Estimating vertical ground reaction forces from plantar pressure using interpretable high-dimensional approximation

Laura Weidensager[1] ⬤ · **Dominik Krumm**[2] · **Daniel Potts**[1] · **Stephan Odenwald**[2]

Abstract

Several studies have reported methods for capturing ground reaction forces in the field. However, these forces are usually measured indirectly with non-interpretable "black box" models. Here we report an interpretable model with a function approximation algorithm. The model uses time series of plantar pressure from instrumented insoles to estimate vertical ground reaction forces. The study included data from 16 persons moving at different speeds on a two-belt treadmill equipped with force plates. The introduced regression model, based on a high-dimensional approximation, using only low-dimensional variable interactions, demonstrated its ability to estimate vertical ground reaction forces. The normalized mean square error for velocities between 3 and $9 \, \mathrm{km \, h^{-1}}$ ranged from 10.6 to 24.4 %. The accuracy of the presented approach, which can be used to analyze and interpret the learned model, was comparable to that reported in the literature. Furthermore, the evaluation of the learned model is particularly suitable for embedded and portable systems and, after a one-time calibration measurement, allows permanent and laboratory-independent measurement of vertical ground reaction forces.

1 Introduction

The generation of sport-specific load-time profiles is often based on the measurement of ground reaction forces (GRFs) and their time profile [8]. GRFs describe how humans interact with their surroundings [5]. They are crucial for understanding individual gait patterns [20]. They also form the basis for calculating inverse dynamics in biomechanical studies [5].

Force plates are often used to record these GRFs because they are considered the "gold standard". Force plates contain three to four interconnected force sensors for measuring GRFs and ground reaction moments in all three spatial directions. A disadvantage of force plates is that they must be hit "correctly", i.e., with the complete foot and ideally in the center, otherwise measurement errors might occur.

This circumstance leads to the undesirable effect of "targeting" in some individuals, i.e., the movement of the person is unconsciously changed in order to hit the force plate and the recorded kinetic parameters are thus "unnatural" [11]. Other disadvantages include the fact that multiple force plates have to be used to record left and right foot contacts [2] and that they are generally not portable as they are usually firmly attached to the foundation of a laboratory [11]. Furthermore, it is not possible to measure consecutive steps, which means that the person being studied has to perform many measurement repetitions to obtain a statistically validated "average step" [14].

The above challenges in using standard force plates are addressed by developing treadmills with integrated force plates and mobile force plates. Nevertheless, these developments are not free of challenges either. Indeed, with instrumented treadmills it is possible to measure GRFs over a longer period of time and not only a specific time point [17]. However, the literature questions whether running on a treadmill is comparable to natural running [18, 24] and whether GRFs can be recorded accurately [25]. The use of mobile force plates, on the other hand, is more likely to result in targeting. This is presumably due to the fact that

✉ Laura Weidensager
laura.weidensager@math.tu-chemnitz.de

1 Applied Functional Analysis, Chemnitz University of Technology, 09126 Chemnitz, Germany

2 Sports Equipment and Technology, Chemnitz University of Technology, 09126 Chemnitz, Germany

they are often placed in a position that is exposed to the individual being tested and thus consciously perceptible.

Given all these existing challenges and according to the findings in literature [1, 3, 5, 8, 22], there is an increasing need for continuous recording of GRFs under field conditions (i.e., in real-life situations). Having more comprehensive and representative measurements from the field, would allow to facilitate the understanding of environmental influence and activity-specific movement characteristics [1]. To overcome the aforementioned challenges of measuring GRFs under laboratory conditions and to exploit the potential of monitoring GRFs in real-world environments, several techniques and technologies using wearable sensors have been developed. An overview of the different approaches, which have been divided into three categories, is provided in the systematic review by Shahabpoor and Pavic [22].

One approach to record GRFs in the field is based on systems with pressure measurement insoles. These systems do not record GRFs directly. Rather, they are calculated indirectly via mathematical relations from the measured plantar pressures. Hence, the evaluation model used has a crucial influence on the accuracy of the calculated GRFs. While older studies mostly used "simple" evaluation models, such as linear function and coordinate transformation, more recent studies increasingly use "complex" evaluation models based on multivariate statistics, time–frequency analysis, skeletal muscle models, and artificial neural networks [8]. The advantage of the "complex" compared to the "simple" evaluation models is that the measurement uncertainty can be reduced and, moreover, the horizontal GRFs and the spatial ground reaction moments can be derived in addition to the vertical GRFs. However, these machine learning models are a "black box". Therefore, one does not know how exactly they work.

In this work, we are therefore pursuing an interpretable function approximation algorithm. This approach should help us to understand the relation between the plantar pressures measured with wireless pressure measurement insoles and the vertical GRFs. Furthermore, we construct a prediction model that, unlike current models, uses only the pressure measurements at a particular time stamp rather than the entire step or gait cycle. Therefore, our approach is qualified for implementation in a real-time system.

2 Methods

2.1 Participants, equipment and procedures

The study was approved by the institution's ethics committee (reference #101525731). Eighteen persons (5 females, 13 males; characteristics (mean ± SD): age 29.5 ± 6.9 years,

height 178.7 ± 9.2 cm, body mass 74.8 ± 12.4 kg, body mass index 23.3 ± 2.3 kg m^{-2}) met the inclusion criteria, volunteered to participate in the study, and provided written informed consent. However, in our study, we only included data from 16 persons because we did not include the data from the person in the pilot study and excluded the data from another person due to data synchronization issues.

To collect data on bipedal locomotion on a treadmill, the Gait Real-time Analysis Interactive Lab (GRAIL, Motek Medical B.V., Houten, NLD), two body-attached sensor networks (BASN) with integrated accelerometers and gyroscopes (DIALOGG Dataloggers, ENVISIBLE GmbH, Chemnitz, DEU; [4]), and a pair of pressure insoles (Smart footwear sensors/HD 002, IEE, Echternach, LUX) were used (Fig. 1). The GRAIL consisted of a two-belt treadmill with integrated force plates (M-Gait, Motek Medical B.V., Houten, NLD) mounted on a motion frame with two degrees of freedom.

Participants' running shoes were equipped with pressure insoles placed under the inner sole of the shoe. Two DIALOGG Dataloggers were attached with tape to the lateral side of each shoe. After ensuring that the dataloggers were operating correctly and the two force plates were set to zero, participants proceeded to the treadmill to begin measurements.

Measurements were divided into three runs with a total duration of 40 minutes. To cover the worst-case scenario and account for participants' varying levels of familiarity with treadmill running, exclusively the data from the first run of each person was used for this work. Run 1 began with a 2-min warm-up at a speed of 3 km h^{-1}, followed by a 1-min rest. The participants then completed three 3-min exercise tasks in succession, i.e., walking at 4 km h^{-1}, hiking at 6 km h^{-1}, and running at 9 km h^{-1}. There was a 1-min break before the end of the measurement. The average number of loading cycles to which the sensors were subjected during this first run was 659, see Table S1 for details. Run 2 lasted 11 minutes and run 3 lasted 16 minutes. In between each run, there was a pause of at least 1 minute.

The data of the pressure sensors (insoles) were sampled with 100 Hz and the GRFs of the force plates with 1000 Hz. There was no synchronization between the pressure and force measurements. For this reason, the experimenter tried to start the measurement systems as simultaneously as possible.

2.2 Analysis

The GRF data in vertical direction $F_v(t)$ is a time series, as well as the data from the pressure insoles, which we denote by $p_1(t), \dots, p_8(t)$. The data-stream of $F_v(t)$ was downsampled to 100 Hz. Both data streams have to be synchronized.

Fig. 1 The experimental setup. Upper left: Pressure insole with eight sensors connected to a body-attached sensor network. Lower left: Running shoe with inserted insole. Right: Person on two-belt treadmill equipped with force plates

This was done best by exploiting the first step after the standing phase from $F_v(t)$ and $p_{\text{sum}}(t) := \sum_{i=1}^{8} p_i(t)$. The vertical components of the GRF, $F_v(t)$, were set to zero when the sum $p_{\text{sum}}(t)$ was zero, i.e., the foot was in the air.

After trimming all time series (pressures $p_i(t)$ and vertical forces $F_v(t)$) by steps, $F_v(t)$ were denoised using a wavelet transform for each step separately. The parameters were adapted to the velocity of the steps. The resting phases were omitted. Since the aim is a model, which is independent of the velocity, the data from all velocities were put together. The right and left foot were treated separately. This approach does not require a normalization of the steps to the same length. Instead, all timestamps are used as data.

In average, we used about $3.6 \cdot 10^4$ time stamps per foot for every person, which are about $1.2 \cdot 10^6$ in total.

2.3 Regression model

We aimed for searching the correlation between the pressure data $p_i(t)$ and the vertical forces $F_v(t)$. This can be modeled by the 8-dimensional function $f : [0,1]^8 \to \mathbb{R}$, where we have

$$F_v(t) \approx f(p_1(t), \ldots, p_8(t)). \tag{1}$$

We used min–max scaling, which scaled features to lie in the interval $[0,1]$. We used an interpretable learning algorithm, which was based on high-dimensional function approximation. The algorithm is available in the Julia package [21].

High-dimensional approximation suffers from the "curse of dimensionality". For that reason, the ANOVA (Analysis of variances) approach in our algorithm uses an iterative least squares algorithm to construct the ANOVA terms of order 1 and 2 by the approximation

$$F_v \approx f := f_\varnothing + \sum_{i=1}^{8} f_i(p_i) + \sum_{i,j=1}^{8} f_{i,j}(p_i, p_j) \tag{2}$$

$$= f_\varnothing + \sum_{i=1}^{8} \sum_{k=1}^{N_1} a_k e_k(p_i) + \sum_{i,j=1}^{8} \sum_{k_1,k_2=1}^{N_2} a_{k_1,k_2} e_{k_1}(p_i) \cdot e_{k_2}(p_j), \tag{3}$$

which has variable interactions up to order 2. The constant $f_\varnothing$ is approximately the mean of the function f, i.e., the mean of the measured force F_v. Using only terms up to order 1 gives slightly worse results, which is why we kept the two-dimensional terms. The ANOVA terms $f_{\boldsymbol{u}}(p_{\boldsymbol{u}})$ for $\boldsymbol{u} \subset \{1, \ldots, 8\}$ are described by solving a least squares system using some basis functions iteratively. Suitable and implemented are the half-periodic-cosine $e_k(x) = \cos(\pi k x)$ and a wavelet basis, where e_k are the periodic Chui–Wang wavelets. Since the choice of the basis does not influence the approximation errors too much, we worked with the wavelet basis. Furthermore, this basis has linear complexity, such that fast algorithms are available. The input values of the algorithm are the pressure measurements $p_i(t)$, $i \in \{1, \ldots, 8\}$ and $F_v(t)$ for all training data $t \in T_{\text{train}}$. After using an iterative least squares algorithm, the output values are the base

coefficients of the function, see references [10, 16] for theoretical foundations and examples. Uniformly distributed data have been shown to work well in theory, but the data here are not uniform. Therefore, the method from Lippert and Potts [9] for transforming the data was applied.

This prediction of F_v at time t only depends on the pressure measurements of the sensors at time t, i.e., $p_i(t)$ and not on the whole gait cycle.

2.4 Assessing model performance

To assess model accuracy, we performed a series of estimations on withheld data and calculated the mean performance measures across intra-person and inter-person groupings. Denote the sample of test data by $\mathcal{T}$. We used the normalized root mean square error

$$\text{NRMSE} = \frac{1}{\overline{F_v}} \sqrt{\frac{\sum_{t \in \mathcal{T}} \left(F_v(t) - f(t) \right)^2}{|\mathcal{T}|}} \tag{4}$$

as measure of model performance, where $\overline{F_v} = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} F_v(t)$ denotes the mean of the test data. Furthermore, we calculated Pearson's correlation coefficient

$$\text{PCC} = \frac{\sum_{t \in \mathcal{T}} f(t) \cdot F_v(t) - |T| \cdot \overline{F_v} \cdot \overline{f}}{\sqrt{\sum_{t \in \mathcal{T}} F_v(t)^2 - |T| \cdot \overline{F_v}^2} \times \sqrt{\sum_{t \in \mathcal{T}} f(t)^2 - |T| \cdot \overline{f}^2}}. \tag{5}$$

We considered different data splitting techniques. In the case considered here, learning a model should be performed person-specific owing to individual differences in the sizes of feet and gait. The first evaluation (E1) used the data from every person separately, i.e., a person-specific model was constructed. For the split into training and testing data, it is important to split with respect to the different velocities to avoid local information and to improve the bias properties of the predictors. Each of the four parts belonging to the different velocities was split into five parts. The testing data consisted of one part from every velocity respectively, i.e., 80/20-splitting into T_{train} and T_{test}. These parts were chosen randomly.

For comparison, we also trained models using data from all persons. To compare the data from different persons, $F_v(t)$ were normalized to the body-weight. The second evaluation (E2) used the testing and training data received from the previous procedure for all persons together. And in the third evaluation (E3) the data from three randomly chosen persons were used as testing data.

We performed the model learning and evaluation five times for each evaluation case for left and right foot respectively, and presented the mean of the NRMSE and PCC. To assess the model performance, we analyzed the prediction error in the testing data with respect to the different velocities v.

2.5 Parameter choice

The bandwidths N_1 and N_2 in (3) are to be chosen depending on the number of samples to avoid over- or underfitting. From a theoretical point of view, logarithmic oversampling is a good choice, i.e., $M \approx N \log N$, where M is the number of total training time stamps in T_{train} and N is the total number of used coefficients in the approximation, which depends on N_1, N_2 and the chosen basis. In the data set considered here, a large oversampling also serves as regularization. The person-specific models were trained using about 2400 coefficients in total.

3 Results

Figure 2 illustrates results for the person-specific models from evaluation E1 for one person (person 15). The first row shows some example data from this person together with the predicted vertical GRFs f. Since for the prediction of $F_v(t)$ only the pressure measurements $p_i(t)$ are used, the resulting prediction $f(t)$ is not as smooth as the measured F_v. An additional smoothing of f after prediction is possible, but this does not retain the property of using only the pressures at the time stamp $p_i(t)$ for predicting $F_v(t)$.

The second row in Fig. 2 shows the mean GRFs (measured and predicted) for all steps in the test data, split into the different velocities. In the third row, we plotted the relative errors between predicted f and measured F_v, where the mean and the standard derivation over all steps in the different velocities are shown. The NRMSEs for this person were 7.0 % ($v = 3 \, \text{km h}^{-1}$), 6.7 % ($v = 4 \, \text{km h}^{-1}$), 10.2 % ($v = 9 \, \text{km h}^{-1}$) and 18.2 % ($v = 12 \, \text{km h}^{-1}$). The lowest NRMSE 6.0 % was reached by person 5 at the speed $v = 4 \, \text{km h}^{-1}$, and the highest NRMSE 39.6 % by person 13 at the highest speed $v = 12 \, \text{km h}^{-1}$. For more details, see supplementary Table S1.

Table 1 summarizes the results of the different evaluation approaches. The results for E1 are averaged over all persons. The person-specific model E1 is better than the other strategies E2 and E3, since the NRMSE are smaller. The NRMSE for velocities between 3 and 9 km h^{-1} ranged from 10.6 to 24.4 % for the person-specific model (E1). Furthermore, in general (with one exception) the lower the speed of the person, the smaller the NRMSE.

The eight one-dimensional ANOVA terms $f_i(p_i)$ from (2) are plotted in Fig. 3 for the person-specific trained model for some of the persons (2, 4, 15). The total prediction f for eight given pressure measurements p_i
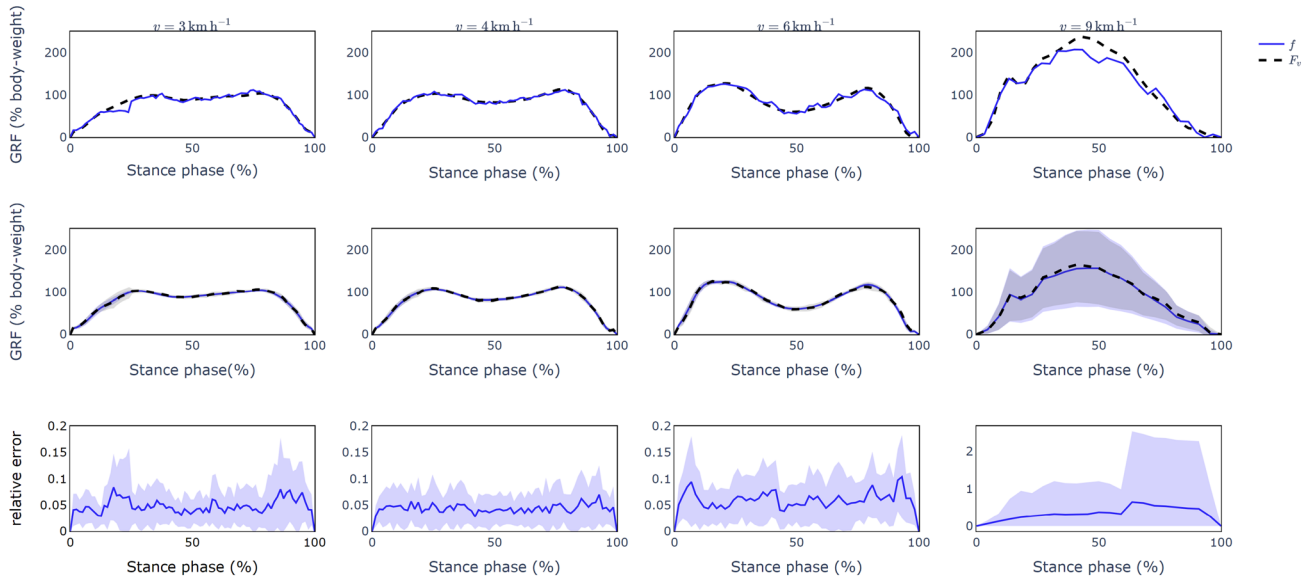
Fig. 2 Approximation results from one exemplary person (15), split in the different velocities. First row: One example step from testing data of every velocity. Plotted is the measured vertical GRF F_v (black) and the predicted force f (blue). Second row: Mean vertical GRFs (measured and predicted) for all steps in the test data. Third row: Mean relative errors between predicted f and measured F_v. The standard deviations are shaded respectively

Table 1 Results of the different evaluation approaches

	NRMSE				PCC			
v (km h^{-1})	3	4	6	9	3	4	6	9
E1	0.112	0.106	0.137	0.244	0.946	0.948	0.933	0.918
E2	0.201	0.195	0.232	0.359	0.843	0.846	0.826	0.838
E3	0.318	0.341	0.394	0.529	0.715	0.675	0.639	0.659

The results for E1 are averaged over all persons. A detailed list for the approximation errors for the different persons can be found in the supplementary Table S1

($i = 1, \ldots, 8$) is the sum in (2), i.e., the sum of the eight plotted terms, all two-dimensional terms $f_{i,j}(p_i, p_j)$ and the constant $f_\varnothing$.

4 Discussion

4.1 Principal findings

This study suggests that there is a connection of the form (1), with a function f of low effective dimensionality (only ANOVA terms up to order 2 involved). In total the best relative NRMSE was reached for the velocity 4 km h^{-1}, this belongs to the PCC 94.8 %.

The relative NRMSE values given are valuable indicators of the quality of the model, enabling users to assess its performance based on their specific objectives. The acceptability of these values will naturally depend on the particular application and the level of precision required.

Four different velocities were studied. Table 1 and Fig. 2 show different results for the different velocities. The most important finding is that the approximation for the highest velocity has the highest errors. But in comparison to the other velocities, there is more deviation in $F_v(t)$ as well as in the approximation f in comparison to the mean step, which makes prediction more difficult in that case. In the work of Masani et al. [12], an increasing trend in the variability of the GRFs with increasing velocity was found. Consequently, we also expected more noise in the measured data at higher velocities. Another aspect is that we had to downsample the measured vertical GRF $F_v(t)$ from 1000 to 100 Hz, because of the lower sampling rate of the pressure sensors. This frequency could be too small for the velocity $v = 9$ km h^{-1}. The prediction model proposed in this study may be poor if the target data is far outside the scope of training data, such as running at much higher velocity than 9 km h^{-1}, but for moving at a velocity intermediate to the training data velocity, we expect good
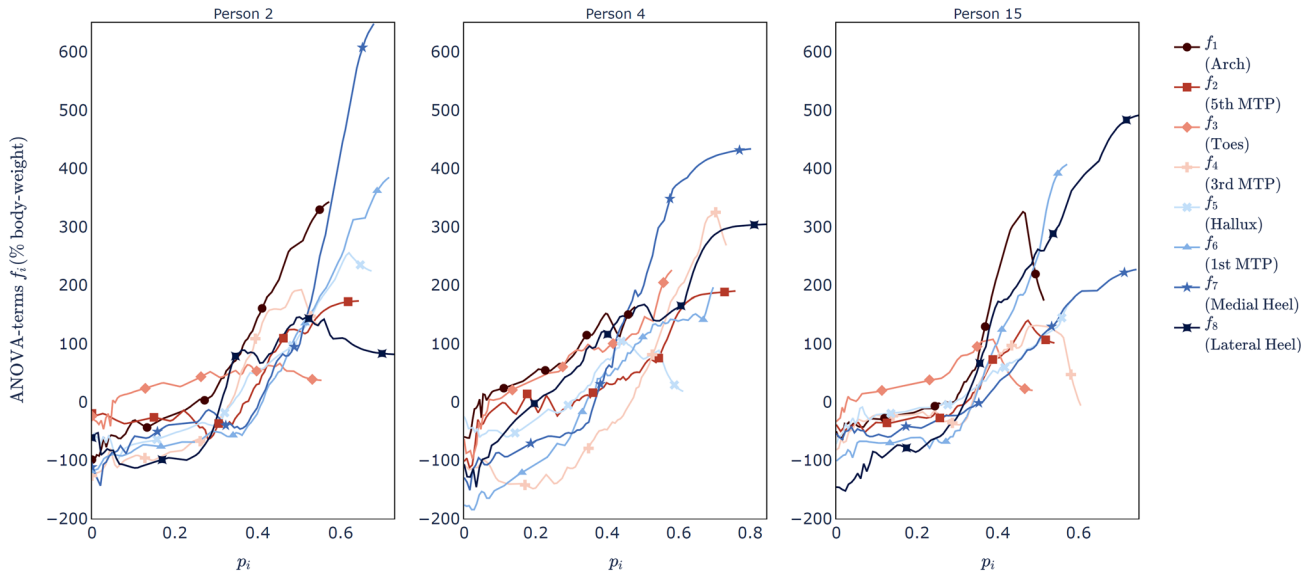
Fig. 3 Learned function for the person-specific models from persons 2, 4 and 15. Plotted are the one-dimensional ANOVA terms $f_i(p_i)$ $i = 1, \ldots, 8$, see (2). The sum over these terms together with the constant and the two-dimensional terms is the learned model $f(p_1, \ldots, p_8)$

results, since this data would be similar to the measured data.

The use of model (2) instead of black box predictions has the advantage that it is possible to analyze and interpret the learned model. Figure 3 shows that there are different gait styles depending on the person, which is reflected in different ANOVA terms. In Park et al. [15] the authors found that the gait of a person is as unique as their facial motion and finger impedance. In Table S1, this leads to different results for different persons. The walking behavior of some persons fits more into our approximation setting than that of other. The individuality is one reason why the evaluation techniques E2 and E3 give clearly bigger NRMSE and smaller PCC as the individual models. Especially the prediction of data from different test persons than the persons for the training data in E3, which is particularly desirable, does not yet work with our approach.

The individuality obviously also leads to different ANOVA terms $f_i(p_i)$ for the different persons (Fig. 3). In general, as it was expected and should be the case, the higher the pressures p_i, the higher the corresponding vertical GRF F_v. Furthermore, analysis of the resulting ANOVA terms provides insight into the behavioral pattern of the persons. Specifically, Fig. 3 shows that person 15 applies more pressure to sensor 8 (i.e., lateral heel). In contrast, persons 2 and 4 generate their pressure more on sensor 7 (i.e., medial heel). Beyond this, sensors 2 (i.e., 5th metatarsophalangeal joint) and 3 (i.e., phalanges III to V) play a more minor role in predicting F_v, as the ANOVA terms for these sensors have lower absolute values. Furthermore, it is possible to cluster

persons by their gait style using the ANOVA terms, which should be verified by a further study with more participants.

Since we work with two force plates, cross-loading yields to difficulties in the data. Especially in the data with velocity $v = 9 \, \text{km h}^{-1}$ the persons often step (partly) on the wrong plate. This shows in the different NRMSE in Table 1. Furthermore, the specific results in Table S1 coincide with the cross-loading, which often occur differently depending on the persons.

4.2 Possible applications

The Julia package ANOVAapprox was used for the model learning [21]. Therefore, only the denoised pressure measurements and the vertical GRF F_v have to be inserted in the algorithm, which could in principle be integrated in a data collection system. The model learning phase took about 30 s. Afterward, the prediction of F_v from field or laboratory pressure data measurements $p_i(t)$ basically consists of the set-up of a matrix of size $N \times M_{\text{test}}$ and one matrix–vector multiplication with the learned coefficient vector. Here N is the number of used coefficients and M_{test} is the number of testing data time stamps. It is also possible to do this evaluation in time, i.e., do the multiplication for every time stamp separately. Unlike other approaches in the literature [5, 23], we do not have to normalize the data with respect to the steps.

We propose the following procedure if a new person wants to capture the vertical GRF. For an individual trained model, the person should record some training data on the instrumented treadmill in the laboratory at different velocities, such that enough training data is available. To

determine the minimal amount of data required, we reduced the size of training data. Our findings indicate that training for just 2 minutes in total at varying velocities is sufficient to achieve testing errors similar to those presented in our results obtained using approximately 9 minutes of training data. Then we train a person-specific model by learning the coefficients a_k and a_{k_1,k_2} in (3) by solving the corresponding least squares system. With this learned model, it is possible to predict $F_v(t)$ from 8 new measurements $p_i(t), i \in \{1, \dots, 8\}$. Furthermore, splitting the training data into training and evaluation data gives the possibility to calculate the expected mean error for new unseen data, like in Table S1.

4.3 Possible improvements

While the investigation demonstrated that the insoles are a valid tool for in-shoe force measurement under the test conditions, several considerations should be taken into account when interpreting the results. Notably, the insoles only capture the vertical GRF, and fail to capture anterior–posterior or medial–lateral GRFs, similar as in Burns et al. [1]. Other literature [5, 7, 23] also considered these GRFs. We decided to not consider these components here. Due to the noise in the captured data, the approximation approach (1) would be less stable in this case. It turned out that these forces did not perform well using the approach (1) together with the low-dimensional approximation (2), so that it might not be a good task to find the connection (1).

The problem of cross-loading reduced the accuracy of our prediction. Especially at higher velocities this led to additional errors. Studying the prediction shows that in cross-loading cases our prediction f represents a more reliable gait curve than the actual measured F_v. On the other hand, the prevention of cross-loading by instructing the persons to hit the force plates "correctly", i.e., completely and ideally in the center, leads to the undesirable effect of "targeting" for some persons, which we wanted to avoid.

Recognizing the laboratory-based treadmill setting of our study, we must acknowledge the inherent simplifications, encompassing directional constraints, absence of natural stops, lack of terrain variability, and omission of environmental influences [19].

The study's capped maximum running speed of $v = 9 \, \text{km} \, \text{h}^{-1}$ could constrain the generalizability of the approach, particularly when extrapolating to faster-paced runners. Additionally, post-data acquisition synchronization was necessary due to hardware limitations, which, posed a moderate methodological limitation. To heighten methodological rigor and broaden applicability across running contexts, future research could encompass a wider speed range relevant to competitive and elite runners and incorporate direct synchronization in future measurements.

4.4 Comparison to other work

The main difference in our approach to other approaches in the literature is that we do not normalize the time with respect to the gait cycles. Furthermore, for our analyses, we omit the data (also in the test data) where all pressure sensors are zero, which belongs to $f(\mathbf{0}) = 0$. Our approximation predicts $F_v(t)$ to 0 during this time. This can lead to a higher NRMSE in comparison to testing on the whole gait cycles, which was done in most other studies. Therefore comparing only the resulting NRMSE can be misleading.

In Honert et al. [5], linear models were compared with recurrent neural networks for the accuracy of predicting the GRFs. For the recurrent neural network, they also used the pressure features at the time stamps directly before and after t for the prediction. They did not train person-specific models, but validated with a leave one out cross-validation. Overall the accuracy of our trained model was worse than the recurrent network, but better than the linear model. Nevertheless, Honert et al. [5] also found that a person-specific model training increases the accuracy.

In Eguchi et al. [3], a machine learning-based estimation of the vertical GRF was done. They used a linear least squares regression that fits the insole measurements during single leg stance to body-weight to learn a model. They also added constraints to the regression so that the estimates of the vertical GRF during walking have proper magnitude. In comparison to this study, we have a much larger data set from a treadmill available, which also includes data from trials with velocities up to $v = 9 \, \text{km} \, \text{h}^{-1}$. Therefore, it is complicated to compare the results. But especially for the lower velocity, our method gives lower NRMSEs.

Sim et al. [23] predicted the GRFs using a wavelet neural network. The insole there consists of 99 plantar sensors, where some principle components were selected. The wavelet neural network is similar to our approach using only ANOVA terms up to order 1, but they used different wavelet functions and learned additionally shift and scale parameters. However, the results are similar even though our pressure insole has only eight pressure sensors. Apart from that, and based on our own experience with both measurement systems, we can say that the pressure measurement insole used here seems to be much more robust. This assertion is supported by previous research by Melakessou [13], who highlighted the robustness of the pressure insole used, notably also to lateral shifts in plantar pressure loading. In addition, Hsiao et al. [6] reported that the insoles used by Sim et al. [23] pose a risk of sensor damage with repeated application of high pressure. They concluded that these insoles should be replaced after a reasonable period of use to ensure a high level of system accuracy and precision.

isea
INTERNATIONAL SPORTS ENGINEERING ASSOCIATION

5 Conclusions

We used pressure insoles for estimating the vertical ground reaction forces of a data set which include walking at 3 and 4 km h $^{-1}$, hiking at 6 km h $^{-1}$ and running at 9 km h $^{-1}$. A person-specific model was able to predict the ground reaction forces. We even provide a model that is interpretable, i.e., one can study the influences of the eight different pressure sensor inputs. Our approach, which allows, after a one-time calibration measurement, for permanent and laboratory-independent vertical ground reaction forces, could be used, for example, to guide training, to monitor rehabilitation progress when an athlete returns to play, or to detect asymmetries that may arise or propagate due to injury-related mechanisms.

Availability of data and materials Krumm, D., Koska, D., Wakode, J., Odenwald, S., & Maiwald, C. (2022, April 29). Long-duration kinetic and kinematic bipedal locomotion data on a treadmill at different speed. https://doi.org/10.17605/OSF.IO/KCQTF

Code availability The used Julia package can be found at https://github.com/NFFT/ANOVAapprox.jl [21].

Declarations

Conflict of interest The authors declare no conflict of interest. S.O. is Senior Advisor for ENVISIBLE GmbH.

References

1. Burns GT, Deneweth Zendler J, Zernicke RF (2019) Validation of a wireless shoe insole for ground reaction force measurement. J Sports Sci 37(10):1129–1138. https://doi.org/10.1080/02640414.2018.1545515
2. Cavanagh PR, Hewitt FG, Perry JE (1992) In-shoe plantar pressure measurement. Foot 2(4):185–194. https://doi.org/10.1016/0958-2592(92)90047-S
3. Eguchi R, Yorozu A, Fukumoto T et al (2020) Estimation of vertical ground reaction force using low-cost insole with force plate-free learning from single leg stance and walking. IEEE J Biomed Health Inform 24(5):1276–1283. https://doi.org/10.1109/JBHI.2019.2937279
4. Hill M, Hoena B, Kilian W et al (2016) Wearable, modular and intelligent sensor laboratory. Procedia Eng 147:671–676. https://doi.org/10.1016/j.proeng.2016.06.270
5. Honert EC, Hoitz F, Blades S et al (2022) Estimating running ground reaction forces from plantar pressure during graded running. Sensors 22(9):3338. https://doi.org/10.3390/s22093338
6. Hsiao H, Guan J, Weatherly M (2002) Accuracy and precision of two in-shoe pressure measurement systems. Ergonomics 45(8):537–555. https://doi.org/10.1080/00140130210136963
7. Jacobs D, Ferris D (2015) Estimation of ground reaction forces and ankle moment with multiple, low-cost sensors. J Neuroeng Rehabil. https://doi.org/10.1186/s12984-015-0081-x
8. Krumm D (2020) Methodische Aspekte bei der Entwicklung mechanischer Simulationen zur Messung der Funktionalitäten eines Handballschuhs [Methodological aspects used in the development of mechanical simulations to measure the functionalities of handball shoes]. Universitätsverlag Chemnitz
9. Lippert L, Potts D (2022) Variable transformations in combination with wavelets and ANOVA for high-dimensional approximation. arXiv: 2207.12826
10. Lippert L, Potts D, Ullrich T (2023) Fast hyperbolic wavelet regression meets ANOVA. Numer Math 154:155–207. https://doi.org/10.1007/s00211-023-01358-8
11. Mann R, Malisoux L, Urhausen A et al (2016) Plantar pressure measurements and running-related injury: a systematic review of methods and possible associations. Gait Posture 47:1–9. https://doi.org/10.1016/j.gaitpost.2016.03.016
12. Masani K, Kouzaki M, Fukunaga T (2002) Variability of ground reaction forces during treadmill walking. J Appl Physiol 92(5):1885–1890. https://doi.org/10.1152/japplphysiol.00969.2000
13. Melakessou F (2017) A new sensor for gait analysis: demonstration of the IEE's smart insole. In: Proceedings of the 15th ACM conference on embedded network sensor systems. Association for Computing Machinery, New York, NY, USA, SenSys'17. https://doi.org/10.1145/3131672.3136955
14. Oriwol D, Milani TL, Maiwald C (2012) Methodological issues associated with the mean value of repeated laboratory running measurements. Footwear Sci 4(3):183–190. https://doi.org/10.1080/19424280.2012.676676
15. Park G, Lee KM, Koo S (2021) Uniqueness of gait kinematics in a cohort study. Sci Rep 11(1):57. https://doi.org/10.1038/s41598-021-94815-z
16. Potts D, Schmischke M (2022) Learning multivariate functions with low-dimensional structures using polynomial bases. J Comput Appl Math 403(113):821. https://doi.org/10.1016/j.cam.2021.113821
17. Racic V, Pavic A, Brownjohn J (2009) Experimental identification and analytical modelling of human walking forces. J Sound Vib 326(1–2):1–49. https://doi.org/10.1016/j.jsv.2009.04.020

18. Rozumalski A, Novacheck TF, Griffith CJ et al (2015) Treadmill vs. overground running gait during childhood: a qualitative and quantitative analysis. Gait Posture 41(2):613–618. https://doi.org/10.1016/j.gaitpost.2015.01.006

19. Sanseverino G, Krumm D, Kilian W et al (2023) Estimation of hike events and temporal parameters with body-attached sensors. Sports Eng 26(18):1460–2687. https://doi.org/10.1007/s12283-023-00411-x

20. Saunders JB, dec M, Inman VT, Eberhart HD (1953) The major determinants in normal and pathological gait. J Bone Jt Surg Am 35(3):543

21. Schmischke M, Lippert L (2022) NFFT/ANOVAapprox.jl: v1.1.7. https://doi.org/10.5281/zenodo.7303514

22. Shahabpoor E, Pavic A (2017) Measurement of walking ground reactions in real-life environments. Sensors 17(9):2085. https://doi.org/10.3390/s17092085

23. Sim T, Kwon H, Oh SE et al (2015) Predicting complete ground reaction forces and moments during gait with insole plantar pressure information using a wavelet neural network. J Biomech Eng 137(9):0910,010. https://doi.org/10.1115/1.4030892

24. van Hooren B, Fuller JT, Buckley JD et al (2020) Is motorized treadmill running biomechanically comparable to overground running? Sports Med 50(4):785–813. https://doi.org/10.1007/s40279-019-01237-z

25. Willems PA, Gosseye TP (2013) Does an instrumented treadmill correctly measure the ground reaction forces? Biol Open 2(12):1421–1424. https://doi.org/10.1242/bio.20136379