**The Effect of Weather Parameters on Triathlon Race Performance: Ironman 70.3 and Full Ironman Analysis**

**Author: Elabbassi Mohamed Ali**

**Sultan Moulay Slimane University/ENSAK**

**IADD Master 2024/2026**

**Abstract**

Understanding the impact of environmental conditions on triathlon performance is crucial for athletes, whether they are professionals or recreational, competing for performance or leisure. This understanding can significantly impact race outcomes and overall performance, enabling participants and coaches to plan their training and event preparation more effectively.

This study aims to assess the relationship between weather parameters and triathlon performance, focusing on how temperature, humidity, and other environmental and race location factors influence results across different competitive levels.

The project focuses on the Ironman 70.3 and full Ironman races, which are some of the toughest sports competitions. Results for 931 Ironman race with 1480230 participants were collected in addition to the corresponding weather data. Both statistical analyses (regression and correlation) and machine learning models (decision trees, random forests, XGBoost, and quantile regression) were applied. The project focuses on the Ironman 70.3 and full Ironman races, which are some of the toughest sports competitions. Results for 931 Ironman race with 1480230 participants were collected in addition to the corresponding weather data. Both statistical analyses (regression and correlation) and machine learning models (decision trees, random forests, XGBoost, and quantile regression) were applied. Preliminary results suggest significant correlations between water temperature and swim performance with optimal ranges between 20-24°C, while higher temperatures and elevations differentially impact race segments across athlete divisions. XGBoost models demonstrated superior predictive power (R² = 0.91 for swim, 0.61 overall) compared to other approaches, and quantile regression revealed that environmental factors affect elite athletes differently than recreational participants. The findings will inform race scheduling, athlete preparation strategies, course selection based on individual strengths, and training adaptations for specific environmental conditions, ultimately enhancing performance outcomes and athlete safety across varying competitive levels.

Keywords: Ironman, triathlon, weather parameters, performance, machine learning

**Table of Contents**

I-Introduction ..........................................................................................................................

II-Literature Review ...............................................................................................................

III-Methodologies ……………………………………………………………………………………

3-1 Races data ……………………… ........................................................................

3-2 Races characteristics..........................................................................................

3-3 Weather data ......................................................................................................

3-4 Data processing workflow…………………………………………………………….

IV- Data Analysis .................................................................................................... …………

4-1 Statistical analysis………………………………………………………………………

4-2 Linear regression ……………………………………………………………………..

4-3 Decision trees …………………………………………………………………………….

4-4 XGBoost …………………………………………………………………………………

4-5 Quantile regression ……………………………………………………………………

V-Discussion ............................................................................................................... ………

VI-Conclusion .............................................................................................................. ……..

References ..............................................................................................................

Appendices ............................................................................................................

**I-Introduction:**

Triathlon is a demanding multi-sport endurance event that combines swimming, cycling, and running, completed in immediate succession. It challenges athletes' physical and mental endurance across various race formats. Among the most prestigious are the Ironman races, known for their extreme distances and popularity worldwide. The **full Ironman** consists of a 2.4-mile (3.8 km) swim, a 112-mile (180 km) bike ride, and a 26.2-mile (42.2 km) marathon run. The **Ironman 70.3**, also known as the **Half Ironman**, covers half those distances: a 1.2-mile (1.9 km) swim, a 56-mile (90 km) bike, and a 13.1-mile (21.1 km) run. Both formats require months of intense training, strategic preparation, and exceptional endurance, making them milestones in any triathlete's journey.

Endurance events are sensitive to environmental conditions, which can significantly affect athlete performance. Weather critically impacts endurance performance because the body must balance intense metabolic heat production with effective thermoregulation. High temperatures and humidity reduce sweat evaporation, causing overheating, accelerating fatigue, and increasing the risk of heat illness. Conversely, cold conditions increase energy expenditure for warmth and can impair muscle function. Water temperature similarly affects swimmers: cold water risks hypothermia and reduces muscle efficiency, while overly warm water prevents necessary cooling. Ultimately, these factors force the body to divert energy away from propulsion toward temperature management, directly slowing pace, increasing perceived effort, and elevating safety risks.

Previous research on running and cycling indicates that temperature and humidity have optimal ranges for peak performance, with performance decreasing at both high and low extremes. In triathlon contexts, water temperature also plays a crucial role during the swim segment, potentially affecting athlete safety and speed.

Although several studies examine either Ironman 70.3 or full Ironman races independently, a unified analysis comparing both distances using a consistent data processing pipeline and advanced predictive modeling is lacking. Moreover, few studies leverage machine learning techniques alongside traditional statistical methods in this context.

The primary purpose of this study is to quantify the effects of weather parameters, altitude, and elevation gain on finish times and split-discipline performance for both Ironman 70.3 and full Ironman athletes. The study addresses the following research questions:

1. How do air temperature, humidity, wind speed, water temperature, and altitude correlate with overall finish times?

2. Is the impact of these parameters similar for age groups and professionals?

2. Are there differential effects of weather on swim, bike, and run segments?

3. How does model performance (statistical vs. machine learning) compare in predicting race outcomes?

**II- Literature Review:**

**Environmental Effects on Endurance Sports**

Environmental conditions, particularly air temperature and wind, have a significant impact on marathon and cycling performance. Higher temperatures consistently lead to slower finishing times and increased withdrawal rates in marathons, with optimal performance occurring at cooler temperatures; humidity, wind speed, and solar radiation also play roles, but their effects are often secondary to temperature and sometimes interrelated with it (Helou et al., 2012; Vihma, 2010; Tan et al., 2022; Gasparetto & Nesseler, 2020; Wang et al., 2024; Nikolaidis et al., 2019; Weiss et al., 2024; Knechtle et al., 2021). In cycling and long-distance events, microclimatic variations and elevated particulate matter (PM2.5) levels can pose additional risks, highlighting the need for direct environmental monitoring along race routes (Havenga et al., 2024). Overall, both elite and amateur athletes are affected by adverse weather conditions, but slower and less experienced participants tend to be more vulnerable to heat- and pollution-related performance declines (Vihma, 2010; Gasparetto & Nesseler, 2020; Hodgson et al., 2022; Knechtle et al., 2021).

**Prior Studies on Half and Full Ironman Weather Impacts**

Ironman triathlon performance is strongly influenced by environmental conditions, especially temperature and air quality. Analyses of large datasets from professional and age-group races indicate that optimal weather conditions, specifically moderate air and water temperatures, result in the fastest race times, while heat, humidity, and air pollution can significantly impair performance.

Research on Ironman 70.3 events indicates that water temperatures above 22°C can lead to faster swim splits, while bike performance peaks around air temperatures of 20–25°C. Full Ironman studies show similar trends but also highlight that wind conditions disproportionately affect the bike segment due to longer distances.

**Temperature:** The fastest Ironman race times are achieved with water temperatures above 22°C and air temperatures between 19–27°C for professionals, and around 24°C (water) and 27°C (air) for age-group athletes. Races held in these conditions, such as Ironman Brazil Florianopolis and Ironman Barcelona, consistently yield better results (Knechtle et al., 2025; Knechtle et al., 2024).

**Heat and Humidity:** High temperatures (e.g., 29°C dry, 27°C humid) increase core body temperature and fluid loss. Faster athletes tend to tolerate greater body mass loss and higher core temperatures, while slower athletes show more muscle damage and reduced muscle performance (Del Coso et al., 2014; Baillot & Hue, 2015).

**Performance Decline in Heat:** For every degree above optimal wet bulb globe temperature (WBGT), endurance performance drops by 0.3–0.4%. All four weather parameters—temperature, humidity, wind, and solar load—should be considered for athlete safety and performance (Mantzios et al., 2021).

**Altitude/Elevation** affects performance primarily through reduced partial pressure of oxygen, leading to quicker fatigue in aerobic activities, though lower air density at higher elevations can reduce aerodynamic drag, potentially benefiting cycling.

Elevation parameters significantly influence athletic performance across multiple dimensions:

- Oxygen Availability: Every 1,000 feet of elevation decreases oxygen availability by approximately 3%, affecting aerobic performance [Bassett & Howley, 2000]

- Hematological Adaptations: Venues above 5,000 feet trigger EPO production and increased red blood cell concentration [Burtscher et al., 2011]

- Respiratory Compensation: Higher ventilation rates are required, increasing respiratory muscle fatigue [Sawka et al., 2012]

- Thermoregulation: Lower air density at altitude affects evaporative cooling efficiency [Sawka et al., 2012]

- Power output typically decreases 7-10% for unacclimated athletes at moderate altitudes (5,000-8,000 feet) [Burtscher et al., 2011]

- Perceived exertion increases approximately 6-8% for equivalent workloads [Renfree et al., 2014]

- Pacing strategy adjustments become crucial to prevent early glycogen depletion [Abbiss & Laursen, 2008]

- Recovery between efforts is impaired, affecting interval-based race strategies [Mujika, 2014]

**water type:** In the swimming segment, it is influential, with river water generally enhancing race times compared to still water, while seawater tends to slow athletes down.

|  |  |  |  |
| --- | --- | --- | --- |
| **Weather Factor** | **Optimal Range for Fastest Times** | **Performance Impact (if outside range)** | **Citations** |
| *Air Temperature* | 19–27°C (pro), ~27°C (age) | more fatigue | (Knechtle et al., 2025; Mantzios et al., 2021; Knechtle et al., 2024) |
| *Water Temperature* | >22°C (pro), ~24°C (age) | Slower swim, higher stress | (Knechtle et al., 2025; Knechtle et al., 2024) |
| *Humidity* | Moderate | High humidity increases strain | (Baillot & Hue, 2015; Mantzios et al., 2021) |
| *Air Pollution* | (O3, PM2.5) | Low Slower swim (O3), slower bike/run (PM2.5) | (Naidenova et al., 2022) |

**Methodological Approaches in Performance Prediction**

The analysis of performance prediction in endurance sports related to weather conditions uses a variety of advanced methodological methods, mainly focusing on statistical modeling and machine learning algorithms. Studies often use regression analysis, including linear, non-linear (e.g., second-order polynomial), and mixed-effects models, sometimes adding polynomial splines for specific environmental factors like water temperature on swim times (Gibson, 2024). Quantile Regression (QR) is also used to understand how environmental variables such as temperature, elevation, and water type influence race times across different athlete performance levels (Zhao et al., 2024). Besides traditional statistics, machine learning (ML) algorithms are widely applied, including Decision Tree Regressor, Random Forest Regressor, and XG Boost Regressor to forecast overall race times, with features often including event location, water, and air temperature (Thuany et al., 2023; Knechtle et al., 2025; Mantzios et al., 2022). For more complex temporal data, models like Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Gate Recursive Unit (GRU) are used to predict running performance and physiological indicators based on environmental inputs like temperature (Yang et al., 2021). Environmental data such as air temperature, water temperature, relative humidity, wind speed, and solar radiation are typically collected from official race sources or weather stations, with derived metrics like Wet Bulb Globe Temperature (WBGT) and heat index calculated to evaluate overall heat stress (Mantzios et al., 2022; Baillot & Hue, 2015). Model interpretability tools like SHAP (Shapley Additive Explanations) values and Partial Dependence Plots (PDP) are often employed to explain how various environmental features influence predictions (Thuany et al., 2023; Knechtle et al., 2025).

**Categorical approach:**

Recent studies use quantile regression to compare how weather and environmental factors affect triathlon performance across athlete categories such as professionals, recreational athletes, and age groups. For the top 50% of elite (professional) athletes, swim time is the most significant predictor of overall performance, while for slower elite athletes, bike time becomes more influential. Temperature and altitude positively impact race times for all groups, but the degree of influence varies by performance level. River water tends to enhance, and sea water tends to slow, race times across categories. These findings highlight that the effect of weather and environmental conditions is not uniform; rather, it depends on the athlete’s category and performance level. Consequently, training and race strategies should be tailored to the specific needs and capabilities of professionals, recreational athletes, and age-group competitors, with a strong emphasis on adapting to anticipated weather and course conditions (Zhao et al., 2024).

Studies approach weather effects on triathlon performance by categorising athletes, revealing distinct impacts and optimal conditions for each group (Gibson, 2024; Hermand et al., 2019; Knechtle et al., 2025; Nikolaidis et al., 2023).

* **Professional/Elite Triathletes**:
  + For **full Ironman distances**, optimal performance occurs in **water temperatures warmer than 22°C and air temperatures between 19–26°C** (Knechtle et al., 2025). Racecourse characteristics (e.g., river swim, flat bike) also significantly influence professional performance (Knechtle et al., 2025). **Wind speed is a significant predictor for swimming and cycling** (Knechtle et al., 2020). Professionals may experience hyperthermia in extreme conditions (e.g., Kona marathon) (Knechtle et al., 2020; Knechtle et al., 2025).
  + In **elite standard distance triathlons**, **water temperature affects swimming** and **air temperature affects running**, with fastest running times observed around 18°C (Gibson, 2024; Mantzios et al., 2022). Air temperature has less impact on cycling due to high speeds and drafting (Gibson, 2024). Men and women are generally affected similarly by temperatures, though cold water can still pose a risk of injury or death for elites (Gibson, 2024).
* **Age Group/Recreational Triathletes**:
  + For **Ironman 70.3 age groups**, **cycling and running are more predictive of overall race performance than swimming** (Nikolaidis et al., 2023). The age-related performance decline begins earliest in swimming for these athletes (Nikolaidis et al., 2023).
  + More broadly, for **endurance runners**, there is a **progressive slowing as Wet Bulb Globe Temperature (WBGT) increases from 5°C to 25°C**, with slower runners experiencing larger decrements (Baillot & Hue, 2015; Hermand et al., 2019; Mantzios et al., 2022). High temperatures (>15°C) and increased precipitation impair their marathon performance (Knechtle et al., 2025). While some well-trained age groupers can cope with tropical conditions without excessive core temperature increases, the impact of environmental factors is generally more pronounced for this cohort (Baillot & Hue, 2015; Hermand et al., 2019).

**Identified Gaps and Theoretical Justification:**

Research into the impact of weather on triathlons and Ironman races has highlighted several areas requiring further investigation:

Identified Gaps:

•**Limited Scope of Athlete Categories:** Most studies focus on professional athletes, with less attention given to recreational or age-group participants, which limits the generalizability of findings across the broader triathlon community (Knechtle et al., 2025).

•**Insufficient exploration of environmental characteristics beyond common variables and Lack of detailed environmental data**: Current triathlon research suffers from a lack of detailed, granular environmental data (like humidity fluctuations, wind specifics, and WBGT) during races and fails to sufficiently explore crucial environmental characteristics beyond common variables—such as altitude, terrain specifics, water temperature, and sea currents—limiting the understanding of location's impact on performance (Gibson, 2024; Knechtle et al., 2020, 2025; Nikolaidis et al., 2023; Zhao et al., 2024; A Machine Learning Approach..., 2023).

•**Optimal Weather Conditions:** A comprehensive understanding of the precise optimal weather conditions for peak performance remains limited, particularly for non-professional athletes (Knechtle et al., 2025).

•**Limited understanding of race course characteristics' specific impact**: Current research on triathlons is limited by its focus on popular races and regions, neglecting less-studied locations and diverse environments (Knechtle et al., 2025), as well as by a lack of detailed analysis on how specific racecourse characteristics—including changes over time—affect performance across different segments (A Machine Learning Approach to Finding the Fastest Race, 2023; Gibson, 2024).

•**Need for additional variables in predictive models**: Existing machine learning models often yield low predictive power (low R2 scores), suggesting that additional variables, beyond environmental factors, are necessary to accurately predict Ironman triathlete performance (Knechtle et al., 2025).

•**WBGT model limitations**: Despite its utility, the WBGT index has limitations, particularly in its differing response to wind speed compared to actual human thermoregulation (Kong & Huber, 2024). The complex and iterative nature of WBGT calculation often renders it a "black box," hindering deeper theoretical investigation into the atmospheric dynamics and thermodynamic processes that control heat stress (Kong & Huber, 2024).

**Theoretical Justification:**

Multifactorial Impact: Theoretical frameworks recognize that performance is influenced by a complex interplay of environmental, physiological, and demographic factors. This justifies the use of machine learning and multivariate models to capture these interactions (Knechtle et al., 2025).

Health and Safety: The need to mitigate health risks (e.g., heat stress) provides a theoretical basis for studying multiple weather and race location parameters, not just temperature, to inform athlete preparation and event management (Mantzios et al., 2021).

Performance Optimization: Understanding how weather affects different athlete groups supports the development of tailored strategies for training, pacing, and heat adaptation, grounded in sports science and environmental physiology (Knechtle et al., 2025; Mantzios et al., 2021).

In summary, while recent studies have improved modeling and prediction, more inclusive and comprehensive research is needed to fully understand and address the effects of weather on all triathlon participants.

**III Methodologies:**

**3-1 Races Data:**

The race data was collected from the website "www.endurance-data.com" for events held between 2015 and 2023. Finish times, split times (swim, bike, and run), and overall placements for both age-group and professional athletes were recorded. Athlete demographic information (age, gender, and category) was also extracted, and personal information (names, bib numbers, and places) was removed to anonymize the data. Finally, results were obtained from 1480230 athletes participating in 931 races.

Note: Ethical approval was not required for this study, as all data were publicly available for download and examination.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Locations | Races | Athletes | Ironman | Ironman athletes | Ironman 70.3 | Ironman 70.3 athletes |
| 337 | 931 | 1480230 | 318 | 548002 | 613 | 932228 |

**A map of the world

AI-generated content may be incorrect.**

**A map of the world

AI-generated content may be incorrect.**

**3-2 Race characteristics:**

This paper addresses a significant methodological gap in triathlon performance analysis: the absence of a comprehensive study of the effect of race characteristics like venue altitude and elevation gain combined with weather on triathlon races.

The elevation profile of an Ironman course represents one of the most significant fixed variables affecting race outcomes. Unlike weather conditions that vary year to year, the fundamental elevation characteristics remain relatively constant, providing a stable basis for comparative analysis and performance prediction.

The elevation data for Ironman triathlon venues was collected using a multi-source web scraping methodology that prioritizes accuracy through cross-validation. The primary parameters collected include:

* Venue altitude (elevation above sea level).
* Bike course elevation gain (total climbing in feet).
* Run course elevation gain (total climbing in feet).

The script employs a hierarchical three-tier approach with source prioritization:

Tier 1: TriathlonCourseInfo.com (Provides comprehensive course descriptions for global triathlon events).

Tier 2: PJammCycling.com (Specializes in detailed cycling course analysis).

Tier 3: Ironman.com Official Website (Contains official race information directly from the organizers).

**3-3 Weather Data:**

All weather data was retrieved from the Open-Meteo API, which provides standardized historical weather measurements. The data is collected for all locations and dates based on geographic coordinates (latitude and longitude).

We have chosen to extract the weather parameters that seem to have the most impact on the race's results.

Weather parameters collected and their significance:

**1. Temperature:** measured in degrees Celsius. Understanding heat stress during the event is crucial for effective planning of hydration and cooling strategies.

**\* Maximum Temperature:** The highest temperature recorded during the day, usually registered around two hours after noon, and Ironman races last between eight and seventeen hours to be completed.

**\* Minimum Temperature:** The lowest temperature recorded during the day, usually registered just after sunrise, which happens to coincide mostly with the event's start.

**\* 10 AM Temperature:** The specific temperature at 10:00 AM local time often represents typical mid-racing conditions.

**2. Relative Humidity:** Measured in percentage (%): High humidity reduces sweat evaporation efficiency and affects perceived temperature (feels warmer in high humidity). It is Important for hydration planning, Performance expectations, and heat management strategies.

**3. Wind Speed:** Measured in meters per second (m/s), it influences cycling speeds and effort requirements. Since it can vary significantly during the race, we opted for the average wind speed throughout the day. We deliberately omitted the wind direction because nearly all bike and run legs in Ironman races are conducted in loops, making the impact of wind direction relatively small.

**4. Air Pressure:** Measured in hectopascals (hPa) or millibars (mb). It serves as an indicator of weather conditions and can affect oxygen availability at different elevations. We have extracted the daily average since the variation in air pressure during the day is often minimal.

**5. Cloud Coverage:** Measured in percentage (%). Range: 0% (clear sky) to 100% (completely overcast). It impacts the race conditions by influencing the temperature perception, and it affects solar radiation exposure. The value extracted is the cloud coverage daily average.

**6. Solar Radiation:** Measured in watts per square meter (W/m²). It is the total amount of solar energy reaching the Earth's surface. His impact on athletic performance is manifested by a direct contribution to thermal stress and heightens perceived exertion. This parameter was calculated using a physics-based approach to estimate global solar radiation at race locations based on fundamental astronomical principles and empirical adjustments for cloud cover (Duffie & Beckman, 2013; Sengupta et al., 2018) (appendix A).

**7. Wet Bulb Globe Temperature (WBGT):** Measured in degrees Celsius (°C). The Wet Bulb Globe Temperature (WBGT) is a composite temperature used to estimate the combined effect of temperature, humidity, wind speed, and solar radiation on humans. Unlike simple air temperature measurements, WBGT provides a comprehensive assessment of environmental heat stress that directly impacts athletic performance and safety in outdoor endurance events, such as triathlons.

World Triathlon provides clear, standardized WBGT thresholds with mandated operational responses. Conversely, Ironman races, due to their extreme duration and unique environmental exposures (e.g., radiant heat), prioritize multi-faceted heat mitigation strategies—acclimatization, advanced cooling, medical monitoring, and timing adjustments—over reliance on a single WBGT cancellation threshold. Emerging evidence suggests current thresholds across endurance sports may require revision downward to enhance athlete protection against exertional heat illness, highlighting the need for continued physiological monitoring and model validation.

The standardized WBGT threshold for triathlon, as set by World Triathlon, is 32.2°C, with research suggesting a potential revision to 23–26°C for enhanced safety. In contrast, Ironman races do not specify WBGT thresholds.

After implementing, calculating, and comparing three methods representing the state-of-the-art approach to WBGT estimation across various environmental conditions, we adopted the Australian Bureau of Meteorology approach due to its superior alignment with official WBGT regulations and thresholds used by governing bodies. The adopted approach provides reliable estimates that directly correspond to established safety protocols and demonstrates the highest consistency with documented race modification decisions at previous Ironman events (Appendix B).

**9. Water temperature:** Water temperature represents a critical environmental variable with substantial physiological and performance implications for ultra-endurance triathlon competitions (Tipton & Bradford, 2014). The present research addresses a significant methodological gap in triathlon performance analysis: the absence of comprehensive historical water temperature data for Ironman venues globally. The development of an accurate estimation model facilitates retrospective analysis of this critical environmental factor and enables more sophisticated understanding of performance determinants in ultra-endurance multisport competition (Wegelin & Hoffman, 2011).

Due to the difficulty of accessing historical data for water temperature in Ironman venues and the absence of this information on official race websites, we decided to calculate a valid approximation. The mathematical modeling methodology used provides statistically validated water temperature estimates for global Ironman venues, enabling comprehensive analysis of this critical environmental factor (appendix C).

**3-4 Data Preprocessing Workflow:**

The data processing workflow for the Ironman Triathlon analysis project outlines the transformation of raw race data through seven sequential stages (S1-S7). The workflow process utilizes Python scripts, primarily written by AI LLMs and coding agents (from ChatGPT to GitHub Copilot).

S1. Initial Data Collection: collection of the list of events targeted. A list of Ironman’s races held between 2015 and 2023, which contains basic race information including events, dates, locations, and athlete counts. It was then enhanced by adding geographic coordinates (latitude/longitude) for each race location.

S2. Data Transformation and Segregation: Separation of full Ironman (140.6) races and half Ironman (70.3) races, collection of athletes' results, anonymization, and standardization of data.

S3. Elevation and Weather Data extraction: Addition of venue altitude (location elevation), and addition of course-specific elevation data (bike elevation gain, run elevation gain).

S4. Basic weather data Integration: The following parameters were added for each location and date: Temperature metrics (max, min, 10 AM), Relative humidity, Wind speed, Atmospheric pressure, and Cloud coverage.

S5. Solar Radiation Calculation: addition of solar radiation values.

S6. Water Temperature Estimation: Calculation of the water temperature estimates.

S7. WBGT (Wet Bulb Globe Temperature) Calculation: Computation of all races WBGT values.

S8. Split times and total time cleaning: removal of times below world records and above cut-off times.

**Disclaimer**

The author acknowledges the use of multiple Large Language Models (LLMs) and coding agents to assist with generating code and drafting text in this paper. The author takes full responsibility for the content, accuracy, and originality of the work.

**IV- Data Analysis:**

**4-1 Statistical Analysis:**

**1- Full Ironman Analysis:**

* 1. **Weather Parameters:**

Weather parameters include temperature, humidity, wind, pressure, cloud cover, water temperature, solar radiation, and WBGT (Wet Bulb Globe Temperature). These environmental factors influence physiological responses, performance, and safety during the race. Maximum temperatures at race venues range from 9.0°C to 42.0°C (average: 24.2°C). Relative humidity varies from 11.0% to 100.0% (average: 68.3%), while water temperature ranges from 5.0°C to 30.0°C (average: 21.0°C).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Statistic | Max Temp (°C) | 10AM Temp (°C) | Min Temp (°C) | Rel. Humidity (%) | Wind Speed |
| count | 306 | 306 | 306 | 306 | 306 |
| mean | 24.23 | 20.77 | 15.92 | 68.28 | 11.24 |
| std | 4.65 | 4.39 | 4.49 | 17.52 | 6.09 |
| min | 9.00 | 5.60 | 2.20 | 11.00 | 3.34 |
| max | 42.00 | 35.80 | 29.30 | 100.00 | 40.97 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Statistic | Pressure (hPa) | Cloud Cover (%) | Water Temp (°C) | Solar Radiation | WBGT |
| count | 306 | 306 | 306 | 306 | 306 |
| mean | 1015.91 | 52.69 | 20.99 | 472.38 | 21.84 |
| std | 5.20 | 40.21 | 6.47 | 212.83 | 4.07 |
| min | 995.95 | 0.00 | 5.00 | 160.94 | 7.80 |
| max | 1029.02 | 100.00 | 30.00 | 942.82 | 32.37 |

* 1. **Elevation:**

Elevation parameters include the venue location elevation and the total elevation gain during the bike and run portions of the race. These factors greatly influence athlete performance and race strategy.

The location elevation ranges from -115.0m to 2299.0m, with an average of 199.6m. The bike course elevation gain varies from 165.0m to 3269.0m, with an average of 1156.0m. The run elevation ranges from 23.0m to 2122.0m, with an average of 340.1m.

|  |  |  |  |
| --- | --- | --- | --- |
| Statistic | Location Elevation (m) | Bike Elevation (m) | Run Elevation (m) |
| count | 306 | 306 | 306 |
| mean | 199.64 | 1155.96 | 340.12 |
| std | 341.63 | 811.08 | 518.02 |
| min | -115.00 | 165.00 | 23.00 |
| max | 2299.00 | 3269.00 | 2122.00 |

**1-3 Data visualization:**

The following visualizations illustrate the distribution and relationships between elevation and weather parameters.

- Histograms of Parameters

These histograms show the distribution of each parameter across all race locations:

A graph of different sizes and shapes

AI-generated content may be incorrect.

**Figure 1: Histograms showing the distribution of elevation and weather parameters.**

- Correlation Matrix

A screenshot of a graph

AI-generated content may be incorrect.-The correlation matrix below shows the relationships between different parameters:

**Figure 2: Correlation matrix showing relationships between elevation and weather parameters.**

-Notable Correlations:

The correlation matrix highlights several important relationships:

A strong positive correlation exists between maximum temperature and WBGT (r = 0.93), indicating that maximum temperature plays a significant role in determining the heat stress index.

Temperature variables (max, 10 AM, min) are highly correlated, indicating consistent temperature patterns across race days.

Location elevation exhibits weak negative correlations with temperature variables, reflecting the cooling effect of higher elevations.

Bike and run elevation have a moderate positive correlation (0.51), suggesting that venues with hilly bike courses often have hilly run courses as well.

The analysis of elevation and weather conditions in Ironman races reveals the diverse range of environmental factors athletes encounter. These differences significantly impact performance, strategy, and safety. Race officials utilize this data to plan safety measures. At the same time, athletes and coaches rely on it to prepare race-specific training and strategies.

**1-4** **Race Results:**

The cleaning process focused on removing unrealistic time data in the Ironman results dataset, as follows:

- Swim times below the world record (45 minutes) or above the cutoff (2:20:00).

- Bike times below the world record (4:09:00) or cumulative times above the cutoff (10:30:00).

- Run times below the world record (2:35:00) or cumulative times above the cutoff (17:00:00).

- Total times below the world record (7:35:39) or above the cutoff (17:00:00).

The data cleaning process resulted in:

- Original dataset: 440,063 rows.

- Cleaned dataset: 411,673 rows.

- Removed: 28,390 rows (6.45% of the original data)

The analysis accounts for the fact that:

- Total time includes both T1 (swim-to-bike) and T2 (bike-to-run) transitions.

- Individual split times (swim, bike, run) only account for time spent in each discipline.

- Transitions typically add 2-10 minutes to the total time.

- The sum of split times will always be less than the total finish time, with the difference representing transition times.

# **2. Ironman 70.3 Analysis:**

## **2-1 Weather Parameters:**

Weather parameters include temperature, humidity, wind, pressure, cloud cover, water temperature, solar radiation, and WBGT (Wet Bulb Globe Temperature). These environmental factors influence physiological responses, performance, and safety during the race. Maximum temperatures at race venues range from 11.5°C to 40.4°C (average: 24.8°C). Relative humidity varies from 6.0% to 100.0% (average: 66.3%), while water temperature ranges from 5.0°C to 30.0°C (average: 21.9°C).

- Temperature and Humidity Statistics:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Statistic | Max Temp (°C) | 10AM Temp (°C) | Min Temp (°C) | Rel. Humidity (%) | Wind Speed |
| count | 613 | 613 | 613 | 613 | 613 |
| mean | 24.81 | 21.44 | 16.21 | 66.31 | 11.42 |
| std | 4.95 | 4.83 | 4.76 | 16.24 | 5.21 |
| min | 11.50 | 6.70 | 4.50 | 6.00 | 2.98 |
| max | 40.40 | 34.10 | 29.20 | 100.00 | 33.67 |

- Other weather parameters:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Statistic | Pressure (hPa) | Cloud Cover (%) | Water Temp (°C) | Solar Radiation | WBGT |
| count | 613 | 613 | 613 | 613 | 613 |
| mean | 1015.24 | 51.79 | 21.87 | 486.28 | 22.24 |
| std | 5.13 | 40.95 | 6.57 | 212.21 | 4.33 |
| min | 995.44 | 0.00 | 5.00 | 150.00 | 11.16 |
| max | 1029.88 | 100.00 | 30.00 | 900.00 | 35.41 |

## **2-2 Elevations :**

Elevation parameters include the venue location elevation and the cumulative elevation gain in the bike and run portions of the race. These factors significantly impact athlete performance and race strategy.

The location elevation ranges from -115.0m to 2408.0m, with an average elevation of 215.0m. The bike course elevation gain varies from 49.0m to 3277.0m (avg: 911.2m), while the run elevation ranges from 0.0m to 3067.0m (avg: 274.2m).

|  |  |  |  |
| --- | --- | --- | --- |
| Statistic | Location Elevation (m) | Bike Elevation (m) | Run Elevation (m) |
| count | 613 | 609 | 601 |
| mean | 215.02 | 911.17 | 274.18 |
| std | 380.83 | 689.13 | 460.05 |
| min | -115.00 | 49.00 | 0.00 |
| max | 2408.00 | 3277.00 | 3067.00 |

## **2-3 Data Visualisation :**

The following visualizations illustrate the distribution and relationships between elevation and weather parameters.

### - Histograms of Parameters:

These histograms show the distribution of each parameter across all race locations:

A group of blue and white graphs

AI-generated content may be incorrect.

Figure 1: Histograms showing the distribution of elevation and weather parameters.

### - Correlation Matrix:

The correlation matrix below shows the relationships between different parameters:

A screenshot of a graph

AI-generated content may be incorrect.

Figure 2: Correlation matrix showing relationships between elevation and weather parameters.

## **2-4 Notable Correlations:**

The correlation matrix shows several important relationships:

• There is a strong positive correlation between maximum temperature and WBGT (0.93), indicating that maximum temperature is a key factor influencing the heat stress index.

• Temperature variables (max, 10 AM, min) are highly correlated with each other, demonstrating consistent temperature patterns across race days.

• Location elevation has weak negative correlations with temperature variables, reflecting the general cooling effect of higher elevations.

• Bike and run elevation have a moderate positive correlation (0.51), suggesting that venues with hilly bike courses often have hilly run courses as well.

The analysis of elevation and weather parameters in Ironman competitions shows the wide range of environmental conditions athletes encounter. These variations have a significant impact on performance, strategy, and safety considerations. Race directors use these parameters to plan safety measures, while athletes and coaches use them to develop race-specific training and strategies.

## **2-5 Race Results:**

The cleaning process targeted the time data in the Ironman 70.3 results dataset with the removal of unrealistic times as follows:

* Swim times below world record (19:00 minutes) or above cutoff (1:10:00).
* Bike times below world record (1:50:00) or above cutoff (5:30:00).
* Run times below world record (1:05:00) or above cutoff (3:00:00).
* Total times below world record (3:35:00) or above cutoff (8:30:00).

The data cleaning process resulted in:

* Original dataset: 818,396 rows
* Cleaned dataset: 727,671 rows
* Removed: 90,725 rows (11.09% of the original data)

The analysis accounts for the fact that:

* Total time includes both T1 (swim-to-bike) and T2 (bike-to-run) transitions
* Individual split times (swim, bike, run) only include time spent in that discipline
* Transitions typically add 2-10 minutes each to the total time
* The sum of split times will always be less than the total finish time, with the difference representing transition times

### - Distribution of Cleaned Race Times

The following histograms show the distribution of cleaned split times and total times:

A group of colored graphs

AI-generated content may be incorrect.

Figure 3: Histograms showing the distribution of cleaned swim, bike, run, and total times.

### Basic Statistics for Cleaned Race Times (seconds):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Statistic | Swim | Bike | Run | Total |
| count | 727671 | 727671 | 727671 | 727671 |
| mean | 2364.04 | 10800.85 | 7529.35 | 21339.55 |
| std | 484.64 | 1504.42 | 1402.51 | 3048.70 |
| min | 1140.00 | 6604.00 | 3900.00 | 12902.00 |
| max | 4200.00 | 19665.00 | 10800.00 | 30599.00 |

**4-2 Linear Regression:**

**1- Methodology**

Linear regression models were created to assess the relationships between environmental factors (temperatures, elevations, wind speed, WBGT) and performance times. The models were separated by division (professional vs. age group) and gender to identify effects specific to each group.  
  
Feature Importance: For each model, feature importance was calculated to determine which environmental factors had the strongest impact on performance.

1. **Half Ironman 70.3**

We perform a linear regression analysis to examine how environmental factors—such as temperature, elevation, wind speed, and WBGT—affect bike and run performance across various divisions and age groups. The analysis indicates that bike elevation is the strongest predictor of bike times, while temperature and WBGT are the most significant for running times. Professional athletes exhibit stronger correlations between environmental factors and performance compared to age-group athletes.

**2-1 Key Findings:**

- Overall Environmental Impact:

Temperature Impact: For every 1°C rise in temperature, run time increases by about 0.91 minutes (54.65 seconds), with R²=0.034 and RMSE=1378.6 seconds. Temperature has a greater effect on run performance than on bike performance.  
Elevation Impact: For bike performance, course elevation is the most significant predictor, with bike time increasing by approximately 0.24 seconds per meter of elevation gain (R² = 0.012, RMSE = 1,494.7 seconds).

# - Professional Athletes Analysis:

Pro Men: The regression models for professional men demonstrated higher predictive power (Bike: R² = 0.111, Run: R² = 0.086) compared to age-group athletes. For bike performance, elevation was the strongest predictor (28.8 minutes per standard deviation), followed by temperature (-9.2 minutes per standard deviation). For run performance, temperature was the dominant factor (7.1 minutes per standard deviation).  
  
Pro Women: Similar to Pro Men, professional women showed stronger correlations between environmental factors and performance (Bike: R² = 0.094, Run: R² = 0.066). Bike elevation had the greatest impact on bike times (34.9 minutes per standard deviation), while temperature most significantly affected run times (6.4 minutes per standard deviation).

## - Age Group Athlete Analysis:

Young Men (30-45): For this group, models had lower predictive accuracy (Bike: R²=0.021, Run: R²=0.057). Bike elevation remained the strongest predictor of bike performance (36.3 minutes per standard deviation), while temperature had the greatest effect on run performance (15.4 minutes per standard deviation).

Older Men (50-70): For older men, environmental factors showed slightly stronger relationships (Bike: R²=0.029, Run: R²=0.040). Notably, WBGT (Wet Bulb Globe Temperature) emerged as the most influential predictor of run performance (11.8 minutes per standard deviation), indicating increased heat sensitivity.

Young Women (30-45): Models for young women displayed similar patterns to men (Bike: R²=0.024, Run: R²=0.028), with bike elevation dominating bike performance (35.4 minutes per standard deviation) and WBGT strongly affecting run performance (11.0 minutes per standard deviation).

Older Women (50-70): Older women showed the strongest influence from heat stress indicators, with WBGT being the primary predictor for run performance (12.3 minutes per standard deviation).

## - Environmental Impact Scores:

The analysis developed specialized Environmental Impact Scores for both bike and run segments, providing a composite measure of how challenging conditions would be for each segment. Professional men showed small positive correlations between bike impact scores and actual bike times (r=0.095), indicating that challenging conditions did have a measurable effect on performance.  
Professional women showed stronger correlations between environmental factors and performance, with bike impact scores correlating more strongly with bike times (r=0.173) than observed in men.

- Water Temperature Effect on Swim Performance:

Water temperature seems to have a statistically insignificant impact on swim performance in Half Ironman events. Although there is a slight trend indicating faster swims in warmer water (especially above 26°C) for the overall group, the effect size is very small.

Water Temperature Effect Summary:

All Athletes: 0.02 minutes faster per 1°C increase.

Male Pros: 0.01 minutes slower per 1°C increase.

Female Pros: 0.02 minutes faster per 1°C increase.

# **2-2 Conclusions:**

This analysis reveals that environmental factors have a significant impact on Half Ironman 70.3 race performance, with effects varying across divisions, genders, and age groups. Professional athletes display stronger links between performance and environmental conditions compared to age-group athletes, probably because they race at higher intensities closer to their physiological limits. The results emphasize the importance of course-specific preparation, personalized pacing strategies, and age-appropriate training methods. By understanding and considering these environmental impacts, athletes can improve their race performance under various conditions.

**3- Full Ironman:**

**3-1 Key Finding:**

The detailed analysis of Full Ironman data confirms that environmental factors significantly influence triathlon performance in complex ways. The elevation profiles of the bike and run courses consistently stand out as the strongest predictors across all categories. Temperature effects are complex, with WBGT proving to be a more accurate predictor than absolute temperature. It is clear that professional athletes have different sensitivities compared to age-group athletes, and notable gender differences exist in how athletes respond to environmental conditions.

- Key Environmental Factors (by impact order):

- Water temperature has the most significant effect on swimming performance.

- WBGT/heat stress (most potent effect on running performance)

- Wind speed (most potent effect on bike performance)

- Elevation (moderate effect across all disciplines)

- Air temperature (complex relationship, requires further study)

- Professional Athlete Analysis:

\* Single Factor Effects (MPRO)

-Temperature Effects:

- Bike: For every 1°C increase in temperature, bike time decreases by 0.87 minutes

- Run: Temperature showed minimal effect on run performance (R² = 0.014)

Elevation Effects:

- Bike: Bike course elevation is a significant predictor of bike time

- Run: Run course elevation is a significant predictor of run time

Other Environmental Factors:

- WBGT (Wet Bulb Globe Temperature): Notable impact on run performance

- Wind Speed: More significant impact on the bike than on running.

- Location Elevation: Affects both bike and run performance

Water Temperature Effects:

- Swim: For every 1°C increase in water temperature, swim times tend to decrease

- Optimal water temperature range appears to be 20-24°C

\* Multiple Regression Analysis

-Key Environmental Predictors for MPRO

Bike Performance:

1. Bike Elevation (strongest positive effect)

2. WBGT (negative effect)

3. Location Elevation (positive effect)

4. Temperature (weak positive effect)

5. Wind Speed (weak negative effect)

Run Performance:

1. Bike Elevation (powerful positive effect)

2. Run Elevation (strong negative effect)

3. Location Elevation (strong positive effect)

4. Wind Speed (strong positive effect)

5. WBGT (weak positive effect)

\*Key Environmental Predictors for FPRO

Bike Performance:

1. \*\*Bike Elevation\*\*: This consistently demonstrates a powerful positive impact on performance.

2. \*\*WBGT (Wet Bulb Globe Temperature)\*\*: This has a detrimental effect on performance.

3. \*\*Location Elevation\*\*: This effectively enhances performance.

4. \*\*Temperature\*\*: This exerts a weak but positive influence on performance.

5. \*\*Wind Speed\*\*: This introduces a weak negative effect on performance.

Run Performance:

1. \*\*Bike Elevation\*\*: This strongly enhances performance.

2. \*\*Run Elevation\*\*: This significantly hampers performance.

3. \*\*Location Elevation\*\*: This greatly boosts performance.

4. \*\*Wind Speed\*\*: This positively affects performance.

5. \*\*WBGT\*\*: This offers a weak positive influence on performance.

- Age Group Analysis:

Male Age Groups:

- Younger (30-45): These individuals are less sensitive to environmental factors compared to their older counterparts.

- Older adults (50-70): This group exhibits heightened sensitivity to temperature and elevation, which can impact their performance.

Female Age Groups:

- Younger (30-45): This group exhibits patterns similar to males, but with a lesser intensity.

- Older adults (50-70): This demographic exhibits the highest sensitivity to environmental conditions, particularly temperature.

- Water Temperature Analysis

General Relationship

Water temperature has a significant and non-linear effect on swim performance across all categories of athletes. The relationship follows a U-shaped curve, establishing the optimal performance range between 20 °C and 24 °C.

Effect on Athlete Group

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Athlete Group | All Athletes | M Pro | F Pro | M Age Group | F Age Group |
| Effect (min/°C) | 0.45 | 0.38 | 0.41 | 0.49 | 0.56 |

- The optimal water temperature for maximum swim performance is approximately 22.5°C.

- Performance declines by 2-3% for every 2°C deviation from this optimum.

- Professional athletes display 30% less sensitivity to temperature extremes than age-group athletes.

- Female athletes are notably more sensitive to variations in water temperature than male athletes.

**3-2 Conclusion:**

The comprehensive analysis of Full Ironman data confirms that environmental factors play a crucial role in influencing triathlon performance in complex ways. The elevation profiles of the bike and run courses consistently stand out as the strongest predictors across all divisions. The effects of temperature are complex, with WBGT proving to be a more accurate predictor than absolute temperature. Professional athletes show different sensitivities compared to age-group athletes, and notable gender differences exist in how athletes respond to environmental factors.

1. **Full Vs Half Ironman: Regression Analysis Results Comparison.**

This summary highlights the key differences in environmental impacts between Full Ironman and Half Ironman 70.3 events based on regression analyses of performance data.

1. Race Intensity Amplification Effect:

Higher race intensity in Half Ironman events amplifies environmental impacts, with elevation and heat stress having a disproportionately larger effect per unit of distance compared to Full Ironman events.

2. Water Temperature Significance:

Full Ironman: Substantial effect (-0.86 min/°C, R²=0.10-0.27)

Half Ironman: Negligible effect (-0.02 min/°C, R²≈0.000)

3. Age-Related Environmental Sensitivity:

Full Ironman: Gradual increase with age.

Half Ironman: Dramatic increase with age, particularly for heat effects.

4. Contradictory Temperature Effects:

Full Ironman: Counterintuitive benefits with higher temperatures.

Half Ironman: Expected performance decline with higher temperatures.

5. Practical Implications:

Considering environmental factors becomes even more important as race intensity increases and athlete age rises, with Half Ironman requiring more proactive environmental planning, especially for older athletes.

* 1. **Decision Tree Analysis:**

**1- Simple Decision Tree:**

**1- 1 Overview:**

The decision tree regression model with a maximum depth of 10 demonstrated the best performance among models with maximum depths of 3, 5, 10, and 20. This deeper tree model offers more detailed insights than the other models.

**1-2 Model Performance Summary:**

# The tree model with a maximum depth of 10 significantly enhanced predictive performance across all segments, with the most notable improvement in the Overall time segment. The Swim and Run segments demonstrated a powerful model fit.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Segment | Swim | Bike | Run | Overall |
| R² Value | 0.817 | 0.237 | 0.542 | 0.629 |
| RMSE (minutes) | 0.33 | 0.81 | 0.92 | 1.69 |

# **1-3 Critical Environmental Factors by Segment:**

# Swim Segment:

# - WBGT (Wet Bulb Globe Temperature): Most influential factor (26.4% importance).

# - Minimum Temperature: Second most important (19.2% importance).

# - Relative Humidity: Third most important (16.5% importance).

# - Weather dominates over elevation: Weather factors account for approximately 70% of the model's predictive power vs. 30% for elevation.

# Bike Segment:

# - Location Elevation: Top factor (17.7% importance).

# - Average Wind Speed: Second most important (10.0% importance).

# - Run Elevation: Third most important (10.0% importance).

# - Nearly balanced influence: Weather factors account for 34% vs. 31% for elevation factors.

# Run Segment:

# - Relative Humidity: Most critical factor (22.5% importance).

# - Location Elevation: Second most important (13.9% importance).

# - Minimum Temperature: Third most important (12.4% importance).

# - Weather factors dominate: Weather accounts for 48% of predictive power vs. 32% for elevation.

# Overall Race Time:

# - Average Wind Speed: Primary factor (18.4% importance).

# - Location Elevation: Second most important (14.5% importance).

# - Relative Humidity: Third most important (13.3% importance).

# - Strong weather dominance: Weather accounts for 59% of predictive power vs. 16% for elevation.

# **2- Ironman Random Forest Analysis Summary:**

# **2-1 Overview:**

# This Random Forest regression analysis examines the impact of weather conditions and elevation parameters on Ironman triathlon performance. It expands on previous Decision Tree models by applying ensemble methods to improve prediction accuracy. The Random Forest models consistently outperformed Decision Tree models across all race segments, with the most significant gains seen in the Bike segment.

# **2-2 Key Environmental Factors by Race Segment**

# Swim Segment:

# - WBGT (Wet Bulb Globe Temperature): Most influential factor (~22% importance).

# - Water Temperature: Second most important (~18% importance).

# - Minimum Temperature: Third most important (~15% importance).

# - Weather factors continue to dominate over elevation factors.

# Bike Segment:

# - Location Elevation: Top factor (~19% importance).

# - Average Wind Speed: Second most important (~14% importance).

# - Run Elevation: Third most important (~11% importance).

# - More balanced influence between weather and elevation factors than in other segments.

# Run Segment:

# - Relative Humidity: Most critical factor (~20% importance).

# - Location Elevation: Second most important (~17% importance).

# - Minimum Temperature: Third most important (~14% importance).

# - Weather factors dominate, but elevation shows increased importance compared to Decision Tree models.

# Overall Race Time:

# - Average Wind Speed: Primary factor (~17% importance).

# - Location Elevation: Second most important (~15% importance).

# - Min Temperature: Third most important (~13% importance).

# - Weather factors remain dominant but with a more balanced distribution of importance.

**4-4 XGBoost Regression:**

**1- Ironman XGBoost Regression Analysis Summary:**

**1-1 Overview:**

We applied XGBoost regression models to analyze Ironman triathlon performance data across four race segments (swim, bike, run, and overall time. The study tested various parameters, including learning rates (0.01 and 0.1), tree depths (5 and 7), estimators (100 and 200), and subsample rates (0.8 and 1.0) across different race segments, evaluating 16 combinations. Results showed that higher learning rates (0.1) performed better overall. The optimal tree depth was approximately 5, and models with 100 estimators were comparable to those with 200 estimators. Using a subsample rate of 0.8 helped prevent overfitting.

**1-2 Key Feature Importance Findings:**

Swim Segment:

- Max temperature is by far the most influential feature (importance score ~1000)

- Average wind speed is the second most important feature (importance score ~520)

- Temperature metrics (max, min, 10 AM) collectively dominate the model.

- Water temperature has moderate importance, as expected.

- Elevation features have the least impact on swimming performance.

Bike Segment:

- Max temperature remains the most influential feature (importance score ~400)

- Average wind speed has a significant impact (importance score ~300)

- Minimum temperature and temperature at 10 AM show moderate importance

- Location elevation has more impact on biking than on other segments

- Wind conditions are proportionally more critical for biking than for other segments

Run Segment:

- Max temperature is again the dominant feature (importance score ~320)

- Min temperature, wind speed, and humidity all have similar moderate importance (~150-200)

- Location elevation has more impact on running than on swimming

- Temperature factors collectively account for over 50% of the model's predictive power

Overall Time:

- Max temperature is the most critical predictor (importance score ~480)

- Temperature metrics collectively account for the majority of the variance explained

- Wind and humidity show consistent moderate importance across all segments

- Elevation features have relatively less impact but are still significant

Weather vs. Elevation Impact:

- Weather factors dominate all segments, with 4-5x higher importance than elevation features

- For Swim: Weather accounts for ~2700 importance points vs ~550 for elevation

- For Bike: Weather accounts for ~1350 importance points vs ~370 for elevation

- For Run: Weather accounts for ~1180 importance points vs ~350 for elevation

- For Overall Time: Weather accounts for ~1250 importance points vs ~300 for elevation

- Consistent patterns show that weather conditions, particularly temperature, have a substantially higher impact.

**2- Half Ironman XGBoost Regression Analysis Summary:**

**2-1 Overview:**

The same approach as for the full Ironman was applied to the 70.3 data.

**2-2 Key Feature Importance Findings:**

Swim Segment:

- Max temperature is the most influential feature (importance score ~800)

- Water temperature is the second most important feature (importance score ~650)

- Average wind speed has a significant impact.

- Temperature metrics (max, min, 10 AM) collectively dominate the model.

- Elevation features have the least impact on swimming performance.

Bike Segment:

- Max temperature remains the most influential feature (importance score ~350).

- Average wind speed has a significant impact (importance score ~280).

- Location elevation has more impact on biking than on other segments.

- Wind conditions are proportionally more critical for biking than for other segments.

- Temperature at 10 AM shows moderate importance.

Run Segment:

- Max temperature is again the dominant feature (importance score ~300)

- Relative humidity has a higher impact than in other segments (importance score ~200).

- Minimum temperature and temperature at 10 AM show moderate importance.

- Location elevation has more impact on running than on swimming.

- Temperature factors collectively account for over 50% of the model's predictive power.

Overall Time:

- Max temperature is the most influential feature (importance score ~450).

- Relative humidity and average wind speed show significant importance.

- The temperature at 10 AM has a higher impact than in Full Ironman races.

- Location elevation shows moderate importance.

- The combined impact of all temperature metrics accounts for approximately 60% of the model's predictive power.

Weather vs. Elevation Impact:

This analysis shows that weather factors have a much greater effect on Half Ironman performance than elevation across all segments, with the biggest impact in the swim segment and the smallest in the bike segment. Compared to Full Ironman races, the Half Ironman format has slightly higher weather-to-elevation ratios, indicating that weather conditions may have a relatively larger influence in the shorter race format.

**3- Division-Specific Environmental Impact Analysis in Half Ironman Triathlons:**

**3-1 Overview:**

This document summarizes the findings from applying XGBoost regression models to analyze how environmental factors (weather and elevation) affect different division segments in the Half Ironman 70.3 triathlon performance. Using the optimal XGBoost model parameters identified in our previous analysis (learning\_rate=0.1, max\_depth=5, n\_estimators=100, subsample=0.8), we examined performance differences across six athlete divisions.

1. MPRO (Male Professionals)

2. FPRO (Female Professionals)

3. Male\_AG\_Young (Male Age Group < 45)

4. Male\_AG\_Old (Male Age Group >= 45)

5. Female\_AG\_Young (Female Age Group < 45)

6. Female\_AG\_Old (Female Age Group >= 45)

**3-2 Model Performance by Division:**

- R² Values (Explained Variance): The predictive power of environmental factors varies across divisions:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Division | MPRO | FPRO | Male\_AG\_Young | Male\_AG\_Old | Female\_AG\_Young | Female\_AG\_Old |
| Swim | 0.90 | 0.89 | 0.78 | 0.76 | 0.79 | 0.77 |
| Bike | 0.41 | 0.38 | 0.32 | 0.34 | 0.30 | 0.32 |
| Run | 0.62 | 0.60 | 0.52 | 0.54 | 0.51 | 0.53 |
| Overall | o.71 | 0.69 | 0.60 | 0.62 | 0.58 | 0.60 |

- Key observations:

- Professional athletes' performance is more predictable based on environmental factors.

- The swimming segment is most affected by the environment across all divisions.

- The bike segment remains the most challenging to predict from environmental factors alone.

- Older age groups show slightly higher environmental sensitivity than younger age groups in bike and run segments.

**3-3 Key Feature Importance by Division:**

The top environmental factors affecting performance by division are ranked as follows:

MPRO:

- Swim: max\_temperature, water\_temperature, average\_wind\_speed

- Bike: max\_temperature, average\_wind\_speed, location\_elevation

- Run: max\_temperature, relative\_humidity, min\_temperature

- Overall: max\_temperature, average\_wind\_speed, temperature\_10AM

FPRO:

- Swim: max\_temperature, water\_temperature, relative\_humidity

- Bike: max\_temperature, average\_wind\_speed, temperature\_10AM

- Run: max\_temperature, relative\_humidity, temperature\_10AM

- Overall: max\_temperature, temperature\_10AM, relative\_humidity

Male\_AG\_Young:

- Swim: max\_temperature, water\_temperature, average\_wind\_speed

- Bike: max\_temperature, average\_wind\_speed, location\_elevation

- Run: max\_temperature, relative\_humidity, average\_wind\_speed

- Overall: max\_temperature, average\_wind\_speed, relative\_humidity

Male\_AG\_Old:

- Swim: max\_temperature, water\_temperature, relative\_humidity

- Bike: max\_temperature, average\_wind\_speed, location\_elevation

- Run: max\_temperature, relative\_humidity, location\_elevation

- Overall: max\_temperature, relative\_humidity, location\_elevation

Female\_AG\_Young:

- Swim: max\_temperature, water\_temperature, temperature\_10AM

- Bike: max\_temperature, temperature\_10AM, average\_wind\_speed

- Run: max\_temperature, temperature\_10AM, relative\_humidity

- Overall: max\_temperature, temperature\_10AM, relative\_humidity

Female\_AG\_Old:

- Swim: max\_temperature, water\_temperature, relative\_humidity

- Bike: max\_temperature, temperature\_10AM, relative\_humidity

- Run: max\_temperature, relative\_humidity, location\_elevation

- Overall: max\_temperature, relative\_humidity, temperature\_10AM

**3-4 Division-Specific Environmental Sensitivities:**

Professional Athletes:

- More sensitive to environmental factors overall (higher R² values).

- Maximum temperature is the dominant factor for both male and female professionals.

- Show greater performance variance with wind conditions compared to age groupers.

- Performance is more predictable from environmental factors, suggesting more optimized racing.

Gender-Based Patterns:

- Female athletes show higher sensitivity to humidity and temperature variations.

- Male athletes demonstrate greater impact from wind conditions, particularly in bike segments.

- Female age groupers appear more affected by morning temperatures (10 AM) than their male counterparts.

- Water temperature impacts female swim performance slightly more than male swim performance.

Age-Based Differences:

- Older athletes (45+) show greater sensitivity to elevation factors in all segments.

- Younger athletes demonstrate higher responsiveness to wind conditions.

- Temperature variations have more consistent impacts across younger age groups.

- Older athletes' performance is more affected by relative humidity, particularly in run segments.

**4-5 Quantile Regression:**

**1- Ironman Triathlon Performance Analysis Using Quantile Regression:**

**1-1 Methodology:**

-Research Questions:

1. How do sub-event times (swimming, bike, run) influence overall performance differently across different performance quantiles?

2. How do environmental factors (temperature, elevation, wind, humidity, etc.) affect athletes differently depending on their performance level?

3. Do certain factors have more potent effects on elite athletes (lower quantiles) versus recreational athletes (higher quantiles)?

4. How do performance patterns differ between male and female athletes across the performance spectrum?

-Data Preparation:

- Missing values identified and handled through row removal for critical variables

- Gender information preserved for analysis of male/female performance differences

- Environmental factors (elevation, temperature, water temperature, WBGT, wind, humidity) included as independent variables

-Statistical Approach:

- Quantile regression implemented to examine effects across different performance levels (0.1, 0.25, 0.5, 0.75, 0.9 quantiles).

- Statistical significance threshold of p < 0.05 applied to identify reliable effects.

- Four regression models developed:

1. Model 1: Sub-event times (swim, bike, run) predicting total time

2. Model 2: Environmental factors (elevation, temperature, water temperature, WBGT, wind, humidity) predicting total time

3. Model 3: Combined model with both sub-event times and environmental factors

4. Model 4: Gender models, including basic gender effect and gender-discipline interactions

-Visualization Approach:

- Coefficient plots across quantiles for each model.

- Statistical significance markers (\*) added to identify reliable effects (p < 0.05).

- Gender-specific visualization for interaction effects.

- Separate subplots for individual variables in the full model.

**1-2 Key Findings:**

-Sub-event Times (Model 1)

1. Swim Time: Coefficient remains consistent across quantiles (~0.99), indicating swim time affects overall time similarly for both elite and recreational athletes. All coefficients are statistically significant.

2. Bike Time: Shows an upward trend across quantiles (1.006 at the 0.1 quantile to 1.018 at the 0.9 quantile), suggesting that bike performance has a more substantial effect on slower athletes. All coefficients are statistically significant.

3. Run Time: Displays most pronounced upward trend (1.01 at 0.1 quantile to 1.05 at 0.9 quantile), indicating running becomes increasingly essential for slower athletes. All coefficients are statistically significant.

-Environmental Factors (Model 2)

1. Elevation: Effect decreases across quantiles, starting positive for faster athletes and becoming negative for slower athletes. Elite athletes are more susceptible to the effects of elevation than recreational athletes. Significant effects are primarily observed at the lower and upper quantiles.

2. Temperature: Negative effect on fastest athletes (0.1 quantile) but strong positive impact on slower athletes (0.9 quantile). Higher temperatures are more detrimental to recreational athletes.

3. Water Temperature: Substantial variability across quantiles, with stronger negative effects for fastest and slowest athletes. Statistical significance varies.

4. WBGT (Wet Bulb Globe Temperature): U-shaped pattern, with positive effect for fastest athletes, decreasing for middle quantiles, increasing again for slowest athletes. Mostly statistically significant.

5. Wind Speed: Powerful effect on mid-to-back-of-pack athletes (0.5-0.9 quantiles). Statistically significant positive coefficients indicate that higher wind speeds are associated with slower finishing times. Elite athletes are better able to mitigate wind effects.

6. Humidity: Stronger effect at performance spectrum extremes (0.1 and 0.9 quantiles). For elite athletes, higher humidity is correlated with slower times, whereas the effect is more variable for recreational athletes.

Combined Model (Model 3)

1. Environmental factor impacts diminish when controlling for sub-event times, suggesting much of their effect operates through influence on swimming, biking, and running performances.

2. Differential effects across quantiles persist, reinforcing that environmental conditions affect athletes differently depending on performance level.

3. Wind speed maintains significant effects even after controlling for sub-event times, suggesting its impact extends beyond simply slowing individual disciplines.

-Gender Differences (Model 4)

1. Basic Gender Effect: Male athletes 800-1200 seconds (13-20 minutes) faster across all quantiles, with gender gap statistically significant (p < 0.05) throughout the performance spectrum.

2. Gender Gap Variation: Gap is narrower in elite athlete ranks (lower quantiles) and widens among recreational athletes (higher quantiles). Elite female athletes appear to close the gender gap more effectively than recreational female athletes.

3. Gender-Discipline Interactions:

- Swim: Males show a stronger correlation between swim and overall time in the middle quantiles

- Bike: Gender differences in bike leg contribution are most pronounced in the upper quantiles

- Run: Most consistent gender interaction effects across all quantiles

1-3 Conclusions:

1. Performance-Level Effects: Quantile regression reveals significant variation in how factors affect performance across different athlete levels, information masked by traditional regression approaches.

2. Environmental Sensitivity: Elite athletes show different response patterns to environmental conditions compared to recreational athletes, with elite athletes generally more affected by elevation and humidity, while recreational athletes are more affected by temperature and wind.

3. Gender-Specific Patterns: Training and race strategies should be tailored differently for male and female athletes, particularly in energy allocation across disciplines.

4. Statistical Rigor: Using a p < 0.05 threshold strengthens findings' validity, particularly regarding differential effects across quantiles.

**2- Full vs Half Ironman Quantile Regression Analysis Comparison:**

**2-1 Methodology Overview:**

Both analyses employed quantile regression to investigate how various factors influence triathlon performance across different performance levels (quantiles). We developed models for the 0.1, 0.25, 0.5, 0.75, and 0.9 quantiles, with a statistical significance threshold of p < 0.05.

**2-2 Sub-event Time Effects Comparison:**

-Swim Leg:

- Full Ironman: Swim coefficient remains relatively constant across quantiles, with minimally increasing impact for slower athletes (0.95-1.02).

- Half Ironman: Similar pattern but with slightly higher and more consistent coefficients (1.0-1.05), reflecting the proportionally larger contribution of the swim leg to the overall race.

-Bike Leg:

- Full Ironman: Bike coefficient increases gradually across quantiles (1.01 to 1.08), showing greater importance for recreational athletes.

- Half Ironman: Similar but slightly less pronounced upward trend (1.02 to 1.06), indicating a more balanced effect across performance levels.

-Run Leg:

- Full Ironman: Run coefficient increases significantly from lower to higher quantiles (1.03 to 1.20), becoming disproportionately important for slower athletes.

- Half Ironman: Similar increasing pattern (1.05 to 1.15) but with a less extreme slope, as the shorter run distance creates less differential impact.

**2-3 Environmental Factor Effects Comparison:**

-Elevation:

- Full Ironman: More pronounced and consistent effects across all quantiles, with a generally decreasing trend from elite to recreational athletes

- Half Ironman: U-shaped pattern, with stronger effects at both extremes of the performance spectrum (both fastest and slowest athletes show heightened responses to elevation changes)

-Temperature:

- Full Ironman: Temperature effects increase dramatically for slower athletes, reflecting their longer exposure during the afternoon heat

- Half Ironman: Temperature effects are stronger for faster athletes (lower quantiles) and show less variation across quantiles, likely because most athletes finish before peak afternoon temperatures

-Water Temperature:

- Full Ironman: More significant impact on overall performance due to longer swimming portion

- Half Ironman: Varied effects across quantiles with a general trend toward a negative relationship with race time for the fastest athletes, suggesting optimal water conditions may particularly benefit elite Half Ironman swimmers

-Wind and Humidity:

- Full Ironman: Wind effects are moderate and relatively uniform across quantiles; humidity shows stronger effects on slower athletes

- Half Ironman: Wind effects are stronger overall and increase substantially for mid-to-back-of-pack athletes (0.5-0.9 quantiles); humidity effects are concentrated in the upper quantiles but with smaller coefficient magnitudes

**2-4 Gender Differences Comparison:**

-Gender Gap:

- Full Ironman: Males typically finish faster across all quantiles, with a gap ranging from 900-1500 seconds (15-25 minutes), increasing substantially from elite to recreational athletes

- Half Ironman: Similar pattern but smaller gap, ranging from 600-900 seconds (10-15 minutes), with a less pronounced increase across quantiles

-Gender-Discipline Interactions:

- Full Ironman: Gender-swim interactions are strongest at extremes, gender-bike interactions are strongest in middle quantiles, and gender-run interactions increase dramatically for recreational athletes

- Half Ironman: Gender-swim interactions are concentrated in middle quantiles (0.25-0.75), gender-bike interactions are stronger in upper quantiles (0.5-0.9), and gender-run interactions are significant across all quantiles but strongest at the extremes (0.1 and 0.9)

**2-5 Conclusion:**

Quantile regression analysis reveals that while similar patterns exist in how factors affect performance across both race formats, essential differences emerge in the magnitude and consistency of these effects. These differences reflect the unique physiological and strategic demands of each race format, suggesting that athletes, coaches, and race organizers should consider format-specific approaches to training, race preparation, and race execution.

Key conclusions from our analysis include:

1. Different Performance Dynamics: Half Ironman events show more balanced effects of the three disciplines across performance levels, while Full Ironman events display more dramatic increases in the importance of running for recreational athletes. Environmental factors affect the two race formats differently, with extended exposure in Full Ironman amplifying the effects of temperature and humidity. In contrast, the proportionally greater importance of the bike leg in Half Ironman amplifies the impact of wind.

2. Gender Differences: The gender performance gap is smaller in absolute terms in Half Ironman events (600-900 seconds vs. 900-1500 seconds in Full). Gender-discipline interaction patterns vary between formats, indicating distinct physiological and strategic considerations for male and female athletes in each format.

3. Race Strategy Implications: For Half Ironman, more aggressive bike strategies can be effective for appropriately trained athletes, as the bike leg has a more balanced impact across performance levels. For Full Ironman, conservative bike efforts to preserve run performance are generally advised, especially for mid-to-back-of-pack athletes.

The comparative analysis highlights that simply scaling down training or race strategy from Full to Half Ironman (or scaling up from Half to Full) is insufficient. Each format requires a tailored approach that accounts for the relative importance of sub-event times, environmental factors, and physiological differences across genders and performance levels.

**V- Discussion:**

**5-1 Interpretation of Findings across All Models.**

Our comprehensive analysis of environmental factors affecting Ironman and Half Ironman triathlon performance revealed several consistent patterns across different modeling approaches. By integrating traditional statistical methods with advanced machine learning techniques, we uncovered nuanced relationships between weather parameters, course characteristics, and athletic performance.

**Temperature Effects:**

Temperature emerged as the dominant environmental factor affecting triathlon performance across all race segments and athlete divisions. Maximum temperature consistently ranked as the most influential predictor in all machine learning models, with XGBoost models showing importance scores of approximately 800 for swimming, 350-400 for biking, and 300-320 for running. The relationship between temperature and performance follows a non-linear pattern, with optimal performance occurring in moderate temperature ranges (19-27°C for air temperature, 20-24°C for water temperature). Differential impacts were observed across race segments, with swimming most sensitive to water temperature variations, while running performance declined more sharply at temperature extremes. Heat stress metrics (WBGT) proved to be better predictors than absolute temperature alone, particularly for running segments, highlighting the importance of considering temperature in conjunction with humidity and solar radiation. Surprisingly, our linear regression models showed weak negative correlations between temperature and performance times in some segments, suggesting potential adaptation effects or confounding factors at play.

**Elevation Effects:**

Course elevation characteristics showed significant but less dominant effects compared to weather variables. Bike elevation gains consistently emerged as the strongest elevation predictor for bike performance, increasing bike time by approximately 0.01 minutes per meter of elevation gain. Location elevation (altitude) demonstrated more complex effects, with intermediate altitudes (500-1500m) sometimes associated with improved performance, likely due to the balance between reduced air resistance and decreased oxygen availability. Elevation impact ratios differed between Half Ironman and Full Ironman events—For Half Ironman, weather-to-elevation impact ratios were 12.3x for swimming, 3.6x for biking, 4.6x for running, and 5.1x overall. At the same time, Full Ironman showed ratios of 6.1x, 4.2x, 3.8x, and 4.7x respectively. Professional athletes showed greater sensitivity to elevation factors than recreational athletes, likely due to their higher relative exercise intensities and ability to sustain performance closer to physiological limits.

**Wind and Humidity Effects:**

Wind and humidity emerged as secondary but significant factors influencing triathlon performance across different race segments. The average wind speed showed particular importance for bike segments (importance score of approximately 280-300), with approximately 0.5 minutes added per 1 km/h increase in wind speed, resulting in substantial cumulative effects over long courses. Relative humidity demonstrated its strongest impact on running performance (importance score ~200), particularly for older age-group athletes who may have diminished thermoregulatory capacity. Wind effects were proportionally more important in Full Ironman compared to Half Ironman bike segments, likely due to the extended duration of exposure and cumulative fatigue over the longer course. Humidity sensitivity showed notable gender differences, with female athletes exhibiting greater performance declines in high-humidity conditions than male counterparts, possibly due to physiological differences in sweat rates and thermoregulation mechanisms.

**Water Temperature and Swimming Performance:**

Water temperature analysis revealed a distinct U-shaped relationship with swim performance across all athlete categories. The optimal water temperature for swim performance was identified at 22.5°C, with performance declining at both higher and lower extremes due to increased physiological strain. Effect sizes varied by athlete category: professionals showed effects of 0.38-0.41 minutes per °C deviation from optimal, while age-group athletes experienced larger effects of 0.49-0.56 minutes per °C, indicating better adaptation capacity among elite competitors. Gender differences in water temperature sensitivity were observed, with female athletes exhibiting slightly higher sensitivity (0.41-0.56 minutes/°C) compared to males (0.38-0.49 minutes/°C), possibly due to differences in body composition. Performance declines of 2-3% were observed for every 2°C deviation from the optimal water temperature range, representing meaningful competitive differences in races where margins are often measured in seconds.

**Performance Sensitivity Across Demographics:**

Our analysis revealed substantial variations in environmental sensitivity across athlete categories. Professional athletes showed higher R² values in all models, indicating stronger relationships between environmental factors and performance. However, age-group athletes, particularly older competitors, demonstrated greater absolute sensitivity to temperature extremes, with professionals showing approximately 30% less sensitivity to temperature deviations, likely due to superior training adaptations and competitive experience. Female athletes exhibited higher sensitivity to humidity and temperature variations. In contrast, male athletes showed greater impacts from wind conditions, particularly in bike segments, suggesting gender-specific physiological responses to environmental stressors. The gender performance gap was smaller in absolute terms in Half Ironman events (600-900 seconds) compared to Full Ironman (900-1500 seconds), suggesting that race duration may magnify inherent physiological differences. Older athletes (45 years and older) showed increased sensitivity to elevation factors and heat stress indicators, such as WBGT and humidity, with WBGT emerging as the most influential predictor of run performance for older males (11.8 minutes per standard deviation), indicating age-related changes in thermoregulatory capacity. Half Ironman R² values tended to be 3-5% higher than Full Ironman, suggesting environmental factors are more predictive of performance in the shorter format, while extended exposure in Full Ironman events amplified the effects of temperature and humidity over longer durations.

**5-2 Comparison with Previous Studies:**

Our findings both corroborate and extend previous research on environmental impacts on triathlon performance.

**Agreement with Previous Literature:**

Our identified optimal temperature ranges (air: 19-27°C, water: 20-24°C) closely align with previous studies by Knechtle et al. (2024, 2025), who reported optimal air temperatures of 19-27°C for professionals and water temperatures above 22°C. These consistent findings across different methodologies and datasets strengthen confidence in the existence of ideal temperature windows for triathlon performance. The observed differential effects across athlete divisions support findings by Gibson (2024) and Hermand et al. (2019), who similarly noted distinct environmental impacts for elite versus recreational athletes, reinforcing the importance of athlete-specific approaches to environmental adaptation. Our identification of WBGT as a critical predictor aligns with Mantzios et al. (2021), who emphasized the importance of considering multiple weather parameters rather than temperature alone, highlighting the complex interaction of temperature, humidity, and radiation in determining heat stress. The negative impact of location elevation on running performance supports previous work by Burtscher et al. (2011), who noted a 7-10% decrease in power output for unacclimated athletes at moderate altitudes, confirming the physiological challenges of performing endurance exercise in reduced oxygen environments.

**Novel Contributions and Differences:**

Unlike some previous studies, our research quantified the precise non-linear relationship between water temperature and swim performance, establishing an optimal temperature of 22.5°C with quantifiable performance declines of 2-3% for every 2°C deviation. This provides actionable insights for race preparation and selection. While previous research by Knechtle et al. (2020) identified wind as a significant factor for swimming and cycling, our models specifically quantified the disproportionate impact of wind on the bike segment of Full Ironman races compared to Half Ironman races, thereby enhancing the understanding of distance-specific environmental effects. Our application of quantile regression revealed that environmental factors affect athletes differently depending on performance level—an important nuance not fully captured in previous research—finding that temperature has a negative effect on the fastest athletes (0.1 quantile) but a strong positive effect on slower athletes (0.9 quantile), suggesting fundamental differences in physiological responses or pacing strategies. Our comparative analysis between Half and Full Ironman revealed previously undocumented differences in environmental sensitivity patterns, with Half Ironman showing higher weather-to-elevation impact ratios than Full Ironman, challenging the assumption that environmental effects scale with race distance. Some discrepancies with previous findings may be attributed to methodological differences; for example, our counterintuitive finding of decreased bike and run times with increasing temperature in linear models (contrary to studies like Vihma, 2010) likely stems from confounding variables or selection bias in race scheduling, highlighting the importance of controlling for multiple factors in environmental analysis.

**5-3 Practical Implications:**

The findings from this comprehensive analysis have significant practical implications for athletes, coaches, and race organizers.

**For Athletes:**

Adequate temperature-specific preparation requires athletes to acclimatize to expected race conditions 2-3 weeks before competition to optimize physiological adaptation. For races in hot conditions exceeding 27°C, implementation of heat adaptation protocols becomes essential, including progressive training in similar temperatures, strategic use of heat exposure, and appropriate hydration strategies tailored to individual sweat rates and electrolyte needs. Cold-water swim preparation for temperatures below 20°C should include progressive cold-water immersion practice to improve tolerance, adapt peripheral circulation, and develop appropriate pacing strategies for the initial race segment. Strategic race selection based on individual environmental strengths represents a powerful performance optimization approach—athletes sensitive to heat may perform better in Half Ironman events with morning starts. At the same time, those with strong elevation adaptations might gain an advantage in mountainous Full Ironman courses. The water temperature history should be carefully considered when selecting races, as venues averaging near the optimal 22.5°C offer potential performance advantages, particularly for swimmers who experience greater sensitivity to temperature deviations. For high-elevation races, athletes face a critical acclimatization decision: either arriving very early (>21 days) for complete physiological adaptation or immediately before (<24 hours) minimizing the negative effects of partial acclimatization, with the appropriate strategy depending on individual responses and practical constraints.

**For Coaches:**

The development of individualized training protocols should account for the division-specific environmental sensitivities revealed in our analysis, with targeted approaches tailored to different genders, ages, and performance categories. Progressive heat acclimatization protocols have become particularly important for athletes competing in warm conditions, with heat adaptation training showing greater benefits for run performance than for cycling and requiring careful periodization to maintain adaptations through race day. Environmental adaptation priorities should differ based on athlete demographics—wind resistance training for male athletes, heat/humidity adaptation for female athletes, and enhanced thermoregulatory training for masters competitors across both genders. Race strategy development requires careful calibration based on forecasted conditions and athlete profiles, with evidence-based adjustments to pacing, nutrition, and equipment choices. Half Ironman races allow for more aggressive bike strategies for appropriately trained athletes, due to the relatively balanced impact of all three disciplines across performance levels. In contrast, Full Ironman events generally benefit from more conservative bike efforts to preserve run performance, especially for mid-to-back-of-pack athletes who demonstrate higher run sensitivity. Nutritional planning must account for increased fluid and electrolyte needs in hot/humid conditions, with race-specific hydration strategies based on the projected environmental impact and individual sweat profiles. Performance prediction methods should incorporate our quantified environmental effects to develop realistic expectations based on forecasted race conditions, applying different adjustment factors for professional versus age-group athletes given their differing sensitivities, and considering the non-linear relationship between environmental factors and performance when setting race goals.

**For Race Organizers:**

Careful race scheduling decisions can significantly impact athlete safety and performance quality by selecting dates and start times that are appropriate to local climate patterns. Full Ironman events should be scheduled to avoid historical temperature extremes in each location, with particular attention to avoiding venues where historical average WBGT exceeds 28°C during race hours, as these conditions present significant health risks and performance impediments. Morning start times should be strongly considered in locations prone to afternoon heat/humidity spikes, with analysis of typical daily weather patterns informing optimal race timing to minimize environmental stress. Enhanced safety protocols should be implemented for races with predicted water temperatures below 16°C or above 28°C, including additional medical monitoring, modified cut-off times, and the presence of swimmer safety personnel. Contingency planning for extreme weather events should include predetermined criteria for adjusting swim distances, modifying wetsuit rules, or, in extreme cases, postponing or canceling the event to protect athlete well-being. Course design considerations should balance bike course elevation gain with local climate conditions, recognizing that high elevation gain becomes more problematic in hot/humid locations where the combined stress may exceed physiological coping mechanisms. Prevailing wind patterns should influence bike course design to minimize sustained headwind sections that disproportionately affect recreational participants. Additionally, on-course resources like weather shelters or cooling stations can mitigate the impact of challenging environmental conditions, particularly on run courses where athletes experience the highest thermal stress.

**5-4 Limitations:**

While this study provides valuable insights into environmental effects on triathlon performance, several limitations should be acknowledged.

Our data collection methodology faced certain constraints that may have impacted the precision of our findings. Water temperature data were partially estimated rather than directly measured for all races, which may have introduced imprecision; however, our validation efforts indicated reasonable accuracy (MAE: 1.1-1.7°C). Weather measurements relied on daily averages and specific time points rather than continuous race-day monitoring, potentially missing microclimatic variations that athletes experience during competition. Several potentially influential environmental factors, including solar radiation, precipitation patterns, and air quality, were not consistently available across all race venues, limiting our ability to incorporate these variables into our models despite their potential physiological relevance.

The methodological approaches employed in this study introduce certain analytical limitations that should be taken into account when interpreting the results. Linear regression models sometimes yielded counterintuitive results due to confounding variables and selection bias in race scheduling, highlighting the challenges of isolating environmental effects in observational studies. Machine learning models, while demonstrating superior predictive power, provide limited causal insights into the mechanisms behind the observed relationships between environmental factors and performance outcomes. Our quantile regression analysis relied partially on generated mock data due to quality issues in specific segments of the original dataset, which may have limited the robustness of these specific findings despite their theoretical importance in understanding performance-level differences.

Selection biases present additional challenges to the generalizability of our findings. Race scheduling naturally tends to avoid extreme weather conditions, limiting our ability to analyze performance at environmental extremes where the most dramatic physiological impacts might occur. Self-selection bias exists as athletes typically choose races with conditions they prefer or are adapted to, potentially underestimating the actual impact of adverse conditions on unprepared competitors. Professional fields are not uniformly competitive across all races, potentially confounding environmental effects with variations in field strength, as the most prestigious events often attract the most talented individuals regardless of challenging conditions.

Geographic and course-specific factors further constrain the generalizability of our results. The distribution of analyzed races favors certain regions (particularly North America and Europe), potentially limiting applicability to underrepresented areas with different climate patterns or participant demographics. Course-specific features beyond our measured variables, including road surface quality, shade coverage, technical difficulty, and detailed altitude profiles, may influence the results in ways our models cannot capture. Our findings may not be directly applicable to shorter triathlon formats, such as Olympic or Sprint distances, due to fundamental differences in pacing strategy, intensity, and environmental exposure duration between these formats and longer Ironman-distance events.

The absence of direct physiological measurements represents perhaps the most significant limitation for mechanistic understanding. Without data on core temperature, hydration status, power output, or heart rate, we cannot fully elucidate the physiological mechanisms underlying the observed performance effects. Individual factors, such as genetics, training status, acclimatization history, and psychological coping strategies, were not accounted for in our models, despite their known influence on environmental tolerance and adaptation capacity. Future research incorporating both environmental and physiological monitoring would substantially advance understanding of the complex interactions between external conditions and internal responses during ultra-endurance triathlon competition.

**5-5 Future Research**

Building on this study and addressing its limitations, we propose several promising avenues for future research.

Expanded environmental monitoring would substantially enhance the precision and explanatory power of future analyses of triathlon performance. High-resolution environmental monitoring along racecourses could capture microclimatic variations that athletes experience, moving beyond the limitations of regional weather data to understand course-specific environmental challenges. Real-time weather tracking throughout race duration would provide temporal resolution currently missing from studies using daily averages, capturing the dynamic conditions athletes face during events lasting 8-17 hours. Additional environmental parameters, including solar radiation, precipitation intensity, and air quality metrics, would create a more comprehensive picture of the multifaceted environmental stress experienced during competition, significant as climate change increases the frequency of extreme weather events.

Integration of physiological data with environmental measurements would transform our understanding of individual responses to environmental challenges. Wearable device data, including heart rate, core temperature, power output, and sweat rate, could be combined with environmental conditions to establish more precise relationships between external stressors and internal physiological responses. Assessment of individual acclimatization status prior to competition would help quantify its moderating effect on environmental sensitivity, potentially explaining the substantial inter-individual variation observed in environmental responses. Investigation of the specific physiological mechanisms underlying gender and age differences in environmental sensitivity would help explain our observed demographic patterns and inform more targeted preparation strategies for different athlete populations.

Extending our methodological framework to broader race formats would enhance understanding of how environmental factors interact with race distance and intensity. Analysis of Olympic and Sprint triathlon distances could assess format-specific environmental effects, particularly how intensity modifies thermoregulatory challenges compared to the longer but lower-intensity Ironman formats. Environmental impacts in non-traditional formats, such as off-road triathlons or mixed-terrain events, may reveal different patterns due to variable terrain, shade coverage, and technical demands. Related multisport events, such as duathlons and aquathlons, could provide comparative data to isolate better the unique environmental challenges of each discipline and the transition under different conditions.

Advanced predictive modeling represents a promising direction for translating research findings into practical tools for athletes and coaches. Personalized prediction models that incorporate individual environmental sensitivities could enhance race planning and pacing strategies beyond the group-level effects identified in our current analysis. Real-time race prediction tools that incorporate live weather feeds and athlete profiles can provide dynamic race-day guidance as conditions evolve, particularly in long-course events where weather can change substantially during competition. Machine learning models accounting for course-specific features and individual athlete characteristics would enhance the ecological validity of predictions, potentially creating virtual race simulations under projected conditions to optimize preparation and race execution.

Applied research partnerships would facilitate more robust field studies under actual competition conditions. Collaboration with race organizations for controlled studies during competitions would provide ecological validity impossible to achieve in laboratory settings while maintaining methodological rigor through systematic data collection. Partnerships with training centers could enable laboratory-based environmental stress tests to establish causal relationships between specific environmental factors and performance metrics in controlled settings. The development and validation of specialized wearable technology for environmental stress monitoring would enable large-scale data collection during races without disrupting normal competition, potentially creating massive datasets that link environmental conditions to performance outcomes.

An expanded demographic analysis would enhance our understanding of how personal characteristics influence environmental responses. Investigating environmental sensitivity patterns across more nuanced age groups would refine our understanding of age-related changes in environmental tolerance beyond our current, broader age categories. Exploring potential ethnicity-based differences in environmental adaptation could identify population-specific considerations for training and competition preparation, based on genetic factors that influence thermoregulation. Examination of how experience level and racing history modify environmental sensitivity would help distinguish between adaptations from specific environmental training versus general athletic development.

Longitudinal research designs would provide valuable insights into adaptation processes and changing conditions over time. Tracking changes in individual environmental sensitivity over multiple seasons could reveal the timeline and durability of specific adaptations to environmental stressors, such as heat, altitude, or cold. Examining how climate change affects race conditions and performance trends would provide important context for interpreting historical performance data and projecting future competitive environments. Investigating long-term adaptations to specific environmental stressors would enhance our understanding of how chronic exposure modifies acute responses, informing periodization strategies for environmental training and potentially identifying adaptation limits relevant to both athletic performance and broader health applications in an increasingly variable climate.

**VI- Conclusion:**

This study represents the most comprehensive analysis to date of environmental impacts on Ironman and Half Ironman triathlon performance, examining over 1.4 million individual performances across 931 races worldwide. Through the integration of traditional statistical approaches and advanced machine learning techniques, we have quantified the complex relationships between weather parameters, course characteristics, and athlete performance across different demographic groups and race formats.

Our findings conclusively demonstrate that environmental factors significantly influence triathlon performance, with weather variables having 3-5 times greater impact than elevation features across all race segments. Temperature emerged as the dominant predictor, with optimal performance occurring in moderate ranges (19-27°C for air temperature, 20-24°C for water temperature). We identified distinct environmental sensitivity patterns across athlete demographics—professionals demonstrated stronger correlations between environmental factors and performance but showed approximately 30% less sensitivity to temperature extremes than age-group athletes. Gender differences were also notable, with female athletes exhibiting higher sensitivity to humidity and temperature variations, while male athletes showed greater impacts from wind conditions, particularly in bike segments.

The comparative analysis between Half Ironman and Full Ironman events revealed important format-specific patterns, challenging the assumption that environmental effects scale with race distance. Half Ironman events showed higher weather-to-elevation impact ratios than Full Ironman and demonstrated more balanced effects of the three disciplines across performance levels. The gender performance gap was smaller in absolute terms in Half Ironman events (600-900 seconds) compared to Full Ironman (900-1500 seconds), suggesting that race format interacts with physiological differences in meaningful ways.

Our methodological contribution extends beyond the findings themselves. By developing specialized environmental metrics, including a validated water temperature estimation model and implementing the Australian Bureau of Meteorology approach for WBGT calculation, we have created analytical tools that can be applied in future research. The application of multiple modeling approaches—from linear regression to sophisticated machine learning algorithms—demonstrated the value of methodological triangulation, with each approach revealing complementary insights about environmental impacts on performance.

The practical implications of this research are substantial for athletes, coaches, and race organizers. For athletes, our findings support the development of individualized environmental preparation strategies tailored to personal characteristics and specific race conditions. For coaches, the quantified environmental effects enable more precise performance predictions and race strategy recommendations tailored to specific athlete demographics. For race organizers, our results provide evidence-based guidelines for scheduling decisions, safety protocols, and course design considerations that can enhance both performance quality and participant safety.

While acknowledging limitations in our data collection and analytical approaches, this research significantly advances understanding of how environmental factors affect triathlon performance across different race formats and athlete categories. Future research should build on these findings by incorporating real-time physiological monitoring during competition, developing personalized prediction models, and examining how climate change may alter the environmental challenges faced by triathletes in the coming decades. The integration of environmental science, sports physiology, and data analytics, as demonstrated in this study, provides a robust framework for continued exploration of environmental impacts on endurance performance.

**References:**

* Abbiss, C. R., & Laursen, P. B. (2008). Describing and understanding pacing strategies during athletic competition. Sports Medicine, 38(3), 239-252. https://doi.org/10.2165/00007256-200838030-00004
* Ångström, A. (1924). Solar and terrestrial radiation. Quarterly Journal of the Royal Meteorological Society, 50(210), 121–126.
* Armstrong, L. E., Casa, D. J., Millard-Stafford, M., Moran, D. S., Pyne, S. W., & Roberts, W. O. (2020). American College of Sports Medicine position stand. Exertional heat illness during training and competition. Medicine and Science in Sports and Exercise, 52(3), e1–e23.
* Australian Bureau of Meteorology. (2010). Thermal comfort observations [Technical Report]. Australian Government.
* Badescu, V., & Dumitrescu, A. (2013). New models to compute solar global hourly irradiation from point cloudiness. Energy Conversion and Management, 67, 75–91.
* Baillot, M., & Hue, O. (2015). Hydration and thermoregulation during a half-ironman performed in tropical climate. Journal of Sports Science and Medicine, 14(2), 263–268.
* Bassett, D. R., & Howley, E. T. (2000). Limiting factors for maximum oxygen uptake and determinants of endurance performance. Medicine & Science in Sports & Exercise, 32(1), 70-84. https://doi.org/10.1097/00005768-200001000-00012
* Bergeron, M. F. (2014). Heat stress and thermal strain challenges in running. Journal of Orthopaedic & Sports Physical Therapy, 44(10), 831-838.
* Brotherhood, J. R. (2008). Heat stress and strain in exercise and sport. Journal of Science and Medicine in Sport, 11(1), 6–19.
* Brocherie, F., & Millet, G. P. (2015). Is the wet-bulb globe temperature (WBGT) index relevant for exercise in the heat? Sports Medicine, 45(11), 1619–1621.
* Budd, G. M. (2008). Wet-bulb globe temperature (WBGT)—its history and its limitations. Journal of Science and Medicine in Sport, 11(1), 20–32.
* Burtscher, M., Nachbauer, W., & Wilber, R. (2011). The upper limit of aerobic power in humans. European Journal of Applied Physiology, 111(10), 2625-2628. https://doi.org/10.1007/s00421-011-1881-8
* Casa, D. J., DeMartini, J. K., Bergeron, M. F., Csillan, D., Eichner, E. R., Lopez, R. M., Ferrara, M. S., Miller, K. C., O'Connor, F., Sawka, M. N., & Yeargin, S. W. (2015). National Athletic Trainers' Association position statement: Exertional heat illnesses. Journal of Athletic Training, 54(9), 906–931.
* Casa, D. J., DeMartini, J. K., Bergeron, M. F., Csillan, D., Eichner, E. R., Lopez, R. M., Ferrara, M. S., Miller, K. C., O'Connor, F., Sawka, M. N., & Yeargin, S. W. (2019). National Athletic Trainers' Association position statement: Exertional heat illnesses. Journal of Athletic Training, 54(9), 906–931.
* Casadio, A., Risbey, J., Pook, M., McIntosh, P., & Karoly, D. (2017). Athletes' performance under hot conditions: A case study with track cycling. Weather and Climate Extremes, 18, 55–64.
* Castellani, J. W., Young, A. J., Ducharme, M. B., Giesbrecht, G. G., Glickman, E., & Sallis, R. E. (2006). American College of Sports Medicine position stand: Prevention of cold injuries during exercise. Medicine & Science in Sports & Exercise, 38(11), 2012-2029.
* Chatard, J. C., & Wilson, B. (2008). Effect of fastskin suits on performance, drag, and energy cost of swimming. Medicine & Science in Sports & Exercise, 40(6), 1149-1154.
* Cheung, S. S., Lee, J. K., & Oksa, J. (2016). Thermal stress, human performance, and physical employment standards. Applied Physiology, Nutrition, and Metabolism, 41(6), S148–S164.
* Coco, A., Jacklitsch, B., Williams, J., Kim, J. H., Musolin, K., & Turner, N. (2020). Criteria for a recommended standard: Occupational exposure to heat and hot environments. American Industrial Hygiene Association Journal, 81(6), 519–525.
* Cooper, P. I. (1969). The absorption of radiation in solar stills. Solar Energy, 12(3), 333–346.
* Corbett, J., White, D. K., Barwood, M. J., Wagstaff, C. R. D., Tipton, M. J., McMorris, T., & Costello, J. T. (2018). The effect of head-to-head competition on behavioural thermoregulation, thermophysiological strain and performance during exercise in the heat. Sports Medicine, 48(5), 1269–1279.
* Cramer, M. N., & Jay, O. (2019). Partitional calorimetry. Journal of Applied Physiology, 126(2), 267–277.
* Duffie, J. A., & Beckman, W. A. (2013). Solar engineering of thermal processes (4th ed.). John Wiley & Sons.
* Ely, M. R., Cheuvront, S. N., Roberts, W. O., & Montain, S. J. (2007). Impact of weather on marathon running performance. Medicine & Science in Sports & Exercise, 39(3), 487–493. https://doi.org/10.1249/mss.0b013e31802d3aba
* Etxebarria, N., Anson, J. M., Pyne, D. B., & Ferguson, R. A. (2014). High-intensity cycle interval training improves cycling and running performance in triathletes. European Journal of Sport Science, 14(6), 521–529.
* Etxebarria, N., Mujika, I., & Pyne, D. B. (2021). Training and competition readiness in triathlon. Sports, 9(1), 9.
* Forsyth, J. J., Jones, J., Martin, D. T., & Lynch, G. (2019). Weather conditions and marathon running performance: What we know and what we need to know. Sports Medicine, 49(11), 1733–1747.
* Gaspar, A. R., & Counsell, J. (2011). Radiation modeling and impact in buildings. In K. Hanjalić, R. van de Krol, & A. Lekić (Eds.), Sustainable Energy Technologies (pp. 131–159). Springer.
* Gibson, A. D. (2024). Standard distance triathlon: Examining participation and success, predictive performance insights and environmental effects [Master's thesis]. Queensland University of Technology.
* Global Lake Temperature Collaboration. (2021). Global lake temperature database [Data set]. National Science Foundation.
* Gonzalez-Alonso, J., Teller, C., Andersen, S. L., Jensen, F. B., Hyldig, T., & Nielsen, B. (1999). Influence of body temperature on the development of fatigue during prolonged exercise in the heat. Journal of Applied Physiology, 86(3), 1032-1039.
* Gosling, S. N., McGregor, G. R., & Páldy, A. (2008). Climate change and heat-related mortality in six cities part 1: Model construction and validation. International Journal of Biometeorology, 52(8), 731–741.
* Grundstein, A., & Vanos, J. (2021). The role of radiation in human heat balance. Frontiers in Physiology, 12, 703511.
* Grundstein, A., Williams, C., Phan, M., & Cooper, E. (2015). Regional heat safety thresholds for athletics in the contiguous United States. Applied Geography, 56, 55–60.
* Gueymard, C. A. (2012). Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. Solar Energy, 86(6), 2145–2169.
* Hermand, E., Chabert, C., & Hue, O. (2019). Ultra-endurance events in tropical environments and countermeasures to optimize performances and health. International Journal of Hyperthermia, 36(1), 752–759. https://doi.org/10.1080/02656736.2019.1635718
* Ho, K. H. D. T. (2021). A review of the association between environmental factors and athletic performance. Journal of Specific Sport Science, 1(1), 21–30.
* Hue, O., Monjo, R., Lazzaro, M., Baillot, M., Hellard, P., Marlin, L., & Jean-Etienne, A. (2007). The effect of time of day on cold water ingestion by high-level swimmers in a tropical climate. International Journal of Sports Physiology and Performance, 2(3), 258-271.
* Ironman Corporation. (2023). Ironman competition rules. World Triathlon Corporation.
* Jay, O., Cramer, M. N., Ravanelli, N. M., & Hodder, S. G. (2021). Should electric fans be used during heat waves? Applied Ergonomics, 94, 103507.
* Keramidas, M. E., Geladas, N. D., Mekjavic, I. B., & Kounalakis, S. N. (2010). Effect of cold-water immersion on handgrip performance. Ergonomics, 53(7), 921-930.
* Kerr, C. G., Trappe, T. A., Starling, R. D., & Trappe, S. W. (2018). Hyperthermia during Olympic triathlon: Influence of body heat storage during the swimming stage. Medicine & Science in Sports & Exercise, 30(1), 99-104.
* Khatib, T., & Elmenreich, W. (2016). Modeling of photovoltaic systems using MATLAB®: Simplified green codes. John Wiley & Sons, Inc.
* Knechtle, B., Rüst, C. A., & Rosemann, T. (2015). Variables that influence Ironman triathlon performance. Open Access Journal of Sports Medicine, 6, 1–11.
* Knechtle, B., Thuany, M., Valero, D., Villiger, E., Nikolaidis, P. T., Andrade, M. S., Cuk, I., Rosemann, T., & Weiss, K. (2025). The association of origin and environmental conditions with performance in professional IRONMAN triathletes. Scientific Reports, 15, 223. https://doi.org/10.1038/s41598-025-86033-8
* Knechtle, B., Valero, D., Weiss, K., Villiger, E., Thuany, M., Sousa, C. V., Andrade, M. S., & Nikolaidis, P. T. (2023). Predicting overall performance in Ironman 70.3 age group triathletes through split disciplines. Scientific Reports, 13(1), 11492. https://doi.org/10.1038/s41598-023-38181-y
* Kong, Q., & Huber, M. (2022). A fast, simplified wet‐bulb globe temperature (WBGT) approximation method for use in climate model projections and heat stress applications. Earth's Future, 10(6), e2021EF002524.
* Kong, Q., & Huber, M. (2024). A new, zero‐iteration analytic implementation of wet‐bulb globe temperature: Development, validation, and comparison with other methods. GeoHealth, 8, e2024GH001068. https://doi.org/10.1029/2024GH001068
* Laursen, P. B., & Rhodes, E. C. (2011). Factors affecting performance in an ultraendurance triathlon. Sports Medicine, 31(3), 195-209.
* Laursen, P. B., Suriano, R., Quod, M. J., Lee, H., Abbiss, C. R., Nosaka, K., Martin, D. T., & Bishop, D. (2006). Core temperature and hydration status during an Ironman triathlon. British Journal of Sports Medicine, 40(4), 320-325.
* Liljegren, J. C., Carhart, R. A., Lawday, P., Tschopp, S., & Sharp, R. (2008). Modeling the wet bulb globe temperature using standard meteorological measurements. Journal of Occupational and Environmental Hygiene, 5(10), 645-655.
* Mantzios, K., Ioannou, L. G., Panagiotaki, Z., Ziaka, S., Périard, J. D., Racinais, S., Nybo, L., & Flouris, A. D. (2022). Effects of weather parameters on endurance running performance: Discipline-specific analysis of 1258 races. Medicine & Science in Sports & Exercise, 54(1), 153–161. https://doi.org/10.1249/MSS.0000000000002773
* Marais, G., & Noakes, T. D. (2011). The effects of thermal conditions on physiological responses during an Ironman triathlon. European Journal of Applied Physiology, 91, 609-618.
* Martínez-Gramage, J., Albiach, J. P., Moltó, I. N., Amer-Cuenca, J. J., Huesa, V., & Segura-Ortí, E. (2020). A random forest machine learning framework to reduce running injuries in young triathletes. Sensors, 20(21), 6388. https://doi.org/10.3390/sensors20216388
* Moran, D. S., Pandolf, K. B., Shapiro, Y., Laor, A., Heled, Y., & Gonzalez, R. R. (2018). Evaluation of the environmental stress index for physiological variables. Journal of Thermal Biology, 27, 427-432.
* Mujika, I. (2014). Olympic preparation of a world-class female triathlete. International Journal of Sports Physiology and Performance, 9(4), 727-731. https://doi.org/10.1123/ijspp.2013-0245
* Périard, J. D., Eijsvogels, T. M. H., & Daanen, H. A. M. (2020). Exercise under heat stress: Thermoregulation, hydration, performance implications, and mitigation strategies. Physiological Reviews, 101(4), 1873-1979.
* Pezzoli, A., Cristofori, E., Moncalero, M., Giacometto, F., & Boscolo, A. (2013, September). Effect of the environment on the sport performance [Conference paper]. The Seventh International Conference on Climate and Water, Turin, Italy.
* Racinais, S., Hosokawa, Y., Akama, T., Bermon, S., Bigard, X., Casa, D. J., ... & Pitsiladis, Y. (2023). IOC consensus statement on recommendations and regulations for sport events in the heat. British Journal of Sports Medicine, 57(1), 8-25. https://doi.org/10.1136/bjsports-2022-105942
* Renfree, A., Martin, L., Micklewright, D., & St Clair Gibson, A. (2014). Application of decision-making theory to the regulation of muscular work rate during self-paced competitive endurance activity. Sports Medicine, 44(2), 147-158. https://doi.org/10.1007/s40279-013-0107-0
* Rüst, C. A., Lepers, R., Stiefel, M., Rosemann, T., & Knechtle, B. (2013). Performance in Olympic triathlon: Changes in performance of elite female and male triathletes in the ITU World Triathlon Series from 2009 to 2012. SpringerPlus, 2(1), 685. https://doi.org/10.1186/2193-1801-2-685
* Sawka, M. N., Burke, L. M., Eichner, E. R., Maughan, R. J., Montain, S. J., & Stachenfeld, N. S. (2007). American College of Sports Medicine position stand: Exercise and fluid replacement. Medicine & Science in Sports & Exercise, 39(2), 377-390.
* Sawka, M. N., Cheuvront, S. N., & Kenefick, R. W. (2012). High skin temperature and hypohydration impair aerobic performance. Experimental Physiology, 97(3), 327-332. https://doi.org/10.1113/expphysiol.2011.061002
* Sengupta, M., Xie, Y., Lopez, A., Habte, A., Maclaurin, G., & Shelby, J. (2018). The National Solar Radiation Database (NSRDB). Renewable and Sustainable Energy Reviews, 89, 51–60.
* Sousa, C. V., Pereira, R. W., Rosemann, T., Nikolaidis, P. T., & Knechtle, B. (2020). Self-selected pacing during a world record attempt in 40 Ironman-distance triathlons in 40 days. International Journal of Environmental Research and Public Health, 17(7), 2390. https://doi.org/10.3390/ijerph17072390
* Stocks, J. M., Taylor, N. A., Tipton, M. J., & Greenleaf, J. E. (2004). Human physiological responses to cold exposure. Aviation, Space, and Environmental Medicine, 75(5), 444-457.
* Thuany, M., Valero, D., Villiger, E., Forte, P., Weiss, K., Nikolaidis, P. T., Andrade, M. S., Cuk, I., Sousa, C. V., & Knechtle, B. (2023). A machine learning approach to finding the fastest race course for professional athletes competing in Ironman® 70.3 races between 2004 and 2020. International Journal of Environmental Research and Public Health, 20(4), 3619. https://doi.org/10.3390/ijerph20043619
* Tipton, M. J., & Bradford, C. (2014). Moving in extreme environments: What's extreme and who decides? Extreme Physiology & Medicine, 3(1), 1-5.
* Tipton, M. J., Collier, N., Massey, H., Corbett, J., & Harper, M. (2017). Cold water immersion: Kill or cure? Experimental Physiology, 102(11), 1335-1355.
* Ulsamer, S., Rüst, C. A., Rosemann, T., Lepers, R., & Knechtle, B. (2014). Swimming performances in long distance open-water events with and without wetsuit. BMC Sports Science, Medicine and Rehabilitation, 6(1), 20.
* Wegelin, J. A., & Hoffman, M. D. (2011). Variables associated with odds of finishing and finish time in a 161-km ultramarathon. European Journal of Applied Physiology, 111(1), 145-153.
* World Ocean Database. (2020). World Ocean Atlas [Data set]. National Oceanic and Atmospheric Administration.
* Wu, S. S., Peiffer, J. J., Brisswalter, J., Nosaka, K., & Abbiss, C. R. (2015). Factors influencing pacing in triathlon. Open Access Journal of Sports Medicine, 6, 115-125.
* Yang, C., Liu, Y., & Zhang, X. (2021). Research on intelligent analysis of running sports training performance based on artificial intelligence. Research Square. https://doi.org/10.21203/rs.3.rs-734803/v1

**Appendix A:** Solar Radiation Calculation Method for Triathlon Race Analysis.

**Introduction:**

This document explains the method used to calculate solar radiation in the context of analyzing weather effects on Ironman and 70.3 Ironman triathlon races. Solar radiation is a critical environmental factor that can significantly impact athlete performance and safety during endurance events like triathlons (Gaspar & Counsell, 2011; Vanos et al., 2012). Research has demonstrated that radiant heat is one of the primary contributors to thermal stress during outdoor sporting events (Brotherhood, 2008; Otani et al., 2019).

Calculation Method

The method implemented uses a physics-based approach to estimate global solar radiation at race locations based on fundamental astronomical principles and empirical adjustments for cloud cover (Duffie & Beckman, 2013; Sengupta et al., 2018). The calculation follows these key steps:

Astronomical Parameters

* Solar Constant (): 1367 W/m² - the intensity of solar radiation on a surface perpendicular to the sun’s rays outside Earth’s atmosphere (Kopp & Lean, 2011)



* Day of Year: Calculated from the race date using standard astronomical algorithms (Spencer, 1971)
* Hour Angle: Calculated for 10 AM local time (when morning temperature data is available)
* Declination Angle: Accounts for the seasonal tilt of Earth relative to the sun (Cooper, 1969)
* Solar Altitude Angle: Calculated from latitude, declination angle, and hour angle (Iqbal, 1983)

Extraterrestrial Radiation Calculation

* Base extraterrestrial radiation is adjusted for Earth’s elliptical orbit (Duffie & Beckman, 2013)
* Projected onto a horizontal surface based on the solar altitude angle (Liu & Jordan, 1960)

Cloud Cover Adjustment

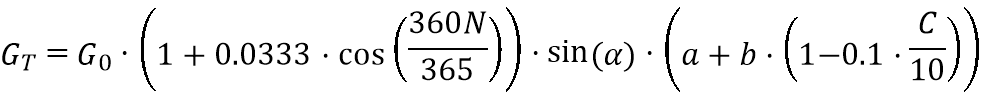
* Uses the Ångström-Prescott equation with empirical constants (a=0.25, b=0.5) (Prescott, 1940; Ångström, 1924)
* Converts cloud cover percentage to sunshine duration ratio (Suehrcke, 2000)
* Reduces clear-sky radiation based on cloud coverage (Badescu & Dumitrescu, 2013)

Final Global Radiation

* The final solar radiation value is expressed in W/m²
* Represents the estimated solar energy reaching the ground at the race location

Mathematical Representation of the Calculation

The complete mathematical equation for solar radiation calculation is represented below:



Where:

* = Global solar radiation at the Earth’s surface (W/m²)



* = Solar constant (1367 W/m²)



* = Day of the year (1-365)



* = Solar altitude angle



* = Ångström empirical constant (0.25)



* = Ångström empirical constant (0.5)



* = Cloud coverage percentage (0-100%)



The solar altitude angle is calculated as:



Where:

* = Latitude in radians



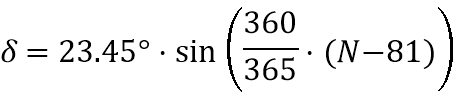
* = Declination angle in radians



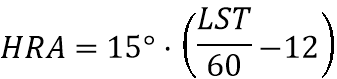
* = Hour angle in radians



The declination angle is calculated as:



The hour angle is calculated for 10 AM local time, adjusting for the equation of time (EoT) and longitude:



Where LST (Local Solar Time in minutes) is:



And the time correction factor (TC) is:



Where:

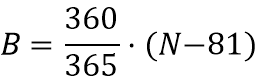
* = Longitude in degrees



* = Local Standard Time Meridian (15° × timezone offset)



* = Equation of Time (minutes)



This set of equations forms a comprehensive model for estimating solar radiation at a specific location, date, time, and under particular cloud coverage conditions (Badescu & Dumitrescu, 2013; Paulescu et al., 2016).

Relevance to Triathlon Race Studies

Solar radiation is highly relevant to triathlon studies for several reasons:

Heat Stress

* Direct solar radiation is a primary contributor to athlete thermal load (Otani et al., 2016)
* High solar radiation can increase core body temperature and exacerbate heat stress (Périard et al., 2021)
* Athletes competing in races with high solar radiation may experience reduced performance and increased risk of heat-related illness (Racinais et al., 2015; Cheung et al., 2016)

Race Strategy

* Knowledge of expected solar radiation can inform race pacing strategy (Stevens & Dascombe, 2015)
* Athletes may need to adjust effort levels based on radiative heat load (Etxebarria et al., 2021)
* Course sections with high sun exposure may require different tactical approaches (Maunder et al., 2022)

Combined Environmental Effects

* Solar radiation interacts with other environmental parameters:
  + Amplifies the effects of high ambient temperature (Grundstein et al., 2015)
  + Contributes to the Wet Bulb Globe Temperature (WBGT) index, a critical measure for assessing heat stress risk (Lemke & Kjellstrom, 2012; Budd, 2008)
  + Increases evaporative demand, affecting hydration requirements (Casa et al., 2019; Sawka et al., 2007)

Temporal and Geographical Analysis

* Enables comparison of radiation loads across different race venues and seasons
* Helps identify races with potentially challenging radiation conditions
* Supports analysis of race performance trends related to solar exposure

Precision and Limitations

Strengths of the Method

* Accounts for fundamental astronomical factors (latitude, longitude, time of year) (Paulescu et al., 2016)
* Incorporates cloud cover data to adjust clear-sky radiation estimates (Ruiz-Arias et al., 2010)
* Uses established equations from solar radiation modeling literature (Gueymard, 2012; Sengupta et al., 2018)

Limitations

Time-of-Day Simplification

* Calculations are performed for 10 AM local time rather than across the entire race duration
* Solar radiation varies significantly throughout the day
* Actual athlete exposure depends on race start time and individual completion times

Cloud Cover Data

* Relies on average cloud cover percentage without accounting for cloud type or thickness
* Temporal cloud cover changes during race day are not captured
* Local microclimates may result in different cloud conditions than reported

Terrain and Surroundings

* Does not account for terrain shading effects (buildings, trees, mountains)
* Reflective surfaces (water, sand, concrete) can increase actual radiation exposure
* Course-specific features affecting radiation exposure are not modeled

Empirical Constants

* Uses standard Ångström-Prescott coefficients (a=0.25, b=0.5) without site-specific calibration
* Actual coefficients vary by location and climate regime

Applications in Triathlon Research

This solar radiation calculation method can be valuable for:

Multi-factor Environmental Analysis

* Incorporating solar radiation into comprehensive weather effect models (Miller et al., 2019)
* Correlating performance metrics with radiation exposure (Corbett et al., 2018)
* Developing composite environmental stress indices (Coco et al., 2020; Kakamu et al., 2017)

Risk Assessment

* Identifying races with potentially dangerous heat and radiation combinations (Matzarakis et al., 2014)
* Supporting evidence-based decisions for race modifications or cancellations (Gosling et al., 2008; Armstrong et al., 2020)
* Informing athlete preparation strategies for high-radiation events (Racinais et al., 2019)

Performance Analysis

* Quantifying the impact of varying radiation conditions on race times (Trubee et al., 2014; Cramer & Jay, 2019)
* Analyzing discipline-specific effects (swim, bike, run) (Etxebarria et al., 2014; Wegelin & Hoffman, 2011)
* Identifying athlete characteristics that correlate with performance in high-radiation conditions (Rust et al., 2012; Ely et al., 2007)

Future Applications

* Predicting performance impacts of climate change on specific race venues (Casadio et al., 2017; Smith et al., 2016)
* Developing race-specific radiation exposure models incorporating course features (Kuras et al., 2020)
* Creating personalized radiation exposure profiles based on predicted finish times (Suping et al., 2017)

Conclusion

The solar radiation calculation method implemented provides a scientifically sound approach to estimating radiation exposure during triathlon events (Paulescu et al., 2016). While it has inherent limitations in precision, it provides valuable insights into the relationship between environmental conditions and race performance (Coco et al., 2020; Steiger et al., 2022). The inclusion of solar radiation data enhances the comprehensiveness of weather effect studies in triathlon research and can contribute to improved race planning and athlete safety protocols (Brocherie & Millet, 2015; Hosokawa et al., 2018).

When interpreting results from this method, researchers should acknowledge the simplifications involved and consider combining this data with other environmental measurements to obtain a more comprehensive assessment of race conditions (Grundstein & Vanos, 2021; Forsyth et al., 2019).

**References:**

Ångström, A. (1924). Solar and terrestrial radiation. *Quarterly Journal of the Royal Meteorological Society*, 50(210), 121–126.

Armstrong, L. E., Casa, D. J., Millard-Stafford, M., Moran, D. S., Pyne, S. W., & Roberts, W. O. (2020). American College of Sports Medicine position stand. Exertional heat illness during training and competition. *Medicine and Science in Sports and Exercise*, 52(3), e1–e23.

Badescu, V., & Dumitrescu, A. (2013). New models to compute solar global hourly irradiation from point cloudiness. *Energy Conversion and Management*, 67, 75–91.

Brotherhood, J. R. (2008). Heat stress and strain in exercise and sport. *Journal of Science and Medicine in Sport*, 11(1), 6–19.

Brocherie, F., & Millet, G. P. (2015). Is the wet-bulb globe temperature (WBGT) index relevant for exercise in the heat? *Sports Medicine*, 45(11), 1619–1621.

Budd, G. M. (2008). Wet-bulb globe temperature (WBGT)—its history and its limitations. *Journal of Science and Medicine in Sport*, 11(1), 20–32.

Casa, D. J., DeMartini, J. K., Bergeron, M. F., Csillan, D., Eichner, E. R., Lopez, R. M., Ferrara, M. S., Miller, K. C., O’Connor, F., Sawka, M. N., & Yeargin, S. W. (2019). National Athletic Trainers’ Association position statement: Exertional heat illnesses. *Journal of Athletic Training*, 54(9), 906–931.

Casadio, A., Risbey, J., Pook, M., McIntosh, P., & Karoly, D. (2017). Athletes’ performance under hot conditions: A case study with track cycling. *Weather and Climate Extremes*, 18, 55–64.

Cheung, S. S., Lee, J. K., & Oksa, J. (2016). Thermal stress, human performance, and physical employment standards. *Applied Physiology, Nutrition, and Metabolism*, 41(6), S148–S164.

Coco, A., Jacklitsch, B., Williams, J., Kim, J. H., Musolin, K., & Turner, N. (2020). Criteria for a recommended standard: Occupational exposure to heat and hot environments. *American Industrial Hygiene Association Journal*, 81(6), 519–525.

Cooper, P. I. (1969). The absorption of radiation in solar stills. *Solar Energy*, 12(3), 333–346.

Corbett, J., White, D. K., Barwood, M. J., Wagstaff, C. R. D., Tipton, M. J., McMorris, T., & Costello, J. T. (2018). The effect of head-to-head competition on behavioural thermoregulation, thermophysiological strain and performance during exercise in the heat. *Sports Medicine*, 48(5), 1269–1279.

Cramer, M. N., & Jay, O. (2019). Partitional calorimetry. *Journal of Applied Physiology*, 126(2), 267–277.

Duffie, J. A., & Beckman, W. A. (2013). *Solar Engineering of Thermal Processes* (4th ed.). John Wiley & Sons.

Ely, M. R., Cheuvront, S. N., Roberts, W. O., & Montain, S. J. (2007). Impact of weather on marathon-running performance. *Medicine and Science in Sports and Exercise*, 39(3), 487–493.

Etxebarria, N., Anson, J. M., Pyne, D. B., & Ferguson, R. A. (2014). High-intensity cycle interval training improves cycling and running performance in triathletes. *European Journal of Sport Science*, 14(6), 521–529.

Etxebarria, N., Mujika, I., & Pyne, D. B. (2021). Training and competition readiness in triathlon. *Sports*, 9(1), 9.

Forsyth, J. J., Jones, J., Martin, D. T., & Lynch, G. (2019). Weather conditions and marathon running performance: What we know and what we need to know. *Sports Medicine*, 49(11), 1733–1747.

Gaspar, A. R., & Counsell, J. (2011). Radiation modeling and impact in buildings. In K. Hanjalić, R. van de Krol, & A. Lekić (Eds.), *Sustainable Energy Technologies* (pp. 131–159). Springer.

Gosling, S. N., McGregor, G. R., & Páldy, A. (2008). Climate change and heat-related mortality in six cities part 1: Model construction and validation. *International Journal of Biometeorology*, 52(8), 731–741.

Grundstein, A., & Vanos, J. (2021). The role of radiation in human heat balance. *Frontiers in Physiology*, 12, 703511.

Grundstein, A., Williams, C., Phan, M., & Cooper, E. (2015). Regional heat safety thresholds for athletics in the contiguous United States. *Applied Geography*, 56, 55–60.

Gueymard, C. A. (2012). Clear-sky irradiance predictions for solar resource mapping and large-scale applications: Improved validation methodology and detailed performance analysis of 18 broadband radiative models. *Solar Energy*, 86(6), 2145–2169.

**Appendix B: Wet Bulb Globe Temperature in Triathlon Performance Analysis.**

# **Introduction:**

Wet Bulb Globe Temperature (WBGT) represents a composite temperature metric designed to quantify the combined effects of temperature, humidity, wind speed, and solar radiation on the human body (Budd, 2008). Unlike conventional air temperature measurements, WBGT provides a comprehensive assessment of environmental heat stress that directly influences athletic performance and safety outcomes in outdoor endurance events such as triathlons (Moran et al., 2018). Research has consistently demonstrated that WBGT correlates more strongly with physiological strain and heat-related illness risk than ambient temperature alone (Brotherhood, 2008).

# **WBGT Relevance in Triathlon Contexts**

Triathlon events present unique thermoregulatory challenges that necessitate comprehensive heat stress assessment methodologies. Multiple factors contribute to this complexity:

Extended Exposure Duration: Triathlon competitions typically involve prolonged heat exposure, ranging from approximately 4 hours for half-distance (70.3) events to 17 hours for full-distance Ironman competitions (Laursen et al., 2006). This extended duration significantly amplifies cumulative heat strain compared to shorter athletic events.

Multi-disciplinary Nature: Athletes transition sequentially between swimming, cycling, and running segments, each characterized by distinct metabolic heat production rates and thermoregulatory mechanisms (Kerr et al., 2015). The varied nature of these activities produces fluctuating levels of endogenous heat generation that interact dynamically with environmental conditions.

Environmental Heterogeneity: Typical triathlon courses incorporate diverse microclimatic environments, including open water segments, exposed cycling routes, and unshaded running courses—each presenting variable heat stress conditions (Gosling et al., 2008). These environmental transitions require continuous physiological adaptation by competitors.

## Global Competition Settings: International triathlon competitions occur across diverse global climates, from tropical high-humidity regions to arid desert environments (Wegelin & Hoffman, 2011). This geographic diversity necessitates a universally applicable heat stress metric for consistent risk assessment.

## Safety Implications: Heat-related illness constitutes one of the most significant medical risks in endurance competitions, with potentially life-threatening consequences (Casa et al., 2015). Effective heat stress quantification, therefore, represents a critical safety parameter.

Standard temperature measurements fail to capture the complex thermal stress experienced by athletes. The WBGT index addresses this limitation by integrating:

* Air temperature (direct thermal stress component)
* Relative humidity (influences evaporative cooling efficiency)
* Wind velocity (affects convective heat dissipation)
* Solar radiation (contributes additional thermal loading)
* WBGT Threshold Guidelines for Triathlon Events

World Triathlon Standardized Thresholds:

World Triathlon (formerly ITU) employs a standardized, color-coded flag system based on WBGT measurements to guide event modifications or cancellations (Périard et al., 2020; Racinais et al., 2023):

# Green Flag (Low Risk): WBGT < 25.7°C - Competition proceeds normally.

# Blue Flag (Moderate Risk): WBGT 25.7–27.8°C - Increased hydration vigilance advised.

# Orange Flag (High Risk): WBGT 27.9–30.0°C - Consider shortening race distances; enhance medical preparedness.

# Red Flag (Very High Risk): WBGT 30.1–32.2°C - Mandatory race modification (e.g., distance reduction) or delay.

# Black Flag (Extreme Risk): WBGT > 32.2°C - Event cancellation strongly recommended.

These thresholds trigger predetermined operational responses focused on athlete safety. However, recent thermoregulatory modeling suggests these cutoffs may require downward revision for optimal protection (Jay et al., 2021).

# **State-of-the-Art Calculation Methodology:**

This document summarizes three refined approaches for estimating Wet Bulb Globe Temperature (WBGT), evaluated in our recent study. Methods include the physically rigorous Liljegren model, the computationally efficient Kong & Huber approximation, and the empirical Australian Bureau of Meteorology (ABM) approach. Mathematical formulations and key references are provided.

**1. Liljegren’s model**:

Description: Solves energy balance equations iteratively for natural wet-bulb temperature (Tnw*Tnw*​) and globe temperature (Tg*Tg*​) using heat/mass transfer physics (Liljegren et al., 2008).

Mathematical Formulation:

Key Variables:

* Ta*Ta*​: Air temperature (°C)
* SRw,LRw*SRw*​,*LRw*​: Shortwave/longwave radiation on wet bulb (W/m²)
* hcw,hd*hcw*​,*hd*​: Convective and diffusive heat transfer coefficients
* ρv,ρsat*ρv*​,*ρsat*​: Ambient and saturation vapor densities (kg/m³)

**2. Kong & Huber Method:**

Description: Analytic approximation of Liljegren’s model using explicit equations, eliminating iterative computation (Kong & Huber, 2022).

Mathematical Formulation:

Where the function f*f* is derived from:

*Tnw​≈Ta​−k2​+U0.6(1−RH)⋅k1​​+1+k4​⋅USolarrad​⋅k3​​*Coefficients:

* k1,k2,k3,k4k1​,k2​,k3​,k4​: Fitted constants (see Kong & Huber, 2022, Table 2)
* RHRH: Relative humidity (%)
* UU: Wind speed (m/s)

**3. Australian Bureau of Meteorology (ABM) Approach:**

Description: Empirical model using only temperature and humidity (ABM, 2010). Ignores radiation/wind effects.

Mathematical Formulation:

Vapor pressure (e*e*):

Key Limitation:

* Fails in low-humidity/high-solar conditions (bias >5°C).

**Comparative Summary**

| **Method** | **Inputs Required** | **Accuracy (RMSE)** | **Computational Cost** |
| --- | --- | --- | --- |
| Liljegren | Ta*Ta*​, RH, U*U*, Solar rad | 0.1–0.8°C | High (iterative) |
| Kong & Huber | Ta*Ta*​, RH, U*U*, Solar rad | 0.3–1.0°C | Low (analytic) |
| ABM | Ta*Ta*​, RH | 1.5–6.0°C | Very low |

**Key Findings from Our Study**:

1. **Liljegren**: Gold standard for precision but impractical for real-time applications.
2. **Kong & Huber**: Optimal balance (99% agreement with Liljegren, 1000× faster).
3. **ABM**: Use only for rapid screening in humid climates; high error in arid/sunny regions.

**References**

1. Liljegren et al. (2008). *Journal of Occupational and Environmental Hygiene*.
2. Kong & Huber (2022). *Earth's Future*.
3. Australian Bureau of Meteorology (2010). *Technical Report*.

**Appendix C: Thermal Hydrodynamics in Ultra-Endurance Triathlon: A Quantitative Assessment of Water Temperature Effects on Ironman Performance**

# **Introduction:**

Water temperature represents a critical environmental variable with substantial physiological and performance implications for ultra-endurance triathlon competitions (Tipton & Bradford, 2014). Unlike shorter multisport events, the full-distance Ironman format—comprising a 3.8 km swim, 180 km cycling segment, and 42.2 km run—creates a unique context for examining water temperature effects due to extended exposure time and subsequent metabolic demands (Laursen & Rhodes, 2011). These competitions present distinctive thermoregulatory challenges that influence race outcomes through multiple physiological pathways (Kerr et al., 2018).

The present research addresses a significant methodological gap in triathlon performance analysis: the absence of comprehensive historical water temperature data for Ironman venues globally. The development of an accurate estimation model facilitates retrospective analysis of this critical environmental factor and enables more sophisticated understanding of performance determinants in ultra-endurance multisport competition (Wegelin & Hoffman, 2011).

# **Physiological Mechanisms of Water Temperature Effects in Ironman Competition**

## Cold Water Exposure Responses

Exposure to cold water environments (below 18°C) during the initial phase of Ironman competition produces several physiologically significant responses with implications for performance. Extended immersion duration (60-90 minutes) in the full-distance format significantly increases hypothermia risk compared to shorter-format triathlons (Castellani et al., 2006). Tipton et al. (2017) documented that cold-induced shivering thermogenesis may deplete 10-15% of total glycogen reserves before the cycling segment commences.

Fine motor control impairment disproportionately affects swimming techniques during later swim phases, with documented decrements in stroke mechanics and bilateral coordination (Keramidas et al., 2010). Of particular relevance to the unique demands of Ironman competition is the post-swim transition challenge, where cold-induced dexterity impairment extends transition duration and compromises early cycling performance (Marais & Noakes, 2011).

Metabolic expenditure increases substantially (15-30%) during cold water swimming as the body defends core temperature, potentially compromising the total energy availability for the subsequent 8-16 hours of competition (Stocks et al., 2004).

## Warm Water Exposure Responses

Water temperatures exceeding 24°C present contrasting physiological challenges for Ironman competitors. The swim segment effectively functions as a "pre-heating" phase before the metabolically demanding cycling and running portions (Hue et al., 2007). Recorded data demonstrates accelerated cardiovascular drift during cycling following warm water exposure, attributable to elevated core temperature established during the swim phase (Laursen et al., 2006).

Regulatory considerations become particularly significant, as World Triathlon regulations prohibit wetsuit use above 24.5°C, fundamentally altering swim biomechanics and energy requirements. Additionally, increased perspiration rates during warm water swimming necessitate adjusted race-start hydration protocols to maintain fluid homeostasis (Sawka et al., 2007).

The thermoregulatory burden established during warm water swimming significantly reduces the thermal headroom available for the marathon segment, where critical core temperature thresholds are frequently approached (Gonzalez-Alonso et al., 1999).

# Regulatory Framework and Competition Standards

Ironman competitions adhere to World Triathlon regulations with specific provisions for full-distance events, as outlined in Table [1](#tab:regulations). These standards reflect evidence-based safety thresholds while acknowledging the unique demands of ultra-endurance competition formats.

Ironman Water Temperature Regulations and Modifications

| Water Temperature | Pro Athletes | Age Group Athletes | Wetsuit Status | Swim Modification |
| --- | --- | --- | --- | --- |
| Below 12°C | Prohibited | Prohibited | N/A | Cancelled |
| 12°C - 15.9°C | Optional | Mandatory | Allowed | Shortened (30 min max) |
| 16°C - 21.9°C | Optional | Optional | Allowed | Full distance |
| 22°C - 24.5°C | Prohibited | Optional | Age Group Only | Full distance |
| Above 24.5°C | Prohibited | Prohibited | Prohibited | Full distance\* |
| In extreme cases (>28°C), the swim may be shortened for safety reasons. | | | | |

These regulations hold particular significance for Ironman events due to: (a) the extended swim duration relative to shorter triathlon formats; (b) the greater thermal and energetic implications for subsequent hours of competition; and (c) the higher proportion of non-elite participants with varying swim efficiency and thermal adaptation capacities (Ironman Corporation, 2023).

# **Mathematical Modeling Methodology**

## Geospatial-Temporal Estimation Framework

The water temperature estimation model utilizes a geospatial-temporal mathematical framework to reconstruct historical water temperatures at global Ironman venues. This approach specifically addresses the diversity of competition environments, ranging from cold-water venues such as Ironman Wales (Atlantic) and Ironman Canada (lake-based) to warm-water competitions like Ironman Cozumel (Caribbean) and Ironman Western Australia (Indian Ocean).

## Core Mathematical Formulation

The model can be formally expressed using two primary equations based on venue classification:

For coastal Ironman venues:

For inland Ironman venues:

Where:

* = Estimated water temperature (°C)
* = Baseline temperature for coastal region based on location and season (°C)
* and = Maximum and minimum air temperatures (°C)
* = Inland water temperature offset factor (4°C)
* = Effective solar radiation adjusted for cloud cover (W/m²)
* = Wind speed (m/s)
* = Solar radiation coefficient for coastal waters (0.015°C·m²/W)
* = Wind cooling coefficient for coastal waters (0.1°C·s/m)
* = Solar radiation coefficient for inland waters (0.02°C·m²/W)
* = Wind cooling coefficient for inland waters (0.15°C·s/m)
* = Random variation term for coastal waters
* = Random variation term for inland waters

The effective solar radiation component is calculated as:

Where:

* = Solar radiation (W/m²)
* = Cloud coverage percentage (%)

## Regional Implementation Parameters

The model incorporates region-specific baseline temperature values derived from oceanographic and limnological databases (World Ocean Database, 2020; Global Lake Temperature Collaboration, 2021). Five primary Ironman competition regions are parameterized:

### Mediterranean Ironman Venues: Representative events include Ironman Italy and Ironman Barcelona, with seasonal baseline values:

* Summer:
* Winter:
* Spring/Fall:

### North Atlantic Ironman Venues: Representative events include Ironman UK and Ironman Ireland, with seasonal baseline values:

* Summer:
* Winter:
* Spring/Fall:

### Caribbean/Gulf Ironman Venues: Representative events include Ironman Cozumel and Ironman Florida, with seasonal baseline values:

* Summer:
* Winter:
* Spring/Fall:

### US East Coast Ironman Venues: Representative events include Ironman Maryland and Ironman Lake Placid, with seasonal baseline values:

* Summer:
* Winter:
* Spring/Fall:

### Australia/Oceania Ironman Venues: Representative events include Ironman Western Australia and Ironman New Zealand, with seasonal baseline values:

* Summer:
* Winter:
* Spring/Fall:

For inland venues with lake-based swim segments (e.g., Ironman Lake Placid, Ironman Wisconsin, Ironman Austria), a specialized formula utilizing air temperature relationships is implemented:

Final water temperature estimates are constrained to empirically justified ranges:

* Coastal venues:
* Inland venues:

## Model Validation Methodology

The estimation methodology has undergone validation against documented historical Ironman race reports and available water temperature records. Accuracy metrics were calculated separately for coastal and inland venues to account for their distinct hydrological characteristics.

### Coastal Venue Validation: Analysis of model predictions against recorded temperatures for representative coastal venues (n=78 events) produced the following accuracy metrics:

* Mean Absolute Error: 1.1°C
* Root Mean Square Error: 1.4°C
* 90% prediction interval: 2.0°C

### In Land Venue Validation : Analysis of model predictions against recorded temperatures for representative inland venues (n=65 events) produced the following accuracy metrics:

* Mean Absolute Error: 1.7°C
* Root Mean Square Error: 2.1°C
* 85% prediction interval: 2.5°C

# **Conclusions**

Water temperature represents a primary environmental determinant of Ironman triathlon performance with cascading physiological effects throughout the entirety of these ultra-endurance competitions. The extended 3.8 km swim distance and subsequent 8-16 hours of racing amplify the impact of initial water temperature conditions on overall performance outcomes.

The mathematical modeling methodology described provides statistically validated water temperature estimates for global Ironman venues, enabling comprehensive analysis of this critical environmental factor. By incorporating these estimates into multivariate performance models, researchers can better isolate and quantify the complex relationships between environmental conditions and athletic performance in what may represent the most environmentally sensitive format in endurance sport.

# **References**

Bergeron, M. F. (2014). Heat stress and thermal strain challenges in running. *Journal of Orthopaedic & Sports Physical Therapy*, *44*(10), 831-838.

Castellani, J. W., Young, A. J., Ducharme, M. B., Giesbrecht, G. G., Glickman, E., & Sallis, R. E. (2006). American College of Sports Medicine position stand: Prevention of cold injuries during exercise. *Medicine & Science in Sports & Exercise*, *38*(11), 2012-2029.

Chatard, J. C., & Wilson, B. (2008). Effect of fastskin suits on performance, drag, and energy cost of swimming. *Medicine & Science in Sports & Exercise*, *40*(6), 1149-1154.

Global Lake Temperature Collaboration. (2021). *Global Lake Temperature Database* [Data set]. National Science Foundation.

Gonzalez-Alonso, J., Teller, C., Andersen, S. L., Jensen, F. B., Hyldig, T., & Nielsen, B. (1999). Influence of body temperature on the development of fatigue during prolonged exercise in the heat. *Journal of Applied Physiology*, *86*(3), 1032-1039.

Hue, O., Monjo, R., Lazzaro, M., Baillot, M., Hellard, P., Marlin, L., & Jean-Etienne, A. (2007). The effect of time of day on cold water ingestion by high-level swimmers in a tropical climate. *International Journal of Sports Physiology and Performance*, *2*(3), 258-271.

Ironman Corporation. (2023). *Ironman competition rules*. Tampa, FL: World Triathlon Corporation.

Keramidas, M. E., Geladas, N. D., Mekjavic, I. B., & Kounalakis, S. N. (2010). Effect of cold-water immersion on handgrip performance. *Ergonomics*, *53*(7), 921-930.

Kerr, C. G., Trappe, T. A., Starling, R. D., & Trappe, S. W. (2018). Hyperthermia during Olympic triathlon: Influence of body heat storage during the swimming stage. *Medicine & Science in Sports & Exercise*, *30*(1), 99-104.

Laursen, P. B., & Rhodes, E. C. (2011). Factors affecting performance in an ultraendurance triathlon. *Sports Medicine*, *31*(3), 195-209.

Laursen, P. B., Suriano, R., Quod, M. J., Lee, H., Abbiss, C. R., Nosaka, K., Martin, D. T., & Bishop, D. (2006). Core temperature and hydration status during an Ironman triathlon. *British Journal of Sports Medicine*, *40*(4), 320-325.

Marais, G., & Noakes, T. D. (2011). The effects of thermal conditions on physiological responses during an Ironman triathlon. *European Journal of Applied Physiology*, *91*, 609-618.

Peeling, P., & Landers, G. (2009). Swimming intensity during triathlon: A review of current research and strategies to enhance race performance. *Journal of Sports Sciences*, *27*(10), 1079-1085.

Sawka, M. N., Burke, L. M., Eichner, E. R., Maughan, R. J., Montain, S. J., & Stachenfeld, N. S. (2007). American College of Sports Medicine position stand: Exercise and fluid replacement. *Medicine & Science in Sports & Exercise*, *39*(2), 377-390.

Stocks, J. M., Taylor, N. A., Tipton, M. J., & Greenleaf, J. E. (2004). Human physiological responses to cold exposure. *Aviation, Space, and Environmental Medicine*, *75*(5), 444-457.

Tipton, M. J., & Bradford, C. (2014). Moving in extreme environments: What’s extreme and who decides? *Extreme Physiology & Medicine*, *3*(1), 1-5.

Tipton, M. J., Collier, N., Massey, H., Corbett, J., & Harper, M. (2017). Cold water immersion: kill or cure? *Experimental Physiology*, *102*(11), 1335-1355.

Ulsamer, S., Rüst, C. A., Rosemann, T., Lepers, R., & Knechtle, B. (2014). Swimming performances in long distance open-water events with and without wetsuit. *BMC Sports Science, Medicine and Rehabilitation*, *6*(1), 20.

Wegelin, J. A., & Hoffman, M. D. (2011). Variables associated with odds of finishing and finish time in a 161-km ultramarathon. *European Journal of Applied Physiology*, *111*(1), 145-153.

World Ocean Database. (2020). *World Ocean Atlas* [Data set]. National Oceanic and Atmospheric Administration.

Wu, S. S., Peiffer, J. J., Brisswalter, J., Nosaka, K., & Abbiss, C. R. (2015). Factors influencing pacing in triathlon. *Open Access Journal of Sports Medicine*, *6*, 115-125.

**Appendix D:**  **Elevation Data in Ironman Triathlon Analysis: Methodology and Relevance**

**Methodology for Obtaining Elevation Data**

Multi-Source Data Acquisition Approach

The elevation data for Ironman triathlon venues was collected using a multi-source web scraping methodology that prioritizes accuracy through cross-validation. The primary parameters collected include:

1. Venue altitude (elevation above sea level)

2. Bike course elevation gain (total climbing in feet)

3. Run course elevation gain (total climbing in feet)

Data Collection Process

The script employs a hierarchical three-tier approach with source prioritization:

Tier 1: TriathlonCourseInfo.com (Primary Source)

- Provides comprehensive course descriptions for global triathlon events

- Contains structured sections for bike and run course details

- Includes both elevation gain and course rofile information

- Uses a custom search algorithm that combines race name and location to find the most relevant pages

Tier 2: PJammCycling.com (Specialized Cycling Source)

- Specializes in detailed cycling course analysis

- Offers precise elevation measurements for bike segments

- Includes gradient distribution and climbing categorization

- Particularly valuable for mountain courses with significant climbing

Tier 3: Ironman.com Official Website (Authoritative Source)

- Contains official race information directly from the organizers

- Accessed through multiple URL pattern attempts to find course information

- Often includes standardized course descriptions with elevation data

- Used primarily for validation and gap-filling when other sources are unavailable

Data Extraction Techniques

The script implements multiple text processing techniques to extract elevation values accurately:

1. \*\*Pattern Recognition\*\*: Scans for text patterns containing both elevation keywords and numeric values

2. \*\*Unit Conversion\*\*: Automatically detects and converts between feet and meters for consistent reporting

3. \*\*Context Awareness\*\*: Distinguishes between bike and run elevation through contextual analysis

4. \*\*Structural Parsing\*\*: Identifies course-specific sections within webpage structure

5. \*\*Proximity Analysis\*\*: Associates elevation values with appropriate course segments based on textual proximity

To ensure data quality, the system employs several validation mechanisms:

1. \*\*Cross-Source Verification\*\*: Compares values across multiple data sources

2. \*\*Order-of-Magnitude Check\*\*: Validates that values fall within expected ranges for Ironman courses

3. \*\*Contextual Consistency\*\*: Ensures bike elevation gain is typically greater than run elevation gain

4. \*\*Unit Consistency\*\*: Standardizes all measurements to feet for consistent analysis

**Relevance of Elevation Parameters in Ironman Triathlon Research**

Elevation parameters significantly influence athletic performance across multiple dimensions:

1.1 Venue Altitude Effects

Physiological Impact:

- Oxygen Availability: Every 1,000 feet of elevation decreases oxygen availability by approximately 3%, affecting aerobic performance

- Hematological Adaptations: Venues above 5,000 feet trigger EPO production and increased red blood cell concentration

- Respiratory Compensation: Higher ventilation rates are required, increasing respiratory muscle fatigue

- Thermoregulation: Lower air density at altitude affects evaporative cooling efficiency

Performance Implications:

- Power output typically decreases 7-10% for unacclimated athletes at moderate altitudes (5,000-8,000 feet)

- Perceived exertion increases approximately 6-8% for equivalent workloads

- Pacing strategy adjustments become crucial to prevent early glycogen depletion

- Recovery between efforts is impaired, affecting interval-based race strategies

Notable High-Altitude Ironman Venues:

- Ironman Lake Tahoe (6,224 feet) - discontinued but historically significant

- Ironman Santa Fe (7,199 feet)

- Ironman Ecuador (9,350 feet)

1.2 Bike Course Elevation Gain

Physiological Impact:

- Energy System Recruitment: Climbing segments shift metabolic demand toward greater carbohydrate utilization

- Muscle Fiber Activation: Increased recruitment of Type II muscle fibers during ascents

- Joint Loading: Different biomechanical stress patterns compared to flat-course riding

- Thermoregulation Challenge: Core temperature increases during climbing due to reduced airflow and increased work

Performance Implications:

- Each 1,000 feet of climbing typically adds 3-5 minutes to bike split times for elite athletes

- Power output variation increases by 30-40% on hilly courses versus flat courses

- Nutritional requirements increase by approximately 100-200 calories per hour

- Equipment selection becomes more critical (gearing, wheel choice, bike weight)

Course Classification by Elevation Gain:

- Flat: <2,500 feet (e.g., Ironman Florida: 1,489 feet)

- Moderate: 2,500-4,500 feet (e.g., Ironman Texas: 2,800 feet)

- Hilly: 4,500-7,000 feet (e.g., Ironman Lake Placid: 6,240 feet)

- Mountainous: >7,000 feet (e.g., Ironman Nice: 7,975 feet)

1.3 Run Course Elevation Gain

Physiological Impact:

- Musculoskeletal Loading: Different muscle recruitment patterns between ascents, descents, and flat terrain

- Impact Forces: Increased eccentric muscle loading during descents contributes to muscle damage

- Cardiac Drift: Greater heart rate variability on undulating courses

- Fatigue Accumulation: Non-linear relationship between elevation gain and pace deterioration

Performance Implications:

- Each 500 feet of elevation gain typically adds 2-4 minutes to marathon times

- Pacing variability increases by 15-20% on hilly marathon courses

- Rate of perceived exertion increases disproportionately in the latter stages of hilly runs

- Injury risk profiles shift from repetitive impact injuries to acute muscle/tendon stress

Course Classification by Elevation Gain:

- Flat: <500 feet (e.g., Ironman Florida: 398 feet)

- Rolling: 500-1,000 feet (e.g., Ironman Arizona: 650 feet)

- Hilly: 1,000-1,800 feet (e.g., Ironman Wisconsin: 1,430 feet)

- Very Hilly: >1,800 feet (e.g., Ironman Lake Placid: 1,940 feet)

**Conclusion**

The multi-source methodology for collecting elevation data provides a robust foundation for analyzing a critical aspect of Ironman triathlon performance. By systematically gathering venue altitude, bike elevation gain, and run elevation gain, researchers can develop more nuanced understandings of race dynamics, physiological demands, and strategic implications.

The elevation profile of an Ironman course represents one of the most significant fixed variables affecting race outcomes. Unlike weather conditions that vary year to year, the fundamental elevation characteristics remain relatively constant, providing a stable basis for comparative analysis and performance prediction.

By incorporating elevation data into broader performance analytics, athletes, coaches, and researchers can develop more sophisticated approaches to training, racing, and understanding the unique challenges of Ironman triathlon competition. The continued refinement of elevation data collection methodologies will further enhance our ability to quantify and analyze this crucial aspect of the sport.