Characterization of Space and Time-Dependence of 3-Point Shots in Basketball
Gabin Rolland, Romain Vuillemot, Wouter J.T. Bos, Nathan Rivière

To cite this version:
Gabin Rolland, Romain Vuillemot, Wouter J.T. Bos, Nathan Rivière. Characterization of Space and Time-Dependence of 3-Point Shots in Basketball. MIT Sloan Sports Analytics Conference, Mar 2020, Boston, United States. hal-02482706

HAL Id: hal-02482706
https://hal.science/hal-02482706
Submitted on 18 Feb 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Characterization of Space and Time-Dependence of 3-Point Shots in Basketball

Gabin Rolland, Romain Vuillemot, Wouter Bos, Nathan Rivière
1548844

Abstract

Understanding characteristics of 3-point shots is paramount for modern basketball success, as in recent decades, 3-point shots have become more prevalent in the NBA. They accounted for 33.6% of the number of total shots in 2017-2018, compared to only 3% in 1979-1980 [1]. In this paper, we aim at better understanding the connections between the type of 3-point shooting (catch-and-shoots and pull-ups) and the timing for shooting, using two distinct space-time models of player motion. Those models allow us to identify individual behavior as a function of specific defensive settings, e.g. shot-behavior when a player is guarded closely. We assess our models using SportVU data for specific NBA players and our code is open-source to enable more players and teams explorations, as well as to support further research and application of those models, beyond basketball and sport.

1. Introduction

When the 3-point-line was introduced in 1979, on average 2.8 shots were taken per game. Since then, this number increased roughly by a factor of ten, up to ~27 shots per game [2]. The influence of 3-point shots has thereby radically changed the NBA and has become a major asset for the most successful teams. An illustrative example are the Golden State Warriors for whom 3-point shots following a counter-attack were decisive during the 2017-2018 season. They won that season taking 654 3-point shots in the first seven seconds after a change of possession for a success rate of 43% [3]. Understanding the factors which influence the 3-point performance is therefore extremely valuable as illustrated by either predictive or descriptive carried out in sports analysis.

Our work relates to players’ influence zone models, and their dependence on players’ velocity. In [4], [5] and [6] for each player a dominant region was calculated using Voronoi tessellation, adapting the metric using different motion models adapted to the players’ dynamics. Fernandez and Bornn used statistical techniques to characterize players’ space occupation [7]. Other investigations focused on events occurring in a match, such as passes or shots. For instance, different methods were developed to predict shooting and passing possibilities in basketball [8], or describing important plays in football matches [9]. A different analysis tool [10] based its assessment of football passes on space occupation models. Several studies are related to shot performance in basketball: [11] proposed an interface allowing to compare shooting performance from different players; [12] introduces a metric associated with the possible number of points scored in an action; [13] describes a statistical study on the influences of different factors which influence 3-point performance and tries to classify shooters behavior; Finally, [14] introduces two metrics to respectively characterize shot quality and shot ability.

In this paper we aim at better understanding 3-point shots in basketball, as those are frequent and decisive situations during NBA games. In particular we focus on catch-and-shoot and pull-up shots
and their dependence on defense settings. We first introduce our methodology to extract 3-point attempts (discrete events) from SportVU spatio-temporal data (players and ball position time-series). We then present our space-time models designed to quantify available space for a single player relative to the space occupied by other players. One of our models measures only the distances between players (static model, Figure 2 (a)), while the other one takes into account the speed and direction of the players (dynamic model, Figure 2 (b)). Those two models allow us to characterize defensive dynamics and to analyze its impact on 3-point efficiency. All our code and analysis are available as an open-source project https://github.com/amigocap/MecaSportStats/ to support further research and application of those models.

2. Extracting Pull-up and Catch-and-Shoot Data from SportVU

Our study relies on motion data of player and ball positions during 632 matches from the 2015-2016 NBA season, supplied by STATS SportVU. We were able to detect 3-point attempts by tracking back the ball paths that crossed a circular horizontal surface of radius 5 feet above the basket (at 10 feet height) and looking at the position of the latest ball owner at shot time. Figure 1 illustrates the distribution of such shots (circles) attempted by Stephen Curry. Shot time was defined as the time just before the ball’s vertical angle was higher than 70 degrees tracking back from the moment ball’s height reaches 10 feet. Figure 1 distinguishes hits and miss shots (respectively blue and red) which were differentiated by evaluating the trajectory of the ball and its intersection (or not) with the plane of the basket.

![Figure 1: 380 Curry's shots from Season 2015-2016. Blue dots represent successful shots location and red ones miss shots location. Black segments represent the distance of the closest defender.](image)

After identifying 3-point attempts, we differentiate two subclasses of shots, catch-and-shoot and pull-up shots. A catch-and-shoot shot consists of the play where a player receives a pass and shoots less than two seconds after the pass without dribbling. Conversely, a pull-up shot is a shot taken after a dribble. To differentiate these shots, we had to set the release time which is the delay between the
time the player catches the ball ($t_{\text{catch}}$) and the ball releases the shooter’s hands ($t_{\text{shot}}$). To do so, we determine the time the shooter catches the ball by looking at when the ball was within 1.6 feet of the player for the first time. Then a shot was considered as catch-and-shoot if release time was less than 2 seconds and if the shooter did not dribble before shooting, e.g. if the ball did not pass under 2 feet.

As a result, we found 26 332 valid shots; we discarded shots with noisy or data gaps. 72% of these shots were catch-and-shoot ones. We found 9136 hits and 17196 misses, corresponding to an overall success rate of 34.7%. We obtained a catch-and-shoot success rate of 36% against 31.5% for the pull-up ones which is consistent with previous studies [15]. For all these shots, we recovered the position of the shooter and his closest defender three seconds before the shot in order to evaluate several features such as the distance of the closest defender at the shot time illustrated by black segments in Figure 1. This quantity is useful to characterize the opening of a shot and allows to compare to the definitions of the NBA stats website [16] defining 4 levels of opening for a shot according to the distance of the nearest defender: from 0 to 2 feet it is a very tight shot, from 2 to 4 feet it is a tight shot, from 4 to 6 feet it is an open shot and more than 6 feet it is a wide-open shot.

3. Players Occupation Models

To identify the occupation zones of a player, one needs to take into account the positions of all players from both teams. The simplest approach is to determine for every point in space the closest player. This allows to subdivide the court into 10 Voronoi cells, defined as follows: a point belongs to a player’s cell if and only if it is closer to the player than to any other player [17]. This approach is however limited by its discrete, binary nature and does not account for other players’ orientation and inertia (i.e. direction and motion).

![Figure 2](https://amigocap.github.io/MecaSportStats/video.mp4)

3.1. Occupation Map

We introduce a continuous measure, by taking into account the relative distance of a point to a player. This quantity, $\delta_{\text{space}}(x,y)$ is calculated as the difference between the distance $d$ from the point $(x,y)$ to the closest player and the distance from the same point to the closest opponent,
\[
\delta_{\text{space}}(x,y) = d_{\text{closest opponent}}(x,y) - d_{\text{closest player}}(x,y)
\]  

(1)

This quantity does therefore not only take into account the distance to a player, but also the fact that only players of the opposite team will dispute the control of a certain area. On the boundaries of a Voronoi cell between players of different teams the value of \(\delta_{\text{space}}\) is per definition \(\delta_{\text{space}} = 0\). By calculating \(\delta_{\text{space}}\) for each point of the court we were able to determine an occupation-map describing the occupation of the court by the two teams. We show an illustration of such an occupation-map in Figure 2 (a).

3.2. Dynamic Players Occupation Model

The previous approach gives a straightforward way to quantify how players occupy a basketball court. However, it does not take into account that players have inertia (i.e. players in motion). Clearly, for a fast-moving player, it is more difficult to control the area behind him, and some delay is induced by the fact that players have a finite force and can therefore not instantaneously change their velocity to a different value. It is thus necessary to develop a model characterizing the acceleration of the players. However, the distance to a point does not fully define the control of a point, but it is rather the time it takes for a player to reach the point which defines the controlled area. Therefore, the boundaries between a defense-controlled zone and an attack-controlled zone are more precisely defined by a quantity \(\delta_{\text{time}}\),

\[
\delta_{\text{time}}(x,y) = t_{\text{closest opponent}}(x,y) - t_{\text{closest player}}(x,y)
\]  

(2)

where \(t_{\text{closest player}}\) is the time it takes for the closest attacker (in seconds) to join the point \((x,y)\), and \(t_{\text{closest opponent}}\) the same time for a defender. In the unphysical case of players with zero mass, or infinite force, this \(\delta_{\text{time}}\) should behave as \(\delta_{\text{space}}\). However, inertia is expected to change this. The model we use to determine the time is similar in spirit, but not in detail and implementation, to previously investigated models \([4], [5]\) and \([6]\). In particular, the fact that we have an analytical solution for the time-interval makes its implementation fast compared to iterative procedures such as the one proposed in \([5]\), which makes the current procedure more tractable for real-time applications. Details on time calculation are available in the Appendix.

With \(\delta_{\text{time}}\) we can again determine an occupation-map, one of which is illustrated in Figure 2(b). When comparing both maps Figure 2(a) and 2(b) clear visual differences are observed. Consider for instance the position on the court indicated by a green circle. In a simple (non-inertial) model considering only the distance of the players to a point, this part of the court is controlled by the defensive team. However, in reality, it is the offensive player (1) which will arrive at this point before the player (2) since the velocity of player (1) is directed towards this point while the defensive player (2) is moving away from this point. This example clearly shows how the inertia of the different players can be important in the determination of the influence zones of the teams.

4. Offset Induced by Inertia Between Space and Time

We now apply our models to characterize space and time free-space dependencies. To achieve this, we first calculate \(\delta_{\text{space}}\) and \(\delta_{\text{time}}\) at the player position \((x_{\text{player}}, y_{\text{player}})\) as
\[ \delta_{space}^* = \delta_{space}(x_{player}, y_{player}) = d_{closest\ opponent}(x_{player}, y_{player}) \quad (3) \]

\[ \delta_{time}^* = \delta_{time}(x_{player}, y_{player}) = t_{closest\ opponent}(x_{player}, y_{player}) \quad (4) \]

Those calculations allow us to represent the evolution of players’ free space (Figure 3). A clear trend is that free space \( \delta_{time}^* \) lags free space \( \delta_{space}^* \). This difference clearly shows how inertia introduces a delay in defensive play. When a defender observes that the offensive player he is guarding changes his position, he needs to adapt his velocity to adjust to the best possible defensive position. The time lag \( \tau_\delta \) is of the order of 0.3 seconds.

\[ \text{Figure 3: Example of Kawhi Leonard’s free space evolution during an 8s time sequence of the game opposing San Antonio Spurs to Washington Wizards that played on November 4, 2015. Comparison of free space measured using } \delta_{space}^* \text{ (red) and } \delta_{time}^* \text{ (blue).} \]

The precise value of the delay is obtained by measuring for which value \( \tau_\delta \) the maximum of the correlation coefficient \( \rho_\delta \) is observed. This correlation coefficient is defined as,

\[ \rho_\delta(t) = \frac{\langle \delta_{space}'(t + \tau) \delta_{time}'(t) \rangle}{\langle \delta_{space}'^2 \rangle^{1/2} \langle \delta_{time}'^2 \rangle^{1/2}} \quad (5) \]

with \( \delta' = \delta - \langle \delta \rangle \) and where \( \langle \cdot \rangle \) denotes a time-average. The maximum of the correlation is shown in Figure 4(a). The maximum correlation is found to be \( \tau_\delta = 0.3 \text{s} \). It thus takes on average 0.3s for a defensive player to adapt his velocity to the velocity of the offensive player just before a 3-point shot.

We can further compare this specific measure of the adaptation time of a defender to approach the shooting attacker to a global, team-based delay. Indeed, the necessity to rapidly approach an
offensive player starting his shot leads probably to a much faster adaptation to the defensive positions than the global delay of the full team with respect to its inertia-corrected position. In order to check this we computed the temporal correlation of $\delta_{space}$ and $\delta_{time}$ as in Equation (5), but now not only for the distance to the closest defender to the holder of the ball but computing the maximum correlation between the full occupation map such as the one in Figure 1(a) and the inertia-corrected occupation map (cf. Figure 1(b)). This was performed for every event during a quarter of the game in Figure 4(b) and the maximum correlation obtained was $\overline{\tau}_{8} = 1.26\,s$, where the overline indicates that a global measure is considered.

### Figure 4: (a) Time-correlation between $\langle \delta_{space}^{*} \rangle$ and $\langle \delta_{time}^{*} \rangle$, representing offensive player’s free space evaluated 3 seconds before a 3-point shot. (b) Global correlation between the occupation map measured with $\delta_{space}$ and $\delta_{time}$ (such as the occupation maps in Figure 1(a) and (b), respectively). The result is obtained by averaging over all the events during one quarter of a game.

## 5. Characterizing Free Space and 3-point Shots

This section uses the previously introduced occupation models to analyze and compare shooters’ behavior before 3-point attempts.

### 5.1. Free Space Evolution Before a 3-point Shot

Figure 5(a) and 5(b) respectively show the evolution of $\delta_{space}^{*}$ and $\delta_{time}^{*}$ three seconds before and one second after 50 randomly chosen 3-point shots (grouped either by pull-up shots or catch-and-shoot ones). Two different trends can be noticed between these shots. Before shooting, a catch-and-shoot shot requires substantial free space for the shooter who needs time to both receive the ball and shoot while pull-up shots only require time to shoot. This is why catch-and-shoot curves generally lay above pull-up ones in this representation. Moreover, pull-up shots reveal an oscillation in the player’s free space as he is dribbling, while it appears that before a catch-and-shoot shot the shooter was already free so his free space only decreases before the shot.
Figure 5: Evolution of 50 shot-free-space indicators calculated with the static model $\delta_{\text{space}}^*$ (a) and dynamic model $\delta_{\text{time}}^*$ (b) during the interval ranging from 3 seconds before to 1 second after the shot. The shot takes place at $t=0$. Pull-up shots and catch-and-shoot ones are differentiated.

While we mainly performed a visual analysis to separate shots, we now divide catch-and-shoot shots into two subclasses thanks to the k-means clustering algorithm. We then plot the average evolution of shooter’s free space for the two catch-and-shoot classes and for pull-up shots. The results are available in Figure 6. As expected, pull-up shots reveal a global decrease as the shooter dribbles over time while catch-and-shoot shooters need some free space to have time to receive the ball. However, two behaviors stand out between players who are already free three seconds before the shot and those who have to move to free themselves. Considering the two ways of evaluating free space we can see that the global trend is the same, but the time offset calculated in section 4 appears again.

Figure 6: Average evolution of a shooter’s free-space measured by (a) $\delta_{\text{space}}^*$ and (b) $\delta_{\text{time}}^*$ before a 3-point attempt. Shots are differentiated into three subclasses: pull-up shots, catch-and-shoot shots where the shooter has free space for a long time and catch-and-shoot shots where the shooter has to free himself from opponents.
5.2. Free Space and 3-point Shot Efficiency
As explained in section 2, the opening of a shot can be quantified thanks to the distance to the closest defender. Does this remark hold if we take into account the free space of the player at the time of the shoot defined as $\delta_{time}(t_{shot})$? We have computed the evolution of success rate with shooter’s free space at a shot time using a moving average procedure and the results are shown in Figure 7. Those results are similar to the results by [18] (for Figure 7(a)) and confirm that 6 feet is a key distance for successful shots. The trend in Figure 7(b) is the same: the more time a defender needs to join the shooter, the more efficient the shooter is. The equivalent of 6 feet in free space calculated with $\delta_{time}(t_{shot})$ is 0.4s: if the closest defender is within 0.5s, the shooter’s accuracy decreases drastically.

![Figure 7](image)

(a)

5.3. Comparing Catch-and-shoot Shooters
We have seen that catch-and-shoot shooters need a certain amount of free space to receive the ball and shoot. One key question is: how does this free space influence the player’s behavior once they receive the ball? The correlation between release time and free space at the catch time is a good indicator. But in that case, it is more valuable to use $\delta_{time}$ to calculate free space to take into account whether a defender is moving in the shooter’s direction at catch time. Figure 8 shows for 16 different players, ordered by success rate, the link between $\delta_{time}$ at catch time and release time. We can remark that players can have a similar success rate but have different behavior. For instance, Korver does not seem to adapt his release time to the free space he has at catch time, while Matthews, Nowitzki or Thompson clearly take more time if they are more free when they catch the ball. Moreover, this figure reveals that players like Porter or Russell only shoot if $\delta_{time}(t_{catch})$ is superior to 0.5s while Curry or Korver take shots even if they are closely guarded. Finally, let’s note that Kobe Bryant’s release time can be really high even if his free space at catch time is low. This represents the fact that he shoots even when a defender is very close to him.
As mentioned in section 2, shot difficulty can be measured by the opening of a shot e.g. the distance of the closest defender $\delta_{\text{space}}$ at the shot time. However, we can refine the shot characteristic looking at release time to compare players’ behavior and check if players shoot quickly or not and if they take open or closed attempts. Figure 9 reveals different behaviors. Players like Ilyasova, Gordon Johnson, and Leonard only take open shots (closest player at least at 4 feet). Players with a success rate higher than 40% globally shoot quicker than the median. This suggests that good shooters need to be able to shoot rapidly after receiving the ball. Among these players, we can focus on Leonard and Thompson noting that Thompson’s attempts can be very tight or wide-open for the same release time while Leonard mainly takes shots when the nearest defender is 4 to 8 feet away from him. Two players can be mentioned as they have a particular behavior. Looking at Russell’s distribution we can note that he misses almost all the shots he tries if the nearest defender is within 6 feet of him. However, he still has a global percentage of 33%, therefore, he would benefit from taking only wide-open shots. Then Bryant has also a particular behavior as most of his shots are taken with the player closer than 6 feet.
Figure 8: Shot release time function of free space at catch time calculated with $\delta_{time}$ for 16 different players ordered by success rate. Dashed lines show median release time and median $\delta_{time}$ of all catch-and-shoot shots. The regression lines show how players adjust their release time according to the free space at the catch time.
**Figure 9:** Distance to the closest defender at shot time $\delta^*_\text{space}$ function of release time. Vertical dashed line shows median release time over all catch-and-shoot shots while horizontal dashed lines delimit degrees of a shot opening defined by NBA stats website [16]: very tight, tight, open and wide open above the thickest dashed line. Hits and miss are respectively associated with blue and red dots.

**5.4. Analysis of a Shot**

Figure 10 illustrates the step-by-step analysis of a single, catch-and-shoot 3-point shot by Thompson. This is an example where the shooter has to free himself from his defender to be able to receive the ball. At the beginning of the action, Thompson is guarded by two players (Figure 10 (a)) which is confirmed by respective values of $\delta^*_\text{space}$ and $\delta^*_\text{time}$ 4 feet and 0.4 second. Then, Thompson frees
himself from defenders not only thanks to his step backward but also to Curry's movement that attracts defenders (Figure 10 (b)). Finally, he attempts the shot and succeeds with a release time of 0.5 second and the closest defender at 5 feet or 0.4 second at shot time. This is, therefore, an open shot taken rapidly.

![Figure 10: Analysis of Thompson catch-and-shoot shot.](https://amigocap.github.io/MecaSportStats/video.mp4)

6. Conclusion and Perspectives

In this section, we summarize our contribution and main takeaways based on the occupation models we introduced in this paper. Visual analysis reveals that the dynamic model is a better way to characterize court occupation. A more detailed quantitative comparison enabled to quantify the inertia with a measurable time-lag in the defensive dynamics. We have computed this time-lag, which is of the order of 0.3s focusing on 3-point attempts and 1.26s for the global team behavior, which is a fairly important amount of time in a high-frequency sport such as basketball.

The study of 3-point shooters free space evolution 3 seconds before shot time revealed three different behaviors. Firstly, we noticed that catch-and-shoot attempts require more free space than pull-up ones as the player needs time to catch the ball before taking his shot. Secondly, we were able to differentiate two types of catch-and-shoot shots as some of them ask the shooter to free himself
from defenders while he can only stay in his position when he is free enough both to receive and shoot the ball three seconds before the attempt.

Then, considering shot accuracy, we were able to determine key values: when the closest defender is within 6 feet or 0.4s the percentage of success decreases drastically.

Eventually, the study of 16 individual behaviors reveals that some players, like Matthews, adapt their release time to the free space they have when they catch the ball while others prefer to keep a constant release time. We were also able to note that some players only take their chance if they have enough free space at time catch while players like Curry can shoot even if they are closely guarded. Finally, a focus on release time indicates that players with a success rate higher than 40% are always able to shoot quicker than average suggesting it is a necessary condition to be a good shooter.

The main perspective for our work is to take into account in our models the players’ characteristics such as their height, speed or endurance. This will require additional meta-dataset to be combined with SportVU data. We sought to look at the correlation between types of shots and de-aggregated data over space and time, e.g. shot clock, current score or decisive game or not. Another improvement of our model is to take into account the direction of the nearest defender in addition to his distance since a defender not on the firing line has less influence on the shooter's performance.

We release code and analysis as an open-source project with a permissive license [https://github.com/AmigoCap/MecaSportStats](https://github.com/AmigoCap/MecaSportStats) to support further research and application of those models. In particular we are interested in applying those models to transportations systems in cities, to let urban planners optimize bus and metro lines locations and schedules based on the share of population they cover.

**Acknowledgments**

The authors are grateful for the pioneering work by Marc Mattis, which constituted the foundation of this study. We are also grateful to STATS LLC for releasing its SportVU player tracking data. This research was partially funded by the M2I project [http://www.mob2i.fr/](http://www.mob2i.fr/), Projet Investissement d’Avenir on urban mobility by the French Environment Agency (ADEME) and by a BQR (Bonus Qualité Recherche) project funding of École Centrale de Lyon.
References


We want to compute, for a moving player on a domain how much time it takes to reach a point \((x, y)\). Without loss of generality, we will determine the time it takes for a player at a given position and with a given speed to reach the point \((x, y) = (0,0)\). Initially the velocity of the player at position \((x(0), y(0))\) is \((u(0), v(0))\). We will consider that the player will use a constant force (per unit mass) in a given direction, of strength \(|F|^2 = F_x^2 + F_y^2\). This is an assumption which allows to find a simple analytical solution. In particular, it allows to consider the two directions separately.

Newton’s law writes:

\[
d^2_x x = F_x
\]

so that we have:

\[
x(t) = x(0) + u(0)t + \frac{1}{2} F_x t^2
\]

We evaluate this expression at \(x = 0\) and want to determine at which time this point is reached. Let us first determine the force per mass \(F_x\),

\[
F_x = -2 \frac{x(0) + u(0)t}{t^2}
\]

Analogous expressions to (6)-(8) are written for \(d_y y, y(t)\) and \(F_y\). Since \(|F|^2 = F_x^2 + F_y^2\), we have,

\[
F^2 = \frac{4}{t^2} \left( \left( \frac{x(0) + u(0)t}{t^2} \right)^2 + \left( \frac{y(0) + v(0)t}{t^2} \right)^2 \right)
\]

Yielding a 4th order polynomial for \(t\),

\[
t^2 - \frac{4}{F^2} \left( \left( \frac{x(0) + u(0)t}{t^2} \right)^2 + \left( \frac{y(0) + v(0)t}{t^2} \right)^2 \right) = 0
\]

This equation has formally 4 solutions. However, only one of these is the shortest physical time for a player to reach the origin. The constraints to choose the correct solution are that the time needs to be the smallest positive and real root of equation (10).

The present model contains one adjustable control parameter, the value of \(F\). A previous investigation [4] suggests that the value of \(F\) should be of order 10m/s.

In principle, since we have not bounded the velocity of a defender and fixed its acceleration, nonphysically large velocities can be developed. Thereto in a previous model [4] the dynamics were
refined introducing a drag, which limits the increase in velocity. However, for the present application, the time it takes for a defender to reach a shooter rarely exceeds 1 second, so that the velocities do not reach non-physical values. Therefore, and for the sake of simplicity, we have chosen not to refine the model any further.