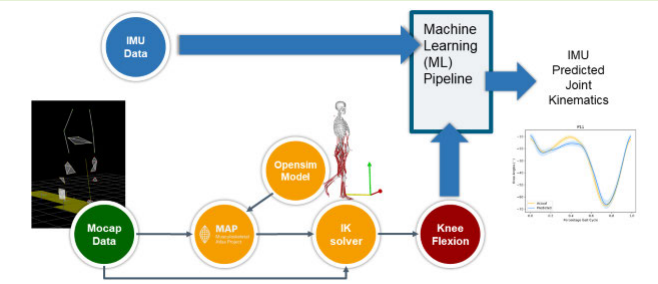
Personalized Machine Learning Approach to Estimating Knee Kinematics Using Only Shank-Mounted IMU

Ted Yeung, Astrid Cantamessa, Andreas W. Kempa-Liehr, *Member, IEEE*, Thor Besier, and Julie Choisne

***Abstract*—Knee kinematics is a valuable measure of knee joint function. However, collecting that data outside the clinic is difficult, especially with a limited number of wearable sensors and when you only use an ankle-mounted inertial measurement unit (IMU) to estimate knee kinematics. Due to the cyclic nature of gait, it is possible to use machine learning to extract joint angles from only ankle-mounted sensors. This study aimed to use time-series feature extraction and a random forest regressor to generate a person-specific surrogate model for estimating knee joint flexion angles from a single-mounted IMU above the ankle. Optical motion capture (OMC) and inertial data from ten healthy participants walking on a treadmill were collected to create ten personalized surrogate models for estimating right knee flexion angles during gait. An additional ten models were created for a leave-one-out analysis to test the generalisability of the models. Temporal cross validation of the personalized models and a leave-one-out analysis was performed on the selected feature set. The personalized models achieved an average root-mean-square error (RMSE) of 2.45° ± 0.65° (R^2 of 0.98) compared to a gold-standard OMC. The generalized models achieved an average RMSE of 6.77° ± 3.38° (R^2 of 0.83) in the leave-one-out analysis. Time-series feature-based personalized surrogate models could be used to accurately estimate knee kinematics by using a single ankle-mounted sensor. However, more data are required to train a generalized model using the presented method.**



***Index Terms*— Health monitoring, inertial measurement unit (IMU), knee kinematics, machine learning, personalized models wearable.**

I. Introduction

WEARABLE sensors, such as inertial measurement units (IMUs), are increasingly used to collect knee kinematics and improve traditional clinical assessments. However, obtaining accurate knee kinematics from IMUs is challenging and currently requires attaching IMUs on each side of the knee and a complex kinematic model with known person-specific parameters (e.g., IMU placements relative to the body) [1], [2], [3], [4], [5], [6]. Improper calibration or IMU placement could lead to misalignment errors, typically called "kinematic crosstalk" [7], [8]. This type of error leads to an overestimation of joint angles. Furthermore, the resulting knee flexion angles are also prone to errors such as numerical drift. This arises from using numerical integration to calculate orientation. Sensor fusion techniques, such as Madgwick et al. [9] and Kalman filters [10], [11], [12], have been used to solve this issue by fusing the gyroscope, accelerometer, and magnetometer data for estimating a stable orientation (pitch, roll, and yaw). However, these values are still sensitive to the amount of linear acceleration and electromagnetic interference, for instance, near or on a treadmill.

One property of gait we can exploit to solve this issue is the cyclic nature of the motions. The repeating cycles can be used with IMU data to train a machine learning model for predicting knee flexion angles directly from raw IMU data (acceleration and angular velocity). This reduces the number needed and solves common IMU errors such as numerical drift. Machine learning approaches have been used to predict gait kinematics, biomechanics, and spatiotemporal parameters

Manuscript received 19 February 2023; revised 4 April 2023; accepted 5 April 2023. Date of publication 19 April 2023; date of current version 31 May 2023. This work was supported by the National Science Challenges: Science for Technological Innovation New Zealand. The associate editor coordinating the review of this article and approving it for publication was Dr. Varun Bajaj. *(Corresponding author: Ted Yeung.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University of Auckland Human Participant Ethics Committee (UAHPEC) under Reference No. 017659.

Ted Yeung, Andreas W. Kempa-Liehr, Thor Besier, and Julie Choisne are with The University of Auckland, Auckland 1010, New Zealand (e-mail: tyeu008@aucklanduni.ac.nz).

Astrid Cantamessa is with the University of Liege, 4000 Liege, Belgium.

Digital Object Identifier 10.1109/JSEN.2023.3267398

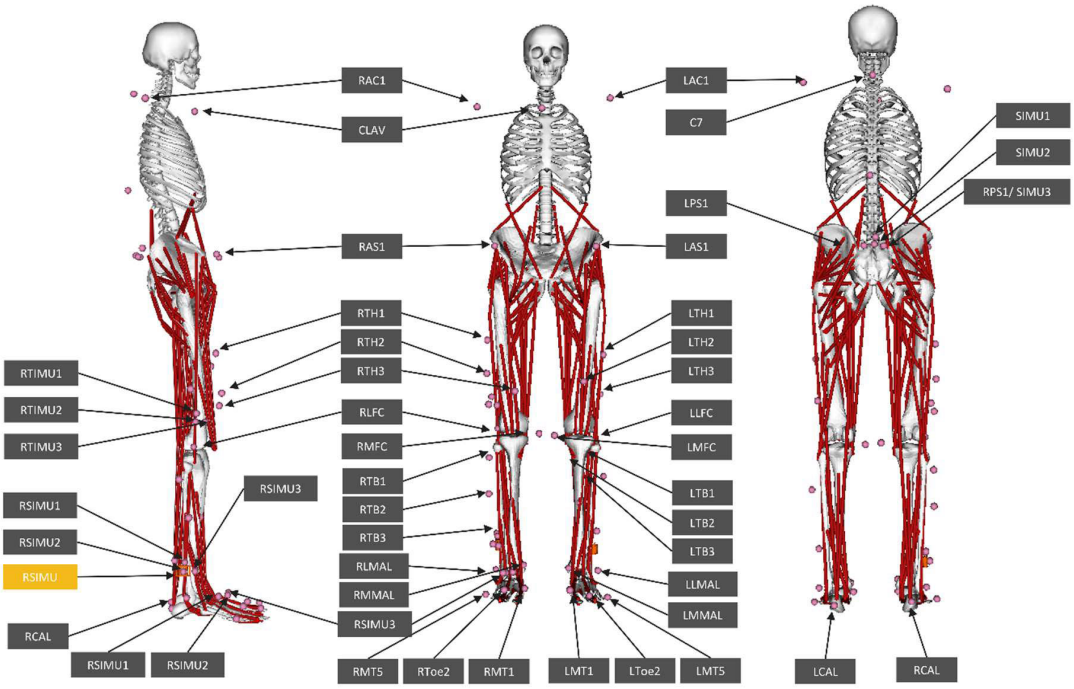


Fig. 1. Marker set (46 Markers) used for treadmill walking. IMU placement (RSIMU) is colored in orange [20].

to bone load [13], [14], [15], [16]. However, most of these methods are based on neural networks requiring an extensive dataset to train and are usually not optimized for time-series datasets. Furthermore, most of these approaches treat the feature extraction processes as a black box, making it difficult to determine the source of error for the resulting model. With the recent trend toward explainable artificial intelligence [7], tools, such as Time-Series FeatuRe Extraction based on Scalable Hypothesis tests (tsFresh, version 0.14.1 [17]), have been developed to aid in extracting features in a statistically meaningful way [17].

In this study, we used an adaptation of the novel approach mentioned above in explainable artificial intelligence to investigate whether a random forest surrogate model could: 1) estimate knee angles with errors under the clinical threshold of 5° [18] and 2) whether it is feasible to estimate knee angles and with one a single IMU mounted on the shank.

II. Materials and Methods

A. Data Collection

Healthy individuals aged between 18 and 40 years with no history of knee injury were selected for the study to remove the effect of joint health as a variable for evaluating the performance of the models. Treadmill walking data from ten healthy adult participants (four females, six males, 27.4 ± 4.7 years old, 1.75 ± 0.08 m, and 72.32 ± 10.33 kg; no history of a knee injury) was collected at a gait laboratory (AUT Millennium, Auckland, New Zealand). Optical motion capture (OMC) was recorded using a VICON motion capture system (8 MX T40-S cameras operating at 200 Hz), and synchronized acceleration (±16 g) and angular velocity (±2000°/s) data were captured at 250 Hz using IMeasureU Blue Thunder sensors [19] and VICON Nexus Software (Nexus 2.8).

TABLE I
Shank-Mounted IMU Data Definitions

Data	Description
Shank_imu_Ox	Angular Velocity in the X-axis
Shank_imu_Oy	Angular Velocity in the Y-axis
Shank_imu_Oz	Angular Velocity in the Z-axis
Shank_imu_X	Acceleration in the X-axis
Shank_imu_Y	Acceleration in the Y-axis
Shank_imu_Z	Acceleration in the Z-axis

We placed 46 retroreflective markers (see Fig. 1 [20]) on the participants with locations set out by the University of Western Australia marker set [21]. In addition, a shank-mounted IMU on the right leg is 1 cm above the lateral malleolus (outside ankle) (RSIMU shown in Fig. 1). A static pose was captured before the participants were asked to complete 1 min of treadmill walking at a self-selected speed. The labels assigned to the IMU data are presented in Table I.

B. Data Processing

An OpenSim lower limb model (gait2392.osim [22]) was customized and scaled to the participant using the MAP client [23], [24]. Inverse kinematics (IKs) was performed

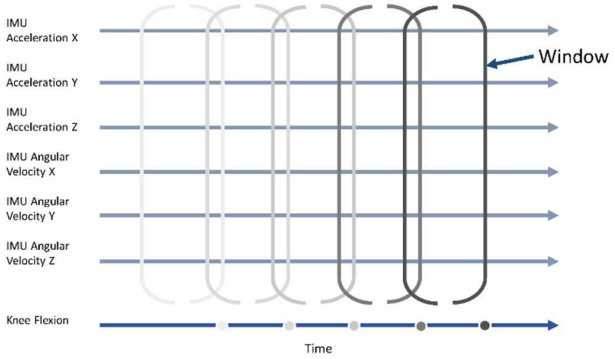


Fig. 2. Windowing IMU data for feature extraction [23].

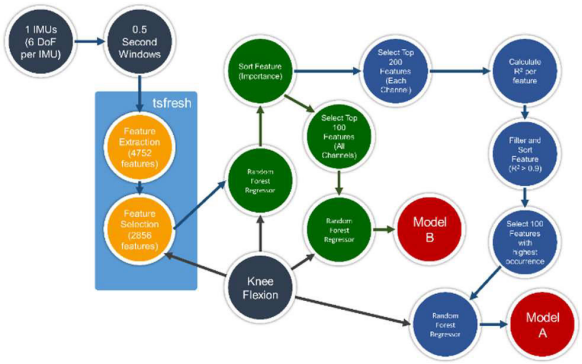


Fig. 3. Features selection algorithm on instrumented single leg and model training workflow: 1) orange circles are the initial feature extraction and selection; 2) green circles are the components of the pipeline that use model-specific features selected by just feature importance to train the final model; and 3) the blue circles are the components of the pipeline that use model-specific features selected by its feature importance and correlation to the target value to train the final model [30].

in OpenSim using this scaled model to track the marker trajectories and calculate lower limb joint kinematics. The resulting knee kinematics from IK were filtered using a third-order Butterworth low-pass filter with a cutoff frequency of 6 Hz [25]. The resulting knee kinematics data are separated into the training and evaluation sets. The training set will be used with the corresponding IMU data for generating a surrogate model to estimate knee angles. The evaluation set will be used to test the performance of the resulting model. Raw IMU data were resampled using univariate spline to 200 Hz to match the joint kinematic data. These blocks of data were used for feature extraction [26], [27]. Zero offsets in gyroscope data were removed by averaging the x, y, and z components of angular velocity separately during the static pose and removing it from the measured angular velocity data giving six channels of time-series data (three for acceleration and three for angular velocity).

C. Machining Learning Workflow and Algorithm

For the machine learning model, we chose the random forest model instead of another deep learning model due to the small sample size in the dataset. It was shown in a previous study that for limited sample sizes, the random forest model had the best chance of producing a usable model [27] and was less likely to overfit [28]. To create a model for estimating knee angles, each time series, x_i, in the set of IMU time series ($T = \{x_i\}_{i=1}^{N}$) needs to map to the feature space. This is accomplished by first windowing the data 0.5-s blocks (see Fig. 2 [29]) with a step size of 0.005 s. Using the windowed data, a feature vector can be created for each of the time series $\vec{x_i} = (f_1(x_1), f_2(x_2), \ldots, f_m(x_i))$ and then combined into an input matrix F for random forest regression. Using an open-source Python package (tsfresh), we calculated 794 features per time series, which for this study equates to a feature set, F, with 4764 features.

Two different methods of feature selection were tested (see Fig. 3 [30]):

1) tsFresh feature selector and sorted by feature importance with R^2 filtering (Set A);
2) tsFresh feature selector and sorted by feature importance (Set B).

The 4764 features were filtered through the tsFresh feature selector, where they were tested for their statistical significance in predicting the target knee angles. This is done by testing the hypothesis: H_0 = the feature is not relevant and cannot be added and H_1 = the feature is relevant and should be kept. Any features that failed the significance testing were removed from the set. This approach ensured that the features set contained relevant features before being used as inputs for the multivariate nonlinear regression [14], [17].

For the collected data, this step reduced the feature set to an average of 2856 features. Then, a temporary model was created by performing a multivariate nonlinear regression using the Random Forest Regressor (from the Python package, scikit-learn). The coefficient of determination and feature importance determined how well each feature correlated with the corresponding knee angles during gait. These features were sorted based on their importance (computed with mean impurity decrease, scikit-learn). Then, a subset of the 100 most important features (Set B) was selected for training the final model.

To further optimize the feature set, a second approach (Set A) was also taken where, instead of using the top 100 features to train a final model, the top 200 features were extracted from each channel of the IMU data. Then, a count of appearance is made of each feature across the different participants to determine its prevalence in their feature set. Using this information, the feature lists from each channel were combined and reordered based on their prevalence. Finally, the combined feature set is filtered to remove features that have R^2 less than 0.9 before selecting the top 100 features for training the final models. This approach aimed to reduce the possibility that useful data would be removed in the final model. The algorithm workflow is presented in Fig. 3. The top ten features in the selected feature set are shown in Section III.

D. Model Evaluation

To evaluate the model estimations, the data were reorganized into gait cycles. This is done by first performing a peak analysis (scipy's find peaks function) on the knee kinematics time series to find the mid-swing, heel strike, and toe-off

TABLE II
FEATURES DESCRIPTION FOR TABLE III

Features Names	Description
agg_linear_trend	Calculates a linear least-squares regression for values of the time series that were aggregated over chunks versus the sequence from 0 up to the number of chunks minus one.
f_agg_"max", f_agg_"mean", f_agg_"min"	It is the max, mean or min of the attribute that were aggregated over chunks versus the sequence
Chunk_len	The number frames in the chunk used in the a linear least-squares regression for the given sequence (i.e. window of data)
attr_"slope", attr_"intercept"	They are the attributes of the fitted line in the linear least-squares regression

time points. Then, the data are windowed from heel strike to heel strike. Next, temporal cross validation was performed to evaluate the predictive performance of the models. The models were trained on the first 70% of the data in this approach and tested on the remaining 30%. This resulted in ∼50 s of training data and ∼20 s of testing data for each individual. Next, we calculated the root-mean-square error (RMSE) and coefficient of determination (R^2) values between the predicted knee kinematics from IMU data and the "gold-standard" kinematics value obtained from the IK analysis using OMC data. A leave-one-out analysis was also performed to investigate the generalizability of the trained model with the selected feature set. Finally, a stability analysis was also performed by generating each model ten times to determine variation in the model generated by the method. This variation is reported as the coefficient of variation in the result section.

III. RESULTS

The final feature selection was based on the percentage of participants that have a particular feature. A brief description of the features can be found in Table II. After feature selection, the top ten features shown in Table III appeared in the top 100 of at least 40% of the participants (see Fig. 4). The most common feature was the maximum angular acceleration seen in the windowed data along the z-axis corresponding to the feature Shank_imu_Oz__agg_linear_trend__f_agg_"max"__chunk_len_5__attr_"slope" in Table III.

In the model stability analyses (see Fig. 5), the results showed (RMSE) a maximum value of 2.5 % variation from the mean model, and when the clinical threshold was considered, the models only varied by 1.2% from the mean model.

The person-specific models presented in Table IV showed low RMSE (average error of 2.45°) with an average correlation coefficient of 0.98. The standard deviation of the RMSE was

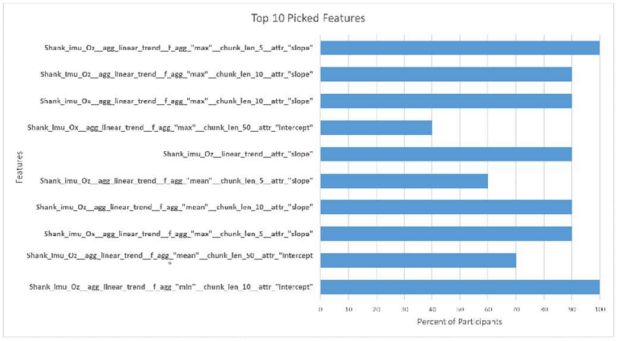


Fig. 4. Top ten selected features based on feature set A and their prevalence in different participant's feature set.

TABLE III
TOP TEN FEATURES AND ITS MAXIMUM REPORTED IMPORTANCE ACROSS TEN DIFFERENT PARTICIPANT-SPECIFIC MODELS

Features Names	Importance
Shank_imu_Oz__agg_linear_trend__f_agg_"max"__chunk_len_5__attr_"slope"	0.80
Shank_imu_Oz__agg_linear_trend__f_agg_"max"__chunk_len_10__attr_"slope"	0.78
Shank_imu_Ox__agg_linear_trend__f_agg_"max"__chunk_len_10__attr_"slope"	0.73
Shank_imu_Ox__agg_linear_trend__f_agg_"max"__chunk_len_50__attr_"intercept"	0.68
Shank_imu_Oz__linear_trend__attr_"slope"	0.21
Shank_imu_Oz__agg_linear_trend__f_agg_"mean"__chunk_len_5__attr_"slope"	0.15
Shank_imu_Oz__agg_linear_trend__f_agg_"mean"__chunk_len_10__attr_"slope"	0.18
Shank_imu_Ox__agg_linear_trend__f_agg_"max"__chunk_len_5__attr_"slope"	0.10
Shank_imu_Oz__agg_linear_trend__f_agg_"mean"__chunk_len_50__attr_"intercept"	0.05
Shank_imu_Oz__agg_linear_trend__f_agg_"min"__chunk_len_10__attr_"intercept"	0.05

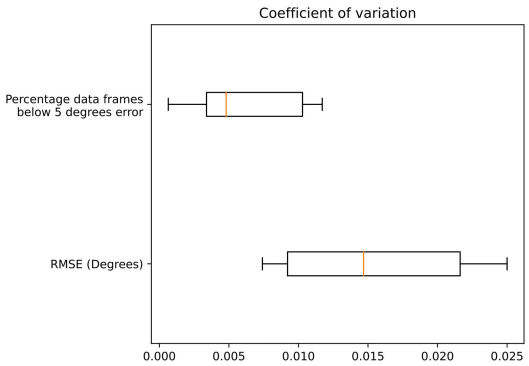


Fig. 5. Result of ten repeats of model generation per participant.

0.65°. The maximum RMSE errors were 3.56° for participant P07.

Due to the better performance of Set A, the features from Set A were used to create the models used for the leave-one-out analysis. In this analysis, seven out of ten models in the study have R^2 values above 0.87, and five models have average RMSE values below 5°. The best case was the model used to

TABLE IV
COMPARISON BETWEEN GOLD-STANDARD KNEE KINEMATICS AND IMU-DERIVED KNEE KINEMATICS USING THE PERSON-SPECIFIC MODEL OF SELECTION SETS A AND B

Participant #	**Set A**		**Set B**	
	Average RMSE (degrees)	*Correlation coefficient (R^2)*	*Average RMSE (degrees)*	*Correlation coefficient (R^2)*
P07	3.56	0.97	3.52	0.97
P08	2.13	0.99	1.98	0.99
P09	2.08	0.99	4.46	0.94
P10	3.33	0.97	3.23	0.97
P11	2.18	0.98	2.03	0.99
P12	2.08	0.99	2.00	0.99
P13	2.67	0.97	2.43	0.98
P15	1.65	0.99	1.70	0.99
P16	1.83	0.99	1.59	0.99
P17	3.02	0.97	2.91	0.97
Mean	2.45	0.98	2.58	0.98
Standard deviation	0.65	0.01	0.93	0.02

TABLE V
LEAVE-ONE-OUT ANALYSIS OF SELECTED FEATURE SET A

Participant #	*Average RMSE (degrees)*	*SD (degrees)*	*Correlation coefficient (R^2)*	*BMI*
P07	12.91	10.70	0.64	25.14
P08	5.70	5.37	0.91	21.86
P09	10.43	6.49	0.70	25.17
P10	6.21	3.47	0.87	20.90
P11	3.41	2.68	0.97	22.46
P12	4.61	4.59	0.94	21.93
P13	4.47	4.33	0.93	20.23
P15	11.01	7.08	0.50	25.08
P16	4.24	4.08	0.94	22.53
P17	4.63	4.36	0.93	28.03
Mean	6.76	5.32	0.83	23.33
Standard deviation	3.38	2.30	0.15	2.42

predict P11's gait, while the worst case was P07's gait (see Fig. 6 [31]), which showed a significant difference.

The highest error (Table V) in the generalized model occurs during the stance phase with the predicted value above the 5° threshold (see Fig. 7 [32]). The lowest errors occur in the middle of the swing phase of gait. For personalized models, the errors are well below the 5° clinical error threshold.

IV. DISCUSSION

This study provides a method for estimating knee flexion range of motion using only a shank-mounted IMU. Furthermore, it eliminates potential sources of error (i.e., numerical drift and crosstalk) by directly relating the time-series features extracted from the IMUs inputs to the desired knee angles. We produced ten personalized surrogate models with good predictive accuracy (average RMSE of 2.45°) of knee flexion angles. This is well below the 5° error threshold [18] and indicates that the features used could be optimal for predicting knee angles during treadmill walking. The pipeline was also found to be able to generate a consistent model with variation between generated models for the same participant to be 2.5% from the mean model.

We believe that the predictive accuracy of this approach can be attributed to relating the time-series features (extracted and selected) from the IMU signals to the knee kinematics without user intervention while still being able to control how the features are sorted. This allows us to generate a good predictive model with a small sample of data. Our result is also comparable to the result presented in previous studies where Mundt et al. [33], Chen et al. [34], and Tan et al. [35] used deep learning approaches where the errors are on average 5°–10° with a recent study reporting similar errors of approximately 2°–3° [36]. Although we did not investigate other gait parameters such as stride time [37] in this study, we believe that our approach could produce similar results given the accuracy shown in our knee kinematics and the reported results in the literature of a similar approach. Our personalized model results generally showed lower errors than those studies but our generalized models with similar accuracy [38]. This indicates that the random forest model is a good alternative method for generating models for estimating knee angles. However, our approach has some limitations evident by examining the results.

The quality of the data and the subsequent extracted features are the limiting factors on how well the model can perform. For instance, in this study, the same feature set was used across all participants, and all participants had similar walking speeds and duration on the treadmill. The variation of the errors seen across different models (see Table IV) could have occurred if the feature set did not fully explain the variation in gait between individuals. This became more evident when the generalizability of the surrogate models was examined using the leave-one-out analysis (Table V). We found that three of the ten participants had RMSE over 10° with an R^2 value below 0.9. This led to the observed errors in the worst case shown in Fig. 6. Removing these outliers, the average rms error dropped to $4.75° \pm 0.9°$. This result is comparable to a similar study on overground walking [38], where they showed an rms error of $5.0° \pm 1.0°$ after removing outliers. Hence, this further supports the notion that the feature set may not have

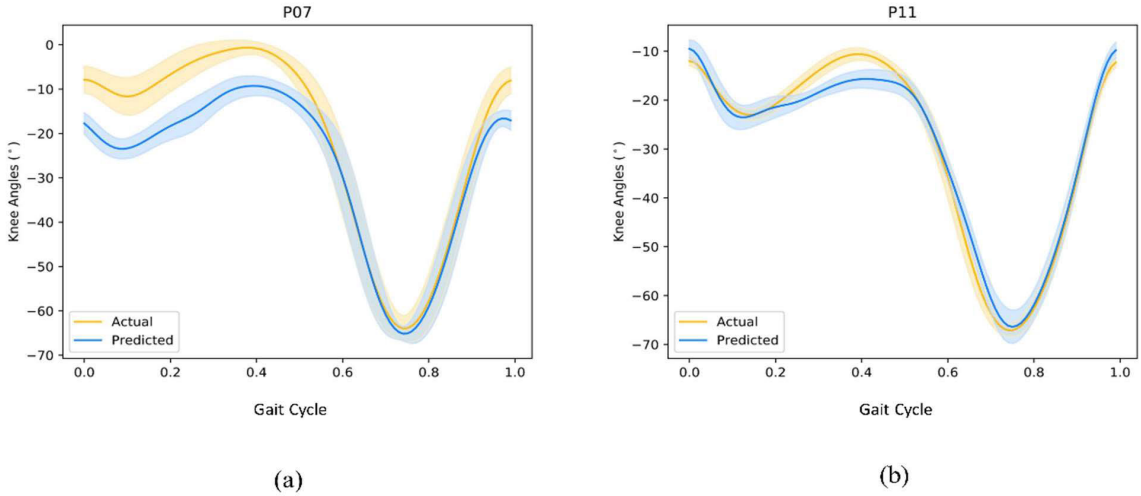


Fig. 6. Leave-one-out results showing (a) worst case and (b) best case across the gait cycle (heel strike to heel strike). The yellow and blue bands represent the standard deviation [27].

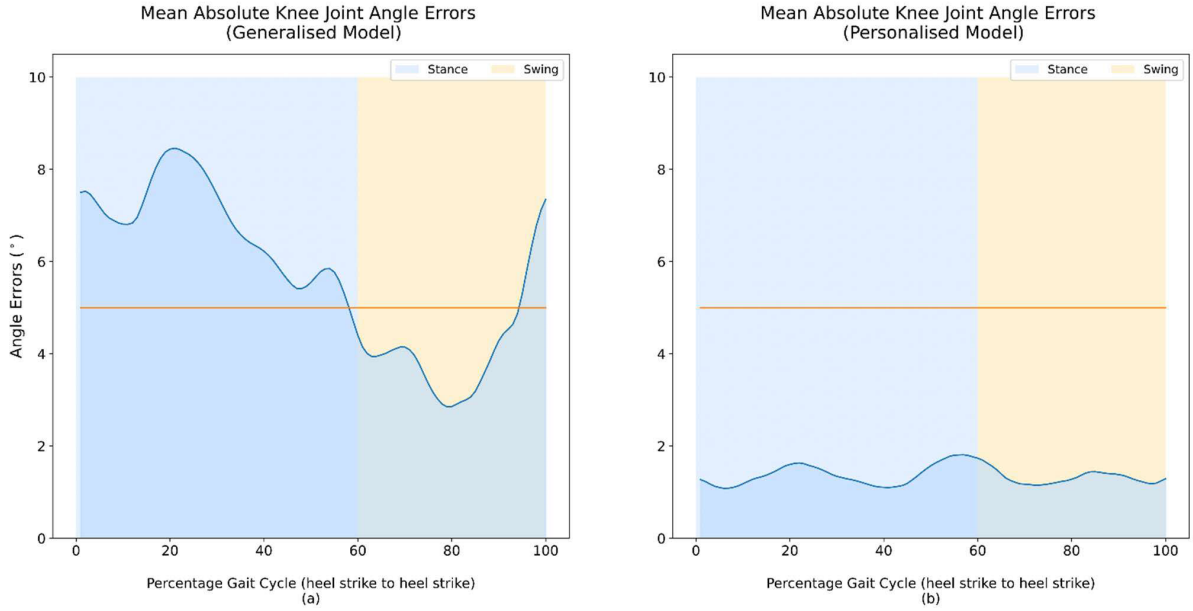


Fig. 7. Absolute difference between the OMC derived joint angles and the models predicted joint angles. (a) Generalized model. (b) Personalized model. The orange line is the 5° threshold for clinical measurements [28].

the right features for the variation seen in the dataset or the training dataset was not diverse enough to capture the variation in gait between individuals, especially during the stance phase, as shown in Fig. 7. This problem is similar to deep learning methods such as artificial neural network (ANN) [33]. In these cases, some authors will use stimulated data to increase the sampled data and add more variability [33], and this might also work for our presented approach. Another method we could use to consider variability between individuals is to create additional channels by combining existing data channels to generate new features, i.e., similar to the way tools such as TSFuse [39]. This was shown to aid in improving model prediction for gait parameters [27], [40].

The computational resource required to extract features is another limitation of this method. The presented method requires a lot of computing resources (i.e., memory space, at least 32 GB of RAM) to compute the features and subsequently train a nonlinear regression model and it is not always possible to have access to a high-powered computer. However, the feature selection and model training steps are only required to generate the model. Once the model was generated, it was possible to run the model on a low-powered computer using a small feature set, for example, Microsoft surface pro tablet with an i5 and 8 Gb of RAM [14].

The number of features chosen could be a limitation as the accuracy of the resulting model depends on the data used for training and the features used in the model. For this study, 100 features were selected as they gave good accuracy for the personalized models. The same feature set is then used to train our generalized models. However, this could have led to overfitting and the subsequent errors seen in some of our generalized models. For future work, it is proposed that an analysis will need to be performed on the features extracted by tsFresh and determine how the importance of these features changes with increasing variation in the data, i.e., walking speed [40]. As another piece of future work, this method will

be extended to estimating angles from different joints of the lower limb and using other wearable sensing technologies such as electromyography (EMG) to investigate muscle activity while walking [27].

Although our method showed results were similar to previous IMU studies (see Argent et al. [16] study), our method also does not require anthropometric variables when training a machine learning model. In addition, we can use a single IMU to estimate knee flexion, a technique that traditional IMU methods (as described in studies [5] and [41]) would not be able to perform while still achieving clinically accurate results with root mean square errors of $<5°$. This method is similar to deep learning models.

We believe that the presented method also has several other benefits when compared to the traditional method of obtaining joint angles from IMUs. In conventional methods, the orientation of the IMU and its relative orientation to the body must be calculated. Since this is computed from angular velocity, numerical errors, such as drift, will reduce the accuracy of the estimated joint angles. Methods, such as the one described by Nazarahari and Rouhani [4], could estimate a calibration transformation from IMU-to-body anatomical coordinate system by using the gravity vector, principal component analysis (PCA), and straight-line walking. However, these methods require the user to perform the calibration motion every time the IMU is placed on the individual, as the quality of the resulting estimation relies heavily on the quality of the calibration. This is especially true for the method presented by Seel et al. [42] and Muller et al. [43] where the user requires balance on one leg and perform a series of dynamic motion. This could be difficult to implement depending on the application, e.g., an elderly patient using the IMU as part of their rehabilitation-monitoring regime. The presented method could also solve this issue by personalizing and tuning the models at a clinical visit, where "ground truth" kinematics can be determined using either an OMC or a clinic-based IMU system. Then, the subsequent generated models could then be used to analyze gait in the real world as a remote monitoring system.

V. Conclusion

Accurate estimation of knee angles using only a shank-mounted IMU is possible from a personalized surrogate model trained using time-series features extracted by tsFresh. In addition, a generalized model was shown to be possible even though the present errors are high. This is due to the study's limitations, such as sample size and lack of variation in gait on the treadmill. Solving this will require more data with greater gait variation to create a robust model for predicting knee kinematics for individuals not represented in the training data.

Acknowledgment

We would like to acknowledge the funding support from National Science Challenges: Science for Technological Innovation NZ.

References

[1] K. Lebel, P. Boissy, H. Nguyen, and C. Duval, "Inertial measurement systems for segments and joints kinematics assessment: Towards an understanding of the variations in sensors accuracy," *Biomed. Eng. OnLine*, vol. 16, no. 1, pp. 1–16, May 2017, doi: 10.1186/s12938-017-0347-6.

[2] L. Vargas-Valencia, A. Elias, E. Rocon, T. Bastos-Filho, and A. Frizera, "An IMU-to-body alignment method applied to human gait analysis," *Sensors*, vol. 16, no. 12, p. 2090, Dec. 2016, doi: 10.3390/s16122090.

[3] T. McGrath, R. Fineman, and L. Stirling, "An auto-calibrating knee flexion-extension axis estimator using principal component analysis with inertial sensors," *Sensors*, vol. 18, no. 6, p. 1882, Jun. 2018, doi: 10.3390/s18061882.

[4] M. Nazarahari and H. Rouhani, "Semi-automatic sensor-to-body calibration of inertial sensors on lower limb using gait recording," *IEEE Sensors J.*, vol. 19, no. 24, pp. 12465–12474, Dec. 2019, doi: 10.1109/JSEN.2019.2939981.

[5] M. Al Borno et al., "OpenSense: An open-source toolbox for inertial-measurement-unit-based measurement of lower extremity kinematics over long durations," *J. NeuroEng. Rehabil.*, vol. 19, no. 1, p. 2021, Dec. 2022, doi: 10.1186/s12984-022-01001-x.

[6] C. A. Bailey, T. K. Uchida, J. Nantel, and R. B. Graham, "Validity and sensitivity of an inertial measurement unit-driven biomechanical model of motor variability for gait," *Sensors*, vol. 21, no. 22, p. 7690, Nov. 2021, doi: 10.3390/S21227690.

[7] A. Baudet et al., "Cross-talk correction method for knee kinematics in gait analysis using principal component analysis (PCA): A new proposal," *PLoS ONE*, vol. 9, no. 7, Jul. 2014, Art. no. e102098, doi: 10.1371/journal.pone.0102098.

[8] N. P. Brouwer, T. Yeung, M. F. Bobbert, and T. F. Besier, "3D trunk orientation measured using inertial measurement units during anatomical and dynamic sports motions," *Scandin. J. Med. Sci. Sports*, vol. 31, no. 2, pp. 358–370, Feb. 2021, doi: 10.1111/sms.13851.

[9] S. O. H. Madgwick, A. J. L. Harrison, and A. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, Zurich, Switzerland, Jun./Jul. 2011, pp. 1–7, doi: 10.1109/ICORR.2011.5975346.

[10] Z. Zhang, W. C. Wong, and J. Wu, "Wearable sensors for 3D upper limb motion modeling and ubiquitous estimation," *J. Control Theory Appl.*, vol. 9, no. 1, pp. 10–17, Feb. 2011, doi: 10.1007/s11768-011-0234-9.

[11] A. Sakai, Y. Tamura, and Y. Kuroda, "An efficient solution to 6DOF localization using unscented Kalman filter for planetary rovers," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2009, pp. 4154–4159, doi: 10.1109/IROS.2009.5354677.

[12] E.-H. Shin and N. El-Sheimy, "An unscented Kalman filter for in-motion alignment of low-cost IMUs," in *Proc. PLANS Position Location Navigat. Symp.*, Oct. 2004, pp. 273–279, doi: 10.1109/PLANS.2004.1309005.

[13] C. Dindorf, W. Teufl, B. Taetz, G. Bleser, and M. Fröhlich, "Interpretability of input representations for gait classification in patients after total hip arthroplasty," *Sensors*, vol. 20, no. 16, p. 4385, Aug. 2020, doi: 10.3390/s20164385.

[14] A. W. Kempa-Liehr, J. Oram, A. Wong, M. Finch, and T. Besier, "Feature engineering workflow for activity recognition from synchronized inertial measurement units," in *Proc. Pattern Recognit., ACPR Workshops*, Auckland, New Zealand, Nov. 2019, pp. 223–231.

[15] E. Halilaj, A. Rajagopal, M. Fiterau, J. L. Hicks, T. J. Hastie, and S. L. Delp, "Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities," *J. Biomech.*, vol. 81, pp. 1–11, Nov. 2018, doi: 10.1016/j.jbiomech.2018.09.009.

[16] R. Argent, S. Drummond, A. Remus, M. O'Reilly, and B. Caulfield, "Evaluating the use of machine learning in the assessment of joint angle using a single inertial sensor," *J. Rehabil. Assistive Technol. Eng.*, vol. 6, Jan. 2019, Art. no. 205566831986854, doi: 10.1177/2055668319868544.

[17] M. Christ, N. Braun, J. Neuffer, and A. W. Kempa-Liehr, "Time series FeatuRe extraction on basis of scalable hypothesis tests (tsfresh—A Python package)," *Neurocomputing*, vol. 307, pp. 72–77, Sep. 2018, doi: 10.1016/j.neucom.2018.03.067.

[18] A. A. Slater, T. J. Hullfish, and J. R. Baxter, "The impact of thigh and shank marker quantity on lower extremity kinematics using a constrained model," *BMC Musculoskelet. Disord.*, vol. 19, no. 1, pp. 1–10, Nov. 2018, doi: 10.1186/S12891-018-2329-7/FIGURES/5.

[19] A. Wong and R. Vallabh. (2018). *Sensor Specification*. IMeasureU. Accessed: Feb. 9, 2022. [Online]. Available: https://imeasureu.com/wp-content/uploads/2018/05/Sensor_Specification_v1.5.pdf

[20] T. Yeung, "Marker set (46 makers)," Univ. Auckland, Auckland, Tech. Rep., Apr. 2023. [Online]. Available: https://wiki.auckland.ac.nz/display/ABI/Basic+Mocap+Protocols, doi: 10.17608/k6.auckland.21588141.v1.

[21] T. F. Besier, D. L. Sturnieks, J. A. Alderson, and D. G. Lloyd, "Repeatability of gait data using a functional hip joint centre and a mean helical knee axis," *J. Biomech.*, vol. 36, pp. 1159–1168, Aug. 2003, doi: 10.1016/S0021-9290(03)00087-3.

[22] S. L. Delp et al., "OpenSim: Open-source software to create and analyze dynamic simulations of movement," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 11, pp. 1940–1950, Nov. 2007, doi: 10.1109/TBME.2007.901024.

[23] J. Zhang et al., "The MAP client: User-friendly musculoskeletal modelling workflows," in *Proc. Int. Symp. Biomed. Simulation*, 2014, pp. 182–192.

[24] D. Bakke and T. Besier, "Shape model constrained scaling improves repeatability of gait data," *J. Biomech.*, vol. 107, Jun. 2020, Art. no. 109838, doi: 10.1016/j.jbiomech.2020.109838.

[25] S. Schreven, P. J. Beek, and J. B. J. Smeets, "Optimising filtering parameters for a 3D motion analysis system," *J. Electromyogr. Kinesiol.*, vol. 25, no. 5, pp. 808–814, Oct. 2015, doi: 10.1016/j.jelekin.2015.06.004.

[26] A. Cantamessa, "Evaluating outcome following knee arthroplasty using inertial measurement units," M.S. thesis, Université de Liège, Liège, Belgium, 2020. [Online]. Available: https://matheo.uliege.be/handle/2268.2/10555

[27] H. Zarshenas, B. P. Ruddy, A. W. Kempa-Liehr, and T. F. Besier, "Ankle torque forecasting using time-delayed neural networks," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 4854–4857.

[28] C. Tang, D. Garreau, and U. von Luxburg, "When do random forests fail?" in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 31, 2018, pp. 2987–2997.

[29] T. Yeung, "Windowing," Univ. Auckland, Auckland, New Zealand, Nov. 2022. Accessed: Apr. 25, 2023. [Online]. Available: https://auckland.figshare.com/articles/figure/Windowing/21588204/1, doi: 10.17608/k6.auckland.21588204.v1.

[30] T. Yeung, "Features engineering," Univ. Auckland, Auckland, New Zealand, Nov. 2022. Accessed: Apr. 25, 2023. [Online]. Available: https://auckland.figshare.com/articles/figure/features_engineering/21588207/1, doi: 10.17608/k6.auckland.21588207.v1.

[31] T. Yeung, "Best and worst case," Univ. Auckland, Auckland, New Zealand, Nov. 2022. Accessed: Apr. 25, 2023. [Online]. Available: https://auckland.figshare.com/articles/figure/best_and_worst_case/21588279/1, doi: 10.17608/k6.auckland.21588279.v1.

[32] T. Yeung, "ML knee joint angle estimation errors—General vs personal," Univ. Auckland, Auckland, New Zealand, Nov. 2022. Accessed: Apr. 25, 2023. [Online]. Available: https://auckland.figshare.com/articles/figure/ML_Knee_joint_angle_estimation_errors_-_General_vs_Personal/21597144/1, doi: 10.17608/k6.auckland.21597144.v1.

[33] M. Mundt et al., "Estimation of gait mechanics based on simulated and measured IMU data using an artificial neural network," *Frontiers Bioeng. Biotechnol.*, vol. 8, p. 41, Feb. 2020, doi: 10.3389/fbioe.2020.00041.

[34] S. Chen, S. S. Bangaru, T. Yigit, M. Trkov, C. Wang, and J. Yi, "Real-time walking gait estimation for construction workers using a single wearable inertial measurement unit (IMU)," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2021, pp. 753–758, doi: 10.1109/AIM46487.2021.9517592.

[35] J.-S. Tan et al., "Predicting knee joint kinematics from wearable sensor data in people with knee osteoarthritis and clinical considerations for future machine learning models," *Sensors*, vol. 22, no. 2, p. 446, Jan. 2022.

[36] W. Teufl, M. Miezal, B. Taetz, M. Fröhlich, and G. Bleser, "Validity of inertial sensor based 3D joint kinematics of static and dynamic sport and physiotherapy specific movements," *PLoS ONE*, vol. 14, no. 2, Feb. 2019, Art. no. e0213064, doi: 10.1371/journal.pone.0213064.

[37] R. Jain, V. B. Semwal, and P. Kaushik, "Stride segmentation of inertial sensor data using statistical methods for different walking activities," *Robotica*, vol. 40, no. 8, pp. 2567–2580, Aug. 2022, doi: 10.1017/S026357472100179X.

[38] B. Oubre et al., "A simple low-cost wearable sensor for long-term ambulatory monitoring of knee joint kinematics," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 12, pp. 3483–3490, Dec. 2020.

[39] A. De Brabandere, P. Robberechts, T. O. D. Beeeck, and J. Davis, "Automating feature construction for multi-view time series data," in *Proc. ECMLPKDD Workshop Automating Data Sci.*, 2019, pp. 1–16.

[40] H. Zarshenas, *EMG-Informed Estimation of Human Walking Dynamics for Assistive Robots*. Auckland, New Zealand: Univ. Auckland, 2022.

[41] C. Monoli, J. F. Fuentez-Perez, N. Cau, P. Capodaglio, M. Galli, and J. A. Tuhtan, "Land and underwater gait analysis using wearable IMU," *IEEE Sensors J.*, vol. 21, no. 9, pp. 11192–11202, May 2021.

[42] T. Seel, J. Raisch, and T. Schauer, "IMU-based joint angle measurement for gait analysis," *Sensors*, vol. 14, no. 4, pp. 6891–6909, Apr. 2014, doi: 10.3390/s140406891.

[43] P. Muller, M.-A. Begin, T. Schauer, and T. Seel, "Alignment-free, self-calibrating elbow angles measurement using inertial sensors," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 2, pp. 312–319, Mar. 2017, doi: 10.1109/JBHI.2016.2639537.

Ted Yeung received the Ph.D. degree in bioengineering from the University of Auckland, Auckland, New Zealand, in 2017.

He is an Emerging Researcher with extensive research experience in 3-D gait analysis and wearable sensor technology. He collaborated with the Fraunhofer Institute, Stuttgart, Germany, from 2017 to 2019, as a Research Fellow looking at using wearable sensors for exoskeleton control, and he worked in industry for VICON IMeasureU, Auckland, as a Research Engineer.

Dr. Yeung recently present a talk at the International Society of Biomechanics (ISB) Congress 2021, where he won the ISB and APAB Abstract Award for the Best Asian Pacific Association for Biomechanics abstracts submitted to the 2021 ISB Congress. In 2020, he returned to academia and joined the Musculoskeletal Modeling Group, Auckland Bioengineering Institute, Auckland, via a seed grant from National Science Challenge—Science for Technological Innovation (awarded in 2020).

Astrid Cantamessa received the bachelor's and master's degrees in biomedical engineering from the University of Liège, Liège, Belgium, in 2018 and 2020, respectively. She is pursuing Ph.D. degree with the Mechanics of Biological and Bioinspired Materials Laboratory.

Ms. Cantamessa received an FRIA Grant to work on the biological and biomechanical characterization of the cement lines in healthy and osteoporotic bones. Her master's thesis and internship were conducted at the Auckland Bioengineering Institute in New Zealand, a world-leading research institute. The aim was to develop and assess a new workflow based on machine learning to quantitatively evaluate the joint function of patients following knee arthroplasty.

Andreas W. Kempa-Liehr (Member, IEEE) received the Ph.D. degree from the University of Münster, Münster, Germany, in 2004.

From 2009 to 2016, he was working in different data science roles at EnBW Energie Baden-Württemberg AG and Blue Yonder GmbH, Karlsruhe, Germany. He is a Senior Lecturer with the Department of Engineering Science, The University of Auckland, New Zealand, and an Associate Member with the Freiburg Materials Research Center (FMF), University of Freiburg, Germany. He continued his research as the Head of the Service Group Scientific Information Processing, FMF.

Thor Besier received the Ph.D. degree in musculoskeletal biomechanics from The University of Western Australia, Perth, WA, Australia, in 2000.

He was a Postdoctoral Fellow with the Bioengineering Department, Stanford University, Stanford, CA, USA, from 2003 to 2006. He is a Professor with the Auckland Bioengineering Institute, Auckland, New Zealand, and has a joint appointment with the Department of Engineering Science, University of Auckland, Auckland. He established the Stanford's Human Performance Laboratory, Stanford, as the Director of Research, where he was a Faculty Member with the Department of Orthopedics from 2006 to 2010. He is a world Leader in musculoskeletal modeling and biomechanics and will lead the atlas-based model generation.

Julie Choisne received the M.Sc. degree in mechanical engineering from Leonardo da Vinci Engineering School, Courbevoie, France, in 2009, and the Ph.D. degree in biomechanics from Old Dominion University, Norfolk, VA, USA.

She is a Researcher with extensive research experience in computational modeling and bone biomechanics. She also has a strong track record of successful grant achievements and awards.

Dr. Choisne was recently awarded the Aotearoa Fellowship, a One Million NZD Grant to create a personalized digital twin representing the pediatric population, Digi Mini Me associated with the recently funded MBIE catalyst grant "12 Labours." Her notable achievements include the Mechanical Engineering Faculty Award for Best Doctoral Thesis in 2013, the MedTech CoRE seed and ART grants in 2015 and 2018, the Science for Technological Innovation (SfTI) seed grants in 2017 and 2020, and the Prestigious HRC Emerging Researcher First Grant in 2019.