UNIVERSITY OF LAS PALMAS DE GRAN CANARIA

DEPARTAMENT OF PHYSICAL EDUCACTION

Doctoral Program:

Motor Praxiology, Physical Education and Sport Training

DOCTORAL THESIS

BASKETBALL FROM THE PERSPECTIVE OF NON-LINEAR COMPLEX SYSTEMS

YVES DE SAÁ GUERRA

LAS PALMAS DE GRAN CANARIA, 2013

BASKETBALL

FROM THE PERSPECTIVE OF NON-LINEAR COMPLEX SYSTEMS

—

DOCTORAL THESIS

By YVES DE SAÁ GUERRA, B.S.

SCIENCES OF PHYSICAL ACTIVITY AND SPORT

UNIVERSITY OF LAS PALMAS DE GRAN CANARIA,

GRAN CANARIA, SPAIN

2013

Anexo II

UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Departamento de Educación Física. Facultad de Educación Física

Programa de doctorado: Praxiología motriz, educación física y entrenamiento deportivo.

Título de la Tesis

BASKETBALL FROM THE PERSPECTIVE OF NON-LINEAR COMPLEX SYSTEMS

Tesis Doctoral presentada por D. Yves de Saá Guerra

Dirigida por el Dr. D. Manuel Navarro Valdivielso

Codirigida por el Dr. D. Juan Manuel García Manso

Codirigida por el Dr. D. Juan Manuel Martín González

El Director, (firma)	**El Codirector ,** (firma)	**El Codirector,** (firma)	**El Doctorando,** (firma)

Las Palmas de Gran Canaria, a _____ de_________________ de 2013

The more I know,

the less I can affirm,

categorically.

Cuanto más sé,

menos puedo afirmar las cosas

categóricamente.

Yves

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the help of so many people in so many ways. It carried out thanks to the commitment and dedication of a research group led by Dr. Manuel Navarro Valdivielso, Dr. Juan Manuel García Manso and Dr. Juan Manuel Martín.

It is an honor be part of this group and I would like to express my sincere gratitude to my thesis director Dr. Manuel Navarro Valdivielso for his unconditional support, his advices and his academic orientation, as well as for his help in the develop of this thesis. Also I would like to express my deepest admiration to my co-directors Dr. Juan Manuel García Manso, because he taught me everything I know about training and advised me about sport in an exquisite way. He has conveyed me all his strength and passion for the sport. I will never be able to remove his seed; and to Dr. Juan Manuel Martín, because he developed the necessary programs in Matlab that allow us to deepen in sports reality and advise with mathematics. I have been fascinated by all of our conversations about the universe and the secrets of life. It is a real pleasure to listen to and learn from him. Without his help this thesis would not been possible. I feel privileged for sharing part of my life with these three great people.

I also would like to acknowledge the collaboration and support of Professor Adrian Bejan, from Duke University, USA. by his passionate discussions and by his outstanding relevance in the scientific community. I was a real pleasure learn from him during my stay at Duke University.

I cannot conclude this chapter without to extend my sincere thanks to my fellows of the Laboratory for Analysis and Planning of Sports Training, of Department of Physical Education of this university, especially to Dr. Samuel Sarmiento Montesdeoca for his support and friendship.

I must also acknowledge the support of the University of Las Palmas de Gran Canaria, for supporting my research by awarding a grant for the development of this thesis.

Finally, and by no means least important, I would like to give thanks to my entire family for their unconditional support every time I needed it. They are my role models. They always have been. And always will be.

Yves

Index

Abstract

We consider basketball as a complex systemic unit. There is a reciprocal relationship between the basketball team and its environment. Therefore it is important to figure out the behavior of the participating agents, as well as their relationships. First at all we carried out an approach to the degree of competitiveness of a sport league as complex phenomenon. In the first study, we used the results of previous seasons as a way to investigate the victory probabilities of each team. We developed a model based on Shannon entropy using two extreme competitive structures, a hierarchical structure and a random structure and applied this model to investigate competitiveness of the NBA (USA) and the ACB (Spain). Both leagues entropy levels are high (NBA mean 0.983; ACB mean 0.980) indicating high competitiveness although entropy of the ACB (from 0.986 to 0.972) demonstrated more seasonal variability than the NBA (from 0.985 to 0.990). This methodology has been useful for investigating sports competitiveness.

The second study deal with score in basketball. Scoring in a basketball game is a highly dynamic, non-linear process. Several mechanisms concur to make the scoring process in the NBA games exciting and hardly predictable. We analyze all the games in five NBA regular seasons (2005-06, 2006-07, 2007-08, 2008-09, 2009-10), for a total of 6150 games. Scoring does not behave uniformly; therefore, we as well analyze the distributions of the differences in points in the basketball games. To further analyze the behavior of the tail of the distribution, we also carry out a scaling analysis in order to verify that distribution. This analysis reveals different areas of behavior related to the score, with specific instances of time that could be considered tipping points of the game. The presence of these critical points suggests that there are phase transitions where the scoring dynamic of the games varies significantly.

As third study we propose the use of the network theory in order to clarify the features of the basketball teams as a player network. Sport as network is a new field of application in sport research. In fact, there are only few papers dealing with sport as network. We carried out an in investigation about player network in order to clarify the behavior of players of the same team when they fight against the other team. We study the game of NBA Finals Chicago Bulls vs. Miami Heat. Regarding player interactions, we measured for each team the number or passes, screens and space creations per play. This analysis point out that teams modify their interactions and are able to behave as small-world network or as scale-free network regarding the game situation (time and score).

Introduction

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

Introduction

In our effort to try to understand the sport reality, we are forced to question everything that we consider true or static. This leads us inexorably, to try to isolate a phenomenon in order to study better. But far from our purpose, when we move towards this point, we realize that we fall into the contradiction of hoping of be able to understand a phenomenon by isolating it from its environment. Following the asseveration: " *I am I and my circumstance"* (Ortega y Gasset, 1933), it is not possible to understand the sport phenomenon by isolating the constituent elements of their relationships with their own universe. There is a duality element–environment. From this relationship might emerge new behaviors, which is known from the perspective of complexity, as *an emergency phenomenon or an emergent behavior.*

I.e. in order to reason this phenomenon, it would be good to move away from the deterministic and reductionist classical model, and move on to study the systems from a more global conception (holistic), allowing us to identify and describe the processes of new forms of organization, which is also useful in sport. Organizing the sport training from a systematic conception and conceive the athlete, or the team in our case, as a system that functions as a whole and that is affected by the surrounding environment (Gambetta, 1989; Martín Acero & Lago Peñas, 2005; García Manso & Martín González, 2008). Chaos theory has provided a new light to observe all these systems, seemingly incomprehensible or random, because it is usual that natural systems are chaotic. Chaos hides an internal order that is possible to find (Prigogine & Holte, 1993).

In complex systems, the processes occurring simultaneously in different scales or levels are important, the intricate and complex behavior of the system as a whole, depends on the units, although not directly, because in complex systems the structures have strong relationships, often in a non-linear manner (Vicsek, 2002; Goodwin, 2002; Amaral & Ottino, 2004; Solé, 2009).

It would be useless to consider the conscious thought as a mere sum of neurons, or reduce the behavior of a team to the sum of the individual abilities of his players individually. This definition also distinguishes the complex from the simple and complicated, as can be any other mechanism such as that determines the operation of a car, an airplane or a computer.

Complexity is a measure of the number of possibilities: the ways in which we act or react to the environment. It is an important point for the study of sport from the point of view of complex systems. However, the true complexity of behavior, as seen in basketball, occurs in the interaction between its elements and response possibilities of each structure.

Sport performance is the result of the combination of many variables that sometimes we know and dominate it through different analytical methods, and we try to understand even better to improve it. Seldom effectiveness in sport shows a linear behavior. It would be very easy to understand and get better. Actually, there are many actions we can consider proper, even whether these actions would repeat, need not be consecutive. The behavior of players, ball, coaches and many more aspects, may condition the outcome. So it works as any complex organization, hence, must be understand as a complex system.

Establishing team Sports as complex systems, our intention is to move away from classic and deterministic models, in order to pass on a new perspective for unifying criteria and analysis and thus to address and resolve issues raised above, and increase performance in different sports collective. This model is based on how its components are related in a precise and determined, and how they react to other complex systems. In fact, all we seek is to recognize and identify patterns of collective behavior and relationships among its components, which make it, resemble self-organizing complex systems non-linear in critical (System Organized Critically). According to Goodwin, the ideal is to investigate the conditions that promote self-organization(Goodwin, 2002), in order to obtain the sporting excellence. Complex systems are the result of an evolutionary process. Darwin's ideas and the study of evolution have focused on the competition as a driving force of evolutionary change. Players for example, when they cooperate, compete better as a team (Bar-Yam, 2003).

Team performance can be postulated as win as many games as possible. It results from the synchronous interaction of certain states of optimization of systems that make up, which also have a reciprocal relationship with the emerging and critical environment: the competition.
The aim of the study is to observe the behavior of the structure (macrostructure) of basketball as a system. We want to find out how basketball elements are interconnected and how they affect each other, analyzing the laws that govern them. Therefore we have established three levels or items to carry out the investigation: league, games and basketball team as network.

Basketball Background

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

3. Basketball Background

3.1 Basketball History

Creation of basketball

Dr. James Naismith was a Canadian physical education professor at the International Young Men's Christian Association Training School (YMCA) in Springfield, Massachusetts, USA. In December of 1891, James Naismith was asked by his director to devise an indoor game for the school´s 40 students, to help keep them physically active between football season and the springtime activities of baseball and track (Muruzábal del Solar, 2012).

Dr. Naismith combined elements of outdoor games like football and lacrosse with the concept of a game he played in childhood, Duck on a Rock. To win Duck on Rock, players threw stones to hit a target placed on top of a large boulder (Naismith, 1941).

He set up peach baskets attached to both ends of a gymnasium balcony onto a 3.05 meters elevated track, and used a ball in order to score on it. The peach baskets retained the ball at bottom and it had to be taken out manually after each point scored until the bottom of the basket was removed. The peach baskets were used until 1906 when they were finally replaced by metal hoops with backboards.

The backboards appeared as protection, to prevent the fan located on the railing of the gallery, where hung the baskets, could hinder the entry of the ball in the basket, which later went on to become a metal ring and a network without holes, for lead in today's networks (Tous Fajardo, 1999).

The early players did not use dribbling of the ball, except for the “bounce pass” to teammates. Passing the ball was the primary means of ball movement. Dribbling was eventually introduced but limited by the asymmetric shape of early balls. Dribbling only became a major part of the game around the 1950s, as manufacturing improved the ball shape.

The first official game was played in the YMCA gymnasium in Albany, New York on January 20, 1892 with nine players. The game ended at 1-0. The shot was made from 7.6 meters, on a court just half the size of a present day National Basketball Association (NBA) court. By 1897-1898 teams of five became standard (Naismith, 1941).

A ball and an elevated goal, those are the simple ingredients of the sport that now have players and rabid fans in nearly every part of the world. According to Alexander Wolff, in his book *100 Years of Hoops* (Wolff, 1991), Naismith drew up the rules for the new game in "about an hour". Nowadays basketball is one of the world's most popular and widely viewed sports (Griffiths, 2010).

The First 13 Rules of Basketball

Naismith and Wheeler wrote the first 13 rules of the game (International Basketball Federation; FIBA, 2012). They were published in the school newspaper, The Triangle, for first time in 1892:

1. The ball may be thrown in any direction with one or both hands.

2. The ball may be batted in any direction with one or both hands (never with the fist).

3. A player cannot run with the ball. The player must throw it from the spot on which he catches it, allowance to be made for a man who catches the ball when running at a good speed if he tries to stop.

4. The ball must be held in or between the hands; the arms or body must not be used for holding it.

5. No shouldering, holding, pushing, tripping, or striking in any way the person of an opponent shall be allowed; the first infringement of this rule by any player shall count as a foul, the second shall disqualify him until the next goal is made, or, if there was evident intent to injure the person, for the whole of the game, no substitute allowed.

6. A foul is striking at the ball with the fist, violation of rules 3, 4, and such as described in rule 5.

7. If either side makes three consecutive fouls, it shall count a goal for the opponents (consecutive means without the opponents in the mean time making a foul).

8. A goal shall be made when the ball is thrown or batted from the grounds into the basket and stays there, providing those defending the goal do not touch or disturb the goal. If the ball rests on the edges, and the opponent moves the basket, it shall count as a goal.

9. When the ball goes out of bounds, it shall be thrown into the field of play by the person first touching it. In case of a dispute, the umpire shall throw it straight into the field. The thrower-in is allowed five seconds; if he holds it longer, it shall go to the opponent. If any side persists in delaying the game, the umpire shall call a foul on that side.

10. The umpire shall be judge of the men and shall note the fouls and notify the referee when three consecutive fouls have been made. He shall have power to disqualify men according to rule 5.

11. The referee shall be judge of the ball and shall decide when the ball is in play, in bounds, to which side it belongs, and shall keep the time. He shall decide when a goal has been made, and keep account of the goals with any other duties that are usually performed by a referee.

12. The time shall be two 15-minute halves, with five minutes rest between.

13. The side making the most goals in that time shall be declared the winner. In case of a draw, the game may, by agreement of the captains, be continued until another goal is made.

3.2 Basketball General Description

Some authors such as Knapp (1981), Hernández (1994) and Ruiz (1999) analyze basketball regarding to formal structure and functionality of the game itself. It also can be classified as a team sport or collective sport, of cooperation-opposition, with functional structure which develops in a common space for both teams and simultaneous intervention on ball (Hernández Moreno, 1994)

Basketball is defined as a team sport played by two teams of five players each on the court, and can be substituted by bench players (each league allow a different number of bench players), previously selected by the coach. The aim of each team is to score in the opponents' basket and to prevent the other team from scoring while following a set of rules. The team that has scored the greater number of points at the end of playing time shall be the winner (Rule 1. Art. 1. FIBA, 2012).

In the leagues we studied (National Basketball Association, NBA; Asociación de Clubs de Baloncesto, ACB; and National College Athletic Association, NCAA); the game is controlled by officials and score officials throughout the Official Basketball Rules (NBA, 2012; NCAA, 2012a; FIBA, 2012b) (ACB uses FIBA Rules). Nevertheless there are some differences between the NBA, FIBA and NCAA rules, regarding court dimensions, time outs, fouls, officials, etc. (see table 1).

We carried out this comparison because these three leagues are the aim of study. In general, rules for the male gender and for the female gender are different. Only the FIBA rules establish the same rules for both genders.

Usually, teams play on a marked rectangular court with a basket at each width end. The dimensions of the court are described by the official rules (NBA, 2012; FIBA, 2012b; NCAA, 2012b) (See Figure 1, Figure 2, Figure 3 and Figure 4).

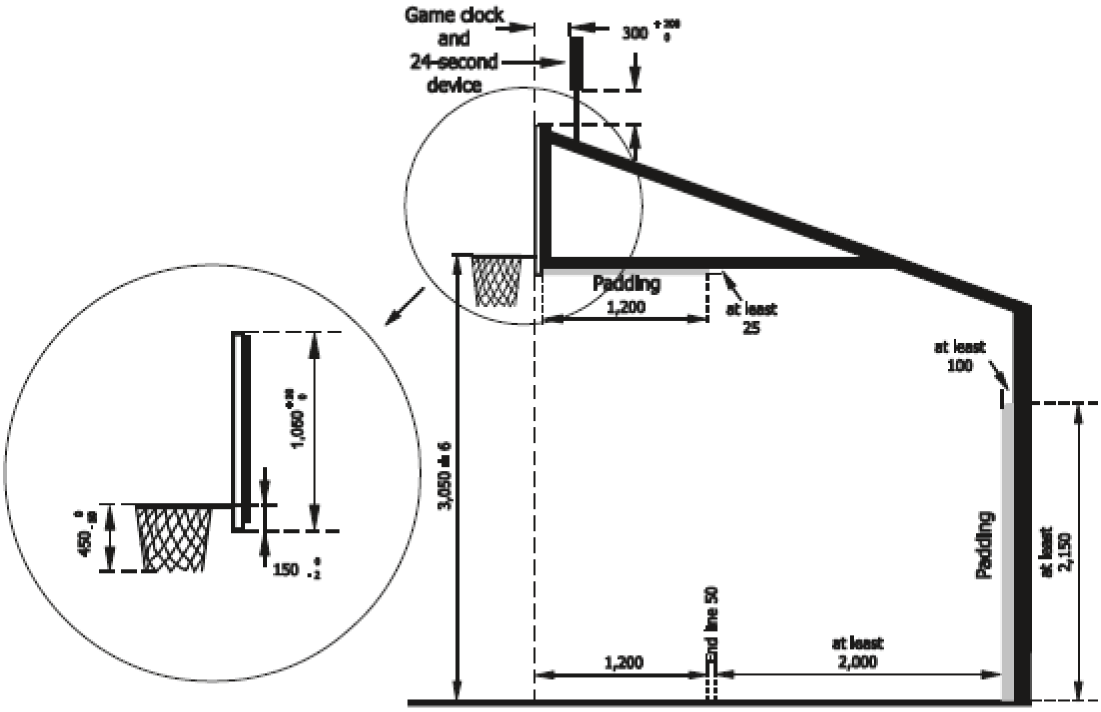


Figure 1. Hoop dimensions. Source: **FIBA.**

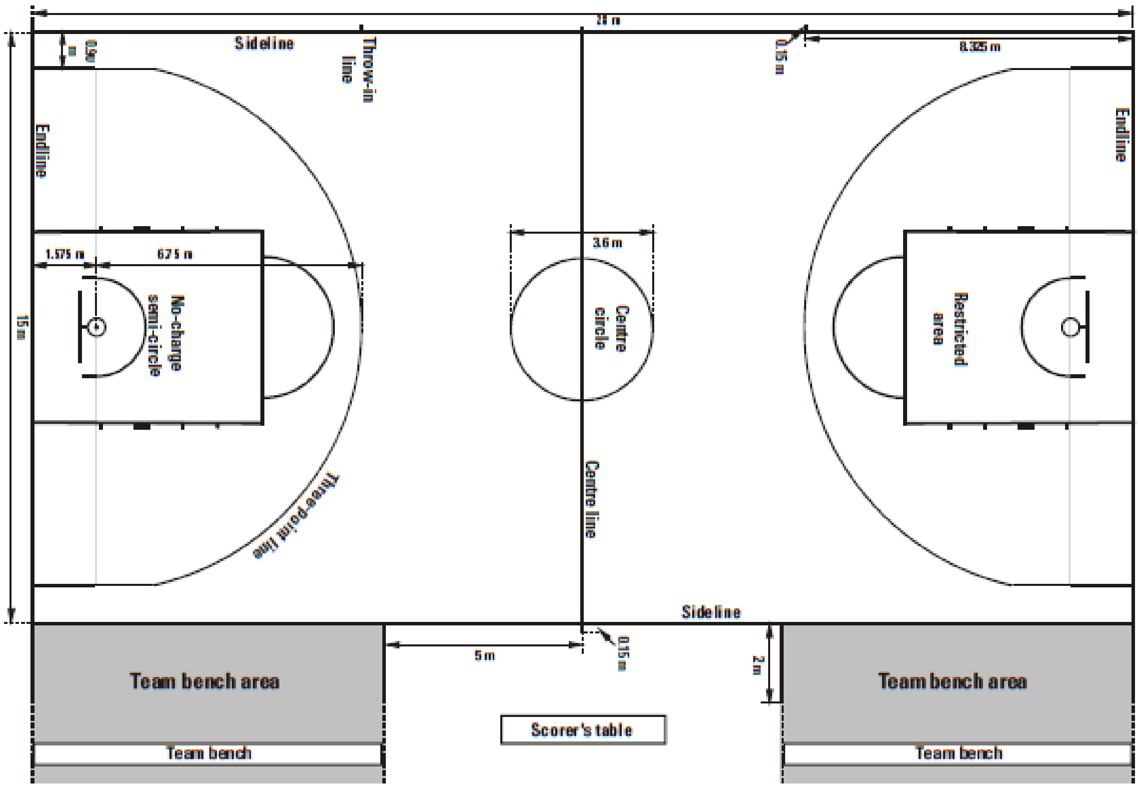


Figure 2: FIBA Basketball Court. Source: **FIBA.**

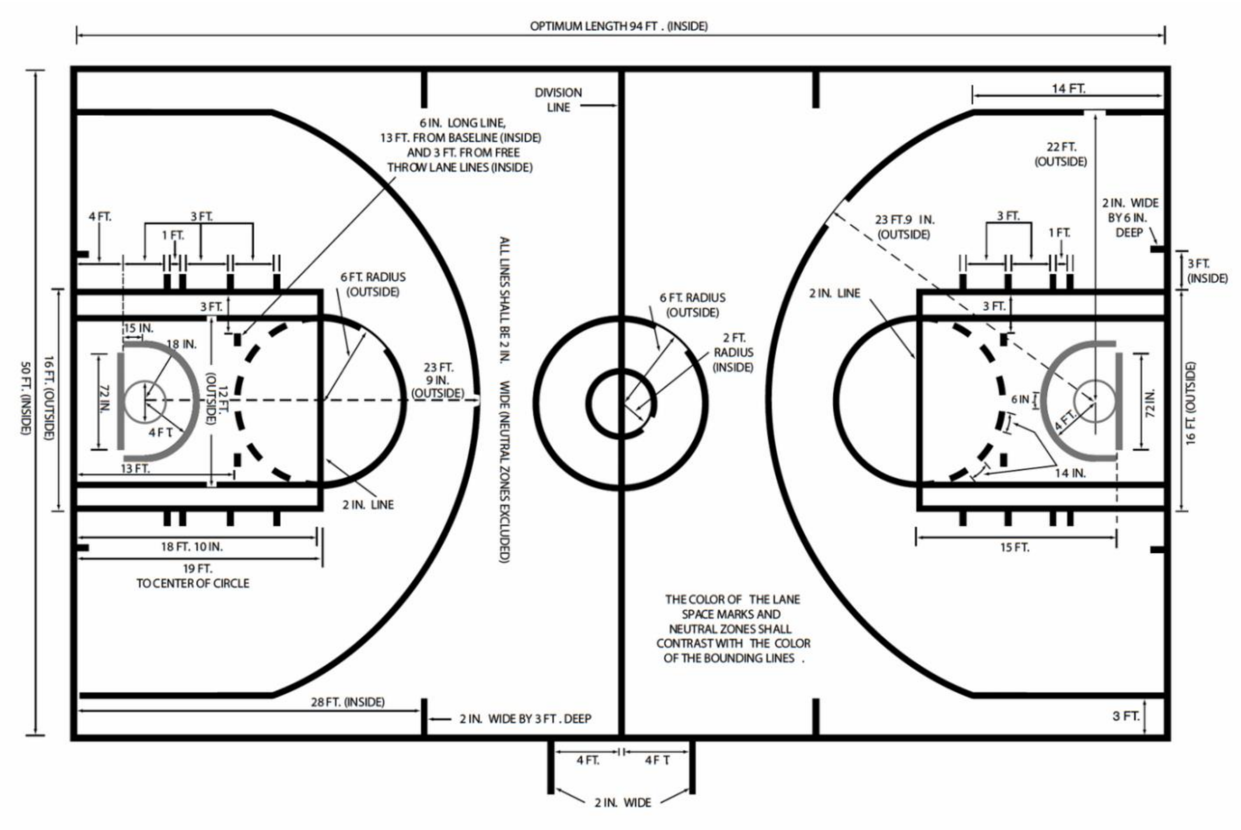


Figure 3. NBA Basketball Court Dimensions. Source: **NBA.**

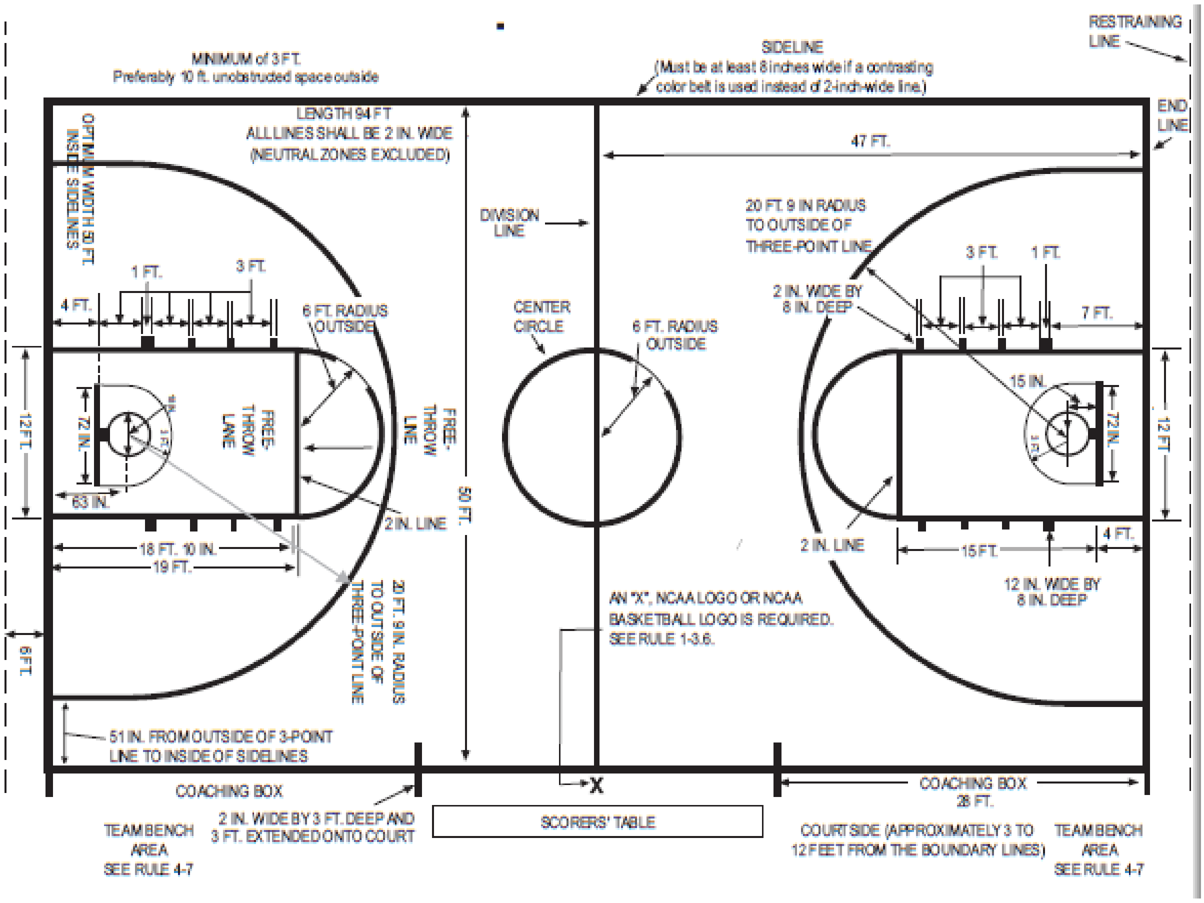


Figure 4. NCAA Basketball Court Dimensions. Source: **NCAA.**

In general, rules (NBA, 2012; FIBA, 2012b; NCAA, 2012b) establish that a team can score a field goal by shooting the ball through the basket during regular play. A field goal scores two points for the shooting team if a player is touching or closer to the basket than the three-point line, or three points if the player is behind the three-point line. The ball can be advanced on the court by bouncing it while walking or running (dribbling) or throwing (passing) it to a teammate. It is a violation to run with the ball without dribbling it (traveling), or to carry out a double dribble (to hold the ball with both hands then resume dribbling). A personal foul is a disruptive physical contact, and it is penalized. A free throw is usually awarded to an offensive player if he is fouled while shooting the ball. A technical foul may also be issued when certain infractions occur, most commonly for unsportsmanlike conduct on the part of a player or coach. A technical foul gives the opposing team a free throw.

Despite these general rules, there some notable differences among different basketball leagues as we exposed at Table 1:

Table 1. Major differences among FIBA Rules, NBA Rules and NCAA Rules.			
	FIBA	**NBA**	**NCAA**
Three Point Line (measured from the middle of the basket)	6.75 m (22' 14") arc.	7.24 m (23' 9") arc, which intersects with lines parallel to the sideline that are 6.7 m (22') from the basket at their closest point.	6.25 m (20' 9") arc.
Restricted Area (a.k.a. "The Key" or "The Lane")	4.88 m (16´) wide rectangle	4.88 m (16´) wide rectangle	3.6 m (12') wide rectangle
Playing Time	4 - 10 minute quarters; OT periods are 5 min each	4 - 12 minute quarters; OT periods are 5 min each	2 - 20 minute halves; OT periods are 5 min each
Game Clock Operation — Last Minutes of Play/Field Goal The clock stops after a successful field goal as follows:	In the last two minutes of the fourth period and any OT period	In last minute of quarters 1, 2, 3, last 2 minutes of quarter 4 and any OT	In the last minute of the second half and any OT period
Shot Clock — Time Allowed to Shoot	24 seconds	24 seconds	35 seconds
Shot clock – Operation	When play resumes with less than full amount on shot clock (e.g., defense taps ball out-of-bounds), shot clock does not start until team establishes control inbounds. Shot clock is reset after most fouls (personal or technical). **Exception:** Shot clock is not reset on a double foul or an alternating possession situation when the same team retains possession.	When play resumes with less than full amount on shot clock, shot clock starts with the first touch in-bounds. The shot clock is reset to 24 seconds on most personal fouls and defensive violations in backcourt (e.g., kicking or punching ball). **Exceptions:** The shot clock remains the same as when play was interrupted or is reset to 14 seconds (whichever is greater) when (1) a personal foul occurs and the throw-in will be in the frontcourt, (2) a jump ball occurs overtime. One 20-second time-out per half; unused	When play resumes with less than full amount on shot clock, shot clock starts with the first touch in-bounds. The shot clock is reset after most fouls (personal or technical).

		20-second time-out in 2nd half may be carried into extra period.	
Time Outs — Number and Duration	2 time outs in first half, 3 in second half, one per overtime period. All time-outs must last 60 seconds. Time-outs do not accumulate.	6 "regular" time outs per regulation time (with some restrictions — some mandatory timeouts for TV are built into the 6), 2 regular timeouts per overtime period. Regular time outs are 60 seconds in duration, except the first two timeouts in each period and the extra mandatory timeout in Quarters 2 and 4, which are 100 seconds. Time outs do not accumulate into overtime One 20 second timeout per half and each overtime period. Maximum 3 regular timeouts in the fourth period. If a team has 2 or 3 regular timeouts remaining when the fourth period or overtime period reaches the 2:00 mark, those will change to one regular timeout and one 20-second timeout. (Thus, a team may never have more than 1 regular and two 20-second timeouts in the last two minutes of a game.)	**Electronic Media Game:** Four 30-second time-outs and one 60-second time-out per game. Maximum of three 30-second time-outs and one 60-second time-out may be carried into 2nd half. One additional 30-second time-out is added per extra period (any time-outs remaining from 2nd half may be carried into extra period). First 30-second time-out of 2nd half is extended to the length of a media time-out. If coach requests 2 consecutive 30-second time-outs, players may sit, so long as the request is made when the first timeout is granted. Normally, players must remain standing and on the floor during a 30-second time-out. **Non-Electronic Media Game:** Four full time-outs per game (75 seconds, with warning after 60 seconds); add 1 time-out per extra period. Two 30-second time-outs per game (used anytime).All time-outs are cumulative.

Player Fouls	Foul out on 5 (personal + technical)	Foul out on 6 personal or 2 technical fouls	Foul out on 5 (personal fouls + non-administrative technical fouls)
Technical Foul — Penalty	2 free throws and possession of the ball at centre; no possession at centre if the foul occurs before the first half (game would still start with a jump ball).	1 free throw per technical foul; play resumes at the point of interruption; foul is charged to individual in question (and automatic fine assessed)	2 free throws, play resumes at point of interruption.

As we can see, the tendency these last years is to reduce differences between FIBA basketball and NBA basketball regarding court dimensions: three point line and restricted area. But time issues such gam time, time outs etc. remains as always; and even player fouls (five for FIBA and six for NBA). Probably, along with time, both regulations become more similar, in matter of quarter durations, fouls, times out, etc. as happened in the past with shot clock (in the ACB past from 30 seconds to 24 seconds). College basketball remains with almost the same differences than previous years (Paulo Ferreira, Ibáñez Godoy, & Sampaio, 2009; Anderson & Pierce, 2009; Wilner & Rappoport, 2012).

3.3. Organizational structure of basketball

In this section we describe those competition formats more commons and the elements which take part in the basketball system:

3.3.1. General structures

There are several common structures which are used in the most part of the countries in order to organize the basketball and its interactions (games, tournaments, federations, players, referees, etc.).

Sport Team. It is the sum of the players, coaching staff and the administrative section. It is the unit which competes in a league. Some teams are linked to a sport club or an institution such as university or a college (Solé Forto, 2002; Mestre, Brotóns, & Álvaro, 2002; Martín Acero & Lago Peñas, 2005; García Manso & Martín González, 2008).

Sport club. They are private associations, composed by natural or legal persons, which are intended to promote one or more types of sports, the practice of these by their members and participation in activities and sports competitions (Ley 10/1990, de 15 de octubre, del Deporte, Art. 13. Ministerio de Educación, Cultura y Deporte).

Club´s teams, often with professional players in their teams, regularly compete against other clubs in professional leagues. Sport clubs may have several brunches of different sports such as football basketball, cricket, volleyball, handball, rink hockey, water polo, rugby, track and field athletics, boxing, baseball, cycling, tennis, rowing, gymnastics, etc. (Mestre et al., 2002; Blanco, 2006).

Teams and athletes´ sport club share also the same resources, administrative section, supporters, facilities, and also the use to have farm teams, which in turn compete in others professional farm leagues.

League. A league is a sport competition system in which each contestant play a game against all other contestants in a constant number of opportunities, generally two times (Quirk & Fort, 1997; Schmidt & Berri, 2001; Humphreys, 2002; Fort & Maxcy, 2003; de Saá Guerra et al., 2012). It also called round-robin tournament or all-play-all tournament. In turn can be played in one single group (depend of the numbers of teams) or can be divided in several groups, conferences and/or divisions, according to a prearranged schedule.

Play-off/Cup. The playoffs or finals (we can also include cup tournaments) in a sport are a game or series of games played after the regular season by the top teams classified, in order to determine the league champion. Teams use to play in a bracket (ACB, 2012; NBA, 2012; NCAA, 2012a).

A bracket is a tree diagram that represents the series of games played during a tournament, named as such because it appears to be a large number of interconnected (punctuational) brackets.

There are several formats for the play-off brackets, meaning the number of games per round (1-1-1-1, 3-3-3, 7-7-7-7, etc.). In college basketball is very famous the fact of filling in brackets, especially in NCAA basketball, is referred to as bracketology.

Federation. Federations are private non-profit organizations composed by administrative section, sports clubs, athletes, coaches, judges and referees and professional leagues, in order to promote, practice or contribute to the development of sport (Ley 10/1990, de 15 de octubre, del Deporte, Art. 30. Ministerio de Educación, Cultura y Deporte).

The functions of the basketball federations are the government, administration, management, organization and regulation of the sport of basketball throughout the territory covered, whether international or domestic, regarding competitions and championships organized.

As well as drawing up of the corresponding licenses that are required to participate as player, coach or referee, in competitions and championships organized.

3.3.2. International structures

In basketball there are several international structures which try to regulate and promote worldwide basketball through the national teams.

FIBA (International Basketball Federation)

The International Basketball Federation (FIBA) is the organization that is dedicated to regulate the rules of basketball worldwide, as well as holding regular competitions and events in the disciplines of basketball (men and women).

The association was founded in Geneva in 1932, two years after the sport was officially recognized by the IOC (International Olympic Committee). The name FIBA came from its French name Fédération Internationale de Basketball, is an association of national organizations which governs international competition in basketball. Originally known as the Fédération Internationale de Basketball Amateur (hence FIBA), the word "Amateur" was dropped in 1986 after the distinction between Amateurs and Professionals. The "BA" now represents the first two letters of basketball. The main aim of the FIBA was to coordinate tournaments and teams. Argentina, Czechoslovakia, Greece, Italy, Latvia, Portugal, Romania and Switzerland were the founder members (FIBA, 2012).

The FIBA Central Board is currently composed of 23 members (22 have the right to vote) and meets twice yearly. The FIBA Central Board has, among other competences, the power to establish the FIBA Internal Regulations. It also assigns the organization of all FIBA Basketball World Cup. FIBA counts with 213 member federations.

FIBA has organized a FIBA World Championship for men since 1950 and a World Championship for Women since 1953. Both events are now held every four years, alternating with the Olympics.

IOC (International Olympic Committee)

The IOC coordinates the activities of the Olympic Movement. It is also responsible for supervising and manages everything about the Olympics. Owns all the rights associated with the Olympic symbols, flag, anthem, lemma, oath and games. It controls the rights to broadcast the games, advertising and other activities according to the Olympic Charter.

It is also the international body responsible for organizing and selecting the cities that will host the Olympic Games every 4 years (IOC, 2012). In detail the role of the IOC, according to the Olympic Charter (September 2004), is:

- To encourage and support the promotion of ethics in sport as well as education of youth through sport and to dedicate its efforts to ensuring that, in sport, the spirit of fair play prevails and violence is banned.
- To encourage and support the organization, development and coordination of sport and sports competitions.
- To ensure the regular celebration of the Olympic Games.
- To cooperate with the competent public or private organizations and authorities in the endeavor to place sport at the service of humanity and thereby to promote peace.
- To take action in order to strengthen the unity and to protect the independence of the Olympic Movement.
- To act against any form of discrimination affecting the Olympic Movement.
- To encourage and support the promotion of women in sport at all levels and in all structures with a view to implementing the principle of equality of men and women.
- To lead the fight against doping in sport.
- To encourage and support measures protecting the health of athletes.
- To oppose any political or commercial abuse of sport and athletes.
- To encourage and support the efforts of sports organizations and public authorities to provide for the social and professional future of athletes.
- To encourage and support the development of sport for all.
- To encourage and support a responsible concern for environmental issues, to promote sustainable development in sport and to require that the Olympic Games are held accordingly.

- To promote a positive legacy from the Olympic Games to the host cities and host countries.
- To encourage and support initiatives blending sport with culture and education.
- To encourage and support the activities of the International Olympic Academy (IOA) and other institutions which dedicate themselves to Olympic education.

Basketball appeared for first time in Olympic Games in San Louis in 1904 as an exhibition game. Basketball was included in Olympic Games in Berlin 1936 as Olympic Sport. Women´s basketball is present in Olympic Games in Montreal 1976.

3.3.3. National Structures

In every country, organizational structures of basketball are designed in different ways. According our investigation we studied the basketball in USA and Spain, hence we describe the organizational structures of basketball in USA and Spain as a follows.

3.3.3.1. Organizational structure of basketball in USA:

Basketball Association of America

The Basketball Association of America (BAA) was a professional basketball league in North America, founded in 1946. The league merged with several leagues such as the National Basketball League (NBL) in 1949, forming the National Basketball Association (NBA). Eleven cities are fortunate to welcome a team to represent them, are essentially cities located on the coast: Nueva York, Chicago, Boston, Providence, Toronto, Cleveland, San Luis, Washington, Detroit, Pittsburgh y Philadelphia. The first game is played in the city of New York and confronts New York Knicks vs. Toronto Huskies. African American players do not start to play until 1950.

There were several attempts to create other professional leagues to overthrow the NBA, highlighting the ABA League.

National Basketball League

The National Basketball League was founded in 1898 in the USA and was the first professional league in the world. Six teams took part in it and the first champions were the Trenton Nationals, followed by the New York Wanderers, the Bristol Pile Drivers and the Camden Electrics. The National League lasted five seasons (1904), but new leagues quickly were formed throughout New England and the Mid-Atlantic States, prominent among them were the Philadelphia Basketball League, Eastern Basket Ball League, Hudson River League, New York State League and the Interstate Basket Ball League.

ABA (American Basketball Association)

The American Basketball Association (ABA) was founded as an alternative to the NBA in 1967. Teams were created in different cities than NBA teams. The ABA was characterized by the color of the ball (red, white and blue) and the manner of play. The ABA also introduced several rules that differed from the NBA. Among them was the three-point line. The NBA adopted the three point line in 1979-80. The ABA competed with the National Basketball Association (the NBA) for players, fans, and media attention. In June 1976, four of the strongest ABA teams (the New York Nets, Denver Nuggets, Indiana Pacers, and San Antonio Spurs) joined the NBA (Silverman, 2012).

NCAA (National Collegiate Athletic Association)

The National Collegiate Athletic Association (NCAA) is an association of several institutions, conferences, organizations and individuals that organizes the athletic programs of many colleges and universities in the United States. It is headquartered in Indianapolis, Indiana. The NCAA was founded in 1906 to protect young people from the dangerous and exploitive athletics practices of the time.

In that period, the football was used in order to formation and gang tackling, but there were numerous injuries and deaths and prompted many college and universities to discontinue the sport. The most part of the fans and people related with football thought that college football should be reformed or abolished.

President Theodore Roosevelt summoned college athletics leaders to two White House conferences to encourage reforms. In December 1905, in New York City, 62 colleges and universities became charter members of the Intercollegiate Athletic Association of the United States (IAAUS). It was officially was constituted March 31, 1906, but in 1910 was renamed as National Collegiate Athletic Association (NCAA). Gradually, more rules committees were formed and more championships were created, including a basketball championship in 1939 (NCAA, 2012c).

One of the keys to success in college basketball was that the BAA had money and good game courts, but lacked of talent and experience players. This lack of players forces the owner to create a new policy to end the problem. This new policy is to recruit the best college players to shape and a championship that combines the speed, talent and experience.

The basketball college championship in USA is divided in three divisions (Division I, Division II and Division III). In turn, every division is divided in conferences of several teams each. Currently, in the Division I are involved a total of 344 teams (the number varies within the season analyzed), divided in 31 conferences through all USA (the number of teams per conference is not homogenous). The competition format of the NCAA described in this thesis makes reference only to the Division I of the men´s basketball.

There is a regular phase (regular league), and a playoff. In the regular phase teams play against the teams of the same conference. In addition, they play extra games (tournaments) in order to get more points for the playoff classification.

Most of these tournaments are the same every season. Some are very important in college basketball community. Some of the most popular are: 2K Sports Classic, Coaches vs. Cancer, Puerto Rico Tip-Off, Paradise Jam, CBE Classic, Maui Invitational, NIT Season Tip-Off, Old Spice Classic, 76 Classic, Legends Classic, ACC/Big Ten Challenge, Big12/Pac10 Hardwood Series or the Jimmy V Basketball Classic.

After the regular phase, the best teams classified play a playoff (only one game per round) for the national championship. There are two ways of qualifying for the play-off tournament. One is directly and the other is by invitation granted by the NCAA.

The direct classification is obtained by a ranking made with the RPI index. The RPI index is a mathematical equation that takes into account the games won, games lost, a series of numerical constants and strength of the schedule. Once obtained this ranking the 31 best will qualify directly for the tournament.

The playoff tournament include **68 universities (31 champions + 37 invited) and is held in March (also called *March Madness*).** The 68 teams are divided into four regions (**South, East, West and Midwest)** and organized into a single elimination bracket. Each team is ranked within its region.

From the 68 teams, 60 of them go directly to the second round remaining the 8 lower-seeded teams, which have received fewer votes from the NCAA; dispute the four remaining places in four small games of the first round (called First Four).

After an initial four games between 8 lower-ranked teams, the tournament takes place over the course of three weekends, at pre-selected neutral sites around the United States. Lower-ranked teams are placed in the bracket against higher ranked teams. Each weekend cuts three-fourths of the teams, from a Round of 64, to a round of 16 also called *Sweet Sixteen*, to a Final with four teams, called *Final Four*. The Final four is usually played on the first weekend in April.

NBA (National Basketball Association)

The NBA is the men's professional basketball league in North America (United States and Canada). The league was founded as the Basketball Association of America (BAA) in New York City on June 6, 1946 (NBA, s. f.). The league adopted the name National Basketball Association (NBA) in 1949 after merging with the rival National Basketball League (NBL). The NBA is currently the most significant professional basketball league in the United States of America, in terms of popularity, salaries, talent, and level of competition (Patterson, 1993; Hausman & Leonard, 1994).

The NBA is a league of closed structure (no promotions and demotions), composed by 30 franchised members, which 29 are located in the United States and one in Canada. The current league organization divides the 30 teams into two conferences of three divisions, with five teams each. The current divisional alignment was introduced in the 2004–2005 season.

During the regular season, each team plays 82 games, 41 at home and 41 away. A team plays against its opponents in its own division for four times per season (16 games), three or four times against teams from the other two divisions in its own conference (36 games) and twice against teams in the other conference, respectively (30 games). This asymmetric structure means that the strength of the schedule varies significantly among teams.

The NBA organization chart is constituted by the CEO and different departments, which directly depend on the CEO. The most remarkable, regarding another professional basketball leagues, is that in the NBA referees are professionals and depend directly on the NBA (do not depend on the Federation). The NBA departments are:

- Basketball Operations
- Broadcast Operations
- Communications
- Community and Player Programs
- Creative Services
- Events and Attractions
- Facilities and Administration
- Finance and Benefits
- Global Marketing Partnerships
- Global Merchandising Group
- Human Resources
- Information Technology
- Interactive Services
- International
- International Media Distribution
- Legal
- Legal and Business Affairs
- Marketing
- D- League (NBA Development League)
- WNBA

- NBA Entertainment Production, Programming and Photos
- Referee Operations
- Security
- Strategic Development
- Team Marketing and Business Operations

Source: NBA

Salary cap and Draft

The NBA has several mechanisms established in order to prevent teams, with large surpluses of profit, can sign the best players available, thereby facilitating the maintenance of equality in the league. The more representative mechanisms are the salary cap and the draft. The lottery draft process has changed along the years, but exists since NBA foundation; unlike the salary cup, which started in the mid-1940s, (it was abolished after only one season); and reinstated in the 1984–85 season.

Salary cap

The North American professional sports leagues (Major League Baseball (MLB), National Basketball Association (NBA), National Football League (NFL), and National Hockey League (NHL) have an agreement or rule that sets a limit on the amount of money that a team can spend on player payrolls, called salary cap (Scott, Long, & Somppi, 1985). The salary cap, in the NBA, started in 1983 (Hill & Groothuis, 2001).

A simple model shows that a salary cap can improve the competitive balance among clubs as well as the salary distribution among players (Késenne, 2000). For that reason, every franchise has to study carefully what market players could be interesting for his project (depending on the team´s project). The salary cup is defined by the league's collective bargaining agreement (CBA). The salary cap ensures that each franchise can only “shield” economically one or two key players, who are often called franchise players. There are three kinds of regulations: hard salary cap, soft salary cap (with luxury tax), and luxury tax (The NBA utilizes a soft salary cap) (Scully, 1989; Késenne, 2000; Fort & Maxcy, 2003):

- ***Hard salary cap.*** A hard salary cap is where the league sets a maximum amount of money allowed for player salaries, and no team can exceed that limit. At the beginning the salary cap was not a hard cap (Hill & Groothuis, 2001).

- ***Soft salary cap.*** A soft salary cap has a set limit to player salaries, but there are several major exceptions that allow teams to exceed the salary cap. For example, in the case of the NBA, teams can exceed the salary cap when keeping players that are already on the team (Dietl, Franck, Lang, & Rathke, 2010).

- ***Luxury tax:*** A luxury tax system does not have a limit to how much money can be spent on player salaries. However, there is a tax levied on money spent above a threshold set by the Collective Bargaining Agreement (CBA) between the players union and the owners. For every dollar a team spends above the tax threshold, they must also pay some fraction to the league. This system is used to discourage teams from greatly exceeding the tax threshold, with the goal of ensuring parity between large and small market teams (Dietl, Lang, & Werner, 2008).

NBA lottery draft

The NBA Draft is for the procedure by which, in late June each year, the NBA franchises join to their teams, players from U.S. universities or leagues in other countries. These players are usually amateur U.S. college basketball players, but international players are also eligible to be drafted. College players who have finished their four-year college eligibility are automatically eligible for selection, while the underclassmen have to declare their eligibility and give up their remaining college eligibility.

The first draft took place in 1950 with the aim to provide good players to the league (see above, NCAA). Teams could forfeit their first-round pick and select a player from their immediate geographical area, commonly known as a "territorial pick" (NBA, 2007). In 1985 they changed to a lottery system, this NBA Lottery system set the order of selection for the non-playoff teams (or the teams holding their picks through trades) for the first round only. Teams picked in inverse order of their records in the second round in all succeeding rounds. In 1990, the NBA changed the format of the lottery to give the team with the worst record the best chance of landing the first pick. Currently NBA Draft consists of two rounds, the first and second round. Each of the thirty league teams has a choice in each. A team may transfer future rounds of the draft during the competition (until the transfer period ends) and the post-season. The order of selection is made according to the following procedure (NBA, 2003):

- In positions 15-30, are located 16 teams qualified for the playoffs, in inverse order of their winning percentage at the end of regular season (the more wins a team gets, the lower the position).

- In the first fourteen positions distributed teams not qualified for a playoff. The ranking is established by a draw (the lottery), which are more likely to occupy the first places the teams that have accumulated more losses.

As we can see actually, only the top three picks are deemed lottery picks. These positions are chosen from the 14 teams that participate in the playoffs. But not all teams have the same chances. I.e. the worst team gets the most chances (more balls in the lottery process). After this process, the remainder of the first-round draft order is in inverse order of the win-loss record for the remaining teams, or the teams who originally held the lottery rights if they were traded. The lottery does not determine the draft order in the subsequent rounds of the draft.

USA BASKETBALL

USA Basketball is a non-profit organization and the governing body for basketball in the United States. It is responsible for representing his country at the FIBA and at the U.S. Olympic Committee in all matters relating to the basketball and as well as for some national competitions. USA Basketball is an organization made up of organizations. There are five member categories:

Professional:

- National Basketball Association
- National Basketball Association Development League
- Women's National Basketball Association

Collegiate:

- National Association of Intercollegiate Athletics
- National Collegiate Athletic Association
- National Junior College Athletic Association

Scholastic:

- National Federation of State High School Associations

Youth:

- Amateur Athletic Union.

Associate:

- Athletes In Action
- Basketball Travelers
- College Commissioners Association
- Gazelle Group
- Harlem Globetrotters
- International Sports Exchange
- Latin-American Basketball League of Los Angeles
- National Amateur Basketball Association
- National Association of Basketball Coaches
- National Basketball Players Association
- National Junior College Basketball Coach Association
- National Junior College Women's Coach Association
- National Wheelchair Basketball Association
- USA Deaf Sports Federation
- United States Armed Forces
- Women's Basketball Coaches Association.

3.3.3.2. Organizational structure of basketball in Spain

It is important to know, in advance, the organizational framework that basketball is developed in Spain, in order to understand its current state. Next we are going to describe briefly the organizational structures related to basketball in Spain, but only the most relevant for this work. The main classification is made by the economic source of the organization. There is a distinction regarding whether the structure depends on government or it is a private organization (Blanco, 2006):

Sport State Organizations

Public Sector

- C.S.D. (Consejo Superior de Deportes).
- Autonomous Regions Sports Department (Direcciones o Consells de Deporte).
- Local Government Services (Concejalias).
- University Services.

Sport Private Organizations

Private sector

Association Network (non-profit)

National level

- Spanish Olympic Committee.
- Sports Federations.
- Clubs Associations.
- Sports promotion organizations.
- Leagues.

Autonomous Regions Level

- Regional Federations.
- Different Forms of Association.

Business sector (for profit).

- Companies of sports services.
- Sports corporations.

CSD (National Sports Council)

The National Sport Council replaces the old National Sports Delegation, created in 1941, under the Secretary General of the Movement. The National Sport Council is defined by the Law act 10/1990, as an Autonomic Structure with administrative functions; throughout the Government manage sport in all the national territory.

The powers of the National Sports Agency are:

- To authorise or revoke the constitution with due justification and to approve the statutes and regulations of Spanish Sports Federations.

- To recognise, to the intents and purposes of this Act, the existence of a sporting discipline.

- To establish, in conjunction with the Spanish Sports Federations, their objectives, sports programmes (especially in high-level sport), their budgets and organic/functional structures, subscribing to the corresponding agreements. These agreements will be of a legal-administrative nature.

- To grant economic subsidies due to Sports Federations and other Sporting Bodies and Associations, carrying out inspections and verifying that they comply with the aims set out in the current legislation.

- To classify official, professional, state competitions.

- To promote and propel scientific research in the field of sport, in accordance with the criteria established in the General Promotion and Coordination Act for Scientific/Technical Research.

- To promote and propel measures to prevent, control and repress the use of prohibited substances and illegal methods aimed at artificially enhancing athletes' physical capacity or at modifying competition results.

- To act in coordination with the Autonomous Regions regarding general sporting activity and to cooperate with these regional governments in the development of the powers vested in them through their respective statutes.

- To authorise or deny, with prior agreement from the Ministry of Foreign Affairs, the celebration of official, international sporting events in Spain, as well as the participation of Spanish teams in international competitions.

- To coordinate school and university sports programmes with the Autonomous Regions, when they have national or international influence.

- To design and put into practice, in collaboration with the Autonomous Regions and, where relevant, local Bodies, plans for the construction and improvement of sports facilities for the development of high-level sport, as well as to update, within the field of its powers, the technical regulations in existence regarding this type of facilities.

- To draw up proposals to establish minimum educational content for qualifications for specialised sports coaches. The Agency should also help in the establishment of study programmes and plans relative to said qualifications, accrediting authorised centres to impart the courses, and monitoring the development of training programmes in those Autonomous Regions that have not assumed powers in the field of education.

- To authorise annual spending of the Spanish Sports Federations as anticipated in the regulation, to determine where the net assets of these federations should go, in case of their liquidation, to control the grants that would have been awarded to them and to authorise the taxing and disposal of their real estate when the federations have been fully or partially financed by public state funds.

- To permanently update the census of sports facilities in collaboration with the Autonomous Regions.

- To authorise the inscription of Sports Companies in the Sports Associations Register, independent of their inscription in the registers of the corresponding Autonomous Regions.

- To authorise the inscription of the Spanish Sports Federations in the corresponding international Sports Federations.

- To collaborate in the field of the environment and countryside protection with other, corresponding public bodies and with the Federations that are specifically linked with them.

- Any other faculty attributed by law or through the regulations that contributes to achieving the goals and objectives set out in current legislation.

Source: CSD (Sports Act/10/90, of 15 October, Art.8)

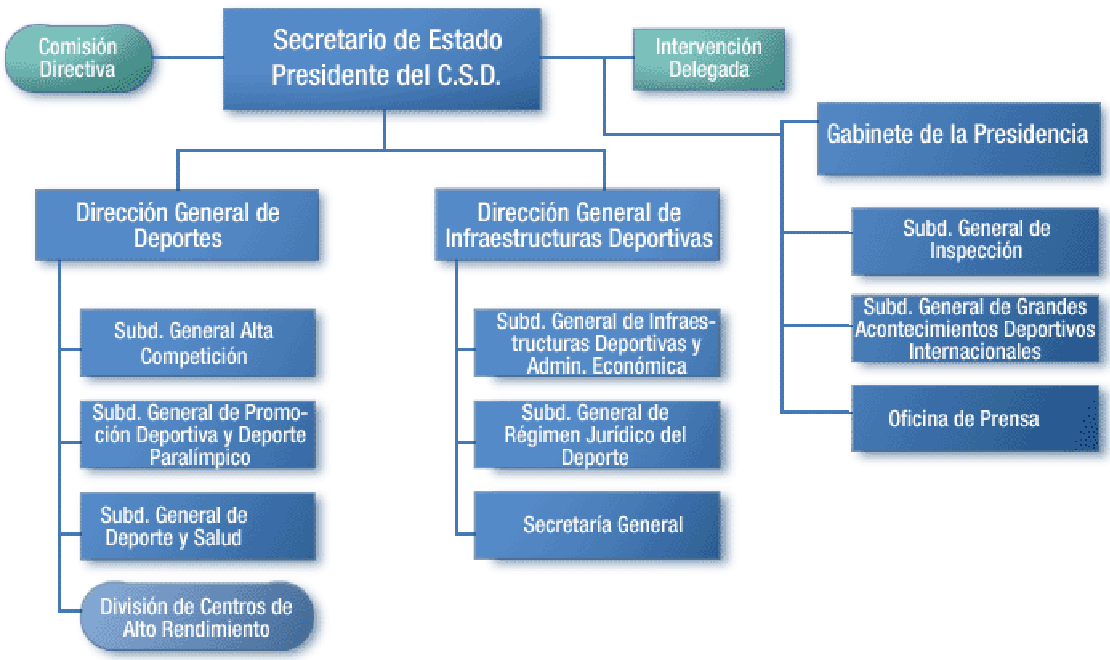


Figure 5. National Sport Council structure. Source: **C. S. D.**

FEB (Spanish Basketball Federation)

It is a non-profit organization which in charge of promotion, management, and coordination throughout the national territory of basketball, in all its manifestations and variations.

The farm leagues are managed by FEB (semi-professionals and farm leagues), and also manage the Autonomous Regions Federations (for amateur competitions), but the professional league of Spain is managed by ACB (association of several sport clubs).

All the participants in competitions organized by FEB are integrated into the federation, such as corporations, sports, sports clubs, athletes, coaches and referees. And also it is responsible for issuing all licenses necessary for participating in its activities.

The FEB is affiliated to FIBA as a member, being obliged, therefore to follow its statutes and regulations in all matters affecting the technical order and international relations. Internationally, the FEB is the representation of Spanish basketball in basketball international official activities and competitions celebrated within and outside the Spanish territory. The FEB elaborates the rosters of national teams (seniors and farm teams). The technique structures of the Basketball Federation are:

1. Administration and representation:
    - General Assembly and Executive Committee
    - The president

2. Management:
    - Executive Committee
    - Committee on Regional Federations
3. Consulting:
    - Area executive committee

4. Internal Management System:
    - General Secretary
    - Management
    - Those that could be created to better fulfill the federative purposes

5. Technical-Sporting
    - General Secretary
    - Competition
    - Referees
    - Coaches
    - Committee on Health and Prevention of Doping

6. Discipline
    - National Competition Committee
    - National Appeals Committee

ACB (Basketball Club Association [Asociación de Clubs de Baloncesto])

The ACB League is nowadays the main men's professional basketball league in Spain. It began in 1957, with the name of National League, and was originally organized by the Spanish Basketball Federation (FEB). In 1983-84 season the ACB was established with its own competition format and replaced the National League. The league is rated as one of the three "A" level European national domestic leagues in the ULEB League Rankings system (ACB, 2012).

In Spain, the professional leagues are private structures that carry out functions of public interest and are supervised by the National Sport Council (Millán Garrido, 2010). The ACB model presents an open structure which means there are promotions and demotions. The top teams classified in the regular season play a championship in a play-off format. The last ranked teams are relegated to a lower division and in turn are replaced for the two top ranked teams of the bottom category. The season 2011-12 have participated a total of 18 teams, but this number has varied in previous seasons (from 13 until 24 teams).

The ACB sport model consists in a regular season of double confrontations (only two games against the same team). The order of each team's first-half fixtures is repeated in the second half of the season.

Concerning to the administrative structure, the ACB is the association of several clubs (which have their own teams (farm teams) and even other different sport teams) with its own organizational structure. It means that some teams can participate in other European leagues simultaneously, such as Euroleague, ULEB, etc. unlike the NBA and NCAA. But the participation in these leagues is related with the standing of the previous season. Other difference with the NBA and NCAA is that the referees belong to the Spanish Basketball Federation (FEB).

The central administration revolves around a Steering Committee composed of five members on direct dependency on the Directorate General. The maximum responsible has the General Assembly as the ultimate authority for the final decision. The five members shall be directors of the six strategic areas plus the Director of referees.

The current organizational structure of the ACB comprises six areas: Competition, Media, Events, Institutional, Commercial and Administration. The last three sections also cover the rest of the organization transversally.

The Competition Area includes the departments of Competition, Scouting and External Relations, including work and relationship with the various international competitions: Euroleague, ULEB and others. In front of this area will be the CEO of the ACB.

The TV and Communication area contains the departments of Communication, ACB.COM, Audiovisuals and Television. This new area was created with the purpose of managing, in a comprehensive manner, the relationship with television operators of the Association.

The areas of Events Management and try to increase brand value and cost of major events in the ACB such as the *Copa del Rey* and the *Super copa*, and also deal with Finance, Information Technology and Statistics issues.

The Institutional Area (General Secretary) is responsible for all legal affairs and documentaries of the Association and manages the Human Resources Department and General Services.

Business Area (Business and Marketing) includes the departments of Marketing and Business Development. These are key in generating income for the organization through national and international agreements, sponsorship plan and fitness centers promoted by the Association.

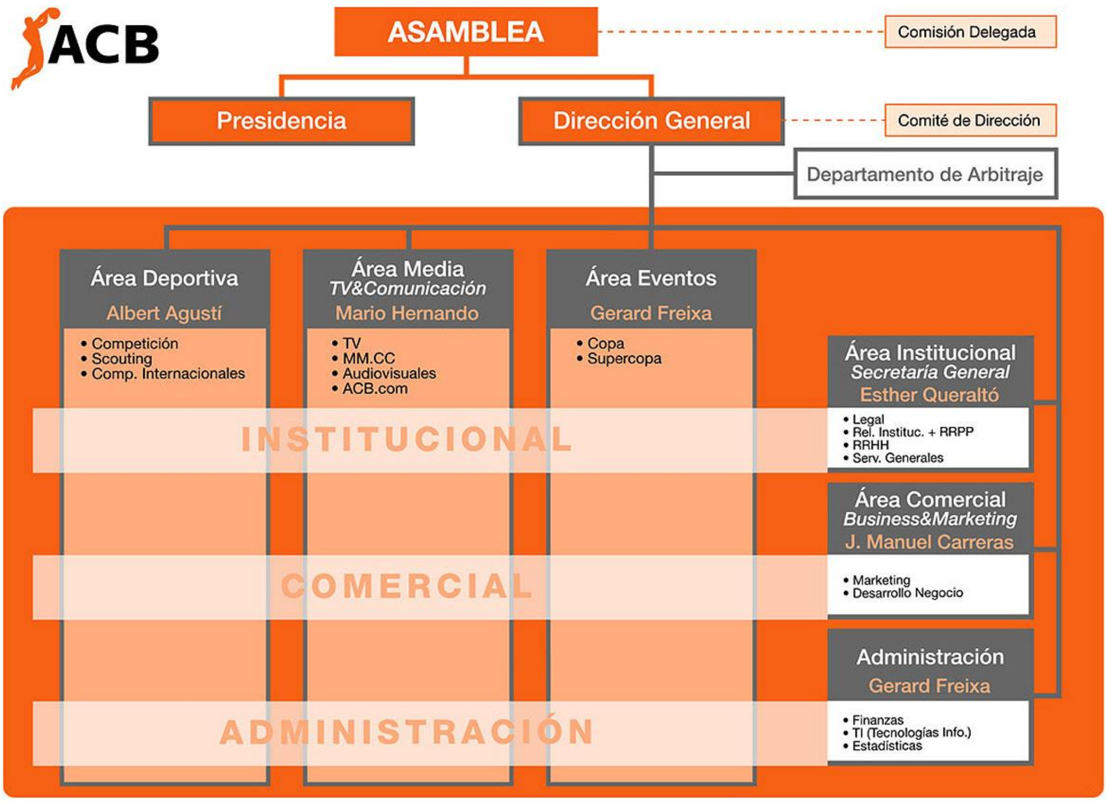


Figure 6. Current organizational structure of the ACB. Source: **ACB.**

3.3.4. Structural comparison between beginning and the present

The architecture of the basketball network can provide us a new and good perspective of how is the organization of basketball. Certain organizational and functional principles in complex systems are universal because some networks have similarities to other biological and technological networks (Solé & Goodwin, 2002). Hence we represented the basketball network (focused on ACB and NBA) in the season 1983-1984 (Figure 7), when the ACB was founded, and in the season 2008-2009 (Figure 8), in order to understand their evolution during all these years. We simplified some structures (such european leagues) for better understanding.

1983-1984

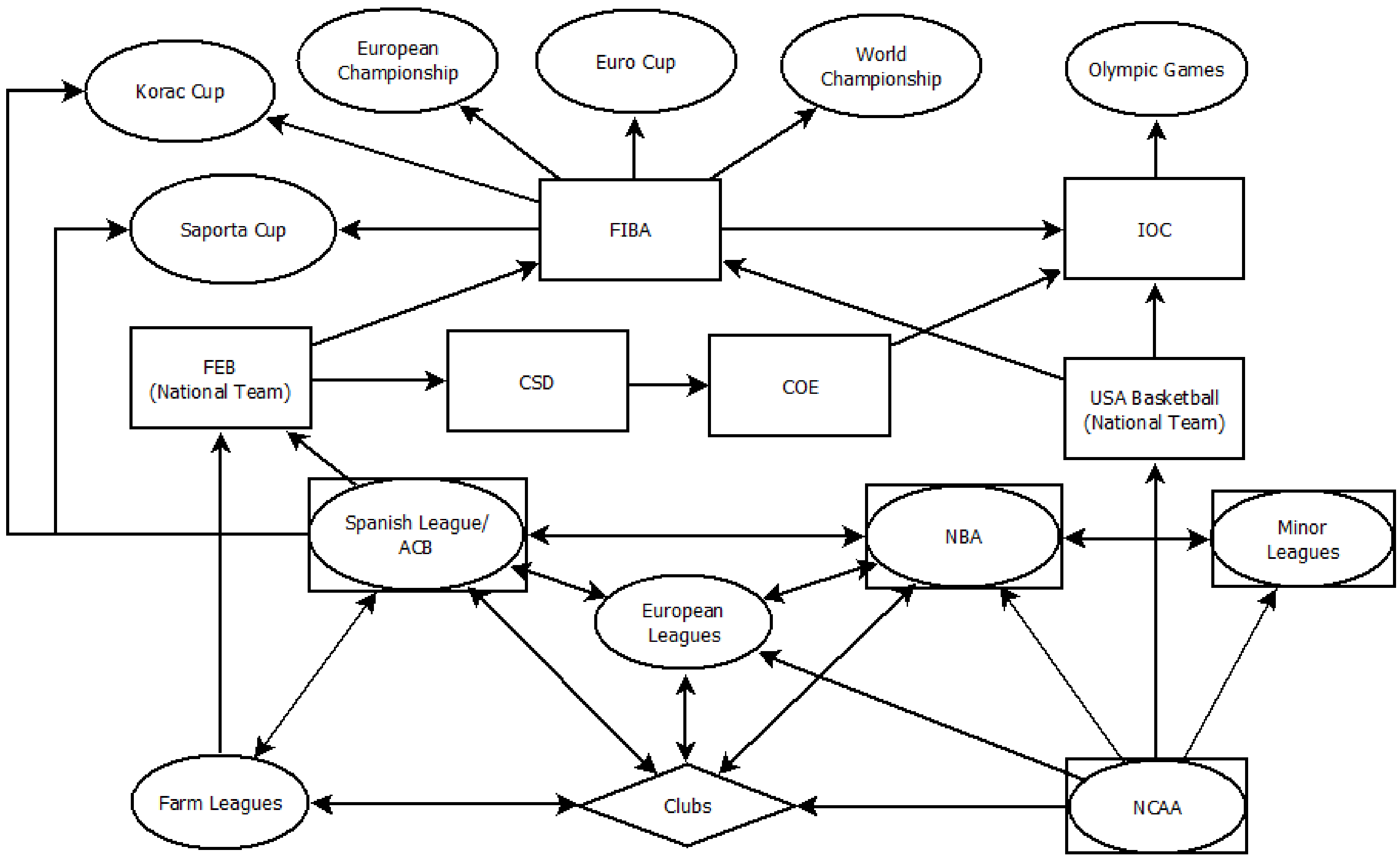


Figure 7 Representation of the network of institutions and basketball competitions. Some structures (such Farm Leagues and European leagues) have been simplified for better understanding. The circles represent the competitions (leagues/tournaments/championships) and the squares represent the institutions. We can see how some structures can be considerate hybrids, because they are competitions and organizations simultaneously. Not all the relationships are the same (some relationships are unidirectional and others bidirectional). There are dissipative structures (such the NCAA) and structures whose goals are to condense or concentrate the resources. But the most active structures of the network are those that consume and produce resources at the same time (such the ACB). We note that the network is based on these kinds of structures. The triangle based on the ACB-NBA - Sport Clubs is the engine of the system.

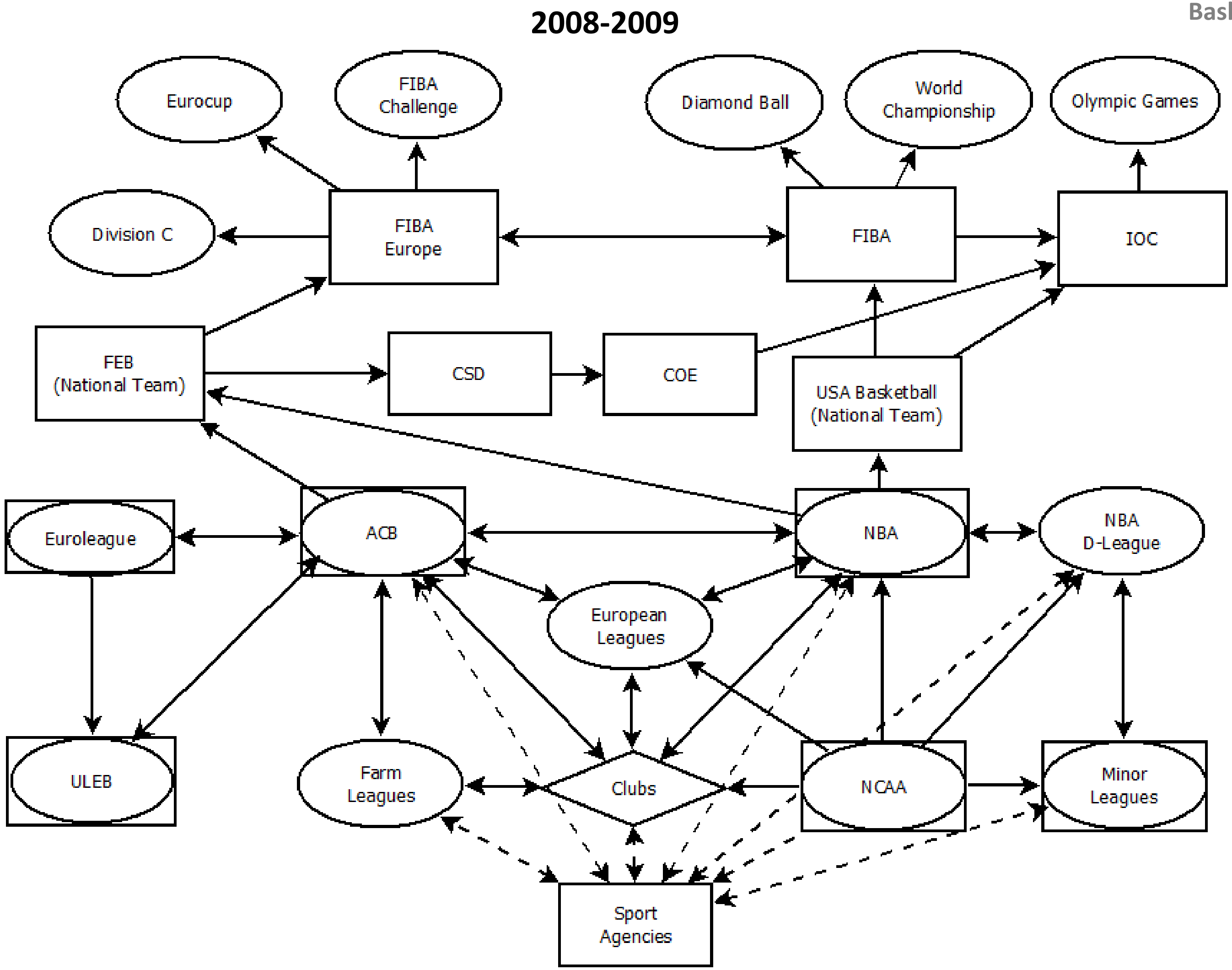


Figure 8. Illustration of basketball network in the season 2008-2009. The representations of the elements are the same as before: circles symbolize the competitions (leagues/tournaments/championships), squares represent the institutions, and finally square plus circles are hybrid structures. The network has evolved: some structures have disappeared and new emerged. Even some structures have adapted to the evolution in time and the grown of the network by a bifurcation of one of its structures (such FIBA in FIBA Europe), in order to preserve the effectiveness of the flow within the network. We are still seeing how the triangle formed by ACB-NBA-Sport Clubs is the key of the network, but whit the emergency of more professional leagues, the relationships have been modified. The network has become more professional.

The Figure 7 and the Figure 8 represent our proposal of the basketball network graph using the structures existing in Europe and USA and the institutional relationships among them. We proposed players as raw material and the arrows as the channels what they follow. Hence we can see, for example, the path that a player can take from a domestic league to the World Championship or the Olympic Games. We simplified some structures such the European leagues (Greece League, Italia League, Croatia League, etc.) and farm leagues in Spain (Gold LEB, Silver LEB, EBA, etc.) inasmuch as we considered them as similar structures with the same properties.

At the graph the circles represent the competitions (leagues/tournaments/championships). The squares represent the institutions, such as the FIBA or the IOC. There some structures which are competitions with its own organizational structure (they can be independent or semi-independent from other structure such as federation), for instance ACB or NBA, hence can be considerate hybrids, because they are competitions and institutions simultaneously. There is a structure which we highlighted as special, because has a composition very different from the others, we refer to the sport clubs in Spain. This element possesses some properties different that a team: are managed by the president, the board and general assembly. There are clubs devoted exclusively to basketball activities and other clubs where the same sport club has several sections: football, basketball, volleyball, handball, hockey, etc.

In addition, clubs have player farms, where they train young players and club teams competing in affiliated minor leagues, in order to create a young sport star, or as a showcase for other clubs. This structure is more wide and developed than a team, and enables the club maximize its resources, sharing common management structures, facilities, human resources, material resources, economic, etc. which enable to prepare rosters much more competitive.

There are some productive/dissipate structures (such the NCAA or FIBA) whose aim is to produce or spread the material in the network. On the other hand there are structures whose goals are to condense or concentrate the resources, for instance the national federations. But the most active structures of the network are those that consume and produce resources at the same time (such the ACB). We note that the network is based on these kinds of structures, and the figure of the sport clubs are the most representative. The union of these structures is the engine of the system, as we can observe at the triangle based on ACB-NBA-Clubs.

Regarding to the connections not all the relationships are the same (some relationships are unidirectional and others bidirectional). They indicate de flow within the network and the structures which are connected somehow.

We can observe how at the first graph (Figure 7), in the season 1983-1984 the most part of the leagues are amateur such as the Spanish Farm Leagues, the ACB (before become professional) and the NCAA, which provide players to the USA national team (Olympic Games). These structures are which support the most part of the international competitions in addition to the NBA (professional league).

The Figure 8 represents our proposal for the basketball network in the season 2008-2009. We can see how some structures have disappeared and have appeared new ones. The most remarkable fact is that the professionalization of the basketball has influenced notably in the architecture of the network. We can appreciate how the players for the USA national team, now are provided by the NBA (professional league), and the same happens in Spain, the ACB (now professional) supplies players for the national Spanish team (even the NBA provides players for the Spanish national team nowadays).

An important fact is that the FIBA has created a European delegation (FIBA Europe was founded in 2001). This process of emergency indicates a high growth of the flow to international competitions between national teams in Europe, in our case. Moreover, the European competitions have changed their format or have been replaced with new ones. As we mentioned, in the season 1983-1984 these competitions were played by sport clubs, in Europe. But in the season 2008-2009 are played by national teams (competitions which are dependent on the FIBA Europe).

The most remarkable fact related with the design of the network, is the appearance of two hybrid structures (as we called), the ULEB (founded in 1991) and the Euroleague (founded in 2000). This emergency is, probably, one of the sources of the network restructuration, because we have to take into account that both of these competitions are played by sport clubs, instead of national teams.

In fact, in 2000, major professional sport clubs on Europe, led by the Spanish, Italians and Greeks, grouped in the Union of European Basketball Leagues (ULEB), split off from the FIBA in order to organize a new Euroleague with modern management criteria. These clubs wanted to receive more revenue from television broadcasting rights and merchandising than the offered by the FIBA.

In addition, the NBA created its own farm league: the NBA Development League (NBA D-League); in which participating teams plays their own league (there is not promotions to the NBA). Some NBA teams share the resources but players. This league depends on the NBA.

We past from 8 competitions, 6 institutions and 4 hybrid structures, in season 1983-1984, to 7 competitions, 8 institutions and 6 hybrid structures, in season 2008-2009. Note that the enlargement is related with the hybrid structures, which indicate us that the professionalization provides the apparition of these kinds of elements. These are good examples of the how a process such as the professionalization of a sport, can change the composition of the its structure. We also can appreciate how the clustering is not homogeneous in the two networks (table 2).

Table 2. Number of connections of the network nodes in the season 1983-1984 (left) and in the season 2008-2009 (right).

1983-1984		2008-2009	
Agent	**Connections**	**Agent**	**Connections**
FIBA	8	ACB	8
ACB	7	NBA	8
Clubs	5	Sport Agencies	7
NBA	5	Clubs	6
NCAA	5	FIBA	6
European Leagues	4	FIBA Europe	6
FEB	4	European Leagues	5
IOC	4	NCAA	5
Farm Leagues	3	FEB	4
USA Basketball	3	IOC	4
COE	2	NBA D-League	4
Korac Cup	2	Farm Leagues	3
Sporta Cup	2	Minor Leagues	3
CSD	2	USA Basketball	3
Minor Leagues	2	COE	2
EuroCup	1	CSD	2
European Championship	1	Euroleague	2
Olympic Games	1	ULEB	2
World Championship	1	Division C	1
		EuroCup	1
		FIBA Challenge	1
		World Championship	1

There are other structures that enhance the movement of players, which no participate directly in the network, but influence strongly in it. We refer to the Player Agencies (or Sport Agents). Their aim is to localize players and move to another structure of the network where is needed; in return for part of the benefit created.

This confirms that the nature of the basketball network has changed and that the flows are not constants.

3.4. Competition Analysis

Several disciplines such as economy, applied statistic, physics, evolutive biology, social sciences, sport sciences, etc. have been tracking the evolution of a system within a controlled environment, often through analysing the interactions of the agents involved.

Our aim was to investigate from an overview, the intern dynamic of professional basketball leagues studying its competitiveness degree. The degree of equality of the playing strengths of teams, competitive degree or competitive balance, is a central concept in the analysis of professional sports leagues.

There is considerable interest in clarify the skewness and fluctuations in competitive balance throughout seasons; and analysing the effects of regulatory, institutional and other changes, as indicated by the extensive literature on the subject (Schmidt & Berri, 2001; Fort & Maxcy, 2003; T. A. Rhoads, 2004; Goossens, 2006) and applied in different sports such as baseball (Scully, 1989; Owen, Ryan, & Weatherston, 2007), American football (Humphreys, 2002), basketball (Noll, 1988; Berri, Brook, Frick, Fenn, & Vicente-Mayoral, 2005), ice hockey (Richardson, 2000), European football (soccer) (Halicioglu, 2006) or golf (T. Rhoads, 2005).

The most part of works deal with this phenomenon with regard to the mechanics of the game itself, meaning the game in isolation, without implications to the competition (league)(Chatterjee & Yilmaz, 1999; McGarry, Anderson, Wallace, Hughes, & Franks, 2002; Lebed, 2006; McGarry & Franks, 2007; Passos et al., 2008; Passos, Araújo, Davids, Milho, & Gouveia, 2009), nevertheless, few works do from the perspective of competition between teams in different sports (Yilmaz & Chatterjee, 2000; Malacarne & Mendes, 2000; Onody & de

Castro, 2004; Mendes, Malacarne, & Anteneodo, 2007; Vaz de Melo, Almeida, & Loureiro, 2008; Ribeiro, Mendes, Malacarne, Jr, & Santoro, 2010).

The competition model (type of confrontation: league, conferences, groups, play-off, schedule, etc.) directly affects on competition development. So small changes can alter the outcome, given the close relationship between the competitive model and competition.

A league is more competitive when there is more random. In fact, the more difficult to predict the final outcome, the harder is the competition. Whether all teams have a similar performance level, an imbalance can be shaped by a small change. However, when the competition is less random, the degree of competitiveness decrease significantly, therefore, we have a competition where there are nice prospects to know the final outcome.

Some authors point out that a sport competition with a high competitive balance is more attractive (Goossens, 2006; Quirk & Fort, 1997). The more competitive the league, the more revenue generates (tickets, sponsors, TV, etc.) and is more attractive for fans and media (Soebbing, 2008; Ribeiro et al., 2010; Watanabe, 2012). Cairns, Jennett and Sloane (1986) introduced the different dimensions of competitive balance. They proposed distinct usages of what they called "uncertainty of outcome". These authors consider four kinds of competitive balance. The first one, called match uncertainty. Secondly, the season uncertainty deals with the uncertainty within a single season. The third kind is the dominance of a few teams over seasons called championship uncertainty. Fourthly, the uncertainty of the outcome: the absence or presence of long-run domination by one club can produce a decrease of interest by other clubs fans and even sponsors. This can depend on the levels of seasonal uncertainty with which it is associated.

1. Match uncertainty.
2. Seasonal Uncertainty.
3. Championship Uncertainty.
4. Absence of long-run domination.

Szymanski (2003), uses the same classification as well. But he only makes reference to the first three distinctions. Berri et al. (Berri et al., 2005) noted that every time a competitor reached a level of dominance, the uncertainty of outcome has been compromised, the demand for the output of this industry is expected to decline. Some authors (Knowles, Sherony, & Haupert, 1992; Rascher, 1999) observed that fan attendance in Major League Baseball is maximized when the probability of the home team winning is approximately 0.6. If the home team presents a higher probability of success, the fan attendance is expected to decline. Consequently, given the importance of fan attendance to a league financial success, leagues are expected to implement rules and institutions design to address the relative strength of teams on the games.

The most part of the literature, which are focus on exploring sports and competitive balance, often refer to mechanisms enacted to improve the distribution of wins within a league, such as salary caps, the rookie draft, reserve clause, etc. Indeed, we observe how these mechanisms and the processes which cause it, can be related with the competitiveness variability during several seasons (de Saá Guerra et al., 2012). (Berri et al., 2005) says that: *whenever one competitor reaches a level of dominance where uncertainty of outcome has been compromised, the demand for the output of this industry is expected to decline*.

3.5. Game analysis

Some authors point out that the general performance of some professional basketball leagues present a non-homogeneous behavior (Yilmaz & Chatterjee, 2000; de Saá Guerra et al., 2011) and even some events seem to match with significant changes in the level of competitiveness in the league. For instance, teams try to incorporate into their rosters the best players available. Leagues modify rules and competition system in order to increase the performance. These modifications can have a directly effect on the level of uncertainty before and during every single game. Hence, the study of the game is a key for understanding of the basketball dynamic and the effects that the external modifications could produce on the game itself.

We focused on the score and the time-scoring as the indicators of the dynamic of the game, because the evolution of the score is a great indicator of the uncertainty in every single game. We have to take into account that there are a lot of mechanisms that make the scoring process an exciting and barely unpredictable phenomenon. Hence, we cannot know, in advance, the behavior of scoring dynamic of a basketball game.

Some authors have studied the data distribution by using statistical models, such us Poisson process, negative binomial, log normal or generalized extreme value distribution (Malacarne &

Mendes, 2000; Greenhough, Birch, Chapman, & Rowlands, 2001; Mendes et al., 2007; Bittner, Nußbaumer, Janke, & Weigel, 2009; Heuer, Mueller, & Rubner, 2010). Some of them highlight the presence of heavy-tailed distributions (*Power Laws*), which are associated with many natural and social phenomena. Some similar distributions appear in other works which relate this to fields such as statistical physics and non-linear complex systems, and ideas related with that theories such as anomalous diffusion by the Zipf-Mandelbrot law (Malacarne and Mendes, 2000), self-organized criticality phenomena, or from non-linear dynamics (McGarry, Anderson, Wallace, Hughes, and Franks, 2002; Bourbousson, Sève, and McGarry, 2010).

Time

The parameter of time is one of the most important elements in basketball. There are several rules related with time and thought years theses have been modified according to the evolution of the game. One of the most representative rules what deal with time is the 24 second shot clock, which oblige teams to make a shoot to the basket before the 24 runs out. The idea of these kinds of rules is to promote high scores in general. Hence the influence of the time in the game is highly remarkable (Grasso, 2010).

In November of 1950 the game Minneapolis vs. Fort Wayne finished with 18-19 for Fort Wayne. This fact caused a crisis, and the commissioner Maurice Podoloff thought that this kind of games could not go on. Danny Biason, president of Syracuse Nats found a solution in 1954. He devised the current 24 seconds shot clock rule based on the mathematic formula (NBA):

4 quarts/game x 12 minutes/quarter x 60 seconds/minute =
= 2880 seconds /game

2880 seconds / 120 possessions or shots (estimated mean per game) =
= 24 seconds / possession

The aim of this idea was to oblige teams to shot. And try to make game more attractive and excitant. The FIBA also adopted a rule regarding shooting time after the game Hungry v.s. URSS in 1956. But the used the formula:

20 minutes/half x 60 seconds /minutes x 2 halfs/game =
= 2400 seconds /game

2400 seconds / 80 possessions or shots (estimated mean per game) =
= 30 seconds / possession

They used 30 shot clock instead. But in 2000-2001 season, the FIBA changed ball possession time from 30 to 24 seconds (Gómez Ruano, Lorenzo Calvo, & Sampaio, 2003; Palao Andrés, Ortega Martín, Piñar López, & Ortega Toro, 2004; FIBA, 2012) with the aim of accelerate the game and obtain higher scores. This modification has a deep effect on the dynamic of the game and it is one of the best examples how the modification of a rule can modify the internal process of the game (an external parameter can modify the evolution of a system).

To carry out this analysis, we focused on the time elapsed between points achieved by any team during the game, in all the games that we analyzed. In order to find out how the time behaves in a basketball game, and whether any pattern exists that help us to clarify how a basketball game works.

Scoring

What really displays the dynamic of a basketball game is the score. The score reflects all the successful events (points) achieved for teams and sets up what team wins. Hence the performance parameter of a basketball game is the score. What really matters is to achieve more points than the other team, so each team use different strategies in order to defeat the opponent. It means that there is not only one strategy. Each team adapts taking into account its resources and to what it is facing.

There are two ways to approach this matter. The first one is to focus on the absolute score. But the absolute score increase until the game is over, so only can indicate us the dimension of the how many have been scored.

The other parameter that indicates us how the game is developing is the point difference. It the reflection of the real game because deal with the points accumulated and the score runs. This is very important because the score runs are regarding with the time; and this the way in which both parameters are related (Chatterjee & Yilmaz, 1999; Kubatko, Oliver, Pelton, & Rosenbaum, 2007; Vaz de Melo et al., 2008; Sampaio, Lago, Casais, & Leite, 2010).

Complex Systems Background

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

4. Complex Systems Background

A complex system is a set of several elements (also called agents) which are related among them and whose links contain information hidden to the observer. The established relationships among them are mainly type non-linear. These interactions are local interactions. That is, affect only the relationship between an agent and to the elements which surround him, but none of them is aware of the collective behavior (Goodwin, 2002; Vicsek, 2002; Amaral & Ottino, 2004; Solé, 2009).

These processes, that take place simultaneously on different levels or scales, are important. In fact, the way in what its units are related, greatly influences in the output of the entire system. That is why the laws that describe the behaviour of a complex system are qualitatively different from those that govern its units (Amaral & Ottino, 2004; Vicsek, 2002).

As a result of these interactions, new properties emerge that cannot be understood from the individual features of each element. These properties are called *emergent properties*. That is why a complex system must be treated as a whole, from a holistic conception, not just the elements that constitute it because in a complex system the whole is greater than the sum of the parts. Complexity is the result of incessant adaptive processes (Holland, 1995).

4.1.No-linear

When the system is linear, the same stimulus always produces the same outcome. Every time the process is repeated, the same outcome will be obtained. On the other hand, if the system is non-linear, a stimulus can yield several results. Although the conditions are the same, the outcome or outcomes cannot be known in advance (Prigogine & Holte, 1993; Solé & Goodwin, 2002; Amaral & Ottino, 2004).

The interrelationships of the components of the complex system are governed by non-linear equations. As mentioned above, not always effectiveness in sport shows a linear behavior, but there are several actions that can be considered effective, and also do not have to be consecutive. Complexity, in itself, is a measure of the number of possibilities. Such equations often have a strong dependence on initial conditions of the system, which makes it even more difficult to assess their behavior

4.2. Self-organization

The idea of self-organization can be expressed as the general tendency of a given system to generate behavior patterns from local interactions of its constituent elements and from the relationships with the environment. It is the essential part of any complex system and allows the system to recover the balance, modified and adapted to the surrounding environment. Usually, the different system elements are self-regulated by themself always seeking to optimize the overall operation of the assembly. The complex network of interdependent systems in which human beings can organize, for example, is changing and readjusting to reality that corresponds to live in each moment (García Manso & Martín González, 2008).

Self-organization is a process in which the internal organization of a system increases in complexity without being guided or managed by an outside source. Self-organizing systems usually display emergent properties.

The order and disorder need each other, mutually occur. They are antagonistic concepts but complementary at the same time. In some cases, some of disorder allows a different order and sometimes, richer. For example, an organism can persist as a result of the death of its cells, or an organization is perpetuated by the dismissal of its members. The variation and change are inevitable and unavoidable stages through which every complex system must travel to grow and develop. When this transformation is achieved without the involvement of external factors to the system, referred to a process of self-organization (Nicolis & Prigogine, 1977).

Self-organization stands out as an essential part of any complex system. It is the form through which the system recovers the balance, changing and adapting to the surrounding environment (it responds to external aggressions that seek to modify its structure).

In this kind of phenomena is essential the idea of levels. The interrelationships among the elements of a level originate new types of elements on another level which behave quite differently, for example, from molecules to macromolecules, macromolecules into cells and from the cells to tissues. Thus, the self-organizing system is built as a result of increasing order space-time which is created in on different levels or layers, one above the other.

To a large extent, complex systems can be understood like a machine that generate order, which requires constant energy intake generated by the chaos that feeds (it is an open system and dissipative). The self-organizing complex systems are considered adaptive because it can react to external stimuli and responding to any situation that threatens its stability as a system. Thus, it experiences fluctuations. This has a limit, of course. It is said that the system settles into a state and when it is away from him tends to make every effort to return to the previous situation. This happens for example with the human body constantly strives to maintain the same body temperature.

4.3.Critical State

The general idea of system in a critical state can be understood as a state close to the boundary of another state (**critical point**). Meaning that any slight perturbation, can lead to a new state (**phase transition**).

One of the most famous example is the sandpile model by Bak–Tang–Wiesenfeld (Bak, Tang, & Wiesenfeld, 1987). The model describes how a sandpile is builds up as grains of sand randomly placed onto a pile. At the beginning small perturbations only cause small responses. Small avalanches take place until the pile reaches a critical state in which its slope fluctuates about a constant angle of repose (threshold or critical point). If we add one sand grain more, this can causes the slope exceeds the critical value and originates a big avalanche. The variation of the local slopes makes it impossible to predict when this phenomenon will take place.

Critical systems are featured by be in a delicately balanced state which, in turn, is linked to the environment, showing a great sensibility (Jost, 2005). This situation gives them a highly unpredictable behavior (chaotic, not random).

The most part of the complex systems are unstable (they are out of the equilibrium state). This implies that the systems cannot sustain themselves unless they receive a constant supply of energy (order needs chaos and chaos needs order. They cannot exist without each other, as mentioned earlier). They demand adjustments following specific patterns. Any minimum variation among composing elements can modify unpredictably, the interrelations and therefore, the behavior of entire system. Thus, the evolution of such systems is characterized by intermittency or fluctuation (situation in which the order and disorder constantly alternate). Their evolutionary states do not pass through continuous and gradual process, but occur through reorganizations and jumps. Each new state is only a transition, a **period of entropic**

rest in the words of Russian-Belgian Nobel Prize Ilya Prigogine (Prigogine & Stengers, 1984; Prigogine & Holte, 1993)**.**

These systems never reach a global optimum, the minimum energy state. In general, grow gradually until they reach the limit of its potential development. At that moment, they suffer a disorder, a kind of rupture that induces a fragmentation of pre-existing order. But then, begin to emerge regularities that organize the system in accordance with new laws, producing another kind of development. This behavior is typical in natural systems: for example, the transit of the insects, from egg to larva and from there to the chrysalis. Consequently, the organization of complex systems is given at different levels. The laws governing the causality of a given level can be totally different from a higher level (Kauffman, 1995; Bak, 1999).

4.4. Self-Organized Systems and Sport

Under these kinds of limit situations, in sport, athletes and their environment have to make a big effort in order to overcome the circumstances. In that moment is when they can really learn. It is at this time when sports systems create new strategies, training plans and, therefore, is when they evolve, change or behave according to the new reality. That is, the rivalry and competitiveness are the elements that generate the critical behavior.

When the system is self-organized critically, information flows better among all parts of the system (Solé, 2009). Moreover, these kinds of systems have memory and regulatory mechanisms that adjusts the response to demand. These systems evolve trying to optimize their resources and tend naturally to be in these states, therefore, serve as attractors of the system (Ivancevic & Ivancevic, 2006). I.e., the operation of system is the key and not the individual features of its elements.

All we know what is really interesting in sport is the competition. Competition attracts large masses of public, media and, frequently, large amounts of financial resources. Usually, this leads sports (especially in elite) to play in a critical area (García Manso & Martín González, 2008), in the edge of the error, risking, competing next to the limit.

This phenomenon promotes that sport evolves. Players change their game style, teams change tactics, game dynamic changes as well, new training methodologies emerge in order to

support competition requirements, etc. And even we can see how some sport introduce new rules (or modify old rules) in order to maintain competition attractiveness.
Some rules such as offside in rugby, 24 seconds shot clock in basketball, three touches in volleyball, a stolen base in baseball, etc. are attempts to lead sports to critical areas. Because sport adapt and the natural tendency leads to a hierarchical structuring more or less defined. Hence, the efforts of some sports in order to avoid these kind of situations.

As mentioned agents, who participate in these sport systems, compete among them; and the natural tendency leads to hierarchical structures, where some teams are clearly superior to others. Theoretically, this situation could be extended in time and hardly be broken by natural means, because best teams would continue hogging the best resources. This phenomenon is known as **Preferential Attachment** (Barabási & Albert, 1999), or **Snowball Effect** or **Saint Matthew Effect**. It is the popular the rich get richer and the poor get poorer.

So, theoretically, we can point out that this situation will continue as long as no external source modifies the environment in which the sport is developing (rules, competition sport model). That is why so important to figure out the operation of the sport system, meaning league, game team, etc. and how modifications (rules, new elements, etc.) affect the entire system.

The creation or modifications of these systems usually follow certain laws, meaning that some of these phenomena present the same features. One of the most important examples is the appearance of *Power Laws* or heavy-tailed distributions. This distribution is followed by many natural phenomena, often fractal, are also evident in many not natural systems. A lot of elements interact to produce a structure of higher level. These systems evolve far from equilibrium and are often highly dissipative (systems far from equilibrium). The *Power Laws* are described by mathematical expressions such as:

$$Y=cX^b$$

Where X and Y are two variables, or observable quantities, *c* is a constant and *b* is the scaling exponent. This kind of expression has two properties:

1) The logarithmic transformation becomes a line (see Figure 9):

$$log(Y) = log(c) + b\ log(X)$$

2) It is invariant to scale changes (scale-free).

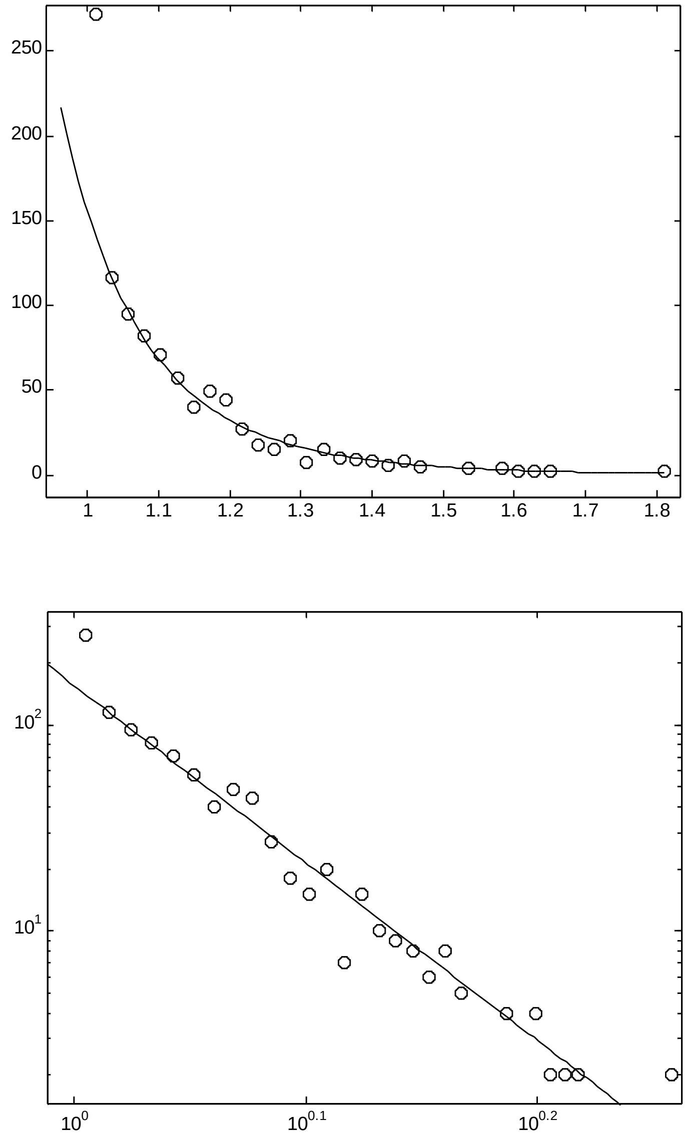


Figure 9. Example of a distribution (upper panel) and its log-log plot transformation (lower panel).

Phenomena with this type of behavior (*Power Laws*) are also called **scale-free**. By scale we mean the spatial and temporal dimension of a phenomenon. The hypothesis of scale that rises in the context of the study of critical phenomena led to two categories of predictions, both of which have been well verified by a large amount of experimental data on various systems. One of the most important is the scaling law we have mentioned; its usefulness lies in linking the various critical exponents that characterize the singular behavior of the order parameter and response functions (Amaral & Ottino, 2004).

Moreover, this kind of distribution can point out phenomena such as fractality (Barabási & Albert, 1999), self-organized criticality (Dhar, 1990; Bak, 1999), clustering (Newman, 2001a; Albert & Barabási, 2002), alometric laws (West, Brown, & Enquist, 1997a), etc. In short, they indicate the possible presence of complex systems.

Here we present several examples of diverse elements which follow *Power Laws*:

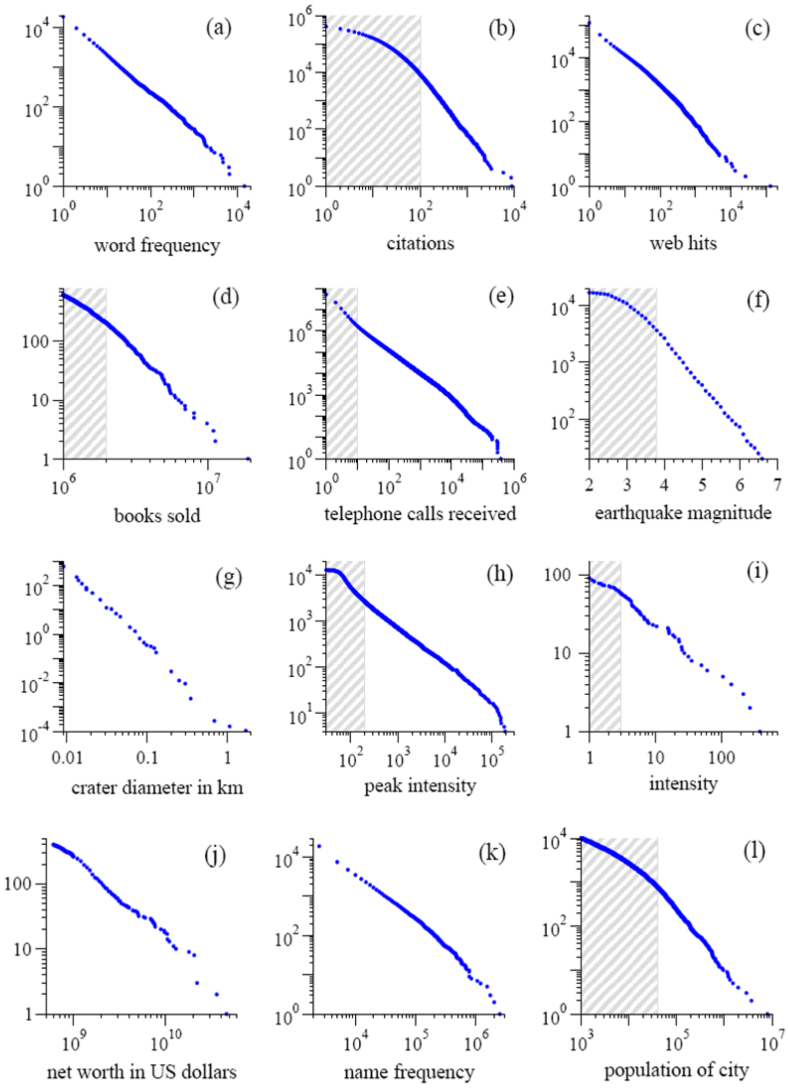


Figure 10. Examples of *Power Laws*: (a) Word frequency, (b) Citations of scientific papers, (c) Web hits, (d) Copies of books sold, (e) Telephone calls, (f) Magnitude of earthquakes, (g) Diameter of moon craters, (h) Intensity of solar flares, (i) Intensity of wars, (j) Wealth of richest Americans, (k) Frequencies of family names and (l) Populations of US cities. Source: **(Newman, 2005)**.

Also in sport there are a lot of examples of this kind of distributions: athletics records (Katz & Katz, 1999; Savaglio & Carbone, 2000), power lifting (García Manso et al., 2008), goals distributions (Malacarne & Mendes, 2000; Mendes et al., 2007), tenure lengths of sports managers (Aidt, Leong, Saslaw, & Sgroi, 2006), scoring in basketball (de Saá Guerra et al., 2013), etc.

As we can see, sport in general is a good example of complexity, therefore we believe accurate to use this methodology in order to analyze basketball.

General Research Design

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

5. General Research Design

5.1 Objectives

The objectives of this thesis are:

1. **To identify the organizational structures (leagues, federations, clubs etc.), participatory structures (players) which make up the SOC and the emergent behavior resulting from the interactions among them.**

2. **Figure out whether some of these patterns follow any known law and the meaning of these behaviors regarding basketball.**

3. **Trying to model institutional, group and / or collective behaviors that may occur in the organization and practice of basketball.**

5.2 Hypothesis

Our hypothesis for this dissertation is the following:

Basketball behaves as a System Self-Organized Critically during the regular phase of the league.

5.3 Research design

Introduction

For several years we have conducted our research in the Laboratory of Analysis for Planning and Athletic Training, which belongs to the Physical Education Department at the University of Las Palmas de Gran Canaria. In this laboratory our workgroup is involved in several research lines, such as research on physical activity and cognitive impairment (Alzheimer´s Disease), exercise in elderly, training for improving athletic performance, use of new technologies applied to sport, and in which I am involved: analysis of team sports from the perspective of non-linear complex systems. Following this research line, we have analyzed basketball from this standpoint, in order to a better understand of the reality of modern basketball and enrich, even further, this sport.

It would be interesting to know why a player increases his sports performance when he moves from league. Why basketball has succeeded in attracting media and fans compared to other sports. Or why the professional league in Spain, is one of the most competitive leagues in Europe. It would also be interesting to see how the play flows during a game. Figure out how score evolves, why the ACB is the second one that brings more players to international competitions and why this trend increasing, what makes a team better than another. Or why there were a significant increase in the flow of players to the NBA without precedent in the Spanish basketball. Or even to know why the success of the Spanish team not happened before with better rosters, and of course, the most interesting: find out how to win games.

Although sport is something seemingly simple, put the ball in a hoop, run faster than my adversaries, lift more weight than my opponents, effectiveness in sport seldom shows a linear behavior. Actually, there are many actions we can consider proper. Performance in basketball does not relate to a single factor, but a compendium of various elements that influence each other. The sport competitive system, the interactions of the players, the ball, the referees, and many other aspects, determine the final outcome.

Sports performance is the result of the combination of countless variables, sometimes known, sometimes not; and that through different analytical methods, we try to understand better in order to enhance them. Our intention with this essay is to expand our horizons of knowledge. To try to solve problems that, not even previously, we were unable to consider. The theory of complexity applied to sports is a great tool that can significantly enhance sport performance.

Study Design. Basketball as a complex system

Complex systems are the result of an evolutionary process. Darwin's ideas and the study of evolution have focused on the competition as a driving force of evolutionary change. Players, when cooperate, compete better as a team(Bar-Yam, 2001). In fact, cooperating or opposing of players (the attractors that form the system), is what gives rise to the different levels that can be found. The coexistence of these two behaviors is what allows basketball (and its elements) evolves.

Players compete for a place in the roster. This makes they improve. But the cooperation among several players is what enables to a team compete (Bar-Yam, 2001). Thereby we have two different levels of behavior. The same element can display two different properties, depending on the type of interaction. When players cooperate is a synergistic relationship. When players compete is an antagonist relationship. Competition occurs only when there is cooperation, and improvement only occurs when there is opposition.

As we mentioned before, a team is not only the result of the interaction of the players. It needs an environment where it can develop and evolve to new states. If it behaves as we assume, as a System Self-Organized Critically, the instabilities and leaps to new forms are the result of internal fluctuations and the interaction with the environment.

Here appears, for example, the figure of the coach as an environmental factor which influences the final outcome. If we keep the same players but we change the coach, we can obtain a completely different behavior in the team, with outcome radically different from the initial.

Thus, whether we change the environment, we modify the interactions among the players; therefore, we alter the final result. Other clear example are the rules. Rules provide an artificial environment and an artificial system of information. They are artificial spatiotemporal boundaries, so any variation influence highly in the dynamic competition. That is why we considered interesting to study basketball from different levels.

Several attempts have been made to understand some modalities of sport from the point of view of complexity. Most of these works deal with this phenomenon with regard to the mechanics of the game itself (in isolation) (Chatterjee & Yilmaz, 1999; Malacarne & Mendes, 2000; McGarry et al., 2002; Lebed, 2006; Kubatko et al., 2007; McGarry & Franks, 2007; Mendes et al., 2007; Bittner et al., 2009; Heuer et al., 2010), nevertheless, less researches has been conducted from the perspective of competition between teams (Yilmaz & Chatterjee, 2000; Onody & de Castro, 2004; Vaz de Melo et al., 2008; Ribeiro et al., 2010; de Saá Guerra et al., 2012).

Basketball Study

Basketball performance does not deal with only one factor, but it is the result of several elements which influence each other. It would be interesting to know why a player increases his sport performance when he goes to a different league, or even when a team hires a new head coach. We tried to observe the behavior of the macrostructure and its performance. We wanted to check out whether the elements we proposed have any correlation among them and the law that governs them. Therefore, in order to clarify these questions and our hypothesis, we divided our investigation in three levels or studies:

- Study 1. Basketball Competition (League)
- Study 2. Basketball game
- Study 3. Basketball team

League

The aim of this part of the study was to assess the competitiveness degree based on the uncertainty level that might exist in each confrontation. We calculated the value of the Shannon entropy, which quantifies the information contained by a variable, in order to determine the degree of uncertainty or randomness of the competition. So we used the concept of competitiveness as a relative indicator of quality.

This analysis provided us a value of the competitiveness level and also allowed us to identify possible causes of the increase or decrease in quality during several seasons (regular phase. Not play-off), and we compared several leagues of high quality among them, such as the NBA (USA) or the ACB (Spain). This analysis also revealed how it is structured the network of basketball, identifying the nodes and how they interact with each other (economic structure, competitive organization, player's source, influence of other networks, etc.).

Game

Scoring in a basketball game is a process highly dynamic and type non-linear. The level of teams improves each season because they incorporate to their rosters the best players they can afford. These and other mechanisms make scoring, in the basketball games be something exciting and hardly predictable.

We studied the behavior of the game as the interaction between two teams in the same environment in isolation. We examined all the games from seasons analyzed always in regular season (home games and away games). We investigated scoring in the games regarding the differences in points. And in turn, we studied the evolution of the time intervals between points.

Team

We want to know the key to success of a team. Our hypothesis is the self-organization. The way which teams accomplish their aims and overcome to external threats. This is achieved through local interactions to make up the disequilibrium, ergo the entire system act properly (all the elements work coordinated and smoothly). We will apply the theory of complex networks. It cannot be understand the team as the mere sum of its players, but as something else. We try to establish a quality rating of a team, to determine the degree of self-organization of a team. The property which we believe is the real key to success.

Team as a SOC presents two distinct behaviors: attack and defense. The goal of an attacking team is to achieve points or cause the opposing team commits a violation of the rules, i.e., a mistake that benefits the attacking team (personal foul). On the other hand, when we deal with the defensive behavior, we can say that its main objective is to try to contain the opposing team. While the offensive team considers the basket or failure as a positive achievement or goal, the defense team seeks to divert the energy of the opponent so that they make a mistake or neutralize their attempts to reach their goals during the game time and/or possession.

5.4 Significance

The important of this work is the novelty of the application of complex systems as a methodological tool to the sport, and above all, the results and conclusions that can be obtained with this type of research techniques.

A second point to note is the extent and depth of study, inasmuch as we want to find out how basketball works and how it behaves at several levels (league-game-team), describing the elements or nodes which form it (Sports Network). Moreover, this allows us to understand the connections to other networks, sports, and their linkages.

Finally we want emphasize the applicability of the study in several fields. One of them could be developing of intervention strategies for creating of competition patterns, management and teaching for players, coaches and sports organizations staff. And even for the staff of entities outside the sport, but with links to the sports network. In the academic field, it is important because we can learn and apply new techniques to other fields of knowledge. To open new research lines that can be follow by other researches, and pass on those discoveries with the methodologies to the university and scientific community.

Following, we present (Figure 11) the general research design, in order to a better understanding of the document:

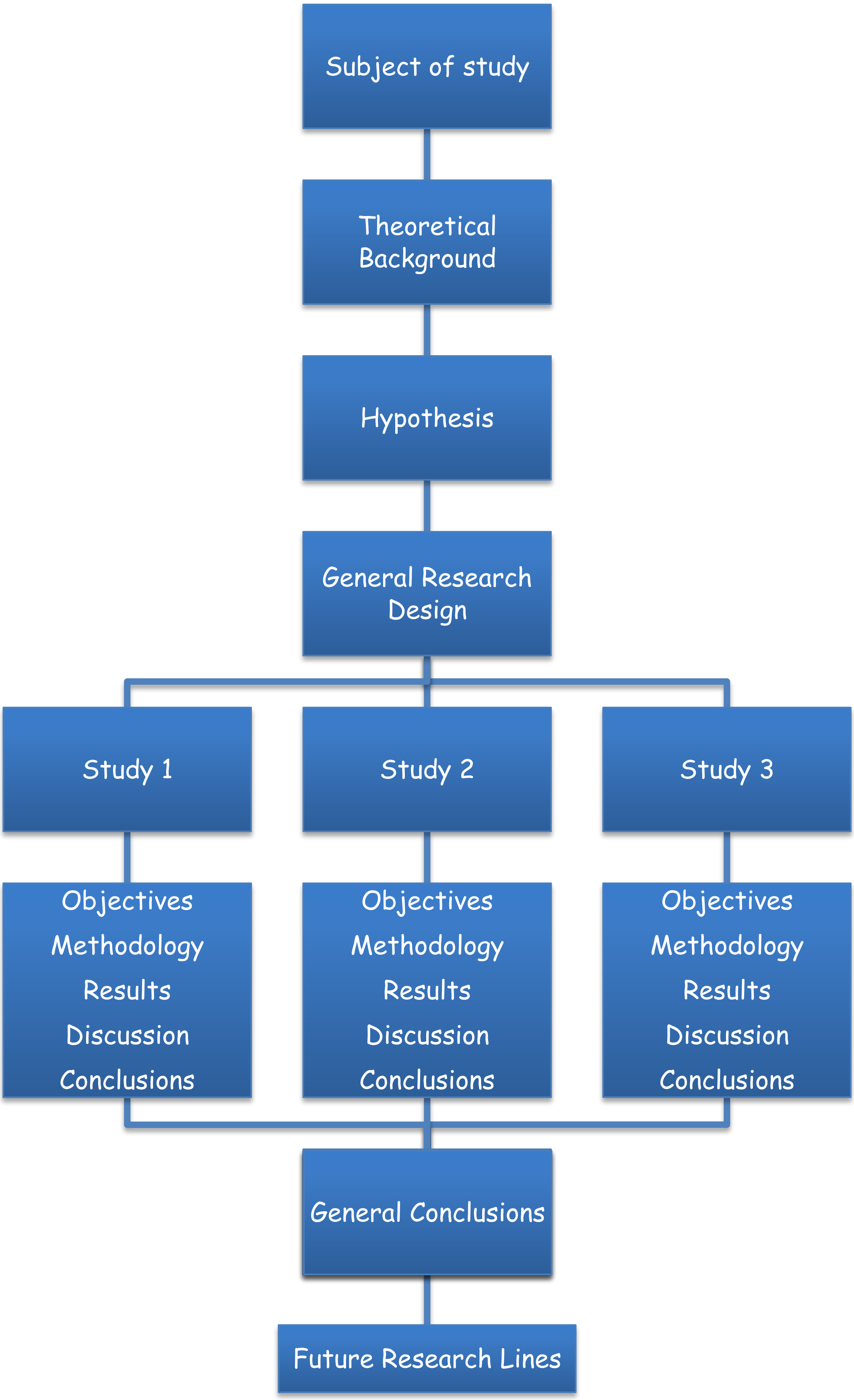


Figure 11: General Design of Research.

Study 1. Basketball league

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

6. Study 1. Basketball league

6.1 Intro

A league is a competitive model. In a sport, the design of the competitive model is as important as the preparation of the subjects involved (players, coaches, officials, etc.).

Team performance can be postulated as winning as many games as possible. Final standing is the result of the way in which all teams interact of in a pre-schedule calendar. That is the competition model. Therefore, the effectiveness of a team is closely conditioned by is the result of certain states of optimization of systems that compose it, which also have a reciprocal relationship with emerging and critical environment: the competition.

The study of the equilibrium between systems which interact within a same environment, is a frequent issue by disciplines such as economy, applied statistic, physics, evolutive biology, social sciences, sport sciences, etc. Given the difficulty of predicting results of the games, and therefore the final standing, we cannot use a linear methodology for analysis, it is necessary to use a methodology that allows us to explore the nature of the competition with as much detail as possible, as is the theory of complexity.

The point is to understand the sport from a systemic conception and conceive the athlete, or team in our case, as a system that works as a whole that is affected by the surrounding environment (Gambetta, 1989; García Manso & Martín González, 2008).

The competitive model has a direct influence on the competition (type of confrontation), development and evaluation so that small changes can dramatically alter the final result, given the close relationship between the competitive model and competition (de Saá Guerra et al., 2012; Lebed, 2006).

The team ability to compete and how the championship works (how competition format is designed: conferences, divisions, game schedule, league, playoff, etc.) determine the level of competitiveness. Competitiveness is a comparative concept of the ability to strive for a goal. The more balanced competition, the greater the degree of competitiveness, and vice versa. This is an interesting because it reflects the reality of the competitive system, e.g. higher budgets allow signing players of better quality. Whereas tighter budgets do not allow hire best players, given their high cost. One of the most widespread ideas to explain the phenomenon of equality among the competitors of the same championship is the concept of **competitive balance**.

Competitive balance represents the degree of equality within a championship. A central concept used in the economic analysis of professional sports leagues, as indicated by the extensive literature on the subject (Schmidt & Berri, 2001; Fort & Maxcy, 2003; T. A. Rhoads, 2004; Goossens, 2006). The idea of competitive balance is to try to measure the degree of global competitiveness in a given league. And specifically has been applied in disciplines such as baseball (Owen et al., 2007; Scully, 1989), American football (Bennett & Fizel, 1995), basketball (Noll, 1988; Berri et al., 2005), ice hockey (Richardson, 2000), football (soccer) (Halicioglu, 2006) or golf (T. Rhoads, 2005). Thus, greater competitive balance should lead to greater demand (Quirk & Fort, 1997; Goossens, 2006). Indeed, the most competitive leagues tend to be more attractive and generate more revenue (tickets, sponsors, TV, etc.)(Szymanski, 2003) and this is closely related with the sport model (Ribeiro et al., 2010).

The key element of the economic success in the professional sport is the increase in competitive balance. Each time a competitor reaches a very high domain level, the competition equilibrium (the way in which teams compete) is broken down, in the sense that the uncertainty declines significantly. In these situations, when the uncertainty of the outcome diminishes, the interest of the competition may reduce considerably. When this happens, the attendance of spectators to the games can decrease and, consequently, access to financial resources may be compromised (Berri et al., 2005). For this reason, sport organizations which design sport competitions models (leagues), try to design structures and rules which enable cope with a decrease of competitiveness in a championship. A certain level of competitive balance seems reasonable to hold the interest of spectators and sponsors for all teams but the determination of the optimal level is very complex.

Some authors (Knowles et al., 1992; Rascher, 1999)noted that fan attendance in Major League Baseball is maximized when the probability of the home team winning is approximately 0,6. If the home team has a higher probability of finding success, we can expect fan attendance to decline. Consequently, given the importance of fan attendance to a league's financial success, leagues are expected to implement rules and institutions designed to address the relative strength of teams on the games.

We studied the results from different seasons of two of the main professional basketball leagues, the NBA (National Basketball Association, USA) and the ACB (Basketball Clubs Association, Spain)and the results of one high level amateur league, the Division I of NCAA Men´s basketball (National Collegiate Athletic Association, USA). Data have been obtained from the official NBA, ACB and NCAA webpages (www.nba.com; www.acb.com and www.ncaa.com).

The Spanish Professional Basketball League (ACB) is an open model league, where every season participating teams are readjusted taken into accounts promotions and demotions to lower categories. The eight top ranked teams play the play-off in order to be proclaimed champion of the league.

The professional North American League (NBA) is a franchise model competition. Participating teams are divided in two conferences (Eastern and Western). In turn, these are divided in three divisions per conference. When the regular season finishes, top ranked teams will meet in the play-off for the title. The NBA is a closed model where there are neither promotions nor demotions.

The NCAA (basketball college championship in USA) is divided in three divisions (Division I, Division II and Division III). In turn, every division is divided in conferences of several teams each. We only used the data from the Division I of the men´s basketball. We must remember that the Division I of NCAA men´s basketball is composed by a total of 344 teams (the number varies in the season analyzed), divided in 31 conferences through all USA (the number of teams per conference is not homogenous).

The aim of this study was to analyze, from an overview, the sport model and the intern dynamic of several basketball leagues (professionals and amateur) by studying its competitiveness degree. Also we tried to develop a model for the competitiveness level analysis in team sport competitions, which would be useful to assess their competitiveness level based on the uncertainty level that might exist for each confrontation.

6.2. Methodology

Confrontation matrices

Our interest is to focus on studying sport leagues, where each team usually plays twice against each other team (once at home, once away) in games according to a prearranged schedule.

A series of games between a number N of teams, can be defined by its matrix of confrontation $\mathbf{A} = [A_{i,j}]_{N\times N}$, with the same number of rows and columns. This is a double entrance matrix where each row and each column correspond to the results of each game between any two teams. We shall use the subscript i or j for teams, $i \neq j$, so we use $Aij = 1$ if team i beats team j, and $Aij = 0$ otherwise. Other options such as ties or different values of 0 or 1 are not considered in this introduction at the moment without loss of generality. From this matrix, at the end of the competition, we obtain the final score **R**. See Table1 for an example of the matrix $N = 4$.

	a	b	c	d	HW	R
a	*x*	*1*	*0*	*1*	*2*	*4*
b	*0*	*x*	*0*	*0*	*0*	*1*
c	*1*	*1*	*x*	*1*	*3*	*5*
d	*0*	*0*	*1*	*x*	*1*	*2*
AL	*1*	*2*	*1*	*2*		
AW=(N-1) - AL	*2*	*1*	*2*	*1*		

Table 3. Example of a confrontation matrix with $N=4$ teams (**a**,**b**,**c** and **d**). The rows represent the *games* played (won or lost) by a team at home. The columns represent the won or lost games played by a team away. **HW** (Home Wins) represents the total number of games won by the teams at home. **AL** (Away Lost) is the lost games away. **AW** represents the total number of games won away. The final score **R** is the sum of the home and away wins, **R**=**HW**+**AW**.

The row i of matrix **A** represents the points for games won or lost by the team i at home, while column j represents the away games won or lost by the same. Therefore the horizontal sum:

$$\sum_{j=1}^{N} A(i,j) = n_i$$

represents the number of games won by the team *i* at home (n_i), where *N* is the total number of teams. Note that *A(i,j)=0*, if *i = j*. Likewise, the vertical sum

$$\sum_{j=1}^{N} A(j,i) = m_i$$

represents the number of away games lost by *i* (n_j). Therefore, the total number of games won by the team *i* will be $R_i = n_i + (N-1) - m_i$

The vector **R** (score vector) represents the results obtained by each team in each season. The result vector **R** behaves randomly, in the sense that we do not know the final result, but the results of previous seasons (historical performance), may provide some clues. The values of **R** historical or previous seasons divided by the sum of all games can be considered to be a discrete probability distribution

$$p_i = \frac{R_i}{\sum_{j=1}^{N} R_j}$$

where p_i indicates the probability that the *i* team gets a certain result and therefore can be considered as a performance indicator.

If the distribution is uniform, all p_i values are equal or similar to each other, and all the teams have approximately the same playing level. This represents a case where it is difficult to predict the final outcome. This may be considered to be highest possible parity among the teams (competitive balance). However, if there are certain values of p_i greater than the rest, it means that there are some teams in the competition with superior performance to other teams.

In the case of a uniform distribution, any team has an equal chance of winning. In terms of statistical mechanics, such distributions are related to equilibrium situations where all structures and gradients have been eliminated. The disorder is maximum; therefore the values of entropy (*S*) are also maximum. Following this analogy, if the system is isolated, cannot exchange matter, energy or information with its environment, all random fluctuations that may occur and thus, all gradients that can be formed, tend to be neglected.

However, when the system is not isolated and interacts with the surrounding environment by exchanging matter or energy, it is possible that some fluctuations in prosperity generates gradients which are more or less permanent, keeping the system away from equilibrium. The local emergence of order (symmetry breaking) is only possible in open systems interacting with their environment (Mainzer, 2005)

When the set of probabilities of a system is known, we can define the Shannon entropy (*S*), which is an average measure of uncertainty and, hence, refers to the average amount of information that is contained in a random variable. It is defined as:

$$S = \sum_{i=1}^{N} \left(p_i \log \frac{1}{p_i} \right)$$

The value of *S* changes with the value of *N*, and if *p* is the probability distribution obtained from a given result matrix **A** for *N* teams, we could not compare different seasons if the number of teams changes. Hence, it is preferable to use the normalized entropy (*Sn*):

$$Sn = \frac{S}{\log(N)}$$

So, the value of *Sn* is bounded between *0* and *1*, where *1* corresponds to the situation in which all values *pi* are equal to each other.

If we define the equilibrium state as the situation of maximum competitiveness, *Sn* provides a numerical value of competitiveness for a given season. From this point of view, if a competition is less random, the degree of competitiveness is lower, which means that we have a competition with less uncertainty about the final result. In this way, we can express the result table as a set of probabilities, and we can calculate the value of *Sn* to obtain a parameter that will measure the extent of the system's departure from equilibrium.

Extreme values of the results matrix

The results matrix **A** range between two extreme values of *Sn.* One is the uniform distribution case, previously mentioned and the other value is unknown. To obtain these results, we use a simple algorithm that permits the acquisition of those settings established in the matrix where the values of *Sn* are maximum or minimum.

1. We start with a results matrix **A** = $[A_{i,j}]_{N \times N}$, for which we calculate the initial value of Shannon entropy Sn_0.
2. We choose a value *A(i, j)* at random, on condition that $i \neq j$, and exchange its value, so that if *A (i, j) = 1*, we substitute for *0*, and vice versa.
3. We calculate the probability distribution and the value of *Sn* for the new situation.
4. If $Sn > Sn_0$, the change is accepted. On the contrary, the matrix **A** remains unchanged and we come back to step 2.

The algorithm will stop when the value of *Sn* reaches its maximum value, in this case *1*. We can use the same algorithm by changing $Sn < Sn_0$ in step 4. In this case, the algorithm will finish when the value of *Sn* reaches a minimum. This algorithm provides two extreme solutions for a matrix of this kind.

Solution 1. In the first case (*Sn* maximum), *Sn* peaked the **theoretical maximum** reflecting a competitive solution in which all teams have an equal chance of winning. We define it as **random competition (RC)**. For this solution there are two particular cases, **A1** and **A2**, corresponding to theoretical situations in which all teams win home games and lose away games, or vice versa**.**

$$\mathbf{A1}=\begin{bmatrix} 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix};\quad \mathbf{R1}=\begin{bmatrix} N-1 \\ N-1 \\ N-1 \\ \dots \end{bmatrix}$$

$$\mathbf{A2}=\begin{bmatrix} 0 & 1 & 1 & \dots \\ 1 & 0 & 1 & \dots \\ 1 & 1 & 0 & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix};\quad \mathbf{R2}=\begin{bmatrix} N-1 \\ N-1 \\ N-1 \\ \dots \end{bmatrix}$$

However, this result would also be valid in the case that the teams were all teams have an equal chance of losing. Without loss of generality we do not consider this solution.

Solution 2. The next matrix is the result of the other process (*Sn* minimum), whose lines have been arranged according to the values of **R**.

$$\mathbf{Am}=\begin{bmatrix} 0 & 1 & 1 & 1... \\ 0 & 0 & 1 & 1... \\ 0 & 0 & 0 & 1... \\ 0 & 0 & 0 & 0... \end{bmatrix}; \quad \mathbf{Rm}=\begin{bmatrix} 2N-2 \\ 2N-4 \\ 2N-6 \\ ... \end{bmatrix}$$

It corresponds to the case that the best team wins every game, the next team wins all games except the two that were lost to the first one, and so on. The resulting **Rm** values, sorted from highest to lowest, are in this case: *2N-2, 2N-4, 2N-6, ..., 0*. This theoretical solution, which we call **hierarchical competition (HC)**, is more predictable and provides the minimum value of *Sn*, for any table of results of this type.

Any vector results **R** obtained at the end of a season in a competition, **ordered decreasingly**, remain somewhere between these two theoretical models. It should be kept in mind that these are extreme theoretical values (solutions 1 and 2). In some sport competitions, such as the NBA, teams play more than two games against the same team, therefore, there are values on matrix greater than 1. In that case, the method is sufficient to find the solutions. In another case, as in football with three possible results by: lose, tie or win, with punctuation 0, 1, 3, the method also works.

Random results

It is nearly impossible, or at least, very improbable for a competition to end with one of these two results (extreme hierarchical result or extreme random result). To be more realistic, we consider that the outcome of each game between two teams is random, but follows one of the two distributions *p* (random or hierarchical). Thus, each element of the matrix A(i,j) takes a value 1 or 0 according to the probability *pij = pi / (pi + pj*). For this, we selected a random number *r*, ranging between 0 and 1, from a uniform distribution. If *r < pij*, we chose the value 1 and 0 otherwise.

By this process, we generated 5000 random matrices of size NxN, around each of the cases (random or hierarchical) and we calculated the probability density values of *Sn* in each case (Figure 12).

Application to basketball

We studied the ACB (Spanish professional basketball league), the NBA (American professional basketball league) and the NCAA (the American basketball college league) as amateur championship. We analysed 14 ACB regular seasons (1996-97 to 2009-10), 18 seasons in the NBA league (1992-93 to 2009-10) and 10 (2001-2011) seasons of the NCAA.

The ACB is a professional league with a season of double confrontations between the 18 teams (306 possible victories). It has an open structure in which there are promotions and relegations, and where the best classified teams in the regular season must compete for the championship in a playoff format. For the analysis, we used the league final standings table from every season.

The NBA is a league of closed structure (neither promotions nor relegations), comprised of thirty franchised member clubs, of which twenty-nine are located in the United States and one in Canada. The current league organization divides thirty teams into two conferences of three divisions with five teams each. The current divisional alignment was introduced in the 2004–05 season. During the regular season, each team plays 82 games, 41 at home and 41 away. A team faces its opponents in its own division four times a year (16 games); teams from the other two divisions in its own conference either three or four times (36 games); and teams in the other conference twice respectively (30 games). In this case, the confrontation matrix can contain different values of zero and one, which has been taken into account when we calculated the extreme value theoretical and random of *Sn.*

The NCAA is a league of closed structure as well (neither promotions nor relegations). Universities are classified into a Division depending on the number of scholarships that can be offered to athletes in each group. All the sport programs of a university have to be included in the same Division. As mentioned, the basketball college championship in USA is divided in three divisions (Division I, Division II and Division III). Currently, the Division I is divided 31 conferences through all USA (the number of teams per conference is not homogenous) and participate a total of 344 teams (the number varies within the season analyzed). There is a regular phase (regular league), and a playoff. In the regular phase teams play against the teams of the same conference. In addition, they play extra games (tournaments) in order to get more points for the playoff classification. After the regular phase, the best teams classified play a playoff (only one game per round) for the national championship.

NOTE: when we carried out the research, the three point line in the NCAA was located in 6.02 meters, and now, with the current rules, was moved to 6.25 meters.

6.3. Results and discussion

6.3.1. Entropy analysis

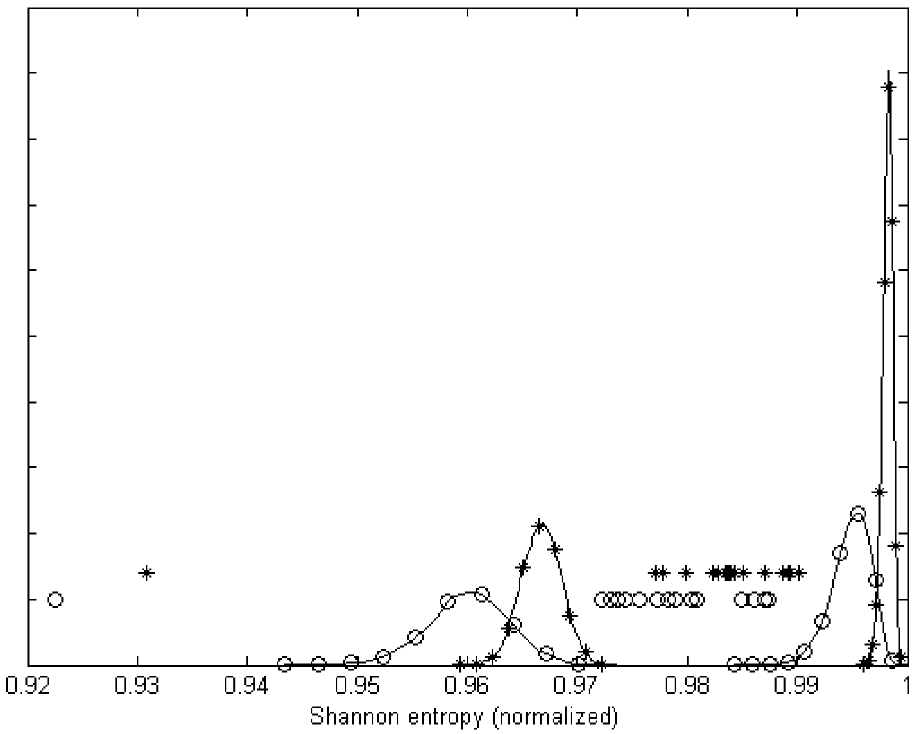


Figure 12. The X axis represents normalized entropy values. The continuous lines represent the probability densities of *Sn* values obtained by generating 5000 random results around the two extreme theoretical distributions for the matrices representing the two leagues: ACB (o) and NBA (*). The y-axis values would be the appropriate ones for the area under each curve is unity. We do not present because this axis is only of interest to note that the distributions retain their relative sizes. The peaks of the curves determine the values of location (usually near the mean) on standardized values of entropy (x-axis). The points between the distributions indicate the values of entropy obtained for both leagues, in the seasons studied (in normalized entropy values). The two isolated values on the left side represent the theoretical extreme values in the hierarchical case. The maximum theoretical value for the random case is 1. We have discarded the results of NCAA in this analysis because the results can be very confusing due to the large number of participating teams.

In Figure 12, the x-axis represents the values of entropy. It shows three different elements. First, the values between the distributions lines are the actual *Sn* values (in arbitrary units) for the ACB (o) and NBA (∗) of obtained for all seasons analysed. Second the distribution lines, obtained from the generation of random matrices for ACB and NBA cases, represent the *Sn* probability densities. The theoretical maximum value for both leagues is 1 (RC), whereas the two values on the left, far from most of the other values in the set of data (''o'' for ACB and ''∗'' for NBA), represent the theoretical minimum in the two cases (HC). We have discarded the results of the NCAA because the results are meaningless due to the league has many teams.

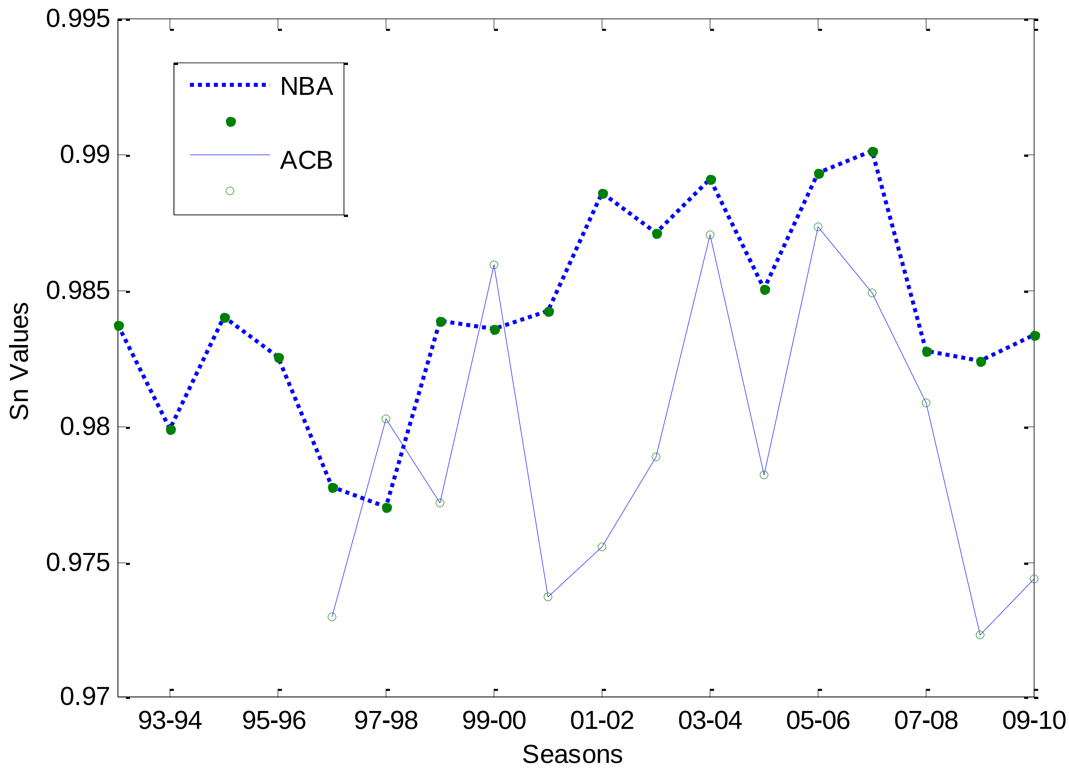


Figure 13. Entropy values of the 18 NBA seasons analyzed (dashed line), and the 14 seasons of the ACB (solid line). The entropy of the NBA is more stable and close to its average values, showing two different periods for this trend. The first one is from 1995–1996 season to 1998–1999 season. The second one is from the 2000–2001 to 2006–2007 season. In the ACB, the entropy values vary further from season to season, presenting both highly competitive years and other years that are less competitive. This behavior is remarkable, especially because it usually occurs cyclically.

Figure 13 shows the values of *Sn* obtained for all seasons analyzed (ACB, solid line; NBA, dashed line). Note that the NBA entropy remains more stable than the ACB entropy (*Sn* NBA mean=0.9842 ± 0.0037; *Sn* ACB mean = 0.9793 ± 0.0053). Nevertheless, we can distinguish two periods well differenced: the first one is from 1995–1996 season to 1998–1999 season, and the second one is from 2000–2001 to 2006–2007 season. In the ACB, the entropy values vary further from season to season. Their values oscillate between two league profiles (seasons that were very competitive and seasons that were less competitive).

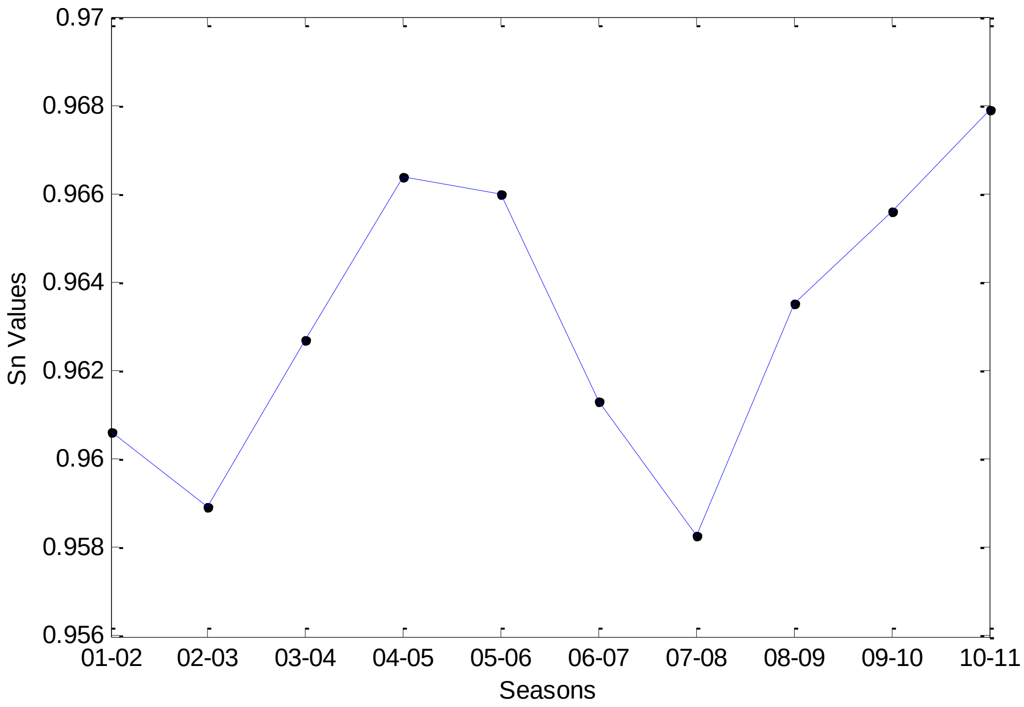


Figure 14. Entropy values of the 10 season of NCAA men´s basketball Division I. Data display a no uniform tendency. Sn values vary throughout the seasons analyzed, showing an upward trend in the last 4 seasons. Even they reach higher values than the previous maximum.

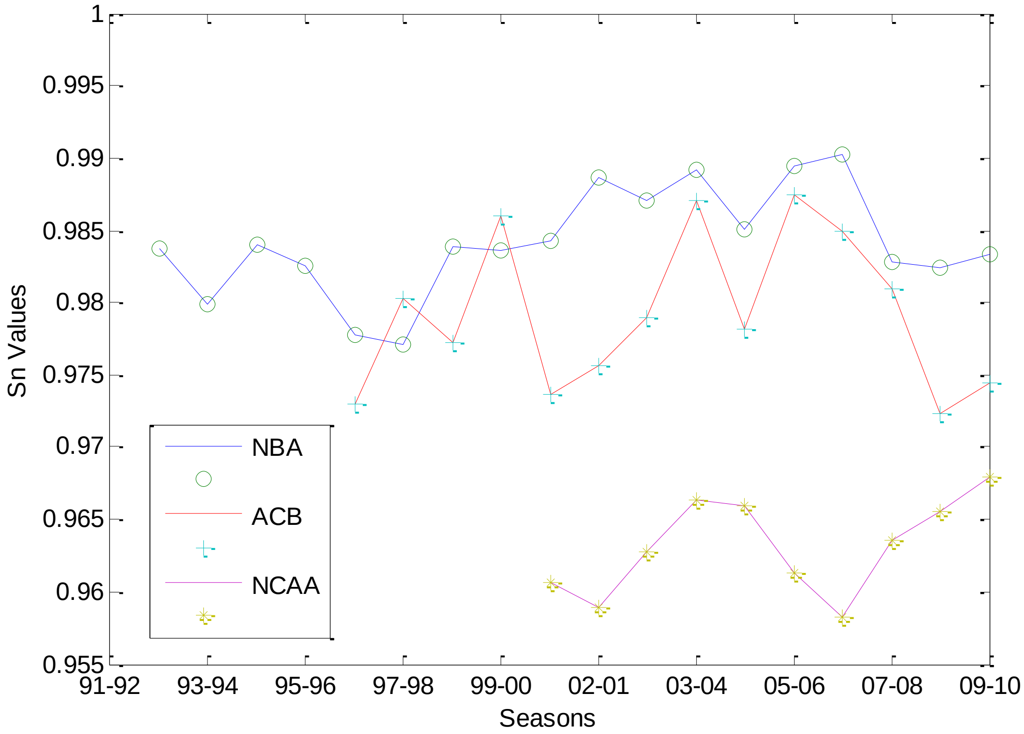


Figure 15. Comparison of the three leagues analyzed. We can note that the professional leagues (the ACB and the NBA) have a higher level of competitiveness than the amateur league (NCAA). But the in the last season, the values are closer than ever. This fact is very relevant and will be interesting to find out the reason for this behavior.

In the analysis of the entropy values of NCAA men´s basketball Division I (Figure 14) we can appreciate that the tendency is not homogeneous. There was a period (from 2004-2005 to 2005-2006) where the competitiveness was maximum (0,967) with slopes very similar to both sides. But the most notable fact is that the last 4 seasons the *Sn* values increase up the highest value of entropy (0,968).

When we compare the three leagues (Figure 15), we can see that the professional leagues present higher values of entropy, what means that, in general, the competitiveness is greater than the amateur league. Neither of them display a homogeneous comportment but all of them oscillate during the years analyzed. Only the NBA presents a most stable behavior. As we mentioned above, these asymmetries may be originated by readjustments in the competitive system, enlargements, promotions and demotions, etc. But possibly, the only fact that seems to coincide in all the leagues is the economic crisis. Nevertheless, it gives the impression to have different effects in the leagues analyzed. In the professional leagues (the ACB and the NBA) the economic crisis lead to a decrease (and to a stabilization at the last seasons) on the general competitiveness level, while on the other hand, the NCAA experiment an increase in the entropy values during the same period.

Indeed, we can observe that in the season 2006-2007, the entropy of the NBA starts to decline while the *Sn* values of the NCAA growth. And the ACB suffers a drop drastically during these same years. It would be interesting to find out the reason for these variations so marked, but it is necessary further deepen into the causes that originate this class of phenomena.

ACB analysis

The temporal evolution of these values can be seen in Figure 13 and 15, and, as the data show, there are some seasons in which the confrontations have a high degree of uncertainty (1999–2000, 2003–2004 and 2005–2006 seasons) and other seasons with a completely different profile (seasons 1996–1997; 2000–2001 and 2008–2009). Issues such as the possibility of relegation, the major budgetary differences of teams, high economic dependence on public institutions, or the high volatility of the rosters are some of the factors that can influence this behavior. This market has become increasingly active, as basketball has been professionalized, so much so that a few players remain more than five seasons with the same team in the ACB (Arjonilla López, 2011)

A team's performance is mainly determined by two factors: economics and the players themselves. Both are closely linked. A large budget provides the ability to sign superior players and to make rosters balanced to a greater extent. Teams with smaller budgets select players with a supposedly lower level; consequently, their rosters will be less balanced and less competitive.

The teams make their rosters based on budget and sporting objectives. These objectives are closely linked to the competitive ACB model (open model). Moreover, the absence of a salary cap and a conspicuous disparity between budgets of teams can lead to large differences in the quality of the rosters (sporting potential gradient). This makes differences in performance insurmountable for some teams in the ACB, especially for promoted teams, whose budgets and rosters are tight. The unstable private and state economic support of some teams and not others, which have a solid economic support and the marketing (trading) of players, may in part explain these oscillations in the entropy, which are characteristic of the ACB. Curiously, the years with minimum values of *Sn* coincide with Olympic years. This may be a topic for future study.

NBA analysis

The NBA has a more stable shape than the ACB. There are seasons with lower competitiveness (≤0,980) than the overall average (0,983), and periods where the competitiveness remains at higher levels (range: 0,985–0,990) (Figure 13 and 15).

In Figure 13 and 15, we can clearly see that there are seasons that do not correspond to the general trend, showing a behavior that is perhaps anomalous. These seasons are 1993–1994 and 1996–1998, with the lower values of *Sn*, and the period from 2001–2002 to 2006–2007 seasons, with higher levels. In the 1995–1998 seasons, Michael Jordan's Chicago Bulls achieved the best win record in the NBA regular season to date (won–lost: 72–10, 69–13, and 62–20 respectively). This performance is probably responsible for the decline in competitiveness in the league during this time. Indeed, some authors mention that the 1990s were the least competitive decade in the history of the NBA(Berri et al., 2005). This seems far from frivolous, as this sort of phenomenon can be accompanied by a heightened attraction to athletes and the general public (Rovell, 2003), as well as private companies and media, resulting in a strong economic impact (Forbes, 2008; Mathur, Mathur, & Rangan, 1997).

Another interesting area covers the period from the 2001–2002 to 2006–2007 seasons. These correspond with the renegotiation of the salary cap (1999–2005) when the Collective Bargaining Agreement (CBA) was signed. This probably had an impact on the overall performance of the league, because the objective of the salary cap was to prevent teams with a large profit surplus from signing the best players available, thereby facilitating the equality of retention in the league. This mechanism, together with the draft, is necessary for each franchise to carefully select which players may be interested in the market for its particular franchise goal (depending on the background of the team). Consequently, each franchise is only able to "shield" economically one or two players, commonly referred to as "franchise players".

The decline of competitiveness in the 2004–2005 season (see Figure 13 and 15) may be attributed to seasonal variability, although it could also be caused by enlargement of the participating teams from 29 to 30, which further restructured the divisions in each conference.

This changed the format of the divisions: instead of having two divisions per conference, there would be three per conference with five teams each. The 2005–2006 season increased levels of competition again, as shown in Figure 13 and 15, possibly because all the Central Division teams qualified for the playoff s and this was the first time that a division managed to place all of its teams in the postseason since the Midwest Division did so 20 years ago.

NCAA analysis

Regarding to the Figures 14 and 15, we can see that Data display a no uniform tendency. *Sn* values vary throughout the seasons analyzed, showing an upward trend in the last 4 seasons. Even they reach higher values than the previous maximum.

The *Sn* value range for all time period analyzed. We can see a period where they reached values significantly elevated and a final period where values reach their maximum.

In the NCAA participates a great number of teams with different performance levels, hence when we analyze the entropy data we have to take into account the heterogeneity of the sample for the analysis of the results. The values of *Sn* range from 0.9679 to 0.9583. These values are quite distant from the values of the professional leagues but even that, the NCAA is the most stable of the three (*Sn* NCAA mean=0.9631 ± 0.0033).

The main point we must bear in mind when we analyze the NCAA is that the teams that qualify for the playoffs using an index called Rating Percentage Index (RPI). The RPI is a quantity used to rank sports teams based upon a team's wins and losses and its strength of schedule. The current used formula for determining the RPI of a college basketball team is as follows.

$$RPI = (WP * 0,25) + (OWP * 0,50) + (OOWP * 0,25)$$

where WP is Winning Percentage, OWP is Opponents´ Winning Percentage and OOWP is Opponents´ Opponents´ Winning Percentage. The WP is calculated by taking a team´s wins divided by the number of games it has played (i.e. wins plus losses).

So NCAA standings are elaborated using the RPI instead of games won, as the rest of leagues analyzed. This index tends to equilibrate the differences among different teams, so that theoretically favors the weaker teams and hampers the teams with better historical trajectory.

This cause that a lot of the teams play extra tournaments in order to improve its index, hence the teams of NCAA do not play the same number of games during the regular phase.

Also, the NCAA has the feature that a single player only can remain in the same university for four years as maximum. There is not a players market as at the professional competitions. This fact conditions the possible advantage that the “best” teams can obtain over the rest of the participants.

Theoretically, the best teams recruit the best players in order to win more games and qualify for the playoff. If this were true, the differences would be insurmountable for the rest of the teams. But in reality, the opposite happens.

The current tendency of the players that come from the high schools is to choose teams where they are going to play a lot of minutes and be the “star” of the team, instead of choose a good basketball program in a good university or college in order to receive a good sport and academic formation.

Coaches mention that a very large number of players conceive the college team as a step in their career to the NBA, what is a big mistake, in the words of the coaches.

As an additional consideration, we observe that in both professional leagues there is a significant decay in *Sn* in the last three seasons to a lower limit, which apparently remains attached to that value. It is possible that this is related to the current economic crisis in which some teams are less affected than others. This is less important in the NBA, as compared to the ACB, indicating the sensitivity of the two different sports models to external economic effects. We must take into account that the NBA is a private league, and that it works as a company, while the ACB depends largely on regional government subsidies, which particularly affects some teams.

Comparison among teams or conferences

We used this protocol (normalized Shannon entropy) in order to determinate the competitive balance in different ACB, NBA and NCAA seasons. This proposal makes a rude analysis of competitive balance without regard to the ranking in which teams complete this phase of the league (regular season), and discriminates well between leagues. For this reason, we compare the values of the win ratio of (**R**) with two extreme theoretical models (random of maximum competitiveness and hierarchical of minimal competitiveness) in order to analyze the actual behavior of competitive balance in the two professional basketball leagues. We have discarded the results of the NCAA because the results are meaningless due to the league has many teams with different performance levels.

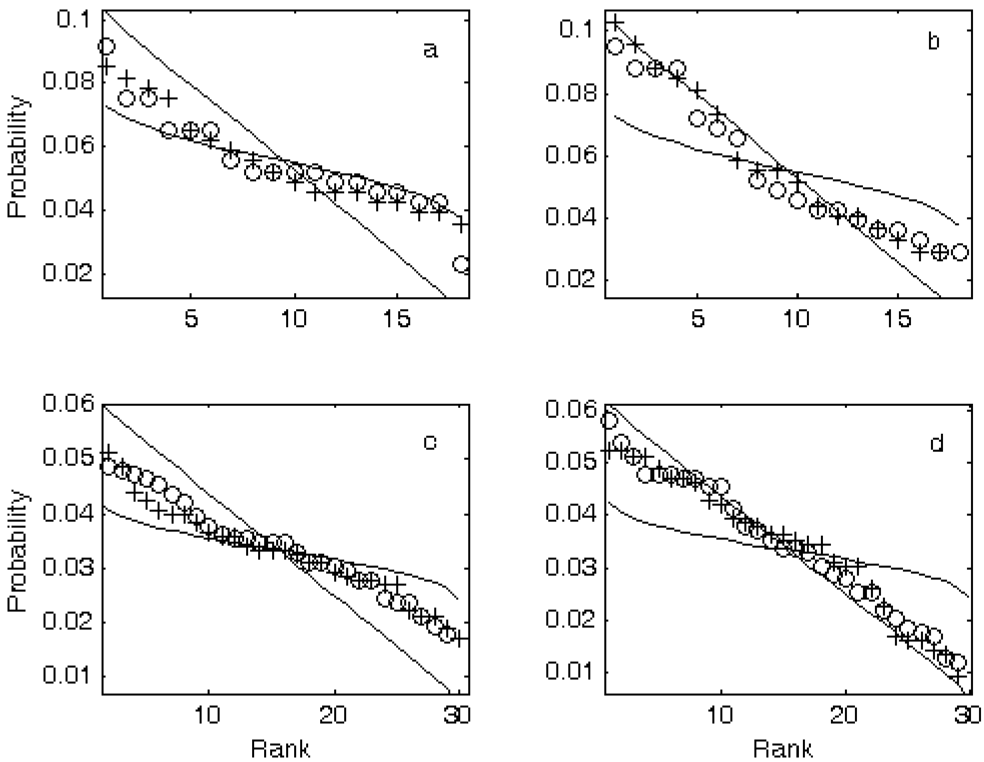


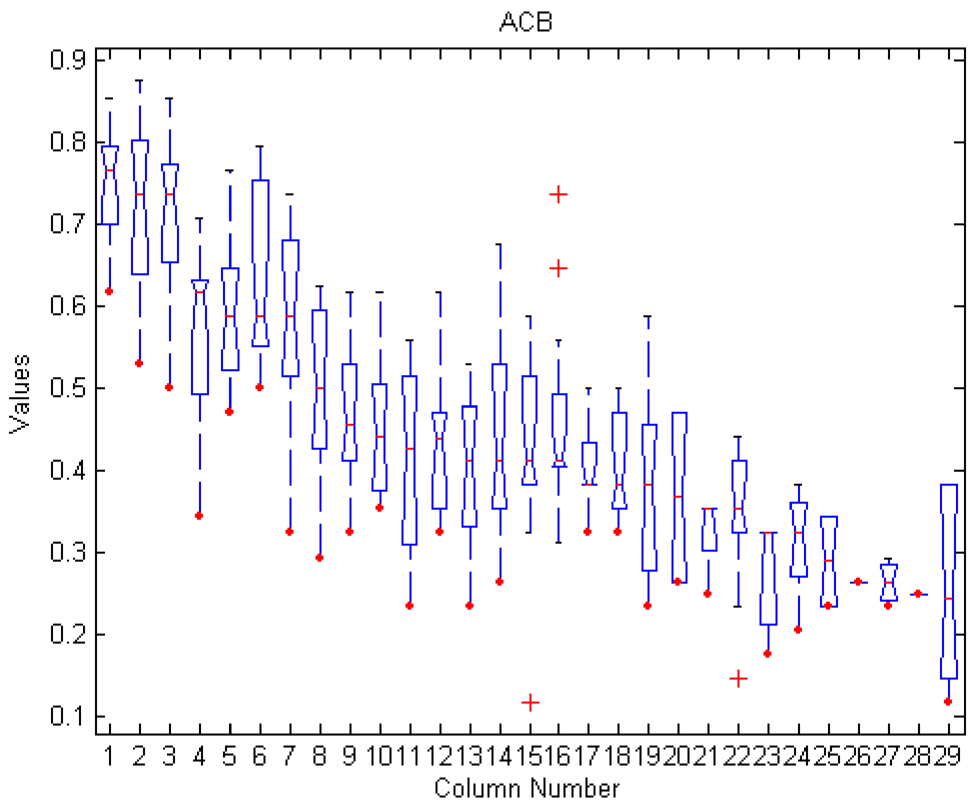
Figure 16. Examples of (a) two ACB seasons (2003–2004, 2005–2006) and (c) two NBA seasons (2003–2004, 2005–2006) that were more competitive compared to the theoretical random extreme. Also shown are examples of hierarchical (b) ACB seasons (2000–2001, 2008–2009) and (d) NBA (1996–1997, 1997–1998), compared with theoretical models of each league. Each point represents the probability value p obtained from the results array R of each season. The solid lines represent the two extreme cases. The straight line represents the hierarchical case, while the other (oblique line) results were used to calculate the probability of the average result obtained in the random process. We have discarded the results of NCAA in this analysis because the results can be very confusing due to the large number of participating teams.

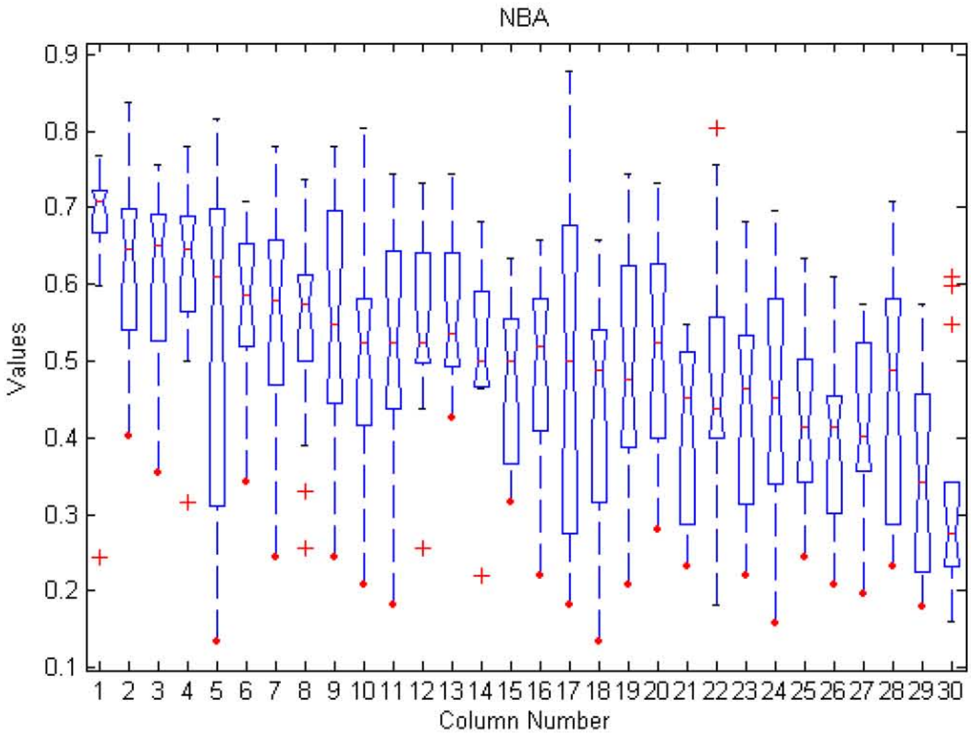
Figure 16 (a) shows the values of the probabilities p of the 2003-04 season (o) and 2005-06 season (+) of the ACB, as an example of a very competitive season. Figure 16 (b) shows the seasons 2000-01 (o) and 2008-09 (+) of the ACB as an example of seasons with lower *Sn* values. The same results are shown in Figure 16 (c) for the NBA, for the seasons of 2003-04 (o) and 2005-06 (+), and 3 (d) for the seasons 1996-97 (o) and 1997-98 (+).We have discarded the results of the NCAA because the results are meaningless due to the league has many teams.

The ACB shows a degree of competitiveness away from the theoretical hierarchical extreme, in which competitiveness is lower. In the most competed seasons, the entropy values are close to the theoretical random distribution, almost at the tail of the distribution. The rest of the values are located between the two distributions, although, in the case of lower entropy, the values are very close to the tail of the hierarchical random distribution. Similar behaviors can be observed in the NBA.

6.3.2. Statistical and cluster analysis

The Figure 17 represents the boxplot of the results **R** of all participating teams, through the seasons analyzed in ratio values (wins/games played).





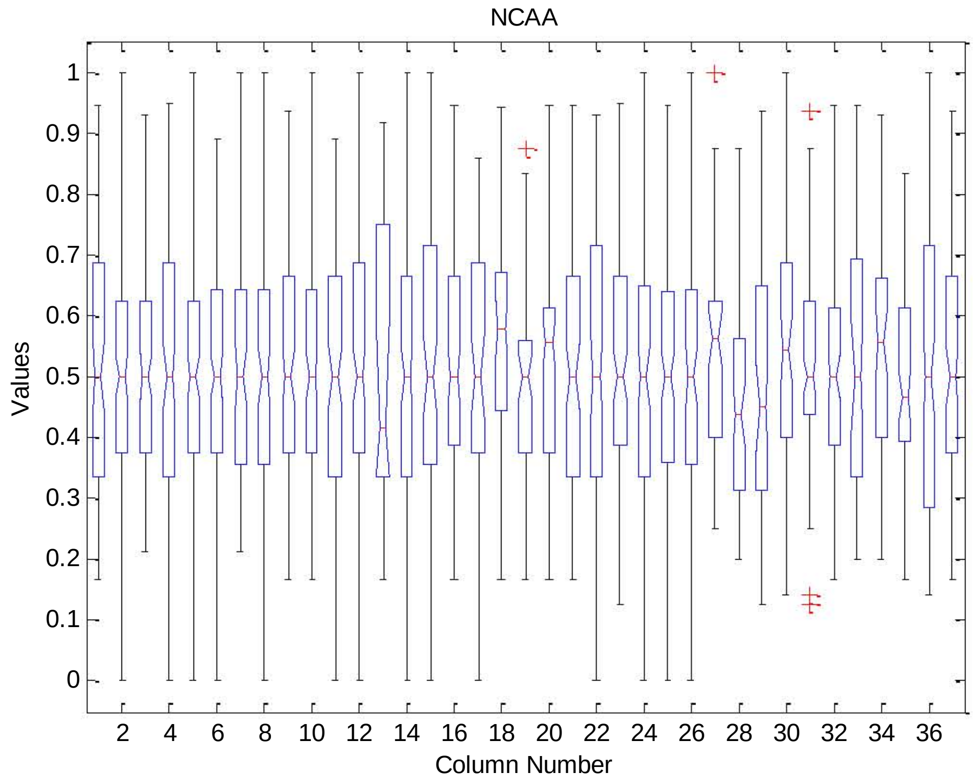


Figure 17. Boxplot of entirely results obtained, expressed in ratio (wins/games played), of the all participating teams in seasons analyzed. Upper and middle plots show up the professional leagues ACB and NBA. The bottom plot represents the same ration of the conferences of the NCAA (not teams) which participate in the seasons analyzed. In the results of ACB we can distinguish three clusters or three zones. And even within some of them, other delimitation on a smaller scale, which are occupied by teams to which can be called transition teams. In the NBA (the middle plot) teams seem to have values more similar to each other. This suggests a higher competitiveness

degree, because virtually any team can reach high levels of performance. On the other side, the NCAA plot (lower plot) shows a high homogeneous performance level and the data cloud covers a very wide spectrum. But despite of that data show this tendency, we have to take into account that teams do not play the same number of games. Hence the boxplot seems point out the same values.

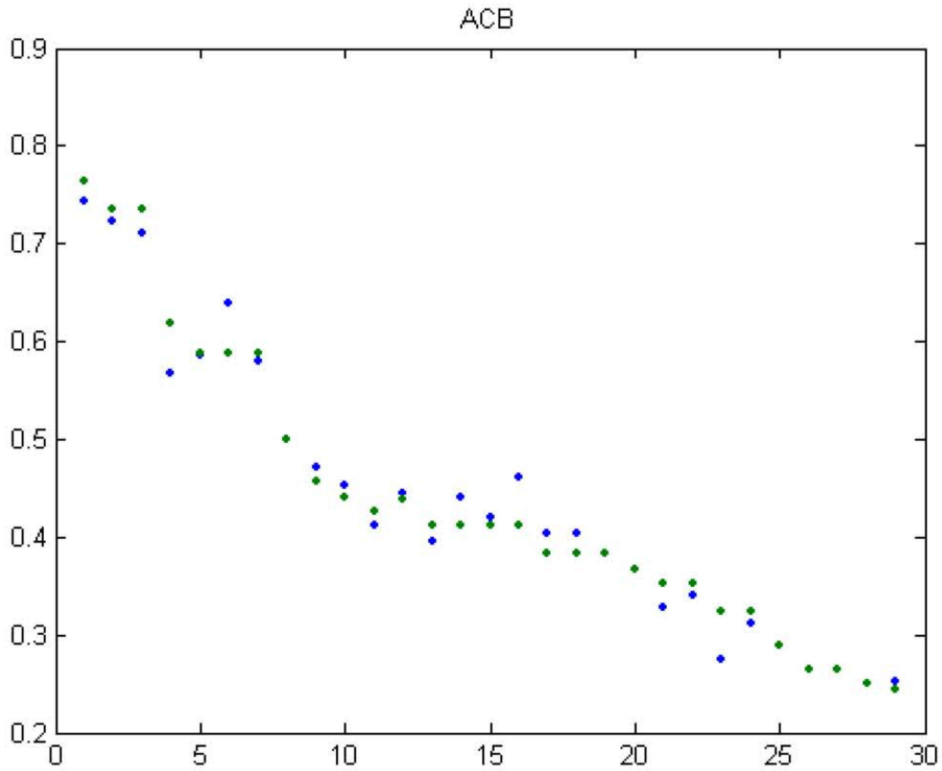
In all professional cases we observe that teams tend is to cluster around their performance level, although their behavior is different. Seemingly in the amateur league (NCAA), the values are much more stable.

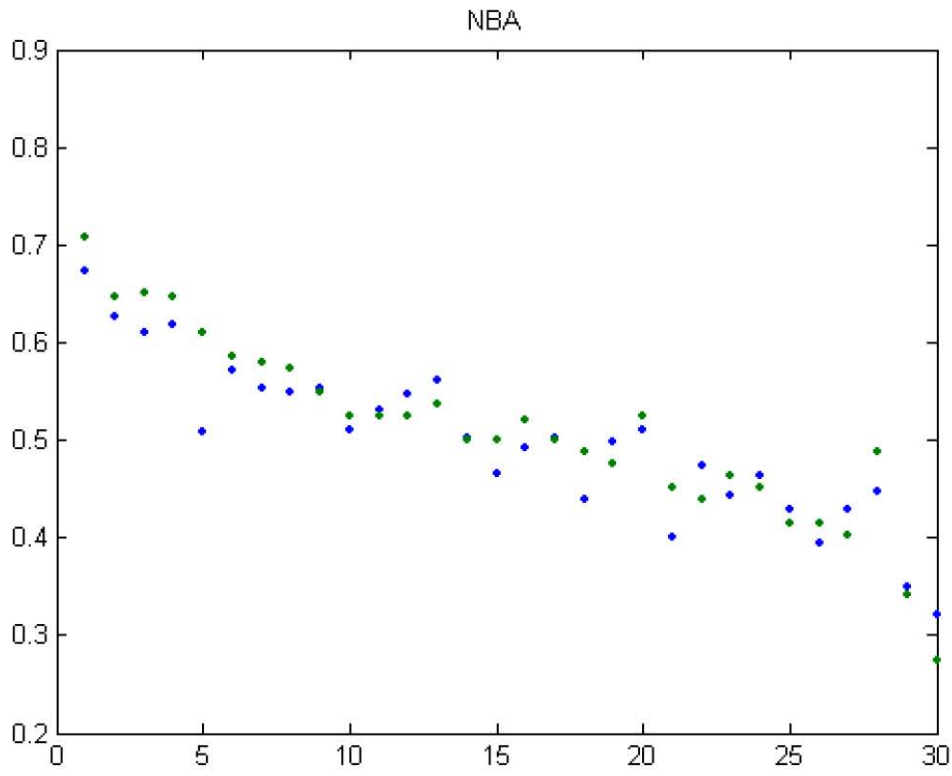
In the ACB (Figure 17) we can distinguish three zones. The first one, with the best performance level, the three best teams are located with their data (data cloud, mean, interquartile ranges and confidence intervals) which are clearly above the rest. These teams are following a model of competitive behavior that we previously defined as random. They cluster and compete to achieve first place (a very high degree of competitiveness between them). Also given the same situation for the second group, where the four teams have a similar performance level (a very high competitiveness among these four teams). The teams situated in the middle sector (third group), struggle among them very even. Moreover, we can talk about *transition teams*, where the level of performance places them in a bordering position with the other two areas of performance. In the last part of this group, the data have little statistical value, because this is the zone which suffers more changes due to the promotions and demotions.

In the NBA the data seems to point out a much more homogeneous behavior. The most part of the data cloud, and the medians, are situated around of mean values of ratio, which indicates a high competitive balance. Occasionally the teams reach unusually high values (high consolidated teams) o low values (low consolidated teams), some with a very strong scattering data, suggesting that they are teams with good results and now they have decreased their performance, or vice versa.

In the NCAA, on the other hand, there is a low variability in the boxplot values. The most part of the conferences reach range from 0 to 1 in the ratio values. But the point is that the teams do not play the same number of games. In the same conference we can find teams with 16 games played against teams with 4 games played. This is because the standings are elaborated by the RPI index, as we mentioned above. This force to several teams to play in tournaments in order to get a better punctuation. That is why the final balance is so heterogeneous. Actually this means that in the NCAA the sport gradients are much more accentuated than the professional leagues.

In the Figure 18 are represented the medians and means of the all the ACB, the NBA and NCAA values, which clarifies the mentioned above.





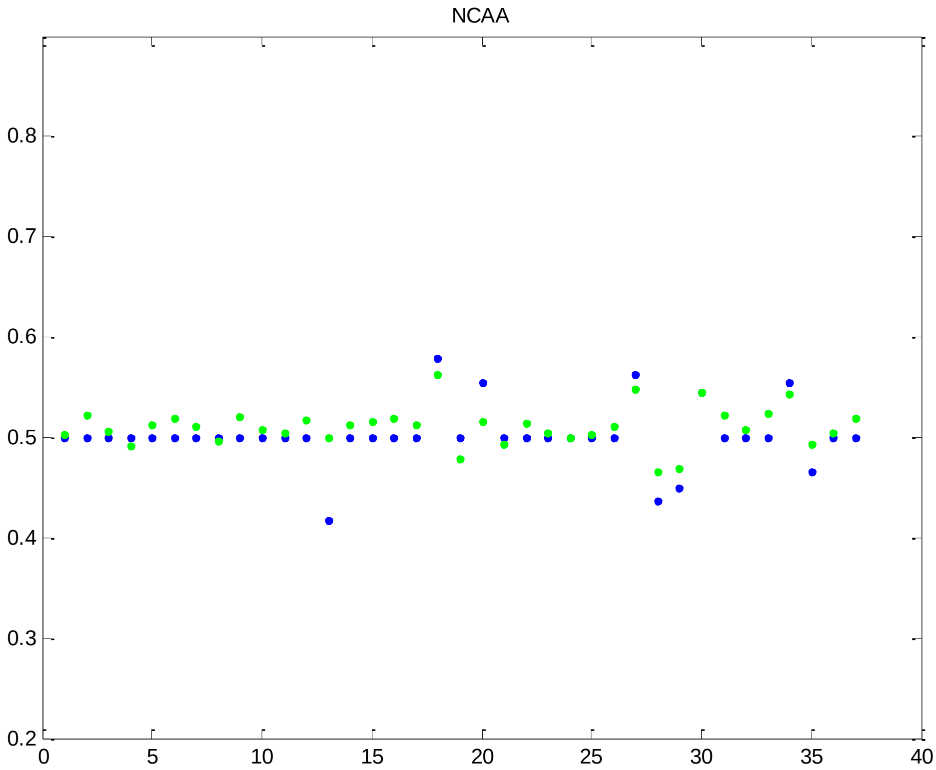
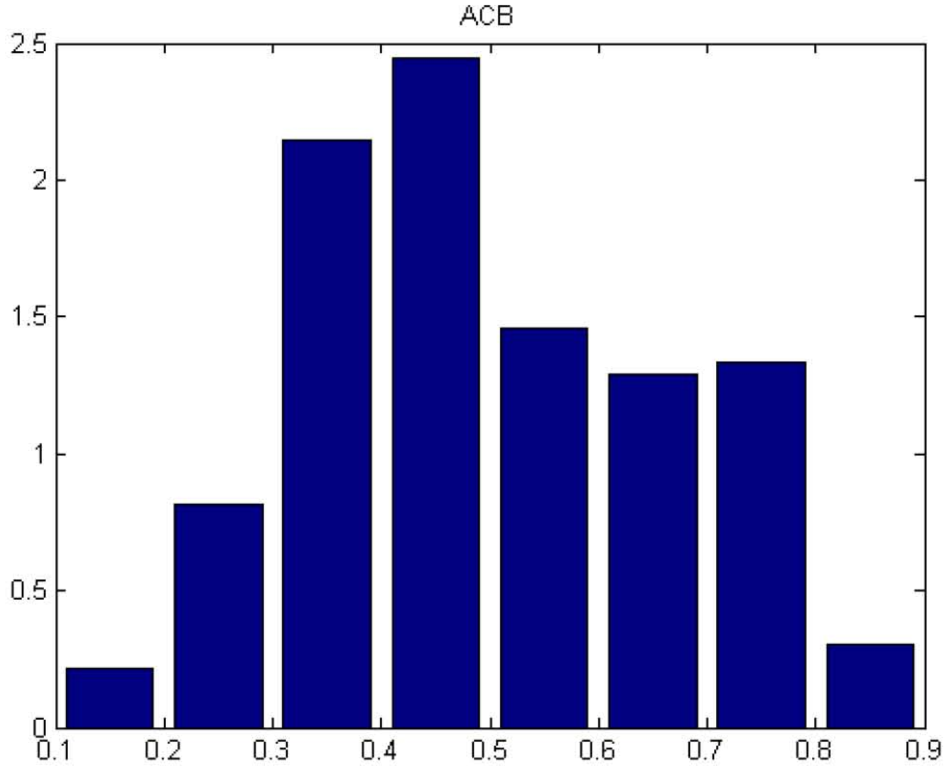


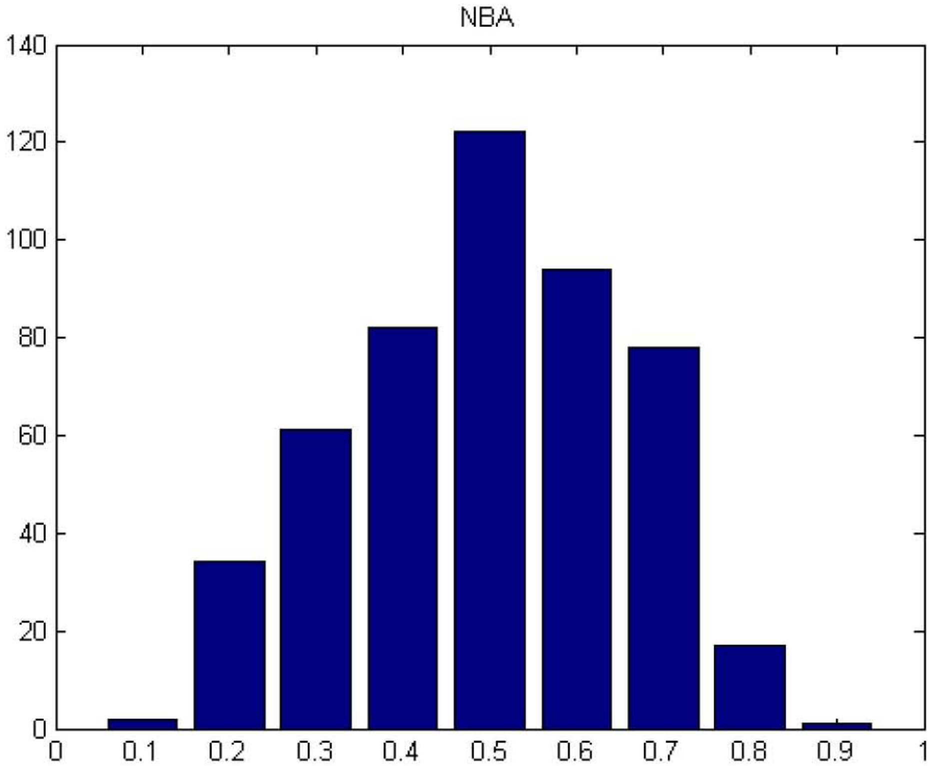
Figure 18. ACB, NBA and NCAA means and medians. Notice how clusters are clearly more visible in the ACB, while the trend in the NBA is much more compact. The NCAA, on the contrary, is much less dispersed than the professional leagues. Values only range from 0.4 to 0.6 and very similar among them.

In the ACB, we can distinguish three clusters clearly. These groups seem to be established according to their performance, taking into account the mean and medians (Figure 18). Meanwhile the NBA does not show such clusters. Its tendency is much more regular than the ACB. The amateur league, the NCAA, displays the lower dispersion of the data.

The ACB data cover a wider segment of ratio (from 0.75 to 0.25 approximately). The NBA by contrast has a lower dispersion range (from 0.70 to 0.35 approximately). The NCAA range from 0.4 to 0.6 approximately and its tendency is much more stable than the professional leagues.

The next histogram displays all the **R** values for the participating teams in the competitions mentioned.


ACB
2.5
2
1.5
1
0.5
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


NBA
140
120
100
80
60
40
20
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

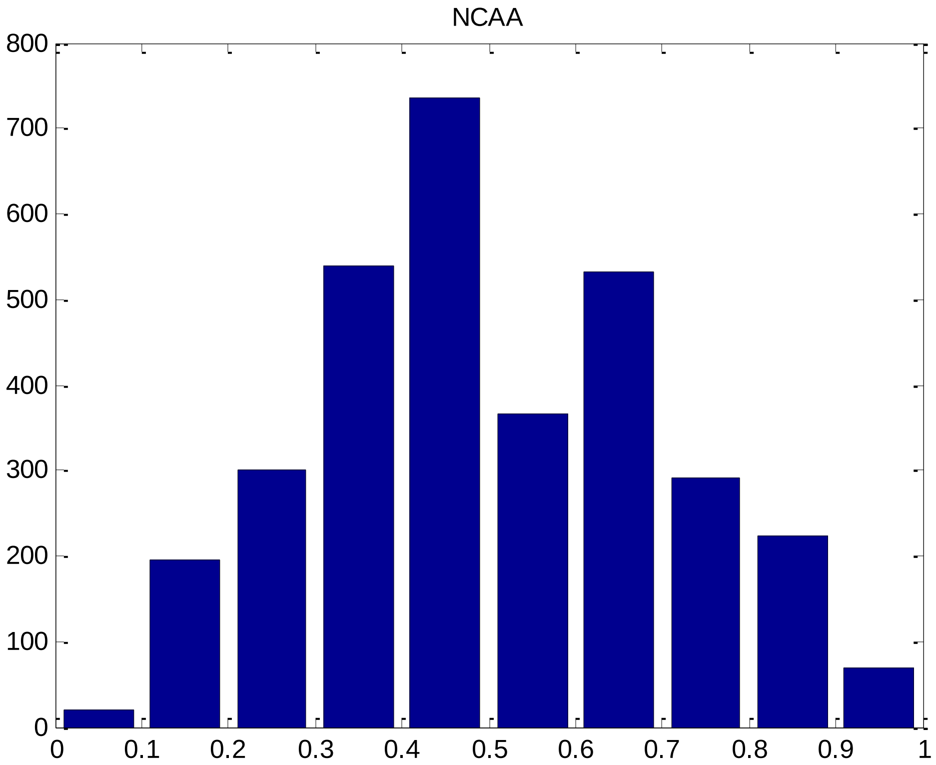


Figure 19. Histogram which represents the ratio values for the entire sample of the ACB (upper), the NBA (middle) and the NCAA (below). Note the asymmetry in the all the distributions. The shape of the distribution seems similar for all leagues but the behavior is completely different.

According to Figure 19, basketball leagues present a ratio peak value and some marked differences in ratio values. These differences indicate the competitiveness degree of each zone, meaning competitiveness as the likelihood of team to achieve a given winning ratio. A smaller difference means that the teams are well distributed throughout all the range of values of ratio, which indicates a tendency to remain in a given ratio value (hierarchical model). In contrast, if the differences are significant, we face to a case of higher competitiveness. This characteristic data distribution, many teams in a single ratio value (peak), and few teams that reach upper bounds (tail), makes us suppose that the data could behave as a *Power Law*, although the amount of data is not significant for such a claim.

In the ACB the peak is around 0.40 of ratio (Figure 19). The teams located below this point, are very irregular and they are not able to overcome the level of performance needed to strive in the middle of the league ranking. This point seems to work as a barrier, understood as a value significantly greater frequency. It is remarkable that most of the teams are placed in intermediate regions of the histogram, while fewer people than the second barrier (0.80 ratio), what can be considered the most competed area and is reserved for the few teams. If we observe the box plot (Figure 17) and the histogram (Figure 19), we can point out that the teams located above the barriers are always the same (except occasionally). The highly competitive area always seems to be occupied by the same teams. That is, teams are clustered around their level of performance. For these teams reach higher levels, it is necessary to overcome certain barriers of performance.

In Table 4, are represented the participating teams in the ACB fourteen seasons analyzed and their distribution in different areas (Figures 17 and 19), in addition to the number of inclusions in that segment. These zones correspond to: the values with a win ratio less than 0.40 (zone 3), data with win ratio between 0.40 and 0.80 (zone 2), ratio values above 0.80 (zone 1).

Table 4. Participating teams in the seasons analyzed and their distribution in zones.

Team	Nº inclusions Zone 1.	Nº inclusions Zone 2	Nº inclusions Zone 3
Alicante		6	1
Barcelona	11	3	
Bilbao Basket		6	
Cáceres CB		5	2
Cantabria Lobos		2	3
Estudiantes	1	12	1
Fuenlabrada		10	2
Gijón Baloncesto			3
Girona		11	1
Gran Canaria		13	1
Granada		8	3
Huelva			1
Joventut	3	10	1
León		3	2
Lleida		3	1
Lugo Breogán		6	1
Manresa		10	1
Menorca		3	1
Murcia		2	5
Ourense			3
Real Madrid	10	4	
San Sebastián		1	3
Sevilla	1	12	1
Tenerife		1	1
Unicaja Málaga	5	9	
Valencia	1	13	
Valladolid CB		11	2
Vitoria	10	4	
Zaragoza			1

We can deduce that the values located over 0.80 correspond to win ratio which are exceptional for the ACB. In our case, these values coincide with highly consolidated teams in this competition (see Table 4) and high performance in several European leagues. When we analyze the anomalous results obtained in some seasons (win ratio lower than 0.40), we observe that these data correspond to teams which, in the seasons analyzed, were poorly consolidated in the ACB (Table 4).

The NBA, in general, has an uncertainty degree higher than the ACB and its dynamic is completely different. In the Figure 19 we observe that the most part of the data are located nearby the 0.5 ratio.

Most striking is that in the extremes the behavior is similar, although with a meaning completely different. The worst results (ratio <0.15) do not demote. In fact, they have some advantage for the season next year, as these teams takeover of the top places of NBA draft, which means in a reinforcement in their roster.

In the other extreme of the histogram, we can see how to reach the best results is very unlikely. It is really difficult to reach values higher than 0.80 during the regular season. We must remember that the NBA seasons are very extensive (82 games) and the playoff classification is very hard-fought. Even we can see how get ratio results higher than 0.70 is infrequent. The most likely is that the most part of the teams are located in medium zones (Figures 17, 18 and 19). The same team can be fighting to get playoff positions and the next year can be located in a ratio <0.5 or vice versa (Figures 17, 18 and 19). I.e. reaching higher values than 0.7 0 or lower than 0.25 is unlikely for the most part of the teams. Note that the most part of the teams presents a similar performance level.

Regarding the analysis of NCAA as an example of amateur league, we can appreciate how the boxplot (Figure 17) and histogram (Figure 19) are completely different from the professional leagues. Standings are based on the RPI index, so the number of games played by the teams is not the same. This causes a marked variability in the number of games played by the participating teams. Probably that is why the distribution presents this shape (Figure 19). The peak is located, approximately, nearby 0.4. The NCAA presents two *Power Laws*; a possible double Pareto as indicated Figure 20.

We have to bear in mind that NCAA participating teams are significantly more (344) than ACB (16) and NBA (30). And the organization of competition is quite different from professional leagues. That is why it seems there are two competitions during regular phase, as indicate the log-log plot of ratio distribution (Figure 20)

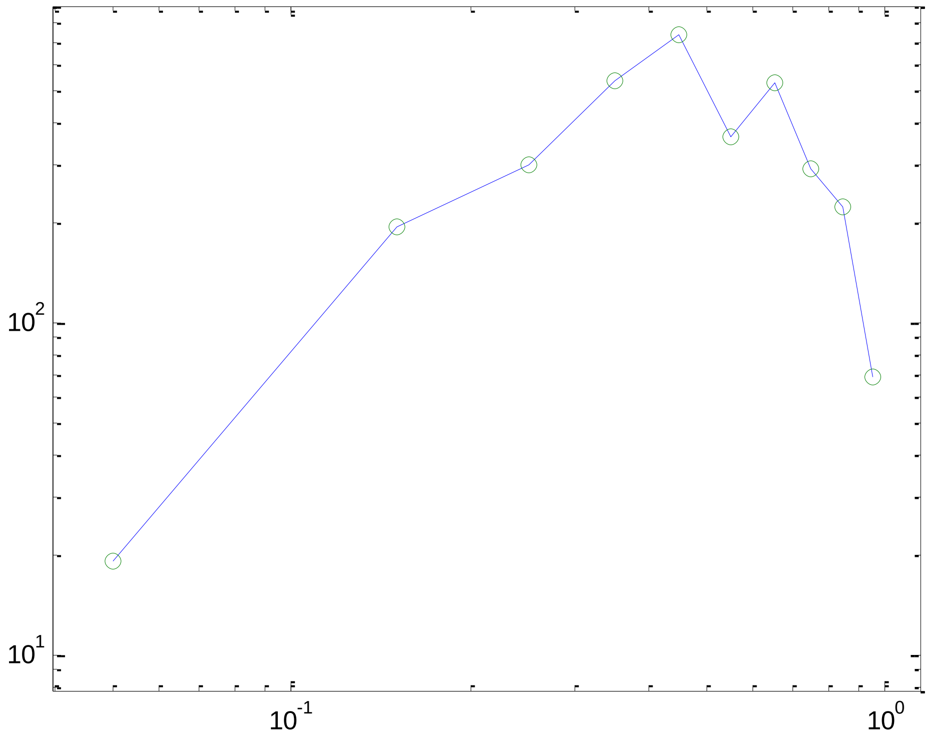


Figure 96. Log-log plot of the NCAA ratio distribution. There are at least two *Power Laws*, which indicate different competition dynamics.

In order to find out whether teams are gathered by their performance, we carried out a **non-hierarchical clustering analysis of partitional reallocation type** (k-means Matlab function) (Figure 21), which places the points in space to be grouped. These points are assigned to the group that is closest to their centroid. It is a method of cluster analysis which aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean. This results in a partitioning of the data space into regions called **Voronoi cells**.

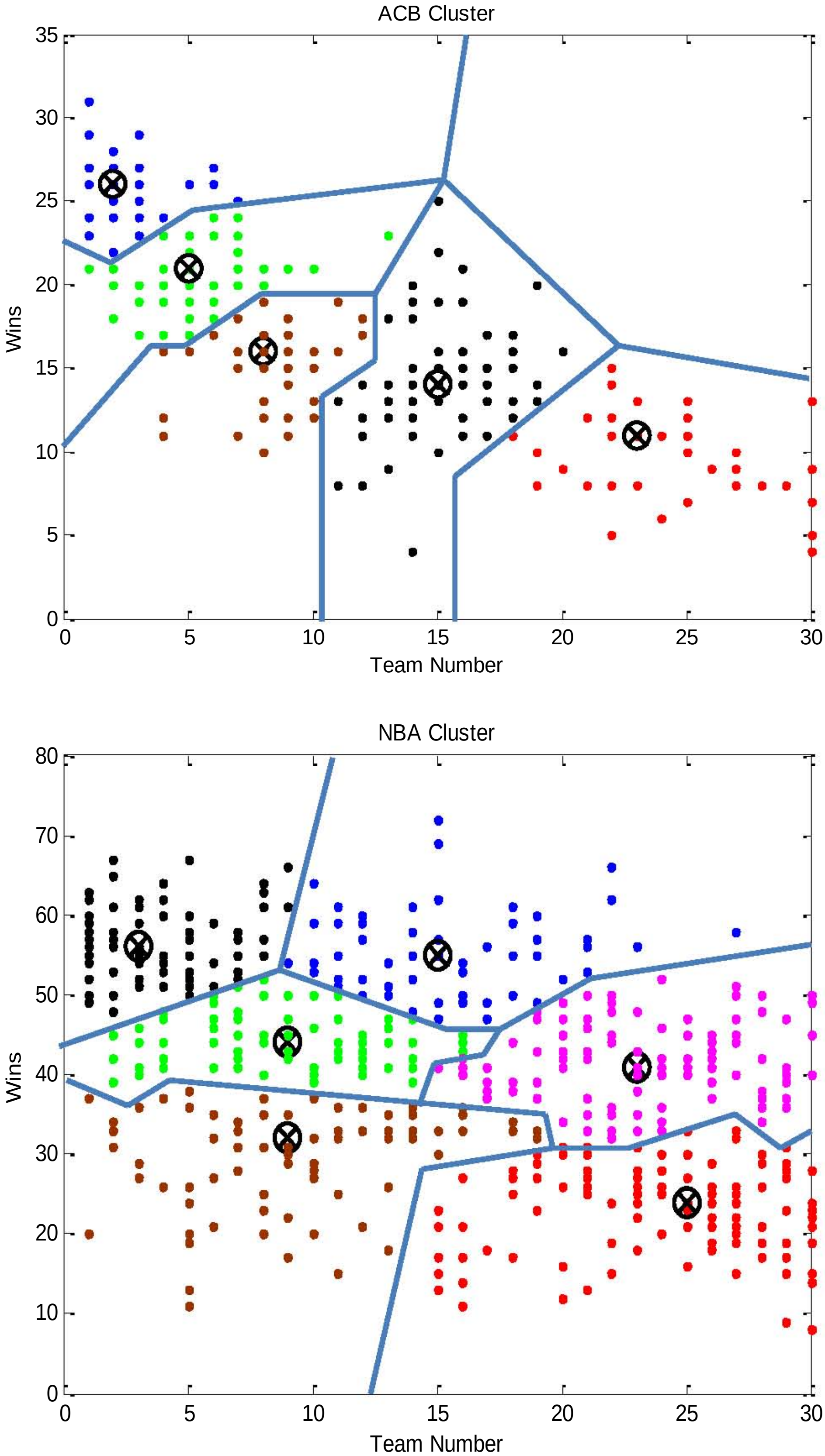


Figure 21. The upper panel represents the ACB clustering. We can observe that show up five regions which are clearly related with the team performance. There are some teams which are clearly located in one region (blue, black and red zones), and occasionally reach a different result. That is, they belong undoubtedly to a region. Teams

located in green and brown areas can be considered transition teams because sometimes reach better or worst results than other seasons. Even we could consider these two regions as a single region, regarding the behavior of teams located on it. The lower panel represents the NBA clustering. It is completely different to the ACB. The results point out four regions (red, magenta, brown and green) with a similar performance. It means that a team can be on top for one or several seasons and subsequent seasons at the lower standing, or vice versa. Moreover, exists an elite located in an own region by their own results (black) but seldom other teams manage to achieve these positions (blue).

At ACB clustering (Figure 21) comes up five regions clearly established by performance. We can see eve how the centroids are positioned hierarchically, which indicates stratification. The two lower regions (red and black) are clearly integrated by certain teams which sometime reached better results but visibly belong to those regions. The other region which is clearly separated from the rest is the blue region. Teams located in this area are markedly superior to the rest. Teams located in green and brown areas can be considered transition teams because at times reach better or worst results than the corresponding results to their zone. Thus we could consider these two regions as a single region, regarding the behavior of teams located on it.

The NBA clustering (Figure 21) points out six regions. The red, magenta, brown and green regions present a similar performance level, but it is not clear what teams are in each one.
It could means that due to the intern mechanisms of NBA teams after a bad season can be competitive for the next one. And teams with good results are obligated to restructure their roster season by season, in order to keep good runs. In fact, we can note that some teams present very good results (black and blue) and seasons not so good (some of them with a very marked data spread). This can point out dynasties. For instance, Chicago Bulls as long as Michael Jordan remained in the roster the team succeeded. But after his retirement, the Chicago Bulls fell into a run of bad results.

The NCAA uses the RPI index in order to elaborate standings, so this kind of analysis for the NCAA does not reflect the reality of the competition system. Hence we considered that the boxplot and clustering analysis of standings do not provide relevant information.

ACB

The ACB presents a structure almost hierarchical. Participating teams are clustered by themself regarding performance level (Figure 21) and this phenomenon creates frequency barriers (wins frequency) for teams less powerful. The ACB peak is around 0.40 ratio (Figures 17 and 19). Teams located below this point are very irregular and are not able to overcome the level of performance needed to strive in the middle of the league ranking. This point seems to work as a barrier, understood as a value significantly greater frequency. It is remarkable that most of the teams are placed in intermediate regions (Figures 17 and 19), and only a few teams are positioned beyond the second barrier (0.80 ratio), which can be considered the most competed area (Figures 17 and 19).

Teams located above the barriers are always the same (except occasionally). So we can point out that the highly competitive area is always occupied by the same teams and so on (Figure 17, Figure18, Figure 19 and Table 4). That is, teams are clustered around their level of performance. Therefore, teams have to overcome certain barriers of performance if they want to achieve higher levels of performance

The best results in the ACB coincide indeed with highly consolidated teams in this competition and with a high performance in several European Leagues. Results lower of ACB data correspond to teams which, in the seasons analyzed, were poorly consolidated (Table 4).

These different performance regions (Figure 21) could be originated because ACB competition model. The ACB is an open league model in which participating teams are adjusted based on promotion and demotions (from or to lower category), and where eight top ranked teams play the play-off. We must take into account that teams elaborate their roster depending on their budget, and that this is, almost always, related to the results obtained. The higher the budget, the better players, coaches and staff can be hire or vice versa. A priori, promoted teams have less competitive rosters, and also tighter budgets.

Given its open structure, some teams (and its underlying structures such as economic network, executive committee, player network, farm of players, etc.) become more experienced throughout all seasons. The existence of these teams has an impact on the rest, and above all on the less experimented teams. Thus, the teams positioned on the extremes are closely related: if the differences among the low ranked teams and the top ranked teams are very high, it is possible that the upturn of the head is more evident, because exist a high probability that the top ranked teams defeat the bottom teams. Thus, the best teams can improve their winning ratio.

It can be deduce that there is a different criticality level for each zone. This sport potential gradient is maintained by energy (players, coaches, money, etc.). This provides that the performance differences for some ACB teams are insurmountable especially to the newly promoted whose budget and templates are tight. The sport model greatly influences the market.

The fact that the teams tend to cluster in zones it is not random, but follows a phenomenon known as *preferential attachment* (Barabási & Albert, 1999), also denominated *Saint Matthew effect* (Bunge, 2001; García Manso & Martín González, 2008), where the strong teams reap more successes and less strong teams will have less wealth. Another mechanism that causes this behavior is *memory effect*, what the systems present. The teams are attached to an attractor as in some areas of the ranking (Figure 17, Figure18, Figure 19 and Table 4).

Others reasons for these differences could be teams´ sport planning, sport aims established for each season, roster, budget, external competition (European Leagues, King´s Cup, tournaments or participations of players in national teams), etc. These aspects may influence substantially.

NBA

Generally, the NBA has an uncertainty degree bigger than the ACB (de Saá Guerra et al., 2012) and its dynamic and structure are completely different. In the Figures 17, 18 and 19 we observe that the most part of the data are located nearby the 0.5 ratio and how teams are dispersed for several regions (Figure 21).

There are not demotions neither promotions. In fact, they worst results (ratio <0.15) (Figures 17 and 19) have some advantage for the next season. These teams takeover top places of NBA draft, which means in a reinforcement in their roster.

We also can see how to reach the best results is very unlikely (Figures 17and 19). We must remember that the NBA seasons are very extensive (82 games) and the play-off classification is very hard-fought. Even to get ratio results higher than 0.70 is infrequent. The most likely is that the majority of the teams are located in medium zones (Figures 17, 18 and 19). The same team can be fighting to get play-off positions and the next year can be located in a ratio <0.5 or vice versa (Figures 17, 18 and 19). I.e. reaching higher values than 0.70 or lower than 0.25 is unlikely for the most part of the teams. Note that the most part of the teams present similar performance levels; that is why clustering present regions similar where teams change their locations during seasons analyzed (Figure 21).

The existence of this performance dynamic could be also due to the sport model employed by the NBA. We must take into account that in the NBA participate much more teams than in the ACB (30 vs. 18), and they play much more games (82 vs. 34). Moreover, the competitive structure is diametrically opposite. There are also mechanisms imposed by the NBA in order to avoid team monopoly (draft, salary cap, reserve clause, etc.). The purpose of these measures is to safeguard always the competitive balance. Therefore it is possible that the most critical parts of the competition are located in the two limits, because they are areas where teams are positioned in them will get rewards (play-off and draft). Perhaps due to its competitive dynamic, the NBA is a good example of Red Queen hypothesis proposed by Van Valen (Van Valen, 1973): *For an evolutionary system, continuing development is needed just in order to maintain its fitness relative to the systems it is co-evolving with*. It is an endless race. All competitors need to improve to keep competing.

ACB and NBA Comparison

The ACB and the NBA seem to present an inverse behavior. In the ACB case, most competed region is the medium ratio area (lowest differences) and green and brown regions (Figure 21). In the NBA, the most competed area is the top and the end of the standings. That is why teams are so disseminated in the cluster analysis (Figure 21). We note that both cases are cases of a highly competitive, but opposite reasons: the ACB is an open model where the last classified is relegated of category, hence the high degree of competitiveness, while in the NBA, the point is qualifying for the play-off for the title or trying to get a good place for the lottery draft.

In the ACB, we observe that teams are clustered throughout their performance level as well (Figure 21). There are teams clearly placed in a particular area of the competition, which could indicate the competitiveness level of the team. The first four positions are occupied almost entirely by the same three teams and, occasionally, some team was able to slip into this elite group (Figure 17, 18 and 19). There is a similar pattern with the play-off positions (the first eight positions), where we can observe clearly how the data cloud and the trust intervals of several teams are encompassed in this area. The last positions are more atypical, because the last two teams pass to a minor league and are replaced by other two different teams. The new promoted teams, a priori, have not the same performance level that the teams of the middle zone.

In the NBA, almost all teams have ever reached the play-off positions, although there are more teams that belong to this area than others (Figure 17). Their data are less scattered and more firmly established in this area. Note that practically all the teams have reached the top five.

Given the high randomness degree present in the ACB and the NBA, we can suppose that the most part of the teams are between the order and chaos, known as critical state. A state of semi equilibrium where the most insignificant variation can produce a change of state or a phase transition, but this is very hard to predict (McGarry et al., 2002; Scheffer et al., 2009). The degree of criticality varies based on the zone of the standing we are studying.

But note that teams, even though the chaotic behavior of the competition, always tend to an attractor (team clustering). Therefore, we can consider competitiveness as an atractor itself. We must remember that complex systems are usually far from the equilibrium. E.g. the living organisms are in continuous fight with the environment in order to remain in a state far from equilibrium, i.e. alive (Amaral & Ottino, 2004). Translated to our context, it means that to preserve the high level of competitiveness, it is necessary to keep fighting against the rivals, and to invest huge amounts of energy in order to survive in the competition.

6.4. Conclusions

The aim of our study was to analyze sport competitions from a general point of view. This study show that the analysis model (matrix results using the Shannon entropy) for the study of competitiveness levels in the system of league competition is a useful and highly sensitive tool to determine the degree of overall competitiveness in the league, and to detect small oscillations in it. This potentially identifies minimum fluctuations in the level of competition, which allows one to focus attention on localized temporal changes and to investigate the mechanisms which cause it.

This model shows that both the ACB and the NBA present a high degree of competitiveness. In both leagues the entropy levels are high (range: 0.985–0.990), although these periods are more stable in the NBA. We can say that both the ACB and the NBA are very competitive leagues whose teams are well balanced within each league. On the contrary, the amateur league, the NCAA, presents a lower competitiveness degree (compared with the professional leagues). But its tendency is to increase during the last seasons.

We can say both the ACB and the NBA are very competitive leagues with a high competitive balance and they are highly conditioned by the sport model. The fact that the ACB is an open league causes that less powerful teams subtract competitiveness to the entirety. We should think about strategies in order to maintain or even increase the degree of global competitiveness of the league, like in the NBA. Despite these issues, the Spanish basketball league (ACB) can be considered very competitive. The NBA has specific mechanisms to ensure high competitiveness, such as the draft, the salary cap, reserve clause, etc. Their aim is to preserve the competitive balance within the system. It is a league with a high uncertainty on the final result; hence, all teams have real possibilities for qualifying for the playoffs.

The case of the NCAA is pretty interesting as well. Despite of the low entropy (compared to the professional leagues), NCAA is an attractive league that generate expectation and attracts thousands of fans and media. Due to its complicated structure and size (number of participating teams), it is complicated carry out an analysis of the entire league. But we can observe that indeed, changes in its sport model, such as enlargement of divisions, rule changes, new punctuation system, etc., alter significantly the competitiveness level throughout years analyzed.

6.5. Practical Proposals

As a practical proposal we suggest to use this methodology in order to study and compare the competitive level of different basketball leagues and their evolution in time. I.e. different basketball leagues such as Euroleague, Eurocup, NBAD-League, ABA, LEB, etc. This process also can point out some events which can be the source of such variations.

As another practical proposal, we put forward to use this methodology in order to figure out how rule changes, enlargement of participating teams, competition format restructuring (open or close, divisions, conferences, etc.), or other league mechanisms such as budget rules, building up teams (hiring), or even practices regulations (i.e. NCAA) can modify the competitive level.

6.6. Limitations of the techniques used

Shannon entropy, as methodology to study the competitiveness in basketball, carries out a coarse analysis of the competitive balance, in the sense that the results do not take into account the team rankings.

On the other hand, if we only use the ratios, boxplots or even clusters in order to analyze the competitiveness degree, we will not know the league overall.

Study 2.
Basketball Game

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

7. Study 2. Basketball Game

7.1 Introduction

Several sources indicate that the degree of competitiveness of a basketball league displays non-linear behavior (Yilmaz & Chatterjee, 2000; de Saá Guerra et al., 2011). The final outcome depends upon the equality between teams and the level of uncertainty before and during each game. The evolution of the score and its final value are what generate uncertainty for each game and for the final standings of a league.

A basketball game can be understood as a clash between two complex systems (teams), both seek to overcome each other within a limited time frame (Chatterjee & Yilmaz, 1999; Bar-Yam, 2001; Vaz de Melo et al., 2008). In complex systems, numerous processes occur simultaneously, exhibiting varying levels or degrees of their respective behaviors. The intricate behavior of a complex system, as a whole, depends on the behavior of all of its units, indirectly, which have strong, often non-linear relationships with one another (Goodwin, 2002; Amaral & Ottino, 2004). The tools and concepts emerging from these new theories of complex non-linear systems can be highly useful for analyzing team sports. Innovative concepts include those of complex networks, co-evolution, game theory, and self-organized criticality. In addition, certain theories of evolutionary ecology, namely the Red Queen hypothesis, a metaphor of a co-evolutionary arms race between species (Solé & Goodwin, 2002), can be used to model economical human behavior within sports analyses (Robson, 2005).

We work under the hypothesis that, theoretically, the score of any given basketball game reflects a random dynamic, a stochastic behavior similar to Brownian motion (random trajectory described by a particle), in the sense that we do not know how big the runs of points will be, or how often points will be made. Hence, we cannot know, in advance, the dynamic behavior of basketball game scoring with absolute certainty.

Far from these assertions, reality shows us that the score of a basketball game is a direct reflection of the dynamic and non-linear interactions of the teams and their components (Kubatko et al., 2007; George, Evangelos, Alexandros, & Athanasios, 2009; Yarrow, Brown, & Krakauer, 2009; Bourbousson et al., 2010; Ziv, Lidor, & Arnon, 2010). However, the evolution of the score seems to follow specific patterns that confer identifiable characteristics to each league. Developing a methodology to identify them thus allows us to know in more detail the internal logic of competition.

The typical way to approach this problem (scoring time in basketball) is to consider it as a random process and treat it as a Poisson process, an arrival process or a point process. Hence, some authors note that the goal distributions in football are not well portrayed by this statistic model, above all at the tails of the distributions. These authors subsequently studied the data distribution by using other statistical models, such us a Poisson process with parameter variable, a negative binomial, a log normal or a generalized extreme value distribution (Malacarne & Mendes, 2000; Greenhough et al., 2001; Mendes et al., 2007; Bittner et al., 2009; Heuer et al., 2010).

Some of these papers emphasize the presence of heavy-tailed distributions (*Power Laws*), which are associated with many natural and social phenomena. These kinds of distributions are related to ideas from statistical physics and non-linear complex systems, such as anomalous diffusion by the Zipf-Mandelbrot law (Malacarne & Mendes, 2000), self-organized criticality phenomena, or from non-linear dynamics (McGarry et al., 2002; Bourbousson et al., 2010).

Our aim was to study the behavior of timing and scoring in basketball games, and we also believe that searching for patterns or basic regularities might be helpful to better understand and quantify the complex systems involved in this sport. We consider the dynamics of the scores in basketball games and, more specifically, their evolution in games of the American professional basketball league (NBA) during 2005-06; 2006-07; 2007-08, 2008-09; 2009-10 seasons. Our interest, in this study, is aimed at the data analysis and the conclusions that can be obtained from possible deviations from random pure behavior, considered as a Poisson process, and from the presence of *Power Laws*.

7.2 Methodology

Team sports have proven to be a good field for analyzing complexity, e.g. random walk, statistics, extreme events, complex statistics, etc. (de Saá Guerra, Martín González, Sarmiento Montesdeoca, et al., 2011; Gabel & Redner, 2012) and also in the construction and exploration of models, in particular basketball.

In this study we analyzed in detail basketball from statistic and game perspective in order to provide an adequate framework where to essay this kind of analysis. It is important to understand the game from statistics, rules, time, scoring, etc.

In general, scoring, in many team sports including basketball, has been considered a Poisson process, although with some restrictions (Heuer et al., 2010; de Saá Guerra, Martín González, Sarmiento Montesdeoca, et al., 2011; Gabel & Redner, 2012). Our interest is to use the framework of Poisson random processes and its limits in order to find out what extent, the Poisson process fails in certain situations and in some extreme situations. Moreover, we try to find complex patterns through scaling analysis or *Power Laws*; or even regularities to improve the understanding of this sport within the completely random framework of the Poissonian model. Hence, in this work the idea was to use two basic framework of reference in order to analyze basketball scoring:

1. Poisson model
2. Scaling analysis

7.2.1 Poisson model

The Poisson distribution is a discrete probability distribution used to model the number of events occurring within a given time interval T

$$p(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

Where p(x; λ) is the probability of observing *x* events in a given time interval, λ is the mean number of events per interval and *e = 2.718282*.

The Poisson model presents these main features:

I. Simplicity: it depends on a parameter which has a precise physic meaning. I.e. arrival rate over time. A single lambda.
II. Matching mean and variance (Index of Dispersion). In real games this allows a quick check of the process.
III. Scoring time intervals follow an exponential distribution, which only depends on a single parameter and also it is easily detectable in a semi-log plot.

Index of dispersion

Index of dispersion can be considered as measure to figure out whether data are modeled by a Poisson process. It is defined as the ratio of the variance to the mean.

$$D = \frac{\sigma^2}{\mu}$$

For Poissonian processes *D = 1*.The variance is equal to the mean. For other results there are two cases:

The first one, when the Index of Dispersion is less than 1, is considered as under-dispersed, which means that the probability data are more clustered around the mean making more predictable. This condition is related with more regular patterns than the randomness associated with Poisson. Points are scattered more regularly.

The second one, when de Index of Dispersion is more than 1, called over-dispersed, indicates larger data dispersion. This case usually treated by testing a negative binomial distribution. This case points out the possible existence of clusters in data throughout time.

In brief, data more clustered or concentrated in time are considered over-dispersed; on the other hand, more regular or periodic data are considered under-dispersed. We have to take into account that an excessive mathematical analysis is useless if do not can be managed by coaches or sport professionals. Hence, this index represents a great tool and easy handling, which point out the phenomenon randomness degree.

7.2.2 Application of Poisson Model to Sport

The distribution limit of data collected from football and basketball cannot be the uniform distribution, which represents the highest entropy values in a general process, because all data have the same occurrence values. This is evident when testing the final standing of a league (de Saá Guerra et al., 2012) where the information entropy help us to distinguish between random and hierarchical. In the scoring case during a game, the scoring probability (0 points, 1 point, 2 points and 3 points) cannot be the same for all the time intervals.

In football, it is easy to find games with zero goals, in contrast, in basketball it is usually to score often. The distribution presents, necessarily, a decreasing long tail, because it is the time provided to score (game time) is finite. In this case, the maximum information entropy could be the Poisson distribution. In the Basketball case we can define relative entropy, but previously normalized with Poisson distribution. Usually this process is carried out using uniform distribution.

7.2.3 Scaling behavior

The Scaling or *Power Law* distributions play an important role in describing non-lineal complex systems, and they are often used to describe many natural and social phenomena. Due to its heavy-tail, *Power Law* distribution suggests that extremely large values occur at higher frequency than in other distributions, such as the normal or exponential. This indicates that common, small events are not qualitatively different from large, extreme events. Scaling or *Power Law* relationships arise commonly as probability or frequency-size distributions, and are characterized by the form $f(x) = Cx^{\alpha}$ where C is a constant and the value $f(x)$ is proportional to some power of the input x (White, Enquist, & Green, 2008). This function is linear when it is plotted on a log-log scale, and the slope of the resultant straight line gives an estimation of the scale exponent α.

The scale-invariance is another characteristic associated with *Power Law* distributions, (Sornette, 2004; M. Newman, 2005), thus such phenomena have the same statistical properties at any scale and are not associated with a particular one. Scaling laws are often generic; they show robustness and do not depend upon details.

A wide number of nature phenomena follows *Power Law* distributions (Schroeder, 1992; M. Newman, 2005), even in extreme natural hazards e.g. earthquakes (Mega et al., 2003), floods (Malamud & Turcotte, 2006), landslides (Li, Ma, Zhu, & Li, 2011) or forest fires (Song, Wang, Satoh, & Fan, 2006). Most of such phenomena show properties of self-similarity like the fractals; objects that show the same structure at all scales. The presence of *Power Laws* have also been suggested as the fingerprint of systems that show self-organized criticality (Bak, 1999).

The detection of a *Power Law* pattern in the empirical data values or in the time interval between them can indicate the presence of unusual underlying mechanisms or processes like feedback loops, random network, self-organization or phase transitions (West, Brown, & Enquist, 1997b; Barabási & Albert, 1999; M. Newman, 2005).

Scaling analysis indicates that the probability of extreme events might be estimated by extrapolating of *Power Law* distributions. Recently some authors (Pisarenko & Sornette, 2012; Sachs, Yoder, Turcotte, Rundle, & Malamud, 2012; Sornette & Ouillon, 2012; Yukalov & Sornette, 2012) have tried to characterize the so-named Dragon-King, extreme events with important social implications, which exceed these extrapolations.

One important limitation of this tool is the occurrence of the *Power Law* behavior at the tail of the distribution where is desirable to have the best accuracy (Stumpf & Porter, 2012). However, the tails are the zones of the empiric distributions where the greatest fluctuations are found. On the other hand, the scaling laws do not always provide the best data fit. Alternative distributions to characterized these environmental phenomena are the lognormal (Mitzenmacher, 2004), the stretched exponential (Laherrère & Sornette, 1998) and other truncated *Power Law* (Redner, 1998; Tsallis & Albuquerque, 2000; Burroughs & Tebbens, 2001).

Despite of these difficulties, there are some advantages of the application of *Power Law* and fractals tools for diagnostic, characterization and even prediction purposes, are the simplicity of the distribution and universality of its self-similarity. This is quite consistent with much of the literature on fractals and scaling in ecologic, geophysics or economics systems. Furthermore, in the latest years; improved statistical tests have provided strong evidence of scaling laws over a substantial (although limited) range of scales (Clauset, Shalizi, & Newman, 2009; Virkar & Clauset, 2012).

One of the goals of this study is to provide a qualitative background on the lineal fit of the power law distributions and particularly the log-log plot analysis. This allows the comparison of analogous phenomena and the characterization of regions over a similar environment. For this purpose we don't need to have absolute certainty that an empirical data set follow a *Power Law*.

We used these two tools (Poisson distribution and *Power Law* distribution) in our analysis. They pointed out the behavior of the system (Basketball).

Problems with Poisson distribution

1. When the index of dispersion is less than 1, the negative binomial distribution can be modeled for discrete data, because presents a longer tail than Poisson distribution. The tail of the distribution could decrease slower than in this case, so we should use the *Power Law* or truncated *Power Law* distribution, which theoretically, present a special meaning regarding extreme events distributions.

2. Another case is when λ is not constant during the game or in each minute of the game. It also known that if lambda follows a gamma distribution, the process is modeled by a negative binomial.

We assume the games as reasonably homogenous events. But, data indicate that not all the games are competed (homogenous), meaning low competitiveness, low team quality, period of the season, roster differences, etc.

The system behavior is neither the same throughout the game time, nor in the first quarter, nor the fourth quarter not even last minutes of the game. Note that depending on the game time, can be considered as different games (substitutions, score differences, faults, etc.). And also is it influenced by tactic decisions (faults, time out, defense, etc.)

7.3 Results and discussion

Application to basketball

We studied a total of 5 seasons (1230 games per season, with a total of 6150 games) of the NBA regular season. In every game we analyzed the game transcription published by the NBA in which are described in detail, all events play by play (NBA). All the statistics reflect the incidences of game ordered by the time in which they occurred (chronological order): two and three point shots, free throws (made and missed), defensive and offensive rebounds, turnovers and steals; violations (out of bounds, fouls, technical, etc.) substitutions, etc. From all this information we focus on the analysis of point time intervals and scoring.

As we described in the methodology, we propose the use of the Index of Dispersion and the λ value in order to analyze basketball games. We based our study to analyze what happens every minute of the game independently. That is, analyze the probability p1(k) that in the first minute k points are scored, p2 (k) in the minute 2, and so on, until the minute 48 for every game. Then we saw how far each of these 48 cases presented Poissonian behavior. To do this we calculated the mean number of points scored in every minute for all games λ_i :

$$\lambda_i = \frac{\sum_{k=1}^{N} n_{i,k}}{N}$$

Being *N* the number of games (N=6150), and *i=1, 2 , … 48*. In addition, we also calculated the Index of Dispersion, which as we saw is an indicator of the extent to which random variable behaves like a Poisson process. As we mentioned above, the Index of Dispersion is the ratio of the variance of the points scored in every game every minute to the mean value of these.

The next Figure represents both values applied to the games studied.

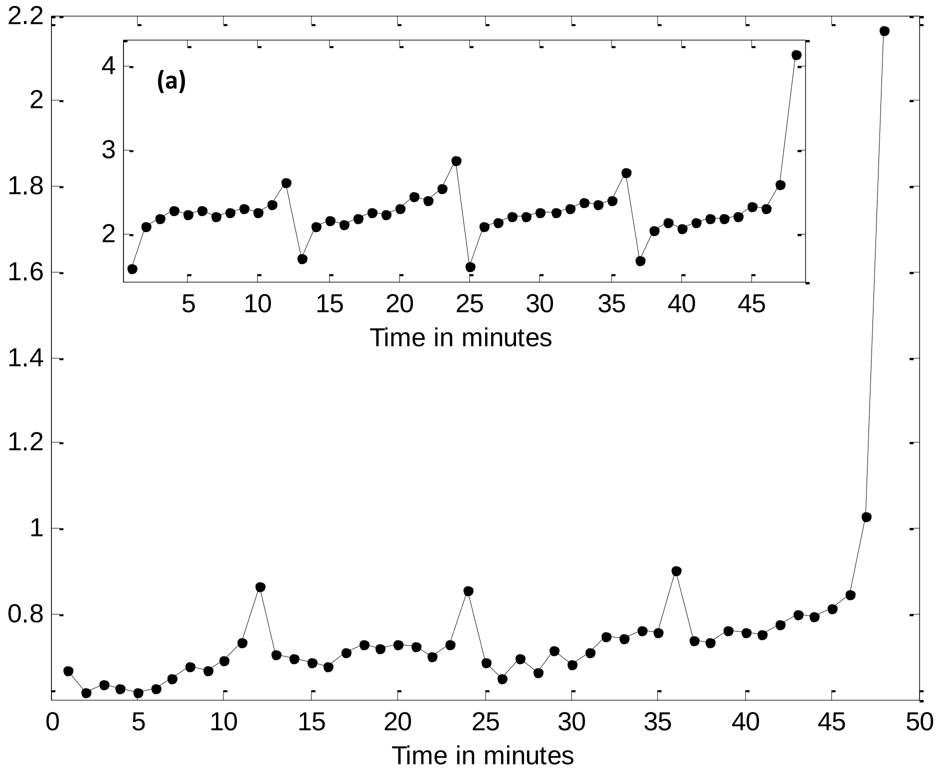


Figure 22. Index of Dispersion of the point scored by minute. We can observe that the trend of the values is to rise over time. Only at the end of each quarter there are a significantly increase, closer to 1, but only at the minute 47 reach the value 1 (pure Poisson). The minute 48 is completely out the range of the rest of the game, reaching values higher than 1. The behavior of this minute is very complex. The upper panel (a) represents the number of points in every minute (λ). This value is low at the beginning of each quarter but note that the value increases along with time, above all at the last quarter.

The most interesting results are obtained when we analyze the game time using the Index of Dispersion and the λ value because points out some very interesting behaviors. The time interval used was 1 minute because we consider that lower frequencies, such as one second, or higher frequencies such as 2 minutes were not clear enough. The profile of these graphs may vary if the selected time interval is different than a minute. But the clarity of interpretation offered by this choice, which includes all the features that we want to emphasize, was what led us to choose the time interval of 1 minute, as the key to playing a basketball game.

The Index of Dispersion (Figure 22) displays that the most part of the quarters remain lower than value 1 (under-dispersed). Only at the end of each quarter the values are higher than in the rest of the quarter. This means that the beginning of every quarter is more predictable than the end. Note that the tendency of all quarters is to rise, to approach to value 1, to become more unpredictable. But only the minute 47 present value 1 in the Index of Dispersion. To be precise, is a pure Poisson process. The game at this stage is completely random.

The minute 48 requires special attention. As we can observe (Figure 22), the minute 48 exceed the value 1 significantly (over-dispersed). This suggests that the last minute in a basketball game is a completely different process than the rest of the game, meaning the game has changed its dynamic.

The upper panel of Figure 22 shows λ (number of events per time) of every quarter. We can observe than the number of events at the beginning of each quarter is low compared to the rest of the quarter, but the tendency is to increase anyway. It is likely that this outcome is because players, as agents of the system, start to interact at the beginning of the game. There are not previous situations (no memory from previous actions, because is the first quarter). We can refer to a zero point o base point from which emerge the characteristic actions of a basketball game. That is, a self-organization problem.

Also it is remarkable the differences after halftime, maybe caused by the adjustments carried out by coaches and technical staff or by the game dynamic itself; as the last minute, where the number of events considerably higher than the rest.

The players of both teams have predefined roles (by player position) and instructions given by technical staff, based on the information about the rival. But is the interaction among them through the game time and the adaptation to the real game what make emerges game patterns: switch roles, motor task resolutions, etc. This is known as attributed role. This fact multiplies the possibilities and the actions carried out by players because players adapt to the environment.

We have to bear in mind that at the beginning of the first quarter the score is 0 for both teams. This does not happen at the rest of the quarters, where the point differences (if exist) may establish future team dynamics. It is possible that these cases are affected by memory processes depending on how big these differences are. But anyway, there are previous situations on which to base strategies.

Moreover, there is a fatigue effect of the system (errors such as fouls, turnovers, etc.) and of the players (physical fatigue, mental fatigue, cognitive fatigue, etc.) which has an accumulative effect along the game time and influences internal processes and emergent behaviors.

Hence the anomalies observed at the end of every quarter are derived by these kinds of processes or mechanisms probably. And above are accentuated in the final stages of the fourth quarter. This may be because accumulated team fouls, but have no direct significance on the score (Figure 22) until the fouling team is in the team bonus (or foul penalty) situation and free throws are awarded.

In order to a better understanding, we based the analysis on the Index of Dispersion. We have selected some cases with representative Index of Dispersion. As the Figure 22 suggests the general tendency is to increase the level of uncertainty during each quarter, above all at the end of the quarter, where the values are closer to 1. Moreover, the minute 47 reach value 1. We analyze two minutes intermediate, minutes 6 and 32; one end of quarter, minute 36 and the values of minute 47, with Index of Dispersion values 0.63, 0.75, 0.90 and 1.02 respectively.

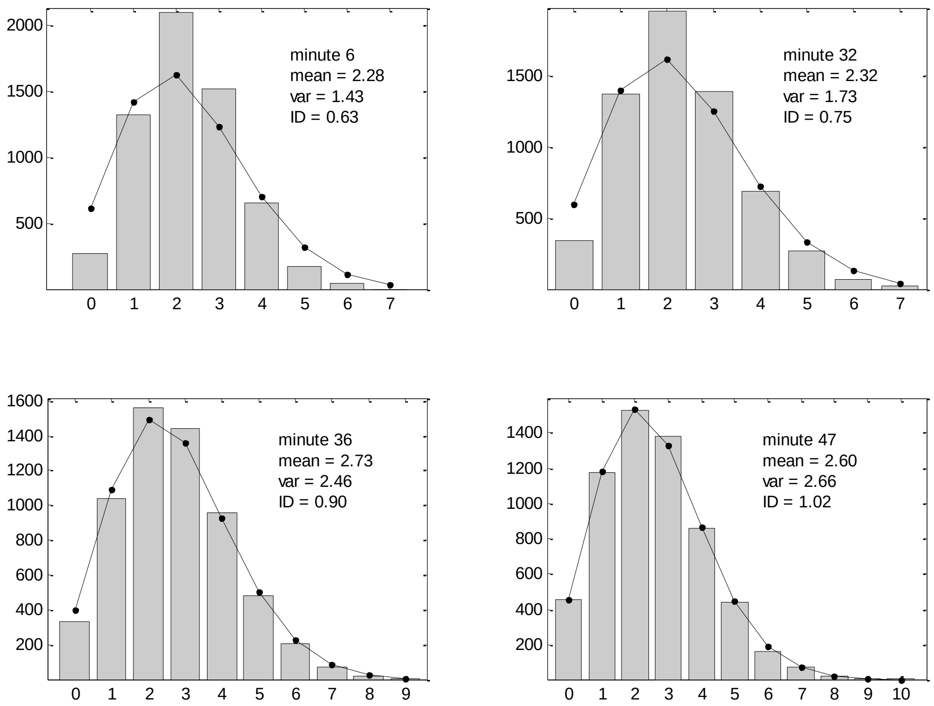


Figure 23. Histograms of the point scored in the minutes 6, 32, 36 and 47, corresponding to Index of Dispersion values 0.63, 0.75, 0.90 and 1.02. The solid line represents the Poisson theoretical distribution. The two upper cases show under-dispersion, whereas the lower cases are cases close to Index of Dispersion=1, with a Poissonian behavior.

We can observe two cases. The upper Figures correspond to the minutes 6 and 32, with Index of Dispersion values 0.63 and 0.75 respectively. We observe that does not fit well to the Poisson distribution. The variance is lower than that corresponding to the Poisson distribution (Index of Dispersion<1), and data are clustered around mean value, with less zeros and with a tail which drops quicker than Poisson, characterizing an under-dispersed Poisson distribution. In general this can means that the moments with an Index of Dispersion lower than 1 are more predictable than the rest of the game.

On the other hand, the two cases below, the end of the quarter (minute 36), fits better than the rest of the quarter (Index of Dispersion 0.90), but what really match with the theoretical Poisson distribution is the minute 47, with an Index of Dispersion 1.02. Note that the number of zeros matches better and decays as Poisson. This represents the most unpredictable moment of the game, except the last minute, which will be treated separately because the nature of the distribution is different.

The results point out that in the two upper cases the number of zeros is lower than the theoretical Poisson distribution, which is the theoretical model we use as a base. Also the tail is reduced, whereas in the two cases below fits better. This seems an indicator of the risk assumed by teams. As we see later in the Figure 34, the number of 3 points and 1 points (fouls) is higher in the end of each quarter, while 2 points are decreasing throughout the game. Teams tend to risk more at these times, and defensive intensity increases (more fouls) which indicates greater likelihood of failures, more zeros than in previous times or greater number of points (longer tail), meaning great randomness and explain its proximity to the Poisson distribution.

In the cases with less risk (two upper subplots), the game seems more predictable and the number of failures is lower; which would justify the least number of zeros and the shorter tail. And also explain its proximity to the Poisson distribution. In this thesis, one of the objectives is to compare the results against the Poisson model. This allows us to better understand the concept of risk and to separate the last minute of each quarter on a basketball game from the rest of the game; and the role of the last minute, as discussed below.

The next figure represents the point scored of the last minute of the game, minute 48, whose Index of Dispersion value is larger than 1, over-dispersed:

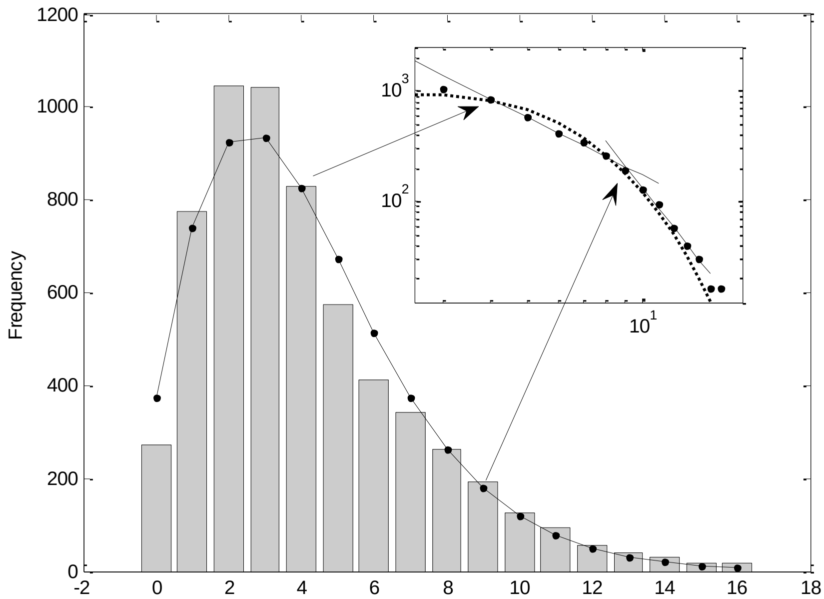


Figure 24. Histogram of point scored in the last minute of the game. The solid line represents negative binomial distribution fit (fitting parameters 3.81, 0.48; STD = 0.039). Note that the values are better fitted at the tail of the distribution. For further analysis, we carried out a log-log plot, in the upper panel, which displays two *Power Laws* with a crossover (straights lines). The dashed line in the upper panel represents the negative binomial fit.

The analysis of the last minute of the game in basketball reveals some game facts very remarkable. The last minute is over-dispersed (Index of Dispersion larger than 1; Figure 22), which is associated with a negative binomial distribution. The presence of a negative binomial distribution points out the existence of clusters of occurrences. Apparently, the point frequency distribution seems to match with the theoretical negative binomial distribution (solid line), particularly at the tail of the histogram (fitting parameters 3.81; 0.48; STD = 0.03. STD is the quadratic difference between the distribution and the data obtained).

The apparent long tail behavior gives the impression of indicating the presence of data far removed from the mean, which might indicate the presence of a truncated *Power Law*. We must take into account that there is only 1 minute of real game time.

To check this, we performed a log-log plot (upper panel) and we observe the values are fitted by two *Power Laws*. This might means that there are scaling phenomena, regarding scoring time. There is a zone 0 - 4 points, with a peak located around 2 - 3 points. But beyond this region, the first *Power Law* appears; from 4 to 9 points approximately. And a second one from 9 to 16 points with a higher slope (truncated). I.e. as the number of points scored increases the playing time is reduced dramatically. The presence of a crossover points out the presence of several scoring dynamics (multi scale behavior).

We have to take into account that there is a rule in basketball designed to provide criticality to the game. We are talking about the 24seconds rule. The aim of this rule is to force teams to shot, what, transferred to the game, provides ideal conditions for a critical situation. In sport there are a lot of examples of rules whose aim is to provide criticality to the game, such as offside in football or rugby, three touches in volley-ball, the D-zone in handball etc.

But as the Figure 22 shows, in the last minute in basketball, the own nature of the game turns critical by itself so significantly, that this rule that gives criticality to the game, makes no sense anymore.

After the examination of game time, we performed an analysis of λ values (number of points per minute) and the Index of Dispersion value of each game (6150 games).

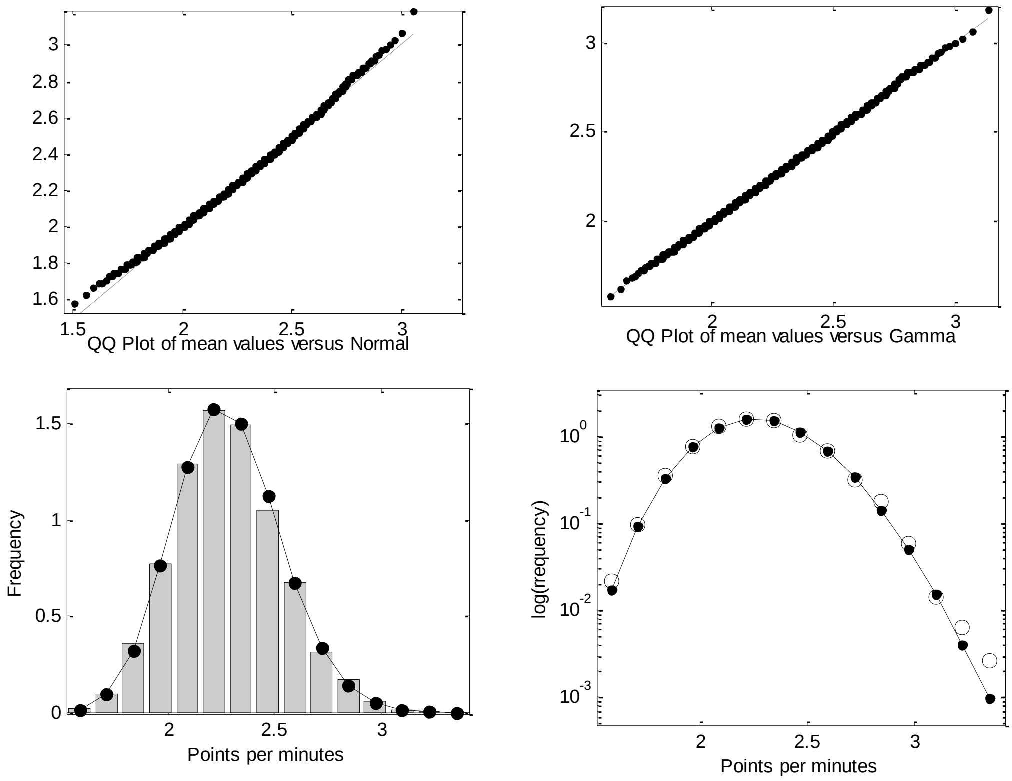


Figure 25. QQ plot of points per minute mean values vs. Normal Distribution (upper left) and QQ plot of mean values vs. Gamma Distribution (upper right). Regarding these plots, note that the data present a better fit by the gamma distribution. With Normal Distribution there are some irregularities in the tails. Down left is represented the number de events per minute with a Gamma fit; and down right the semi-log plot of the previous Figure.

The Figure 25 represents the λ values. The histogram (down left) seems to be almost a Normal Distribution, but fits better with a Gamma Distribution (upper plots). We performed a visual test (QQ-plot) with several distributions: Normal Distribution (QQ-plot up left), Exponential Distribution, Weibull Distribution, etc. but the data fits better by a Gamma Distribution (up right). The statistical values of the λ distribution are: mean = 2.280; STD= 0.251; variance = 0.063; skewness = 0.284 and Kurtosis = 3.174; For the Gamma distribution: Shape parameter k = 82.73 and Scale Parameter θ = 0.027.

The Gamma Distribution is an accurate distribution for modeling the behavior of continuous random variables with positively skewed; i.e. variables that present a greater density of events to the left of the mean than to the right.

In our case we can observe that the number of points per game do not follow a Normal Distribution but is skewed to the right, meaning that there are more probabilities to score more points than the mean (more than 2.28 points; up 3.3 points per minute), although this probability is low compared to the rest, it can take place.

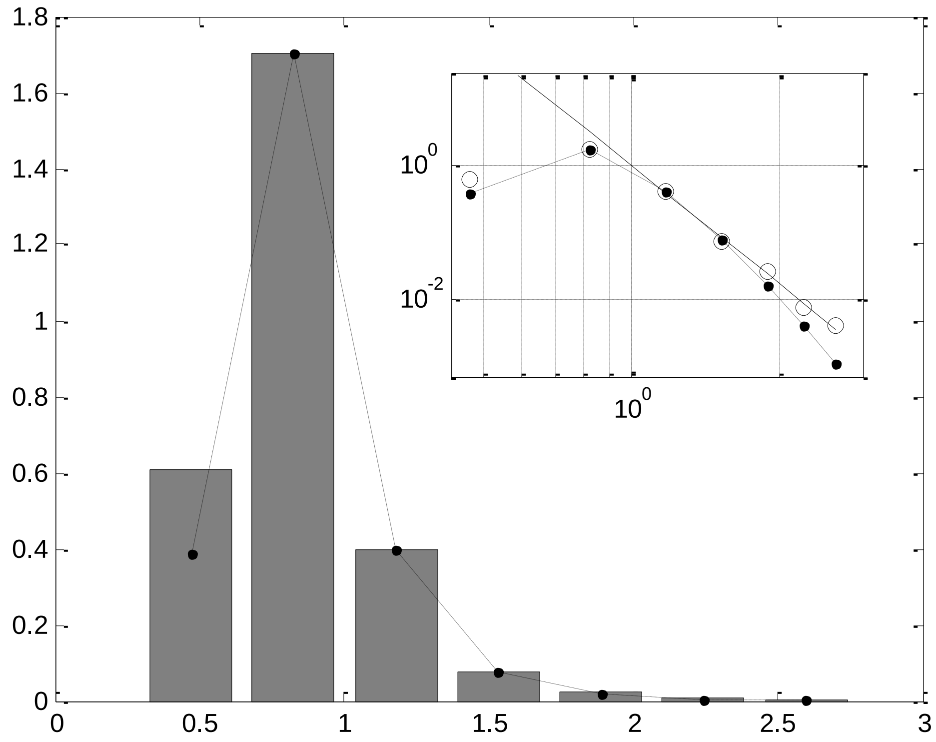


Figure 26. Histogram of Index of Dispersion per game. The dash line represents the Generalized Extreme Value (GEV) Distribution. It seems that fits well, but when we perform the log-log plot (upper panel) we note that the tail in not well fitted ((o) represents the real values, (•) represents the GEV values. We can observe that the most part of the games are located around 1, which means they are very unpredictable. But even we can find some games with values larger than 1 (over-dispersed).

The Figure 26 shows the Histogram of Index of Dispersion per game, the most part of the values are located close to 1, which points out a Poisson process. This means that the number of points scored in the most part of the games follows a Poisson process. But moreover there are some cases where the distribution is over-dispersed. This indicates the presence of extreme events. In order to check this we fitted by a Generalized Extreme Value distribution (GEV).

The dash line represents the Generalized Extreme Value (GEV) Distribution with shape parameter 0.0680; scale parameter 0.1786 and location parameter 0.7150; and seems to fit well except at the first values. To find out whether the data follows a GEV in the tail, we carried out a log-log plot (Figure 26 upper panel) with the theoretical GEV distribution (dash line against real values (o)). Note that the data do not fit well. On the other hand, seems to fit better with a *Power Law* (straight line: -5.82; -0.014), although the slope is big, drops quickly with a scaling behavior.

Therefore we can point out three different game profiles according the Index of Dispersion: between 0 and 0.85 there 3855 games (62.88%); between 0.85 and 1,15 there are 1714 games (27.96%) and higher than 1,15 there are 562 games (9.17%).

According to this results, we use these three game profiles in order to examine some relevant aspects of the game, such as the last minute (minute 48), the final score difference and the maximal difference reached along game time.

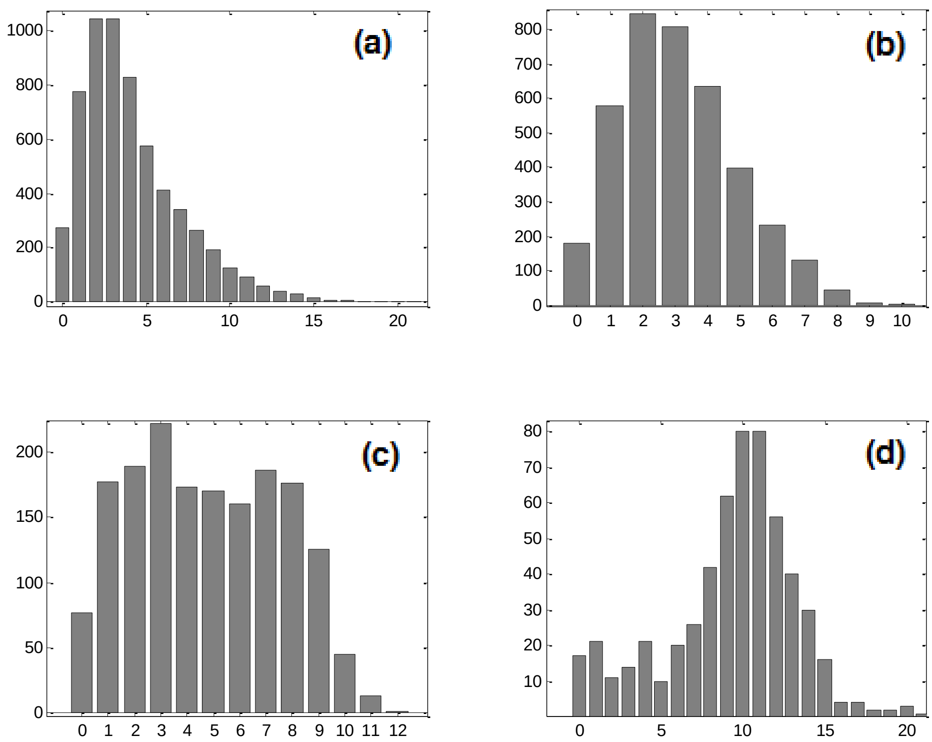


Figure 27. The subplot (a) represents the number of points in the last minute of the game (same as Figure 34). The subplot (b) contains all data lower than 0.85 of Index of Dispersion; (c) between 0.85 and 1.15; and (d) higher than 1.1 5 of Index of Dispersion. The case (b) is under-dispersed (Index of Dispersion <1). There is a frequency peak around 2 or 3 points. The case (c) is the most Poissonian because the Index of Dispersion is closer to 1. Note that the distribution is almost homogeneous from 1 to 9. The case (d) is over-dispersed (Index of Dispersion larger than 1), the peak is located around 11 and 12 points, but we can find games with up 21 points.

The Figure 27 displays in detail the number of points in the last minute of the game. The first subplot (Figure 27 (a)) represents the histogram of all data collected. The second subplot, Figure 27 (b) shows the histogram of the data lower 0.85 Index of Dispersion; under-dispersed.

This case present a peak around 2 or 3 points, which match with a normal rhythm of scoring. Even though there are games with score runs up 10 points, but are rare situations. The case (c) in Figure 27 displays the histogram for the case next to 1 of Index of Dispersion. Between 0.85 and 1.15 to be precise, therefore is more Poissonian than the rest of cases. It is remarkable that the distribution is more homogeneous between 1 and 9 points. That confirms that we cannot predict the outcome at this stage because probabilities are similar, hence score runs are more random. The last subplot (Figure 27 (d)) is the over-dispersed case, with Index of Dispersion values higher than 1.15. The distribution looks like the reverse of case (b) with the peak next to values 10 – 11 points, which is significantly higher than case (d). In fact this value of 10 – 11 points match with the end of the distribution (b), and even its maximum value is the double: 21points. This game profile can be a game with a lot of points at the end, meaning a lot of fouls, free throws, violations (turnovers), time outs to planning the last shot, etc.

The Index of Dispersion regarding score differences is:

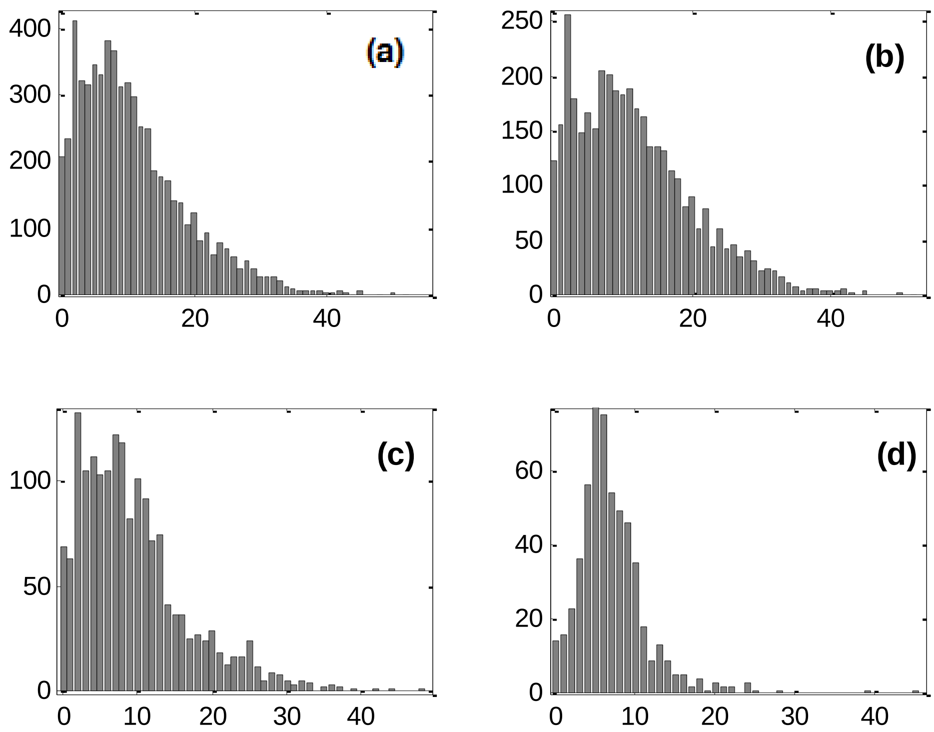


Figure 28. Final score differences in the game. X-axis is expressed in points. The subplot (a) represents the value for the differences of the entire sample analyzed. The subplot (b) shows data lower than 0.85 of Index of Dispersion; (c) between 0.85 and 1.15; and (d) higher than 1.15 of Index of Dispersion. The case (b) is under-dispersed (Index of Dispersion <1). It is almost homogeneous until 12 points, then drops. In the case (c) the data are more clustered form2 to 10 points and there are clear differences with the rest. The case (d) is over-dispersed (Index of Dispersion larger than 1), the peak is located between 3 and 10 points. It is very narrow which means that the differences are very located.

When the Index of Dispersion regarding score differences (Figure 28) we note some remarkable facts. The subplot (a) is the histogram of the final score differences of the entire sample analyzed. But we discuss this later. The subplot (b) represents the values of final score differences for games with an Index of Dispersion lower than 0,85 (under-dispersed). Note that the distribution is not uniform. There is a main region where the most part of the data are located, around 0 – 12 points. But it is a bimodal distribution. The first peak is for 3 point difference and the second on is for 12 points approximately. Further than this point the distribution decays. This case is the most predictable compared to the rest (still unpredictable anyway). The 3 points differences are significant. But the rest of differences are also numerous, including more than 10 points. Differences higher than 20 also can take place.

We can deduce that the game at this case, when the Index of Dispersion is lower than 0.85 the score runs are more regular. This means that if both teams are scoring often, the final difference will be small. Perhaps this is the reason of the high values of 3 points meaning that the game ended by a three point shot at the last second. But at the same time if one of the teams present a scoring time intervals shorter than the other team (one team score more often that the other), the final difference will high.

In the case (c) note that the data are more clustered form2 to 10 points and there are clear differences with the rest. This is the most Poissonian case because the Index of Dispersion is next to 1. In (d) the peak is located between 3 and 10 points and the distribution is very narrow. In this case the last minutes of the game are more fractioned: more fouls, more free throws, more time outs, etc. hence these are the most competed games.

The same approach for the maximal score differences:

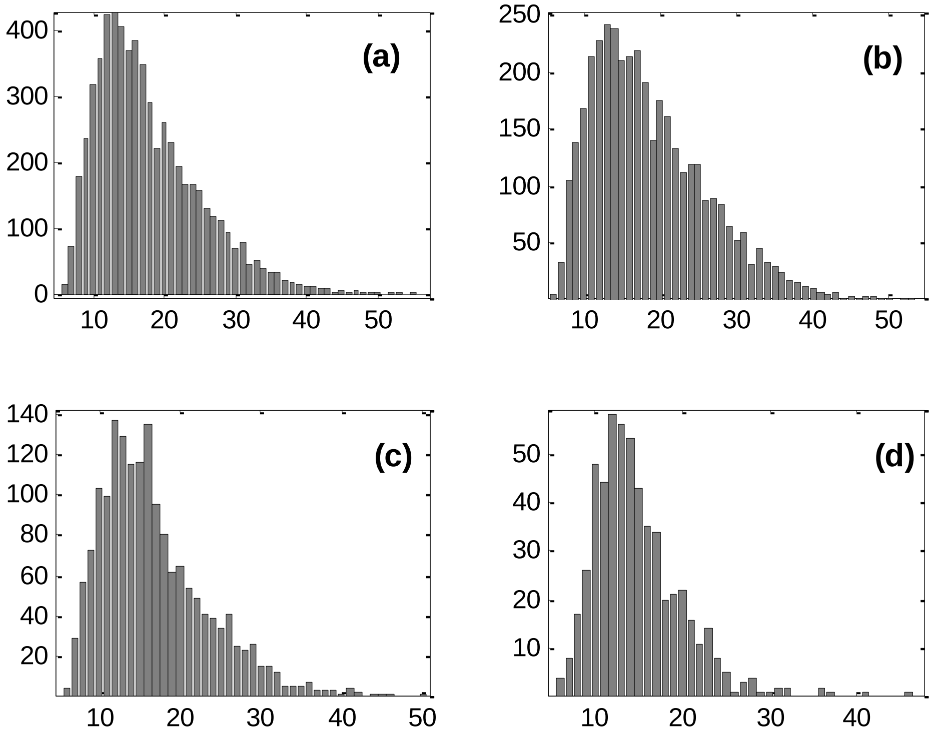


Figure 29. Maximal score differences. (a) The total distribution. (b) The distribution for 0.85 Index of Dispersion. (c) Values between 0.85 and 1.15; and (d) higher than 1.15 of Index of Dispersion.

Now we present the same line of attack for the maximal score difference reached in the game (Figure 29). We can observe that for the subplot (a), the entire sample, the peak is located in 15 points, but higher differences can also take place. When the Index of Dispersion is lower than 0.85 (subplot (b)) the distribution is quite similar to the previous. For the case (c), with an Index of Dispersion close to 1, the data are more clustered around 11-15 points and the distribution drops more quickly than the previous case. And once again we can note that the last case is the more narrow, meaning that these differences have been reached at the last minute, probably.

Time in basketball

Following the methodology proposed in this studio, using Poisson model as a basis for analysis, we studied the final property of Poisson process indicated above: Scoring time intervals follow an exponential distribution, which only depends on a single parameter and also it is easily detectable in a semi-log plot.

After the analysis using the Index of Dispersion and the λ value, we carried out an approach to using the time interval between points distribution. The Figure 30 represents the point interval time distribution (two point, three points and the first free throw, we must bear in mind that on the second free throw, time still stopped).

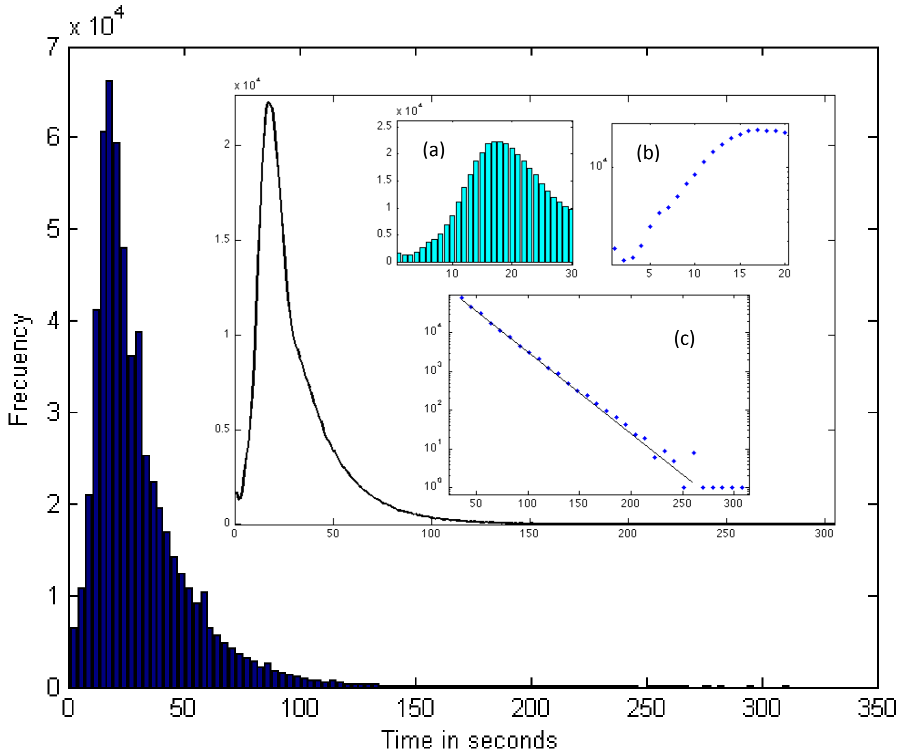


Figure 30. Histogram of field goals time intervals (field goals scored). We can observe that the distribution is not homogeneous. There is a peak around 20 seconds, and long tail behavior is also present. In the sub plot are represented the distribution smoothing, the distribution until 30 seconds (a), the first 20 seconds (b) where we can observe that the first value frequency is greater than the second one and the third, and the semi-log plot beyond 30 seconds (c), where the data distribution fit to a straight line, but the last data (long tail).

At the X-axis are represented the time intervals in seconds, in which the points are produced. At the Y-axis are represented the frequency of the intervals. The first thing that stands out is that the data show a not symmetric behavior. The distribution has a tail long tail apparently with a maximum value of 310 seconds and a frequency peak around the value of 20 seconds.

The first values are not progressive as we can think in advance. In fact points with one second difference are more numerous than the two and three seconds (Figure 30 (a), (b)). And even the slope varies around 6 seconds, until reach the peak, located around 20 seconds (Figure 30, (a), (b)). This can be related with fast breaks (until 6 seconds) and regular rhythm play (until 24 seconds). If we take semi logarithmic of time interval beyond 30 seconds (after the curve-peak), we obtain Figure 30 (c). On the X axis is represented the time differences and the Y-axis logarithmic scale. We note that the data distribution becomes in a straight line, which means that this timeframe points out the presence of an exponential distribution, indicating a Poissonian behavior but only in this region (>30 seconds).

The behavior of the tail is best seen by taking logarithms: the next Figure (Figure 31) shows the point time intervals (X-axis) and the logarithm of the frequencies (Y-axis). The time intervals follow an exponential distribution beyond 24 seconds, approximately (see also Figure 30 (c)). It becomes in a straight line (semi-log plot). The upper panel of Figure 31 represents the log-log plot of the data set from 100 seconds to the end.

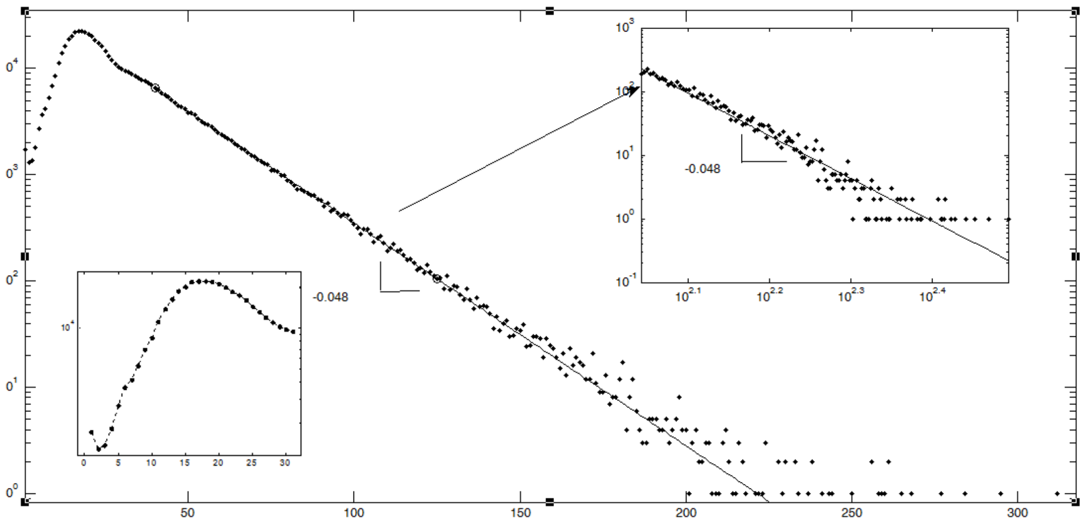


Figure 31. Semi-log plot of point time interval frequency. Apparently, there are three different behaviors. The data follow a distribution with a maximum (peak) around 30 seconds. Above 30 seconds follows an exponential distribution. To further analyze the behavior of the distribution tail (from 100 seconds), we also carried out a log-log plot (upper panel) to verify whether this trend is approaching a *Power Law* behavior.

In the Figure 31 we can observe in detail the entire data by taking semi-log, for a better understand. We can see how the data do not follow a unique distribution. The data distribution fit to a straight line after the peak region (more than 24 seconds approximately) until 100 seconds. Further than this point the data are more scattered, and are better fitted by a log-log plot (Figure 31 upper panel).

Therefore, we can say that there is no single behavior regarding point time intervals. The distribution is complex. The values correspond to different game moments and game situations. A key problem in a basketball game is the different behaviors according to game time. This is important to take into account when we want to model or to predict in order to a better understanding.

Our results point out that we can consider three zones at least. The first one extends from the beginning (0 seconds) up approximately 24 seconds. This area has a bell-shaped distribution, with a maximum around 20 seconds, with an irregular behaviour before 6 seconds and truncated beyond 24 seconds. The Figures 30 and 31 show us that in a basketball game, the most likely time between goals is around 20 seconds (this seems logical considering the 24 seconds of possession). The behaviour of the time intervals in this area deals with a kind of rhythm of play, ball possession, and scoring related with 24 seconds of possession, but below 6 seconds, presents some particularities (see Figures 30 and 31) as we mentioned above:

The most part of time differences of one or two seconds are produced at minute 48. Points with one second difference only in the minute 48. The source of this particularity can be the free throws scored due to the high number of fouls made at the end of the game (for further detail see the foul section); but also because time outs called and strategies to score quickly. This can be the foundation of the shape of the distribution at the beginning as we can see in Figure 30 (b). The same Figure displays that the slope varies in second 6, approximately. Hence we can deduce that the first array corresponds with fast breaks and the second slope with plays more elaborated (from 6 to 20 seconds). In fact, the most common is to score around 20 seconds.

This may be related with rebounds: defensive rebounds, because allow building fast breaks quickly, and offensive rebounds because allow scoring quickly with a high success rate, and further elaborating successive attacks. Hence, the strategies of many teams are to make fouls to avoid these situations, which create serious disadvantages between a team and the other.

The second area, from 24 seconds up 100 seconds approximately, shows a decrease in a straight line (Figures 30 (c) and 31), which corresponds to an exponential distribution, as we saw in the case III), suggests that the distribution follows a Poisson process, i.e., completely random, without memory, for time intervals larger than 24 seconds.

This is an interesting result because if it is a Poisson phenomenon, it could have a feature called *memorylessness* (also called evolution without after-effects): the number of goals occurring in any bounded interval of time after time *t* is independent of the number of goals occurring before time *t*. It means that the time in which each point is scored is independent of the previous. The score becomes more random.

Beyond 100 seconds approximately, as third region, we can see how the data are scattered and the final biased. It can be considered as rare phenomena (low probability). For values over 100 seconds is possible that it behaves as a *Power law*. To verify the actual behavior of the data, we performed a log-log plot (upper panel Figure 31). The results point out that the data fits better by a log-log plot, are less scattered, which means that can be considered a *Power Law*.

From an overview we can observe that when the point time intervals reach high values, behaves as a *Power Law*; therefore we can say that the events data has memory, whereas when it behaves as Poisson distribution does not present such feature. This is important because can be related with the scoring differences between the winning team and losing team, and how the establish their strategies for each situation. Winning team tends to waste time, to waste possessions (longer time intervals). While the losing team, tends to the opposite. The strategy of long time intervals is typically of teams that want to play to low scores, where behaves as a *Power Law*, thus they are more likely to win the game because the systems is more critical (SOC).

We are facing two diametrically opposed tendencies. This, as we have seen, can present a directly influence in the game. Moreover, if we extrapolate this to a higher level, to a league for instance, it can provide us some clue about how is the internal dynamic of the league: the fact that exists different dynamics well defined regarding point time intervals reflects stages of the game, profiles games or tactics employed by teams, which shows the sport reality in basketball. A league where there are large differences between their teams, the score differences in their games will be more pronounced, because theoretically weaker teams tend to employ defensive strategies through lower scores in order to increase their chances of victory. While on the other hand, a more balanced league, where the differences between its components are not so marked, this trend will not be so pronounced.

All team sports are based on fan attendance, particularly basketball, which tries to provide exciting scores. Hence the number of points per minute is as determinant factor when we analyze basketball (Figure 999).

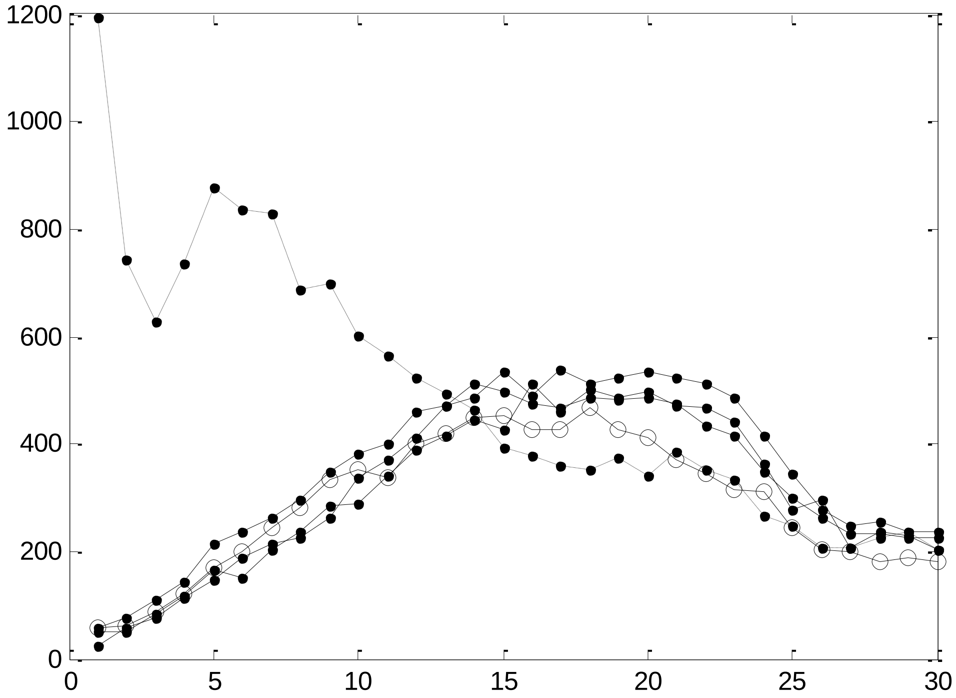


Figure 32. Number of point scored (Y-axis) with a time difference of 1 seconds, 2 seconds, 3 seconds,..., up 30 seconds (X-axis). The dash line represents points scored in the last minute of the game, where it is clear that most points with one second difference is given in this period. The other values correspond to the last minute of the remaining quarters (solid line (•)) and to the y al minute 47 (o), whose behavior was similar**.**

When we study the number of points corresponding to each time intervals in the last minute for each quarter (Figure 32), meaning number of points scored with one second difference between them, two second difference and so on until 30 seconds, we are we note that there are some significant differences. It seems that the last minute in the three first quarters solid line (•), behaves similar. The values increase up 14 seconds; then remain stable until 23 seconds. It can have sense by itself, in the sense of this the most probably time between points in these periods, and probably is related with the 24 seconds of possession. Beyond 23 seconds drops up 27 seconds. Note that from 27 to 30 the frequency presents similar values. This stability points out a slight tendency to extend the time between points at the end of every quarter. The minute 47, solid line (o), follow a similar dynamic as the previous cases. The frequency increases up 14 seconds and stabilizes until 18 seconds. After that, the frequency drops up 26 seconds where presents the same frequency values until 30 seconds.

The last minute of the game, dash line, minute 48 is completely different to the rest of the last minutes of previous quarters. The highest frequency value is for points with 1 second differences with a significant difference from other quarters. The frequency falls up 3 seconds, but increase again until 5 seconds and is still high compared to the rest of sample. The array form 5 to 8 seconds is the more stable region of this minute. But beyond this area, the frequency declines until the end. Even the range from 14 to 23 seconds is lower than the rest of quarters as we can see in Figure 32. It seems that the rule of 24 seconds make no sense here. The short intervals are numerous than the large.

For the entire game time the shots with 1 second difference are: 1 point shot 1525 (89%); 2 point shots 137 (8%) and 3 point shots 48 (3%). The source of this tendency is the fouls and free throws, probably. In fact, if we analyze the minute 48 we observe that the shots with 1 second difference are: 1 point shot 1128 (94.55%); 2 point shots 48 (4.02%) and 3 point shots 16 (1.34%). we realize that the issue of fouls is accentuated. Therefore it is interesting to check out what were the final score when there were shots with one second difference. Histogram of final score differences when there is some 1 second scoring time interval in the last quarter:

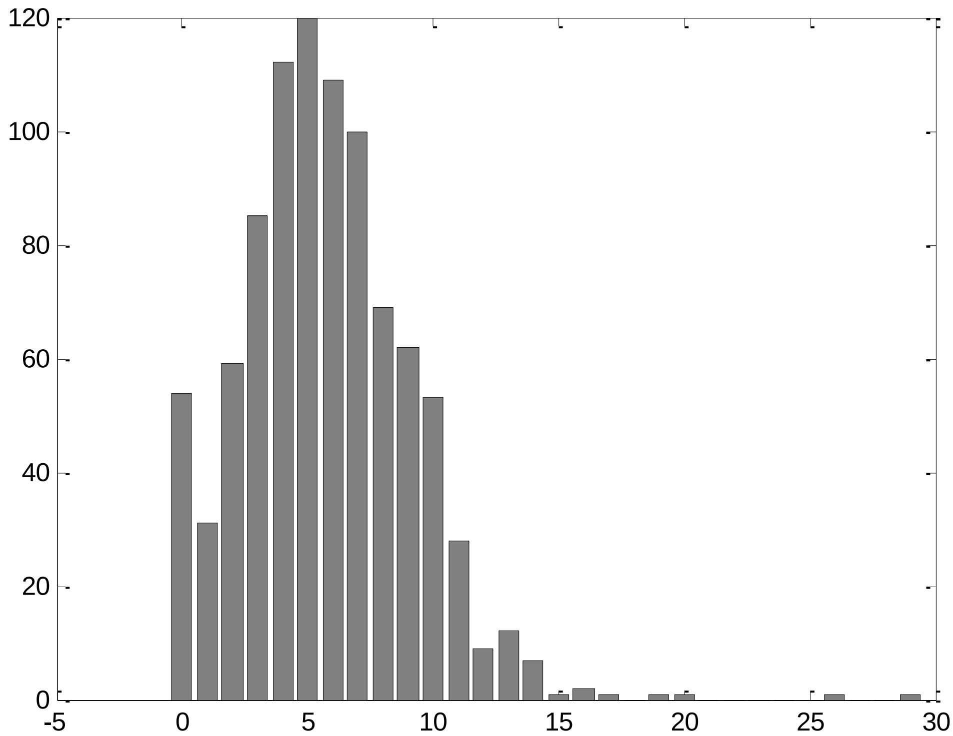


Figure 33. Final score differences of games with 1 second difference in the last quarter. The data are clustered from 0 to 10 points approximately. More than 10 points is a rare event. The peak value is 5 points.

We note that the most likely score difference is between 0 (we did not count overtime) and 10 points. But we can find differences up 29 points. The 5.88% finish in a tie (0 points; with over time). A total of 526 games of 918 (58%) finish with a difference between 3 and 7 points. And the 93% finish with less than 11 points

From an overview, there were 690 games with at least a case of points with 1 second difference. In 184 games there were two cases. In 41 there were three cases and in four games there were only three cases.

Scoring

Regarding to score, the absolute value of score (result) always grows along with game, but do not evolve uniformly. This is a reality which is maintained on all basketball games. Score runs and maximum values achieved by the teams may vary, but always does incrementally. But what that really sets the dynamics of the game is the point differences between a team and another during the game time and, above all, at the end of the game.

For that reason, we analyzed the differences on the final score of the whole sample analysis (6150 NBA games). The result (Figure 34) point out that most of the games (65%) ended with a difference between 1 and 11 points, 33% had a difference between 11 and 28 points, and only 2% did so with a difference of 28 or more points. To verify whether the data followed a *Power Law* type distribution, we performed a log-log plot whose result can be seen in the upper panel of Figure 34.

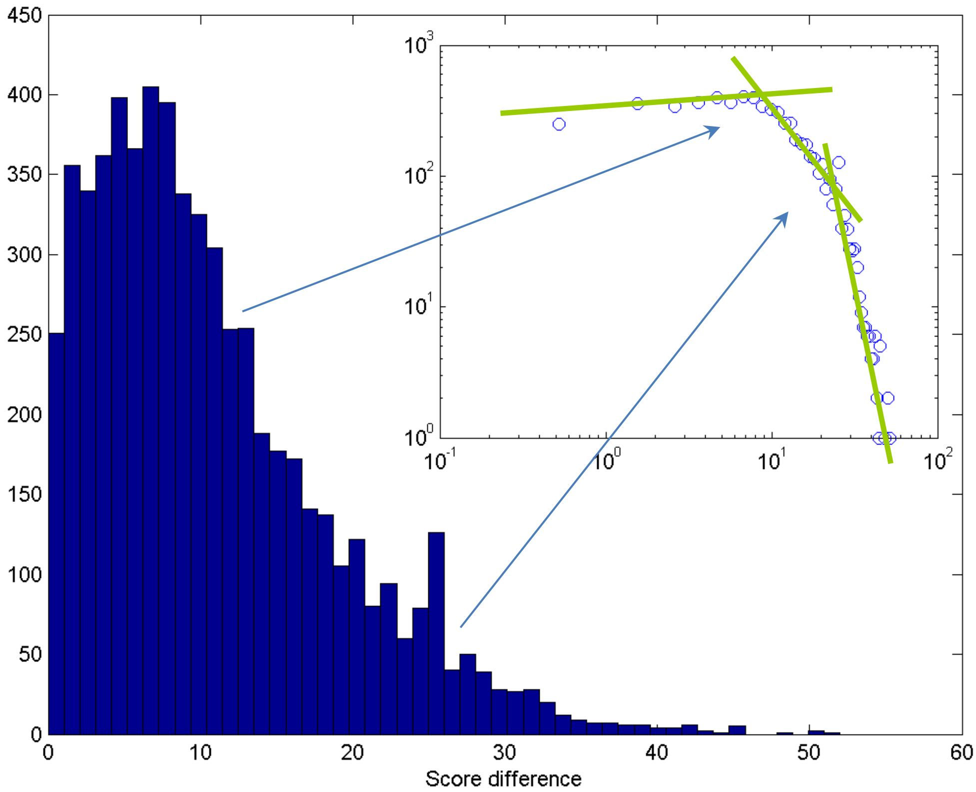


Figure 34. Point difference histogram existing in the final score of each game studied. The distribution is approximately uniform for values less than 10-12. Further than this value the distribution shows a possible behavior of long tail. Log-log plot of data point difference and frequency. We can see that the first array present a homogeneous tendency. Around the value of 10 points, an interruption in this trend takes place; and a second one at a value around 25-28 points. This suggests the presence of more than one *Power Law*.

0-10 score difference

From 1 point to 10 points approximately, the distribution is almost uniform, which corresponds with situations of high uncertainty. If we exceed this score, from 10 to 28 points, the behavior appears to follow a *Power Law*. This indicates that the nature of the game has changed. Finally, over 28 points (a second *Power Law*), the essence of the game changes radically, and the final outcome is more predictable. In brief, results between 0 and 10 points are similar: the game is hard-fought. This points out that as long as the game remains between these values, the final result is unpredictable. This dynamic suggests that this area of point difference (0-11 points) works as an attractor, because the system (the game) tries to remain within this narrow area throughout the game time. In fact, in around 20% (about 1174) of the games assessed, teams did not exceed the maximum score difference of 11 points for the entirety of the game.

Regarding the final result, the number of games that finished with a point difference lower or equal to 11 points was 3846 games (62% of the games analyzed), which corresponds with the first cut in the log-log (Figure 34). 2324 of these games reached a maximum point difference between 11 and 20 points, and 43% did it during the last quarter of the game. In 578 games (almost 25% of the 2324 games), a team was able to overcome the difference (between 11-20 points) and win the game. This means that there were teams able to overcome a significant difference (between 11-20 points) and even win the game. It is possible that by achieving good score runs, the game is able to reach the critical area, and, joined with strategy at the end of the game (final quarter, fouls, free throws, time-outs, etc.), the combination needed for that team to win the game can be achieved.

1831 games reached a maximum point difference of more than 20 points. In 348 games (20%), this situation was overcome and the game was located in the area of a 0-11 point difference. 34 games (9.7%) of these cases won the game. In these games, the maximum difference was reached between the 9-minute and 44-minute mark. The results point out that it is indeed very difficult to overcome a 20-point difference in the last 4 minutes of the game.

11-28 score difference

Out of the 2285 games with a score difference higher than 11 points, only 27 were able to win and overcome the differences between 12 and 22 points. Note that in this case, these differences were reached before the 24th minute (most of them in the second half), except a case in which it was done in the 33rd minute, but in that case the team was losing by only 15 points. Hence, beyond 10 points the dynamic is completely different and is more predictable.

Larger than 28 score difference

Neither team was able to either overcome the difference or achieve the region of an 11-point difference. This means that if there is more than a 28-point difference, then there is a clear superiority of one team over another, so much so that the game is quite predictable.

We must remember that there is not a fixed criterion to identify non-linear complex systems or self-organized criticality behaviors in sports. But whether a *Power Law* appears, it is possible we are dealing with a non-linear complex system (Savaglio & Carbone, 2000; García Manso et al., 2008). The log-log plot of the distribution of point difference is broken into several *Power Laws*; for certain characteristic values that can be considered thresholds or critical points, which means that game dynamic, can be characterized by several critical phenomena, or with several scales (multi-scale). The presence of crossovers in *Power Laws* is an indicator of changes in the underlying dynamic and suggests that perhaps we are dealing with a phase transition and critical exponents (McGarry et al., 2002; Scheffer et al., 2009).

Returning to the general distribution of point differences, it also can be modeled by a Negative Binomial (1.94; 0.15), as we can see in the next Figure:

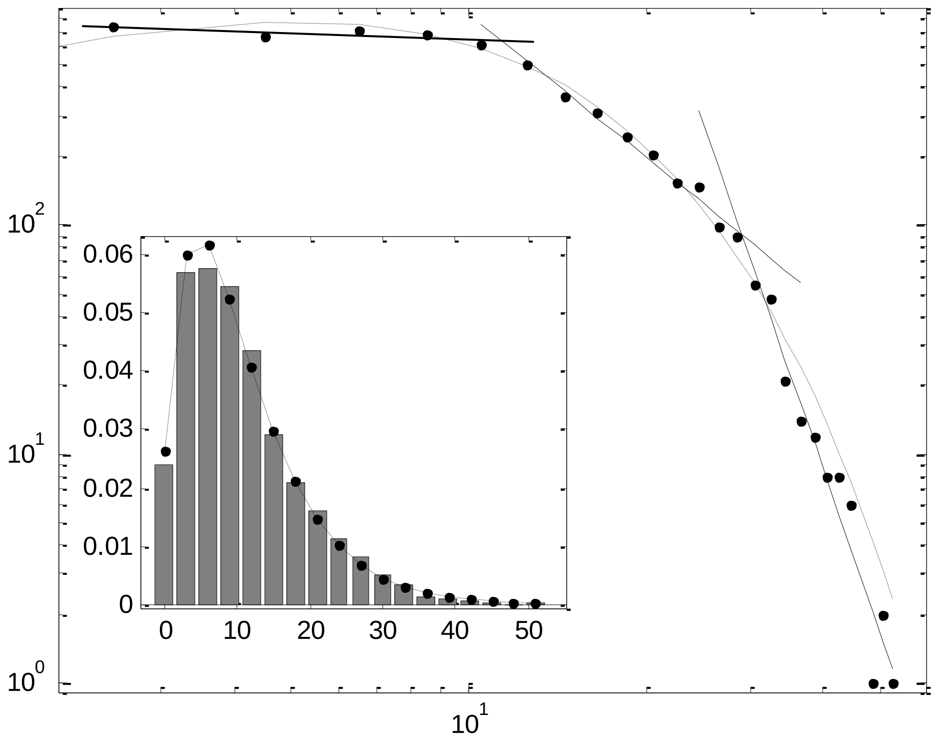


Figure 35. Histogram of the point differences with a dash line as theoretical negative binomial fit. The main figure represents the log-log plot of the same data, also with a negative binomial distribution (dash line). Note that the histogram does not fit well by the negative binomial distribution at the beginning but it does at the tail. The log-log plot displays two cross over. The first one around 10 points and the second one around 28 points.

We can observe that the histogram of the point differences is relatively well fitted by a negative binomial distribution (Figure 35) especially at the end. Next it is shown the end of games in detail. If the point difference is less or equal to 10 points the distribution of the points scored at the last minute of the game follows a negative binomial distribution (Figure 36).

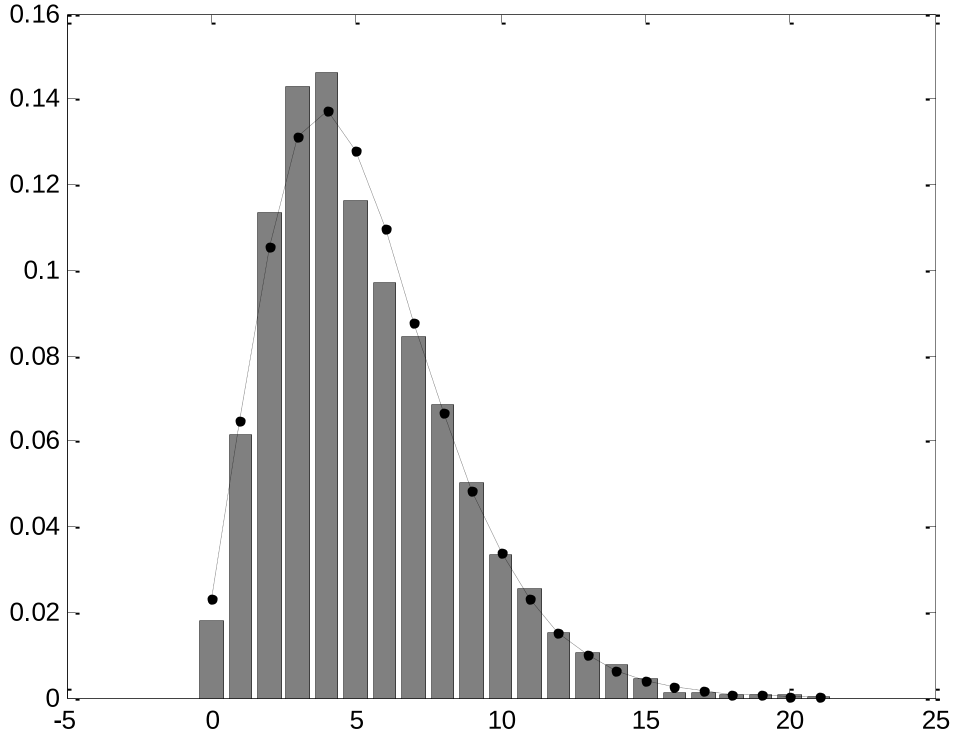


Figure 36. Histogram for the points scored at the last minute for games ended by 10 or less points difference. The dash line represents the theoretical negative binomial distribution.

The histogram points out that the most likely, for games ended by 10 or less points difference, is to score 3 or 4 points in the last minute. But the distribution seems to fit well to a negative binomial distribution with parameters 5.9301 and 0.5306; mean = 5.2471; median = 5.0000; variance = 9.9922 and STD = 3.1610. Note that there is a tail, but the most part of the points scored at the last minute are located bellow 10 points which point toward a great competitiveness. The distribution for shots scored was 1 point shots = 12008; 2 points shots = 4954 and 3 points shots = 1772.

For the case of games ended with a difference between 11 and 28 points, the histogram is:

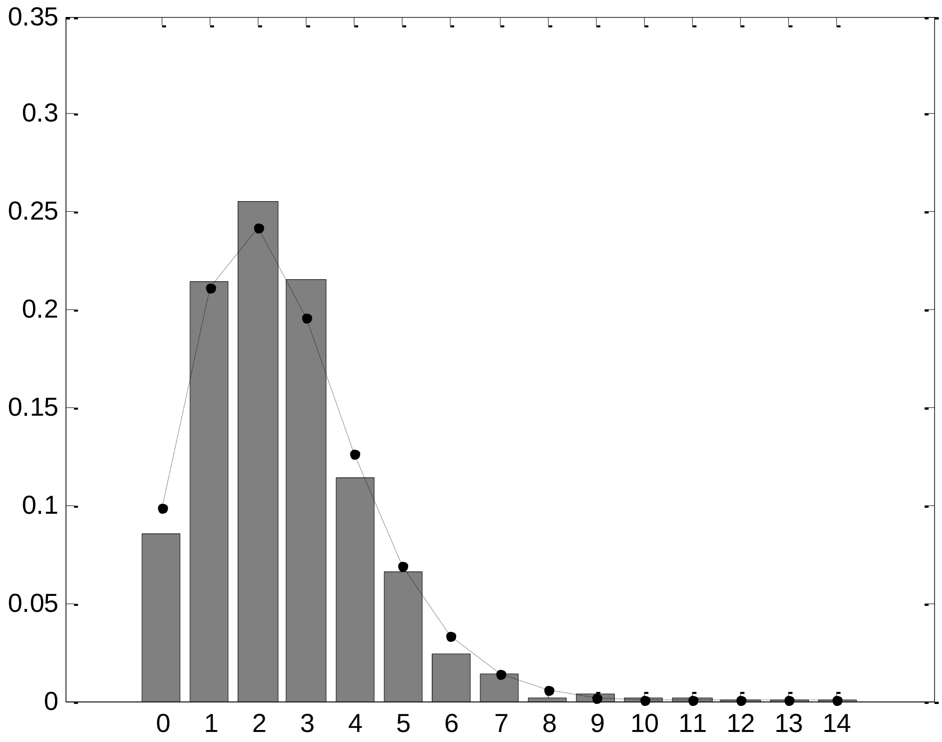


Figure 37. Histogram for points scored in the last minute for games with a difference between 10 and 28 points. The dash line represents the theoretical negative binomial distribution.

The Figure 37 represents the histogram of games ended with a point difference from 10 to 28 points. We can observe that the highest probability is for 1, 2 and 3 points and presents a tail as well. The statistical values are: mean = 2.5049; median = 2.0000; variance = 3.0034 and STD = 1.7330

The Index of Dispersion for these data was 1.19 which points out an over-dispersed Poisson distribution. The dash line represents the negative binomial distribution with parameters 14.9755; 0.8567. The distribution for shots scored was 1 point shots = 1732; 2 points shots = 1693and 3 points shots = 517. Note that in this case the number of points is more clustered than the previous case. And the number of shots scored is lower as well.

For the last case, the case for games ended with a point difference higher than 28 points (Figure 38), the statistic was: mean = 2.3212; median = 2.0000; variance = 2.1241 and STD = 1.4574. The Index of Dispersion is 0.91 (lower than 1. Under-dispersed), ergo it is not negative binomial, is more Poissonian.

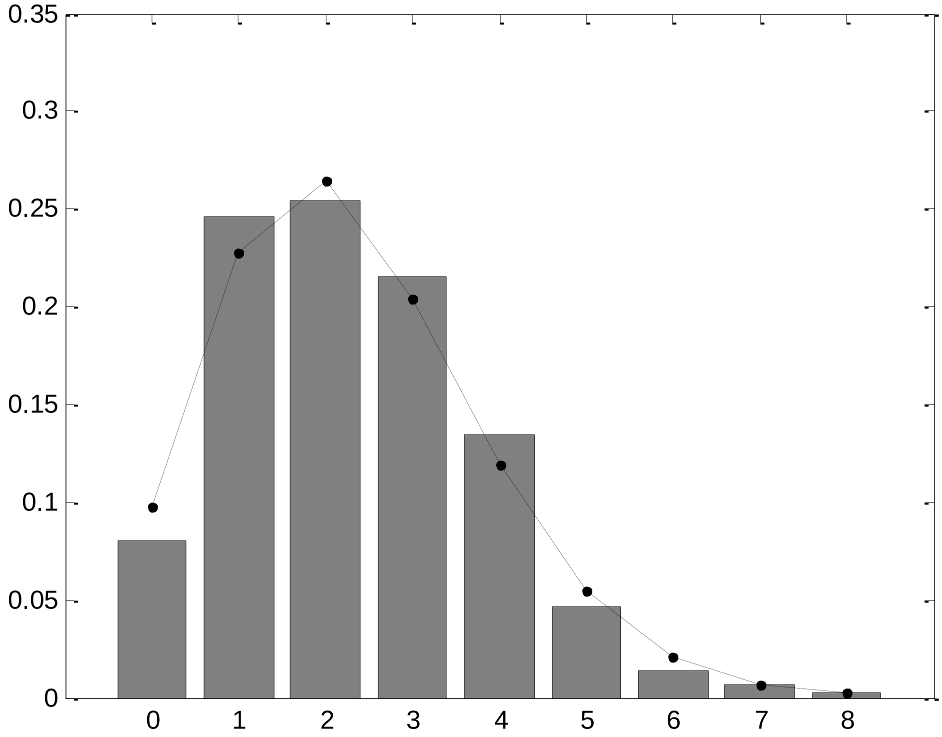


Figure 38. Histogram of the point difference for games ended by more than 28 points. The dash line represents a Poisson theoretical distribution.

Note that in this case the dash line symbolizes a Poisson theoretical distribution with L = 2.32; the distribution of the shots scored was1 point shots = 626; 2 points shots = 889 and 3 points shots = 235. Notice how the percentages now resemble the overall average. And the number of points is small. We can say that in this case the game does not change their behavior in the last minute.

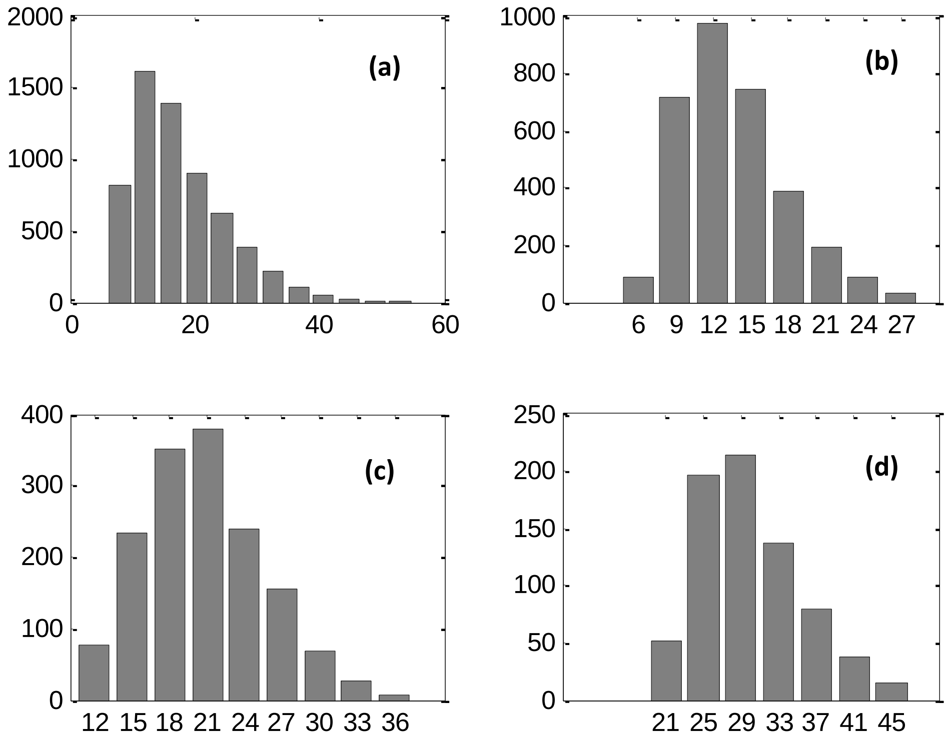


Figure 39. Histograms of maximum point differences throughout the game (a). Histograms for cases that end with less than 11 points (b), between 11 and 28 points (c) and for more than 28 points (d).

The Figure 39 represents the histograms of maximum point differences throughout the game: The entire sample (a), for cases that end with less than 11 points (b), between 11 and 28 points (c) and for more than 28 points (d). The case for differences lower than 11 points, the maximal differences are between 9 and 18 points. The subplot (c), the maximal differences are between 15 and 27 points. And the case (d) the maximal differences are more than 25 points, implying the superiority of one team over another and when they play the last minute, there is not change; they do not fight for winning. Clearly, the *Power Law* discriminates these cases.

Substitutions

The number of substitutions per quarter is 44980 first quarter, 75494 second quarter, 46335 third quarter and 68059 fourth quarter. Contrary to what we might think, the second quarter presents the highest number of substitution throughout game time. Graphically:

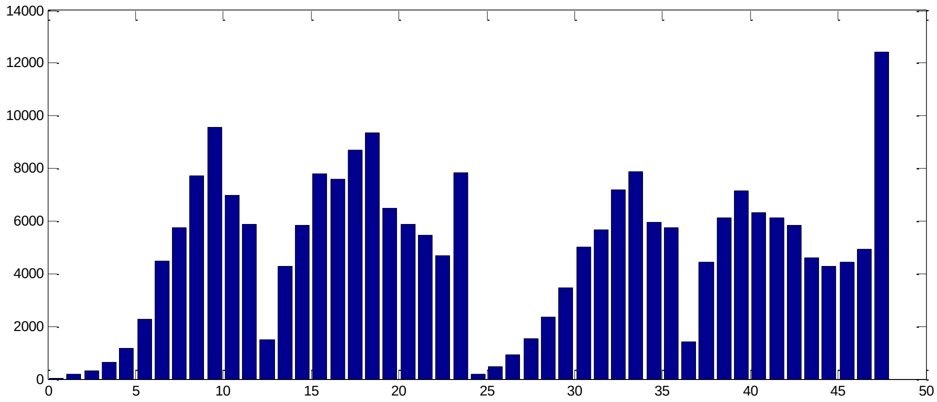


Figure 40. Histogram of substitution per minute.

The histogram of substitutions shows us that the number substitutions are more numerous at in the second quarter. The statistics are: mean= 0.7981; variance = 1.4695; skewness = 2.5365; kurtosis = 18.4818. The maximal number of substitutions performed by a team was 30, in the sample analyzed. We also note that there is similarity between the first and third quarter, and between the second and the forth respectively, but the last minute (substitutions caused by fouls or close ends).

The total number of substitutions per minute follows a *Power Law* from 4 substitutions, as the next figure shows

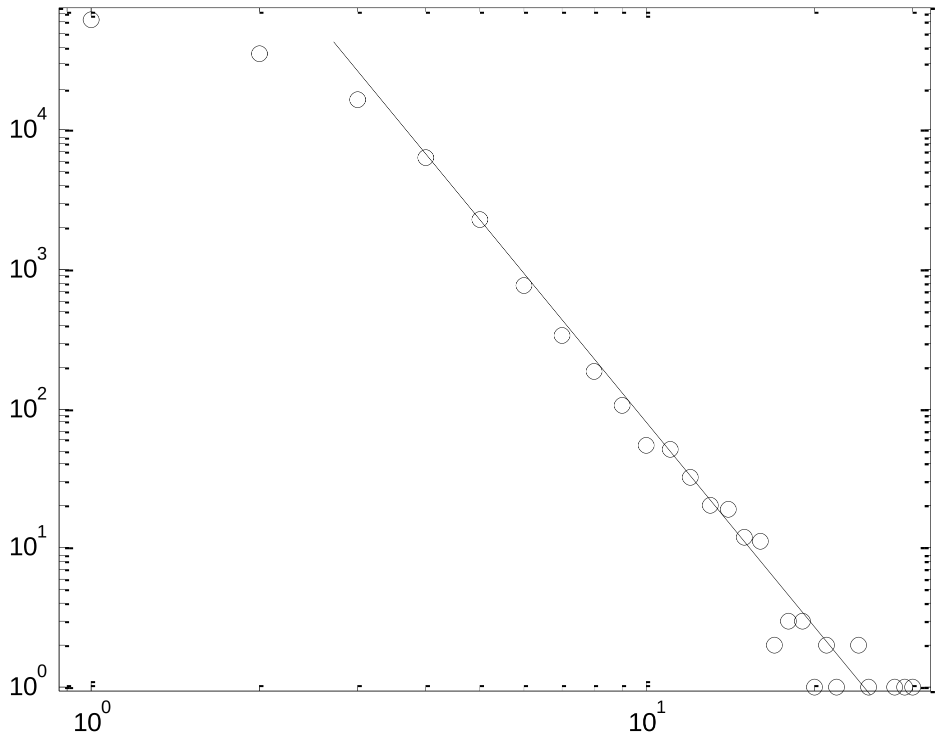


Figure 41. Log-log plot of the total number of substitutions. Note that from 4 substitutions the data becomes in a straight line, which points out a *Power Law* behavior.

We can see how from one to four substitutions present a logical distribution, one substitutions is the highest value, followed by two and so on. This can be related with the number of players at the court, five, in the sense of coaches do not substitute all players at the court. But beyond four substitutions, the distribution behaves as a *Power Law*. This can be interpreted as there are some situations where couches carry out a great number of substitutions per minute, and probably this are related with situations where they want to perform some concrete strategies, such as strategies in the last minutes of the games.

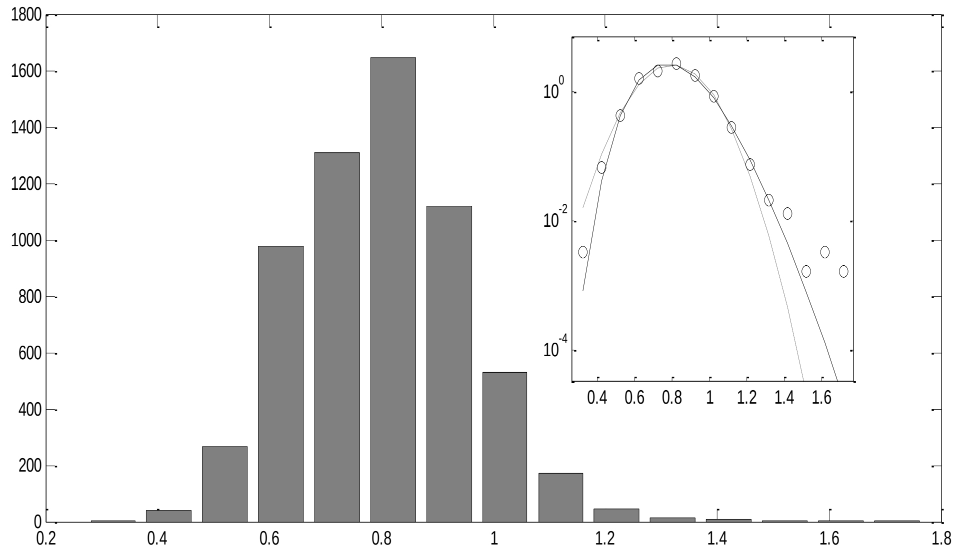


Figure 42. Index of Dispersion for the substitutions. The subplot represents the semi-log plot with Normal Distribution fit (dash line) and Gamma Distribution fit (solid line). The most part of the data are located under 1 (under-dispersed). It seems to fit better to a Gamma Distribution.

In the Figure 42 the Index of Dispersion for substitution is shown. Note that the most part of the data are lower than 1, which indicates that the data are under-dispersed; therefore we can deduce that there are some situations where substitutions are expected. Only a few data are over-dispersed, in fact, when we performed a semi-log plot (upper panel Figure 42) we can see how data are better fitted by a Gamma distribution (solid line) than by a Normal Distribution (dash line). The statistics (mean per game and per minute: λ mean) are mean = 0.7981; variance = 0.0224; skewness = 0.4741; kurtosis = 3.9436 min= 0.2708; max =1.7708. Thus we can consider it as a Gamma Distribution with fit parameters 28.5784 and 0.0279 (Quadratic Difference: 0.47).

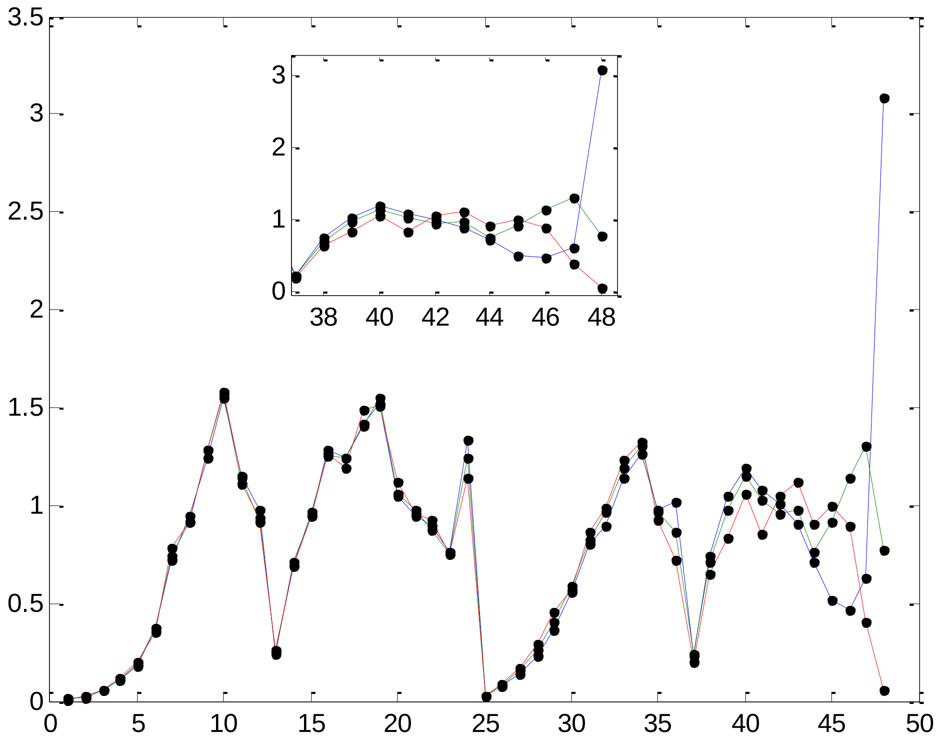


Figure 43. Number of substitutions per minute for games with point differences lower than 10 points (blue), between 11 to 28 (green) and more than 28 points (red). The upper panel displays the last quarter in detail.

The Figure 43 shows the number of substitutions per minute for the three gam profiles: with point differences lower than 10 points (blue), between 11 to 28 (green) and more than 28 points (red). We can observer that at the tendency is almost the same for all profiles but in the last quarter. It seems that the number substitutions increase at the beginning of the quarter, reach a peak and decrease at the end of the quarter. This behavior does not fulfill in the last minute of second quarter neither in the last quarter.

In the last quarter we can see how the substitutions vary depends on the situation. For the most competed situation (less than 10 point difference) is similar to the other cases at the beginning, but drops more than the others. This can be caused because teams play with the starter players and only make substitutions promoted by fouls or some specific strategies. The middle case when score finish between 10 to 28 points, on the contrary, tends to increase the number of substitutions as long as the end is approaching, perhaps seeking the victory, but finally, in the last minute this behavior is reduced.

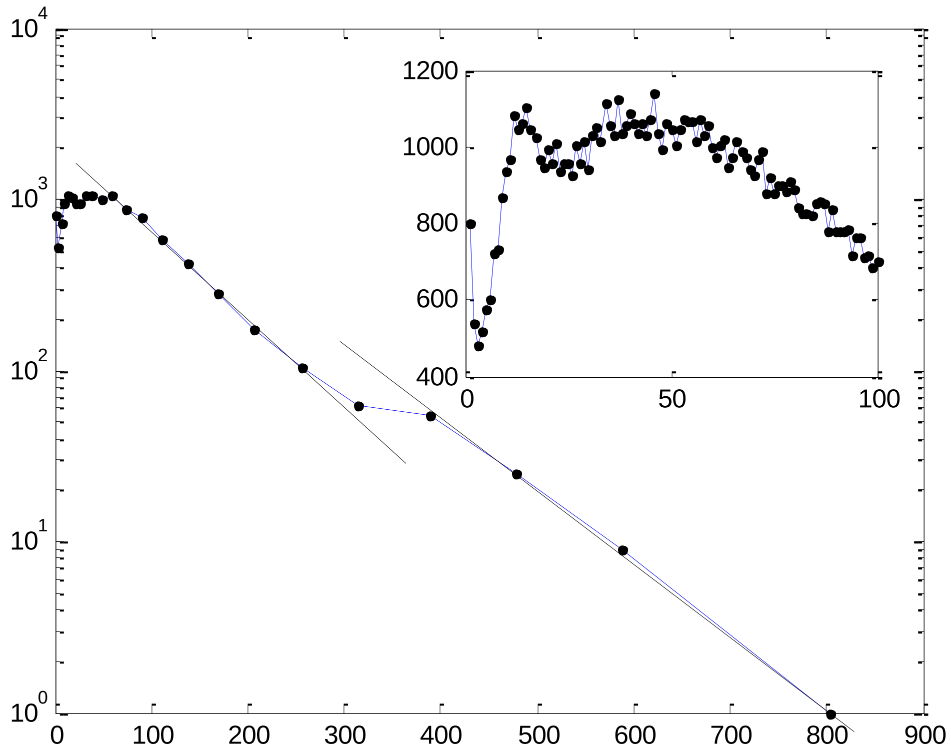


Figure 44. Semi-log plot of the time differences between substitutions. The upper panel represents the detail of time differences until 100 seconds of the total data.

In the Figure 44 we can observe the time intervals between substitutions. The number of substitutions with 1 second differences is significantly (see upper panel Figure 44). From 12 to 50 seconds there is a region that we can consider stable, compared to the rest. Beyond 60 seconds the data decrease until the end. Hence we can see how the time elapsed from a substitution to the next one is around from 12 to 60 seconds. Moreover, the end of the game probably related with a probability so high for one second time intervals. It is also remarkable the behavior around 300 seconds, were the slope changes. Further than 400 the tendency is different to the previous data.

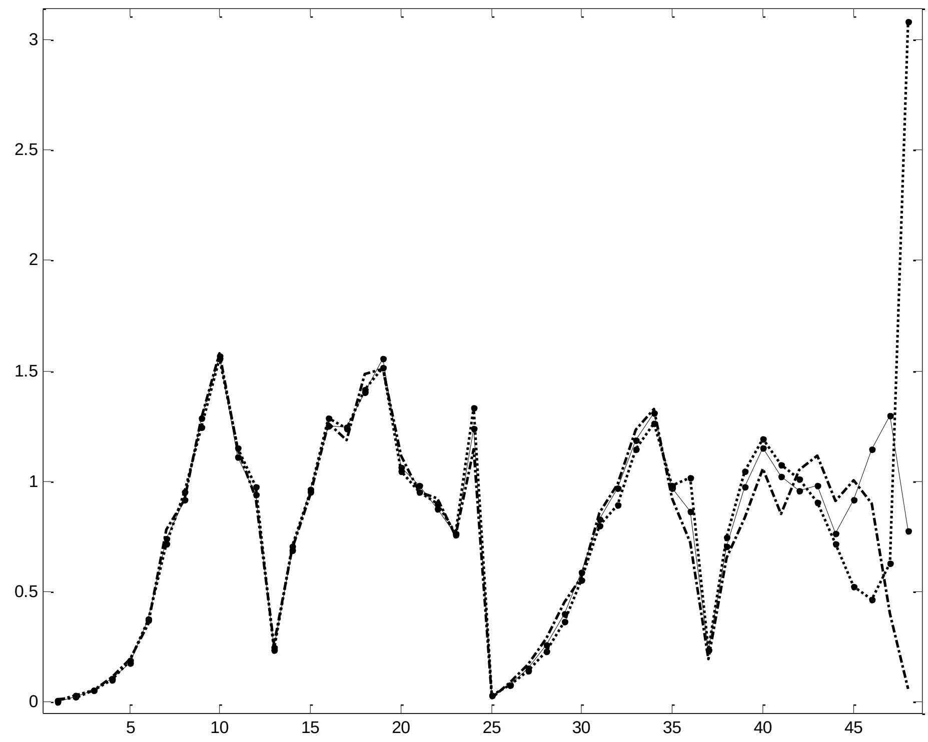


Figure 45. Here are showed the substitutions by quarter. The dash line represents the games ended by a point difference lower than 10 points. The solid line represents the games ended by 10 – 28 points. The —·— line the cases ended by more than 28 points. We can observe that the behavior among the three cases is quite similar but the last quarter. The most competed games (dash line) experience an increase at the end of the quarter. It is also remarkable the decrease in previous minutes (40 to 46). The intermediate cases (10-28 points differences) increase until the last minute, where drops. The last case, games ended by more than 28 points decrease from the minute 45.

The Figure 45 corresponds to the substitutions by quarter for the three cases regarding the final score. The dash line matches up to the most competed games, ended by less than 10 points, the solid line represents the games ended between 10 – 28 points and the —·— line the cases ended by more than 28 points. Note the significantly differences are only manifested at the last quarter. It is also remarkable that the first and the third quarter are quite similar. The substitutions tend to increase along with time, and in the last instants tend to decrease. The second and fourth quarter present also some similarities, but the last quarter displays more skewness.

The most competed games tend to carry out less number of substitutions in the last 5 minutes, approximately; perhaps they want to keep on the court the best players available at that moment, but in the last minute the last minute the tendency increase considerably. This is probably because some players are foul out, or because some last minute strategies. The case of games with a 10-28 points differences increase the number of substitutions in the last instants of the game, seeking the victory, but finally they give up to the score difference, as the decrease of substitutions shows. The case with the higher score differences displays less substitutions in the last minutes of the game time, probably because they have not enough time to overcome the situation.

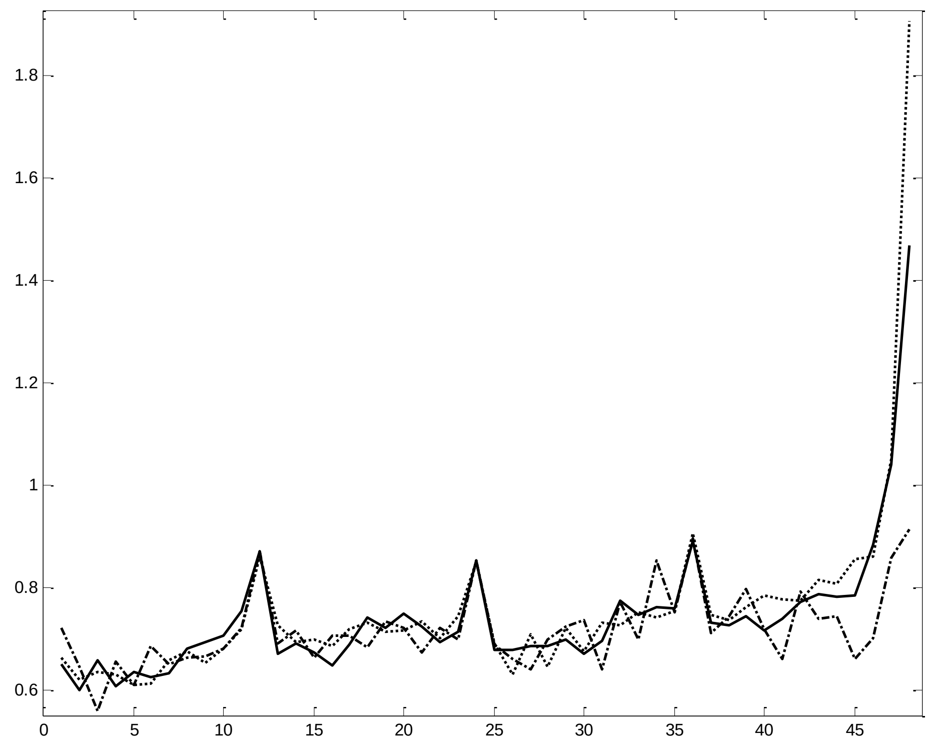


Figure 46. Score Index of Dispersion for the three cases proposed: games ended by a point difference lower than 10 points (dash line), games ended by 10 – 28 points (solid line) and ended by more than 28 points (—·— line).

The Index of Dispersion of score shows that if we perform an analysis by the three cases regarding the final score, we can see how all them tend to the value 1 at the end of the quarter. Only in the last minute, the game profiles of score differences lower than 10 points (dash line) and games ended by 10 – 28 points (solid line) exceed the value 1 (over-dispersed). Only the last case, games with score differences higher than 28 points, remains below 1. This point out that, regarding the final score difference, games are quite similar until the last quarter. In this quarter, the last minute is especially interesting; in the sense of games with low score differences present a final much more complex concerning points per minute, score runs, three point shots, fouls, free throws, etc.

Something similar happen when we perform the Index of dispersion of substitutions, as the next figure shows:

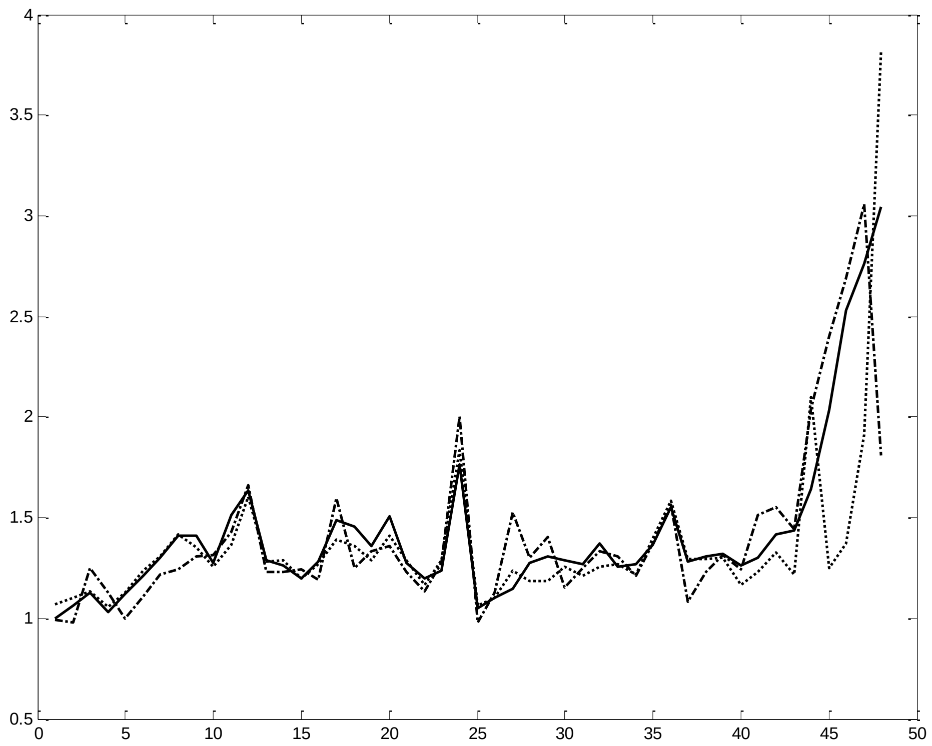


Figure 47. Index of Dispersion of the substitutions for the three cases referred. The dash line represents the games ended by a point difference lower than 10 points. The solid line represents the games ended by 10 – 28 points. The —·— line the cases ended by more than 28 points.

We can see that the Index of Dispersion of substitutions is always higher than 1 (over-dispersed). Even in the last quarter can reach values four times more. For all cases the Index of Dispersion tends to increase at the end of the game time. Note that with this item, the break between quarters is not so clear as before. Only the half time seems to be clear.

If we perform an analysis of kind of shots by quarter, we obtain the next figure:

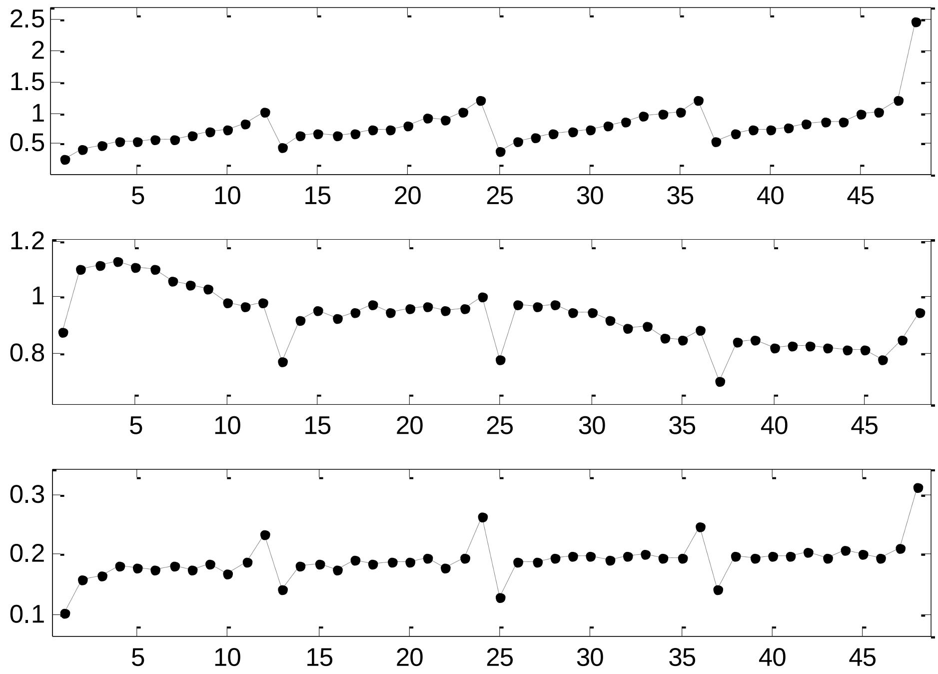


Figure 48. λ value per minute of 1 point shots (upper subplot), 2point shots (middle panel) and 3 point shots (lower panel) of the sample analyzed.

Note that for the free throws the λ value (number of points per minute) tends to increase as we approach to the end of each quarter, and it is reduced at the beginning. It seems logical taking to account that the most likely source of this behavior is the fouls accumulation (but is not the unique). This is clear at the last minute, where the λ value is more than the double than the rest. Surprisingly, the 2 point shots decrease along with time. Only at the first minute of each quarter is low; and the second quarter is more stable than the rest. Moreover, the last two minutes presents a proliferation in the 2 point shots. Regarding 3 point shots, they remain stable during game time. Only increase at the end of the each quarter (above all the last quarter); and note that in the beginning of each quarter there are less 3 point shots.

It is important to bear in mind that the great difference between free throws (1 point shots) and the 2 or 3 point shots is the intentionality. Players can choose to shoot from 2 or 3 point line. But the 1 point shot (a free throw) comes up from the game. It is an emergency. They do not choose to shoot a free throw. That is why the fouls issue is so important and so determinant in the last minutes of the quarter as well as in the last minutes of the game time.
On the other hand, we can see how players tend to shoot less as the game time progress. They tend to risk less from 2 point area, but they still risk from 3 point line, instead. This is probably

related with the score runs. The higher the score runs are, the greater the tendency to stop them, meaning more fouls or more 3 points shots, in order to reduce the differences quickly.

The next figure combines the plot of points and substitutions by second, showing the details of the quarter ends:

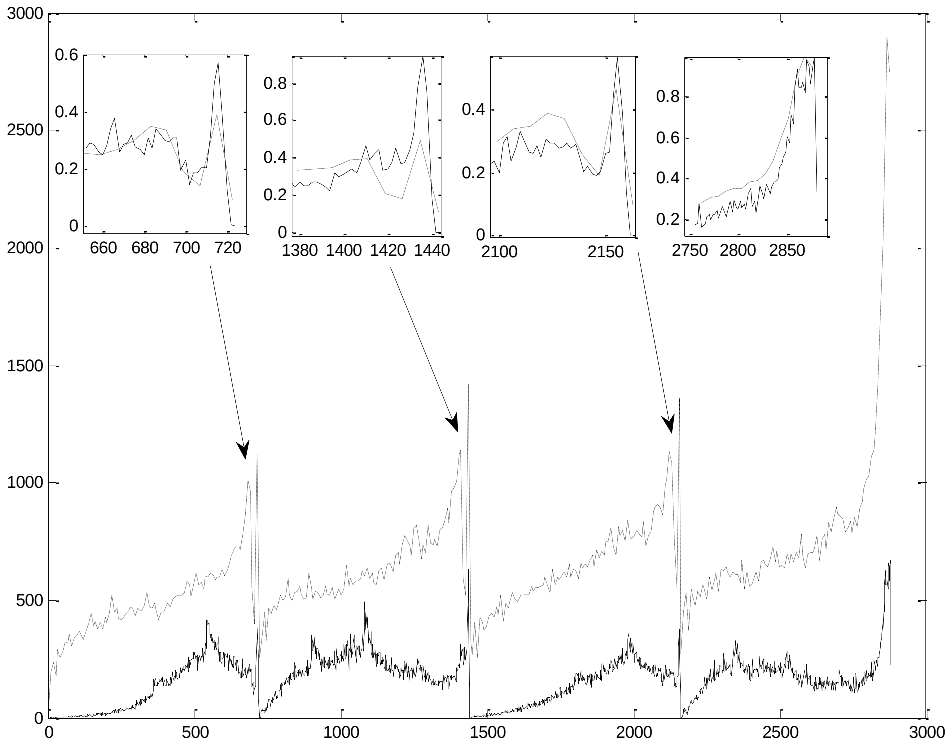


Figure 49. Detail of every quarter for points scored and substitutions. The dash line symbolizes the number of points per second and the solid line the number of substitutions. The upper boxes represents the final minutes (last 60 seconds for the first, second and third quarter and 120 seconds for the fourth quarter) of each quarter of points and substitutions normalized: dividing by the maximum value in each case.

The dash line corresponds to the number of points per second and the solid line the number of substitutions. Note that the first and the third quarter vary in the same way, points increment is similar but decrease at the end but substitutions remain increasing; whereas the second and fourth quarter are also similar (excluding the minute 48). They are more stable than the other two.

If we observe in detail the last 60 seconds of the first three quarters (see upper boxes), we see that there is a depression of the data just before the larger peak, except for points in the second quarter. This might be because teams try to spend the time possession in order to perform the last shot. In the last quarter on the contrary both, points and substitutions, increase until the end, which is quite normal if we assume that games are very competed.

In fact, if we observe the behavior of the 1 point shots (meaning fouls) and substitutions there is a correlation between the first and third quarter (Figure 50):

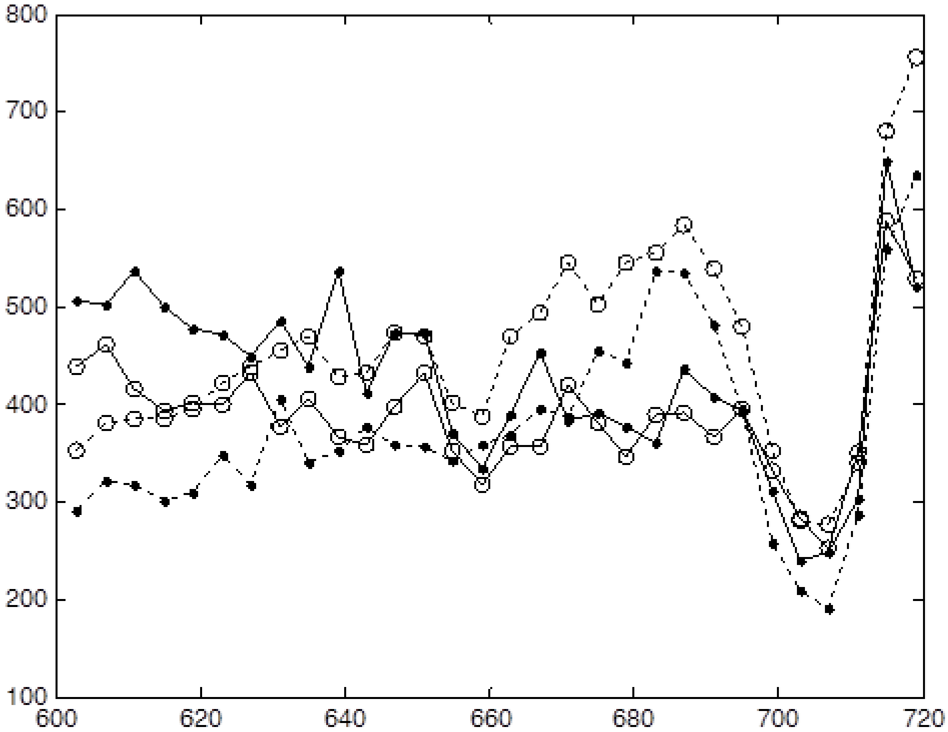


Figure 50. The dash lines represent fouls. The solid lines represent substitutions of (•) first quarter and (o) third quarter. The sampling was two second for better understanding.

We can note that the tendency is very similar. There is a significantly drop just before the peak at the end for both elements. But moreover, fouls in both periods are high just before this descent. This can be caused because teams tend to spend the last possession just before the break, as we pointed out before.

Something similar happen in the second period but not in the fourth, as we can see in the next Figure:

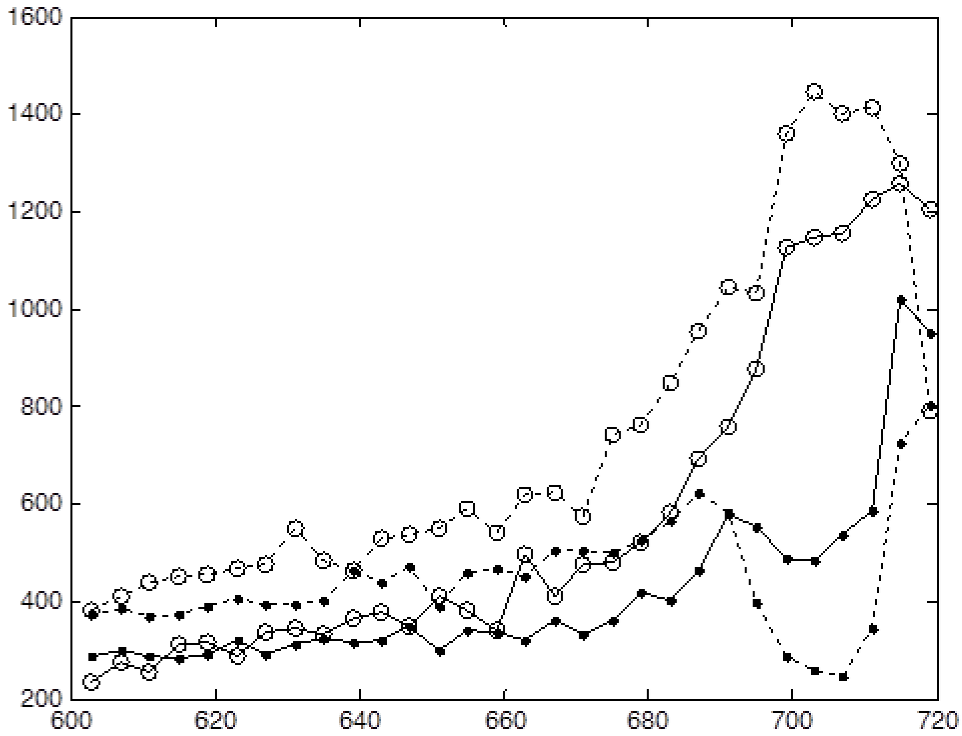


Figure 51. The dash lines represent fouls. The solid lines represent substitutions of (•) second quarter and (o) fourth quarter. The sampling was two second for better understanding.

The second quarter is similar to the first and the third, with a decent in fouls and points (no so significant) just before the half time. We can deduce that indeed the breaks have some effects regarding team behavior, meaning they rather spend the entire last possession in order to carry out the last shot. Not so in the fourth period, because the end of the game is coming, therefore teams tend to use all their resources.

7.4 Conclusions

As we can see, the use of a single probabilistic parameter in order to figure out the behavior of basketball games presents serious deficiencies. The Poissonian model and the scaling analysis provide an adequate framework in order to figure out the basketball game dynamics. Using simple tools such as the Index of dispersion and the λ value we are able to detect some concrete patterns.

Using these methods we established several games profiles: competed games, less competed games and very predictable games. Even we identified remarkable moments in the game such as the last minute of the game.

And based upon these results, we conclude that the research involving basketball games, should be carry out based on the point difference (these three game profiles) and not only distinguishing between games won and games lost.

Moreover, time in basketball does not follow a uniform behavior, but there are different behaviors in terms of time ranges. These temporal asymmetries indicate that the basketball score behavior has a non-linear nature. The score is a reflection of the different actions and behaviors resulting from the teams clash. It seems that teams generate complex behaviors that are manifested in the way the score evolves.

The competitive dynamics in the NBA can be considered an example of the Red Queen hypothesis proposed by Van Valen (1973): *For an evolutionary system, continuing development is needed just in order to maintain its fitness relative to the systems it is co-evolving with*. It is a race without end. All competitors need to improve to remain competitive.

As future research lines, it would be interesting to see whether the teams which are complex systems, possess or generate phenomena of learning and memory. And the degree of randomness that exists on the scoring of basketball is due to the chaos that reigns during the basketball game.

7.5. Practical Proposals

As a practical proposal we suggest to use this methodology in order to study and compare different basketball games took from other leagues, in order to asses them according to the score difference rather than win-lost games.

As another practical proposal, we suggest to us this methodology with very high games, such as Olympic Games, World Championship, European Championship, etc. This methodology enables us to detect some particularities in international games.

7.6. Limitations of the techniques used

It is necessary a big data base to obtain relevant conclusions, hence there are problems derived from working with big data bases.

Study 3.
Basketball Team

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

8. Study 3. Basketball Team.

8.1. Introduction

As last stage at our investigation, we wanted to make an approach to the dynamic of a basketball team. Our intention was to figure out the internal process of a basketball team during a real basketball game. We are talking about the flow of the game and the design of the structure, which can be described as in many others natural systems (Newman, 2001b; Bejan & Lorente, 2011; Bejan & Zane, 2013). We can see how structure, shape and functionality are closely related in several sport systems (García Manso et al., 2008; Yarrow, Brown, & Krakauer, 2009; Charles & Bejan, 2009; Ribeiro, Mendes, Malacarne, Jr, & Santoro, 2010; Bejan, Jones, & Charles, 2010).

Therefore, we propose the use of the network theory in order to clarify the features of the basketball team as a players network. This methodology can provide us the chance to find out how a basketball team works through the players´ behavior. Through their real interactions.

Most studies deal with this phenomenon with regard to external aspects, for instance the inclusion of a player in a team or a league or tournament (Onody & de Castro, 2004; Radicchi, 2011). We wanted to tackle the problem from a collaboration point of view. From internal process, not from external aspects.

8.2. Methodology

8.2.1. Network theory

Network theory is applied in several fields of knowledge and study such as biology, mathematics, economy, ecology, physics, sociology, engineering and of course in sports science. The first antecedent of that has certainty in the scientific field is the famous problem of the seven bridges of Königsberg, proposed by the mathematician Leonhard Euler in 1736. Euler described mathematically the vertex and links necessary to resolve the problem of finding a path through the city that would cross each bridge once and only once, and return to the starting point. So, the graph theory was established, a branch of mathematics which studies the properties of network structures.

In more recent times, noteworthy was the work of Jacob Moreno, a disciple of Sigmund Freud, who in the 20s broke with his mentor's ideas about social-emotional problems of the individual. He believes that were related to current relationships with family, friends, acquaintances, workmates, etc. In order to study these relationships he developed the sociograms and he used some concepts such as centrality and isolation.

Another relevant work in this field, and perhaps one of the most famous, was the one carried out by the North American psychologist Stanley Milgram in 1967. Milgran devised a theory that he called “Small-world theory”, which attracted the interest of many researchers. The general conclusion of Milgram works was that a Small-world network presents a high clustering coefficient and short path length. On average, any person in the world is separated from any other for only six intermediaries or six degrees of separation. Also he found a curious fact; most of the transmissions studied went through the same four people. This type of structure promotes creativity and collaboration because information flows through many nodes far apart in a few steps (Uzzi & Spiro, 2005; Fleming & Marx, 2006).

In fact affiliation in networks has shown a significant effect on performance. So much so that in the masterpiece of sociology of philosophies written by Randall Colins, describes whether the inventions in art, science and philosophy are the result of individual work or of those who were part of teams, work networks or communities. Collins proved that except for three individuals: the Taoist Wang Chung, the buddhist of 14th century Bassui Tokusho and for the Arabic philosopher Ibn Khaldun, of 14th century as well, the rest of great advances, including Freud, Hegel, de Medicci, Smith, Hutchinson, Watson and Crick, and Darwin were carried out by individuals who were part of a network (Uzzi, Amaral, & Reed-tsochas, 2007).

Regarding network study, another relevant finding was in 1960 when Paul Erdos and Alfréd Rényi published their random graph theory (Erdös & Rényi, 1960). They studied how the network topology changes depending on the number of connections. When the number of connections (m) is small, the networks looks fragmented in small groups or nodes, which they called components (n). When the number of connections increases, the isolated nodes start to get connected. And later, the nodes that were already connected will be connected to others what were not. A phase transition takes place when $m4n/2$, where many of these crosslinking groups are joined spontaneously to form a single giant component.

This theory has been studied in depth by other mathematicians and scientists from other areas. They have also served as an archetype for coupling dynamic models of gene networks, ecosystems and the spread of infectious diseases and viruses (Kauffman, 1995; May, 2001; Strogatz, 2001).

Another significant advance in the study of network topology came when Duncan Watts modeled mathematically the small-world networks. This model is known as Watts-Strogatz model in honor of J. Duncan Watts and Steven Strogatz (Watts & Strogatz, 1998). The model was designed as simple as possible which addresses some limitations of Erdos-Rényi model, for example, model-Rényi Erdos establishes a constant probability, random and independent of which two nodes are connected, and also have a low clustering coefficient. On the other hand, the Watts-Strogatz model sets small average distances and high values of clustering coefficient. It also differs in that the Erdos-Renyi model follows a Poisson degree distribution, rather than a *Power Law* as observed in most real networks.

The concept of scale-free network was introduced in 1999 by Albert-László Barabási and Reka Albert, which is known as Barabási–Albert model (Barabási & Albert, 1999). This model explains how scale-free networks are built randomly by a mechanism denominated preferential attachment or Saint Mathew effect. The scale-free networks are widely observed in natural and artificial systems, including the Internet, bibliographic citations networks and some social networks.

What is a network?

A network is just a graphic representation (graph) of some nodes (network elements) connected by links (real o illustrative).

That is, we can make a graphical representation of reality through a graph that represents the agents involved and their relationships and communication channels between them. In this way we can study their behavior and properties.

But why is it so important to know about the network topology. Just because the structure affects the function (Montoya Terán, Solé, & Rodríguez Fernández, 2001; Strogatz, 2001;

Amaral & Ottino, 2004; Uzzi et al., 2007; Solé, 2009). As discussed below, networks can take many forms that directly affect their properties.

There are several criteria of classification for studying networks. Perhaps most famous are those which make reference to their functions, such as social networks, transport networks, neuronal networks, telecommunication networks, etc. Thus we can clearly discern the general features of the network. But in science, the study of the networks is performed from the perspective of the network properties, which are also closely related to their architecture.

The classifications most used are:

- Simple networks
- Bayesian network
- Complex networks

Simple networks are those in which relationships among their elements are carry out linearly and these links are well defined. A great example is the Army chain of command, where the levels are linearly related and perfectly defined.

A **bayesian network**, influence diagram or probabilistic causal network, is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph. This enables to find out the underlying causal structure in a data set. For example, a bayesian network could represent the probabilistic relationships between diseases and symptoms. Given symptoms, the network can be used to compute the probabilities of the presence of various diseases.

Complex networks have some specific features, such as high clustering coefficient, specific degree distributions and non-linear relationships usually. Within this kind of network, there are several structures. This is the kind of network classification most common in our field. That is why we are going to describe more about it.

We have to take into account that study of networks is a thrilling subject and despite to be relatively new, it has attracted the attention of many scientists from different disciplines, Therefore we can observe that many of the concepts and classifications are very recent and the papers are published in journals of high international quality (Riolo, Cohen, & Axelrod, 2001; Strogatz, 2001; Amaral & Ottino, 2004; Uzzi et al., 2007; Guimera, Sales-Pardo, & Amaral, 2007), hence, sometimes there is no consensus in dealing with certain aspects of network theory. But perhaps this is a symptom of that research in this field is in full swing

8.2.2. Complex Networks

Some authors, in fact, have highlighted the difficulties involved the study of complex networks. Strogatz emphasizes the following aspect (Strogatz, 2001): the most common mistakes is to assume that the networks are static, when actually, from a functional point of view self-organization would be the most desirable response.

Other authors also point out that it is very difficult, if not impossible, to obtain information from a network with hundreds or thousands of nodes and connections, unless the information about the nodes and connections is presented in the context of specific scale (Guimera & Amaral, 2005).

As we can see, complex networks appear in many biological and technological contexts, indicating universality of certain functional and organizational principles in complex systems (Montoya et al., 2001; Solé & Goodwin, 2002).

As mentioned, there are several kinds of networks, depending on their properties. In complex networks are included two that we mentioned in the introduction and which most of the literature refers: the small-world networks and scale-free networks.

Small-world networks

Small-world networks present two main features:

- High clustering coefficient.

- Short paths length. It is possible to connect two vertices of the network through only a few links. Therefore the mean distance should be small. The local connectivity suggests that the network is of finite dimension

The structure of small-world network facilitates the transmission of information, because far-away nodes can be linked in a few steps. This high speed in transmission through the communication channel between the distal parts of the system facilitates any dynamic process (such as synchronization or calculation), which requires the overall coordination and information flow. There are numbers examples: metabolic (Jeong, Tombor, Albert, Oltvai, & Barabási, 2000), scientific collaborations (Newman, 2001b), internet (Adamic, 1999), etc.

A small-world network can preserve its inherent structure despite a considerable number of external aggressions. This suggests the need for a massive amount of restructuring to transform a small-world network in other structure, an important finding for the understanding of how to measure and evaluate the structure behind an industry and economy. Sometimes the small-world network structure affects performance and sometimes not affects performance in ways comparable (Uzzi et al., 2007).

Montoya and collaborators (Montoya et al., 2001) point out that, in ecology, may not matter to have a single network with many species or to have many small subnets with few species each. But it is not. They point out that the risk of extinction is much higher in the second case, due to what they named *effect of biological insurance*. This mechanism ensures to the ecosystem a better chance of overcoming strong shocks, because there are species that can adapt to new conditions and / or assume the roles of other species that have disappeared.

Scale-free networks

A scale-free network is one whose degree distribution of connections follows a *Power Law*. In the real world, some nodes are more connected than others. If network connections were random, the degree of connections would be well defined by a Poisson distribution. But for most real networks, the distribution is well defined by a distribution in the form of a *Power Law* (Barabási & Albert, 1999; Strogatz, 2001). There are many examples such as Internet (Montoya et al., 2001; Strogatz, 2001; Amaral & Ottino, 2004), telephone calls (Barabási & Albert, 1999), metabolic reactions (Jeong et al., 2000), transport network (Amaral & Ottino, 2004) and social networks such as collaborations between Hollywood actors (Barabási & Albert, 1999) and citations in scientific journals (Seglen, 1992; Redner, 1998).

Such networks are often associated with fractal phenomena, phase transitions and other situations where *Power Law* show up (Newman, 2001b; Strogatz, 2001; Amaral & Ottino, 2004). But as these arguments and considering the preferential attachment mechanism described for such networks, a key issue appears. Why these networks are not increased to infinity? i.e. Why there are not new nodes and new connections are not constantly created between new nodes and old. Amaral and colleagues collaborators (Amaral, Scala, Barthélémy, & Stanley, 2000) propose two interesting points in response to this dilemma:

1. The aging of the nodes. The elements that compose the network are not eternal and reached a moment longer constitute part of the network. Because it disappear or become part of another network. That is why the nodes aging mechanism limits the preferential attachment in these networks.

2. The cost of adding new nodes or links. Amaral and colleagues cite the example of an airport. It would be impossible for an airport become a hub for all airlines in the world, simply by a question of space and time. Each airport has some limitations on air traffic and passenger traffic. Hence, the network capabilities limit adding new nodes.

These same authors also note that there is an analogy with critical phenomena, because in the critical points of phase transitions appear *Power Laws*, and once this point is exceeded, no longer observed *Power Laws*. Often the distributions become exponential or Gaussian (Stanley, 1987; Amaral et al., 2000).

8.2.3. Sport Application

There are only a bunch of papers dealing with sport as network. This is due to it is a new field of application in which this methodology is applied, probably. The application spectrum is wide because exists a lot of levels where this theory can be used. For instance we can study the behavior of a player within a team, or we can analyze the trade of players in a league or between sport clubs, or even the dynamic of the teams in a league or tournament. Therefore, it is important to define the background where we are to carry out our investigation.

For instance, Park and Newman (2005) propose a network-based ranking system for US college football. The method has a free parameter and they provide empirical evidence indicating the typical range of the best values for this parameter and a method for choosing a value in any particular case.

Bejan and Haynsworth (2012) studied the natural design of the ranking of college basketball, and its relationship with academics. Both present hierarchical structure but are independent flows because there is no correlation between the two rankings. The movement of basketball players from high school to the professional level is a flow with its own architecture.

But this kind of analysis is not exclusive for team sports. It is true network analysis in sport is often associated with team sports, but in individual sport network analysis can be performed as well, for instance tennis. Radicchi (2011) try to find out the quality of participating players in the ATP ranking between 1968 and 2010. He studied the contact network of professional players ranked in the Association of Tennis Professionals (ATP). This is restricted only to those players who have been number one in the ATP ranking (24 players). Moreover, Radicchi point out that in general, networks obtained from the aggregation of a sufficiently high number of matches have topological complex features consistent with the majority of networked social systems so far studied in literature (Albert & Barabási, 2002; Newman, 2003).

Another interesting example is the study of players as social networks, meaning sport clubs as structures which contain players. And therefore study the movement of players from a team to another, or even the time they remain in the same team. A good example is described by Onody and de Castro (2004), who studied the Brazilian players network of 32 editions of the Brazilian football championship. They point out the existence of a growing segregationist pattern, where the players transfer occurs, preferentially, between teams of the same size. And the average shortest path length values may suggest that it is size independent but, most probably, but they mention that this conclusion is misled by the presence of only some few generations of players in the growing Brazilian players network. They say that it is 190 times more probable to find someone who has played for only two clubs than for eight clubs.

Another level in the study of sports networks is the game itself. In this case there are different ways to face the problem. In football, for example, the flow of the game can be study through players (Hughes & Franks, 2005; Brillinger, 2007).

8.2.4. Application to basketball

As an approach for using network theory in basketball, we carried out an in investigation about players network. Our intention was try to clarify how is the behavior of members of the same team when the fight against the other team. Remember that collaboration among the players of the same team allows team to compete against other teams. But at the same time confrontation between teams cause a critical situation which allows teams to improve their performance. It can seem simple but it does not. Therefore it is very important the selection of the team (roster, coach, staff) and the process of hiring new players (for instance the combination between veterans and rookie players). In fact, a long-standing problem in biological and social sciences is to understand the conditions required for the emergence and maintenance of cooperation in evolving populations (Riolo et al., 2001; Guimera, Uzzi, Spiro, & Amaral, 2005). The aim should be cooperation without reciprocity, without selfishness. As Phil Jackson said: *to put the "me" in service of the "we"* (Jackson & Delehanty, 2006).

But the problem does not end here. We have to bear in mind that there is another team with opposite purpose; ready to stop every thrust. If teams are competitive enough, the system (basketball game) becomes critical; even more because they have to resolve the situation in a bounded time (24 second shot clock and game time) and always regarding to scoring, which works as order parameter (de Saá Guerra et al., 2013). Therefore, we have to take into consideration the surrounding conditions: time and scoring.

Hence, we can see how environment influences over players, and players try to dominate the situation. For that reason actions of players can provide us some clues about their real interactions; and scoring indicates whether they succeeded or not. Players actions contain information, i.e. coach indications, pre-prepared play, strategy, etc. those provide us a sign over the profile of each team and how they face different game situations.

8.3. Results and discussion

As example of this methodology, we study the last game of 2011 Eastern Conference NBA Finals Chicago Bulls vs. Miami Heat. Regarding player interactions, we measured for each team the number or passes, screens and space creations per play.

Passes represent the clearest example of player interaction, because the fact to pass the ball to a teammate enables to make a shot or another favorable situation. Screens also represent an interesting example of player interaction because the aim of a screen is to neutralize a teammate defense, in order to provide a superiority situation such as drive to the basket, a clear shot, good pass creation, space creation, etc. This point is related whit next.

Space creation is an interaction among players as well. Great example of this phenomenon is the pick and roll situation: A two-person play in which an offensive player sets a screen (pick) on the ball handler's defender and cuts (rolls) to the basket after the ball handler drives by the screen. He now has the small guy on him (called a switch) and can scores easily. Also, when a player is occupying a court location, he can leave that space to perform a clear out play, or a curl cut play (e.g. the famous UCLA cut).

Here are represented the graph of passes, screens and space creation for both teams:

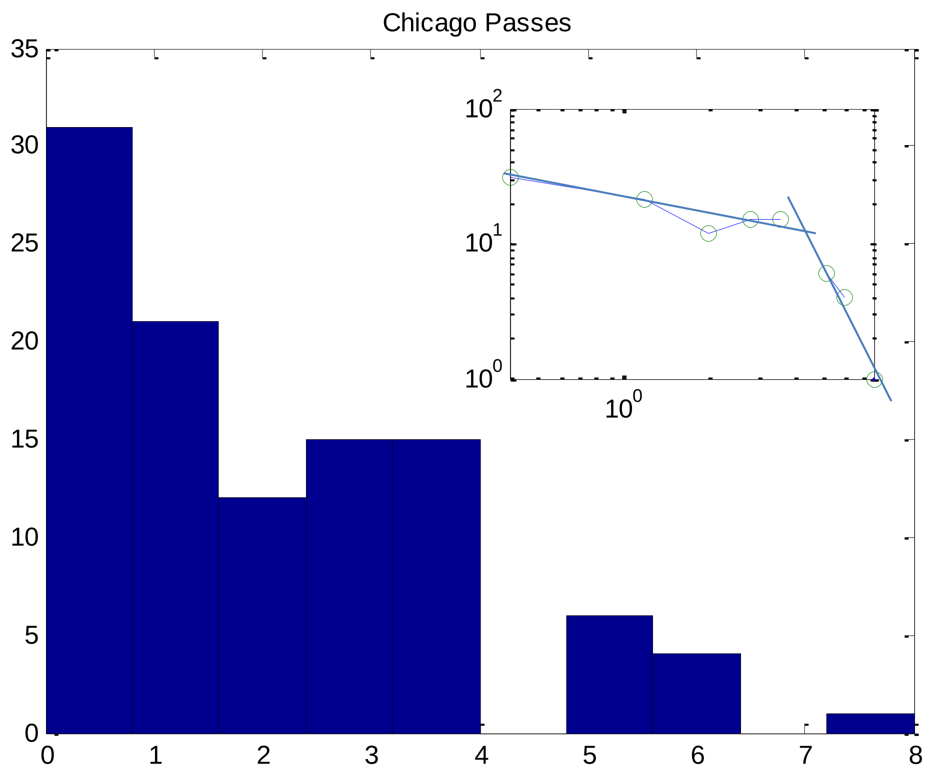


Figure 52. Histogram of Chicago passes. X-axis represents the number of passes and the y-axis represents the number of plays. The distribution is not uniform. There are more plays with zero passes than one. Plays with two, three and four passes are relatively homogeneous, but beyond four passes the distribution decays. When we carry out a log-log plot (upper panel) we observe there is a tipping point around 4 passes.

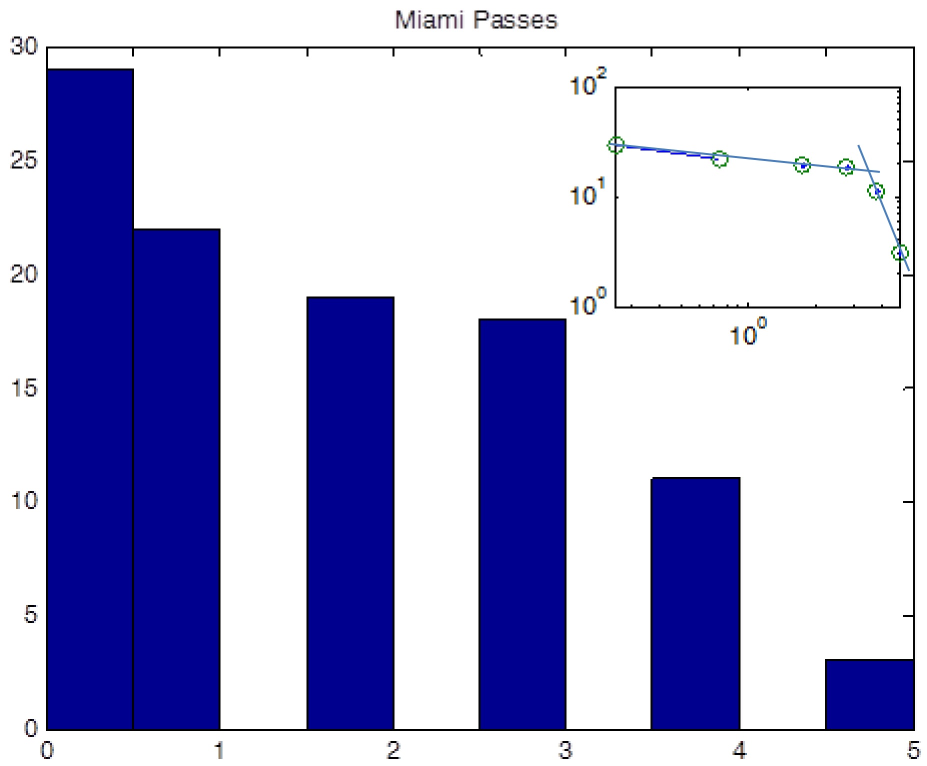


Figure 53. Histogram Miami passes. X-axis represents the number of passes and the y-axis represents the number of plays. It seems to be no large differences between the number of passes per play. The distribution always is decreasing but the log-log plot (upper panel) indicates a cut beyond three passes.

The Figures 52 and 53 both represent the histogram of the number of passes per play. The x-axis represents the number of passes and the y-axis represents the number of plays with that number of passes happened in the game. Both distributions are not uniform but decrease. Chicago, in this game, in some plays gives up to eight passes whereas Miami only reaches five passes maximum. Even the differences in Miami are lower than Chicago, generally. Both upper panels are the log-log plot of the distribution. Note that the cut for Chicago is located in four passes, but for Miami is located around three.

The presence of this truncated *Power Laws* point out different dynamic regarding passing. For the two teams once reached the tipping point changes, the passing dynamic substantially. This is important as we point out later.

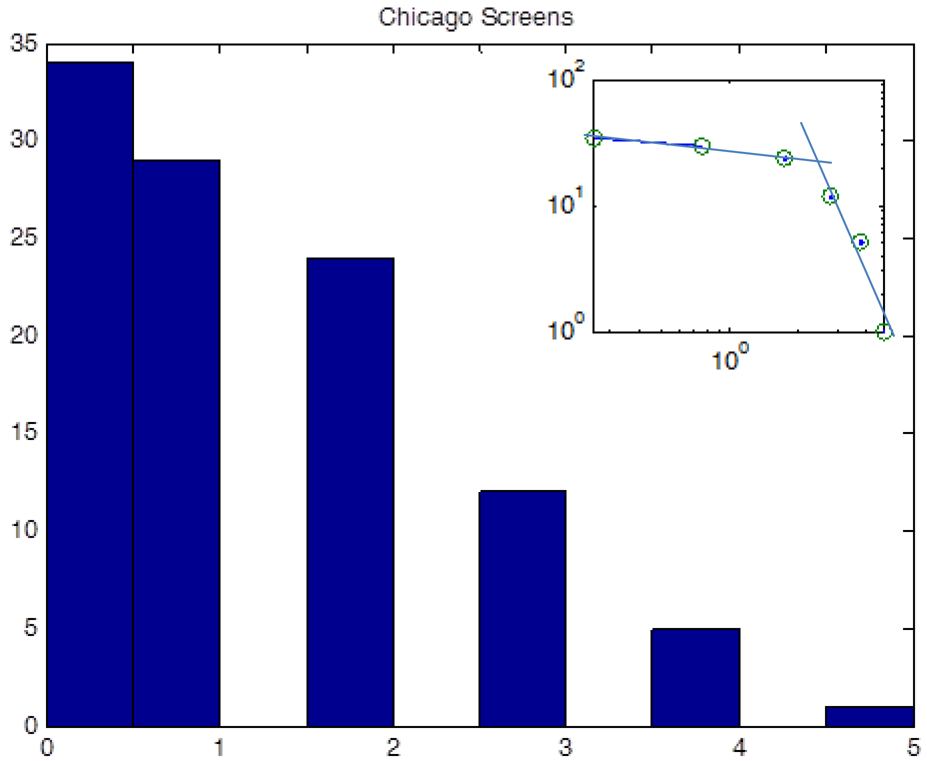


Figure 54. Histogram Chicago screens. X-axis represents the number of screens and the y-axis represents the number of plays. The distribution is not uniform. There are more plays with zero screens than one and so on. Plays with one and two screens are relatively homogeneous. Then log-log plot (upper panel) show us that beyond two screens the distribution decays, there is a tipping point located here.

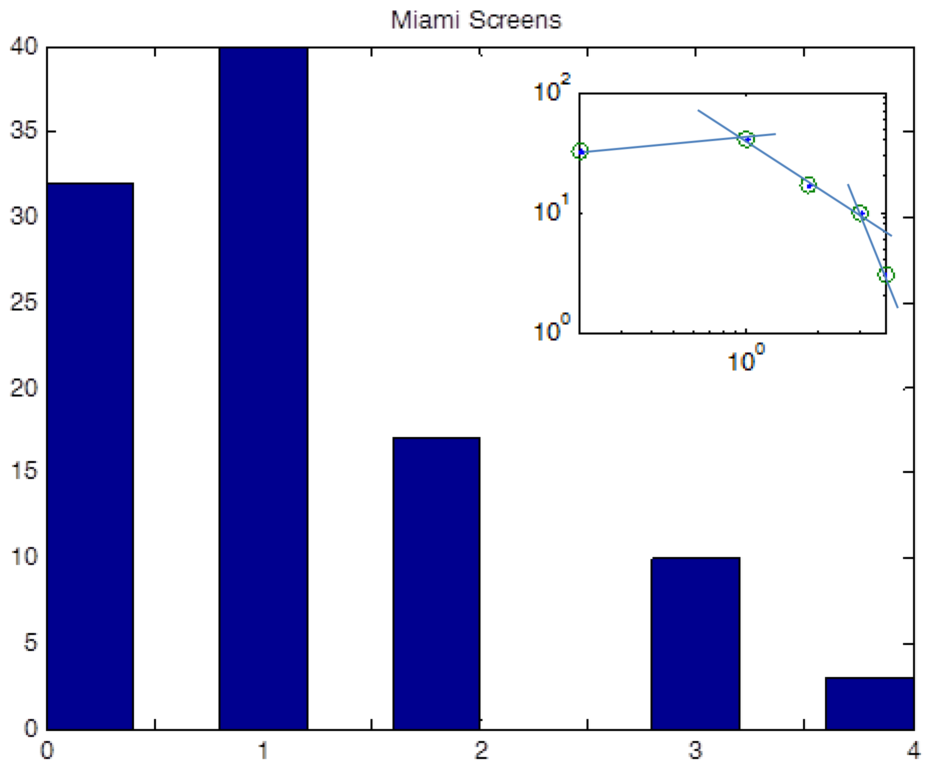


Figure 55. Histogram Miami screens. X-axis represents the number of screens and the y-axis represents the number of plays. The number of plays with a single screen is more numerous than the rest. This is interesting because when we perform a log-log plot (upper panel) points out the presence of three different areas.

The Figures 54 and 55 both represent the histogram of the number of **s**creens per play. We can see that the histograms are not similar. The histogram of Chicago decays always while the predominant situation in Miami is with one single screen. Moreover, sometimes Chicago performs plays with up to 5 screens; one more than Miami. When we carried out a log-log plot for both histograms (upper panels), note a cut in two screens for Chicago and something curious happen in Miami. There is not a single cut, but two. The first one located in one screen and the second one located in three screens.

It is remarkable that in Miami appears two different cutting points. This fact gives us a lot of information about the behavior regarding screens, and for analogy, about the Miami game style. The core situation for Miami is with one screen mainly, or with no screens. If they are not able to resolve the situation, their screening dynamic change appearing two or three screens. And when the situations become very complicated, they reach up to four screens, but this situation is very rare compared with the previous.

Chicago, on the other hand present a homogeneous behavior with zero, one and two screens. Once reached this point, Chicago team profile says that the internal dynamic have been modified and they can be given up to five screens in the same play.

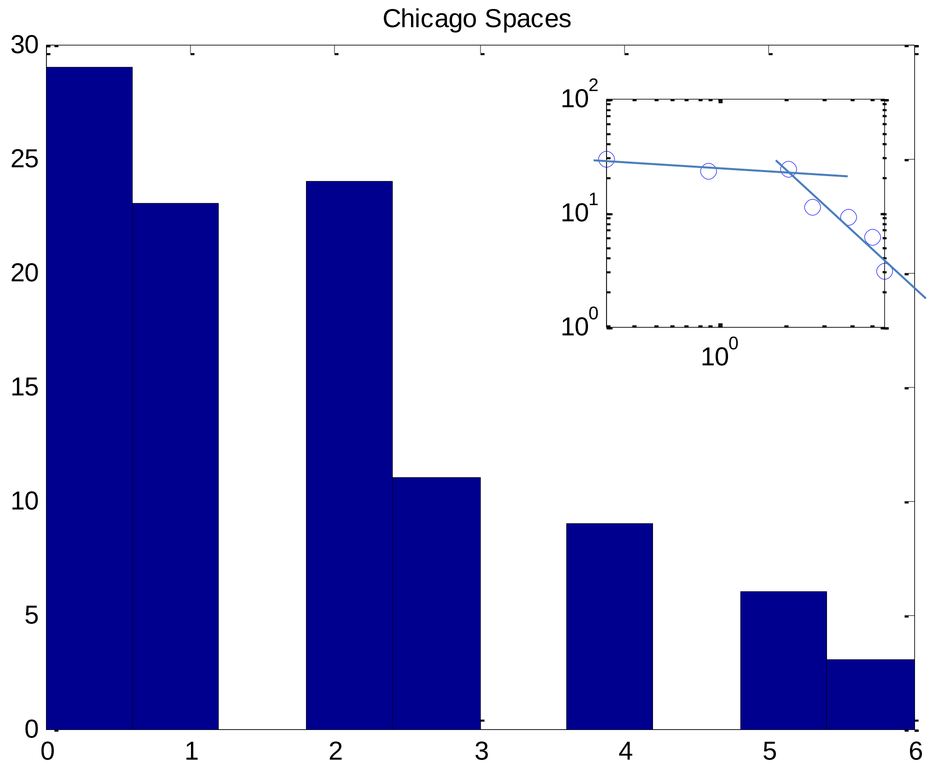


Figure 56. Histogram Chicago spaces creation. X-axis represents the number of space creations and the y-axis represents the number of plays. The most common situation is with zero, one and two space creations. Beyond this point the histogram decays up to six space creation. The log-log plot (upper panel) indicates a cut beyond two space creations.

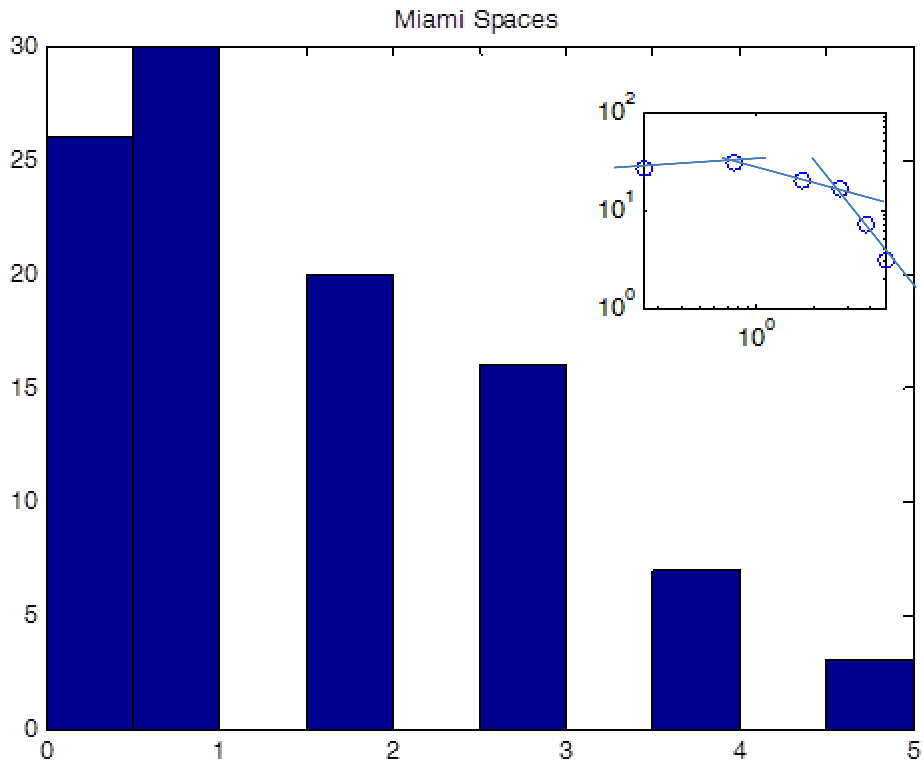


Figure 57. Histogram Miami spaces creation. X-axis represents the number of space creations and the y-axis represents the number of plays. Here the behavior is similar to screens. The number of plays with a single spaces creation is more numerous than the rest, followed by zero spaces creation. Another kind of dynamic is given by plays with two and spaces creation, as we can see when we take the log-log plot (upper panel), and a third zone points out the presence a different performance.

As we can observe in the Figures 56 and 57, Chicago and Miami display different profiles regarding space creations. Chicago seems to behave a homogeneous region from zero to two space creations, and the values drops from this point. In contrast, Miami shows a predominance of one space creation even though zero space creations values are similar. The log-log plot (upper panels) reveals two different areas in Chicago and three in Miami.

This analysis point out the presence of several truncated *Power Laws* (different for each).Some papers (Malacarne & Mendes, 2000; Greenhough et al., 2001; Mendes et al., 2007; Bittner et al., 2009; Heuer et al., 2010) emphasize the presence of heavy-tailed distributions (*Power Laws*), which are associated with many natural and social phenomena. These kinds of distributions are related to ideas from statistical physics and non-linear complex systems, such as anomalous diffusion by the Zipf-Mandelbrot law (Malacarne & Mendes, 2000), self-organized criticality phenomena, or non-linear dynamics (Bourbousson et al., 2010; McGarry et al., 2002). But as we mentioned, in basketball game dynamic (team behavior) is closely related to environment because environment conditions the team outcome. Hence, these phenomena cannot be understood by isolation.

The only parameter that set the performance in a basketball game is the score, but it is a non-linear process. In fact its dynamic can tell if the game is competed or not (de Saá Guerra et al., 2013). Therefore to understand team dynamic is necessary to study the system in its own background.

The game analyzed, in particular, is a high competed game. The most part of the time, point difference was lower or equal to ten points, and when we analyze all the parameters together, we obtain the following Figures:

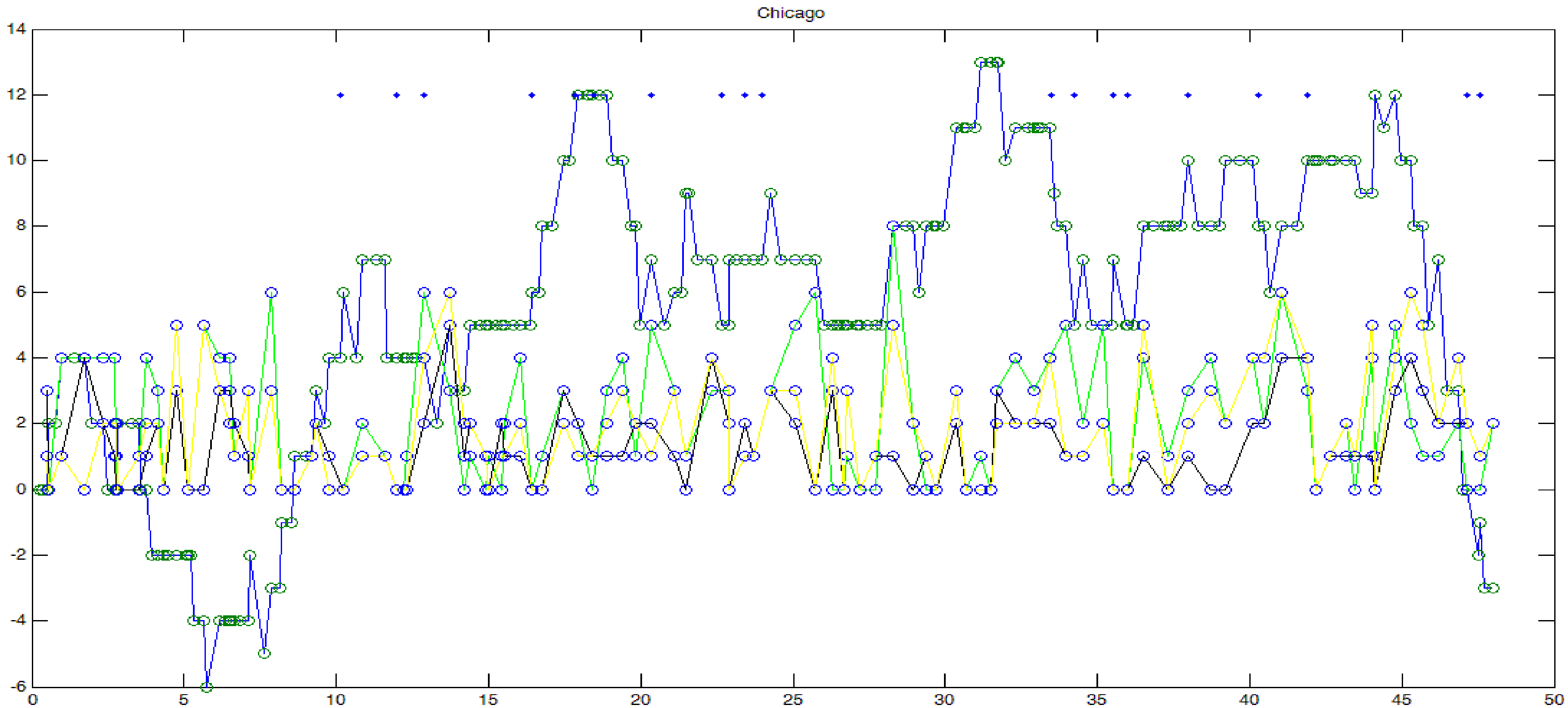


Figure 58. The x-axis displays the game time. The y-axis is the point difference: positive for Chicago, negative for Miami. The blue line corresponds to the point difference throughout game time. The green line represents the passes, the black one the screens and the yellow one the space creations. The blue points above represent the player substitutions. The most remarkable aspect is that the frequencies of these phenomena match with some key game situations and some score runs. In the Chicago case, we can observe that in disadvantage situations or/and after negative score runs, the frequency of passes, screens and space creations increases. E.g. minutes five, fourteen, twenty, forty two, forty five, etc. And the reverse situation: When Chicago takes advantage, frequencies of passes, screens and space creation tend to stabilize (low values).

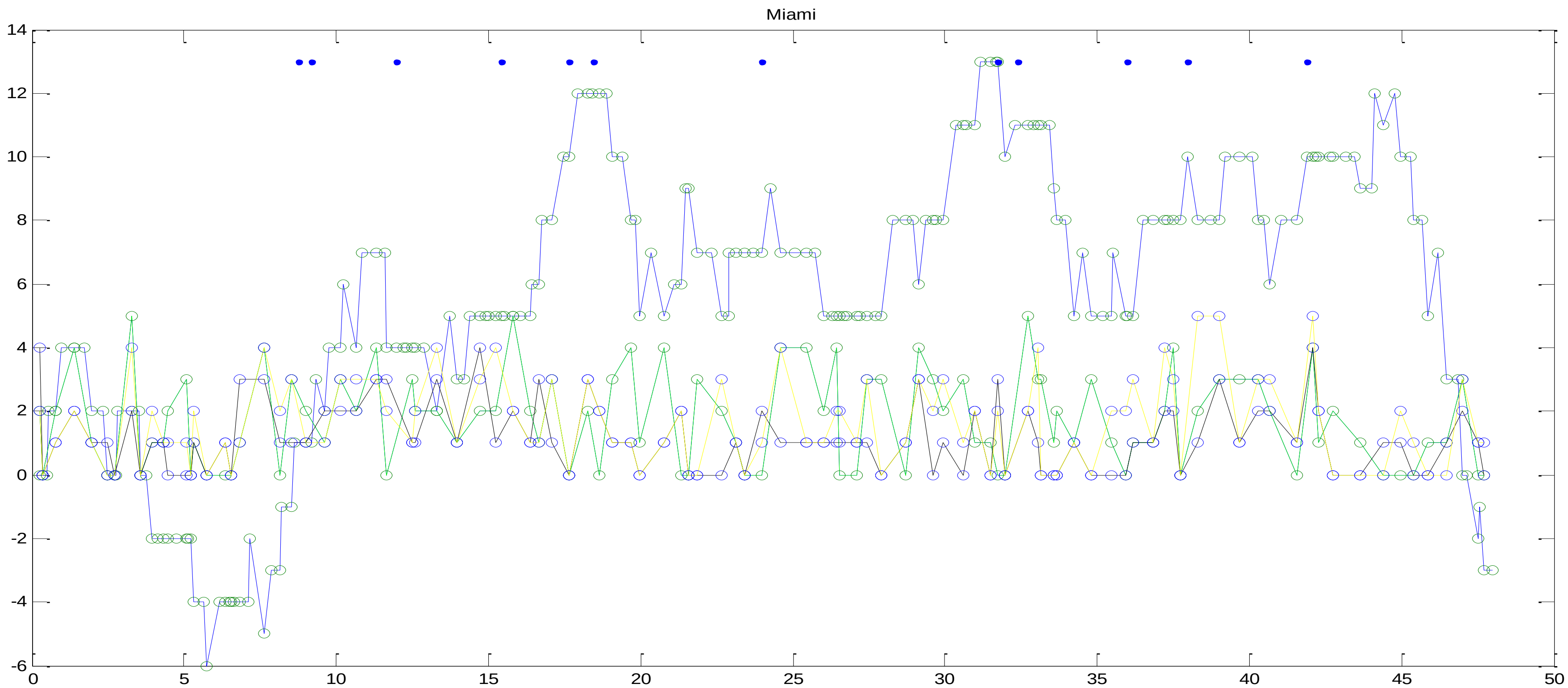


Figure 59. The x-axis represents the game time. The y-axis, the point difference: positive for Chicago, negative for Miami. The blue line corresponds to the point difference throughout game time. The green line represents the passes, the black one, screens and the yellow one space creations. The blue points above represent the player substitutions. As in the previous case, the note that the frequencies of passes, screens and space creations match with some key game situations and some score runs. In the Miami case, we can observe the most part of the game are in a disadvantage situation, so the frequencies are high. But in some occasion, such as around five, fourteen, twenty, forty five minutes, etc. values are low (lower frequencies) and match with a positive score run or score advantage.

As we can see in the Figures 58 and 59 the frequency of passes, screens and space creations do not remain stable but change according game situations (score evolution) throughout game time. This is a remarkable behavior, and point out a possible example of self-organization. Moreover, accomplish a pass is not so easy. The aim of defense is to avoid the progress of the ball, or rival players to the basket. It similar to slider-block models of earthquakes where to elements try to fight for a space through local interactions and should therefore display self-organized criticality.

In fact, the team success (attack or defense) depends on those action sequences. The Figures 52, 53, 54, 55, 56, and 57 show us that there are some truncated *Power Laws* regarding players´ interactions (passes, screens and space creations). If we consider team as a player network, we can deduce that if the team remains in the first part of the *Power Law* (before the first tipping point) team behaves as a small-world network. They network is connected by a short path length and they are able to resolve the situation despite opposition of rival team. The small-world networks can preserve its inherent structure despite a substantial number of shocks or attacks or disturbances (Uzzi et al., 2007). This can be interpreted as that, indeed, when the team is in a stable situation, is suffering constantly attack the other team, but maintains its structure throughout the game.

But when team exceeds the first tipping point, the *Power Law* indicates that there are nodes more connected than others. So, game flow is focused in some players. The team becomes from a diffuse flow to a concentrated flow, even hierarchical. Now they behave as a scale-free network.

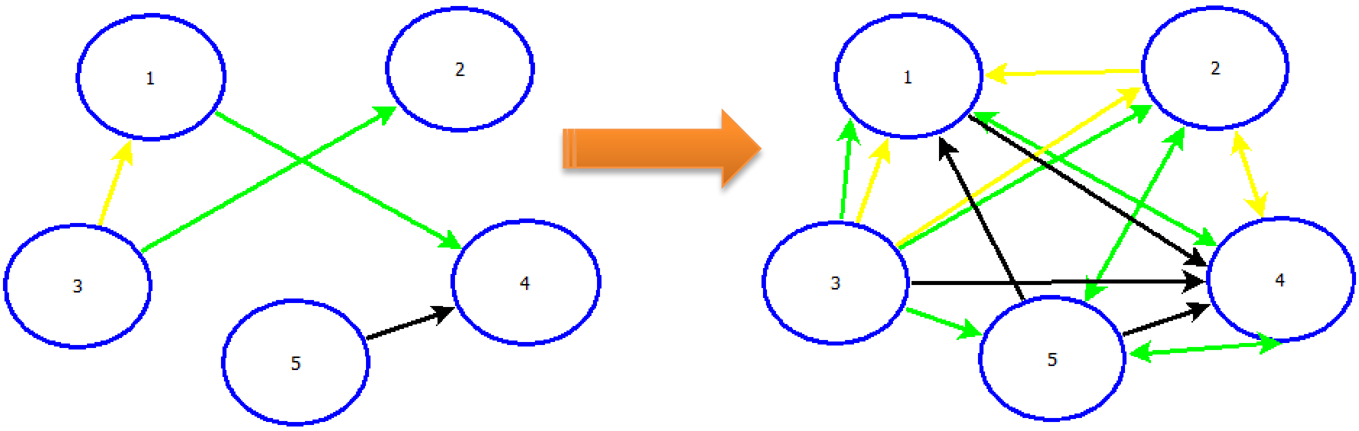


Figure 60. Evolution of teams from Small-world network to Scale-free network. They resolve the game situation with only a few steps, but if the situation becomes more critical, they modify the game flow to a Scale-free network.

This does not necessarily means that the player, where the game is focused, is the player who carries out shoots. It also can be a game distributor or game creator. Or just an emergency from the game (caused by time or specific game situation).

Hence, we can see how scoring influences in game dynamic so much so that team modifies its internal flow to try overcome the situation. Team behaves as a System Organized Critically (SOC).

8.4 Conclusions

Some parameters such as passes, screens and space creations, interpreted as interactions among players, can be used in order to analyze the dynamic and the flow of the game in a real basketball game, understood as a dynamic network.

These parameters follow a homogeneous distribution until a certain value. Once this value is exceeded, these dynamics turn into a *Power Law* distribution. The fact that some *Power Laws* are disclosed can mean that player interactions change substantially. The variations in these factors are closely related to scoring and its evolution in time (score difference), which works as order parameter. The behavior as dynamic network can point out some team features, and even the possible team profile, meaning team strategies, game systems, etc. Therefore, using network theory we can observe changes in internal team dynamic, regarding game situation (score mainly).

Teams try to defeat their opponents. A basketball game can be interpreted as a critical self-organized system because game flow does not remain stable and teams (players as networks) try to overcome the situation by self-organization.

8.5. Practical Proposals

As a practical proposal we suggest to use this methodology in order to study basketball teams as a whole. This methodology can provide some ideas in order to develop tactical systems, opponent analysis, players trustworthiness, etc. because it can provide us a precisely idea about how game flow is built and how is performed.

Also, it can be used to make a basketball scouting report. To define teams profiles. Or even define some behaviors used in concrete situations, such as fastbreaks, last second strategies, end-game strategies, etc.

8.6. Limitations of the techniques used

We think that the main difficult to use this methodology is counting the items we want to evaluate. Sport software can be used, but is still needed counting the number of times that a situation is produced. Moreover, the most part of these actions take place at the same time, simultaneously. Hence, it is necessary a lot of time to carry out this analysis, as well as to interpret the results.

General Conclusions

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

9. General conclusions

The final conclusions obtained through the studies that form this thesis are as follow:

1. Taking into account the behavior of the different levels analyzed (league, game and team), we can say that basketball can be considered as a Complex System.

2. The competitiveness degree does not remain stable during the course of time. It varies through season analyzed and these variations match with relevant event concerning basketball structures (league enlargement, rule modifications, important sport events, etc.). This fact points out that modifications in the participating structures in basketball as a system, modify the entire system and its operation.

3. Some structures, which take part in basketball as system, try to rule over the others, meaning they try to make system be hierarchical (teams). But the emerging systems (leagues) try to keep equality among them (be more random), in order to preserve attractiveness.

4. Score behaves as an order parameter. The presence of critical dynamics: the accelerations, critical slowing downs (pertubations), and the high degree of randomness that exists in most of the games, suggests that we are facing to phase transitions and critical exponents.

5. Scoring time interval on basketball does not follow a uniform behavior, but there are patterns in terms of time ranges. These temporal asymmetries indicate that basketball score has a non-linear nature. It seems that teams generate complex outputs, which are manifested as the ways that score evolves.

6. Teams try to defeat its opponent. A basketball game can be interpreted as a critical self-organized system. Game flow does not remain stable. Using network theory we can observe changes in internal team dynamic, regarding game situation (score mainly).

Future Research Lines

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

10. Future Research Lines

The future research lines proposed from this study are:

- Figure out whether the events that match with the significant skewness are actually the source of these modifications in the competitiveness degree.

- Find out whether teams, as complex systems, generate phenomena of learning and memory.

- Investigate the sources of randomness degree predominant in basketball scoring that reign during the basketball game.

- Figure out if the variations according network theory, are caused by game strategy or another source.

Spanish Abstract. Resumen en español

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

11. Resumen en español

Introducción

Sería interesante conocer por qué un jugador aumenta su rendimiento deportivo cuando cambia de equipo o cuando cambia de liga. O por qué el baloncesto ha tenido éxito en la atracción de medios de comunicación o aficionados respecto a otras modalidades deportivas. O incluso a averiguar por qué la liga profesional de baloncesto española es una de las más competidas de Europa. También sería interesante descifrar el flujo de juego durante un partido. Averiguar cómo evolucionan los marcadores, o por qué la liga ACB es la segunda que más jugadores aporta a competiciones internacionales, y por qué esta tendencia va en aumento. Averiguar a su vez qué hace a un equipo mejor que otro.

Aunque el deporte pueda parecer algo simple, meter la pelota en un aro, correr más rápido que un adversario, levantar más peso que mi contrincante, la efectividad rara vez muestra un comportamiento linear. En realidad, hay muchas acciones que pueden considerarse adecuadas. El rendimiento en baloncesto no atañe a un solo factor, sino a un compendio de numerosos elementos que influencian los unos sobre los otros. El sistema deportivo, la interacción de los jugadores, el balón, los árbitros y otros muchos aspectos determinan el resultado final.

El rendimiento deportivo, entonces, es el resultado de la combinación de numerosas variables. Algunas veces conocidas y otras no. En ocasiones, y a través de métodos analíticos tratamos de entender mejor este fenómeno con la intención de mejorar. Nuestra intención con este trabajo es expandir nuestros horizontes de conocimiento. Tratar de resolver cuestiones que ni tan siquiera nos habíamos planteado anteriormente. La teoría de la complejidad ha demostrado ser una herramienta de gran ayuda para la consecución de estas premisas.

Sistemas complejos

Un sistema complejo resulta de la conjunción de varios elementos, también llamados agentes, los cuales están relacionados entre ellos, y cuyas conexiones contienen información oculta al observador. Estas relaciones establecidas entre ellos son de manera no lineal. Y suelen ser además, de tipo local, es decir un elemento con el o los elementos colindantes. Estas relaciones afectan a las relaciones entre ellos pero ninguno es consciente del comportamiento colectivo (Goodwin, 2002; Vicsek, 2002; L. A. N. Amaral & Ottino, 2004; Solé, 2009).

Estos procesos que tienen lugar simultáneamente a diferentes escalas, son importantes. De hecho, la manera en que las sus unidades se relacionan influye sustancialmente en la respuesta de todo el sistema. Es por ello que algunos autores apuntan que las leyes que describen el comportamiento de un sistema complejo, son cualitativamente diferentes de aquellas que gobiernan sus unidades (Amaral & Ottino, 2004; Vicsek, 2002).

Como resultado de estas interacciones, emergen nuevas propiedades que no pueden ser entendidas desde las características individuales de cada elemento. Estas nuevas propiedades son llamadas *propiedades emergentes*. Es por esto que un sistema complejo debe ser tratado como un todo, desde una concepción holística. No simplemente como los elementos que lo constituyen, porque en un sistema complejo el todo es más que la suma de sus partes. La complejidad es el resultado de incesantes procesos adaptativos (Holland, 1995).

No linear

Cuando un sistema es linear, el mismo estímulo siempre produce la misma respuesta. Cada vez que el proceso se repite, la misma respuesta tiene lugar. Por otro lado, si el sistema es no linear, un estímulo puede conducir a varios resultados. Aunque las condiciones sean las mismas, la respuesta o respuestas no pueden ser conocidas de antemano (Prigogine & Holte, 1993; Solé & Goodwin, 2002; Amaral & Ottino, 2004).

Las interrelaciones de los componentes del sistema complejo son gobernadas por ecuaciones de tipo no linear. Recordemos que la efectividad en deporte, rara vez muestra un comportamiento lineal, porque existen multitud de acciones que pueden considerarse efectivas, y además no tienen que ser consecutivas. La complejidad en sí misma es una medida del número de posibilidades. Existe, además, una gran dependencia de las condiciones iniciales, lo cual hace más difícil si cabe, predecir y evaluar su comportamiento.

Autoorganización

La idea de la autoorganización puede ser expresada como la tendencia general de un sistema dado a generar patrones de comportamiento a partir de las interacciones locales de sus elementos constitutivos y de las relaciones con el medio ambiente. Es la parte esencial de cualquier sistema complejo y permite que el sistema se recupere el equilibrio, modificado y adaptado al medio ambiente circundante. Por lo general, los diferentes elementos del sistema se autorregulan por sí mismos, siempre buscando optimizar el funcionamiento global del conjunto. La compleja red de sistemas interdependientes en que los seres humanos se pueden organizar, por ejemplo, está cambiando y readaptando a la realidad que les corresponde vivir en cada momento (García Manso and Martín González, 2008).

La autoorganización es un proceso en el que la organización interna de un sistema aumenta en complejidad sin ser guiado o dirigido por una fuente externa. Sistemas de autoorganización suelen mostrar *propiedades emergentes*.

El orden y el desorden se necesitan mutuamente, se producen mutuamente. Son conceptos antagónicos pero complementarios, al mismo tiempo. En algunos casos, alguno de desorden permite un orden diferente y, a veces, más rico. Por ejemplo, un organismo puede persistir como resultado de la muerte de sus células, o una organización es perpetuada por la destitución de sus miembros. La variación y el cambio son etapas inevitables e ineludibles a través del cual todo sistema complejo debe viajar a crecer y desarrollarse. Cuando se logra esta transformación sin la intervención de factores externos al sistema, se habla de un proceso de autoorganización (Nicolis & Prigogine, 1977).

La autoorganización se destaca como una parte esencial de cualquier sistema complejo. Es la forma a través del cual el sistema se recupera el equilibrio, el cambio y la adaptación al medio ambiente circundante (que responde a las agresiones externas que tratan de modificar su estructura).

En este tipo de fenómenos es esencial la idea de los niveles. Las interrelaciones entre los elementos de un nivel originan nuevos tipos de elementos en otro nivel que se comportan de manera muy diferente, por ejemplo, de moléculas a las macromoléculas, de macromoléculas a las células y de las células a los tejidos. Por lo tanto, el sistema autoorganizado se construye como resultado del orden creciente espacio-tiempo que se crea en diferentes niveles o capas, una encima de la otra.

En gran medida, los sistemas complejos pueden entenderse como una máquina que genera orden, lo que requiere la ingesta de energía constante generada por el caos que alimenta (que es un sistema abierto y disipativo). Los sistemas complejos autoorganizados se consideran adaptativos, ya que pueden reaccionar a los estímulos externos y responder a cualquier situación que amenace su estabilidad como sistema. Por lo tanto, experimentan fluctuaciones. Esto tiene un límite, por supuesto. Se dice que el sistema se asienta en un estado y cuando está lejos de él tiende a hacer todo lo posible para volver a la situación anterior. Esto sucede por ejemplo con el cuerpo humano se esfuerza constantemente por mantener la misma temperatura corporal.

Estado Crítico

La idea general del sistema en un estado crítico puede ser entendida como un estado cercano a la frontera de otro estado (**punto crítico**). Lo que significa que cualquier pequeña perturbación, puede dar lugar a un nuevo estado (**transición de fase**).

Uno de los más famoso ejemplo es el modelo de pila de arena por Bak-Tang-Wiesenfeld Wiesenfeld (Bak et al., 1987). El modelo describe cómo un montón de arena se acumula en forma de granos de arena colocados al azar en una pila. En las pequeñas perturbaciones que comienzan sólo causan respuestas pequeñas. Avalanchas pequeñas tienen lugar hasta que la pila llega a un estado crítico en el que la pendiente oscila alrededor de un ángulo constante de reposo (umbral o punto crítico). Si añadimos un grano de arena más, esto puede causar que la pendiente supere el valor crítico y se origine una gran avalancha. La variación de las pendientes locales hace imposible predecir cuando este fenómeno se llevará a cabo.

Los sistemas críticos se caracterizan por estar en un estado delicado equilibrio que, a su vez, está vinculada con el medio ambiente, mostrando una gran sensibilidad (Jost, 2005). Esta situación les confiere un comportamiento altamente impredecible (caótico, no al azar).

La mayor parte de los sistemas complejos son inestables (están alejados del estado de equilibrio). Esto implica que los sistemas no pueden sostenerse a menos que reciban un suministro constante de energía (orden necesita el caos y el caos necesita el orden. Ellos no pueden existir el uno sin el otro, como mencionamos anteriormente). Exigen ajustes siguiendo patrones específicos. Cualquier variación mínima entre los elementos que lo componen puede modificar de manera impredecible, las interrelaciones y por lo tanto, el comportamiento de todo el sistema. Por consiguiente, la evolución de este tipo de sistemas se caracteriza por la intermitencia o la variación (situación en la que el orden y el desorden se alternan constantemente). Sus estados evolutivos no pasan por proceso continuo y gradual, sino que se producen a través de reorganizaciones y saltos. Cada nuevo estado es sólo una transición, un período de reposo entrópico, en palabras del Premio Nobel ruso-belga Ilya Prigogine (Prigogine & Stengers, 1984; Prigogine & Holte, 1993).

Estos sistemas nunca llegan a un óptimo global, al estado de energía mínima. En general, crecen gradualmente hasta que alcanzan el límite de su potencial de desarrollo. En ese momento, sufren un trastorno, una especie de ruptura que induce una fragmentación del orden pre-existente. Pero entonces, comienzan a surgir regularidades que organizan el sistema de acuerdo con las nuevas leyes, produciendo otro tipo de desarrollo. Este comportamiento es típico en los sistemas naturales: por ejemplo, el tránsito de los insectos, de huevo a larva y de allí a la crisálida. En consecuencia, la organización de los sistemas complejos se da en diferentes niveles. Las leyes que rigen la causalidad de un nivel dado pueden ser totalmente diferentes a las de un nivel más alto o bajo (Kauffman, 1995; Bak, 1999).

Sistemas autoorganizados y deporte

En este tipo de situaciones límite, en el deporte, los deportistas y su entorno tienen que hacer un gran esfuerzo con el fin de superar las circunstancias. En ese momento es cuando realmente se puede aprender. Es en este momento cuando los sistemas deportivos crean nuevas estrategias, planes de entrenamiento y, por lo tanto, es cuando se desarrollan, cambian o se comportan de acuerdo a la nueva realidad. Es decir, la rivalidad y la competitividad son los elementos que generan el comportamiento crítico.

Cuando el sistema está autoorganizado críticamente, la información fluye mejor entre todas las partes del sistema (Solé, 2009). Por otra parte, este tipo de sistemas tienen memoria y los mecanismos de regulación ajustan la respuesta a la demanda. Estos sistemas evolucionan tratando de optimizar sus recursos y tienden naturalmente a ser en estos estados, por lo tanto, sirven como atractores del sistema (Ivancevic and Ivancevic, 2006). Es decir, el funcionamiento del sistema es la clave y no las características individuales de sus elementos.

Todos sabemos lo que es realmente interesante en el deporte es la competición. La competencia atrae a grandes masas de público, medios de comunicación y, con frecuencia, grandes cantidades de recursos financieros. Por lo general, esto conduce a los deportes (especialmente en elite) a jugar en un área crítica (García Manso and Martín González, 2008), en el límite del error, arriesgando, compitiendo cerca del límite.

Este fenómeno favorece que el deporte evolucione. Los jugadores cambian su estilo de juego, los equipos cambian las tácticas, la dinámica de juego cambia, así, nuevas metodologías de formación surgen con el fin de apoyar los requisitos de competición, etc. E incluso podemos ver cómo algunos deporte introducir nuevas reglas (o modificar las reglas antiguas) con el fin de mantener el atractivo de la liga.

Algunas reglas, como el fuera de juego en rugby, los 24 segundos de posesión en baloncesto, los tres toques en el voleibol, el robo de bases en el béisbol, etc. son intentos de llevar a los deportes hacia las áreas críticas. Porque el deporte se adapta y la tendencia natural conduce a una jerarquización más o menos definida. De ahí, los esfuerzos de algunos deportes con el fin de evitar este tipo de situaciones.

Como hemos mencionado anteriormente, los agentes que participan en estos sistemas deportivos, compiten entre ellos, y la tendencia natural conduce a estructuras jerárquicas, donde algunos equipos son claramente superiores a los demás. Teóricamente, esta situación podría extenderse en el tiempo y casi no puede revertirse por medios naturales, debido a mejores equipos seguirán acaparando los mejores recursos. Este fenómeno se conoce como **ventaja acumulativa [Preferential Attachment]** (Barabási & Albert, 1999), efecto **bola de nieve** o efecto **San Mateo**. Es el popular el rico se hace más rico, y el pobre se hace más pobre.

Por lo tanto, teóricamente, podemos señalar que esta situación se mantendrá mientras que no haya ninguna fuente externa que modifique el ambiente en el que el deporte se está desarrollando (reglas, modelo deportivo de competición). Es por eso que es tan importante entender el funcionamiento del sistema deportivo, es decir, la liga, el equipo de juego, etc. y cómo las modificaciones (reglas, nuevos elementos, etc.) afectan a todo el sistema.

La creación o modificación de estos sistemas suelen seguir ciertas leyes, lo que significa que algunos de estos fenómenos presentan las mismas características. Uno de los ejemplos más importantes es la aparición de **leyes de potencia** o distribuciones de cola pesada. Muchos fenómenos naturales siguen este tipo de distribución, a menudo fractal, que también son evidentes en muchos sistemas no naturales. Una gran cantidad de elementos interactúan para producir una estructura de nivel superior. Estos sistemas se desarrollan lejos del equilibrio y son a menudo altamente disipativos (sistemas lejos del equilibrio). Las leyes de potencia presentan dos características principales: su transformación logarítmica se trasforma en una línea recta y es invariante de escala.

Es por ello que son llamados también libres de escala. Entendiendo por escala la dimensión espacial y temporal del fenómeno. La hipótesis de escala que se eleva en el contexto del estudio de los fenómenos críticos condujo a dos categorías de predicciones, ambos de los cuales han sido bien verificada por una gran cantidad de datos experimentales en varios sistemas. Uno de los más importantes es la ley de escala que hemos mencionado; su utilidad radica en la vinculación de los diversos exponentes críticos que caracterizan el comportamiento singular del parámetro de orden y funciones de respuesta (Amaral & Ottino, 2004).

Además, este tipo de distribución puede señalar fenómenos tales como fractalidad (Barabási & Albert, 1999), autoorganización (Dhar, 1990; Bak, 1999), agrupamientos (clustering) (Newman, 2001a; Albert & Barabási, 2002), leyes alométricas (West et al., 1997a), etc. En resumen, indican la posible presencia de sistemas complejos.

También en el deporte hay una gran cantidad de ejemplos de este tipo de distribuciones: records en atletismo (Katz & Katz, 1999; Savaglio & Carbone, 2000), power lifting (García Manso et al., 2008), distribuciones de goles (Malacarne & Mendes, 2000; Mendes et al., 2007), permanencia de directivos (Aidt et al., 2006), anotación en baloncesto (de Saá Guerra et al., 2013), etc.

Como podemos ver el deporte, en general, es un buen ejemplo de complejidad, por lo tanto, creemos adecuado utilizar esta metodología para analizar el baloncesto.

Baloncesto como sistema complejo.

Los sistemas complejos son el resultado de un proceso evolutivo. Las ideas de Darwin y el estudio de la evolución se han centrado en la competición como fuerza conductora de los cambios evolutivos. Los jugadores cuando cooperan, compiten mejor como equipo. De hecho, la cooperación o la oposición de los jugadores (los atractores que forman el sistema), es lo que produce diferentes niveles o escalas en las que el deporte se construye. La presencia de estos dos comportamientos es lo que permite al baloncesto, y a sus elementos, evolucionar.

Los jugadores compiten por un puesto en el equipo. Esto hace que mejoren. Pero la cooperación entre ellos es lo que permite al equipo competir. Por tanto, tenemos dos comportamientos posibles. El mismo elemento puede mostrar dos propiedades diferentes, dependiendo del tipo de interacción. Cuando un jugador coopera es una relación sinergista. Cuando los jugadores compiten, es una relación antagónica. La competición se da solamente cuando hay cooperación; y la mejora se da exclusivamente cuando hay oposición.

Como hemos mencionado anteriormente, un equipo no es sólo el resultado de la interacción de sus jugadores. Necesita un entorno donde desarrollarse y evolucionar hacia nuevos estados. Si se comporta como asumimos, como un sistema crítico auto-organizado, las inestabilidades y saltos hacia nuevas formas son el resultado de fluctuaciones internas y de interacciones con el entorno.

Llegados a este punto, aparece la figura del entrenador como elemento que influencia directamente sobre el resultado final. Un cambio de entrenador o un cambio en la plantilla, puede resultar en una respuesta completamente diferente. Otro ejemplo claro de modificación del entorno son las modificaciones reglamentarias. Las reglas proveen un entorno artificial y un sistema artificial de información. Existen límites espaciotemporales, así como de interacción, reguladas por el reglamento, de manera que cualquier modificación de las reglas puede conducir a modificaciones sustanciales en la dinámica de juego. Es por ello que hemos considerado interesante estudiar el baloncesto desde tres niveles diferentes:

- Estudio 1. La competición (liga).
- Estudio 2. El partido.
- Estudio 3. El equipo.

Liga

El objetivo de este estudio fue desarrollar una herramienta que nos permitiese evaluar el grado de competitividad basado en la incertidumbre que existe en cada confrontación. Nosotros calculamos el valor de la entropía de Shannon, la cual cuantifica la información contenida por una variable, para determinar el grado de incertidumbre o aleatoriedad que existe en la competición. De esta manera usamos el concepto de competitividad como indicador relativo de calidad.

Este análisis nos permite identificar las posibles causas del incremento o descenso de la competitividad durante varias temporadas, e incluso comparar diferentes ligas entre si. Nosotros hemos analizado la fase regular (no play-off) de la NBA y la ACB. Este análisis además puede señalar posibles influencias que otros sistemas pueden tener en el desarrollo de las mismas, tales como estructura económica, organización competitiva, fuente de jugadores, influencias de otros sistemas no deportivos

Partido

La anotación en baloncesto es un proceso altamente dinámico y de tipo no-linear. El nivel de los equipos tiende a mejorar cada temporada, ya que tratan de incorporar a sus plantillas los mejores jugadores disponibles, o los que su presupuesto les permite. Esto y otros mecanismos hacen que la anotación en baloncesto sea algo excitante e impredecible.

Hemos estudiado el comportamiento de la evolución de los marcadores en 5 temporadas de la NBA, con un total de 6150 partidos. Más concretamente las diferencias de puntos, los intervalos temporales de anotación, así como su evolución en el tiempo.

Equipo

Hemos querido realizar una aproximación científica hacia el funcionamiento de los equipos, tratándolos como redes de jugadores. Partimos de la idea de que la clave del éxito de un equipo es la autoorganización, la manera en la que los equipos logran sus objetivos y se sobreponen a amenazas externas. Esto se consigue a través de interacciones locales para compensar el desequilibrio ocasionado por estos agentes externos. De manera que el equipo funciona como un todo. Todos sus elementos trabajan de manera coordinada y eficaz. No podemos entender al equipo como la mera suma de sus jugadores, sino como algo más.

Hemos estudiado la manera de relacionarse de los jugadores en la cancha en relación a parámetros del juego tales como pases, bloqueos y creaciones de espacio. Además hemos relacionado estas acciones con la evolución temporal del marcador.

Objetivos

Los objetivos de esta tesis son:

1. Identificar las estructuras organizativas (ligas, federaciones, clubs, etc.), las estructuras participativas que conforman el sistema y los comportamientos resultantes de sus interacciones.

2. Averiguar si alguno de estos patrones sigue alguna ley conocida y determinar el significado que estas tienen en el baloncesto.

3. Tratar de modelar comportamientos institucionales, grupales y/o colectivos que se puedan dar en la organización y la práctica del baloncesto.

Hipótesis

Nuestra hipótesis para esta disertación es la siguiente:

El baloncesto, desde un punto de vista institucional y práctico, se comporta como un sistema complejo crítico autoorganizado durante la fase regular de la liga.

Relevancia del trabajo

La novedad de esta tesis reside en la aplicación de los sistemas complejos como herramienta al deporte, pero sobretodo, los resultados y las conclusiones que se pueden obtener con este tipo de técnicas de investigación.

El segundo punto que queremos destacar es la extensión y profundidad de la tesis, ya que hemos abordado el tema a estudiar desde varios niveles (liga-partido-equipo), lo que nos da una perspectiva integral del sistema. Esto nos permite conocer los tipos de relaciones que pueden existir entre los elementos que participan en el sistema.

Finalmente nos gustaría enfatizar la aplicabilidad de esta tesis en varios campos del conocimiento. Uno de ellos puede ser el desarrollo de estrategias de intervención para la creación de sistemas competitivos, administración y enseñanza dirigida a los profesionales de este campo. E incluso para personal ajeno al deporte en sí, pero con vinculación al deporte, tales como sistemas económicos, administrativos, etc. En el campo académico también es relevante debido a que podemos aprender y aplicar estas nuevas técnicas a otros campos del conocimiento. Abrir nuevas líneas de investigación que pueden ser seguidas por otros investigadores y transmitir esos descubrimientos y metodologías a la comunidad universitaria y científica.

Estudio 1. Liga de Baloncesto

Hemos estudiado los resultados de diferentes temporadas de dos de las mejores ligas profesionales de baloncesto, la NBA (National Basketball League, USA) y la ACB (Asociación de Clubes de Baloncesto, España), y los resultados de una liga amateur de alto nivel, la División I de la liga universitaria norteamericana (NCAA, National Collegiate Athletic Association, USA).

La liga ACB es un modelo abierto donde existen ascensos y descensos de categoría. Cada año los equipos participantes varían en función de estos ascensos y descensos de categoría. Los ocho equipos mejor clasificados juegan una eliminatoria para proclamarse campeones de liga.

La NBA es un modelo de franquicias. Los equipos participantes son divididos en dos conferencias (Este y Oeste), y a su vez estas son organizadas en tres divisiones por conferencia. Al término de la fase regular, los mejores equipos compiten en una eliminatoria dividida por conferencias. La NBA es un modelo cerrado donde no hay ascensos y descensos de categoría.

La liga universitaria norteamericana está dividida en tres divisiones (División I, II y III). A su vez, cada división está dividida en conferencias de varios equipos cada una. Nosotros sólo usamos los datos de la División I. Hemos de recordar que la División I de la NCAA está compuesta por un total de 344 equipos (aunque el número varía ligeramente en las temporadas estudiadas), divididos en 31 conferencias a través de todo el territorio de Estados Unidos (el número de equipos por conferencia no es homogéneo).

El objetivo de este estudio fue analizar, desde una perspectiva global, diferentes modelos deportivos y la dinámica interna de varias ligas (profesionales y amateur), mediante el análisis del grado de competitividad. También tratamos de desarrollar un modelo para el estudio del nivel de competitividad en competiciones deportivas, el cuál puede ser útil para evaluar tales competiciones mediante la incertidumbre que debe existir para cada confrontación.

Como criterio metodológico utilizamos, en cada caso, una matriz de confrontaciones en las que los resultados de los enfrentamientos pueden ser múltiples, es decir, el número de victorias o derrotas de cada equipo puede presentar diferentes combinaciones, de esta manera, podemos calcular el valor de entropía de Shannon para determinar el grado de incertidumbre.

El vector de resultados (***R***) representa el resultado obtenido por cada equipo en cada temporada. ***R***, en principio, se comporta de manera aleatoria, en el sentido que no conocemos el resultado final, pero los resultados de temporadas anteriores (histórico de resultados), nos pueden brindar algún indicio. Los valores de ***R*** históricos o de temporadas anteriores divididos por la suma del total de partidos, pueden representar también una distribución de probabilidad discreta.

Cuando el conjunto de probabilidades de un sistema es conocido, podemos definir la entropía de Shannon (***S***), que es una medida de la incertidumbre promedio y, por tanto, hace referencia a la cantidad media de información que contiene una variable aleatoria. Siendo máxima cuando todos los valores p_i sean iguales. El valor de ***S*** cambia con el valor de *N*, número de equipos, y por tanto no son comparables si en temporadas diferentes el número de equipos cambia. Por ello es preferible utilizar la entropía normalizada (***Sn***). De manera que el máximo valor de *Sn* está acotado entre 0 y 1, donde 1 corresponde a la situación en la que todos los valores p_i son iguales.

Una liga es más competitiva cuando es más aleatoria. Cuando es más complicado dilucidar el resultado final. Sin embargo, cuando la competición es menos aleatoria, el grado de competitividad decrece significativamente. Tanto la ACB como la NBA muestran un alto grado de competitividad. En ambas ligas los niveles de entropía son elevados (rango: 0.9851 a 0.9902). Aunque estos períodos son más estables en la NBA. En cuanto a la liga NCAA, hemos de tener en cuenta el gran número de equipos participantes, de ahí, la gran heterogeneidad de la liga y de los equipos (presupuestos, jugadores, instalaciones, etc.). Los valores de *Sn* oscilan entre 0.9679 hasta 0.9583. Estos valores se encuentran bastantes alejados de las ligas profesionales, pero a pesar de este hecho, la liga NCAA es la más estable de las tres (NCAA *Sn* media=0.9631 ± 0.0033).

En conclusión podemos decir que tanto la ACB como la NBA son ligas muy competitivas y cuyos equipos están muy equilibrados entre ellos. Mientras que la liga universitaria norteamericana, a pesar de presentar unos valores inferiores, muestra una tendencia mucho más constante a lo largo de las temporadas analizadas.

Estudio 2. El partido de Baloncesto

Como hemos visto en el estudio anterior, el grado de incertidumbre de una liga de baloncesto muestra un comportamiento no lineal. Esto puede depender de la igualdad que existe entre los equipos participantes. La evolución del marcador y el resultado final son lo que generan incertidumbre para cada partido, y por ende, para la clasificación en la liga.

El marcador de un partido de baloncesto es el reflejo directo de la dinámica y de la interacción no lineal de los equipos y sus componentes. Sin embargo, la interacción de los elementos que componen el deporte parece poseer ciertos patrones o propiedades que le confieren características propias de cada deporte. Nosotros hemos tratado de identificarlas con la intención de conocer en profundidad la lógica interna de la competición.

Hemos estudiado un total de 6150 partidos (5 temporadas) de la fase regular de la liga NBA. Estudiando los tiempos de anotación, los cuales pueden ser considerados como un proceso aleatorio. Esto significa que pueden ser tratados como un proceso de Poisson, es decir, como un proceso de llegadas. Nuestro objetivo fue estudiar los intervalos de tiempo, así como la anotación, con la idea de identificar patrones o regularidades básicas que nos fuesen útiles para entender mejor la dinámica de los partidos de baloncesto.

En baloncesto el tiempo entre canastas sigue un proceso aleatorio, tal y como sugieren nuestros resultados. Esta tendencia está definida por λ, que es la relación entre el número de eventos y el tiempo en el que tiene lugar. La diferencia de los tiempos entre canastas presenta un pico en torno a los 20 segundos, y un valor máximo de 310 segundos, en la muestra analizada. De esto podemos deducir que el rango temporal más probable entre puntos está en torno a los 20 segundos. Y que casos de 5 minutos sin anotar son muy extraños y poco probables, pero pueden suceder.

El índice de dispersión (la relación entre la varianza y la media), muestra que la mayor parte los cuartos permanecen por debajo del valor 1 (sub-disperso). Solamente al final de cada curato se aprecia un incremento significativo en esto valores, pero siempre por debajo del valor 1. Esto significa que el comienzo de cada cuarto es más predecible que el final. Tan sólo en el minuto 47 se alcanza el valor 1, lo que indica un proceso puro de Poisson.

El minuto 48 requiere especial atención, ya que excede significativamente el valor 1. Esto sugiere que el último minuto del partido sigue una dinámica completamente diferente que el resto del partido.

En cuanto al resultado final, la mayor parte de los partidos acabaron con una diferencia menor o igual a 10 puntos (65%), un 33% de los partidos finalizaron con una diferencia de entre 11 y 28 puntos, y tan solo un 2% lo hizo con diferencias superiores a 28 puntos. Por consiguiente, podemos distinguir tres perfiles de partidos, en base al resultado final. Los más competidos, con una diferencia menor de 11 puntos, los más impredecibles. Otros menos competidos, entre 11 y 28 puntos, y partidos completamente desequilibrados, de más de 28 puntos de diferencia, donde existe una clara superioridad de un equipo sobre el otro. En este caso el resultado final del partido es bastante predecible.

Podemos concluir que el la diferencia de puntos puede usarse como un indicador de la dinámica del partido, ya que funciona como parámetro de orden. El marcador es un reflejo de las diferentes acciones y comportamientos que resultan del juego y de la interacción de los jugadores. Dado el alto grado de aleatoriedad que existe en la mayor parte de los partidos con menos de diferencia 11 puntos, se puede deducir que la mayoría de los partidos tienen un alto grado de incertidumbre. Por lo tanto, es muy difícil saber de antemano quién será el ganador.

De esta manera la NBA puede ser considerada como un ejemplo de *Hipótesis de la Reina Roja*, propuesta por Van Valen (1973): “Para un sistema evolutivo, la mejora continua es necesaria para sólo mantener su ajuste a los sistemas con los que está co-evolucionando”. Esto es, una carrera sin fin. Donde todos los competidores necesitan mejorar sólo para permanecer compitiendo.

Estudio 3. El equipo de Baloncesto

Como última etapa de nuestra investigación, hemos querido hacer una aproximación a la dinámica de un equipo de baloncesto. Nuestra intención era averiguar el proceso interno de un equipo de baloncesto durante un partido de baloncesto real. Estamos hablando del flujo del juego y el diseño de la estructura, que puede ser descrito como en muchos otros sistemas naturales. Podemos ver cómo la estructura, la forma y la funcionalidad están estrechamente relacionadas en varios sistemas del deporte.

La teoría de redes se aplica en diversos campos del conocimiento y de estudio como la biología, matemáticas, economía, ecología, física, sociología, ingeniería, y como no, en las ciencias del deporte. El primer antecedente del que se tiene constancia en el campo científico es el famoso problema de los siete puentes de Königsberg, planteado por el matemático Leonhard Euler en 1736. Euler describió matemáticamente los vértices y conexiones necesarias para resolver el problema, de manera que se estableció la teoría de grafos, una rama de las matemáticas que estudia las propiedades de las estructuras de red.

En 1999 se introdujo el concepto de red libre de escala (scale-free network) propuesta por Albert-László Barabási y Reka Albert, al cual se le conoce como el modelo Barabási–Albert (Barabási & Albert, 1999). Este modelo explica cómo se forman al azar redes libres de escala mediante un mecanismo denominado ventaja acumulativa (preferential attachment). Las redes libres de escala son ampliamente observadas en los sistemas naturales; y provocadas por el hombre, incluido Internet, las redes de citas bibliográficas y algunas redes sociales.

Por lo tanto, nosotros proponemos el uso de la teoría de redes como medio para averiguar las características del equipo de baloncesto, entendido como una red de jugadores. Esta metodología nos puede proporcionar la oportunidad de descubrir cómo funciona un equipo de baloncesto a través del comportamiento de los jugadores. A través de sus interacciones reales.

La mayoría de los estudios frente a este fenómeno en lo que respecta a los aspectos externos, por ejemplo la inclusión de un jugador en un equipo o una liga o un torneo. Hemos querido abordar el problema desde un punto de vista de colaboración. Desde el proceso interno y no desde aspectos externos.

No son muchos los ejemplos de la aplicación del estudio de redes en el deporte. Esto es debido quizás a que es un campo de aplicación en el que esta metodología es bastante reciente. Lo que debemos tener claro es que debe quedar bien definido el nivel en el que se aplica esta técnica de estudio, ya que no es lo mismo estudiar el comportamiento de los jugadores de un equipo, como la dinámica de los equipos en una liga.

Como objeto de estudio, tomamos el último partido de la final de conferencia Este de la NBA de 2011, que enfrentó a Chicago Bulls y a Miami Heat. Definimos como variables de estudio la interacción de los jugadores en la cancha, es por eso que medimos, como medio de comunicación de los jugadores en la cancha y como indicador del flujo de juego, el número de pases, los bloqueos y las creaciones de espacio para cada jugada.

Los pases representan el ejemplo más claro de la interacción de los jugadores en la cancha, porque el hecho de pasar el balón a un compañero de equipo permite crear una situación de tiro u otra situación favorable. Los bloqueos también representan un ejemplo interesante de la interacción de los jugadores porque el objetivo de un bloqueo es neutralizar a un defensor y/o buscar una situación de superioridad tras el bloqueo, tal como una penetración clara hacia canasta, un tiro claro, un aclarado, etc. Este parámetro está relacionado con el siguiente.

La creación de espacio representa también la interacción entre jugadores. Un buen ejemplo de este fenómeno es la situación de pick and roll (bloqueo y continuación), donde un jugador bloquea la defensor de un compañero, e inmediatamente, tras el bloqueo, avanza hacia la canasta.

Los resultados muestran que la frecuencia de pases, bloqueos y creaciones de espacio no permanecen estables, sino que varían en función de la situación del partido (evolución del marcador). Los equipos modifican su estructura de red en base a la situación a la que se enfrentan en una misma jugada o durante el partido, tal y como muestra la presencia de leyes de potencia truncadas. Estos puntos de inflexión coinciden perfectamente con situaciones bien definidas del partido.

En conclusión podemos decir que algunos parámetros, tales como pases, bloqueos y creaciones de espacio, pueden ser interpretados como interacciones entre los jugadores. Además, se pueden utilizar con el fin de analizar la dinámica y el ritmo del juego en un partido de baloncesto real, entendida como una red dinámica.

Los equipos tratan de derrotar a sus oponentes. Un partido de baloncesto se puede interpretar como un sistema auto-organizado en estado crítico porque el flujo de juego no se mantiene estable y los equipos (red de jugadores) tratan de superar la situación mediante la auto-organización.

Bibliography

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

Bibliography

ACB. (2012). ACB.COM. Recuperado 8 de mayo de 2012, a partir de http://www.acb.com/

Adamic, L. A. (1999). The Small World Web. En S. Abiteboul & A.-M. Vercoustre (Eds.), *Research and Advanced Technology for Digital Libraries* (pp. 443-452). Springer Berlin Heidelberg. Recuperado a partir de http://link.springer.com/chapter/10.1007/3-540-48155-9_27

Aidt, T. S., Leong, B., Saslaw, W. C., & Sgroi, D. (2006). A power-law distribution for tenure lengths of sports managers. *Physica A: Statistical Mechanics and its Applications*, *370*(2), 697-703. doi:10.1016/j.physa.2006.03.016

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47-97. doi:10.1103/RevModPhys.74.47

Amaral, L. A. N., & Ottino, J. M. (2004). Complex networks. *The European Physical Journal B - Condensed Matter*, *38*(2), 147-162. doi:10.1140/epjb/e2004-00110-5

Amaral, L. a. N., Scala, A., Barthélémy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences*, *97*(21), 11149-11152. doi:10.1073/pnas.200327197

Anderson, K. J., & Pierce, D. A. (2009). Officiating bias: The effect of foul differential on foul calls in NCAA basketball. *Journal of Sports Sciences*, *27*(7), 687-694. doi:10.1080/02640410902729733

Arjonilla López, N. (2011, abril 13). *Incidencia de los factores distancia, tiempo, fatiga y concentración de la efectividad en el baloncesto*.

Bak, P. (1999). *How Nature Works: the science of self-organized criticality.* (2.ª ed.). New York: Copernicus. Springer-Verlag.

Bak, P., Tang, C., & Wiesenfeld, K. (1987). Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, *59*(4), 381-384. doi:10.1103/PhysRevLett.59.381

Bar-Yam, Y. (2001). Introducing complex systems, presented at the International Conference on Complex Systems, Nashua, NH, 2001.

Bar-Yam, Y. (2003). Complex Systems and Sports. New England Complex Systems Institute. Recuperado a partir de http://necsi.edu/projects/yaneer/SportsBarYam.pdf

Barabási, A.-L., & Albert, R. (1999). Emergence of Scaling in Random Networks. *Science*, *286*(5439), 509-512. doi:10.1126/science.286.5439.509

Bejan, A., & Haynsworth, P. (2012). The natural design of hierarchy: Basketball versus academics. *International Journal of Design and Nature and Ecodynamics*, *7*(1), 14-26. doi:10.2495/DNE-V7-N1-14-26

Bejan, Adrian, Jones, E. C., & Charles, J. D. (2010). The evolution of speed in athletics: Why the fastest runners are black and swimmers white. *International Journal of Design and Nature and Ecodynamics*, *5*(3), 199-211. doi:10.2495/DNE-V5-N3-199-211

Bejan, Adrian, & Lorente, S. (2011). The constructal law and the evolution of design in nature. *Physics of Life Reviews*, *8*(3), 209-240. doi:10.1016/j.plrev.2011.05.010

Bejan, Adrian, & Zane, J. P. (2013). *Design in Nature: How the Constructal Law Governs Evolution in Biology, Physics, Technology, and Social Organization*. Knopf Doubleday Publishing Group.

Bennett, R. W., & Fizel, J. L. (1995). Telecast Deregulation and Competitive Balance. *American Journal of Economics and Sociology*, *54*(2), 183–199. doi:10.1111/j.1536-7150.1995.tb02689.x

Berri, D. J., Brook, S. L., Frick, B., Fenn, A. J., & Vicente-Mayoral, R. (2005). The short supply of tall people: competitive imbalance and the National Basketball Association. *Journal of Economic Issues*. Recuperado a partir de http://www.highbeam.com/doc/1G1-140240254.html

Bittner, E., Nußbaumer, A., Janke, W., & Weigel, M. (2009). Football fever: goal distributions and non-Gaussian statistics. *The European Physical Journal B - Condensed Matter and Complex Systems*, *67*(3), 459-471.

Blanco, E. (2006). *Manual de la OrganizaciÓn Institucional Del Deporte*. Editorial Paidotribo.

Bourbousson, J., Sève, C., & McGarry, T. (2010). Space–time coordination dynamics in basketball: Part 2. The interaction between the two teams. *Journal of Sports Sciences*, *28*(3), 349-358. doi:10.1080/02640410903503640

Brillinger, D. R. (2007). A Potential Function Approach to the Flow of Play in Soccer. *Journal of Quantitative Analysis in Sports*, *3*(1). doi:10.2202/1559-0410.1048

Bunge, M. (2001). El efecto San Mateo. *Polis, Revista de la Universidad Bolivariana*, (002). Recuperado a partir de http://redalyc.uaemex.mx/src/inicio/ArtPdfRed.jsp?iCve=30500225

Burroughs, S. M., & Tebbens, S. F. (2001). Upper-truncated Power Laws in Natural Systems. *pure and applied geophysics*, *158*(4), 741-757. doi:10.1007/PL00001202

Cairns, J., Jennett, N., & Sloane, P. J. (1986). The Economics of Professional Team Sports: A Survey of Theory and Evidence. *Journal of Economic Studies*, *13*(1), 3-80. doi:10.1108/eb002618

Charles, J. D., & Bejan, A. (2009). The evolution of speed, size and shape in modern athletics. *The Journal of Experimental Biology*, *212*(15), 2419-2425. doi:10.1242/jeb.031161

Chatterjee, S., & Yilmaz, M. R. (1999). The NBA as an Evolving Multivariate System. *The American Statistician*, *53*(3), 257-262. doi:10.2307/2686106

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM Review*, *51*(4), 661-703. doi:10.1137/070710111

De Saá Guerra, Y., Martín González, J. M., Arjonilla López, N., Sarmiento Montesdeoca, S., Rodríguez Ruiz, D., & García Manso, J. M. (2011). Competitiveness analysis in the NBA regular seasons. *Education. Physical Training. Sport.*, *1*(80). Recuperado a partir de http://www.lkka.lt/get.php?f.11008

De Saá Guerra, Y., Martín Gonzalez, J. M., Montesdeoca, S. S., Rodriguez Ruiz, D., Arjonilla López, N., & García Manso, J. M. (2013). Basketball scoring in NBA games: An example of complexity. *Journal of Systems Science and Complexity*, *26*(1), 94-103. doi:10.1007/s11424-013-2282-3

De Saá Guerra, Y., Martín González, J. M., Sarmiento Montesdeoca, S., Rodríguez Ruiz, D., García-Rodríguez, A., & Juan Manuel García-Manso. (2012). A model for competitiveness level analysis in sports competitions: Application to basketball. *Physica A: Statistical Mechanics and its Applications*, *391*(10), 2997-3004. doi:10.1016/j.physa.2012.01.014

De Saá Guerra, Y., Martín González, J., Sarmiento Montesdeoca, S., Rodríguez Ruiz, D., Arjonilla López, N., & García Manso, J. M. (2011). Basketball scoring in NBA games: an example of complexity. *arXiv:1108.0779*. Recuperado a partir de http://arxiv.org/abs/1108.0779

Dhar, D. (1990). Self-organized critical state of sandpile automaton models. *Physical Review Letters*, *64*(14), 1613-1616. doi:10.1103/PhysRevLett.64.1613

Dietl, H., Lang, M., & Werner, S. (2008). The Effect of Luxury Taxes on Competitive Balance, Club Profits, and Social Welfare in Sports Leagues. *Institute for Strategy and Business Economics University of Zurich. Working Paper*. Recuperado a partir de http://papers.ssrn.com/abstract=1237737

Dietl, H. M., Franck, E. P., Lang, M., & Rathke, A. (2010). Organizational Differences between U.S. Major Leagues and European Leagues: Implications for Salary Caps. *Norh American Association of Sport Economists. Working Paper*. Recuperado a partir de http://econpapers.repec.org/paper/isowpaper/0122.htm

Erdös, P., & Rényi, A. (1960). On the evolution of random graphs. *Math. Inst. Hung. Acad. Sci.*, *5*, 17-61.

FIBA. (2012a). FIBA - history - 80th Anniversary. *FIBA.COM*. Recuperado a partir de http://fapd.fiba.com/CMS/inc_boxes/templates/cms_base_set/v3/News/p/newsArti.ajax.html

FIBA. (2012b). FIBA Rules. *Official Basketball Rules 2012*. Recuperado a partir de http://www.fiba.com/pages/eng/fc/FIBA/ruleRegu/p/openNodeIDs/897/selNodeID/897/baskOffiRule.html

Fleming, L., & Marx, M. (2006). Managing Creativity in Small Worlds. *California Management Review*, *48*(4), 6-27. doi:10.2307/41166358

Forbes. (2008). Forbes.com #18 Michael Jordan. Recuperado a partir de www.forbes.com/lists/2008/53/celebrities08_Michael-Jordan_UGGU.html.

Fort, R., & Maxcy, J. (2003). «Competitive Balance in Sports Leagues: An Introduction». *Journal of Sports Economics*, *4*(2), 154-160. doi:10.1177/1527002503004002005

Gabel, A., & Redner, S. (2012). Random Walk Picture of Basketball Scoring. *Journal of Quantitative Analysis in Sports*, *8*(1). doi:10.1515/1559-0410.1416

Gambetta, V. (1989). New trends in training theory. *New Studies in Athletics*, *4*(3), 7-10.

García Manso, J. M. G., & Martín González, J. M. (2008). *La formación del deportista en un sistema de rendimiento deportivo: Autoorganización y emergencia, entre el orden y el caos*. Editorial Kinesis.

García Manso, J. M., Martín González, J. M., Da Silva-Grigoletto, M. E., Vaamonde, D., Benito, P., & Calderón, J. (2008). Male powerlifting performance described from the viewpoint of complex systems. *Journal of Theoretical Biology*, *251*(3), 498-508. doi:10.1016/j.jtbi.2007.12.010

George, M., Evangelos, T., Alexandros, K., & Athanasios, L. (2009). The inside game in World Basketball. Comparison between European and NBA teams. *International Journal of Performance Analysis in Sport*, *9*(2), 157-164.

Goodwin, B. (2002). *How the Leopard Changed Its Spots. [Las Manchas Del Leopardo]*. Tusquets Editor.

Goossens, K. (2006). Competitive balance in european football: comparison by adapting measures: national measure of seasonal imbalance and Top 3. *Rivista di Diritto ed Economia dello Sport*, *2*(2), 77-122.

Grasso, J. (2010). *Historical Dictionary of Basketball*. Scarecrow Press.

Greenhough, J., Birch, P. C., Chapman, S. C., & Rowlands, G. (2001). Football goal distributions and extremal statistics. *arXiv:cond-mat/0110605*. doi:10.1016/S0378-4371(02)01030-0

Griffiths, S. (2010, septiembre 20). The Canadian who invented basketball. *BBC*. Recuperado a partir de http://www.bbc.co.uk/news/world-us-canada-11348053

Guimera, R., & Amaral, L. A. N. (2005). Cartography of complex networks: modules and universal roles. *Journal of statistical mechanics (Online)*, *2005*(P02001), P02001-1-P02001-13. doi:10.1088/1742-5468/2005/02/P02001

Guimera, R., Sales-Pardo, M., & Amaral, L. A. N. (2007). Classes of complex networks defined by role-to-role connectivity profiles. *Nat Phys*, *3*(1), 63-69. doi:10.1038/nphys489

Guimera, R., Uzzi, B., Spiro, J., & Amaral, L. A. N. (2005). Team Assembly Mechanisms Determine Collaboration Network Structure and Team Performance. *Science*, *308*(5722), 697-702. doi:10.1126/science.1106340

Halicioglu, F. (2006). The impact of football point systems on the competitive balance: evidence from some european football leagues. *Rivista di Diritto ed Economia dello Sport*, *2*(2), 67-76.

Hausman, J. A., & Leonard, G. K. (1994). Superstars in the NBA : economic value and policy. Recuperado a partir de http://18.7.29.232/handle/1721.1/64241

Hernández Moreno, J. (1994). *Fundamentos del deporte: análisis de las estructuras del juego deportivo*. INDE.

Heuer, A., Mueller, C., & Rubner, O. (2010). Soccer: is scoring goals a predictable Poissonian process? *arXiv:1002.0797*. doi:10.1209/0295-5075/89/38007

Hill, J. R., & Groothuis, P. A. (2001). The New NBA Collective Bargaining Agreement, the Median Voter Model, and a Robin Hood Rent Redistribution. *Journal of Sports Economics*, *2*(2), 131-144. doi:10.1177/152700250100200203

Holland, J. H. (1995). *Hidden Order: How Adaptation Builds Complexity*. Perseus Books Group.

Hughes, M., & Franks, I. (2005). Analysis of passing sequences, shots and goals in soccer. *Journal of sports sciences*, *23*(5), 509-514. doi:10.1080/02640410410001716779

Humphreys, B. R. (2002). Alternative Measures of Competitive Balance in Sports Leagues. *Journal of Sports Economics*, *3*(2), 133-148. doi:10.1177/152700250200300203

IOC. (2012). olympic.org. Recuperado a partir de http://www.olympic.org/

Ivancevic, V. G., & Ivancevic, T. T. (2006). *High-Dimensional Chaotic and Attractor Systems: A Comprehensive Introduction*. Springer.

Jackson, P., & Delehanty, H. (2006). *Sacred Hoops: Spiritual Lessons of a Hardwood Warrior*. Hyperion Books.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., & Barabási, A. L. (2000). The large-scale organization of metabolic networks. *Nature*, *407*(6804), 651-654. doi:10.1038/35036627

Jost, J. (2005). *Dynamical Systems: Examples of Complex Behaviour*. Springer Science & Business.

Katz, S. J., & Katz, L. (1999). Power laws and athletic performance. *Journal of Sports Sciences*, *17*(6), 467-476. doi:10.1080/026404199365777

Kauffman, S. A. (1995). *At Home in the Universe: The Search for Laws of Self-Organization and Complexity*. Oxford University Press.

Késenne, S. (2000). The Impact of Salary Caps in Professional Team Sports. *Scottish Journal of Political Economy*, *47*(4), 422–430. doi:10.1111/1467-9485.00171

Knapp, B. (1981). *La habilidad en el deporte* (Miñón.). Valladolid.

Knowles, G., Sherony, K., & Haupert, M. (1992). The Demand for Major League Baseball: A Test of the Uncertainty of Outcome Hypothesis. *The American Economist*, *36*(2), 72-80.

Kubatko, J., Oliver, D., Pelton, K., & Rosenbaum, D. T. (2007). A Starting Point for Analyzing Basketball Statistics. *Journal of Quantitative Analysis in Sports*, *3*(3). Recuperado a partir de http://econpapers.repec.org/article/bpjjqsprt/v_3a3_3ay_3a2007_3ai_3a3_3an_3a1.htm

Laherrère, J., & Sornette, D. (1998). Stretched exponential distributions in nature and economy: «fat tails» with characteristic scales. *The European Physical Journal B - Condensed Matter and Complex Systems*, *2*(4), 525-539. doi:10.1007/s100510050276

Lebed, F. (2006). System approach to games and competitive playing. *European Journal of Sport Science*, *6*(1), 33-42. doi:10.1080/17461390500422820

Li, C., Ma, T., Zhu, X., & Li, W. (2011). The power–law relationship between landslide occurrence and rainfall level. *Geomorphology*, *130*(3–4), 221-229. doi:10.1016/j.geomorph.2011.03.018

Lorenzo Calvo, A., Gómez Ruano, M. A., & Sampaio, J. (2003). Análisis descriptivo de las posesiones de 24 segundos en baloncesto. *Lecturas: Educación física y deportes*, (67), 28-.

Mainzer, K. (2005). *Symmetry And Complexity: The Spirit And Beauty Of Nonlinear Science*. World Scientific.

Malacarne, L. ., & Mendes, R. . (2000). Regularities in football goal distributions. *Physica A: Statistical Mechanics and its Applications*, *286*(1–2), 391-395. doi:10.1016/S0378-4371(00)00363-0

Malamud, B. D., & Turcotte, D. L. (2006). The applicability of power-law frequency statistics to floods. *Journal of Hydrology*, *322*(1–4), 168-180. doi:10.1016/j.jhydrol.2005.02.032

Martín Acero, R., & Lago Peñas, C. (2005). Complejidad y rendimiento en los deportes sociomotores de equipo (DSEQ): dificultades de investigación. *Lecturas: Educación física y deportes*, (90), 3-.

Mathur, L. K., Mathur, I., & Rangan, N. (1997). The wealth effects associated with a celebrity endorser: The Michael Jordan Phenomenon. *Journal of Advertising Research*, *37*(3), 67-73.

May, R. M. (2001). *Stability and Complexity in Model Ecosystems*. Princeton University Press.

McGarry, T., Anderson, D. I., Wallace, S. A., Hughes, M. D., & Franks, I. M. (2002). Sport competition as a dynamical self-organizing system. *Journal of Sports Sciences*, *20*(10), 771-781. doi:10.1080/026404102320675620

McGarry, T., & Franks, I. M. (2007). System approach to games and competitive playing: Reply to Lebed (2006). *European Journal of Sport Science*, *7*(1), 47-53. doi:10.1080/17461390701216831

Mega, M. S., Allegrini, P., Grigolini, P., Latora, V., Palatella, L., Rapisarda, A., & Vinciguerra, S. (2003). Power-Law Time Distribution of Large Earthquakes. *Physical Review Letters*, *90*(18), 188501. doi:10.1103/PhysRevLett.90.188501

Mendes, R. S., Malacarne, L. C., & Anteneodo, C. (2007). Statistics of football dynamics. *IEEE Transactions on Automatic Control*, *363*(3), 357-363. doi:10.1140/epjb/e2007-00177-4

Mestre, J. A., Brotóns, J. M., & Álvaro, M. (2002). *La Gestión Deportiva: Clubes y Federaciones*. INDE.

Millán Garrido, A. (2010). *La reforma del régimen jurídico del deporte profesional*. Editorial Reus.

Mitzenmacher, M. (2004). A Brief History of Generative Models for Power Law and Lognormal Distributions. *Internet Mathematics*, *1*(2), 226-251. doi:10.1080/15427951.2004.10129088

Montoya Terán, J. M., Solé, R. V., & Rodríguez Fernández, M. Á. (2001). La arquitectura de la naturaleza: complejidad y fragilidad en redes ecológicas. Recuperado a partir de http://rua.ua.es/dspace/handle/10045/11245

Muruzábal del Solar, J. M. (2012). *El baloncesto navarro: historia de un proyecto compartido*. Federación Navarra de Baloncesto.

Naismith, J. (1941). *Basketball: Its Origin and Development*. U of Nebraska Press.

NBA. (2003). NBA.com: Evolution of the Draft and Lottery. Recuperado 7 de mayo de 2012, a partir de http://www.nba.com/features/inside_lottery_050524.html

NBA. (2007). NBA.com: Let the Ping-Pong Balls Fall. Recuperado 7 de mayo de 2012, a partir de http://www.nba.com/features/inside_lottery_050524.html

NBA. (2012). NBA.com: Rules. Recuperado a partir de http://www.nba.com/analysis/rules_index.html

NBA. (s. f.). NBA.com: The First Game. Recuperado 7 de mayo de 2012, a partir de http://www.nba.com/history/firstgame_feature.html

NCAA. (2012a). Rules- NCAA.org. Recuperado 8 de mayo de 2012, a partir de http://www.ncaa.org/wps/wcm/connect/public/NCAA/Playing+Rules+Administration/NCAA+Rules+Sports/Basketball/Mens/Playing+Rules/Mens+Basketball+Index

NCAA. (2012b). NCAA Rules. Recuperado a partir de http://www.ncaa.org/wps/wcm/connect/public/NCAA/Playing+Rules+Administration/NCAA+Rules+Sports/Basketball/Mens/Playing+Rules/Mens+Basketball+Index

NCAA. (2012c). History - NCAA.org. Recuperado 8 de mayo de 2012, a partir de http://www.ncaa.org/wps/wcm/connect/public/NCAA/About+the+NCAA/Who+We+Are/About+the+NCAA+history

Newman, M. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, *46*(5), 323-351. doi:10.1080/00107510500052444

Newman, M. E. J. (2001a). Clustering and preferential attachment in growing networks. *Physical Review E*, *64*(2), 025102. doi:10.1103/PhysRevE.64.025102

Newman, M. E. J. (2001b). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, *98*(2), 404-409. doi:10.1073/pnas.98.2.404

Newman, M. E. J. (2003). The Structure and Function of Complex Networks. *SIAM Review*, *45*(2), 167-256. doi:10.1137/S003614450342480

Nicolis, G., & Prigogine, I. (1977). *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations* (1.ª ed.). John Wiley & Sons.

Noll, R. (1988). Professional basketball. *Stanford University Studies in Industrial Economics, Working Paper*, *144*.

Onody, R. N., & de Castro, P. A. (2004). Complex network study of Brazilian soccer players. *Physical Review E*, *70*(3), 037103. doi:10.1103/PhysRevE.70.037103

Ortega y Gasset, J. (1933). El hombre y la gente. *Revista Occidente para Alianza Editorial. Madrid*. Recuperado a partir de http://html.rincondelvago.com/el-hombre-y-la-gente_ortega-y-gasset.html

Owen, P., Ryan, M., & Weatherston, C. (2007). Measuring Competitive Balance in Professional Team Sports Using the Herfindahl-Hirschman Index. *Review of Industrial Organization*, *31*(4), 289-302. doi:10.1007/s11151-008-9157-0

Palao Andrés, J. M., Ortega Martín, V., Piñar López, M. I., & Ortega Toro, E. (2004). Incidencia de la modificación del tiempo de posesión de 30 a 24 segundos sobre el tanteo en baloncesto y propuesta de modificación en categorías de formación. *Lecturas: Educación física y deportes*, (79), 7-.

Park, J., & Newman, M. E. J. (2005). A network-based ranking system for US college football. *Journal of Statistical Mechanics: Theory and Experiment*, *2005*(10), P10014. doi:10.1088/1742-5468/2005/10/P10014

Passos, P., Araújo, D., Davids, K., Gouveia, L., Milho, J., & Serpa, S. (2008). Information-governing dynamics of attacker–defender interactions in youth rugby union. *Journal of Sports Sciences*, *26*(13), 1421-1429. doi:10.1080/02640410802208986

Passos, P., Araújo, D., Davids, K., Milho, J., & Gouveia, L. (2009). Power Law Distributions in Pattern Dynamics of Attacker-Defender Dyads in the Team Sport of Rugby Union: Phenomena in a Region of Self-Organized Criticality? *Emergence: Complexity & Organization*, *11*(2), 37-45.

Patterson, S. (1993). National Basketball Association Salary Cap, Dead or Alive, The. *Thurgood Marshall Law Review*, *19*, 535.

Paulo Ferreira, A., Ibáñez Godoy, S. J., & Sampaio, J. (2009). Las reglas y la casualidad en Baloncesto: una aproximación histórica. *Retos: nuevas tendencias en educación física, deporte y recreación*, (15), 9-13.

Pisarenko, V. F., & Sornette, D. (2012). Robust statistical tests of Dragon-Kings beyond power law distributions. *The European Physical Journal Special Topics*, *205*(1), 95-115. doi:10.1140/epjst/e2012-01564-8

Prigogine, I., & Holte, J. (1993). *Chaos: The New Science*. University Press of America.

Prigogine, I., & Stengers, I. (1984). *Order out of chaos: man's new dialogue with nature*. Bantam Books.

Quirk, J. P., & Fort, R. D. (1997). *Pay Dirt: The Business of Professional Team Sports*. Princeton University Press.

Radicchi, F. (2011). Who Is the Best Player Ever? A Complex Network Analysis of the History of Professional Tennis. *PLoS ONE*, *6*(2), e17249. doi:10.1371/journal.pone.0017249

Rascher, D. A. (1999). A Test of the Optimal Positive Production Network Externality in Major League Baseball. *SSRN eLibrary*. Recuperado a partir de http://papers.ssrn.com/sol3/papers.cfm?abstract_id=81914

Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B - Condensed Matter and Complex Systems*, *4*(2), 131-134. doi:10.1007/s100510050359

Rhoads, T. (2005). *A Measure of Competitive Imbalance for the PGA Tour.* (Towson University.).

Rhoads, T. A. (2004). Competitive balance and conference realignment in the NCAA. *The 74th Annual Meeting of Southern Economic Association, New Orleans, LA.*

Ribeiro, H. V., Mendes, R. S., Malacarne, L. C., Jr, S. P., & Santoro, P. A. (2010). Dynamics of tournaments: the soccer case - A random walk approach modeling soccer leagues. *The European Physical Journal B*, *75*(3), 8. doi:10.1140/epjb/e2010-00115-5

Richardson, D. H. (2000). Pay, Performance, and Competitive Balance in the National Hockey League. *Eastern Economic Journal*, *26*(4), 393-417.

Riolo, R. L., Cohen, M. D., & Axelrod, R. (2001). Evolution of cooperation without reciprocity. *Nature*, *414*(6862), 441-443. doi:10.1038/35106555

Robson, A. J. (2005). Complex Evolutionary Systems and the Red Queen*. *The Economic Journal*, *115*(504), F211-F224. doi:10.1111/j.1468-0297.2005.01002.x

Rovell, D. (2003). Cashing in on the ultimate cash cow. *ESPN.com*. Recuperado a partir de espn.com http://sports.espn.go.com/espn/print?id=1538744&type=story.

Ruiz, L. M. (1999). Rendimiento deportivo, optimización y excelencia en el deporte. *Revista de psicología del deporte*, *8*(2), 235-248.

Sachs, M. K., Yoder, M. R., Turcotte, D. L., Rundle, J. B., & Malamud, B. D. (2012). Black swans, power laws, and dragon-kings: Earthquakes, volcanic eruptions, landslides, wildfires, floods, and SOC models. *The European Physical Journal Special Topics*, *205*(1), 167-182. doi:10.1140/epjst/e2012-01569-3

Sampaio, J., Lago, C., Casais, L., & Leite, N. (2010). Effects of starting score-line, game location, and quality of opposition in basketball quarter score. *European Journal of Sport Science*, *10*(6), 391-396. doi:10.1080/17461391003699104

Savaglio, S., & Carbone, V. (2000). Human performance: Scaling in athletic world records. *Nature*, *404*(6775), 244. doi:10.1038/35005165

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., … Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, *461*(7260), 53-59. doi:10.1038/nature08227

Schmidt, M. B., & Berri, D. J. (2001). Competitive Balance and Attendance The Case of Major League Baseball. *Journal of Sports Economics*, *2*(2), 145-167. doi:10.1177/152700250100200204

Schroeder, M. (1992). *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise*. Henry Holt and Company.

Scott, 1, 2Frank A., Long, 1, 2James E., & Somppi, K. (1985). Salary Vs. marginal revenue product under monopsony and competition: The case of professional basketball. *Atlantic Economic Journal*, *13*(3), 50-59. doi:10.1007/BF02304294

Scully, G. W. (1989). *The Business of Major League Baseball* (1St Edition.). University Of Chicago Press.

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, *43*(9), 628–638. doi:10.1002/(SICI)1097-4571(199210)43:9<628::AID-ASI5>3.0.CO;2-0

Silverman, D. (2012). *Basketball*. ABDO.

Soebbing, B. P. (2008). Competitive Balance and Attendance in Major League Baseball: An Empirical Test of the Uncertainty of Outcome Hypothesis. *International Journal of Sport Finance*, *3*(2), 119-126.

Solé Forto, J. (2002). *Fundamentos del entrenamiento deportivo: libro de ejercicios*. Ergo.

Solé, R. (2009). *Complex networks [Redes complejas]* (1st. ed.). Barcelona: Tusqets.

Solé, R., & Goodwin, B. (2002). *Signs of life: how complexity pervades biology*. Basic Books.

Song, W., Wang, J., Satoh, K., & Fan, W. (2006). Three types of power-law distribution of forest fires in Japan. *Ecological Modelling*, *196*(3–4), 527-532. doi:10.1016/j.ecolmodel.2006.02.033

Sornette, D. (2004). *Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization, and Disorder : Concepts and Tools*. Springer.

Sornette, D., & Ouillon, G. (2012). Dragon-kings: Mechanisms, statistical methods and empirical evidence. *The European Physical Journal Special Topics*, *205*(1), 1-26. doi:10.1140/epjst/e2012-01559-5

Stanley, H. E. (1987). *Introduction to Phase Transitions and Critical Phenomena*. Oxford University Press.

Strogatz, S. H. (2001). Exploring complex networks. *Nature*, *410*(6825), 268-276. doi:10.1038/35065725

Stumpf, M. P. H., & Porter, M. A. (2012). Critical Truths About Power Laws. *Science*, *335*(6069), 665-666. doi:10.1126/science.1216142

Szymanski, S. (2003). The Economic Design of Sporting Contests. *Journal of Economic Literature*, *41*(4), 1137-1187. doi:10.1257/002205103771800004

Tous Fajardo, J. (1999). *Reglamento de baloncesto comentado*. Editorial Paidotribo.

Tsallis, C., & Albuquerque, M. P. de. (2000). Are citations of scientific papers a case of nonextensivity? *The European Physical Journal B - Condensed Matter and Complex Systems*, *13*(4), 777-780. doi:10.1007/s100510050097

Uzzi, B., Amaral, L. A., & Reed-tsochas, F. (2007). *Small-world networks and management science research: a review*.

Uzzi, B., & Spiro, J. (2005). Collaboration and Creativity: The Small World Problem. *American Journal of Sociology*, *111*(2), 447-504. doi:10.1086/432782

Van Valen, L. (1973). A New Evolutionary Law. *Evolutionary Theory*, *1*, 1-30.

Vaz de Melo, P. O. S., Almeida, V. A. F., & Loureiro, A. A. F. (2008). Can complex network metrics predict the behavior of NBA teams? En *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 695–703). New York, NY, USA: ACM. doi:10.1145/1401890.1401974

Vicsek, T. (2002). Complexity: The bigger picture. *Nature*, *418*(6894), 131. doi:10.1038/418131a

Virkar, Y., & Clauset, A. (2012). Power-law distributions in binned empirical data. *arXiv:1208.3524*. Recuperado a partir de http://arxiv.org/abs/1208.3524

Watanabe, N. M. (2012). Japanese professional soccer attendance and the effects of regions, competitive balance, and rival franchises. *International Journal of Sport Finance*, *7*(4), 309+.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of /`small-world/' networks. *Nature*, *393*(6684), 440-442. doi:10.1038/30918

West, G. B., Brown, J. H., & Enquist, B. J. (1997a). A General Model for the Origin of Allometric Scaling Laws in Biology. *Science*, *276*(5309), 122-126. doi:10.1126/science.276.5309.122

West, G. B., Brown, J. H., & Enquist, B. J. (1997b). A General Model for the Origin of Allometric Scaling Laws in Biology. *Science*, *276*(5309), 122-126. doi:10.1126/science.276.5309.122

White, E. P., Enquist, B. J., & Green, J. L. (2008). ON ESTIMATING THE EXPONENT OF POWER-LAW FREQUENCY DISTRIBUTIONS. *Ecology*, *89*(4), 905-912. doi:10.1890/07-1288.1

Wilner, B., & Rappoport, K. (2012). *The Big Dance: The Story of the NCAA Basketball Tournament*. Taylor Trade Publications.

Wolff, A. (1991). *100 years of hoops*. Oxmoor House.

Yarrow, K., Brown, P., & Krakauer, J. W. (2009). Inside the brain of an elite athlete: the neural processes that support high achievement in sports. *Nature Reviews Neuroscience*, *10*(8), 585-596. doi:10.1038/nrn2672

Yilmaz, M. R., & Chatterjee, S. (2000). Patterns of NBA team performance from 1950 to 1998. *Journal of Applied Statistics*, *27*(5), 555-566. doi:10.1080/02664760050076399

Yukalov, V. I., & Sornette, D. (2012). Statistical outliers and dragon-kings as Bose-condensed droplets. *The European Physical Journal Special Topics*, *205*(1), 53-64. doi:10.1140/epjst/e2012-01561-y

Ziv, G., Lidor, R., & Arnon, M. (2010). Predicting team rankings in basketball: The questionable use of on-court performance statistics. *International Journal of Performance Analysis in Sport*, *10*(2), 103-114.

Annexes

Basketball from the perspective of non-linear complex systems

Yves de Saá Guerra
2013

Annexes Index

1. **Paper 1.** De Saá Guerra, Y., Martín González, J. M., Arjonilla López, N., Sarmiento Montesdeoca, S., Rodríguez Ruiz, D., & García Manso, J. M. (2011). Competitiveness analysis in the NBA regular seasons. Education. Physical Training. Sport., 1(80).

2. **Paper 2.** De Saá Guerra, Y., Martín González, J. M., Sarmiento Montesdeoca, S., Rodríguez Ruiz, D., García-Rodríguez, A., & Juan Manuel García-Manso. (2012). A model for competitiveness level analysis in sports competitions: Application to basketball. Physica A: Statistical Mechanics and its Applications, 391(10), 2997-3004.

3. **Paper 3.** De Saá Guerra, Y., Martín Gonzalez, J. M., Montesdeoca, S. S., Rodriguez Ruiz, D., Arjonilla López, N., & García Manso, J. M. (2013). Basketball scoring in NBA games: An example of complexity. Journal of Systems Science and Complexity, 26(1), 94-103.

4. **Paper 4.** De Saá Guerra, Y., Martín González, J. M., García Manso, J. M., Navarro Valdivielso M., & García Rodríguez, A. (2013). Competitive balance and clustering in professional basketball NBA and ACB. Journal of Sport Management Review. (submitted).