

CROSS-COMPARISON OF THE PERFORMANCE OF DISCRETE, PHASE AND FUNCTIONAL DATA ANALYSIS TO DESCRIBE A DEPENDENT VARIABLE

Chris Richter¹, Leonardo Gualano¹, Noel E. O'Connor¹ and Kieran Moran²

¹CLARITY: Centre for Sensor Web Technologies, Dublin, Ireland

²School of Health and Human Performance, Dublin City University, Dublin, Ireland

The aim of this study was to assess and contrast the ability of discrete point, functional principal component analysis (fPCA) and analysis of characterizing phases (ACP) to describe a dependent variable (jump height) from vertical ground reaction force curves captured during the propulsion phase of a countermovement jump. A stepwise multiple regression analysis was used to assess the ability of each data analysis technique. The order of effectiveness (high to low) was ACP, fPCA and discrete point analysis. Discrete point analysis was not able to generate strong predictors and detected also erroneous variables. FPCA and ACP detected similar factors to describe jump height. However, ACP performed better than fPCA because it considers the time and magnitude domain separately and in combination and it examines key-phases, without the influence of non-key-phases.

KEY WORDS: discrete point analysis, functional principal component analysis, analysis of characterizing phases, dependent variable

INTRODUCTION: While discrete point analysis has significant limitations, the majority of studies in biomechanics almost exclusively use discrete point analysis. Limitations in discrete point analysis are the need to pre-select discrete measures (e.g. maximums, minimums, overall duration) and the loss of possibly extremely important information (Donoghue et al., 2008, Doná et al., 2009, Richter et al., 2012). In recent years, researchers have sought new ways to analyze data as a continuous signal rather than discrete measures; introducing techniques originating from a variety of fields - e.g. computer science, psychology, cognitive science, physics and engineering (Chau 2001a; Chau 2001b; Kelso 1995). One of the most promising techniques is functional principal component analysis (fPCA), which reduces the dimensionality of a set of signals while preserving the information needed to describe a data set (Jolliffe 2002). FPCA is part of the functional data analysis family, and treats a signal as a single function rather than as a series of individual data points (Ramsay 2006). However, fPCA: (a) can mask important features of the data because functions are analyzed as a whole rather than as key-phases, with the non-key-phases of the functions inappropriately altering results, (b) cannot identify the specific phase of features (subjective visual inspection is needed), and (c) cannot examine the combined magnitude-time domain, which can hold important information. To overcome these limitations, Analysis of Characterizing Phases (ACP) has been recently proposed (Richter et al. 2012). The idea behind ACP is to detect key-phases (phases of variance) within a sample of curves, which are used to examine differences between populations or a relationship to a dependent variable (i.e. performance determining factor) in the time, magnitude and magnitude-time domain. To date, no study has directly measured and compared the performance of discrete point analysis, fPCA and ACP to evaluate their ability to identify those variables explaining the movement outcome. In order to compare the performance of the abovementioned analysis techniques it is necessary to examine a movement signal that fully determines the movement outcome. This criterion is met in countermovement jumps by examining jump height in relation to the vertical ground reaction force (force) trace of the propulsion phase (impulse momentum relationship). The aim of this study was to measure and contrast the performance of discrete point analysis, fPCA and ACP to describe jump height (dependent variable) using the propulsion phase of the force-time history of a countermovement jump.

METHODS: A stepwise multiple regression analysis was used to assess the ability of discrete point analysis, fPCA and ACP to explain a dependent variable.

Dependent Variable: The jump height of a countermovement jump (CMJ) was chosen as the dependent variable. Force curves of 60 athletes captured during CMJs were used for data analysis. All athletes were free from any injury and experienced in performing a CMJ. The University Ethics Committee approved the study and all subjects were informed of any risk and signed an informed consent form before participation. Prior to data collection, every subject completed a standard warm-up routine. The subjects performed 15 maximum effort CMJs without an arm swing, standing with each foot on a force platform, and rested for 30 seconds between the trials. Two force plates (BP-600900, AMTI, MA, USA), each with a frequency of 250Hz, recorded the produced force. Jump height was calculated by the impulse momentum relationship. Based on jump height, the best jump performance of each subject was chosen for data analysis. All curves were normalized to body mass and only the propulsion phases were used for analysis.

Discrete Point Analysis: Based on previous literature (Cormie et al., 2009; Dowling et al., 1993; Sheppard et al., 2008; Petushek et al., 2010, Newton et al. 1999) the following discrete features were identified and used for statistical analysis: a) initial force, b) maximum force, c) initial-to-maximum rate of force development, d) time from initial-to-maximum force, e) position maximum force (in %), f) time from maximum force to take-off, and g) propulsion phase duration. The initial-to-maximum rate of force development was calculated (Equation 1) from the initial force to the point i at which the maximum force occurred (Cormie et al., 2009). The mean force was not included in the selected variables. Mean force was discarded because it reflects the impulse momentum relationship in jumps and does not easily relate to the underlying neuromuscular control nor can it indicate specific training interventions to enhance performance.

$$RoFD(i) = (force(i) - force(1)) / \Delta time \quad Eq. (1)$$

Functional Principal Component Analysis: FPCA was performed to generate principal component scores, which were used for statistical analysis (Ramsay 2006). The used principal component scores retained 99% of the variance of the data and were VARIMAX rotated to optimize their interpretability (Ramsay 2006).

Analysis of Characterising Phases: ACP was applied to detect key-phases within the captured force curves. Key-phases were used to calculate participant scores capturing the time, magnitude and magnitude-time domain (Richter et al., 2012). VARIMAX rotated functional principal components (retaining 99% of the data's variance) were used to identify key-phases (Richter et al., 2012). Participant scores were used for statistical analysis and generated by calculating the area between a participant's curve (p) and the mean curve across the data set (q) for every point (i) within a key-phase (Equation 2 and 3).

$$score = \int p_i - q_i \quad Eq. (2)$$

$$score = \int 0.5 * (\Delta_{time} p_{i,i+1} + \Delta_{time} q_{i,i+1}) * \Delta_{magnitude} p_i q_i \quad Eq. (3)$$

The separated force features (discrete points, principal component scores and participant scores) were input into a stepwise multiple regression analysis to assess the ability of discrete point analysis, fPCA and ACP to explain jump height. Data processing was performed in MATLAB and statistical analyses were performed using SPSS 20.

RESULTS: The discrete point analysis was able to explain 26.3% of the variance in jump height (Table 1). Predictor variables were peak force and phase duration. FPCA explained 82.8% of the variance in jump height using principal component scores generated from four principal components as predictor variables (Table 1). ACP explained 99.3% of the variance in jump height using five key-phases as predictor variables utilizing information from the time domain, magnitude domain and the magnitude-time domain (Table 1).

Table 1: Generated regression models using features generated by discrete point analysis, fPCA and ACP

Discrete point analysis		Functional principal component analysis			Analysis of Characterizing Phases		
feature	standardized coefficient	feature	estimated phase	standardized coefficient	feature	domain phase	standardized coefficient
peak force	0.73	PCscore5	80-88%	-0.74	phase 5	comb. 82-87%	0.52
phase duration	0.50	PCscore1	1-13%	0.33	phase 2	comb. 28-42%	1.28
		PCscore2	24-46%	-0.39	phase 3	time 91-94%	1.95
		PCscore3	90-95%	-0.32	phase 4	comb. 57-69%	0.71
					phase 1	comb. 1-9%	0.75
					phase 3	comb. 91-94%	0.38
					phase 3	mag. 91-94%	-0.22
$r^2 = 0.263$		$r^2 = 0.828$			$r^2 = 0.993$		

DISCUSSION: Direct comparison between discrete point analysis, fPCA and ACP indicates that discrete point analysis has the lowest ability to describe jump height. Discrete point analysis generated the lowest r^2 -values because it was not only incapable to generate strong predictors from the force curves, also it generated erroneous variables. For example, the predictor variable maximum force used in the regression model is erroneous because it does not describe a specific neuromuscular capacity. The force curves examined were non-, uni- or bi-modal, with maximum force potentially occurring at either peak¹. Consequently, maximum force does not represent a consistent measure across subjects and has to be discarded. In contrast to predictor variables generated by discrete point analysis, predictor variables in fPCA and ACP are strong predictors and do relate to the underlying neuromuscular control. Therefore, findings of fPCA and ACP may help to indicate effective training interventions. fPCA identified forces at the beginning (ca. 1-13%) the middle (ca. 24-46%) and the end (ca. 80-88% and ca. 90-95%) of the movement cycle as predictor variables. Predictor variables used in the ACP regression model were phases in the beginning (1-9%) and the middle (28-42% and 57-69%) utilizing the force-time domain; and the end (91-94%) of the movement cycles utilizing the time, force and force-time domains (Table 1). The fPCA and ACP regression model out-perform the discrete point regression model in respect to (a) their ability to describe jump height, and (b) generate meaningful information to develop or improve training programs.

The comparison of ACP to fPCA indicates that ACP has the greater ability to describe jump height than fPCA². fPCA generated lower r^2 -value than ACP because it does consider only the magnitude domain, discarding both the time and magnitude-time domain. The time domain should always be included during an analysis as it holds important information – the standardized coefficient of the time domain was one of the largest in the discrete point and ACP regression model. Consequently, including the variable 'phase duration' to the fPCA regression model increases its ability to explain variances in jump height by 12% ($r^2 = 94.8\%$). This highlights the importance of the time domain. Further, the importance of examining the magnitude-time domain is indicated by the ACP regression model. Selected variables are frequently from the magnitude-time domain. Based on the fact that the regression analysis chose variables capturing the combined magnitude-time information rather than the magnitude information alone, highlights the need for examining the combined domain.

¹ Mean maximum force position = 39%; standard deviation maximum force position = 29%;

² This was also the case when the number of predictor variables in the ACP model was reduced. An ACP regression model with four predictor variables already described 97.3% of the variance in the jump height.

CONCLUSION: FPCA and ACP showed a greater ability to describe jump height compared to discrete point analysis because FPCA and ACP: (a) analysed only related phases of the force curve and hence examine comparable neuromuscular capacities, (b) analysed the whole force curve rather than prior selected discrete data points, and (c) can identify features that occur solely as a phase. The comparison between fPCA and ACP highlighted a greater ability of ACP to describe jump height. ACP out-performed fPCA, because it (a) analyzes the time, magnitude and magnitude-time domain of a curve and, (b) examines specifically only key-phases rather than the full movement cycle.

REFERENCES:

- Chau, T. (2001a). A review of analytical techniques for gait data. part 1: Fuzzy, statistical and fractal methods. *Gait & Posture*, 13(1), 49-66.
- Chau, T. (2001b). A review of analytical techniques for gait data. part 2: Neural network and wavelet methods. *Gait & Posture*, 13(2), 102-120.
- Cormie P, McBride JM, McCaulley GO. (2009). Power-time, force-time, and velocity-time curve analysis of the countermovement jump: Impact of training. *The Journal of Strength & Conditioning Research*, 23(1), 177-186.
- Donà G, Preatoni E, Cobelli C, Rodano R, Harrison AJ. (2009). Application of functional principal component analysis in race walking: An emerging methodology. *Sports Biomechanics*, 8(4), 284-301.
- Donoghue OA, Harrison AJ, Coffey N, Hayes K. (2008). Functional data analysis of running kinematics in chronic achilles tendon injury. *Medicine & Science in Sports & Exercise*, 40(5), 1323-1335.
- Dowling JJ, Vamos L. (1993) Identification of kinetic and temporal factors related to vertical jump performance. *Journal of Applied Biomechanics*, 9(4), 95-110.
- Jolliffe IT. (2002). Principal component analysis, Springer Verlag, New York, USA
- Kelso, J. A. S. (1995). Dynamic patterns: The self-organization of brain and behavior (1st ed.). Massachusetts: The MIT Press.
- Newton RU, Kraemer WJ, Häkkinen K. (2012). Effects of ballistic training on preseason preparation of elite volleyball players. *Medicine & Science in Sports & Exercise*, 31(2), 323-330.
- Petushek E, Garceau L, Ebben W. (2012). Force, velocity, and power adaptations in response to a periodized plyometric training program. *International Symposium of Biomechanics in Sports*, 28(1), 262-265.
- Ramsay JO. (2006). Functional data analysis, Springer Verlag, New York, USA
- Richter, C., O'Connor, N.E., Moran, K. (2012) Comparison of discrete point and continuous data analysis for identifying performance determining factors. In *Proceedings of XXX Congress of the International Society of Biomechanics in Sports*, 384-387.
- Sheppard JM, Cronin JB, Gabbett TJ, McGuigan MR, Etxebarria N, Newton RU. (2008). Relative importance of strength, power, and anthropometric measures to jump performance of elite volleyball players. *The Journal of Strength & Conditioning Research*, 22(3), 758-765.

Acknowledgements

This work is supported by Science Foundation Ireland under grant 07/CE/I114.