UNITED STATES MILITARY ACADEMY

LITERATURE REVIEW

INITIAL WRITTEN PRODUCT

Aerobic Training and its Influence on Long-Distance Race Performance

Author
Creighton MORGENFELD

Supervisor
Dr. Margaret REYNOLDS

April 21, 2023

CHAPTER

1 STATISTICAL MODELS

Statistical models draw conclusions from real-world data. Researchers exploring a question may design an experiment or utilize existing sources to collect their data. They can then use a wide array statistical methods to evaluate the data and arrive at a solution. Statistical models gain credibility from their direct link to real-world data, although the model's descriptive ability is limited by the quality of the data or study.

Tanda [6]

Tanda studied 22 runners (21 male, 1 female) between ages 28-54 for a duration of 5 years. They ran 46 marathons, with times ranging from 2:47 to 3:36. The participants provided daily exercise records, typically with 3-month train-ups, including run distance and time. Various training indicators were considered, but ultimately not included in the final regressions, to include previous marathons, training days per week, mean workout distance per day and maximum workout distance. The two training indices included in the regression are mean workout distance per week and mean training pace.

The graphs provided in this paper could potentially lend themselves to be scraped using online tools and allow for additional analysis.

Billat [1]

Billat et al examined the physiological and training differences between high-level (HL) ($<$ 2:16 male and $<$ 2:38 female) and top-class (TC) ($<$ 2:11 male and $<$ 2:32 female) marathon runners. Specifically, the authors were interested in testing the independence of three factors that inform marathon running speed (vMarathon): energy cost of running (Cr), maximal oxygen consumption (VO2max), and that fraction that can be sustained throughout the race (FR). In a 1986 paper, di Prampero demonstrated the following relationship: $vMarathon = FR * VO_2max * Cr^{-1}$. Despite a reliance on physiological measures outside the scope of this study, Billat et al found that TC male athletes ran further than HL male athletes (by 40 km/wk) however, the load distribution was the same (18% faster than vMarathon, 4% at vMarathon, and 78% slower than vMarathon). Billat et al also found the training volume was not statistically different for the females, but TC female runners had more training sessions per week, and more at their 3,000m race pace.

Hagan [3]

Hagan et al studied 50 male volunteers to determine various indicators of marathon performance time (MPT). Hagan et al found that while VO2 max measurements were the most predictive of MPT, other physical and training factors also provide an informed prediction of MPT. Over the 9 week training period, various training indicators were collected including total distance, average pace, average workout length, total workouts, and total workout days.

$$MPT = 42.8 + 6.62 * \text{fastest5Mile} - 0.05 * \text{notFirstMarathon} - 1.45 * \text{longestTrainingRun}, R = 0.89$$

Doherty [2]

Doherty et al conducted a meta-analysis of 85 articles to included data from 8,945 runners (25% female). The authors sought to collect the many studies conducted on training indices' and their correlation with marathon performance to better inform training for runners from many different groups, as many studies focused on a specific group or demographic (elite male runners, for example). Doherty et al examined the impact of average weekly distance, number of weekly runs, number of runs greater than 32 km in length, maximum running distance in a single week, longest training run, average training pace (% of eventual race pace), and weekly training hours.

CHAPTER

2 MECHANISTIC MODELS

Mechanistic models ground potential solutions in the "mechanics" of real-world phenomenon. Mathematicians may look to biology, engineering, physics, or other fields to understand the processes and laws governing certain processes. While a statistical approach may struggle to provide data providing both internal and external validity to a model, a mechanistic approach instead works to determine the solve a general form of the problem before adjusting the values of constants to match real-world occurrences. The resulting model is more flexible and generalizable than a statistical model, although it may sacrifice complexity for mathematical elegance.

Woodside [7]

Woodside examines world track records for available men's and women's races ranging from 50 m to 275 km. Keller's 1974 model of running relies on a physiological approach to understanding the mechanics of running. His model involves the conversation of linear momentum, the conservation of energy, and that oxygen consumption yields energy. However, Keller's model only describes races with distances up to 10 km, races of greater length involve paces so slow, the rate of energy production, σ is greater than the rate of energy consumption, $f(t)v(t)$. To extend this model to races of greater length, Woodside includes an additional fatigue factor, $1 + \gamma T$. Like Keller, Woodside uses race data to determine appropriate values for γ, and achieves remarkably accurate results.

Keller [4,5]

Short article (3 pages w/figures) detailing the theory behind Keller's model. Running adheres to Newton's second law which, in combination with an understanding of a number of physiological quantities (runner's maximum force, resistive force, and rate at which oxygen metabolism creates energy), allows for the creating of a mechanistic model to determine the optimal race pace. Keller break races (those long enough) into three distinct phases: acceleration, maintenance, and deceleration. The first phase of a race involves a runner exhibiting maximum force to reach the optimal velocity as quickly as possible. Once that optimal velocity is reached, the runner should maintain that speed until the runner does not have enough energy to maintain that speed. The final phase, deceleration, occurs once the runner has run out of energy, and "coasts" through the finish. While the deceleration phase is absent from most races, this is most likely due to a sub-optimal race pace, as the runner could have exerted more effort during the maintenance phase. Keller's approach yields impressive results, the error between the Keller's theoretical time and the record was 2.1% for short sprints and 3.1% for longer races.

Keller also poses some issues with his model. He failed to account for vertical (wasted) motion, different efficiencies of energy sources, the accumulation of waste products (i.e. lactic acid), or to distinguish between internal and external resistance. These additions would require further measurements. Keller also proposes adapting his model for use in other sports, as well as attempting to understand the impacts of terrain on changes in the optimal race pace.

Keller builds upon his previous article to solve differential equations for the optimal race velocity at various points throughout the competition. The article follows the same setup from above, and yields the following piecewise

$$v(t) = \begin{cases} F\tau(1 - e^{-t/\tau}), & 0 \leq t \leq t_1 \\ \tau/\lambda, & t_1 \leq t \leq t_2 \\ \sqrt{\sigma\tau + [v^2(t_2) - \sigma\tau]e^{-2(t_2-t)/\tau}}, & t_2 \leq t \leq T \end{cases}$$

CHAPTER

3

LINKING THE MODELS

While both modelling approaches suffer from unique challenges, a combination of the techniques can yield powerful and insightful results. A mechanistic approach lends to a model both a firm grounding in the laws that govern reality, as well as "future proofing" the work by maintaining a generalizability untethered to imperfect data. A statistical approach allows for an informed estimates of constants and links past measurements to future predictions.

Differences in Modelling Approach

Many mechanistic models rely of an athlete's physiological attributes (see chapter 2) to draw conclusions about their fitness or race potential. Statistical models have moved towards using training indices (mileage, duration, training pace, etc.) that are easier to measure and offer greater accessibility to the growing population of amateur athletes.

The broad applicability of those models relying on training indices has led to an increased interest in existing statistical models from the broader community of non-professional runners. However, the susceptibility of statistical approaches to propagate errors in data leaves much to be desired for in terms of the rigor of modelling approach.

Future Work

Building upon the work of Keller and Woodside, among others, we hope to develop a mechanistic model of long-distance race performance utilizing those training indices that have been focuses of statistical modelling for many years. Specifically, we will utilize submaximal VO_2 max tests in conjunction with statistical modelling of the impact of VO_2 max of energy production through the oxidative energy system.

BIBLIOGRAPHY

[1] Veronique Billat, ALEXANDRE DEMARLE, Jean Slawinski, Mario Paiva, and Jean Pierre Koralsztein, *Physical and training characteristics of top-class marathon runners*, Medicine and science in sports and exercise **33** (200112), 2089–97.

[2] Cailbhe Doherty, Alison Keogh, James Davenport, Aonghus Lawlor, Barry Smyth, and Brian Caulfield, *An evaluation of the training determinants of marathon performance: A meta-analysis with meta-regression*, Journal of Science and Medicine in Sport **23** (2020), no. 2, 182–188.

[3] R. Donald Hagan, M G Smith, and Larry R. Gettman, *Marathon performance in relation to maximal aerobic power and training indices.*, Medicine and science in sports and exercise **13 3** (1981), 185–9.

[4] Joseph B. Keller, *A theory of competitive running*, Physics Today **26** (1973), 43–47.

[5] ———, *Optimal velocity in a race*, The American Mathematical Monthly **81** (1974), no. 5, 474–480.

[6] Giovanni Tanda, *Prediction of marathon performance time on the basis of training indices*, Journal of Human Sport and Exercise **Volume 6** (201109), 521–520.

[7] William Woodside, *The optimal strategy for running a race (a mathematical model for world records from 50 m to 275 km)*, Mathematical and Computer Modelling **15** (1991), no. 10, 1–12.