

Using Bootstrapping Techniques to Improve the Classification of Healthy and Abnormal Gait Patterns

Morris H. Morgan III¹, Yannis Halkiadakis², Helia Mahzoun Alzakerin², Kristin D. Morgan²

¹MECK Limited, LLC, Williamsburg, VA 23185

²Department of Biomedical Engineering, University of Connecticut, Storrs, CT 06269

Abstract

Gait asymmetry is often observed in populations with varying degrees of altered neuromuscular control. Traditionally, gait asymmetry is evaluated using mean peak vertical ground reaction forces (vGRF). However, mean peak vGRF does not have the level of sensitivity needed to detect subtle changes in gait movement patterns. Biomechanically, it appears that time series generated from peak vGRF waveforms may be suitable to assess changes in abnormal gait patterns. Previously, autoregressive (AR) modeling has been used successfully to delineate between healthy and pathological gait during running; but has been little explored in walking. In running, the coefficients of the AR model fit to peak vGRF data were used to differentiate between healthy controls and post-anterior cruciate ligament reconstruction (ACLR) individuals. To expand upon this work, we will investigate the ability of AR modeling to both quantitatively and graphically delineate between healthy and abnormal gait patterns in healthy controls. We will use an asymmetric walking protocol to simulate abnormal gait patterns and fit the AR model to both the normal and asymmetric walking patterns. To address the scarcity of our limited dataset, classical bootstrapping techniques will be employed to increase the number of peak vGRF time series and improve our ability to construct viable and robust AR models to aid in the statistical based-classification of healthy and abnormal gait.

Introduction

Autoregressive (AR) modeling is a data-driven statistical method that characterizes time series as a function of their prior correlated values.^{1,2} In biological settings AR modeling has been used extensively, for example to extract respiratory rates and detect propulsive force asymmetries and motor control strategies from physiological signals.^{3,4} In such studies, AR model coefficients capture relevant temporal features. Herein, a second-order AR(2) model, was used to fit peak vertical ground reaction force (vGRF) running data and evaluate gait dynamic stability. These AR(2) models were found to provide the best fit based on the analysis of autocorrelation function and partial autocorrelation function plots that modeled the current vGRF peak value as a function of the two preceding vGRF peaks.^{1,2} The derived AR model coefficients were important because they provided both a quantitative and visual way of denoting between limb stability using the stationarity triangle plot devised by Box and Jenkins¹. This stationarity triangle could be quickly used to detect stable or unstable time series dynamics based on the AR(2) coefficients placement within or outside the triangle boundaries.^{1,2} Those AR(2) coefficients that resided inside the triangle characterized stable time series function, whereas those coefficients that resided outside of the triangle are indicative of unstable time series dynamics. Furthermore, this movement of AR(2) coefficients between different regions within the triangle, also captures changes in underlying gait pattern dynamics.^{1,2} The peak vGRF was the variable selected to assess between-limb stability because vGRF generated asymmetries are known to remain for years after ACLR and rehabilitation and are strongly correlated with altered limb loading and elevated injury risk.^{5,6} Given that a goal of post-ACLR rehabilitation is to track the restoration of functional stability, AR modeling should be a valuable tool for both assessing and visually tracking changes in dynamic stability in post-ACLR individuals.

The objective of the current study is to use bootstrapping to enhance our ability to establish criteria that delineate differences in dynamic stability between control and post-ACLR individuals from peak vGRF running data. Typically, peak vGRF datasets are not of sufficient size for classification purposes. Hence, alternative methods, such as bootstrapping, must be used to achieve a better understanding of the underlying distributional feature of small datasets. In this study we will use the dataset recently produced by Morgan (2019)⁸ summarized below to determine the underlying distributional features of running gait data and to devise robust classification criteria.

Methods

Participant Demographics

A total of 15 healthy controls (height 1.7 [0.1] m; mass 63.5 [13.5] kg; age 21.0 [4.4] yr; speed 2.6 [0.3] m/s) and 15 post-ACLR (height 1.7 [0.1] m; mass 66.8 [9.6] kg; age 21.1 [8.4] yr; average running speed 2.7 [0.3] m/s) participants performed a running protocol on a split-belt treadmill. Each group needed 15 participants to identify a moderate effect (0.70) and achieve adequate statistical power ($\alpha = .05$; $1 - \beta = 0.80$). The participant's ages ranged from 18 to 40 years. Written consent was obtained from each individual prior to their participation in the study in accordance with the institutional review board. The post-ACLR participant's reconstruction surgery was performed by physicians within the same orthopedic practice, where participants received either a bone-patellar bone or a hamstring graft. Each post-ACLR participant was cleared for return-to-sport drills by their physician prior to their participation. All control participants had no previous knee surgery and were injury free for at least 6 months prior to the study.

Instrumented Gait Analysis

For the running protocol, participants ran at a self-selected speed to get acclimated to the instrumented split-belt treadmill (Bertec Corp, Columbus, OH). Once acclimated, participants were instructed to jog at a comfortable pace. Participants wore WR662 running sneakers (New Balance, Brighton, MA). GRF data were collected at 1200 Hz and filtered using a zero-lag, fourth-order Butterworth filter with a 35-Hz low-pass cutoff frequency. Peak vGRF data were extracted from each limb for each running stride during a 10-second trial. A time series was created by extracting alternating vGRF peaks from the individual's right and left limbs. This meant that the vGRF was first extracted from the right limb and then was followed by the vGRF peak for the left limb. This process of compiling the alternating peak vGRF for the right and left limbs was repeated for the entire running interval. This process was repeated for the ACLR participants. Their time series were constructed from the vGRF peaks from their reconstructed and non-reconstructed limbs.

Autoregressive Modeling Analysis

First, a linear trend was removed from the vGRF time series. Then, an AR(2) was fit to the detrended vGRF peak time series. An AR(2) model was selected based on the results of the autocorrelation function and partial autocorrelation function plots, which illustrate how a time series is strongly correlated with a lagged version of itself.^{1,2} An AR(2) model was deemed appropriate because the autocorrelation function plots were shown to slowly decay and the partial autocorrelation function plots displayed a drop-off after a lag of 2.^{1,2} To validate the model fit, a histogram of the residuals, obtained from subtracting the AR(2) generated model value from the original time series, was plotted. The model was a good fit because the residuals were normally distributed about zero.^{1,2} An AR(2) model is appropriate here because the time series model errors are not independent but serially correlated.^{1,2} The AR(2) model generated 2 coefficients, AR1 and AR2, for each time series, which were plotted on a stationarity triangle. The stationarity triangle indicates the behavior of the time series; more specifically, it indicates if the time series is stable.^{1,2} The coefficients of stable time series lie inside of the stationarity triangle.^{1,2} Coefficients that reside outside of the triangle are unstable. As coefficients move to the edge of the triangle, they indicate the time series is less stable.^{1,2} The dimensionless AR(2) model coefficients, AR1 and AR2, acted as the x- and y-coordinates, respectively, and were plotted on the stationarity triangle. The distance of the dimensionless AR(2) coefficients from the centroid of the triangle, located at $(0, -1/3)$, was used to quantify the stationarity of the model. The analyses were conducted using a custom MATLAB code (MATLAB R2018a; The MathWorks, Inc, Natick, MA).

Results

The AR(2) coefficients for the post-ACLR and healthy controls indicated that both groups resided in the regions 1 and 2 on the AR(2) stationarity triangle thus indicating that both groups exhibited an overdamped response (Fig. 1). Moreover, the results indicated that the post-ACLR individuals reside predominantly in region 1 while the healthy controls encompassed regions 1 and 2 (Fig. 1).

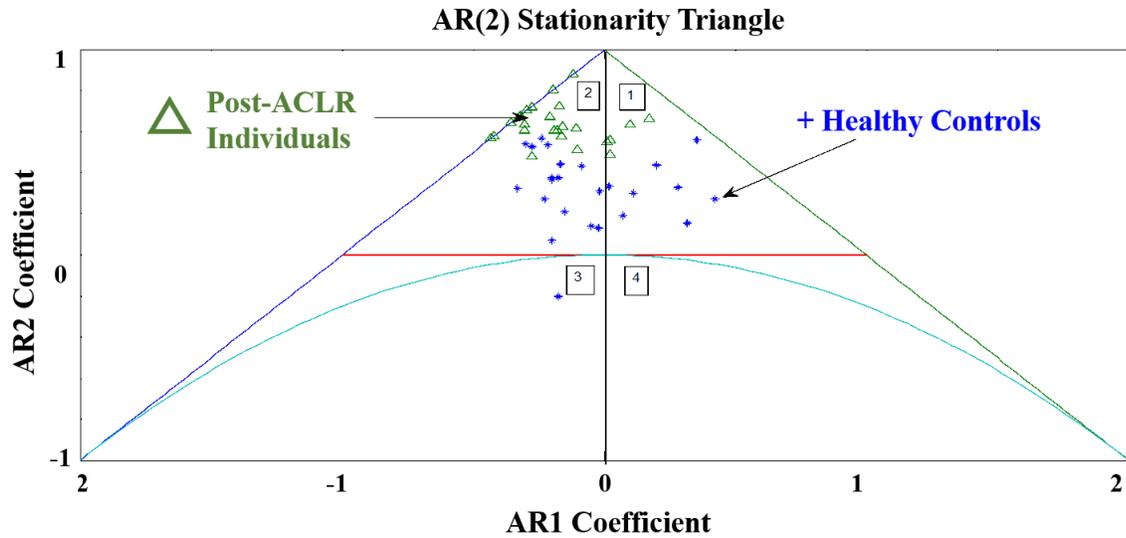


Figure 1. Comparison of the post-ACLR and healthy control individuals AR model coefficients displayed on a AR(2) stationarity triangle.

Bootstrapping was used to obtain 400 independent estimates of the two separate AR1 and AR2 coefficients under the provision that the values of the respective AR1 and AR2 coefficients must lie within the boundaries of the stationarity triangle. The discriminant value (DC) were also evaluated for each paired set of AR1 and AR2 coefficients (Eq. 1) and used to plot discriminant curve contour lines. Normal probability plots of the respective AR coefficients were found to exhibit normal behavior (Fig. 2).

$$DC = AR1^2 + 4 * AR2 \quad (1)$$

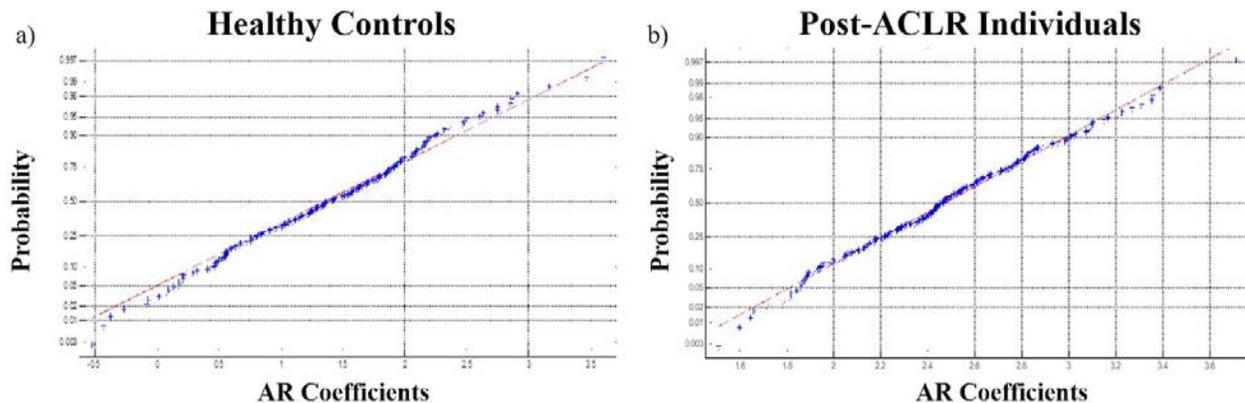


Figure 2. (a) Normal probability plot for the AR(2) model coefficients for the healthy controls. (b) Normal probability plot for the AR(2) model coefficients for the post-ACLR individuals.

The results below display how the healthy controls and post-ACLR groups are distributed in various

regions of the triangle with the post-ACLR subjects clustered mostly in region 1. The healthy controls mainly resided in regions 1 and 2 but can also be found dispersed in the underdamped regions 3 and 4 (Fig. 3).

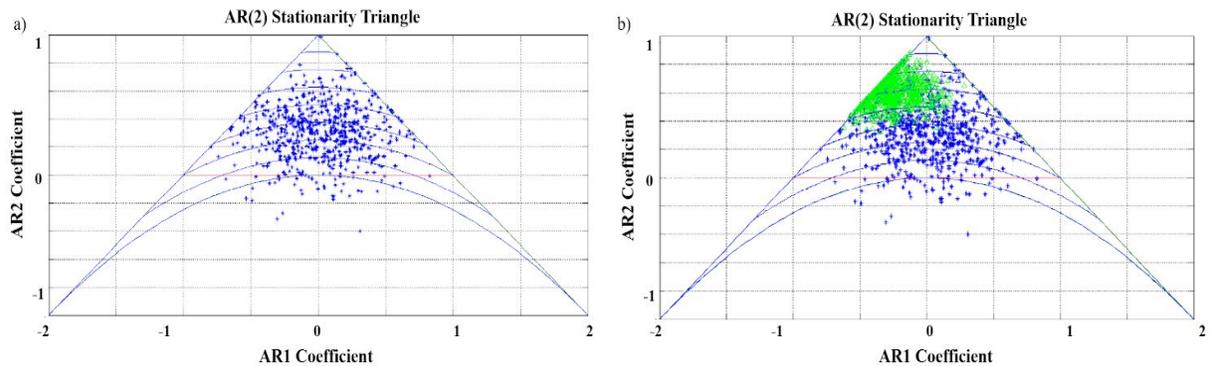


Figure 3. (a) The bootstrapped AR(2) model coefficients for the healthy controls (blue). (b) Comparison of the bootstrapped AR(2) model coefficients for the healthy controls (blue) and post-ACLR individuals (green).

Three simple geometrical classification criteria were devised using an objective function that minimizes the sum of the total number of false positive and false negative observations in each case. One criterion was a best symmetric system (BSS) that used the optimal placement of intersecting horizontal ($AR_2=0.3$) and vertical ($AR_1=0$) lines to locate the critical classification pathway. A second criterion was a best symmetric alternative (BSA) that used the optimal placement of an intersecting horizontal line ($AR_2=0.3$) with the left boundary line ($AR_2=AR_1+0.7$) of the stationarity triangle to identify the critical classification pathway. The remaining and final criterion was a quadratic based system (QBS) that used the optimal placement of the discriminant curve to local the desired critical classification pathway. This final criterion satisfied the interval constraint of $1.4 < DC < 4$. To quantify the ability of the three criteria to accurately classify the individuals we generated a separate confusion matrix (Tables 1-3) for each case. A confusion matrix consists of four quadrants that summarizes all possible error states. A practical goal of any effective classification scheme is to minimize the total number of type 1 and 2 errors. Type 1 errors are false- positives while type 2 errors are true-negatives. Interestingly enough, all three criteria yield essentially the same level of performance. Each had a high sensitivity ($> 94\%$), an excellent accuracy level ($>90\%$) and a low error rate ($<10\%$). Although more sophisticated classification schemes can be employed for this type of analysis, in light of the performance of our simple geometrical schemes this analysis was sufficient.

Table 1. Confusion Matrix for Criterion 1: BSS Method

N=400		Predicted No	Predicted Yes	
Actual No		TN=172	FP=28	200
Actual Yes		FN=12	TP=188	200
		184	216	400

Table 2. Confusion Matrix Criterion 2: BSA Method

N=400		Predicted No	Predicted Yes	
Actual No		TN=164	FP=36	200
Actual Yes		FN=2	TP=198	200
		166	234	400

Table 3. Confusion Matrix for Criterion 3: QBS Method

N=400	Predicted No	Predicted Yes	
Actual No	TN=159	FP=41	200
Actual Yes	FN=0	TP=200	200
	159	241	400

Table 4. Summary Table of Classification Models for the Three Methods

	Method 1: BSS	Method 2: BSA	Method 3: QBS
Sensitivity	0.94	0.99	1.00
False Alarm Rate	0.14	0.09	0.20
Accuracy	0.90	0.91	0.90
Precision	0.87	0.85	0.93
Error Rate	0.10	0.09	0.10

Discussion

The purpose of this study was to employ bootstrapping to aid in identifying optimal criteria to delineate between healthy controls and post-ACLR individuals based on their running dynamics. Gait biomechanics studies are often hampered by their limited sample sizes, which make drawing conclusions about larger general populations difficult. However, bootstrapping allowed us to increase our sample size and, together with our three newly derived criteria we were able to classify the healthy controls and post-ACLR with greater than 90% accuracy. We were able to accurately classify individuals based on our ability to use AR modeling to characterize gait dynamics. AR modeling has not been extensively used in gait biomechanics; however, previous work has shown that it has potential to identify differences in gait biomechanics between healthy controls and those suffering from neuromuscular or cognitive impairments.^{8,18} These findings works highlight the robustness of this technique and support its use as a clinical tool.

Conclusion

Bootstrapping of the AR(2) model coefficients showed how post-ACLR and healthy individuals dispersed across regions 1 and 2 of the AR (2) stationarity triangle. While several simple classification schemes lead to effective delineation of group behavior. Effective longitudinal studies are needed to access individual group migration rates that would aid in the development and tracking of rehabilitation protocols for post-ACLR populations.

References

1. Box GEP, Jenkins GM. Time Series Analysis: Forecasting and Control, revised ed. San Francisco, CA: Holden-Day; 1976.
2. Montgomery DC, Jennings CL, Kulahci M. Introduction to Time Series Analysis and Forecasting. Hoboken, NJ: John Wiley & Sons; 2008.
3. Koontz AM, Cooper RA, Boninger ML. An autoregressive modeling approach to analyzing wheelchair propulsion forces. Med Eng Phys. 2001;23(4):285–291. PubMed ID: 11427366 doi:10.1016/S1350-4533(00)00082-5
4. Lee J, Chon KH. An autoregressive model-based particle filtering algorithm for extraction of respiratory rates as high as 90 breaths per minute from pulse oximeter. IEEE Trans Biomed Eng. 2010;57(9): 2158–2167.

doi:10.1109/TBME.2010.2051330

5. Winter SL, Challis JH. Classifying the variability in impact and active peak vertical ground reaction forces during running using DFA and ARFIMA models. *Hum Move Sci.* 2017;51:153–160. doi:10.1016/j.Humov.2016.12.003
6. Boggess G, Morgan K, Johnson D, Ireland ML, Reinbolt JA, Noehren B. Neuromuscular compensatory strategies at the trunk and lower limb are not resolved following an ACL reconstruction. *Gait Posture.* 2018;60:81–87. PubMed ID: 29169096 doi:10.1016/j.gaitpost.2017.11.014
7. Efron, B., and Tibshirani, R. J. (1994). *An Introduction to the Bootstrap.* CRC press.
8. Morgan KD. Autoregressive Modeling as Diagnostic Tool to Identify Postanterior Cruciate Ligament Reconstruction Limb Asymmetry. *Journal of Applied Biomechanics.* 2019,35,388-392. <https://doi.org/10.1123/jab.2018-0414>.
9. Morgan KD, Donnelly CJ, Reinbolt JA. Elevated gastrocnemius forces compensate for decreased hamstrings forces during the weight acceptance phase of single-leg jump landing: implications for anterior cruciate ligament injury risk. *J Biomech.* 2014;47(13):3295–3302. PubMed ID: 25218505 doi:10.1016/j.jbiomech.2014.08.016
10. Hamill J, van Emmerik RE, Heiderscheit BC, Li L. A dynamical systems approach to lower extremity running injuries. *Clin Biomech.* 1999;14(5):297–308. doi:10.1016/S0268-0033(98)90092-4
11. Robertson GE, Caldwell GE, Hamill J, Kamen G, Whittlesey S. *Research Methods in Biomechanics.* Champaign, IL: Human Kinetics; 2013.
12. van Emmerik RE, Rosenstein MT, McDermott WJ, Hamill J. A nonlinear dynamics approach to human movement. *J Appl Biomech.* 2004;20(4):396–420. doi:10.1123/jab.20.4.396
13. Georgoulis AD, Moraiti C, Ristanis S, Stergiou N. A novel approach to measure variability in the anterior cruciate ligament deficient knee during walking: the use of the approximate entropy in orthopaedics. *J Clin Monit Comput.* 2006;20(1):11–18. PubMed ID: 16523229 doi:10.1007/s10877-006-1032-7
14. Davison and Hinkley. *Bootstrap Methods and Their Applications*
15. Vinod and Lopez-de-Lacalle. *Journal of Statistical Software.* Jan 2009.
16. Kline PW, Morgan KD, Johnson DL, Ireland ML, Noehren B. Impaired quadriceps rate of torque development and knee mechanics after anterior cruciate ligament reconstruction with patellar tendon autograft. *Am J Sports Med.* 2015;43(10):2553–2558.
17. Morgan KD, Zheng Y, Bush H, Noehren B. Nyquist and Bode stability criteria to assess changes in dynamic knee stability in healthy and anterior cruciate ligament reconstructed individuals during walking. *J Biomech.* 2016;49(9):1686–1691.
18. Alzakerin, HM, Halkiadakis Y, Morgan KD. Autoregressive modeling to assess stride time pattern stability in individuals with Huntington's disease. *BMC neurology.* 2019; 19(1), 316.