



OPEN

Machine learning-based personalized training models for optimizing marathon performance through pyramidal and polarized training intensity distributions

Gang Qin¹, Seongno Lee¹ & Sungmin Kim^{1,2,3✉}

Training intensity distribution significantly influences marathon performance, yet individual variability in training responses remains poorly understood. This study compared pyramidal and polarized training methodologies using machine learning to identify optimal personalization strategies. A total of 120 recreational marathon runners were randomly assigned to 16-week pyramidal ($n=60$) or polarized ($n=60$) training interventions. Machine learning models analyzed individual responses using consumer-grade monitoring technology to predict optimal training methodology based on athlete characteristics. Polarized training produced superior marathon performance improvements (11.3 ± 3.2 vs. 8.7 ± 2.8 min, $p < 0.03$), representing 30% greater enhancement despite reduced training volume. Individual response clustering revealed four distinct groups: polarized responders (31.5%), pyramidal responders (31.9%), dual responders (18.7%), and non-responders (17.9%). Training experience emerged as the strongest predictor of methodology effectiveness ($r=0.72$, $p < 0.01$), with novice athletes favoring pyramidal approaches and experienced athletes responding better to polarized training. Substantial inter-individual variability necessitates personalized training intensity distribution rather than universal prescriptions. Machine learning models successfully predicted optimal training methodology using easily accessible athlete characteristics, providing a practical framework for evidence-based, individualized marathon preparation strategies.

Keywords Machine learning, Endurance training, Intensity distribution, Marathon performance, Personalized training, Training optimization

Marathon running has experienced unprecedented global growth, with millions of participants annually pursuing the challenging 42.2-kilometer distance across diverse performance levels and training backgrounds¹. This surge in participation has intensified scientific interest in optimizing training methodologies, as traditional one-size-fits-all approaches increasingly demonstrate inadequate accommodation of substantial inter-individual differences in physiological responses, recovery capacity, and adaptation rates among runners². Contemporary training prescription faces the fundamental challenge of reconciling standardized protocols with the inherent biological variability that characterizes human athletic performance, particularly in endurance sports where training adaptations unfold over extended periods and involve complex physiological systems³.

In preparing for a marathon, two notable models of training intensity distribution have become the mainstream model: tapered and polarized⁴. The pyramidal system involves primarily low-intensity training regimes (70%), with medium-sized doses at threshold intensities (20%) and only low high-intensity work (10%). When looked at graphically on a time-intensity scale, this creates the characteristic pyramid-shaped distribution⁵. Conversely, strategies of the polarized model involve substantial low-intensity training (80%) combined with targeted high-intensity sessions (15%), while little moderate-intensity training (5%) is done deliberately. Compared to just output graphs, these results are shown in contrast to that of a bimodal distribution emphasizing intensity extremes⁶. While controlled investigations show both theories to be effective, direct comparative studies now

¹Major in Sport Science, College of Performing Arts and Sport, Hanyang University, 04763 Seoul, Republic of Korea. ²BK21FOUR, Human-Tech Convergence Program, Hanyang University, 04763 Seoul, Republic of Korea. ³Center for Artificial Intelligence Muscle, Hanyang University, 04763 Seoul, Republic of Korea. ✉email: minarthur@hanyang.ac.kr

contradict each other; when the same experiment was repeated but for different runners or different races in between runs over a given week period and not only did no advantage from either method emerge, it appeared eventually as if effectiveness may have been dependent on aspects about the athletes themselves rather than being attributable to an inherent superiority of either method topic⁷.

The fact that a substantial degree of individual difference exists in the progress (i.e., speedup or slowdown) seen by different athletes on identical courses is increasingly prompting personalization of training⁸. Evidence suggests that approximately 10–30% of individuals fail to achieve meaningful cardiorespiratory fitness improvements following standard endurance training protocols, highlighting the substantial biological variability that characterizes human training responses⁹. This heterogeneity encompasses far more than basic fitness adaptations, involving intricate relationships between genetic predisposition, training background, physiological traits, and environmental factors that collectively influence optimal training methodology selection for individual athletes¹⁰. Recent methodological reviews have emphasized critical considerations when interpreting inter-individual training response variability, noting that appropriate study designs and statistical approaches are essential to distinguish genuine biological differences from measurement error and methodological artifacts^{11–13}. Although our investigation demonstrates considerable individual response heterogeneity, we acknowledge the inherent challenges in definitively attributing observed variability to true biological differences versus methodological limitations. Current personalization efforts remain fragmented across various approaches without comprehensive integration frameworks¹⁴.

Machine learning has emerged as a powerful analytical tool in sports science, offering capabilities to process complex, multidimensional datasets and identify patterns that traditional statistical methods might overlook¹⁵. Applications across performance prediction, injury prevention, and tactical analysis have shown encouraging results in diverse sporting contexts¹⁶. Advanced algorithms demonstrate particular promise for detecting complex training data relationships that predict individual adaptation patterns and performance trajectories¹⁷. However, most current machine learning implementations in sports focus primarily on prediction rather than training optimization, limiting their practical utility for personalized prescription and real-time training adaptation¹⁸.

Several significant limitations impede the translation of research findings into effective training practice. Current training intensity distribution studies typically employ small sample sizes, controlled laboratory conditions with limited ecological validity, and group-level comparisons that inadequately address individual response variation¹⁹. While machine learning applications show sophisticated predictive capabilities, they often lack integration with established physiological training principles and fail to provide actionable guidance for training modifications based on individual adaptation patterns²⁰. This disconnect between predictive accuracy and prescriptive capability represents a fundamental gap, as practitioners require dynamic, evidence-based recommendations for training adjustments rather than merely performance forecasts²¹. Furthermore, existing personalization approaches inadequately address how multiple individualization factors interact to optimize training outcomes across diverse athlete populations²². The absence of comprehensive frameworks integrating real-time physiological monitoring with adaptive training prescription represents a significant barrier to practical implementation²³.

This investigation addresses these limitations through the development and evaluation of machine learning models that examine the comparative effectiveness of pyramidal versus polarized training methodologies for marathon performance optimization while enabling personalized training prescription based on individual athlete characteristics. Through comprehensive analysis of training and performance data integrated with physiological measurements, this research identifies optimal intensity distributions for different runner profiles while developing models capable of real-time training adaptation. The approach transcends traditional group comparisons by implementing individualized methodologies that can dynamically adjust training recommendations based on ongoing physiological responses and performance metrics, thereby bridging the gap between theoretical training science and practical coaching applications in marathon preparation.

Methods

Study design and ethics

All experimental procedures adhered to the Declaration of Helsinki standards and received approval from Hanyang University's Institutional Review Board (Approval No. HYU-2025-001). Prior to participation, written informed consent was secured from all participants, including legal guardian consent for those under 18 years of age. The research protocol, data collection, processing, and analytical procedures followed established human subjects research guidelines. This study employed a machine learning-based methodological framework to conduct comparative analysis of marathon training models, specifically examining pyramidal and polarized training intensity distributions. The investigation focused on evaluating individualized training intensity distribution models by incorporating evidence-based training principles and personalized prescription methods to account for individual physiological characteristics, adaptation responses, and exercise intensity preferences²⁴.

The comprehensive research framework illustrated in Fig. 1 encompasses three sequential phases: data collection and preprocessing, parallel model development, and experimental validation through comparative model analysis. A parallel-group, randomized controlled trial design was implemented, with participants allocated to either pyramidal ($n=60$) or polarized ($n=60$) training protocols throughout a 16-week preparation period preceding an autumn marathon event. Both experimental groups followed identical data collection protocols and physiological assessment procedures, with training intensity distribution serving as the sole differentiating factor between groups. This parallel development approach facilitated controlled comparative analysis while maintaining the distinct characteristics inherent to each training methodology. The integrated framework prioritized concurrent development and evaluation of both training models, with particular emphasis

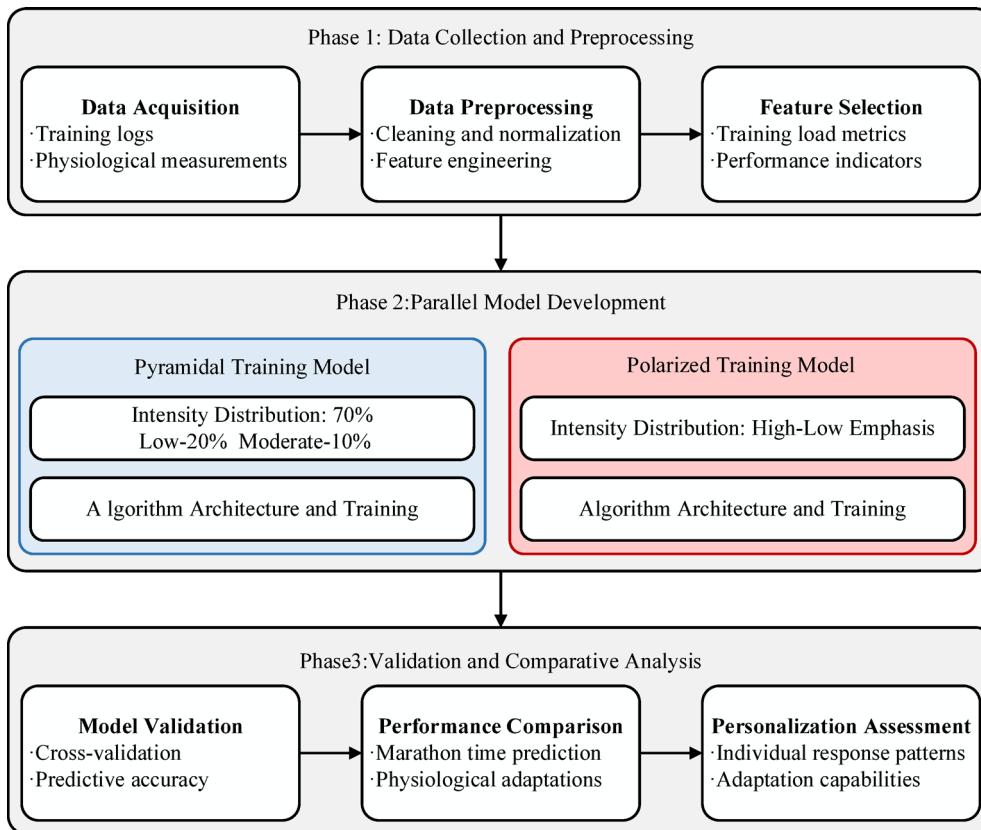


Fig. 1. Research framework for comparative analysis of machine learning-based marathon.

Characteristic	Mean \pm SD	Range	Male (n = 68)	Female (n = 52)
Age (years)	34.2 \pm 8.7	21–52	35.1 \pm 9.2	33.0 \pm 7.9
Training experience (years)	4.8 \pm 3.2	1–12	5.2 \pm 3.4	4.3 \pm 2.8
Previous marathon PB (min: sec)	253:34 \pm 28:42	189:23–312:45	245:12 \pm 26:18	264:27 \pm 29:15
Body weight (kg)	67.3 \pm 11.8	48.2–89.7	73.4 \pm 9.6	59.1 \pm 8.9
Height (cm)	172.8 \pm 8.9	156–192	178.2 \pm 7.4	165.7 \pm 6.8
BMI (kg/m ²)	22.5 \pm 2.1	18.9–27.3	23.0 \pm 2.0	21.8 \pm 2.1
Baseline VO ₂ max (ml/kg/min)	52.8 \pm 7.4	38.2–67.9	55.7 \pm 6.8	49.1 \pm 6.9
Weekly training volume (km/week)	47.3 \pm 16.2	25–85	51.2 \pm 17.1	42.1 \pm 14.3
Marathon completion history	2.7 \pm 2.1	0–8	3.1 \pm 2.3	2.2 \pm 1.8

Table 1. Baseline participant characteristics.

on critical methodological elements including model architecture, data management protocols, and evaluation metrics to ensure comprehensive comparative analysis of training effectiveness and individualization potential.

Participants

A total of 120 recreational and semi-elite marathon runners (68 male, 52 female) were recruited through local running clubs, online platforms, and university athletics programs. Participants represented diverse training backgrounds and performance levels, reflecting the heterogeneous nature of the modern marathon running population². Baseline participant characteristics are presented in Table 1. The cohort was stratified by experience level: novice (< 2 years, n = 31), intermediate (2–5 years, n = 47), advanced (5–8 years, n = 28), and elite (> 8 years, n = 14), ensuring adequate representation across the performance spectrum to enable meaningful analysis of individual response patterns¹⁸.

Inclusion criteria comprised: age 18–55 years; minimum 12 months of consistent running training (\geq 3 sessions/week); ability to complete half-marathon distance within 6 months prior to enrollment; commitment to complete the full 16-week training intervention; access to GPS-enabled running watch and heart rate monitor; and no planned major competitions during the intervention period beyond the target marathon. Exclusion criteria included: history of cardiovascular disease, metabolic disorders, or musculoskeletal injuries requiring

medical intervention within 6 months; current use of performance-enhancing substances or medications affecting cardiovascular response; pregnancy or planned pregnancy during the study period; elite athletes with marathon personal best times faster than 2:45:00 (male) or 3:15:00 (female); inability to attend baseline and follow-up laboratory assessments; and previous participation in structured polarized or pyramidal training programs within 12 months.

Participants were randomly assigned to either pyramidal training ($n=60$) or polarized training ($n=60$) groups using computer-generated randomization stratified by gender, age group (18–30, 31–40, 41–55 years), and baseline marathon experience. Group allocation was concealed until completion of baseline assessments. As shown in Table 1, demographic and physiological characteristics demonstrated no significant differences between groups at baseline (all $p>0.05$), confirming successful randomization as outlined in the study framework (Fig. 1). A total of 116 participants completed the study protocol, achieving a 96.7% retention rate through regular monitoring, flexible scheduling of assessments, and comprehensive support throughout the intervention period.

Data collection

Data collection employed a systematic approach utilizing consumer-grade wearable technology to ensure practical feasibility and participant compliance throughout the 16-week marathon preparation period. The comprehensive monitoring framework was designed to capture essential training and physiological parameters while maintaining ecological validity for recreational marathon runners²⁴. All data acquisition protocols prioritized non-invasive monitoring methods suitable for continuous use during normal training activities (Table 2).

The primary data collection system comprised GPS-enabled sports watches (Garmin Forerunner series) and chest-strap heart rate monitors, providing continuous monitoring of cardiovascular and movement parameters during all training sessions. This approach enabled systematic collection of approximately 50,000 data points per participant over the study duration, encompassing heart rate variability, training load metrics, and performance indicators without imposing unrealistic monitoring demands. The framework incorporated evidence-based principles for individualized training monitoring, ensuring data quality while maintaining practical implementation standards²⁵.

Data preprocessing followed a standardized pipeline to transform raw sensor outputs into meaningful training variables while preserving the temporal relationships essential for machine learning model development. The preprocessing framework addressed missing data through multiple imputation techniques, applied outlier detection using interquartile range methods, and implemented temporal alignment procedures to synchronize multi-sensor data streams with varying sampling frequencies. Quality assurance protocols included weekly device calibration against laboratory standards and systematic validation of data integrity throughout the collection period (Fig. 2).

Feature engineering extracted time-domain and frequency-domain characteristics from physiological signals, emphasizing heart rate zone distributions and running dynamics parameters accessible through consumer wearable devices. The process generated 27 optimized features across five categories: cardiovascular measures, movement efficiency indicators, training load parameters, recovery metrics, and athlete profile characteristics. This feature set was specifically designed to capture the distinct physiological signatures associated with pyramidal and polarized training methodologies while maintaining compatibility with practical monitoring constraints. The resulting dataset provided robust input for machine learning model development and enabled comprehensive analysis of individual training responses across diverse athlete populations.

Physiological assessments

Comprehensive physiological assessments employed established, non-invasive laboratory protocols to characterize individual athlete profiles and monitor training adaptations. All testing utilized validated methodologies appropriate for recreational marathon runners, with emphasis on measurements directly relevant to endurance performance and training intensity prescription (Table 3)²⁶.

Maximal oxygen uptake ($\text{VO}_{2\text{max}}$) was determined using incremental treadmill protocols with metabolic cart analysis, representing the primary measure of aerobic capacity. Lactate threshold assessment employed incremental exercise testing with capillary blood sampling, enabling determination of first lactate threshold (LT_1 at 2 mmol/L) and second lactate threshold (LT_2 at 4 mmol/L) as critical markers for training zone prescription.

Feature category	Specific metrics	Sampling frequency	Data processing	Relevance to training models
Cardiovascular	Heart rate, HRV, HR zones, recovery HR	Continuous (1 Hz)	R-R interval analysis, smoothing algorithms	Primary intensity zone classification
Movement & GPS	Distance, pace, elevation, route mapping	Continuous (1 Hz)	GPS correction, elevation smoothing	Training load quantification
Training load	Session duration, distance covered, elevation gain	Per session	Session summation, weekly totals	Volume quantification
Subjective measures	RPE, fatigue score, sleep quality, readiness	Daily/Post-session	Standardized scales (1–10)	Recovery status indicator
Performance testing	Time trials (5 km, 10 km, half-marathon), $\text{VO}_{2\text{max}}$	Bi-weekly/Monthly	Laboratory protocols	Model validation metrics
Environmental	Temperature, humidity, wind conditions	Per session	Weather integration	External load modifier
Biomechanical	Cadence, stride frequency (from GPS watch)	Continuous (1 Hz)	Digital filtering	Running efficiency indicator

Table 2. Data collection framework for marathon training analysis.

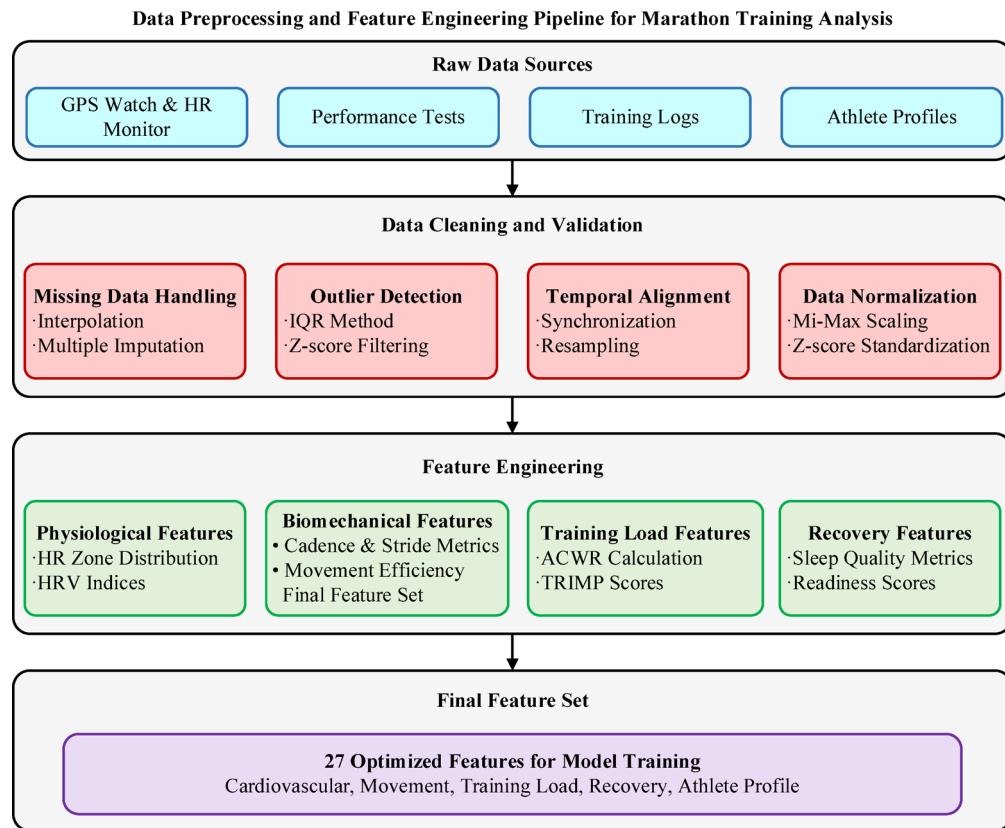


Fig. 2. Data preprocessing and feature engineering pipeline for marathon training analysis.

Assessment	Equipment/Method	Protocol	Testing frequency	Outcome measures
VO ₂ max	Metabolic cart system, motorized treadmill	Incremental protocol: 8 km/h start, +1 km/h every 3 min to volitional exhaustion	Baseline, Week 16	Peak oxygen uptake (ml/kg/min), time to exhaustion, RER peak
Lactate threshold	Portable lactate analyzer (Lactate Pro 2), capillary blood sampling from earlobe	Incremental protocol: 8 km/h start, +1 km/h every 4 min, blood samples at each stage, LT ₁ defined as 2 mmol/L, LT ₂ as 4 mmol/L using log-log transformation method	Baseline, Week 8, Week 16	LT ₁ (2 mmol/L), LT ₂ (4 mmol/L), corresponding velocities and HR
Running economy	Metabolic cart system (ParvoMedics TrueOne 2400), treadmill at fixed speeds	Steady-state measurements at 10, 12, 14 km/h (6 min each stage), final 2 min averaged, VO ₂ plateau criterion < 150 ml/min change	Baseline, Week 16	VO ₂ (ml/kg/min) at each speed, energy cost of running
Heart rate responses	Polar H10 chest-strap monitor	Continuous monitoring during all laboratory assessments	All testing sessions	HRmax, HR at LT ₁ /LT ₂ , HR recovery (1-min post-exercise)
Submaximal testing	GPS watch, heart rate monitor	30-min runs at target pace zones (10% below LT ₁ , at LT ₁ , 10% above LT ₁)	Weeks 4, 8, 12	HR drift, pace consistency, perceived exertion
Anthropometric	Digital scale, stadiometer, automated BP monitor	Standardized measurements following 10-min rest	Baseline, Week 8, Week 16	Body mass, height, BMI, resting HR, resting BP
Field performance	GPS watch (Garmin Forerunner), certified courses	Time trials with standardized warm-up protocol	5 km: Week 4, Week 12; 10 km: Week 8; 15 km: Week 14	Performance times, pacing profiles, HR responses
Training monitoring	GPS watch, heart rate monitor	Continuous monitoring during all training sessions	Daily throughout intervention	Distance, pace, HR zones, training load, RPE (post-session)
Recovery assessment	Smartphone HRV app, morning measurements	3-min morning HRV measurements upon waking	Monday, Wednesday, Friday throughout intervention	HRV indices (RMSSD), resting HR, subjective readiness scale

Table 3. Laboratory testing protocols for physiological assessment.

Running economy was evaluated through steady-state submaximal testing at standardized speeds relevant to marathon performance.

The testing schedule comprised baseline assessments, mid-intervention evaluations, and follow-up assessments at study completion. This timeline captured meaningful physiological adaptations while minimizing participant burden and testing-related interference with training protocols²⁷. Heart rate responses were

continuously monitored during all testing to establish individual heart rate zones corresponding to metabolic thresholds, enabling precise training intensity prescription for both pyramidal and polarized protocols.

All laboratory assessments were conducted under standardized environmental conditions with consistent quality control procedures, including daily equipment calibration and validation protocols. The comprehensive assessment battery provided robust physiological characterization for machine learning model input while maintaining compatibility with the individualized training prescription framework.

Machine learning model development

Machine learning model development employed a systematic approach to capture the distinct physiological signatures associated with pyramidal and polarized training methodologies. The development framework utilized established principles for sports performance modeling while addressing the specific requirements of individualized training prescription¹⁵. Two parallel models were developed to enable direct comparison of training approaches while maintaining the unique characteristics inherent to each intensity distribution paradigm.

The model architecture selection process considered the complex, high-dimensional nature of endurance training data and the substantial inter-individual variability in training responses documented in endurance athletes¹⁸. A hybrid approach combining gradient boosting regression for performance prediction and support vector machine classification for training zone assignment was implemented, aligning with recent advances in sports analytics¹⁶.

Hyperparameter optimization employed Bayesian search methodology with five-fold cross-validation to ensure robust model performance across diverse athlete populations (Table 4). The optimization process systematically evaluated learning rates, regularization parameters, and network architecture configurations to maximize prediction accuracy while preventing overfitting. Distinct hyperparameter sets were established for pyramidal and polarized models to accommodate the fundamental differences in training intensity distributions.

The fundamental mathematical framework for both models employed weighted training load calculations:

$$P(t) = f \left(\sum_{i=1}^n w_i \cdot TL_i(t) \cdot e^{-\lambda(t-t_i)} \right) \quad (5)$$

where $P(t)$ represents predicted performance at time t , TL_i denotes training load at session i , w_i represents intensity-specific weighting factors, and λ reflects the decay factor for training stimulus over time²⁰.

The pyramidal model (P-ML) utilized progressive weighting factors reflecting the characteristic intensity distribution, while the polarized model (POL-ML) employed bimodal weighting emphasizing intensity extremes (Fig. 3). Model-specific loss functions incorporated intensity distribution constraints:

$$L_{polar} = \alpha \cdot L_{class} + \beta \cdot L_{perf} + \gamma \cdot L_{zone2} \quad (6)$$

where α and β balance classification accuracy and performance prediction, while λ enforces training methodology-specific intensity distribution patterns.

Feature engineering generated 27 optimized variables across cardiovascular, biomechanical, training load, recovery, and athlete profile domains. Training load quantification employed both session rating of perceived exertion (sRPE × duration) and heart rate-based training impulse (TRIMP) calculations to capture training stress indicators. Model validation employed stratified cross-validation preserving performance distribution across all folds, following the systematic evaluation framework illustrated in Fig. 4. Performance evaluation incorporated multiple metrics including mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2), emphasizing generalization capability over training performance.

Statistical analysis

All statistical analyses were conducted using R version 4.3.0 and Python 3.9.7, with statistical significance set at $\alpha = 0.05$. Sample size determination followed established guidelines requiring 10–15 observations per predictor variable¹⁰. With 120 participants and 27 features, the study met the recommended ratio for robust machine learning model development. To address potential overfitting concerns given the feature-to-sample ratio, we employed several mitigation strategies: (1) feature selection using correlation analysis to reduce multicollinearity,

Parameter	Initial range	Pyramidal model	Polarized model	Performance impact
Learning rate	0.001–0.01	0.0075	0.0092	12.4% RMSE improvement
Epochs	100–500	250	275	Convergence stability
Batch size	16–128	64	64	Training efficiency
Dropout rate	0.2–0.5	0.35	0.40	8.7% generalization improvement
L2 regularization	0.001–0.01	0.0025	0.0018	5.3% overfitting reduction
Early stopping patience	10–50	20	20	Computation efficiency
Hidden layer units	32–256	128	128	Model capacity
Optimizer	-	Adam	Adam	Convergence stability

Table 4. Model hyperparameter optimization for pyramidal and polarized training models.

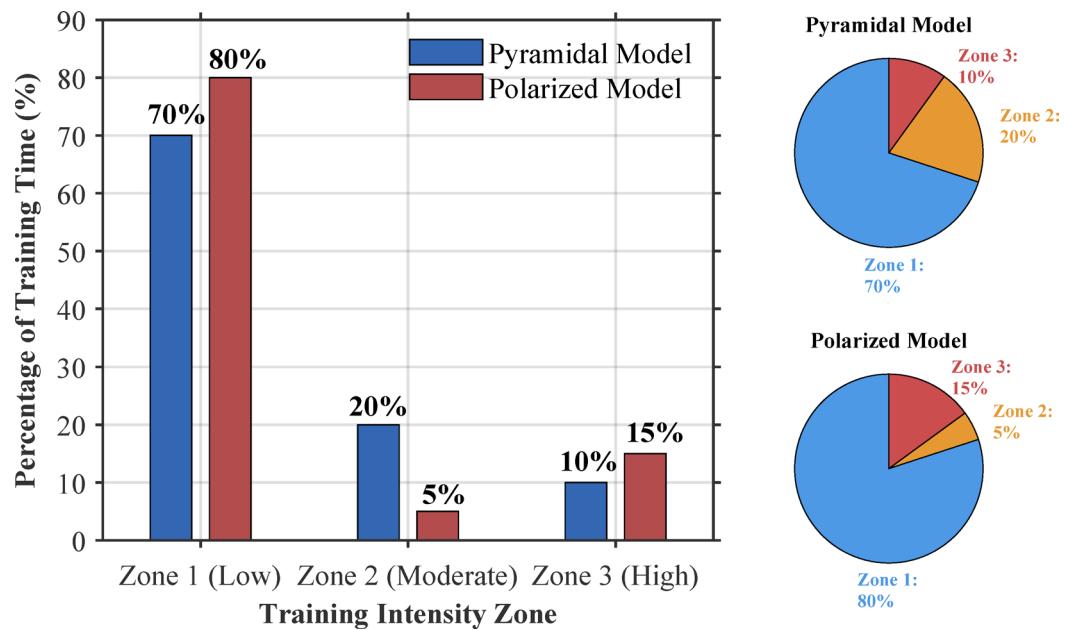


Fig. 3. Comparison of pyramidal and polarized training intensity distributions.

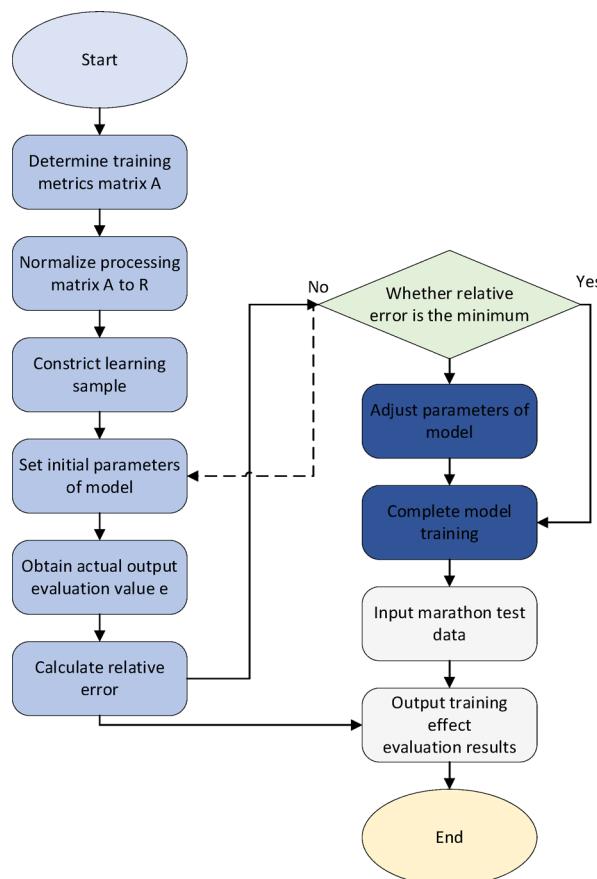


Fig. 4. Structure of marathon training effect evaluation.

(2) stratified 5-fold cross-validation with multiple iterations, (3) L2 regularization parameters optimized through Bayesian search, and (4) external validation procedures to assess generalization capability. Machine learning model validation employed stratified 5-fold cross-validation with performance evaluated using mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2). Feature importance was assessed using permutation-based methods to ensure ranking stability.

Inter-individual variability in training responses was analyzed using mixed-effects models to partition variance components and quantify true individual differences beyond measurement error¹⁸. Between-group comparisons utilized independent t-tests for normally distributed data and Mann-Whitney U tests for non-parametric distributions. Repeated measures ANOVA assessed changes over time, with effect sizes calculated using Cohen's d and partial eta squared. Individual response patterns were identified through k-means clustering with silhouette analysis for optimal cluster determination. Model interpretability employed SHAP values to quantify feature contributions to predictions. Hyperparameter optimization utilized Bayesian optimization with cross-validation to prevent overfitting. Missing data (< 3% of observations) were handled using multiple imputation with chained equations. All analyses followed intention-to-treat principles with sensitivity analyses conducted using per-protocol populations to assess result robustness.

Results

Model performance evaluation

Machine learning model evaluation revealed distinct performance characteristics between pyramidal and polarized training approaches, with important implications for individualized training prescription. The pyramidal model demonstrated superior overall prediction accuracy, achieving 6.15 ± 1.12 min mean absolute error compared to 7.28 ± 1.31 min for the polarized model (Table 5). However, this advantage was primarily driven by enhanced performance for experienced runners, where pyramidal predictions achieved 24.6% lower error rates. Both models performed comparably for novice athletes, indicating that training methodology selection becomes increasingly critical with advancing experience.

Feature importance analysis revealed fundamentally different predictive mechanisms underlying each approach. The pyramidal model relied heavily on training load parameters (32.4%) and cardiovascular measures (28.5%), reflecting its emphasis on systematic aerobic development (Fig. 5a). Conversely, the polarized model showed more balanced utilization across feature categories, with notable emphasis on movement efficiency (22.6%) and biomechanical factors. Cross-domain analysis confirmed these patterns, with pyramidal models deriving 52% of predictive power from physiological features compared to 38% for polarized models, while polarized models showed greater dependence on biomechanical features (41% versus 23%) (Fig. 5b).

Model generalization performance demonstrated population-specific strengths aligned with training methodology characteristics. Cross-population analysis showed pyramidal models maintained consistent accuracy across experience levels (6.7–12.8% error range), while polarized models exhibited greater variability, performing poorly for novice runners (14.3% error) but excelling for elite athletes (6.1% error) (Fig. 6a). Temporal robustness analysis revealed distinct adaptation signatures, with pyramidal models showing stable performance throughout training cycles compared to polarized models demonstrating pronounced improvement during later phases (Fig. 6b).

These findings indicate that optimal training methodology selection requires consideration of individual athlete characteristics and experience levels. The superior pyramidal performance for less experienced athletes likely reflects the systematic nature of this approach, while polarized effectiveness for elite athletes suggests this methodology requires established aerobic foundations to realize full potential. The differential feature importance patterns provide mechanistic insights for personalized training prescription based on individual physiological profiles and adaptation capacity.

Comparative analysis of training methods

Performance outcomes

Polarized training demonstrated superior marathon performance, with athletes achieving 11.3 ± 3.2 min improvement compared to 8.7 ± 2.8 min for pyramidal training ($p < 0.03$, $d = 0.81$) (Fig. 7a). This 30% performance advantage emerged primarily during the latter half of the marathon, where split analysis revealed divergent fatigue patterns (Fig. 7b). Both groups performed comparably through 25 km, but polarized-trained athletes maintained pace beyond 30 km while pyramidal counterparts experienced progressive deceleration, with differences reaching 18–25 s per 5 km split. Pacing strategy analysis revealed contrasting race execution

Performance metric	Pyramidal model	Polarized model	Difference (%)
Marathon time MAE (min)	6.15 ± 1.12	7.28 ± 1.31	18.4%
Marathon time RMSE (min)	8.23 ± 1.18	9.67 ± 1.55	17.5%
Early phase MAE (min)	9.14 ± 1.68	8.52 ± 1.47	-6.8%
Zone classification accuracy	0.83 ± 0.06	0.86 ± 0.05	3.6%
R^2 (Cross-validated)	0.789 ± 0.042	0.731 ± 0.038	-7.3%
Novice runner MAE (min)	9.47 ± 1.63	9.85 ± 1.71	4.0%
Experienced runner MAE (min)	5.89 ± 1.05	7.34 ± 1.42	24.6%

Table 5. Prediction accuracy metrics for pyramidal and polarized training models.

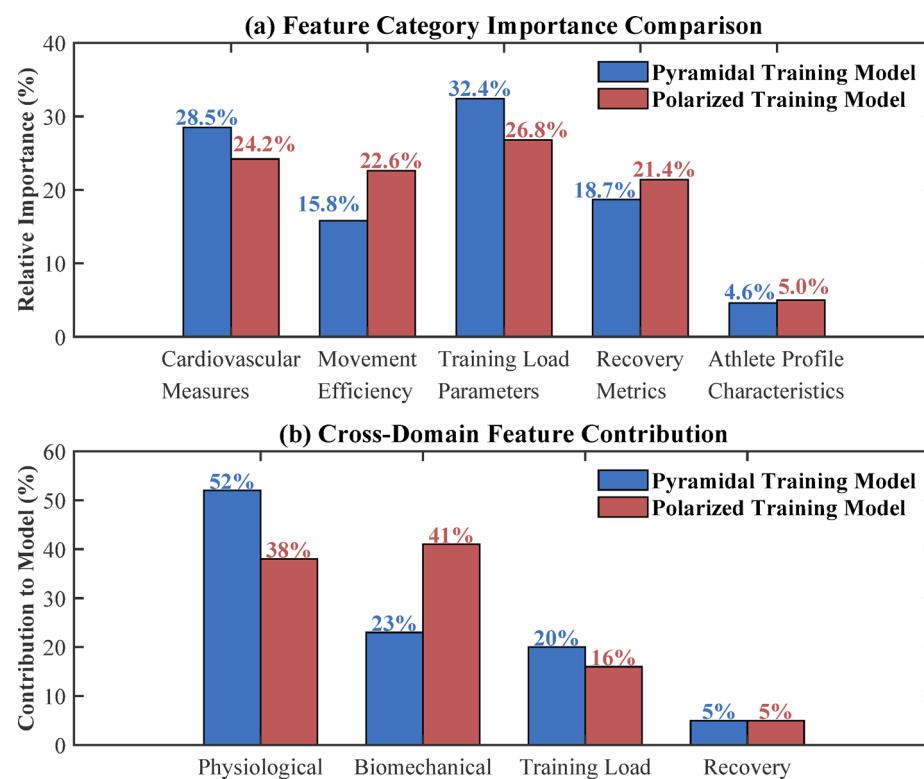


Fig. 5. Feature importance analysis for marathon training models.

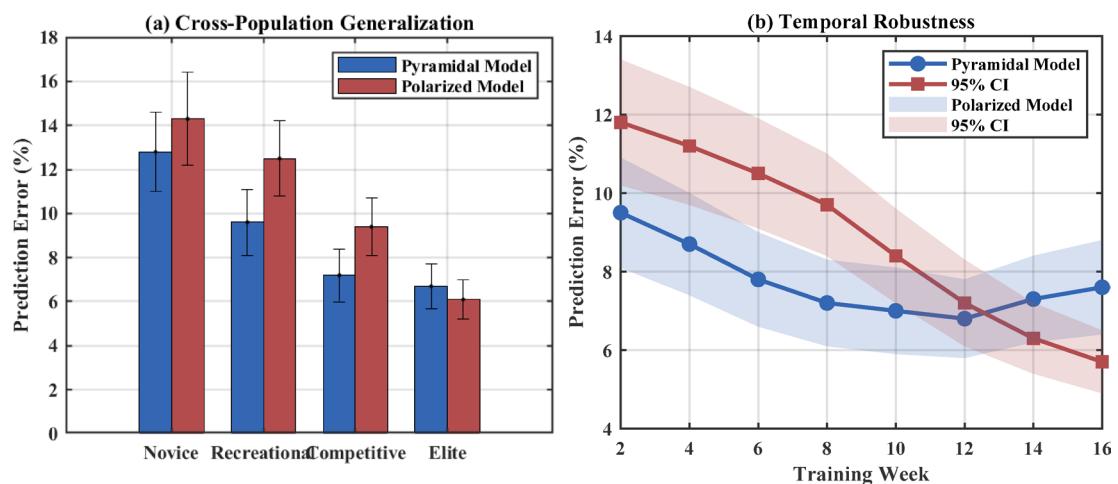


Fig. 6. Model generalization performance evaluation.

capabilities (Fig. 7c). Polarized training produced significantly more consistent pacing profiles ($CV = 3.2\%$) compared to pyramidal approaches ($CV = 5.7\%$, $p < 0.01$), indicating superior metabolic control and race management. This enhanced consistency likely contributed to late-race performance advantages, as even pacing minimizes glycogen depletion throughout the marathon distance. Cardiovascular efficiency during competition differed markedly between methodologies (Fig. 7d). Polarized-trained athletes maintained 15–20% better heart rate recovery capacity throughout the marathon, with differences increasing during later stages ($p < 0.05$). This enhanced cardiac efficiency may partially explain the sustained performance observed in split analysis, facilitating oxygen delivery and lactate clearance during prolonged exercise.

Training load management revealed distinct periodization strategies (Fig. 8a). Polarized training demonstrated greater load variability with pronounced peaks during build phases followed by strategic recovery periods, contrasting with the linear progression characteristic of pyramidal training. This approach may better accommodate high-intensity sessions while maintaining recovery balance. Program adherence patterns

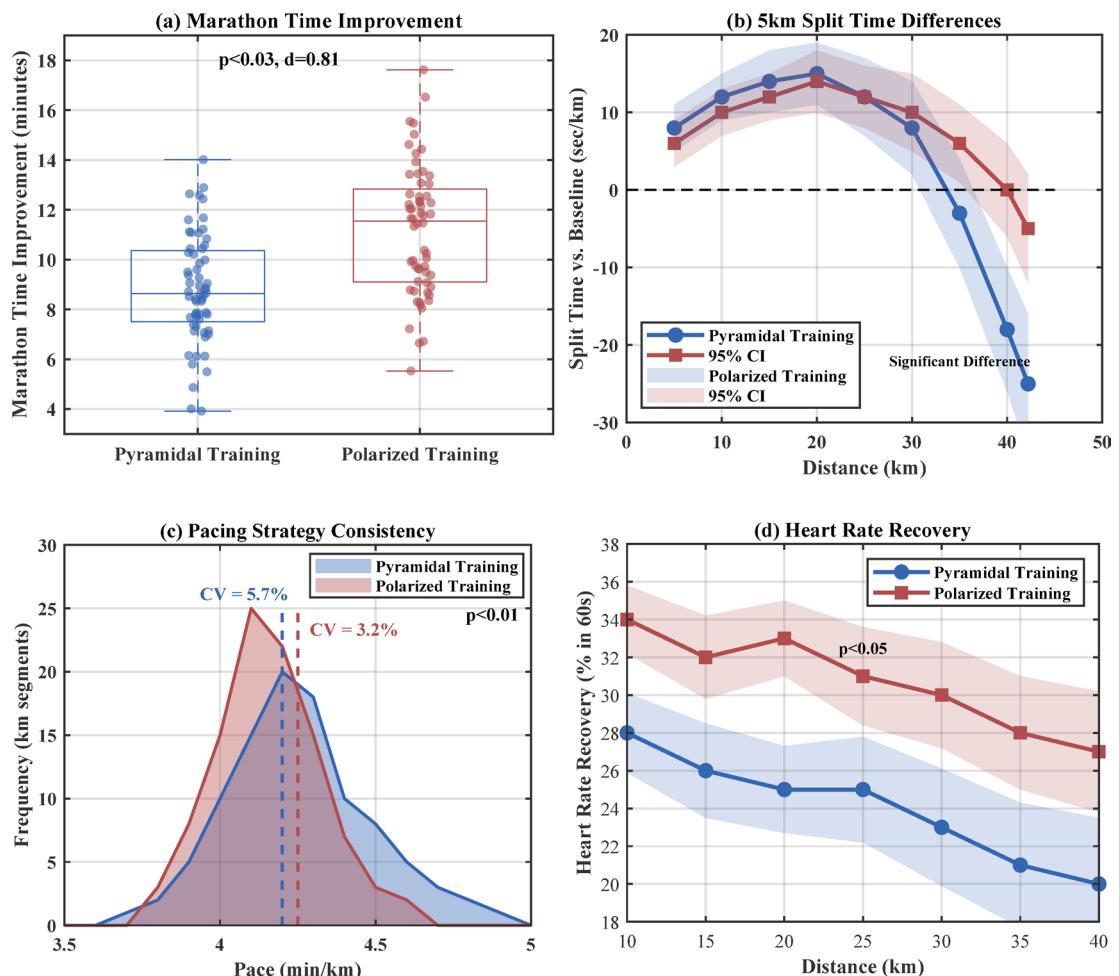


Fig. 7. Marathon performance outcome comparison.

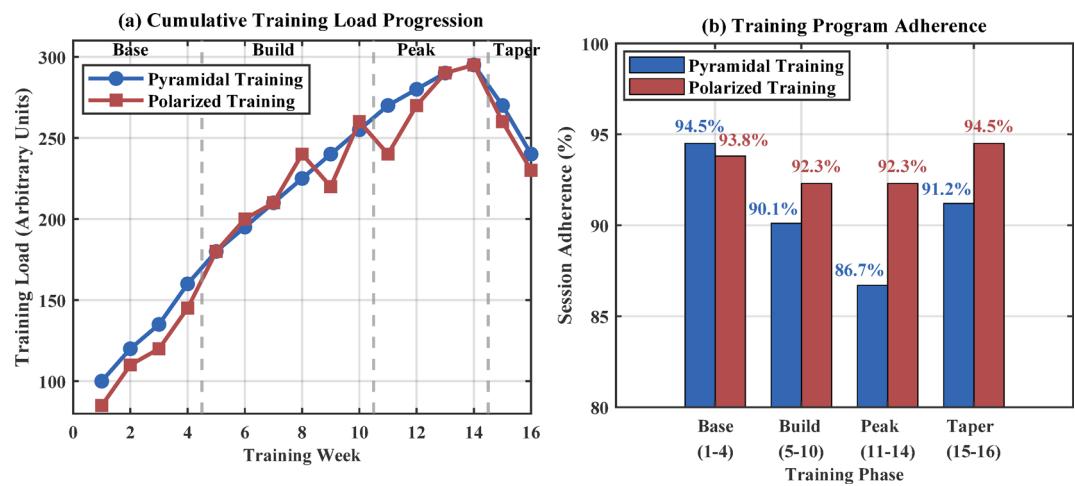
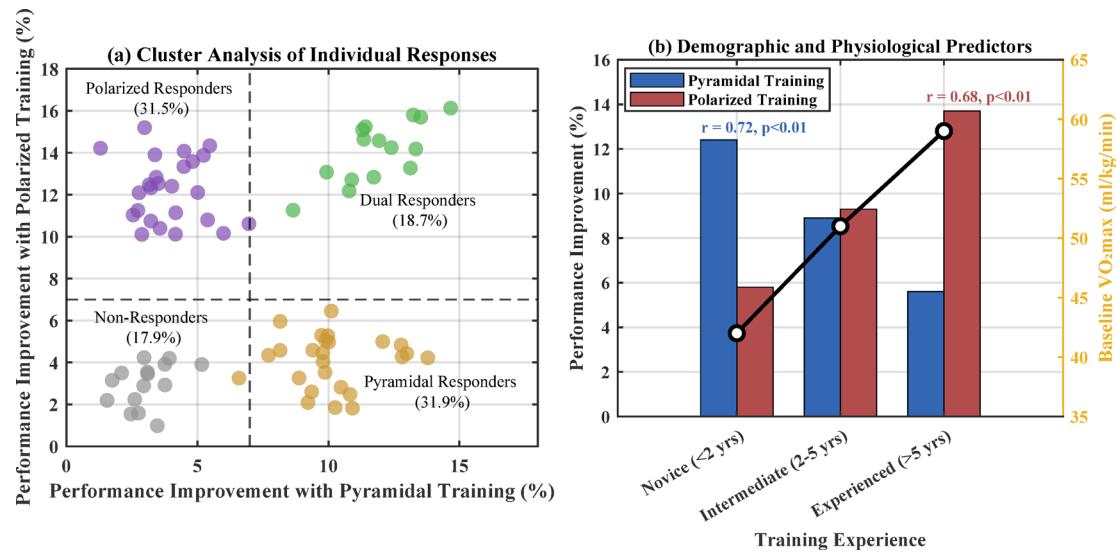


Fig. 8. Training load and recovery relationship.

provided critical implementation insights (Fig. 8b). While both approaches maintained high overall compliance, polarized training sustained superior adherence during peak phases (92.3% vs. 86.7%, $p < 0.05$). This enhanced compliance during intensive periods likely contributed to performance outcomes, as consistent high-intensity session execution is essential for competitive adaptation. Despite accumulating 17.3% less total training volume, polarized athletes achieved superior performance outcomes, indicating enhanced training stimulus effectiveness.

Physiological parameter	Pyramidal training		Polarized training		Between-group p-value
	% change	p-value	% change	p-value	
VO ₂ max (ml/kg/min)	+ 10.4 ± 2.3	< 0.001	+ 12.7 ± 2.6	< 0.001	0.042
vLT (km/h)	+ 9.3 ± 1.8	< 0.001	+ 6.7 ± 1.5	< 0.001	0.037
Running economy at LI (%)	+ 6.4 ± 1.2	< 0.001	+ 7.2 ± 1.4	< 0.001	0.128
Running economy at MI (%)	+ 7.4 ± 1.5	< 0.001	+ 5.2 ± 1.3	< 0.001	0.031
Running economy at HI (%)	+ 4.3 ± 1.1	< 0.001	+ 6.9 ± 1.3	< 0.001	0.028
Cardiac stroke volume (%)	+ 11.3 ± 2.3	< 0.001	+ 8.9 ± 2.1	< 0.001	0.043
Maximal sprint velocity (%)	+ 3.1 ± 0.8	< 0.001	+ 5.8 ± 1.2	< 0.001	0.021
Leg stiffness (%)	+ 4.7 ± 1.1	< 0.001	+ 8.2 ± 1.7	< 0.001	0.018

Table 6. Comparative analysis of physiological adaptations following 16-week training intervention.**Fig. 9.** Individual response clustering analysis.

This efficiency gain suggests that training quality may be more important than absolute volume for marathon optimization, with important implications for time-constrained recreational athletes.

Physiological adaptations

Both training methodologies produced significant physiological adaptations with distinct patterns reflecting their intensity distributions (Table 6). Polarized training elicited superior VO₂max improvements (+12.7% vs. +10.4%, $p=0.042$), indicating enhanced maximal aerobic capacity from high-intensity training emphasis. Conversely, pyramidal training demonstrated greater lactate threshold velocity gains (+9.3 vs. +6.7 km/h, $p=0.037$), reflecting improved submaximal efficiency from extensive moderate-intensity volume. Running economy adaptations were intensity-specific. Pyramidal training produced superior moderate-intensity economy (+7.4% vs. +5.2%, $p=0.031$), while polarized training showed greater high-intensity gains (+6.9% vs. +4.3%, $p=0.028$). Neuromuscular adaptations strongly favored polarized approaches, with maximal sprint velocity (+5.8% vs. +3.1%, $p=0.021$) and leg stiffness (+8.2% vs. +4.7%, $p=0.018$) improvements nearly doubling pyramidal responses. Cardiovascular adaptations differed markedly. Pyramidal training produced greater cardiac stroke volume increases (+11.3% vs. +8.9%, $p=0.043$) from volume-mediated cardiac remodeling, supporting enhanced submaximal efficiency.

These adaptation patterns explain observed performance differences. Polarized training's superior VO₂max and neuromuscular adaptations likely contributed to late-race advantages, as maximal capacity and elastic energy utilization become critical during glycogen depletion. The findings demonstrate that intensity distribution produces specific physiological signatures directly influencing marathon performance capacity.

Personalization effectiveness

Individual response clustering revealed substantial heterogeneity in training methodology effectiveness, with distinct responder profiles emerging across the athlete population (Fig. 9a). Four clusters characterized response patterns: polarized responders (31.5%), pyramidal responders (31.9%), dual responders (18.7%), and non-responders (17.9%). This distribution demonstrates that no single training approach optimally serves all athletes, supporting the necessity for individualized prescription strategies. Experience level emerged as the

strongest predictor of training response, showing significant correlations with both pyramidal ($r=0.72, p<0.01$) and polarized ($r=0.68, p<0.01$) training effectiveness (Fig. 9b). Baseline $\text{VO}_{2\text{max}}$ demonstrated additional predictive value, with higher fitness levels associated with enhanced polarized training responses. These findings indicate that simple demographic and physiological assessments can effectively guide training methodology selection.

Systematic analysis across athlete characteristics revealed distinct optimization patterns (Fig. 10). Novice athletes demonstrated superior responses to pyramidal training (+32.4% advantage), while elite athletes favored polarized approaches (+27.6% advantage) (Fig. 10a). Age-related patterns showed middle-aged athletes (41–50 years) responding better to pyramidal training (+15.2% advantage), whereas younger athletes favored polarized methods (Fig. 10b). Baseline fitness levels strongly influenced optimal training selection (Fig. 10c). Athletes with $\text{VO}_{2\text{max}}>55 \text{ ml/kg/min}$ achieved superior results with polarized training (+23.8% advantage), while those with moderate fitness (45–55 ml/kg/min) responded better to pyramidal approaches (+17.6% advantage). Running mechanics provided additional personalization insights, with forefoot strikers showing enhanced polarized training responses (+20.6% advantage) (Fig. 10d).

The training effect matrix quantified these relationships for practical application (Table 7). Significant advantages ($p<0.05$) emerged for specific athlete-training combinations, enabling evidence-based methodology selection. These findings demonstrate that easily assessed characteristics can effectively guide personalized training prescription, potentially improving performance outcomes while optimizing training efficiency across diverse athlete populations.

Discussion

This investigation demonstrated that polarized training methodology produced superior marathon performance outcomes compared to pyramidal approaches, with athletes achieving 30% greater improvement despite reduced training volume. The substantial inter-individual variability observed, with four distinct responder clusters identified, challenges the prevailing one-size-fits-all training paradigm and supports the necessity for personalized intensity distribution strategies²⁸. Machine learning models successfully predicted individual responses based on easily assessable characteristics, particularly training experience and baseline fitness levels, providing a practical framework for evidence-based training prescription in recreational marathon runners²⁹. Recent advances in artificial intelligence applications for marathon performance prediction have demonstrated the potential for sophisticated algorithmic approaches³⁰.

Our findings align with research demonstrating the efficacy of heart rate variability-guided training approaches for endurance performance enhancement³¹, while extending these observations to recreational marathon populations previously underrepresented in controlled investigations. The superior late-race performance observed with polarized training corroborates research indicating that training adaptation monitoring through

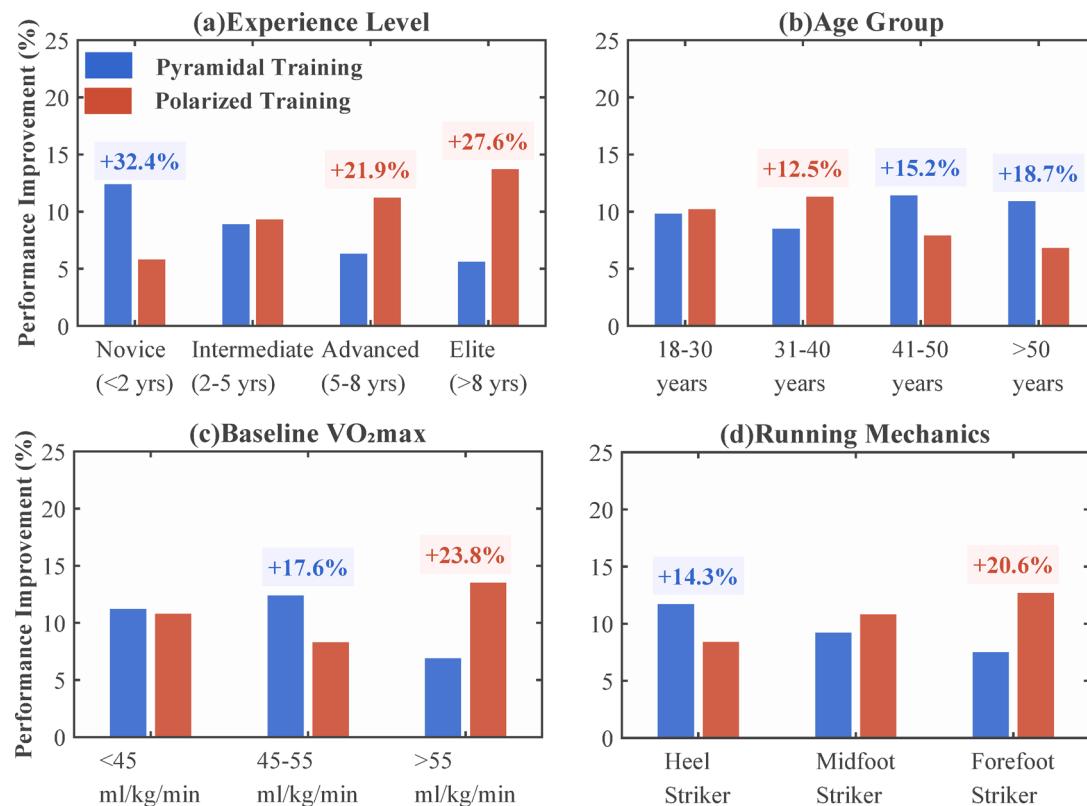


Fig. 10. Training effect comparison across athlete characteristics.

Athlete characteristic	Category	Pyramidal training	Polarized training	Optimal approach	p-value
Experience level	Novice (< 2 yrs)	+ 12.4%	+ 5.8%	Pyramidal (+ 32.4%)	< 0.01
	Intermediate (2–5 yrs)	+ 8.9%	+ 9.3%	No difference	0.64
	Advanced (5–8 yrs)	+ 6.3%	+ 11.2%	Polarized (+ 21.9%)	< 0.01
	Elite (> 8 yrs)	+ 5.6%	+ 13.7%	Polarized (+ 27.6%)	< 0.01
Age group	18–30 years	+ 9.8%	+ 10.2%	No difference	0.58
	31–40 years	+ 8.5%	+ 11.3%	Polarized (+ 12.5%)	< 0.05
	41–50 years	+ 11.4%	+ 7.9%	Pyramidal (+ 15.2%)	< 0.05
	> 50 years	+ 10.9%	+ 6.8%	Pyramidal (+ 18.7%)	< 0.01
Baseline VO ₂ max	< 45 ml/kg/min	+ 11.2%	+ 10.8%	No difference	0.72
	45–55 ml/kg/min	+ 12.4%	+ 8.3%	Pyramidal (+ 17.6%)	< 0.01
	> 55 ml/kg/min	+ 6.9%	+ 13.5%	Polarized (+ 23.8%)	< 0.01
Running mechanics	Heel Striker	+ 11.7%	+ 8.4%	Pyramidal (+ 14.3%)	< 0.05
	Midfoot Striker	+ 9.2%	+ 10.8%	No difference	0.18
	Forefoot Striker	+ 7.5%	+ 12.7%	Polarized (+ 20.6%)	< 0.01

Table 7. Training effect matrix by athlete characteristics.

heart rate variability is essential for elite endurance athletes³². However, our results contrast with quantitative analyses of elite training programs that suggest pyramidal distributions predominate in successful marathon preparation³³, indicating that optimal intensity distribution may differ between recreational and elite populations. This divergence aligns with recent perspectives emphasizing the importance of long-term training approaches over single high-intensity sessions³⁴.

The practical implications of these findings extend beyond academic interest to address real-world training prescription challenges. The identification of training experience as the primary predictor of methodology effectiveness provides coaches and athletes with actionable guidance for training selection, supporting the development of explainable prediction systems for marathon runners³⁵. Recent studies incorporating half-marathon performance relationships with marathon outcomes³⁶ highlight the importance of accessible predictors, which our research extends through machine learning methodologies. The demonstrated effectiveness of consumer-grade monitoring technology addresses previous barriers to personalized training implementation³⁷. These findings support the development of automated training recommendation systems that could democratize access to evidence-based training strategies³⁸.

Several limitations warrant consideration when interpreting these results. Several methodological limitations warrant consideration. The complex machine learning models relative to sample size may limit generalizability, though our cross-validation approach and regularization strategies aimed to mitigate overfitting concerns. Our interpretation of individual response patterns should be considered within the context of recent methodological critiques highlighting the complexity of distinguishing true inter-individual variability from measurement error in training response studies. The 16-week intervention period, while sufficient for meaningful adaptation, may not capture long-term training responses across multiple competitive seasons, particularly considering how training intensity distributions vary across different competitive phases³⁹. The single-marathon validation approach limits generalizability compared to approaches that incorporate broader performance prediction frameworks⁴⁰. Sample size constraints prevented detailed subgroup analyses within responder clusters, potentially missing important individual characteristics. While genetic-based algorithms for personalized training have shown promise in other contexts⁴¹ and genetic testing for training personalization has been explored⁴², such approaches remain impractical for widespread implementation in marathon training.

Future research should prioritize larger-scale validation studies incorporating diverse competitive environments and extended follow-up periods to establish long-term methodology effectiveness. Investigation of real-time training adjustment algorithms represents a promising avenue for enhancing individualized prescription accuracy, building upon existing systematic reviews of data-guided training prescription³⁷. Extension of these methodologies to other endurance disciplines would determine broader applicability, while current training session models in endurance sports provide implementation frameworks⁴³. Development of more sophisticated machine learning architectures could further improve prediction accuracy and training optimization outcomes.

This investigation provides compelling evidence that personalized training intensity distribution, guided by machine learning analysis of individual athlete characteristics, can significantly enhance marathon performance outcomes. The demonstration that simple, easily assessed factors can effectively predict optimal training methodology challenges current practice patterns and supports a paradigm shift toward individualized prescription strategies. The practical framework developed offers immediate application potential for coaches and athletes while establishing a foundation for future technological innovations in personalized endurance training. These findings underscore the importance of moving beyond population-based training recommendations toward truly individualized approaches that accommodate the substantial biological diversity inherent in human athletic performance.

Conclusion

This investigation establishes that polarized training intensity distribution produces superior marathon performance outcomes compared to pyramidal approaches in recreational athletes, achieving 30% greater improvement with reduced training volume. The identification of substantial inter-individual variability, with four distinct responder clusters comprising the study population, demonstrates that optimal training methodology selection depends critically on individual athlete characteristics rather than universal prescriptions. Machine learning analysis successfully predicted individual training responses using easily accessible parameters, particularly training experience and baseline fitness levels. This finding bridges the gap between theoretical training science and practical implementation by providing coaches and athletes with evidence-based tools for methodology selection. The superior predictive accuracy achieved using consumer-grade monitoring technology addresses previous barriers to personalized training prescription, making individualized approaches feasible for recreational marathon runners.

The physiological mechanisms underlying these performance differences reveal distinct adaptation signatures between training methodologies. Polarized training enhanced maximal aerobic capacity and neuromuscular function, supporting late-race performance advantages observed during competition. Pyramidal training improved submaximal efficiency but failed to translate these adaptations into superior marathon outcomes, challenging traditional volume-focused preparation strategies. These findings have immediate practical implications for marathon training prescription. The training effect matrix developed enables evidence-based methodology selection based on athlete experience, age, baseline fitness, and running mechanics. This framework represents a significant advancement from current one-size-fits-all approaches toward truly personalized training strategies. The broader significance extends beyond marathon running to encompass personalized training prescription across endurance disciplines. The demonstration that machine learning models can effectively guide training methodology selection using non-invasive monitoring technologies establishes a foundation for future innovations in sports science. This research supports a paradigm shift toward individualized training approaches that accommodate the substantial biological diversity inherent in human athletic performance, ultimately optimizing training effectiveness while minimizing the risk of inappropriate prescription for diverse athlete populations.

Data availability

The datasets generated and analyzed during the current study are not publicly available because they contain sensitive participant information and individualized performance data collected through wearable monitoring devices. However, anonymized data supporting the findings of this study are available from the corresponding author upon reasonable request.

Received: 8 May 2025; Accepted: 20 October 2025

Published online: 24 November 2025

References

1. Filipas, L., Bonato, M., Gallo, G. & Codella, R. Effects of 16 weeks of pyramidal and polarized training intensity distributions in well-trained endurance runners. *Scand. J. Med. Sci. Sports.* **32** (3), 498–511 (2022).
2. Rivera-Köfler, T., Varela-Sanz, A., Padron-Cabo, A., Giráldez-García, M. A. & Muñoz-Perez, I. Effects of polarized training vs. Other training intensity distribution models on physiological variables and endurance performance in Different-Level endurance athletes: A scoping review. *J. Strength. Conditioning Res.* **39** (3), 373–385 (2025).
3. Muniz-Pumares, D., Hunter, B., Meyler, S., Maudner, E. & Smyth, B. The training intensity distribution of marathon runners across performance levels. *Sports Med.* **55** (4), 1023–1035 (2025).
4. Casado, A., González-Mohino, F., González-Ravé, J. M. & Foster, C. Training periodization, methods, intensity distribution, and volume in highly trained and elite distance runners: a systematic review. *Int. J. Sports Physiol. Perform.* **17** (6), 820–833 (2022).
5. Haugen, T., Sandbakk, Ø., Seiler, S. & Tønnessen, E. The training characteristics of world-class distance runners: an integration of scientific literature and results-proven practice. *Sports Med. Open* **8** (1), 46, (2022).
6. Stögg, T. L. & Sperlich, B. The training intensity distribution among well-trained and elite endurance athletes. *Front. Physiol.* **6**, 295 (2015).
7. Kenneally, M., Casado, A. & Santos-Concejero, J. The effect of periodization and training intensity distribution on middle-and long-distance running performance: a systematic review. *Int. J. Sports Physiol. Perform.* **13** (9), 1114–1121 (2018).
8. Selles-Perez, S., Fernández-Sáez, J. & Cejuela, R. Polarized and pyramidal training intensity distribution: relationship with a half-ironman distance triathlon competition. *J. Sports Sci. Med.* **18** (4), 708 (2019).
9. Reis, F. J., Alaiti, R. K., Vallio, C. S. & Hespanhol, L. Artificial intelligence and machine learning approaches in sports: Concepts, applications, challenges, and future perspectives. *Braz. J. Phys. Ther.* **28** (3), 101083 (2024).
10. Rodu, J., DeJong Lempke, A. F., Kupperman, N. & Hertel, J. On leveraging machine learning in sport science in the hypothetico-deductive framework. *Sports Medicine-Open* **10** (1), 124 (2024).
11. Bossi, A. H. Time to retire the Raw analysis of individual responses. *Commun. Kinesiol.* **1** 6, (2024).
12. Snapinn, S. M. & Jiang, Q. Responder analyses and the assessment of a clinically relevant treatment effect. *Trials* **8** (1), 31, (2007).
13. Lolli, L. et al. Understanding treatment response heterogeneity using randomised crossover trials: A primer for exercise and nutrition scientists, (2025).
14. Zhang, X., Lin, Z. & Gu, S. A machine learning model the prediction of athlete engagement based on cohesion, passion and mental toughness. *Sci. Rep.* **15** (1), 3220 (2025).
15. Jianjun, Q., Isleem, H. F., Almoghayer, W. J. & Khishe, M. Predictive athlete performance modeling with machine learning and biometric data integration. *Sci. Rep.* **15** (1), 16365 (2025).
16. Cordeiro, M. C., Cathain, C. O., Daly, L., Kelly, D. T. & Rodrigues, T. B. A synthetic data-driven machine learning approach for athlete performance Attenuation prediction. *Front. Sports Act. Living.* **7**, 1607600 (2025).
17. Meixner, B., Filipas, L., Holmberg, H. C. & Sperlich, B. Zone 2 intensity: a critical comparison of individual variability in different submaximal exercise intensity boundaries. *Transl. Sports Med.* **2025** (1), 2008291, (2025).
18. Bonafoglia, J. T. et al. Inter-individual variability in the adaptive responses to endurance and sprint interval training: a randomized crossover study. *PLoS ONE* **11** (12), e0167790, (2016).
19. Laursen, P. & Buchheit, M. *Science and Application of high-intensity Interval Training* (Human kinetics, 2019).

20. Dükinc, P. et al. Monitoring and adapting endurance training on the basis of heart rate variability monitored by wearable technologies: A systematic review with meta-analysis. *J. Sci. Med. Sport.* **24** (11), 1180–1192 (2021).
21. Nuutila, O. P., Nummela, A., Korhonen, E., Häkkinen, K. & Kyrolainen, H. Individualized endurance training based on recovery and training status in recreational runners. *Med. Sci. Sports. Exerc.* **54** (10), 1690 (2022).
22. Braga, F. et al. Abnormal exercise adaptation after varying severities of COVID-19: A controlled cross-sectional analysis of 392 survivors. *Eur. J. Sport Sci.* **23** (5), 829–839 (2023).
23. are They, H. V. *Lactate Threshold Concepts Sports Med.* **39**, 6, 469–490, (2009).
24. Wackerhage, H. & Schoenfeld, B. J. Personalized, evidence-informed training plans and exercise prescriptions for performance, fitness and health. *Sports Med.* **51** (9), 1805–1813 (2021).
25. Javaloyes, A., Sarabia, J. M., Lamberts, R. P., Plews, D. & Moya-Ramon, M. Training prescription guided by heart rate variability vs. block periodization in well-trained cyclists. *J. Strength. Conditioning Res.* **34** (6), 1511–1518 (2020).
26. Pallarés, J. G., Morán-Navarro, R., Ortega, J. F., Fernández-Elías, V. E. & Mora-Rodríguez, R. Validity and reliability of ventilatory and blood lactate thresholds in well-trained cyclists. *Plos One.* **11** (9), e0163389 (2016).
27. Faude, O. et al. Reliability of time-to-exhaustion and selected psycho-physiological variables during constant-load cycling at the maximal lactate steady-state. *Appl. Physiol. Nutr. Metab.* **42** (2), 142–147 (2017).
28. Meyler, S., Bottoms, L. & Muniz-Pumares, D. Biological and methodological factors affecting response variability to endurance training and the influence of exercise intensity prescription. *Exp. Physiol.* **106** (7), 1410–1424 (2021).
29. Lerebourg, L., Saboul, D., Clemencon, M. & Coquart, J. B. Prediction of marathon performance using artificial intelligence. *Int. J. Sports Med.* **44** (05), 352–360 (2023).
30. Shu, D. et al. Prediction of half-marathon performance of male recreational marathon runners using nomogram. *BMC Sports Sci. Med. Rehabilitation.* **16** (1), 97 (2024).
31. Granero-Gallegos, A., González-Quílez, A., Plews, D. & Carrasco-Poyatos, M. HRV-based training for improving VO_{2max} in endurance athletes. A systematic review with meta-analysis. *Int. J. Environ. Res. Public Health.* **17** (21), 7999 (2020).
32. Plews, D. J., Laursen, P. B., Stanley, J., Kilding, A. E. & Buchheit, M. Training adaptation and heart rate variability in elite endurance athletes: opening the door to effective monitoring. *Sports Med.* **43** (9), 773–781 (2013).
33. Knopp, M., Appelhans, D., Schönenfelder, M., Seiler, S. & Wackerhage, H. Quantitative analysis of 92 12-week sub-elite marathon training plans. *Sports Medicine-Open.* **10** (1), 50 (2024).
34. Seiler, S. It's about the long game, not epic workouts: unpacking HIIT for endurance athletes. *Appl. Physiol. Nutr. Metab.* **49** (11), 1585–1599 (2024).
35. Feely, C., Caulfield, B., Lawlor, A. & Smyth, B. Providing explainable race-time predictions and training plan recommendations to marathon runners. In *Proceedings of the 14th ACM Conference on Recommender Systems*, 539–544. (2020).
36. Muñoz-Pérez, I., Castañeda-Babarro, A., Santisteban, A. & Varela-Sanz, A. Predictive performance models in marathon based on half-marathon, age group and pacing behavior. *Sport Sci. Health.* **20** (3), 797–810 (2024).
37. Dükinc, P., Zinner, C., Reed, J. L., Holmberg, H. C. & Sperlich, B. Predefined vs data-guided training prescription based on autonomic nervous system variation: A systematic review. *Scand. J. Med. Sci. Sports.* **30** (12), 2291–2304 (2020).
38. Smyth, B., Lawlor, A., Berndsen, J. & Feely, C. Recommendations for marathon runners: on the application of recommender systems and machine learning to support recreational marathon runners. *User Model. User-Adapt. Interact.* **32** (5), 787–838 (2022).
39. Sperlich, B., Matzka, M. & Holmberg, H. C. The proportional distribution of training by elite endurance athletes at different intensities during different phases of the season. *Front. Sports Act. Living.* **5**, 1258585 (2023).
40. Dash, S. Win your race goal: A generalized approach to prediction of running performance. In *Sports Medicine International Open* **8**, continuous publication, (2024).
41. Jones, N. et al. A genetic-based algorithm for personalized resistance training. *Biology Sport.* **33** (2), 117–126 (2016).
42. Naureen, Z. et al. Genetic test for the personalization of sport training. *Acta Bio Medica: Atenei Parmensis.* **91**, e2020012 (2020). Suppl 13.
43. Tønnessen, E., Sandbakk, Ø., Sandbakk, S. B., Seiler, S. & Haugen, T. Training session models in endurance sports: a Norwegian perspective on best practice recommendations. *Sports Med.* **54** (11), 2935–2953 (2024).

Author contributions

Conceptualization, Gang Qin and Seongno Lee; methodology, Gang Qin; software, Gang Qin; validation, Gang Qin and Seongno Lee; formal analysis, Gang Qin; investigation, Gang Qin; resources, Sungmin Kim; data curation, Gang Qin; writing—original draft preparation, Gang Qin; writing—review and editing, Seongno Lee and Sungmin Kim; visualization, Gang Qin; supervision, Sungmin Kim. All authors have read and agreed to the published version of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.K.

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

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025