

Estimating Vertical Ground Reaction Force during Running with 3 IMUs

Bouke L. Scheltinga^{1,2}, Hazal Usta¹, Jasper Reenalda^{1,2}, Jaap H. Buurke^{1,2}

¹Biomedical signals and systems, University of Twente
Drienerlolaan 5, Enschede, The Netherlands
b.l.scheltinga@utwente.nl;

²Roessingh Research & Development, Enschede, the Netherlands
Roessinghsbleekweg 33B, Enschede, The Netherlands
j.reenalda@rrd.nl; j.buurke@rrd.nl

Abstract – Quantification of biomechanical load is crucial to gain insights in the mechanisms causing running related injuries. Ground reaction forces (GRF) can give insights into biomechanical loading, however, measuring GRF is restricted to a gait laboratory. Developments in inertial sensor technology make it possible to measure accelerations outside the lab in the runners' own environment. Using Newton's second law, accelerations of body masses can be used to estimate forces. The main objective of this study is to estimate vertical GRF with three inertial measurement units using a generic algorithm based on Newton's second law. This is achieved using an optimization for filtering properties of the acceleration data and body masses. 10 Heel strike runners performed 9 trials at three different velocities (10, 12 and 14km/h) and three different stride frequencies (low, preferred, high), on an instrumented treadmill. Inertial measurement units were placed at sternum, pelvis, upper legs, tibias and feet. Optimization was performed for different sensor configurations. The root mean squared error (RMSE) between the estimated GRF and measured GRF was used as loss function in the optimization. As performance measure of the algorithm, RMSE, active peak error and Pearson's correlation coefficient were used. The setup with sensors on the tibia and pelvis showed the best result, with an average RMSE of 0.179 bodyweight, peak error of 3.7% and Pearson's correlation coefficient of 0.98. Using leave-one-subject-out cross validation, it is shown that the algorithm is generalizable for heel strike runners. Model performance decreases with velocity but increases with stride frequency. The main error of the algorithm is seen in the first 25% of the stance phase, however, the general performance is comparable or better than what is described in current literature.

Keywords: Ground reaction force, running, inertial measurement units, estimation

1. Introduction

Worldwide, running is a very popular sport. Unfortunately, a lot of runners are plagued by injuries [1]. Monitoring biomechanical load and especially change in biomechanical load can help in understanding the development of running injuries [2]. Ground reaction force (GRF) is seen as an important measure to quantify biomechanical load during running [3]. Furthermore, by linking GRF with kinematics, inverse dynamics can be applied to estimate structure specific body load. However, the measurement of GRF is restricted to the laboratory setting as a force plate or an instrumented treadmill is needed to measure the forces. A sensor set-up with a minimal amount of sensors to quantify GRF as a measure of biomechanical load would be of great help in (longitudinal) studies to get a better understanding in the development of running injuries [2], [4].

Inertial measurement units (IMUs) can be used to estimate GRFs, for example, it is shown that there is a relation between peak tibial acceleration and the peak vertical ground reaction force [5]. More recently, a Kalman based method was used to estimate the GRF during the whole stance phase, however there was still an error in the estimated peak value (6.1-7.1%) and relatively high error in the loading rate (>24%) [6]. Furthermore, it is shown that an artificial neural network is capable in estimating vertical GRFs from IMU data [7]. However, it appears to be difficult to create general models, applicable to a large group of subjects when using machine learning methods [7]. Furthermore, it appears to be hard to estimate the absolute values of the GRF [8]. For methods based on a physical model, it is shown that IMU filter settings do influence the GRF estimations [9]. Besides filtering, the mass distribution plays a role in the performance of the GRF estimation algorithm [10]. Therefore, the filter frequencies, filter orders and mass distribution will be optimized in this study.

An algorithm intended to be used in the runners' environment should be robust to variations in different parameters. It is already shown that accuracy can be dependent on running speed [10], but the algorithm should also work at a range of stride frequencies. Hence, the effect of running velocity and stride frequency on the performance of the algorithm is quantified.

The primary objective of this work is to estimate vertical GRF with three IMUs using a generic model based on Newton's second law. The secondary objective is to validate the model for different velocities and stride frequencies.

2. Material and Methods

2.1. Experimental Design and Measurement Setup

Ten healthy experienced heel strike runners participated in this study (4 F, 6 M; age $31.9y \pm 9.7y$, height $1.78m \pm 0.10m$, mass $73.7kg \pm 19.2kg$). All participants signed a written informed consent and the protocol was approved by the medical ethical committee (CCMO Arnhem-Nijmegen).

Participants were equipped with sensors following the lower body configurations of the Xsens MVN Link system (sampling frequency of 240Hz) using double sided tape. Sensors were placed on the sternum, pelvis, both feet, both tibias and both upper legs. On top of the sensor, additional tape was placed and leg sleeves were pulled over the tibia sensors to prevent movement artefacts??. After sensor placement, the system was calibrated according to the manufacturer's instructions.

Ground reaction force data was collected using a 3-dimensional force plate instrumented dual-belt treadmill (sampling frequency 2048Hz).

First, the subject mass was determined during a static trial on the force plate. Then, the participants ran at three different speeds in random order (10, 12 and 14km/h) and three different stride frequencies, relative to their preferred stride frequency (100, 110 and 90%) for 90 seconds. For each speed, the preferred stride frequency was determined during the first trial. In the other two trials, the stride frequency was imposed using a metronome. Each trial started with three jumps for synchronization of the IMUs and treadmill. There was a 3-minute break between the trials to prevent the participants from fatiguing.

2.2. Data Processing

Measured GRF (mGRF) data was filtered using a sixth-order, zero-phase shift low pass Butterworth filter with a cut-off frequency (F_c) of 30 Hz. Next, this data was resampled to 240Hz to match the IMU system. Then, the GRF data was normalized for body weight. From the IMUs, the sensor free acceleration is used. This is the gravity subtracted acceleration in the global frame. Next, the cross correlation between the sensor free acceleration of the tibia in vertical direction with the vertical mGRF was calculated. The lag value corresponding to the maximum in the correlation was used to synchronize the treadmill and IMUs.

40 strides were extracted from the middle part of each trial and used for the fitting and evaluation of the estimated GRF (eGRF). Data was labelled as flight-phase if the mGRF was $<25N$ for longer than 0.05s and as stance-phase if the measured GRF was $>25N$ for longer than 0.05s.

2.3. GRF Estimation Model

Six different sensor configurations were created, combining two sensors on the leg (tibias, thighs or feet) and one on the trunk (pelvis or sternum). To obtain eGRF from acceleration using three sensors, Newton's second law is used as shown in Eq (1).

$$eGRF = (m_b \times g) + \sum_{i=1}^3 m_b \times WF_i \times (a_{z,i}) \quad (1)$$

m_b Represents the body mass of the subject, g the gravitational acceleration, i the sensor number, WF the weight factor and a_z the sensor free acceleration in vertical direction. The eGRF was normalized for body weight.

WF should correspond to the fraction of body mass corresponding to the segment where the sensor is attached. Since we have two symmetrically placed sensors on the legs, we assume the corresponding weight factors to be equal. Also, the sum of all weight factors should equal one to ensure the full body mass being used. Hence, only one parameter for WF was optimized in the case of three sensors.

Besides the weight factor, it is shown that IMU filter properties influence eGRF [9]. Therefore, the IMU filter cut-off frequencies and orders of the used sensors are included in the optimization. Again, due to symmetry, the filter settings for the leg sensors were assumed to be equal.

2.4. Model Optimization

An optimization was performed where one weight factor WF, two filter orders and two cut-off frequencies were optimized. This optimization is performed using the Optuna optimization framework [11]. The loss function is defined as the root mean squared error (RMSE) between the eGRF and mGRF during stance-phase. This value was minimized using an optimization with 1000 trials. The optimization was initiated 5 times and the best results were used. The weight factor was optimized between 0 and 1, cut-off frequencies between 0.5 and 50Hz and the orders between 2 and 10 with increments of 2.

2.5. Model Evaluation

The main performance metric is RMSE between mGRF and eGRF during stance phase. This metric is indicative for the accuracy over the full stance phase. Further, Pearson's correlation coefficient is calculated between mGRF and eGRF. Also, the absolute (active) peak error is calculated in bodyweights and as a relative value. To evaluate the generalizability of the models, leave-one-subject-out-cross-validation was performed for the best sensor configuration. The same principle was used to investigate the generalizability per velocity and stride frequency. Each time, one velocity or stride frequency was left out while the parameters were optimized for the other velocities or stride frequencies. Filter orders were fixed for this purpose based on the best results with the full data set.

Lastly, a sensitivity analyses on the optimized parameters is performed to inspect the sensitivity to changes in the cut-off frequencies and the weight factor. All combinations of cut-off frequencies between 4-20Hz (sensor on legs) and 4-12Hz (sensor trunk) were analysed, and the performance metrics were shown in a heatmap. For the weight factor, the values between 0.3 and 0.8 were analysed.

All data processing and optimizations were done in Python 3.8.

3. Results

With a RMSE of 0.179 body weight (BW), the pelvis-tibias configuration had the best performance (Table 1). This configuration also shows the lowest peak error (3.66%) and a very strong correlation between the eGRF and mGRF ($r=0.98$). From the WF, it can be concluded that the sensor on the trunk has a higher contribution to the estimate than the sensors on the lower body. For the configurations involving the feet, it is seen that there is hardly any contribution from the feet sensors (WF=0.93 and 0.97).

The leave-one-subject-out-cross-validation shows variation in the performance per subject, a RMSE between 0.130 and 0.235BW is achieved (Table 2). For all subjects, there is a very strong correlation between eGRF and mGRF. The model parameters are relatively stable, with WF between 0.52 and 0.58, F_c pelvis between 5.73 and 6.59Hz and F_c tibias between 8.19 and 9.11Hz.

Table 1: Effect of sensor configuration on estimated ground reaction force and model parameters

Sensor configuration	RMSE (BW)	Absolute mean peak error (BW)	Relative mean peak error (%)	Pearson's r	WF pelvis / sternum	Fc pelvis / sternum	Fc tibias / thighs / feet	Order pelvis / sternum	Order tibias / thighs / feet
Pelvis + tibias	0,179	0,089	3,68	0,98	0,54	6,24	8,62	4	2
Sternum + tibias	0,183	0,128	5,30	0,97	0,56	6,74	7,47	4	2
Pelvis + thighs	0,214	0,118	4,88	0,97	0,60	4,21	5,69	4	4
Sternum + thighs	0,200	0,096	4,00	0,97	0,46	6,68	4,58	8	2
Pelvis + feet	0,233	0,124	5,15	0,96	0,93	4,49	1,40	6	6
Sternum + feet	0,224	0,123	5,12	0,96	0,97	5,23	31,50	4	10

Table 2: Leave-one-subject-out-cross-validation using the pelvis-tibias configuration. Orders were fixed at 4 (pelvis) and 2 (tibias)

Subject	RMSE (BW)	Absolute mean peak error (BW)	Relative mean peak error (%)	Pearson's r	WF pelvis	Fc pelvis (Hz)	Fc tibia (Hz)
02	0,187	0,051	1,87	0,98	0,58	6,01	9,29
04	0,130	0,052	2,30	0,99	0,57	6,17	9,11
05	0,158	0,090	3,57	0,99	0,59	5,87	9,22
07	0,168	0,062	2,58	0,98	0,56	6,24	8,81
08	0,168	0,066	2,80	0,98	0,58	6,08	8,93
09	0,228	0,176	6,83	0,97	0,62	5,51	9,34
12	0,137	0,116	4,82	0,99	0,57	6,24	9,04
13	0,194	0,054	2,21	0,98	0,56	6,25	8,95
14	0,235	0,051	2,10	0,95	0,56	6,26	9,10
15	0,204	0,084	4,04	0,96	0,55	6,53	8,82
Mean	0,181	0,080	3,31	0,98	0,57	6,11	9,06

Model performance decreases with increasing speed, however, still a RMSE of 0.20BW is achieved at 14km/h (Table 3, Fig. 1). Also, the model parameters remain relatively constant. Performance of the model is best at high and preferred stride frequency (Table 4). At a lower stride frequency, the impact peak (at ~15% stance) is more prominent and also the active peak (mid-stance) is slightly higher in the mGRF (Fig. 2). The largest deviation between eGRF and mGRF is in the first 25% of the stance phase.

The sensitivity analyses shows that there is a surface between 7-10Hz for the tibias and 5.5-7.5Hz for the pelvis with a low RMSE, relative mean peak error and high Pearson's correlation coefficient (Fig. 3). It is seen that the optimization curves for the different outcome measures differ. A weight factor for the pelvis between 0.45 and 0.65 gives a RMSE <0.19BW, a relative mean peak error <4% and a Pearson's correlation coefficient >0.97 (Table 5).

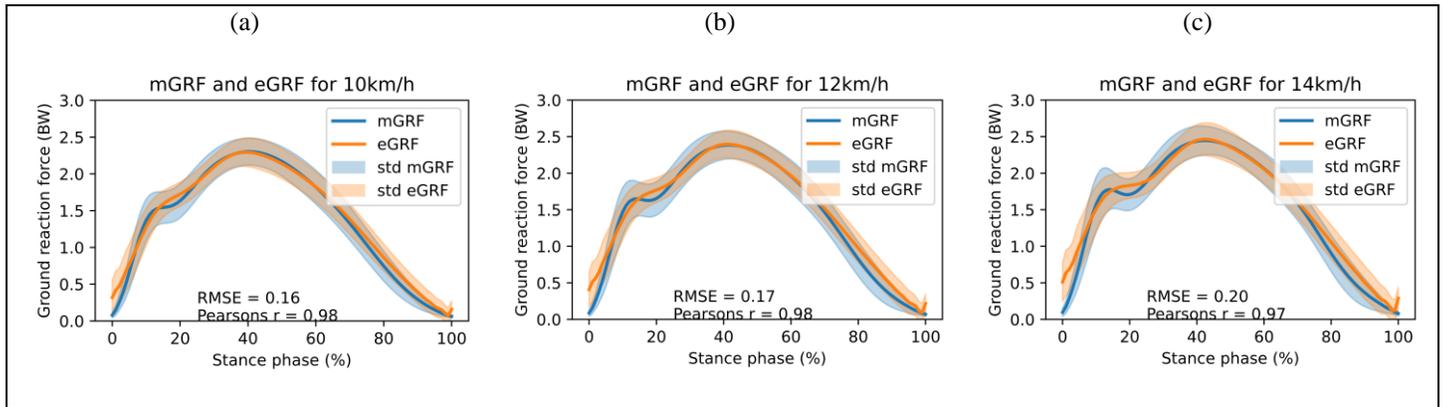


Figure 1: Measured and estimated GRF for 10, 12 and 14km/h, using the pelvis-tibias configuration. Filter orders were fixed at 4 (pelvis) and 2 (tibias)

Table 3. Performance per velocity using the pelvis-tibias configuration. Filter orders were fixed at 4 (pelvis) and 2 (tibias)

Speed (km/h)	RMSE (BW)	Absolute mean peak error (BW)	Relative mean peak error (%)	Pearson's r	WF pelvis	Fc pelvis (Hz)	Fc tibias (Hz)
10	0,163	0,085	3,63	0,98	0,55	6,11	8,42
12	0,171	0,084	3,47	0,98	0,56	6,07	8,64
14	0,204	0,099	3,99	0,97	0,52	6,37	8,66

Table 4: Performance per stride frequency using the pelvis-tibias configuration. Orders were fixed at 4 (pelvis) and 2 (tibias)

Relative stride frequency	RMSE (BW)	Absolute mean peak error (BW)	Relative mean peak error (%)	Pearson's r	WF pelvis	Fc pelvis (Hz)	Fc tibias (Hz)
Low (90%)	0,196	0,093	3,78	0,97	0,54	6,03	8,48
Preferred (100%)	0,171	0,084	3,47	0,98	0,54	6,34	8,44
High (110%)	0,168	0,087	3,74	0,98	0,56	6,19	8,50

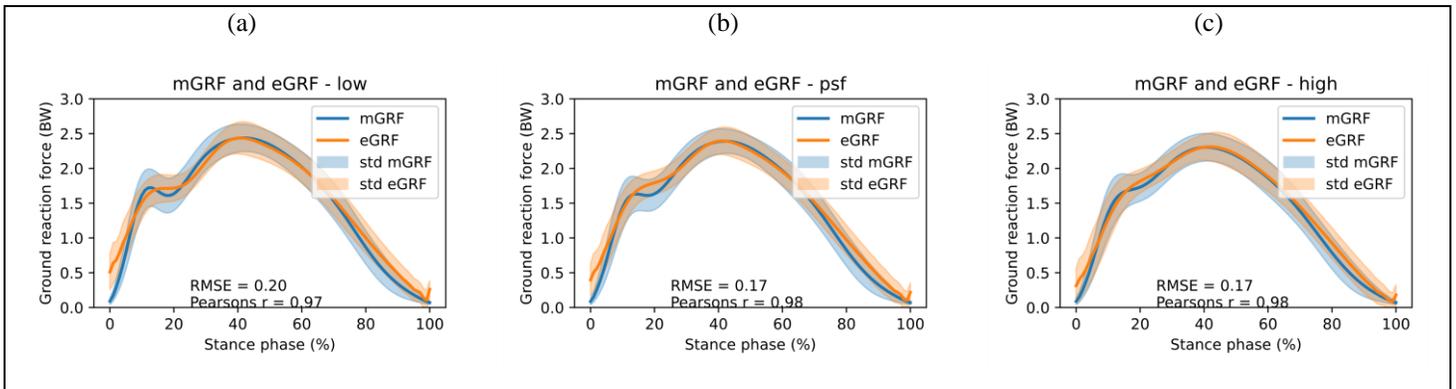


Figure 2: mGRF and mGRF for low (a), preferred (b) and high (c) stride frequency, using the pelvis-tibias configuration at all the speeds. Filter orders were fixed at 4 (pelvis) and 2 (tibia)

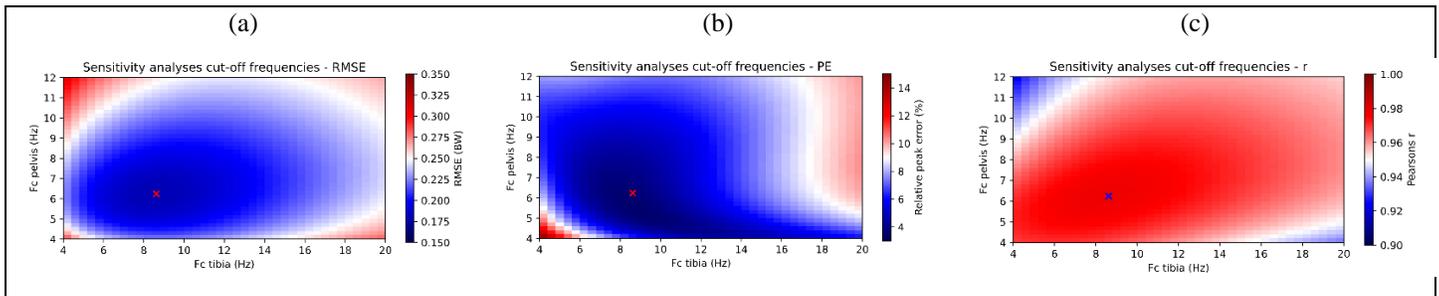


Figure 3: Sensitivity analyses with cut-off frequencies. (a) represents the effect on the RMSE, (b) the effect on peak error and (c) the effect on Pearson's correlation coefficient. Weight factor was fixed at optimum value for the whole data set (0.54). The cross (x) indicates the optimum value for the RMSE value.

Table 5: Sensitivity analyses with the weight factor. Cut-off frequencies were fixed at the optimum values for the whole data set (6.24Hz for pelvis and 8.62Hz for tibia). The optimum was found at 0.54

Weight factor pelvis	0,30	0,35	0,40	0,45	0,50	0,55	0,60	0,65	0,70	0,75	0,80
RMSE (BW)	0,220	0,206	0,194	0,185	0,180	0,179	0,182	0,188	0,198	0,210	0,225
Relative mean peak error (%)	4,52	4,22	3,97	3,79	3,70	3,70	3,78	3,95	4,22	4,60	5,13
Pearson's r	0,963	0,968	0,972	0,975	0,977	0,978	0,977	0,976	0,974	0,970	0,966

4. Discussion

The purpose of this study was estimate vertical GRF with three IMUs. Using the pelvis-tibias configuration, a lower RMSE ($<0.235\text{BW}$, 0.179BW averaged over all data) compared to other work (0.27BW [7] and 0.28BW [12]) is achieved. The model shows a very strong correlation (Pearson's $r=0.98$) with the measured GRF, comparable to other work ($r=0.90-0.99$ [7]). It is seen that accuracy decreases with speed but increases with stride frequency, indicating that as forces become larger, the estimation becomes less accurate.

As seen from the weight factor, the major contribution for the GRF estimation came from the sensor on the trunk (except for the sternum-thighs configuration). This is explained by the fact that the sensor on the trunk is closer to the centre of mass of the body. With the sternum-feet configuration, there was hardly any contribution from the feet sensors (WF=0.97), but there was still a RMSE of 0.224BW. It is expected that a configuration with only one sternum sensor will have very similar results. Future research could investigate the accuracy of GRF estimation with only one IMU.

As shown with the leave-one-subject-out cross validation, there are differences in performance per subject. So is the lowest RMSE of 0.130BW found for S04, the largest value found for S014 and is the RMSE 0.181BW on average. If the algorithm is used on newly measured heel strike runners, it is expected that the RMSE falls within that range.

As the parameters of the optimization in the leave-one-subject-out cross validations are within a small range, it can be assumed that adding more subjects would not have much impact on the optimum parameters of the algorithm. The sensitivity analyses also indicates that, if the optimized values are changed (for example because more subjects are added) it will not have much influence on the overall performance.

Although the majority of runners has a heel strike [13], the model should be validated for other foot strike patterns as well to be applicable to non-heel strikers. It is known that GRF shape is dependent on the foot strike pattern, thus it might be the case that a different cut-off frequencies or weight factors are needed. Furthermore, the subjects included in this study were all experienced runners (>15km of running per week for 6 months). The algorithm should be validated on non-experienced runners as well.

The plots in Fig. 2 indicate that most of the error is seen in the first 25% of the stance phase (impact peak). Although peak GRF is a measure of interest [14], also loading rate or impact peak are used [15], where the latter two are extracted from the first 25% of the stance phase. This is a limitation of this algorithm. A solution could be to redefine the loss function to prioritize the first 25% of the stance phase. However, this will go with a decrease in performance over the complete stance phase.

The current research was performed on a treadmill. Before applying this method in practice, the implications of outdoors effects should be investigated. For example, the effects of ground surface and running on slopes should be investigated.

In this study, the same algorithm is used for male and female runners. However, it is known that the mass distribution is different between male and female. With the used algorithm, optimum WF is influenced by the mass distribution. This implies that it can be beneficial to have a separate model for male and female runners.

A future direction to improve the estimate of vertical GRF can be by combining a physical model with a machine learning model in a so-called hybrid model [16].

4. Conclusion

The main goal of this study was to develop a model to estimate vertical GRF using three IMUs, robust for different speeds, stride frequencies and subjects. This study shows that it is possible to estimate vertical GRF outside the lab for heel strike runners at different velocities and stride frequencies with only 3 IMUs. This could help in the future to quantify biomechanical load in the runners' own environment.

References

- [1] R. N. Van Gent, D. Siem, M. Van Middelkoop, A. G. Van Os, S. M. A. Bierma-Zeinstra, and B. W. Koes, "Incidence and determinants of lower extremity running injuries in long distance runners: A systematic review," *Br. J. Sports Med.*, vol. 41, no. 8, pp. 469–480, 2007, doi: 10.1136/bjsm.2006.033548.
- [2] Bertelsen, M. L., Hulme, A., Petersen, J., Brund, R. K., Sørensen, H., Finch, C. F., Parner, E. T., & Nielsen, R. O. (2017). A framework for the etiology of running-related injuries. *Scandinavian Journal of Medicine & Science in Sports*, 27(11), 1170–1180. <https://doi.org/10.1111/sms.12883>
- [3] J. Verheul, N. J. Nedergaard, J. Vanrenterghem, and M. A. Robinson, "Measuring biomechanical loads in team sports—from lab to field," *Sci. Med. Footb.*, vol. 4, no. 3, pp. 246–252, 2020, doi: 10.1080/24733938.2019.1709654.
- [4] S. S. Yeung, E. W. Yeung, and L. D. Gillespie, "Interventions for preventing lower limb soft-tissue running injuries,"

Cochrane Database Syst. Rev., no. 7, 2011, doi: 10.1002/14651858.cd001256.pub2.

- [5] E. Charry, W. Hu, M. Umer, A. Ronchi, and S. Taylor, “Study on estimation of peak Ground Reaction Forces using tibial accelerations in running,” in *Proceedings of the 2013 IEEE 8th International Conference on Intelligent Sensors, Sensor Networks and Information Processing: Sensing the Future, ISSNIP 2013*, 2013, vol. 1, pp. 288–293, doi: 10.1109/ISSNIP.2013.6529804.
- [6] B. LeBlanc, E. M. Hernandez, R. S. McGinnis, and R. D. Gurchiek, “Continuous estimation of ground reaction force during long distance running within a fatigue monitoring framework: A Kalman filter-based model-data fusion approach,” *J. Biomech.*, vol. 115, p. 110130, Nov. 2021, doi: 10.1016/j.jbiomech.2020.110130.
- [7] Wouda, F. J., Giuberti, M., Bellusci, G., Maartens, E., Reenalda, J., van Beijnum, B.-J. 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, 1–14. <https://doi.org/10.3389/fphys.2018.00218>
- [8] A. Ancillao, S. Tedesco, J. Barton, and B. O’flynn, “Indirect measurement of ground reaction forces and moments by means of wearable inertial sensors: A systematic review,” *Sensors (Switzerland)*, vol. 18, no. 8. 2018, doi: 10.3390/s18082564.
- [9] E. M. Day, R. S. Alcantara, M. A. McGeehan, A. M. Grabowski, and M. E. Hahn, “Low-pass filter cutoff frequency affects sacral-mounted inertial measurement unit estimations of peak vertical ground reaction force and contact time during treadmill running,” *J. Biomech.*, vol. 119, Apr. 2021, doi: 10.1016/J.JBIOMECH.2021.110323.
- [10] Komaris, D. S., Perez-Valero, E., Jordan, L., Barton, J., Hennessy, L., O’Flynn, B., & Tedesco, S. (2020). Effects of segment masses and cut-off frequencies on the estimation of vertical ground reaction forces in running. *Journal of Biomechanics*, 99, 109552. <https://doi.org/10.1016/J.JBIOMECH.2019.109552> [11] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A Next-generation Hyperparameter Optimization Framework,” *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 2623–2631, Jul. 2019, doi: 10.1145/3292500.3330701.
- [12] Y. Ohtaki, K. Sagawa, and H. Inooka, “A Method for Gait Analysis in a Daily Living Environment by Body-Mounted Instruments,” *JSME Int. J. Ser. C*, vol. 44, no. 4, pp. 1125–1132, 2001, doi: 10.1299/jsmec.44.1125.
- [13] Larson, P., Higgins, E., Kaminski, J., Decker, T., Preble, J., Lyons, D., McIntyre, K., & Normile, A. (2011). Foot strike patterns of recreational and sub-elite runners in a long-distance road race. <https://doi.org/10.1080/02640414.2011.610347>, 29(15), 1665–1673.
- [14] Kiernan, D., Hawkins, D. A., Manoukian, M. A. C., McKallip, M., Oelsner, L., Caskey, C. F., & Coolbaugh, C. L. (2018). Accelerometer-based prediction of running injury in National Collegiate Athletic Association track athletes. *Journal of Biomechanics*, 73, 201–209. <https://doi.org/10.1016/j.jbiomech.2018.04.001>
- [15] C. D. Johnson, A. S. Tenforde, J. Outerleys, J. Reilly, and I. S. Davis, “Impact-Related Ground Reaction Forces Are More Strongly Associated With Some Running Injuries Than Others,” *Am. J. Sports Med.*, vol. 48, no. 12, pp. 3072–3080, Oct. 2020, doi: 10.1177/0363546520950731.
- [16] R. D. Gurchiek, N. Cheney, and R. S. McGinnis, “Estimating biomechanical time-series with wearable sensors: A systematic review of machine learning techniques,” *Sensors (Switzerland)*, vol. 19, no. 23, 2019, doi: 10.3390/s19235227.