FRACTAL ANALYSES OF GAIT VARIABILITY DURING A MARATHON

Ben Hunter[1], Bettina Karsten[2], Andrew Greenhalgh[1], Mark Burnley[3], Aliaksandr Leuchanka[4], Tim Clark[5], and Daniel Muniz-Pumares[1].

School of Life and Medical Sciences, University of Hertfordshire, Hatfield, UK[1]
European University of Applied Sciences (EUFH), Berlin, Germany[2]
School of Sport and Exercise Sciences, University of Kent, Chatham Maritime, University of Kent, UK[3]
Altra Running, a VF Company, Denver, CO, USA[4]
Scribe Labs Inc., San Francisco, CA, USA[5]

Detrended fluctuation analysis (DFA) and Higuchi's fractal dimension (HG) have previously been used to characterise motor control during gait. However, there is limited evidence of either being applied to running gait within a race environment. The aims were to: i) examine statistical persistence and fractal dimension of stride dynamics during a marathon, and ii) explore the relationship between DFA and HG for running gait. Therefore, DFA and HG were applied to stride interval series of each km of the 2018 TCS New York Marathon. Results showed consistent persistence, variability, and fractal dimension of stride interval series throughout the marathon with no significant differences observed between the beginning, middle, and end of the Marathon. Moreover, HG was shown to correlate strongly with DFA, which may be useful in monitoring motor control using fractal analyses in real time, by decreasing computation time and improving robustness to changing time series lengths.

KEYWORDS: running, accelerometery, variability, nonlinear dynamics, gait.

INTRODUCTION: Healthy running gait, although stable, is characterised by variable stride dynamics. Further to linear measures including standard deviation, average, and coefficient of variation, methods such as detrended fluctuation analysis (DFA) have been used to reveal important information about the time-dependent fluctuations within a time series. Briefly, DFA can be used to quantify the statistical persistence or anti-persistence of a time series returning a scaling exponent (DFA-α). A time series can be considered persistent if deviations in a time series are followed by deviations in the same direction (e.g., a short stride interval followed by a short stride interval), and anti-persistent if deviations are more likely to occur in the opposite direction (e.g., a short stride interval followed by a long stride interval). A DFA-α value approaching 1 indicates the presence of statistical persistence (i.e., stride times that persist over many strides that result in a predictable pattern), <0.5 indicates an uncorrelated series (i.e., a short stride time followed by a long stride time), and >1.0 indicates statistical persistence (i.e., long-range correlations that do not decay with increasing time lag). Importantly, these measures can be used to elucidate underlying motor control function.
Previously, DFA has been used to quantify changes in gait dynamics in response to fatigue (Meardon et al., 2011; Mo & Chow, 2018; Bellenger et al., 2019). Statistical persistence has been shown to decrease over the course of a run performed at 5 km race pace (Meardon et al., 2011) and over a prolonged run, at a velocity equivalent to anaerobic threshold (Mo & Chow, 2018). Moreover, reductions in the persistence of stride intervals were noted after an overreaching protocol (Bellenger et al., 2019). These findings demonstrate statistical persistence of stride intervals, and therefore motor control, is affected by fatigue. Further to estimation of DFA-α, methods have been developed to determine the fractal dimension including Higuchi's Fractal Dimension (HG). However, HG has seldom been applied to gait research and has not previously been applied to running gait. HG is less affected by time series length when compared to DFA (Phinyomark et al., 2020). This has implications for stride dynamics as individuals will exhibit variable stride frequencies over a given distance. Briefly, HG computes the fractal dimension (D_{HG}), which is the extent to which the curve that represents the analysed signal on a plane and returns a value between 1 and 2 (Phinyomark et al., 2020). A value of 1 corresponds to a regular time series, whereas Gaussian-type noise returns different values: 1.5 for Brownian, 1.8 for pink, and 2.0 for white noise. In contrast to other methods including DFA, HG may be more suitable for analysis of short time series of a

non-stationary signal (i.e., stride dynamics) and only requires one input parameter. The robustness and speed of HG may suit it to implementation of real-time feedback systems in consumer wearable devices during running. This may provide runners with an indication of fatigue status without the need for invasive measures, or the use of more lengthy algorithms prone to changes dependent on time series length (i.e., DFA).

The popularity of footworn accelerometers, commonly known as foot pods, capable of sampling at high frequencies, gives the possibility of analysing stride interval characteristics in the field. Measures of fractal scaling have been applied to stride time series for longer events including a half marathon and marathon (Hoos et al., 2014; Norris et al., 2016). However, the utility of analysing the whole time series (Hoos et al., 2014) or sections equivalent to a third of the event (Norris et al., 2016), may be limited when considering the effects of fatigue. Given the onset of fatigue may be manifested at various points throughout a marathon due to dynamic pace fluctuations (Casado et al., 2020) and other processes, and stride interval dynamics are sensitive to fatigue (Meardon et al., 2011; Mo & Chow, 2018), a more granular analysis is warranted. The use of shorter time series may permit ongoing feedback with smaller computational demand. Interestingly, Norris et al. (2016) demonstrated the potential utility of DFA in a real-time monitoring system using 8 min epochs. However, this method resulted in differing numbers of strides per epoch, and thus different time series lengths being analysed. This has been shown to affect the accuracy of DFA-α estimation. Therefore, the aim of this study was twofold. The primary aim was to examine variability, statistical persistence, and fractal dimension of stride dynamics over the course of a marathon. Given the robustness and rapid computability of HG, a secondary aim was to compare the estimation of DFA-α and D_{HG}.

METHODS: The anonymised dataset for this study consisted of activity logged by 14 participants (10 M; 4F, age: 38.3 ± 8.5 yr; finishing time: 210.49 ± 35.33 mins) running the 2018 TCS New York Marathon. The study was reviewed by the ethics board at the University of Hertfordshire and deemed to be exempt from ethical approval since no data were collected. Each participant wore RunScribe^{TM} IMUs (Scribe Labs Inc., San Francisco, CA, USA) on trainers in line with manufacturer's instructions. RunScribe^{TM} IMUs contain a tri-axial accelerometer, magnetometer and gyroscope, and a barometric altimeter. Data were recorded bilaterally at 500 Hz continuously during the entire 42.2 km. Data from the right foot were used for further analysis. Interstride interval (S_{INT}) was defined as the time between consecutive foot strikes by a proprietary algorithm (Scribe Labs Inc., San Francisco, CA, USA).

For each km, DFA-α, D_{HG}, standard deviation (SD), coefficient of variation (CV%), and means were computed for S_{INT} using MATLAB (v2021a, Mathworks, Cambridge, UK). The DFA and HG algorithms have been described elsewhere (Phinyomark et al., 2020). The parameters for DFA and HG were determined *a priori* by analysing 20 signals of 240 data points using the *ColoredNoise System* object in MATLAB. The signals included white noise (DFA-$\alpha \simeq 0.5$, $D_{HG} \simeq 2.0$), pink noise (DFA-$\alpha \simeq 1.0$, $D_{HG} \simeq 1.8$), and Brown noise (DFA-$\alpha \simeq 1.5$, $D_{HG} \simeq 1.5$). The parameters that exhibited the smallest deviation from the theoretical values and lowest variability were used in subsequent analysis. Therefore, DFA-α was calculated using box sizes ranging from 4 to 60 with an arithmetic progression of 6, and a polynomial order of 1, and D_{HG} was computed using a maximum interval length of 16.

As DFA-α is sensitive to changes in series length, only the 240 consecutive strides corresponding to the midpoint of each km, were analysed for each participant. The first 20 strides of the race were omitted to avoid start up effects and the last 195 m were also omitted. Outliers from each time series were removed if they were outside 1.5 times the inter-quartile range. This allowed for stride dynamics attributable to intrinsic factors to be analysed, rather than datapoints affected by external factors such as slowing at aid stations or Bluetooth dropout. The median number of outliers removed was 10. A repeated measures analysis of variance (ANOVA) was used to test for differences between the beginning (5 km), middle (23 km), and end (37 km), representing the flattest sections of the race. Data were tested for normality using the Kolmogorov–Smirnov test. Non-normal data were log transformed. Sphericity was tested using Mauchly's test, with Huynh–Feldt corrections made for violations ($P < 0.05$). All statistical analyses were performed in SPSS (v27, IBM, NY, USA). An alpha level for significance was set at $P < 0.05$. Data are presented as means ± SD. Associations between D_{HG} and DFA-α were explored by using Spearman's correlation coefficient.

RESULTS & DISCUSSION: There were no significant effects for section on measures of variability, statistical persistence, and fractal dimension of S_{INT} between the beginning, middle and end of the marathon (Table 1). Figure 1 shows that measures of statistical persistence and fractal dimension, as well as traditional measures of variability, exhibited steady state behaviour and with little variation throughout a marathon. The DFA-α and D_{HG} values showed a predictable and strongly persistent time series across all time points. Our results confirm the findings of Norris et al. (2016), which showed no significant change in DFA-α during a marathon. However, the current study showed medium to large effect sizes, suggesting that persistence, fractal dimension, and mean S_{INT}, may be affected by fatigue experienced during a marathon. The lack of significance may be a consequence of the small sample size.

Table 1. Measures of DFA-α, D_{HG}, standard deviation, coefficient of variation, and mean of interstride interval at distance intervals corresponding to the beginning (5 km), middle (23 km) and end (37 km) of the marathon. Mean ± SD.

	Beginning	Middle	End	η^2	*P*
DFA-α	0.804 ± 0.134	0.802 ± 0.128	0.759 ± 0.134	0.063	0.427
D_{HG}	1.928 ± 0.047	1.928 ± 0.063	1.953 ± 0.045	0.133	0.156
SD	8.893 ± 2.207	8.544 ± 1.988	8.545 ± 1.868	0.019	0.775
CV	1.294 ± 0.322	1.237 ± 0.301	1.238 ± 0.305	0.023	0.736
Mean (s)	0.688 ± 0.351	0.692 ± 0.399	0.695 ± 0.396	0.148	0.125

Previously, significantly reduced DFA-α has been shown during running tasks (Meardon et al., 2011; Mo and Chow, 2018). The differences in findings may be due to disparities in task demands, with both studies designed to induce high levels of metabolic perturbation. As changes to motor control have been shown to be dependent on metabolic rate (Pethick et al., 2016), stride dynamics may also be affected as a system output. Marathons are typically performed in the moderate and heavy intensity domains, where a metabolic steady state can be attained (Poole et al., 2016). It is unlikely that participants performed a significant portion of the race in the severe intensity domain, as to do so would result in accelerated fatigue and a marked decrease in pace, not seen in the current data set. However, given previously documented fatigue mediated changes to stride interval persistence (Bellenger et al., 2019) it is likely that central fatigue, as well as muscle damage, also mediate stride interval dynamics. Further research should examine whether stride dynamics change prior to a noticeable deterioration in pace during the marathon. The increase in mean S_{INT}, evidenced by a large effect size, may be an adaptation to fatigue to potentially alter the degrees of freedom to mitigate the possibility of injury or harm.

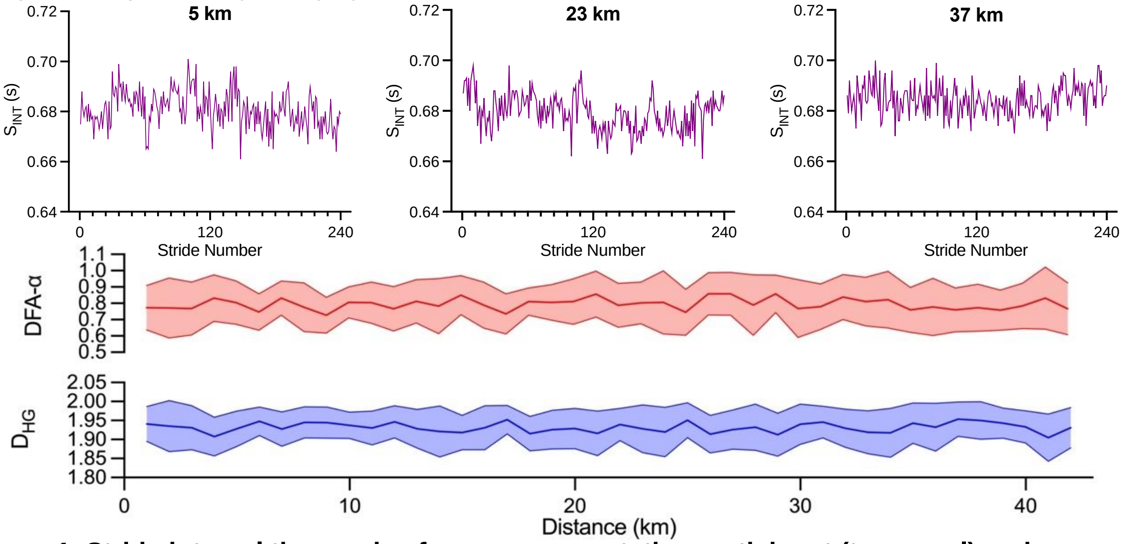

Figure 1. Stride interval time series from a representative participant (top panel) and mean changes to DFA-a and DHG of interstride intervals throughout the marathon of all participants (bottom panel). Shaded areas denote SD.

Correlation coefficients ρ between estimations of DFA-α and D_{HG} from all km intervals and participants were -0.866 (P < 0.001) as shown in Figure 2. The strong negative linear relationship between DFA-α *and* D_{HG} evident in the current study is congruent with the findings of Phiyomark et al. (2020), which showed a similar relationship (ρ > 0.8) during walking. Therefore, D_{HG} could be a viable measure of stride interval characteristics during running, providing practitioners and researchers with a quick and robust measure of non-linear stride interval properties. This could be applied to wearable technology where simple data processing methods would be desirable to reduce computational demand and provide prompt real-time feedback. However, it has been suggested that estimations of both fractal exponent (e.g., DFA-α) and fractal dimension (e.g., D_{HG}) could be in conjunction to provide a complete classification of gait (Dierick et al., 2017). Further research should examine this potential complimentary method of analysis and its utility during running gait.

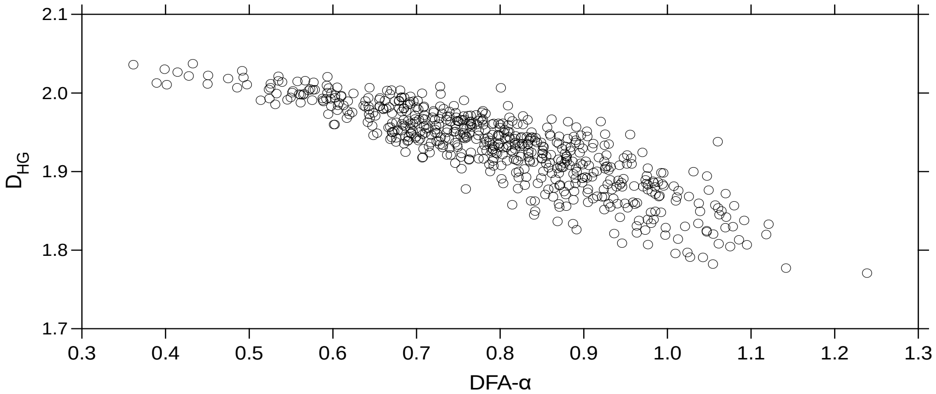

Figure 2. Correlation coefficients ρ between estimated DFA-α and D_{HG} on the entire dataset.

CONCLUSION: Our findings show that stride time dynamics remain consistent throughout a marathon. However, further research should elucidate whether these parameters change in line with maladaptive pacing and resultant accumulation of fatigue. The strong negative correlation between DFA-α and D_{HG}, suggests that D_{HG} may provide practitioners and researchers with a viable alternative to monitor stride time dynamics with shorter data sets. Moreover, the fast computation of HG can provide real-time analysis of data which may be useful for real time monitoring of motor control in athletes in conjunction with foot pods.

REFERENCES

Bellenger, C. R., Arnold, J. B., Buckley, J. D., Thewlis, D., & Fuller, J. T. (2019). Detrended fluctuation analysis detects altered coordination of running gait in athletes following a heavy period of training. *Journal of science and medicine in sport*, *22*(3), 294-299.

Casado, A., Hanley, B., Jiménez-Reyes, P., & Renfree, A. (2020). Pacing profiles and tactical behaviors of elite runners. *Journal of Sport and Health Science*.

Dierick, F., Nivard, A. L., White, O., & Buisseret, F. (2017). Fractal analyses reveal independent complexity and predictability of gait. *PLoS one*, *12*(11), e0188711.

Hoos, O., Boeselt, T., Steiner, M., Hottenrott, K., & Beneke, R. (2014). Long-range correlations and complex regulation of pacing in long-distance road racing. *International journal of sports physiology and performance*, *9*(3), 544-553.

Meardon, S. A., Hamill, J., & Derrick, T. R. (2011). Running injury and stride time variability over a prolonged run. *Gait & posture*, *33*(1), 36-40.

Mo, S., & Chow, D. H. (2018). Stride-to-stride variability and complexity between novice and experienced runners during a prolonged run at anaerobic threshold speed. *Gait & posture*, *64*, 7-11.

Norris, M., Amado, A., Hamill, J., Kenny, I. C., & Anderson, R. (2016). The application of detrended fluctuation analysis in running and its integration into a real-time system. In *2016 International Society of Biomechanics in Sports Conference*. International Society of Biomechanics in Sport.

Pethick, J., Winter, S. L., & Burnley, M. (2016). Loss of knee extensor torque complexity during fatiguing isometric muscle contractions occurs exclusively above the critical torque. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *310*(11), R1144-R1153.

Poole, D. C., Burnley, M., Vanhatalo, A., Rossiter, H. B., & Jones, A. M. (2016). Critical power: an important fatigue threshold in exercise physiology. *Medicine and science in sports and exercise*, *48*(11), 2320.