ISSN 2616-5805 Vol. 5, Issue 2: 45-50, DOI: 10.25236/AJMS.2024.050208

Exploring a Quantitative Evaluation System for Young Chinese Footballers: A Preliminary Approach

Ge Heming[1,a,*]

[1]*Beijing No.4 High School, Beijing, China*
[a]*13901130349@163.com*
[*]*Corresponding author*

Abstract: *This paper establishes a quantitative evaluation system for young football players in different positions and selects important evaluation indicators using a multiple linear regression model. Our results show that physical attributes are significant for players in all positions, while technical attributes are only significant for midfielders and central defenders. Mental attributes are significant for full-backs and forwards, and negative images directly affect the evaluation of forwards.*

Keywords: *Football, Outstanding Athletes, Talent Selection, Indicators System*

1. Introduction

1.1. Problem Background

In recent years, the global football stage has seen a steady supply of talented young stars partially due to the successful implementation of overseas systems to select and train young talents. Therefore, that the Chinese football community must enhance its competitiveness by improving its talent selection systems, the core of which is a quantitative evaluation system.

Moneyball, a movie, recounts a real-life story of events that unfolded in Major League Baseball in the United States during the early 21st century. It shows how a middle-tier team selected players using quantitative methods and set a record winning streak. Although baseball's complexity and unpredictability cannot be compared to football, quantitative methods are being used extensively to select exceptional players in developed football countries. These models can be co-opted into China's football talent selection.

1.2. Literature Review

The Elite Youth Player Identification Manual issued by the Chinese Football Association and the Internal Player Reports of some elite football clubs in China were mainly referenced as the basis for China's evaluation system for young footballers. The main flaws of such primitively quantitative evaluation systems are summarized below.

1.2.1. Incomplete data indicators and lack of correlation between data

The mere 20 to 30 indicators still need to meet the complexity of football. The data has yet to be subject to a rigorous correlation analysis. For example, it is even more important for a player to anticipate the movement of the ball and the opponents than simply running and reacting when the ball is challenged. In other words, anticipation can effectively compensate for speed.

1.2.2. Equally weighted indicators that are not true to football

Since footballers in different positions have different technical attributes, their potential must be analyzed from different perspectives. However, all indicators in the referenced systems are weighted equally, overlooking the uniqueness of players' abilities or the requirements of different positions. Therefore, we have designed the weights of secondary indicators for players in different positions and obtained the weights of primary indicators through regression analysis.

1.2.3. Availability of positive indicators but no negative indicators

Youth athletes' growth relates to negative indicators, such as injury vulnerability and negative

ISSN 2616-5805 Vol. 5, Issue 2: 45-50, DOI: 10.25236/AJMS.2024.050208

image, as much as positive indicators, such as physical, technical, tactical, and mental attributes. If positive indicators facilitate their growth, negative indicators frustrate their growth.

There are numerous theoretical studies on quantifying talent selection criteria for youth football players both at home and abroad. For example, *Psychometric Attributes of Sport Diagnostics in the German Football Talent Identification and Development Program (Oliver Höner, Andreas Votteler, Markus Schmid, Florian Schultz & Klaus Roth, 2014)*, a widely referenced paper, emphasizes the importance of scientifically evaluating the performance of young athletes and underpins the future examination of the prognostic validity of success in adulthood. The study found that in measuring the abilities of football players, the football-specific technical skills such as ball control and dribbling are more effective in differentiating between players than the tests of speed abilities such as sprint. The researchers also provide detailed testing guidelines and recommended scoring criteria, offering better tools to enhance the effectiveness of Germany's talent identification and development program.

Similarly, *Research Status and Reference of Foreign Youth Football Player Selection Criteria(Liu Meihan ,2017)*, a domestic paper, has investigated foreign selection criteria systems and suggested holistically reflecting the potential of young players using multifactorial and multidimensional assessments. The authors argue that although this method is relatively complex, demanding, and laborious, it represents the direction of talent selection research in China. In another domestic paper entitled *Constructing a Multidimensional Selection Criteria System for Outstanding Young Male Football Players(Liu Weimin, Zhou Shu, Peng Yanming ,2005)*, the authors sampled 16 players from Henan Jianye Football Club's reserve team and 16 players from Wuhan Sports University's college team, analyzed numerous indicators of football-specific abilities using metrics and expert questionnaires, and found that the optimal mix of abilities encompasses agility, 30-meter sprint and task orientation, and 1v1 anticipation.

In China, these theoretical studies often face practical challenges such as insufficient data and a need for authoritative assessment indicators, leading to a scarcity of research papers that can establish differentiated quantification criteria systems for players in different positions..

2. Project and Methods

2.1. Project

This study aimed to explore the main factors enabling player success in different positions and forms using quantitative methods.

In designing the indicator system, references were made to documents such as *The Elite Youth Player Identification Manual* issued by the Chinese Football Association and the *Internal Player Reports* of some elite football clubs in China.

Players' market value was the main observed variable for ease of quantitative analysis.

Since young players' statistics are not easily accessible elsewhere, data were selected from the official Football Manager Editor 2022 (FM) software, an alternative source that is deemed valuable and compelling because all its data has been reviewed by over one thousand football experts. However, the data selection process might be subject to survivorship bias and thus overlook important explanatory factors for the exceptional quality of the sample whose subjects are all world-class players. Furthermore, the data analysis might be affected by data insufficiency due to our limitations in data processing.

The talent selection process was optimized using regression as a mathematical tool, one that will become more effective as mathematical tools evolve with the more extensive application of machine learning and large-scale models. The present study is just a catalyst, hoping to inspire new methodologies for football development.

2.2. Framework and Methods

The present study mainly aims to build a quantitative analysis system for young players and provide a data basis for selecting star players by analyzing and evaluating their potential with redesigned evaluation indicators and indicator weights. The fundamental ideas are summarized below.

ISSN 2616-5805 Vol. 5, Issue 2: 45-50, DOI: 10.25236/AJMS.2024.050208

2.2.1. Given the changing positions of and requirements on players as modern football evolves, the field positions are classified into four categories: central defenders, wingbacks, midfielders, and forwards. Each of these four positions is built with an analytical evaluation model, as each must possess significantly different abilities

Youth athletes' growth relates to negative indicators, such as injury vulnerability and negative image, as much as positive indicators, such as physical, technical, tactical, and mental attributes. If positive indicators facilitate their growth, negative indicators frustrate their growth.

2.2.2. Correlations between player success and the indicators

As global football is fully marketized, players' market value can be taken as a proxy for their worth: an increase in market value can be seen as a relative success, and a decrease as a relative failure. For each position, 50 observation samples were randomly selected, with 35 randomly selected from the top 70 players with the biggest increase in market value and 15 randomly selected from the bottom 30 players with the biggest decrease in market value, as measured in million Euros.

Primary and secondary indicator systems were built in tandem. While the former system was generated from the latter, the weights of secondary indicators were designed based on experience to obtain primary indicators, which were normalized using Z-Score normalization. In the future, an expert ratings system can complete this process, thus yielding better results.

Regression analysis was used to find the relationship between primary indicators and levels of player success.

2.3. Data Source and Collection

2.3.1. Selection of sample players

A sample of 200 players was randomly selected based on the above description with their market value mainly from the TransferMarkt website (TM).

2.3.2. Data on players in different positions

The primary indicators system of players in different positions was designed into five positive indicators (technical, tactical, physical, mental, and communication attributes) and two negative indicators (injury vulnerability and negative image). The primary indicators were divided into secondary ones weighted differently based on their position characteristics. The weighted secondary indicators were then summed up as a score of their corresponding primary indicator.

The player data were mainly collected from TM and FM for two reasons. First, their data is deemed objective and consistent with the mainstream perception of the industry both because TM claims that all its information comes from real data that has been systematically reviewed by experts and has been disclosed and because FM is a famed game software whose parent company, Sports Interactive, employs more than 1,000 investigators to update the annual player data of all global leagues. The second reason concerns the purpose of this study: to explore whether our methodology works for our evaluation system, one that will be refined as our data improves should the methodology work.

2.3.3. Data collection

The secondary indicators were collected for the 200 players and the scores of primary indicators were calculated according to the predetermined rules.

3. Research Process

3.1. Research Steps

Multiple linear regression was used to analyze how key indicators for players in different positions affect their latest market value and changes in their market value. The sequence of steps is summarized below.

3.1.1. Data processing and preliminary analysis

All the modeling data with no obvious anomalies were ensured to be ready before a multiple linear regression equation with all independent variables was established, and its significance was subjected to preliminary analysis.

ISSN 2616-5805 Vol. 5, Issue 2: 45-50, DOI: 10.25236/AJMS.2024.050208

3.1.2. Stepwise regression analysis

The Akaike Information Criterion (AIC) was used to choose a combination of variables that would best explain variations in the target variable, simplify the model, and improve its predictive accuracy.

3.1.3. Multicollinearity test

The variance expansion factor (VIF) was used to assess the model's multicollinearity and ensure non-correlation between variables.

3.1.4. Models and conclusions

The multiple linear regression models were established, and their variables were evaluated using the F-test and t-test.

3.2. Linear Regression Models Established

3.2.1. For Central Defenders

Their model results are summarized below:

*Y*Latest Market Value= **13.2003** + **5.9958*X***Physical Attributes+ **3.8157*X***Technical Attributes

This equation means that if a central defender improves by one standard deviation in technical or physical attributes (i.e., showing significant improvement compared to other players), their market value will increase accordingly. Specifically, their value will increase by 3.8 million euros for one standard deviation increase in technical ability. In contrast, their value will increase by 6 million euros for one standard deviation improvement in physical attributes.

3.2.2. For Wingbacks

Their model results are summarized below:

*Y*Latest Market Value= **10.609** + **5.9537*X***Physical Attributes+ **3.6018*X***Technical Attributes

The analysis results indicate that:

Specifically, their value will increase by 3.8 million euros for each standard deviation increase in technical ability. In contrast, their value will increase by 6 million euros for onestandard deviation improvement in physical attributes.

3.2.3. For Midfielders

Our results are as follows:

*Y*Latest Market Value= **28.963** + **12.1066*X***Physical Attributes+ **12.0014*X***Technical Attributes

The results indicate that when midfielders score one standard deviation above the average level in technical or physical attributes, their market value will increasecorrespondingly. Specifically, for one standard deviation increase in technical and physical attributes, a player's market value will increase by12 million euros.

3.2.4. For Forwards

Our results are as follows:

*Y*Latest Market Value= **26.3120** + **8.2331*X***Physical Attributes+ **8.5360*X***Mental Attributes - **6.7562*X***Mental Attributes

Theseresults show that when forward players score one standard deviation above the average level in mental or physical attributes, their market value will increase by 8.53 million euros and 8.23 million euros, respectively. Conversely, if their negative image score increases by one standard deviation, their market value will decrease by 6.76 million euros.

4. Summary and Conclusion

The following table tabulates the regression results for young players in the four positions.

ISSN 2616-5805 Vol. 5, Issue 2: 45-50, DOI: 10.25236/AJMS.2024.050208

Table 1: The regression results for young players

Position	Major influencing factors on the latest market value
Forwards	"Physical Attributes," "Mental Attributes," "Negative Image"
Midfielders	"Physical Attributes," "Technical Attributes"
Wingbacks	"Physical Attributes," "Mental Attributes"
Central Defenders	"Physical Attributes," "Technical Attributes"

Notably, "Physical Attributes" consistently emerged as an influential indicator across all positions.

"Technical Attributes" was a key indicator of a football player's fundamental skills and a significant predictor of the latest market value for both center-backs and midfielders. Based on our design of secondary indicators, there are differences in the composition of technical attributes for players in different positions. Specifically, the market value of central defenders is more influenced by their tackling, heading, and passing skills. At the same time, midfielders' market value depends more on their passing and dribbling skills and ability to connect and transfer the ball in the front and back fields and defeat pressing in a small area.

"Mental Attributes" as a primary indicator clearly distinguishes the market value of attacking players, with secondary indicators such as "concentration, stability, work rate, pressure management, composure, and decisions" being its crucial components.

On the other hand, a "negative image" emerged as a negative predictor in the model for young forwards' market value. This finding aligns with the understanding that a player's market value can be significantly impacted by their inability to effectively control their emotions and uphold good sportsmanship, which could result in bans and other disciplinary actions.

5. Recommendations

Based on the preceding analysis, we propose the following recommendations for the selection of young talented players:

- Since different positions require different core capabilities, designing indicators and corresponding weights is crucial when evaluating young players.

- A notable negative factor is the negative image associated with forwards, significantly impacting their market value. In contrast, this negative image does not significantly affect players in other positions, which may have something to do with the tactical discipline required.

- Various explicit indicators influence a player's market value differently. For central defenders, physical attributes have a coefficient of 6.00, whereas technical abilities have a coefficient of 3.82. This result indicates that physical attributes contribute over 1.5 times more to a player's market value than technical abilities. In the case of midfield players, physical attributes (coefficient of 12.10) and technical abilities (coefficient of 12.00) have nearly equal contributions to a player's market value. Similarly, for wingbacks, physical attributes (coefficient of 5.94) contribute around 1.5 times more to their market value than mental attributes (coefficient of 3.60). For forwards, physical attributes (coefficient of 8.23) and mental attributes (coefficient of 8.53) have almost equal contributions, with mental attributes even slightly more essential.

- Our analysis is based on world-class players under 22 years old with a market value exceeding 1 million euros. That some factors are insignificant in this analysis does not mean they are unimportant. For instance, forwards' technical attributes are insignificant, possibly because the analyzed players excel in this aspect, and thus, their values cannot be differentiated by these attributes.

- The analysis is more of a methodological attempt, with significant data limitations as only one year of data was used due to data collection constraints and subjectivity in the scoring system. Future improvements in building an expert system, collecting detailed and multi-year data, and enhancing analytical tools can lead to a more precise evaluation system.

Acknowledgements

Acknowledgments are extended to Associate Professor Han Bo at Beijing Sport University, Professor Lang at Beijing Normal University, and Director Wei Kexing at Beijing Guoan Football Club

ISSN 2616-5805 Vol. 5, Issue 2: 45-50, DOI: 10.25236/AJMS.2024.050208

for their support and advice.

References

[1] Liu Meihan (2017). Research Status and Reference of Foreign Youth Football Player Selection Criteria. Science & Technology of Stationery & Sporting Goods, 6: 14-15.
[2] Liu Weimin, Zhou Shu, Peng Yanming (2005). Constructing a Multidimensional Selection Criteria System for Outstanding Young Male Football Players. China Sport Science and Technology, Vol 41, No.3: 83-85.
[3] Oliver Höner, Andreas Votteler, Markus Schmid, Florian Schultz & Klaus Roth (2014). Psychometric properties of the motor diagnostics in the German football talent identification and development programme. Journal of Sports Sciences, June 2014:145-159.