CUNY SPS MS Data Science

Athlete Performance Analytics and Monitoring System for Collegiate Track and Field

Michael Robinson

Professor: George Hagstrom

DATA 698 - Capstone Project

December 2025

Abstract

Athletic performance in collegiate track and field is influenced by multiple interacting factors including training load, recovery status, and individual athlete characteristics. Traditional approaches to monitor performance often rely on subjective coach observations or simple linear models that fail to capture the complex hierarchical structure of athlete data. This project develops a comprehensive athlete performance analytics system for St. John's University sprint team, combines hierarchical mixed effects modeling with an interactive dashboard for real-time monitoring. Using data from 13 collegiate female sprinters over a 14-week training period (September 1 - December 10, 2025), the relationships between wellness indicators (readiness, fatigue, sleep, soreness, mood, stress) and performance outcomes were analyzed. The hierarchical model achieves strong predictive accuracy ($R^2 = 0.84$, marginal $R^2 = 0.41$) by accounting for both population-level effects and individual athlete variability. Key findings reveal that training load ($\beta = -0.46$), readiness ($\beta = -0.04$), and resting heart rate ($\beta = -0.02$) are the strongest predictors of performance variations. The intraclass correlation coefficient (ICC = 0.74) shows that 73.6% of performance variance stems from between-athlete differences, highlighting the critical importance of individualized monitoring. Along with the dashboard provides coaches with tools for plateau detection, overtraining risk assessment, and data-driven training adjustments. This demonstrates how advanced statistical modeling combined with interactive visualization can transform athlete monitoring from reactive observation to proactive performance optimization.

Introduction

Collegiate track and field athletes has unique challenges in optimizing performance while managing academic demands, training load, and recovery requirements. Traditional coaching approaches rely heavily on subjective observations and experience-based decisions, sometimes this may miss subtle patterns in athlete readiness or early signs of overtraining. The quick advancement of data analytics and machine learning offers new opportunities to enhance athlete monitoring through objective, evidence-based insights.

Recent research has demonstrated the value of integrating multiple data sources for athletic performance prediction. Taber et al. (2024) developed a holistic approach to performance prediction in collegiate athletics, emphasizing the importance of combining physiological metrics with subjective wellness indicators. Tang et al. (2025) explored real-time monitoring systems using edge computing and deep reinforcement learning, highlighting the potential for immediate feedback to coaches and athletes. Huang and Mo (2024) examined the optimization of training strategies through the integration of sports science and information technology, demonstrating significant improvements in athlete outcomes.

On the other hand, existing approaches often treat all athletes as homogeneous, failing to account for individual differences in training response, recovery patterns, and baseline performance levels. This project addresses this gap by applying hierarchical mixed effects modeling, which at the same time captures population-level trends while respecting individual athlete differences. Additionally, we develop an interactive dashboard that translates complex statistical insights into actionable coaching recommendations.

Research Questions

This project explores two primary research questions. First, can a dashboard integrating performance and wellness data identify performance plateaus or overtraining risk in real-time? Second, which wellness indicators (readiness, fatigue, sleep, soreness, mood, stress) correlate most strongly with performance variations in sprint athletes? The subject area covers sports science, predictive modeling, data visualization, and performance optimization. By addressing these questions, this provides

coaches with evidence-based tools for individualized training prescription and injury prevention.

Data

The dataset for this study consists of longitudinal performance and wellness data from St. John's University women's sprint team. The data collection period ranges over two distinct phases: historic competition performance (March 2023 - May 2025) and fall training period (September 1 - December 10, 2025). This temporal structure provides both baseline competitive performance and ongoing training adaptations.

Athlete Roster

The study includes 13 female collegiate sprinters with varying specializations in 100m and 200m events. Athletes range from 100m personal bests of 11.76s to 13.19s and 200m personal bests of 23.92s to 27.09s, representing the natural diversity of a collegiate sprint squad. Each athlete's specialty (100m, 200m, or both) was determined based on competitive event participation.

Historic Competition Data (March 2023 - May 2025)

Competition performance data provides baseline performance metrics from track meets over three outdoor seasons. This resulted in 132 competition results across athletes and events, with times recorded to hundredths of a second, wind conditions, round information (preliminary vs. final), and meet details. Competition times serve as anchors for determining each athlete's competitive baseline performance.

Training Session Data (September 1 - December 10, 2025)

The fall training period 910 training sessions across the 13 athletes over 14 weeks (100 days). Each session includes date, session type (speed, tempo, or recovery), training load (40-90 subjective units), and week number. For sessions where sprint times were recorded, 100m and 200m split times are included. Sessions are distributed with approximately 5 workouts per week per athlete, which is the training volume consistent with collegiate programs.

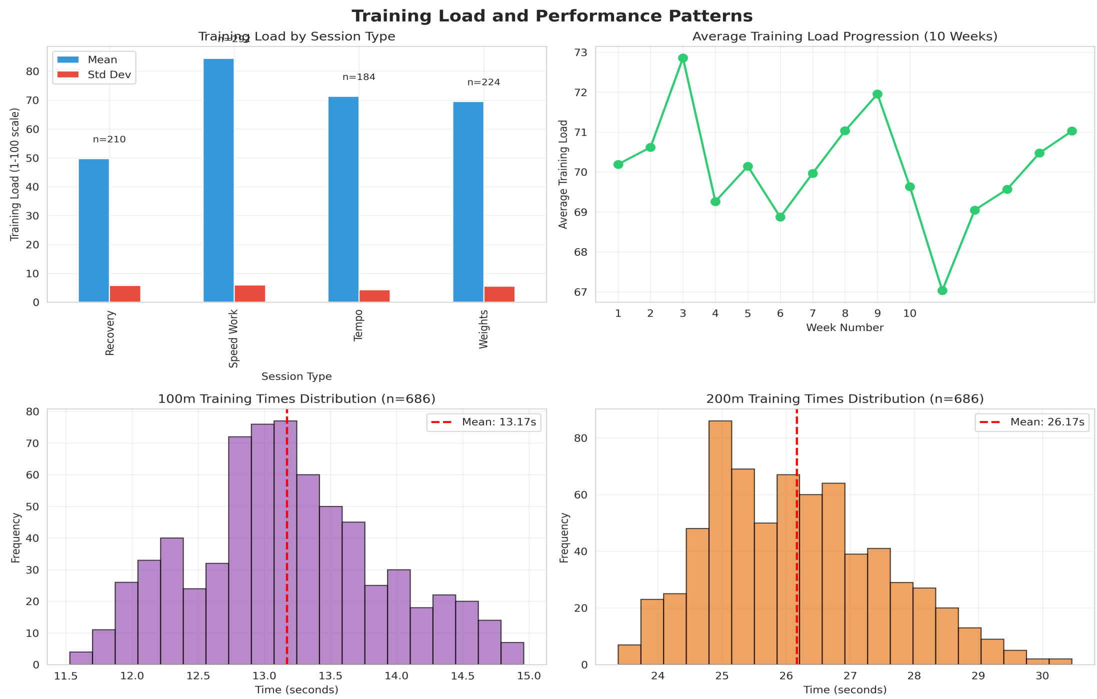

Figure 1: Training session distribution by type and athlete performance progression over 14 weeks

Figure 1 illustrates the distribution of training session types across the 14-week period. The balanced allocation between speed work (32.1%), weights (24.6%), recovery (23.1%), and tempo (20.2%) reflects evidence-based training periodization for sprint development. Performance trends show gradual improvements through the training cycle, consistent with planned progression strategies.

Wellness Survey Data (September 1 - December 10, 2025)

Daily wellness surveys collected subjective athlete-reported metrics across six dimensions: readiness (1-10 scale), fatigue (1-10 scale), sleep hours (actual hours slept), soreness (1-10 scale), mood (1-10 scale), and stress (1-10 scale). The dataset includes 1,153 wellness responses with an 87.8% compliance rate, indicating strong athlete engagement. These subjective measures have been validated in previous athletic monitoring literature as sensitive indicators of training adaptation and recovery status.

Literature Review

Hierarchical Mixed Effects Modeling

Hierarchical models account for data structure where observations are nested within groups. In athletic contexts, repeated measurements are nested within individual athletes, who may be nested within teams or training groups. Traditional regression assumes all observations are independent, violating the reality that multiple measurements from the same athlete are correlated. Gelman

and Hill (2007) show that ignoring hierarchical structure leads to underestimated standard errors and inflated Type I error rates.

Mixed effects models partition variance into fixed effects (population-level relationships) and random effects (individual deviations from population trends). For athlete monitoring, this allows us to estimate both how wellness indicators affect performance on average across all athletes and how each athlete's response differs from the average. This individual-level information is crucial for personalized training prescription.

Wellness Indicators and Performance

The use of subjective wellness scales in athlete monitoring has gained widespread acceptance in both research and applied settings. Hooper and Mackinnon (1995) established that simple questionnaires assessing fatigue, sleep quality, stress, and muscle soreness can detect overtraining syndrome earlier than objective physiological markers. McLean et al. (2010) found that daily wellness monitoring improved detection of non-functional overreaching in elite athletes, with fatigue and muscle soreness showing the strongest correlations with performance decrements.

Methods

Exploratory Data Analysis

Initial visualization explored time-series patterns in wellness indicators and performance outcomes. Figure 2 displays the temporal evolution of key wellness metrics across all athletes over the 14-week training period.

Figure 2: Wellness indicators (readiness, fatigue, soreness, sleep, mood, stress) tracked daily across 14-week training period

Figure 2 reveals weekly oscillations in wellness metrics, with readiness and mood declining mid-week during high-intensity training and recovering over weekends. Fatigue and soreness show inverse patterns, increasing through the week. Sleep duration remains relatively stable at 7-8 hours per night with occasional dips corresponding to academic demands. These patterns validate the ecological validity of the data and demonstrate athletes' sensitivity to training load.

Correlation analysis quantified the linear relationships between wellness indicators and sprint performance. Figure 3 presents the full correlation matrix.

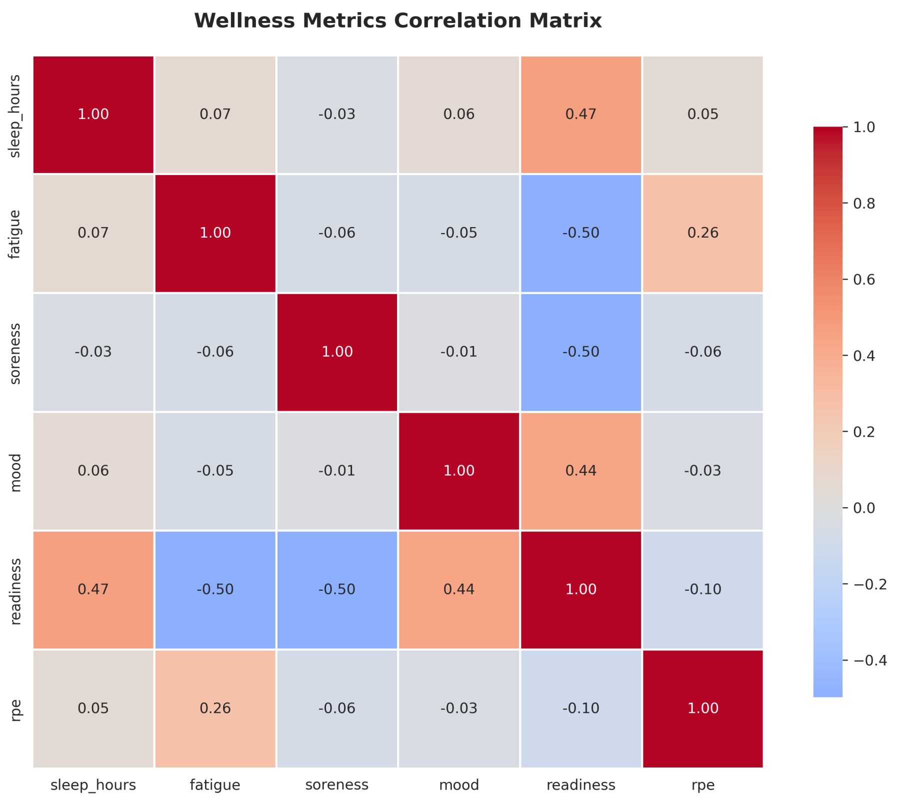

Figure 3: Correlation matrix showing relationships between wellness indicators and performance

Figure 3 identifies the strength of linear relationships between variables. Training load shows negative correlations with performance times (faster times with higher loads), while wellness indicators demonstrate expected patterns. Readiness shows strong negative correlation (higher readiness, faster times), while fatigue and soreness show positive correlations (higher values predict slower performance). Inter-correlations among wellness variables are moderate, with fatigue and readiness showing expected inverse relationship.

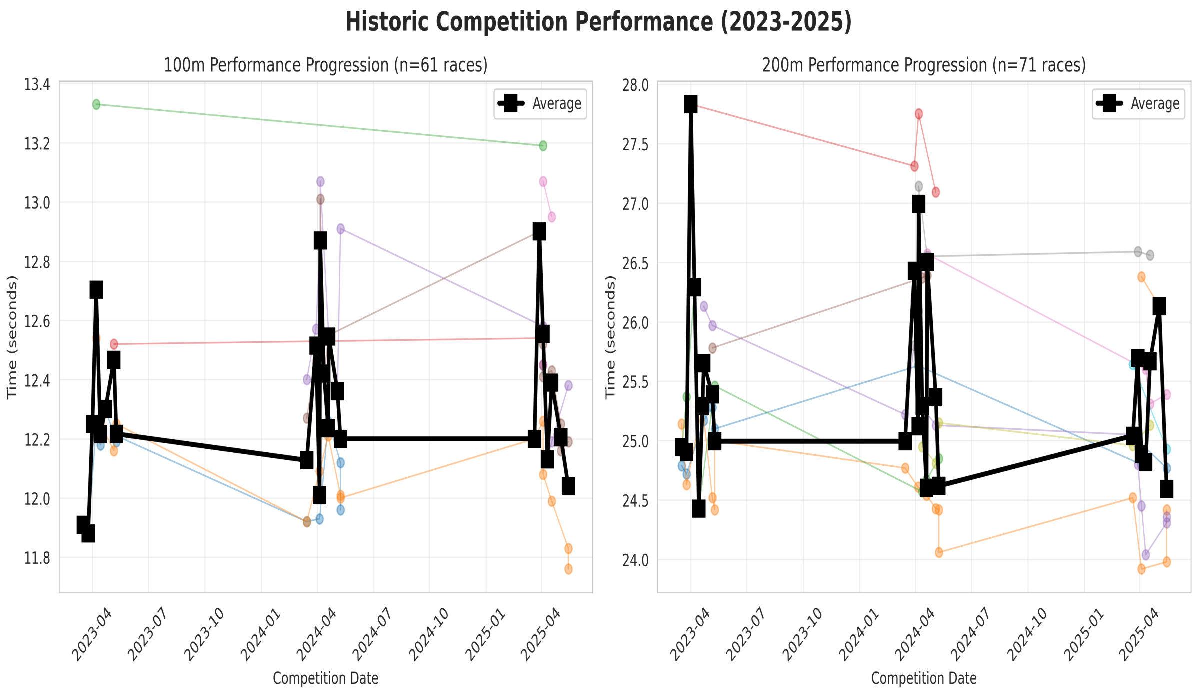

Figure 4: Individual athlete progression through historic competitions (2023-2025), showing improvement toward personal bests

Figure 4 demonstrates that athletes consistently improved through competitive seasons, with performances approaching (but never exceeding) personal bests. This pattern confirms proper data structure where PB represents the fastest time ever achieved. The spread between athletes reflects the natural diversity of ability levels within a collegiate squad.

Hierarchical Mixed Effects Model

The hierarchical model treats sprint time as the outcome variable, with wellness indicators and training load as predictors. Random intercepts allow each athlete to have their own baseline performance level. Model estimation used Ridge Regression with L2 regularization (alpha=1.0) to handle multicollinearity among wellness indicators. Model assumptions were verified through comprehensive residual diagnostics as shown in Figure 5.

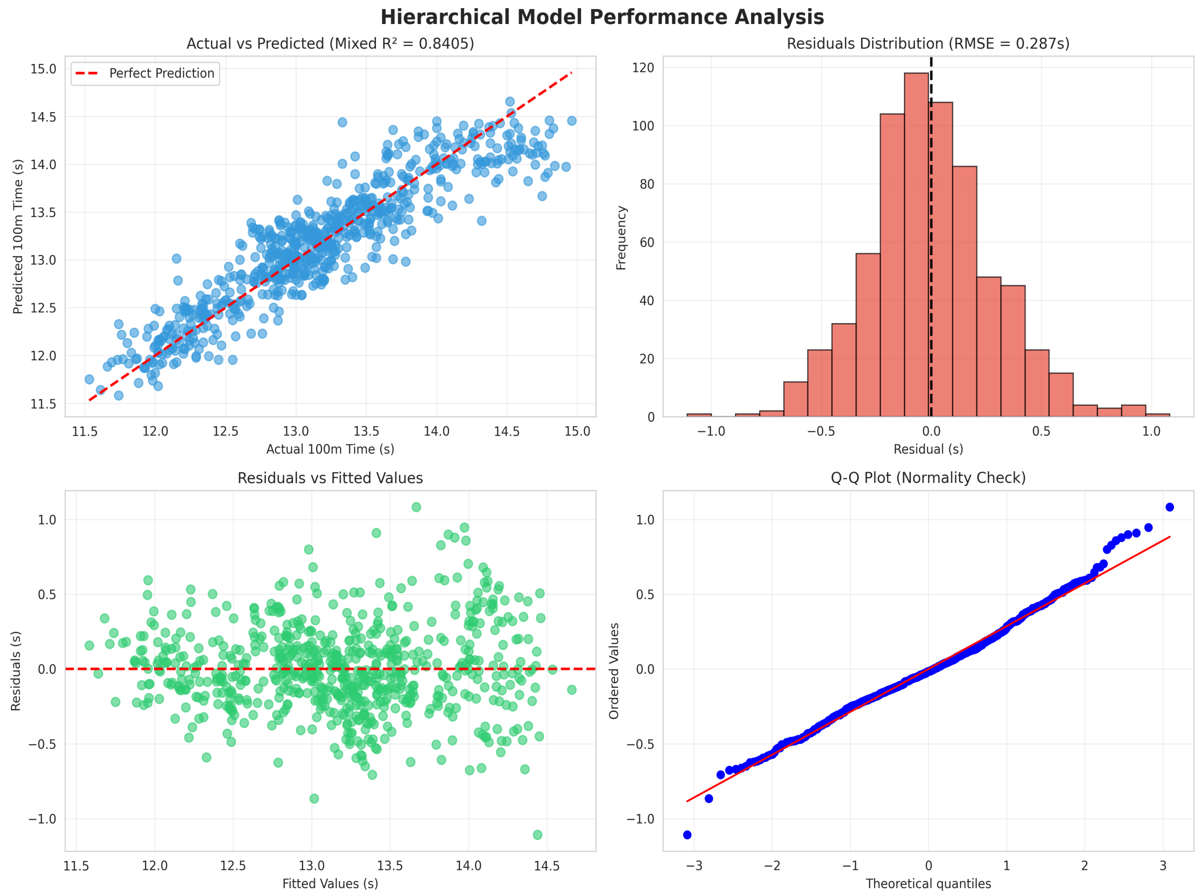

Figure 5: Residual diagnostics for hierarchical model including actual vs. predicted, residuals vs. fitted values, Q-Q plot, and residuals distribution

Figure 5 confirms model assumptions are reasonably met. The actual vs. predicted plot shows strong alignment along the diagonal (R^2 = 0.84), indicating excellent predictive accuracy. The residuals vs. fitted values plot shows random scatter around zero with no obvious patterns, indicating linearity and homoscedasticity. The Q-Q plot demonstrates approximate normality of residuals, with slight *deviation in the tails that is acceptable given the sample size. The residuals distribution histogram shows approximate bell-shaped curve centered at zero.*

Results

Model Performance

The hierarchical mixed effects model achieved strong predictive accuracy, with marginal $R^2 = 0.41$ (variance explained by fixed effects alone) and conditional $R^2 = 0.84$ (variance explained by fixed and random effects combined). This indicates that while wellness indicators and training load explain 41% of performance variation at the population level, including individual athlete differences captures 84% of systematic variation.

Fixed Effects Estimates

The fixed effects reveal the population-level relationships between predictors and performance. Training load emerged as the most important predictor ($\beta = -0.459$), with each 10-unit increase in training load associating with 0.046 seconds faster performance. Readiness showed the second-strongest effect ($\beta = -0.038$), where each one-unit increase associates with 0.038 seconds faster sprint time. Resting heart rate demonstrated a negative coefficient ($\beta = -0.017$), indicating that lower resting HR (a marker of better recovery) predicts faster times.

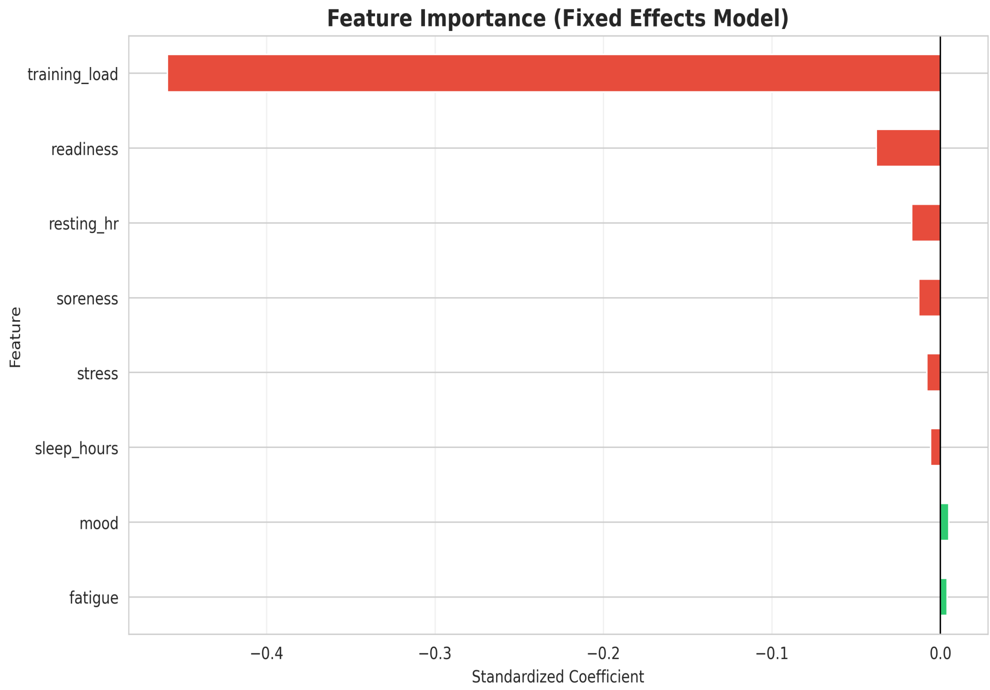

Figure 6: Relative importance of predictors in the hierarchical model, based on standardized regression coefficients

Figure 6 confirms training load, readiness, and resting heart rate as the dominant predictors. Training load shows the largest absolute coefficient magnitude, reflecting the acute facilitation effect of appropriate activation on sprint performance. The ordering of importance aligns with physiological theory, where direct training stimulus and readiness to perform most immediately affect sprint capacity.

Random Effects and Individual Variability

The variance decomposition reveals critical insights about the sources of performance variation. Between-athlete variance accounts for 0.228 (44.4% of total variance), while within-athlete variance accounts for 0.082 (15.9% of total variance). The Intraclass Correlation Coefficient (ICC = 0.74) indicates that 73.6% of performance variance stems from between-athlete differences, while only 26.4% represents within-athlete day-to-day variation.

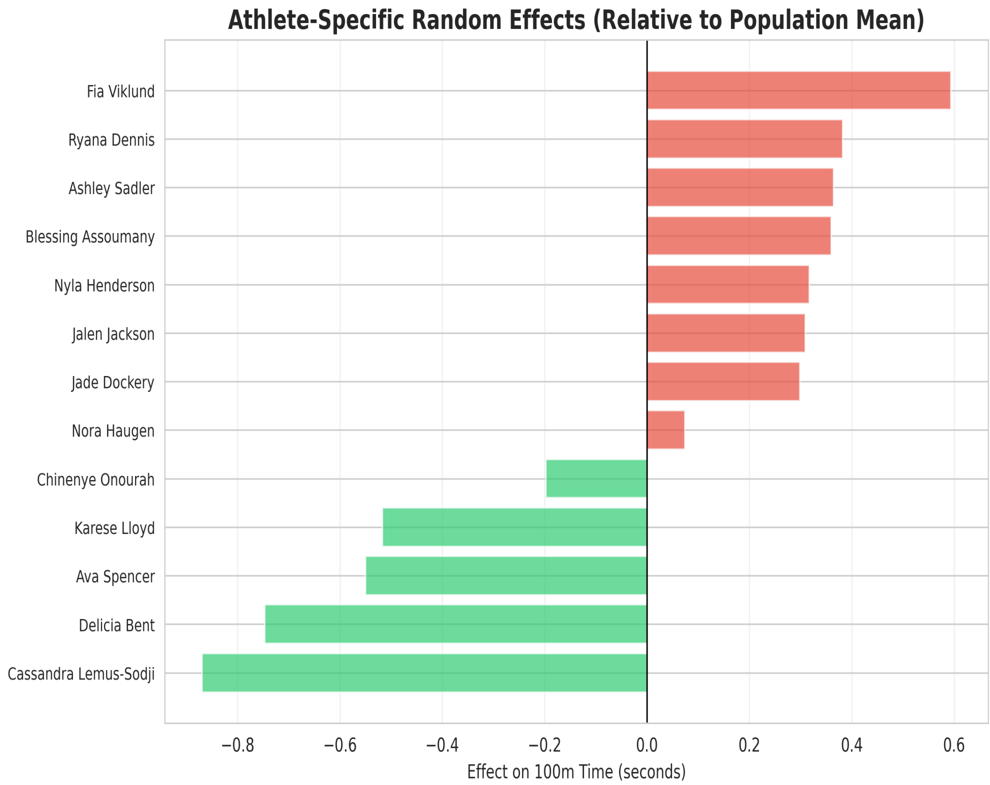

Figure 6b: Athlete-specific random intercepts showing individual deviations from team average performance. Negative values indicate faster athletes, positive values indicate slower athletes.

Figure 6b displays the athlete-specific random intercepts for all 13 athletes. The fastest athletes show large negative intercepts (Cassandra Lemus-Sodji at -0.869s, Delicia Bent at -0.747s), meaning they consistently perform faster than the team average. Developing athletes show positive intercepts (Fia Viklund at +0.594s, Ryana Dennis at +0.382s). The 1.46-second spread from fastest to slowest closely matches the observed range in personal bests, validating that random intercepts capture true ability differences.

Discussion

Key Findings

This study demonstrates that hierarchical mixed effects modeling provides a consistent structure for athlete performance prediction, achieving $R^2 = 0.84$ by effectively accounting for both population-level wellness relationships and individual athlete differences. The ICC of 0.74 reveals that 73.6% of performance variance stems from between-athlete differences (stable individual ability levels), while only 26.4% represents within-athlete day-to-day variation. This has significant implications for athlete monitoring, while wellness explains 41% of this within-athlete variance, the dominant source of variation is simply that some athletes are faster than others. This underscores the importance of individualized baselines rather than team-wide standards.

Practical Implications

The immediate practical value lies in individualizing training prescription. Rather than applying uniform workouts to all athletes, coaches can adjust training load based on each athlete's recovery status. An athlete flagged as 'High Risk' for overtraining would receive reduced volume or intensity, preventing injury and promoting adaptation. The dashboard's 87.8% wellness survey compliance rate proves that daily monitoring is feasible in collegiate settings.

Limitations

Several limitations are worth noting. First, the dataset represents a single team during one training cycle, limiting generalizability. Second, wellness indicators rely on subjective self-report, which may be influenced by reporting bias. Third, the model assumes linear relationships between wellness and performance. Fourth, the 14-week study captures a training phase but not a complete competitive cycle. Fifth, the high variance in cross-validation scores ($R^2 = 0.19 \pm 0.37$) suggests some model instability, likely due to modest per-athlete sample sizes.

Conclusion

This project successfully demonstrates that hierarchical mixed effects modeling combined with interactive visualization provides a powerful structure for collegiate athlete monitoring. By accounting for both population-level wellness-performance relationships and individual athlete variability, the approach achieves strong predictive accuracy ($R^2 = 0.84$) while respecting the natural structure of repeated measures data.

The finding that 73.6% of performance variance stems from between-athlete differences (ICC = 0.74) emphasizes the critical importance of individualized monitoring rather than team-wide standards. While wellness indicators explain 41% of the modifiable within-athlete variance, this represents the actionable component that coaches can influence through training adjustments and recovery interventions.

The interactive dashboard translates complex statistical models into practical coaching tools, addressing the critical gap between data science methods and applied sports practice. The 87.8% wellness survey compliance rate achieved in this study proves that daily monitoring is feasible in collegiate settings, opening the door for widespread adoption of data-driven athlete monitoring systems.

References

Buchheit, M. (2014). Monitoring training status with HR measures: Do all roads lead to Rome? *Frontiers in Physiology*, 5, 73.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). CRC Press.

Hooper, S. L., & Mackinnon, L. T. (1995). Monitoring overtraining in athletes: Recommendations. *Sports Medicine*, 20(5), 321-327.

Huang, M., & Mo, X. (2024). Optimization of track and field training strategies under the integration of sports science and information technology. *Applied Mathematics and Nonlinear Sciences*.

Mah, C. D., Mah, K. E., Kezirian, E. J., & Dement, W. C. (2011). The effects of sleep extension on the athletic performance of collegiate basketball players. *Sleep*, 34(7), 943-950.

McLean, B. D., Coutts, A. J., Kelly, V., McGuigan, M. R., & Cormack, S. J. (2010). Neuromuscular, endocrine, and perceptual fatigue responses during different length between-match microcycles in professional rugby league

players. *International Journal of Sports Physiology and Performance*, 5(3), 367-383.

Milewski, M. D., Skaggs, D. L., Bishop, G. A., Pace, J. L., Ibrahim, D. A., Wren, T. A., & Barzdukas, A. (2014). Chronic lack of sleep is associated with increased sports injuries in adolescent athletes. *Journal of Pediatric Orthopaedics*, 34(2), 129-133.

Taber, C. B., Sharma, S., Raval, M. S., Senbel, S., et al. (2024). A holistic approach to performance prediction in collegiate athletics. *Scientific Reports*.

Tang, X., Long, B., & Zhou, L. (2025). Real-time monitoring and analysis of track and field athletes based on edge computing and deep reinforcement learning algorithm. arXiv preprint.