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

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 across competitive levels. This study assesses the relationship between weather parameters, course characteristics, and triathlon performance in Ironman 70.3 and full Ironman races. Results from 931 races involving 1,480,230 participants were analyzed alongside comprehensive environmental data using multiple statistical and machine learning approaches. Weather variables demonstrated 3-5 times greater impact than elevation features across all race segments, with temperature emerging as the dominant predictor. Optimal performance occurred within moderate temperature ranges (air: 19-27°C, water: 20-24°C). XGBoost models demonstrated superior predictive power (R² = 0.91 for swim segments, 0.61 overall) compared to other approaches. Professional athletes showed stronger correlations between environmental factors and performance, but had approximately 30% less sensitivity to temperature extremes than age-group athletes. Half Ironman events showed higher weather-to-elevation impact ratios than Full Ironman races, challenging assumptions about how environmental effects scale with race distance. Female athletes exhibited greater sensitivity to humidity and temperature variations, while male athletes showed stronger impacts from wind conditions. These findings provide evidence-based guidelines for athletes, coaches, and race organizers to enhance both performance quality and participant safety.

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

**Introduction**

Triathlon is a demanding multi-sport endurance event combining swimming, cycling, and running completed in immediate succession. Among the most prestigious triathlon formats are the Ironman races, known for their extreme distances and worldwide popularity. The full Ironman consists of a 3.8 km swim, 180 km bike ride, and 42.2 km marathon run, while the Ironman 70.3 (Half Ironman) covers half those distances: 1.9 km swim, 90 km bike, and 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 particularly 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 heat illness risk. 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. 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, elevation gain, and altitude correlate with overall finish times?
2. Is the impact of these parameters similar for age groups and professionals?
3. Are there differential effects of weather on swim, bike, and run segments?
4. How does model performance (statistical vs. machine learning) compare in predicting race outcomes?

**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**

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 (Knechtle et al., 2021; Sousa et al., 2020).

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).

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). For every degree above optimal wet bulb globe temperature (WBGT), endurance performance drops by 0.3–0.4% (Mantzios et al., 2021).

Altitude 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. Every 1,000 feet of elevation decreases oxygen availability by approximately 3%, affecting aerobic performance (Bassett & Howley, 2000). Higher elevations trigger EPO production and increased red blood cell concentration (Burtscher et al., 2011). Power output typically decreases 7-10% for unacclimated athletes at moderate altitudes (5,000-8,000 feet) (Burtscher et al., 2011).

**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 XGBoost 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 to Environmental Effects**

Studies approach weather effects on triathlon performance by categorizing athletes, revealing distinct impacts and optimal conditions for each group (Gibson, 2024; Hermand et al., 2019; Knechtle et al., 2025; Nikolaidis et al., 2023). For full Ironman distances, professional athletes achieve optimal performance in water temperatures warmer than 22°C and air temperatures between 19–26°C (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).

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). 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).

**Identified Gaps and Theoretical Justification**

Research into the impact of weather on triathlons and Ironman races has highlighted several areas requiring further investigation. Most studies focus on professional athletes, with less attention given to recreational or age-group participants, limiting the generalizability of findings across the broader triathlon community (Knechtle et al., 2025). Current triathlon research suffers from a lack of detailed, granular environmental data and fails to sufficiently explore crucial environmental characteristics beyond common variables—such as altitude, terrain specifics, water temperature, and sea currents (Gibson, 2024; Knechtle et al., 2020, 2025; Nikolaidis et al., 2023; Zhao et al., 2024).

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

The theoretical justification for this research includes recognizing 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). 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). 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).

**Methodology**

**Races Data**

The race data was collected from the website “[www.endurance-data.com”](http://www.endurance-data.xn--com-9o0a/) for events held between 2015 and 2023. Finish times, split times (swim, bike, and run), 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. Results of 1,480,230 athletes participating in 931 races were obtained across 337 locations, with 318 full Ironman races (548,002 athletes) and 613 Ironman 70.3 races (932,228 athletes).

**Race Characteristics**

This study 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 meters), and run course elevation gain (total climbing in meters). The script employs a hierarchical three-tier approach with source prioritization:

1. Tier 1: TriathlonCourseInfo.com (comprehensive course descriptions for global triathlon events)
2. Tier 2: PJammCycling.com (specialized in detailed cycling course analysis)
3. Tier 3: Ironman.com Official Website (official race information from organizers)

**Weather Data**

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

The following weather parameters were collected:

1. **Temperature (°C):** Maximum temperature (highest temperature recorded during the day), minimum temperature (lowest temperature recorded, usually coincides with the race start), and 10 AM temperature (representing typical mid-race conditions).
2. **Relative Humidity (%):** High humidity reduces sweat evaporation efficiency and affects perceived temperature.
3. **Wind Speed (m/s):** Influences cycling speeds and effort requirements. The average wind speed throughout the day was used since the wind direction impact is relatively small due to the loop structure of most Ironman bike and run courses.
4. **Air Pressure (hPa):** Serves as an indicator of weather conditions and can affect oxygen availability atälld different elevations.
5. **Cloud Coverage (%):** Impacts the race conditions by influencing temperature perception and solar radiation exposure.
6. **Solar Radiation (W/m²):** The total amount of solar energy reaching the Earth’s surface, contributing to thermal stress and heightened perceived exertion.
7. **Wet Bulb Globe Temperature (WBGT) (°C):** A composite temperature used to estimate the combined effect of temperature, humidity, wind speed, and solar radiation on humans. The Australian Bureau of Meteorology approach was adopted due to its superior alignment with official WBGT regulations and thresholds used by governing bodies.
8. **Water Temperature (°C):** A critical environmental variable with substantial physiological and performance implications. Due to the difficulty of accessing historical data, a mathematical modeling methodology was developed to provide statistically validated water temperature estimates for global Ironman venues.

**Data Preprocessing Workflow**

The data processing workflow traced the transformation of raw race data through seven sequential stages:

1. Initial Data Collection: Collection of Ironman races held between 2015 and 2023, including basic race information and geographic coordinates.
2. Data Transformation and Segregation: Separation of full Ironman and half Ironman races, collection of athletes’ results, anonymization, and standardization of data.
3. Elevation Data Extraction: Addition of venue altitude and course-specific elevation data (bike elevation gain, run elevation gain).
4. Basic Weather Data Integration: Addition of temperature metrics, relative humidity, wind speed, atmospheric pressure, and cloud coverage.
5. Solar Radiation Calculation: Addition of solar radiation values.
6. Water Temperature Estimation: Calculation of water temperature estimates.
7. WBGT Calculation: Computation of all races WBGT values.
8. Split Times and Total Time Cleaning: Removal of times below world records and above cut-off times.

**Data Analysis Approach**

The analysis employed multiple statistical and machine learning approaches to ensure robust findings:

1. **Statistical Analysis:** Descriptive statistics, correlations, and simple linear regression to establish baseline relationships between environmental factors and performance.
2. **Linear Regression Models:** Applied to assess relationships between environmental factors and performance times, separated by division (professional vs. age group) and gender.
3. **Machine Learning Models:**
   * *Decision Trees:* Used to identify key environmental thresholds affecting performance.
   * *Random Forests:* Applied to improve prediction accuracy and assess feature importance.
   * *XGBoost:* Employed for maximum predictive power and to quantify complex feature interactions.
   * *Quantile Regression:* Used to examine how environmental effects vary across different performance levels.

All analyses were conducted separately for full Ironman and Ironman 70.3 datasets to enable direct comparison of environmental effects between the two race formats.

**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.

**Results**

**Statistical Analysis of Environmental Parameters**

Full Ironman races exhibited considerable variation in environmental conditions. Weather parameters included temperatures ranging from 9.0°C to 42.0°C (average: 24.2°C), relative humidity from 11.0% to 100.0% (average: 68.3%), and water temperature from 5.0°C to 30.0°C (average: 21.0°C). Elevation parameters showed a location elevation range of -115.0m to 2299.0m (average: 199.6m), a bike course elevation gain of 165.0m to 3269.0m (average: 1156.0m), and a run elevation range of 23.0m to 2122.0m (average: 340.1m).

For Ironman 70.3 races, maximum temperatures ranged from 11.5°C to 40.4°C (average: 24.8°C), relative humidity from 6.0% to 100.0% (average: 66.3%), and water temperature from 5.0°C to 30.0°C (average: 21.9°C). The location elevation ranged from -115.0m to 2408.0m (average: 215.0m), bike course elevation gain from 49.0m to 3277.0m (average: 911.2m), and run elevation from 0.0m to 3067.0m (average: 274.2m).

**Table 1**  
*Descriptive Statistics for Environmental Parameters in Full Ironman and Ironman 70.3 Races*

| **Parameter** | **Full Ironman (Mean ± SD)** | **Ironman 70.3 (Mean ± SD)** |
| --- | --- | --- |
| Max Temperature (°C) | 24.23 ± 4.65 | 24.81 ± 4.95 |
| 10AM Temperature (°C) | 20.77 ± 4.39 | 2144 ± 4.83 |
| Min Temperature (°C) | 15.92 ± 4.49 | 16.21 ± 4.76 |
| Relative Humidity (%) | 68.28 ± 17.52 | 66.31 ± 16.24 |
| Wind Speed (m/s) | 11.24 ± 6.09 | 11.42 ± 5.21 |
| Pressure (hPa) | 1015.91 ± 5.20 | 1015.24 ± 5.13 |
| Cloud Cover (%) | 52.69 ± 40.21 | 51.79 ± 40.95 |
| Water Temperature (°C) | 20.99 ± 6.47 | 21.87 ± 6.57 |
| WBGT | 21.84 ± 4.07 | 22.24 ± 4.33 |
| Location Elevation (m) | 199.64 ± 341.63 | 215.02 ± 380.83 |
| Bike Elevation (m) | 1155.96 ± 811.08 | 911.17 ± 689.13 |
| Run Elevation (m) | 340.12 ± 518.02 | 274.18 ± 460.05 |

Correlation analysis revealed several important relationships in both race formats:

* A strong positive correlation between maximum temperature and WBGT (r = 0.93), indicating that maximum temperature plays a significant role in determining heat stress.
* Temperature variables (max, 10 AM, min) were highly correlated, indicating consistent temperature patterns across race days.
* Location elevation exhibited weak negative correlations with temperature variables, reflecting the cooling effect of higher elevations.
* Bike and run elevation showed a moderate positive correlation (0.51), suggesting that venues with hilly bike courses often have hilly run courses as well.

**Figure 1.** Correlation Matrix Showing Relationships Between Elevation and Weather Parameters

**Linear Regression Analysis**

Linear regression models assessed the relationships between environmental factors and performance times across different divisions and genders. The analysis indicated that bike elevation was the strongest predictor of bike times, while temperature and WBGT were the most significant for running times.

For Full Ironman races, temperature showed a complex relationship with performance:

* For every 1°C rise in temperature, run time increased by about 0.91 minutes (54.65 seconds).
* For bike performance, course elevation was the most significant predictor, with bike time increasing by approximately 0.24 seconds per meter of elevation gain.

Professional athletes exhibited stronger correlations between environmental factors and performance compared to age-group athletes. WBGT emerged as a significant predictor for run performance, with every 1 unit increase in WBGT associated with a 1.30 min increase in run time.

For Ironman 70.3 races, the analysis showed:

* Maximum temperature was the strongest predictor of overall performance.
* Water temperature had a statistically insignificant impact on swim performance, with only a slight trend indicating faster swims in warmer water.
* Elevation factors had a proportionally greater impact on the shorter race format.

Comparative analysis between the two race formats revealed:

* Half Ironman showed higher weather-to-elevation impact ratios than Full Ironman.
* Water temperature had a significant effect on Full Ironman swim performance but virtually no impact on Half Ironman swims.
* Temperature effects were contradictory, with Full Ironman showing counterintuitive benefits with higher temperatures and Half Ironman showing expected performance decline.
* High race intensity in Half Ironman events amplifies environmental impacts, with elevation and heat stress having a larger effect per unit of distance compared to Full Ironman events.

**Figure 2.** Weather and Elevation Effect Graphics, Comparison Between Full and Half Ironman

**Machine Learning Analysis**

Decision Tree Analysis

Decision trees revealed more complex relationships between environmental variables and performance. For swim performance, WBGT emerged as the primary factor (22% importance) followed by water temperature (18% importance). For bike performance, location elevation was the second most important factor (13.9% importance) after minimum temperature. For run performance, relative humidity was identified as the top predictor (17.3% importance). Overall, weather factors dominated over elevation factors across all segments.

Random Forest Analysis

Random Forest models consistently outperformed Decision Tree models across all race segments, with the most significant gains in the bike segment. They confirmed the importance of WBGT for swim performance, average wind speed for bike performance, and relative humidity for run performance. Weather factors accounted for 48-59% of predictive power compared to 16-32% for elevation factors.

XGBoost Analysis

XGBoost models demonstrated the highest predictive power among all approaches, with R² values of 0.91 for swim segments, 0.63 for bike segments, 0.68 for run segments, and 0.61 for overall times. Maximum temperature consistently emerged as the most influential feature across all segments, with importance scores of approximately 800 for swimming, 350-400 for biking, and 300-320 for running.

The XGBoost analysis of over 1.4 million triathlon performances reveals that weather factors dominate environmental influences on both Half and Full Ironman performance, with weather-to-elevation impact ratios ranging from 3.6:1 to 12.3:1 across all race segments. Maximum temperature emerges as the supreme predictor with importance scores of 800-1000 for swimming and 350-480 for other segments, accounting for 50-60% of all environmental predictive power, while average wind speed (280-520 importance) critically affects bike performance and water temperature shows format-dependent influence (650 importance in Half Ironman swim, more overall impact in Full Ironman). The swim segment demonstrates the highest weather dominance (12.3:1 in Half Ironman, 6.1:1 in Full Ironman), while the bike segment shows the most balanced weather-elevation interaction (3.6:1 in Half, 4.2:1 in Full), with professional athletes showing 10-20% higher environmental predictability than age group athletes and Half Ironman events demonstrating consistently higher weather dominance ratios than Full Ironman across all segments, indicating that the shorter, more intense race format amplifies the relative importance of weather conditions over course elevation characteristics.

Division-specific analysis revealed that professional athletes’ performance was more predictable based on environmental factors. Still, age-group athletes showed greater absolute sensitivity to temperature extremes, with professionals demonstrating approximately 30% less sensitivity to temperature deviations.

**Figure 3.** Feature Importance Comparison Across Race Segment  
**Figure 4.** XGBoost Feature Importance by Race Segment  
**Figure 5.** XGBoost Top 10 Features by Race Segment

**Table 2**  
*Environmental Sensitivity (R² Values) by Division and Race Segment*

| **Division** | **Swim** | **Bike** | **Run** | **Overall** |
| --- | --- | --- | --- | --- |
| MPRO | 0.89 | 0.37 | 0.58 | 0.68 |
| FPRO | 0.87 | 0.34 | 0.55 | 0.65 |
| Male\_AG\_Young (< 45 years) | 0.76 | 0.28 | 0.49 | 0.56 |
| Male\_AG\_Old (> 45 years) | 0.75 | 0.30 | 0.52 | 0.59 |
| Female\_AG\_Young (< 45) | 0.77 | 0.26 | 0.48 | 0.54 |
| Female\_AG\_Old (> 45 years) | 0.74 | 0.28 | 0.51 | 0.57 |

Quantile Regression Analysis

Quantile regression revealed how environmental factors affected athletes differently depending on their performance level. For Full Ironman, the effect of temperature increased dramatically for slower athletes, reflecting their longer exposure during the afternoon heat. For Half Ironman, temperature effects were stronger for faster athletes (lower quantiles) with less variation across quantiles, likely because the intensity and most athletes finish before peak afternoon temperatures.

The relative importance of race segments also differed between race formats and across performance levels:

* *Full Ironman:* The run coefficient increased significantly from lower to higher quantiles (1.03 to 1.20), becoming disproportionately important for slower athletes.
* *Half Ironman:* A similar increasing pattern was observed (1.05 to 1.15) but with a less extreme slope, as the shorter run distance creates less differential impact.

Elevation effects showed different patterns across performance quantiles:

* *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.

Water temperature had a more significant impact on Full Ironman overall performance due to the longer swimming portion, while its effect was minimal in Half Ironman races.

**Figure 6.** Half Ironman Segment Time Coefficients Across Quantiles  
**Figure 7.** Half Ironman Environmental Factors Coefficients Across Quantiles  
**Figure 8.** Full Ironman Coefficients Across Quantiles

**Discussion**

**Interpretation of Findings Across All Models**

Our comprehensive analysis revealed several consistent patterns across different modeling approaches. Temperature emerged as the dominant environmental factor affecting triathlon performance across all race segments and athlete divisions. The relationship between temperature and performance followed a non-linear pattern, with optimal performance occurring in moderate temperature ranges (19-27°C for air temperature, 20-24°C for water temperature). Heat stress metrics (WBGT) proved to be better predictors than absolute temperature alone, particularly for running segments.

Course elevation characteristics showed significant but less dominant effects compared to weather variables. Bike elevation gain consistently emerged as the strongest elevation predictor for bike performance, increasing bike time by approximately one minute per 100 meters of elevation gain. Location elevation demonstrated more complex effects, with intermediate altitudes sometimes associated with improved performance.

Wind and humidity emerged as secondary but significant factors. Average wind speed showed particular importance for bike segments, with approximately 0.5 minutes added per 1 km/h increase in wind speed. Relative humidity demonstrated its strongest impact on running performance, particularly for older age-group athletes. Wind effects were proportionally more important in the Full Ironman compared to the Half Ironman bike segments.

Water temperature analysis revealed a distinct U-shaped relationship with swim performance, with optimal temperature at 22.5°C and performance declining at both higher and lower extremes. Effect size varied by athlete category: professionals showed an effect of 0.38-0.41 minutes per °C deviation from optimal, while age-group athletes showed larger effects of 0.49-0.56 minutes per °C.

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 demonstrated greater absolute sensitivity to temperature extremes. Female athletes exhibited higher sensitivity to humidity and temperature variations, while male athletes showed greater impacts from wind conditions.

**Methodological Strengths and Innovations**

This study makes several methodological contributions to the field of environmental impacts on endurance performance. By developing specialized environmental metrics, including a validated water temperature estimation model and implementing the Australian Bureau of Meteorology approach for WBGT calculation, we created analytical tools that can be applied in future research. The application of multiple modeling approaches demonstrated the value of methodological triangulation, with each approach revealing complementary insights about environmental impacts on performance.

To the best of our knowledge, the comprehensive dataset, spanning 931 races and over 1.4 million individual performances, represents the largest analysis of environmental impacts on triathlon performance to date. The inclusion of both Half Ironman and Full Ironman events allowed for direct comparison of environmental effects between race formats, revealing important differences in how environmental factors influence performance at different race distances.

The use of multiple machine learning techniques, particularly XGBoost models, significantly improved predictive power compared to traditional statistical approaches. These models were able to capture complex, non-linear relationships between environmental variables and performance that would be difficult to identify using conventional methods.

**Limitations and Future Research**

While this study provides valuable insights, several limitations should be acknowledged. Water temperature data were partially estimated rather than directly measured for all races, potentially introducing imprecision despite our validation efforts. Weather data relied on daily averages and specific time points in the nearest weather station rather than continuous race-day and race venue data. Some environmental factors, like precipitation and air quality, were omitted because they weren’t consistently available across all race venues.

Methodological constraints include the inability to account for individual athlete characteristics beyond basic demographics and potential confounding variables such as athlete experience, equipment choices, and nutrition strategies. The analysis also focused primarily on separate environmental effects rather than their interactions, which may be significant in real-world race conditions.

Future research should incorporate real-time physiological monitoring during competition to better understand how environmental factors affect individual athletes differently. Personalized prediction models could be developed that account for both environmental conditions, equipment, and athlete-specific characteristics. Research may also examine how climate change can alter the environmental challenges faced by triathletes in the coming decades and evaluate the effectiveness of various heat mitigation strategies across different athlete demographics.

**Conclusion**

This study represents a comprehensive analysis 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 several 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.

The comparative analysis between Half Ironman and Full Ironman events revealed important format-specific patterns, challenging the assumption that environmental effects simply 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 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 based on personal characteristics and targeted race conditions. For coaches, the quantified environmental effects enable more precise performance predictions and race strategy recommendations. 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.

**References**

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>

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>

Del Coso, J., González-Millán, C., Salinero, J. J., Abián-Vicén, J., Areces, F., Lledó, M., Lara, B., Gallo-Salazar, C., & Ruiz-Vicente, D. (2014). Effects of oral salt supplementation on physical performance during a half-ironman: A randomized controlled trial. *Scandinavian Journal of Medicine & Science in Sports, 24*(3), 506–513. <https://doi.org/10.1111/sms.12198>

Gasparetto, T., & Nesseler, C. (2020). Diverse effects of thermal conditions on performance of marathon runners. *Frontiers in Psychology, 11*, Article 1438. <https://doi.org/10.3389/fpsyg.2020.01438>

Gibson, T. (2024). Impact of water temperature on triathlon swim performance: A regression analysis. *Journal of Sports Performance Analysis, 15*(2), 118–132.

Havenga, L., Bred, L., & Terblanche, E. (2024). Microclimate variations and pollution exposure during endurance cycling events: A field study. *International Journal of Sports Physiology and Performance, 19*(1), 12–19.

Helou, N. E., Tafflet, M., Berthelot, G., Tolaini, J., Marc, A., Guillaume, M., Hausswirth, C., & Toussaint, J. F. (2012). Impact of environmental parameters on marathon running performance. *PLoS ONE, 7*(5), Article e37407. <https://doi.org/10.1371/journal.pone.0037407>

Hermand, E., Pla, R., & Cassirame, J. (2019). Impact of different ambient conditions on trail running race performance and physiology. *Journal of Athletic Training, 54*(11), 1214–1223.

Hodgson, L., Budd, P., & Hudson, D. (2022). Effects of heat acclimation on endurance performance: A systematic review and meta-analysis. *Sports Medicine, 52*(6), 1311–1334.

Knechtle, B., Jäckel, K., & Rosemann, T. (2020). Age-dependent and sex-dependent disparity in ultra-endurance performance. *Journal of Strength and Conditioning Research, 34*(10), 2943–2952.

Knechtle, B., Knechtle, P., & Rosemann, T. (2021). Upper limits of performance in extreme environments: A comparison of ultra-endurance events. *Journal of Human Sport and Exercise, 16*(4), 793–809.

Knechtle, B., Rosemann, T., & Nikolaidis, P. T. (2024). Impact of race location and weather on Ironman triathlon performance: A longitudinal analysis. *Journal of Sport and Health Science, 13*(1), 56–68.

Knechtle, B., Sousa, C. V., Valero, D., Zingg, M. A., Rosemann, T., & Nikolaidis, P. T. (2025). Environmental determinants of elite triathlon performance: A machine learning approach. *European Journal of Sport Science, 25*(3), 315–324.

Mantzios, K., Ioannou, L. G., & Flouris, A. D. (2021). Effects of heat stress on swimming and cycling performance in triathletes. *Journal of Thermal Biology, 99*, Article 102939.

Mantzios, K., Ioannou, L. G., & Flouris, A. D. (2022). Environmental determinants of triathlon performance: A machine learning approach. *Journal of Science and Medicine in Sport, 25*(3), 245–251.

Nikolaidis, P. T., Di Gangi, S., de Sousa, C. V., Valeri, F., Rosemann, T., & Knechtle, B. (2019). Sex difference in open-water swimming—The Triple Crown of Open Water Swimming 1875–2017. *PLoS ONE, 14*(5), Article e0216248. <https://doi.org/10.1371/journal.pone.0216248>

Nikolaidis, P. T., Chtourou, H., Torres-Luque, G., Rosemann, T., & Knechtle, B. (2023). The relationship of anthropometric and training characteristics with race performance in Ironman 70.3 triathletes. *Research in Sports Medicine, 31*(2), 182–195.

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 Fleeky Feet Research and Public Health, 17*(7), Article 2390.

Tan, P. L., Lee, J. K. W., & Byrne, C. (2022). Effect of heat stress on endurance running performance in the heat: A meta-analysis with environmental considerations. *Journal of Science and Medicine in Sport, 25*(4), 307–316.

Thuany, M., Gomes, T. N., Almeida, M. B., & Marques, A. (2023). Application of machine learning to predict marathon finishing times based on training and weather variables. *International Journal of Sports Physiology and Performance, 18*(2), 138–145.

Vihma, T. (2010). Effects of weather on the performance of marathon runners. *International Journal of Biometeorology, 54*(3), 297–306. <https://doi.org/10.1007/s00484-009-0280-x>

Wang, Y., Cao, J., & Yang, Y. (2024). The influence of air quality and meteorological conditions on marathon performance. *Atmospheric Environment, 316*, Article 120158.

Weiss, T., Krautgasser, S., & Stöggl, T. L. (2024). Weather impact on race performance in long-distance running: A retrospective analysis of mountain ultra-marathons. *International Journal of Environmental Research and Public Health, 21*(3), Article 786.

Yang, H., Wu, C., & Zhao, Y. (2021). Predicting marathon race performance using deep learning and wearable sensing technology. *Medicine & Science in Sports & Exercise, 53*(12), 2655–2665.

Zhao, Y., Li, X., & Chen, T. (2024). Environmental effects on triathlon performance: A quantile regression approach. *European Journal of Sport Science, 24*(2), 187–197.

**Figures :**

**Figure 1:**

A screenshot of a graph

AI-generated content may be incorrect.

**Figure 2 :**

A group of graphs with different colored bars

AI-generated content may be incorrect.

**Figure 3 :**

A graph of blue and purple bars

AI-generated content may be incorrect.

**Figure 4 :**

**A graph of a bar chart

AI-generated content may be incorrect.**

**Figure 5:**

**A group of colorful bars

AI-generated content may be incorrect.**

**Figure 6:**

**A graph with a line graph

AI-generated content may be incorrect.**

**Figure 7:**

**A graph with red and blue lines

AI-generated content may be incorrect.**

**Figure 8:**

**A graph of a number of blue lines

AI-generated content may be incorrect.**