López, F.A., Martínez, J.A. y Ruiz, M. (). Análisis espacial de lanzamientos en baloncesto; el caso de L.A. Lakers / Spatial pattern analysis of shot attempts in basketball;the case of l. A. Lakers. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 13 (51) pp. 585-613. Http://cdeporte.rediris.es/revista/revista51/artanalisis409.htm

## ORIGINAL

# SPATIAL PATTERN ANALYSIS OF SHOT ATTEMPTS IN BASKETBALL; THE CASE OF L.A. LAKERS

# ANÁLISIS ESPACIAL DE LANZAMIENTOS EN BALONCESTO; EL CASO DE L.A. LAKERS

**López, F.A.[1], Martínez, J. A.[2],& Ruiz, M.[3]**

**English version:** López, F.A.[1], Martínez, J.A.[2] y Ruiz, M.[3]

[1] fernando.lopez@upct.es, Profesor Titular. Departamento de Métodos Cuantitativos e Informáticos. Universidad Politécnica de Cartagena. España
[2] josean.martinez@upct.es, Profesor Contratado Doctor. Departamento de Economía de la Empresa. Universidad Politécnica de Cartagena. España
[2] manuel.ruiz@upct.es, Profesor Titular. Departamento de Métodos Cuantitativos e Informáticos. Universidad Politécnica de Cartagena. España

**Acknolwdegments:** We are in debt to Ryan J. Parker, from www.basketballgeeek.com, for freely sharing his database about shot locations. Thanks also to www.basketball-reference.com. Finally, we thank Julián Felipo, Serhat Ugur and Martí Casals for their helpful comments in previous drafts of this article.
Fernando A. López (Fernando.lopez@upct.es has been partially supported by the project ECO2009-10534/ECON from Ministerio de Ciencia y Tecnología of Spain, and the project 11897/PHCS/09 de la Fundación Séneca Agencia de Ciencia y Tecnología of the Región de Murcia
Jose A. Martínez (josean.martinez@upct.es) has been partially supported by the project 11860/PPC/09 from Fundación Seneca of the Región de Murcia
Manuel Ruiz Marín (manuel.ruiz@upct.es) has been partially supported by the project MTM2009-07373 from Ministerio de Ciencia e Innovación of Spain, Fundación para la Formación e Investigación Sanitarias and Fundación Seneca of the Región de Murcia

**Código UNESCO / UNESCO Code:** 5899 Educación Física y Deporte / Physical Education and Sport
**Clasificación Consejo de Europa / Council of Europe Classification:** 17 (Otras: Análisis cuantitativo del deporte) / Other: Quantitative analysis of sport

**Recibido** 22 de julio de 2011 / **Received** July 22, 2011
**Aceptado** 15 de diciembre de 2011 **Accepted** December 15, 2011

## ABSTRACT

The importance of quantitative analysis in sports using objective data (such as game statistics), has been had prominent in recent years. In this paper we have shown an application of spatial statistics to understand more thoroughly the game of basketball. This methodology has been rarely used in sports research, specifically in basketball. We have depicted how a spatial clustering technique, such as the Kulldroff test, which is widely employed in epidemiology, can be applied to analyze basketball data. This test detects low and high incidence clusters of shots, and therefore it better characterizes the game of teams and individual players. In addition, we have also used a test based on entropy, the V-test, which serves to statistically compare shooting maps. We illustrate the interesting contribution of this methodological perspective in the case of the analysis of the Lakers' performance, showing the transformation of this team from a medium-level NBA franchise into a champion team, because of, among other factors, the incorporation of two key players in the 2007-08 season: Pau Gasol and Derek Fisher.


**KEYWORDS:** Basketball, Cluster of shots, Spatial analysis, NBA

**RESUMEN**

La importancia del análisis cuantitativo del deporte a través del uso de estadísticas de juego ha sido un área de creciente interés en los últimos años. En esta investigación mostramos una aplicación de la estadística espacial al entendimiento del juego en baloncesto. Esta metodología ha sido infrautilizada en la investigación en deporte hasta la fecha, especialmente en baloncesto. Así, describimos una técnica de segmentación espacial, como el test de Kulldroff, que es ampliamente utilizado en epidemiología, y puede ser aplicado para analizar estadísticas de juego en baloncesto. Este test detecta clusters de lanzamientos de baja y alta incidencia, caracterizando mejor el juego de los equipos y de los jugadores. Además, desarrollamos un test basado en entropía, el test V, que permite comparar los mapas de lanzamiento. Para ilustrar empíricamente el atractivo de esta aproximación metodológica, aplicamos estos análisis al rendimiento del equipo de la NBA: Los Angeles Lakers. Este equipo vivió una enorme transformación en las últimas temporadas, pasando de ser una franquicia de rendimiento medio a ganar la NBA, debido a, entre otros factores, la incorporación de dos jugadores clave en la temporada 2007-08: Pau Gasol y Derek Fisher.

**PALABRAS CLAVES:** Baloncesto, clusters de lanzamientos, análisis espacial, NBA.

## 1. INTRODUCTION

The importance of quantitative analysis in sports using objective data such as game statistics has been prominent in recent years. In fact, academic and professional attention to sports analysis has grown exponentially since the appearance of the "Moneyball" phenomenon (Lewis, 2003). In the specific field of basketball, the significance of this topic is outstanding, and has been the subject matter of articles in academic journals (e.g. Berri, 1999; Berri & Bradbury, 2010; Esteller-Moré & Eres-García, 2002; Kubatko, Oliver, Pelton & Rosenbaum, 2007), seminal books (e.g. Oliver, 2004; Winston, 2009), prospective books (e.g. Doolittle & Pelton, 2009; Hollinger, 2005), and a plethora of specialized websites (e.g. www.apbrmetrics.com, www.hoopsstats.com, www.nbastuffer.com, www.basketball-reference.com, www.82games.com, www.basketballvalue.com).

There are three main streams of research when using quantitative analysis in basketball to better understand the game through statistics: The first one is related to the valuation of teams and players performance, in order to get a more objective view about productivity, efficacy, efficiency and the value of the players in the game. Some of the most outstanding research materials about this topic are: Berri (1999; 2008), Berri & Eschker (2005), Fernández, Camerino, Anguera & Jonsson (2009), Hoon-Lee & Berri (2008), Esteller-Moré & Eres-García (2002), Mavridis, Tsamourtzis, Karipidis & Laios (2009), Rimler, Song & Yi (2010), Piette, Annand & Zang (2010). The second one is concerned with obtaining accurate predictions, in order to minimize the risk of decision making units (managers, coaches, etc.). Some of the more relevant studies on this topic are the following: Alferink, Critchfield, Hitt & Higgins (2009), Berri, Brook & Schmidt (2007), Berri & Schmidt (2002), Berry, Reese & Larkey (1999), Hitt, Alferink, Critchfield & Wagman (2007), Romanowich, Bourret & Vollmer (2007), Sánchez, Castellanos & Dopico (2007), Skinner (2010), Tauer, Guenther & Rozek, (2009), Trininic, Dizdar & Luksic (2002), Vollmer & Bourret (2000). Finally, a third stream of research has focused on the analysis of controversial themes such as competition imbalance, manipulation of games, salary determination, race discrimination, and other miscellaneous topics, such as the momentum of teams or player's hot-hand. A sample of these studies are: Arkes, (2010), Berri, Brook, Frick, Fenn & Vicente-Mayoral (2005), Balsdon, Fong & Thayer (2007), Fort, Hoon-Lee & Berri (2008), Fort & Maxcy (2003), Humphreys (2000; 2002), Michaelides (2010), Price &Wolfers (2010), Vergin (2000), Zimmer & Kuethe (2007).

However, few studies have analysed basketball games from a spatial perspective, beyond stats appearing in box-scores. The progressive inclusion of shot-location coordinates in play-by-play data in the most important basketball competitions around the world (see Martínez, 2010) facilitates data analysis using spatial statistics. Nevertheless, as far as we know,a few studies, such as Hickson & Waller (2003) and Reich, Hodges, Carlin & Reich (2006) have used this perspective. Both studies only analysed the performance of a single player (Michael Jordan and Sam Cassell, respectively). As Piette, Sathyanarayan &

Kai (2010) explain, the first study models each shot chart as an instance of some Poisson process and estimated the corresponding nonparametric functions relating to each event. The second research applies a Bayesian multivariate logit model to spatial data along with an added set of covariates. To determine the model parameters, sampling is done via a Monte Carlo Markov Chain method. The results from these two studies are helpful examples of the capabilities of this type of approach.

Although spatial statistics are progressively being incorporated to analyse sports data (e.g. Mulrooney, 2007), and there is a continuous improvement in generating shot location data in basketball (Chen, Tien, Chen, Tsai & Lee, 2009), it is necessary to rely on a powerful tool to understand spatial patterns of shooting in order to help coaches and analysts to evaluate the game and to make low-risk decisions. In addition, this may complement other research about space-time coordination dynamics of basketball teams (Bourbousson, Sève & McGarry, 2010; Jäger & Schöllhorn, 2007), or shooting abilities (Piette, Sathyanarayan & Kai, 2010). Consequently, the procedure we introduce in this study is a novel approach to enrich information obtained from play-by-play data of basketball games or video tracking of players.

In this research, we use a different approach in order to analyse spatial data from shot attempts location.. Therefore, the new contributions of our research are the following: Firstly, we apply our analysis to the shots attempted by whole teams and specific players. Specifically, we centre our analysis on the NBA team Los Angeles Lakers (LAL). As a result, we are able to detect differences in shot location patterns from the LAL to the whole league and from specific LAL players within their own team. Secondly, we use the Kulldorff test (Kulldorff, 1997) to analyse spatial data, in order to detect low and high incidence clusters of shot locations, which is a novel approach in sport science. This tool allows us to find significant different shooting patterns, and to visually show these areas in order to compare intuitively the performance of disparate teams and players. Obviously, the statistical approach of this clustering process provides much more information regarding shooting performance than a mere descriptive analysis of shot location (such as, for example, analysis made in specialized sites such as www.82games.com/shotzones.htm or http://hoopdata.com/shotstats.aspx). As we show in our empirical application, the utility of spatial analysis can help to analyse the transformation of the LAL from a middle NBA team (2006-07 season) into a NBA champion team (2008-09 season), because of (among other factors) the incorporation of two key players in the 2007-08 season: Pau Gasol and Derek Fisher.

Therefore, in this research we provide answers to questions such as: *Does the LAL team have a different spatial pattern of shooting than the rest of the NBA teams? Or similarly, does player P have a different spatial shooting pattern than the rest of his team?* In addition, when analyzing the spatial pattern of shots attempted by player P, an interesting question is: *Does there exist a high frequency (low frequency respectively) spatial shooting cluster of the LAL (resp. player P) differing to the one expected by chance with underlying population all*

*NBA shots (resp. team shots)?* Notice that by obtaining the high and low frequency shooting clusters one is also able to find out whether the incorporation of player P has changed the spatial shooting pattern of his team (or other players within the team) by comparing the cluster between the season in which player P has been incorporated to the team and the previous season

## 1.1. THE LAL TRANSFORMATION

In the season 2006/07 the LAL was a team living in a transition period after 5 years of success (four NBA finals and three championships from 2000 to 2004), and then 2 years of failures (2004/05 and 2005/06). In these 2 unsuccessful seasons, the LAL did not qualify for play-offs in the former, and was eliminated in the first round in the latter. Therefore, the LAL began the 2006/07 season with several doubts regarding team performance. Note that the LAL are one of the most glamorous franchises in NBA history, and always have to perform to high rates of expectation. Therefore, these poor results were not consistent with the LAL expectations. In that season, the first season of our analysis, the LAL performed more poorly than the prior season (0.51 of Win-Loss record against 0.54), and also was eliminated in the first-round of the play-offs.

Next season (2007/08) the LAL made important changes in its roster. They signed guard Derek Fisher in July 20, 2007; forward Tevor Ariza in November 20, 2007; and power-forward Pau Gasol in February 1, 2008. These were the most important player movements in that season, because other signed players had a marginal presence in the roster (see all these player movements and stats in www.basketball-reference.com). In contrast, one of the most important players of the previous season, Smush Parker, signed in July, 2007 as a free agent with Miami Heat, and other important player as Maurice Evans were traded in order to sign Trevor Ariza.

Finally, in the following season (2008/09), there were no relevant changes in the roster, because the few player movements had very little effect in the distribution of minutes per game on the remaining players. In these three seasons, Kobe Bryant and Lamar Odom were always in the top three of the team in minutes played, so they can be considered as reference factors in team performance for the period of time analyzed. Obviously, it is pertinent to remember that Kobe Bryant is one of the most important players in the League, and one of the best players in NBA history.

The Lakers had a 0.51 Win-Loss record in the 2006/07 season, a 0.69 in 2007/08 and a 0.79 in 2008/2009. In the two latter seasons, the LAL played and won the NBA final. The transformation of team performance was evident, with the 2007/08 season being the peak of the point of change. Specialized analysts agree that the incorporation of Fisher and Gasol was crucial for such a transformation (e.g. Bresnahan, 2010; Kleeman, 2009; Manning, 2009; Sanderson, 2010). Note too that Trevor Ariza only played 24 games in that season because of an injury. Before signing Gasol, the LAL had a 0.65 Win-

Loss record in the first 46 games of the season. After signing Gasol, and considering only the 27 games where Gasol played (he missed the remaining 9 games of the season due to an injury), the LAL achieved a 0.84 Win-Loss record.

In this research we analyse whether this obvious change in the LAL performance is reflected in disparate patterns of shot locations, or if the arrival of Gasol and Fisher spatially changed the LAL game, and whether players such Kobe Bryant or Lamar Odom changed their shooting model.

Normal box-score stats partially reflect such a transformation (Table 1). The LAL improved both offensive and defensive performances, but the percentages of field goals and free throws were much the same in the three seasons (with a slight improvement). Therefore, changes in the LAL offensive-defensive differential were mainly caused by the increment of games. As the LAL increased rebounds, steals and decreased turnovers, they could shoot with more assiduity, and then they improved points per game. Note that, in the second season, attempted field goals increased by 2.89 per game, and in the third season by 1.98. Consequently, offensive play changed in a more relevant form than the defensive game.

**Table 1.** Box-score stats of the LAL and Opponents in the three regular seasons.

| | | FGA | FG% | 3PA | 3P% | FTA | FT% | TRB | AST | STL | BLK | TOV | PTS | PTS/G | W-L% |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2006/07* | LAL | 6581 | 0.466 | 1702 | 0.353 | 2193 | 0.747 | 3339 | 1827 | 593 | 416 | 1257 | 8368 | 102 | 0.512 |
| | Opponents | 6635 | 0.461 | 1444 | 0.358 | 2291 | 0.76 | 3420 | 1776 | 641 | 404 | 1193 | 8374 | 102 | |
| 2007/08 | LAL | 6818 | 0.476 | 1751 | 0.378 | 2270 | 0.769 | 3620 | 2003 | 654 | 438 | 1156 | 8904 | 108.6 | 0.695 |
| | Opponents | 7022 | 0.445 | 1531 | 0.362 | 1995 | 0.752 | 3509 | 1793 | 634 | 368 | 1168 | 8309 | 101 | |
| 2008/09 | LAL | 6981 | 0.474 | 1516 | 0.361 | 2087 | 0.77 | 3602 | 1908 | 718 | 420 | 1103 | 8768 | 106.9 | 0.793 |
| | Opponents | 6825 | 0.447 | 1700 | 0.345 | 1931 | 0.753 | 3399 | 1854 | 635 | 392 | 1275 | 8140 | 99.3 | |

*FGA: Field Goal Attempted; FG%: Field Goal Percentage; 3PA: Three Points Attempted; 3P%: Three Points Percentage; FTA: Free Throw Attempted; FT%: Free Throw Percentage; TRB: Total Rebounds; AST: Assists; STL: Steals; BLK: Blocks; TOV: Turnovers; PTS: Points; PTS/G: Pointes Per Game; W-L%:Win-Loss Percentage.*

* The LAL played 250 extra-minutes in this season compared with the following two seasons, due to overtime. Therefore, we have averaged the data of that season in order to make numbers comparable.

Source: www.basketball-reference.com

Although these stats are very useful to understand the change in the LAL performance, we will show how spatial analysis can be a powerful tool to complement information provided by basic stats, offering new insights about how the LAL changed their offensive and defensive game, through change in shooting locations.

## 2. METHOD

We downloaded the data base of play-by-play statistics including shot location from www.basketballgeek.com. This is the only free source of this type of data, because other play-by-play sources do not include shot location coordinates, and sources of shot charts are not coded in readable data base format. Data of three whole regular seasons (from 2006/07 to 2008/09) were downloaded. As this data is not official stats, there are some missing games. Therefore, a total of 3,509 from the 3,690 games played were available. This represents less than 5% of games missing, which we consider an admissible number for not affecting the results. Data was read and filtered using the MATLAB 2010b package. As all games were coded in the play-by-play format, a validation process was achieved in order to ascertain that attempted shots codified by www.basketballgeek.com and the official stats matched. More than 99% of concordance was found among the downloaded data base and official stats of the LAL and about 95% of the whole NBA. Therefore, we consider the data base to be reliable. Regarding the coordinates (x,y) of shot location, the interpretation is as follows: If you are standing behind the offensive team's hoop then the X axis runs from left to right and the Y axis runs from bottom to top.

We consider the basketball court to be a regular 51 $\times$ 94 lattice. Therefore, each cell has an approximate size of 30 $\times$ 30$cm^2$. We will concentrate the analysis on the shots taken from just before the mid court and excluding the farthest shots and those in the three lines behind the basket. Thus our lattice will have a total of R=51 $\times$ 33 = 1683 locations. We only consider the nearest 33 lines to the rim, once the two first lines where there is a negligible number of shots are excluded.

### *2.1. KULLDORFF TEST*

To detect spatial non-random clusters we will use the Kulldorff test. The procedure of this test is to impose a window on the map and move the window centre over each point location so that the window includes different sets of neighbouring points at different positions. By adjusting the centre location and radius, the method generates a large number of distinct windows, each including a different set of neighbouring points. At each point location, the size of the window is increased continuously from '0' to a user-defined maximum size. The maximum-size parameter sets an upper bound on the window's radius in one of two ways: (1) by specifying the maximum percentage of the total population within the window or (2) by specifying the geographic extent of the circle. Option (1) is used in the research reported here. Due to the court shape we have used elliptic Windows with a maximum size of a 5% of the total shots attempted.

The null hypothesis tested by the Kulldorff test is that in all locations the probability of an attempted shot is the same while the alternative hypothesis is that there exists a window W so that the probability of an attempted shot inside W is different from the one outside W. Now we will introduce some notation which is needed to follow the mathematical description of the test.

Let n be the total number of shots attempted by player P. Let $n_s$ and $n_{W\ be}$ the total number of shots attempted by player P at location s and in window W respectively. Let $N_s$, $N_W$ and N be the total number of the team shots attempted at location s, window W and in the whole basketball court respectively. Notice that the variable $X_s$ counting the number of shots attempted by player P at location s distributed as a Binomial distribution B($N_s$,$p_s$) where $p_s$ is the shooting probability of player P at location s. Then the null hypothesis and the alternative hypothesis can be stated as

$H_0$: $p_s = p$ for all s∈ ☹

$H_1$: There exists a window W such that $p_s = p_W$ for all s∈W and $p_s = q_W$ for all $s \notin W$ with $p_W \neq q_W$

respectively.

Therefore under the null hypothesis $H_0$ the joint distribution of the R variables $(X_1, X_2, \ldots, X_R)$ is a multinominal distribution with likelihood function $\frac{N!}{n_1! \cdots n_R!}\left(\frac{n}{N}\right)^n$ while under $H_1$ the likelihood function remains as $\frac{N!}{n_1! \cdots n_R!}\left(\frac{n_W}{N_W}\right)^{n_W}\left(\frac{n-n_W}{N-N_W}\right)^{n-n_W}$. Then the likelihood ratio statistic in the window W is:

$\lambda_W = \left(\frac{n_W}{E_W}\right)^{n_W}\left(\frac{n-n_W}{n-E_W}\right)^{n-n_W}$ where $E_W = \frac{N_W n}{N}$ is the expected value under $H_0$ of shots attempted in window W.

Then the Kulldorff statistic for high and low frequency attempted shots is defined as the maximum of the values $\lambda_W$ with W running all possible windows in the lattice ☹, that is,

$$Ku_{high} = \sup_W \left\{ \left(\frac{n_W}{E_W}\right)^{n_W}\left(\frac{n-n_W}{n-E_W}\right)^{n-n_W} I\left(\frac{n_W}{E_W} > \frac{n-n_W}{n-E_W}\right) \right\}$$

$$Ku_{low} = \sup_W \left\{ \left(\frac{n_W}{E_W}\right)^{n_W}\left(\frac{n-n_W}{n-E_W}\right)^{n-n_W} I\left(\frac{n_W}{E_W} < \frac{n-n_W}{n-E_W}\right) \right\}$$

where I(x) is the indicator function taking the value 1 if the logic function x is true and 0 otherwise. To evaluate the statistical significance of the *primary cluster*, a large number of random replications of the data set are generated under the null hypothesis. The p-value is obtained through a Monte Carlo hypothesis testing (Dwass 1957), by comparing the rank of the maximum likelihood from the real data set with the maximum likelihoods from the random data sets. If this rank is r, then the p-value = r/(1 + # simulations). By repeating this procedure and eliminating the selected window we can detect secondary clusters. There is a free software available to run the Kulldorff test, SatScan, which can be downloaded from www.satscan.org.

In a similar way, using a likelihood ratio test one can design a nonparametric test to detect global differences in the frequency of the spatial shooting pattern. This statistic test is based on entropy measures and tests for the null hypothesis that the spatial shooting frequency of team A is equal to the spatial shooting frequency of team B, against any other alternative. The statistic is

$\hat{V} = 2N\ h(A,B) + h(A \cup B) - h(A) - h(B)$ that asymptotically follows a $\chi^2$ distribution and the construction can be found in the Appendix section.

In order to help the reader to understand the statistical procedure we illustrate the Kulldorff test with an easy example. Consider a 4×4 regular lattice. Assume that we have a team composed only of two players, player A and player B with the attempted shot distribution shown in Figure 1.

**Figure1**. Basic example of the Kulldorff test

Player A

| | | | |
|---|---|---|---|
| 10 | 8 | 2 | 0 |
| 8 | 4 | 0 | 0 |
| 2 | 0 | 1 | 1 |
| 0 | 0 | 1 | 1 |

Player B

| | | | |
|---|---|---|---|
| 1 | 1 | 1 | 1 |
| 1 | 1 | 0 | 0 |
| 1 | 0 | 4 | 6 |
| 1 | 0 | 6 | 8 |

Team %

| | | | |
|---|---|---|---|
| 0.16 | 0.13 | 0.04 | 0.01 |
| 0.13 | 0.07 | 0.00 | 0.00 |
| 0.04 | 0.00 | 0.07 | 0.10 |
| 0.01 | 0.00 | 0.10 | 0.13 |

Under this setting we have N=70 with $n_A$=38 and $n_B$=32. Then we can estimate the shooting probability of a player just by dividing the total number of shots at location s divided by N, for instance at location 1 we have the shooting probability at 0.16=(10+1)/70 (see third column of Figure 1). Also under $H_0$ the shooting probability of player A at any location is $p_A$=38/70=0.54 and of player B is $p_B$=32/70=0.46

Now assume that window W under scrutiny is the shadow one. Then, under $H_0$ the expected value of shots attempted by player A is $E_W=34 \cdot p_A$= 18.45 while the real shot frequency is 30. Therefore window W is a high frequency cluster for player A. On the other hand, under $H_0$ the expected value of shots attempted by player B is $E_W=34 \cdot p_B$= 15.54 while the real shot frequency is 4 and therefore window W is a low frequency cluster for player B. The Kulldorff statistic scans all possible windows in the basketball court looking for the maximum difference between the expected shot frequency and the real one and afterwards obtains the p-value through a Monte Carlo hypothesis test.

For this example one can compute the entropies and the V-statistic values obtaining h(A)=1.38, h(B)=1.31, h(AUB)=2.31, h(A,B)=0.69 and V=41.85. Notice that, as expected, V=41.85 rejects the null hypothesis of equal shooting distribution.

## 3. RESULTS

In order to know the spatial pattern of shots attempted we have considered all the games played in the regular seasons 06-07; 07-08; 08-09. A total of 3,509 games with 563,740 shots attempted.

Figure 2a shows the spatial pattern of shots attempted in the NBA. Figure 2b shows the spatial pattern of shots attempted by the LAL in the same three seasons. The locations with 0 shots are shown in dark blue, which change to red as the frequency of shots increase.

**Figure 2**. Shooting frequency of the NBA and the Lakers in the 3 seasons considered

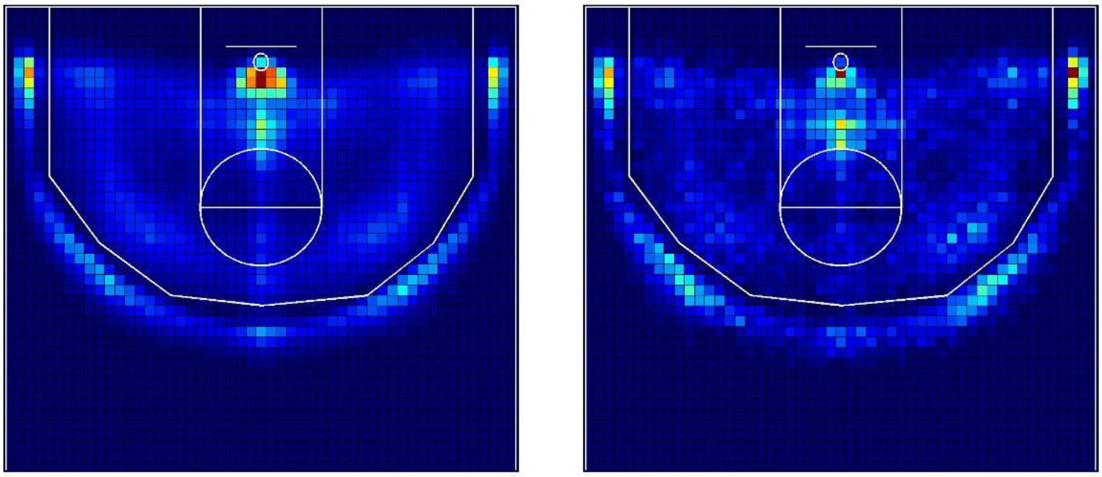


Figure 2a: the NBA    Figure 2b: the Lakers

One location exists with coordinates (26,7) with higher shot frequency, with a large difference to the remaining locations. This cell contains approximately 26% of the shots attempted for the whole NBA teams and almost 30% for the LAL. In addition, both maps show that the three point zones placed at either end of the court have high shot frequency. This is a logical finding because this three point zone is nearer to the rim than the three point zone placed in front of the rim, due to the non-symmetrical characteristic of the three point line. In the first place, we may say that both maps show a similar pattern of shot locations. However, after applying the V test, we detect significant differences between both maps (V:2627.5; df:1589; p-value:<0.001). We may generally say that the Lakers have been less oriented to the game into the paint that the aggregate NBA teams.

We are also interested in the spatial pattern of shots attempted by four of the most important players in the LAL, Kobe Bryant (KB), Pau Gasol (PG), Derek Fisher (DF) and Lamar Odom (LO). As can be expected the spatial pattern is different among the 4 players. Again these differences are logical, because of the different characteristics of these players, their team role, and their player style. Figure 3 shows the spatial pattern of shots attempted by these four players, and Table 2 shows the results of the V test.

Kobe Bryant is a total player from the offensive viewpoint. He shoots from all the zones, although he slightly prefers to be oriented about 60-75 degrees to the right of the rim. On the other hand, Pau Gasol plays as a power-forward and as a center, and he shoots near the rim, inside the paint, and preferably oriented to the left. Dereck Fisher is specialized in three point shooting, especially from low angles at the right and the left of the rim, beyond the three-point line. And Lamar Odom, who is also a power-forward as is Gasol, is a more versatile player, where the game outside the paint is very important, including shooting from the three point zone.

**Figure 3:** Shooting frequency of the four most important players of the LAL

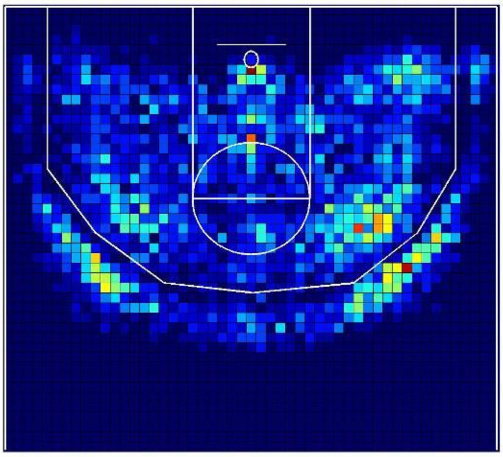


Figura 2a: KB

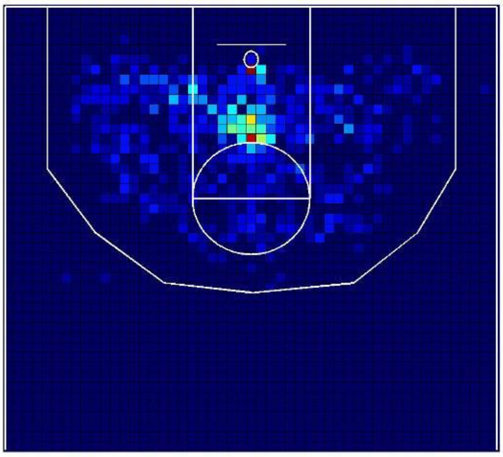


Figura 2b: PG

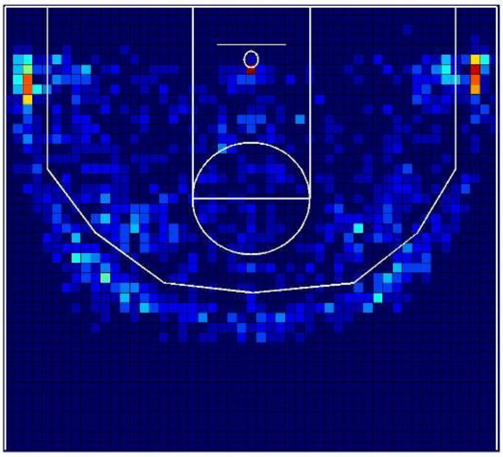


Figura 2c: DF

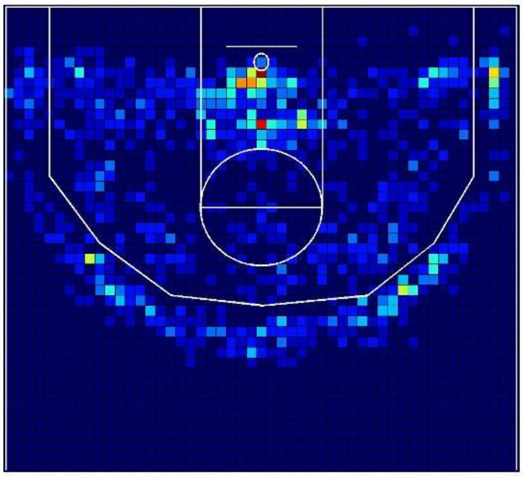


Figura 2d: LO

**Table 2:** Values of the V test to detect global differences in spatial pattern of shooting attempts

| | | LAL | KB | PG | DF | LO |
|---|---|---|---|---|---|---|
| | V | 2627.05 | 2523.93 | 2010.77 | 1943.78 | 1802.93 |
| NBA | p-value | <0.001 | <0.001 | <0.001 | <0.001 | <0.0011 |
| | df | 1589 | 1589 | 1589 | 1589 | 1589 |
| | V | | 2662.96 | 2015.17 | 1876.64 | 1696.24 |
| LAL | p-value | | <0.001 | <0.001 | <0.001 | <0.001 |
| | df | | 1344 | 1344 | 1344 | 1344 |
| | V | | | 2019.81 | 1859.39 | 2088.97 |
| KB | p-value | | | <0.001 | <0.0010 | 0.00 |
| | df | | | | 1199 | 1185 |
| | V | | | | 2158.18 | 1424.41 |
| PG | p-value | | | | <0.001 | <0.001 |
| | df | | | | 852 | 785 |
| | V | | | | | 1798.22 |
| DF | p-value | | | | | <0.001 |
| | df | | | | | 901 |

Note: Chi-square degrees of freedom have been adjusted deleting cells containing zeros.

As we previously explained, analysts agree that the incorporation of Gasol and Fisher in the 2007/08 season were  key factors in the transformation of the LAL. Therefore, we are going to analyse the previous and the following seasons in order to see if this hypothetical change has been reflected in the structure of shooting.  Figure 4 shows the high and low incidence clusters of the LAL against the NBA in the 2006/07 and 2008/09 seasons. Recall that high incidence clusters can be interpreted as the preferred shooting locations, where shots attempted by the LAL are statistically above the expected shots of the whole NBA, and the opposite for the low incidence clusters.

**Figure 4.** Low and high incidence clusters for the LAL against the NBA

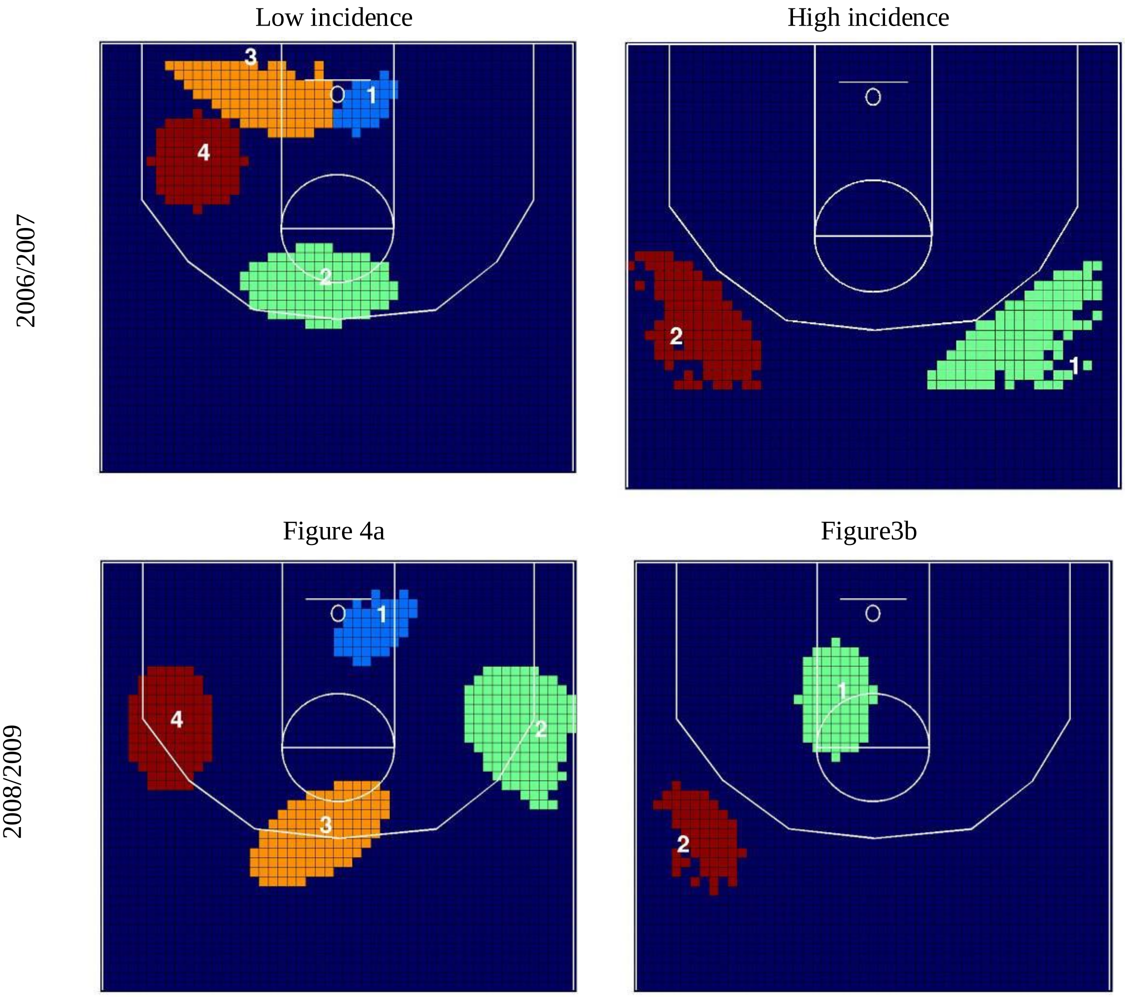



Figure 4a Figure3b

Figure 4c Figure 4d

As Figure 4a shows, in the 2006/07 season, the Lakers had two low incidence clusters near to the rim (clusters 1 and 3). The second low incidence cluster is placed between the free-throw line and the three-point line, and the fourth cluster is sited in a large two-point zone in the middle left of the court. It seems that the game inside the paint is not the LAL's preferred location for shooting. However, in the 2008/09 season there was an important change in the spatial pattern of shots (Figure 4c), because clusters 2 and 4 of the Figure 4a move toward the 3 point-line (clusters 3 and 4 of Figure 4c), and most importantly, cluster 2 moves to another distinct location, in the middle-right of the court opposite to cluster 4.

Regarding high incidence clusters, in the 2006/07 season, the Lakers had two clusters in the middle angle of the three-point line (clusters 1 and 2). However, in the 2008/09 season there was another important change, because cluster 1 moves to inside the paint.

Considering that Pau Gasol usually plays in a zone located on the left side of the rim, and obviously inside the paint, changes in both low and high incidence clusters seem to reflect the importance of Gasol in the LAL shooting pattern. Therefore, the incorporation of Gasol has made the Lakers a more powerful team inside the paint with respect to the aggregated NBA teams. Table 3 shows some statistical properties of clusters.

**Table 3.** Statistics of low and high incidence clusters for the LAL against the NBA

| Season | Cluster | Nº | Size | $O_W$ | $E_W$ | Ku | p-value |
|---|---|---|---|---|---|---|---|
| 2006/2007 | High | 1 | 119 | 371 | 230.56 | 35.56 | 0.001 |
| | | 2 | 107 | 331 | 202.34 | 34.93 | 0.001 |
| | Low | 1 | 29 | 100 | 247.83 | 57.71 | 0.001 |
| | | 2 | 109 | 153 | 289.70 | 38.28 | 0.001 |
| | | 3 | 98 | 185 | 308.16 | 28.33 | 0.001 |
| | | 4 | 81 | 191 | 272.95 | 14.33 | 0.004 |
| 2008/2009 | High | 1 | 46 | 187 | 325.72 | 35.74 | 0.001 |
| | | 2 | 132 | 198 | 310.06 | 23.76 | 0.001 |
| | | 3 | 111 | 215 | 321.44 | 19.68 | 0.001 |
| | | 4 | 99 | 233 | 327.06 | 15.43 | 0.001 |
| | Low | 1 | 78 | 493 | 311.95 | 46.26 | 0.001 |
| | | 2 | 60 | 151 | 92.31 | 15.89 | 0.002 |

*Size: Number of cells in the cluster; $O_W$ = Number of attempted shots observed in the cluster. $E_W$ = Expected shots in cluster W. Ku = Kulldorff statistic value.*

### 3.1 LAKERS OPPONENT GAME IN THE SEASONS 2006/2007 AND 2008/2009

As basketball is 50% attack and 50% defence, we also analyse the opponents pattern of shot locations, in order to obtain a more complete picture of the LAL transformation. Following the same methodology, we show results in Figure 5. Recall that we are comparing the shooting pattern of the Lakers opponents in the 2006/07 season with the shooting pattern of the Lakers opponents in the 2008/09 season, i.e. after the incorporation of Gasol and Fisher. If we focus only on cluster changes, cluster 2 of low incidence in the 2006/07 moves to the zone where Gasol is usually placed in defence in the 2008/09 season. In addition, cluster 1 becomes bigger. On the other hand, cluster 3 of high incidence (placed inside the paint) disappears. Again, globally these results seem to indicate the important contribution of Gasol to the LAL, because the opponents pattern of shot location have significantly changed in the main area of influence of the Gasol game. Table 4 depicts some statistical properties of clusters.

**Figure 5**. Low and high incidence clusters for the LAL' opponents

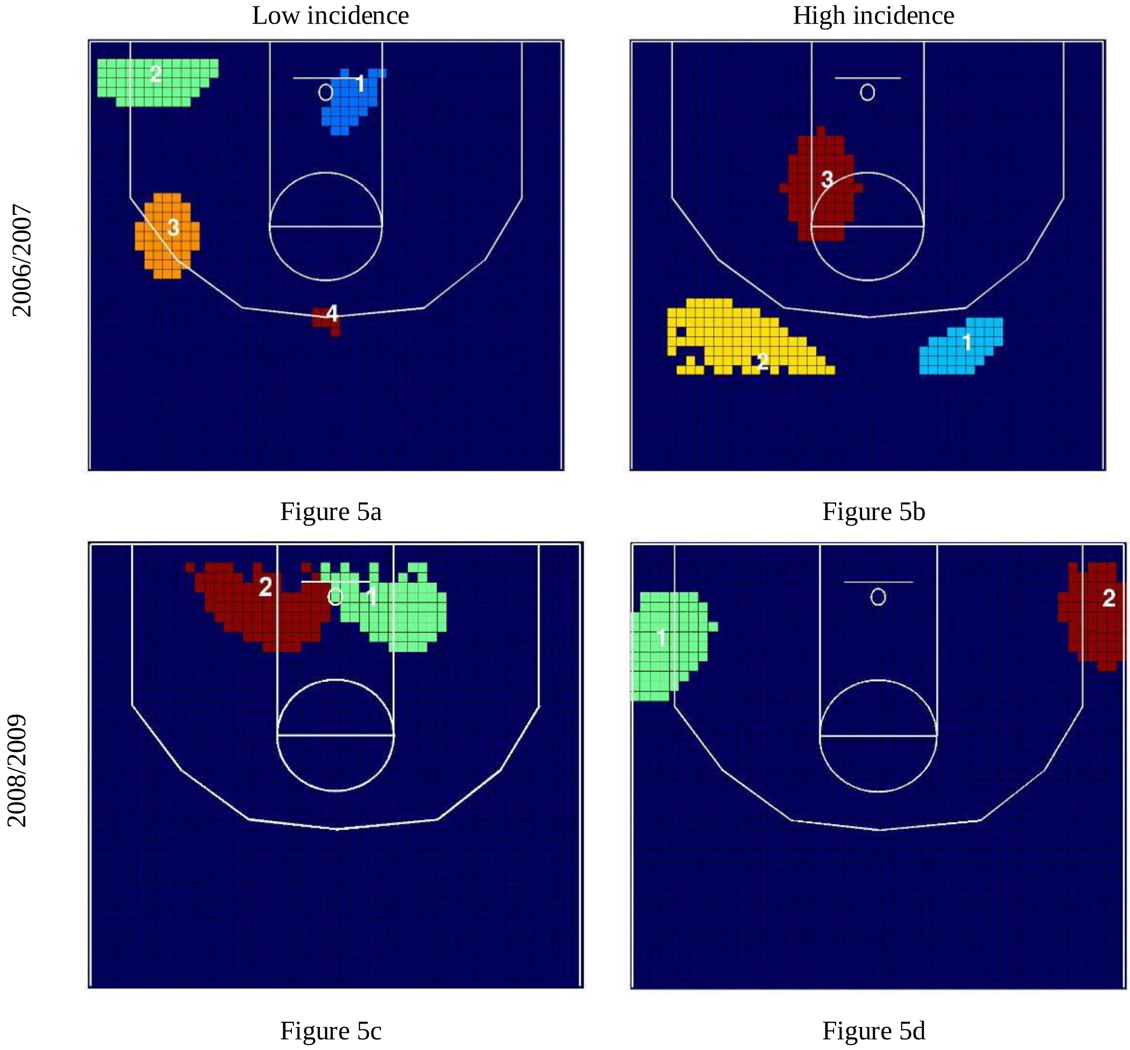



**Table 4**. Statistics of low and high incidence clusters for the LAL' opponents

| Season | Cluster | Nº | Size | $O_W$ | $E_W$ | Ku | p-value |
|---|---|---|---|---|---|---|---|
| 2006/2007 | High | 1 | 111 | 217 | 130.25 | 23.51 | 0.001 |
| | | 2 | 39 | 111 | 53.56 | 22.48 | 0.001 |
| | | 3 | 72 | 269 | 180.25 | 19.52 | 0.001 |
| | | 4 | 27 | 100 | 60.49 | 10.73 | 0.054 |
| | Low | 1 | 34 | 87 | 209.76 | 47.07 | 0.001 |
| | | 2 | 57 | 106 | 167.35 | 13.13 | 0.017 |
| | | 3 | 6 | 0 | 11.82 | 11.16 | 0.078 |
| 2008/2009 | High | 1 | 76 | 418 | 312.56 | 16.64 | 0.002 |
| | | 2 | 61 | 306 | 223.94 | 13.74 | 0.003 |
| | Low | 1 | 76 | 179 | 307.55 | 32.34 | 0.001 |
| | | 2 | 87 | 183 | 294.41 | 24.89 | 0.001 |

*Size: Number of cells in the cluster; $O_W$ = Number of attempted shots observed in the cluster. $E_W$ = Expected shots in cluster W. Ku = Kulldorff statistic value.*

### *3.2 THE GAME OF THE MOST IMPORTANT LAKER'S PLAYERS BEFORE AND AFTER THE INCORPORATION OF PAU GASOL*

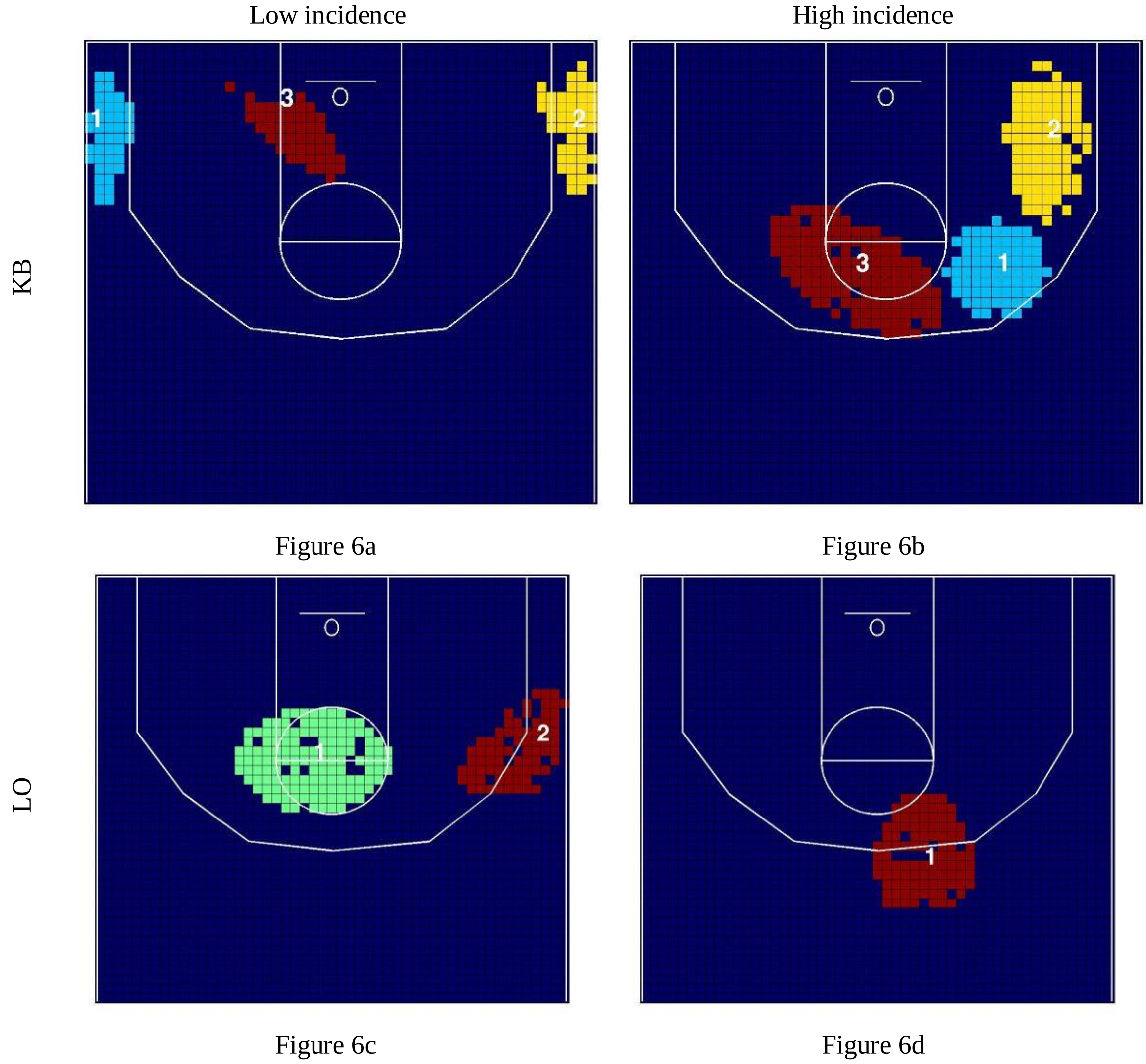
Beyond the influence of Gasol on the Lakers team, it is also interesting to study if there is any change in the game of specific LAL players. As commented previously, Bryant and Odom were the most prominent players in the pre-Gasol period. We thus analyse the spatial shooting patterns of these two players for the 2006/07 season and the games of the 2007/08 season played before the incorporation of Gasol. Results of the Kulldorff test and the graphic visualization of clusters are shown in Table 5 and Figure 6, respectively.

**Table 5.** Statistics of low and high incidence clusters for KB and LO in the Pre-Gasol period.

| Season | Cluster | Nº | Size | $O_W$ | $E_W$ | Ku | p-value |
|---|---|---|---|---|---|---|---|
| KB | Alta | 1 | 74 | 244 | 126.22 | 45.99 | 0.001 |
| | | 2 | 91 | 214 | 123.56 | 27.17 | 0.001 |
| | | 3 | 137 | 207 | 125.42 | 22.19 | 0.001 |
| | Baja | 1 | 42 | 19 | 86.03 | 37.00 | 0.001 |
| | | 2 | 46 | 34 | 92.14 | 23.51 | 0.001 |
| | | 3 | 49 | 48 | 105.42 | 17.61 | 0.001 |
| LO | Baja | 1 | 11 | 0 | 15.15 | 12.21 | 0.028 |

*Size: Number of cells in the cluster; $O_W$ = Number of shots attempt observed in the cluster; $E_W$ = Expected shots in cluster W; Ku = Kulldorff statistic value.*

**Figure 6**. Low and high incidence clusters for KB and LO in the pre-Gasol period.




Therefore, in the pre-Gasol period, Bryant prefers to shoot from the centre and right side of the court, whilst he shoots relatively less from the three-point zone located near the court corners. On the other hand, Odom prefers to shoot around the three-point zone in front of the rim instead of shooting from a nearer zone.

Nevertheless, the most interesting analysis comes from the comparison of the pre-Gasol period with the Gasol period (from February of 2008 to the end of the 2009 season), i.e. whether changes have been produced in teammates of Gasol after his incorporation to the team. Results of the Kulldorff test and the graphic visualization of clusters are shown in Table 6 and Figure 7, respectively.

**Table 6**. Statistics of low and high incidence clusters for KB, PG, DF and LO in the Gasol period.

| Season | Cluster | Nº | Size | $O_W$ | $E_W$ | Ku | p-value |
|---|---|---|---|---|---|---|---|
| KB | Alta | 1 | 95 | 187 | 99.31 | 32.04 | 0.001 |
| | | 2 | 90 | 172 | 100.50 | 21.83 | 0.001 |
| | | 3 | 88 | 133 | 72.87 | 16.66 | 0.001 |
| | Baja | 1 | 60 | 19 | 86.69 | 39.21 | 0.001 |
| | | 2 | 53 | 29 | 78.83 | 20.24 | 0.001 |
| | | 3 | 31 | 45 | 99.54 | 19.59 | 0.001 |
| PG | Alta | 1 | 28 | 166 | 62.07 | 62.59 | 0.001 |
| | | 2 | 85 | 124 | 59.17 | 27.01 | 0.001 |
| | | 3 | 40 | 59 | 27.37 | 11.27 | 0.027 |
| | Baja | 1 | 72 | 0 | 64.52 | 66.19 | 0.001 |
| | | 2 | 79 | 0 | 64.52 | 64.85 | 0.001 |
| | | 3 | 120 | 0 | 64.52 | 64.85 | 0.001 |
| | | 4 | 73 | 0 | 61.00 | 61.23 | 0.001 |
| | | 5 | 70 | 1 | 61.77 | 54.80 | 0.001 |
| | | 6 | 63 | 15 | 54.89 | 18.24 | 0.001 |
| DF | Alta | 1 | 74 | 105 | 42.44 | 34.28 | 0.001 |
| | | 2 | 104 | 99 | 40.60 | 30.13 | 0.001 |
| | | 3 | 60 | 89 | 42.14 | 21.04 | 0.001 |
| | | 4 | 83 | 84 | 42.85 | 15.49 | 0.002 |
| | Baja | 1 | 27 | 8 | 41.83 | 20.06 | 0.001 |
| | | 2 | 92 | 13 | 42.75 | 14.51 | 0.002 |
| LO | Alta | 1 | 23 | 37 | 13.02 | 14.71 | 0.002 |
| | Baja | 1 | 31 | 7 | 40.78 | 21.68 | 0.001 |
| | | 2 | 87 | 9 | 40.07 | 17.17 | 0.001 |
| | | 3 | 86 | 8 | 34.07 | 11.91 | 0.028 |
| | | 4 | 83 | 15 | 42.51 | 11.63 | 0.037 |

*Size: Number of cells in the cluster; $O_W$ = Number of attempted shots observed in the cluster; $E_W$ = Expected shots in cluster W; Ku = Kulldorff statistic value.*

**Figure 7**. Low and high incidence clusters for KB and LO in the Gasol period.

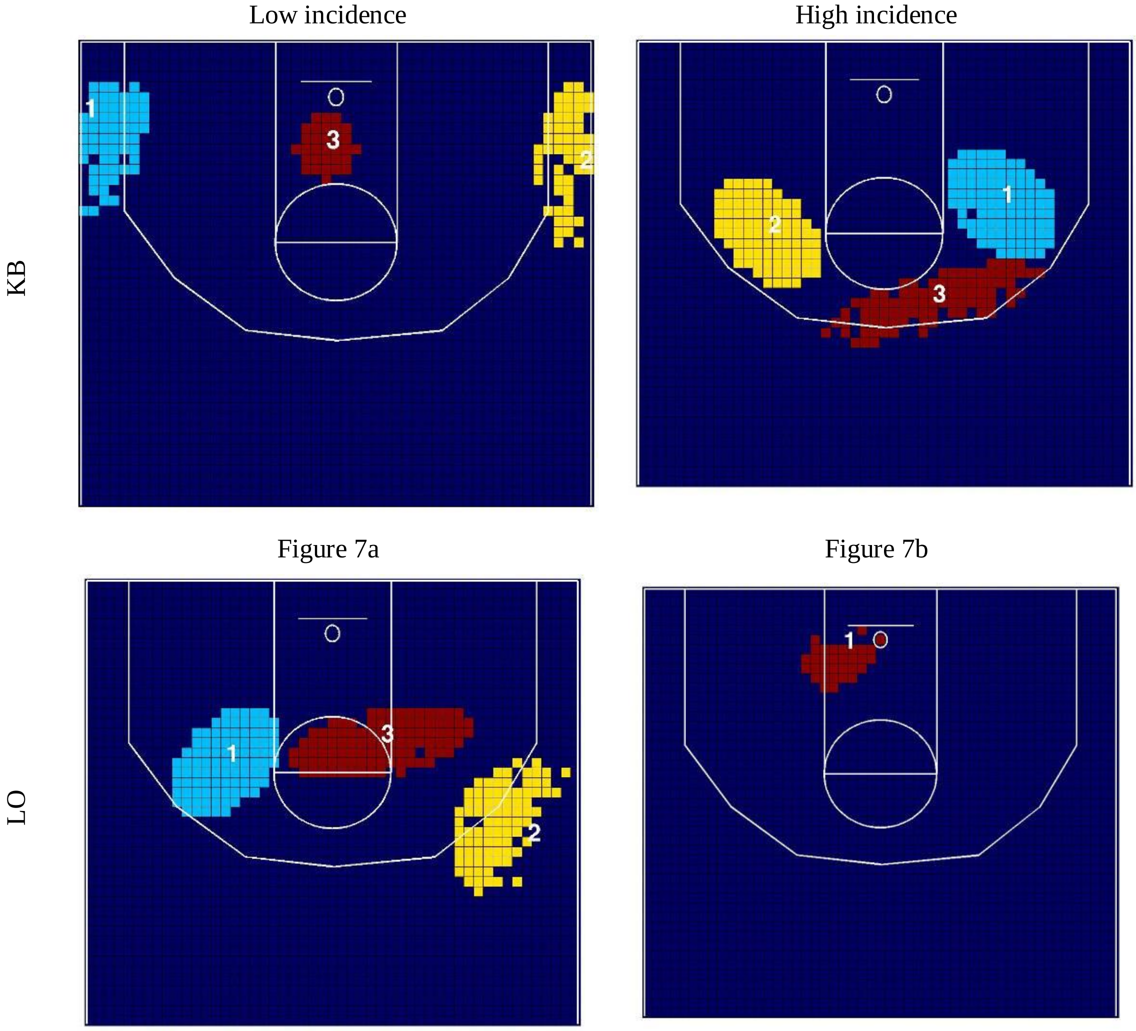



Figure 7a Figure 7b

Figure 7c Figure 7d

As can be seen, significant changes are found. Regarding Bryant, the high incidence zone moves to the left of the court, being more oriented to the centre of the court than before, where there is very little difference from the left and the right angles (though slightly prefering the right).

With regard to Odom, he has translated the high incidence cluster inside the paint, so the incorporation of Gasol has made Odom shoot nearer the rim. One possible explanation to this increase of shots attempted can be attributed to the augment of offensive rebounds achieved by Odom since Gasol's incorporation. Note that Odom averaged 1.8 offensive rebounds per game in the 2006/07 season, and 2.14 in the portion of the 2007/08 season played without Gasol. Since Gasol arrival, Odom got 3.02 offensive rebounds in the 2007/08 season and 2.30 in the 2008/09 season. It seems evident the increase of the Odom's contribution to this part of the Lakers game since the Gasol trade.

Regarding Gasol, there are three high incidence clusters located inside the paint and in the 4-5 meters zone. It seems clear that Gasol slightly prefers to shoot from the left side of the rim. Finally Fisher shows a shooting pattern consistent with his speciality, the three-point shoot. Figure 8 shows the clusters of Gasol and Fisher.

**Figure 8**. Low and high incidence clusters for PG and DF in the Gasol period.

Low incidence | High incidence

PG

DF

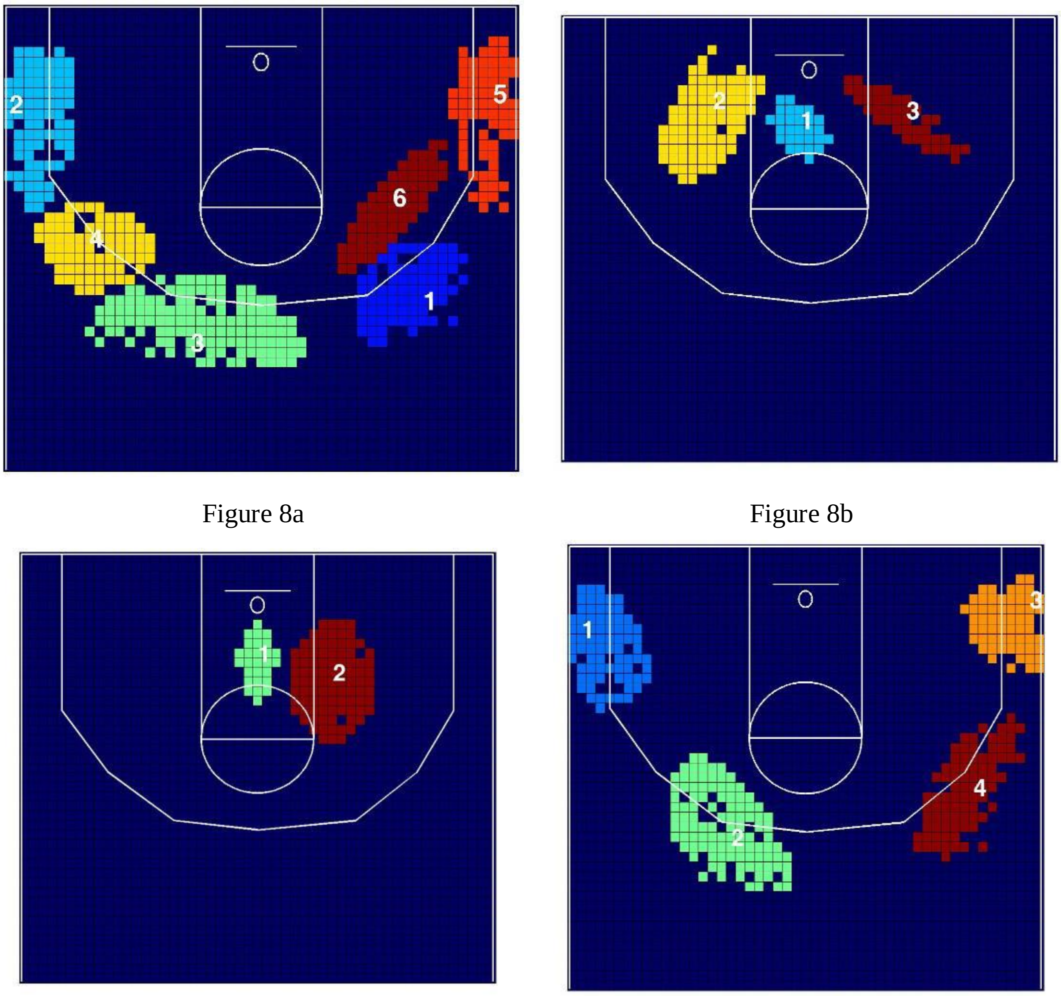


Figure 8a | Figure 8b

Figure 8c | Figure 8d

## 4. DISCUSSION, LIMITATIONS AND FURTHER RESEARCH

In this paper we have shown an application of spatial statistics to understand the game of basketball more thoroughly. This methodology has been underutilized in sports research, and specifically in basketball. We have shown the interesting contribution of this methodological perspective in the case of the analysis of the Lakers performance, and the transformation of this team from a medium-level NBA franchise into a championship team.

We have depicted how a spatial clustering technique, such as the Kulldorff test, which is widely used in epidemiology, can be applied to analyse basketball data. This test detects low and high incidence clusters of shots, and therefore characterizes the game of teams and individual players much better. In addition, we have also used a test based on entropy, the V-test, which serves to compare shooting maps statistically.

Combining both methods, we have found how the incorporation of two players: Dereck Fisher, and especially Pau Gasol, has been associated with the change in the shooting pattern of the Lakers. In addition, not only the offensive game has been affected, but the defensive game too, because the shooting pattern of opponents has also changed. The figure of Pau Gasol emerges as the main reason for these changes, because of the appearance of some clusters in the zone on the court where Gasol is used to play. The Lakers have intensified the game inside the paint after Gasols arrival, and have probably caused opponents to shoot less than expected in a zone where Gasol use to play in defence. In addition, some particular players, such as the superstar Kobe Bryant has changed his game, because, in the case of Bryant, high incidence clusters have moved by some degrees to the left of the court.

All the information derived from this spatial analysis should complement other basic and advanced stats which can be freely found in specialized websites. These stats, based on box-scores and play-by-play data, together with spatial data, must serve to get a complete picture of the performance of teams and players. In fact, as Ballard (2009) explains, some teams such as Houston Rockets, use it to inform its players about the opponents style of play (e.g. they inform the specialized defensive player Shane Battier about the play of opponents stars). Using spatial statistics as we have throughout this research, may provide very useful information for that purpose, because of the distinction among high and low incidence clusters of shots.

This research can be extended to other teams by applying a similar approach. An example could be the nemesis of the Lakers, the Boston Celtics, who made a similar transformation in the 2007/08 season, winning the NBA against the Lakers. Recall that Boston had a very poor 0.29 Win-Loss percentage (the second worst in its history) in the 2006/07 season. Then, this franchise signed two All-star players: Ray Allen and Kevin Garnett, and the performance of the next season was completely the opposite; Celtics got a 0.80 Win-Loss percentage (the third best in its history). In addition, comparisons of a team against another team or a group of similar teams can also be achieved. For example, it could be interesting to achieve a kind of benchmark analysis by comparing the Lakers with a cluster of teams in the same “strategic group”, in order to avoid the noise caused by teams low performance teams, such as teams who have not got play-offs. We have done such an analysis (it is available from the authors upon request) and we found similar results when comparing it to the whole NBA. Again a high incidence cluster appeared in the 2008/09 season in the zone where Gasol used to play.

The perspective we adopt in this research is one of several ways to incorporate spatial analysis into the understanding of the game. For example, other research may analyse the shooting patterns of shots made, in order to detect high and low incidence clusters of high-percentage and low-percentage locations. In addition, size of cells can be increased in order to obtain more data per location, but accepting the risk of being less than exact in the assignment of shots to spatial locations. However, simulations may be achieved in order to explore the consistency of the clusters obtained under different cells size.

Therefore, this research is a new contribution to sports sciences and basketball analysis, where only some studies such as Hickson & Waller (2003) or Reich, Hodges, Carlin & Reich (2006) have tried to study this theme, although from a different perspective. Other research such as Piette, Sathyanarayan y Kai, (2010) deserves to be highlighted, regarding shooting analysis in basketball. However, it is necessary to note that not all the shoots are equally difficult, and the probability to make may vary for the same location, depending on the circumstances of the shot. This is one of the reasons our study focuses only on the shot attempts.

## 5. CONCLUSIONS

In this research, we have depicted how a spatial clustering technique, such as the Kulldroff test, which is widely used in epidemiology, can be applied to analyse basketball data. This test detects low and high incidence clusters of shots, and therefore better characterizes the game of teams and individual players. In addition, we have also used a test based on entropy, the V-test, which serves to statistically compare shooting maps. We show the interesting contribution of this methodological perspective in the case of the analysis of a Lakers performance, and the transformation of this team from a medium-level NBA franchise into a champion team because of, among other factors, the incorporation of Pau Gasol in the 2007/08 season.

Therefore, this transformation not only have been manifested in the box-score statistics, and other performance indexes or basic shot location data, but in the game itself, in the pattern of shots of the team and in the pattern of shots of the opponents. Consequently, we advocate for continue researching using this approach in order to apply a tool which allows to overcome descriptive analysis of shot location.

**REFERENCES**

Alferink, L. A., Critchfield, T. S., Hitt, J. L., y Higgins, W. J. (2009). Generality of the matching law as a descriptor of shot selection in basketball. *Journal of Applied Behavior Analysis*, *42* (3), 595-608.

Arkes, J. (2010). Revisiting the Hot Hand Theory with free throw data in a multivariate framework. *Journal of Quantitative Analysis in Sports*, 6 (1), Article 2.

Arkes, J. y Martinez, J. A. (2011). Finally, Evidence for a Momentum Effect in the NBA. *Journal of Quantitative Analysis in Sports*, *7* (3) Article 13.

Ballard, C. (2009). *The art of a beautiful game*. Simon y Shuster: New York

Balsdon, E., Fong, L., y Thayer, M. (2007). Corruption in College basketball? Evidence of tanking in postseason conference tournaments. *Journal of Sports economics, 8* (1), 19-38

Bar-Elia M, Avugos S. y Raab M. (2006). Twenty years of “hot hand” research review and critique. *Psychology of Sport and Exercise*, *7*, 525-553.

Berri, D. J. (1999). Who is 'most valuable'? Measuring the player's production of wins in the National Basketball Association. *Managerial and Decision Economics*, *20*, 411-427.

Berri, D. J. (2008). A simple measure of worker productivity in the National Basketball Association. in *The Business of Sport*, eds. Brad Humphreys and Dennis Howard, editors, 3 volumes, Westport, Conn.

Berri, D. J., y Bradbury, J. C. (2010). Working in the land of metricians. *Journal of Sports Economics*, *11* (1), 29-47.

Berri, D. J., y Eschker, E. (2005). Performance when it counts? The myth of the prime time performer in professional basketball. *Journal of Economic Issues*, *1*, 798-807

Berri, D. J., y Schmidt, M. B. (2002). Instrumental versus bounded rationality: a comparison of Major League Baseball and the National Basketball Association. *Journal of Socio-Economics*, *31*, 191-214.

Berri, D. J., y Schmidt, M. B. (2006),. On the road with the National Basketball Association's superstar externality. *Journal of Sports Economics*, *7* (4), 347-358.

Berri, D. J., Brook, S. L, Frick, B, Fenn, A. J., y Vicente-Mayoral, R. (2005) The short supply of tall people: competitive imbalance and the National Basketball Association. *Journal of Economic Issues*, *39* (4), 1029-1041

Berri, D. J., Brook, S. L., y Schmidt, M. B. (2007). Does One Simply Need to Score to Score? *International Journal of Sport Finance, 2* (4), 190-205

Berry, S. M, Reese, C. S., y Larkey, P. L. (1999) Bridging different eras in sports. *Journal of the American Statistical Association*, *94*, 661-686.

Bodvarsson, O. B., y Brastow, R. T. (1998). Do employers pay for consistent performance?: Evidence from the NBA. *Economic Inquiry*, *36* (1), 145-160.

Bourbousson, J., Sève, C., y McGarry, T. (2010). Space-time coordination dynamics in basketball: Part 2. The interaction between the two teams. *Journal of Sports Sciences, 28* (3), 349-58

Breshnahan, M. (2010, 31 January). Pau Gasol's trade to Lakers changed the NBA. Descargado desde http://articles.latimes.com/2010/jan/31/sports/la-sp-lakers-gasol31-2010jan31

Chen, H.-T., Tien, M.-C., Chen, Y.-W., Tsai, W.-J., y Lee, S.-L. (2009). Physics-based ball tracking and 3D trajectory reconstruction with applications to shooting location estimation in basketball video. *Journal of Visual Communication and Image Representation 20* (3), 204-216.

Dwass, M. (1957). Modified randomization tests for nonparametric hypotheses. *Annals of Mathematical Statistics 28*, 181–187.

Doolittle, B. Y Pelton, K. (2009). *Pro Basketball Prospectus 2009-10*. Prospectus Entertainment Ventures LLC.

Esteller-Moré, A., y Eres-García, M. (2002). A note on consistent players' valuation. *Journal of Sports Economics, 3* (4), 354-360.

Fernández, J., Camerino, O., Anguera, M. T., y Jonsson, G. K. (2009). Identifying and analyzing the construction and effectiveness of offensive plays in basketball by using systematic observation. *Behavior Research Methods*, *41*, 719-730.

Fort, R., y Maxcy, J. (2003). Competitive balance in sports leagues: An introduction. *Journal of Sports Economics*, *4* (2), 154-160.

Fort, R., Hoon-Lee, Y., y Berri, D. J. (2008). Race, technical efficiency and retention. The case of NBA coaches. *International Journal of Sport Finance*, *3*, 84-97.

Gilovich, T., Vallone, R. y Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology 17*, 295-314.

Hickson, D. A., y Waller, L. A. (2003). Spatial analyses of basketball shot charts: An application to Michael Jordan's 2001–2002 NBA season, *Technical Report*, Department of Biostatistics, Emory University.

Hitt, J. L., Alferink, L. A., Critchfield, T. S., y Wagman, J. B. (2007). Choice behavior expressed in elite sport competition: Predicting shot selection and game outcomes in college basketball. In L. A. Chiang (Ed.), *Motivation of exercise and physical activity* (pp.79–91). Hauppauge, NY: Nova Science.

Hollinger, J. (2005). *Pro Basketball Forecast*. Washington, D.C.: Potomac, Inc.

Hoon-Lee, Y., y Berri, D. J. (2008). A re-examination of production functions and efficiency estimates for the National Basketball Association. *Scottish Journal of Political Economy*, *55* (1), 51-66.

Humphreys, B. R. (2000). Equal pay on the hardwood: the earnings gap between male and female NCAA. Division I Basketball coaches. *Journal of Sports Economics*, *1* (3) 299-307.

Humphreys, B. R. (2002). Alternative measures of competitive balance in sports leagues. *Journal of Sports Economics*, *3* (2), 133-148

Jäger, J. M., y Schöllhorn, W. I. (2007). Situation-oriented recognition of tactical patterns in volleyball. *Journal of Sports Sciences, 25* (12), 1345-1353.

Kleeman, R. (2009, 12 February). Pau Gasol Powers Runaway Lakers: How One Trade Changed the Title Race. Descargado desde http://bleacherreport.com/articles/123235-pau-gasol-powers-runaway-lakers-how-one-trade-changed-the-title-race

Kubatko, J.; Oliver, D., Pelton, K, y Rosenbaum, D. T. (2007). A starting point for analyzing basketball statistics. *Journal of Quantitative Analysis in Sports*, *3* (3), Article 1.

Kulldorff M. (1997) A spatial scan statistic. *Communications in Statistics: Theory and Methods*, 26, 1481-1496.

Lehmann,E.L. (1986) *Multivariate Linear Hypothesis. Testing statistical hypothesis*. 2nd edn. John Wiley y Sons, Inc, New York.

Lewis, M. M. (2003) *Moneyball: The art of winning an unfair game.* W.W. Norton y Company Inc.

Manning, (2009, 14 December). Popovich: "Pau Gasol changed the landscape of basketball in the NBA" Descargado desde http://thelakersnation.com/blog/2009/12/14/popovich-pau-gasol-changed-the-landscape-of-basketball-in-the-nba/

Martínez, J. A. (2010). Una revisión de los sistemas de valoración de jugadores de baloncesto (II). Competiciones oficiales y ligas de fantasía. *Revista Internacional de Derecho y Gestión del Deporte, 11*, 48-68.

Mavridis, G., Tsamourtzis, E., Karipidis, A., y Laios, A. (2009). The inside game in World Basketball. Comparison between European and NBA teams. *International Journal of Performance Analysis of Sport*, *9*, 157-164.

Michaelides, M. (2010). A new test of compensating differences: Evidence on the importance of unobserved heterogeneity. *Journal of Sport Economics, 11* (5), 475-495.

Mulrooney, T. (2010). Mapping the landscape of collegiate wrestling using spatial statistics. Timothy J. Mulrooney, 29 Nov 2007. Web. 18 Nov 2010. <http://myweb.wssu.edu/mulrooneyti/research.htm>.

Oliver, D. (2004). *Basketball on paper. Rules and tools for performance analysis.* Washington, D. C.: Brassey's, INC.

Piette, J., Sathyanarayan, A., y Kai, Z. (2010). Scoring and shooting abilities of NBA players. *Journal of Quantitative Analysis in Sports*, *6* (1), Article 1.

Price, J., y Wolfers, J. (2010). Racial discrimination among NBA referees. *Quarterly Journal of Economics, 125* (4), 1859-1887

Reich, B. J., Hodges, J. S. Carlin, B. P., y Reich, A. M. (2006). A spatial analysis of basketball shot chart data. *The American Statistician, 60* (1), 3-12.

Rimler, M. S., Song, S., y Yi, D. T. (2010). Estimating production efficiency in men's NCAA college basketball: A bayesian approach. *Journal of Sports Economics*, *11* (3), 287-315.

Romanowich, P., Bourret, J., y Vollmer, T. R. (2007). Further analysis of the matching law to describe two- and three-point shot selection by professional basketball players. *Journal of Applied Behavior Analysis*, *40*, 311–315.

Sánchez, J. M., Castellanos, P., y Dopico, J. A. (2007). The winning production function: Empirical evidence from Spanish basketball. *European Sport Management Quarterly*, *7* (3), 283-300.

Sanderson, C. (2010, 13 July). Why Derek Fisher Is So Important To the L.A. Lakers' Success. Descargado desde http://bleacherreport.com/articles/419819-why-derek-fisher-is-so-important-to-the-lakers-success

Skinner, B. (2010). The price of anarchy in basketball. *Journal of Quantitative Analysis in Sports*, *6* (1), Article 3.

Tauer, J. M., Guenther, C. L., y Rozek, C. (2009). Is there a home choke in decisive playoff basketball games?. *Journal of Applied Sport Psychlogy, 21*, 148-162.

Trininic, S., Dizdar, D., y Luksic, E. (2002). Differences between winning and defeated top quality basketball teams in final tournaments of European club championship. *Collegium Antropologicum, 26* (2), 521-531.

Vergin, R. C. (2000). Winning streaks in sports and the misperception of momentum. *Journal of Sport Behavior 23*, 181-197.

Vollmer, T. R., y Bourret, J. (2000). An application of the matching law to evaluate the allocation of two and three-point shots by college basketball players. *Journal of Applied Behavior Analysis*, *33*, 137–150.

Winston, W. L. (2009). *Mathletics.* New Yersey: Princeton University Press

Zimmer, T., y Kuethe, T. H. (2007). Testing for bias and manipulation in the National Basketball Association playoffs. *Journal of Quantitative Analysis in Sports*, *5* (3), Article 4.

**Referencias totales / Total references:** 58 (100%)
**Referencias propias de la revista / Journal's own references**: 0

## APPENDIX

In order to give an answer to this question we will assume that the basketball court is a $w_1 \times w_2$ regular lattice L. At each location s∈$\mathcal{L}$ we will denote by $n_s$ the total number of shots attempted at location s done by player (resp. team) A. Similarly we denote by $m_s$ the total number of shots attempted at location s done by player (resp. team) B. Denote by $N_s = n_s + m_s$ , $n_A = \sum_{s\in\mathcal{L}} n_s$ , $m_B = \sum_{s\in\mathcal{L}} m_s$ and $N = \sum_{s\in\mathcal{L}} N_s$ the total number of shots attempted at location s and total number of attempted shots taken by A, B and A∪ B players (resp. teams) respectively. Then one could easily compute the relative frequency at location s and total shots attempted of A and B by $p_s = \frac{n_s}{N}$ , $q_s = \frac{m_s}{N}$ , $p_A = \frac{n_A}{N}$ and $q_B = \frac{m_B}{N}$ respectively. Hence the total attempted shots relative frequency at location s is $r_s = p_s + q_s$.

Now under this setting we can define the *total attempted shots entropy*. This entropy is defined as the Shannon's entropy of the distribution of $r_s$ as follows:

$$h(A \cup B) = -\sum_{s\in\mathcal{L}} r_s \ln(r_s)$$

Total attempted shots entropy, h(A∪B), is the information contained in comparing the distribution of r among all locations in L.

Similarly we have the A, B and A versus B attempted shots entropy

$$h(A) = -\sum_{s\in\mathcal{L}} p_s \ln(p_s)$$

$$h(B) = -\sum_{s\in\mathcal{L}} q_s \ln(q_s)$$

and

$$h(A,B) = -p_A \ln(p_A) - q_B \ln(q_B)$$

respectively.

Once we have introduced the basic definitions and notations we will construct a statistical test to check whether the distribution of attempted shots by A is equal to the distribution of B. To this end we consider the following null hypothesis:

$H_0$: The distribution of attempted shots is the same for the LAL and for the opponents

that is,

$$H_0 : q_s = \frac{m_B}{n_A} p_s \text{ for all } s\in L.$$

Assume that the lattice $\mathcal{L}$ is of cardinality R. Notice that the variable number of shots attempted at location s is a Binomial random variable that can split in two binomial distributions $Y_S$=B(N,$p_S$) and $Z_S$=B(N,$q_S$) corresponding to A and B respectively.

Therefore the joint probability density function of the 2R variables is

$$p(Y_1=a_1,\ldots,Y_R=a_R,Z_1=a_{R+1},\ldots,Z_R=a_{2R})=\frac{(a_1+\cdots+a_{2R})!}{a_1!\cdots a_{2R}!}p_1^{a_1}\cdots p_R^{a_R}q_1^{a_{R+1}}\cdots q_R^{a_{2R}}$$

where $a_1+\cdots+a_{2R}=N$ and its likelihood function is:

$$L(p(Y_1=a_1,\ldots,Y_R=a_R,Z_1=a_{R+1},\ldots,Z_R=a_{2R}))=\frac{N!}{n_1!\cdots n_R!m_1!\cdots m_R!}p_1^{n_1}\cdots p_R^{n_R}q_1^{m_1}\cdots q_R^{m_R}$$

.

It is straightforward to see that the maximum likelihood estimators of $p_s$, $q_s$ and $r_s$ are $\hat{p}_s=\frac{n_s}{N}$, $\hat{q}_s=\frac{m_s}{N}$ and $\hat{r}_s=\frac{n_s+m_s}{N}$ respectively.

Then, under the null $H_0$ we have that $H_0: q_s=\frac{m_B}{n_A}p_s$ and thus, $r_s=p_s+q_s=p_s+\frac{m_B}{n_A}p_s=\frac{N}{n_A}p_s$. Therefore, under the null $H_0$, the likelihood ratio statistic is (see Lehmann (1986))

$$\lambda=\frac{\left(\frac{m_B}{n_A}\right)^{m_{Tot}}\left(\frac{n_A}{N}\right)^{N}r_1^{n_1+m_1}\cdots r_R^{n_R+m_R}}{\prod_{s\in \text{☹}}\left(\frac{n_s}{N}\right)^{n_s}\prod_{s\in \text{☹}}\left(\frac{m_s}{N}\right)^{m_s}}.$$

On the other hand V= -2 ln(λ) asymptotically follows a Chi-squared distribution with R-1 degrees of freedom (see for instance Lehmann (1986)). Hence, we obtain that the estimator $\hat{V}$ of V is

$$\hat{V}=-2N\left[\frac{m_B}{N}\ln\left(\frac{m_B}{n_A}\right)+\ln\left(\frac{n_A}{N}\right)+\sum_{s\in \text{☹}}\frac{n_s+m_s}{N}\ln\left(\frac{n_s+m_s}{N}\right)-\sum_{s\in \text{☹}}\frac{n_s}{N}\ln\left(\frac{n_s}{N}\right)-\sum_{s\in \text{☹}}\frac{m_s}{N}\ln\left(\frac{m_s}{N}\right)\right]=$$

$$=2N\left[h(A,B)+h(A\cup B)-h(A)-h(B)\right]$$

which is $\chi^2_{R-1}$ distributed.