Send Orders of Reprints at reprints@benthamscience.org

The Open Sports Science Journal, 2012, *5*, 181-187

Open Access

Tracking Performance in Football – An Example Using Goal Scoring Data

Pedro Moreira Silva[*,1], Júlio Garganta[1], José António Maia[1] and Pedro Manuel Santos[2]

[1]*Centre of Research, Education, Innovation and Intervention in Sport, CIFI²D, FADEUP, Portugal*

[2]*Faculdad de Ciencias de la Educación y del Deporte, Vigo University, Spain*

Abstract: Performance analysis contribution to game understanding is a widely accepted premise among the broad community of team ball sports, including football. In the literature it is possible to find a considerable amount of cross-sectional studies addressing performance evaluations over time. However, longitudinal studies aiming to evaluate the performance of the same participants (individual or teams) are scarce. Recognizing the interest to address this problem, the present paper focuses on the application of "tracking" with the purpose of characterizing performance behavior in football over eight different competitive periods (seasons). To this effect, the tracked performance indicator selected was the average number of goals scored per season on ten representative leagues of European football. Through the analysis of this variable, a detailed look at goal scoring tendencies is provided resorting to the explanation of different statistical procedures. Among these, Cohen´s kappa (*K*) and Foulkes & Davies gama (γ) were emphasized and interpreted as measures of individual and global tracking, respectively. With this example, the usefulness of tracking performance is highlighted, opening new perspectives to explore the field of team and individual performance analysis.

Keywords: Football, goals, performance analysis, tracking.

INTRODUCTION

Despite being considered that mischance may exert some influence on the result of a football match, it has been shown that, in general, the quality of individual and team performance seems to be the main factor in deciding the winning and losing teams [1]. Therefore, the last decade was marked by the emergence of performance analysis as a research subject in sports sciences [2]. Accordingly, in recent years, technological advances have supported the substantial growth of performance analysis systems [see 3 for a review]. The introduction of multi-camera systems along with computer specific software has contributed to the improvement of detailed individual and collective information analysis [4]. These improvements helped to overcome some methodological barriers [5] and to develop the game's body of knowledge [6].

Regardless of being conceptualized under different paradigms, performance analysis' contribution to game understanding has been undeniable. Having evolved from traditional paper and pencil systems, nowadays, the sophisticated field of notational analysis is starting to recognize important issues like the complexity of dynamic interactions [see, 7, 8, 9 for more information] that occur in the game, as an effort to deal with its underlying variability [e.g., 5, 10, 11]. Past information is of paramount importance to design predictive models of performance [12] and provide the coach and athlete(s) the opportunity to reflect on their specific activities. Thus, if "the essence of the coaching process is to instigate observable changes in behavior" [13, p. 12] then monitoring performance indicators over time shall be considered a necessary methodological procedure. In this sense, developmental research addresses the important and challenging issue of monitoring change in performance over a long period of time [14]. However, research to evaluate performance longitudinally does not seem a priority when it comes to evaluate the same participants. The majority of research available has used cross-sectional studies with different participants, teams or groups to monitor changing trends in variables between cohorts. This fact was perceived, for example, by Maia and collaborators [15] when noticing the lack of specific methodological and substantive agenda and may be considered an unjustifiable gap given the existing techniques for the analysis of longitudinal information.

It is argued that performance indicators may have a distinct predictive power on different teams or even on the same team in different competitive moments [16]. Accordingly, the influence of the competitive season in tactical performance indicators demands care in the extrapolation of conclusions from one competitive period to another [17]. The magnitude of change (if any) is an important feature to be considered by researchers and coaches when evaluating performance over a given time period. Performance oscillations and/or stability may then, be identified and analyzed according to possible reasons.

Designing longitudinal studies to investigate stability and change in individual or collective performance may constitute an opportunity to provide an objective look at specific performance indicators over time. In this sense, tracking has been widely used in different domains (e.g., motor control,

*Address correspondence to this author at the Faculdade de Desporto da Universidade do Porto (FADEUP), CIFI²D; Rua Dr. Plácido Costa, 91, 4200-450 Porto, Portugal; Tel: +351 939 629 306; Fax: +351 225 500 689; E-mail: silva.p.m@gmail.com

1875-399X/12

epidemiology, etc.) to analyze the tendency of an individual or group of individuals to remain in a specific growth channel [15]. The term tracking refers to the maintenance of the individual's relative position or, in other words, to the concept of stability when data is analyzed longitudinally [18].

This technique can be extended to performance analysis with the purpose of characterizing performance behavior over different competitive periods (during a season or several seasons) or the response of different groups submitted to distinct training protocols. Considering the first possibility, this study aims to describe different ways to analyze performance stability using accessible statistical and graphical tools based on the concept of tracking. Although a wealth of performance indicators is available to identify many facets of individual and collective performance, we will use a highly intuitive one – the average number of goals scored. The fact that the ultimate game objective is to score goals [19], which is paramount for any team's success [20], also influenced our choice.

Since the present paper constitutes a simple example of the application of this statistical procedure, the authors have decided not to focus on any particular team or club. It is hoped that this technique is recognized as a worthy promising endeavor leading to the appearance of specific studies aiming to evaluate longitudinally relevant performance variables of devoted international teams or athletes, or simply to be one more tool for the coach. New research possibilities are delivered to the readers.

METHODS

The average number of goals scored per match of ten professional European soccer leagues have been analyzed over eight competitive seasons (between 2001/02 and 2008/09), including the English, French, German, Dutch, Italian, Spanish, Portuguese, Czech, Russian and Greek leagues.

These leagues have been selected for having maintained uninterrupted and/or simultaneously, in the analyzed period, at least one team competing on European UEFA championships (UEFA Champions League and/or UEFA Cup). According to this criterion, these were considered representative leagues of European football. Data was collected from the official web sites of the corresponding leagues. Scored goals of all teams were summed and divided by the number of played matches. This procedure was used in every league and season.

Main descriptive statistics, repeated measures ANOVA and paired samples t-test were firstly used. Tracking is very often analyzed with Pearson or Spearman auto-correlations depending on the sample size and violations to bivariate normality. Given the small sample size, we used a resampling technique, the percentile bootstrap method, with 100 samples of size 10 to have more precise estimates of Spearman auto-correlations as well as their standard errors. All these computations were done in SPSS 18 and SYSTAT 13.

The average number of goals scored were then tracked individually, for each league. To this effect, performance channels were initially plotted. These channels, along with each league trajectory, provided a more detailed analysis of tracking. Finally, to assess the degree of individual and global tracking, Cohen´s kappa (*K*) and Foulkes & Davies gama (γ) were calculated. These latter descriptions and analysis were conducted in the Longitudinal Data Analysis (LDA) software [21].

RESULTS

Descriptive Statistics

Table **1** summarizes mean values of scored goals per match for each specified league from season 2001/02 to season 2008/09 (total goals and matches played per league and season are summarized in the **APPENDICE**). In all seasons and all leagues, a minimum mean score of two goals

Table 1. Mean Values (±SD) of Goals Scored Per Match – Seasons 2001/02 to 2008/09

Leagues	2001/02	2002/03	2003/04	2004/05	2005/06	2006/07	2007/08	2008/09	Mean	SD
English	2.63	2.63	2.66	2.57	2.48	2.45	2.64	2.48	2.57	0.09
French	2.34	2.2	2.33	2.17	2.13	2.25	2.28	2.26	2.25	0.07
German	2.92	2.68	2.97	2.91	2.81	2.74	2.81	2.92	2.85	0.10
Dutch	2.82	2.95	2.98	3.1	2.98	2.94	3.11	2.86	2.97	0.10
Italian	2.63	2.58	2.67	2.53	2.61	2.55	2.55	2.6	2.59	0.05
Spanish	2.53	2.67	2.67	2.58	2.46	2.48	2.69	2.9	2.62	0.14
Portuguese	2.67	2.63	2.38	2.32	2.23	2.31	2.3	2.3	2.39	0.16
Czech	2.46	2.47	2.41	2.21	2.36	2.23	2.32	2.42	2.36	0.10
Russian	2.42	2.54	2.49	2.25	2.45	2.34	2.41	2.5	2.43	0.09
Greek	2.88	2.57	2.61	2.32	2.33	2.33	2.38	2.17	2.45	0.22
Mean	2.63	2.59	2.62	2.50	2.48	2.46	2.55	2.54		
SD	0.2	0.19	0.23	0.31	0.26	0.23	0.27	0.27		

per match was obtained. Only the Dutch league has registered in seasons 2004/05 and 2007/08, more than two goals per match (3.1 and 3.12, respectively).

Mean Differences Across Seasons

ANOVA repeated measures with a Greenhouse-Geisser correction revealed significant differences among sport seasons [F(7.63)=3.15; p=0.048]. A posteriori comparisons, using Bonferroni correction showed that the average number of goals scored presented significant differences between seasons 2001/02 and 2006/07 (p=0.018), 2002/03 and 2006/07 (p=0.014), 2003/04 and 2004/05 (p=0.009), 2003/04 and 2005/06 (p=0.001), 2003/04 and 2006/07 (p<0.001), 2006/07 and 2007/08 (p=0.006).

Auto-Correlation Coefficients

Bootstraping Spearman's correlation coefficients (±standard errors) are presented in Table **2** and illustrated in a correlation matrix plot on Fig. (**1**).

Given the auto-correlations, the period of greater stability among seasons 2003/04 and 2007/08 was identified. This tendency is confirmed, visually, on the correlation matrix plot on Fig. (**1**) and in the cluster three presented in Fig. (**2**).

Performance Channels and Spaghetti Plot of Performance Trajectories

Fig. (**3**) describes three channels representing goal scoring performances across all seasons. Channels are considered in the context of this study as different performance zones. It is obvious that in each season two cut-off points (percentiles 33 and 66) divided the series into three classes or channels. Table **3** defines the corresponding cut points for each channel.

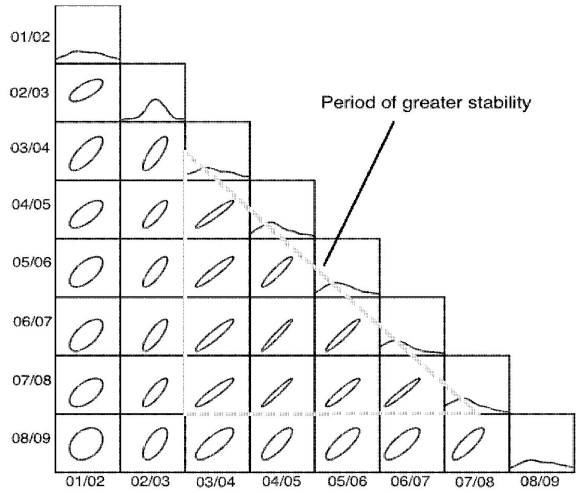

Fig. (1). Correlation matrix plot (the main diagonal contains a Kernel density curve).

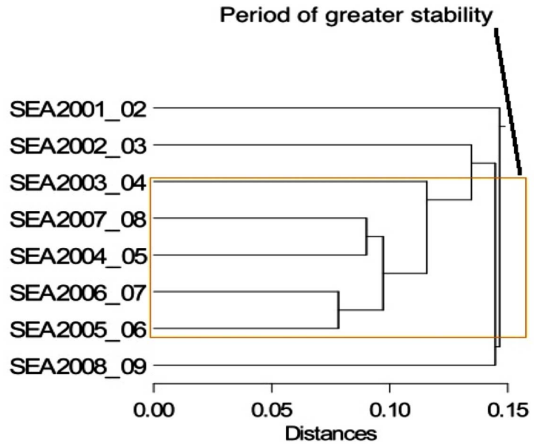

Fig. (2). Cluster tree.

Table 2. Bootstrap Spearman Auto-Correlation Coefficients with Standard-Errors

Seasons	2001/02	2002/03	2003/04	2004/05	2005/06	2006/07	2007/08	2008/09
2001/02	1.00							
2002/03	0.65 ±0.22	1.00						
2003/04	0.55 ±0.24	0.79 ±0.21	1.00					
2004/05	0.64 ±0.23	0.96 ±0.06	0.90 ±0.11	1.00				
2005/06	0.41 ±0.35	0.70 ±0.25	0.92 ±0.09	0.80 ±0.17	1.0			
2006/07	0.50 ±0.25	0.80 ±0.17	0.93 ±0.08	0.89 ±0.09	0.87 ±0.15	1.00		
2007/08	0.46 ±0.29	0.80 ±0.21	0.96 ±0.04	0.90 ±0.12	0.92 ±0.09	0.91 ±0.08	1.00	
2008/09	0.18 ±0.41	0.65 ±0.18	0.74 ±0.18	0.69 ±0.19	0.80 ±0.12	0.73 ±0.17	0.79 ±0.14	1.00

Table 3. Average Number of Goals Scored Percentiles and Cut Points for Each Season

Centiles	2001/02	2002/03	2003/04	2004/05	2005/06	2006/07	2007/08	2008/09
66,67	2.67	2.63	2.67	2.57	2.48	2.48	2.64	2.6
33,33	2.53	2.57	2.49	2.32	2.36	2.33	2.38	2.42

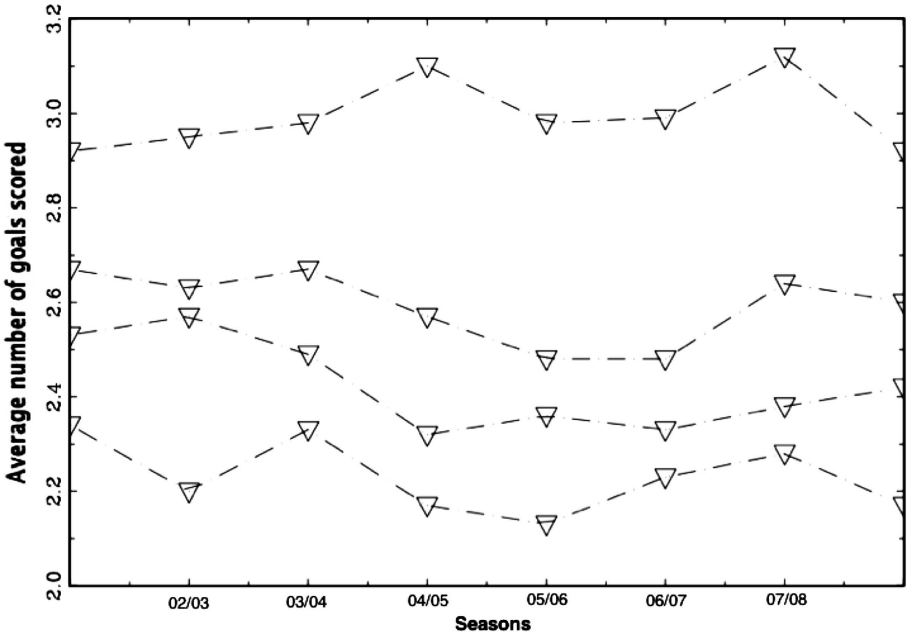

Fig. (3). Channels for the average number of goals scored on all leagues within seasons (computed with the LDA software).

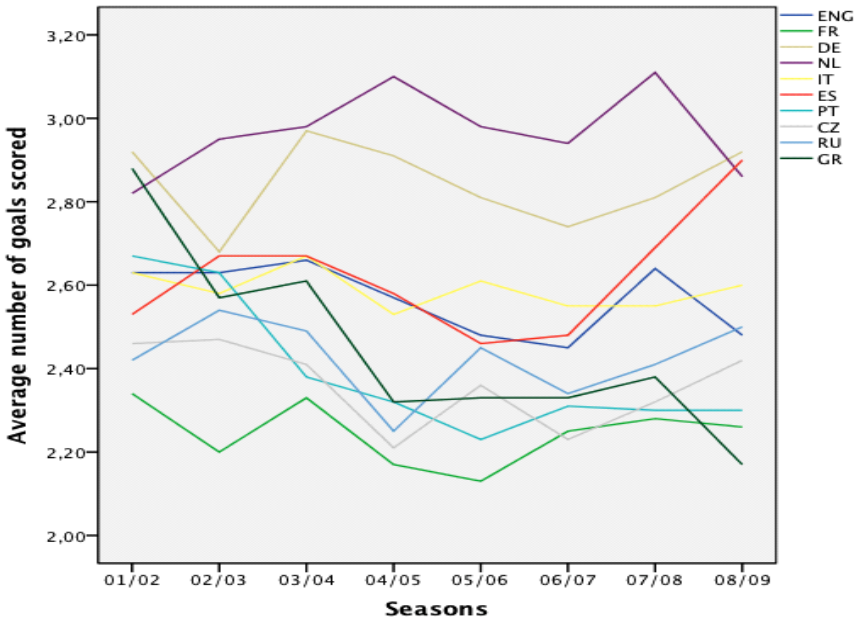

Fig. (4). Trajectories (spaghetti plot) of all leagues' average number of goals scored across seasons.

The spaghetti plot of the leagues' performance trajectories (Fig. **4**) shows some variability on the average number of goals scored per match in some of the leagues.

To analyze the tendency of each league to remain inside of one performance channel, their trajectories and performance channels were plotted together (Fig. **5**).

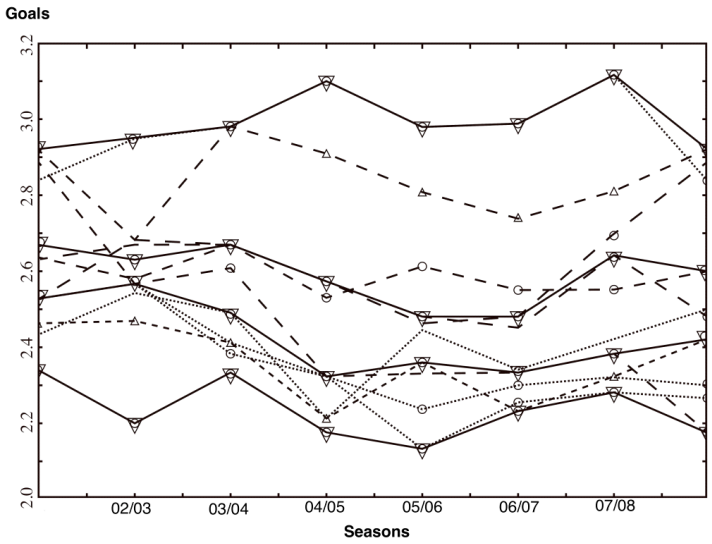

Fig. (5). Channels and "spaghetti plot"across seasons(computed with the LDA software).

Table 4. Kappa and Frequency Distribution Over each Channel

Leagues	K	Channel 1	Channel 2	Channel 3
English	1.00	0	8	0
French	1.00	8	0	0
German	1.00	0	0	8
Dutch	1.00	0	0	8
Italian	0.57	0	6	2
Spanish	0.32	1	3	4
Portuguese	0.57	6	2	0
Czech	1.00	8	0	0
Russian	0.42	4	4	0
Greek	0.53	6	1	1

Fig. (**5**) shows that most of intra-individual trajectories remain inside the same channel or oscillate between two channels.

Cohen's Kappa (*K*) and Foulkes& Davies Gama (γ)

Kappa values are interpreted as a measure of intra-individual change and inter-individual differences in leagues´ season's stability. According to the values presented in Table **4**, the Spanish league had the lowest performance stability over the studied time period (*K*=0.32) with scores oscillating, randomly, between the three performance channels.

The Russian, Portuguese and Italian leagues had moderate levels of stability (*K* between 0.4 and 0.6) with scores alternating, randomly, between two performance channels. Despite having oscillated between the three channels, the Greek league also presented moderate stability (*K* = 0.53) since its trajectory remained inside the lowest channel on six occasions.

The English, French, German, Dutch and Czech leagues had their scores inside the same performance channel throughout the eight seasons, indicating high stability (*K*=1).

The overall kappa for all leagues is 0.61±0.043 (95%CI= 0.52;0.69).

Foulkes and Davies γ examines the number of leagues that maintain the same relative rank. It estimated a measure of γ of 0.87±0.036 (95%CI=0.79;0.94).

DISCUSSION

The aim of the present study was to present an alternate statistical technique to analyze the stability of team performance using simple procedures and graphical tools based on the concept of tracking. It was chosen, as an example, a very intuitive variable– the average number of goals scored. To this purpose, a procedure for the statistical analysis of goal scoring tendencies over eight football seasons in ten European leagues was presented.

According to data presented in Table **1**, in the analyzed leagues a minimum mean of two goals per match were scored and there was not a clear trend in this variable variation. Values decrease and increase apparently with no discernible frame. An ANOVA repeated measures with a Greenhouse-Geisser correction revealed significant differentces among sport seasons, meaning that differences between seasons were important despite being represented as decimals.

These previous analysis only showed mean changes and trends in goals within and across seasons, as well as statisti-

cal significant differences in means across time. Therefore, we had to move to other views of this data set to address other questions and further interpret the richness of the data.

One very simple statistic to describe tracking is autocorrelations. If seasons' performances are highly correlated, it may be suggested that the mean scored goals tend to be stable throughout seasons. Thus, these correlation measures can be interpreted as tracking values and suggest the existence of performance stability or instability (a usual subjective cut-off value to indicate tracking is 0.50). A period of greater stability could, then, be identified between 2003/04 and 2007/08.

One important illustration that clearly isolates performance stability or instability is the idea of building performance channels. Based on the concept suggested by Waddington [22] in Theoretical Biology, and widely used in displaying and analyzing growth charts (for example, body height or weight), tracking can be assumed whenever repeated measures of a given subject (in our case, team performance) remain between a pair of adjacent percentiles (e.g., P_{50}-P_{75}) over time or do not deviate more than one adjacent channel [18]. It is important to note that these channels are empirically suggested by the software used and are expected to be based on substantive grounds.

The channels' delimitation did not indicate a clear increasing or decreasing tendency on scored goals over the seasons for the overall leagues. However, another look should be directed to the trajectories of each league, individually (but see Fig. **4**). Through visual inspection of these trajectories some variability on the average number of goals scored per match in some of the leagues was noticeable, while other leagues presented more stable patterns.

For a better graphical representation of the tendency of each league to remain inside of one performance channel, their trajectories and performance channels were plotted together. This option constitutes, indeed, the best graphical representation of performance tracking.

In general, intra-individual trajectories remained inside the same channel or oscillated between two channels, indicating performance stability and tracking.

For a more objective measure of intra-individual change and inter-individual differences in leagues´ seasons stability, a count was made on how many times each league score remained inside the same channel. Based on this assumption, Cohen's *K* provided a measure of tracking that can be interpreted according to Landis and Koch [23] suggestions: $K \geq 0.75$=excellent; $K \geq 0.40$ and $K < 0.75$=moderated to good; $K < 0.40$=poor. These cut-off values allowed us to classify each league level of stability.

The degree of performance stability of each league, once identified, makes it important to investigate which performance channel(s) lodge the corresponding performance trajectories. The more prominent examples in Table **4** are the German and Dutch leagues. These remained constantly inside the highest performance channel, meaning that more goals per match occurred in these championships. The opposite happened in the Czech and French leagues that maintained their performance trajectories inside the lowest performance channel. Another relevant fact observed was the attainment of the highest performance channel in four random seasons by the Spanish league, despite having registered the lowest degree of stability. The overall kappa result obtained for all leagues indicated the existence of moderate stability in the average number of goals scored per match in European football leagues during the analyzed period.

Finally, Foulkes and Davies γ constitutes the last procedure to tracking. It examined the number of leagues that maintain the same relative rank. This probability was computed after modeling the best linear, or non-linear, individual trajectory of each league. If $\gamma = 0.50$, there is no tracking; if $y = 1$, tracking is perfect; if $\gamma < 0.50$, tracking is irrelevant [15]. Results indicated high tracking in the leagues´ average number of goals scored.

In short, through the application and analysis of this statistical procedure it was possible to verify that the average number of goals scored per match in major European leagues has not changed since season 2001/02, being scored a minimum average of two goals per football match. The Dutch and German leagues presented the best scores while the Czech and French leagues were those presenting less number of scored goals per match. The Spanish league registered unstable scoring patterns. It was not verified a pronounced trajectory in the average number of goals scored for any league since the few oscillations between the distinct performance zones (or channels) were random.

Even though the selected variable used to exemplify the application of this statistical procedure provided a very small and restricted picture of the wide and complex performance landscape of football, it should be regarded as a quite good example of the utility and relevancy of this technique to link important performance information gathered throughout time.

CONCLUSIONS

The field of longitudinal data analysis is very broad and varied, namely in sports performance. Although a whole range of techniques are available and user-friendly in many statistical software, not much research deals with the enthralling and challenging issues of stability and prediction of team performance. We tried to fill a gap in the literature suggesting the novel use of the idea of tracking and some of its main statistics and meanings. The illustrated example highlights the usefulness of tracking performance measures longitudinally.

Although longitudinal data analysis presents many challenges within the framework of individual and/or team performance, the concept of tracking and its various statistical modes of graphical display and numerical summaries appear to be useful for many purposes. Yet, a further research avenue would be to model team performance over time conditional on time-varying or time invariant covariates as well as tracking. A more thought provoking avenue would be to model systems of interacting behaviors that change over time in linear or non-linear ways.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGEMENT

Declared none.

APPENDICE

Appendice – Total Goals and Matches Per League and Season

Leagues	2001/02		2002/03		2003/04		2004/05		2005/06		2006/07		2007/08		2008/09	
	G	M	G	M	G	M	G	M	G	M	G	M	G	M	G	M
English	1001	380	1000	380	1012	380	975	380	944	380	931	380	1002	380	942	380
French	716	306	838	380	884	380	826	380	811	380	855	380	868	380	858	380
German	893	306	821	306	909	306	890	306	861	306	837	306	860	306	894	306
Dutch	847	300	903	306	911	306	948	306	912	306	960	326	1001	322	893	312
Italian	806	306	789	306	816	306	963	382	991	380	969	380	970	380	988	380
Spanish	961	380	1016	380	1015	380	980	380	936	380	942	380	1021	380	1101	380
Portuguese	818	306	806	306	729	306	711	306	681	306	554	240	553	240	552	240
Czech	590	240	592	240	578	240	531	240	567	240	534	240	556	240	582	240
Russian	584	241	609	240	598	240	539	240	587	240	562	240	578	240	600	240
Greek	524	182	617	240	627	240	557	240	560	240	558	240	601	252	547	252

G – Total Goals; M – Total matches

REFERENCES

[1] Lago C. Ganar o perder en el fútbol de alto nivel. Una Cuestión de Suerte? Motricidad Eur J Hum Mov 2005; 14: 135-50.

[2] Glazier PS. Game, Set and Match? substantive issues and future directions in performance analysis. Sports Med 2010; 40(8): 625-34.

[3] Carling C, Reilly T, Williams AM. Performance assessment for field sports. Oxon: Routledge 2009.

[4] Drust B. Performance analysis research: Meeting the challenge. J Sports Sci 2010; 28(9): 921-2.

[5] Rampinini E, Coutts AJ, Castagna C, Sassi R, Impellizzeri FM. Variation in top level soccer match performance. Int J Sports Med 2007; 28(12): 1018-24.

[6] Reilly T, Gilbourne D. Science and football: a review of applied research in the football codes. J Sports Sci 2003; 21(9): 693-705.

[7] McGarry T, Anderson DI, Wallace SA, Hughes MD, Franks IM. Sport competition as a dynamical self-organizing system. J Sports Sci 2002; 20(10): 771-81.

[8] McGarry T. Applied and theoretical perspectives of performance analysis in sport: scientific issues and challenges. Int J Perform Anal Sport 2009; 9: 128-40.

[9] Davids K, Araújo D, Shuttleworth R. Applications of dynamical systems theory to football. In: Reilly T, Cabri J, Araújo D, Ed. Science and Football V. London: Routledge 2005; pp. 556-69.

[10] Gregson W, Drust B, Atkinson G, Di Salvo V. Match-to-match variability of high-speed activities in premier league soccer. Int J Sports Med 2010; 31(4): 237-42.

[11] Lago C, Dellal A. Ball possession strategies in elite soccer according to the evolution of the match-score: the influence of situational variables. J Hum Kinet 2010; 25: 93-100.

[12] Franks I, McGarry T. The science of match analysis. In: Reilly T, Ed. Science and soccer. Liverpool: E & FN Spon 1996; pp. 363-75.

[13] Hughes M, Franks I. Notational analysis of sport. 2nd ed. Oxon: Routledge 2004.

[14] O'Donoghue P. Research methods for sports performance analysis. Oxon: Routledge 2009.

[15] Maia J, Silva R, Seabra A, Lopes V. The importance of tracking (stability and prediction) in longitudinal designs: an applied study to the epidemiology of physical activity and sports performance. Rev Port Cienc Desporto 2002; 2(4): 41-56.

[16] Lago C, Martín R, Seiru-lo F. Team performance in soccer. Modelling the main variables of F.C. Barc Apunts Educ Fís Deportes 2007; 4: 51-8.

[17] Taylor J, Mellalieu S, James N, Barter P. Situation variable effects and tactical performance in professional association football. Int J Perform Anal Sport 2010; 10: 255-69.

[18] Lopes V. O tracking da actividade física. Tema de estudo e inquietação no âmbito do desenvolvimento motor. Rev Port Cienc Desporto 2004; 4(2): 62-5.

[19] Armatas V, Yiannakos A, Zaggelidis G, Skoufas D, Papadopoulou S, Fragkos N. Differences in offensive actions between top and last teams in greek first soccer division. A retrospective study 1998-2008. J Phys Educ Sport 2009; 23(2): 1-5.

[20] Hughes M, Churchill S. Attacking profiles of successful and unsuccessful teams in Copa America 2001. In: Thomas R, Cabri J, Araújo D, Eds. Science and Football V. London: Routledge 2005; pp. 222-8.

[21] Schneiderman E, Kowalski C. Longitudinal data analysis. Release Notes, Version 3.2. Dallas 1993.

[22] Waddington C. Strategy of the genes: a discussion of some aspects of theoretical biology. London: Allen and Unwin 1957.

[23] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977; 33(1): 159-74.

Received: December 30, 2012 Revised: June 20, 2012 Accepted: June 21, 2012