ESTIMATION OF GROUND REACTION FORCE DURING RUNNING USING CONSUMER-LEVEL WEARABLE INSOLES AND MACHINE LEARNING

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

Department for Health[1] and Department of Comupter Science[2], University of Bath (UK), & NURVV Ltd[3]

Data from NURVV Run, a consumer-level wearable technology product, embedding pressure insoles and inertial transducers, were used as an input into a deep learning model for the estimation of vertical ground reaction forces (vGRF) during running. Force data were collected from an instrumented treadmill during a running protocol of mixed gradients and speeds, serving as the gold standard to evaluate the model accuracy. Mean difference in peak vGRF was 0.36 ± 0.26 BW across participants and mean root mean squared error was 0.27 ± 0.15 BW. Model accuracy varied considerably between participants; it would be expected that a larger dataset with a greater variety of input variables would improve on this. A future version of this model could allow continual assessment of load accumulation during distance running, helping identify early signs of elevated injury risk.

KEYWORDS: gait analysis, running biomechanics, wearable sensors, injury prevention

INTRODUCTION: Millions of people around the world run regularly for sport and recreation. Despite the many physiological and psychological benefits associated with participation in regular distance running, there remains high incidence rates of lower extremity injuries (Videbæk et al., 2015). A proportion of running related injuries are acute in nature, but many more are ‘overuse’, and occur as a result of excessive or sub-optimal loading patterns over many strides. The exact causes of overuse injuries are not well understood, but there is agreement about their multifactorial, diverse, and varying nature between individuals (Winter et al., 2021).

For many years, researchers have looked to establish links between biomechanical variables from lab-based equipment, such as force plates and motion capture systems, and an individual’s risk of developing a running related injury (RRI). However, there are many inherent limitations with this approach. The requirement for large amounts of financial resource and technical expertise to collect data with lab-based systems limits both the quantity and regularity with which biomechanical data can be collected from runners. The collection of biomechanical data in a lab also limits the ecological validity of the results. If similar biomechanical variables, believed to be linked to risk of RRI, could be continuously collected in-field, with an unobtrusive and accessible system, many of the barriers that limit the collection of a large dataset of risk factors could be overcome. Machine learning is proving increasingly successful at identifying underlying relationships between large noisy datasets, such as those from wearables, and target biomechanical variables (Halilaj et al., 2018). The purpose of this study was to take the first step towards creating a continuous load monitoring tool for runners, making use of commercially available and affordable sensors. We developed a deep learning model that estimated vertical ground reaction force (vGRF) using only pressure insole data. Despite mixed findings in the literature, large scale reviews in the area suggest there may be links between vGRF derived metrics and risk of lower limb injury (van der Worp et al., 2016). However, the lack of accurate ‘real world’ assessment tools, which this study primarily aims to improve upon, has potentially limited the identification of such associations.

METHODS: *Data Collection:* Eighteen healthy runners (13 males) of mixed abilities completed a sub maximal running protocol on a split belt instrumented treadmill (Bertec, OH, USA; 1000 Hz). All participants were equipped with a NURVV Run system (NURVV, London, UK), comprised of pressure insoles and an inertial measurement unit. The NURVV insole collects pressure data from 16 sensors evenly distributed across the foot at 1000 Hz, downsampled to 50 Hz before being saved. The treadmill protocol was split into three sections: flat (1% gradient), uphill (6% gradient), and downhill (-4% gradient). During the flat and downhill stages participants ran at their self-selected typical easy run speed, as well as 10% faster and 10% slower than this pace. During the uphill stage participants just ran at their chosen easy pace and 10% below this. All participants also ran a flat stage at 12 km/h.

Data Processing: To align NURVV and instrumented treadmill timelines, stride times were first calculated using a pressure/force thresholding approach. Then, the stride times from each system were scrolled over each other to identify the alignment that elicited the highest correlation value. The result of this was a synchronised set of ground contacts from both left and right feet, in the form of pressure insole data and treadmill force data. Due to vibrational noise in treadmill force data the stance phase was isolated and filtered with a low pass 4th order Butterworth filter with a 15 Hz cut off. The isolated contact periods from each system were then normalised to 400 data points. The average contact period contained 361 data points (when sampling at 1000 Hz), 400 was therefore chosen as not to lose any signal resolution. This processing was repeated for all participants at all usable running speeds and gradients. The subsequent dataset was made up of over 54,000 foot contacts with normalised insole pressure traces and filtered target vGRF traces.

Data Analysis: The deep learning model used to estimate the vGRF data was a Long Short Term Memory network (LSTM), a type of recurrent neural network particularly well specialised for sequential data. For a given ground contact, the model takes 16 pressure channels as 1 by 400 (time) vectors as input and outputs a single 1 by 400 vector (estimated vGRF). All pressure data was scaled to Z-scores prior to model input. The model was made up of two layers of 128 node bidirectional LSTM units, followed by three fully connected neural network layers. The accuracy and generalisability of the network was initially evaluated using “leave one subject out” cross validation, in which one participant’s data makes up the validation set and the data from the rest of the participants make up the training set. During the training process the model would pass 256 ground contacts (batch size) through the model before evaluating the root mean squared error (RMSE) between the estimated and measured vGRF. This RMSE (loss) was then used to inform an update of the model parameters in a gradient descent fashion (ADAM optimiser). This was repeated for all ground contacts in the training dataset (epoch) and repeated 10 times per validation participant. The trained model was then used to estimate vGRF from the pressure data for all ground contacts in the validation set. The absolute difference in the peak vGRF and RMSE for each stance phase in the validation set was then calculated.

RESULTS AND DISCUSSION: The mean absolute difference in peak force ranged from 0.06 ± 0.04 BW to 0.85 ± 0.14 BW (mean ± standard deviation). The average absolute difference in peak force across all participants was 0.36 BW (Figure 1). Similar to the peak vGRF comparisons, large variation in RMSE results across participants was seen (0.09 ± 0.03 to 0.62 ± 0.21 BW) (Figure 2). The mean RMSE across all participants was 0.27 BW.

One explanation for the large variation in results across participants is the inclusion of only pressure data in the input dataset. Additional variables, such as footstrike type, shoe size, running speed, and gradient likely confound the relationship between insole pressure and vGRF, but were not included in this first version of the model. Previous work by Alcantara et al. (2021) used a similar LSTM approach, but combined accelerometer data with discrete variables such as footstrike type, mass, gradient, and running speed into the model input. Whilst the results in the current study had some participants with lower mean RMSE results, the inclusion of discrete variables as input likely contributed to all mean RMSE values being below 0.30 BW in the work conducted by Alcantara et al. (2021). Upcoming work on this project will focus on the influence of including these variables into the model input, as well as alternative model architectures and further model refinement, such as hyperparameter optimisation.

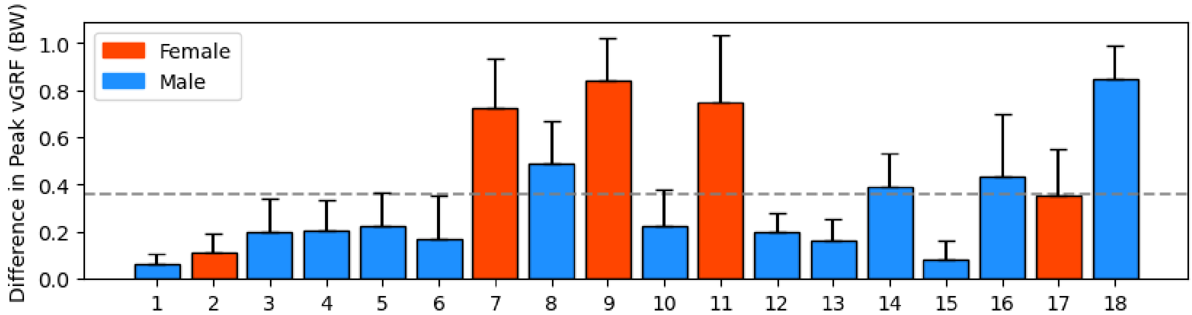

Figure 1. Mean absolute difference (and standard deviation) in peak vGRF (normalised to body weight), between estimated and measured force traces, for all participants.

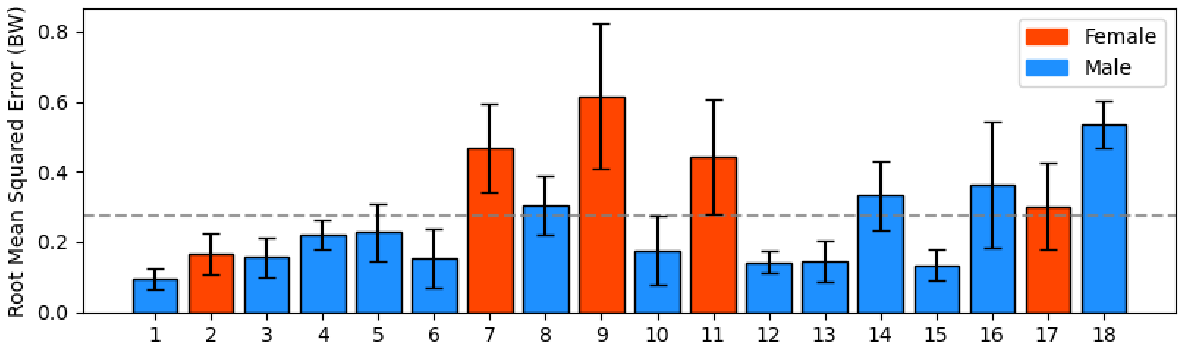

Figure 2. The mean (± one standard deviation) RMSE (normalised to body weight) between the estimated and measured vGRF, for all participants.

As mean RMSE data for participant 5 matched most closely to the mean across all participants their results were deemed as most representative of the current vGRF estimation model. In Figure 3, the estimated and measured vGRF traces for the foot contact closest to the mean RMSE at each gradient was plotted to highlight the expected results for the current model.

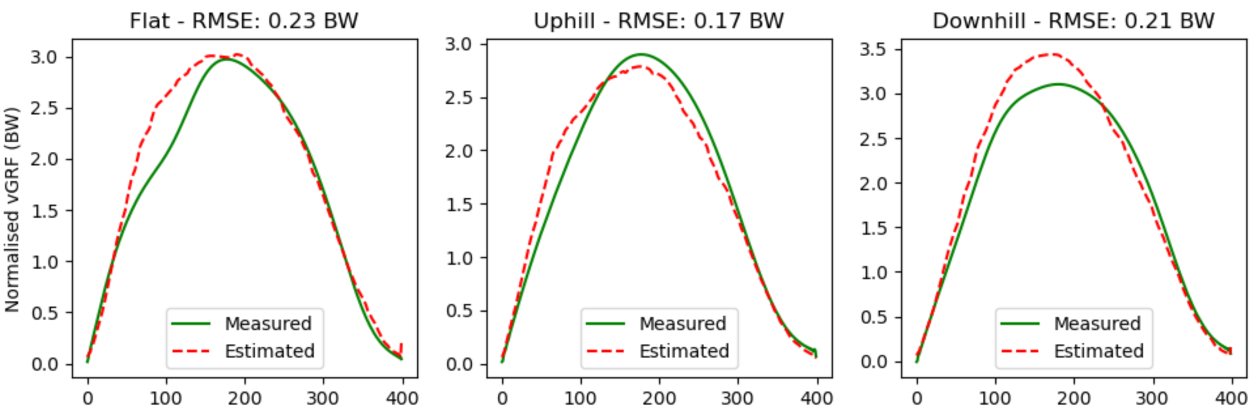

Figure 3. Typical estimated vGRF traces for participant 5 at the same running speed during each of the three gradient stages.

The dataset used in this study contained a population of varying sex-footstrike combinations. Underrepresented subgroups, particularly female participants, showed lower vGRF estimation accuracy in comparison to more heavily represented groups. The most populous subgroup in the dataset was males with a midfoot footstrike pattern (n=6), achieving an average RMSE of 0.23 BW. For those participants with a heel strike pattern the current model was still able to estimate the presence of separate impact peaks and active peaks, as shown by some of the more accurate results from participant 12 (Figure 4). The characteristic is an important model trait. As this would allow accurate calculation of vGRF loading rates for all footstrike types, a metric commonly assessed amongst distance runner populations (Yong et al. 2018).

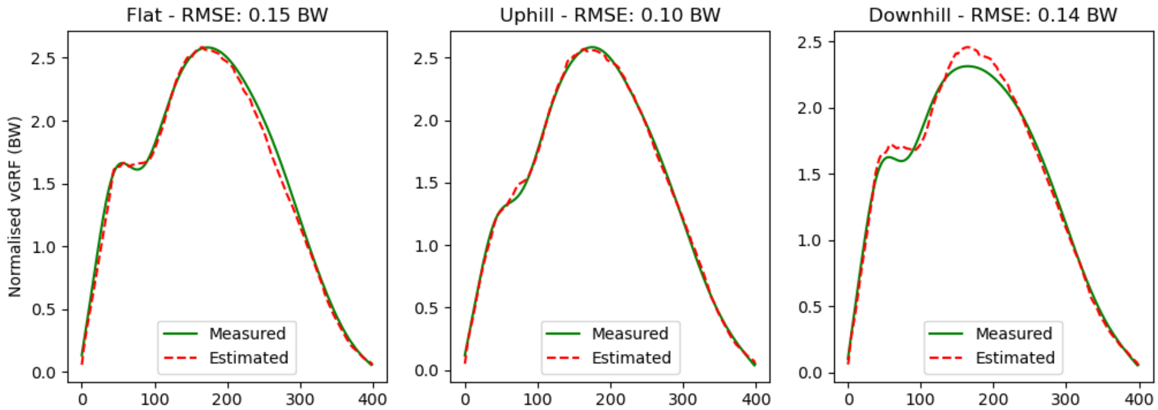

Figure 4. Estimated vertical vGRF traces for participant 12 at the same running speed during each of the three gradient stages.

CONCLUSION: The vGRF estimations produced by the deep learning model in this study showed varying results. Estimation results for many participant's data could be considered accurate enough to monitor load accumulation outside the lab. However, several other participant's data did not meet this standard in this first version of the model. Future work will look to improve on this consistency by training the model on a larger more balanced dataset, adding discrete variables as input, and refining model architecture. Once an appropriate level of accuracy is reached, the low-cost nature of the NURVV system could allow kinetic data to be estimated from thousands of runners throughout their training. Combining this data with regular remote reporting of runner's experiences of pain and injury may allow the complex relationships between running biomechanics and risk of RRI to be more easily unpicked.

REFERENCES:

Alcantara, R.S., Edwards, W.B., Millet, G.Y., Grabowski, A.M., 2021. Predicting continuous ground reaction forces from accelerometers during uphill and downhill running: A recurrent neural network solution. *PeerJ. 10.7717/peerj.12752*

Halilaj, E., Rajagopal, A., Fiterau, M., Hicks, J.L., Hastie, T.J., Delp, S.L., 2018. Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities. *Journal of Biomechanics* 81, 1-11. 10.1016/j.jbiomech.2018.09.009

van der Worp, H., Vrielink, J.W., & Bredeweg, S.W., 2016. Do runners who suffer injuries have higher vertical ground reaction forces than those who remain injury-free? A systematic review and meta-analysis. *British Journal of Sports Medicine, 50,* 450-457. 10.1136/bjsports-2015-094924

Videbæk, S., Bueno, A.M., Nielsen, R.O., Rasmussen, S., 2015. Incidence of Running-Related Injuries Per 1000 h of running in Different Types of Runners: A Systematic Review and Meta-Analysis. *Sports Medicine 45*, 1017-1026. 10.1007/s40279-015-0333-8

Winter, S.C., Gordon, S., Brice, S.M., Lindsay, D., Barrs, S., 2021. Overuse injuries in runners of different abilities—a one-year prospective study. *Research in Sports Medicine 29*, 196-212. 10.1080/15438627.2019.1616548

Yong, J., Slider, A., Montgomery, K., Fredericson, M., Delp, S., 2018. Acute changes in foot strike pattern and cadence affect running parameters associated with tibial stress fractures. *Journal of Biomechanics 76,* 1-7. 10.1016/j.jbiomech.2018.05.017