The authors are with the Instituto de Biomecanica de Valencia, Universidad Politecnica, Valencia, Spain.

Comparison of Functional Regression and Nonfunctional Regression Approaches to the Study of the Walking Velocity Effect in Force Platform Measures

Juan V. Durá, Juan M. Belda, Rakel Poveda, Álvaro Page, José Laparra, José Das, Jaime Prat, and Ana C. García

The effect of walking velocity on force platform measures is examined by means of functional regression and nonfunctional regression analyses. The two techniques are compared using a data set of ground reaction forces. Functional data analysis avoids the need to identify significant points, and provides more information along the waveform.

Keywords: functional analysis, gait analysis, patterns

Force platform measures are mainly presented as temporal waveforms. One of the most commonly used methods for analyzing these waveforms involves the extraction of specific parameters: the waveform value at significant points, such as the maximum ground forces and the peak joint kinematic and kinetic parameters (Goble et al., 2003; Lelas et al., 2003; White et al., 1999).

One of the clinical applications of movement analysis is to identify normal patterns, such as ground reaction force patterns. These patterns are used to establish deviations or differences in patients (Simon, 2004). By doing so, different patterns are used to detect problems or injuries in patients, or to confirm the benefits of rehabilitation treatments. The normal pattern is the mean value measured using healthy subjects. The identification of problems or injury detection is based on significant differences between patient values and normal patterns.

To obtain normal patterns it is necessary to reduce the data variability not affected by pathological issues. The normal pattern is for similar subjects: children, adults, and so on. The variability produced by the subject's weight is reduced using body weight units. Time is normalized to reduce the effect of the duration of the movement. Velocity also increases variability. For this reason it is necessary to compare data measured at similar velocities (Keller et al., 1996; Lelas et al., 2003; Stansfield et al., 2006). But patients walk at different velocities than healthy counterparts, usually lower velocities (Lelas et al., 2003).

One of the most commonly presented solutions to the velocity problem that frequently appears in the literature is regression analysis. The aim of the regression analysis is to predict normal pattern values for the walking velocity of patients using measurements taken from healthy subjects over an appropriate velocity range (Hanlon & Anderson, 2006; Lelas et al., 2003). This type of analysis does not refer to all the information included in waveforms. Furthermore, it is difficult to extract parameters from pathological waveforms (Chau, 2001; White et al., 1999) in that the waveforms of patients may differ greatly from those of healthy people, with the possibility that it may prove very difficult to identify the same parameters obtained from normal waveforms.

Instead of extracting discrete parameters from waveforms, it is possible to analyze the complete waveform by means of functional data analysis (FDA). The FDA technique applies methods of multivariate statistics, such as ANOVA or regression analysis, but works with the whole function as a piece of data.

The aim of this paper is to compare the results obtained using nonfunctional regression analysis (regression with discrete parameters) with those obtained with functional regression analysis, and hence demonstrate the advantages of functional analysis.

Methods

To illustrate the application of the functional linear regression, we have used a data set published by Lafuente et al. (2000). The data set analyzed in our study is the vertical ground reaction force of 27 healthy male subjects walking at three different but self-selected speeds: a normal freely selected speed and two imposed cadences whose rhythms were suggested to the subject by asking him to walk slightly faster and slightly slower than usual. The subject wore his normal street shoes for the test.

The velocity obtained ranged from 0.9 to 1.9 $m \cdot s^{-1}$. The vertical force was normalized with the subject's body weight, and the ground reaction force was expressed as a percentage of that body weight.

To compare the functional regression and nonfunctional regression analyses, we selected three parameters: Fz1max (first maximum vertical force), Fz2max (second maximum vertical force), and Fzmin (minimum vertical force between Fz1max and Fz2max).

Functional and standard regression analyses are carried out following the same smoothing process. The smoothing process is conducted using a least-squares fitting technique, which selects the best coefficient of the 60 B-spline functions, as described in Page et al. (2006).

The functional data analysis performed comprises two steps: (A) linear time normalization and (B) functional regression analysis.

Linear time normalization consisted of assigning time values from 0 to 100, where 0 is the initial contact (heel strike) on the force platform, and 100 is the end of the forefoot contact. These values were selected using a threshold of 5% of the maximum vertical force. Linear time normalization is appropriate for reducing variability when the correlation between phase and timing variables is high (Page & Epifanio, 2007). In this case, the correlation was >0.9.

The functional regression analysis for the vertical ground reaction force (Z) is based on the following model:

$$Z_i = B0(t) + B1(t) \times V_i + \varepsilon_i(t) = \hat{Z}_i(t) + \varepsilon_i(t) \qquad (1)$$

where V_i is the velocity and $\varepsilon_i(t)$ the residual function or error. In this case, t is not the real time, but the normalized time (0 to 100). the terms $B0$ and $B1$ are constants of the function calculated in the regression, whereas in FDA $B0$ and $B1$ are time dependent.

The model fit is assessed using the squared correlation function $RSQ(t)$ and the F-ratio function $F(t)$. $RSQ(t)$ is equivalent to the R^2 parameter of normal linear regression. The $RSQ(t)$ values range from 0 to 1, as in R^2. $F(t)$ and $RSQ(t)$ assess the fit of the model[AUQ1].

$$RSQ(t) = 1 - \frac{\sum_{i=1}^{N}\left(Z_i(t) - \bar{Z}(t)\right)^2}{\sum_{i=1}^{N}\left(Z_i(t) - \hat{Z}(t)\right)^2} \qquad (1)$$

where $\bar{Z}(t)$ is the mean function.

The 5% significance level for the F distribution is calculated using the following degrees of freedom (df) (Ramsay 1997):

$$df(\text{error}) = \text{N} - \text{b}$$

$$df(\text{regression}) = \text{N} - df(\text{error}) - 1$$

where N is the number of functions and b the number of independent functions included in the model (2 in this lineal model). For a detailed explanation of the functional linear model, see Ramsay's book (Ramsay, 1997).

The MATLAB functions used for this purpose are included in the package *fda* developed by J.O. Ramsay. The package *fda* can be freely downloaded from http://ego.psych.mcgill.ca/misc/fda/.

Results

The standard regression analysis shows a significant influence in velocity (V) for the three parameters ($p < .05$). The equations, F ratio, and R^2 obtained are

Fz1max = 0.6632 + 0.3315 × V; F = 140.45, R^2 = .406

Fz2max = 0.8292 + 0.1857 × V; F = 67.90, R^2 = .263

Fzmin = 1.1060 – 0.3065 × V; F= 261.72, R^2 = .567

Keller (1996[AUQ2]) obtained slightly higher maximum R^2 values (0.65) for male adults; however, here the velocity range is different and bigger: 1.5 to 3.5 $m \cdot s^{-1}$. Stansfield et al. (2006) finds similar or lower R^2 values for children.

By way of example, the regression of Fzmin is shown in Figure 1. The components of the functional model can be displayed graphically (Figure 2, left). The continuous line is the mean waveform. The line labeled "Velocity" is $B1(t)$, that part of the functional model dependent on velocity. The line labeled "Constant" is $B0(t)$, the part that is equal for all trials.

The two components $B0(t)$ and $B1(t)$ could be interpreted as the "constant" and "dynamic" components of force. The constant component is similar to the curve obtained when the subject slowly positions himself over the force platform as if being weighed. The dynamic component resembles a spring damping a falling mass. Increased velocity means more energy damping, as if the mass falls from the highest point, and hence increased force oscillation.

\ insert Figures 1, 2, and 3 \

The functional model enables a waveform to be predicted for each velocity. Peak and valley forces become more pronounced the greater the velocity (Figure 2, right). The first peak force to appear is in response to faster velocities, but the second peak is only slightly delayed with faster velocities.

Significant levels (F ratio) change throughout the force curves (Figure 3, right). The mean waveform is represented by a dashed line. The regression is always statistically significant, except at the beginning, the end, and at the transitions from peaks to valleys.

The R^2 or RSQ values have the same behavior as F-ratio (Figure 3, left). Although peak values appear near the moments of Fz1max, Fz2max, or Fzmin, it is misleading to talk about R^2 values for single points, such as Fz1max, Fz2max, or Fzmin. It would be more fitting to say that we have three local maximum values in three different zones: R^2 = .430 and F = 150.8 for the Fz1max zone, R^2 = .345 and F = 105.3 for the Fz2max zone, and R^2 = .577 and F = 273.0 for the Fzmin zone. These maximum R^2 and F values are higher than the values obtained using standard linear regression.

In addition, the statistical analysis of the functional model (Figure 3) provides more information than the single values obtained with the standard linear regression. We can observe how both the statistical significance level and the R^2 values evolve along the force curve. Lower values are obtained at the beginning and end of the curve, and also during the transitions from peaks to valleys.

Discussion

The functional model not only provides estimates of discrete values, but also facilitates the interpretation of the change in the shape of the curve (Figure 2, right). Thus, the increase of the speed produces an increase of peak forces and a decrease in the minimum of force. These results are consistent with those obtained in the standard linear regression. The contribution of the functional approach is that it also provides information on changes in the temporal pattern. The increase of speed produces a slight advance of the first peak force and a slight delay in the onset of the second. This represents a change in strategy of the movement that was not appreciated with a standard linear regression.

This study highlights the advantages of functional regression as a useful technique for studying the effect of walking velocity in force platform measures. As stated in the introductory paragrphs, functional data analysis considers all the information embedded in the waveforms and avoids the need to identify single points (discrete parameters). With regard to discrete value analysis, this is extremely beneficial due to the difficulty involved in extracting parameters from pathological waveforms (Chau 2001, Page & Epifanio, 2007).

We have demonstrated that functional linear regression provides not only more, but also better information than standard linear regression using parameters. Changes in waveforms can be visualized numerically and graphically (Figure 2). The components of the functional regression can be conducive to a clearer mechanical interpretation (constant and dynamic or spring components) than offered by nonfunctional regression. Furthermore, the statistical parameters and significance levels are slightly improved and their evolution can be observed along the waveform. Functional regression models also allow working with multiple variables (e.g., gender, age, height) to reduce variability and improve the curve fitting.

Acknowledgments

This work was funded by the Spanish Government and cofinanced by EU FEDER funds (grants DPI2006-14722-C02-01, DPI2006-14722-C02-02, and DPI2008-06772-C03-03).

References

Chau, T. (2001). A review of analytical techniques for gait data. part 1: Fuzzy, statistical and fractal methods. *Gait & Posture, 13*(1), 49–66.

Goble, D.J., Marino, G.W., & Potvin, J.R. (2003). The influence of horizontal velocity on interlimb symmetry in normal walking. *Human Movement Science, 22*(3), 271–283.

Hanlon, M., & Anderson, R. (2006). Prediction methods to account for the effect of gait speed on lower limb angular kinematics. *Gait & Posture, 24*(3), 280–287.

Keller, T., Weisberger, A., Ray, J., Hasan, S., Shiavi, R., & Spengler, D. (1996). Relationship between vertical ground reaction force and speed during walking, slow jogging, and running. *Clinical Biomechanics (Bristol, Avon), 11*(5), 253–259.

Lafuente, R., Belda, J.M., Sánchez-Lacuesta, J., Soler, C., Poveda, R., & Prat, J. (2000). Quantitative assessment of gait deviation: Contribution to the objective measurement of disability. *Gait & Posture, 11*(3), 191–198.

Lelas, J.L., Merriman, G.J., Riley, P.O., & Kerrigan, D.C. (2003). Predicting peak kinematic and kinetic parameters from gait speed. *Gait & Posture, 17*(2), 106–112.

Page, A., Ayala, G., León, M.T., Peydro, M.F., & Prat, J. (2006). Normalizing temporal patterns to analyze sit-to-stand movements by using registration of functional data. *Journal of Biomechanics, 39*(13), 2526–2534.

Page, A., & Epifanio, I. (2007). A simple model to analyze the effectiveness of linear time normalization to reduce variability in human movement analysis. *Gait & Posture, 25*(1), 153–156.

Ramsay, J.O. (1997). *Functional data analysis*. New York: Springer.

Simon, S. (2004). Quantification of human motion: Gait analysis: Benefits and limitations to its application to clinical problems. *Journal of Biomechanics, 37*(12), 1869–1880.

Stansfield, B.W., Hillman, S.J., Hazlewood, M.E., & Robb, J.E. (2006). Regression analysis of gait parameters with speed in normal children walking at self-selected speeds. *Gait & Posture, 23*(3), 288–294.

White, R., Agourisa, I., Selbie, R.D., & Kirkpatrickb, M. (1999). The variability of force platform data in normal and cerebral palsy gait. *Clinical Biomechanics (Bristol, Avon), 14*(3), 185–192.

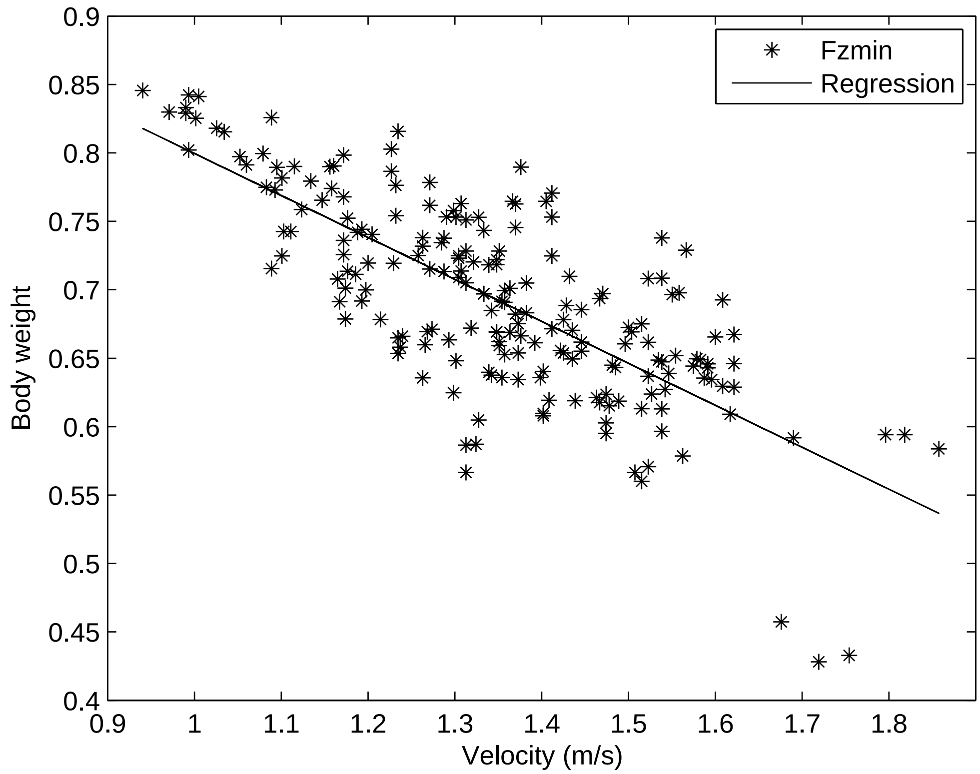

Figure 1 — Regression of Fzmin.

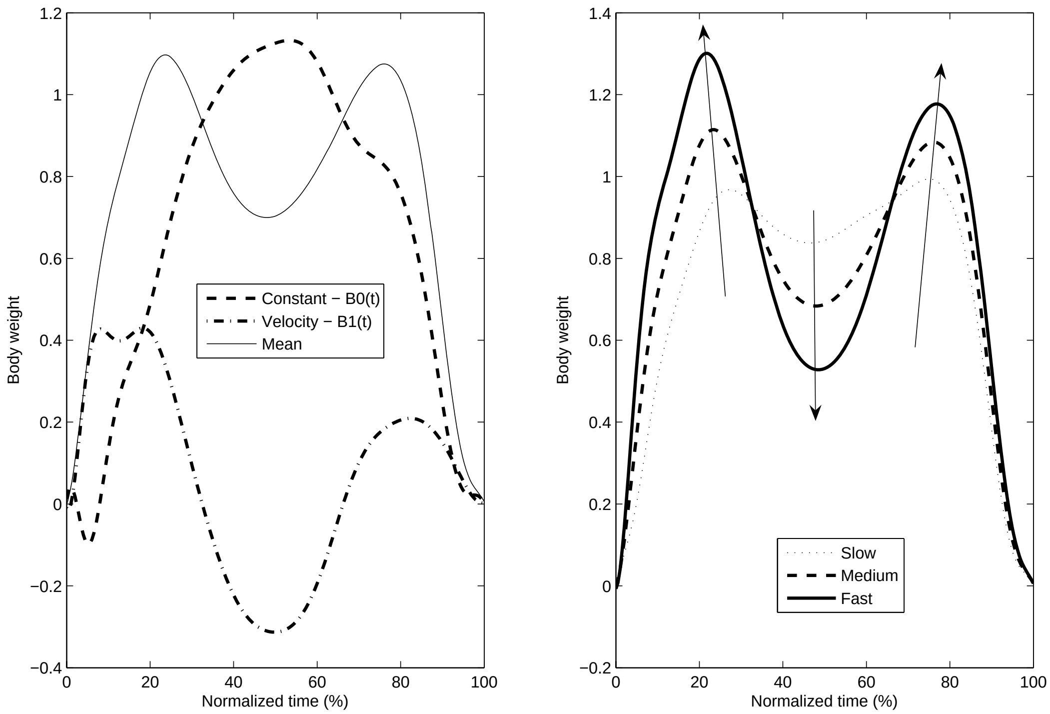

Figure 2 — Linear model (left). Velocity evolution (right).

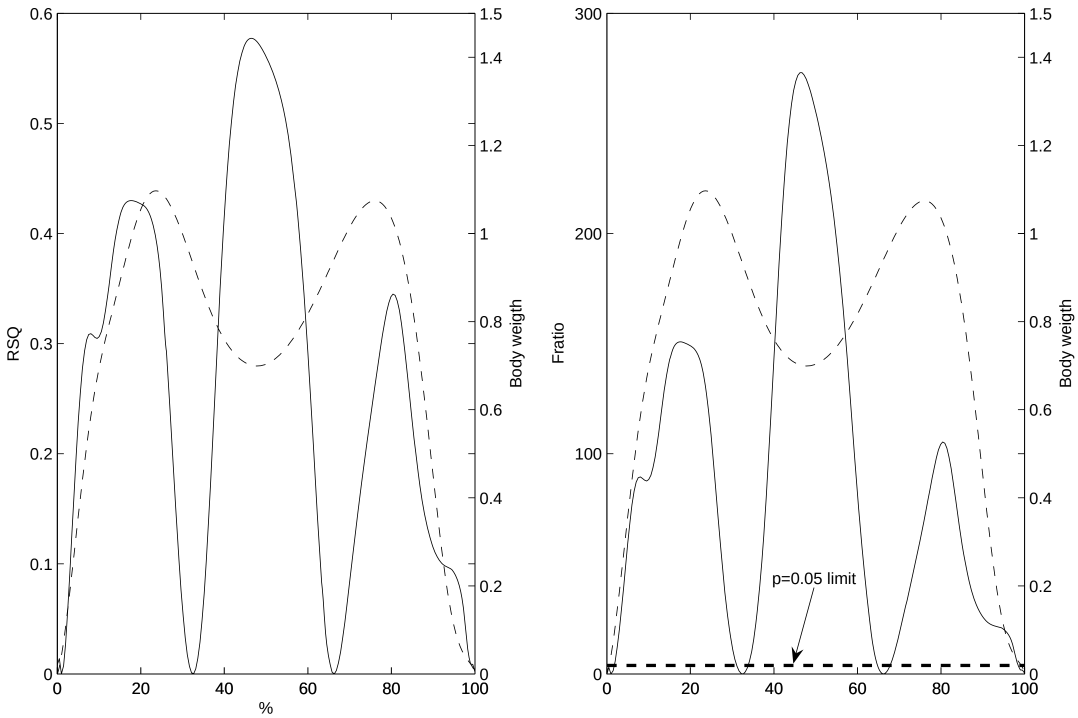

Figure 3 — *RSQ* evolution (left). *F* ratio (right). Dashed lines correspond with mean vertical force (*y*-axis left).

Author Queries

[AUQ1] Should the following be Equation No. 2?

[AUQ2] The in-text citation "Keller 1996" is not in the reference list. Should this be Keller et al., 1996?