Physica A 515 (2019) 240–247

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Detecting the marathon asymmetry with a statistical signature

Véronique Billat [a], Thomas Carbillet [b,*], Matthieu Correa [a], Jean-Renaud Pycke [c]

[a] Unité de Biologie Intégrative des Adaptations à l'Exercice, Université Evry, Université Paris-Saclay, 91025, Evry cedex, France
[b] LJK - Laboratoire Jean Kuntzmann, UJF - Université Joseph Fourier, 38401, Saint-Martin d'Hères, France
[c] Laboratoire de Mathématiques et Modélisation d'Evry, CNRS, Université Evry, Université Paris-Saclay, 91037, Evry, France

HIGHLIGHTS

- 70% of the runners tend towards having a strong asymmetric behavior.
- Runners are spending more time above their marathon average speed than below.
- The statistical signature of a marathon keeps the same order $(--)(-+)(++)(+-)$.

ARTICLE INFO

Article history:
Received 13 July 2018
Available online xxxx

Keywords:
Running
Runners
Endurance
Kendall
Strava
Signature

ABSTRACT

Lately, the sub two-hour marathon attempt in Monza was still based on the belief that constant speed is the best way of running. This idea is relayed by marathon organizers who offer pace-group leaders to help the runners to maintain a target race speed. The purposes of this study are to verify the hypotheses that 1. The mass runners try to maintain a constant speed without succeeding. 2. Marathoners run in an asymmetric way and this turns out to be visible in the speed time series. Those two points are independent of the gender, the level of performance (2h30–3h40) and the profile of the race (Paris vs Berlin). Before considering a predictive running strategy for optimizing personal marathon running performance, here we shed light on some significant statistical features by analyzing speed time series data recorded by 273 runners' GPS. We started with looking for a trend in the speed time series. By means of Kendall's non-parametric rank correlation coefficient we exhibited a decreasing trend in speed data, whichever the level of performance, gender (Male and Female) and race profile (Berlin and Paris marathons). Going deeper in the study we applied a systematic analysis of the asymmetry of speed via classical statistical measures of skewness. Among them the quantiles of the average speed, i.e. the proportion of the race run above or below the final average The combination of the trend and the asymmetry lead to building up a statistical signature for the speed time series which is identical regardless the level of performance, gender and race profile.

1. Introduction

Nowadays the number of runners in the USA equals the whole French population, i.e. 65 millions with an increasing rate of 50% over the last ten years.[1] In the same way, the number of marathon finishers increased, from 353,000 finishers in the

* Correspondence to: 700 Avenue Centrale, Campus de Saint Martin d'Hères, 38401 Domaine Universitaire de Saint-Martin-d'Hères, France.
E-mail address: thomas.carbillet@billatraining.com (T. Carbillet).

1 https://www.statista.com/statistics/227423/number-of-joggers-and-runners-usa/.

https://doi.org/10.1016/j.physa.2018.09.159
0378-4371/

Table 1
Number of marathons run divided into two categories: All Marathons (273 Marathons) and Strongly Asymmetric Speed (191 races) each subdivided in seven groups: All, Men, Women, Less than 3 h marathons, Greater-than and equal to 3 h marathons, Paris and Berlin.

	All	Men	Women	< 3 h	≥ 3 h	Paris	Berlin
Whole dataset	273	230	43	130	143	140	133
Strong asymmetry	191	162	29	91	100	97	94

early 2000s to 507,600 in 2016 [1]. In spite of an increasing volume of data split available on running community websites no study has systematically investigated a marathon running pacing strategy for the various levels of performance, from 2h30 to 3h40.

Indeed variability in pacing has been studied in respect of short and middle-distance running (e.g. 3000 m to 10 km) [2–7], these studies having especially focused on the influence of pacing on metabolic and performance measures. Ely et al. (2008) reported that elite runners completing a marathon had few changes in pace during a marathon suggesting low speed variability [8]. Even lately, the sub two-hour marathon attempt in Monza was still based on the belief that constant speed is the best way of running following the prior idea that optimal is the even pace according to the seminal model of Keller (1973) [9]. In the same way recreational marathon runners adopt the paradigm of constant speed by running with a pace-group leader from 3h00 to 4h30 every 15 min provided by the marathon organizers. Therefore the idealized paradigm of constant speed prevails in theory and perverts the practice. However, considering the physiological limitation (glycogen availability) that means that the runners have to choose the ideal speed allowing them to get the optimal performance without hitting the set of fatigue 10 km or less before the arrival ("the famous Marathon wall")!

By applying methods from mathematical statistics we test the two following hypotheses: first, most marathoners run at a speed that is significantly increasing or decreasing ; second, they run in a significantly asymmetric way leading them, for instance, to run most of the time at a speed above their average speed including the final speed fall. In addition we test a third hypothesis that these features above are independent of 1. the gender, 2. the level of performance (2h30–3h40) and 3. the profile of the race (Paris vs. Berlin) by analyzing speed time series data currently recorded by 273 runners' GPS.

More precisely, the following three points have been studied:
(1) The first one is that we notice a decreasing trend for the speed time series while running a marathon. We therefore characterize this trend statistically via Kendall's τ [10,11].
(2) Beyond the decreasing trend we point out the presence of strong positive or negative skewness in speed time series. In other words, in the case of negative asymmetry, runners spend much more time and run a longer distance above the overall marathon average speed. The combination of the trend and the asymmetry lead to building up a statistical signature for the speed time series which appears to be identical regardless the level of performance, gender and race profile.
(3) Finally we introduce a new statistical signature combining the trend and the asymmetry analyses to help runners and coaches to be able to go beyond the sole performance analysis through chronometers results. In addition before considering a predictive running strategy for optimizing personal marathon running performance, here we shed light on some significant hidden running patterns.

2. Methods

2.1. Methodology

To achieve this study we examined the time and average speed per kilometer run by 273 marathons finishers all during 2017. These data were collected from the public platform Strava.[2]

We say that a marathon has a strong positive (resp. strong negative) asymmetry if the runner spent at least 54% of the time below (resp. above) marathon average speed. More formally in a sample of speeds measured at regular intervals, a strong positive (resp. negative) asymmetry corresponds to a percentile of the average speed greater than 54% (resp. less than 46%.)

We study the marathons through different layers introducing two categories and six groups of runners. In the first category, we take the whole dataset (273 marathons), in the second one, we only take the strongly asymmetric marathons (191 marathons).

Then we split each of the two categories into seven groups: All (A), Men (M), Women (W), Less than 3 h (-3), Greater-than or equal 3 h ($+3$), Paris (P) and Berlin (B). Table 1 gives the frequencies of the categories and groups.

All races were run using a Global Positioning System (GPS) device.

2 https://www.strava.com/.