Quantifying and Comparing NBA Player Career Momentum Using Statistical Methods

Ross Lauterbach
CUNY Hunter College, New York, USA
rosslauterbach1@gmail.com

Abstract

Momentum is one of the most widely referenced yet poorly defined concepts in sports. In the NBA, commentators and fans routinely describe players as "heating up" or "catching fire," often attributing shifts in performance to an intangible momentum factor. Despite its prominence in narrative and analysis, momentum is a measure that has been hard to verify empirically. This paper introduces a statistical approach to capture player momentum throughout an NBA career using smoothed performance trajectories. By constructing game-by-game momentum data and powerful visualizations, we aim to identify sustained periods of elevated or diminished performance and quantify the uncertainty around them. We also take a deep dive into methods of calculation and modeling using momentum.

1 Introduction

While there have been numerous attempts to capture momentum at a specific point in time, particularly within individual games, few have sought to define and quantify it across a player's entire career. In the realm of basketball analytics, much of the momentum literature has focused on short-term phenomena such as hot streaks and game-to-game variability. A study by Gilovich et al. (1985) challenged the widely held belief in the "hot hand," arguing that perceived shooting streaks were simply cognitive illusions rather than statistical realities. More recent work, however, has re-evaluated this conclusion; Miller and Sanjurjo (2018) showed that earlier studies underestimated the likelihood of streaks due to statistical bias, providing evidence that hot-hand effects are both real and measurable.

Beyond in-game performance, researchers have also explored whether momentum carries over between games. Arkes and Martinez (2011) used an econometric framework to assess team-level momentum in the NBA, finding that recent success modestly improves the probability of future wins, even after controlling for team strength. These studies demonstrate a growing interest in quantifying momentum, but they remain largely confined to short-term patterns and team dynamics. This paper aims to extend that line of inquiry by shifting the focus to long-term, player-level momentum. Rather than capturing momentary flashes of brilliance, we propose a model that smooths game-by-game performance data to trace career-long trends. This enables us to identify sustained periods of elevated or diminished output and to quantify uncertainty around those trends. In doing so, we contribute a new tool for understanding player consistency and for empirically validating or challenging popular narratives about career arcs.

2 Data Description

The dataset consists of player-level game logs from NBA regular season games compiled from the official NBA API and also available on Kaggle. Each observation represents a single player's performance in a single game and includes a wide range of traditional and advanced statistics: points, assists, rebounds, steals, blocks, turnovers, shooting percentages, and more. Data was collected starting in 1980, when the 3-point line was introduced, all the way until the end of the 2022-2023 season, resulting in approximately 1.4 million entries. We filtered the dataset to include only regular-season games to avoid postseason variability and ensure comparability across players.

3 Calculation of Momentum

To construct a player's momentum curve, we indexed their games chronologically and computed a game number variable to serve as a time axis. We removed duplicate records and ensured consistent game tracking by identifying each player-game instance via a unique combination of player ID and game ID. A validation test was also run against publicly available data to ensure aggregates agreed with each other. This cleaned dataset serves as the foundation for our momentum calculations.

Figure 1 contains a scatter plot and correlation coefficient for each variable interaction. The variables that were initially under consideration for the momentum calculation were used: points, rebounds, assists, steals, blocks, turnovers, and various efficiency measures.

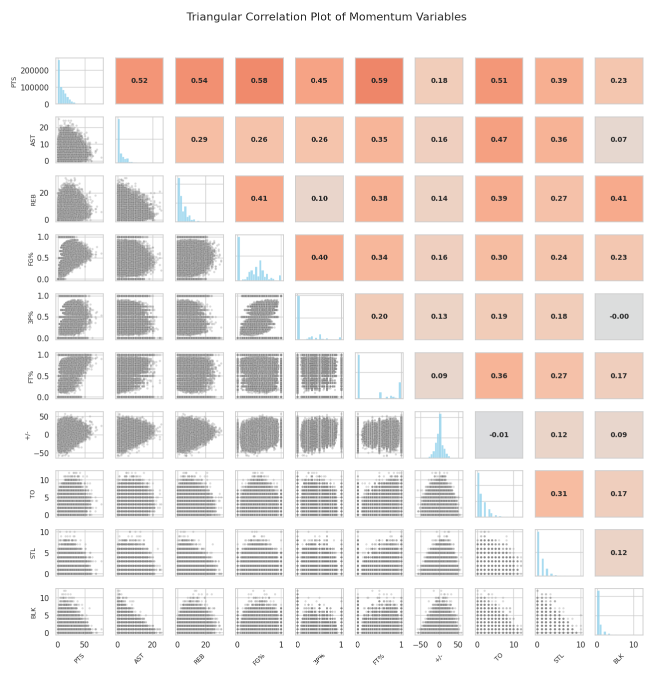

Figure 1: Triangular correlation plot of momentum variables.

Given the lack of extreme multicollinearity, we continued with our analysis as planned. However, the moderate correlation between output and efficiency led us to focus solely on player output as a measure of momentum. This specific weighting below was used given its interpretability and the distribution of the data, but was overall an arbitrary choice based on input from basketball fans and analysts. The score aims to capture a player's all-around influence in a given game while only incorporating the 6 major simple performance metrics. We then smoothed these scores over time using an exponentially weighted moving average (EWMA) to reflect recent performance trends while preserving long-term stability. This smoothed score serves as the core of our momentum metric. Initially, we also incorporated a team indicator representing recent team success, defined as a scaled 10-game rolling win count, to capture potential psychological or contextual effects on a player's performance. This led to the general momentum equation, which is the sum of the EWMA-based performance score and a weighted team indicator. The equation for the performance score, S_i is shown below.

$$S_i = \text{Points}_i + 2 \cdot (\text{Rebounds}_i + \text{Assists}_i) + 5 \cdot (\text{Steals}_i + \text{Blocks}_i - \text{Turnovers}_i)$$

The win indicator I_i is calculated as one fifth of the sum of the binary win/loss indicator W_k over the previous 10 games, with a subtraction of 5 to standardize it:

$$I_i = \frac{1}{5} \sum_{k=i-9}^{i} W_k - 5$$

where W_k is a binary indicator for win (1) or loss (0) at game k.

Finally, the momentum M is calculated by applying an exponential smoothing factor α and a decay parameter γ as follows:

$$M = (1-\alpha) \sum_{j=1}^{i-1} \alpha (1-\alpha)^{i-j-1} S_j + \gamma \cdot I_i$$

4 Momentum Optimization

While the primary goal of this metric is not predictive accuracy, we initially explored whether momentum could be tuned to enhance its alignment with future or current performance outcomes. The underlying hypothesis was that a player's momentum score could offer predictive value beyond its descriptive nature, potentially serving as a useful indicator of future game performance or for assessing a player's impact in the current game. To test this hypothesis, we focus on optimizing two key parameters in the momentum formula: the smoothing parameter, alpha, and the weighting factor of the team, gamma. The idea was to minimize the mean squared error (MSE) between the momentum score and either a player's next-game performance score or their current game plus-minus, which quantifies a player's overall contribution relative to the game outcome. By doing so, we aimed to validate the strength of the momentum signal in predicting a player's performance and to investigate whether the momentum metric could be improved by incorporating team performance factors alongside individual statistics.

The process involved systematically adjusting the parameters alpha and gamma and observing their impact on the predictive accuracy of the momentum score. After a series of trials, the optimal parameter values suggested an intriguing result: the best-performing momentum score excluded team context entirely, with a gamma value of 0, indicating that recent team success did not add any predictive power. This finding was particularly interesting, as it suggested that a player's momentum might be more directly tied to their own individual performance trajectory rather than the broader team performance. Additionally, the value of alpha that minimized the MSE was relatively low, at 0.1, signifying that a shorter smoothing window, one that emphasizes more recent games, was most effective in predicting future performance. Figure 2 shows a heatmap of the optimization process, with darker colors indicating higher performance under the simple linear model with one predictor variable.

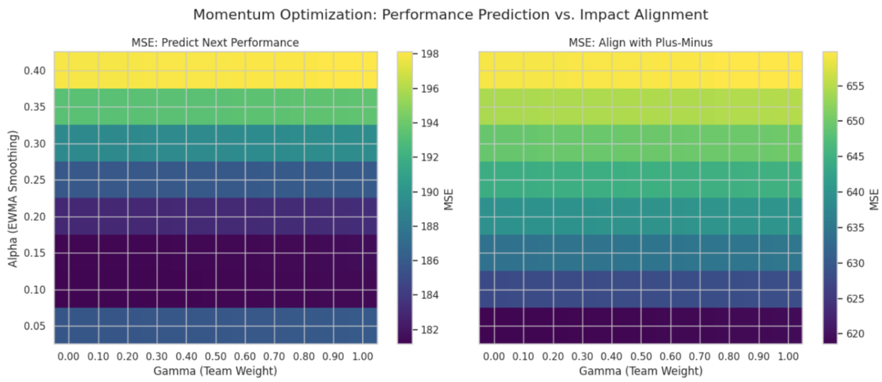

Figure 2: Heatmap of output from grid search momentum optimization using MSE.

5 Momentum Visualization

Now that we can calculate player momentum at a given point in time, we can track and visualize an athlete's performance throughout their career. In particular, momentum curves capture the fluctuations in a player's contributions during a season or throughout their career, providing a dynamic view of their consistency and impact. Rather than simply looking at traditional statistics, we can gain a better understanding of how a player evolves and maintains consistency throughout their career. The previously used methods are meant to filter out the noise of random fluctuations in performance and highlight underlying trends in the data. These smoothed trajectories reveal not only the magnitude of a player's performance but also the stability or volatility of their impact over time.

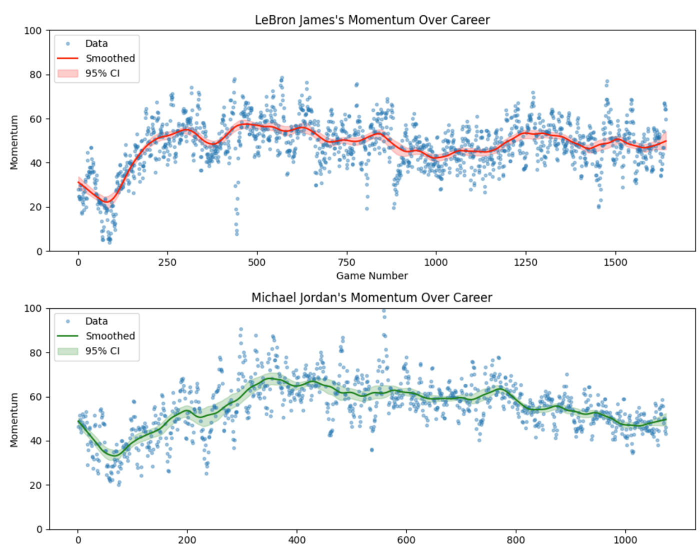

Figure 3: Momentum curves for Lebron James and Michael Jordan with 95% confidence bands.

6 Career Trajectory Clustering

The next objective was to classify players' careers by clustering them on the basis of the shape of their momentum trajectories, independent of the era in which they played. To do this, we created a fixed length vector for each player by interpolating their momentum scores over the first 300 games of their career. This interpolation allowed for a consistent representation of momentum between players, enabling an analysis of how their career trajectories evolved. Only players with at least 300 games were included in the analysis to ensure that each player had a sufficient sample size for comparison and to maintain consistency across the dataset.

However, during the interpolation process, some players had missing values in their momentum curves, often due to irregular game participation, injuries, or gaps in the available data. To ensure that the clustering analysis was based on clean and interpretable data, we excluded any momentum vectors that contained missing values. This filtering step was critical, as it prevented incomplete data from distorting the clustering results and ensured that the dimensionality reduction and clustering algorithms, such as principal component analysis (PCA) and k-means, operated on reliable and consistent input. By removing incomplete data points, we were able to minimize bias and instability in the clustering process, resulting in more accurate groupings of players based on the overall shape of their momentum trajectories. Figure 4 shows both the mean career trajectories in each cluster and the class distinction using PCA components.

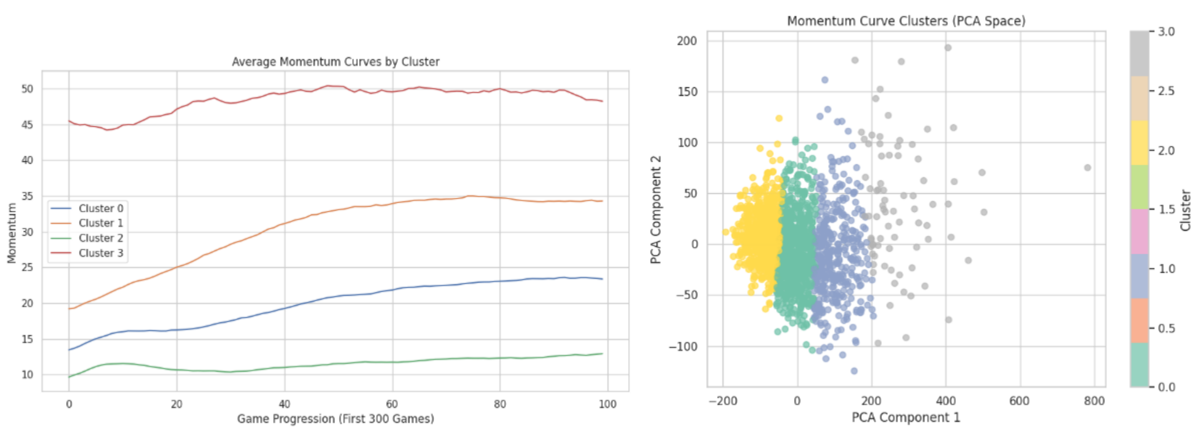

Figure 4: Plot of average career momentum within clusters and decision boundaries based on PCA components.

The benefit of momentum clustering is that it allows for comparison across different eras, something that has been difficult to do from an objective, data-driven standpoint. We can see that any player that is part of Cluster 3 is essentially an outlier in the projected 2D principal component space. Although claims of inflated statistics in the modern era may have some merit, members of Cluster 3 range from modern, durable stars like Shaquille O'Neal, Tracy McGrady, LeBron James, and Chris Paul, to pioneers of the game from the mid to late twentieth century such as Bill Walton, George McGinnis, Dave Cowens, and Wilt Chamberlain. Despite inconsistencies in output throughout different stages of the NBA, this gives us a way to compare performance trends throughout history. Because we ran k-means clustering on large vectors containing the first 300 games, we can visualize the 'average trajectory' of players in a cluster. We can also make inferences about career trajectory.

For example, we see that on average in cluster 3, star players will have a small but pronounced dip at the beginning of their career before figuring things out. Cluster one players, on the other hand, see a fast ascension to contributions but are never able to take the next step to stardom.

7 Conclusion

In conclusion, this study introduces a novel momentum metric to analyze NBA players' careers, focusing on sustained performance over time. Through careful data cleaning and interpolation of momentum scores, we ensured that the analysis was based on reliable and comparable data. The optimization of key parameters, such as the smoothing factor and team context weight, revealed that individual performance trends were the most significant predictors of future performance, while team success did not enhance short-term performance prediction.

The clustering of players based on their momentum trajectories provided valuable insights into career progression, grouping players across eras by the shape of their performance curves. This approach allows for meaningful comparisons of player careers, independent of historical context. In the future, it would be interesting to analyze the curvature of the momentum curves, as well as any insights that can be drawn from the gradient itself. Overall, this work offers a quantitative framework for evaluating momentum in player performance, laying the groundwork for future studies on career trajectories and player evaluation.

References

- Arkes, J., and Martinez, J.A. (2011). Finally, evidence for a momentum effect in the NBA. *Journal of Quantitative Analysis in Sports*, 7(3), Article 10.
- Gilovich, T., Vallone, R., and Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17(3), 295–314.
- Miller, J.B., and Sanjurjo, A. (2018). Surprised by the gambler's and hot hand fallacies? A truth in the law of small numbers. *Econometrica*, 86(6), 2019–2047.