A COMPARISON OF THREE METHODS FOR ESTIMATING VERTICAL GROUND REACTION FORCES IN RUNNING

Dimitrios-Sokratis Komaris[1], Eduardo Perez-Valero[1], Luke Jordan[2], John Barton[1], Liam Hennessy[2], Brendan O'Flynn[1] and Salvatore Tedesco[1]

Tyndall National Institute, University College Cork, Cork, Ireland[1]
Setanta College Ltd, Thurles, Ireland[2]

The purpose of this study was to compare different approaches for the estimation of biomechanical loads in running. A neural network, a biomechanical model, and a two-mass model were tested on the same data set. The predictions of the neural network were highly accurate for all considered running speeds (average RMSE, 0.11 BW). The biomechanical model returned statistically similar results (p=0.113, 0.14 BW), but with increasing RMS errors at high running speeds. Finally, the two-mass model estimates were independent of running speed, but were the least accurate (RMSE, 0.18 BW).

KEYWORDS: Artificial neural networks, biomechanics, motion analysis, kinematics.

INTRODUCTION: Vertical ground reaction forces are usually linked to overuse running injuries (Hreljac, 2004) and their accurate measurement in running is typically carried out with the use of costly instrumented treadmills in laboratory grounds. In an attempt to indirectly measure VGRFs on open field, different methodological approaches that rely on wearable inertial measurement units were suggested (as reviewed by Ancillao et al., 2018). Such attempts typically employ artificial neural networks (ANN), biomechanical models, or mass-spring-damper systems (Komaris et al., 2019a), with progressively accurate results. Yet, comparisons between the reported outcomes of different methods is unfeasible, since the efficiency of an algorithm is directly associated to the used dataset. The purpose of this work was to compare different techniques for the estimation of VGRFs on the same body of data. Three methods were tested: an ANN, a biomechanical model, and a two-mass model.

METHODS: All three analyses in this study were carried out on the same public dataset (Fukuchi et al., 2017) of 28 subjects (age: 34.8 ± 6.6 years; height: 176 ± 6.7 cm; mass: 69.6 ±7.6 kg; gender: 27 males) running at 2.5, 3.5 and 4.5 m/s, for 30s at each condition. Running was captured with a twelve-camera system and an instrumented dual-belt treadmill, with sampling frequencies of 150 and 300 Hz, respectively. Raw recordings, along with further information on recruitment, laboratory configuration and testing are detailed by the authors of the dataset. Details on the three methodological approaches are reported below.

ANN: a supervised, feed-forward ANN with backpropagation was built in Python (in line with Komaris et al., 2019b). The developed network consisted of a hundred-neuron input layer, a hidden layer of ten neurons utilising hyperbolic tangent as an activation function, and an output layer with a hundred linear neurons. Regulation was undertaken using dropout, while root mean square error (RMSE) was used as a loss function. Only the acceleration of the shanks, as computed by the markers' position, was used as an input. Marker and force signals were filtered with a low-pass, second order, zero-phase shift Butterworth filter with a cut-off frequency of 25 Hz. Acceleration and GRF signals were both normalised to participants' body weight, and scaled to 100 points extending from heel-strike to toe-off. The dataset was then divided into training (subjects 1 to 16), validation (subjects 17 to 22) and test sets (subjects 23 to 28), and the network was trained and evaluated in the prediction of VGRFs. The combination of hypermeters (training epochs, batch size, and dropout rate) that optimised predictions were also identified by a grid search on the training set.

Biomechanical Model: A straightforward biomechanical approach (in keeping with Komaris et al., 2019a) was implemented, with GRFs being estimated as the summation of two products of masses and accelerations: $VGRF = m_{thigh} \cdot (a_{V,thigh} + g) + (bm - m_{thigh}) \cdot (a_{V,pelvis} + g)$. A grid-search was carried out to reduce the overall prediction RMS error by optimising the calculation of the thigh masses ($\mathrm{m_{thigh}}$), and the vertical accelerations of the thighs ($a_{V,thigh}$) and pelvis ($a_{V,pelvis}$). For this, the dataset was divided into grid-search (subjects 1 to 22) and test sets (subjects 23 to 28). First, force waveforms were filtered as previously described, while marker data were up-sampled to reach the sampling rate of the instrumented treadmill. Next, during the grid search, spatial data were filtered with the same filter that was employed for the forces but with a broad range of cut-off frequencies (fc): 2−15 Hz for the pelvis and 10−50 Hz for the thigh markers, with 1 Hz intervals. Following their filtering, vertical marker positions were double differentiated, resulting in an approximation of the segments' vertical acceleration. Along with the accelerations of the thighs and pelvis, the $\mathrm{m_{thigh}}$ that reduced $VGRF$ prediction errors was also explored, with values extending from 5−30% of the participant's total body mass (bm), with 1% increments. Therefore, a total of 14,924 combinations per trail were examined at this stage. Finally, two-way ANOVAs and Tukey HSD tests were implemented to determine the effect of running speed and foot-strike pattern (forefoot or rearfoot) on the fc and thigh masses that reduced the RMSE between measured and predicted GRFs. Optimum parameters for each speed and landing condition were then used to estimate loads in the test set (subjects 23 to 28).

Two-Mass Model: The two-mass model (Clark et al., 2017) assumes that the vertical GRF waveform is formed by two overlapping cosine bell curves that are the products of the collision of the lower-limb mass (m_1) and the remaining body ($m_2 = bm - m_1$) with the ground. Accordingly, m_1 and m_2 collisions are responsible for the impact and active peaks of the waveform, respectively. The model assumes constant-speed level running, and that the net vertical centre of mass displacement from step to step is equal to zero. Predictions are generated using body masses ($m_1 = 0.08bm$, as optimised by the original authors) and stride-specific measures: contact time (t_c), aerial time (t_a), and the lower-limb vertical acceleration during impact (determined from Δ_{v1}/Δ_{t1}, **Table II**). For the computation of these measures, the position of the heel markers was used. Each bell curve is based on the raised-cosine function dependent on the aforementioned model parameters. In consonance with the preceding approaches, the predictions were calculated only for the recordings of subjects 23 to 28.

Finally, a one-way ANOVA and post-hoc tests were conducted to determine statistically significant differences in the GRF predictions made from all three computational models.

RESULTS: Recorded VGRFs and their estimates from the three considered approaches were averaged and graphed for each separate running speed of the test set (**Figure 1**).

Table I: RMSE from the prediction of vGRFs as calculated by three computational models.

Computational models	RMSE (BW)				R^2
	2.5 m/s	3.5 m/s	4.5 m/s	All speeds	All speeds
1: ANN	0.09±0.02	0.10±0.03	0.13±0.04	0.11±0.04	0.99±0.01
2: Biomechanical model	0.09±0.01	0.13±0.01	0.19±0.01	0.14±0.04	0.99±0.01
3: Two-mass model	0.18±0.05	0.17±0.04	0.18±0.04	0.18±0.04	0.98±0.03

ANN: Both RMSE (**Table I**) and graphical representation (**Figure 1**, upper row) indicate that the predictions made from the ANN were highly accurate with an average error of 0.11 BW.

Biomechanical Model: A grid search in the trials of 22 subjects was undertaken during the second methodological approach to optimise the filter's fc and the thigh mass allocation as expressed in bm percentage. During the process, values that resulted in minimum prediction errors were logged and subjected to two-way ANOVAs and pairwise post-hoc comparisons to determine if the considered parameters are affected by running speed and foot-strike pattern. The effect of running speed to pelvis fc was statistically significant (p=0.05), and so was the foot-strike pattern to the fc of the thigh markers (p=0.035). There were no other statistically significant main effects or interactions. Average values for the fc of the pelvis (5, 6 and 7 Hz for the 2.5, 3.5 and 4.5 m/s, respectively) and thigh markers (17 and 25 Hz for the forefoot and rearfoot strikers, respectively) along with the thigh mass (0.143 of bm for all conditions) were then used to predict VGRFs in the latter part of the dataset (subjects 23 to 28) with satisfactory graphical (**Figure 1**, middle row) and numerical results (0.14 BW, **Table I**).

Two-Mass Model: Contrary to the former two approaches, the accuracy of the predictions by the two-mass model does not decline as speed increases (**Table I**). Model parameters are inherently dependant to running speed (**Table II**). Although the model overestimated the peak of the force, it accurately captured the unloading phase of the curve (**Figure 1, bottom row**).

Table II: Input parameters for the two-mass model

Speed	t_c (s)	t_a (s)	Δ_{t1} (s)	Δ_{V1} (m/s)
2.5 m/s	0.305±0.006	0.081±0.006	0.037±0.008	0.467±0.062
3.5 m/s	0.258±0.005	0.104±0.005	0.035±0.005	0.668±0.082
4.5 m/s	0.225±0.004	0.108±0.006	0.036±0.004	0.878±0.089
All speeds	0.263±0.005	0.098±0.006	0.036±0.006	0.671±0.078

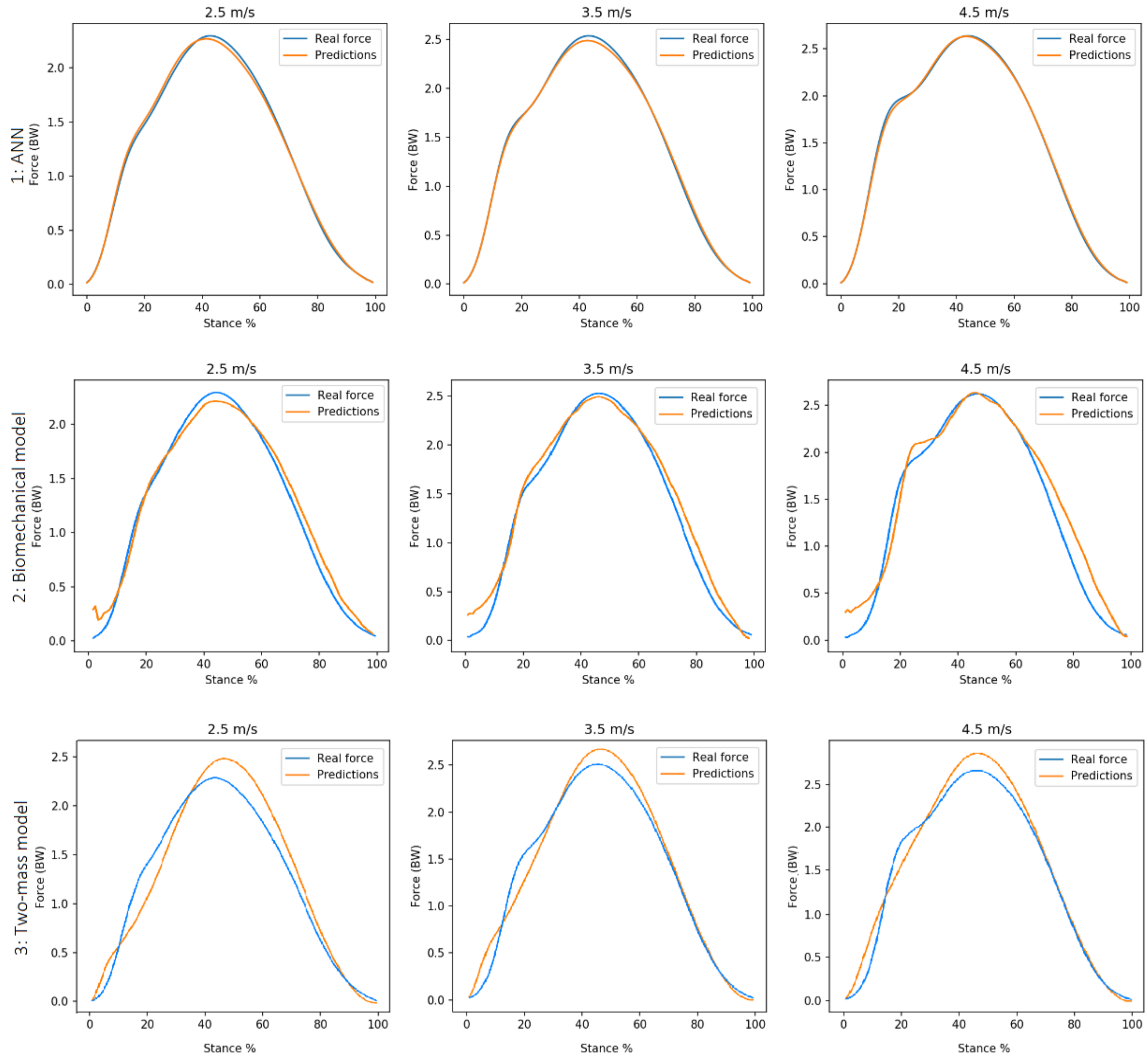

Figure 1: Average predicted (orange line) and measured (blue line) VGRFs at different speeds (2.5, 3.5 and 4.5 m/s), as calculated by an ANN, a biomechanical, and a two-mass model.

Even though the goodness-of-fit agreement between measurements and predictions approached unity for all methods (**Table I**), results by both the biomechanical and two-mass models are inferior to those obtained via ANN. A one-way ANOVA determined that the difference in the RMS prediction errors from the examined methodological approaches was statistically significant (p=0.001). Pairwise post-hoc comparisons further verified that both the ANN (p=0.001) and biomechanical model (p=0.015) were accompanied by lower errors as against the two-mass model. Even though the biomechanical approach failed to accurately predict the loading phase of the GRF waveform, the difference in the RMSE between this method and the ANN was non-significant (p=0.113).

DISCUSSION: To date, several authors researched methods to approximate running loads (as reviewd by Ancillao et al., 2018); yet, no direct comparisons among dissimilar methodological approaches were possible, since the quality of the results is conditioned by the employed dataset. Thus, the present manuscript details the first work to directly compare diverse methods for VGRF predictions in running based on body accelerometry.

The estimations of the ANN in this study were of excellent accuracy (average speed: 0.11 BW) and in proportion with the results previously reported by the authors (0.13 BW, Komaris et al., 2019b). As regards the biomechanical model, the authors previously applied the algorithm on a different data split of the same dataset (Komaris et al., 2019a), and reported comparable optimum values for all considered parameters: 6-7 Hz for the pelvis fc; 20-24 Hz for the thigh fc; 14% of BM for the thigh mass. Furthermore, GRF estimation errors in this study (0.14 BW) were also in agreement to the previously reported statistics (0.14 BW, Komaris et al., 2019a). Likewise, the RMSEs of the two-mass model in this work (2.5 to 4.5 m/s: 0.18 BW) were in accordance to the values reported by the original authors of the model (3 to 6 m/s: 0.17 BW, Clark et al., 2017). As regards to the model parameters (**Table II**), the Δ_{V1} values were approximately half in value compared to the ones reported by Clark et al. (2017). In the authors' opinion, this discrepancy may be due to the ankle (as originally proposed by Clark et al.) being substituted by heel markers; this, was due to the absence of malleoli markers on the used motion capture dataset. Finally, the limited capacity of the two-mass model to effectively predict the impact peak of the waveform (**Figure 1**) may be addressed by optimising the mass of the lower limb (m_1).

Both the ANN and biomechanical model accurately predicted VGRFs at lower speeds (2.5 m/s: 0.09 BW); yet, ANN estimation errors at higher running speeds grew at a lower rate (4.5 m/s: 0.13 compared to 0.19 BW). GRF predictions from the two-mass model were not related to running speed but were followed by higher errors (**Table I**). Statistical tests and graphical illustration dictate that ANNs provide significantly more accurate predictions; however, given their "black-box" behaviour, if insights on the structure are required, biomechanical models may return fair waveforms with RMSEs comparable to those of an ANN (p=0.113).

CONCLUSION: This is the first study that reports on the estimation accuracy of different algorithms in the prediction of VGRFs in running. Three methodological approaches were put to a test on the same dataset: an ANN, a biomechanical model, and a two-mass model. The ANN model returned the most precise illustration of the double-peaked curve, with an excellent depiction of the loading and unloading phases. Irrespective of the less appealing waveform estimation, the biomechanical model returned predictions with RMSEs statistically comparable to those of the ANN. Statistical tests also established that the results of the two-mass model are second to those obtained from both the ANN and biomechanical model.

REFERENCES

Ancillao, Andrea, et al. (2018). "Indirect Measurement of Ground Reaction Forces and Moments by Means of Wearable Inertial Sensors: A Systematic Review." *Sensors, 18*(8).

Clark, Kenneth P., et al. (2017). "A general relationship links gait mechanics and running ground reaction forces." *J Exp Biol, 220*.2, 247-258.

Fukuchi, Reginaldo K., et al. (2017). "A public dataset of running biomechanics and the effects of running speed on lower extremity kinematics and kinetics." *PeerJ, 5*, e3298.

Hreljac, Alan. (2004). "Impact and overuse injuries in runners." Medicine and science in sports and exercise 36.5: 845-849.

Komaris, Dimitrios-Sokratis, et al. (2019a). "Effects of Segment Masses and Cut-off Frequencies on the Estimation of Vertical Ground Reaction Forces in Running." *J Biomech*: 109552.

Komaris, Dimitrios-Sokratis, et al. (2019b). "Predicting Three-Dimensional Ground Reaction Forces in Running by Using Artificial Neural Networks and Lower Body Kinematics." *IEEE Access, 7*, 156779-156786.

ACKNOWLEDGEMENTS: This work was supported by Enterprise Ireland and Setanta College (IP 2017 0606), and by Ireland's European Structural and Investment Programme 2014-2020. This publication was financed in part by Science Foundation Ireland under Grant number 12/RC/2289-P2 which is co-funded under the European Regional Development Fund.