Evaluating the influence of free-space on three-point shot performance in basketball

Gabin Rolland, Nathan Rivière, Wouter J.T. Bos, Romain Vuillemot

March 2019

Abstract

We evaluate the data of 584 NBA basketball matches to assess the influence of the time of which a player disposes to shoot, on his three-point shot performance. To measure the available shooting time we introduce a procedure which takes into account the velocity of the players and we determine the minimum time it takes for a defender to reach the shooting player. We show a significant influence of the available time on the performance. We further show how this result changes when the velocity of the players is not taken into account.

1 Introduction

Nowadays, data analysis has become essential in most branches of the professional sports community. A common principle in many sports is to collect data and analyze them in order to predict or improve performance. For instance, during the preparation for the man's football World Cup of 2014, Germany used, for that time, sophisticated data analysis. At the end, the German football team became world champion, which triggered the attention given to data analytics in sports. The importance of datat analysis is especially acknowledged in the the American National Basketball Association (NBA) [1] which appears as a pioneer in the field. Examples in the open literature are investigations focusing on defense [2] or shot performance [3].

Data can be processed in two different ways. Firstly, it can be used in a predictive perspective, thanks to machine learning algorithms which analyze a large amount of data, recognize patterns, and try to foresee strategies, strengths and weaknesses of adverse teams. Secondly data analysis can be used in a descriptive way, yielding statistics and insights in the past plays of a team. Many studies, either predictive or descriptive, have been carried out in sports analysis. We can mention investigations aiming at the definition of a players' influence zone, and its dependence on the players velocity [4, 5]. In a different way [6], a model was developed considering the players' velocity in a probabilistic way. These ideas were further developed [7] in order to find a measure of the pressure exerted on a football player by other players.

Other investigations focus on events occurring in a match, such as passes or shoots. For instance, different interfaces related to sports data visualization were developed, allowing 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 basketball passes on space occupation models. Finally, several studies concern shot performance in basketball [3, 11]. The first of these studies proposed an interface allowing to compare shooting performance from different players, the second one created a metric associated with the possible number of points scored in an action.

We focus on basketball. In particular in recent years, the number of 3-points shots, and thereby its importance, has increased. When the 3-point line was introduced in 1979, 2.8 shots were taken on average per game against 27 attempts per game today [12]. The influence of three-point shots has thereby radically changed and has become a major asset of the most successful teams (such as the Golden State Warriors) in the recent past. Having insights on how to improve the three-point performance and to understand the different important factors which influence the three-point performance is therefore extremely valuable. Therefore, we investigate in this study one particular factor which influences the performance of shooters. More precisely we try to define the time of which a player disposes to shoot, and consider how the shot performance depends on this time. Specifically, to compute the available time, we introduce a metric which takes into account the velocity of the players. Indeed, the inertia of a moving player will influence the time to reach a point. Intuitively in 3-points shoots, a player which experiences less defensive pressure will, on average, outperform a shooter which is guarded more closely. A question is however whether we can measure this.

The first objective of this study is to develop a model of a players' influence zone. Voronoï theory [13] can be used to measure the spatial ditribution and clustering of objects. It is currently used in different branches of physics (fluid mechanics for instance [14]) and its principle has been applied to analyze the players positions and collective dynamics in teamsports such as football [15] and basketball [16, 17]. Such analysis allows to distinguish in a formal way which parts of a domain are closest to a player on a domain. A direct application of Voronoï tessalation does not take into account the velocity of a player. Clearly if one wants to define an influence zone of a player, i.e., the part of a domain which is controlled by a player of a game such as football or basketball, it might be important to take into account the velocity, or more precisely the inertia of a player. We will show qualitatively that a refined model for a players' influence can be developed by taking into account the player's velocity in the determination of an influence zone. Once we have shown the performance on this definition, we will use this approach to compute statistics and find a statistical description of three-point performance as a function of the size of the player's influence zone.

To summarize our objectives: in this paper, we introduce a model that visually describes space occupation by basketball players. Our study allows to quantify how 'free' a player is and how this influences his 3-points shot performance.

The remainder of this manuscript will be constructed as follows. In the following section we will explain the methods we used to define the player's influence zone, taking into account inertia. In section 3 the data and its analysis is discussed. Then, in section 4 we present different visualization of our free-space model and results of different statistical studies on the influence of free-space on 3-point shot performance. In section 5 we conclude this manuscript.

2 Determination of the free-space and time of a team-player

2.1 Free-space or influence zone of a player

As our work is focusing on the free-space of a player, we need a quantitative measure of space occupation on a basketball court. Obviously to determine the influence zones of the different players, one needs to take into account the positions of all the players of the two teams. Formally, we are considering the problem of a surface occupied by 10 particles, each one associated to a player on the court. The first step in defining the influence zones is to determine for every point in space the closest particle. This allows to subdivide the court into 10 Voronoï cells, defined as follows: a point belongs to a cell if and only if it is closest to the particle in this cell than any other particle. Detailed properties on Voronoï tessalation can be found in literature [18]. To refine our description beyond the discrete nature of Voronoï diagrams, we introduce a continuous measure, taking into account the relative distance of a point to a player. This quantity, δ_d 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,

$$\delta_d(x, y) = d_{\text{closest player}} - d_{\text{closest opponent}}. \quad (1)$$

This quantity does therefore not only take into account the distance to a player, but also the fact that only players of the other team will dispute the control of a certain area. On the boundaries of a Voronoï cell between players of different teams the value of δ_d is per definition $\delta_d = 0$. The continuous nature of δ_d allows to determine a *heat-map* describing the occupation of the court by the players and thereby we are able to evaluate the free-space that each player has as a function of his position. Furthermore since the definition of δ_d distinguishes between the two different teams we also measure how "free" a player is.

2.2 influence of inertia on a player's influence zone

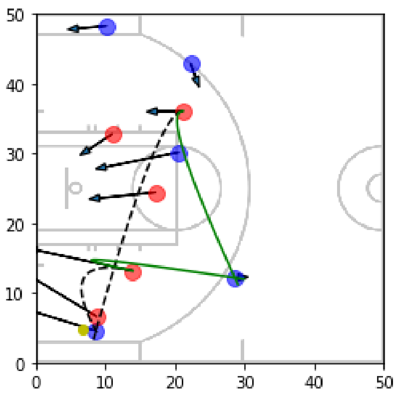


Figure 1: Example of trajectories of 2 defenders (red) to join two possible shooters (blue), computed using a model taking into account the inertia of the defenders.

The approach in the last section gives a straightforward way to quantify the way players occupy a basketball court. However, this approach does not take into account that players have inertia. 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. To take the influence of inertia into account in the definition of players' influence zones, it is necessary to develop a model characterizing the acceleration of the players. Clearly it is not the distance to a point which fully determines the control of a point, but rather the time it takes for a player to reach the point. Therefore, the boundaries between a defense-controlled zone and an attack controlled zone are more precisely defined by a quantity δ_t,

$$\delta_t(x,y) = t_{\text{attack}} - t_{\text{defense}}, \tag{2}$$

where t_{attack} is the time it takes for the closest attacker (in time) to join the point (x,y), and t_{def} the same time for a defender. In the unphysical case of players with zero mass of infinite force, this δ_t should behave as δ_d. However, inertia is expected to change this. We will now discuss the model we used, which is similar in spirit to a previously investigated model [5].

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 :

$$d_t^2 x = F_x \tag{3}$$

so that we have :

$$x(t) = x(0) + u(0)t + \frac{1}{2}F_x t^2. \tag{4}$$

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,

$$F_x = -2\frac{x(0)+u(0)t}{t^2} \tag{5}$$

Analogous expressions to (3-5) are written for $d_t^2 y$, $y(t)$ and F_y. Since $|F|^2 = F_x^2 + F_y^2$, we have,

$$F^2 = \frac{4}{t^2}((\frac{x(0)+u(0)t}{t^2})^2 + (\frac{y(0)+v(0)t}{t^2})^2) \tag{6}$$

Yielding a 4^{th} order polynomial for t,

$$t^2 - \frac{4}{F^2}((\frac{x(0)+u(0)t}{t^2})^2 + (\frac{y(0)+v(0)t}{t^2})^2) = 0 \tag{7}$$

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 (7). This model contains one adjustable control parameter, the value of F. A previous investigation [5] suggests that the value of F should be closed to 10m/s. We used this value for our work. Figure 1 shows an example of trajectories of basketball defenders to join strikers simulated with our model.

This approach refines the description of free-space in section 2.1 which was based on the position of the players only, since we now take into account the inertia of the players. 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 [5] 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.

3 Data analysis

In this investigation we used the data supplied by the company STATS and their technology SPORTSVU. The 584 NBA games, from seasons 2013-2014 to 2016-2017, in this collection, are described in a *play-by-play* format and stored into *JavaScript Object Notation* (JSON).

In the data we had no information on shot-attempts. However we have detailed data on ball-possession and the three-dimensional position of the ball. Using this we determined when a shot was taken, whether it was a three-point shot or not and whether it succeeded or not.

In particular, taking into account that the vast majority of three-point shots reaches the vicinity of the basket, we were able to detect three-point attempts by tracking back the balls that crossed a circular horizontal surface above the hoop of radius 1.5m and height 3.05m . Successful shots were detected by the presence of the ball in a small circular space underneath the hoop. We have checked manually by evaluating the video of a large number of events the efficiency of this method and we have managed to tune the parameters of our routines to successfully detect all events we considered.

After determining the three-point attempts, we differentiate a particular subclass of shots, *catch-and-shoot* shots. For these shots, which are particular in that the shooter shoots just after he receiving the ball the free-space is a determinant factor.

The aim of our work was to determine how the free-space of a shooter evolves before launching a 3-points shot. Using our routines, in 584 matches, we found 26868 shots. For each 3-points shot, we determined whether the shot succeeded or not. We found 9941 hits and 16927 misses. For all these shots we evaluated the free-space determined by δ_d and time δ_t, determined by expressions (1) and (2), respectively, and traced back in time, for three seconds, the evolution of these quantities. Our study aims at assessing the influence of "free-space" on 3-points efficiency. With the information given by our study, we try to answer specifically the following questions :

- Is there a link between free-space at the moment the shooter receives the ball and 3-points efficiency ?
- Is there a link between free-space 1 second before shot and 3-points efficiency ?

4 Results

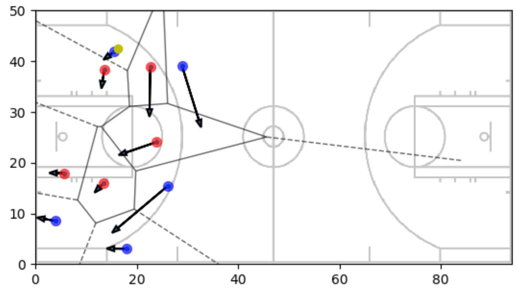
4.1 Visual comparison of different models for free-space

Figure 2 presents the visualizations of the three different types discussed in section 2. In each figure, the attackers are represented by blue dots and the defenders by red ones. The ball is colored in yellow. The velocity vector of each player is indicated in black.

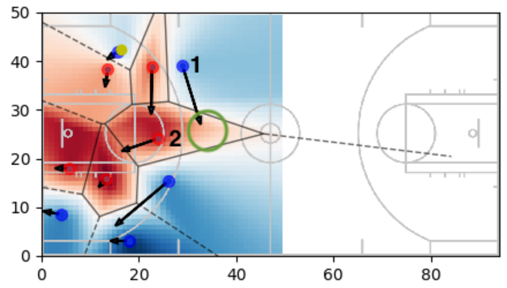
In figure 2(a), the full black lines and dotted black lines are representing the boundaries of the Voronoï cells. Figure 2(b) represents the heat-map associated with the model considering δ_d, representing in some sense the continuous refinement of a Voronoï analysis, distinguishing between attackers and defense. Finally, figure 2(c) represents the heat-map associated with δ_t, generated with the model taking into account the inertia of the players.

Clear visual differences are observed. See for instance the location on the court indicated by a green circle. In a simple (non-inertial) model, this part of the terrain is controlled by the defensive team. However, in reality it is offensive player (1) which will arrive at this point before player (2) since the velocity of player (1) is directed towards this point while defensive player (2) is moving away from this point.

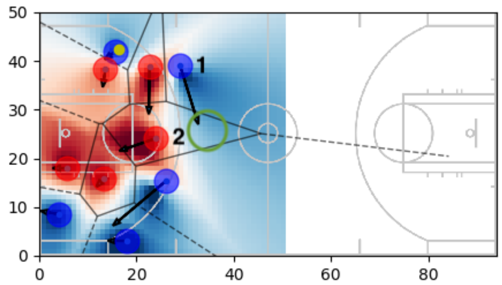
This example clearly shows how the inertia of the different players can be important in the determination of the influence zones of the teams. It is



(a)



(b)



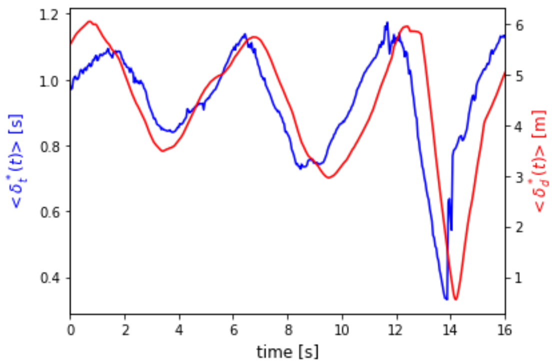
(c)

Figure 2: Analysis of a basketball situation using different visualization tools. (a) Voronoï analysis. (b) Heat map representing the *influence* of the different players as a function of distance δ_d to a point. (c) *Influence* of the players measured by δ_t as a function taking into account inertia.

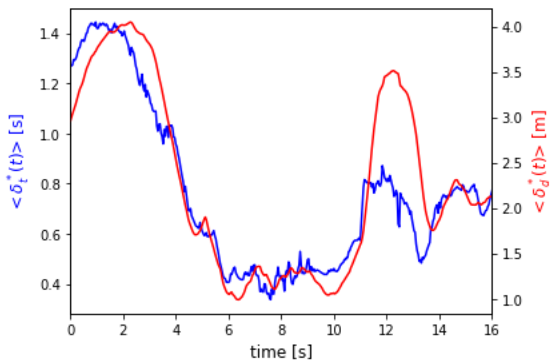
not clear *a priori* how this influences 3 point shot performance.

4.2 General evolution of free-space during a sequence of a play

Figure 3 represents free-space evolution for two random players during a specific sequence. In this figures, δ_d^* and δ_t^* are respectively defined by $\delta_d^* = \delta_d((x,y)_{shooter})$ and $\delta_t^* = \delta_t((x,y)_{shooter})$. A clear trend, in particular in figure 3(a) is that the available time for a player lags the distance. This difference in δ_d^* and δ_t^*, clearly shows how inertia introduces a delay into 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 τ_δ is of the order of 0.3 second. The precise value is obtained by measuring for which value τ_δ the maximum of the correlation coefficient ρ_δ is observed. This correlation coefficient



(a)



(b)

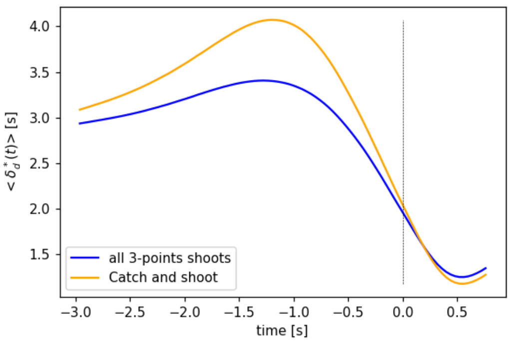
Figure 3: Evolution of free-space for two attackers during a sequence of a play. Comparison of the freespace measured using the distance to the attacker (in red) and the available time taking into account inertial (blue). Figures (a) and (b) are representing two randomly chosen 16s time intervals for two randomly chosen players.

is defined as

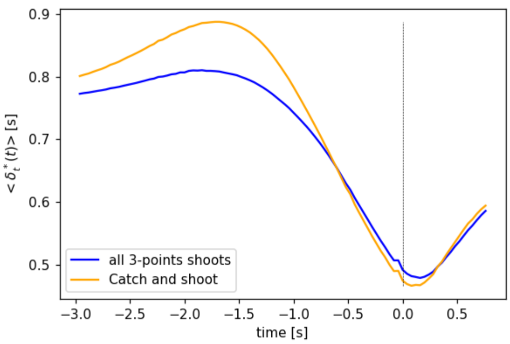
$$\rho_\delta(\tau) = \frac{\langle\langle\delta_d(t+\tau)\rangle\langle\delta_t(t)\rangle\rangle}{\langle\langle\delta_d\rangle^2\rangle^{1/2}\langle\langle\delta_t\rangle^2\rangle^{1/2}}, \quad (8)$$

where $\langle . \rangle$ denotes a time-average. The computation of this quantity will be done in future work.

4.3 General evolution of free-space before a 3-point shot



(a)



(b)

Figure 4: Average evolution of shooter's (a) free-space and (b) time before a three-point attempt. The shoot takes place at 0s. The time in (b) is computed taking into account the inertia of the player. The results are obtained by averaging over all attempts in our data. A difference is observed between the behaviour of all data, compared to *catch-and-shoot* attempts.

Figure 4(a) presents the evolution of available space δ_d^* for a three-point shooter three seconds before a 3 point shot attempt. δ_d^* is defined as the distance between the closest opponent and the shooter, averaged over all attempts of all matches (or only the catch-and-shoot attempts). Indeed, in particular the catch-and-shoot events are interesting in the present framework since those attempts need a substantial free space for the shooter, who needs time to both receive the ball and shoot.

In figure 4(b) we show the influence of the inertia on these results. In this figure we show δ_t^*, the time it takes for the closest defender to reach a shooter, taking into account his inertia, again averaged over all attempts or the subset of catch-and-shoot attempts. Even though evidently the qualitative trend is similar differences are observed in the precise shape of these graphs. In particular, it is observed that at $t = 0$, the time of the attempt, δ_t^* is closer to its minimum value than δ_d^*. This observation needs some further explanation.

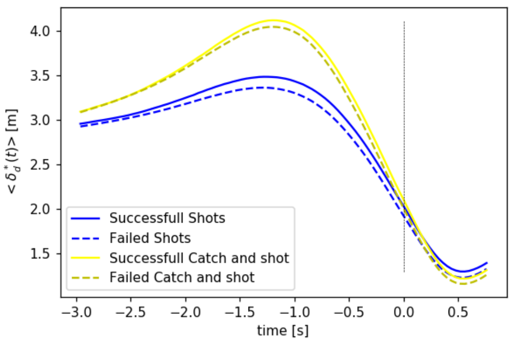
4.4 Free-space and shot efficiency

Intuitively, if a player is closely guarded by a defender, his shot performance deteriorates. Figure 5 allows a comparison between successful shots and failed ones, calculated with both methods (distance and time). Catch-and-shoots are differentiated to see whether catch-and-shoot efficiency is more influenced than normal shots.

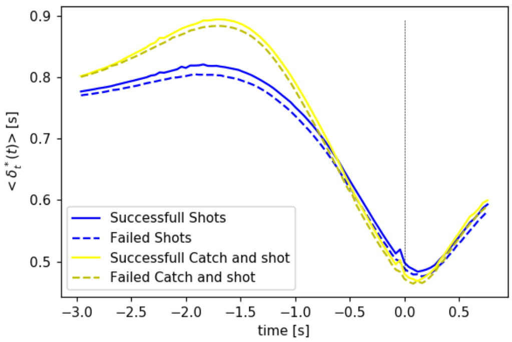
A first observation in all graphs is a small but significant systematic difference, demonstrating that having more time or space improves shot performance. This answers thus to one of our questions, i.e., whether it is possible to measure the influence of free-space on shot performance.

These results show thus an average trend of the evolution of the free-space just before a shot attempt. Obviously, however, not for all attempts the initial distance of the defense is the same. Is it than possible to see as a function of the distance of the defender at a given time before an attempt, how the shot-efficiency evolves. In practice this means: when a shooter decides to shoot, how does his performance depend on the distance of the defense 1 second before this shot, or at the time of the shot itself? We have computed this and the results are shown in Figure 6.

A clear trend is observed in all graphs. In particular it is clear that the shot performance increases

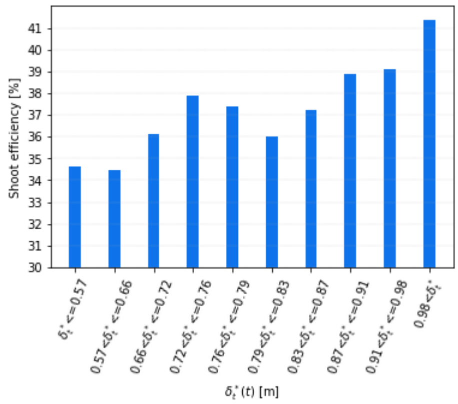


(a)

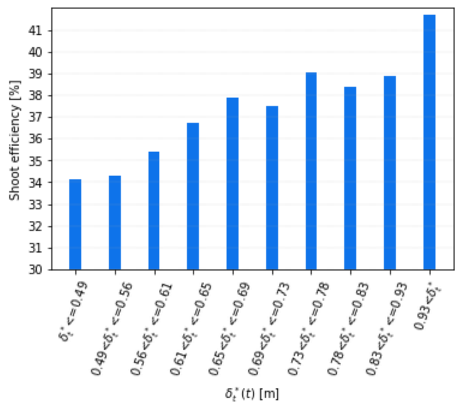


(b)

Figure 5: Influence of free-space on shot performance. Comparison of succesful and missed shots (a) Evolution of δ_d^*, the average distance to the closest defender. (b) Evolution of δ_t^* the average time it takes for the closest defender to reach the shooter.



(a)

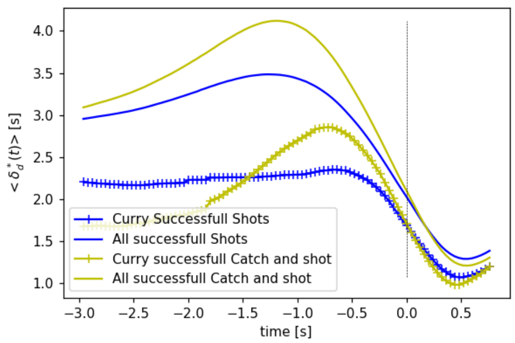


(b)

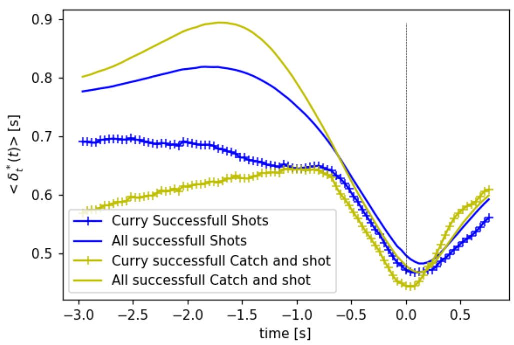
Figure 6: (a) Shot efficiency as a function of the available time δ_t^* (free space)evaluated 1 second before a shot attempt. (b) The same evaluated at the moment of reception of the ball.

from around 34% when a player is closely guarded to a maximum value slightly exceeding 40% when the defense is far away at the time of reception of the ball.

4.5 The influence of individual differences: Stephen Curry's statistics



(a)



(b)

Figure 7: Evolution of free-space (a) δ_d^* and (b) δ_t^*. Comparison between Curry's shots and global statistics.

All these statistics give a statistical overview and do not take into account the personal skills of the players. To illustrate how the statistics can be influenced by individual behaviour, we have chosen to evaluate the performance of an exceptional 3-point specialist. Stephen Curry, playing currently (and in the time where have analyzed the data for) for the Golden State Warriors is known to have a particular way to shoot. He distinguishes himself by having a rapid shot and a particular skill in shooting from far behind the 3-point line [19]. Figures 7(a) and (b) allow to compare the difference between Curry's behaviour and global 3-points shooters' behaviour before shooting. Clearly the statistics are very different and both values of δ_d^* and δ_t^* are always smaller than the global statistics. This reflects in particular the fact that Curry is always very closely guarded by the defense since the opponents know that when he is given too much space, the chance of a successful 3-point shot is very large.

5 Conclusion

Our study investigated the evolution of free space and its influence on three-point shot performance in basketball. We have assessed quantitatively the intuitive observation that a player disposing of more free-space or time when attempting a 3-point shot is more likely to succeed.

We have analyzed two different space occupation models taking into account, or not, the inertia of players on the occupation and control of a basketball court. We have shown that this inertia induces a measurable time-lag in the defensive dynamics. We have computed this time-lag, which is of the order of $0.3s$, which is a fairly important amount of time in a high-frequency sport such as basketball.

The qualitative trend of the evolution of free space of shooters does not dramatically depend on the inertia of the players, but the time lag is clearly observable at the time of shoot.

We have also quantified how the amount of free-space determines the 3-point average and it is observed that this quantity evolves from 34% to over 40% for shots which are taken in the limits of small and large amounts of available free-space, respectively.

Acknowledgements. The authors are grateful for the pioneering work by Marc Mattis, which constituted the foundation of this study, and interesting discussions with Mickael Bourgoin.

Appendix: visualization tool

We have designed and implemented a visualization tool to efficiently analyze the results obtained using our approach and to communicate the outcome with non-scientists experts. Thereto we developed a court-visualization, where players who can potentially shoot are distinguished from the others. The marking in time is represented by circles around each attacker. The more free a player is, the larger the circle is. For high values of free-space, the circle is colored in green, while for lower values it is colored in yellow. Thereby, for a 3-points action, one can see whether the shooter has enough free-space to attempt a shoot. One visualization can be found at the following link `https://raw.githubusercontent.com/AmigoCap/MecaFootCo/master/vid.mp4`.

References

[1] Moreyball: The houston rockets and analytics. `https://digit.hbs.org/\submission/moreyball-the-houston-\rockets-and-analytics/`, 2018.

[2] Luke Bornn an al. Alexander Franks, Andrew Miller. Counterpoints: Advanced defensive metrics for nba basket ball. *MIT : Sports analytics conference*, February 2015.

[3] Kirk Goldsberry. Courtvision: New visual and spatial analytics for the nba. In *2012 MIT Sloan sports analytics conference*, volume 9, pages 12–15, 2012.

[4] M. Stein, H. Janetzko, T. Breitkreutz, D. Seebacher, T. Schreck, M. Grossniklaus, I. D. Couzin, and D. A. Keim. Director's cut: Analysis and annotation of soccer matches. *IEEE Computer Graphics and Applications*, 36(5):50–60, Sep. 2016.

[5] Akira Fujimura and Kokichi Sugihara. Geometric analysis and quantitative evaluation of sport teamwork. *Systems and Computers in Japan*, 36(6):49–58, 2005.

[6] Javier Fernandez, F Barcelona, and Luke Bornn. Wide open spaces: A statistical technique for measuring space creation in professional soccer. In *Sloan Sports Analytics Conference*, 2018.

[7] Gennady Andrienko, Natalia Andrienko, Guido Budziak, Jason Dykes, Georg Fuchs, Tatiana von Landesberger, and Hendrik Weber. Visual analysis of pressure in football. *Data Mining and Knowledge Discovery*, 31(6):1793–1839, November 2017.

[8] Yisong Yue, Patrick Lucey, Peter Carr, Alina Bialkowski, and Iain Matthews. Learning Fine-Grained Spatial Models for Dynamic Sports Play Prediction. In *2014 IEEE International Conference on Data Mining*, pages 670–679, Shenzhen, China, December 2014. IEEE.

[9] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. SoccerStories: A Kick-off for Visual Soccer Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, December 2013.

[10] Manuel Stein, Thorsten Breitkreutz, Johannes Haussler, Daniel Seebacher, Christoph Niederberger, Tobias Schreck, Michael Grossniklaus, Daniel Keim, and Halldor Janetzko. Revealing the Invisible: Visual Analytics and Explanatory Storytelling for Advanced Team Sport Analysis. In *2018 International Symposium on Big Data Visual and Immersive Analytics (BDVA)*, pages 1–9, Konstanz, October 2018. IEEE.

[11] Dan Cervone, Alexander D'Amour, Luke Bornn, and Kirk Goldsberry. Pointwise: Predicting points and valuing decisions in real time with nba optical tracking data. In *Proceedings of the 8th MIT Sloan Sports Analytics Conference, Boston, MA, USA*, volume 28, page 3, 2014.

[12] Konstantinos Kotzias. The evolution of 3-point shot and its influence on the game. `https://statathlon.com/the-evolution-of-3-point-shot-and-its-\influence-on-the-game/`, November 2017.

[13] Georges Voronoi. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. deuxième mémoire. recherches sur les parallélloèdres primitifs. *Journal für die*

reine und angewandte Mathematik, 134:198–287, 1908.

[14] Romain Monchaux, Mickaël Bourgoin, and Alain Cartellier. Preferential concentration of heavy particles: a voronoï analysis. *Physics of Fluids*, 22(10):103304, 2010.

[15] Tsuyoshi Taki, Jun-ichi Hasegawa, and Teruo Fukumura. Development of motion analysis system for quantitative evaluation of teamwork in soccer games. In *Proceedings of 3rd IEEE international conference on image processing*, volume 3, pages 815–818. IEEE, 1996.

[16] António Lopes, Sofia Fonseca, Roland Lese, and Arnold Baca. Using voronoi diagrams to describe tactical behaviour in invasive team sports: an application in basketball. *Cuadernos de Psicología del Deporte*, 15(1), 2015.

[17] Sofia Fonseca, João Milho, Bruno Travassos, and Duarte Araújo. Spatial dynamics of team sports exposed by voronoi diagrams. *Human movement science*, 31(6):1652–1659, 2012.

[18] Atsuyuki Okabe, Barry Boots, Kokichi Sugihara, and Sung Nok Chiu. *Spatial tessellations: concepts and applications of Voronoi diagrams*, volume 501. John Wiley & Sons, 2009.

[19] Highnessfsk. Sport Science: Stephen Curry. `https://www.youtube.com/watch?v=HOiH1eVCggw`, May 2013.