



ELSEVIER

Contents lists available at ScienceDirect

## Expert Systems With Applications

journal homepage: [www.elsevier.com/locate/eswa](http://www.elsevier.com/locate/eswa)

## iSports: A web-oriented expert system for talent identification in soccer

Francisco Louzada<sup>a,\*</sup>, Alexandre C. Maiorano<sup>a</sup>, Anderson Ara<sup>b,1</sup><sup>a</sup> Universidade de São Paulo, Institute of Mathematical Science and Computing, São Carlos, Brazil<sup>b</sup> Universidade Federal de São Carlos, Departamento de Estatística, São Carlos, SP, Brazil

## ARTICLE INFO

## Keywords:

Sport evaluation  
Talent identification  
Z-CELAFISCS methodology  
Principal component analysis  
Factor analysis  
Copula theory

## ABSTRACT

Nowadays soccer is the most practiced sport in the world and moves a multimillionaire market. Therefore, a club that is able to recruit and develop talented players to their fullest potential has a lot of advantages and economic benefits. However, in most clubs the players are selected through scouts and coaches recommendation, with predictive success based mostly on intuition than other objective criteria. In addition, it is known that talent development and identification is a multifactorial process involving many characteristics. To this end, this paper proposes the creation of performance indicators based on multivariate statistical analysis. Usual principal components and factor analysis are performed to construct physical, technical and general score and copula modeling is proposed to create the consistency index, which generalizes the Z score method. With these indicators, a web-oriented expert system for analyzing sport data in real time via R software is proposed as a powerful tool for talent identification in soccer. This system, the so called iSports, allows the monitoring and continuous comparison of athletes in a simple and efficient way, taking into account essential aspects, as well as identifying candidate talented that have above the average performance, that is, who stand out from the studied population of soccer players. In order to promote and popularize the access of information and the statistical science applied in the sports context, the iSports system can be used in any training center of the country, impacting the increase of knowledge of the athletes in training phase at any school, city or region.

© 2015 Elsevier Ltd. All rights reserved.

## 1. Introduction

Nowadays soccer is the most practiced sport in the world, with about 265 million players (men and women). If we include referees and officials, we have about 270 million people actively involved, which represents 4% of the world's population (Kunz, 2007). Besides evolving this large number of people, soccer moves a multimillionaire market (Elferink-Gemser, Huijgen, Coelho-E-Silva, Lemmink, & Visscher, 2012) and the economic benefits of a team being able to recruit talented players and develop them to their fullest potential are obvious (Reilly, Bangsbo, & Franks, 2000a). For example, if we observe the three most valuable squad currently: Real Madrid, FC Barcelona and Bayern Munich, we have a value of 688.80, 620.00 and 568.15 million of Euros, respectively (Markt, 2014).

In this sense, the process of detection, selection, identification and development of talents becomes essential. For Williams and Reilly (2000) these are the main stages involved in the talents recruitment, in which detection can be defined as the discover of people that are

not involved with the sport of matter; selection is defined by the choice of one player, inside a group of people, more suitable for the practice; identification is no more than the discover of potential players, who already practice the sport and development is the learning process that allow the player to accomplish it's potential.

According to Williams (2000), in most clubs the players are selected through scout's recommendation that observe players that standout on performance during training and matches. This process is speculated, with predictive success based mostly on intuition other than objective criteria, making necessary the use of a more objective form of selection. Therefore, in the attempt to produce young talented player, sports scientists have the job to play a conjunct work with coaches, player scouts and "administrators" to find key elements of identification and development in the process (Williams & Reilly, 2000).

In a general scope of an athlete's composing, we have physical condition as indispensable so that their income in a given sport is excellent (Balaguer et al., 2012; F. Helsen, Hodges, Winkel, & Starkes, 2000; le Gall, Carling, Williams, & Reilly, 2010; Reilly et al., 2000a; Vaeyens et al., 2006; Williams & Reilly, 2000). An athlete with a physical condition determined tends to obtain more satisfying results when compared with those not in the same condition, because a high physical condition is related with high performance indicators, a

\* Corresponding author. Tel.: +55 1697863979.

E-mail addresses: [louzada@icmc.usp.br](mailto:louzada@icmc.usp.br) (F. Louzada), [ale\\_maiorano@hotmail.com](mailto:ale_maiorano@hotmail.com) (A.C. Maiorano), [anderson@ufscar.br](mailto:anderson@ufscar.br) (A. Ara).<sup>1</sup> Accept for publication in 2015: Expert System With Applications.

minor game fatigue and even the team's performance (Walker & Turner, 2009). Besides, the best players must adapt to the game's physical demand, which is multifactorial. They don't need to present an extraordinary characteristic in only one area, but a high level in all of them (Reilly et al., 2000a).

However, this is not the only factor that will influence the performance. The capacity that a player/team possess is related with many others variables, as the psychological capacity to support high pressure, mostly when it involves negative consequences (Balaguer et al., 2012) and personality characteristics that facilitate learning, training and competition (Williams & Reilly, 2000). We can also observe the existence of cognitive factors and game intelligence, as anticipation, decision-making and even intelligence and creativity (Williams, 2000; Williams & Ericsson, 2005; Williams & Reilly, 2000). Moreover, we must consider how mature the athlete is (F. Helsen et al., 2000; Hirose, 2009), genetic factors (Howe, Davidson, & Sloboda, 1998; Williams & Reilly, 2000) and sociological (Williams & Reilly, 2000), for example the family influence, the coach's part and the prevention and detection of injuries. Therefore, we can notice that the talent selection process must regard many characteristics (Buekers, Borry, & Rowe, 2015; Gonaus & Müller, 2012; Reilly et al., 2000a; Reilly, Williams, Nevill, & Franks, 2000b; Vaeyens, Lenoir, Williams, & Philippaerts, 2008; Vaeyens et al., 2006; Williams & Reilly, 2000).

With the attempt to observe possible talents in soccer, specific field tests can be used and empirically analyzed, to detect individuals who stand out among the rest. These tests are considered a satisfactory alternative to laboratory tests (Castagna, Manzi, Impellizzeri, Weston, & Alvarez, 2010), as they are usually simple, popular between players and coaches and do not require equipment for their application (Walker & Turner, 2009).

Another contributing factor in the use of these tests is that the laboratory tests involves exercises that are not specific to soccer and the protocols used impact on the player's motivation (Castagna et al., 2010). Furthermore, these tests are usually expensive (Walker & Turner, 2009).

Accordingly to Reilly et al. (2000b), to identify talents for field games at a young age is far from being a mechanic process, being more complex in group sports than in individuals. Therefore, it is not surprising that there are few, if any, models to identify talents that are globally accepted. According to the authors, many of the proposed models are, in the best case, descriptive and schematic and there are few attempts to identify its validity.

Evaluating studies that report the talents selection process, we verify that in many cases they present comparisons between skilled or not skilled teenagers in certain tests/measures, but don't present a way to show, list or rank the best in certain characteristics evaluated taking into account mainly a multivariate proposal.

We observe the use of multivariate methodologies in some studies applied to soccer, like Factor Analysis (Gonaus & Müller, 2012), Discriminant Analysis (Gonaus & Müller, 2012; Huijgen, Elferink-Gemser, Lemmink, & Visscher, 2014; Morris, 2000; Reilly et al., 2000b; Vaeyens et al., 2006), Regression (Reilly et al., 2000a) and MANOVA/MANCOVA (Balaguer et al., 2012; Carling, Le Gall, & Malina, 2012; le Gall et al., 2010; Gonaus & Müller, 2012; Huijgen et al., 2014; Reilly et al., 2000b; Vaeyens et al., 2006; Veale, Pearce, Koehn, & Carlson, 2008). Factor Analysis was used to analyze the relationship between a set of tests; the MANOVA/MANCOVA to compare a set of measures and Discriminant Analysis or Regression to verify prediction from some measures collected using prior knowledge of the player being considered or not considered talented. Although, it is notable that in none of these cases it is possible to evaluate and compare the subject with the others according to its performance in certain specific tests.

For this purpose some field tests can be applied and by collecting data it is possible to create and use indicators that reflect the physical

and technical aptitude reality. As one of the main characteristics of a good indicator, it must be easily calculated and explainable.

In this paper, we use the methodology proposed by Matsudo, Rivet, and Pereira (1987), the construction of indicators using multivariate analysis, such as Principal Component Analysis and Factor Analysis, propose a new indicator that generalizes the Matsudo et al. (1987) approach by a multivariate Copula and develop an online and novel system based on free software.

At our best knowledge, there are not studies in the literature that work systematically procedures of specific tests, beginning with data collection, theoretical development, application and developing a software that allows an integrated structure focused on the selection and development of talents specifically in soccer. It leads us to an innovating structure in applied statistics on sports that possibly will be common in the near future, including in other areas of knowledge.

A related study was presented by Papić, Rogulj, and Pleština (2009), where is developed an online system using Fuzzy methods to identify the individual's aptitude in a specific list of sports. Our study however includes uni and multivariate statistical methods, different languages and softwares and the exclusive use of the R Software (R Core Team, 2014) to calculate the results in a dynamic connection with the server system. Moreover, while their system directs athletes to a particular sport, indicating which sport is most appropriate for a particular athlete, our system aims at the discovery of talents in soccer.

In a recent paper, Ballı and Korukoğlu (2014) proposed a framework that almost match our research, but they consider basketball players, evaluating physical and technical aspects and determining the overall ranking of the candidate players by Fuzzy method like Papić et al. (2009). However, they do not use a web based approach.

To exemplify the methodologies used, we apply tests to evaluate the performance of athletes at a soccer school in São Carlos, Brazil, and present the different reports generated by the proposed system, the iSports, available on <http://www.mwstat.com/isports>.

In general, the web based system proposed in our study has the following characteristics: focused on soccer, built on cloud and evaluates the physical, technical and general ability of players ordering individuals by their performance. Also considers indicators created from multivariate statistical techniques and presents individual and general reports. In addition, has a structure that instantly connects with the R software to generate all necessary calculations and it is built from free software.

The rest of this paper is organized as follows. In Section 2, the applied field tests and statistical procedures used are presented. In Section 3, the system architecture is described and explained. Section 4 shows the implementation and evaluation of the iSports system, including the development of indicators by proposed statistical methods, and, finally, Section 5 concludes the paper.

## 2. System elements and methodology

In this section, we present the performance practice tests and statistical tools used to analyze the information about soccer players. A set of univariate and multivariate statistical techniques were applied in order to create indicators to evaluate and compare the skills of individuals.

### 2.1. Field tests

As noted earlier, specific field tests has been widely used for assess performance in soccer (Ali, 2011; Bradley, Carling, Gomez Diaz, Hood, Barnes, Ade, Boddy, Krustup, & Mohr, 2013; Castagna et al., 2010; Elferink-Gemser et al., 2012; Jankovic, Matkovic, & Matkovic, 1997; Janssens, VAN RENTHERGEM, Bourgois, & Vrijens, 2002; Mujika, Santisteban, Impellizzeri, & Castagna, 2009; Rebelo, Brito, Seabra, Oliveira, & Krustup, 2014; Reilly et al., 2000a; Reilly et al., 2000b; Rösch et al., 2000; Vaeyens et al., 2006; Walker & Turner, 2009).

متن کامل مقاله

دریافت فوری ←

**ISI**Articles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات