Astronomy & Astrophysics manuscript no. aa
December 7, 2025

Predicting NBA Team Performance: From Historical Standings to Player-level Forecasts

Weihao Li[1,⋆], Shuhao Gao[1], C. Brassac[1], and C. Estievenart[1]

McCormick School of Engineering, Northwestern University
e-mail: {weihaoli2027, shuhaogao2027, author3, author4}@u.northwestern.edu

Received December 8, 2025

ABSTRACT

Context. In recent years, publicly available NBA play-by-play and box-score data have made it possible to build increasingly detailed models of team performance. At the same time, many forecasting systems still rely on simple team-level summaries such as past wins or point differential.
Aims. We aim to compare a simple standings-based baseline with a player-centric forecasting pipeline for predicting future regular-season wins at the team level.
Methods. We construct two models. The first directly regresses future team wins on historical win totals over the past two decades. The second learns generic player evolution patterns from a long-horizon panel of both retired and active players, uses these patterns to predict next-season box-score statistics for active players, and aggregates player-level projections into team-level features.
Results. Across held-out seasons, the player-based pipeline improves mean absolute error relative to the standings baseline, especially for teams undergoing large off-season roster changes. For more stable teams, historical standings are surprisingly competitive and sometimes match the player-based forecasts.
Conclusions. Our results suggest that simple team-level baselines remain strong when rosters are stable, but player-level modeling adds value in high-turnover regimes and provides a more interpretable connection between roster moves and expected team performance.

Key words. basketball – sports analytics – machine learning – time series forecasting – player development

1. Introduction

In this project, we study how well different approaches can predict NBA team performance. Our goal is straightforward: given past data, can we forecast how many games each team will win in a future season? To explore this question, we compare two models that rely on fundamentally different sources of information.

The first model is a simple team-level baseline that uses only historical standings. For any pair of seasons within the past twenty years, we examine how a team's previous win total relates to its win total in another season. This provides a direct, standings-based method for predicting future performance without considering player-level factors.

The second model takes a player-centered perspective. We collect long-term statistics for both retired and still-active NBA players and train a model to learn how player performance tends to evolve over time. Using these learned patterns, we predict next-season box-score statistics for current active players and then aggregate the projected player output to estimate each team's future win total.

By comparing these two approaches, we aim to understand when a simple standings-based baseline is sufficient and when a more detailed, player-based forecasting pipeline provides an advantage, particularly in seasons affected by significant roster changes.

⋆ Corresponding author: wli@u.northwestern.edu

2. Prior Literature

2.1. Team-level rating and expectation models

A large body of work in sports analytics has focused on predicting team performance using aggregate team-level statistics. One influential family of models is the *Pythagorean expectation*, which estimates a team's theoretical winning percentage from points scored and points allowed via a power-law relationship. Originally proposed for baseball and later adapted to basketball, this idea underlies many simple baselines that relate scoring margins to wins and losses (see, e.g., Oliver 2004; Sarlis & Tjortjis 2020). These models provide strong, easy-to-interpret benchmarks but depend solely on aggregate scoring margins and do not incorporate information about individual players.

Another common approach is to model team strength with rating systems inspired by Elo. In basketball applications, Elo-style ratings are updated after each game based on the result, home-court advantage, and the margin of victory, and then used to forecast future game and series outcomes. Public-facing systems such as FiveThirtyEight's NBA model illustrate how dynamic ratings can track changes in team strength over a season and provide reasonably accurate probabilistic predictions for both games and playoff series (FiveThirtyEight 2015). Together, Pythagorean and Elo-style systems represent static or quasi-static team-level baselines that are closely related to our first model, which relies on historical team wins and standings. However, because these methods operate on aggregate team outcomes, they are inherently limited in their ability to anticipate abrupt changes in performance driven by roster turnover or player development.

2.2. Machine learning for NBA game and season prediction

Beyond analytic formulas and rating systems, many studies have applied machine learning to predict basketball results using team-level features. Early work by Loeffelholz et al. (2009) used neural networks to predict single-game NBA outcomes from box-score statistics and contextual variables, demonstrating that nonlinear models can capture interactions between basic team statistics. More recent surveys review a wide range of approaches, including logistic regression, support vector machines, tree-based ensembles, and deep neural networks, and typically find that machine-learning models outperform simpler statistical baselines when sufficient historical data are available (Sarlis & Tjortjis 2020).

Researchers have also moved from game-level prediction to season-level tasks, such as forecasting a team's final win total or playoff qualification. For example, Yang (2015) regress regular-season wins on team-level and aggregated player statistics to study which factors are most predictive of team success. These season-level models again treat each team as the unit of analysis and usually rely on summary statistics from the current or previous season.

2.3. Player-level prediction and its link to team performance

Complementary to team-level approaches, a growing literature studies player evaluation and performance prediction using detailed box-score and tracking data. Sarlis & Tjortjis (2020) review many of these methods, including regression-based models for player efficiency metrics, clustering techniques for grouping players with similar playing styles, and rating systems that quantify individual contribution to team success. In practice, coaches and analysts often combine such player-level models with domain knowledge to support decisions about rotations, matchups, and roster construction.

Some work, including Yang (2015), aggregates player statistics into team-level features to predict season outcomes, effectively creating a simple player-to-team pipeline. However, existing player-focused models typically train and evaluate on overlapping time periods for the same set of players, and they rarely combine long-run player evolution with explicit team-level baselines.

In contrast, our project is designed to learn generic player evolution patterns from a historical cohort that includes both retired and still-active players, and then apply these patterns to predict the next-season performance of currently active players. We subsequently aggregate predicted player statistics to the team level and compare the resulting win forecasts to those from a simple standings-based baseline over roughly two decades of NBA data. This setup allows us to evaluate how a player-based forecasting pipeline and a standings-based model perform side by side, especially in seasons where teams undergo substantial roster changes.

3. Example notation and basic formulas

4. Conclusions

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. In hac habitasse platea dictumst. Integer tempus convallis augue. Etiam facilisis.

Acknowledgements. We thank the open basketball-reference and NBA Stats communities for making historical data accessible, and the Northwestern MSCS program for general support of this work.

References

FiveThirtyEight. 2015, How Our NBA Predictions Work, https://fivethirtyeight.com/features/how-our-nba-predictions-work/, accessed: 2025-12-07

Loeffelholz, B., Bednar, E., & Bauer, M. 2009, Journal of Quantitative Analysis in Sports, 5

Oliver, D. 2004, Basketball on Paper: Rules and Tools for Performance Analysis (Washington, DC: Potomac Books)

Sarlis, P. & Tjortjis, C. 2020, Information Systems, 93, 101562

Yang, J. 2015, Predicting Regular Season Results of NBA Teams Based on Regression Analysis of Common Basketball Statistics, UC Berkeley Statistics Project