ESTIMATING JOINT MOMENTS DURING TREADMILL RUNNING USING VARIOUS CONSUMER BASED WEARABLE SENSOR LOCATIONS

Joshua Carter[1], Xi Chen[2], Dario Cazzola[1], Grant Trewartha[3], Ezio Preatoni[1]

University of Bath, UK, Department of Health[1]
University of Bath, UK, Department of Computer Science[2]
Teesside University, UK, School of Health and Life Sciences[3]

We estimated lower limb sagittal plane joint moments during treadmill running using wearable sensors and different commonly used locations. We compared outcomes from supervised recurrent neural network machine learning (ML) models to criterion values from motion capture and inverse dynamics. The normalised root mean squared error between outcomes from the ML model fed with the entire wearable dataset (pressure insoles and inertial measurement units at the foot, wrist, T10, and sacrum) was 8.9%, 13.5%, and 18.2% for the ankle, knee, and hip joint respectively. Removal of any two upper body sensors did not decrease the accuracy of the estimations. This work is a springboard to providing biomechanical feedback to runners to help improve performance and minimise injury risk.

KEYWORDS: running gait, pressure sensors, machine learning, inertial sensors

INTRODUCTION: Lower limb joint moments are a biomechanical metric dependent on both the body's position and the forces being applied during the foot-floor interaction. The calculated metric has utility for both performance and injury prevention focused feedback. Typically, joint moments are calculated using inverse dynamics, where the Newton-Euler equations are resolved for each articulated body segment sequentially (Kingma et al. 1996). Usually, lab based optical motion capture and force plates are paired to perform inverse dynamics. More recently, several studies have attempted to overcome the inherent limitations of lab-based data collection, such as expensive equipment, required technical skillset, and extensive collection time, by evaluating the effectiveness of training machine learning models from wearable sensor data for the estimation of joint moments (Lee et al. 2022).

Machine learning models seek to leverage many datapoints that contain both an input signal and a target output signal, with the goal of modelling an existing underlying relationship. In the current context, data from wearable sensor systems serve as the input with joint moments output from inverse dynamics being the target output. Mundt et al. (2020) trained a long short term memory neural network (LSTM) using simulated inertial measurement unit (IMU) data collected from seven anatomical locations to estimate joint moment signals at the ankle, knee, and hip during walking. The trained models had an average Normalised Root Mean Squared Error (NRMSE) of 14.8%. Whilst this work showcases the possibility of using such an approach, a recent survey of 663 runners found that less than 5% of runners currently wear 'body-worn' sensors other than a wrist-based sensor or heart rate monitor (Clermont et al. 2019). Therefore, the purpose of this study was to not only evaluate how accurately lower limb joint moments could be estimated using wearable sensor signals and machine learning, but also to identify how the accuracy of the model changed when using wearable sensor setups more acceptable for runners daily training.

METHODS: *Data Collection:* 20 healthy runners (7 females) with varying levels of experience completed a sub maximal running protocol on a split belt instrumented treadmill (Bertec, OH, USA; 1000 Hz). All participants wore the NURVV Run system (NURVV, London, UK), made up of a pair of insoles with 16 force sensitive resistor pressure sensors under each foot, and an IMU attached to the lateral aspect of the left foot. The pressure sensors collected data at 1000 Hz to ensure accurate identification of foot contact events, but transmitted data at 50 Hz. The IMU was securely fastened to the side of the shoe and collected data at 1125 Hz. An additional three IMU's (Delsys Trigno, Massachusetts, USA), collecting data at 519 Hz, were secured to the wrist, T10, and sacrum using double sided tape and Velcro straps. These locations were chosen to reflect common commercial wearable locations, such as a running watch, chest heart rate monitor, and the waistband of shorts. Twelve Miqus cameras (Qualisys, Gothenburg, Sweden) surrounded the treadmill and collected position data from a full body reflective marker set. The treadmill protocol (22 minutes of running) was split into three sections: flat (1% gradient, 12 min), uphill (6% gradient, 4 min), and downhill (-4% gradient, 6 min). During the flat and downhill stages, participants ran at their self-selected easy run speed, as well as 10% faster and 10% slower than this pace. During the uphill stage participants ran at their chosen easy pace and 10% below this. All participants also ran a flat stage at 12 km/h.

Data Processing: The Delsys API within Qualisys Track Manager allowed synchronous collection of Delsys IMU, treadmill force, and marker data. Data from the NURVV system was synchronised with the lab systems by correlating the stride times identified from each system, further detail of this method can be found in Carter et al. (2023). OpenSim Software (Stanford, USA) was used to perform inverse kinematics and then inverse dynamics, with sagittal plane hip, knee, and ankle net joint moments extracted. During this processing, both the kinematics and force data were filtered with a low pass Butterworth filter with a 10 Hz cut off. All data were segmented into contact periods using a thresholding approach (50 N) to identify initial contact and toe-off events. These contact periods were then normalised to 101 data points.

Data Analysis: The deep learning model used to estimate the net joint moment signals was an LSTM, a recurrent neural network that is effective for regression tasks on time series data. As input, the model took a matrix of size X-by-101, where X represented the number of input features. Before input to the model, each feature was converted to z-scores independently, using the training means and standard deviations. The bi-directional LSTM layer had a hidden size of 256. The LSTM output vector at each time point was then mapped to a single estimated joint moment value by passing it through three linear layers of sizes 512, 256, and 128. Once repeated for each time point, the estimated joint moment trace was constructed. Dropout layers were added before and after the LSTM layer to avoid overfitting to the training set. A 'leaky_ReLu' activation function was applied to the output of the first two linear layers to add additional non-linearity to the model. The RMSE between the moment signal output from the inverse dynamics processing and the estimated moment signal output from the machine learning model was considered the loss value. The loss was calculated across a batch of 16 samples (foot contacts) and an ADAM optimiser utilised back propagation of the loss through the model to update the model parameters in an attempt to minimise RMSE loss. This training process was repeated over three epochs.

A baseline model was created that utilised all available (47) input features: five discrete characteristics (ground contact time, runner body mass, running speed, incline, and insole length), a time series signal from each of the 16 pressure sensors in the insole, a signal for each of the two dimensions of the centre of pressure trace (calculated from the insole), a signal for each of the three dimensions from the accelerometer sensor, and a signal for each of the three dimensions from the gyroscope sensor from the IMUs located on the shoe, wrist, sacrum, and T10. Three alternative models were then compared to this baseline model; the alternative models focused more on the practicality of real-world data collection, utilising data from only one of the three upper body IMUs (wrist, T10, sacrum). Each model was evaluated using Leave One Subject Out validation (LOSO), where the training is repeated as many times as there are participants in the dataset. With each repetition a different participant is left in the validation set and the remaining participants (19) make up the training set.

RESULTS AND DISCUSSION: The baseline model trained on all available input features had a NRMSE of 8.9 ± 2.2%, 13.5 ± 4.2%, and 18.2 ± 7.4% for the estimation of ankle, knee, and hip moments respectively (mean ± standard deviation) (Figure 1B). Across the participant population, there was a large level of variation in the accuracy that the baseline model achieved, and this pattern was consistent across all three joint moment estimations. Mean NRMSE ranged from 3.5% to 19.9% at the ankle, from 6.1% to 26.7% at the knee, and from

7.5% to 36.0% at the hip. The order of the participants from most accurate to least accurate remained relatively consistent across the three joints estimated. This could suggest that those participants that the models consistently performed poorly on had some characteristics within their running style that was not well represented in the data from the rest of the participants. One approach to try to remediate the model performing poorly on certain participants could be simply by expanding the size of the dataset, thus increasing the variation in the training data, either with more collected data or even using data synthesis/augmentation.

The alternative three models that were trained each included only one of the three upper body IMU sensors, included with the intention of evaluating how a more realistic commercial sensor setup faired in comparison to the baseline model. Across all three joints estimated, a similar pattern was seen; the alternative models occasionally negatively impacted estimation accuracy within individual participants, but did not considerably decrease mean performance across all participants. This suggests that with the current model architecture and training process, a more feasible data collection setup would not necessarily be detrimental for the accuracy of lower limb joint estimations. Of the three alternative models tested, the model that just used sensor data from the NURVV Run system and an IMU at the T10 (location of a heart rate monitor) performed best on average, even slightly outperforming the baseline model with mean NRMSE of 8.4%, 12.8%, and 16.9% for the ankle, knee, and hip moments respectively.

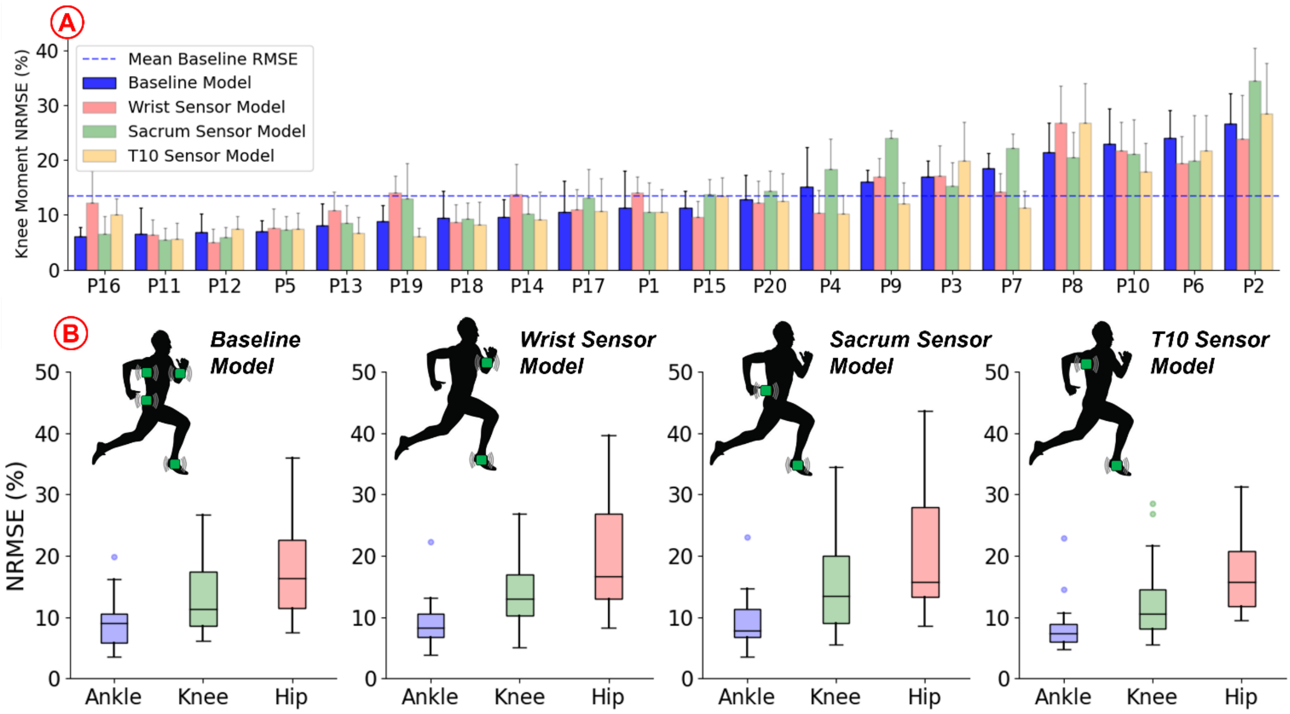

Figure 1A. Bars show the mean Normalised Root Mean Squared Error (NRMSE) for knee moment estimation displaying each participant and model type independently, calculated as the RMSE / the range of the moment values. Error bars show plus one standard deviation.
Figure 1B. Box and whisker plot showing the average NRMSE across all participants for each of the four models. The solid line within each box represents the median error across all participants. The box spans from the 25th to the 75th percentile (interquartile range), and the whiskers encompass 1.5× the interquartile range. Participants outside of this range are plotted as scatter points.

Figure 2 shows examples of estimated and measured net joint moment signals for three independent foot contacts from different participants and during different running conditions. These estimations are output from the baseline model. These examples were chosen to illustrate the average level of accuracy that could be expected with the models trained in this study.

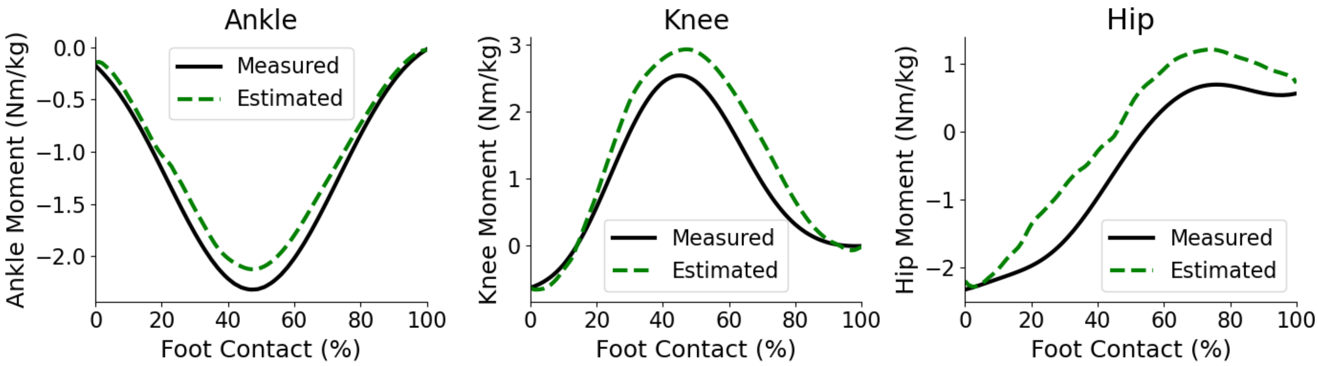

Figure 2: A set of measured and estimated sagittal joint moment signals, chosen based on their similarity to the mean error across all participants and trials. The black lines show the measured moments (from OpenSim), and the dashed green lines are outputs from the trained machine learning models. Ankle: P18, 10.9 km/h, flat Knee: P20, 10.0 km/h, uphill Hip: P3, 14.6 km/h, flat

Tan et al. (2022) recently trained an LSTM model to estimate knee flexion moment during walking using data from eight IMU locations, achieving an average NRMSE of 8.9%.

CONCLUSION: An LSTM model can be used to estimate lower limb joint moments during treadmill running when using wearable sensor data as input. The accuracy with which the joint moments could be estimated varied considerably between participants. The ≤5% accuracy achieved on many of the participants, especially for ankle and knee, could be considered a promising springboard for further developing the model, expanding the sample, and eventually transferring to overground monitoring and then testing outside a lab. However, the current version of the models are still suboptimal for a sub-group of participants. Future work should look to improve on the generalisability of the model, ensuring a more consistent level of error across different participants and running styles. Another finding in this work was that restricting the input to data from fewer upper body wearable sensor locations did not considerably reduce the accuracy of joint moment estimation. This suggests that a sensor setup more acceptable to runners for regular use, such as a shoe-based system and a heart rate monitor, could be as effective as more elaborate and “biomechanics oriented” sensor setups that have been used in previous validation studies. A future version of this model could be used to provide continuous lower limb joint moment data to a much larger pool of runners. Providing runners with this feedback within their typical commercial wearable experience can be used to further inform the runner’s understanding of their biomechanical load, aiming to reduce their risk of injury and maximise training benefit.

REFERENCES:

Carter, J., Chen, X., Cazzola, D., Trewartha, G., & Preatoni, E., 2023. Estimation of ground reaction force during running using consumer-level wearable insoles and machine learning. *ISBS Proceedings Archive* 41, 19.

Clermont, C. A., Duffett-Leger, L., Hettinga, B. A., & Ferber, R., 2020. Runners’ perspectives on ‘smart’ wearable technology and its use for preventing injury. *International Journal of Human–Computer Interaction* 36, 31-40. 10.1080/10447318.2019.1597575

Kingma, I.; de Looze, M.P. Toussaint, H.M.; Klijnsma, H.G. Bruijnen, T.B., 1996. Validation of a full body 3-D dynamic linked segment model. *Human Movement Science.* 15, 833–860. 10.1016/S0167-9457(96)00034-6

Lee CJ, Lee JK., 2022. Inertial motion capture-based wearable systems for estimation of joint kinetics: a systematic review. *Sensors* 22, 2507. 10.3390/s22072507

Mundt, M., Thomsen, W., Witter, T., Koeppe, A., David, S., Bamer, F., ... & Markert, B., 2020. Prediction of lower limb joint angles and moments during gait using artificial neural networks. *Medical & Biological Engineering & Computing*, 58, 211-225. 10.1007/s11517-019-02061-3

Tan, T., Wang, D., Shull, P., & Halilaj, E., 2022. IMU and Smartphone Camera Fusion for Knee Adduction and Knee Flexion Moment Estimation During Walking. *IEEE Transactions on Industrial Informatics*, 19, 1445-1455. 10.1109/TII.2022.3189648.