INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

SEMINAR: CURRENT TOPICS IN DATA SCIENCE

University of Applied Science - Online

Study-branch: Master of Science Data Science

Data Science Applications in Basketball Analytics

Mohammad-Marko Jibrini

Matriculation number: 32103720

79 Dryden Road, Gateshead, United Kingdom, NE9 5TR

TUTOR: Prof Tim Schlippe

Delivery date: 22.10.2023

Contents

I List of Figures

1 Introduction

Basketball is a team sport where two teams compete to score the most points by shooting a ball through the opposing team's hoop. Professional basketball is played in organized leagues across the world. In organized basketball, each team has five active players on the court at any time. Players play different positions and work together to score as many points as possible and limit the number of points scored by the opposition.

The National Basketball Association (NBA) is the world's most popular and competitive basketball league. Based in North America, it consists of thirty teams composed of the best basketball players from across the globe. It started in 1949 and has grown from a humble organization to a multi-billion dollar business.

In this hypercompetitive world, statistical analysis compares players, measures their specific abilities, and helps optimize strategies. However, traditional basketball statistics had limitations, and teams relied more on the instincts and intuitions of players and coaches. Furthermore, experienced observers preferred to rely on their subjective evaluation gained by observing the game. This approach to evaluating players is commonly referred to as the ”eye test”.

However, modern advancements in data science have changed this trend. Player tracking gathers previously unseen data on a completely different level and scale. Moreover, modern AI systems show an ”understanding” of a game comparable to the sharpest minds in basketball. Those systems can uncover hidden patterns, optimize strategies, and improve decision-making in every aspect of the game.

This essay will start by giving an overview of the use of traditional statistics in basketball. Their limitations gave rise to advanced statistics and motivated the advances in data science applications. To explain the evolution of basketball analysis, we will attempt to answer the same basketball question with different analytical frameworks and then compare the answer with the eye test approach. We will conclude by explaining how modern, state-of-the-art systems can impact players' on-court decision-making in different aspects of play.

2 Statistics in Basketball analysis

Statistics have been an integral part of NBA basketball since its inception. Various statistics, commonly called ”stats,” have been recorded since the first game. In this section, we will discuss the evolution of traditional statistics, from the humble box score stats of the 1950s to the advanced player performance metrics of the 2000s.

Before diving into the statistics, let’s look at how they were historically recorded. This was done manually by an observer who used numbers and symbols to represent players and actions. Dean Oliver gives an example in his book ”Basketball on Paper Rules and Tools for Performance Analysis” [Oliver, 2004]. The following line shows the first possession of game four of the 1997 NBA finals between the Chicago Bulls and Utah Jazz.

$$0 \; UTA \; 3\,12\,3\,14\,12^{D}\,32\text{—}_{B}\,3^{R}\,14\,12^{D}\,3\text{——}_{Y}$$

Here, the 0 indicates the number of points at the end of the possession, and UTA stands for Utah Jazz. The possession starts with number 3 winning the jump ball, then a series of passes from 3 to 12, back to 3, then to 14, and finally to 12. Number 12 then dribbles the ball (12^{D}) before passing it to 32, who takes a shot between the foul line and the three point line and misses (13—_{B}). Number 3 grabs the rebound, passes to 14, then to 12, who again dribbles before passing it back to 3, who takes a three-point shot straightway from the hoop, missing(3——_{Y}).

This way of displaying the events is convoluted and not accessible to a broader audience. The following section will discuss a more fan-friendly way of displaying basketball statistics.

2.1 Box Score Stats

The most common way to represent basketball statistics of a game is in a tabular format where the column headers are the different statistics, and different rows represent different players- the ”Box Score”. For example, the table 2.1 shows the Box score for the team ”Denver Nuggets” in the final game of the 2023 NBA finals. Box score stats are often talked about in the context of player averages. For example, Nikola Jokic averaged 24.5 points, 11.8 rebounds, and 9.8 assists for the 2022-2023 regular season[basketball reference, 2023a]. Teams, fans, and analysts all rely on these stats. They can influence whether a player gets an award, a contract by a specific team, and even their salary.

Tabel 2.1 contains the basic information any modern NBA fan would expect to find after a game. However, this was not always the case. While the NBA had its first season in 1949-1950, it counts the three years of Basketball Association of America competitions as part of its history. During those three years (1946-1949), only field goals (FG), free throws (FT), points (PTS), assists (AST), and personal fouls (PF) were recorded [basketball reference, 1947]. In the 1950-1951 season, the league started recording rebounds (REB) and the following season minutes played (MIN). Steals (STL) and Blocks (BLK) became recorded ”stats” only in the 1973-1974 season. The same year, rebounds were split into offensive (OREB) and defensive (DREB). Turnovers (TO) appeared in the box score starting

Player name	MIN	FG	3PT	FT	OREB	DREB	REB	AST	STL	BLK	TO	PF	+/-	PTS
Starters														
A. Gordon	36	7-10	0-1	2-2	2	4	6	1	0	1	0	0	+15	16
M. Porter Jr.	43	5-16	2-11	2-2	2	11	13	1	0	2	0	0	+20	16
N. Jokic	40	8-12	1-2	10-12	0	10	10	14	1	1	2	1	+14	27
J. Murray	44	11-22	2-7	2-2	1	5	6	10	1	0	3	2	+9	26
K.Caldwell-Pope	36	3-8	1-3	0-2	1	2	3	0	1	0	1	2	+5	7
Bench														
J. Green	11	2-3	0-0	0-0	0	1	1	0	0	0	0	PF	-8	4
B.Brown	21	4-7	2-3	0-0	0	5	5	2	1	0	1	PF	+1	10
C. Braun	8	0-1	0-0	0-0	0	1	1	0	0	0	0	PF	-1	0
Team total		40-79	8-27	16-20	6	39	45	29	4	4	10	8		104
		50.6%	29.5%	80.0%										

Table 2.1: Denver Nuggets BOX score, Game 5 of 2023 NBA final [ESPN, 2023]

the 1977-1978 season, and 3-point field goals (3FG) in 1979-1980. Note that the 3-point field goal did not exist until the 1979-1980 season. The plus-minus (+/-) is the difference between the number of points scored(+) and the number of points conceded (-) while the player is on the court. The NBA first published it in the BOX score in the 2007-2008 season [NBA.com, 2023].

While the box score can summarize each player's contribution to a game, it has some significant limitations. First, there is a lack of context. Scoring a game-winning three-pointer at the end of the fourth quarter is not the same as scoring three points in the first quarter. Furthermore, there is no information on the difficulty of shots. Was offense generated through good team play, breakdown in opposition defense, or individual heroics? Regarding play-making, not all assists are equal. Sometimes, a shooter will make a shot (score) despite a bad pass by a teammate; other times, the pass leads to easy points. Furthermore, other play-making activities that lead to high-quality shots, such as good screens, cross-court passes, and hockey assists, are not recorded.

Defence is even harder to represent in the box score. Steals and blocks are recorded, but they are relatively rare events that tell a minimal story of a player's defensive impact. Competent defenders force opposition players into taking low-percentage shots, get deflections, and force turnovers. None of those metrics are present in the box score. Therefore, the "eye test" reigns supreme as an experienced observer would catch those actions.

To exemplify these limitations, let us ask a simple, commonly asked question- "Who is the best shooter in the NBA?" Is it the player that averages the most points per game? That player could be inefficient and hurt his team's chances of winning. Is it the player with the best field-goals made to field-goals attempted ratio? That could be a player that takes very few high-quality shots. Box score stats do very little to answer that question. This motivated the emergence of advanced stats derived from box score stats to give more depth to such answers.

2.2 Advanced Stats

2.2.1 Shooting Efficiency

In the previous section, we asked who is the best shooter in the NBA. We could answer this question using the field goal percentage (FG%), which is calculated as $FG\% = \frac{FGM}{FGA}$, where FGM stands for field goals made, and FGA stands for field goals attempted. Note that field goals are shots taken in

active play, as opposed to free throws.
However, in the 1979-80 season, the NBA adopted a three-point shot[NBA.com, 2023]. Shots made from beyond the three-point line count for three points instead of those inside, which count for two. Now, FG% does not necessarily give an accurate representation of players scoring efficiency. Shooting 50% from beyond the three-point line would yield more points than the same percentage from inside the line while using the same number of possessions. Therefore, another stat was introduced- effective field goal percentage (eFG%), which adjusts the value of a three-pointer by counting it as 1.5 field goal, or $eFG\% = \frac{FGM+0.5(3PTM)}{FGA}$ [stuffer, 2023a], where 3PTM is the number of three-pointers made.

Neither FG% nor eFG% accounts for free throws. Free throws are shots awarded to a player, usually when they are fouled in the act of shooting or when the opposing team has reached its allowed number of fouls for a quarter. They are taken from the free-throw line. Each counts for one point, and a player is awarded two free throws when fouled on a two-point shot and three for three-pointers. If a player makes the shot despite being fouled, they will be awarded one free throw on top of counting the points from the shot they made; this shot is popularly called "and one." True shooting percentage (TS%) factors in free throws into the efficiency equation. The formula is $TS\% = \frac{total\ points}{2*(FGA+0.44*FTA)}$[stuffer, 2023c], where FTA is the number of free throws attempted. The .44 factor is because not all free throws take up a possession, as in and-ones.

Using advanced stats, we get closer to answering who is the best shooter. At least we have a reliable measure of efficiency. However, we still do not have an answer, as we lack the context of how those field goals are made. We will get back to this question later. For now, let us look at advanced stats that summarize the totality of a player's impact in one number.

2.2.2 Player Ratings

The player efficiency rating (PER) is the first attempt at a rating that boils down the total impact of a player to one number. It was developed by John Hollinger, a basketball analyst and former vice president of basketball operations for the Memphis Grizzlies [Hollingert, 2005]. The unadjusted player efficiency rating (uPER) is calculated with a system of linear weights assigned to each box score stat. For example, part of the equation regarding FGs is given by $[2 - (\frac{2}{3} - [(0.5 * \frac{lgAST}{lgFG}) \div (2 * \frac{lgFG}{lgFT})]) * \frac{tmAST}{tmFG}] * FG$. The lg prefix indicates a league average stat, and the tm prefix indicates team stats. The equation uses team and league averages to compare the importance of different stats. The full calculation incorporates the value of possession (VOP), an estimated average number of points per possession. Then, the rating is adjusted to league and team pace and normalized to set the average PER to 15. The full equation for the PER can be found at https://www.basketball-reference.com/about/per.html

Player impact estimate (PIE) is a stat published by the NBA to replace the PER[stuffer, 2023b]. It has a simplified formula and focuses on what percentage of the events in a game each player contributed. Ultimately, both the PER and PIE suffer from the same issues as the box score stats they are derived from. They lack context and undervalue defense.

Win shares are another advanced stat. This time, it divides the team wins into shares to be shared by the players. It estimates a player's contributions to their team winning. However, it falls short when comparing players from different teams, as a better player playing on a team with fewer wins could end up with fewer win shares than a worse player on a better team.

3 Data Science in Basketball Analytics

While statistics provide a good summary of events during a basketball game, they have some significant limitations. One of those limitations is that they record events as binaries. The shot is a hit or a miss, and a pass is an assist or not recorded. This is even more evident on the defensive end. Only blocks and steals are recorded, which entirely end the opposition attack. In reality, most good defensive possessions end with a low-percentage shot. This principle is true in general. On a strategy level, basketball is a game of probabilities. The attacking team seeks to get the highest percentage shot in each possession. At the same time, the defense tries to force the attackers into taking bad (low percentage) shots. On top of that, the rebounding battle is where extra possessions can be won. Ultimately, the math is simple: the team will win if they have more possessions and score more points per possession (TS%) than the opposition.

This chapter will discuss the technological advancement that let us go beyond traditional statistics. We will show how data science can provide much-needed context to traditional stats—namely, the context of space and time. Furthermore, we will see how current state-of-the-art systems can analyze the game at a probabilistic level.

3.1 Early Spatial Analysis

The 2012 paper titled: ”CourtVision: NBA Visual and Spatial Analytics” [Goldsberry, 2012]by Kirk Goldsberry attempts to solve some of the issues in basketball analysis by providing more context. Namely, as Goldsberry said in his presentation of CourtVision, Basketball is a spatial sport. In order to answer the question of who is the best shooter, we need to provide the context of positioning. The authors used NBA game data sets of every game played between 2016 and 2011 and recorded each field goal attempted alongside its Cartesian coordinates, player name, and shot outcome. Figure 3.1 shows the number of field goals from different court areas and their corresponding mean points per attempt. With this data, we can answer the question of the best shooter in the NBA with more context. For example, Nene Hilario and Dwight Howard led the league in FG% in the 2010-2011 season. Both tall centers are not considered great shooters but only take shots near the basket. The authors of the paper introduce two new ”stats.” First, ”Spread” is a count of how many different spots a player shot from. The player with the highest spread was Kobe Bryant; this aligns with his reputation as a dangerous scorer from anywhere on the court. Second is ”Range,” or the number of spots from which a player averages at least one point per attempt. In this metric, Kobe comes third. The two players above him are Steve Nash and Ray Allen, both considered among the best shooters in NBA history.

By adding space as context, we can derive an analysis that matches the ”eye test” much better. Nevertheless, more insight can be drawn from figure 3.1. The most valuable shots, with the highest points per attempt, are from right next to the basket, followed by three-point shots, particularly the corner three. This insight revolutionized how basketball is played. Even though the three-point shot has existed since the eighties, its real value was not understood until the last decade. The average number of three-point attempts by NBA teams has increased from 18.4 in the 2011-2012 season to 34.2 in the 2021-2022 season[basketball reference, 2023b].

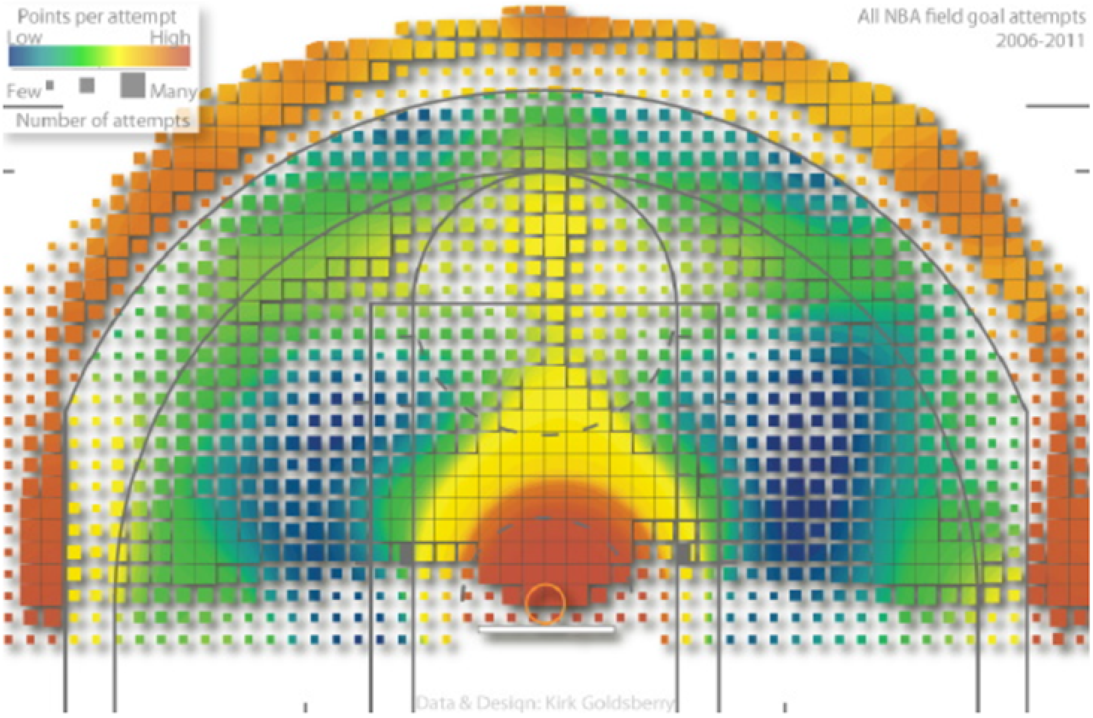

Figure 3.1: NBA field goals 2006-2011 [Goldsberry, 2012]

Furthermore, good spot-up shooters and good defenders, commonly referred to as ”3 and D” players, became a highly sought-after commodity for NBA teams. They can contribute on the offensive end just by standing in the corner behind the three point line. Their defender is ”glued” to them, freeing up space around the basket. If their defender leaves to help the rest of the team's defense, he risks an open, high-value shot from the corner three. Therefore, ”3 and D” players do not need a versatile offensive skillset to contribute. They do not need to be good dribblers or have the ability to create offense themselves. Section 4.1 explains why such players are highly valuable.

Adding space as a context in NBA statistics has both improved how we evaluate the player's ability and changed how teams run their offense. In the next section, we will look at how spatial tracking has evolved in the NBA and how it opened a whole new way to analyze the game.

3.2 Sports UV

SportVU is a camera system that collects data in 25 frames per second and tracks player and ball movement. It was developed by scientists with a background in missile tracking[Journal, 2013]. It was showcased to NBA executives during the 2009 NBA finals. At the start of the 2010-2011 season, four NBA teams started using this technology, and since the 2013-2014 season, SportUV cameras have been installed in all NBA arenas. The exact technology used by SportUV is a closely guarded secret. They rely on deep learning and computer vision to extract actionable data from the raw footage. Output data is time-stamped spatial data (x, y) for all ten players on the court and (x, y, z) for the ball. An example of the output data, used to generate figures 3.2, and 3.3, is taken from repo [Linou, 2016]. Figure 3.2 shows a fastbreak lead by Russell Westbrook #0. Figure 3.3 is seconds later, and moments after Kevin Durant #35 shoots the ball. Here, the size of the ball indicates its position on the z-axis.

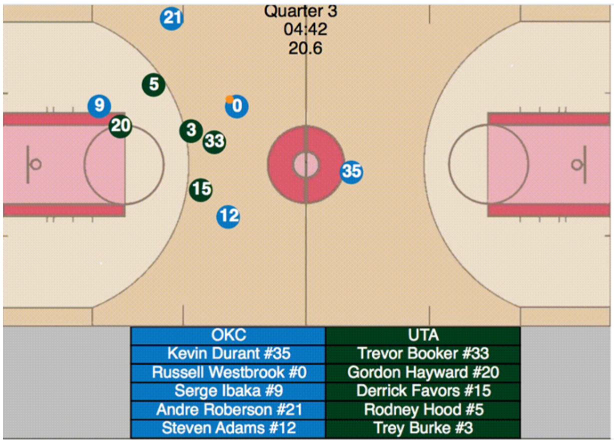

Figure 3.2: Fastbreak

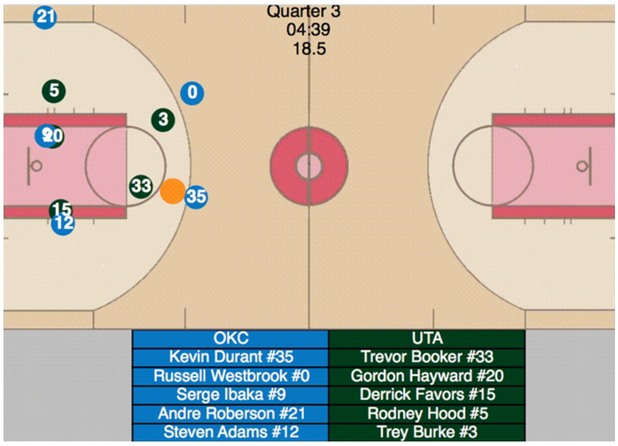

Figure 3.3: Shot

3.3 Second Spectrum

In the 2017-2018 season, Second Spectrum replaced SportUV to become the Official Optical Tracking Provider of the NBA. They pledged to leverage "unique technology for machine understanding of sports for comprehensive analysis of players and teams". Second, Spectrum improves player tracking and tracks more data. Most importantly, it uses Artificial intelligence to "teach" machines basketball terms[Maheswaran, 2016a]. As Rajiv Maheswaran explains, at the beginning, the AI was taught basic terms, such as pass, shot, and rebound. It learned to recognize specific offensive actions such as post-ups, pick and rolls, and isolations using Spatiotemporal pattern recognition. By 2015, it "understood" events that go beyond most fans' knowledge, with the precision that professional NBA coaches trusted. It was used by many of the top NBA teams to inform their strategy.

Going back to the question of the best shooter. We have talked about high- and low-percentage shots, but we can quantify those percentages using Second Spectrum data. Quantified Shot Quality (qSQ) is a metric that takes into account the type of shot, location of the shot, and nearby defenders when describing the quality of a shot [Chang et al., 2005].

Figure 3.4: Quantified Shot Quality from Second Spectrum [Chang et al., 2005]

To answer the question of the best shooter, we can derive the Quantified Shooter Impact (qSI). This metric measures the impact of a shooter on the probability of a shot going in, compared to the league average. Now, we can differentiate between good shooters who take bad shots and bad shooters who take good shots, even when they have the same FG%. The white line in figure 3.5 shows players who shot 47% FGs.

For the 2015-2016 regular season, the highest qSI belonged to Steph Curry[Maheswaran, 2016b]. He not only passes the "eye test" but is widely recognized as the best shooter ever. Finally, we have a metric that takes into account all relevant information. This approach matches the "eye test" and goes above and beyond in its ability to give exact numbers to previously unquantifiable measures. For example, Steph has a qSI of +13%. This means that any shot gets an extra 13% added to its qSQ when taken by him.

As stated at the beginning of this chapter, basketball is a game of probabilities. Now, we have the tools to study it in that manner. In the next chapter, we will discuss how this technology can impact court decisions in the three major aspects of play in basketball.

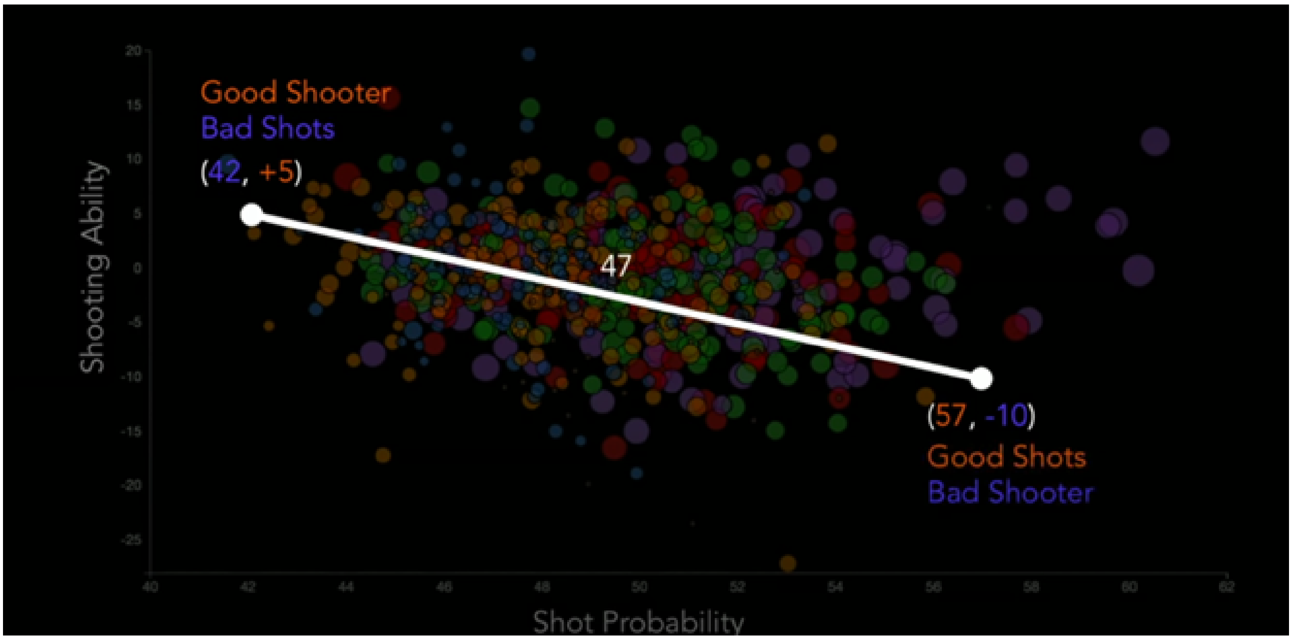

Figure 3.5: Shooting ability VS Shot Probability [Maheswaran, 2016a]

4 Data Science use in On-Court Decision Making

In this chapter, we will examine how the developments in player tracking technology can be used to inform on-court decision making. The remainder of this chapter will be divided according to the three main areas of play. Offense, where teams aim to maximize points per possession; defense, where they limit the opposition's points per possession; and the rebounding battle for more possessions. We will examine the same play from the offensive and defensive perspective for continuity.

4.1 Offensive Application of Data Science

Lets go back to figure 3.1. Here, we can see that shots from right behind the basket and the corner three (C3) provide the best point per possession. A basketball play called drive-and-kick takes advantage of both of these. Before diving deeper, let us understand what makes those two parts of the court particularly fruitful.

It is obvious why shots near the basket return the most points per possession. On the other hand, C3s require a deeper dive. On average, C3 generates 12 more points, per 100 shots, than non-corner threes (nC3)[Pelechrinis and Goldsberry, 2021]. The first assumption would be that the difference is due to distance from the basket. The average distance for a corner three in the dataset used by [Pelechrinis and Goldsberry, 2021] is 7.01 meters, compared to 7.65 meters for nC3 threes. According to the logistic regression model, based solely on distance from the basket, that would lead to a 1.8 % increase in FG%. However, researchers find that the actual FG% for C3 is 38.8%, compared to 34.7%, a difference of 4.1%. So, how can we explain this difference?

Authors of "The Anatomy of Corner 3s in the NBA: What makes them efficient, how are they generated and how can defenses respond?" [Pelechrinis and Goldsberry, 2021] utilized player tracking to study C3s.They used k-mean clustering on the shooter's and defender's movement to find ten major patterns that lead to C3s, see figure 4.1. They find that C3 shots are less contested. The average distance to the nearest defender on C3s is 1.98 m, compared to less than 1.83 m in nC3s. Furthermore, C3 are more often assisted, over 90% compared to under 70%. The conclusion is that C3s are results of good ball movement, resulting in an assisted, less contested shot. Furthermore, over a third of the passes to corner threes come from near the basket. This brings us back to the drive-and-kick. In this play, the ball handler "penetrates" the defense and attacks the basket, trying to get as close to it as possible; this is the "drive." The opposing team brings in extra defenders to stop him from taking a high-percentage shot at the basket, leaving an open spot-up shooter in the corner. The ball handler passes the ball to the shooter, the "kick," for an uncontested corner three.

The ball handler can prepare for such situations by utilizing Second Spectrum data. They can look at their own qSQ at the basket and the shooter's qSQ from the corner when guarded by a different defensive scheme. Then, add to the qSI of both offensive players to determine whether to take the shot themselves or pass the ball.

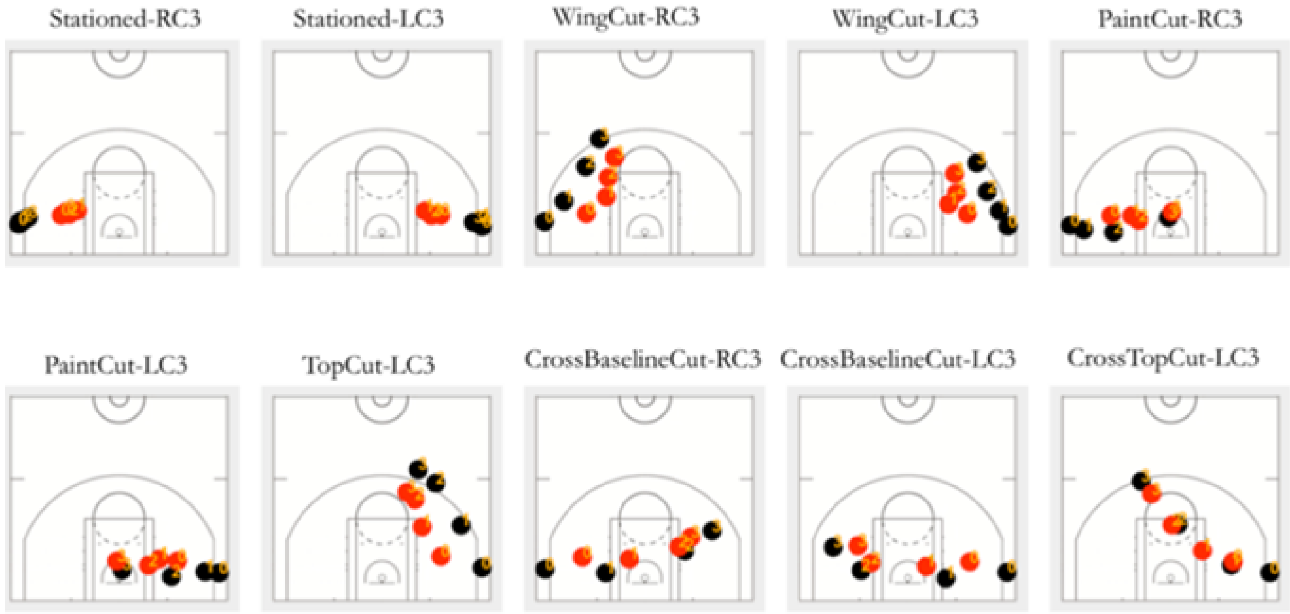

Figure 4.1: Corner three clusters [Pelechrinis and Goldsberry, 2021]

4.2 Defensive Application of Data Science

In this section, we will take the drive-and-kick action and examine it from the defensive point of view. More specifically, we will focus on the defender guarding the spot-up shooter and examine how they could use Second Spectrum data to optimize their decision-making.

Said defender has three options. One: they can stick to the shooter and prevent an open three but risk an easy shot at the basket. Two: help defend the penetrating player and risk an open three. Three: try an in-between strategy where they move closer to the ball handler to try and bother him while staying close enough to the shooter so they can get back to them in enough time once a pass is made. In practice, this decision boils down to the distance "D" they stand between the basket and the shooter.

We can write the qSQ for both the C3 and the shot at the basket in terms of D. For the first, D will determine the distance of the primary defender, and for the latter, 1-D is the distance of the secondary defender, see figure 3.4. The authors of [Pelechrinis and Goldsberry, 2021] paper applied game theory to this system and solved for Nash equilibrium.

The results are that the player should always commit to the shooter or the ballhandler. The in-between strategy is found to be less effective. A smart defender will alternate between the two to keep the offense guessing. The exact proportion of where to commit can be further derived from qSQ data.

4.3 Data Science use in the Rebounding Battle

Let us continue with the analogy from the previous section. Regardless of who took the shot, the ballhandler or the spot-up shooter, the rebounding battle starts when they release the ball. As all coaches love to remind their players, a successful defensive possession ends with a defensive rebound. Otherwise, if an offensive player grabs the ball after a miss, they can go again. Here, teams fight for extra possessions that can tip the game in their favor. To understand this battle better, the authors of "The Three Dimensions of Rebounding" [Maheswaran et al., 2014], break it down into distinctive phases.

First is *positioning* when the ball is released from the shooter's hand. To assess the value of each player's position, the authors of [Maheswaran et al., 2014] first assign to each player all the points on the court that are close to them in the form of a Voronoi diagram. Next, each area on the court is assigned a value corresponding to the probability of a rebound landing there. The total value of a player's "real estate" measures their probability of getting a rebound if no one moves.

Next is the crashing and blocking out phase. This is when players are either trying to get to the ball, "crashing," or blocking opposing players from getting there. Each player's positional value is recalculated when the ball is closest to the rim's center. Players with good initial positioning will struggle to gain value as other players swarm the priced real estate. To account for this, researchers looked at the difference between the two probabilities against the probability at the shot, figure 4.2. The line plot is a linear regression of the variables. The positive slope for offensive players indicates that players with good initial positioning will fight for the offensive rebound, while others will not attempt. A player's "crash" metric is their distance from the regression line.

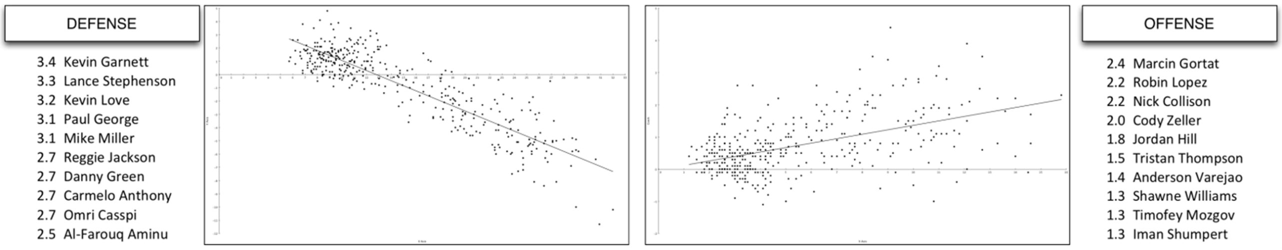

Figure 4.2: Initial positioning value vs positioning delta [Maheswaran et al., 2014]

Hustle is the ability for a player to generate opportunities for a rebound beyond their initial positioning. Opportunity is when the player is closest to the ball when the ball's z-axis is lower than 3 meters. The ability to generate opportunities could correlate with initial positioning. So, to derive the hustle metric, the authors performed the same regression and adjustments.

Finally, raw *conversion* is measured by dividing the number of rebounds by the number of opportunities a player generates. Again, the same methodology is applied to form a regression line of conversion percentage to initial positioning. Here, there was a positive slope for both offense and defense. This makes sense, as players with good positioning are likelier to convert on their first opportunity.

Before player tracking, teams could only prepare for the rebounding battle by labeling opposition "good" and "bad" rebounders based on their rebounding numbers. Now, teams can understand the value of the positioning battle for each area of the court. Furthermore, defenses can use the crashing metric to identify which offensive players to box out. Finally, teams can access more information when acquiring new players when they need more possessions.

5 Conclusion

Basketball analysis is built on the foundation of traditional and then advanced statistics. Data science, specifically computer vision, revolutionized basketball analytics by adding the context of space. After that came the advanced AI systems that study the game at a probabilistic level and ”understand” different actions happening on the court.

Currently, the Second Spectrom system is trusted by all NBA teams to provide a competitive advantage against their opponents. It can reveal the patterns and tendencies of the opponents and maximize offensive positions against specified defenses. Furthermore, it can identify which skills are needed for a team when making new player acquisitions.

In this essay, we have examined this analytical revolution through a simple question: ”Who is the best shooter in the NBA?” then subjecting the answers to the ”eye test”. Different historical approaches used different metrics to answer this question, from points-per-game of traditional statistics to true shooting percentage and more abstract metrics such as spread and range. Finally, we had a metric that took in all relevant information in the form of the quantified shooter impact. Advancements in data science allow fans, analysts, and teams alike to answer basketball questions with more nuance than ever.

That is not to say that AI systems have complete supremacy over the eye test approach. Quantified shooter impact is still not a perfect measure of a player's shooting ability. A more comprehensive measure of players shooting would be a heat map with different values in different areas of the court. For example, while a short shooter could be more likely than an average player to make a contested three-point shot, but they might struggle more when challenged by tall defenders near the basket.

Furthermore, the impact of nearby defenders will vary depending on who is taking shots. For example, Kevin Durant is praised for his ability to be unbothered by defenses. His long frame, wingspan, and shooting ability make it so that he can shoot over most defenders.

We cannot say whether the NBA and Second Spectrum have implemented similar advances, as we can only review publically available information. Regardless, it is hard to believe that AI systems will make the eye test obsolete. Perhaps, in the current state of basketball analysis, the eye test can be used as a guide for future advances instead of being treated as an adversary.

Bibliography

[basketball reference, 1947] basketball reference (1947). Baa 1947 stats. https://www.basketball-reference.com/leagues/BAA_1947.html,.

[basketball reference, 2023a] basketball reference (2023a). Nikola jokic 2022-2023 regular season stats. https://www.basketball-reference.com/players/j/jokicni01/gamelog/2023,.

[basketball reference, 2023b] basketball reference (2023b). Tree point attempts per year. https://www.basketball-reference.com/leagues/NBA_stats_per_game.html,.

[Chang et al., 2005] Chang, Y.-H., Maheswaran, R., Su, J., Kwok, S., Levy, T., Wexler, A., and Squire, K. (2005). *Quantifying Shot Quality in the NBA*. Second Spectrums.

[ESPN, 2023] ESPN (2023). Nba finals game five. https://www.espn.co.uk/nba/boxscore/_/gameId/401544850,.

[Goldsberry, 2012] Goldsberry, K. (2012). *CourtVision : New Visual and Spatial Analytics for the NBAt*. Harvard University.

[Hollingert, 2005] Hollingert, J. (2005). *Pro Basketball Forecast*. Potomac Books.

[Journal, 2013] Journal, S. B. (2013). Sportvu tracking system could find reach beyond court for nba teams. https://www.sportsbusinessjournal.com/Daily/Issues/2013/10/29/NBA-Season-Preview/Tech.aspx,.

[Linou, 2016] Linou, K. (2016). Github repo. https://github.com/linouk23/NBA-Player-Movements,.

[Maheswaran, 2016a] Maheswaran, R. (2016a). The math behind basketball's wildest moves. https://www.youtube.com/watch?v=66ko_cWSHBU/,.

[Maheswaran, 2016b] Maheswaran, R. (2016b). What advanced tracking data reveals about nba shooters. https://www.espn.com/nba/playoffs/2016/story/_/id/15530589/what-advanced-tracking-data-reveals-nba-shooters,.

[Maheswaran et al., 2014] Maheswaran, R., Chang, Y.-H., Su, J., Kwok, S., Levy, T., Wexler, A., and Hollingsworth, N. (2014). *The Three Dimensions of Rebounding*. Second Spectrum.

[NBA.com, 2023] NBA.com (2023). Stat recording history. https://www.nba.com/stats/help/faq,.

[Oliver, 2004] Oliver, D. (2004). *Baseketball on Paper Rules and Tools for Performance Analysis*. Potomac Books.

[Pelechrinis and Goldsberry, 2021] Pelechrinis, K. and Goldsberry, K. (2021). *The Anatomy of Corner 3s in the NBA: What makes them efficient, how are they generated and how can defenses respond?* arXiv e-prints.

[stuffer, 2023a] stuffer, N. (2023a). Effective field goal percentage. https://www.nbastuffer.com/analytics101/effective-field-goal-percentage-efg/,.

[stuffer, 2023b] stuffer, N. (2023b). Player impact estimate. https://www.nbastuffer.com/analytics101/player-impact-estimate-pie/,.

[stuffer, 2023c] stuffer, N. (2023c). True shooting percentage. https://www.nbastuffer.com/analytics101/true-shooting-percentage/,.