2023 IEEE International Conference on Robotics and Automation (ICRA 2023)
May 29 - June 2, 2023. London, UK

Anticipation and Delayed Estimation of Sagittal Plane Human Hip Moments using Deep Learning and a Robotic Hip Exoskeleton

Dean D. Molinaro, *Student Member*, *IEEE*, Ethan O. Park, and Aaron J. Young, *Member*, *IEEE*

***Abstract*— Estimating human joint moments using wearable sensors has utility for personalized health monitoring and generalized exoskeleton control. Data-driven models have potential to map wearable sensor data to human joint moments, even with a reduced sensor suite and without subject-specific calibration. In this study, we quantified the RMSE and R^2 of a temporal convolutional network (TCN), trained to estimate human hip moments in the sagittal plane using exoskeleton sensor data (i.e., a hip encoder and thigh- and pelvis-mounted inertial measurement units). We conducted three analyses in which we iteratively retrained the network while: 1) varying the input sequence length of the model, 2) incorporating noncausal data into the input sequence, thus delaying the network estimates, and 3) time shifting the labels to train the model to anticipate (i.e., predict) human hip moments. We found that 930 ms of causal input data maintained model performance while minimizing input sequence length (validation RMSE and R^2 of 0.141±0.014 Nm/kg and 0.883±0.025, respectively). Further, delaying the model estimate by up to 200 ms significantly improved model performance compared to the best causal estimators ($p<0.05$), improving estimator fidelity in use cases where delayed estimates are acceptable (e.g., in personalized health monitoring or diagnoses). Finally, we found that anticipating hip moments further in time linearly increased model RMSE and decreased R^2 ($p<0.05$); however, performance remained strong ($R^2>0.85$) when predicting up to 200 ms ahead.**

I. Introduction

Estimating human joint moments using wearable sensors is a critical component in enabling real-time, personalized health monitoring and generalized exoskeleton control. Currently, human joint moments are computed *post hoc* using inverse dynamics models informed by motion capture and ground reaction force (GRF) measurements collected from stationary, in-lab equipment [1]. This method enables in-lab biomechanical analyses of isolated movements; however, it lacks the modularity to enable personalized analyses, such as long-term health monitoring of daily mobility. Further, recent research in lower-limb exoskeletons has demonstrated that these devices may be used to further improve human mobility in both able-body and clinical populations [2]–[4]. Currently, these systems are largely constrained to well-defined ambulatory contexts (e.g., constant speed treadmill walking), as they rely on discrete modes of operation to provide predefined assistance trajectories. Real-time estimates of the user's joint moments could enable broadly generalizable exoskeleton controllers effective at assisting the user "in-the-wild," as the exoskeleton could provide assistance based on the user's physiological state instead of relying on hand-engineered parameterizations of human gait [5].

Though estimating human joint moments using wearable sensors could provide major benefits in health monitoring and exoskeleton control, computing these values analytically is challenging. One approach is to use an inverse dynamics model informed by wearable sensors opposed to conventional motion capture and GRF measurement systems [6], [7]. For instance, human joint kinematics can be measured by electrogoniometers or encoders [8] or estimated using accelerometers and gyroscopes placed on each lower-limb segment [7], while GRFs can be measured using instrumented footwear [6], [7], [9]. Alternatively, previous research has also investigated the use of forward dynamics models to estimate human joint moments based on measured muscle activations and musculoskeletal models [10]–[13]. While these methods often generalize well across human movements, they require complicated and often cumbersome measurement devices, subject-specific calibration, and detailed setup, largely limiting their practicality.

Opposed to analytical methods, machine learning models trained using supervised learning can approximate the mapping between wearable sensors and human joint moments [14]–[18]. Given the nature of data-driven approaches, these estimators can accurately estimate human joint moments with a reduced number of sensors compared to analytical methods. Further, machine learning models designed to estimate gait parameters can leverage patterns in the time history of the input sensors to further improve model performance. This concept was initially implemented using hand-engineered features of raw time series data [16], [19], [20] and since has progressed into latent representations learned by the model using recurrent and convolutional neural networks [17], [18]. For instance, in our previous work, we demonstrated that a temporal convolutional network (TCN) could accurately estimate human joint moments during level ground, ramp, and stair walking using input trajectories from an encoder and thigh- and trunk-mounted inertial measurement units (IMUs) [18], outperforming alternative deep learning models.

While our previous work optimized a purely causal hip moment estimator, it is likely that removing the requirement of causality in the model input sequence (i.e., inducing delay in the estimate) would further improve model performance. In

This work was supported in part by the National Science Foundation Graduate Research Fellowship Program under Grant Number DGE-2039655, in part by the NSF NRI under Award Number 1830215, and in part by the NSF SURE Robotics REU under Award Number 1757401. *(Corresponding Author: Dean Molinaro)*.

Dean D. Molinaro and Aaron J. Young are with the George W. Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30309 USA and with the Institute for Robotics and Intelligent Machines (IRIM), Georgia Institute of Technology, Atlanta, GA 30309 USA (e-mail: dmolinaro3@gatech.edu; aaron.young@me.gatech.edu).

Ethan O. Park is with the Grainger College of Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61820 USA (e-mail: ethanop2@illinois.edu).

this case, applications that are not penalized by short delays in joint moment estimates but may heavily rely on the detection of small changes in mobility (e.g., health monitoring) could greatly benefit from a noncausal implementation of a joint moment estimator. Thus, in this study, we analyzed the effects of delaying the hip flexion/extension moment estimates with respect to the model input sequence using a TCN. We hypothesized that a noncausal estimator (i.e., one that generates an estimate $\hat{y}_{t-\Delta t}$ corresponding to a time Δt before the latest value in the input sequence $x_{0:t}$) would improve model performance in R^2 and RMSE compared to a purely causal estimator, as the kinematic input sequence after-the-fact should further allow the estimator to approximate the underlying dynamics of the system.

In contrast, a model that anticipates (i.e., predicts) the human's joint moments may have benefits in other contexts. For instance, to promote a seamless interaction between an exoskeleton and human user, the exoskeleton controller could be designed such that the controller commands assistance based on the user's anticipated movements (i.e., the user's intent) [21], [22]. In this case, anticipating the user's joint moments *a priori* may allow the exoskeleton to better transition with changes in the user's gait (e.g., during ambulation mode transitions) and could have general benefits from a controls perspective. In our previous work, we identified the effect of anticipation time on joint moment estimation; however, the model used kinematic and electromyographic data collected across the entire limb and the model was trained using subject-specific data [22]. In this study, we have extended this analysis by evaluating the effect of anticipation time on hip moment estimation when trained on a subject-independent basis using kinematic sensor data localized around the hip joint.

Thus, this study comprises a detailed analysis to extend the capability of our previously optimized subject-independent hip moment estimator. Specifically, we have 1) quantified the potential benefits of extending the input sequence of the TCN, 2) analyzed the potential improvement in model performance from noncausal input sequences, and 3) extended our previous work to understand the potential of this system to anticipate future hip flexion/extension moments. Additionally, the analyses in this study were conducted using input sensor data recorded from a robotic hip exoskeleton, further verifying the real-world viability of the model. Therefore, this study provides critical information for the future design of data-driven joint moment estimators applicable in multiple domains, including exoskeleton control and personalized health monitoring, using a reduced sensor suite and without the need for subject-specific calibration.

II. Experimental Data Collection

A. Protocol & Measurements

Nine able-body participants (5 males, 4 females, height of 1.75±0.07 m, body mass of 70.4±8.7 kg, age of 22±3 years) participated in this study. Each participant provided written, informed consent according to the protocol approved by the Georgia Tech Institutional Review Board for this study. During the protocol, each participant completed 20 walking bouts over level ground, ramp inclines/declines of ±7.8°, ±9.2°, ±11.0°, and ±12.4°, and stairs of height 10.2 cm, 12.7 cm, 15.2 cm, and 17.8 cm (Fig. 1a). Inclines and stair heights were set based on available presets. During each trial, motion capture data were collected at 200 Hz using a 36-camera Vicon motion capture system (Oxford Metric, Oxford, U.K.), and 6-axis GRF data were collected at 1000 Hz using 10 Bertec force plates (Bertec, Columbus, Ohio, USA). Based on the force plate configuration, each walking bout recorded either 6 level ground strides, 2 ramp ascent and descent strides, or 5 stair ascent and descent strides.

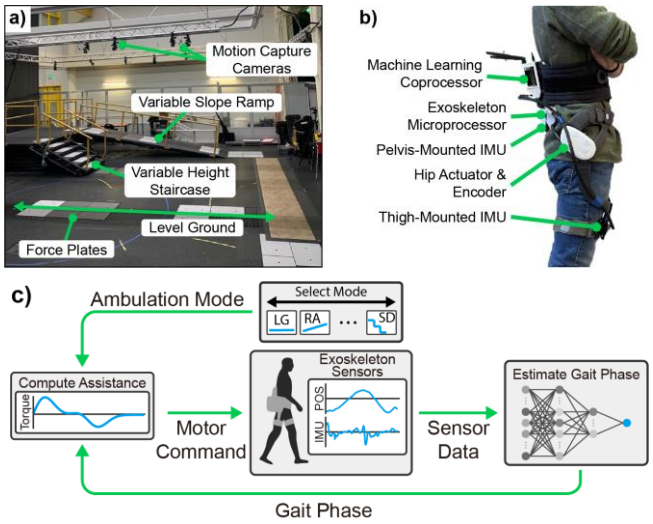

Fig. 1. a) The experimental setup is shown. Participants walked over level ground, ramps, and stairs while motion capture and force plate data were collected. b) The robotic hip exoskeleton is shown. The exoskeleton assisted the participants while walking and recorded wearable sensor data. c) A block diagram of the exoskeleton controller is shown. Commanded torque was computed as a function of gait phase and ambulation mode.

B. Robotic Hip Exoskeleton

While completing each trial, participants wore the Gait Enhancing and Motivating System (GEMS), a robotic hip exoskeleton developed by Samsung Electronics (Suwonsi, Gyeonggi-do, South Korea) [23], able to provide a peak torque of 12 Nm to the user's hips in the sagittal plane (Fig. 1b). The exoskeleton controller ran at 200 Hz and recorded sagittal hip encoder and pelvis-mounted IMU data. An onboard coprocessor (Jetson Nano, NVIDIA, CA, USA) was also added to the exoskeleton to provide real-time gait phase estimates using a convolutional neural network developed and validated in our previous work [24] and to record data from two MPU-9250 IMUs (TDK, CA, USA) mounted on the thigh struts. The coprocessor (server) provided gait phase estimates to the exoskeleton (client) asynchronously using a wired connection and TCP/IP. In a separate process, the coprocessor also recorded the thigh IMU data at 200 Hz. The thigh IMU data measured by the coprocessor were resampled using linear interpolation based on the exoskeleton timestamps to align the coprocessor data with the exoskeleton data. The total mass of the exoskeleton and coprocessor system was 3.7 kg.

During the walking bouts, the exoskeleton provided hip flexion/extension assistance to the user based on predefined spline trajectories that were a function of gait phase and ambulation mode. The splines were parameterized using six nodes corresponding to flexion assistance onset, peak flexion assistance, flexion assistance offset, extension assistance onset, peak extension assistance, and extension assistance

offset, computed using piecewise cubic hermite interpolating polynomials. To collect a diverse dataset, two assistance trajectories were used per ambulation mode: 1) a trajectory designed to mimic the average biological hip flexion/extension moment for each ambulation mode based on the results from Camargo *et al.* [25] and 2) a trajectory based on previously optimized assistance trajectories [26], [27] and further tuned based on the preference of the experimenters during initial piloting. Each subject completed 10 walking bouts for each condition using each controller (20 total walking bouts per ambulatory condition). During exoskeleton operation, an experimenter manually selected the ambulation mode and controller type from a remote connection.

C. Ground-Truth Hip Moment Labeling

The opensource, musculoskeletal modeling software, OpenSim [28], [29], was used to compute ground-truth joint moments from the motion capture and GRF data. The Gait2392 lower-limb model was first scaled to each participant's anthropometry using motion capture data from a static trial while the participant did not wear the exoskeleton. Since markers around the pelvis were occluded by the hip exoskeleton, a second static trial was collected while the participant wore the exoskeleton with the pelvis markers placed on the exoskeleton. The subject-scaled OpenSim model was then rescaled using the second static trial without modifying the segment properties of the model (i.e., only adjusting model pose and marker locations). Finally, the mass of the exoskeleton was manually added to the pelvis of the scaled OpenSim model as a point mass. Thus, this scaling approach provided an OpenSim model scaled to the participant's anthropometry but with marker locations that matched those when wearing the exoskeleton.

Using the scaled OpenSim model, subject joint kinematics and moments were computed using the OpenSim Inverse Kinematics and Inverse Dynamics Tools, respectively. Before running these tools, the motion capture and GRF data were lowpass filtered using a zero-lag, Butterworth filter with 6 and 20 Hz cutoff frequencies, respectively. Additionally, the resulting joint kinematics and moments were filtered using a zero-lag, Butterworth filter with a 6 Hz cutoff frequency. The hip flexion/extension moments were then normalized to each participant's body mass and used as the label when training and evaluating the hip moment estimators in this study.

III. Hip Moment Estimation & Analyses

A. Hip Moment Estimator

In this study, we implemented the hip moment estimator using a temporal convolutional network (TCN) [18], [30], [31]. TCNs leverage 1D convolutional layers to learn latent feature representations of time series input data. Further, by exponentially dilating the kernel through the model layers, TCNs can explicitly input long sequences of data with a relatively small number of learnable parameters. Given these benefits, TCNs have yielded competitive performance in many sequence modeling problems [18], [31]. For instance, our previous work demonstrated that the TCN significantly outperformed alternative neural network structures when estimating human biological hip moments [18]. Given these results, we implemented the TCN in this study using a similar framework. Specifically, the hip moment estimator in this study was trained using supervised learning and the ground-truth, mass normalized hip flexion/extension moments as the label. The input channels of the TCN were the exoskeleton ipsilateral encoder position and velocity and the 6-axis accelerometer and gyroscope data measured by the ipsilateral thigh-mounted and pelvis-mounted IMUs. Contralateral limb sensor data were not included to minimize model input dimensionality and to prevent overfitting to limb symmetries specific to healthy ambulation. Before being input to the neural network, the encoder velocity was also filtered using a 2nd order 10 Hz low pass filter to smooth the signal. To train the model for estimating both left and right hip moments, the dataset was split into two groups – one corresponding to the data required for estimating left side hip moments and the other corresponding to right side hip moments. To align the data, such that model inputs were agnostic to the corresponding side – the left leg input data was mirrored before being input to the model. This was done by negating the left thigh accelerometer data and rotating both the left thigh accelerometer and gyroscope data 180° along the medial/lateral axis. Additionally, the same transform was applied to the pelvis IMU data for left side estimates. The TCN was then trained by randomly combining both the right side and left side datasets, such that the model could output hip flexion/extension moment estimates regardless of side.

To validate the hip moment estimator, each condition was trained and evaluated using 9-fold leave-one-subject-out validation, in which the model was iteratively trained on 8 subjects in the dataset and validated on the hold-out subject. Thus, the results of our analyses are based on models trained without subject-specific data in the training set. Each model was trained using mean-squared-error loss with the Adam optimizer in PyTorch on the Georgia Tech Pace Computing Cluster. For each fold, the model was trained from random initialization for a minimum of 100 epochs and maximum of 500 epochs. Early stopping based on the validation subject loss was used to prevent model overfitting and reduce overall training time. The hyperparameters of the TCN were selected based on our previous optimization [18] (i.e., 50 channels per hidden layer, dropout probability of 0.3, and learning rate of 0.0005); however, the kernel size and number of network levels were reoptimized based on our analysis of model input sequence length (see Section III.B).

B. Analysis of Causal Input Sequence Length

Given the structure of the TCN, the model input sequence length is dictated by the kernel size, number of network levels (i.e., residual blocks), and dilation. As outlined by Bai *et al.* [31], we designed the TCN to have a constant kernel size for each layer of the network. Further, kernel dilation was exponentially increased with each level of the TCN, starting with a dilation of one. This strategy of dilation exponentially increases the input sequence length (i.e., receptive field) h of the model with a linear increase in the number of levels l of the network, computed as,

$$h = 1 + \sum_{i=0}^{l-1} 2(k-1)d_i, \quad (1)$$

where the dilation at level i is computed as

$$d_i = 2^i. \quad (2)$$

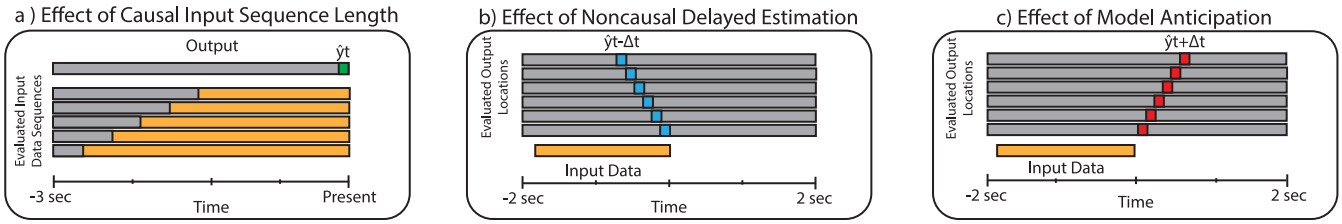

Fig. 2. Diagrams depicting the a) analysis of varying model input sequence length (input sequences depicted in orange), b) analysis of delaying model estimation (timing of estimate depicted in blue), and c) analysis of model anticipation (timing of prediction depicted in red) are shown.

In our previous optimization of the TCN hip moment estimator [18], we did not evaluate hyperparameter combinations resulting in input sequence lengths greater than 1000 ms due to constraints in the dataset. Given that the optimized model of our previous work resulted in an input sequence length of 930 ms, which was close to the maximum boundary, it is possible that increasing the network capacity to include a longer input sequence could further improve performance. Thus, we conducted an extended hyperparameter sweep including kernel sizes on the closed interval of 3 to 10 while also varying the number of network levels on the closed interval of 3 to 7 (Fig. 2a). In this analysis, we only excluded hyperparameter combinations that exceeded an input sequence length of 3000 ms, resulting in a total range in candidate input sequences from 140 to 2790 ms.

From the hyperparameter sweep, we reported the best validation RMSE and R^2 of the evaluated networks; however, since it was likely that several combinations of hyperparameters would have similar validation results to the best network but may have drastically different input sequence lengths, we introduced a second criteria in our analysis, the network of minimally sufficient input sequence length. This was defined as the network which had the smallest input sequence length that resulted in a RMSE within 3% of the best model. Since TCN input sequence length correlates with overall model size, the network of minimally sufficient input sequence length provides a practical context into which set of hyperparameters maintained performance while gaining common benefits of a smaller network (e.g., faster training/inference times and robustness to overfitting).

C. Analysis of Delayed Hip Moment Estimation

To quantify the effect of delaying the hip moment estimate with respect to the input sequence, we iteratively retrained the TCN from random initialization while decrementing the time instance of the corresponding hip moment labels when training the model (Fig. 2b). The total model delay was swept from 50 to 1000 ms in increments of 50 ms. To prevent the noncausal input data from overwriting important data points of the causal input sequence, we conducted this analysis on an expanded TCN (5 network levels with a kernel size of 7), resulting in a total input sequence length of 1860 ms. This expanded model had an input sequence twice the length of our previously optimized model [18], thus preventing the likelihood that key information in the causal input sequence would be lost when incorporating the noncausal input data. To determine the minimum delay that maximized model performance, we computed the maximally beneficial delay, defined as the minimal estimator delay resulting in a RMSE within 3% of the best result.

D. Analysis of Hip Moment Anticipation

We conducted an additional analysis to quantify the ability of the TCN to anticipate (i.e., predict) future hip moments. Specifically, the TCN was iteratively retrained from random initialization while the corresponding hip moment labels were shifted forward in time, thus training the model to anticipate future hip moments (Fig. 2c). The evaluated anticipation times spanned the closed interval of 50 ms to 1000 ms with increments of 50 ms. To interpret the results among both delayed estimation and anticipation, we again used the expanded TCN structure defined in Section III.C when running this analysis.

E. Statistical Analyses

All statistical tests were completed using MATLAB (Mathworks, Natick, MA, USA) with a 0.05 alpha level of significance. Statistical trends identifying the effects of delayed model estimation and model anticipation on validation RMSE and R^2 were conducted using linear regression. Separate linear regression models were fit for delays in model estimation less than and greater than the maximally beneficial delay (see Section III.C). Additionally, a one-way repeated measures ANOVA was used to identify a main effect among the validation RMSE and R^2 results of the best causal estimator, the causal estimator corresponding to the minimally sufficient input sequence length, and the noncausal estimator corresponding to the maximally beneficial delay. Finally, a *post hoc* multiple comparisons test with a Bonferroni correction was used to identify statistical differences among the three models.

IV. Results

A. Effects of Causal Input Sequence Length

The hyperparameter sweep produced TCN architectures with input sequence lengths ranging from 140 to 2790 ms (Fig. 3). The best validation RMSE was 0.138±0.013 Nm/kg, resulting from the architecture with an input sequence length of 2520 ms (6 network levels with kernel size of 5). The minimally sufficient network (i.e., the network with the smallest sequence length resulting in a RMSE within 3% of the best network) had a sequence length of 930 ms (5 network levels with kernel size of 4), resulting in a RMSE of 0.141±0.014 Nm/kg. Further, the best validation R^2 was 0.888±0.027, resulting from the model with input sequence length of 2790 ms (5 network levels with kernel size of 10), whereas the minimally sufficient model resulted in a R^2 of 0.883±0.025.

B. Effects of Delayed Hip Moment Estimation

The results of iteratively delaying the estimated hip moments with respect to the input data are shown in Fig. 4

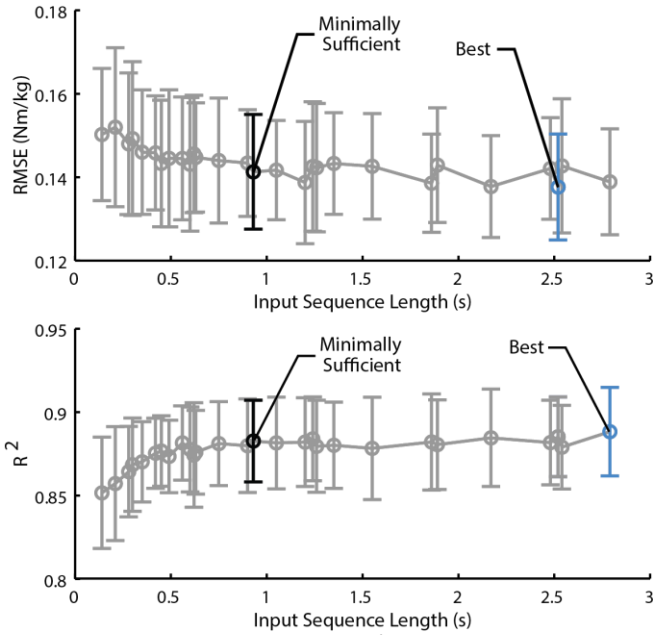

Fig. 3. The resulting RMSE (top) and R^2 (bottom) of the hip moment estimator are shown with respect to the model input sequence length. The input sequence length was varied by sweeping a range of kernel sizes and number of layers in the network. The blue data point indicates the sequence length resulting in the best model performance. The black data point represents the minimally sufficient sequence length computed as the smallest sequence length resulting in a RMSE within 3% of the minimum. Error bars represent ±1 standard deviation.

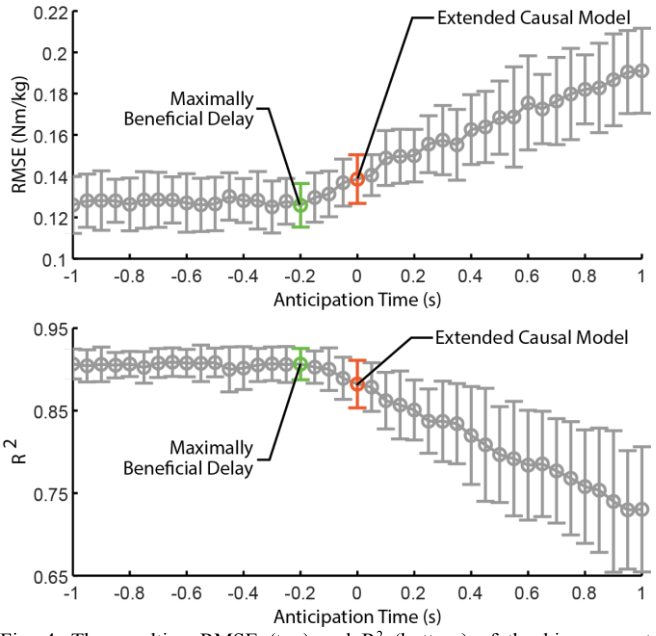

Fig. 4. The resulting RMSE (top) and R^2 (bottom) of the hip moment estimator are shown with respect to anticipation time. Positive anticipation times indicate model prediction of future hip moments. Negative anticipation times indicate the delay of the model estimate with respect to the input sequence. The red data point represents an anticipation time of zero (i.e., the TCN was trained to estimate hip moments at the same instance in time as the most recent input data). The green data point represents the maximally beneficial delay (i.e., the minimal amount of delay resulting in a RMSE within 3% of the best result). Error bars represent ±1 standard deviation.

and 5. The maximally beneficial delay (i.e., the resulting sequence length of noncausal data that minimized estimator delay but maintained a validation RMSE within 3% of the best result) was 200 ms, resulting in a RMSE and R^2 of 0.126±0.011 Nm/kg and 0.906±0.019, respectively (Fig. 4). The line of best fit within an anticipation time of 0 ms and the maximally beneficial delay (green trendline in Fig. 5) resulted in first order coefficients of 0.065 Nm/kg-s ($p=0.007$) and -0.124 /s ($p=0.015$) when fit to the validation RMSE and R^2 results, respectively. Further, the line of best fit between the maximally beneficial delay and maximum delay evaluated in this study (black trendline in Fig. 5) resulted in first order coefficients of -0.001 Nm/kg-s ($p=0.855$) and -0.001 /s ($p=0.930$) when fit to the validation RMSE and R^2, respectively.

Additionally, a significant effect was found when comparing the validation RMSE and R^2 results among the best causal estimator, the minimally sufficient causal estimator, and the noncausal estimator of maximally beneficial delay (one-way repeated measures ANOVA: $p<0.001$ for both RMSE and R^2). Further, the noncausal estimator significantly improved RMSE and R^2 compared to the causal estimators (multiple comparisons: $p<0.05$); however, no statistical differences between the two causal estimators were found.

C. *Effects of Hip Moment Anticipation*

Anticipating the hip flexion/extension moments diminished model performance (Fig. 4 and 5). Specifically, the line of best fit among positive anticipation times (i.e., from 0 to 1000 ms) resulted in a first order coefficient of 0.052 Nm/kg-s ($p<0.001$) when fit to the validation RMSE. Similarly, the first order coefficient of the linear regression model fit to the R^2 results was -0.154 /s ($p<0.001$).

V. Discussion

This study provides a multi-faceted analysis, in which we 1) quantified the effect of TCN input sequence lengths on hip moment estimation performance via an extended hyperparameter sweep of model kernel size and number of levels, 2) determined the benefits in model performance gained by leveraging noncausal sensor information via delayed estimation, and 3) quantified the trend in model performance when retraining the TCN to anticipate, or predict, the hip moments in time. Based on the results of the hyperparameter sweep, we found that the minimally sufficient input sequence length was 930 ms, meaning further increases in model input sequence length had little to no impact in the resulting validation RMSE. Further, we did not find any statistical differences in RMSE or R^2 between the best performing network and the model of minimally sufficient sequence length (multiple comparisons test: $p>0.05$).

Alternatively, delaying the model estimate by adding noncausal data to the input sequence posed an additional method for improving model performance. This could be especially beneficial in applications that can accept small delays in joint moment estimates (e.g., out-of-lab biomechanical analyses or monitoring daily effort). In this study, delaying the model estimates statistically improved model RMSE and R^2 by 0.065 Nm/kg-s and 0.124 /s when delaying the estimate by up to 200 ms (linear regression analysis: $p<0.05$). However, we found that further delaying the model estimate with respect to the input sequence beyond 200 ms did not yield further improvement to model performance (linear regression analysis: $p>0.05$). This is an important finding for applications of wearable sensor-based

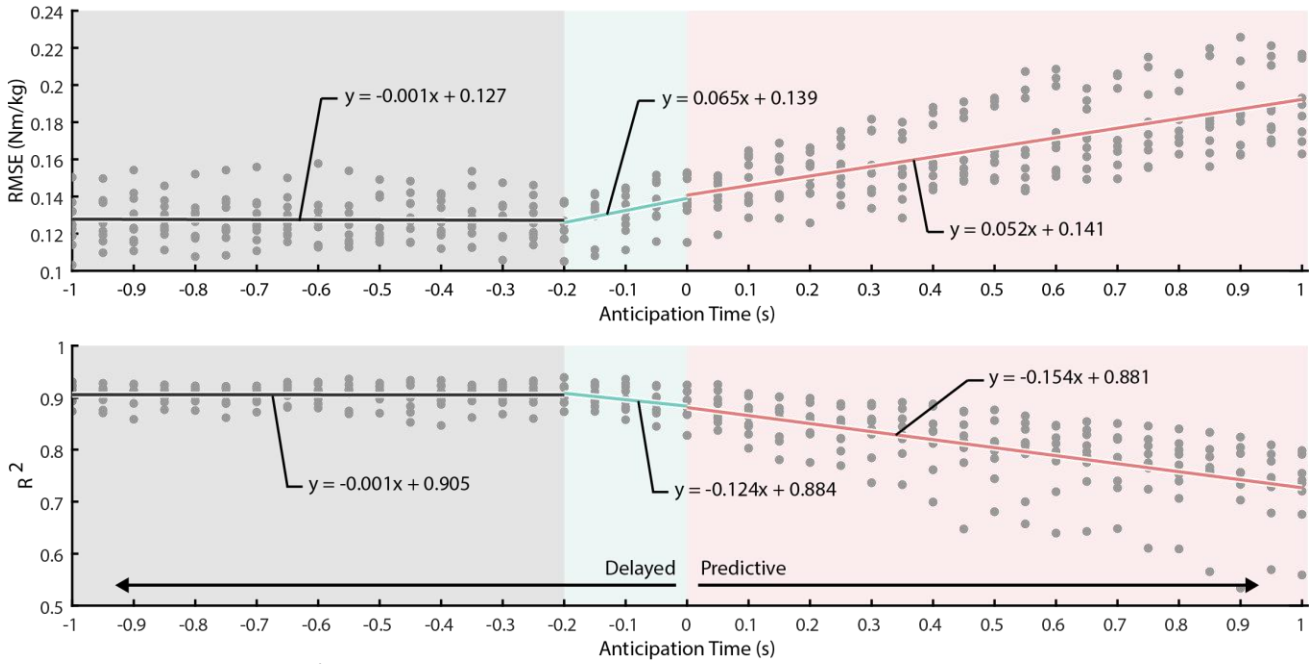

Fig 5. The resulting RMSE (top) and R^2 (bottom) of each model (grey data point) evaluated during the anticipation time analyses are shown. Linear regression models were fit to three intervals of anticipation time: 1) the interval between the maximum evaluated delay and maximally beneficial delay, shown in black; 2) the interval between the maximally beneficial delay and an anticipation time of zero, shown in green; and 3) the interval between an anticipation time of zero and the maximum evaluated anticipation time, shown in red.

joint moment estimators, able to accommodate small delays in the estimate. For instance, the ~10% improvement in RMSE of using the delayed estimator could improve the predictive power of a joint moment estimator deployed to diagnose the onset of mobility disorders in aging populations.

The results of the causal model of minimally sufficient input sequence length (RMSE and R^2 of 0.141±0.014 Nm/kg and 0.883±0.025, respectively) and the noncausal model of maximally beneficial delay (RMSE and R^2 of 0.126±0.011 Nm/kg and 0.906±0.019, respectively) were either similar to or better than previous subject-independent, data-driven hip moment estimators in the literature. For instance, our previous work of optimizing a subject-independent TCN to estimate biological hip moments using wearable sensors resulted in an average RMSE and R^2 of 0.131 Nm/kg and 0.880, respectively [18], but was trained and evaluated using simulated IMU data. Additionally, Dorschky *et al.* implemented an IMU-based convolutional neural network, with a hip moment estimation RMSE of 0.17 Nm/kg during walking [17]. Further, Forner-Cordero *et al.* reported a sagittal plane hip moment estimation RMSE of 0.15 Nm/kg using an inverse dynamics model informed by pressure insoles [6], indicating that the causal and noncausal estimators of this study can outperform analytical approaches, even with a reduced set of sensors.

As expected, anticipating hip moments further in time significantly reduced model performance (linear regression analysis: $p<0.05$). However, we found that predicting hip moments up to 200 ms into the future resulted in a validation RMSE below 0.150 Nm/kg and R^2 above 0.850, suggesting the model has strong predictive power within this anticipation window. Future analyses should explore the effects on joint moment anticipation when deployed online, especially when the estimated result could impact the resultant joint moments (e.g., in an exoskeleton controller that leads the user's joint dynamics based on the anticipated results); however, this analysis was outside the scope of this study.

This study contained multiple limitations. Although the models were trained using measured sensor data from an actuated exoskeleton, the performance of these models was evaluated offline. It is possible that deploying these models online (e.g., in the loop of an exoskeleton controller) could result in further error [32]. Additionally, the scope of this analysis was limited to hip moment estimation in the sagittal plane. Future work should include the extension of this analysis to incorporate the remaining lower-limb joints.

VI. Conclusion

This study analyzed the relative effects of time history information in the delayed estimation and anticipation of human hip moments using a temporal convolutional network. In general, we found that benefits to model performance saturated after adding more than 930 ms of previous sensor data as input to the model. Further, we found that delaying the estimator by 200 ms maximized model performance while minimizing total estimator delay. Additionally, anticipating future hip moments up to 200 ms in the future did not degrade model performance beyond that of an alternative analytical model, suggesting the TCN was able to maintain high levels of fidelity during short windows of anticipation, which may have value for control applications.

Acknowledgment

The authors would like to thank Inseung Kang for his technical advice and assistance throughout this study. We would also like to thank Gayeon Choi and Heejoo Jin for their immense help in data collection and post-processing. Finally, we would like to thank Samsung Electronics for providing the exoskeleton technology used in this study.

REFERENCES

[1] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[2] G. S. Sawicki, O. N. Beck, I. Kang, and A. J. Young, "The exoskeleton expansion: improving walking and running economy," *Journal of NeuroEngineering and Rehabilitation*, vol. 17, no. 1, pp. 1–9, 2020.

[3] H.-J. Lee *et al.*, "Training for walking efficiency with a wearable hip-assist robot in patients with stroke: a pilot randomized controlled trial," *Stroke*, vol. 50, no. 12, pp. 3545–3552, 2019.

[4] M. K. Ishmael, D. Archangeli, and T. Lenzi, "Powered hip exoskeleton improves walking economy in individuals with above-knee amputation," *Nature Medicine*, pp. 1–6, 2021.

[5] G. M. Gasparri, J. Luque, and Z. F. Lerner, "Proportional joint-moment control for instantaneously adaptive ankle exoskeleton assistance," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 4, pp. 751–759, 2019.

[6] A. Forner-Cordero, H. J. F. M. Koopman, and F. C. T. van der Helm, "Inverse dynamics calculations during gait with restricted ground reaction force information from pressure insoles," *Gait & Posture*, vol. 23, no. 2, pp. 189–199, Feb. 2006, doi: 10.1016/j.gaitpost.2005.02.002.

[7] T. Khurelbaatar, K. Kim, S. Lee, and Y. H. Kim, "Consistent accuracy in whole-body joint kinetics during gait using wearable inertial motion sensors and in-shoe pressure sensors," *Gait & posture*, vol. 42, no. 1, pp. 65–69, 2015.

[8] K. Kanjanapas and M. Tomizuka, "7 Degrees of Freedom Passive Exoskeleton for Human Gait Analysis: Human Joint Motion Sensing and Torque Estimation During Walking," *IFAC Proceedings Volumes*, vol. 46, no. 5, pp. 285–292, 2013.

[9] A. M. Howell, T. Kobayashi, H. A. Hayes, K. B. Foreman, and S. J. M. Bamberg, "Kinetic Gait Analysis Using a Low-Cost Insole," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 12, pp. 3284–3290, Dec. 2013, doi: 10.1109/TBME.2013.2250972.

[10] D. G. Lloyd and T. F. Besier, "An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo," *Journal of biomechanics*, vol. 36, no. 6, pp. 765–776, 2003.

[11] T. S. Buchanan, D. G. Lloyd, K. Manal, and T. F. Besier, "Neuromusculoskeletal modeling: estimation of muscle forces and joint moments and movements from measurements of neural command," *Journal of applied biomechanics*, vol. 20, no. 4, pp. 367–395, 2004.

[12] G. Durandau, D. Farina, and M. Sartori, "Robust real-time musculoskeletal modeling driven by electromyograms," *IEEE transactions on biomedical engineering*, vol. 65, no. 3, pp. 556–564, 2017.

[13] M. Sartori, M. Reggiani, D. Farina, and D. G. Lloyd, "EMG-driven forward-dynamic estimation of muscle force and joint moment about multiple degrees of freedom in the human lower extremity," *PloS one*, vol. 7, no. 12, p. e52618, 2012.

[14] D. A. Jacobs and D. P. Ferris, "Estimation of ground reaction forces and ankle moment with multiple, low-cost sensors," *J NeuroEngineering Rehabil*, vol. 12, no. 1, p. 90, Dec. 2015, doi: 10.1186/s12984-015-0081-x.

[15] M. E. Hahn and K. B. O'Keefe, "A neural network model for estimation of net joint moments during normal gait," *Journal of Musculoskeletal Research*, vol. 11, no. 03, pp. 117–126, 2008.

[16] D. D. Molinaro, I. Kang, J. Camargo, and A. J. Young, "Biological Hip Torque Estimation using a Robotic Hip Exoskeleton," in *2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)*, 2020, pp. 791–796.

[17] E. Dorschky, M. Nitschke, C. F. Martindale, A. J. van den Bogert, A. D. Koelewijn, and B. M. Eskofier, "CNN-Based Estimation of Sagittal Plane Walking and Running Biomechanics From Measured and Simulated Inertial Sensor Data," *Frontiers in bioengineering and biotechnology*, vol. 8, p. 604, 2020.

[18] D. D. Molinaro, I. Kang, J. Camargo, M. C. Gombolay, and A. J. Young, "Subject-Independent, Biological Hip Moment Estimation during Multimodal Overground Ambulation using Deep Learning," *IEEE Transactions on Medical Robotics and Bionics*, vol. 4, no. 1, pp. 219–229, Feb. 2022, doi: 10.1109/TMRB.2022.3144025.

[19] H. A. Varol, F. Sup, and M. Goldfarb*, "Multiclass Real-Time Intent Recognition of a Powered Lower Limb Prosthesis," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 3, pp. 542–551, Mar. 2010, doi: 10.1109/TBME.2009.2034734.

[20] A. J. Young, A. M. Simon, N. P. Fey, and L. J. Hargrove, "Intent Recognition in a Powered Lower Limb Prosthesis Using Time History Information," *Ann Biomed Eng*, vol. 42, no. 3, pp. 631–641, Mar. 2014, doi: 10.1007/s10439-013-0909-0.

[21] H. C. Siu, J. Sloboda, R. J. McKindles, and L. A. Stirling, "A neural network estimation of ankle torques from electromyography and accelerometry," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 1624–1633, 2021.

[22] J. Camargo, D. Molinaro, and A. Young, "Predicting biological joint moment during multiple ambulation tasks," *Journal of Biomechanics*, vol. 134, p. 111020, Mar. 2022, doi: 10.1016/j.jbiomech.2022.111020.

[23] K. Seo, J. Lee, Y. Lee, T. Ha, and Y. Shim, "Fully autonomous hip exoskeleton saves metabolic cost of walking," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 4628–4635.

[24] I. Kang, D. D. Molinaro, S. Duggal, Y. Chen, P. Kunapuli, and A. J. Young, "Real-Time Gait Phase Estimation for Robotic Hip Exoskeleton Control During Multimodal Locomotion," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3491–3497, Apr. 2021, doi: 10.1109/LRA.2021.3062562.

[25] J. Camargo, A. Ramanathan, W. Flanagan, and A. Young, "A comprehensive, open-source dataset of lower limb biomechanics in multiple conditions of stairs, ramps, and level-ground ambulation and transitions," *Journal of Biomechanics*, vol. 119, p. 110320, 2021.

[26] P. W. Franks, G. M. Bryan, R. M. Martin, R. Reyes, and S. H. Collins, "Comparing optimized exoskeleton assistance of the hip, knee, and ankle in single and multi-joint configurations," *bioRxiv*, 2021.

[27] P. W. Franks, G. M. Bryan, R. Reyes, M. P. O'Donovan, K. N. Gregorczyk, and S. H. Collins, "The Effects of Incline Level on Optimized Lower-Limb Exoskeleton Assistance," *bioRxiv*, 2021.

[28] S. L. Delp *et al.*, "OpenSim: open-source software to create and analyze dynamic simulations of movement," *IEEE transactions on biomedical engineering*, vol. 54, no. 11, pp. 1940–1950, 2007.

[29] A. Seth *et al.*, "OpenSim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement," *PLoS computational biology*, vol. 14, no. 7, p. e1006223, 2018.

[30] A. van den Oord *et al.*, "Wavenet: A generative model for raw audio," *arXiv preprint arXiv:1609.03499*, 2016.

[31] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," *arXiv preprint arXiv:1803.01271*, 2018.

[32] L. J. Hargrove *et al.*, "Intuitive control of a powered prosthetic leg during ambulation: a randomized clinical trial," *Jama*, vol. 313, no. 22, pp. 2244–2252, 2015.