Supplementary material for the manuscript: "Is the Maturity Offset Equation Valid and Useful? A Simulation-Based and Longitudinal Evaluation with Implications for Youth Sport"

Humberto M. Carvalho, Federal University of Santa Catarina*

Luciano G. Galvão, Federal University of Santa Catarina

Fábio C. Karasiak, Federal University of Santa Catarina

Ahlan B. Lima, Federal University of Santa Catarina

Carlos E. Gonçalves, University of Coimbra

September 11, 2025

Bibliometric analysis

We conducted a bibliometric analysis using the Web of Science Core Collection, including all manuscripts (excluding reviews) citing the original or revised maturity offset equation (Mirwald et al., 2002; Moore et al., 2015). Bibliographic data were exported in `BibTeX` format and processed with the `bibliometrix` package in R (R Core Team, 2024). The search, last updated in August 2025, served as a proxy for the use of the maturity offset equation in youth research.

In total, 2,092 articles cited these references, with the majority published in journals classified under Sport Sciences. Some journals were assigned to multiple Web of Science research areas, resulting in 3,113 entries overall. Table S1 summarizes the top 10 research areas, showing that applied sport science predominates (1,151 entries; 37.0%), followed by contributions from Physiology, Pediatrics, Public, Environmental & Occupational Health, and Hospitality, Leisure, Sport & Tourism. A the-

*Corresponding author:Campus Trindade,88040-900, Florianópolis, Santa Catarina, Brasil

matic analysis of article titles and keywords indicates that, even in journals outside Sport Sciences, much of the research focuses on sport or exercise, highlighting the interdisciplinary application of the maturity offset equation. To provide a visual overview of thematic emphasis, a word cloud based on author keyword frequency was produced, where larger terms indicate higher frequency across the citing articles (Figure S1).

Table S1: Top 10 Web of Science research areas for articles citing the maturity offset equation (n = 3,113 entries, as some journals are identified in multiple research areas)

Rank	WOS Research Area	Articles (n)	% of Total (n = 3,113)
1	Sport Sciences	1151	37.0%
2	Physiology	300	9.6%
3	Pediatrics	233	7.5%
4	Public, Environmental & Occupational Health	73	2.3%
5	Nutrition & Dietetics	57	1.8%
6	Multidisciplinary Science	94	3.0%
7	Endocrinology & Metabolism	92	3.0%
8	Hospitality, Leisures, Sport & Tourism	83	2.7%
9	Environmental Sciences	78	2.5%
10	Biology	70	2.2%

A total of 348 journals cited the original or revised maturity offset equation, with 2,092 articles included in the analysis. The majority of citations were concentrated in sport science journals, led by the Journal of Strength and Conditioning Research (133 articles) and the Journal of Sports Sciences (113). Other frequently cited journals included Pediatric Exercise Science (91), International Journal of Environmental Research and Public Health (73), and PLOS ONE (57). The top 20 journals collectively accounted for a substantial proportion of all citations, reflecting the strong focus of maturity offset research within applied sport science, physiology, and related interdisciplinary fields.

Figure S1: Word cloud of author keywords from articles citing the maturity offset equation. The size of each term reflects its frequency across the 2,092 citing articles, with larger terms indicating more commonly used keywords.

Table S2: Top 20 journals citing the maturity offset equation (n = 2,092 articles; 348 journals in total)

Rank	Journal Title	Articles (n)
1	Journal of Strength and Conditioning Research	133
2	Journal of Sports Sciences	113
3	Pediatric Exercise Science	91
4	International Journal of Environmental Research and Public Health	73
5	PLOS ONE	57
6	Scandinavian Journal of Medicine & Science in Sports	51
7	European Journal of Applied Physiology	48
8	Frontiers in Physiology	48
9	Medicine and Science in Sports and Exercise	47
10	Sports	40
11	International Journal of Sports Medicine	32
12	European Journal of Sport Science	31
13	International Journal of Sports Physiology and Performance	30
14	American Journal of Human Biology	29
15	Frontiers in Sports and Active Living	29
16	Journal of Human Kinetics	29
17	Journal of Science and Medicine in Sport	28
18	Journal of Sports Medicine and Physical Fitness	27
19 - 21	Applied Physiology Nutrition and Metabolism; Applied Sciences – Basel; Children – Basel	26

Figure S2 presents a world map of manuscript frequencies citing the maturity offset equation by corresponding author's country. Table S2 provides absolute counts and relative percentages of the 2,092 articles. The United Kingdom emerged as the leading contributor (16.5%), followed by Brazil (11.0%) and Spain (10.8%), indicating strong outputs from Europe and South America. North America (Canada 7.3%; USA 6.8%) and Portugal (6.3%) also featured prominently. Collectively, these countries account for nearly 60% of all citations, underscoring a regional concentration of research leadership. At the same time, smaller but consistent contributions from countries such as Australia, Germany, Chile, and Tunisia illustrate the broader global use/citation of the maturity offset equation.

Table S3: Frequency of manuscripts citing the maturity offset equation by corresponding author's country (n = 2,092).

Rank	Country	Frequency	% of Total (n = 2,092)
1	UK	346	16.5%
2	Brazil	229	11.0%
3	Spain	225	10.8%
4	Canada	152	7.3%
5	USA	143	6.8%
6	Portugal	131	6.3%
7	Australia	90	4.3%
8	Germany	88	4.2%
9	Chile	67	3.2%
10	France	53	2.5%
11	Italy	50	2.4%
12	Belgium	46	2.2%
12	New Zealand	46	2.2%
14	Switzerland	40	1.9%
15	Tunisia	37	1.8%
16	Greece	35	1.7%
16	Qatar	35	1.7%
18	Poland	28	1.3%
19	Finland	22	1.1%
19	Japan	22	1.1%

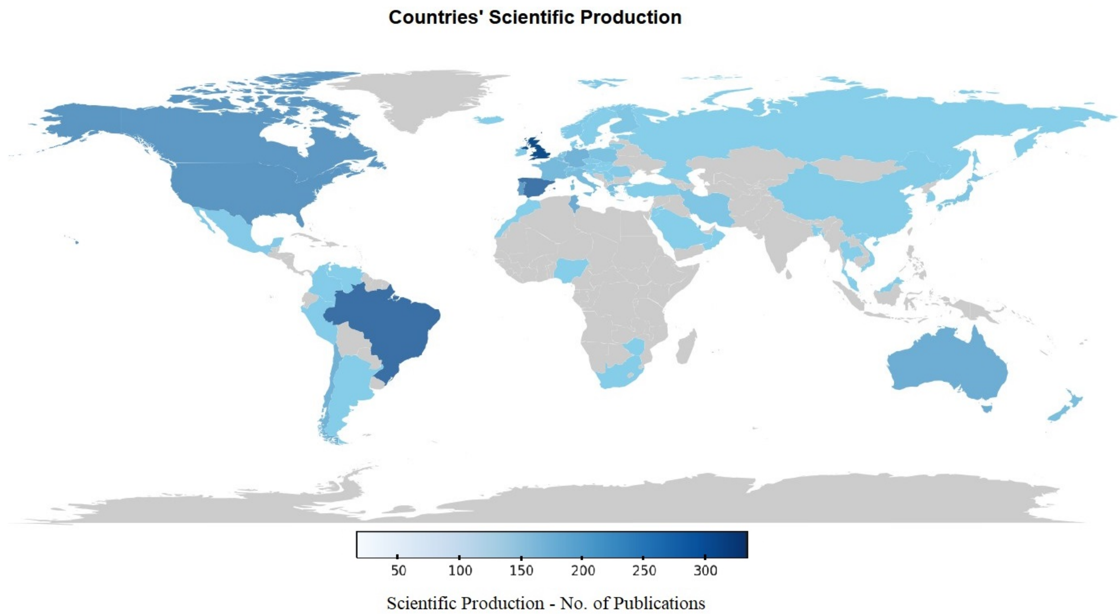

Figure S2: Worldwide geographic distribution of research citing the maturity offset equation.

Simulated Growth Data

We simulated longitudinal stature data to evaluate the accuracy of the maturity offset prediction equation under varying maturity conditions. The simulation was based on Model 2 from Preece & Baines (1978), which incorporates biologically plausible variation in growth timing and tempo by assigning individuals to early, on-time, and late maturing groups. This approach reflects typical maturity-related variability observed in youth populations and follows recommendations from previous simulation-based validation studies of age at peak height velocity (PHV) prediction methods (Cao et al., 2018; Sayers et al., 2013; Simpkin et al., 2017).

Simulation Design

The dataset consists of 500 individuals (250 males and 250 females), with stature measured at 0.5-year intervals between ages 7.0 and 18.0 years. Each individual was randomly assigned to one of three sex-specific maturity timing groups—early (16%), on-time (68%), or late (16%)—approximating a normal distribution of maturity timing.

Growth trajectories were generated using a modified Preece–Baines-like logistic growth model, parameterized by biologically informed values for initial stature, adult stature, growth tempo, and age at PHV. Stature at each age statures at each age was computed from a sigmoidal curve centered on the individual's age at PHV.

Simulation Model Specification

Individual stature at age t, denoted stature_{it}, was modeled as:

$$\text{stature}_{it} = stature_{i,7} + \frac{stature_{i,\text{adult}} - stature_{i,7}}{1 + \exp\left[-k_i \cdot (t - \text{APHV}_i)\right]}$$

Where:

- stature_{it} is the simulated stature (cm) for individual i at age t
- $stature_{i,7}$ is the baseline stature at age 7, sampled from a sex-specific normal distribution
- $stature_{i,\text{adult}}$ is adult stature, sampled from a sex-specific normal distribution
- APHV_i is the age at peak height velocity, sampled from a normal distribution conditional on sex and maturity group
- k_i is the growth tempo (steepness) parameter, sampled from a log-normal distribution allowing individual variability in growth rate

This simulation follows the framework of Model 2 from the Preece–Baines growth function family (Preece & Baines, 1978), which captures the sigmoidal growth curve with individualized timing and tempo parameters. The true age at peak height velocity (PHV) was determined according to the analytical approach described by Sayers et al. (2013).

Model Parameters

Simulated parameter values were drawn as follows:

Table S4: Simulation parameters used for growth data generation

Parameter	Females	Males
$stature_{i,7}$ (cm)	$\mathrm{N}(124, 3^2)$	$\mathrm{N}(126, 3^2)$
$stature_{i,\mathrm{adult}}$ (cm)	$\mathrm{N}(165, 6^2)$	$\mathrm{N}(176, 6^2)$
Growth rate k_i	$\log\text{-normal}(\log(1.0), 0.15^2)$	Same for both sexes
Age at PHV APHV_i (years).	Early: $\mathrm{N}(10.5, 0.4^2)$ On-time: $\mathrm{N}(11.8, 0.4^2)$ Late: $\mathrm{N}(13.0, 0.4^2)$	Early: $\mathrm{N}(12.5, 0.4^2)$ On-time: $\mathrm{N}(14.0, 0.4^2)$ Late: $\mathrm{N}(15.2, 0.4^2)$

Data Generation

Individuals were assigned sex and maturity groups with approximate proportions of 16% early, 68% on-time, and 16% late maturers per sex, reflecting empirical distributions in pediatric growth research (Malina et al., 2004).

Repeated stature measurements were simulated every 6 months between ages 7 and 18 years, resulting in dense longitudinal data for each individual. True age at PHV values were recorded and used as the reference standard for evaluating estimation accuracy.

The resulting dataset, which mirrors realistic adolescent growth patterns and inter-individual variability in maturation, was used to assess the performance of the maturity offset equation across sex and maturity timing groups.

The generated simulation data is displayed in Figure S3, and the distribution of age at PHV is shown in Figure S4.

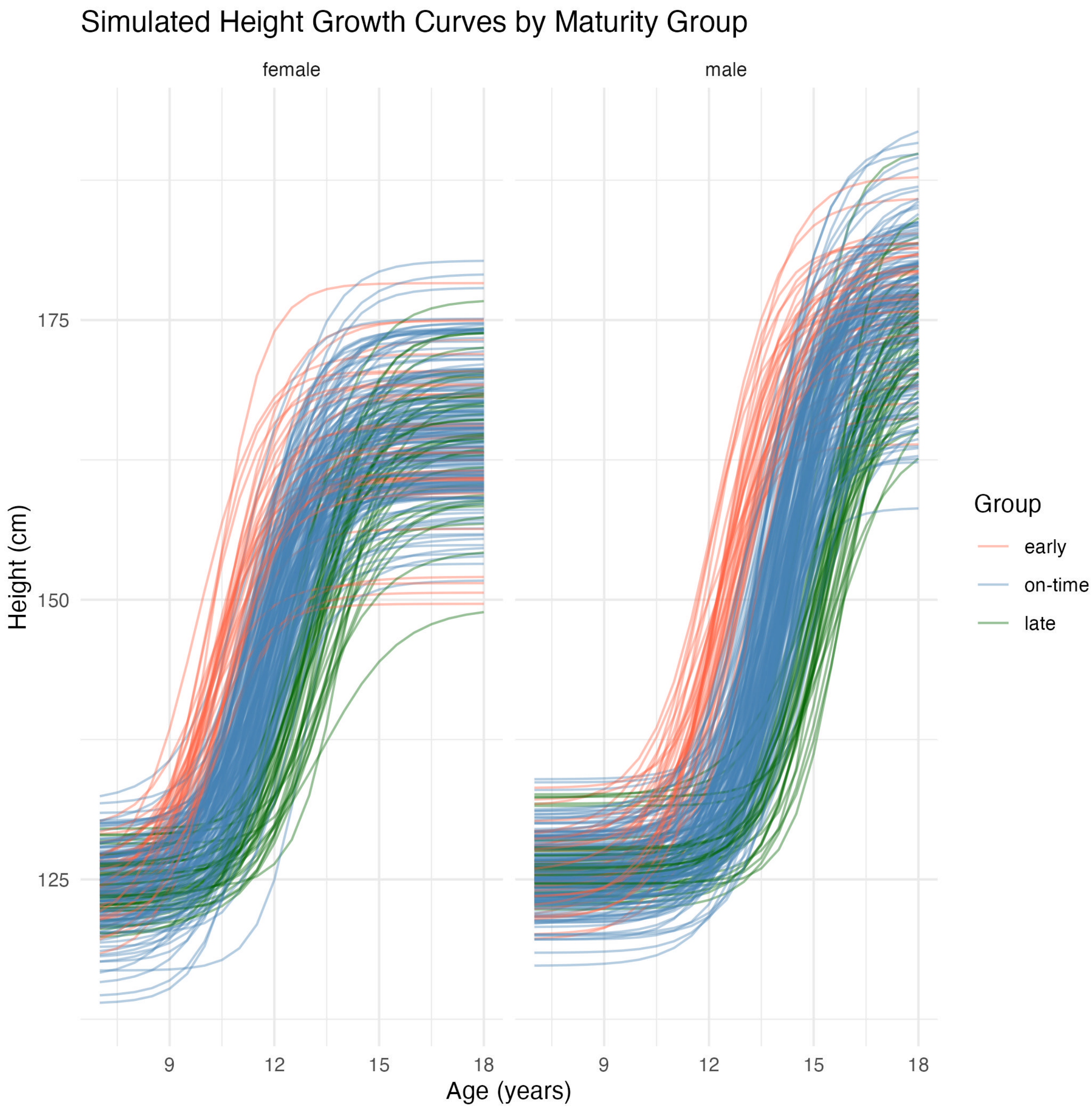

Figure S3: Simulated stature growth trajectories of individuals in the study dataset.

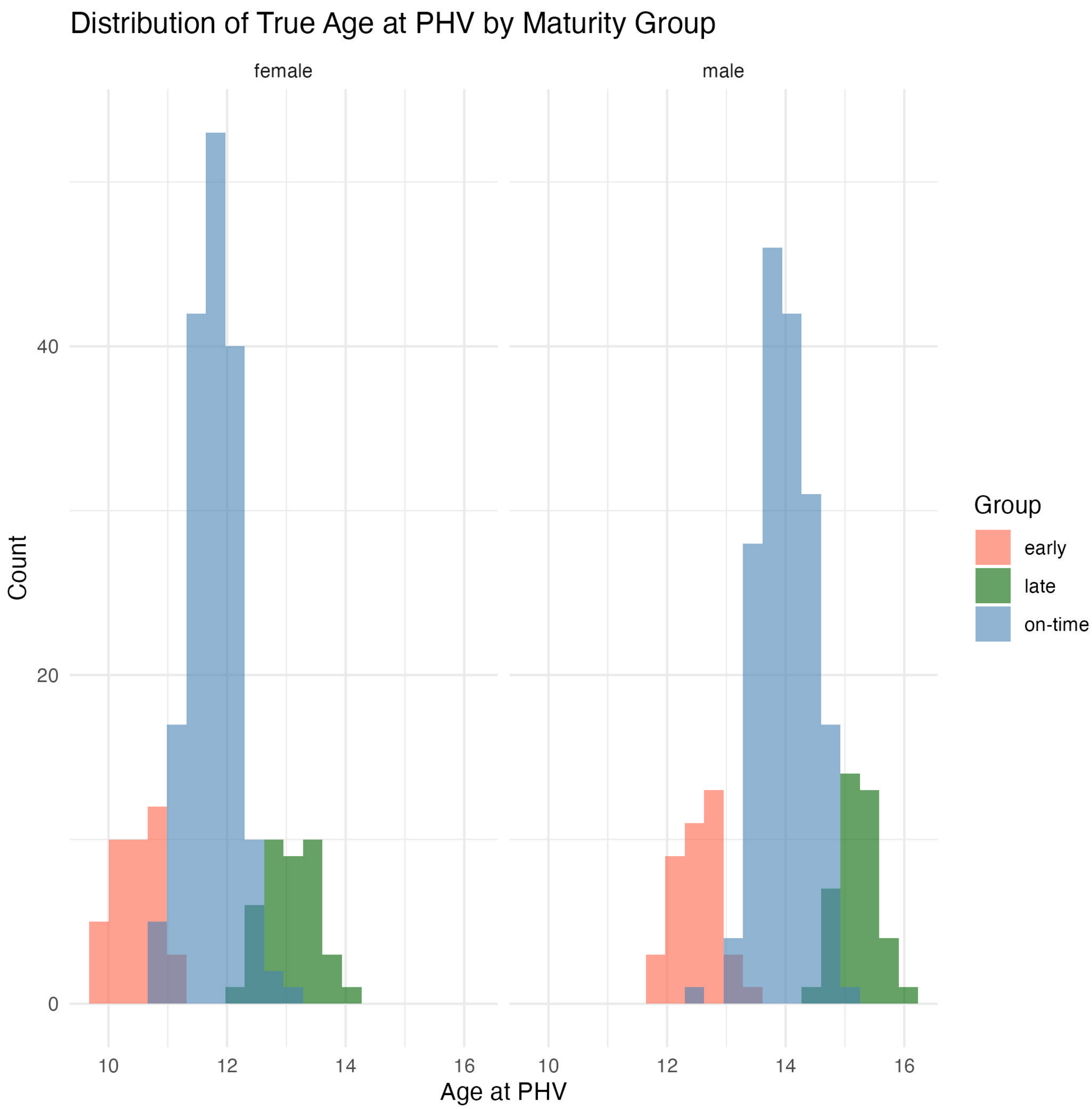

Figure S4: Distribution of true age at peak height velocity (PHV) by maturity status group in the simulated dataset.

Real-world data

We analyzed repeated measures collected annually between 1997 and 2010 from children aged 7 to 16 years at a public federal school in Florianópolis, Santa Catarina, Brazil (Galvão et al., 2025). The dataset included 3,568 observations (girls: $n = 1,925$; boys: $n = 1,643$) from 589 students (girls: $n = 290$; boys: $n = 297$). Measurements were collected by trained physical education teachers under standardized conditions, with retrospective access approved by the local ethics committee (protocol 80129124.0.0000.0121).

Model Error Metrics

To evaluate the accuracy of predicted values (e.g., predicted age at PHV), we used two commonly reported error metrics: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).

The MAE measures the average magnitude of the errors between predicted and true PHV, disregarding their direction, given as:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i|$$

Where, $\hat{y}_i$ is the predicted PHV value for observation i, y_i is the true PHV value for observation i, n is the total number of observations, and $|\cdot|$ denotes the absolute value.

The RMSE also quantifies prediction error but gives more weight to larger errors by squaring the residuals before averaging:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}$$

MAE provides a straightforward indication of average error magnitude, while RMSE reflects both the variance and magnitude of prediction errors. RMSE is more sensitive to outliers than MAE, which can be particularly relevant when evaluating the consistency of model predictions across a diverse sample. Both MAE and RMSE are reported in the same units as the outcome (years). Lower values of both metrics indicate better predictive accuracy.

Classification performance was assessed using confusion matrices comparing predicted versus true or observed maturity status (early, on-time, late) and developmental status (>2 yr pre-PHV, 1–2 yr pre-PHV, circa-PHV ±1 yr, 1–2 yr post-PHV, >2 yr post-PHV), categorized based on age at PHV relative to sex-specific distributions as described earlier. Confusion matrices were used to estimate overall accuracy, as well as class-specific metrics—precision and recall (sensitivity)—to assess the model's ability to correctly identify each group.

Precision, defined as the proportion of individuals predicted in a given group who truly belong to that group, is calculated as:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

Sensitivity (recall), defined as the proportion of true group members correctly identified, is calculated as:

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

These metrics complement overall accuracy by highlighting the model's performance across groups, especially since misclassification can have different practical implications in youth sports. By including both maturity status and developmental status, we provide a more detailed assessment of prediction accuracy across the full range of growth and maturation.

Confusion matrices summarizing the classification of participants into maturity status and developmental status groups based on predicted versus true age at PHV are presented in Tables S2–S5. From a practical perspective, misclassification assigns a child or adolescent athlete to the wrong maturity or developmental group, which may have implications for training, monitoring, or talent identification.

For the maturity status classification (early, on-time, late), most on-time individuals were correctly classified in both girls and boys, while early and late maturers were often misclassified as on-time. In girls (Table S2), 2,818 on-time individuals were correctly identified, whereas 570 early and 552 late individuals were assigned to the on-time category. In boys (Table S3), 2,782 on-time individuals were correctly classified, but a substantial number of early and late maturers were systematically misclassified into the on-time group. These confusion matrices (Tables S2–S3) highlight that misclassification primarily affects individuals at the extremes of maturity, whereas average maturers are classified with high accuracy.

Table S5: Confusion matrix showing the classification of girls into maturity status groups based on predicted versus true age at PHV.

True / Predicted	Early	On-time	Late
Early	216	570	134
On-time	573	2818	519
Late	121	552	247

Table S6: Confusion matrix showing the classification of boys into maturity status groups based on predicted versus true age at PHV.

True / Predicted	Early	On-time	Late
Early	147	744	29
On-time	612	2782	511
Late	148	400	372

For the developmental status classification (>2 yr pre-PHV, 1–2 yr pre-PHV, circa-PHV $\pm$1 yr, 1–2 yr post-PHV, >2 yr post-PHV), misclassification was concentrated in categories near PHV. In girls (Table S4), individuals >2 years pre-PHV (n = 1475) and >2 years post-PHV (n = 1949) were classified with accuracy, whereas intermediate groups (1–2 years pre-PHV, circa-PHV, 1–2 years post-PHV) showed substantial misclassification across adjacent stages. In boys (Table S5), extreme categories (>2 years pre-PHV, n = 2567; >2 years post-PHV, n = 956) were accurately classified, but individuals around PHV were often misclassified (e.g., circa-PHV, n = 665 correctly classified but 334 misclassified).

Table S7: Confusion matrix showing the classification of girls into developmental status groups based on predicted versus true age at PHV.

True/ Predicted	>2yr pre-PHV	1–2yr pre-PHV	circa-PHV (±1yr)	1–2yr post-PHV	>2yr post-PHV
>2yr pre-PHV	1475	35	0	0	0
1–2yr pre-PHV	277	208	15	0	0
circa-PHV (±1yr)	38	279	636	46	1
1–2yr post-PHV	0	0	204	226	70
>2yr post-PHV	0	0	32	259	1949

Table S8: Confusion matrix showing the classification of boys into developmental status groups based on predicted versus true age at PHV.

True/ Predicted	>2yr pre-PHV	1–2yr pre-PHV	circa-PHV (±1yr)	1–2yr post-PHV	>2yr post-PHV
>2yr pre-PHV	2567	37	0	0	0
1–2yr pre-PHV	251	240	9	0	0
circa-PHV (±1yr)	23	248	665	63	1
1–2yr post-PHV	0	0	127	274	99
>2yr post-PHV	0	0	17	173	956

References

Cao, Zhiqiang, Hui, L. L., & Wong, M. Y. (2018). New approaches to obtaining individual peak height velocity and age at peak height velocity from the SITAR model. Computer Methods and Programs in Biomedicine, 163, 79–85. https://doi.org/https://doi.org/10.1016/j.cmpb.2018.05.030

Galvão, Luciano G., Karasiak, Fábio C., Conceição, Victor J. S., Silva, Diego Augusto Santos, & Carvalho, Humberto M. (2025). Modeling pubertal growth variability in schoolchildren from florianópolis, santa catarina, brazil: A bayesian analysis. American Journal of Human Biology, 37. https://doi.org/10.1002/ajhb.70129

Malina, R. M., Bouchard, C., & Bar-Or, O. (2004). Growth, maturation, and physical activity (2nd ed.). Human Kinetics.

Mirwald, R. L., Baxter-Jones, A. D., Bailey, D. A., & Beunen, G. P. (2002). An assessment of maturity from anthropometric measurements. Med Sci Sports Exerc, 34, 689–694. http://www.ncbi.nlm.nih.gov/pubmed/11932580

Moore, S. A., McKay, H. A., Macdonald, H., Nettlefold, L., Baxter-Jones, A. D., Cameron, N., & Brasher, P. M. (2015). Enhancing a somatic maturity prediction model. Med Sci Sports Exerc, 47, 1755–1764. https://doi.org/10.1249/MSS.0000000000000588

Preece, M. A., & Baines, M. J. (1978). A new family of mathematical models describing the human growth curve. Ann Hum Biol, 5, 1–24. https://doi.org/10.1080/03014467800002601

R Core Team. (2024). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Sayers, Adrian, Baines, Mike, & Tilling, Kate. (2013). A new family of mathematical models describing the human growth curve—erratum: Direct calculation of peak height velocity, age at take-off and associated quantities. Annals of Human Biology, 40, 298–299. https://doi.org/10.3109/03014460.2013.772655

Simpkin, Andrew J., Sayers, Adrian, Gilthorpe, Mark S., Heron, Jon, & Tilling, Kate. (2017). Modelling height in adolescence: A comparison of methods for estimating the age at peak height velocity. Ann Hum Biol, 44, 715–722. https://doi.org/10.1080/03014460.2017.1391877