

# NRCD: An Open Database of Collegiate Running with Unified Performance Standardization

Jonathan A. Karr Jr  
jkarr@nd.edu  
nationalrunningclubdatabase@gmail.com  
University of Notre Dame  
Notre Dame, Indiana, USA  
National Running Club Database  
Michigan, USA

Nicholas Pell  
npell@alumni.unc.edu  
University of North Carolina at  
Chapel Hill  
Chapel Hill, North Carolina, USA

Ramzi K. Bualuan  
rbualuan@nd.edu  
University of Notre Dame  
Notre Dame, Indiana, USA

Ryan M. Fryer  
rfryer@nd.edu  
University of Notre Dame  
Notre Dame, Indiana, USA  
University of Virginia  
Charlottesville, Virginia, USA

Kayla Ambrose  
kambrose@alumni.nd.edu  
University of Notre Dame  
Notre Dame, Indiana, USA

Ben Darden  
bdarden1205@vt.edu  
Virginia Tech  
Blacksburg, Virginia, USA

Evan Hall  
ehall9@alumni.nd.edu  
University of Notre Dame  
Notre Dame, Indiana, USA

Nitesh V. Chawla  
nchawla@nd.edu  
University of Notre Dame  
Notre Dame, Indiana, USA

## Abstract

Collegiate running in the United States generates thousands of race results annually in cross country and track and field, yet no large-scale dataset has been publicly available for research. Existing websites such as Athletic.net, MileSplit, and TFRS host results but do not support bulk download, restricting prior analyses to ~500 performances, often skewing studies toward male athletes. We introduce the **National Running Club Database (NRCD)**, the first openly available collegiate running dataset at scale: 128,963 approved performances from 28,913 athletes across 1,336 meets in four sports (cross country (XC), indoor and outdoor track, and road races), 36.3% women, spanning 2004 through 2026. Within that single export, meets from August 2023 onward carry comprehensive course distance, elevation gain and loss, weather at race time, and track venue metadata (97.7% of XC rows with weather fields); earlier seasons back to 2004 are included with sparser metadata. NRCD is **community-governed** through open submission and expert approval and is maintained as a live database whose meet volume has grown yearly. We release a unified performance standardization framework that operationalizes established distance, elevation, and heat adjustments in one pipeline. Furthermore, **we recommend gender-stratified modeling**. On XC, full standardization lowers median within-athlete cross-meet variability by 51.0% (women) and 34.4% (men) versus raw times. We release the dataset and pipeline

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
CIKM '26, Rome, Italy

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-XXXX-X/2018/06  
<https://doi.org/XXXXXXX.XXXXXXX>

with a python package 'nrcd' under FAIR principles, supporting longitudinal athlete modeling, environmental-confounder studies, and gender-equity research in collegiate sport.

## CCS Concepts

• **Information systems** → **Data mining**; • **Applied computing** → *Health informatics*.

## Keywords

sports analytics, running, performance standardization, longitudinal modeling, community-contributed data, collegiate athletics

## ACM Reference Format:

Jonathan A. Karr Jr, Ryan M. Fryer, Ben Darden, Nicholas Pell, Kayla Ambrose, Evan Hall, Ramzi K. Bualuan, and Nitesh V. Chawla. 2026. NRCD: An Open Database of Collegiate Running with Unified Performance Standardization. In *The 35th ACM International Conference on Information and Knowledge Management (CIKM '26)*, November 7-11, 2026, Rome, Italy. ACM, New York, NY, USA, 14 pages. <https://doi.org/XXXXXXX.XXXXXXX>

Dataset with Datasheet: <https://zenodo.org/records/20578527>

Code: [https://github.com/National-Running-Club-Database/data\\_paper](https://github.com/National-Running-Club-Database/data_paper)

Python package 'nrcd': <https://pypi.org/project/nrcd/>

GitHub for Python package: <https://github.com/National-Running-Club-Database/nrcd>

Website: <https://www.nationalrunningclubdatabase.com>

## 1 Introduction

Collegiate distance running is organized across World Athletics, NCAA, NAIA, NJCAA, and club bodies including the National Intercollegiate Running Club Association (NIRCA) [48]. Officiated competition in NIRCA ranges from NCAA Division I-caliber performances to times of community 5K road races, so an open resource

must serve a full band of runners rather than a single subfield or ability tier. Thousands of results are published annually on Athletic.net, MileSplit, and TFRS [2, 41, 72], yet none of these platforms support bulk relational export. Prior open analyses of running performance have therefore relied on hand-curated samples of a few hundred records [42], with documented male skew [1, 31, 69]. A separate barrier is *comparability*: cross country courses differ in measured distance, terrain, and weather; track meets differ in surface, banking, and venue altitude; road races vary in distance and conditions. Adjustment formulas exist in coaching and biomechanics literature [60, 38, 14, 23] but have not been operationalized at scale in a single public collegiate corpus.

NRCD enables information-retrieval and data-mining research that access constraints have blocked: longitudinal athlete modeling across seasons, benchmarking nonlinear predictors [61, 13], fairness analysis across environmental confounders, and reproducible sports-informatics benchmarks. The resource pairs a public dataset of NIRCA athletes, a validated conversion formula, a pip-installable ‘nrcd’ Python package, and a live, moderated, contributed website on <https://www.nationalrunningclubdatabase.com>, addressing future sustainability. By supporting equitable participation research and evidence on training and competition under environmental stress, NRCD aligns with UN SDG 5 (Gender Equality) and UN SDG 3 (Good Health and Well-Being) [66].

**Contributions and scope:** This resource paper describes the **dataset** and a **standardization framework**, with empirical validation on cross country (XC): each fall meet is the same aerobic race for a given gender (8k/6k). Track mixes distances and event counts per athlete per meet throughout the season; therefore, we restrict it from validation. We do *not* report predictive benchmarks, feature-importance studies, or race-frequency findings here. Those belong in separate analytical papers. We present:

- (1) **NRCD** - 128,963 results (36.3% women) across four sports;
- (2) **Community-governed curation** - results added by community and approved by domain experts;
- (3) **Environmentally aware standardization** - converted-only and standardized modes with a derived heat law ( $k=0.0016$ ) and the ‘nrcd’ Python package;
- (4) **Gender-stratified empirical checks** - separate analyses for men and women.

## 2 Related Work

**Open running data:** Prior open collegiate work uses less than a thousand hand-curated results [42]. No prior resource offers close to 100,000+ results, with course and weather metadata.

**Standardization.** Riegel’s law [60], Maurer elevation factors [38], and weather–performance work [16, 77, 63] exist separately; NRCD integrates them with a quadratic heat surrogate fit to Hadley’s coaching bands [78], excluding AQI adjustment [16].

**Community sports data:** Fitness apps record training, not officiated meets [22, 39]. NRCD provides governed meet results with unique ids for user tracking.

| Sport         | Results        | Athletes <sup>†</sup> | Meets        |
|---------------|----------------|-----------------------|--------------|
| Cross Country | 70,285         | 21,630                | 524          |
| Indoor Track  | 22,809         | 7,207                 | 326          |
| Outdoor Track | 33,689         | 9,607                 | 428          |
| Road Race     | 2,180          | 1,627                 | 58           |
| <b>Total</b>  | <b>128,963</b> | <b>28,913</b>         | <b>1,336</b> |

**Table 1: Results by sport**

<sup>†</sup>Each *Athletes* cell counts distinct ids. Sport rows are not summable: 28.9% of athletes are multi-sport.

| Sport         | Era           | Results | Athletes <sup>†</sup> | Weather (%) | Alt. (%) | Course (%) |
|---------------|---------------|---------|-----------------------|-------------|----------|------------|
| Cross Country | historical    | 46,930  | 15,899                | 4.4         | 97.5     | 3.4        |
| Cross Country | comprehensive | 23,355  | 7,083                 | 97.7        | 100.0    | 38.5       |
| Indoor Track  | historical    | 4,171   | 2,179                 | —           | 100.0    | —          |
| Indoor Track  | comprehensive | 18,638  | 5,461                 | —           | 99.4     | —          |
| Outdoor Track | historical    | 12,321  | 4,817                 | 0.0         | 100.0    | —          |
| Outdoor Track | comprehensive | 21,368  | 5,474                 | 36.9        | 98.8     | —          |
| Road Race     | historical    | 1,332   | 1,049                 | 34.5        | 100.0    | 0.0        |
| Road Race     | comprehensive | 848     | 607                   | 67.0        | 100.0    | 27.1       |

**Table 2: Total coverage by sport and era (Alt. = altitude, Course = gain/loss and/or short/long course details)**

| Sport                  | Women         | Men           | Women (%)   |
|------------------------|---------------|---------------|-------------|
| Cross Country          | 27,112        | 43,173        | 38.6        |
| Indoor Track           | 7,674         | 15,135        | 33.6        |
| Outdoor Track          | 10,989        | 22,700        | 32.6        |
| Road Race              | 982           | 1,198         | 45.1        |
| <b>Total</b>           | <b>46,757</b> | <b>82,206</b> | <b>36.3</b> |
| <b>Unique athletes</b> | <b>11,381</b> | <b>17,532</b> | <b>39.4</b> |

**Table 3: Results by sport and gender**

## 3 The NRCD Dataset

NRCD covers *Cross Country*, *Indoor Track*, *Outdoor Track*, and *Road Race*. Tables 1–2 summarize the approved export; Table 2 splits **comprehensive** (meet date  $\geq$  August 2023, richer metadata) from **historical** (2004–July 2023). Of 128,963 results, 64,209 are comprehensive and 64,754 historical. For longitudinal use, 10,335 athletes (35.7%) have results in  $\geq 2$  school years (August 1–July 31) and 4,687 (16.2%) in  $\geq 3$ .

## 4 Community-Governed Curation

NRCD is **open-contributed**: coaches, athletes, and volunteers submit meets; no entry is added until an admin with domain expertise approves it against public postings. Admins also review public meet results and registered teams nationwide to add missing meets, so coverage is broader than user submissions alone. Approved meets grew from 68 to 142, 340, and 281 per year (2022–2025). When names share a last name and first initial from the same team (John Smith & Johnny Smith) or exceed 90% similarity, the system *suggests* athlete merges; an admin approves or denies each suggestion. Public exports will continue to be updated **yearly**. Exports strip PII and include datasheets [75, 20]; weather via OpenWeatherMap is included where course records exist [52].

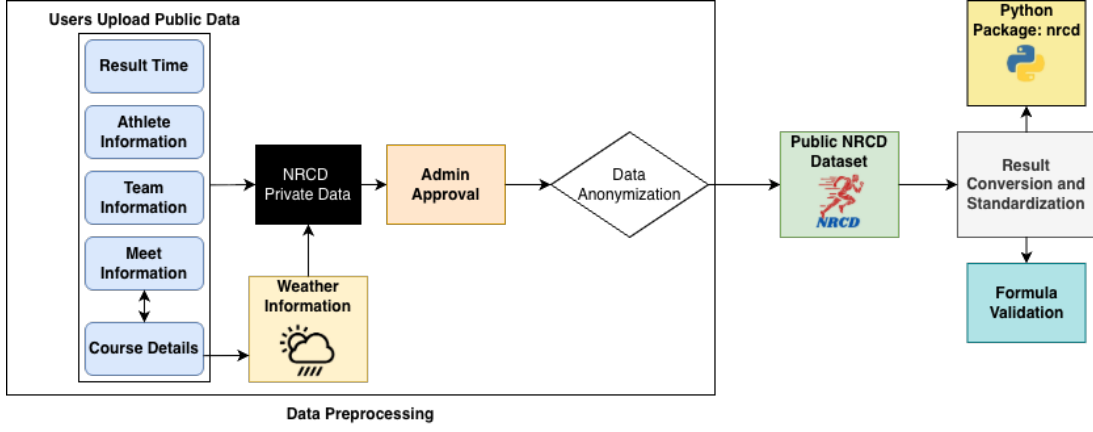


Figure 1: NRCD data and standardization pipeline

| Sport         | Gender | Sprint        | Mid          | Dist.         | Relay        | Field        | Total          |
|---------------|--------|---------------|--------------|---------------|--------------|--------------|----------------|
| Cross Country | F      | 0             | 0            | 27,112        | 0            | 0            | 27,112         |
| Cross Country | M      | 0             | 0            | 43,173        | 0            | 0            | 43,173         |
| Indoor Track  | F      | 2,419         | 908          | 2,642         | 628          | 1,077        | 7,674          |
| Indoor Track  | M      | 5,320         | 1,831        | 5,314         | 1,085        | 1,585        | 15,135         |
| Outdoor Track | F      | 3,650         | 1,405        | 3,468         | 653          | 1,813        | 10,989         |
| Outdoor Track | M      | 7,268         | 3,111        | 8,205         | 1,408        | 2,708        | 22,700         |
| Road Race     | F      | 0             | 0            | 982           | 0            | 0            | 982            |
| Road Race     | M      | 0             | 0            | 1,198         | 0            | 0            | 1,198          |
| <b>All</b>    | -      | <b>18,657</b> | <b>7,255</b> | <b>92,094</b> | <b>3,774</b> | <b>7,183</b> | <b>128,963</b> |

Table 4: Results by category. Sprint.: ≤400 m and hurdles; mid: 401-1000 m; dist.: &gt; 1000 m and steeplechase.

## 5 Unified Standardization Framework

Figure 1 summarizes the pipeline: approved results join athletes, teams, events, and course metadata; sport-specific factors (Table 5) yield comparable times *within* sport and gender. For XC, factors apply in fixed order: weather, grade, venue altitude ( $f_{alt}$ ), measured-vs.-reported course length, then Riegel conversion to  $d_{target}$  (Equation 2; heat in Section 6.1).

Raw time  $t_{raw}$  maps to standardized time via multiplicative factors and Riegel conversion [60]:

$$t_{std} = t_{raw} \cdot \prod_i f_i \cdot \left( \frac{d_{target}}{d_{actual}} \right)^b, \quad b \in \{1.055 \text{ (M)}, 1.08 \text{ (F)}\}. \quad (1)$$

<sup>1</sup> For XC,

$$t_{std} = t_{raw} \cdot f_{weather} \cdot f_{elev} \cdot f_{alt} \cdot \left( \frac{d_{actual}}{d_{reported}} \right)^b \cdot \left( \frac{d_{target}}{d_{actual}} \right)^b, \quad (2)$$

with  $d_{target} \in \{8000 \text{ m}, 6000 \text{ m}\}$  for men and women (NIRCA defaults) [48]. The same exponent  $b$  is used twice:  $(d_{actual}/d_{reported})^b$

<sup>1</sup>Exponents follow Riegel’s fatigue-factor table [60] ( $\approx 1.05$  men,  $\approx 1.08$  women); we implement rounded 1.055/1.08. Later predictors use a single population exponent (1.06), including the classical Riegel simplification and the rank-1 baseline [7]. Refits on modern world records give discipline-specific values (e.g.  $\approx 1.04$  women’s road,  $\approx 1.12$  track). Unified  $b=1.06$  on NRCD XC shifts validation percentages by  $<0.5$  pp.

corrects short or long measured courses relative to the listed distance, and  $(d_{target}/d_{actual})^b$  converted to gender reference (8km/6km).

**Gender-stratified analysis:** We recommend fitting, evaluating, and reporting all models separately by gender (Section 6), never pooling raw or adjusted seconds across men and women.

**Temperature:** Heat index  $H$  is temperature plus dew point (°F) [63]. We fit a batchable quadratic surrogate to Hadley’s coaching bands [78] (not peer-reviewed physiology):

$$p(H) = k(H - 100)^2, \quad f_{weather} = 1 - \frac{p(H)}{100}, \quad H > 100, \quad (3)$$

with  $f_{weather}=1$  for  $H \leq 100$ , for aerobic events per Table 5. We fit  $k=0.0016$  to Hadley midpoints only (Table 6), not finish times.

**AQI:** Measured where available but omitted from  $f_i$  since it is confounded with temperature and dew point [16].

**Elevation and course length:**  $f_{elev} = 1.04^g \cdot 0.9633^l$  for gain/loss grades  $g, l$  [38] (ft→grade% via measured length). Measured course length differing from reported distance adds  $(d_{actual}/d_{reported})^b$  before converting to  $d_{target}$ .

**Venue altitude:**  $f_{alt} = t_{sea}/t_{alt}$  from the gender-stratified Péronnet-Thibault model [55, 54] at  $d_{actual}$ : MAP reduced at meet elevation; air density from barometric pressure when recorded.

**Converted only vs. standardized:** We present *converted* (distance and course-length, no weather/elevation) and *standardized* adjustments. Converted-only times can absorb fall-season cooling and grade, inflating apparent within-season fitness gains (Section 6.3).

**Track:** Equation 1 with NCAA facility-indexing multipliers [47, 46, 3, 14] (tabulated  $\alpha$ ; repo lookup). Outdoor sprint/hurdle wind uses per-result wind with event-specific calm-equivalent  $\Delta t$  [35, 59, 58, 44, 70]: 100 m (Linthorne); 200 m (Quinn Table 2, straight-gauge); 110 m hurdles (Spiegel & Mureika); 400 m (Quinn’s curved-track simulation); same  $f_{alt}$  when  $z$  is set; heat on outdoor mid/distance.

**XC and Road:** Equation 1 with Riegel,  $f_{weather}$ ,  $f_{elev}$ , course-distance accuracy, and  $f_{alt}$  when metadata exist.

## 6 Formula Validation (XC)

We validate for XC with three checks that exercise the formulas, not predictive benchmarks. All summaries are **gender-stratified**; never pool absolute seconds across 6 km and 8 km targets.

| Factor                        | Cross Country | In. track | Out. track | Road |
|-------------------------------|---------------|-----------|------------|------|
| Riegel distance               | ✓             | ✓         | ✓          | ✓    |
| Heat ( $f_{\text{weather}}$ ) | ✓             | —         | mid+       | ✓    |
| Wind (per-result)             | —             | —         | sprint     | —    |
| Elevation gain/loss           | ✓             | —         | —          | ✓    |
| Course distance accuracy      | ✓             | —         | —          | ✓    |
| AQI*                          | ✓             | —         | ✓          | ✓    |
| Altitude                      | ✓             | ✓         | ✓          | ✓    |
| Track length & banking        | —             | ✓         | —          | —    |

**Table 5: Adjustment factors by sport (AQI only measured)**

| $H$ | Hadley (%) | Mid (%) | Quad. (%) | $f_{\text{weather}}$ |
|-----|------------|---------|-----------|----------------------|
| 110 | 0.0-0.5    | 0.25    | 0.16      | 0.9984               |
| 120 | 0.5-1.0    | 0.75    | 0.64      | 0.9936               |
| 145 | 3.0-4.5    | 3.75    | 3.24      | 0.9676               |
| 160 | 4.5-6.0    | 5.25    | 5.76      | 0.9424               |
| 175 | 8.0-10.0   | 9.00    | 9.00      | 0.9100               |

**Table 6: Hadley bands vs. quadratic surrogate ( $k=0.0016$ ).**

| $k$    | W reduc. (%) | M reduc. (%) | W infl. (%) | M infl. (%) |
|--------|--------------|--------------|-------------|-------------|
| 0.0010 | 50.7         | 34.0         | 9           | 14          |
| 0.0016 | 51.0         | 34.4         | 12          | 21          |
| 0.0020 | 51.0         | 34.0         | 16          | 25          |

**Table 7: Sensitivity of XC validation to heat coefficient  $k$ . Reduc. = median within-athlete cross-meet SD; inflat. = converted-only vs. standardization.**

## 6.1 Heat Surrogate

(1) **Band fidelity:**  $k=0.0016$  matches Hadley slowdown ranges in percent (Table 6), with RMSE 0.34 pp vs. band midpoints and 70% of  $H \in [101, 180]$  inside published ranges (fidelity to the heuristic [63], not independent physiological validation). (2) **Structure:**  $f_{\text{weather}}=1$  for  $H \leq 100$ ; monotone decrease as  $H$  rises. (3) **Plausibility:** on 35,909 XC rows with weather, mean factor 0.990 when  $H > 100$ . Finish-time recalibration of  $k$  is unstable and rejected. Validation metrics are stable for  $k \in \{0.001, 0.0016, 0.002\}$  (Table 7).

## 6.2 Within-Athlete Variance

For athletes with  $\geq 3$  regular-season XC meets per season (nationals excluded; calendar and school years coincide for fall-only XC), we computed median cross-meet standard deviation per athlete-season on *raw* and *standardized* finish times (6 km / 8 km per gender). Percent reduction differs by gender (51.0% vs. 34.4%), driven mainly by a wider women’s raw distribution and stronger compression from distance conversion, not by larger heat-index swings (median cross-meet  $H$  SD is similar when metadata exist).

## 6.3 Within-Season Improvement

For athletes with  $\geq 2$  regular-season races, *improvement* is first-race minus best later finish times on the gender reference distance (positive = faster later). Table 9 compares *converted only* versus

| Group        | $n$   | Raw    | Full std. | Reduct. |
|--------------|-------|--------|-----------|---------|
| Women (6 km) | 2,167 | 108.21 | 52.97     | 51.0%   |
| Men (8 km)   | 3,929 | 85.08  | 55.82     | 34.4%   |

**Table 8: Median within-athlete cross-meet SD (cross country,  $\geq 3$  meets/season).**

| Group | $n$   | Conv. only        | Full std.         | Inflat. (%) |
|-------|-------|-------------------|-------------------|-------------|
| Women | 5,548 | 45.1 [42.7, 47.9] | 40.4 [36.9, 42.6] | 12 [10, 18] |
| Men   | 8,986 | 41.4 [39.1, 43.9] | 34.3 [32.2, 36.6] | 21 [17, 24] |

**Table 9: Median within-season improvement (sec) with 95% bootstrap CIs. Inflat. (%) =  $100 \times (\text{median}_{\text{conv}}/\text{median}_{\text{std}} - 1)$ .**

*standardization* on the same athlete-seasons; brackets give 95% percentile-bootstrap CIs (2,000 replicates).

Converted-only inflates median within-season improvement by 12% (women) and 21% (men) versus full standardization because fall cooling and grade remain in converted-only times. Illustratively, gender-stratified random forest and gradient boosting (train 2023, test 2024) predict within-season improvement rate (s/day) from season-level XC features. Converted-only finish times yield higher test  $R^2$  than standardized (RF men 0.74 vs. 0.72, women 0.74 vs. 0.72; GB men 0.61 vs. 0.57, women 0.58 vs. 0.52); environment enters via finish times, not weather columns, so higher  $R^2$  signals confounding, not fitness. **Without full standardization, models risk predicting external factors, not the athlete.**

**Limitations:** Statistics only reflect NIRCA club competition. Pre-August 2023 entries lack complete metadata. Results are adjusted for environmentally aware standards, not individual training plans. Variance and improvement validation only uses XC.

## 7 Conclusion

NRCD is the first open large-scale collegiate running dataset with documented standardization. A derived heat surrogate is validated against Hadley’s piecewise bands, not finish-time regression. Gender-stratified XC checks show that full standardization reduces within-athlete cross-meet variability (51.0% women, 34.4% men vs. raw) and avoids overstating within-season improvement relative to converted-only times. We recommend gender-stratified modeling on all derivative work. Finally, we release the dataset and reproducible scripts to remove the bulk-export barrier for sports informatics.

## Ethics Statement

All of our data is obtained from the NRCD, which is publicly accessible. We remove all PII such as links and persons’ names and have permission to use and publish NRCD’s dataset. Notre Dame’s IRB office classified our research as not human subject research.

## Acknowledgments

We greatly appreciate Tom Meurer for supporting the initial project in Notre Dame’s Advanced Database class. We also thank the members and staff of NIRCA, the CRC Coaches Poll, and the Notre Dame Running Club for their support.

## Generative AI Usage Disclosure

Moderate AI use (Cursor, Claude) was used for code generation and paper editing, though all content is validated as our own. No AI was used for result upload or validation.

## References

- [1] Nash Anderson et al. 2023. Under-representation of women is alive and well in sport and exercise medicine: what it looks like and what we can do about it. (2023).
- [2] Athletic.net. 2026. Athletic.net.
- [3] Kyle R Barnes and Rita Malcata. 2017. Conversion index for running on different indoor track and field facility types. *International Journal of Performance Analysis in Sport*, 17, 4, 375–384.
- [4] Ling Beisecker, Patrick Harrison, Marzia Josephson, and JD DeFreese. 2024. Depression, anxiety and stress among female student-athletes: a systematic review and meta-analysis. *British journal of sports medicine*, 58, 5, 278–285.
- [5] Thibault Besson, Robin Macchi, Jeremy Rossi, Cedric YM Morio, Yoko Kuni-masa, Caroline Nicol, Fabrice Vercruyssen, and Guillaume Y Millet. 2022. Sex differences in endurance running. *Sports medicine*, 52, 6, 1235–1257.
- [6] Cheri A Blauwet, Emily M Brook, Adam S Tenforde, Elizabeth Broad, Caroline H Hu, Eliza Abdu-Glass, and Elizabeth G Matzkin. 2017. Low energy availability, menstrual dysfunction, and low bone mineral density in individuals with a disability: implications for the para athlete population. *Sports Medicine*, 47, 9, 1697–1708.
- [7] Duncan AJ Blythe and Franz J Király. 2016. Prediction and quantification of individual athletic performance of runners. *PLoS One*, 11, 6, e0157257.
- [8] Laurent Bosquet, Jonathan Montpetit, Denis Arvisais, and Inigo Mujika. 2007. Effects of tapering on performance: a meta-analysis. *Medicine & Science in Sports & Exercise*, 39, 8, 1358–1365.
- [9] John R Brotherhood. 2008. Heat stress and strain in exercise and sport. *Journal of science and medicine in sport*, 11, 1, 6–19.
- [10] Georgie Bruinvels et al. 2021. Prevalence and frequency of menstrual cycle symptoms are associated with availability to train and compete: a study of 6812 exercising women recruited using the strava exercise app. *British Journal of Sports Medicine*, 55, 8, 438–443.
- [11] Richard J Burden, Katie Morton, Toby Richards, Gregory P Whyte, and Charles R Pedlar. 2015. Is iron treatment beneficial in, iron-deficient but non-anaemic (idna) endurance athletes? a systematic review and meta-analysis. *British journal of sports medicine*, 49, 21, 1389–1397.
- [12] Robert F Chapman, Jonathon L Stickford, and Benjamin D Levine. 2010. Altitude training considerations for the winter sport athlete. *Experimental physiology*, 95, 3, 411–421.
- [13] Tianqi Chen et al. 2015. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1, 4, 1–4.
- [14] Brandon Corts. 2017. *Quantitative analysis of 200 meter track times*. Ph.D. Dissertation. Massachusetts Institute of Technology.
- [15] Kerry S Courneya and Albert V Carron. 1991. Effects of travel and length of home stand/road trip on tie home advantage. *Journal of Sport and Exercise Psychology*, 13, 1, 42–49.
- [16] Marika Cusick, Sebastian T Rowland, and Nicholas DeFelice. 2023. Impact of air pollution on running performance. *Scientific reports*, 13, 1, 1832.
- [17] Kirsty J Elliott-Sale, Kelly L McNulty, Paul Ansdell, Stuart Goodall, Kirsty M Hicks, Kevin Thomas, Paul A Swinton, and Eimear Dolan. 2020. The effects of oral contraceptives on exercise performance in women: a systematic review and meta-analysis. *Sports medicine*, 50, 10, 1785–1812.
- [18] Peter M Fowler, Wade Knez, Stephen Crowcroft, Amy E Mendham, Joanna Miller, Charli Sargent, Shona Halson, and Rob Duffield. 2017. Greater effect of east versus west travel on jet lag, sleep, and team sport performance. *Medicine and science in sports and exercise*.
- [19] Hugh HK Fullagar, Sabrina Skorski, Rob Duffield, Daniel Hammes, Aaron J Coutts, and Tim Meyer. 2015. Sleep and athletic performance: the effects of sleep loss on exercise performance, and physiological and cognitive responses to exercise. *Sports medicine*, 45, 2, 161–186.
- [20] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. Datasheets for datasets. *Communications of the ACM*, 64, 12, 86–92.
- [21] Nanci S Guest et al. 2021. International society of sports nutrition position stand: caffeine and exercise performance. *Journal of the international society of sports nutrition*, 18, 1, 1.
- [22] Dorothea Gute, Stephan Schlögl, and Aleksander Groth. 2022. Keep on running! an analysis of running tracking application features and their potential impact on recreational runner's intrinsic motivation. In *International Conference on Human-Computer Interaction*. Springer, 359–373.
- [23] Héctor Gutiérrez, Eduardo Piedrafitá, Pablo Jesús Bascuas, Irela Arbonés, César Berzosa, and Ana Vanessa Bataller-Cervero. 2025. Real-time performance prediction in long-distance trail running: a practical model based on terrain difficulty and pacing variability. *Sports*, 13, 11, 385.
- [24] Brian Hanley. 2018. Pacing profiles of senior men and women at the 2017 iaaf world cross country championships. *Journal of sports sciences*, 36, 12, 1402–1406.
- [25] Brian Hanley. 2014. Senior men's pacing profiles at the iaaf world cross country championships. *Journal of sports sciences*, 32, 11, 1060–1065.
- [26] Florentina J Hettinga, Andrew M Edwards, and Brian Hanley. 2019. The science behind competition and winning in athletics: using world-level competition data to explore pacing and tactics. *Frontiers in Sports and Active Living*, 1, 11.
- [27] Christina R Hollis, Rachel M Koldenhoven, Jacob E Resch, and Jay Hertel. 2021. Running biomechanics as measured by wearable sensors: effects of speed and surface. *Sports biomechanics*.
- [28] Wouter Hoogkamer, Shalaya Kipp, Barry A Spiering, and Rodger Kram. 2016. Altered running economy directly translates to altered distance-running performance. *Medicine & Science in Sports & Exercise*, 48, 11, 2175–2180.
- [29] Kieran Howle, Adam Waterson, and Rob Duffield. 2020. Injury incidence and workloads during congested schedules in football. *International journal of sports medicine*, 41, 02, 75–81.
- [30] Sandra K Hunter et al. 2023. The biological basis of sex differences in athletic performance: consensus statement for the american college of sports medicine. *Translational Journal of the American College of Sports Medicine*, 8, 4, 1–33.
- [31] Jessica J James et al. 2023. Underrepresentation of women in exercise science and physiology research is associated with authorship gender. *Journal of Applied Physiology*, 135, 4, 932–942.
- [32] Dina C Janse van Rensburg et al. 2021. Managing travel fatigue and jet lag in athletes: a review and consensus statement. *Sports Medicine*, 51, 10, 2029–2050.
- [33] Jie Kang, Nicholas A Ratamess, Avery D Faigenbaum, Jill A Bush, Caitlyn Finnerty, Michael DiFiore, Andy Garcia, and Noah Beller. 2023. Time-of-day effects of exercise on cardiorespiratory responses and endurance performance—a systematic review and meta-analysis. *The Journal of Strength & Conditioning Research*, 37, 10, 2080–2090.
- [34] Thierry M Lejeune, Patrick A Willems, and Norman C Heglund. 1998. Mechanics and energetics of human locomotion on sand. *Journal of Experimental Biology*, 201, 13, 2071–2080.
- [35] Nicholas P Linthorne. 1994. The effect of wind on 100-m sprint times. *Journal of Applied Biomechanics*, 10, 2, 110–131.
- [36] Emily Ann Lundstrom, Nancy I Williams, Heather CM Allaway, Ana Carla Chierighini Salamunes, and Mary Jane De Souza. 2025. Pre-season energy deficiency predicts poorer performance during a competitive season in collegiate female long-distance runners. *European Journal of Sport Science*, 25, 3, e12261.
- [37] P Macdermid and Philippa Grayling. 2021. A comparative analysis of pace, work, and gait during national championship cross-country and road running events. *J Sport Exerc Sci*, 5, 4, 230–235.
- [38] Julian Maurer. 2018. *Race Time Prediction on Individual Historical Training Data for Hilly and Non-Hilly Courses*. Master's thesis. Saarland University, German Research Center for Artificial Intelligence, 66123 Saarbrücken, Germany, (Mar. 2018). Master's Thesis, Master's Program of Media Informatics.
- [39] Grant McKenzie and Krzysztof Janowicz. 2017. Ised: constructing a high-resolution elevation road dataset from massive, low-quality in-situ observations derived from geosocial fitness tracking data. *PloS one*, 12, 10, e0186474.
- [40] Kelly Lee McNulty, Kirsty Jayne Elliott-Sale, Eimear Dolan, Paul Alan Swinton, Paul Ansdell, Stuart Goodall, Kevin Thomas, and Kirsty Marie Hicks. 2020. The effects of menstrual cycle phase on exercise performance in eumenorrheic women: a systematic review and meta-analysis. *Sports medicine*, 50, 10, 1813–1827.
- [41] MileSplit. 2026. MileSplit.
- [42] Matt Millett and TJ Melanson. 2015. Predicting running times from race history.
- [43] Margo Mountjoy et al. 2018. Ioc consensus statement on relative energy deficiency in sport (red-s): 2018 update. *British journal of sports medicine*, 52, 11, 687–697.
- [44] Jonas R Mureika. 2001. The legality of wind and altitude assisted performances in the sprints. *arXiv preprint physics/0102039*.
- [45] National Collegiate Athletic Association. 2025. Ncaa goals time commitments report. [https://ncaaorg.s3.amazonaws.com/research/goals/2026RES\\_GOALS-TimeCommitmentsReport.pdf](https://ncaaorg.s3.amazonaws.com/research/goals/2026RES_GOALS-TimeCommitmentsReport.pdf). (2025).
- [46] National Collegiate Athletic Association. 2012. NCAA indoor track and field facility indexing conversion summary. [https://ncaaorg.s3.amazonaws.com/championships/sports/crosstrack/common/XTF\\_FacilityIndexingConversionSummary.pdf](https://ncaaorg.s3.amazonaws.com/championships/sports/crosstrack/common/XTF_FacilityIndexingConversionSummary.pdf). (2012).
- [47] National Collegiate Athletic Association. 2012. NCAA indoor track size conversion charts. <https://ustfccca.org/assets/NCAA-Indoor-Track-Size-Conversion-Charts.pdf>. (2012).
- [48] National Intercollegiate Running Club Association. 2026. NIRCA.

- [49] Nicole L Nelson and James R Churilla. 2016. A narrative review of exercise-associated muscle cramps: factors that contribute to neuromuscular fatigue and management implications. *Muscle & nerve*, 54, 2, 177–185.
- [50] Rasmus Oestergaard Nielsen, Lotte Rønnow, Sten Rasmussen, and Martin Lind. 2014. A prospective study on time to recovery in 254 injured novice runners. *PLoS one*, 9, 6, e99877.
- [51] Rasmus Østergaard Nielsen, Erik Thorlund Parner, Ellen Aagaard Nohr, Henrik Sørensen, Martin Lind, and Sten Rasmussen. 2014. Excessive progression in weekly running distance and risk of running-related injuries: an association which varies according to type of injury. *Journal of orthopaedic & sports physical therapy*, 44, 10, 739–747.
- [52] [n. d.] Openweathermap api. <https://openweathermap.org/api>. ()
- [53] Richard Michael Page, Adam Field, Ben Langley, Liam David Harper, and Ross Julian. 2023. The effects of fixture congestion on injury in professional male soccer: a systematic review. *Sports Medicine*, 53, 3, 667–685.
- [54] FRANCOIS Peronnet, G Thibault, and DL Cousineau. 1991. A theoretical analysis of the effect of altitude on running performance. *Journal of Applied Physiology*.
- [55] François Péronnet and Guy Thibault. 1989. Mathematical analysis of running performance and world running records. *Journal of applied physiology*, 67, 1, 453–465.
- [56] Hugh C Pinnington and Brian Dawson. 2001. The energy cost of running on grass compared to soft dry beach sand. *Journal of science and medicine in sport*, 4, 4, 416–430.
- [57] Lewis GC Evans Pugh. 1971. The influence of wind resistance in running and walking and the mechanical efficiency of work against horizontal or vertical forces. *The Journal of physiology*, 213, 2, 255–276.
- [58] Mike Quinn. 2004. The effects of wind and altitude in the 400-m sprint. *Journal of sports sciences*, 22, 11–12, 1073–1081.
- [59] Mike D Quinn. 2003. The effects of wind and altitude in the 200-m sprint. *Journal of applied biomechanics*, 19, 1, 49–59.
- [60] Peter S Riegel. 1981. Athletic records and human endurance: a time-vs.-distance equation describing world-record performances may be used to compare the relative endurance capabilities of various groups of people. *American Scientist*, 69, 3, 285–290.
- [61] Steven J Rigatti. 2017. Random forest. *Journal of insurance medicine*, 47, 1, 31–39.
- [62] Jeffrey A Rothschild, Tom Stewart, Andrew E Kilding, and Daniel J Plews. 2024. Predicting daily recovery during long-term endurance training using machine learning analysis. *European Journal of Applied Physiology*, 124, 11, 3279–3290.
- [63] Maximum Performance Running. 2013. Temperature and dew point. <https://maximumperformancerunning.blogspot.com/2013/07/temperature-dew-point.html>. (2013).
- [64] Martin P Schwellnus, Nicola Drew, and Malcolm Collins. 2011. Increased running speed and previous cramps rather than dehydration or serum sodium changes predict exercise-associated muscle cramping: a prospective cohort study in 210 ironman triathletes. *British journal of sports medicine*, 45, 8, 650–656.
- [65] MP Schwellnus. 2009. Cause of exercise associated muscle cramps (eacm)—altered neuromuscular control, dehydration or electrolyte depletion? *British journal of sports medicine*, 43, 6, 401–408.
- [66] [n. d.] Sdg resource document — targets overview. [https://sdgs.un.org/sites/default/files/2020-09/SDG%20Resource%20Document\\_Targets%20Overview.pdf](https://sdgs.un.org/sites/default/files/2020-09/SDG%20Resource%20Document_Targets%20Overview.pdf). ()
- [67] Stephen Seiler. 2010. What is best practice for training intensity and duration distribution in endurance athletes? *International journal of sports physiology and performance*, 5, 3, 276–291.
- [68] Lisa M Sinclair and Pamela Sue Hinton. 2005. Prevalence of iron deficiency with and without anemia in recreationally active men and women. *Journal of the American Dietetic Association*, 105, 6, 975–978.
- [69] Ella S Smith, Alannah KA McKay, Megan Kuikman, Kathryn E Ackerman, Rachel Harris, Kirsty J Elliott-Sale, Trent Stellingwerff, and Louise M Burke. 2022. Auditing the representation of female versus male athletes in sports science and sports medicine research: evidence-based performance supplements. *Nutrients*, 14, 5, 953.
- [70] J Spiegel and JR Mureika. 2003. A model of wind and altitude effects on 110-m hurdles. *Sportscience*, 7, 1–12.
- [71] D Travis Thomas, Kelly Anne Erdman, Louise M Burke, et al. 2016. Nutrition and athletic performance. *Med Sci Sports Exerc*, 48, 3, 543–568.
- [72] Track & Field Results Reporting System. 2026. TFRS: Track & Field Results Reporting System.
- [73] Ross Tucker and Timothy D Noakes. 2009. The physiological regulation of pacing strategy during exercise: a critical review. *British journal of sports medicine*, 43, 6, e1–e1.
- [74] Alexandra S Voloshina and Daniel P Ferris. 2015. Biomechanics and energetics of running on uneven terrain. *The journal of experimental biology*, 218, 5, 711–719.
- [75] Mark D Wilkinson et al. 2016. The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3, 1, 1–9.
- [76] Kevin J Wilson and Nina Wilson. 2024. Assessing course difficulty and the effect of weather in amateur cross country running races. *arXiv preprint arXiv:2405.09865*.
- [77] E Zender-Świercz, B Galiszewska, M Telejko, and M Starzomska. 2024. The effect of temperature and humidity of air on the concentration of particulate matter-pm2.5 and pm10. *Atmospheric Research*, 312, 107733.
- [78] Weichen Zhao. 2015. The effect of weather on track performances: a panel data analysis on carolina godiva track club summer track races.

## Appendix

The appendix is an extension of the CIKM paper submission.

## A Supplementary Validation

### A.1 Stepwise variance decomposition (XC)

For athlete-seasons with  $\geq 3$  regular-season XC meets (nationals excluded), Table 10 reports median cross-meet standard deviation for four result-time variants: *raw*; *converted-only* (Riegel + measured course length); *pre-weather env.* (grade, altitude, course length on top of converted); and *fully standardized* (+ heat). Raw→full is percent SD reduction from raw to the final column.

**What drives the gender gap?** Most reduction occurs at the *distance-conversion* step, for women (raw 108 s to converted 55 s) due to incorrect course length. The pre-weather column (grade + altitude + length, no heat) moves medians only  $\sim 0.3$  to 0.5 s; heat adds another  $\sim 1.6$  to 2.9 s ( $\sim 3$  to 5% incremental compression beyond converted-only). Women’s larger raw→full percent (51.0% vs. 34.4%) reflects a wider pre-adjustment distribution on heterogeneous course distances (6 km targets with variable measured lengths), not larger heat-index swings across meets.

### A.2 Environmental factor distributions (comprehensive XC)

Table 11 summarizes adjustment factors on comprehensive-era XC results ( $n=20,095$  after nationals exclusion). Medians near 1.0 confirm that population-level corrections are small on average; tails matter for individual meets.

**Elevation check.** After converting stored gain/loss from feet to grade percent before Maurer factors, the elevation median is exactly 1.0 with P95 1.002, consistent with the corrected unit convention. Weather drives the only material tail mass among active factors at the median athlete-season.

### A.3 Within-season improvement direction

Converted-only result times overstate not only *magnitude* of within-season improvement (12% / 21% median inflation for women / men) but also the *share* of athlete-seasons classified as improving (first race slower than later best): 70.2% / 68.2% under converted-only vs. 67.8% / 65.9% under full standardization ( $\sim 2.3$  pp gap). Directional bias is modest but aligned with inflation: environmental drift in unadjusted result times mimics fitness gains.

### A.4 Illustrative ML $R^2$ and bootstrap uncertainty

The illustrative analysis is a *demonstration*, not a benchmark: gender-stratified models predict within-season improvement *rate* (s/day) from season-level features derived from the same result times (train 2023, test 2024;  $\geq 2$  XC meets). Table 12 reports random forest and

| Group        | <i>n</i> | Raw   | Conv. | Pre-wx | Full | Raw→full |
|--------------|----------|-------|-------|--------|------|----------|
| Women (6 km) | 2,167    | 108.2 | 54.9  | 54.6   | 53.0 | 51.0%    |
| Men (8 km)   | 3,929    | 85.1  | 59.1  | 58.7   | 55.8 | 34.4%    |

**Table 10: Median within-athlete cross-meet SD (seconds) by adjustment stage.**

| Factor subset             | Median | Mean  | P05 to P95     | % outside [0.95, 1.05] |
|---------------------------|--------|-------|----------------|------------------------|
| Weather (all rows)        | 0.994  | 0.990 | 0.969 to 1.000 | 0.4                    |
| Weather ( $H > 100$ only) | 0.990  | 0.988 | 0.966 to 1.000 | 0.5                    |
| Elevation (grade)         | 1.000  | 1.000 | 1.000 to 1.002 | 0.0                    |
| Altitude (Peronnet)       | 0.996  | 0.996 | 0.993 to 1.000 | 0.0                    |

**Table 11: Factor  $f$  such that adjusted time =  $f \times$  prior result time (comprehensive XC).**

| Model | Gender | Conv. | Std.  | $\Delta R^2$ 95% CI |
|-------|--------|-------|-------|---------------------|
| RF    | Men    | 0.737 | 0.720 | [−0.003, +0.039]    |
| RF    | Women  | 0.735 | 0.723 | [−0.029, +0.051]    |
| GB    | Men    | 0.610 | 0.571 | [+0.010, +0.073]    |
| GB    | Women  | 0.581 | 0.515 | [+0.023, +0.111]    |

**Table 12: Test  $R^2$  by result-time adjustment.  $\Delta R^2 = R^2_{\text{conv}} - R^2_{\text{std}}$ .**

gradient boosting; 95% CIs are percentile bootstrap on the test set (2,000 replicates).

**Why report  $R^2$  at all?** Features include within-season time summaries (first\_time, best\_time, slope, etc.) built from the result times under test, so high  $R^2$  is expected and is *not* external predictive validity. The relevant comparison is *paired*: same features, same splits, two result-time variants. Converted-only times embed fall cooling and course grade; standardized times partial those out. Higher  $R^2$  under converted-only therefore signals that the model can exploit environmental structure left in the label (confounding, not added athlete signal).

**Why RF  $\approx$  GB?** Direction is identical (converted  $\geq$  standardized). RF gaps are small and not significant at 95%; GB gaps are larger and significant for both genders. This suggests that a flexible but regularized fit still exploits environmental structure left in converted-only labels more than RF does. Neither model uses weather columns directly; environment enters only through result times.

## A.5 Extended robustness checks

**A.5.1 Riegel exponent sensitivity.** Gender-specific exponents ( $b=1.055$  men,  $b=1.08$  women) vs. a unified  $b=1.06$  shift the XC variance-validation percentages by at most 0.6 percentage points (Table 13), consistent with the main-text footnote.

**A.5.2 Finisher-only export and minimum-race sensitivity.** The public export contains approved *finisher* times only; DNS/DNF are absent as rows. XC improvement inflation is stable when requiring  $\geq 2$ ,  $\geq 3$ , or  $\geq 4$  races per athlete-season (men 18.6 to 20.7%, women 8.7 to 13.4%; Table 14). Only 19.0% of XC athlete-seasons have  $\geq 3$  races, so sparse-season athletes dominate the  $\geq 2$  race demo.

| Riegel config             | Men raw→std | Women raw→std | Max $ \Delta _{\text{pp}}$ |
|---------------------------|-------------|---------------|----------------------------|
| Gender-specific (default) | 34.4%       | 51.0%         | 0.0                        |
| Unified $b=1.06$          | 34.5%       | 50.7%         | 0.3                        |
| Unified $b=1.055$         | 34.4%       | 50.5%         | 0.5                        |
| Unified $b=1.08$          | 35.0%       | 51.0%         | 0.6                        |

**Table 13: Median within-athlete SD reduction under alternate Riegel exponents ( $\geq 3$  meets/season).**

| Min races | Men <i>n</i> | Men infl. | Women <i>n</i> | Women infl. |
|-----------|--------------|-----------|----------------|-------------|
| 2         | 8,986        | 20.7%     | 5,548          | 11.6%       |
| 3         | 3,979        | 18.9%     | 2,192          | 8.7%        |
| 4         | 1,290        | 18.6%     | 705            | 13.4%       |

**Table 14: Converted-only median improvement inflation vs. full standardization by minimum races per athlete-season.**

| Sport         | <i>n</i> | Median factor | % outside [0.98, 1.02] |
|---------------|----------|---------------|------------------------|
| Cross Country | 31,354   | 0.986         | 51.5                   |
| Outdoor Track | 17,180   | 0.998         | 10.4                   |
| Indoor Track  | 15,124   | 1.000         | 0.1                    |
| Road Race     | 848      | 0.998         | 1.5                    |

**Table 15: Comprehensive-era running results: full sport-appropriate adjustment pipeline.**

**A.5.3 Outdoor track wind and venue factors.** Comprehensive outdoor track ( $n=18,438$  parseable results): sprint/hurdle events ( $\leq 400$  m) with wind logic applicable on 6,939 rows, but wind *recorded* on only 6.4% of those. Among applicable rows the wind-factor median is 1.0 (P95 1.042). NCAA venue length/banking factors are unity at the median for this corpus (lap metadata sparse). The full track pipeline (wind when present, heat, altitude) yields median factor 0.998 with 10.6% of rows outside [0.98, 1.02]; meet-level weather coverage remains 3.8%.

**A.5.4 Cross-sport standardization magnitude.** Table 15 summarizes comprehensive-era adjustment factors (adjusted/raw) for running events. XC shows the largest tail mass (51.5% outside [0.98, 1.02]) because distance conversion and course metadata move times materially. Indoor track is tight at the median (0.10% outside band); outdoor track and road sit between.

**A.5.5 Meet-redacted temporal ML split.** The illustrative ML demo uses train 2023 / test 2024. To stress meet-level leakage, we rebuild 2024 test features after randomly withholding 25% of 2024 meets (25 meets) from each athlete’s race history; labels are unchanged. Random-forest  $R^2$  shifts slightly but the conv. vs. std. gap remains small (men  $\Delta R^2 = -0.012$ ; women  $+0.011$ ; Table 16), comparable in magnitude to the temporal bootstrap CIs in Table 12.

## B Cross Country Performance Band and NCAA Context

NIRCA club competition spans NCAA Division I-caliber athletes to recreational 5K-equivalent efforts. To contextualize ability spread,

| Gender | $R^2$ conv. | $R^2$ std. | $\Delta R^2$ |
|--------|-------------|------------|--------------|
| Men    | 0.721       | 0.733      | -0.012       |
| Women  | 0.753       | 0.742      | +0.011       |

**Table 16: Meet-redacted temporal split (train 2023; test 2024 with 25% of meets removed from feature construction).**

| Gender       | $n$    | P1    | P5    | P10   | P25   | P50   | P75   | P90   | P95   | P99   |
|--------------|--------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Men (8 km)   | 12,812 | 25:05 | 26:11 | 26:52 | 28:19 | 30:22 | 33:08 | 36:25 | 39:13 | 47:27 |
| Women (6 km) | 8,818  | 21:52 | 23:08 | 23:54 | 25:23 | 27:21 | 29:47 | 32:40 | 34:38 | 40:35 |

**Table 17: Career PR for *all* athletes with  $\geq 1$  parseable XC result (mm:ss); nationals included.**

we report each athlete’s career *personal record* (PR): the fastest **fully standardized** XC result time across all school years and meets in the export, **including the NIRCA nationals championship** (8 km men, 6 km women). Tables 20–21 use these standardized PRs, not raw meet race times. Nationals are excluded from formula-validation summaries in the main text because they are a single end-of-season championship, but they belong in PR tables because many athletes peak there. One row per athlete, not per result or per season. Table 17 summarizes the full export; Table 18 restricts to **longitudinal subsets** used in season-trajectory and improvement work ( $\geq 2$  or  $\geq 3$  distinct school years with XC results; optional  $\geq 2$  races in each qualifying year to mirror the main-text improvement demo, which still excludes nationals from within-season slopes). Multi-season athletes are faster at every percentile (retention bias: committed runners return; one-meet newcomers drop out).

**Anchor definition.** Collegiate comparison times are *team scoring depth*, not individual place at nationals or whole-roster medians. XC teams score their top five finishers; sixth and seventh runners displace opponents but do not add points. We therefore anchor on the **5th runner** among the **top-15 teams** at each division’s national championship—the slowest counting scorer on a top-tier team. This requires only five finishers per team (nearly all championship teams field at least five; only  $\sim 95\%$  field seven). NIRCA anchors are fit from standardized NRCD nationals results (2015+); NCAA, NAIA, and NJCAA anchors are approximate from published championship team results. Anchors are illustrative, not formal equivalence tests.

**NIRCA vs. collegiate team depth.** Table 19 orders benchmarks; Table 20 counts exceedance; Table 21 lists the *between-anchor* bands that explain where NIRCA club runners sit relative to varsity nationals depth. For men, NIRCA ( $\sim 27:00$ ) sits between D1 and D2; 7.2% of athletes are in the NIRCA–D1 gap. For women, D2 and NIRCA share  $\sim 24:00$ , so the salient gap is to D1 ( $\sim 22:00$ ): only 1.3% of career PRs reach D1 5th-runner pace vs. 10.8% at NIRCA/D2.

**Interpretation.** About 25% of career PRs occur at NIRCA nationals meets, so excluding them materially understated the elite tail. The full export (Table 17) still mixes one-meet recreational starters with multi-year club runners; P99 (men 47:27, women 40:35) reflects that long tail. Longitudinal cohorts (Table 18) are the relevant ability reference for papers that filter to  $\geq 2$  school years (35.7% of XC athletes in Section D) or require  $\geq 2$  regular-season races per season before estimating improvement: median PR shifts  $\sim 1:20$

| Cohort                         | Gender | $n$    | P5    | P50   | P95   |
|--------------------------------|--------|--------|-------|-------|-------|
| All athletes                   | Men    | 12,812 | 26:11 | 30:22 | 39:13 |
| All athletes                   | Women  | 8,818  | 23:08 | 27:21 | 34:38 |
| $\geq 2$ school years          | Men    | 4,783  | 25:37 | 28:58 | 35:04 |
| $\geq 2$ school years          | Women  | 2,947  | 22:29 | 25:57 | 31:16 |
| $\geq 3$ school years          | Men    | 2,199  | 25:20 | 28:19 | 33:33 |
| $\geq 3$ school years          | Women  | 1,262  | 22:08 | 25:18 | 29:50 |
| $\geq 2$ yr, $\geq 2$ races/yr | Men    | 2,980  | 25:22 | 28:24 | 33:43 |
| $\geq 2$ yr, $\geq 2$ races/yr | Women  | 1,790  | 22:10 | 25:23 | 29:56 |

**Table 18: Career PR by longitudinal cohort (Aug 1 school-year boundary; Section D). Anchor exceedance by cohort is in Table 20.**

| Benchmark                             | Men (8k)     | Women (6k)   |
|---------------------------------------|--------------|--------------|
| NCAA D1 All-American                  | $\sim 24:00$ | $\sim 20:00$ |
| NCAA D1 5th runner (top-15 teams)     | $\sim 26:00$ | $\sim 22:00$ |
| NIRCA nationals 5th runner (NRCD fit) | $\sim 27:00$ | $\sim 24:00$ |
| NCAA D2 5th runner                    | $\sim 28:00$ | $\sim 24:00$ |
| NAIA 5th runner                       | $\sim 28:30$ | $\sim 24:30$ |
| NCAA D3 5th runner                    | $\sim 29:30$ | $\sim 25:30$ |
| NJCAA 5th runner                      | $\sim 28:30$ | $\sim 25:30$ |

**Table 19: Standardized anchor *times* (fastest→slowest). “5th runner” = fifth counting finisher on each of the top-15 teams at that division’s nationals (Section B).**

| Anchor <sup>§</sup>        | Men   | Women |
|----------------------------|-------|-------|
| NCAA D1 All-American       | 0.1%  | 0.0%  |
| NCAA D1 5th runner         | 4.1%  | 1.3%  |
| NCAA D2 5th runner         | 21.2% | 10.8% |
| NCAA D3 5th runner         | 39.6% | 26.6% |
| NAIA 5th runner            | 27.2% | 15.1% |
| NJCAA 5th runner           | 27.2% | 26.6% |
| NIRCA nationals 5th runner | 11.3% | 10.8% |

**Table 20: Share of athletes whose standardized career PR is at or faster than each anchor (Table 19). <sup>§</sup>Threshold labels shortened; times in companion table.**

to 1:30 faster, and the share at NIRCA 5th-runner pace roughly doubles (e.g. men 11.3% to 21–29%; women 10.8% to 20–27%). Even in those subsets, NCAA D1 5th-runner pace remains uncommon ( $\leq 12\%$  men,  $\leq 4\%$  women).

## C Metadata Missingness: Results vs. Meets

Result-level coverage counts every row with a join to course metadata; meet-level coverage counts a meet if *any* result in that meet has the field. Meet-level rates are lower when large invitationals contribute many results from a single weather-enriched row, and are often the more honest unit for season-trajectory analysis (one environment per competition day).

| Band (faster anchor → slower) | Men   | Women | Interpretation         |
|-------------------------------|-------|-------|------------------------|
| D1 5th → NIRCA 5th            | 7.2%  | 9.5%  | NIRCA depth, not D1    |
| NIRCA 5th → D2 5th            | 9.9%  | —     | D2 depth, not NIRCA    |
| NIRCA 5th → D3 5th            | 28.3% | 15.8% | D3 depth, not NIRCA    |
| NIRCA 5th → NAIA 5th          | 15.9% | 4.3%  | NAIA depth, not NIRCA  |
| NIRCA 5th → NJCAA 5th         | 15.9% | 15.8% | NJCAA depth, not NIRCA |

**Table 21: Athletes between adjacent anchors (standardized career PR). Read row as: faster than the slower anchor but slower than the faster one. “—”: anchors coincide (women’s D2 and NIRCA both ~24:00).**

| Sport      | Era           | Weather (%) |      | Course feat. (%) |      |
|------------|---------------|-------------|------|------------------|------|
|            |               | Res.        | Meet | Res.             | Meet |
| XC         | comprehensive | 97.7        | 90.7 | 38.5             | 18.2 |
| XC         | historical    | 4.4         | 3.3  | 3.4              | 0.8  |
| Out. track | comprehensive | 36.9        | 3.8  | —                | —    |
| Road       | comprehensive | 67.0        | 40.5 | 27.1             | 16.7 |

**Table 22: Metadata coverage: share of results vs. share of meets (altitude omitted; >97% throughout).**

**Big-meet skew.** Among comprehensive-era XC meets ( $n=280$ ), the largest 10 meets contain 28.0% of all results; the largest 25 contain 45.5%. Median meet size is 25 results (max 1,361). Result-level weather coverage therefore overstates how uniformly metadata blanket the calendar.

**Future releases.** Planned extensions include backfilling course gain/loss and measured length for high-volume invitationals, expanding outdoor-track weather joins, and documenting per-meet completeness in the datasheet. Historical rows (pre-August 2023) will remain sparse by design.

### C.1 Geolocation, altitude, and weather provenance

Comprehensive-era metadata is attached to meets and courses, not to individual athletes or GPS tracks. When Enrichment looks up the meet city and state through OpenWeather Geocoding and uses the first US match. One point represents the whole meet or course row, not points along the race route.

Venue altitude on the meet and course records is in feet from USGS EPQS at those coordinates. It reflects the geocoded host location (often near city center), not elevation sampled along the course. A rural park can be several miles from that point.

Weather comes from OpenWeather One Call 3.0 timemachine at the same coordinates. The listed race date and start time are converted to local time (via TimeZoneDB, or the meet timezone when stored). OpenWeather returns one hourly snapshot, not a minute-level reading. For example, a 10:55 am start uses the 10:00 am hour block. Conditions at the API point may differ from a wooded or exposed part of the course. Air-quality history is only available for meets on or after 27 November 2020. Indoor track is not weather-enriched. See Table 22 for coverage rates and Section G.16 for other race-context factors standardization does not capture.

| Sport             | Athletes      | ≥ 2 yr       | P50 meets  | P90 meets  | P50 res.   | P90 res.   |
|-------------------|---------------|--------------|------------|------------|------------|------------|
| Cross Country     | 21,630        | 35.7%        | 2          | 4          | 2          | 4          |
| Outdoor Track     | 9,607         | 30.3%        | 1          | 3          | 2          | 5          |
| Indoor Track      | 7,207         | 27.2%        | 1          | 3          | 2          | 5          |
| Road Race         | 1,627         | 20.3%        | 1          | 1          | 1          | 1          |
| <b>All sports</b> | <b>28,913</b> | <b>35.7%</b> | <b>n/a</b> | <b>n/a</b> | <b>n/a</b> | <b>n/a</b> |

**Table 23: Per sport: athletes with results, % with ≥ 2 school years, and within-season meets/results per athlete-season (medians and P90). All-sports row from full export.**

## D Longitudinal Depth by Sport

Longitudinal modeling requires repeated athlete-seasons. Table 23 summarizes school-year spans (August 1 boundary) and within-season meet/result counts.

**Interpretation.** XC offers the densest repeated-meet structure (median 2 meets/season, P90 4) and the highest multi-year retention among sports (35.7% of XC athletes with ≥ 2 school years; 16.0% with ≥ 3). Track athletes accumulate more *results* per season (multiple events per meet) but fewer *meets*. Road racing is sparse (median one result per athlete-season).

## E Public Export Structure and Zenodo Release

The analysis in this paper uses the **approved public CSV export** (v2.0.0, May 2026). Zenodo (<https://zenodo.org/records/17917357>) is permanently linked to the GitHub source repository [https://github.com/National-Running-Club-Database/national\\_running\\_club\\_database\\_public\\_dataset](https://github.com/National-Running-Club-Database/national_running_club_database_public_dataset); each Zenodo release is a snapshot of that repo. The live submission site (<https://nationalrunningclubdatabas.e.com>) accepts new meets continuously. Tables 24 to 26 summarize tables, joins, and *key fields* as documented in DATASHEET.md and README.md on GitHub [20]; full CSV schemas are in README.md, not repeated here.

**Release filtering.** Direct identifiers (legal names, profile URLs, contact fields) are removed before Zenodo upload. Athletes receive stable pseudonymous *athlete\_id* values; teams retain institution names because they are already public on meet results pages. DNS/DNF rows are absent; the export is finisher-only.

**Re-identification risk and out-of-corpus linkage.** Meet times, team, and approximate date are often already *public* on host-university results sites, Milesplit, TFRS, and NIRCA meet pages linked from the project website. NRCD therefore does not introduce a wholly new class of sensitive facts for most rows; it aggregates and standardizes information that motivated runners and coaches can already look up. Residual risk remains: a motivated party could join *athlete\_id* across meets within NRCD, or link a small number of distinctive performances back to public web results (*out-of-corpus re-identification*). We treat this as a misuse scenario rather than a product goal: researchers should not attempt to recover identities, publish deanonymized subsets, or merge NRCD with social-media graphs. IRB review classified the public export as non-human-subjects research with PII removed; downstream work should cite the Zenodo DOI and respect the datasheet use constraints (DATASHEET.md in the GitHub dataset repository).

| Table                        | Primary key                         | Role                                                      |
|------------------------------|-------------------------------------|-----------------------------------------------------------|
| result.csv                   | result_id                           | Finisher time, athlete, meet, event                       |
| meet.csv                     | meet_id                             | Date, sport, venue altitude, nationals flag               |
| athlete.csv                  | athlete_id                          | Gender; <i>no names in public export</i>                  |
| team.csv                     | team_id                             | Institution / club label                                  |
| athlete_team_association.csv | (athlete_id, team_id)               | Roster membership by season                               |
| running_event.csv            | running_event_id                    | Event name and distance                                   |
| sport.csv                    | sport_id                            | XC, indoor/outdoor track, road                            |
| course_details.csv           | (meet_id, running_event_id, gender) | Weather, grade, measured length                           |
| joined.csv                   | result_id                           | Denormalized merge of result, meet, team, athlete, course |

**Table 24: NRCD public export tables (v2.0.0: 128,963 approved results, 28,913 athletes with  $\geq 1$  result, 1,336 meets, 187 teams). Only approved=True meets and finishers are released.**

| File                         | Key fields                                                                                                   |
|------------------------------|--------------------------------------------------------------------------------------------------------------|
| athlete.csv                  | athlete_id, gender, is_hidden                                                                                |
| athlete_team_association.csv | athlete_id, team_id                                                                                          |
| meet.csv                     | meet_name, start_date, meet_city, meet_state, altitude, sport_id, nirca_only, regionals, nationals, approved |
| result.csv                   | athlete_id, meet_id, running_event_id, result_time, grade, wind; relay athlete_id_2 to 4 and splits          |
| running_event.csv            | event_name                                                                                                   |
| sport.csv                    | sport_name                                                                                                   |
| team.csv                     | team_name, region, city, state                                                                               |

**Table 25: Public-export key fields (relational tables).**

| File               | Key fields                                                                                                                                                                |
|--------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| course_details.csv | elevation_gain, elevation_loss, estimated_course_distance, date_of_event, temperature, dew_point, humidity, barometric_pressure, wind_speed, weather_conditions, altitude |
| joined.csv         | Denormalized key fields from result, meet, team, athlete, and course tables; wind, result_time, grade, relay IDs/splits                                                   |

**Table 26: Public-export key fields**

**Recommended acquisition path.** Download the Zenodo archive (<https://zenodo.org/records/17917357>) or clone / create a submodule of the GitHub dataset repository ([https://github.com/National-Running-Club-Database/national\\_running\\_club\\_database\\_public\\_dataset](https://github.com/National-Running-Club-Database/national_running_club_database_public_dataset)). Read DATASHEET.md and README.md there for column semantics, era splits, and intended uses.

## F Empowering Users with the nrCD Python Package

Standardization formulas in the main text are implemented in the standalone **nrCD** package (PyPI: <https://pypi.org/project/nrCD/>; source: <https://github.com/National-Running-Club-Database/nrCD>), decoupled from the Zenodo CSVs. Athletes, coaches, and researchers can apply the same adjustments to private or newly collected results. API details are in PARAMETERS\_DOC and the package README.

**Use case 1: fair cross-meet XC comparison.** A woman runs 24:10 on a listed 6K course that actually measures 6.2 km, with ~120 m of climb at 78 °F; `standardize_xc` converts distance, grade, heat, and altitude to a comparable 6 km finish time. Distances accept km, mi, or meters; weather defaults to °F or `temp_unit="C"`; grade accepts vertical m, feet, or percent:

**Use case 2: outdoor wind vs. indoor banked track.**

`standardize_outdoor_track` applies `wind_mps` on outdoor sprints only; `standardize_indoor_track` uses `lap_length_m` and banked indoors with no wind term.

**Use case 3: reproducible research on Zenodo.** Load approved CSVs with the package dataset helpers and regenerate paper statistics from the same formulas and `StandardizeConfig` coefficients used in Appendix A.5.1.

**Use case 4: metadata backfill for new meets.**

`enrich_race_context` can populate temperature, dew point, and U.S. venue altitude from city, state, and race datetime; standardization itself requires no API keys.

The package is MIT-licensed open source on GitHub: open an issue when something is missing or incorrect, or fork and submit pull requests with improvements. Cite the resource paper when publishing work that relies on its formulas.

## G Factors Not Covered by NRCD Standardization

Researchers designing studies with NRCD should treat the factors below as *potential covariates*, *stratification variables*, or *sensitivity analyses*. They are not defects in the export, but are latent modifiers

```

from nrcd.standardize import format_time, standardize_xc

std = standardize_xc(
    "24:10",
    gender="F",
    reported_distance="6k",          # alt: 6.2, "6.2", reported_distance_m=6000
    distance_unit="km",             # alt: "mi"; or pass *_distance_m in meters
    actual_distance=6.2,            # alt: actual_distance_m=6200
    temperature=78, dew_point=70,   # °F default; alt: temp_unit="C"
    elevation_gain=120,
    elevation_loss=120,
    grade_input="m",                # alt: "feet", or omit for % grade
    target_distance=6000,           # alt: "6k", "6000m", target_distance_m=6000
)
print(format_time(std)) # e.g. 21:45.67 on standardized 6 km

```

that standardized times do not encode. NRCD adjusts officiated meet times for distance, course length, grade, heat, altitude, and (track) venue and wind where metadata exist; the subsections below cite literature on what remains unmodeled so downstream work can bound or explicitly model it [60, 38].

### G.1 Training load, aerobic fitness, and recovery markers

**Training volume and periodization:** Weekly mileage, intensity distribution, taper length, and recency of hard sessions all influence race readiness on a given weekend [67]. Longer horizons matter too: summer base miles, off-season work, and pre-season preparation predict in-season performance in collegiate distance runners, not just the prior week [36]. Meta-analytic evidence supports structured tapers: about two weeks with 41 to 60% volume reduction while maintaining intensity yields measurable performance gains [8].

**Aerobic capacity:** Meet results do not encode lab- or field-derived fitness such as  $\dot{V}O_{2\max}$ , running economy, or lactate and ventilatory thresholds, even though these traits explain much of the spread in endurance performance [5]. Standardized times, therefore, conflate a one-day outcome with an unobserved aerobic profile built over months of training.

**Heart-rate variability:** HRV from wearables or morning resting measures is used to flag autonomic recovery and training readiness across multi-week blocks [62]. NRCD has no HRV, sleep-stage, or training-log fields, so day-to-day freshness is latent even when meet dates are known.

**Body composition:** Seasonal changes in body mass and lean mass shift running economy even when aerobic fitness is stable [28]. NRCD does not record weight, body-fat percentage, or composition scans.

NRCD records *when* athletes raced, not *how* they trained beforehand, their underlying aerobic capacity, or their recovery state that week; two athletes with identical standardized times may differ sharply in summer training load,  $\dot{V}O_{2\max}$ , HRV, glycogen stores, and neuromuscular readiness.

### G.2 Within-season race frequency and cumulative fatigue

Meet density over the prior two to three weeks is distinct from chronic training load: repeated competitions compress recovery

even when weekly mileage is unchanged [29, 53]. NRCD lists every finisher time but not how many races an athlete started in the surrounding fortnight, so short-term staleness and fixture congestion are latent when interpreting a single slow result or a within-season improvement slope (Section A.3).

### G.3 Injury and illness

Acute and overuse injuries alter biomechanics, limit training continuity, and often appear in meet data as DNS/DNF rather than slow finish times [50, 51]. Since NRCD only records finish times, this is not taken into account. Prospective cohort studies link rapid weekly distance progression (for example, >30% week over week) to elevated hazard of patellofemoral pain, iliotibial band syndrome, and related distance-running injuries [51]. Upper-respiratory illness, concussion return-to-play protocols, and subclinical illness similarly decouple observed performance from underlying capacity. Without injury logs or training diaries, latent health state remains unobserved.

### G.4 Sleep

Sleep loss is common before competition and can impair sport-specific performance, cognitive function, and autonomic balance even when maximal strength tasks are preserved [19]. Partial sleep restriction elevates perceived exertion and slows decision-relevant processing, both relevant to mass-start XC and tactical track races. NRCD has no sleep-duration or sleep-quality fields; late-night travel, exam weeks, and early bus departures for away meets are invisible confounders.

### G.5 Nutrition and energy availability

Carbohydrate availability, pre-race fueling, and day-of-meet energy intake influence glycogen stores, perceived exertion, and late-race pace regulation. Joint position statements recommend individualized plans for macronutrient timing, recovery nutrition, and safe supplementation [71]. Low energy availability, whether from intentional restriction, poor meal access on travel weekends, or academic schedule disruption, can impair performance independently of fitness and overlaps with RED-S pathways in women (Section G.14.4) and men (Section G.15) [43]. NRCD records no dietary logs, body mass on race day, or supplement use; two athletes with identical

standardized times may differ in prerace carbohydrate loading or micronutrient status (e.g. iron, vitamin D).

## G.6 Caffeine and legal ergogenic aids

Caffeine is widely used in collegiate racing and can improve endurance performance at moderate doses, independent of weather and grade corrections [21, 71]. Beetroot nitrate, sodium bicarbonate, and other legal aids show smaller or event-specific effects in position-stand summaries [71]. NRCD stores no supplement timing or dose fields, so identical standardized times may reflect different prerace ergogenic protocols.

## G.7 Hydration, dehydration, and exercise-associated cramping

Fluid deficits can raise cardiovascular strain and degrade endurance performance in warm conditions, especially when athletes underdrink relative to sweat loss [71]. Exercise-associated muscle cramping (EAMC) during hard mass-start XC is often attributed to dehydration or sodium loss, but prospective endurance studies find cramping more closely associated with higher running intensity, prior cramp history, and neuromuscular fatigue than with abnormal serum electrolytes or greater body-mass loss at the time of cramp [65, 64, 49]. NRCD does not record fluid intake or cramp events; a single slow finish time cannot distinguish cramp-shortened effort from fitness limits. Heat standardization adjusts expected thermal stress but not hydration behavior.

## G.8 Travel and jet lag

Long-haul travel induces travel fatigue (discomfort, mild hypoxia, disrupted routines) and, when crossing multiple time zones, jet lag (circadian misalignment, sleep disruption, gastrointestinal symptoms) [32]. Controlled studies report impaired sprint and intermittent endurance performance for several days after eastward transmeridian flights, with larger effects than westward travel of equal duration [18]. Even same-day car trips to regional meets can shorten sleep and disrupt prerace fueling [32, 71]. Travel therefore remains a latent performance modifier for away meets throughout the season. XC validation and within-season improvement summaries in the main text **exclude NIRCA nationals**, which *partially* limits how much travel confounds those checks; it does not remove travel effects on regular-season results. NRCD encodes meet and team location but not hours in transit or arrival time relative to race start.

## G.9 Academic load and student-athlete time demands

NCAA student-athletes report median in-season commitments near 66.5 hours per week combined across athletics and academics [45], with exam periods, lab deadlines, and missed class time adding stress independent of training load. Systematic reviews document elevated depression and anxiety prevalence relative to non-athlete students, often under-reported because of stigma and performance culture [4]. Cognitive fatigue and all-nighter study sessions can blunt race-day effort regulation even when physical training is unchanged. NRCD links athletes to institutions but not to course

schedules, GPA pressure, or academic calendar events. Additionally, several NIRCA clubs are student run, meaning that certain runners also need to deal with race-day logistics.

## G.10 Circadian rhythm and time of day

Morning versus afternoon start times shift core temperature, hormone profiles, and perceived effort even when weather metadata are held constant [33]. `time_of_event` is stored for comprehensive-era rows, but standardization does not adjust for circadian phase; a cool 8:00 am race and a warm 4:00 pm race on the same course are not equated for body-clock effects. Travel-related circadian disruption is covered separately (Section G.8).

## G.11 Precipitation, footing, and solar load

Rain, snow, mud, and saturated grass change grip and energy cost in ways that temperature and dew point alone do not capture [34, 56, 76]. Comprehensive-era exports may include precipitation fields from OpenWeather, but the standardization pipeline does not apply a footing factor; wet-course effects remain in the residual (Section G.13). Direct solar radiation and UV exposure on open courses add radiant heat load beyond the heat-index surrogate used for air temperature and humidity [9, 78]. Cloud cover and UV index are sometimes present in the course details but are not part of the published adjustment formulas.

## G.12 Altitude acclimatization

Venue altitude enters through Peronnet-style corrections (main text), but days at elevation before the race are not modeled [54, 12]. An athlete who lives or trains at altitude for weeks differs from one who drives to a mountain meet the morning of the race, even at the same starting line and barometric pressure. NRCD encodes meet elevation, not acclimatization history.

## G.13 Footwear and surface conditions

Shoe mass, cushioning, plate stiffness, and spike length alter economy and traction; roughly 1% higher metabolic cost per 100 g added mass slows time trials proportionally [28]. Compliant surfaces (mud, saturated grass, snow, ice) dissipate elastic energy and raise stabilizer demand [34, 56, 74, 37, 27]. Precipitation-specific effects are discussed in Section G.11. Meet results record neither footwear model nor spike choice nor post-race soil moisture.

## G.14 Gender-specific factors in women

NRCD standardizes times with gender-specific Riegel exponents and separate XC target distances (6 km women, 8 km men), but does *not* encode physiology that disproportionately affects women. For equity research using NRCD, within-gender modeling is essential; pooling or comparing raw seconds across genders without distance conversion remains inappropriate. The subsections below summarize latent factors documented primarily in women that can shift race-day performance or race availability independently of weather, grade, or altitude.

*G.14.1 Menstrual-cycle phase and endogenous hormones.* Fluctuations in oestradiol and progesterone across the menstrual cycle can alter substrate use, thermoregulation, and perceived exertion. A

2020 meta-analysis of eumenorrheic women found a *trivial* average reduction in endurance and strength performance in the early follicular phase relative to other phases (median pooled effect  $\approx -0.06$ ), with low study quality and large between-study heterogeneity [40]. The largest phase contrast (early vs. late follicular) remained small (effect  $\approx -0.14$ ). Population guidelines by cycle phase are therefore not supported; individual tracking is recommended when performance matters. NRCD stores no cycle phase, hormonal contraceptive status, or symptom diary fields, so cycle-aligned meet dates are an unobserved confounder for within-woman season trajectories.

**G.14.2 Menstrual symptoms and race availability.** Cycle-related symptoms can affect performance even when laboratory tests of “phase” show only trivial mean shifts. In a survey of 6,812 exercising women (Strava cohort), 84 to 91% reported cramps, fatigue, mood changes, or breast pain across the cycle; higher symptom burden predicted missing or altering training and missing competitions (odds ratios  $\approx 1.07$  to  $1.09$  per unit on a 54-point symptom index) [10]. For NRCD, this matters in two ways: (i) a slow time may reflect symptom-limited effort rather than fitness, and (ii) women who skip meets during symptomatic weeks are *censored from the dataset* (finisher-only export), biasing observed improvement rates and participation counts. DNS/DNF rows are absent, so symptom-driven non-starts are invisible.

**G.14.3 Iron deficiency and anemia.** Iron deficiency is more prevalent in female than male endurance athletes (often cited 15 to 35% vs. 3 to 11% in training populations) because of menstrual losses, lower dietary intake, foot-strike hemolysis, and exercise-induced inflammation [68, 11]. Even without anemia, low ferritin can reduce  $\dot{V}O_{2\max}$  and increase fatigue [11]. NRCD has no ferritin, hemoglobin, or supplementation fields; two women with identical standardized XC times may differ in oxygen-carrying capacity.

**G.14.4 Low energy availability and menstrual dysfunction (RED-S).** Relative energy deficiency in sport (RED-S) arises when energy intake does not cover exercise and physiological needs. It impairs metabolic rate, bone health, immunity, cardiovascular function, and menstrual function in women [43]. Functional hypothalamic amenorrhoea and oligomenorrhoea are common warning signs in female distance runners and can precede stress fractures and prolonged performance plateaus. RED-S is not captured by meet times; affected athletes may race poorly, reduce start frequency, or leave the sport entirely. Coaches and clinicians screen with energy-availability and menstrual-history tools that NRCD does not contain.

**G.14.5 Oral contraceptives and hormonal contraception.** Many collegiate women use combined oral contraceptives (OCs), which suppress ovulation and create “pseudo-phases” of synthetic estrogen and progestin distinct from the natural cycle [17]. A 2020 meta-analysis found OCP users may perform slightly worse on average than naturally menstruating women, but the group effect is trivial and highly variable; performance was stable across OCP consumption vs. withdrawal weeks [17]. NRCD does not record contraceptive type or timing, so researchers cannot stratify by hormonal profile.

**G.14.6 Implications for women’s XC in NRCD.** Women constitute 39.4% of unique athletes in the NRCD export (11,381 women vs. 17,532 men; 36.3% of result rows) and show larger raw-to-standardized variance compression than men (51.0% vs. 34.4% median cross-meet SD reduction), driven mainly by distance normalization rather than larger weather swings. Gender-stratified validation in the main resource reflects this. Analyses that treat standardized times as complete fitness proxies for women should explicitly acknowledge menstrual physiology, iron status, energy availability, and symptom-driven meet absence as unmeasured moderators [40, 10, 43].

## G.15 Latent factors in men’s physiology

Men constitute 60.6% of unique athletes in the export and 63.7% of result rows, yet meet times do not encode male-specific physiology either. Endogenous testosterone and training-response differences between men and women are summarized in recent consensus work [30, 5]. Iron deficiency without anemia is common in recreationally active men, not only women, and can limit aerobic capacity [68, 11]. Low energy availability and RED-S also occur in male distance runners (bone stress, hormonal suppression, performance decline) but are under-screened relative to women’s sport [43, 6]. Gender-stratified modeling remains essential; these gaps mean standardized men’s times still confound latent health and hormonal state with environment-adjusted performance.

## G.16 Psychology, pacing, and race tactics

Effort is regulated anticipatorily from expected duration, prior experience, and afferent feedback [73]; championship stakes, team-scoring position, and rivalry meets further shift risk tolerance and early surges. Running in a pack reduces effective wind resistance and aerobic demand compared with leading or racing alone [57, 26], distinct from sit-and-kick or surge tactics. Mass-start XC is tactically heterogeneous: World Championships analyses show positive (fade) pacing for most finishers, with medalists separating only late [25, 24]. Hill surges, lane congestion, and mid-race position change instantaneous speed in ways we cannot capture [26]. NRCD stores one aggregate time per result, not splits, pack position, or within-race tactical profile. Field depth, course familiarity, and home-context effects are selection issues (Section G.17).

## G.17 Selection and roster composition

Who appears in NRCD depends on club vs. varsity thresholds at a given university, athletes choosing to DNS, and self-selection [1, 31]. A fast time from a thin roster meet is not exchangeable with a depth-loaded invitational. Competitive field quality shifts early pace and risk taking even when an athlete’s fitness is unchanged [24, 26]. Athletes who train on or know a course, or compete near home with familiar travel and support, may gain an edge not captured by weather and grade metadata [15, 76]. Meet-level regionals and nationals flags (Section E) mark post-season championship context but do not encode race-day hype: crowd energy, broadcast pressure, and “big-meet” arousal can lift or disrupt performance beyond weather and grade adjustments, and often align with the nationals-heavy career PRs in Section B. Club runners sometimes sacrifice their own effort to pace a teammate toward a PR or fill a supporting team-scoring role; that choice is invisible in one finisher time and

can bias within-season improvement rates (Section A.3), separate from the within-race tactics in Section G.16. Spectators (family, friends, and home-club supporters) can further raise arousal and perceived support, interacting with academic stress (Section G.9) and away-meet travel (Section G.8). Standardization cannot correct for who chose to start, if they got excited for nationals, or whether they raced for a personal outcome vs. a teammate's PR.

### **G.18 Course geometry beyond aggregate grade**

Recorded gain/loss and measured length summarize certain course features, but not turn frequency, terrain, or choke points. Uneven micro-terrain raises energy cost by several percent even on modest height variation [74]; footing effects (Section G.13) can add further separation between courses with identical elevation metadata.

### **G.19 Officiating, timing, and wind on non-sprint track events**

Photo-finish latency, course-cutting penalties, and starter variability introduce measurement error unrelated to fitness. Wind readings exist for sprints and jumps where recorded, but middle-distance track races often lack per-lane wind metadata; non-sprint wind corrections are not applied in NRCD [35, 59].

**Researchers using NRCD for causal or predictive modeling should treat standardized times as environment-adjusted outcomes and explicitly model or bound the latent factors above.**