SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(1), 1083.
https://doi.org/10.62617/mcb1083

Article

Data-driven insights into basketball performance: Unveiling the impact of advanced analytics on player and team efficiency

Zeyu Liang[1,*], Yucai Gao[2], Jiuyuan Wang[3], Zhihao Liu[4,*]

[1] School of Chinese Basketball, Beijing Sport University, Beijing 100084, China
[2] School of Physical Education, Shandong Normal University, Jinan 250061, China
[3] School of Exercise and Health Sciences, Beijing Sport University, Beijing 100084, China
[4] Kyungil University Graduate School, Gyeongbuk 38428, Korea
*** Corresponding authors:** Zeyu Liang, zeyu.liang@bsu.edu.cn; Zhihao Liu, 15238466466@163.com

CITATION

Liang Z. Gao Y, Wang J, Liu Z. Data-driven insights into basketball performance: Unveiling the impact of advanced analytics on player and team efficiency. Molecular & Cellular Biomechanics. 2025; 22(1): 1083. https://doi.org/10.62617/mcb1083

ARTICLE INFO

Received: 11 December 2024
Accepted: 20 December 2024
Available online: 7 January 2025

Abstract: This study investigates the impact of data science on basketball performance, comparing key performance indicators (KPIs) across NCAA Division I collegiate basketball and NBA games. Using a dataset of 180 games over three seasons, the study examines metrics such as Player Efficiency Rating (PER), True Shooting Percentage (TS%), and Defensive Rating (DRtg). Machine learning algorithms, including logistic regression, decision trees, and support vector machines, were employed to predict game outcomes and evaluate the relationships between KPIs and team success. The results reveal that in collegiate basketball, elevated shooting accuracy (TS%) and defensive metrics (DRtg) are strong predictors of success, while in the NBA, PER plays a more significant role. The findings highlight the importance of integrating data-driven insights into coaching strategies and performance enhancement, with practical recommendations for teams at both competitive levels. This study fills a gap in the literature by offering a comparative analysis of basketball KPI usage in different competitive environments.

Keywords: basketball performance; data science; key performance indicators; machine learning algorithms

1. Introduction

Performance Analysis (PA) has become an important factor in professional sports that provides benefits regarding coaching, training, and the selection of the players [1]. In the management of sports, the use of advanced data analysis helps the coaches in decision-making and enhances several tactical strategies and the overall performance of games [2]. In basketball, there is another tool that has been considered very useful for the evaluation of efficiency and decision-making, namely Key Performance Indicators, which are quantitative values that relate to different characteristics of the players and the team [3].

As pointed out in the current literature on the subject, the application of KPIs is crucial in predicting match outcomes and in improving the performance of the teams [4]. For example, PER and T's percentages have become measures of players' and teams' efficiency in men's basketball [5]. PER combines different statistical data into one to predict the average productivity of a player and TS% considers the value of shots in shooting accuracy [6]. The following KPIs provide important figures that could potentially have a significant influence on the understanding of players and team performances and the planning of coaching and the games in general [7].

There is also the fact that while basketball analytics has advanced even more, there is still a considerable shortage of scientific sources regarding the systematic application of KPIs at various levels [8,9]. Prior studies on KPIs in sports have been conducted on rugby and soccer and have explained KPIs as tackles, possession, and set-pieces [10]; however, there is very limited literature on basketball analysis.Several studies argue that most of the existing research is conducted at the elite professional level, highlighting a gap in information regarding its application in collegiate and lower-tier leagues [11]. therefore, there is a dearth of information on its use in collegiate and below leagues. Subsequently, the use of KPI has led to inadequacy in understanding available results related to basketball performance, especially at lower league levels. Yet, understanding basketball performance indicators is necessary for the development of measures to improve individual and team performance in the various KPIs.

To address this gap, this study employs a mixed-method research design to evaluate the advantages of performance analysis in basketball. The purpose of this study is to determine which of the college performance indicators and NBA games performance indicators are more related to and can influence changes in performance and outcomes [12]. Prior research has identified certain factors such as shooting accuracy and defense as having important roles to play in the team's performance [13]. However, it is crucial to assess the effectiveness of these KPIs in other settings and competitive tiers.

Recent researchers have also proved that the SAP application does not remain invariant with references to context factors including the type of the team, the competition, or the strategy [14]. For example, a survey on NBA franchises found that the integration of analyses into game plans can lead to substantial performance improvement [15]. However, there is a visible paucity of research on similar applications in collegiate basketball thus pointing to a potential area for future research.

Therefore, the following paper aims to provide a systematic experimental analysis that examines how data-driven performance indices could enhance basketball performances at collegiate and professional levels. It is designed to give useful and informative recommendations through KPIs that may assist coaches and teams in general in improving their strategies and results [16]. Besides, it removes the gap in performance analytics in basketball between the research findings and practice.

2. Methods

2.1. Design & sample

The study focuses on NCAA Division I and NBA games because these levels represent two distinct competitive environments with unique styles of play and varying levels of player skill. The KPIs selected—PER, TS%, and DRtg—are well-established metrics in basketball performance analysis. PER is a composite metric that captures a player's overall contributions, TS% accounts for scoring efficiency, and DRtg assesses a player's defensive impact. These KPIs provide a comprehensive picture of individual and team performance.

- Total Games Analyzed: 180 games from the 2021/2022, 2022/2023 and/or 2023/2024 seasons.

- March Madness: NCAA Division I Men's Basketball Tournament:
    a. Number of Games: The total retrospective of all the games ever created is 90.
    b. Coverage: They have thirty games from each season's tournament. This includes games in the preliminary stage as well as the quarter-final, the semifinal and the final match hence covering all levels of competition and the level of difficulty in the tournament.
    c. Teams Represented: It has games of 45 different collegiate teams in total. It provides a general picture of their performance in various institutions and in various competitive conditions.
- NBA Regular Season:
    a. Number of Games: The total retrospective of all the games ever created is 90.
    b. Coverage: The regular season of each season consists of 8 games from each of the ten different countries. This is done to get a random sample of the various teams in the leagues so that no category is favored over another. This approach considers the strengths of the two teams, the strategies that are likely to be employed, and the environment of the match.
    c. Teams Represented: The sample includes 90 different NBA teams, which provides an overall picture of the professional team processes and results during the regular season.

Sampling Rationale:

- Enhanced Robustness: This improvement is because instead of examining one season's performance in isolation, the study assists in establishing the variations in performance in three successive seasons. It is a more holistic approach to analyzing performance indicators and a team's interactions throughout the season.
- Diverse Competitive Situations: Thus, the tournament and other games, in which opponents were eager to win, show all the competitive factors and their impact on the efficiency indicators using a large number of different games. The NCAA tournaments are single elimination hence, they bring different forms of pressure than the NBA's long, regular season.
- Comprehensive Team Analysis: As can be observed from the above, the study will be able to capture trends and patterns that would have otherwise gone unnoticed due to the large number of games and teams. This affords a broader perspective on the evaluation of the teams and individuals in the various categories or levels of the competition.

Data Collection and Sampling:

- NCAA Division I Men's Basketball Tournament (90 games): It involves the collection of data from various levels of competition from elimination rounds to the final bouts. This breadth allows an investigation of performance in both significant and insignificant / less significant games.
- NBA Regular Season (90 games): These criteria assist in selecting these games as they ensure a wide coverage of the various performances of the team – both at home and when playing away and against different opponents.

Analytical Tools:

- Video Analysis Software: The action was described by using Synergy Sports Technology (version 3.1) to identify the detailed actions of the individual players and the teams. It also enables the assessment of the performance in terms of scoring, defensive actions, and turnovers among other activities in the game.

Because of the large sample data collected over three seasons, the study can offer crucial information on changes in basketball performance at the college and professional levels and how the competition environment affects team and individual performances.

2.1.1. Data processing and integration of artificial intelligence

1) Data Preprocessing: Data cleaning and structural cleaning were also conducted on the data obtained from Synergy Sports Technology. It involved removing data that was conflicting, redundant, insufficient, or perhaps incomplete, or restructuring the data in a way that made it coherent.
2) Feature Extraction: The KPIs were calculated from the data obtained from the field studies. These KPIs were specific to basketball and included Per, TS%, DRtg, and so on—which was used as an index of team or player performance.
3) Data Normalization: To prevent differences in scaling and units in the case of joined variables, the data were normalized. In this step of the analysis, the values of KPIs were scaled to a predetermined range to have a better fit for further procedures.
4) Machine Learning Algorithms: The choice of logistic regression, decision trees, and support vector machines (SVM) was motivated by their proven effectiveness in handling classification problems, which is the primary focus of this study—predicting game outcomes based on performance metrics. Logistic regression is well-suited for binary classification, decision trees provide easily interpretable models, and SVM excels in high-dimensional spaces, making it ideal for capturing complex relationships between KPIs.
    - Algorithm Selection: The models and the predictions of the games were created by using several algorithms of machine learning. These were machine learning algorithms such as the logistic regression model, the decision tree model, and support vector machines.
    - Training and Validation: The dataset was divided into training and validation sets. The training set was used in building the predictive models while the validation set was used for testing and model tuning. To improve the generalization of the results and avoid overfitting, cross-validation techniques were applied.
    - Model Evaluation: The evaluation of the model's performance was conducted with the help of general characteristic parameters such as accuracy, precision, recall, and F1 measure. From these, the KPIs and models that will be used in the evaluation of performances of games were determined in this evaluation.
5) Discriminant Analysis: This was done in a bid to identify the KPIs that were useful in making the distinction between the performing and the non-performing teams. The values obtained from discriminant analysis supported the hypothesis of the existence of the correlation between teams and KPIs with regard to their

importance in demonstrating discriminant validity and the role of one KPI in the discrimination of another KPI.

To these ends, the study will employ advanced statistical and machine learning techniques, focusing on a large dataset from three seasons, and aim to compare the collegiate and professional levels of basketball in terms of performance metrics across different competitions and leagues.

2.1.2. Research protocol and procedure

This research is based on a structured research protocol, that focuses on quantitative data analysis using advanced analytics tools. The research process is made up of three main stages:

1) Data Collection:
    - Data were collected using Synergy Sports Technology (version 3.1), enabling the collection of detailed in-game action records like shots, rebounds, turnovers, and fouls.
    - A basketball analyst with over seven years of experience manually reviewed the games and collected data to ensure data accuracy and consistency.
    - 10% of the games were tested for inter-rater reliability, which yielded excellent agreement (ICC > 0.80) on all KPIs.
2) Data Processing:
    - Preprocessing of the raw data was conducted to remove incomplete or redundant entries and organized using Microsoft Excel after successful transfer.
    - All KPIs were scaled to the same range of data normalization, hence improving the comparability of different variables.
    - SPSS for statistical analysis was then conducted for the processed data, to ensure a rigorous performance evaluation methodology.
3) Feature Extraction and KPI Analysis:
    - The calculation of specific KPIs (such as PER, TS%, and DRtg) was conducted to represent key metrics in basketball performance analysis.
    - The evaluation of these KPIs was aimed at determining their relevance and predictive power in distinguishing between winning and non-winning teams.

2.2. Performance indicators

From the literature review and consultation with the experts in the field, the available quantitative measures were used to define the KPIs. The KPIs separated individual and team performance, including offense and defense layers. As presented in **Table 1**, the KPIs included:

Table 1. KPIs definitions.

KPI	**Definition**
Player Efficiency Rating (PER)	Composite metric reflecting overall player efficiency by integrating various statistics.
True Shooting Percentage (TS%)	Measures shooting efficiency, including field goals, three-pointers, and free throws.
Defensive Rating (DRtg)	Estimated points allowed per 100 possessions, assessing defensive impact.
Defensive Win Shares (DWS)	Contribution to team defense in terms of wins.

Table 1. (*Continued*).

KPI	Definition
Total Rebounds	The sum of both offensive and defensive rebounds.
Turnovers	A number of times the ball is lost to the opposing team due to errors.
Assist-to-Turnover Ratio	The ratio of assists to turnovers reflects ball distribution efficiency.
Total Possession Time	The total duration of ball control by a team during a game.
Field Goal Percentage (FG%)	Percentage of successful field goals out of total attempts.
Three-Point Shooting Percentage (3P%)	Percentage of successful three-point shots out of total attempts.

2.3. Procedure

Data Collection: The data collection process was administered by a competent basketball analyst with over seven years of working experience. When performing the match analysis, the analyst used Synergy Sports Technology to input and record various in-game actions such as shots made, rebounds, fouls committed, and turnovers. This was done in the same manner for each game to ensure that nothing was missed in the data collection process.

Data Processing: The results were then exported from SPSS and transferred to Microsoft Excel for further analysis. To enhance the reliability of the findings, inter-rater reliability checks were conducted on 10% of the games selected randomly. The ICC was employed to assess inter-rater reliability, and the findings revealed that all the KPIs exhibited excellent reliability with a coefficient of more than 0.80.

Contextual Analysis: To evaluate the contextual factors, a two-step cluster analysis was employed to categorize the contexts based on the success probability and intensity. The analysis made a clear distinction between the games that were very sensitive like the playoff games and other games, thus making a clear comparison of the performance of the team during regular episodes of the competition.

2.4. Statistical analysis

Data Normality: The assumption of normality was checked by the Shapiro-Wilk test and the calculated p-value was greater than 0.05. This was useful in determining the next statistical tests that were relevant for the data distribution.

Descriptive Statistics: Descriptive statistics in the form of mean and standard deviations were calculated for all the KPIs that were discussed in the current study for both collegiate and NBA and high stakes and regular-season games. Such statistics provided an overall view of the performance and any changes or fluctuations that may have happened during the period.

Inferential Statistics: To compare the score values between the winners and the losers, the independent samples t-test was used. This was done by discriminant analysis to determine which of the KPIs was most appropriate in discriminating between the winning and the non-winning team. The significance of the Structural coefficients (SC) was tested using the criterion of ≥ 0.30.

Model Validation: The discriminant models were determined using the leave-one-out cross-validation technique. To check model fit and reliability, the percentage of correct classification of the data points was also calculated.

Statistical Software: All statistical tests were conducted using the IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY) and the level of significance was set at $p < 0.05$.

2.5. Qualitative research

2.5.1. Purpose of qualitative research

In the qualitative research, we explore the experiences, perspectives, and insights of key stakeholders in basketball, such as coaches, players, and performance analysts. This will help in gathering rich, contextual information on the impact of KPIs on real-life decision-making processes, and the influence of data-driven insights on strategic planning, player development, and in-game adjustments.

2.5.2. Data collection

The data collection process will involve Semi-structured interviews with each participant. The interview guide will include open-ended questions to explore:

- How coaches, players, and analysts understand and interpret KPIs (e.g., Player Efficiency Rating, True Shooting Percentage, etc.).
- The perceived influence of KPIs on game preparation, training adjustments, and in-game strategies.
- Any challenges or limitations faced in the practical use of basketball KPIs.
- The extent to which qualitative insights (e.g., player motivation, team chemistry) complement or contradict KPI data.

The interviews will last between 30-45 minutes and will be audio-recorded with participant consent. The interviews will be transcribed verbatim for subsequent analysis.

2.5.3. Data analysis

The interview transcripts will be analyzed using thematic analysis. The method enables the identification of recurring themes and patterns within the data. The analysis will follow these steps:

1) Familiarization with the data: Reading and re-reading the transcripts to become deeply familiar with the content.
2) Coding: Assigning labels or codes to significant phrases, sentences, or ideas that relate to the use of KPIs in basketball performance.
3) Theme development: Grouping similar codes into broader themes such as "Impact of KPIs on Strategic Decision-Making", "Challenges of KPI Implementation", and "Integration of Qualitative and Quantitative Data".
4) Review and refinement: Refining the themes to ensure they accurately reflect the participants' experiences and insights.

2.5.4. Trustworthiness and rigor

The following strategies will be employed to ensure the rigor of the qualitative findings:

- Triangulation: Comparing insights from coaches, players, and analysts to identify consistent themes or diverging perspectives.

- Member checking: Participants will be given the opportunity to review and confirm the accuracy of their interview transcripts and the interpretation of key themes.
- Peer debriefing: Colleagues familiar with basketball analytics will review the findings to provide feedback and validate the interpretation of the results.

2.6. Data processing calculation methods

This section details the calculation processes for the key performance indicators (KPIs) used in the analysis. These formulas represent the underlying computational methods used to derive insights from the data collected from basketball games at both the collegiate (NCAA Division I) and professional (NBA) levels.

Player Efficiency Rating (PER)

The formula for PER is complex but can be simplified as follows:

PER = (uPER/league pace) × (league average PER//team minutes) × (adjusted player minutes/game minutes)

Where:

- uPER is the unadjusted player efficiency rating (calculated using individual stats such as points, rebounds, assists, steals, blocks, and turnovers).
- league pace adjusts for the pace at which the team plays.
- team minutes refers to the total minutes played by the team.
- game minutes are the total minutes played by a team in the game.

True Shooting Percentage (TS%)

The formula is as follows:

TS% = Points Scored/2 × (Field Goal Attempts + 0.44 × Free Throw Attempts)

This formula adjusts shooting efficiency to consider the differing values of two-point shots, three-point shots, and free throws.

Defensive Rating (DRtg)

The simplified formula is:

DRtg = 100 × (Points Allowed/Possessions Faced)

Where:

Points Allowed refers to the total points conceded while the player is on the court.

Possessions Faced is the number of defensive possessions during which the player/team is on the court.

Defensive Win Shares (DWS)

The formula for DWS is derived as:

DWS = Team Defensive Rating/(Points Allowed) × Player Minutes Played

Assist-to-Turnover Ratio

The formula is:

Assist-to-Turnover Ratio = Total Assists/Total Turnovers

Field Goal Percentage (FG%)

The formula is:

FG% = (Field Goals Made/Field Goals Attempted) × 100

Data Normalization

The formula for normalization is:

$X_{norm} = X - X_{min}/X_{max} - X_{min}$

Where:

- XXX is the original value of the KPI.
- Xmin and Xmax are the minimum and maximum values of the KPI, respectively.
- Xnorm is the normalized value of the KPI.

Machine learning algorithm calculations

For example, the logistic regression model is calculated as follows:

$P(y = 1|X) = 1/1 + e - (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n)$

Where:

$P(y = 1|X)$ is the probability of a team winning the game based on the KPIs X_1, X_2, …, X_n.

β_0, β_1, …, β_n are the coefficients representing the influence of each KPI.

3. Results

This section presents the results of the research based on the provided performance evaluation of collegiate and NBA basketball games and the use of KPIs in increasing the degree of differentiation between the success of the teams. The tables with the comparison of different KPIs and their discriminant features support the presented results.

3.1. Descriptive statistics and univariate analysis are used to give a general idea about the collected data and present it in tabular or graphical form

3.1.1. Below are the formulas used to derive

Table 2 presents the descriptive of KPIs for winners and losers throughout all the observed games. The descriptive analysis at the event level reveals several metrics that are statistically distinct between the two groups.

Table 2. The Average and variability of KPIs—and the uni-variate differences of the two teams that won and lost in collegiate and NBA matches.

Performance Indicator	Winning Teams (Mean ± SD)	Losing Teams (Mean ± SD)	*p*-value
Player Efficiency Rating (PER)	21.35 ± 4.76	18.47 ± 5.32	< 0.001#
True Shooting Percentage (TS%)	54.12% ± 7.65	48.23% ± 8.21	< 0.001#
Defensive Rating (DRtg)	104.76 ± 6.12	110.52 ± 7.84	< 0.001#
Defensive Win Shares (DWS)	2.36 ± 0.87	1.78 ± 0.92	0.002**
Total Rebounds	44.12 ± 5.84	39.45 ± 6.91	< 0.001#
Turnovers	12.34 ± 4.22	16.45 ± 4.76	< 0.001#
Assist-to-Turnover Ratio	1.95 ± 0.28	1.60 ± 0.31	< 0.001#
Total Possession Time (mins)	23.47 ± 3.76	21.32 ± 4.12	0.008**
Field Goal Percentage (FG%)	46.58% ± 5.92	42.45% ± 6.81	< 0.01*
Three-Point Shooting Percentage (3P%)	37.22% ± 7.05	31.98% ± 7.90	< 0.001#

*Significance codes: $P < 0.05$; $P < 0.01$; $P < 0.001$.

Cohesion was found to be positively correlated with Skype and higher values of numerous KPIs such as PER, TS%, and TRB were characteristic of the winning teams.

On the other hand, teams that the players of which lost a particular game had higher turnover rates than assist rates, as depicted by their lower Assist/Turnover ratios.

3.2. Contextual analysis: High stake games as a comparison to the normal regular games

Table 3 outlines indices for performance evaluation of the high stakes and regular season games as per the winning and the losing teams.

Table 3. The analysis that has been done includes using mean and standard deviations on KPIs and univariate analysis of differences between the winning and losing teams in high-stakes and regular-season games.

Performance Indicator	High-Stakes Games (Winning)	High-Stakes Games (Losing)	*p*-value	Regular-Season Games (Winning)	Regular-Season Games (Losing)	*p*-value
Player Efficiency Rating (PER)	22.45 ± 5.10	19.32 ± 5.45	< 0.001#	20.25 ± 4.30	17.62 ± 5.15	< 0.01*
True Shooting Percentage (TS%)	56.35% ± 8.12	49.40% ± 8.32	< 0.001#	53.10% ± 7.02	46.10% ± 7.85	< 0.01*
Defensive Rating (DRtg)	102.12 ± 6.45	115.34 ± 7.65	< 0.001#	106.98 ± 5.90	108.45 ± 6.22	0.027*
Defensive Win Shares (DWS)	2.55 ± 0.90	1.65 ± 0.87	< 0.001#	2.15 ± 0.85	1.85 ± 0.96	0.041*
Total Rebounds	46.23 ± 6.12	37.76 ± 7.22	< 0.001#	43.65 ± 5.45	41.12 ± 6.95	0.055
Turnovers	11.85 ± 4.12	17.22 ± 4.56	< 0.001#	13.40 ± 4.32	15.65 ± 4.82	0.041*
Assist-to-Turnover Ratio	2.10 ± 0.30	1.50 ± 0.29	< 0.001#	1.90 ± 0.27	1.65 ± 0.35	0.002**
Total Possession Time (mins)	25.12 ± 4.12	22.45 ± 4.30	0.024*	22.23 ± 3.65	20.45 ± 4.05	0.032*
Field Goal Percentage (FG%)	48.45% ± 6.78	42.78% ± 7.45	< 0.01*	44.32% ± 6.00	41.20% ± 6.95	0.031*
Three-Point Shooting Percentage (3P%)	39.10% ± 7.15	32.00% ± 8.05	< 0.001#	35.45% ± 6.75	30.80% ± 7.45	0.015*

*Significance codes: $P < 0.05$; $P < 0.01$; $P < 0.001$.

In the critical match, the winners had higher values of the PER, TS%, and DWS. These differences persisted but were somewhat less pronounced in the regular-season games.

3.3. Discriminant analysis

The discriminant analysis was then used to identify which of the KPIs were discriminant between the winning and the losing teams. These results are shown in the form of structure coefficients and cross-validation percentages in **Table 4** below.

Table 4. Discriminant analysis structure coefficients of kpis of winning and losing team.

Performance Indicator	All Games	High-Stakes Games	Regular-Season Games
Player Efficiency Rating (PER)	0.42*	0.45*	0.35*
Actual Shooting Percentage (TS%)	0.48*	0.52*	0.44*
Defensive Rating (DRtg)	−0.39*	−0.43*	−0.36*
Defensive Win Shares (DWS)	0.37*	0.38*	0.35*
Total Rebounds	0.31*	0.30*	0.28*
Turnovers	−0.31*	−0.33*	−0.29*

Table 4. (*Continued*).

Performance Indicator	All Games	High-Stakes Games	Regular-Season Games
Assist-to-Turnover Ratio	0.34*	0.32*	0.36*
Total Possession Time (mins)	0.25*	0.27*	0.22*
Field Goal Percentage (FG%)	0.33*	0.31*	0.30*
Three-Point Shooting Percentage (3P%)	0.29*	0.28*	0.26*

For discriminant value, structure coefficients should equal or exceed *0.30*.

Based on the discriminant analysis, it was found that TS%, PER, and DRtg were the most important KPIs that were most suitable for identifying the winners from the losers. In playoffs, new regularity was established, shooting efficiency and defensive trends were strong while in regular-season games, shooting efficiency and rebound trends were strong.

3.4. Cross-validation and model performance

Cross-validation results confirm the robustness of the KPIs in predicting game outcomes:

- All Games: 82.4%;
- High-Stakes Games: 85.1%;
- Regular-Season Games: 78.9%.

These high accuracy rates therefore demonstrate the effectiveness of the selected KPIs in categorizing the winning and the losing teams regardless of the stipulated game conditions.

4. Discussion

The objective of this study was to identify possible KPIs that are most suitable in identifying the winners from the losers within the collegiate and NBA basketball games. In the framework of the cup and league matches the study was designed to find out the set of KPIs affecting the play-off option among all the factors most significant for the match result. The findings suggest that winners are better than losers in some aspects that are beneficial for coaches and analysts to develop strategies for the game.

4.1. KPI and games result in performance indicators and match result expected results expected outcome 1

Thus, based on the objective data of the study carried out, it can be stated that the analyzed teams have higher effective r performance in the context of the selected KPIs [17]. Specifically, the winning teams had higher Player Efficiency Ratings (PER) and True Shooting Percentages (TS%), which indicates the ability of individual players and shooting efficiency to decide the game's outcome[18]. These findings are in line with previous research on the correlation between high players' efficiency and favorable outcomes [19]. The average PER of players and TS%, which reflects the efficiency of different types of field goals, were overall higher in the winners which confirms the importance of the discussed parameters in the game.

Apart from shooting efficiency, the winners had better Defence, with lower DRtg and much higher DWS. This means that strong defense has been found to be a significant factor in winning since it reduces the opponent's chances of getting a goal and improves the team's performance [20]. Lower DRtg indicates that teams with better efficiency are stopping offenses of their opponents and higher DWS means those teams are contributing to defensive stops and rebounds to reduce an opportunity for an opponent to score the ball.

Turnovers and Assist-to-Turnover Ratios (AST/TO) also emerged as another significant discriminant factor. The winning teams committed fewer turnovers and had higher AST/TO ratios, which indicates that possession and ball handling are critical in determining the outcome of games [21]. This is in agreement with the argument made earlier in this paper that teams that win the space and can convert possessions into shot attempts to gain an advantage [22]. Less turnovers imply fewer opportunities for the opponent team to score goals and a high AST/TO ratio depicts better co-ordination and decisions made in the field of play of the team [23].

Free Throw Rate (FT Rate) supported the hypothesis while Field Goal Percentage (FG%) and Three-Point Shooting Percentage (3P%) provided support to the hypothesis that shooting efficiency determined the probability of winning a game. The importance of shooting accurately was also emphasized in recent studies by [24], as shooting is the basis of scoring and winning games for a particular team. Field goals and 3-pointers indicate that the teams are able to put pressure on the opposite team and capitalize on the chances that are available in the offense.

4.2. Contextual analysis: Superbowl games as a comparison to the normal regular games

Some differences were defined by the comparison of KPI indicators of high-stakes and regular-season games. The index of higher and better comparative values of PER, TS%, and DWS of the losing teams was more inflated in high-stakes games while the successful teams had better statistical performances. This goes a long way in suggesting that performance metrics are emphasized when under pressure since all aspects of team play are amplified [25]. Therefore, games of high stakes, where the change in the patterns of games is noted, reveal the significance of the performance of certain players and defense as factors that define the outcome of the game.

Yes, in the regular-season games, the KPIs were also there but the difference between the winning team and the losing team was not that much [26]. This means that while such performance indicators are relevant irrespective of other games in the corresponding season, their impact could be less certain in those fairly important games [27]. Since the regular-season games include games of different levels and significance, there would be significant variations in the KPI impact compared to what is observed in the high-stakes games.

4.3. Discriminant analysis and cross validation

The discriminant analysis data also aligned with the intended conclusion by indicating that the KPIs with the highest discriminant coefficients were TS%, PER, and DRtg. The application of these metrics was also established to possess good levels of predictive validity for game outcomes and hence, the validity of their application in real-life practices of coaching and performance evaluation [28]. The high discriminant power of these KPIs proved that they are valuable as measures to differentiate between the winning and the losing team in terms of strategic planning and the evaluation of games.

Cross-validation analysis revealed the high accuracy of the model especially in high-risk games. The overall accuracy rate for all games was 82.4%, and for high-stake games, the accuracy rate was slightly higher and was 85.1% This fact proves the effectiveness of the selected KPIs and their ability to differentiate between winners and losers [29]. This confirms the relevance of these KPIs as valuable tools for evaluating and improving the performance of the team. The high degree of accuracy also suggests that these KPIs are not only beneficial but are also essential in predicting game results and coaching decisions.

4.4. Conclusion and recommendations

It would be possible to use some of the findings from this study in the training of coaches and in performance evaluations. He should improve on his player's shooting accuracy, defense, and ball possession for a team to score or for the team not to concede. These KPIs should be incorporated into the training programs as well as the strategies developed in the context of the game to improve the performance of the team and get better outcomes [30]. For instance, shooting drills and defensive postures, and avoiding 'turnovers' would be beneficial in improving the performance of the team.

However, there are some of the limitations that have contributed to the completion of this study as follows: They only watched collegiate and NBA basketball games; therefore the results of the study cannot be generalized to other basketball leagues or any other level of basketball. However, the study did not take into account other factors such as social relations with other team members and one's health state that may impact the performance metrics [31]. Further studies appear to confirm the relevance of the pointed KPIs to other leagues and levels of competition. However, it is also possible that expanding the range of variables, such as team relations, players' condition, and the opponent team factors, might provide a better understanding of efficiency indicators in basketball [32].

5. Conclusion and recommendations

5.1. Summary of findings

This study contributes to the advancement of performance analytics in basketball by highlighting the significance of metrics such as FG%, TS%, and defensive statistics in enhancing player and team performances. The analysis demonstrates that winning teams exhibit higher shooting efficiency rates, defensive effectiveness indices, and ball control indices compared to losing teams. Through detailed case studies, it becomes evident how leveraging these analytics can lead to the formulation of superior game strategies and improved performance, ultimately translating into victories on the court.

5.2. Implementing recommendations for enhanced performance

Building on these findings, it is imperative for training programs, strategic planning, and decision-making processes to integrate analytical insights from both teams and coaches. The following recommendations are proposed for practical implementation:

1) Emphasize Shooting Efficiency: Coaches should prioritize drills and strategies aimed at improving FG% and TS%. This can be achieved through targeted shooting drills, player analysis to identify shooting prospects, and enhancing shot precision within key areas of the court.
2) Enhance Defensive Capabilities: Discussions should focus on improving Defensive Rating (DRtg) and Defensive Win Shares (DWS). Coaches can implement defensive strategies that reduce opponents' shooting percentages, increase defensive rebounds, and enhance defensive stops.
3) Optimize Ball Management: Teams should concentrate on reducing turnovers, with a key metric being the Assist-to-Turnover Ratio (AST/TO). Training sessions should incorporate activities that enhance ball handling skills, decision-making under pressure, and creating scoring opportunities for the team.
4) Leverage Analytical Tools: Utilize analytical tools to monitor player performance and overall gameplay trends. Video analysis and data visualization tools, including performance indices, enable a deeper understanding of performance patterns, facilitating informed decision-making to enhance team performance.
5) Integrate Real-Time Data: Coaches and teams can leverage real-time data analytics during games to receive immediate feedback and make strategic adjustments on the fly. Enhancing data collection and utilization during games can be achieved through automated data collection technologies.

5.3. Future research directions

Future research in basketball analytics should explore the following key areas to advance the application of analytics in the sport:

1) Real-Time Data Analytics and AI: Investigate the application of real-time data analytics and artificial intelligence in refining game strategies, potentially providing alerts for in-game modifications based on dynamic factors.
2) Longitudinal Studies: Conduct post-event surveys to assess the long-term effects of performance analytics on teams and players across multiple seasons, uncovering patterns and career impacts.
3) Cross-League and International Comparisons: Extend research to different leagues and countries to understand how performance indicators vary across varying competition levels and cultures, offering insights for broader application in basketball platforms.
4) Impact of Player Health and Team Dynamics: Analyze the relationship between player health and team cohesion using analytical statistics, providing insights into performance enhancement and necessary adjustments for personal and team development.

5) Integration of Qualitative Data: Combine quantitative data with qualitative insights from players, coaches, officials, and stakeholders to enrich the understanding of performance metrics and their implications on game outcomes and team dynamics.

By advancing research in these areas, basketball analytics can continue to evolve, providing teams and coaches with better tools and actionable insights for improved performance outcomes.

Author contributions: Conceptualization, ZL (Zeyu Liang) and YG; methodology, ZL (Zeyu Liang); software, ZL (Zeyu Liang); validation, ZL (Zeyu Liang), YG, JW and ZL (Zhihao Liu); data curation, ZL (Zhihao Liu); writing—original draft preparation, ZL (Zeyu Liang); writing—review and editing, YG; visualization, JW; supervision, ZL (Zeyu Liang); project administration, ZL (Zeyu Liang); funding acquisition, YG. All authors have read and agreed to the published version of the manuscript.

Ethical approval: Not applicable.

Declaration of competing interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability: The datasets generated during and analyzed during the current study are available from the corresponding author upon reasonable request.

Conflict of interest: The authors declare no conflict of interest.

References

1. K. Mikołajec, A. Maszczyk, and T. Zając, 'Game Indicators Determining Sports Performance in the NBA', J Hum Kinet, vol. 37, pp. 145–151, 2013, doi: 10.2478/hukin-2013-0035.
2. F. W. Otte, K. Davids, S.-K. Millar, and S. Klatt, 'Specialist role coaching and skill training periodisation: A football goalkeeping case study', International Journal of Sports Science & Coaching, vol. 15, no. 4, pp. 562–575, Aug. 2020, doi: 10.1177/1747954120922548.
3. T. J. Mosey and L. J. G. Mitchell, 'Key performance indicators in Australian sub-elite rugby union', Journal of Science and Medicine in Sport, vol. 23, no. 1, pp. 35–40, Jan. 2020, doi: 10.1016/j.jsams.2019.08.014.
4. J. O. C. Coyne, A. J. Coutts, R. U. Newton, and G. Gregory Haff, 'Relationships Between Different Internal and External Training Load Variables and Elite International Women's Basketball Performance', Int J Sports Physiol Perform, vol. 16, no. 6, pp. 871–880, Jun. 2021, doi: 10.1123/ijspp.2020-0495.
5. Y. Wu et al., 'OBTracker: Visual Analytics of Off-ball Movements in Basketball', IEEE Trans. Visual. Comput. Graphics, pp. 1–11, 2022, doi: 10.1109/TVCG.2022.3209373.
6. Y. Fu and J. Stasko, 'HoopInSight: Analyzing and Comparing Basketball Shooting Performance Through Visualization', IEEE Trans. Visual. Comput. Graphics, pp. 1–11, 2023, doi: 10.1109/TVCG.2023.3326910.
7. S. C. Canuto and M. Bezerra De Almeida, 'Determinants of Basketball Match Outcome Based on Game-related Statistics: A Systematic Review and Meta-Analysis', EJHM, vol. 48, pp. 4–20, 2022, doi: 10.21134/eurjhm.2022.48.2.
8. Y. Li, S. Geng, L. Wang, and F. Li, 'A novel robust data envelopment analysis with asymmetric uncertainty and an application to National Basketball Association', Expert Systems with Applications, vol. 239, p. 122341, Apr. 2024, doi: 10.1016/j.eswa.2023.122341.
9. B. Ma, Y. Wang, and Z. Li, 'Application of data mining in basketball statistics', Applied Mathematics and Nonlinear Sciences, vol. 8, no. 1, pp. 2179–2188, Jan. 2023, doi: 10.2478/amns.2021.2.00182.

10. K. McGuigan and K. Collins, 'Understanding the impact of pitch location on shot outcome in Gaelic football – where is the scoring zone?', International Journal of Performance Analysis in Sport, vol. 21, no. 4, pp. 491–506, Jul. 2021, doi: 10.1080/24748668.2021.1917978.
11. R. Mandić, S. Jakovljević, F. Erčulj, and E. Štrumbelj, 'Trends in NBA and Euroleague basketball: Analysis and comparison of statistical data from 2000 to 2017', PLoS One, vol. 14, no. 10, p. e0223524, 2019, doi: 10.1371/journal.pone.0223524.
12. A. Fox, M. Manisera, M. Sandri, and P. Zuccolotto, 'Analyzing Basketball Data with BasketballAnalyzeR', CHANCE, vol. 35, no. 3, pp. 42–56, Jul. 2022, doi: 10.1080/09332480.2022.2123161.
13. M. Rogers, A. J. Crozier, N. K. Schranz, R. G. Eston, and G. R. Tomkinson, 'Player Profiling and Monitoring in Basketball: A Delphi Study of the Most Important Non-Game Performance Indicators from the Perspective of Elite Athlete Coaches', Sports Med, vol. 52, no. 5, pp. 1175–1187, May 2022, doi: 10.1007/s40279-021-01584-w.
14. E. Alonso Pérez-Chao, R. Portes, C. Ribas, A. Lorenzo, A. S. Leicht, and M. Á. Gómez, 'Impact of Spectators, League and Team Ability on Home Advantage in Professional European Basketball', Percept Mot Skills, vol. 131, no. 1, pp. 177–191, Feb. 2024, doi: 10.1177/00315125231215710.
15. V. Sarlis and C. Tjortjis, 'Sports analytics — Evaluation of basketball players and team performance', Information Systems, vol. 93, p. 101562, Nov. 2020, doi: 10.1016/j.is.2020.101562.
16. N. Zhu and Q. Dai, 'Basketball Data Analysis Based on Spark Framework and K-means Algorithm', in The 2021 International Conference on Machine Learning and Big Data Analytics for IoT Security and Privacy, vol. 98, J. Macintyre, J. Zhao, and X. Ma, Eds., in Lecture Notes on Data Engineering and Communications Technologies, vol. 98. , Cham: Springer International Publishing, 2022, pp. 853–857. doi: 10.1007/978-3-030-89511-2_116.
17. G. Csataljay, P. O'Donoghue, M. Hughes, and H. Dancs, 'Performance indicators that distinguish winning and losing teams in basketball', International Journal of Performance Analysis in Sport, vol. 9, no. 1, pp. 60–66, Apr. 2009, doi: 10.1080/24748668.2009.11868464.
18. L. Torres-Ronda, I. Gámez, S. Robertson, and J. Fernández, 'Epidemiology and injury trends in the National Basketball Association: Pre- and per-COVID-19 (2017–2021)', PLoS ONE, vol. 17, no. 2, p. e0263354, Feb. 2022, doi: 10.1371/journal.pone.0263354.
19. R. Seshadri, 'Improving Player Efficiency Rating in Basketball through Machine Learning', Mar. 03, 2024. doi: 10.58445/rars.1000.
20. Y. Chen, 'Research on the Team Operation and Management Strategy of the Mainstream Basketball Leagues in China and America - Based on the Comparison Between NBA and CBA', AEMPS, vol. 51, no. 1, pp. 259–267, Dec. 2023, doi: 10.54254/2754-1169/51/20230643.
21. T. Vencúrik et al., 'Determinants of Dribbling and Passing Skills in Competitive Games of Women's Basketball', IJERPH, vol. 18, no. 3, p. 1165, Jan. 2021, doi: 10.3390/ijerph18031165.
22. University of the Philippines, College of Human Kinetics, Department of Physical Education, Quezon City, Philippines, K. M. R. Raval, J. C. Pagaduan, and University of Tasmania, College of Medicine, School of Health Sciences, Hobart, Australia, 'Factors that Differentiate Winning and Losing in Men's University Basketball', Monten. J. Sports Sci. Med., vol. 10, no. 2, pp. 13–17, Sep. 2021, doi: 10.26773/mjssm.210902.
23. Á. Bustamante-Sánchez, J. J. Salinero, and J. Del Coso, 'Upper body injuries and Key Performance Indicators in professional basketball players', Arch Med Deporte, vol. 37, no. 6, pp. 387–392, Dec. 2020, doi: 10.18176/archmeddeporte.00013.
24. L. Danilevičius and R. Kreivytė, 'SHOOTING ANALYSIS OF LITHUANIAN NATIONAL BASKETBALL MEN'S TEAMS IN DIFFERENT AGE CATEGORIES', Baltic Journal of Sport and Health Sciences, vol. 1, no. 112, pp. 13–18, 2019, doi: 10.33607/bjshs.v112i1.774.
25. F. S. A. Brown, J. B. Fields, A. R. Jagim, R. E. Baker, and M. T. Jones, 'Analysis of In-Season External Load and Sport Performance in Women's Collegiate Basketball', Journal of Strength and Conditioning Research, vol. 38, no. 2, pp. 318–324, Feb. 2024, doi: 10.1519/JSC.0000000000004636.
26. C. Ayán, J. Cancela, and B. Fernández, 'Changes in Wheelchair Basketball Performance Indicators throughout a Regular Season: a pilot study', International Journal of Performance Analysis in Sport, vol. 14, no. 3, pp. 852–865, Dec. 2014, doi: 10.1080/24748668.2014.11868763.

27. D. Ferioli, A. T. Scanlan, D. Conte, E. Tibiletti, and E. Rampinini, 'The Business End of the Season: A Comparison Between Playoff and Regular-Season Workloads in Professional Basketball Players', International Journal of Sports Physiology and Performance, vol. 16, no. 5, pp. 655–662, May 2021, doi: 10.1123/ijspp.2020-0405.
28. J. Sampaio, T. McGarry, J. Calleja-González, S. Jiménez Sáiz, X. Schelling I Del Alcázar, and M. Balciunas, 'Exploring Game Performance in the National Basketball Association Using Player Tracking Data', PLoS ONE, vol. 10, no. 7, p. e0132894, Jul. 2015, doi: 10.1371/journal.pone.0132894.
29. G. Papageorgiou, V. Sarlis, and C. Tjortjis, 'Evaluating the effectiveness of machine learning models for performance forecasting in basketball: a comparative study', Knowl Inf Syst, vol. 66, no. 7, pp. 4333–4375, Jul. 2024, doi: 10.1007/s10115-024-02092-9.
30. W. Rangel, G. Fellingham, F. Santana, and L. Lamas, 'Integrated evaluation of team strategy, training practices and game performance of a basketball team', International Journal of Sports Science & Coaching, vol. 18, no. 1, pp. 197–206, Feb. 2023, doi: 10.1177/17479541221076621.
31. D. Ferioli, D. Conte, and A. T. Scanlan, 'Editorial: Optimizing player health, recovery, and performance in basketball', Front. Psychol., vol. 13, p. 1101052, Dec. 2022, doi: 10.3389/fpsyg.2022.1101052.
32. B. Demirci et al., 'Association of ACTN3 R577X Polymorphism with Elite Basketball Player Status and Training Responses', Genes, vol. 14, no. 6, p. 1190, May 2023, doi: 10.3390/genes14061190.