Residual Analysis of Ground Reaction Forces Simulation during Gait Using Neural Networks with Different Configurations*

Gustavo Leporace, *Student Member, IEEE*, Luiz Alberto Batista, Leonardo Metsavaht and Jurandir Nadal, *Member, IEEE*

Abstract— **The aim of the study was to analyze and compare the residuals obtained from ground reaction force (GRF) models developed using two different neural network configurations (one network with three outputs; and three networks with one output each), based on accelerometer data. Seventeen healthy subjects walked along a walkway, with a force plate embedded, with a three dimensional accelerometer attached to the shank. Multilayer perceptron networks (MLP) models were developed with the 3D accelerometer data as inputs to predict the GRF. The residuals of these models were evaluated graphically and numerically to verify the fitting. A visual analysis of the simulated signals suggests the model was able to adequately predict the GRF. The errors and correlations found in the MLP models for the 3D GRF is at least similar to other studies, although some of them showed higher errors. There was not difference between the two MLP configurations. However, despite the high correlation coefficient and closeness to a normal probability distribution, the residual analysis still presented a higher kurtosis and skewness, suggesting that the inclusion of other variables and the increase of the validation sample size could increase the fitting of the simulation.**

I. INTRODUCTION

Progression, maintenance of stability and shock absorption during gait are performed by the complex interaction of muscle activations that ends with forces being applied by the foot (propulsion phase) or at the foot (shock absorption phase). These forces must be dealt by muscle-skeletal system to avoid injuries and spare energy [1]. So far there is no means of directly measure these muscles forces during walking. Computational mathematical systems have been developed to estimate these forces by optimization algorithms based on ground reaction forces and kinematic data [2].

To accomplish with that it is usually used an optoelectronic motion analysis system and force plates. However, these devices limit the performance of the gait analysis to a laboratory environment, besides its high cost [3], reducing the application of quantitative biomechanical tests in the clinical practice. In the past few years, kinematic analysis has been proved feasible by the use of low cost inertial sensors, but the measurement of kinetic variables are still poorly explored.

Recent studies developed models using plantar pressure insoles [4,5] and accelerometers attached to the distal skin [6] to estimate GRF, presenting promising results. Accelerometer models have the advantage of being low cost and the gait analysis can be done barefoot, which is not possible with foot pressure insoles. Nevertheless, none of the studies presented the validation of the model, so it is not possible to affirm that results of these studies are acceptable descriptions of the data. There are many statistical tools for model validation, but the primary tool for most process modeling applications is graphical residual analysis [7].

Therefore, the purpose of the study was to analyze and compare the residuals obtained from GRF models developed using two different neural networks configurations, based on 3D accelerometer data.

II. MATERIALS AND METHODS

A. Subjects

Seventeen healthy subjects (11 males; 27.1 ± 3.4 years old; 84.3 ± 4.5 kg) participated in the study. All of them provided written consent approved along with the experimental protocol by the Institutional Ethics and Research Committee. Inclusion criteria were: (i) to have no history of ligament injuries, nor pain at the time of the tests; (ii) to be between 20 and 40 years old; and (iii) to walk independently, without the need of any orthoses or braces.

B. Data Collection

Subjects were instructed to walk at a self-selected speed along an 8 m long walkway. Each subject performed six laps with an accelerometer (± 6 g, model MMA7260Q, Freescale, USA) attached to the distal and anterior part of the shank (Figure 1). A force plate (AccuGait, AMTI, USA) was embedded into the middle of the walkway. The first two laps were not collected to allow familiarization with the task. The last four laps were collected in order to capture, during four gait cycles the acceleration of the shank. All data were collected simultaneously using a BIOPAC system (UIM, MP100 Systems, BIOPAC, USA) with a sample rate of 1 kHz.

C. Data Processing

Initially, data were filtered by a 2nd order Butterworth low pass filter, applied in the direct and reverse directions to avoid phase shift. The cutoff frequency for GRF and

** This study was partially supported by the Brazilian Research Council (CNPq), and FAPERJ and CAPES Foundations.

G. Leporace and J. Nadal are with Programa de Engenharia Biomédica, COPPE, Universidade Federal do Rio de Janeiro, P. O. Box 68.510, 21941-972 Rio de Janeiro RJ, Brasil, guleporace@yahoo.com.br, jn@peb.ufrj.br.

L. A. Batista is with Laboratório de Biomecânica e Comportamento Motor, Universidade do Estado do Rio de Janeiro, Rio de Janeiro RJ, Brasil, batista.l.a@gmail.com.

Leonardo Metsavaht is with Instituto Brasil de Tecnologias da Saúde, Rio de Janeiro RJ, Brasil, leonardo@brasilsaude.org.

978-1-4244-9270-1/15/$31.00 ©2015 IEEE

accelerometer data was 25 Hz. Stance phase was defined when vertical ground reaction forces were above 10 N.

To fit anteroposterior (AP), vertical (Vert) and mediolateral (ML) GRF curves, fourteen inputs were selected based on previous studies [6, 8, 9], as follows:

i to iii: the 3D accelerometer data normalized by body weight;

iv to vi: the 3D relative velocity of the subject's center of mass, represented by the single integration of the acceleration curve;

vii to ix: 3D relative displacement of the subject's center of mass, represented by the double integration of the acceleration curve;

x to xii: the first derivate of the 3D accelerometer signal

xiii: the stance duration (s); and

xiv: time point (%), expressed as a percentage of the stance time.

Multilayer perceptron (MLP) neural networks with one hidden layer were selected to simulate the ground reaction forces. The number of neurons in the hidden layer was chosen by testing the fitting of signals with different sizes, always seeking the most parsimonious model, i.e. few hidden neurons and good generalization power. Networks were modelled using from five to 12 neurons in the hidden layer. Two different network configurations were used to simulate 3D GRF: one network with three outputs; and three networks with one output each.

To assess the network fitting a k-fold cross-validation strategy was used. The subjects were divided into four groups; three containing the four gait cycle from four subjects, and one containing the four gait cycle from five subjects. The MLP was trained using always three groups. The signal of the group left out of the training was simulated using the developed network to test the fitting. For each repetition, the network was trained using the Levenberg-Marquardt backpropagation algorithm according to Favre et al. [8] and Liu et al. [9].

In the hidden and the output layers the used activation transfer function was the hyperbolic tangent and linear, respectively. This typology was used since previous biomechanical studies showed that these transfer functions produced the best signal prediction [6, 8, 9]. To avoid overfitting, the generalization error obtained for the validation set during the training process and the minimum gradient were used as stop criteria.

The network with lower mean absolute deviation (MAD) among all subjects was selected as the best number of neurons in the hidden layers. MAD was calculated as:

$$\mathbf{MAD} = \frac{1}{\mathrm{N}}\sum_{\mathrm{t}=1}^{\mathrm{N}}\left|\widehat{\mathrm{GRF}(\mathrm{t})} - \mathrm{GRF}(\mathrm{t})\right| \tag{1}$$

where $\widehat{\mathrm{GRF}(\mathrm{t})}$ represents the simulated ground reaction force, $\mathrm{GRF}(\mathrm{t})$ corresponds to the collected GRD and N is the vector size.

The normalized mean absolute deviation (MAD%) was also calculated as [8]:

$$\mathbf{MAD\%} = \frac{\mathrm{MAD}}{\mathrm{range\ (GRF(t))}} * 100 \tag{2}$$

D. Statistical Analysis

A data mining was initially performed to verify if the distribution of the error had a Gaussian shape. To accomplish with that, a residual plot was analyzed. A residual plot is a graph that shows the residuals on the vertical axis and the simulated signals on the horizontal axis. If the points in this graph are randomly dispersed around the abscissa, a linear regression model is appropriate for the data; otherwise, a non-linear model is more appropriate. The standardized residuals were calculated to verify if 95% of the values would fall between ± 2 standard deviations in the residual plots.

The histograms of the residuals were also plotted to check whether the variance is normally distributed. A symmetric bell-shaped histogram which is evenly distributed around zero indicates that the normality assumption is likely to be true. Gaussian quartile plots were also calculated to verify the assumption of normality of the residuals.

To compare the MAD of each direction (AP, Vert and ML) between the simulated and collected signals a paired t-test was used. The correlation between the simulated and the collected signal was also calculated for each GRS component, using the Pearson's correlation coefficient. The significance level was set at 0.05.

III. RESULTS

For the single network with three outputs, the number of neurons in the hidden layer with lowest MAD and highest correlation coefficient was eight (Fig. 1, Table 1). The second configuration showed that lowest error was using eight, five and ten neurons for AP, Vert and ML, respectively. There were no statistical differences between the MAD of the two MLP network configurations (p value: 0.243; 0.322; 0.823 for AP, Vert and ML, respectively).

All residual plots showed that there are no trends in the residuals, presenting a stochastic pattern, although it can be seen some outliers (Fig. 2). More than 94% of the standardized residuals were within ± 2 standard deviations (Table 2), therefore it allows concluding that the model had, in general, adequate fit.

TABLE I. PEARSON'S CORRELATION COEFFICIENT (**r**) AND THE MEAN ABSOLUTE ERROR (**MAD**) BETWEEN THE COLLECTED SIGNAL AND THE SIGNALS SIMULATED WITH BOTH A SINGLE MLP (***1 MLP***) AND THREE DIFFERENT NETWORKS (***3MLP***)

GRF	**r**		**MAD (%BW)**		**MAD%**	
	3 MLP	***1 MLP***	***3 MLP***	***1 MLP***	***3 MLP***	***1 MLP***
AP	0.971	0.969	1,83 ± 0,68	1,9 ± 0.6	5.2 ± 2.0	5.4 ± 1.8
Vert	0.969	0.968	5,1 ± 1,64	5,2 ± 1.7	4.7 ± 1.4	4.8 ± 1.5
ML	0.792	0.801	1,35 ± 0,46	1,4 ± 0.5	12.8 ± 5.6	13.0 ± 6.1

GRF: Ground Reaction Force; %BW: % body weight AP: Anteroposterior; Vert: Vertical; ML: Mediolateral

TABLE II. PERCENTAGE OF STANDARDIZED RESIDUAL VALUES THAT FELL WITHIN ± 2 STANDARD DEVIATIONS.

GRF Component	Std Res	
	3 MLP	*1 MLP*
Anteroposterior	94.2 %	94.3 %
Vertical	93.6 %	94.4 %
Mediolateral	95.1 %	94.6 %

Std Res: Standardized Residuals; GRF: Ground Reaction Force

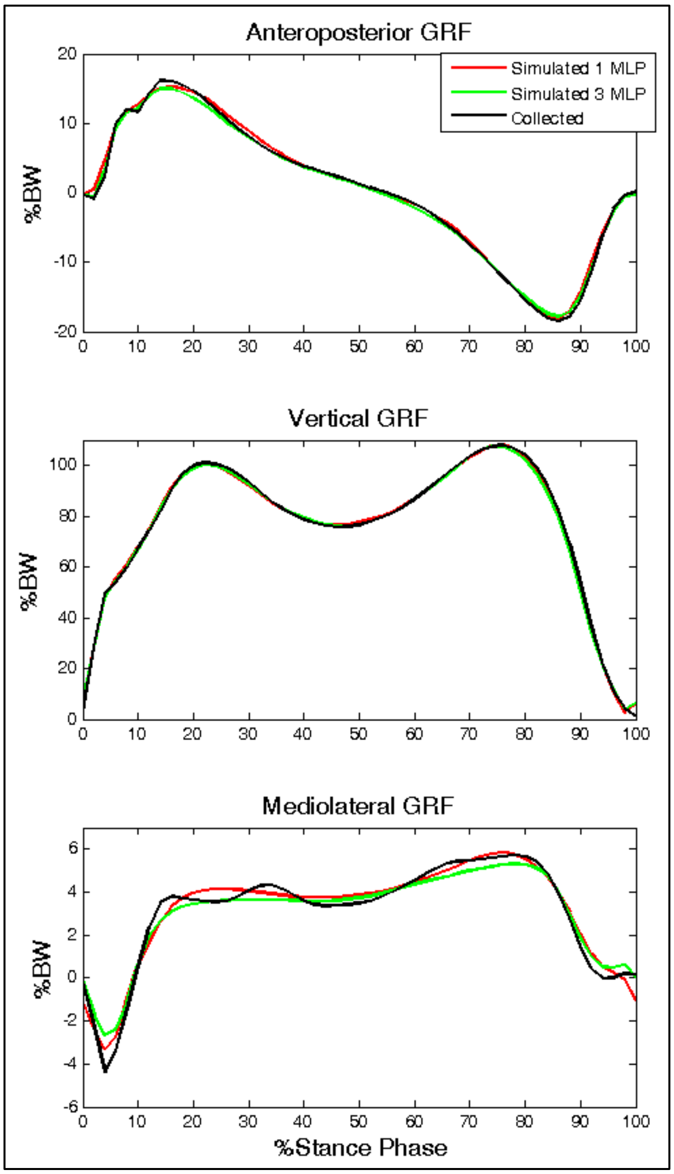

Figure 1. Collected and simulated ground reaction forces (GRF) in both configurations: one network with three outputs (Simulated 1 MLP), three network with one output each (Simulated 3 MLP). The GRF graphs represent the average of the 17 subjects. %BW: Percentage of the Body Weight.

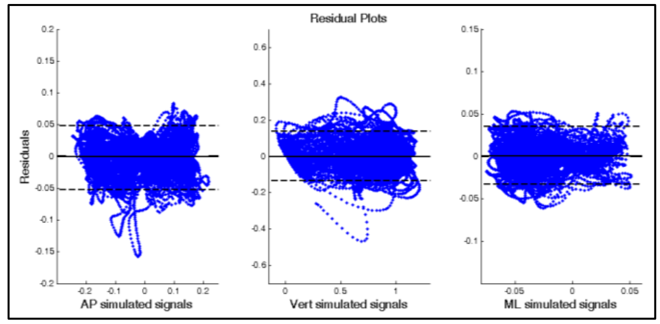

Figure 2. Residual plots of the ground reaction forces simulation. Horizontal continous lines represent zero in the ordinates values and dashed lines represents two standard deviaitions of the ordinates values. AP: Anteroposterior component; Vert: Vertical component; ML: Mediolateral component. Residuals: Difference between the collected and simulated signals.

The histograms and Gaussian quartile plots demonstrated that the residuals resemble a Gaussian distribution, however still presented a left-tail and sharper peak (Fig. 3 and 4). There was no difference in the shape of the plots between the two MLP configurations used in the study.

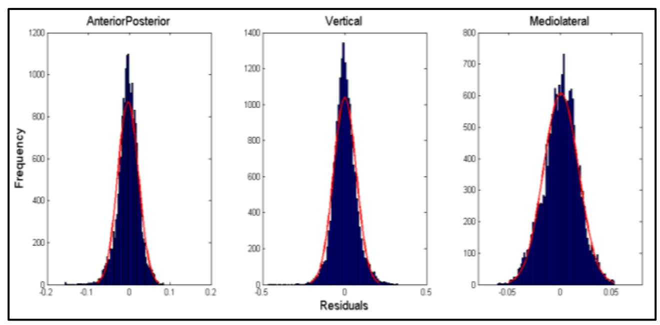

Figure 3. Histograms of the residuals of the simulated ground reaction forces. The red lines represent the expected probability density function of a normal distribution.

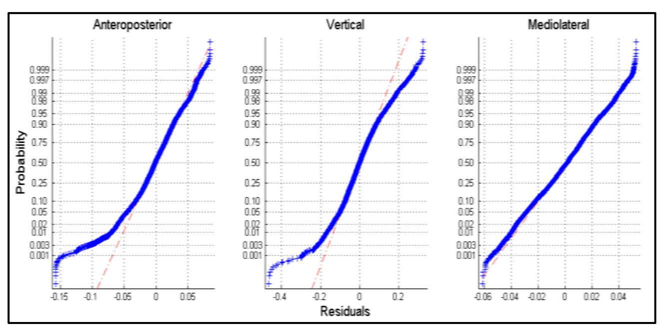

Figure 4. Gaussian quartile plots of the simulated ground reaction forces. The red dashed lines represent the expected behavior for a normal distribuition.

IV. DISCUSSION

The aim of this paper was to develop different neural network model configurations to simulate the ground reaction forces obtained from accelerometer data and analyze the residuals to describe the fitting of the model. Several studies have developed models to simulate biomechanical time series [4-6, 8-9]. Most of them simulate ground reaction forces using as input variables kinematic and/or foot pressure data collected from optoelectronic motion analysis systems and foot pressure insoles [4-6]. In spite of the advance in terms of the absence of force plates, these models still require a high biomechanical instrumentation cost or limit the use to a laboratory environment. Differently, the methodological design presented in this paper only require a low cost accelerometer and can be performed easily in the clinical setting, allowing the widespread of its use.

A visual analysis of the simulated signals (Fig. 1) suggests the model was able to adequately predict the GRF. The errors and correlations found in the MLP model for the 3D GRF is at least similar to other studies, although some of them showed higher errors [4-6]. ML GRF presented the lowest error, followed by AP and Vert forces. This is in agreement with the studies that used pressure insoles and kinematic data to estimate GRF. It seems that the higher the range of the force component, the greater is the absolute error. However, the relative errors and correlations coefficient showed lower influence of these on AP and Vert. There was not statistical difference between the errors of both network configurations.

Most of the studies on the same subject concluded that their model has a high accuracy and could be used to estimate GRF in the clinical setting [4-6]. However, none of them assessed the validity of their model. Validation is one of the most important steps in the model building sequence, but it is also one of the most overlooked. The assumption of validity of previous studies consisted of nothing more that quoting the correlation and errors between the simulated and collected time series.

It is not possible to assume that the fitting was adequate without exploring the residuals of the model. If the residuals have not a stochastic characteristic, it could be assumed the input variables are missing some of the predictive information. In a well fitted model, the residuals must be independent, have a constant variability, have a Gaussian probability density function and have no outliers [7]. In the present study, the residual plots (Fig. 2) showed that the residuals are constantly distributed over the entire range of the simulated forces, confirming the assumptions of homocesdacity, despite a higher amount of outliers than expected (Table 2). There was not difference between both network configurations.

The histograms (Fig. 3) and Gaussian quartile plots (Fig. 4) suggested the residuals resemble a normal probability distribution, although it is possible to verify that the AP and Vert present a subtle left tail, indicating a negative skew, and also, an increased kurtosis (sharper peak). These results, altogether with the residual plots, indicate that the input variables included in the AP and Vert modelling may be leaking some information to the residuals. The inclusion of new parameters, like the shank angles obtained from inertial sensors, could contribute to the solution of this small deviation from the Gaussian distribution of the residuals. Another proposal is to apply statistical techniques to select the most relevant input variables, like the mutual information among all the inputs and outputs [7].

The ML GRF simulation had the probability distribution closest to a Gaussian one. The moderate correlation coefficient (Table 1) found and the visual difference of the simulation and collected signal (Fig. 1) could be related to the high inter- and intra-subject variability in this time series [10]. As the training sample was small, it may have not provided enough information to train the network to deal with this variability. It is proposed the development of new networks with higher validation sample sizes to increase the fitting of the model proposed in this study.

V. Conclusion

The models developed in the present study had an error at least similar to other studies, although some of them showed higher errors. There was no difference between the two MLP configurations. However, despite the high correlation coefficient and closeness to a normal probability distribution, the residual analysis still presented a higher kurtosis and skewness, suggesting that the inclusion of other variables and the increase of the validation sample size could increase the fitting of the simulation.

Acknowledgment

The authors would like to thank Felipe Alvim, Aluizio Netto, and Igor Jesus for the help in the data collection.

References

[1] J. Perry, J. M. Burnfield, *Gait Analysis: Normal and Pathological Function*, Thorofare, NJ: Slack, 2nd ed., 2010.

[2] S. L. Delp, F. C. Anderson, A. S. Arnold, P. Loan, A. Habib, C. T. John, E. Guendelman, D. G. Thelen.,"OpenSim: open-source software to create and analyse dynamic simulations of movement" , *IEEE Trans. Biomed. Eng.*, vol. 54, pp. 1940-1950, November 2007.

[3] R. Baker, "Gait analysis methods in rehabilitation", *J. Neuroeng. Rehabil.*, vol. 2, pp. 3-4, 2006.

[4] H. Rouhani, J. Favre, X. Crevoisier, K. Aminian, "Ambulatory assessment of 3D ground reaction force using plantar pressure distribution", *Gait Posture*, vol. 32, pp. 311-316, July 2010.

[5] Y. Jung, M. Jung, K. Lee, S. Koo, "Ground reaction force estimation using an insole-type pressure mat and joint kinematics during walking", *J. Biomech.*, vol. 22, pp. 2693-2699, August 2014.

[6] G. Leporace, L. A. Batista, F. Alvim, I. Jesus, F. Teixeira, A. Netto, F. F. Nobre, J. Nadal, L. L. Menegaldo, "Simulation of ground reaction forces during gait based on accelerometer data using a neural network model", in *Proc. XIV International Symposium on Computer Simulation in Biomechanics*, Natal, 2013, pp. 1-2.

[7] H. Motulsky, *Fitting Models to Biological Data Using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*, New York, NY: Oxford University Press, 1st ed., 2014.

[8] J. Favre, M. Hayoz, J. C. Erhart-Hledik, T. P. Andriacchi, "A neural network model to predict knee adduction moment during walking based on ground reaction force and anthropometric measurements", *J. Biomech.*, vol. 23, pp. 692-698, February 2012.

[9] Y. Liu, S. M. Shih, S. L. Tian, Y. J. Zhong, L. Li, "Lower extremity joint torque predicted by using artificial neural network during vertical jump", *J. Biomech.*, vol. 41, pp. 906-911, May 2009.

[10] C. T. John, A. Seth, M. H. Schwartz, S. L. Delp, "Contributions of muscles to mediolateral ground reaction force over a range of walking speeds.", *J. Biomech.*, vol. 21, pp. 2438-2443, September 2012.