PEAK VERTICAL GROUND REACTION FORCE PREDICTION FROM KINEMATICS IN MALE RUNNERS USING MACHINE LEARNING ALGORITHMS

Cody Dziuk and Janelle A. Cross

Department of Orthopaedic Surgery, Medical College of Wisconsin, Milwaukee, WI, USA

The purpose of this study was to examine if peak vertical ground reaction forces during treadmill running can be predicted from kinematic input using machine learning models. Eighteen healthy male runners' hip, knee, and ankle sagittal angles, with subject metadata, were input into random forest, support vector, and multi-layer perceptron regressors. Thirty strides per side at three speeds were pulled for the dataset. Random forest performed the best with a correlation coefficient of 0.950 and a root mean squared error of 0.456, while multi-layer perceptron was the worst with values of 0.948 and 0.462 respectively. The study showed machine learning models can predict peak vertical ground reaction forces.

KEYWORDS: running, gait analysis, support vector, artificial neural network, random forest

INTRODUCTION: Vertical ground reaction force (VGRF) is an important metric of running analyses that provides insight into running mechanics. VGRF characteristics have been shown to change in response to a variety of different factors including sex, fatigue, and footwear (Bazuelo-Ruiz et al., 2018, Logan et al., 2010). Also, excessive ground reaction forces and knee joint loads have been identified as potential risk factors to the occurrence of running-related injuries (Messier et al., 2008). In the laboratory, ground reaction forces are measured via force plates or force transducers. This isn’t always practical depending on the situation. Kinematic variables are easier to measure than kinetics, with current technology providing measurements of sagittal kinematics using a phone or tablet (Mousavi et al., 2020). If such an app could be validated against optical three-dimensional (3D) motion capture to provide accurate kinematic data, a method of VGRF prediction based on kinematic input could be established for runners. A prediction model using a validated gold-standard data collection method needs to be established first.
Prior studies have attempted to predict VGRFs by means of neural network models trained on accelerometer data (Ngoh et al., 2018), 3D kinematic data with stacked machine learning models (Ong et al., 2020), neural network models trained on inertial sensor data (Wouda et al., 2018), and various other implementations. These methods require the use of complex electronic equipment. The goal of this study was to determine if peak VGRFs during stance phase of running gait can be predicted based on discrete lower body sagittal kinematics combined with subject metadata.

METHODS: A public dataset of running C3D files from the Laboratory of Biomechanics and Motor Control was analyzed for this study (Fukuchi et al., 2017). Eighteen healthy male runners (age: 34 ± 6 years, mass: 70.0 ± 7.4 kg, height: 174.8 ± 6.8 cm) were selected. Trials consisted of treadmill running recorded at three running speeds (2.5, 3.5, 4.5 m/s). Kinematics were collected at 150 Hz using a 3D optical motion capture system. Kinetics were collected via a dual-belt force-instrumented treadmill at 300 Hz. FS and FO events were marked using Visual 3D software (C-Motion, Germantown, MD), then kinematic and kinetic data for each step normalized to 101 points of stance in order to compare left versus right sides. Ground reaction forces were normalized to bodyweight (N/kg). Lower body kinematic input included hip, knee, and ankle sagittal plane angles at foot strike (FS) and foot off (FO). Subject metadata consisted of height, weight, and running

speed. Support vector regression (SVR), an artificial neural network (ANN), and random forest (RF) were the chosen prediction algorithms.
Statistical parametric mapping (SPM) was performed on the stance phase sagittal angles between sides at each speed. No significant results were found (α=0.05) so the steps from left and right sides were combined. Figure 1 shows the SPM results at the 3.5 m/s speed for the knee sagittal angles.

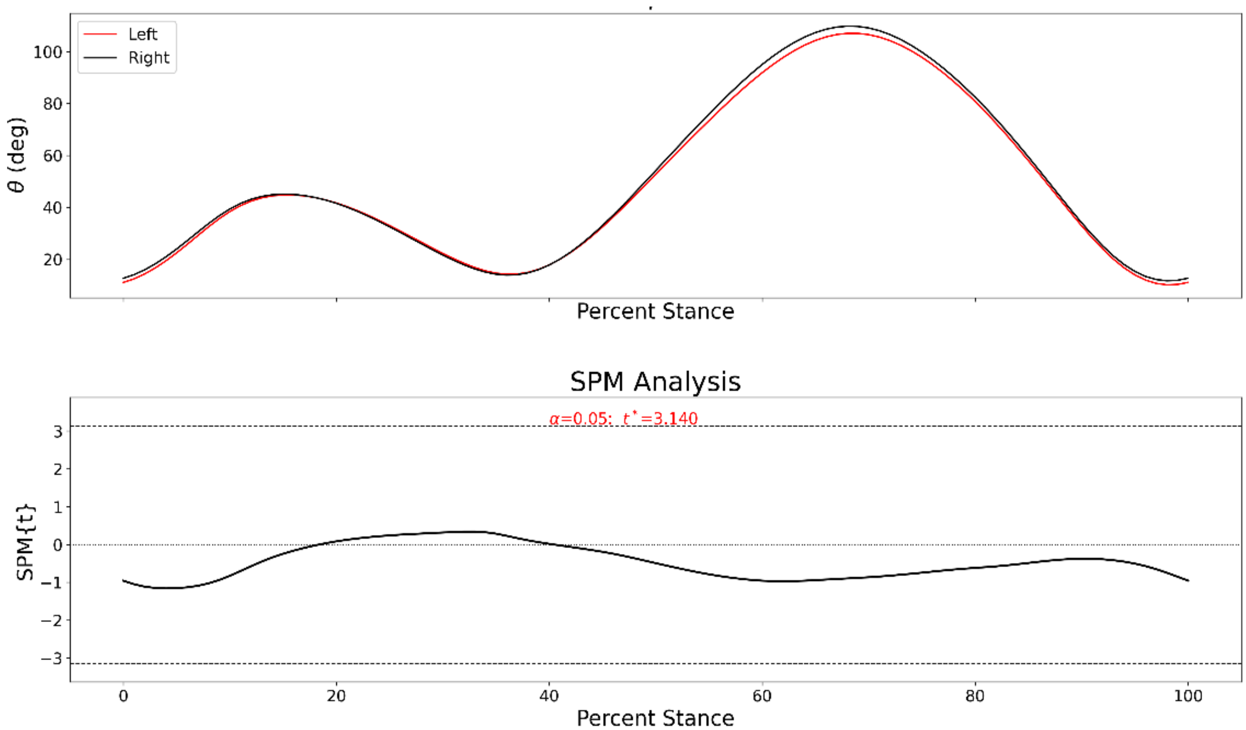

Figure 1. Top: Left and right knee sagittal angles; Bottom: SPM results between sides.

Two discrete time points were chosen to create the input vector data, sagittal angles at 0 and 100 percent of stance, corresponding to FS and FO. Each input vector consisted of 9 metrics: hip, knee, and ankle angles of the side analyzed at the timepoints; as well as subject mass, height, and running speed. Thirty strides per side at all three speeds were pulled for a total of 180 strides per subject, and an input dataset of 3240 steps (18 subjects x 2 sides x 3 speeds x 30 steps). To prepare for machine learning analysis, the trials were shuffled, and a standard 80-20 train-test split was performed. The training vectors were fit transformed, and test vectors were z-transformed. A custom python code was used to implement the machine learning algorithms. The training data was fed into SVR, ANN, and RF regressors from the scikit-learn python library (https://scikit-learn.org/stable/). Grid search cross validation (cv = 10) was performed to determine optimal parameters for each regressor. After training, the optimal regressors were used to predict peak GRFs of the test data.
SVR parameters for grid search included the kernel function applied to the data, regularization parameter (C), the kernel coefficient gamma, and epsilon. For the RF the number of estimators (trees) in the forest, criterion function for split quality, minimum number of samples per split, and max features for a split were chosen. For the ANN activation functions for the hidden layer, solvers for weight optimization, learning rate, and a single hidden were used. Test data was input into

each optimal regressor, and correlation coefficients (R^2) plus root mean squared error (RMSE) recorded.

RESULTS: The results from the optimal regressors can be seen in Table 1. All three methods were able to predict peak VGRF with high accuracy. The RF regressor had the best performance with the highest correlation coefficient (0.950) and lowest RMSE (0.456). The ANN regressor had the worst performance with the lowest correlation coefficient (0.945) and highest RMSE (0.477).

Table 1: Correlation coefficients (R^2) and root mean squared error (RMSE) of the optimal regressors.

Regressor	R^2	RMSE
RF	0.950	0.456
SVR	0.948	0.462
ANN	0.945	0.477

DISCUSSION: The results showed that machine learning regressors can be used to predict peak VGRF. All regressors showed high accuracy, with the RF regressor at 95%, and the lowest RMSE. Support vector regression also showed good results with an accuracy of 94.8%, it is less computationally expensive than other methods, and can be used on small datasets. Prior studies have predicted peak VGRF with greater accuracy, or smaller RMSE (Ngoh et al., 2018, Ong et al., 2020), but these require the use of complex equipment like accelerometers and motion capture labs. For the everyday runner this is not always an option. While the current study used data collected from a 3D motion capture system, only sagittal plane variables were used to determine if predicting the peak VGRF was possible.

If a motion analysis app could provide validated data for sagittal plane kinematics, peak VGRF could be predicted using optimal regressors and a phone or tablet. Mousavi et al. (2020) investigated the app Coach’s Eye, which was found to be reliable for some kinematics, but not all, and they recommend measuring 2D sagittal hip angle with alternative methods. A prior study found that larger knee joint loads were related to poor hamstring flexibility and fatigue, among other factors (Messier et al., 2008). Adjustments made to flexibility and training schedule could be investigated with VGRF prediction to determine if these are being mitigated, thus possibly reducing joint loads during running.

CONCLUSION: The study showed that sagittal plane kinematics input into regression models can be used to predict peak vertical ground reaction forces. As VGRFs are an important running metric, the ability to predict peak VGRFs from 2D sagittal kinematics warrants further investigation. An adequate app to measure sagittal plane kinematics needs to be created to provide reliable training data. If peak VGRFs can be easily predicted with a smartphone or tablet, the everyday runner would have more detailed information for evaluating their running mechanics and training plan.

REFERENCES:

Bazuelo-Ruiz, B., Durá-Gil, J. V., Palomares, N., Medina, E., & Llana-Belloch, S. (2018). Effect of fatigue and gender on kinematics and ground reaction forces variables in recreational runners. *PeerJ*, *6*, e4489. https://doi.org/10.7717/peerj.4489

Fukuchi, R. K., Fukuchi, C. A., & Duarte, M. (2017). A public dataset of running biomechanics and the effects of running speed on lower extremity kinematics and kinetics. *PeerJ*, *5*, e3298. https://doi.org/10.7717/peerj.3298

Logan, S., Hunter, I., J Ty Hopkins, J. T., Feland, J. B., & Parcell, A. C. (2010). Ground reaction force differences between running shoes, racing flats, and distance spikes in runners. *Journal of sports science & medicine*, *9*(1), 147–153.
Messier, S. P., Legault, C., Schoenlank, C. R., Newman, J. J., Martin, D. F., & DeVita, P. (2008). Risk factors and mechanisms of knee injury in runners. *Medicine and science in sports and exercise*, *40*(11), 1873–1879. https://doi.org/10.1249/MSS.0b013e31817ed272
Mousavi, S., Hijmans, J., Moeini, F., Rajabi, R., Ferber, R., van der Worp, H., Zwerver, J. (2020). Validity and reliability of a smartphone motion analysis app for lower limb kinematics during treadmill running. *Physical Therapy in Sport,* 43, 27-35. https://doi.org/10.1016/j.ptsp.2020.02.003.
Ngoh, K. J., Gouwanda, D., Gopalai, A. A., & Chong, Y. Z. (2018). Estimation of vertical ground reaction force during running using neural network model and uniaxial accelerometer. *Journal of biomechanics*, *76*, 269–273. https://doi.org/10.1016/j.jbiomech.2018.06.006
Ong, A., Sujae, I.H., & Hamill, J. (2020). ESTIMATING THE PEAK VERTICAL GROUND REACTION FORCE COMPONENT AND STEP TIME IN TREADMILL RUNNING USING MACHINE LEARNING-A PILOT STUDY. *ISBS Proceedings Archive,* 38(1), 352-355.
Wouda, F. J., Giuberti, M., Bellusci, G., Maartens, E., Reenalda, J., van Beijnum, B. F., & Veltink, P. H. (2018). Estimation of Vertical Ground Reaction Forces and Sagittal Knee Kinematics During Running Using Three Inertial Sensors. *Frontiers in physiology*, *9*, 218. https://doi.org/10.3389/fphys.2018.00218