Research on the Statistical Model System of Each Balanced Game Based on Big Data

Li Yan
Tianjin Ren'ai College, Tianjin, China
liyan2019@tjrac.edu.cn

Abstract. The article uses the holistic, comprehensive and dynamic diversified and complex thinking mode of big data system theory. The article creates a generalized linear model for the value of each technical and tactical performance index and the game result in each score-balanced game. To define the linear relationship between the technical and tactical performance indicators of the game and the probability of winning the game; finally, adopt the data series inference method to define the significance of the linear relationship between each technical and tactical performance indicator and the probability of winning the game. The research found that the information provided by the model created in the paper can be used in football practice such as game performance evaluation, opponent information detection, selection and modification of training preparation plans.

Keywords: Big data, football, training, statistical model.

1. Introduction
Football technical and tactical training is a complex, multi-input, multi-output and dynamic control system. Multi-input and multi-output refer to the input and output of information between coaches, athletes, researchers, media, etc. The existence and transmission of information is based on the movement of materials, and the characteristics of materials are not only expressed through a type of information. It is expressed in multiple forms [1]. The dynamic control system refers to the understanding of the football training process as having different states at different moments, in different states, applying different technical training focuses.

With the rapid development of electronic information technology and video analysis technology, the acquisition and update of football technical and tactical performance analysis data have become more and more convenient. Professional football game video analysis system can generate massive real-time game data. It can be said that football technical and tactical performance analysis has entered the "big data" era. However, just mastering the "big data" and failing to properly explain and apply it will inevitably limit the practical application value of football game performance analysis [2]. The best solution is to establish an appropriate mathematical model to define the causal relationship between football technical and tactical index values and football game results. Unfortunately, the current analysis of football technical and tactical performance is still mainly at the level of descriptive and comparative research, using complex mathematical models to define technical and tactical performance-related variables (such as game technical and tactical indicators, etc.) and game outcome
variables (such as game wins). The research on the relationship between losses, goals, advancements, etc., is relatively lacking. The generalized linear model is considered to be a very effective and complex mathematical model that can be used to define the correlation between technical and tactical performance-related variables in team sports events and competition outcome variables. At present, this model is rarely used in football technical and tactical performance analysis. Based on this research background, the paper introduces K-type cluster analysis, generalized linear model and data series inference method to mathematically model the technical and tactical performance indicators and game results in all football matches with balanced scores, and its game techniques and tactics. The performance is discussed in order to provide solutions for the application of "big data" in football technical and tactical performance analysis.

2. Research on the Control Elements of Football Skills and Tactics Training System

2.1. The control element model of the system

Football technical and tactical training is an extremely complex system and practical process. It must be guided by a variety of basic theories such as system theory and cybernetics in order to achieve the best results. The system is mainly composed of two major control elements: offensive elements and defensive elements. The distribution of each specific element is shown in Figure 1.

Many control elements have their own specific status and functions in training. Offensive elements and defensive elements must be paid equal attention to. It is not biased to be conservative with offense, to promote offense with defines, and to develop in a balanced manner. However, the application of each specific control element is not isolated. Both offense and defines often need to use several control elements comprehensively.

2.2. Research and design of control elements

Among the many control elements mentioned above, scoring, assists, steals, passing and other elements are particularly important. Since these elements can be quantified, statistical analysis and research can be done conveniently, so this element group is selected for research design.
2.3. Research objects and research methods
The research object is the 100 goals and 10 award-winning goals nominated for the International Football Awards in 2019-2019. The video observation and statistics method are used to observe and count 100 goals and 10 winning goals [3]. The standard of statistics is three core concepts, namely, observation and statistics of each goal's attacking method, the three-dimensional space of the attack and the attacking goal. Technical actions, and descriptive statistical analysis of the classification results.

2.4. Methods of Mathematical Statistics
A generalized linear model is created for the value of each technical and tactical performance index and the game result in each score-balanced game to define the linear relationship between the game’s technical and tactical performance index and the probability of winning the game. In the generalized linear model, the following models are created:

\[
\ln(\text{Odds}) = a + bx + e
\]  

Among them, \(\text{Odds} = P/(1-P)\) and \(P\) are the team’s winning probability, \(x\) is the value of a certain technical and tactical performance index, and \(a\), \(b\), and \(e\) are constants. According to this model, the following derivation can be made: Suppose that when \(x = x_0\) is \(P_0 = 50\%\), \(\text{Odds}_{0} = P_0/(1-P_0) = 1\).

\[
\ln(\text{Odds}_0) = 0 + a + bx_0 + e
\]  

When \(x_1 = x_0 + \Delta x\),

\[
\ln(\text{Odds}_1) = a + b(x_0 + \Delta x) + e = b\Delta x (1)
\]  

When \(x_2 = x_0 - \Delta x\),

\[
\ln(\text{Odds}_2) = a + b(x_0 - \Delta x) + e = b\Delta x (2)
\]

When \(\Delta x = 0.5\), \(P_2 - P_1\) represents the change in the probability of winning of a certain team when the value \((x)\) of a certain technical and tactical performance index of a certain team changes by one unit. When \(\Delta x\) = the standard deviation (SD) of \(x\), \(P_2 - P_1\) represents the value of a certain technical and tactical performance index of a certain team, from a typical small value (-SD) to a typical large value (+SD): The change in the team's probability of winning when the standard deviation is increased by 2 standard deviations. K-type cluster analysis and generalized linear model creation are completed by data statistics software SPSS20.0. The order of magnitude inference method is calculated by Excel2007.

3. Results and discussion

3.1. Analysis of key technical indicators
The correlation analysis method is used to find the correlation coefficients between 26 technical indicators and the number of goals, and the key technical indicators that affect the number of goals of the World Cup football team are found through logical analysis. The correlation matrix is composed of \(r\) between the columns of the matrix [4]. The matrix indicates that the values of the \(i\) row and the \(j\)-th column are \(r\) and \(P\) between \(i\) and \(j\). Use SPSS22 software to find the \(r\) and \(P\) of each technical indicator and the number of goals, and find out the two key technical indicators that affect the number
of goals. Now the correlation coefficient and significance level of each key technical indicator and the number of goals are counted. In Table 1 Correlation Analysis Matrix.

Table 1. Correlation matrix between key technical indicators and the number of goals.

|       | Goal | Shoot right | Key pass |
|-------|------|-------------|----------|
| Goal  | 1    | 0.954*      | 0.965*   |
| Shoot right | 0.954* | 1           | 0.889    |
| Key pass | 0.965* | 0.889      | 1        |

It can be seen from Table 1 above that there is a clear correlation between the two technical indicators of shooting on target (r=0.954, P=0.046<0.05) and key pass (r=0.965, P=0.035<0.05) and the number of goals, has significant statistical significance.

3.2. Analysis of shooting index

Shooting refers to in football when the opponent's offensive player shoots the ball into the opponent's goal, including hitting the goal post and the crossbar. Sun Zhiyi's "Comparative Study on the Elements of Goals in the Last Three World Cups" draws from multiple technical indicators that affect goals, the shooting effect is mainly reflected in the number of shots, the rate of shots on target, and the ratio of shots to goals. Liu Hongyou and others pointed out in the "Analysis of Key Indicators for Winning the Chinese Football Super League" that the only significant key variable that determines the outcome of the game is to shoot the goal [5]. This study is based on the conclusion that: the correlation coefficient between the shooting index and the number of goals r=0.954, P=0.046<0.05, there is a significant correlation, which is consistent with the previous research conclusions of experts. In order to explore the factors that affect the shooting index, the relevant analysis method is used to analyze, and it is concluded that the assist is the key factor affecting the assist index, see Table 2.

Table 2. Correlation analysis matrix of shooting index and assist index.

|       | Shoot right | Assist |
|-------|-------------|--------|
| Shoot right | 1           | 0.966* |
| Assist  | 0.966*      | 1      |
|        | 0.034       |

Assists belong to one (or two) players of the scoring team. He gave the ball to the goal player in any way, including any form of passing, or a shot that bounced off the lintel or post, which meant that he (we) "assisted" the goal, which was the one before the goal. The key pass with the foot. Table 2 the correlation coefficient between shooting and assists is r=0.966, P=0.034<0.05, there is a significant correlation, and it has significant statistical significance.

3.3. Analysis of key passing indicators

In football, a key pass refers to a forward pass that can threaten the opponent's defensive team's goal or shot, and can effectively break through the opponent's defines line to obtain a shot, which is the so-called pass threat. In the "Analysis of the Passing Characteristics of Asian Teams in the 20th World Cup", Song Fei pointed out that passing is the link that connects the whole team in football matches and is an important means of tactical play. In the "Comparative Analysis of the Main Technical Indicators of the 19th World Cup and the 2012 Chinese Super League", Guo Shuai analyzed that in the comparison between the World Cup teams and the Chinese Super League teams, there are very
significant differences in the number of threats. This study concluded that the key passing index is the key index that affects the number of goals scored by the World Cup teams (r=0.965, P=0.035<0.05), which is very similar to the above conclusions [6]. As a key technical indicator that directly determines the number of goals, key passes will inevitably provide a greater theoretical basis for sports training. Correspondingly, the indicators that determine the key passes are also particularly important. The relevant analysis is shown in Table 3.

Table 3. Correlation matrix of key passing indicators and other indicators.

|                  | Key pass | take the chance | Shot  | Break into the big penalty area |
|------------------|----------|----------------|-------|-------------------------------|
| Key pass         | -0.99    | 0.976*         | 0.961 |                               |
| take the chance  | 0.01     | -0.993         | -0.99 |                               |
| Shot             |          | 0.007          | 0.01  |                               |
| Break into the big penalty area | 0.039 | 0.01 | 0.014 |                               |

3.4. Analysis of characteristics of attacking technical actions
The FIFA Awards began in 2009 and is an award established by former FIFA President Blatter to commemorate the legendary Hungarian star Puskas. The criteria for the selection of candidate goals are: 1) Appreciation, must be a beautiful goal, able to show personal superb skills, physical fitness or team cooperation goals; 2) Non-discriminatory, regardless of the nature of the match, regardless of the gender of the goalkeeper With nationality and popularity, they all have equal chances of selection; 3) Inevitability, scoring cannot rely on luck or the opponent's mistakes, not accidental, and must be repeatable; 4) The competition is fair, and the competition process must conform to the principle of fair play. Among them, 1) and 3) play a pivotal role in the development of football. Take Messi, Neymar, Ibrahimović and Cristiano Ronaldo (hereinafter referred to as Ronaldo) as an example. They have been nominated many times in the 10th International Football Awards [7]. Judging from the nominations of the 4 people’s goal-making technical moves, the common feature is that the participating scoring technical moves have a certain degree of repetition, such as "excellent shot" was nominated 7 times (Messi 4 times, Neymar 3 times), "Sideways volley" 2 times (Ibrahimović 1 time, Neymar 1 time), "dribbling volley" 2 times (Messi 1 time, Neymar 1 time), "upside down golden hook" 2 Times (Ibrahimović 1 time, Cristiano Ronaldo 1 time).

It can be seen that all the goals that participated in the selection have a repetitive pattern. All nominated and award-winning technical moves are not based on luck and are repeatable, even the extremely difficult inverted golden hook, heel, and vigorous volleys that exceed half of the time are all repetitive. Since the action is repeatable, it must be able to use high-tech means to carry out scientific analysis, which provides strong support for the next step of football training, such as using a 3D motion capture system to record each moment of the action, and build a biological model to reveal the main points of the action, Regularity, thereby improving the technique of attacking the door.

4. Conclusion
The generalized linear model can effectively define the causal relationship between various game technical and tactical performance indicators in a football game and the outcome of the game, and thus can be used to determine which game technical and tactical indicators are the winning indicators. The
information provided by the created model can be used in football practice such as game performance evaluation, opponent information detection, selection and modification of training preparation plans.

References
[1] Abd Razak, A. N., Karim, Z. A., Hashim, A., & Gengatharan, K. Effects of Small-Sided Game on Offensive Effectiveness Using Tactical Periodization in Youth Football Player. European Journal of Molecular & Clinical Medicine, 7 (2) (2020) 5954-5963.
[2] Savrievich, F. F., & Shodievich, T. M. TECHNICAL AND TACTICAL TRAINING OF FOOTBALL PLAYERS. Web of Scientist: International Scientific Research Journal, 2 (04) (2021) 259-266.
[3] Özcan, İ., Eniseler, N., & Şahan, Ç. Effects of small-sided games and conventional aerobic interval training on various physiological characteristics and defensive and offensive skills used in soccer. Kinesiology, 50 (1) (2018) 104-111.
[4] Sarmento, H., Figueiredo, A., Lago-Peñas, C., Milanovic, Z., Barbosa, A., Tadeu, P., & Bradley, P. S. Influence of tactical and situational variables on offensive sequences during elite football matches. The Journal of Strength & Conditioning Research, 32 (8) (2018) 2331-2339.
[5] Batista, J., Gonçalves, B., Sampaio, J., Castro, J., Abade, E., & Travassos, B. The influence of coaches’ instruction on technical actions, tactical behaviour, and external workload in football small-sided games. Montenegrin Journal of Sports Science and Medicine, 8 (1) (2019) 29-32.
[6] Machado, J. C., Barreira, D., Teoldo, I., Travassos, B., Júnior, J. B., Santos, J. O. L. D., & Scaglia, A. J. How does the adjustment of training task difficulty level influence tactical behavior in soccer? Research quarterly for exercise and sport, 90 (3) (2019) 403-416.
[7] Folgado, H., Gonçalves, B., & Sampaio, J. Positional synchronization affects physical and physiological responses to preseason in professional football (soccer). Research in Sports Medicine, 26 (1) (2018) 51-63.