scientific reports

OPEN

The fastest nonprofessional age group IRONMAN triathletes in the world originate from Europe

Beat Knechtle[1,2✉], Elias Villiger[2], Katja Weiss[2], David Valero[3], Thomas Rosemann[2], Pantelis T. Nikolaidis[4], Rodrigo Luiz Vancini[5], Marilia Santos Andrade[6] & Mabliny Thuany[7]

It is well known that elite athletes of specific ethnicities and/or nationalities dominate certain sports disciplines (*e.g.*, East Africans in marathon running). However, we do not know the nationalities of the fastest non-professional IRONMAN triathletes. Therefore, this study intended to identify the fastest athletes by country of origin competing in IRONMAN triathlon events, focusing on non-professional age group triathletes. Data from all IRONMAN age group athletes competing worldwide between 2002 and 2022 in all official IRONMAN races were collected. Sex, age group, country of origin of the athletes, location and year of the event, split times, overall race times, and transition times were obtained. Additionally, the dataset was augmented with specific data (*i.e.* event characteristics such as temperatures for water and air and course characteristics for all three split disciplines) related to the different race locations. We limited the analysis to the top 150 countries by participation (*i.e.* countries with at least 13 successful finishers records in the sample). A total of 677,320 records of IRONMAN age group triathletes originating from 150 different countries and participating in 443 races over 65 different locations were analyzed. European countries such as Germany, Austria, Denmark, Belgium, Switzerland, Norway, Czechia, Estonia, and Slovenia have the fastest IRONMAN age group athletes. IRONMAN Hawaii, IRONMAN Vitoria-Gasteiz and IRONMAN Hamburg are the fastest races. Hilly running and cycling race courses led to slower race times, while flat surfaces, rolling cycling and ocean swimming led to faster race times. Optimal water temperatures were found at 23–25 °C and optimal air temperature ranged between 19–21 and 25–28 °C. The fastest IRONMAN age group triathletes from European countries such as Germany, Austria, Denmark, Belgium, Switzerland, Norway, Czechia, Estonia, and Slovenia. With the presented results for optimal air and water temperatures and description of the optimal cycling and running course characteristics, IRONMAN age group athletes might be able to select an IRONMAN race with the best conditions in order to achieve a fast IRONMAN race time.

Keywords Swimming, Cycling, Running, Multisport, Nationality

Comparative studies have been recurrent in sports science[1,2]. Generally, these studies are used to verify the performance of countries in different sports disciplines, including the world championships and Olympic Games[2]. Previous literature has indicated that some countries represent a higher likelihood of being successful in specific sports disciplines, such as basketball in the United States (US)[3], Kenya and Ethiopia in long-distance running events[4], and Canada in hockey events[5]. These higher probabilities are related to several factors, including social, economic, cultural, and environmental characteristics[6–8].

Usually, these approaches associate the performance of these countries in the most important championships worldwide, to verify those achieving outstanding performance[2]. Even though most of the studies are developed in the context of elite sports, the popularity of sports among countries also presents important indicators regarding the popularity of sports among nonprofessional athletes[9], and the importance of the sport among residents. For

[1]Medbase St. Gallen Am Vadianplatz, Vadianstrasse 26, 9001 St. Gallen, Switzerland. [2]Institute of Primary Care, University of Zurich, Zurich, Switzerland. [3]Ultra Sports Science Foundation, Pierre-Benite, France. [4]School of Health and Caring Sciences, University of West Attica, Athens, Greece. [5]MoveAgeLab, Physical Education Sport Center of Federal, University of Espirito Santo, Vitoria, ES, Brazil. [6]Department of Physiology, Federal University of Sao Paulo, Sao Paulo, Brazil. [7]Department of Physical Education, States University of Pará, Pará, Brazil. ✉email: beat.knechtle@hispeed.ch

example, for endurance sports, previous literature has shown that running is part of the culture of East African countries, also influencing the practice among nonprofessional athletes[10,11].

For endurance sports, triathlon events have increased among nonprofessional participants[12], especially influenced by events held in the United States, such as the Ironman Hawaii[13]. The sport became popular among adults of both sexes and different age groups, engaging in this activity with different purposes, including health and well-being benefits, as well as to improve physical capacities and participate in competitions[14,15]. Previous literature investigating the most successful countries among elite triathletes showed that athletes from the United States presented the best performance, as well as the highest frequency of competitors, and age of peak performance[16,17].

Understanding the age demographics of world-class IRONMAN triathletes who emerge victorious and stand out as the fastest is crucial for several reasons[18]. First, it provides valuable information on the optimal age range for peak athletic performance in long-distance triathlons, offering guidance to both aspiring and experienced athletes on when their training efforts may yield the best results. Additionally, such knowledge helps sports scientists, coaches, and trainers tailor training regimens that consider age-specific physiological changes, helping athletes maximize their potential while minimizing the risk of injury. Moreover, recognizing the age groups dominating IRONMAN competitions contributes to a deeper understanding of the sport's evolving dynamics and may influence the development of age-specific talent pipelines or training programs[19,20]. In general, investigating the age demographics of top-performing IRONMAN athletes enhances our understanding of the physiological nuances of the sport and has practical implications for optimizing training strategies across different age cohorts.

Despite the evidence that indicates higher participation and performance indicators among professional athletes from the USA[16], these results present important limitations. The most important aspect refers to the different methodological approaches used among different studies, which impair the generalization of the findings; also, the time frame should be considered, in association with the greater interest in studying professional athletes[21–23]. Understanding the nuanced interplay of biopsychosocial factors and the intricate balance between intrinsic and extrinsic motivation is crucial when delving into the realms of elite performance in both professional and nonprofessional athletes[24]. The assertion that these factors differ significantly between the two categories raises pertinent questions about the dynamics that influence athletic peak achievements in diverse contexts[25]. In professional athletes, factors such as genetic predispositions, access to high-level coaching, and optimal training environments may take precedence, while nonprofessional athletes may be driven by personal goals, societal expectations, and the pursuit of holistic well-being.

Researching and identifying the countries from which the fastest non-professional IRONMAN triathletes emerge, along with their age groups, is of significant importance for several reasons. Firstly, such information provides valuable insight into the global distribution of talent in the sport, allowing for a more comprehensive understanding of the geographical patterns of high-performance triathletes outside the professional realm[26]. This knowledge can be instrumental in the formation of training programs, talent identification strategies, and the allocation of resources within different nations. Second, analyzing the age groups of the fastest non-professional IRONMAN triathletes offers critical data on the optimal stages of life for achieving peak performance in this demanding endurance sport. This information can guide coaches, trainers, and athletes in tailoring training regimens that consider age-related physiological changes and potential peak performance windows[27,28]. It also helps in the development of age-specific training methodologies to optimize athletic potential at various stages of life. In addition, understanding the demographics of nonprofessional IRONMAN triathletes contributes to a wider promotion of sport and the adoption of a healthy and active lifestyle. Highlighting the diverse age groups and nationalities of successful participants encourages a broader population to participate in triathlons, fostering a sense of inclusivity and inspiration for aspiring athletes.

In summary, researching the countries and age groups of the fastest nonprofessional IRONMAN triathletes is essential for shaping targeted training approaches, fostering global sporting development, and promoting the sport's inclusivity and accessibility to individuals of all ages and backgrounds[26,29]. Despite the importance of professional athletes for the representativeness of the countries at the national level, nonprofessional athletes should be studied to amplify the evidence regarding the fastest countries. Therefore, the purpose of this study was to identify the age group of athletes of the fastest countries competing in IRONMAN events between 2002 and 2020. Based upon existing knowledge we hypothesized that the fastest IRONMAN age group triathletes would also originate from the USA.

Methods

Ethical approval

This study was approved by the Institutional Review Board of Kanton St. Gallen, Switzerland, with a waiver of the requirement for informed consent of the participants as the study involved the analysis of publicly available data (EKSG 01/06/2010). The study was conducted following recognized ethical standards according to the Declaration of Helsinki adopted in 1964 and revised in 2013.

Data set and data preparation

Race data from all official IRONMAN races was downloaded from the official IRONMAN website (www.ironman.com) using a Python script. The sex, age group, country of origin of the athletes, location and year of the event, and times for swimming, running, cycling, overall race times, and transition were therefore obtained. The data was inspected for consistency, removing duplicate and/or incomplete records. Similarly, the event location variable was harmonized to map generic values to their actual location. Race times were re-calculated to hours and are expressed with two decimal digits. Additional location specific data was added and merged with the race data, including average air and water temperatures in °Celsius (in integer form, that is, without any decimals), and the type of race course in each split discipline as categorical variables, including the values

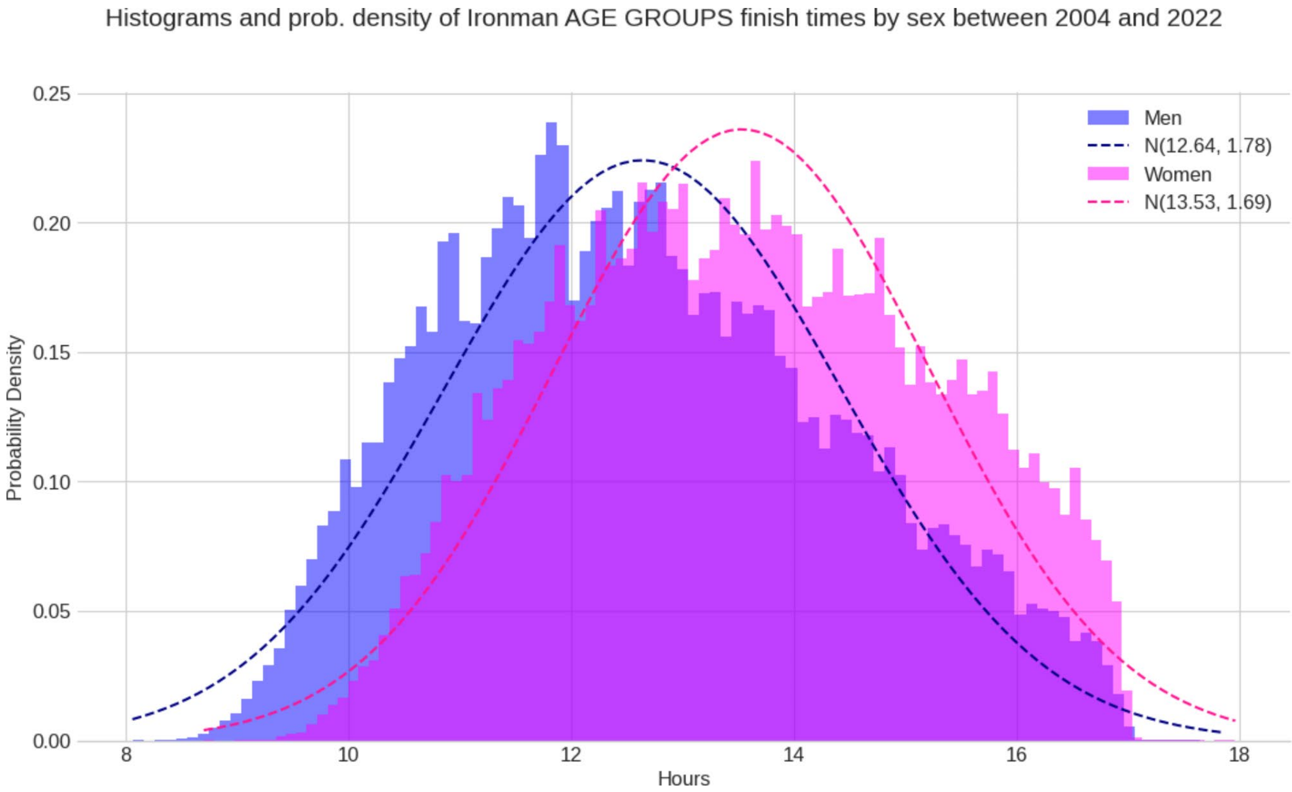

Fig. 1. Distribution of age groups by sex.

of rolling, hilly, and flat for the bike and run race courses, and lake, ocean, river, bay, reservoir for the different swim courses. The race records were separately aggregated by event location and by country, to produce two large ranking tables sorted by number of race records (*i.e.* participation). In doing so, we identified up to 228 different countries in the original data sample, many of them with 1, 2 or 3 records and hence with no statistical interest. We then decided to limit the analysis to the top 150 countries by number of records, which account for 99.94% of the full sample and includes countries with at least 13 race records, while eliminating noise, which is in turn good for computing, interpretation, and overall interest. After the pre-processing and merging of the data, the final dataset consisted of a total of 677,320 finishers´ records (*i.e.* 544,632 from men and 132,688 from women) from the top 150 countries by number of records, participating in 443 events over 65 different locations between 2002 and 2022.

Statistical analysis

Normality of the race time distributions by sex was checked by visual inspection of race time histograms and calculation of Gaussian overlapping curves. The statistical values of the overall race times (*i.e.* mean, std, max, and min) were calculated for each of the 65 locations and 150 countries and are displayed in the large ranking tables. The event location ranking table includes specific race course data, including the average air and water temperatures, and the type of race course. An analysis of the race performance by type of race course was done and the results are shown in the form of boxplot charts and accompanying 2-way ANOVA tests that suggest the statistical significance of the different types of swim, bike, and run race courses. The significance level was set at 0.05 in all cases. Several predictive modelling algorithms were tested, including a Multivariate Linear Regressor (MLR) and three Machine Learning (ML) Regressors, a Decision Tree, a Random Forest and a XG Boost Regressor. The predicted variable (*i.e.* target) was the overall race time (in hours) whilst a total of 17 predictors (*i.e.* features or factors) were used, including a number of categorical variables that had to be encoded before they could be used with the models. The variable sex (men/women) is encoded as 0 = women and 1 = men. The age group variable is encoded as an integer, representing 5-year groups, with group 18 representing less than 20 years, group 20 from 20 to 24, group 25 from 25 to 29, etc. until group 75 which includes any triathletes older than 75 years of age. The country and event location variables are encoded based on their position in the ranking lists, sorted by participation, and starting with zero. The event location average air and water temperature variables, of numerical type, are used as they are reported on the website of the organizers. The three categorical variables (*i.e.* swim, bike and run) are converted into dummy variables (binary flags) indicating the presence with 1 and absence with 0, becoming a set of 11 binary variables. Given the large size of the dataset, a hold-out test strategy was used, with 25% of the dataset reserved for model evaluation: 507,076 race records were used for model training, and 169,026 for model evaluation. For each of these models, the Mean Absolute Error (MAE) and the coefficient of determination (R^2) were calculated. After training, evaluating, and comparing the three models, XG Boost emerged as the best performer and we present only the results from XG Boost. Model interpretability

Country	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	Country_ID
United States	274521	13.36	1.75	8.29	17.96	0
United Kingdom	55410	13.04	1.65	8.53	17.08	1
Canada	38264	13.03	1.73	8.40	17.85	2
Australia	37571	12.26	1.67	8.65	17.06	3
Germany	32662	11.74	1.51	8.21	17.59	4
France	27873	12.11	1.54	8.32	17.03	5
Spain	17009	12.07	1.58	8.12	16.97	6
Sweden	14640	12.09	1.48	8.52	17.84	7
Brazil	14408	11.86	1.59	8.50	17.54	8
Austria	12792	11.68	1.50	8.62	16.94	9
Italy	12702	12.09	1.48	8.60	16.84	10
Mexico	11941	13.23	1.65	8.61	17.09	11
Denmark	11874	11.59	1.40	8.38	17.13	12
New Zealand	11092	12.69	1.75	8.06	17.64	13
Japan	10249	13.45	1.82	8.33	17.01	14
South Africa	9675	13.29	1.65	8.70	16.97	15
Belgium	8444	11.48	1.47	8.27	17.00	16
Switzerland	7570	11.62	1.49	8.40	16.89	17
Ireland	6336	12.54	1.65	8.43	17.04	18
Argentina	5470	12.10	1.59	8.56	17.32	19

Table 1. List of origin of the athletes sorted by country with the number of finishers. The race times are color-coded where darker fields represent faster race times.

tools like SHAP or PDP libraries were used to further understand how, according to the models, each predictor influenced the race finish time. All data processing and analysis were performed using Python (www.python.org/) and a Google Colab notebook (https://colab.research.google.com/).

Country	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	Country_ID
Netherlands	4750	11.83	1.50	8.28	16.99	20
Finland	3497	11.68	1.40	8.54	16.59	21
Poland	3438	12.01	1.52	8.78	16.75	22
Russia	3163	11.91	1.51	8.37	17.05	23
Israel	2952	12.74	1.59	8.67	16.89	24
Philippines	2725	14.31	1.59	9.22	17.00	25
Norway	2314	11.77	1.53	8.46	16.79	26
Singapore	2250	14.13	1.68	9.01	17.25	27
South Korea	2237	13.69	1.65	9.38	17.00	28
Czechia	1709	11.48	1.49	8.54	16.61	29
China	1702	13.70	1.78	8.50	16.99	30
Malaysia	1635	14.83	1.48	9.96	17.15	31
Colombia	1397	12.60	1.79	8.12	17.05	32
Portugal	1284	11.87	1.54	8.64	16.46	33
Chile	1219	12.04	1.60	8.79	16.70	34
Taiwan	1163	13.97	1.85	9.46	17.26	35
Costa Rica	1037	12.39	1.59	8.41	16.90	36
Puerto Rico	965	13.39	1.65	8.29	17.01	37
Estonia	912	11.36	1.36	8.34	16.79	38
Venezuela	868	12.55	1.65	8.56	16.94	39

Figure 1. (continued)

Country	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	Country_ID
Thailand	862	14.49	1.50	9.57	16.99	40
Slovenia	804	11.33	1.45	8.70	16.85	41
Hong Kong	759	13.27	1.92	9.00	17.00	42
India	697	14.27	1.57	9.67	16.95	43
Guatemala	665	12.94	1.67	8.84	17.01	44
Ecuador	653	12.62	1.58	8.48	16.73	45
Ukraine	631	11.79	1.48	8.34	16.60	46
Greece	617	12.17	1.65	8.95	16.97	47
Peru	568	12.81	1.68	8.86	16.92	48
Hungary	559	11.98	1.66	8.88	16.97	49
Slovakia	526	11.97	1.54	9.00	16.57	50
Türkiye	398	12.04	1.59	9.04	16.90	51
Iceland	382	12.16	1.57	8.66	16.72	52
Panama	376	12.10	1.90	8.76	16.68	53
Croatia	346	12.10	1.67	8.98	16.42	54
Romania	334	12.64	1.71	8.90	16.75	55
Luxembourg	302	11.46	1.47	8.69	15.94	56
Indonesia	301	14.36	1.56	9.49	16.96	57
Kazakhstan	293	12.36	1.61	8.81	16.87	58
Uruguay	247	12.55	1.68	9.24	16.85	59

Figure 1. (continued)

Country	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	Country_ID
Latvia	233	11.39	1.47	8.50	16.88	60
Lithuania	227	12.08	1.53	9.51	16.99	61
Morocco	227	12.56	1.46	9.92	16.92	62
Malta	204	12.06	1.61	9.01	16.79	63
Dominican Republic	179	13.08	1.73	9.41	17.40	64
El Salvador	178	13.01	1.82	8.78	16.88	65
Paraguay	172	12.31	1.81	9.02	16.87	66
Andorra	166	11.56	1.34	9.62	16.48	67
Bolivia	147	13.09	1.77	9.76	16.82	68
United Arab Emirates	140	12.54	1.66	9.10	16.76	69
Serbia	134	12.06	1.72	8.95	16.23	70
Egypt	115	13.25	1.63	9.82	16.85	71
Bermuda	111	11.98	1.53	9.82	16.10	72
Bulgaria	110	13.22	1.67	9.48	16.97	73
Vietnam	103	14.44	1.60	10.74	16.99	74
New Caledonia	95	12.29	1.59	9.60	16.99	75
Cyprus	95	12.25	1.43	9.88	16.65	76
Honduras	90	12.90	1.70	9.70	16.64	77
Cayman Islands	89	13.30	1.70	9.20	16.56	78
Reunion	84	12.64	1.48	9.89	16.25	79

Figure 1. (continued)

Country	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	Country_ID
Bosnia and Herzegovina	74	12.52	1.60	9.50	15.84	80
Lebanon	74	12.87	1.56	9.53	16.86	81
Aruba	70	13.44	1.48	10.99	16.81	82
Belarus	68	11.86	1.62	8.77	16.35	83
Armenia	68	13.10	1.70	9.67	16.85	84
Zimbabwe	67	13.20	1.73	9.64	16.78	85
Eswatini	67	11.75	1.69	9.34	16.04	86
Bahrain	63	12.70	2.00	9.00	16.94	87
Iran	62	13.50	1.79	9.27	16.73	88
Guam	61	13.93	2.18	9.55	16.87	89
Brunei Darussalam	55	13.57	1.70	9.39	16.90	90
Monaco	54	11.74	1.48	9.46	15.61	91
Jersey	52	12.55	1.73	8.91	15.74	92
Guadeloupe	51	12.68	1.40	10.37	15.68	93
Sierra Leone	50	13.18	1.51	10.16	16.50	94
Northern Mariana Islands	50	11.95	1.46	9.39	16.00	95
United States Minor Outlying Islands	48	13.35	1.87	9.66	16.83	96
Moldova	43	12.79	1.72	9.48	16.75	97
Trinidad and Tobago	43	14.74	1.31	11.03	16.79	98

Figure 1. (continued)

Results

A total of 677,320 IRONMAN finishers´ records (544,632 from men and 132,688 from women) from the top 150 countries by number of records, participating in 443 IRONMAN events over 65 different locations between 2002 and 2022 were analyzed.

Distributions of race finish times by sex

Figure 1 shows the original histograms (bins = 100) of the IRONMAN overall race times by sex, along with calculated (overlapped) Gaussian envelopes.

Country	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	Country_ID
Cuba	43	13.58	1.83	9.87	17.01	99
Samoa	42	14.14	1.92	11.02	16.95	100
Qatar	40	12.21	1.75	9.47	15.82	101
Tunisia	39	13.42	1.74	9.91	16.93	102
DR Congo	39	13.64	1.77	9.86	16.67	103
Botswana	38	13.31	1.58	10.06	17.01	104
Saudi Arabia	37	12.95	1.63	9.97	16.77	105
Isle of Man	36	12.83	2.08	8.79	16.25	106
Nicaragua	36	14.55	1.59	11.07	16.89	107
Uzbekistan	36	13.71	2.22	9.96	16.73	108
Kyrgyz Republic	35	11.60	0.93	9.92	13.27	109
United States Virgin Islands	35	13.74	1.48	10.08	16.67	110
Azerbaijan	34	13.01	1.61	10.26	16.60	111
French Polynesia	33	12.31	1.47	10.22	15.52	112
Bahamas	32	12.52	1.88	8.85	16.76	113
Gibraltar	31	12.46	1.36	10.32	15.68	114
Guernsey	28	12.81	1.96	9.86	16.57	115
Albania	27	12.95	1.61	10.58	16.22	116
Jamaica	26	13.19	2.15	9.33	16.69	117
Kuwait	25	13.85	1.89	10.60	16.83	118

Figure 1. (continued)

Country	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	Country_ID
Faroe Islands	24	12.05	1.64	9.78	14.74	119
Martinique	24	12.09	2.05	9.46	16.02	120
Liechtenstein	23	12.00	1.62	9.96	16.27	121
Georgia	22	12.67	1.77	10.08	16.81	122
Montenegro	22	11.66	1.61	9.08	16.00	123
Sri Lanka	22	12.72	1.78	10.21	15.90	124
Angola	21	12.20	1.58	10.10	15.68	125
Bangladesh	21	13.77	1.64	11.54	16.92	126
Kenya	21	12.87	1.66	10.48	16.46	127
Jordan	21	13.23	1.80	10.84	15.78	128
Antarctica	20	13.01	1.90	9.35	16.89	129
Cook Islands	20	13.46	1.63	10.75	16.62	130
Pakistan	20	14.84	1.76	11.90	16.88	131
Mauritius	20	12.95	1.99	9.89	16.47	132
British Virgin Islands	19	13.66	1.23	12.05	16.82	133
American Samoa	19	14.09	1.85	10.05	16.89	134
North Macedonia	19	12.29	1.41	10.16	15.05	135
Vatican	18	14.42	1.71	10.77	16.63	136
Fiji	18	13.78	1.65	10.47	16.58	137
Greenland	17	12.43	1.59	9.96	15.10	138

Figure 1. (continued)

Country	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	Country_ID
Barbados	17	13.33	1.85	9.66	16.29	139
Syria	16	13.98	1.25	11.59	15.84	140
San Marino	15	11.88	0.98	10.29	13.52	141
Algeria	15	12.14	1.67	10.19	15.36	142
Afghanistan	15	13.66	2.07	9.33	16.62	143
Ghana	14	14.04	1.66	10.44	16.77	144
French Guiana	14	11.69	1.40	9.35	14.81	145
Curacao	13	13.29	1.73	10.89	16.62	146
Mongolia	13	12.54	0.94	10.89	14.20	147
Myanmar	13	14.12	1.53	11.08	16.36	148
Gabon	13	13.08	1.84	10.24	15.07	149

Figure 1. (continued)

Ranking tables of event locations and tri-athletes' countries of origin
The country (Table 1) and the event location (Table 2) ranking tables below show the statistical details of the 150 countries of origin and the 65 race locations considered in this study. The tables are sorted by number of records, so countries (or event locations) with higher participation are at the top. Overall race time statistical values (mean, std, min and max) are displayed in hours, and the mean overall race time column is color coded, with darker cells indicating lower (better) race times. The event location table additionally shows the characteristics of each race course, through the categorical variables swim, bike, run, and the average temperatures for water (°C) and air (°C).

Most of the athletes originated from the USA, followed by United Kingdom, Canada, Australia, Germany, France, Spain, Sweden, Brazil, Austria, and Italy for the 10 first countries.

Figure 2 shows the difference between women and men regarding the race course characteristics for swimming with swimming in a bay, in the ocean, a river, a lake or a reservoir where men were always faster than women. Also for cycling and running (*i.e.* flat, rolling, or hilly), men were always faster than women.

Multi linear regression (MLR) ordinary least squares (OLS) regressor
The OLS MLR model (Table 3) obtains a R^2 score of 0.128 and the analysis indicates all features have a statistically significant effect on the finish time variable ($p < 0.05$). No differentiation is made between test and train subsets for the MLR, but the full sample is used for training and evaluation (in-sample testing).

Decision tree and random forest regressors
The first two ML models, attempted over several configurations, obtained best values of MAE of 1.27 and 1.26 h, respectively. They both obtain a best R^2 score of 0.25. This is an improvement over the linear MLR model, but still a low R^2 score (weak effect) by any measure.

XG boost regressor
The XGB model is built with 50 estimators, max depth of 9 and learning rate of 0.3, to obtain a R^2 score of 0.27 over the 25% held-out test set. Whilst the highest score of the three ML models tested, this is still a low R^2 score meaning the model can only explain 27% of the variability in the predicted variable, and that more predictors should be considered, if any predictive expectations are to be had. But we can still use the tools in our model interpretability toolset to look for insights into what the model learnt.

EventLocation	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	EventLocation_ID	Swim	Bike	Run	Water (°C)	Air (°C)
IRONMAN® Wisconsin	38540	13.44	1.67	9.00	17.37	0	lake	rolling	rolling	21	22
IRONMAN® Florida	38142	12.96	1.80	8.54	17.47	1	ocean	flat	flat	22	19
IRONMAN® Lake Placid	34328	13.35	1.68	8.94	17.17	2	lake	hilly	rolling	22	17
IRONMAN® Arizona	34237	13.35	1.75	8.39	17.20	3	lake	rolling	rolling	16	17
IRONMAN® Hawaii	32143	11.72	1.86	8.41	17.84	4	ocean	rolling	rolling	24	27
IRONMAN® Austria	30953	11.92	1.59	8.56	17.18	5	lake	rolling	flat	23	27
IRONMAN® France	29277	12.39	1.53	8.34	16.84	6	ocean	hilly	flat	24	28
IRONMAN® Canada	27126	13.25	1.73	9.09	17.10	7	lake	hilly	rolling	21	26
IRONMAN® Coeur d'Alene	24522	13.47	1.71	9.09	17.65	8	lake	hilly	rolling	20	24
IRONMAN® Louisville	21651	13.36	1.70	8.81	17.61	9	river	rolling	rolling	27	29
IRONMAN® Texas	21119	13.31	1.93	8.12	17.00	10	lake	flat	rolling	28	23
IRONMAN® Frankfurt	20052	12.03	1.38	8.73	15.90	11	lake	rolling	flat	23	25
IRONMAN® Cozumel	17881	13.06	1.77	8.12	17.58	12	ocean	flat	flat	26	26
IRONMAN® Copenhagen	16732	11.68	1.38	8.32	16.12	13	bay	rolling	flat	18	21
IRONMAN® New Zealand	16608	12.85	1.75	8.06	17.64	14	lake	rolling	flat	18	21
IRONMAN® Mont-Tremblant	16293	13.06	1.66	8.52	17.13	15	lake	hilly	flat	17	26
IRONMAN® UK	15211	13.39	1.61	8.63	17.04	16	river	hilly	rolling	18	22
IRONMAN® Kalmar	14638	12.08	1.47	8.49	16.26	17	ocean	flat	flat	20	22
IRONMAN® Lanzarote	14523	13.08	1.67	8.95	17.18	18	ocean	hilly	rolling	19	24
IRONMAN® Brazil Florianopolis	14009	11.80	1.52	8.61	17.54	19	ocean	flat	flat	20	21
IRONMAN® Australia - New South Wales	13655	12.58	1.65	8.65	17.03	20	river	rolling	flat	23	21
IRONMAN® Zurich Switzerland	13194	12.36	1.53	8.88	16.03	21	lake	hilly	flat	19	23

Table 2. List of races sorted by the number of finishers. The race times are color-coded where darker fields represent faster race times.

SHAP values and features importances for the XG boost model

The SHAP aggregated values chart in Fig. 3 shows how each predicting variable influences the model output, with the factors rated as more important at the top. The country of origin is the most important predictor. The variable age group comes up as the second most important feature but the one that best separates data points. Red dots (*i.e.* high or older age groups) contribute positively to the race time, whilst shades of purples and blues increasingly move to the left, deducting from the race times. Sex and the location of the event are the next

EventLocation	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	EventLocation_ID	Swim	Bike	Run	Water (°C)	Air (°C)
IRONMAN® Wales	12491	13.54	1.56	9.48	17.05	22	ocean	hilly	hilly	17	16
IRONMAN® Barcelona	11125	11.78	1.43	8.37	16.02	23	lake	hilly	rolling	23	21
IRONMAN® Chattanooga	10781	13.46	1.62	8.89	17.09	24	river	rolling	rolling	25	24
IRONMAN® Western Australia	10643	12.42	1.78	8.63	17.25	25	bay	flat	flat	20	27
IRONMAN® South Africa	10396	13.16	1.71	8.70	17.01	26	bay	rolling	rolling	20	23
IRONMAN® Boulder	8490	13.58	1.78	8.67	17.03	27	reservoir	hilly	hilly	17	26
IRONMAN® Maryland	8389	13.32	1.80	8.64	17.25	28	river	flat	flat	22	24
IRONMAN® Cairns	8243	12.49	1.69	8.52	17.01	29	ocean	rolling	flat	23	26
IRONMAN® Emilia Romagna	8061	11.96	1.56	8.58	15.98	30	ocean	flat	flat	22	25
IRONMAN® Vichy	7417	12.32	1.62	8.40	16.93	31	lake	hilly	flat	25	28
IRONMAN® Hamburg	6653	11.85	1.50	8.21	15.81	32	lake	flat	rolling	21	23
IRONMAN® Mallorca	6257	11.90	1.57	8.28	16.40	33	ocean	hilly	flat	17	25
IRONMAN® Malaysia	5400	13.94	1.88	9.08	17.00	34	ocean	hilly	flat	28	31
IRONMAN® Melbourne	4665	12.01	1.70	8.80	17.06	35	bay	flat	flat	17	18
IRONMAN® Santa Rosa	4654	13.34	1.85	8.70	16.99	36	reservoir	rolling	flat	18	23
IRONMAN® Maastricht	4063	12.16	1.53	8.93	16.67	37	river	rolling	flat	20	27
IRONMAN® St. George	3704	13.84	1.69	9.44	16.96	38	reservoir	hilly	hilly	16	23
IRONMAN® Japan	3514	13.88	1.72	9.65	17.00	39	ocean	flat	flat	21	18
IRONMAN® Tallinn	3327	11.73	1.60	8.46	16.99	40	lake	rolling	flat	18	17
IRONMAN® Regensburg	2823	11.76	1.48	8.71	15.92	41	lake	rolling	flat	22	26
IRONMAN® Los Cabos	2644	13.34	1.77	9.05	17.01	42	ocean	rolling	flat	29	25
IRONMAN® Gurye Korea	2607	13.81	1.66	9.35	16.99	43	lake	hilly	flat	23	23
IRONMAN® Lake Tahoe	2604	14.03	1.71	9.66	16.99	44	lake	hilly	hilly	18	22

Figure 2. (continued)

most important predictors. Further down but with clear separation of red and blue data points, hilly running and cycling race courses add to race time (*i.e.* slower race times), whilst flat surfaces, rolling cycling and ocean swimming deduct from it (*i.e.* faster race times).

Partial dependence plots (PDP) of the XG boost model

The PDP chart is another tool we have to look into our model. PDP charts show how the output of the model varies for each numerical predicting variables (features or factors). According to the XG Boost model PDP charts, men are on average ~ 0.8 h faster than women (Fig. 4), and the fastest athletes are aged 25—34 years

EventLocation	Records count	Race time (mean)	Race time (std)	Race time (min)	Race time (max)	EventLocation_ID	Swim	Bike	Run	Water (°C)	Air (°C)
IRONMAN® Tulsa	2309	13.73	1.91	9.08	17.96	45	lake	rolling	flat	23	23
IRONMAN® New York	1955	13.03	1.64	9.37	16.86	46	bay	flat	flat	18	21
IRONMAN® Taiwan	1939	13.87	1.84	9.27	17.26	47	ocean	rolling	flat	20	22
IRONMAN® Brazil Fortaleza	1754	12.73	1.70	9.20	17.01	48	ocean	flat	flat	28	31
IRONMAN® Vineman	1697	13.48	1.73	9.23	16.99	49	river	rolling	rolling	19	24
IRONMAN® Vitoria-Gasteiz	1572	11.59	1.36	8.77	15.71	50	lake	rolling	flat	21	22
IRONMAN® Mar del Plata	1491	12.16	1.61	8.78	16.91	51	ocean	rolling	rolling	19	20
IRONMAN® Portugal-Cascais	1389	12.31	1.57	8.94	15.95	52	bay	hilly	rolling	17	18
IRONMAN® Subic Bay Philippines	1261	14.54	1.64	9.56	17.00	53	bay	hilly	flat	28	33
IRONMAN® Indiana	1223	13.32	1.73	8.75	16.43	54	reservoir	flat	rolling	23	28
IRONMAN® Muskoka	1046	13.20	1.74	9.32	16.84	55	lake	rolling	hilly	20	26
IRONMAN® Switzerland Thun	883	11.93	1.56	8.48	16.45	56	lake	hilly	flat	17	20
IRONMAN® Des Moines	843	13.82	2.01	8.86	17.00	57	lake	rolling	flat	22	22
IRONMAN® Haugesund Norway	825	12.37	1.54	8.81	16.26	58	lake	rolling	flat	17	18
IRONMAN® Weymouth	699	13.08	1.53	9.52	16.44	59	ocean	rolling	flat	17	15
IRONMAN® Waco	639	14.02	1.77	8.69	17.33	60	river	flat	rolling	18	18
IRONMAN® Finland Kuopio-Tahko	622	11.98	1.42	9.15	15.83	61	lake	rolling	rolling	18	18
IRONMAN® Gdynia	590	12.30	1.56	8.88	16.33	62	ocean	hilly	hilly	17	16
IRONMAN® China	415	13.24	1.83	9.48	16.90	63	ocean	flat	flat	24	32
IRONMAN® Pays D'Aix	383	12.34	1.58	8.78	15.98	64	lake	hilly	rolling	18	20

Figure 2. (continued)

(Fig. 5). The XGBoost model shows that a representative set of European countries including Germany, Austria, Denmark, Belgium, Switzerland, Norway, Czechia, Estonia, and Slovenia are the fastest. The USA and a group of Asian countries including Philippines, Malaysia, and Thailand appear to be the slowest (Fig. 6). IRONMAN Hawaii is the IRONMAN race location with the fastest race times, but also IRONMAN Vitoria-Gasteiz and IRONMAN Hamburg are singled out by the XG Boost model among the fastest race courses (Fig. 7). Regarding temperatures, optimal air temperature ranged at 19–21 or 25–28°Celsius (Fig. 8), and optimal water temperatures at 23–25°Celsius (Fig. 9).

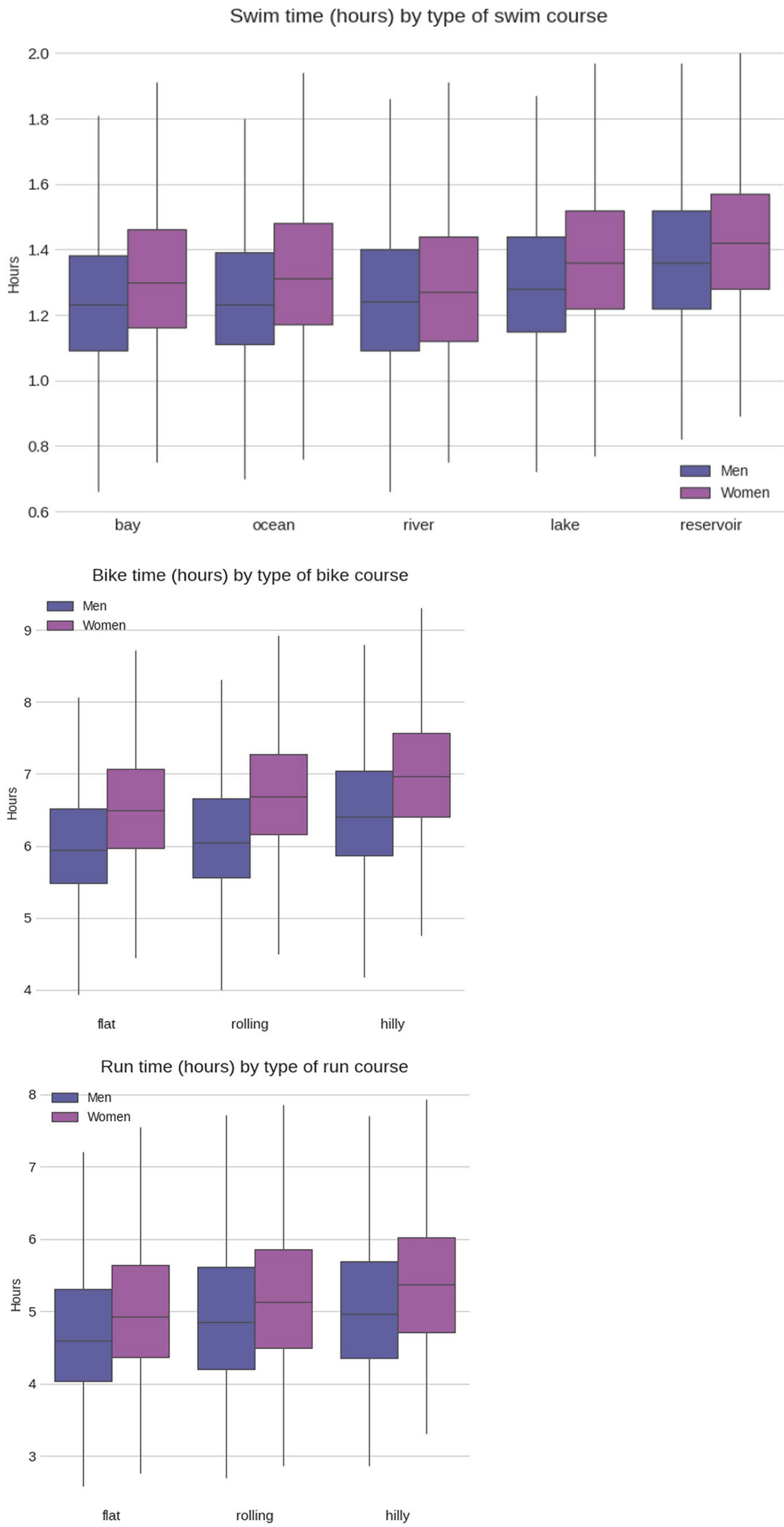

Fig. 2. Differences between women and men considering race course characteristics.

Dep. Variable:	FinishTime	R-squared:	0.128
Model:	OLS	Adj. R-squared:	0.128
Method:	Least Squares	F-statistic:	7079
Prob (F-statistic):	0.00		
No. Observations:	677,320		
Df Residuals:	677,305		
Df Model:	14		
	coef	**t**	**P > |t|**
const	6.8208	655.161	0.000
Sex_ID	−0.8765	−168.384	0.000
AgeGroup_ID	0.0385	177.219	0.000
Country_ID	−0.0043	−22.570	0.000
EventLocation_ ID	−0.0004	−2.226	0.026
Water (°C)	0.0240	25.732	0.000
Air (°C)	−0.0392	− 51.199	0.000
Swim_bay	1.1064	148.235	0.000
Swim_lake	1.3009	261.018	0.000
Swim_ocean	1.0703	198.373	0.000
Swim_reservoir	1.6712	130.301	0.000
Swim_river	1.6719	234.717	0.000
Run_flat	1.9114	300.527	0.000
Run_hilly	2.6612	290.393	0.000
Run_rolling	2.2482	360.031	0.000
Bike_flat	2.3145	417.152	0.000
Bike_hilly	2.4471	508.297	0.000
Bike_rolling	2.0591	451.106	0.000

Table 3. Results of the OLS linear regressor.

Air and water temperature versus race time—3D interaction charts by model
Figure 10 is an interesting set of 3D charts comparing how each of the three ML models "has learnt" the partial dependence of the race time (target) with the air and water temperature variables together. The third and best performing XG Boost model shows a more detailed (higher resolution) PDP chart.

Discussion
This study aimed to identify the dominant nationalities for nonprofessional IRONMAN triathlon competitions between 2002 and 2020 with the hypothesis that the fastest IRONMAN age group triathletes would originate from the USA. The most important findings were (i) European countries (*i.e.* Germany, Austria, Denmark, Belgium, Switzerland, Norway, Czechia, Estonia, and Slovenia) have the fastest athletes, (ii) IRONMAN Hawaii, IRONMAN Vitoria-Gasteiz and IRONMAN Hamburg are the fastest races, (iii) optimal air temperature for cycling and running ranged between 19 °C and 21 °C or at 25–28 °C and optimal water temperature for swimming was at 23–25 °C, (iv) the fastest athletes were 25–34 years old, and (v) men were ~ 0.9 h faster than women. The discussion of these findings is challenging, especially due to the lack of evidence in the scientific literature. However, the main finding highlights the importance of adopting similar approaches in order to identify the most successful countries in sports competitions.

European athletes were the fastest
Our first important finding was that European IRONMAN triathletes were the fastest and we could, therefore, not confirm our hypothesis. Unlike the present findings, previous studies investigating the amateur IRONMAN triathletes competing in IRONMAN Hawaii showed that those from North America showed the best results, finishing among the top five athletes[23]. In this study, the authors considered the total number of participants and the number of athletes in the top five for both sexes, which can be influenced by the place of competition[23]. The hosting effect has been discussed in the scientific literature as an important performance determinant, especially for family support, fans, and familiarization with environmental characteristics[30,31]. However, hosting effects are specific for some sports[32], demanding additional efforts to understand the hosting effect in triathlon competitions.

We considered only the 150 best countries. These results should be considered in light of the limitations of the study. For example, we consider the mean values to determine the fastest country, which does not represent the totality of the athletes. Therefore, adopting a country-level analysis has important practical implications for

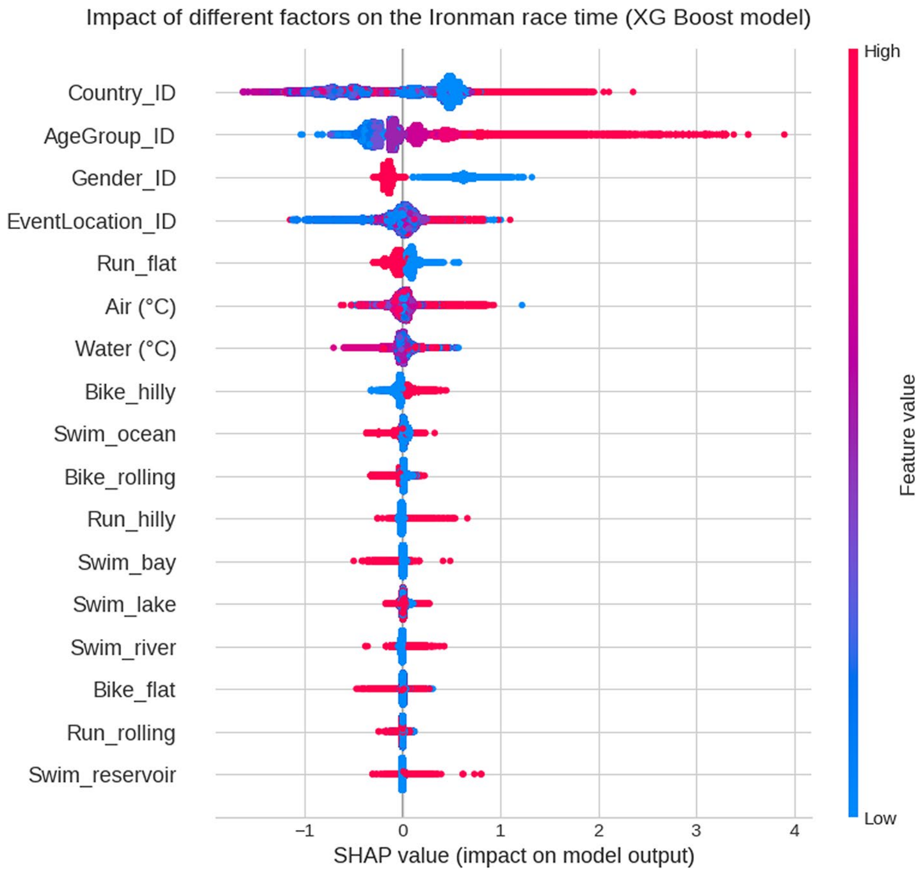

Fig. 3. SHAP aggregated values for XG Boost Regression model.

the sports sectors in these countries, especially considering the interest of the population in the practice. No previous studies that investigated the fastest nationalities among IRONMAN triathletes found similar results[23,33], which impairs the comparisons. However, some similarities between the countries should be considered: the three countries are part of the Northern region, presenting similar population size and economic characteristics. These similarities can influence the sports practice among the adult population and reflect the results achieved at triathlon competitions.

For the top countries by the number of age group records, Germany, Austria, Denmark, Belgium, Switzerland, Norway, Czechia, Estonia, and Slovenia were the fastest. These results are similar to those found by professional IRONMAN 70.3 triathletes[21]. The emergence of Germany, Austria, Denmark, Belgium, Switzerland, Norway, Czechia, Estonia, and Slovenia within the IRONMAN context may be attributed to a combination of unique factors such as geography, climate, and culture[34,35]. Many participants are drawn to IRONMAN triathlons as a means to improve their physical capabilities, seeking personal growth and the satisfaction of overcoming a multifaceted challenge[36]. The competitive aspect also plays a role, with participants aiming to test their limits, set personal records, and, in some cases, compete in organized events. This broader inclusivity, coupled with the diverse motivations behind participation, underscores the evolving and accessible nature of triathlons in contemporary society.

The fastest race courses

A further important finding was that IRONMAN Hawaii was the fastest race course, followed by European race courses such as IRONMAN Vitoria-Gasteiz and IRONMAN Hamburg. The finding that IRONMAN Hawaii is the fastest race course for age group athletes is explained by the fact that IRONMAN Hawaii is the World Championship for IRONMAN triathletes[37] where only the best triathletes can compete after qualification for the IRONMAN World Championship[38].

The aspect of environmental conditions

An important finding was that optimal air temperature ranged between 19 °C and 21 °C or at 25–28 °C and optimal water temperature was at 23–25 °C. To date, we have no specific knowledge for the 'best' temperatures to compete in an IRONMAN triathlon. It is, however, well known that environmental conditions have a considerable influence on endurance performance in running[39] and triathlon[40] where especially high temperatures impair endurance performance[41]. Regarding IRONMAN Hawaii, it has been reported that body core temperature increased during the marathon where an increase in body core temperature appeared to make triathletes run

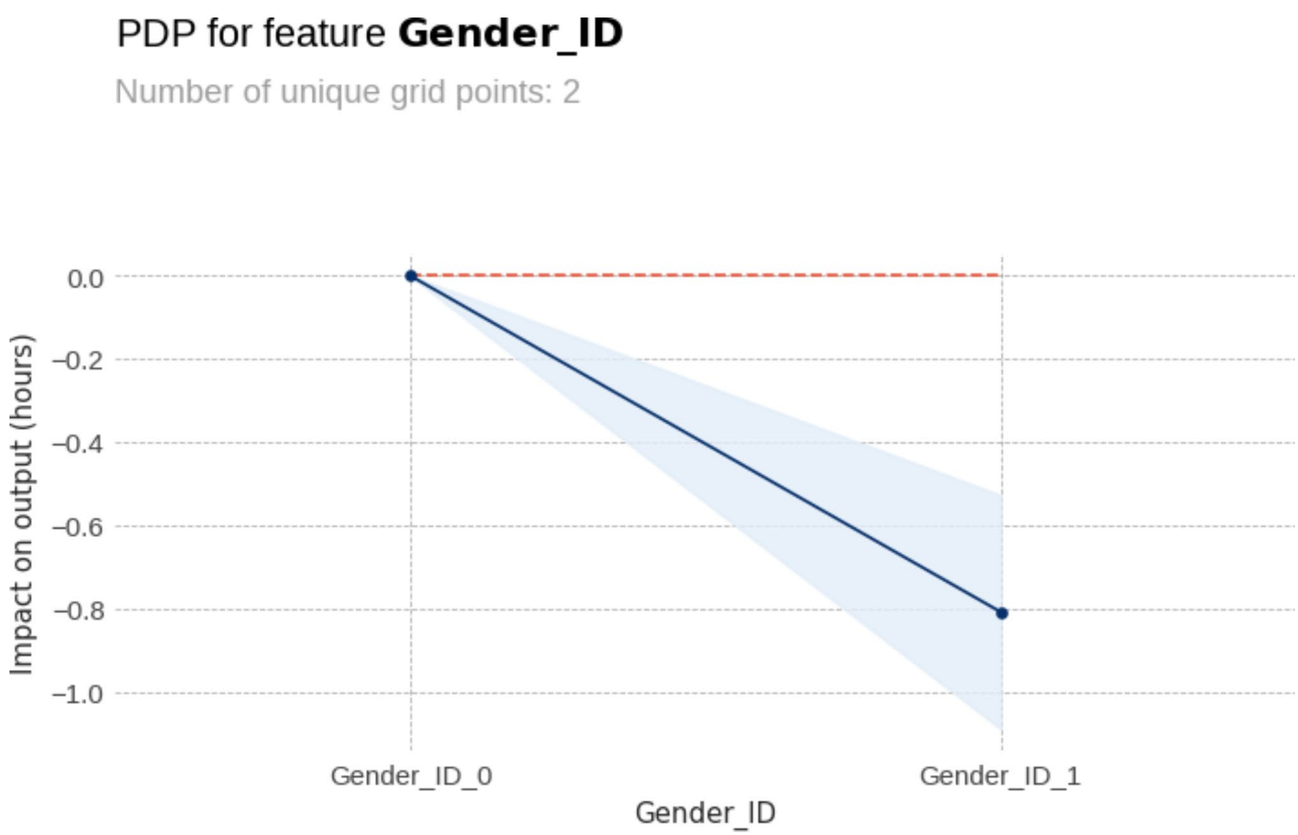

Fig. 4. PDP charts for sex.

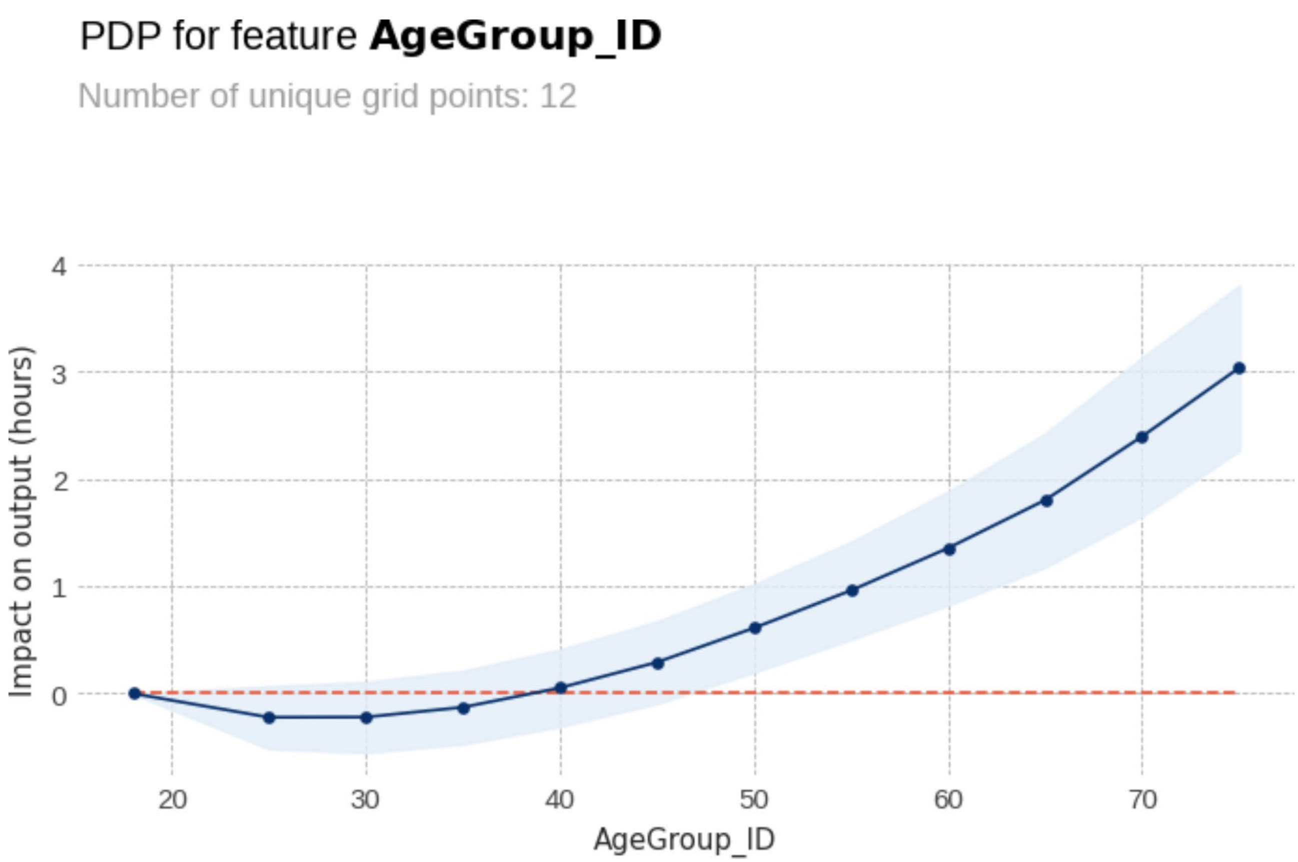

Fig. 5. PDP charts for age group.

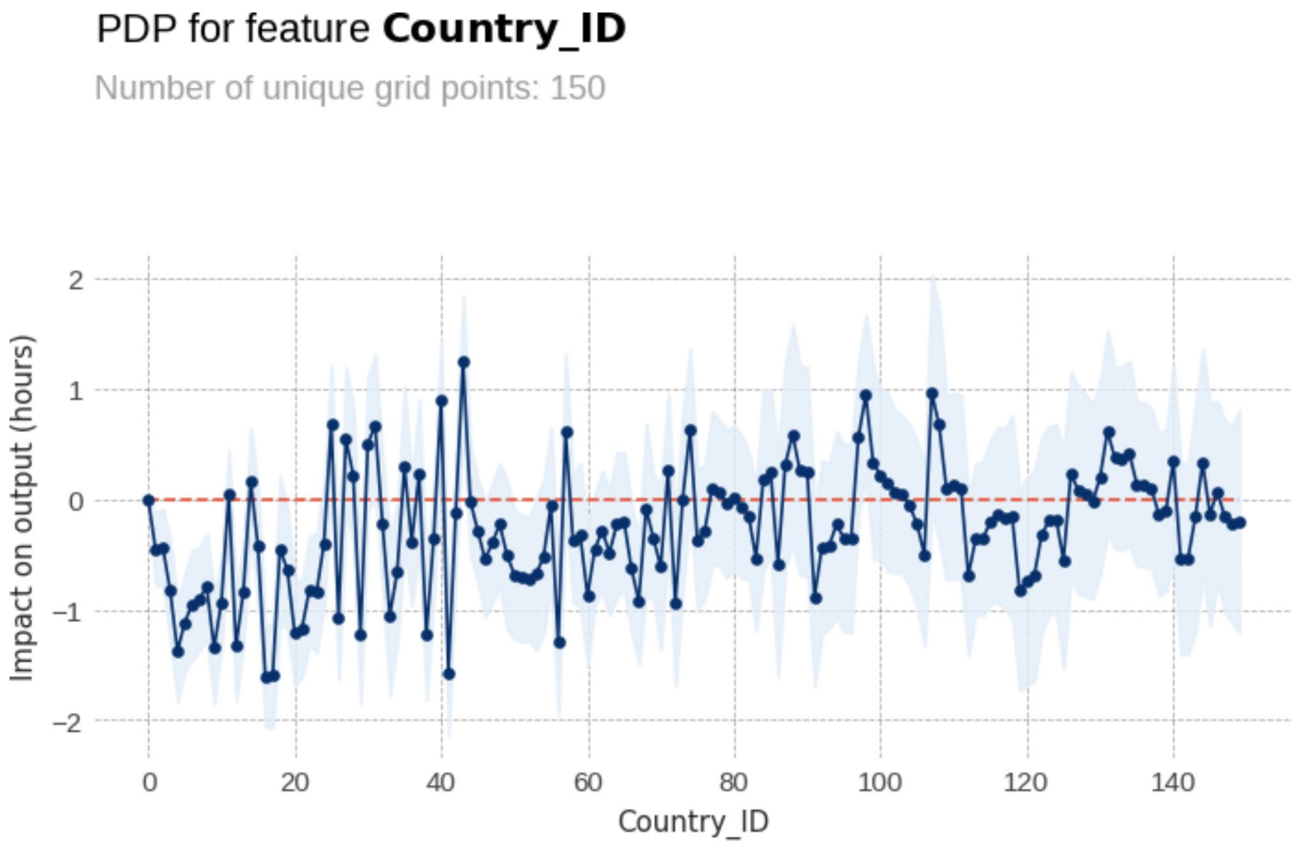

Fig. 6. PDP charts for country of origin of the athletes.

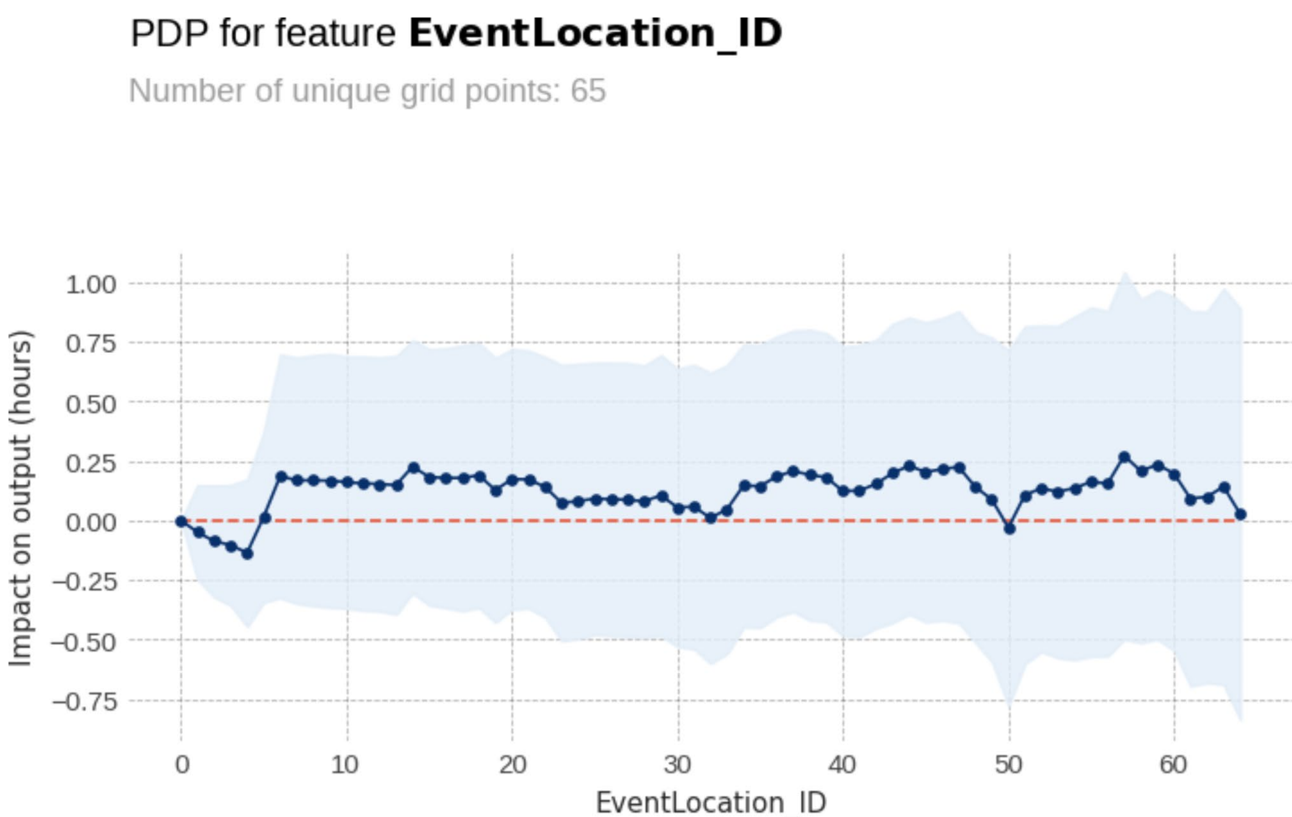

Fig. 7. PDP charts for the location where the race was held.

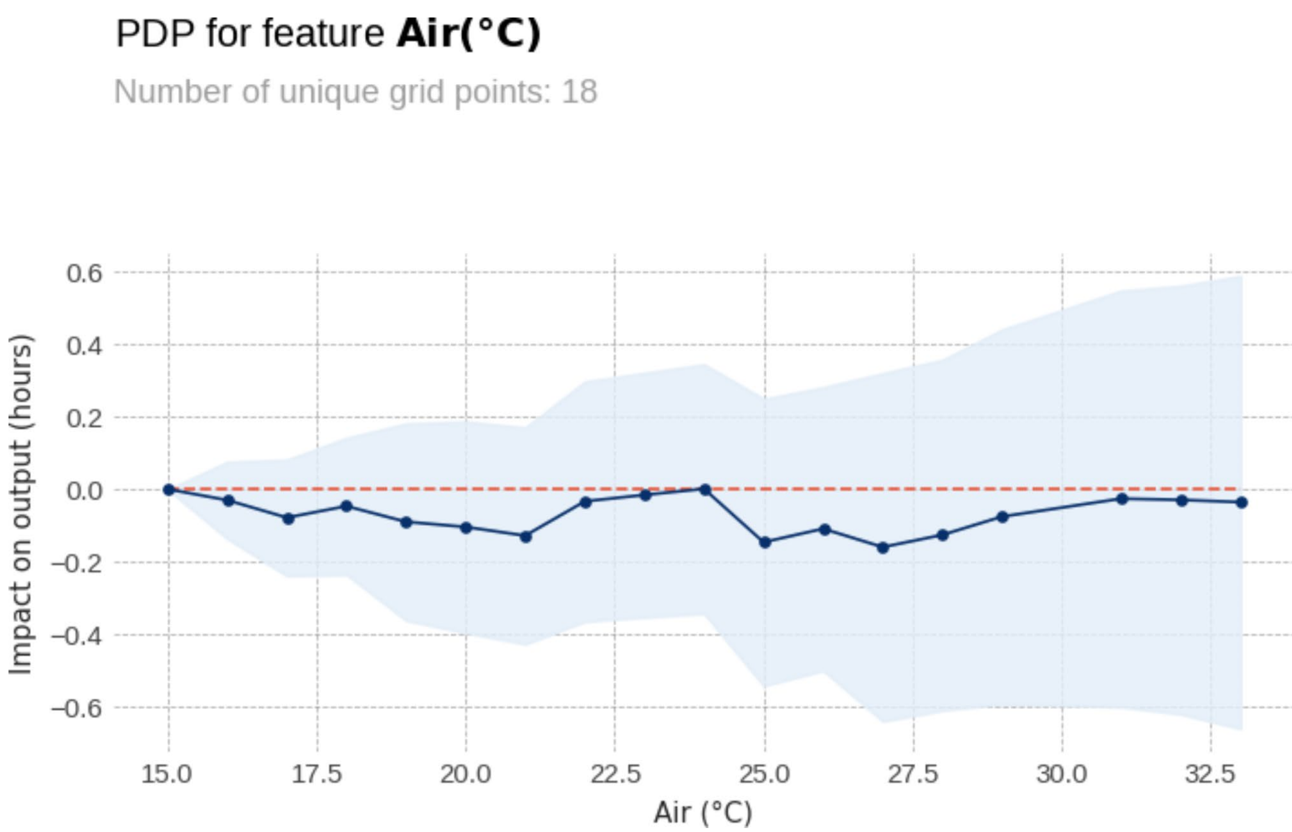

Fig. 8. PDP charts for ambient air temperature during race day.

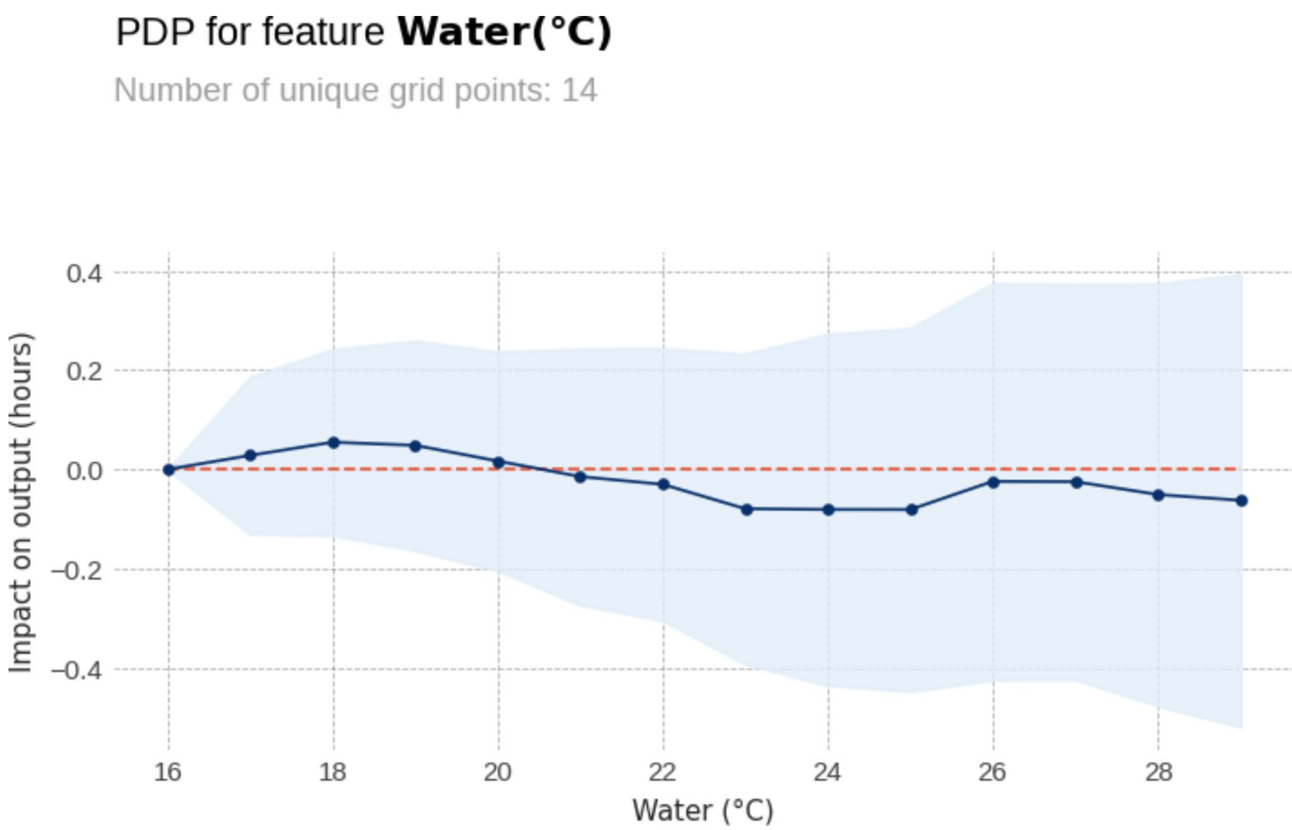

Fig. 9. PDP charts for water temperature in the swim course during race day.

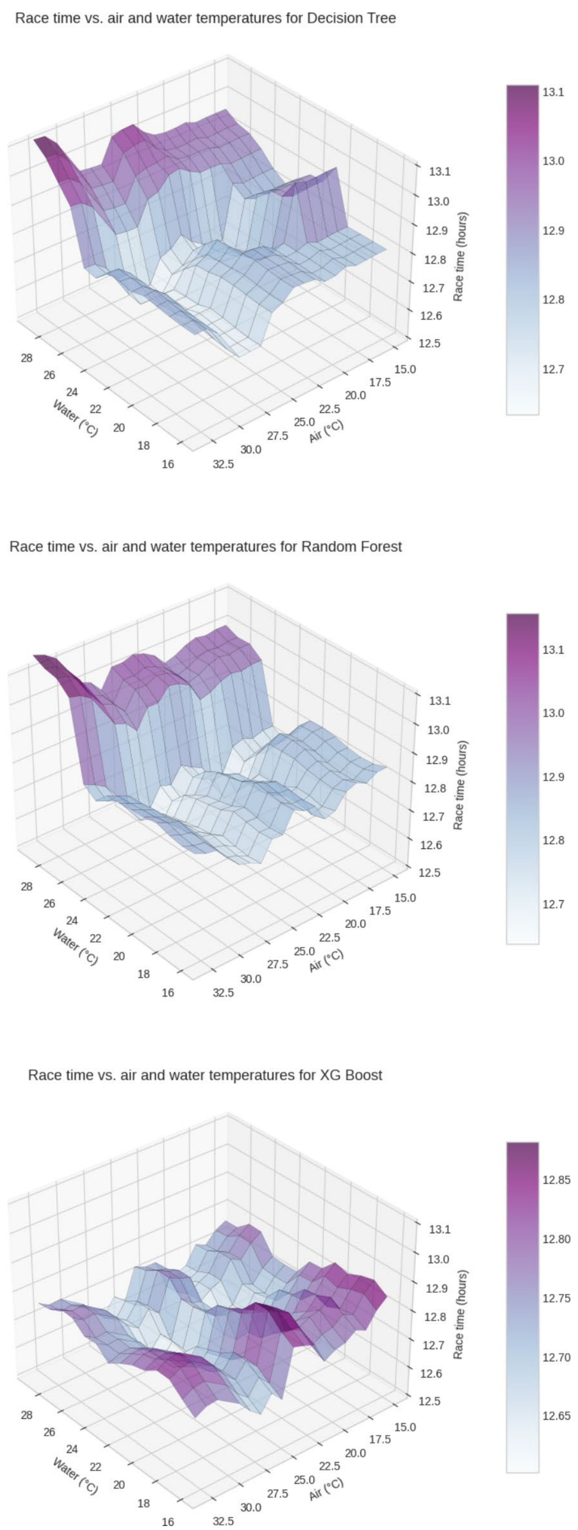

Fig. 10. 3D interaction air and water temperature with race time.

more slowly[42]. The present study shows the optimum race temperatures for both cycling and running where athletes can now select the most appropriate race course for a fast IRONMAN race time.

Regarding water temperature, it is well known that water temperature has a direct effect on swimming performance[43]. However, little is described in the scientific literature. A very recent study reported that swimming in a river had in female triathletes a greater effect on overall race time than cycling or running[40]. Whit the present findings, triathletes can also better select an IRONMAN race regarding the swim course temperatures and the influence on overall race time.

The influence of race course characteristics

A last important finding was the IRONMAN race course characteristic had a considerable influence on overall race time where slower race times were achieved with hilly cycling and running courses whilst flat surfaces, rolling cycling and ocean swimming were leading to faster race times. It is well known that the running surface has an influence on running performance especially regarding running-related injuries[44,45]. Also in cycling, race course characteristics show an influence on race performance[46] where especially ascents slow cyclists down[47]. Changes in elevation during an IRONMAN race have also an influence on pacing during the cycling split[48] where downhill segments show an important influence[49]. The present findings may help IRONMAN triathletes to select an appropriate IRONMAN race for their personal achievements.

Limitations

Despite the use of a data set of nearly 700,000 IRONMAN triathletes and the large time frame of 20 years, we are not sure whether all races correctly measured their split distances. Furthermore, since thousands of athletes compete in an IRONMAN race, athletes can cycle in packages during the cycling split[50], although drafting in cycling in an IRONMAN race is not allowed. Drafting while cycling can considerably improve performance, reduce the cycling split time, and improve the subsequent running split performance[51]. The biopsychosocial factors and intrinsic and extrinsic motivation that led to elite performance in professional and nonprofessional athletes are different. It is extremely important to differentiate and highlight this in the research problem and justification. This could even be the objective of future research. In another way, by analyzing and comparing the athletic achievements of different countries in these events, researchers can discern patterns, trends, and potential influencing factors that contribute to superior athletic performance[52]. These findings not only contribute to understanding the competitive dynamics within the IRONMAN community but also have broader implications for sports science, training methodologies, and potentially national-level athletic development strategies[27]. Further, the study may serve as a foundation for future research, helping athletes, coaches, and sports enthusiasts improve their understanding of the factors that contribute to success in this challenging endurance sport.

Conclusion

In summary, we found that IRONMAN age group athletes from European countries such as Germany, Austria, Denmark, Belgium, Switzerland, Norway, Czechia, Estonia, and Slovenia have the fastest IRONMAN race times although IRONMAN Hawaii is the fastest race course worldwide, most likely due to the fact that IRONMAN Hawaii is the IRONMAN World Championship. With the presented results for optimal air and water temperatures and the description of the optimal cycling and running course characteristics, IRONMAN age group athletes might be able to select an IRONMAN race with the best conditions in order to achieve a fast IRONMAN race time.

Data availability

For this study, we have included official results and split times from the official IRONMAN® website (www.ironman.com) The data sets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 22 February 2024; Accepted: 19 December 2024
Published online: 06 January 2025

References

1. Carpenter, J., Frank, R. & Huet-Vaughn, E. Gender differences in interpersonal and intrapersonal competitive behavior. *J. Behav. Exp. Econ.* **77**, 170–176. https://doi.org/10.1016/j.socec.2018.10.003 (2018).
2. Choi, Y. H., Wei, Q., Zhang, L. & Choi, S.-J. The impact of cultural distance on performance at the summer olympic games. *SAGE Open* **12**(1), 21582440221085264. https://doi.org/10.1177/21582440221085265 (2022).
3. Berri, D. J., Deutscher, C. & Galletti, A. Born in the USA: National origin effects on time allocation in US and Spanish professional basketball. *Natl. Inst. Econ. Rev.* **232**, R41–R50. https://doi.org/10.1177/002795011523200105 (2015).
4. World Athletics. What it takes to become a Kenyan distance champion: World Athletics, 2022 [cited 2022 20 August]. Available from: https://worldathletics.org/be-active/performance/kenyan-distance-running-reasons-success
5. Wattie, N., Schorer, J. & Baker, J. Seeing the forest but not the trees: Heterogeneity in community size effects in Canadian ice hockey players. *J. Sports Sci.* **36**(4), 436–444. https://doi.org/10.1080/02640414.2017.1313444 (2018).
6. Cote, J., Macdonald, D., Baker, J. & Abernethy, B. When, "where" is more important than "when": Birthplace and birthdate effects on the achievement of sporting expertise. *J. Sports Sci.* **24**(10), 1065–1073. https://doi.org/10.1080/02640410500432490 (2006).
7. Rossing, N. N., Stentoft, D., Flattum, A., Côté, J. & Karbing, D. S. Influence of population size, density, and proximity to talent clubs on the likelihood of becoming elite youth athlete. *Scand. J. Med. Sci. Sports* **28**(3), 1304–1313. https://doi.org/10.1111/sms.13009 (2018).
8. Hancock, D. J., Coutinho, P., Côté, J. & Mesquita, I. Influences of population size and density on birthplace effects. *J. Sports Sci.* **36**(1), 33–38. https://doi.org/10.1080/02640414.2016.1276614 (2018).
9. Kasović, M., Škrinjarić, B. & Štefan, L. Macro and meso indicators of success pertaining to European countries in elite triathlon. *Int. Sci. J. Kinesiol.* **13**(2), 49–56 (2020).

10. Bale, J. Kenyan running before the 1968 Mexico Olympics. In *East African Running* (ed. Francis, T.) (Routledge, 2007).
11. Bale, J. & Sang, J. *Kenyan Running: Movement Culture, Geography and Global Change* 1st edn. (Routledge, 1996).
12. Poczta, J. & Malchrowicz-Mosko, E. Mass triathlon participation as a human need to set the goals and cross the borders. How to understand the triathlete?. *Olimpianos J. Olympi. Stud.* **8**, 9. https://doi.org/10.30937/2526-6314.v4.id114 (2020).
13. Dähler, P., Rüst, C., Rosemann, T., Lepers, R. & Knechtle, B. Nation related participation and performance trends in 'Ironman Hawaii' from 1985 to 2012. *BMC Sports Sci. Med. Rehabil.* **6**, 16. https://doi.org/10.1186/2052-1847-6-16 (2014).
14. Poczta, J. & Malchrowicz-Mośko, E. From half-marathon to triathlon: sports participation as a reflection of post-materialism processes in polish society. *Olimpianos J. Olympic Stud.* **5**, 141–151. https://doi.org/10.30937/2526-6314.v5.id128 (2021).
15. Sousa, C. V., Nikolaidis, P. T. & Knechtle, B. Ultra-triathlon-Pacing, performance trends, the role of nationality, and sex differences in finishers and non-finishers. *Scand. J. Med. Sci. Sports* **30**(3), 556–563. https://doi.org/10.1111/sms.13598 (2020).
16. Rüst, C. A. et al. Nation related participation and performance trends in 'Norseman Xtreme Triathlon' from 2006 to 2014. *SpringerPlus* **4**(1), 469. https://doi.org/10.1186/s40064-015-1255-5 (2015).
17. Stiefel, M., Knechtle, B., Rüst, C. A., Rosemann, T. & Lepers, R. The age of peak performance in Ironman triathlon: A cross-sectional and longitudinal data analysis. *Extreme Physiol. Med.* **2**(1), 27. https://doi.org/10.1186/2046-7648-2-27 (2013).
18. Haugen, T. A. et al. Peak age and performance progression in world-class track-and-field athletes. *Int. J. Sports Physiol. Perform.* **13**(9), 1122–1129 (2018).
19. Stevenson, J. L., Song, H. & Cooper, J. A. Age and sex differences pertaining to modes of locomotion in triathlon. *Med. Sci. Sports Exerc.* **45**(5), 976–984. https://doi.org/10.1249/MSS.0b013e31827d17eb (2013).
20. Piacentini, M. F., Vleck, V. & Lepers, R. Effect of age on the sex difference in Ironman triathlon performance. *Movement Sport Sci.* https://doi.org/10.1051/sm/2019030 (2019).
21. Thuany, M. et al. A machine learning approach to finding the fastest race course for professional athletes competing in Ironman® 703 races between 2004 and 2020. *Int. J. Environ. Res. Public Health* **20**(4), 3619. https://doi.org/10.3390/ijerph20043619 (2023).
22. Nikolaidis, P. T. et al. Predicting overall performance in Ironman 70.3 age group triathletes through split disciplines. *Sci. Rep.* **13**(1), 11492. https://doi.org/10.1038/s41598-023-38181-y (2023).
23. Puccinelli, P. et al. Origin and age group of the fastest amateur triathletes competing in 'Ironman Hawaii' between 2003 and 2019. *Sports Med. Health Sci.* https://doi.org/10.1016/j.smhs.2023.07.008 (2023).
24. Thuany, M., Viljoen, C., Gomes, T. N., Knechtle, B. & Scheer, V. Mental health in ultra-endurance runners: A systematic review. *Sports Med.* **53**(10), 1891–1904. https://doi.org/10.1007/s40279-023-01890-5 (2023).
25. Georgiades, E., Klissouras, V., Baulch, J., Wang, G. & Pitsiladis, Y. Why nature prevails over nurture in the making of the elite athlete. *BMC Genom.* **18**(8), 835. https://doi.org/10.1186/s12864-017-4190-8 (2017).
26. Collins, M. et al. The ACE gene and endurance performance during the South African Ironman Triathlons. *Med. Sci. Sports Exerc.* **36**(8), 1314–1320. https://doi.org/10.1249/01.mss.0000135779.41475.42 (2004).
27. González-Ravé, J. M., Hermosilla, F., González-Mohíno, F., Casado, A. & Pyne, D. B. Training intensity distribution, training volume, and periodization models in elite swimmers: A systematic review. *Int. J. Sports Physiol. Perform.* **16**(7), 913–926. https://doi.org/10.1123/ijspp.2020-0906 (2021).
28. Nieto-Jimenez, C., Ruso-Álvarez, J., Pardos-Mainer, E. E. & Orellana, J. N. La variabilidad de la frecuencia cardiaca en el control del entrenamiento en un corredor de Ironman. Estudio de caso (Heart Rate Varibility in the training monitoring of an Ironman runner. A case study. *Retos* **37**, 339–343. https://doi.org/10.47197/retos.v37i37.73873 (2020).
29. Nikolaidis, P. T., Kach, I., Rosemann, T. & Knechtle, B. The role of nationality on the pacing of ironman triathletes. *Asian J. Sports Med.* **8**(4), e57130. https://doi.org/10.5812/asjsm.57130 (2017).
30. Mölenberg, F. J. M., de Waart, F., Burdorf, A. & van Lenthe, F. J. Hosting elite sport events to target recreational sport participation: An interrupted time series analysis. *Int. J. Sport Policy Polit.* **12**(4), 531–543. https://doi.org/10.1080/19406940.2020.1839530 (2020).
31. Asgari, B. & Khorshidi, R. The effect of hosting on performance of host countries in summer and winter Olympic Games. *Eur. J. Sports Exercise Sci.* **2**(2), 1–6 (2013).
32. Forrest, D., McHale, I. G., Sanz, I. & Tena, J. D. An analysis of country medal shares in individual sports at the Olympics. *Eur. Sport Manag. Quart.* **17**(2), 117–131. https://doi.org/10.1080/16184742.2016.1248463 (2017).
33. Jürgens, D. et al. An analysis of participation and performance by nationality at "Ironman Switzerland" from 1995 to 2011. *J. Sci. Cycl.* **1**(2), 10–20 (2012).
34. Foley, R. Swimming in Ireland: Immersions in therapeutic blue space. *Health Place* **35**, 218–225. https://doi.org/10.1016/j.healthplace.2014.09.015 (2015).
35. Britton, E. & Foley, R. Sensing water: Uncovering health and well-being in the sea and surf. *J. Sport Social Issues* **45**(1), 60–87. https://doi.org/10.1177/0193723520928597 (2021).
36. Çetin, A. Triathletes' motivational dimensions for participation in ironman, and comparison of these factors with their life satisfaction and achievement perceptions. *Int. J. Hum. Move. Sports Sci.* **7**(3), 43–50. https://doi.org/10.13189/saj.2019.070301 (2019).
37. Knechtle, B. et al. Performance and pacing of professional IRONMAN triathletes: The fastest IRONMAN World Championship ever-IRONMAN Hawaii 2022. *Sci. Rep.* **13**(1), 15708. https://doi.org/10.1038/s41598-023-42800-z (2023).
38. Stiefel, M., Rüst, C. A., Rosemann, T. & Knechtle, B. A comparison of participation and performance in age-group finishers competing in and qualifying for Ironman Hawaii. *Int. J. Gen. Med.* **6**, 67–77. https://doi.org/10.2147/IJGM.S40202 (2013).
39. Mantzios, K. et al. Effects of weather parameters on endurance running performance: discipline-specific analysis of 1258 races. *Med. Sci. Sports Exerc.* **54**(1), 153–161. https://doi.org/10.1249/MSS.0000000000002769 (2022).
40. Zhao, J., Wu, Y. & Zhang, J. A study of triathletes' race strategies in different competition environments. *Heliyon* **10**(8), e29454. https://doi.org/10.1016/j.heliyon.2024.e29454 (2024).
41. Bouscaren, N., Millet, G. Y. & Racinais, S. Heat stress challenges in marathon versus ultra-endurance running. *Front Sports Act. Living* **13**(1), 59. https://doi.org/10.3389/fspor.2019.00059 (2019).
42. Olcina, G., Crespo, C., Timón, R., Mjaanes, J. M. & Calleja-González, J. Core temperature response during the marathon portion of the ironman world championship (Kona-Hawaii). *Front. Physiol.* **3**(10), 1469. https://doi.org/10.3389/fphys.2019.01469 (2019).
43. Mougios, V. & Deligiannis, A. Effect of water temperature on performance, lactate production and heart rate at swimming of maximal and submaximal intensity. *J. Sports Med. Phys. Fitness* **33**(1), 27–33 (1993).
44. Jafarnezhadgero, A., Eskandari, S., Sajedi, H. & Dionisio, V. C. Long-term effects of running exercises on natural grass, artificial turf, and synthetic surfaces on ground reaction force components in individuals with overpronated feet: A randomized controlled trial. *Gait Posture* **109**, 28–33. https://doi.org/10.1016/j.gaitpost.2024.01.016 (2024).
45. Mitchell, C., McDonnell, S., Oganezova, K., Mockler, D. & Fleming, N. The effect of surface compliance on overground running biomechanics. A systematic review and meta-analysis. *Sports Biomech.* https://doi.org/10.1080/14763141.2023.2236058 (2023).
46. Phillips, K. E. & Hopkins, W. G. Determinants of cycling performance: A review of the dimensions and features regulating performance in elite cycling competitions. *Sports Med. Open* **6**(1), 23. https://doi.org/10.1186/s40798-020-00252-z (2020).
47. Wells, M. S. & Marwood, S. Effects of power variation on cycle performance during simulated hilly time-trials. *Eur. J. Sport Sci.* **16**(8), 912–918. https://doi.org/10.1080/17461391.2016.1156162 (2016).
48. Pryor, J. L. et al. Pacing strategy of a full ironman overall female winner on a course with major elevation changes. *J. Strength Cond. Res.* **32**(11), 3080–3087. https://doi.org/10.1519/JSC.0000000000002807 (2018).

49. Johnson, E. C. et al. Bike and run pacing on downhill segments predict Ironman triathlon relative success. *J. Sci. Med. Sport* **18**(1), 82–87. https://doi.org/10.1016/j.jsams.2013.12.001 (2015).
50. Allebosch, G., Van den Bossche, S., Veelaert, P. & Philips, W. Camera-based system for drafting detection while cycling. *Sensors* **20**(5), 1241 (2020).
51. Hausswirth, C. et al. Effect of two drafting modalities in cycling on running performance. *Med. Sci. Sports Exerc.* **33**(3), 485–492. https://doi.org/10.1097/00005768-200103000-00023 (2001).
52. Knechtle, B., Rosemann, T. & Nikolaidis, P. The role of nationality in ultra-endurance sports: The paradigm of cross-country skiing and long-distance running. *Int. J. Environ. Res. Public Health* https://doi.org/10.3390/ijerph17072543 (2020).

Author contributions
Conceptualization: Beat Knechtle. Data curation: Beat Knechtle, Elias Villiger. Formal analysis: David Valero. Methodology: Beat Knechtle. Writing – original draft: Beat Knechtle, Mabliny Thuany. Writing – Editing: Katja Weiss, Thomas Rosemann, Pantelis T. Nikolaidis, Rodrigo Luiz Vancini, Marilia Santos Andrade.

Declarations

Ethics approval
This study was approved by the Institutional Review Board of Kanton St. Gallen, Switzerland, with a waiver of the requirement for informed consent of the participants as the study involved the analysis of publicly available data (EKSG 01/06/2010). The study was conducted in accordance with recognized ethical standards according to the Declaration of Helsinki adopted in 1964 and revised in 2013.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to B.K.

Reprints and permissions information is available at www.nature.com/reprints.