The Quest for Perfection:

Finding the Optimal Marathon Pacing Strategy

**Executive Summary**

A well-defined pacing strategy is essential for achieving peak performance on race day. However, determining the optimal approach remains a significant challenge for many marathon runners. Common advice often advocates for strategies such as negative splits, despite limited research supporting their effectiveness (Fernandes & Maldonado), or the 10/10/10 method, which advises athletes to run the first 10 miles 15–30 seconds per mile slower than their goal pace (Milne). Both strategies hinge on the athlete feeling strong during the final 10 kilometers (6.2 miles) of the race to meet their goal time, an outcome that is far from guaranteed. An analysis of over 90,000 race results revealed that less than 2% of runners successfully executed a negative split. This underscores the complexity of pacing and the need for strategies tailored to individual athletes.

This paper explores the challenges of developing personalized pacing strategies and presents two distinct methods to generate optimized solutions: a weighted average model and a genetic algorithm. Each method leverages race data to produce tailored pacing strategies based on an athlete’s age, gender, and goal finish time. The findings and methodologies discussed in this research contribute to the broader understanding of marathon pacing and offer actionable insights for athletes and coaches seeking to optimize performance.

**Business Understanding**

To thoroughly address the primary research question—*what is the optimal pacing strategy for a marathon runner, given their age, gender, and desired finish time*—a comprehensive review of prior studies and related research was conducted. This review provided key insights into pacing strategy determinants, modeling techniques, and relevant factors influencing marathon performance.

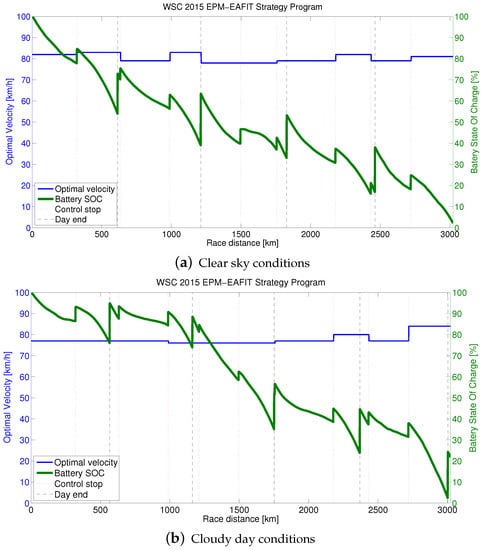
Race strategy optimization techniques are widely used in motorsports, where decisions such as pit stop timing, tire changes, and fuel management are critical to minimizing race time. An intriguing parallel exists between motorsports and marathon pacing strategy, particularly in the context of solar car racing. In solar car competitions, vehicles powered by solar energy must complete courses exceeding 3,000 kilometers in the shortest possible time, requiring careful energy management and pace optimization. This mirrors the challenge faced by marathon runners, who must strategically expend their limited energy to minimize overall race time.

However, there is a critical distinction: whereas solar car racing aims for the absolute fastest time, marathon runners are constrained by personal capabilities and aim to achieve specific goal times. For example, suggesting a pacing strategy for a 2:30:00 marathon to an athlete targeting a sub-3:00:00 finish would be unrealistic and beyond their physical limits. Thus, this research adapts optimization methods, such as genetic algorithms, to focus on realistic and personalized pacing strategies.

Previous studies, such as those by Esteban et al., provide additional insights into pacing strategy optimization. Their findings highlight how conditions impact pacing decisions in solar car races. For instance, under favorable conditions (e.g., clear skies), a car can start faster and gradually deplete its energy over time. In contrast, under poor conditions (e.g., cloud cover), energy conservation is prioritized early in the race, followed by acceleration in the latter stages. This concept is analogous to marathon pacing strategies, where environmental factors such as temperature and humidity play a significant role in determining the most effective pace distribution. Despite these findings, environmental factors were not included in this study due to the unavailability of reliable weather data linked to individual race results. Future work could incorporate such variables to further refine the suggested pacing strategies.

**Figure 1**

Optimal Speed and Energy Depletion based on Sky Condition



When determining which variables to collect, the focus was placed on those shown to have a meaningful impact on pacing strategies. Prior studies have identified *age*, *gender*, and *finish time* as key determinants (Oficial-Casado et al.; Nikolaidis & Knechtle, 2019; March et al.). While some research suggests that age and gender are independent predictors of pacing strategy, others have found significant interactions between the two variables (Nikolaidis & Knechtle, 2018). This discrepancy highlights the importance of investigating these relationships within the current dataset.

To simplify the scope of this research, course elevation changes were excluded, and only flat marathons—such as the Chicago Marathon, Berlin Marathon, and Indianapolis Monumental Marathon—were considered. While elevation changes present unique challenges (e.g., faster muscle breakdown on hilly courses), modeling for these factors is beyond the current scope due to data collection limitation discussed later.

Additionally, psychological and physiological variables, such as motivation, aerobic capacity, and neuromuscular fitness, have been identified as significant influences on pacing strategies (Nikolaidis & Knechtle, 2018). For example, women have been shown to exhibit higher motivation in several areas compared to men, while men tend to display greater competitiveness and risk-taking behaviors. Unfortunately, due to the reliance on historical race results, such data was not available for this study.

**Key Conclusions from Prior Research:**

* Women, older athletes, and faster athletes tend to adopt a more even pacing strategy (March et al.; Nikolaidis & Knechtle, 2018; Oficial-Casado et al.; Deaner et al.).
* Faster starts negatively impact performance due to the early production of blood lactate, which results from the body switching to anaerobic glycolysis in the absence of sufficient oxygen (March et al.).
* Genetic algorithms, as explored by Esteban et al., offer a promising approach to identifying optimal pacing strategies by dividing the race into segments and optimizing pace at each stage.

**Figure 2**

Pacing Profiles Among M 19-24 vs M35-39 for those finishing within 10 minutes of 3:00:00

A graph of different colored lines

Description automatically generated

**Figure 3**

Difference in 1st and 2nd half splits by Gender and Finish Group

A graph of blue and pink bars

Description automatically generated

This foundational understanding informed both the variable selection and modeling approaches used in this research. By focusing on *age*, *gender*, *finish time*, and flat marathon courses, the study aims to produce realistic and actionable pacing strategies while recognizing the limitations of the available data.

***Final Research Question:*** Given an athlete’s age, gender, and desired finish time, what is the optimal pacing strategy to achieve this time?

**Data Understanding/Preparation**

The dataset used in this study comprises race results from three major marathons: the 2022 and 2023 Indianapolis Monumental Marathon, the 2022 Berlin Marathon, and the 2024 Chicago Marathon. These marathons were specifically chosen due to their reputation as flat courses, ensuring that elevation changes—a known factor impacting pacing strategies—could be excluded as a variable in the analysis. While initially the inclusion of course characteristics was considered, the complexities and variations in data collection across different marathons led to a decision to focus solely on flat marathons for consistency and simplicity.

Data collection varied across races, introducing some challenges. Specifically, results for the Berlin and Chicago marathons required scraping split times for individual athletes rather than retrieving results for the entire field at once, making the process more time-consuming.

Once the data was collected, critical preparation steps were undertaken to clean and transform the raw results into a usable format for model development. A set of helper functions was created to automate these tasks. For example, split times recorded in HH:MM:SS format were converted to seconds, and paces listed in MM:SS format were also converted to seconds.

New columns, both continuous and categorical, were derived during this phase:

* **Continuous Columns:** Included metrics such as the time difference between key splits (e.g., start to half, half to finish) and time intervals between successive splits (e.g., time between the 5k and 10k split).
* **Categorical Columns:** Key variables such as *gender*, *age group*, and *finish category* were created. As highlighted in the business understanding section, prior research demonstrates that age and gender are critical determinants of pacing strategy. These variables were therefore stored separately to enable analysis of potential interactions between them.

The *finish category* variable was computed to classify pacing strategies into one of three categories:

1. **Negative Split:** If the percent difference between the first and second half relative to the total time was greater than or equal to +2.5%.
2. **Even Split:** If the percent difference fell within -2.5% and +2.5%.
3. **Positive Split:** If the percent difference was less than or equal to -2.5%.

The computation process involved first determining the time taken to complete the second half of the race and then calculating the percentage difference relative to the total finish time. This classification allows for further exploration of pacing patterns and their relationship with finish times.

Data Cleaning Steps:

* Removed total of 102 elite runners from 2022 and 2023 Indianapolis Monumental Marathon to focus on non-elite runners.
* Removed 31 runners from the 2022 and 2023 Indianapolis Monumental Marathon results due to missing gender or age group.
* Removed 7 runners from the 2022 Berlin Marathon results that didn’t specify gender.
* Removed 65 runners from the 2024 Chicago Marathon results due to missing gender or age group.
* Removed any runner with one or more missing split times to ensure all pacing data was complete and reliable

These cleaning steps ensured the final dataset was both robust and consistent, enabling reliable modeling and analysis. The resulting data allows for exploration of pacing strategies relative to demographic characteristics (e.g., age and gender) and performance metrics (e.g., finish group and split times).

***Variable Table***

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Definition** | **Relationship to RQ** | **Source** |
| race | Name of race result is from | Used to build athlete\_id | Race results |
| year | Year of race result | Used to build athlete\_id | Race results |
| bib | Bib number of athlete | Used to build athlete\_id | Race results |
| athlete\_id | Unique ID for athlete (race\_year\_bib) | Used to ensure data quality and no replication of data. Allows data to be anonymous. | Computed from race results |
| gender | Gender of athlete | Key determinant of pacing strategy | Race results |
| age\_group | Age group/division of athlete | Key determinant of pacing strategy | Race results |
| finish\_group | Category based on official time at finish | Key determinant of pacing strategy | Computed from race results |
| second\_half\_split\_seconds | Duration of the second half of the race in seconds | Used to categorize pacing strategy used by athlete | Computed from race results |
| %\_difference\_second\_half | Percent faster or slower during the second half relative to finish time | Used to categorize pacing strategy used by athlete | Computed from race results |
| pacing\_strategy | Categorized pacing strategy used by athlete | Used to create more detailed strategies in weighted average model | Computed from race results |
| Time (hh:mm:ss)  5k, 10k, 15k, 20k, half, 25k, 30k, 35k, 40k, finish  ex) 5k\_time | Official time at each distance | Used to compute all other time and based variables | Race results |
| Time (seconds)  5k, 10k, 15k, 20k, half, 25k, 30k, 35k, 40k, finish  ex) 5k\_seconds | Official time converted to seconds at each distance | Used to compute split times and half\_split\_difference | Computed from race results |
| Pace (mm:ss)  5k, 10k, 15k, 20k, half, 25k, 30k, 35k, 40k, finish  ex) 5k\_pace | Official pace at each distance | Used to compute pace\_seconds | Race results |
| Split Time (seconds)  10 splits\*  ex) split\_1\_seconds | Computed time between distances converted to seconds | One option for response variable | Computed from race results |

\* Split 1 = (start – 5k), Split 2 = (5k – 10k), Split 3 = (10k – 15k), Split 4 = (15k – 20k),

Split 5 = (20k – Half), Split 6 = (Half – 25k), Split 7 = (25k – 30k), Split 8 = (30k – 35k),

Split 9 = (35k – 40k), Split 10 = (40k – Finish)

**Modeling and Evaluation**

Different fields use various techniques to solve these kinds of optimization problems based on the data available. The models that made the most sense to try in this case were a simple weighted average model, which takes heavily weighs records similar to the user, and a genetic algorithm, which looks at each split time as a particular “gene” of the overall race. A genetic algorithm was implemented after research by Betancur et al., which successfully implemented a genetic algorithm to solve the optimal speed of a solar car during different sections.

***Weighted Average Model***

**Methodology and justification.** The weighted average model serves as a straightforward and foundational method for determining an optimal pacing strategy for marathoners, tailored to their age, gender, and desired finish time. Upon receiving the input parameters, the model segments the data into three groups based on similarity to the user. The first group includes athletes from the same age group, gender, and finish group. The second group consists of athletes from neighboring age groups, of the same gender, and within the same finish group. Finally, the third group includes athletes across all age groups and genders, but within neighboring finish groups.

To calculate the pacing strategy, a weighted linear combination is applied as follows:

This formula assigns the highest weight to the group most similar to the user, while progressively reducing the weight for less similar groups. This approach ensures that the model prioritizes data from athletes with comparable characteristics, while reducing the influence of outliers and broader trends that may not be relevant to the user.

**Figure 4**

Weighted Average Model General Suggestion Output

A white background with blue lines

Description automatically generated

An additional feature of the weighted average model is its ability to provide users with optimal pacing strategies across three commonly used techniques: positive split, even split, and negative split. To achieve this, the data is first segmented by pacing strategy before being divided into the aforementioned three similarity groups. This results in a total of nine subgroups (e.g., Positive Split Group 1, Positive Split Group 2, etc.), and a separate pacing strategy is calculated for each technique. This flexibility allows experienced users to select a strategy that aligns with their personal preferences or training plans.

**Figure 5**

Weighted Average Model Specific Pace Strategy Output

A black and red line

Description automatically generated

**Evaluation.** The weighted average model is a prescriptive model, and as such, traditional evaluation metrics are not directly applicable. Evaluating its effectiveness requires knowledge of the user’s intent or goal time, which was not available in the dataset. For instance, an athlete finishing in 3:01:00 may have either narrowly missed a sub-3-hour goal or significantly exceeded a goal of 3:05:00. Without this context, objective evaluation is challenging.

To address this limitation, a personal example was used to assess the model’s recommendations. During the 2024 Indianapolis Monumental Marathon, the model was applied to generate an even-split pacing strategy for a 23-year-old male with a target finish time under 2:55:00. The model’s suggestion resulted in a predicted finish time of 2:52:47, which matched the actual finish time. While some deviations occurred between 25 km and 35 km, where the user exceeded the suggested pace, these differences did not adversely impact the final outcome. This alignment between the suggested and actual pacing strategies demonstrates the model’s potential to provide effective and realistic recommendations.

**Figure 6**

Weighted Average Even Split Suggestion vs Personal Result

A diagram of a computer

Description automatically generated

**Limitations.** The primary limitation of the weighted average model lies in the inherent constraints of averaging and the Law of Large Numbers. As demonstrated in the age group breakdown (see Table 1), the distribution of data across age groups is uneven. For example, larger age groups (e.g., 16–29) have more data points, resulting in greater stability in their suggested pacing strategies. Conversely, smaller age groups (e.g., 65–69 or 70+) may lack sufficient data, reducing the reliability of the recommendations. To mitigate this issue, a threshold was established to ensure only groups with adequate data points produce results.

Table 1

Age Group Breakdown

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **age group** | **16-29** | **30-34** | **35-39** | **40-44** | **45-49** | **50-54** | **55-59** | **60-64** | **65-69** | **70+** |
| **count** | 16,365 | 14,456 | 14,071 | 14,579 | 11,951 | 10,011 | 6,147 | 3,460 | 1,258 | 542 |

***Genetic Algorithm***

**Methodology and justification.** The genetic algorithm was implemented as an advanced optimization technique to determine an athlete's optimal pacing strategy, inspired by the findings of Bentancur et al. The methodology begins by filtering the dataset to include only relevant records—specifically, athletes from the user’s desired finish group or neighboring finish groups, as well as individuals in neighboring age groups and of the same gender. This filtering ensures that the population used for optimization consists of pacing behaviors that are contextually similar to the user’s goal.

The process starts by randomly initializing a population of 250 potential pacing strategies derived from the filtered dataset. Each pacing strategy is evaluated based on a defined fitness function to assess its quality. Fitness penalties are applied if a strategy produces a total time outside of the user's target range or if individual splits deviate by more than 3% from the previous split, ensuring the strategies remain physiologically realistic and consistent.

The algorithm proceeds iteratively, evolving the population toward an optimal solution. In each generation, parent strategies are selected using a combination of elitism and tournament selection. The two best-performing strategies are automatically carried forward (elitism), while additional parents are probabilistically selected through a tournament process of size five. Selected parents undergo a crossover operation, where components of their pacing strategies are combined to generate new offspring. A crossover rate of 0.5 determines the probability of this recombination occurring. To maintain diversity and prevent the algorithm from converging too quickly on a suboptimal solution, a mutation operation is applied with a mutation rate of 0.05, introducing small random changes to the new pacing strategies. After crossover and mutation, the fitness of the updated population is recalculated, and the best-performing pacing strategy from the final generation is selected as the optimal solution. This iterative evolutionary process allows the genetic algorithm to dynamically search for pacing strategies that balance consistency and efficiency, aligning with the athlete’s goal time.

**Evaluation.** As a prescriptive model, the genetic algorithm cannot be evaluated using traditional performance metrics. To assess its practical effectiveness, a personal case study was conducted. The model was tasked with generating a pacing strategy for a 23-year-old male aiming for a sub-2:55:00 finish time. The genetic algorithm produced a pacing plan, which recommended a slight reduction in pace between 25km and 35km to manage fatigue when compared to my personal race performance. Overall, this pacing strategy aligned closely with real-world performance, where small deviations from 25km to 35km did not prevent achieving the target finish time of 2:52:47.

**Figure 7**

Genetic Algorithm Suggest vs Personal Result

A diagram of a computer

Description automatically generated with medium confidence

**Limitations.** Despite its strengths, the genetic algorithm has notable limitations. The algorithm's performance is dependent on the quality and size of the dataset used for initialization. Limited data for specific age groups, genders, or finish times may reduce population diversity, impacting the algorithm’s ability to converge on an optimal solution. Nonetheless, the genetic algorithm provides a robust and dynamic framework for generating personalized pacing strategies, offering a significant improvement over simpler, static models.

**Solution Deployment**

To ensure the practical application of this study's findings, a publicly accessible Streamlit web application has been developed. This application serves as a tool for marathoners, coaches, and other athletes seeking to identify the optimal pacing strategy to achieve a target marathon time.

The web application allows users to input key demographic and performance variables, including age group, gender, and desired finish group. Based on this input, the application generates personalized pacing strategies in two user-friendly formats. The first output consists of pacing charts that visually represent the suggested split times over the course of the marathon. The second output is a detailed table that includes both individual split times and cumulative time at each checkpoint. The inclusion of cumulative times is particularly valuable for runners who rely on on-course timing clocks, as it allows them to monitor progress without needing specialized configurations on personal devices, such as GPS watches.

Another feature of the application is its flexibility in model selection. Users can choose between the *weighted average model* and the *genetic algorithm* to generate their optimal pacing strategy. Both models demonstrated the ability to produce effective strategies, as discussed in earlier sections, ensuring that users receive reliable and well-calibrated recommendations.

**Discussion**

The results of this study, derived from both models and supported by personal experience, suggest that the most effective pacing strategy for marathon performance is a *controlled positive split*. This strategy involves starting slightly slower than the target average pace, progressing to a faster pace over the middle 20 kilometers, and accounting for the natural onset of fatigue around the 30-kilometer mark. At this point in the race, the majority of marathoners experience a decline in pace; however, by banking time earlier in a controlled manner, this strategy mitigates the risk of missing the target finish time.

The controlled positive split promotes physiological efficiency by allowing the body to gradually increase pace, minimizing excessive energy expenditure that could result from pacing inconsistencies. While this strategy could technically be classified as an "even split" under conventional pacing terminology, the slight deceleration in the latter half of the race—remaining within 2.5% of the first half—distinguishes it as a controlled positive split. For example, an athlete running a 1:26:01 first half and a 1:26:47 second half demonstrates the controlled positive split approach with near-optimal execution.

Both the genetic algorithm and the weighted average model produced pacing strategies that align closely with this description. Although classified as "even pacing" from a high-level perspective, the models' outputs consistently revealed a slight decline in pace after the 30-kilometer mark, reinforcing the viability of the controlled positive split as a robust and practical pacing strategy. This pacing approach addresses the balance between performance optimization and the physiological realities of endurance running, offering a realistic framework for achieving goal times.

**Limitations**

The biggest limitation of the project was not knowing each athlete’s goal or intention. For example, if a runner positively split the race. Did they still hit their goal time? Was this their plan going into the race? Not knowing this information makes it difficult to properly evaluate each model.

A second limitation was the lack of access to psychological and physiological data, which have been shown to significantly impact pacing strategies. Psychological factors such as motivation, risk tolerance, and competitive mindset, as well as physiological variables like neuromuscular fitness, aerobic capacity, and body composition, play a crucial role in determining an athlete’s performance and pacing approach. Incorporating these variables could provide deeper insights and enhance the accuracy of the suggested pacing strategies; however, this data was unavailable due to the nature of publicly accessible race results.

Finally, the project was constrained by a limited dataset. The ideal solution would allow athletes to select a course type or even a specific course. However, the data collection process for each individual race was slightly different. Due to these differences, there was not enough data to include a wide variety of courses or model for elevation changes. This limits the generalizability of the results to primarily flat marathon courses, such as the Chicago, Berlin, and Indianapolis Monumental marathons.

**Learning Implications**

During this process I was able to implement a lot of new ideas and techniques that allowed me to solve this problem. The key ideas used in this process surrounded the modeling phase. Prior to this project I learned about the idea behind genetic algorithms, but had never come across a use case that fits what this algorithm could offer. Learning to implement this kind of model played a key role in determining what the true best strategy was.

Overall, the techniques that were used throughout this project are a combination of skills gained throughout my academic career. However, there were three specific skills that I used, which were learned during my graduate coursework. The three courses and skills used from those courses are listed below:

* Cloud Computing
  + During the data collection phase I had to collect data person by person. For a race like the Chicago Marathon this means looking individually at over 50,000 different sites. To scrape this data in a timely manner I used pySpark. This allowed my computer to parallelize this process and made the data collection much faster. Without implementing pySpark I would not have been able to collect data from as many races as I was able, which would have been a severe limitation on the success of the models.
* Intro to Informatics
  + The final solution deployment was key to the overall success of the project. This allows others to interact with my project and find their optimal pacing strategies. Using a public Streamlit app I was able to make this possible. It was during Intro to Informatics that I first learned how to accomplish this and make an application publicly accessible. Without the Streamlit application being publicly accessible the project has little value to anyone besides myself.
* Data Visualization
  + Visualizing information was a key step throughout my project. I used visualizations to explore data, determine if my data followed past research, and finally output information in a readable format. During a past Data Visualization course, I learned a large variety of techniques and visualization methods to effectively communicate information. These techniques allowed me to answer key questions and communicate my findings.

**Conclusion**

This project provides a comprehensive framework for determining the optimal pacing strategy for marathon runners, grounded in both robust data analysis and innovative modeling techniques. By utilizing race results from three well-known flat marathons and implementing two distinct methods—a weighted average model and a genetic algorithm—the study delivers practical solutions that are both data-driven and adaptable to athletes’ personal goals.

The weighted average model offers a simple yet effective approach, leveraging similarities between athletes to prescribe pacing strategies while accounting for individual preferences like positive, negative, or even splits. On the other hand, the genetic algorithm introduces a more dynamic and adaptable optimization process. Both models align with prior research, affirming that women, older runners, and faster athletes generally favor even pacing strategies, while a controlled positive split emerges as an efficient and realistic approach for achieving optimal results.

While the absence of physiological and psychological data introduces limitations, the insights gained from this study provide actionable value to athletes, coaches, and the broader running community. The publicly accessible Streamlit web application ensures these findings are practical and user-friendly, enabling individuals to tailor pacing strategies to their age, gender, and target finish time.

Ultimately, this work bridges data science and sports performance, illustrating how strategic pacing can help athletes unlock their potential on race day. Future work could explore course-specific strategies and incorporate additional performance metrics to further enhance the models’ accuracy and applicability. By combining innovative algorithms, real-world data, and user-focused solutions, this study contributes meaningful advancements to the science of marathon pacing.

**References**

Betancur, E.; Osorio-Gómez, G.; Rivera, J.C. Heuristic Optimization for the Energy Management and Race Strategy of a Solar Car. *Sustainability* 2017, *9*, 1576. https://doi.org/10.3390/su9101576

Deaner, R. O., Carter, R. E., Joyner, M. J., & Hunter, S. K. (2015). Men Are More Likely than Women to Slow in the Marathon. *MEDICINE AND SCIENCE IN SPORTS AND EXERCISE*, *47*(3), 607–616. https://doi-org.proxy.ulib.uits.iu.edu/10.1249/MSS.0000000000000432

Fernandes, G. , & Maldonado, V. (2024). Against Negative Splitting: The Case for a Counterintuitive Pacing Strategy for Elite Marathon Athletes in Official Events. *American Journal of Sports Science and Medicine*, *12*(1), 1-6.

Jungong Sha, Qing Yi, Xin Jiang, Zhengwei Wang, Houwen Cao, & Shan Jiang. (2024). Pacing strategies in marathons: A systematic review. *Heliyon*, *10*(17). https://doi-org.proxy.ulib.uits.iu.edu/10.1016/j.heliyon.2024.e36760

March, Daniel S; Vanderburgh, Paul M; Titlebaum, Peter J; Hoops, Mackenzie L. Age, Sex, and Finish Time as Determinants of Pacing in the Marathon. Journal of Strength and Conditioning Research 25(2):p 386-391, February 2011. | DOI: 10.1519/JSC.0b013e3181bffd0f

Milne, Keeley. “Try the 10/10/10 Marathon Method for Perfect Pacing - Canadian Running Magazine.” *Canadian Running Magazine*, 12 Apr. 2024, runningmagazine.ca/sections/training/try-the-10-10-10-marathon-method-for-perfect-pacing/. Accessed 18 Dec. 2024.

Nikolaidis PT and Knechtle B (2018) Pacing Strategies in the ‘Athens Classic Marathon’: Physiological and Psychological Aspects. *Front. Physiol.* 9:1539. doi: 10.3389/fphys.2018.01539

Nikolaidis, Pantelis Theodoros; Knechtle, Beat (2019). Do Fast Older Runners Pace Differently from Fast Younger Runners in the “New York City Marathon”? Journal of Strength and Conditioning Research, 33(12):3423-3430.

Oficial-Casado, F., Uriel, J., Pérez-Soriano, P., & Priego Quesada, J. I. (2021). Effect of marathon characteristics and runners’ time category on pacing profile. *European Journal of Sport Science*, *21*(11), 1559–1566.