THE APPLICATION OF DETRENDED FLUCTUATION ANALYSIS IN RUNNING AND ITS INTEGRATION INTO A REAL-TIME SYSTEM

Michelle Norris[1], Avelino Amado[2], Joseph Hamill[2], Ian C. Kenny[1] and Ross Anderson[1].

Biomechanics Research Unit, University of Limerick, Limerick, Ireland[1]

Department of Kinesiology, University of Massachusetts, Amherst, MA 01003, USA[2]

Detrended fluctuation analysis (DFA) provides valuable information regarding both training and injury when applied to running time series. However, there is limited information when applied to recreational runners, or within a real-time environment. Firstly, DFA was applied to the stride time series of select training runs and competitive runs for recreational runners completing a half marathon and full marathon. Results indicate recreational runners maintain similar stride time dynamics in a half marathon, compared to training, however, stride time variability becomes increasingly deterministic during a marathon, compared to training. Secondly, we explore the implementation of DFA in a real-time system and provide evidence to support the use of DFA in running feedback.

KEY WORDS: accelerometry, feedback, running, stride time, variability.

INTRODUCTION: Detrended fluctuation analysis (DFA) has become popular in quantifying the temporal structure of fluctuations within biological time series, such as heartbeat (Peng et al., 1995) and walking stride time (Hausdorff et al., 1995). DFA ascertains that fluctuations do not happen randomly, but exhibit self-similarity and long range correlations which can reveal important information related to disease and ageing. Recently, DFA has been utilised within a sporting domain when applied to running stride time series, for example DFA has distinguished the training status of runners (Nakayama et al., 2010), and identified previously injured runners from their previously non-injured counterparts (Meardon et al., 2011). Progress has been made to investigate stride time variability within an ecological environment as Hoos et al. (2014) investigated long range correlations and pacing during a half marathon race. However, their runners were experienced distance runners and the study did not investigate stride time correlations. Of greater interest are recreational runners as they have been show to incur alterations to cadence and tibial acceleration during periods of prolonged running (Mizrahi et al., 2000). Also, recreational runners are affected by fatigue and reach functional limits in a race environment, to a much greater extent than experienced runners (Bertram & Prebeau-Menezes, 2013). Therefore, in the present study, firstly DFA was applied to recreational runners' stride time series for competitive and training distance runs, to investigate the presence of long range correlations during free-paced running. Secondly, whilst the information provided by DFA may be of significant importance to researchers and runners alike, analysis requires extended data collection. Studies investigating DFA on running populations have typically collected greater than 500 strides requiring increased post-processing capabilities, along with advanced statistical knowledge for correct implementation. This has led to DFA being underutilised within laboratory settings, and never utilised in consumer wearable running devices. Therefore, the second part of the current study discusses the implementation of a real-time DFA feedback system, with representative data supporting the practical application of DFA in the running domain.

METHODS: For the first part of this study comparing DFA applied to training and competitive runs, four recreational runners training for and completing a half marathon (age: 32.5 ± 6.1 years, height: 1.68 ± 0.09 m, mass: 68.4 ± 14.0 kg, personal best 10 km time: 52.2 ± 4.9 min) and three recreational runners training for and completing a marathon were recruited (age: 39.7 ± 3.1 years, height: 1.68 ± 0.13 m, mass: 67.0 ± 18.4 kg, personal best 10 km time: 50.7 ± 4.2 min). Participants were classed as recreational runners as they had not received specialised running training and were undertaking their first half/full marathon distance. Prior to half and full marathon completion runners undertook a self-led Hal Higdon running training programme, appropriate to their ability and end running goal. Runners were required to attach a tri-axial Shimmer 2r accelerometer (SHIMMER Ltd, Dublin, Ireland), to their anterio-medial distal tibia bi-laterally for each training run and the competitive distance event. Accelerometer placement was controlled via a demonstration and manual provided to participants prior to the beginning of data collection. When attached to the tibia a positive vertical acceleration was directed proximally, positive medio-lateral acceleration was directed laterally and positive anterio-posterior acceleration directed posteriorly. Data were sampled at 204.8 Hz (±6 g, sensitivity range of 200 mV/g). The runners' half or full marathon race data were analysed as competitive races, whilst their longest recorded training run was analysed as a non-competitive comparison (Table 1). Data processing was performed for right leg tibial accelerometry files using custom built MATLAB™ (Mathworks, Cambridge, UK) algorithms. Anterior-posterior accelerometry data were filtered at 2 Hz with a 2nd order Butterworth low-pass reverse filter and stride time identified as the time between peaks. Stride time series were subsequently visually checked for outliers, which were manually removed. Each stride time series was divided into three, even time length; a section representing the beginning, middle and end of each run, and the "overall" run was regarded as the whole stride time series. Stride time long range correlations were calculated using DFA, quantified with the scaling exponent, outlined by Peng et al., 1995. In brief, an α value closer to 1 indicates increased dependency of a stride to a previous stride at any given time, an α value closer to 0.5 indicating decreased dependency of a stride to a previous stride at any given time and an α value of less than 0.5 indicates a loss of correlation among different time scales (Meardon et al., 2011). Friedman tests were used to identify statistical significance between run sections (beginning, middle and end) and Wilcoxon-Sign Rank tests was used to identify statistical significance between training and competitive run sections. An alpha value of 0.05 was used to identify statistical significance. Effect sizes (ES) for α values were calculated between matching run sections across training and competitive run, in both half and full marathon groups. Due to small sample size Hedge's G, a modified version of Cohen's D, was employed with ES interpreted as small (0.2), medium (0.5) and large (0.8) (Hedges, 1981). For the second part of the current study an advanced running analysis system comprised of a Shimmer 2r accelerometer, a laptop equipped with Bluetooth capability, MATLAB (Mathworks, Cambridge, UK) and a PhysioNet C+ DFA programme. To provide representative data generated via the running analysis system a healthy active participant (female, age: 26.6 years, height: 1.80 m, mass: 70.1 kg) performed a treadmill running protocol whilst completing the advanced analysis system. The participant ran for an 18 minute period at their preferred running speed (PRS), which was then repeated at 80% of their PRS and 120% of their PRS. The participant was allowed as long as necessary to rest between runs to mitigate the effect of fatigue. Accelerometer attachment, data processing and stride time calculation was performed as outlined for study part one. However, accelerometry data were transmitted via Bluetooth in real-time, which allowed processing during each 18 minute run, at eleven user defined time points (Analysis numbers A1 – A11, Table 2) and at run cessation (A12, Table 2). DFA α values were then relayed on the laptop screen to the researcher. To verify the system met the specified requirement, real-time output of repeated DFA α values, the time difference in seconds (Δt) between the user time point selected and the related α value display time were calculated. To verify the system produced reliable α values over a range of running speeds α values were also recorded and compared to previous literature.

RESULTS & DISCUSSION: The number of strides calculated for each run were, half marathon training and competitive runs = 9,091 ± 1,771 and 9,854 ± 1,189, and marathon training and competitive runs = 15,689 ± 1,920 and 21,109 ± 3,069. There was no significant difference between run sections both within runs (beginning, middle and end), or across training and competitive runs, in both half and full marathon groups. DFA α values ranged between 0.84 and 1.07 across all run sections and types (training\competitive, half\full) (Table 1). Whilst our α values, calculated from participants in training and competitive runs, are higher than those previously found in treadmill running (0.70 – 0.90, Jordan et al., 2006), this may be due to increased visual cues and proprioceptive feedback in outdoor running strengthening correlations. Trivial and small effect size differences were identified between the training and competitive run in the half marathon group (beginning ES = 0.25, middle ES = 0.12, end ES = 0.06, overall ES = 0.40), whilst medium and large effect size differences were identified in the marathon group (beginning ES = 0.82, middle ES = 0.58, end = 0.58, overall = 0.51) (Table 1). It is possible that the large ES present within the half marathon group is due to the influence of a designated pacer or the selection of an incorrect pacing strategy causing a "biological stressor" upon runners. Stress conditions placed on gait have previously been found to increase persistency and therefore strengthen stride time long range correlations in running (Jordan et al., 2006). Whilst it could be argued that pacers are also utilised in half marathon running the participants within the current study performed similar distance to the competitive half marathon within their training programme (19 km v 21.1 km) and therefore would have had an accurate estimate of pacing when performing the competitive run, perhaps explaining trivial and small ES within the half marathon group.

Table 1. Group training and competitive run variables and DFA α values. Average results are presented with range in parentheses.

	Half Marathon (n = 4)		Marathon (n = 3)	
	Training Run	**Competitive Run**	**Training Run**	**Competitive Run**
Run Time (minutes)	106 (77 - 122)	114 (99 - 130)	176 (166 -192)	238 (218 – 267)
Distance (km)	19 (16.1 – 20.0)	21.1 (0.0 – 0.0)	31.6 (30.5 – 32.2)	42.2 (0.0 – 0.0)
Velocity (m/s)	2.99 (2.68 – 3.47)	3.10 (2.71 – 3.55)	3.00 (2.65 – 3.24)	2.98 (2.63 – 3.23)
α Beginning	0.95 (0.91 – 1.00)	0.93 (0.88 - 0.98)[b]	0.92 (0.84 – 0.96)	0.98 (0.92 – 1.03)[e]
α Middle	0.92 (0.84 - 0.98)	0.91 (0.85 – 0.99)[a]	0.89 (0.89 – 0.90)	0.93 (0.88 – 1.01)[c]
α End	0.95 (0.84 - 1.05)	0.94 (0.85 – 1.04)[a]	0.95 (0.91 – 1.00)	0.97 (0.95 – 0.99)[c]
α Overall	0.97 (0.89 - 1.07)	0.94 (0.87 – 0.99)[b]	0.94 (0.89 – 1.01)	0.97 (0.96 – 0.68)[c]

**Effect size between training and competitive run, [a] = trivial, [b] = small, [c] = medium, [e] = large.*

DFA does not investigate discrete values which are susceptible to immediate adjustments. Therefore, longer periods of data collection prior to result output is suggested as more appropriate within this system when referring to "real-time" DFA α value output. All DFA α values were displayed via the real-time analysis system within 0.83 – 2.19 seconds of user identified time points (average of 1.49 ± 0.41 seconds 80% of PRS; average of 1.55 ± 0.34 seconds PRS; and average of 1.28 ± 0.32 seconds 120% PRS) (Table 2). Post overall run DFA α value output occurred within 5 seconds, across all running speeds (average 3.61 ± 1.03 seconds). DFA α values from the real-time analysis system ranged 0.70 – 0.86, within the data analyses epochs (A1- A11), across all running velocities. Our results are similar to those found by Jordan et al. (2006), who identified α values of 0.70 – 0.90 whilst running at similar percentages of PRS. This may verify that our advanced running analysis system generates valid DFA α values over a range of running speeds. Interestingly, we also found that our participants' overall run α was lowest at 100% of PRS (0.80, compared to 0.85 at 80% of PRS and 0.92 at 120% of PRS). This was previously identified by Jordan et al. (2006) and is explained as a runner being most adaptable and therefore less predictable in their stride time, at their PRS. This further supports our system within a training and skill level identification setting, as the system is able to detect α value differences previously identified within the literature.

Table 2. Difference in user defined time point and α value output time (sec) Δt, and DFA α values over three running conditions at 80% PRS, PRS and 120% PRS.

	Δt			A		
Analysis No.	**80% PRS**	**PRS**	**120% PRS**	**80% PRS**	**PRS**	**120% PRS**
A1	2.19	1.13	0.93	0.82	0.75	0.83
A2	0.95	1.02	0.87	0.77	0.74	0.84
A3	0.83	1.76	0.88	0.74	0.72	0.85
A4	1.07	1.38	1.14	0.74	0.70	0.86
A5	1.30	1.18	1.02	0.75	0.71	0.82
A6	1.57	1.44	1.41	0.78	0.74	0.84
A7	1.51	1.76	1.49	0.79	0.73	0.80
A8	1.57	1.77	1.46	0.73	0.79	0.80
A9	1.75	1.69	1.59	0.77	0.81	0.83
A10	1.70	1.80	1.50	0.74	0.83	0.84
A11	1.90	2.10	1.80	0.78	0.81	0.75
Average (± stdev)	1.49 (± 0.41)	1.55 (± 0.34)	1.28 (± 0.32)	N|A	N|A	N|A
Overall Run A12	2.68	4.72	3.44	0.85	0.80	0.92

CONCLUSION:

DFA results here indicate that recreational runners maintain similar stride time dynamics when completing a half marathon distance compared to a long training run. However, possible biological stressors may impact stride variability within a marathon run. Also, the implemented running analysis system provides real-time output of advanced variability information, which previously required extensive data processing and analysis. This provides access to information important in both a training and injury prevention context, for coaches and researchers alike. As advances in Bluetooth technology occur further development of the system will allow advanced stride variability analysis in an ecologically valid environment.

REFERENCES:

Bertram, J. E. A., and Prebeau-Menezes, L. P. (2013) Gait characteristics over the course of a race in recreational marathon competitors. *Research Quarterly for Exercise and Sport*, **84**, 6-15.

Hausdorff, J. M., Peng, C.K., Ladin, Z., Wei, J. Y., and Goldberger, A. L. (1996). Is walking a random walk? Evidence for long-range correlations in stride interval of human gait. *Journal of Applied Physiology*, **78**, 349-358.

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational and Behavioural Statistics,* **6**, 107-128.

Hoos, O., Boeselt, T., Steiner, M., Hottenrott, K., and Beneke, R. (2014). Long-range correlations and complex regulation of pacing in long-distance road racing. *International Journal of Sports Physiology and performance,* **9**, 544-553.

Jordan, K., Challis, J. H., & Newell, K. M. (2006). Long range correlations in the stride interval of running. *Gait & Posture*, **24**, 120-125.

Mizrahi, J., Verbitsky, O., Isakov, E., and Daily, D. (2000). Effect of fatigue on leg kinematics and impact acceleration in long distance running. *Human Movement Science*, **19**, 139-151.

Nakayama, Y., Kudo, K., and Ohtsuki, T. (2010). Variability and fluctuation in running gait cycle of trained runners and non-runners. *Gait & Posture,* **31**, 331-335.

Meardon, S. A., Hamill, J., and Derrick, T. R. (2011). Running injury and stride time variability over a prolonged run. *Gait & Posture*, **33**, 36-40.

Peng, C. K., Havlin, S., Stanley, H. E., and Goldberger, A. L. (1995). Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science,* **5**, 82-87.