 30 USD) of the selected miniature temperature logger makes it an affordable solution for healthcare professionals to gain insight into the wear time of any prescribed (commercial) upper limb orthosis.

This study investigated whether the wear time can be better predicted when 2 sensors (dual sensor configuration) are used, compared with one sensor (single sensor configuration). For the 15–60 min interval use case, the algorithm performed equally well for the single and dual sensor configuration. For the 8-h continuous use case, the dual sensor configuration performed significantly better for the arm (1 and 5 min sampling times) and chest (1 min sampling time), and worse for the arm (15 min sampling time) compared with the single sensor configuration. In the other conditions, the

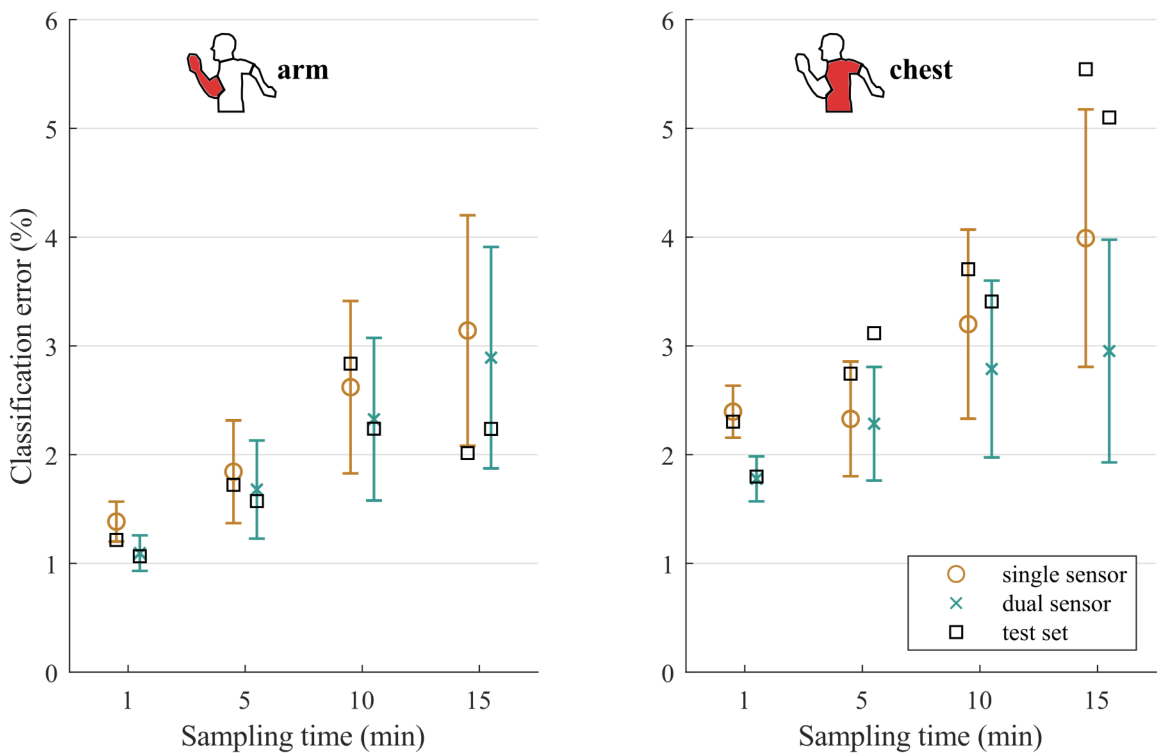

Fig. 5. Mean classification errors for the arm (*left*) and chest (*right*) obtained after cross-validating the training set. Results are shown for the single (o) and dual (x) sensor configuration as well as for different sampling times. The error bars represent the 95% confidence intervals (95% Cis). Classification errors resulting from the test set are indicated with (□).

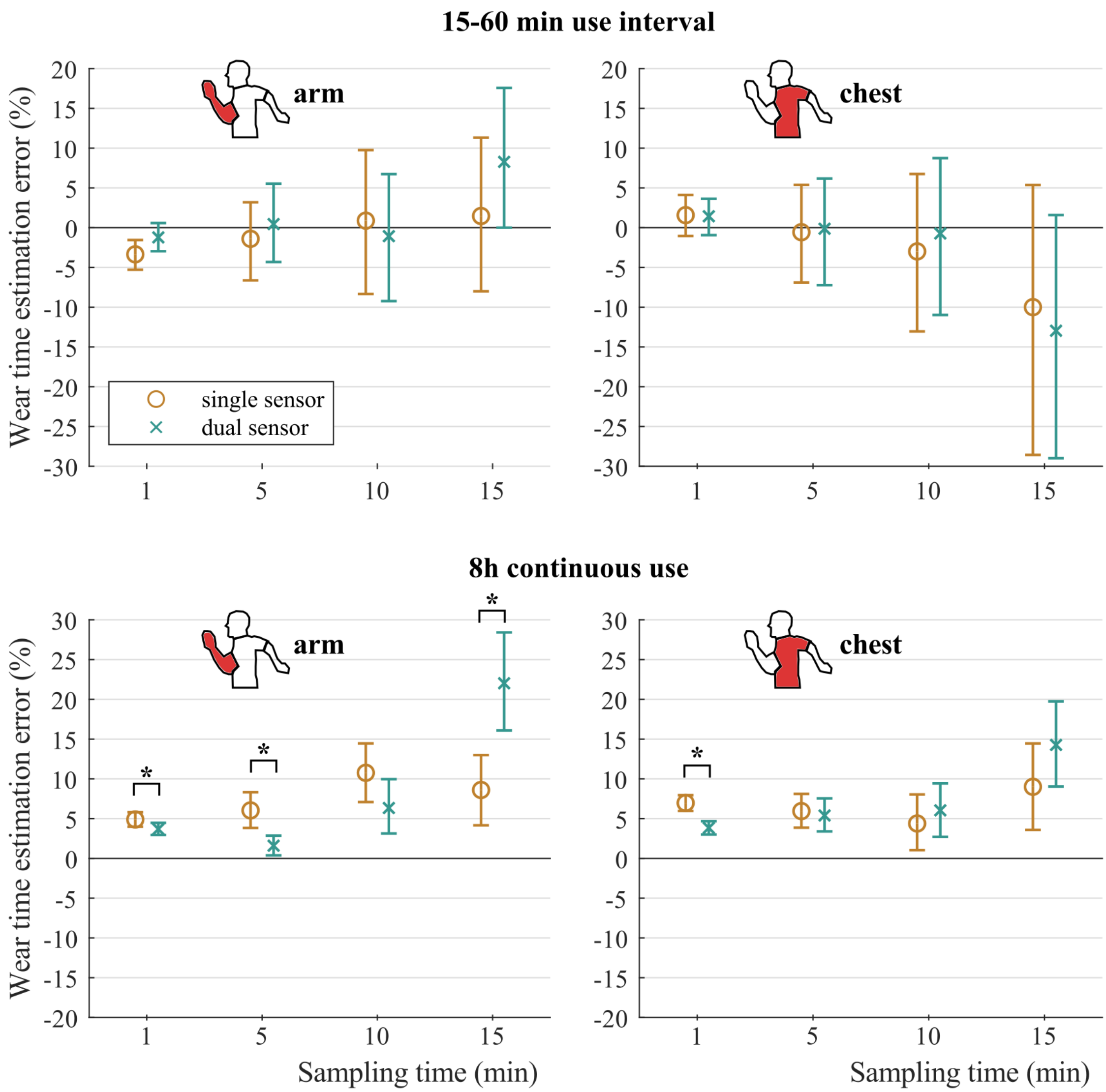

Fig. 6. Test set performance for the single (o) and dual (x) sensor configuration. The mean wear time errors were calculated from sensors that were worn during 15–60-min use intervals (*top*) and 8 h of consecutive use (*bottom*), on the arm (*left*) or chest (*right*). Error bars represent the 95% CIs of the wear time error, obtained after bootstrapping the test set. Significant differences between the single and dual sensor configuration are indicated with (*).

algorithm performed equally well for the single and dual sensor configuration. The added value of using 2 sensors is thus dependent on the use case, sampling time and body location. In general, we can conclude that, in most conditions, there is no practical benefit of using 2 sensors, as this will only increase the volume and cost of the solution, while not resulting in a significantly better wear time estimation.

Compared with subjective alternatives, such as diaries or questionnaires where patients tend to overestimate their wear time by as much as 200% (14), the proposed method is a preferable choice. A direct comparison between the results of this study and other studies using objective methods based on temperature sensors is difficult, as the use cases differ to a great extent. However, some general remarks can be made, as follows.

Firstly, we have shown that we can accurately estimate wear times without the need for direct contact between the sensor and the skin. This enables a wide application of our method, as many upper limb orthoses, such as arm slings, are often worn on top of clothing instead of directly on the skin. The temperature sensor can therefore be easily applied to many types of (commercial) upper limb orthoses. In other studies, the sensor was embedded in the thermoplastic of an orthosis or insole of a shoe, limiting the potential use for different orthoses.

Secondly, we have shown that our algorithm is able to accurately estimate wear time, even during periods of frequent donning and doffing. Temperature sensors need time to warm up when donned or cool down when doffed. These periods of transition may be difficult to detect otherwise, but our algorithm was able to capture these effects, as indicated by the reported classification errors. In other studies, results during frequent donning and doffing are not available as their use case was different (prolonged, continuous use).

Thirdly, the performance of our algorithm was evaluated with unseen test data, leading to unbiased results that show the actual classification and wear time estimation errors. Other studies trained and evaluated their model on the same data-set, leading to overestimated algorithm accuracies.

Depending on the use case, the wear time prescription generally includes a range of several hours, e.g. 2 times a day for half an hour, to 20–23 h/day consecutively. As an example, a wear time error of 5% results in an absolute error of 12 min when the orthosis is worn for 4 h during a 24-h period. The choice for an acceptable level of wear time estimation accuracy is up to the physician or therapist, but, in general, knowledge of wear time to within 90% accuracy is sufficient. For short sampling intervals, reported wear time estimation errors in this work are well within this requirement.

Our method allows monitoring of the patient's adherence for a prolonged period (up to a few months), without supervision, return to the clinic, or recharging of the device, depending on the chosen sampling time. This will further increase the chance of successful implementation in the current clinical practice. For longer sampling times, the amount of data that can be stored on the sensor increases, but the estimated wear time error and estimation error range (indicated by the 95% CIs) also increase. The therapist or physician using this technology should be aware of these implications for the margin of error when choosing a sampling time.

The timestamps logged by the smartphone app were used to label the measurement data ("on" or "off"). Participants were instructed to respond in a timely manner to the notification and immediately confirm their action ("don" or "doff") once done. Theoretically, there could have been a difference between the actual and reported timestamps, which could have negatively affected the classifier performance. However, only 1 participant reported a discrepancy between 1 donning action and subsequent confirmation in the app during 1 interval. Therefore, it was assumed that any discrepancy between the actual and reported timestamps was negligible.

The data for this study were collected in December and January in the Netherlands. As no data were recorded on warm days, the model was not specifically trained for conditions with a high ambient temperature. However, a large variation was still present in the recorded data-sets, to allow for a proper training of the classification model. During the measurements, the ambient temperature outside was in the range 5–10°C. Therefore, participants often wore a coat on top of their clothing when they went outside. Indoors, the ambient temperature was limited, and was approximately in the range 18–21°C. All participants wore long sleeves, but the thickness of their clothing varied, enabling us to train the model for a wide variety of clothing types. We believe that the data presented here represent a worst-case, as (thick) clothing in between the sensors and skin make it more difficult to correctly estimate the wear time. Future studies should address the effect of a higher number of parameters (high ambient temperature, sensors worn directly on skin) on the classifier performance and wear time estimation. By adding these data to the data-set, the classification and subsequent wear time estimation may be improved.

The authors have no conflicts of interest to declare.

REFERENCES

1. Scherer M, Galvin J. Matching people with technology. Rehabil Manag 1994; 7: 128–130.
2. Ehrmann D, Spengler M, Jahn M, Niebuhr D, Haak T, Kulzer B, et al. Adherence Over time: the course of adherence to customized diabetic insoles as objectively assessed by a temperature sensor. J Diabetes Sci Technol 2018; 12: 695–700.
3. Helfenstein A, Lankes M, Ohlert K, Varoga D, Hahne HJ, Ulrich, HW, et al. The objective determination of compliance in treatment of adolescent idiopathic scoliosis with spinal orthoses. Spine 2006; 31: 339–344.
4. Benish BM, Smith KJ, Schwartz MH. Validation of a miniature thermochron for monitoring thoracolumbosacral orthosis wear time. Spine 2012; 37: 309–315.
5. Havey R, Gavin T, Patwardhan A, Pawelczak S, Ibrahim K, Andersson GB, et al. A reliable and accurate method for measuring orthosis wearing time. Spine 2002; 27: 211–214.
6. Lou E, Hill D, Hedden D, Mahood J, Moreau M, Raso J. An objective measurement of brace usage for the treatment of adolescent idiopathic scoliosis. Med Eng Phys 2011; 33: 290–294.
7. Hunter LN, Mitell SW, Mendoza MM, McDonald CM, Molitor F, Mulcahey MJ, et al. The validity of compliance monitors to assess wearing time of thoracic-lumbar-sacral orthoses in children with spinal cord injury. Spine 2008; 33: 1554–1561.
8. Chadwell A, Kenney L, Granat M, Thies S, Head JS, Galpin A. Visualisation of upper limb activity using spirals: a new approach to the assessment of daily prosthesis usage. Prosthet Orthot Int 2018; 42: 37–44.
9. Lutjeboer T, Van Netten JJ, Postema K, Hijmans JM. Validity and feasibility of a temperature sensor for measuring use and non-use of orthopaedic footwear. J Rehab Med 2018; 50: 920–926.
10. Schwarze M, Horoba L, Block J, Putz C, Alimusaj M, Wolf SI, et al. Wearing time of ankle-foot orthoses with modular shank

supply in cerebral palsy: a descriptive analysis in a clinically prospective approach. Rehabil Res Pract 2019; 2019.
11. Takemitsu M, Bowen JR, Rahman T, Glutting JJ, Scott CB. Compliance monitoring of brace treatment for patients with idiopathic scoliosis. Spine 2004; 29: 2070–2074.
12. Efron B. Bootstrap methods: another look at the jackknife. Ann Statistics 1979; 7: 1–26.
13. Cumming G, Finch S. Inference by eye: confidence intervals and how to read pictures of data. Am Psychol 2005; 60: 170–180.
14. Nicholson GP, Ferguson-Pell MW, Smith K, Edgar M, Morley T. The objective measurement of spinal orthosis use for the treatment of adolescent idiopathic scoliosis. Spine 2003; 28: 2243–2250.

JRM Journal of Rehabilitation Medicine JRM Journal of Rehabilitation Medicine JRM