Fuzzy Study on the Winning Rate of Football Game Betting

Woo-Joo Lee[1], Hyo-Jin Jhan[2], Seung-Hoe Choi[3,*]

[1]Department of Mathematics, Yonsei University, Seoul, Korea

[2]School of Aerospace and Mechanical Engineering, Korea Aerospace University, Goyang, Korea

[3]School of Liberal Arts and Science, Korea Aerospace University, Goyang, Korea

Abstract

This study aims to find variables that affect the winning rate of the football team before a match. Qualitative variables such as venue, match importance, performance, and atmosphere of both teams are suggested to predict the outcome. Regression analysis is used to select proper variables. In this study, the performance of the football team is based on the opinions of experts, and the team atmosphere can be calculated with the results of the previous five games. ELO rating represents the state of the opponent. Also, the selected qualitative variables are expressed in fuzzy numbers using fuzzy partitions. A fuzzy regression model for the winning rate of the football team can be estimated by using the least squares method and the least absolute method. It is concluded that the stadium environment, ELO rating, team performance, and importance of the match have effects on the winning rate of Korean National Football (KNF) team from the data on 118 matches.

1. Introduction

For the member countries of International Federation of Association Football (FIFA), their primary interest is the athletic performance in the World Cup qualifiers and the finals [1]. Predicting the results of football matches as well as other popular sport matches is also important for fans [2]. Thus, the methods for forecasting football results have been studied in various ways [3-10]. Researchers can take advantages of the long history of football because the football results have been continuously recorded. Since modern football has been first introduced to Korea, KNF team has played a total of 941 games, with a winning percentage of 54.3% [11].

Presently, the primary methods for predicting a football game are sorted into two types: 1) quantitative methods and 2) qualitative methods. The quantitative approach is used to predict the outcome of a match by analyzing the previous datasets of both teams competing in the match [12-16]. On the other hand, the qualitative approach is used to predict the outcome of a match by expressing the characteristics that cannot be expressed numerically. These might include the athletic performance of the teams, the match location, and the physical condition of the players [17]. In addition, there is also a combination of two approaches. However, while the research on quantitative methods of predicting the outcome of sport events is more prominent, the research on qualitative and combined methods seems to be scarce [18-19]. Furthermore, the variables that affect the performance of each team before the match include the condition of the players' participation, the condition of teamwork, and the starting squads. It can be quite exciting for football fans to predict the outcome of matches using variables that can be inferred before the match [3].

* Corresponding author. E-mail address: shchoi@kau.ac.kr

 Tel.: +82-2-300-0073; Fax: +82-2-300-0492

The objective of this study is to propose the quantitative and qualitative variables affecting the winning rate of the football team before the match, and to present a statistical model of the winning rate of the team using the selected variables. Accordingly, we express qualitative variables in fuzzy numbers and use the selected quantitative and qualitative variables to infer a fuzzy regression model of the football team's winning rate. For this purpose, we analyze 118 games of Korean National Football (KNF) Team. The overall workflow can be sorted into three steps:

(1) Collecting raw data:

Public football in-game data have long been recorded and published online. The quantitative data provided by the Korea Football Association (KFA) on 118 matches from September 2, 2011 to November 14, 2019 are collected. Also, the qualitative data are primarily extracted from reviews of football experts.

(2) Pre-processing and transformation:

Numerically derived variables such as the winning rate, difference in ELO rating, and atmosphere are added, and the details of those are described by Eqs. (7)-(12). Also, the qualitative variables such as the performance of the team and the importance of a match are transformed into fuzzy numbers.

(3) Fitting the data into the fuzzy regression model and estimate the coefficients:

Section 2 presents fuzzy regression model for the winning rate of the football game betting. Section 3 gives the numerical results of KNF team using the absolute deviation method. Section 4 concludes the study.

2. Fuzzy Regression Model for the Winning Rate of the Football Game Betting

This section introduces a fuzzy regression model for the winning rate using triangular fuzzy numbers and fuzzy partitions. Regression model (Eq. (1)) is presented to analyze the impact of match location (p_g), ELO rating difference (d_e), atmosphere of the national football team (t_a), performance (A_a), and importance (I_g) of the match on the winning rate (P_w) of the football team.

$$P_w = F\left(p_g, d_e, t_a, A_a, I_g\right) \tag{1}$$

To estimate a given regression model in the classical statistical method, the residuals, which are the differences between observations and estimated values, must meet Gauss-Markov conditions. In other words, the residuals must satisfy the independence, the normality, and the heteroscedasticity [20-21]. If there is a correlation between independent variables, such as atmosphere, athletic performance and match importance, the use of traditional regression will result in multicollinearity problems. In addition, all variables used in the regression model can be expressed as quantitative variables, but sometimes the independent or dependent variables can be expressed as qualitative or categorical. Although qualitative data can be quantitatively expressed using probability variables to estimate regression models, this can represent a reduction or omission of the information contained in the qualitative data. Thus, it may be more efficient to use qualitative or linguistic data for the regression analysis. The proposed statistical analysis method to account for the causal relationship of such linguistic or qualitative data is the fuzzy regression model.

In 1980, Professor Tanaka introduced the first fuzzy regression model, which applied fuzzy numbers to the regression model introduced by Professor Zadeh in 1965 [22-23]. Since then, fuzzy regression models have been used in many areas [24-26]. Data used for fuzzy regression can be expressed in triangular fuzzy numbers. The triangular fuzzy number, denoted by numbers A $= (l_a, a, r_a)_T$, consists of the mode (a), left endpoint (l_a), and right endpoint (r_a) [26]. The membership functions of the right and left sides of the triangular fuzzy number are shown in Fig. 1.

$$L_A(x) = \begin{cases} \dfrac{1}{a-l_a}(x-l_a), & l_a \le x \le a \\ \\ 0, & o.w \end{cases} \tag{2}$$

and

$$R_A(x) = \begin{cases} \dfrac{1}{a-r_a}(x-r_a), & a \le x \le r_a \\ \\ 0, & o.w \end{cases} \tag{3}$$
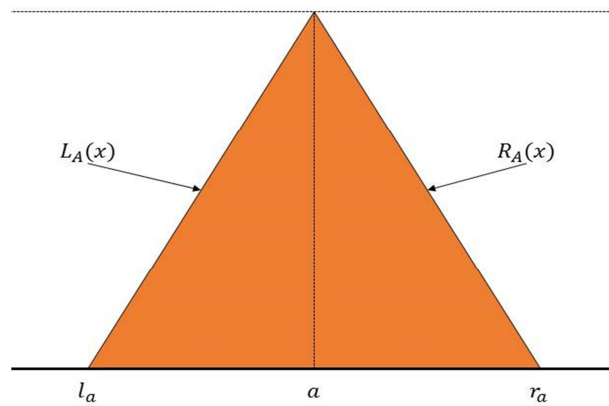

Fig. 1 Fuzzy number

In this study, we apply a fuzzy partition to represent the winning rate, performance, and importance of the football team in triangular fuzzy numbers. If $[A_i : A_i(0) \subseteq X, i = 1, \dots, n+1]$ is set as a fuzzy partition of the universal set X, it satisfies for each i and any $x \in X$:

$$A_i(0) \ne \phi \tag{4}$$

$$U_{i=1}^{n+1} A_i(0) = X \tag{5}$$

$$\sum_{i=1}^{n+1} A_i(x) = 1 \tag{6}$$

where $A_i(0)$ is a 0-level set of a fuzzy set A_i [27].

Fig. 2 implies that the sum of the membership value at one point is 1. Fig. 2 shows a fuzzy segmentation consisting of five points: $a_1, a_2, a_3, a_4,$ and a_5. We use the least absolute deviation method suggested by Choi and Buckley [24] to estimate the regression coefficients of the fuzzy regression model (Eq. (1)). They first transformed the fuzzy number into crisp number to estimate the regression model for the center [24].
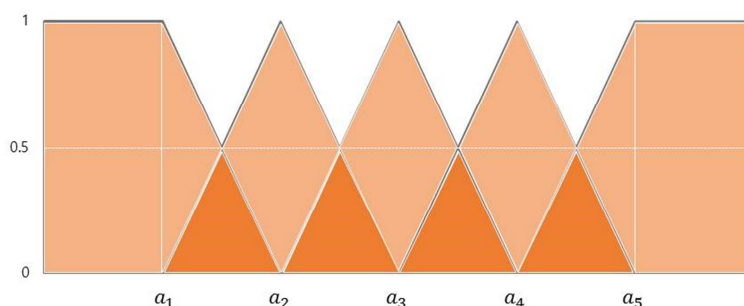

Fig. 2 Fuzzy partition

3. Regression Analysis for the Winning Rate of KNF Team

This section examines the variables that affect the winning rate of KNF team in the betting, and estimates the fuzzy regression model for KNF team. This study uses the data provided by KFA on 118 matches from September 2, 2011 to November 14, 2019, along with the columns written by the experts who analyzed KNF team's performance [2]. This section explains various quantitative and qualitative variables that have influenced the winning rate of KNF team, and it subsequently produces a fuzzy regression model for the winning rate of KNF team using the selected variables.

3.1. Quantitative data for the model

We study 118 sets of KNF team's data to forecast the result of the football matches. The ANOVA results show that KNF team's winning rates not only vary with match types such as friendly matches, World Cup qualifiers, East Asian Cup finals, and World Cup matches, but also vary with the region in which the matches take place ($p-value < 0.05$) [28]. KNF team has its lowest winning rate in the World Cup matches and the highest winning rate in the World Cup qualifiers. Also, the winning rate is higher at home games than at neutral games and away games (Table 1).

Table 1 Results according to the match type and region

Match Type	Number of matches	Winning rate	Location	Number of matches	Winning rate
Friendly	60	0.42	Home	50	0.56
World Cup qualifier	35	0.63	Neutral	40	0.43
East Asian Cup	6	0.48	Away	28	0.45
Asian Cup	11	0.57			
World Cup	6	0.23			
Total	118	Mean: 0.49			

The quantitative data on the football team can be classified into two categories: before the match and after the match. The quantitative data that are available before the match are the team's FIFA ranking, ELO rating, and the dividend rate for the football team announced by the betting company. The FIFA / Coca-Cola World Ranking, which was released for the first time in August 1993, can be obtained from the website of the International Football Confederation [29-30]. The ELO rating, which was developed by Arpad Elo in 1997, can be found online [31-32]. The sample correlation (ρ_{ef}) between the FIFA and ELO rankings for KNF team is positive ($\rho_{ef} = 0.48, p-value < 0.000$). Since ELO rating is instantly published after every match and FIFA ranking is updated every month, it may be more efficient to predict KNF team's winning rate by using ELO rating than FIFA ranking.

KFA releases the results of every match through the KNF team website. There are 17 types of data on game scoreboards published by KFA, including goals for (GF), goals against (GA), shots (x_{i3}), shots on target (x_{i4}), fouls (x_{i5}), yellow cards, red cards, offsides, corner kicks (x_{i9}), saves, clears, interceptions, aerial ball winning rates, pass success rates (x_{i14}), ground ball winning rates (x_{i15}), pass blocks, and ball possession. In order to estimate the regression model proposed in this section, we use the least squares method. This study is conducted based on records provided by KFA from 2011 to 2019. KFA has announced that the online update of 2017-2019 season data is in progress.

With the variable selection method commonly used in regression analysis [28], our variables that influence the goal difference ($d_g(i)$) of KNF team are derived as shown in the following equation ($p-value < 0.000, adjR^2 = 0.61$):

$$d_g(i) = -2.613 + 0.507 d_{e_i}^{*} + 0.115 x_{i3} + 0.248 x_{i6} + 0.016 x_{i15} \tag{7}$$

where $d_{e_i}{}^*$ is the difference in the ELO rating. The above regression equation shows that the variations in KNF team's score difference are statistically significant with shots (x_{i3}), ground ball winning rates (x_{i15}), and ELO ratings ($d_{e_i}{}^*$). In addition, the positive results for the estimated coefficients of shots, ground ball winning rates, and ELO ratings indicate that as the ranking of the opposing team decreases and the team's performance increases, the score increases in relation to the KNF team's losses. Moreover, the estimated regression equation shows that as the aggressiveness of KNF team's defense increases, the number of losses decreases. The variables used in this study are not standardized because they have different characteristics. As a result, the estimated regression model (Eqs. (7) and (8)) show the intercept values.

As shown in Table 1, KNF team's performance can vary according to match importance such as World Cup matches, World Cup qualifiers, Asian Cup finals, and friendly matches. Therefore, in this study, we choose to analyze the data on KNF team's World Cup qualifier matches, which are held regularly ($p-value < 0.000, adjR^2 = 0.877$).

$$d_g\left(i\right) = -9.47 + 0.644 g_{p_i} - 0.133 x_{i5} - 0.269 x_{i9} + 0.095 x_{i14} + 0.03 x_{i15} \tag{8}$$

The estimated regression equation shows that the location of the match (g_{p_i}), pass success rates (x_{i14}), and ground ball winning rates (x_{i15}) are important variables for KNF team's performance. In addition, the estimated regression equation indicates that a high number of fouls (x_{i5}) in a match negatively affect the team's chances of victory because the Korean team is one of the top teams in Asia. This is the same for the number of corner kicks (x_{i9}).

3.2. Qualitative data for the model

Before a football match, the media and fans are interested in predicting the outcome. Although football results do not seem easy to predict, prediction methods can be implemented by using variables that can affect the outcome of a match, including the ranking of the two teams, the match location, the atmosphere and teamwork of each team, the physical condition of players, athletic performance, and the importance of the match. The media provides important information about the game through a preview with football experts, and then a betting company sets a dividend on the victory, draw, and defeat of the match according to the customer's bet. The winning rate of the football team can be predicted before the match based on the opinions of football experts and the information announced by the betting company [33-34]. In this study, we use expert columns and dividend rates published by the betting company, which both provide predictions and analyses of various sports in addition to football to predict the winning rate of KNF team. The dividend rate for a match announced by a betting company cannot be unrelated to the outcome of the match [35-36]. The winning rate calculated by the dividends of winning (d_w), drawing (d_d), and losing (d_l) for a team in a particular match are equal to:

$$p_w = \frac{d_d d_l}{d_d d_l + d_w d_l + d_w d_d} \tag{9}$$

The above expression is given by odds-makers, who estimate the odds in betting or competing [37-38]. The sample correlation (ρ_{wg}) between the winning rate of KNF team and the difference between the scored and lost goal of KNF team is:

$$\rho_{wg} = 0.616\left(p-value < 0.000\right) \tag{10}$$

The regression equation for the winning rate of KNF team (p_w) based on KNF team's goal difference (g_d) is as follows ($adjR^2 = 0.374, p-value < 0.000$):

$$p_w = 0.436 + 0.06 g_d \tag{11}$$

This shows that the winning rate based on the dividends is related to the outcome of KNF team. Other important variables that can determine the outcomes of matches are the performances in previous matches, the current teamwork, the current athletic performance, the physical condition of the players competing in the game, and the importance of the game.

The current atmosphere of the football team can change depending on the outcome of previous game results. When a team goes on a losing streak, the atmosphere of the team will deteriorate, and negative public opinions can cause significant changes, including the replacement of the coach. Therefore, the outcome of the national team's recent matches is an important factor in predicting the national team's winning rate. Based on the recent five matches, the current atmosphere of the national team is defined as:

$$t_{a_i} = 0.3e_{(i-1)}{}^s + 0.3e_{(i-2)}{}^s + 0.2e_{(i-3)}{}^s + 0.1e_{(i-4)}{}^s + 0.1e_{(i-5)}{}^s \tag{12}$$

where $e_{1-t}{}^s$ stands for the ELO rating the KNF team acquired at t matches before the current match.

The importance of the match can affect its outcome. As with other team sports, the importance of the match determines the starting players in the match. While the rookie players are usually assigned in a squad of friendly matches, the best players are assigned for important matches that determine whether the team is qualified for a tournament. In this study, the importance of matches is classified into "very important", "important", "normal", "ordinary", and "very ordinary".

KNF team's athletic performance is another variable that can affect the outcome of a match. A team's athletic performance can be evaluated by its ability to score, the ability to defend, and the organizational ability that expresses the organic movement of the defense, offense, and midfield. In addition, the physical strength and condition of the teams participating in the match are also the factors of the team's performance.

This study utilizes the columns of three experts who regularly analyze KNF team's performance and publish their opinions on team organization prior to KNF matches. These columns enable us to tally all of the expert-derived positive and negative remarks regarding KNF team's offensive, defensive, and organizational abilities, as well as the strength and physical conditioning of KNF team. Team performance is expressed as one of the five categories: "very good", "good", "normal", "bad", or "very bad".

3.3. *Fuzzy regression model for the winning rate of KNF team*

In this study, the following fuzzy partitions are used to represent the winning rate of the football team as fuzzy numbers.

(1) $P_w{}^1$: Defeat with large margin, denoted by $P_w{}^1 = (0, 0, 0.25)$

(2) $P_w{}^2$: Defeat with small margin, denoted by $P_w{}^2 = (0, 0.25, 0.5)$

(3) $P_w{}^3$: Draw, denoted by $P_w{}^3 = (0.25, 0.5, 0.75)$

(4) $P_w{}^4$: Win with small margin, denoted by $P_w{}^4 = (0.5, 0.75, 1)$

(5) $P_w{}^5$: Win with large margin, denoted by $P_w{}^5 = (0.75, 1, 1)_T$

Similar to the winning rate, we apply the following fuzzy partitions for the importance of the match.

(1) $I_g{}^1$: Very ordinary, denoted by $I_g{}^1 = (0, 0, 0.25)_T$

(2) $I_g{}^2$: Ordinary, denoted by $I_g{}^2 = (0, 0.25, 0.5)_T$

(3) $I_g{}^3$: Normal, denoted by $I_g{}^3 = (0.25, 0.5, 0.75)_T$

(4) $I_g{}^4$: Important, denoted by $I_g{}^4 = (0.5, 0.75, 1)_T$

(5) $I_g{}^5$: Very important, denoted by $I_g{}^5 = (0.75, 1, 1)_T$

In this study, we also express KNF team's performance along with the winning rate and importance as fuzzy numbers. We present the national team's performance in a combination of offensive ability, defensive ability, organizational ability, and physical conditioning evaluated by football experts. The overall competition performance of the four categories is equal to 2. Thus, the universal set of the KNF team's ability to apply all four categories is equal to $X = [0.5, 1.5]$. Fig. 3 shows the fuzzy division of a football team into five stages, including "very bad", "bad", "normal", "good", and "very good".
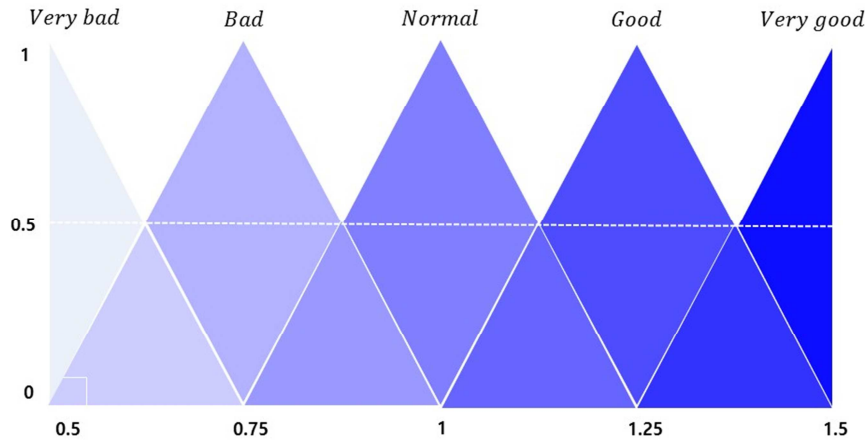

Fig. 3 Fuzzy partition for the national team's performance

We describe the national team's performance as fuzzy numbers using the following divisions:

(1) A_a^1: Very bad, denoted by $A_a^1 = (0.5, 0.5, 0.75)_T$

(2) A_a^2: Bad, denoted by $A_a^2 = (0.5, 0.75, 1)_T$

(3) A_a^3: Normal, denoted by $A_a^3 = (0.75, 1, 1.25)_T$

(4) A_a^4: Good, denoted by $A_a^4 = (1, 1.25, 1.5)_T$

(5) A_a^5: Very good, denoted by $A_a^5 = (1.25, 1.5, 1.5)_T$

The football experts' opinions are combined with the national team's athletic ability and match importance. The betting company publishes the opinions of three or four experts on its website before national team [2, 34].

Table 2 Data of KNF team

Number	Opponent	Location (p_g)	Goal difference (d_g)	ELO rating difference (e_d)	Atmosphere (t_a)	Athletic performance (A_a)	Match importance (I_g)	Winning rate (P_w)
1	Lebanon	Home	6	627	0.6	A_a^4	I_g^4	P_w^4
2	Kuwait	Away	0	195	-0.5	A_a^3	I_g^4	P_w^3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
26	Croatia	Home	-1	-166	-3.3	A_a^4	I_g^3	P_w^3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
76	China	Away	-1	271	-0.2	A_a^3	I_g^4	P_w^3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
117	Lebanon	Away	0	371	-4.3	A_a^5	I_g^4	P_w^5
118	Brazil	Neutral	-3	-320	-8.3	A_a^4	I_g^3	P_w^1

The following is a summary of the expert opinions on a match against Croatia in 2013: "South Korea's recent A-match performance has been bad. However, all of the players in good condition are going to be in the match, and we are expecting much from the Europa League players in particular. This friendly match is an important test match ahead of the final qualifying round." In addition, the expert opinions on the World Cup Asian qualifying rounds against China in 2017 are as follows: "The

Korean team is currently ranked second in its group for the World Cup Asia qualifying round. While winning points in away matches in China is important, key players from the Europa League are absent due to warnings and injuries. The opposing team, China, will put forth a strong effort because the new coach has yet to win his first victory."

We have compiled the national team's data based on the expert opinions above. Table 2 shows the quantitative data and the fuzzy data of KNF team since September 2011. In Table 2, the performance, the game importance, and the winning percentage are expressed in the triangle fuzzy numbers. The aim of this study is to infer the regression equation for the winning rate for the national team using the data given in Table 2. The most common method is to estimate the fuzzy regression model using the least squares method. However, it may be more efficient to use the least absolute deviation method instead of the least squares method, which is sensitive to outliers. In this study, the three levels of least absolute deviation method proposed by Choi and Buckley [24] are used to estimate the fuzzy linear regression model for the winning rate of KNF team given in Table 2.

$$P_{w_j} = c_1 p_{g_j} + c_2 d_{e_j}^{\ *} + c_3 t_{a_j} + c_4 A_{a_j} + c_5 I_{g_j} \tag{13}$$

Step 1: Consider the centroids of the fuzzy number.

To find the center of support for the winning rate, the performance and match importance are expressed in the triangular fuzzy number. The triangular fuzzy number used in this study is symmetrical, so each fuzzy number mode becomes a centroid.

Step 2: Estimate the regression coefficient based on the centroid data.

The pseudo-estimate is estimated based on the centroid data $[(p_{w_j}, p_{g_j}, d_{e_j}, t_{a_j}, a_{a_j}, i_{g_j}) : j = 1, \dots, 118]$ given in Table 2 and the least absolute deviation method, which minimizes the following equations.

$$\overline{p}_{w_j} = \hat{c}_1 p_{g_j} + \hat{c}_2 d_{e_j}^{\ *} + \hat{c}_3 t_{a_j} + \hat{c}_4 a_{a_j} + \hat{c}_5 i_{g_j} \tag{14}$$

$$\sum_{j=1}^{118} \left| p_{w_j} - c_1 p_{g_j} - c_2 d_{e_j}^{\ *} - c_3 t_{a_j} - c_4 a_{a_j} - c_5 i_{g_j} \right| \tag{15}$$

Step 3: Estimate the spread of the fuzzy number.

Using the pseudo-estimate $\overline{P}_{w_j}$ obtained in the second step, the estimate of the winning rate for the KNF team $\hat{P}_{w_j}$ is as follows.

$$\hat{P}_{w_j} = \hat{c}_1 p_{g_j} + \hat{c}_2 d_{e_j}^{\ *} + \hat{c}_3 t_{a_j} + \hat{c}_4 a_{a_j} + \hat{c}_5 i_{g_j} + (-\hat{l}, 0, \hat{r})_T \tag{16}$$

where $\hat{l}$ and $\hat{r}$ are, respectively, the values minimizing the following subject to $l > 0$ and $r > 0$:

$$\sum_{j=1}^{118} \left| l_{p_{w_j}} - \hat{c}_1 p_{g_j} - \hat{c}_2 d_{e_j}^{\ *} - \hat{c}_3 t_{a_j} - \hat{c}_4 l_{a_{a_j}} - \hat{c}_5 l_{i_{g_j}} + l \right| \tag{17}$$

$$\sum_{j=1}^{118} \left| r_{p_{w_j}} - \hat{c}_1 p_{g_j} - \hat{c}_2 d_{e_j}^{\ *} - \hat{c}_3 t_{a_j} - \hat{c}_4 r_{a_{a_j}} - \hat{c}_5 r_{i_{g_j}} - r \right| \tag{18}$$

The fuzzy regression model for the winning rate of KNF team using the data given in Table 2 and the three-stage absolute deviation method is:

$$\hat{P}_{w_j} = 0.078 p_{g_j} + 0.066 d_{e_j}^{\ *} + 0.039 A_{c_j} + 0.288 I_{g_j} + (-0.239, 0.251)_T \tag{19}$$

The estimated model indicates that the match location (p_g), ELO rating (e_d), athletic performance (A_a), and match importance (I_g) affect the national team's winning rate. The results are in line with the thoughts of both experts and fans. In other words, KNF team has a high chance of winning in matches when its performance is good, the importance of match is high, and its Elo rating is high. However, the atmosphere of KNF team does not seem to have a significant impact on the KNF team's winning rate. We think this result requires further study.

4. Conclusions

In this study, we suggested the fuzzy regression model for the winning rate of the football team betting, and estimated the regression model of KNF team using the data of 118 matches of KNF team from 2011 to 2019.

Before the football match, the qualitative data on the basis of football experts' reviews were collected to express KNF team's offensive, defensive, and organizational abilities, as well as the match importance, as a fuzzy number. We subsequently examined the impact of the fuzzy numbers defined by using the fuzzy partition and ELO rating on the football team's winning rate. It was found that the winning rate for KNF team was influenced by the match location, ELO rating, team performance, and match importance. Consequently, as all of the coefficients are positive, it turned out that home court advantage increases the possibility of winning. The better the KNF team's performance reviewed in previous games is and the more important the upcoming match is, the higher the winning rate of KNF team is expected. However, the atmosphere of KNF team was not found to influence the match results.

The results of 118 matches of KNF team were analyzed with a limited amount of information in this study. However, in the future, it would be useful to extend this research to include more games and extrapolate a model of the winning rate depending on the match type, e.g., World Cup matches, regional qualifiers, and friendly matches.

Acknowledgements

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] "FIFA," http://www.fifa.com/index.html, July 05, 2019.
[2] H. O. Steller, D. Sandor, and R. Verlander, "Issues in Sports Forecasting," International Journal of Forecasting, vol. 26, no. 3, pp. 606-621, July-September 2010.
[3] F. Wunderlich and D. Memmert, "Analysis of the Predictive Qualities of Betting Odds and FIFA World Ranking: Evidence from the 2006, 2010, and 2014 Football World Cups," Journal of Sports Sciences, vol. 34, no. 24, pp. 2176-2184, August 2016.
[4] M. Carpita, E. Ciavolino, and P. Pasca, "Exploring and Modelling Team Performances of the Kaggle European Soccer Database," Statistical Modelling, vol. 19, no. 1, pp. 74-101, February 2019.
[5] Y. F Alfredo and S. M. Isa, "Football Match Prediction with Tree Based Model Classification," International Journal of Intelligent Systems and Applications, vol. 11, no. 7, pp. 20-28, July 2019.
[6] T. G. Omomule, A. J. Ibinuolapo, and O. O. Ajayi, "Fuzzy-Based Model for Predicting Football Match Results," International Journal of Scientific Research in Computer Science and Engineering, vol. 8, no. 1, pp. 70-80, February 2020.
[7] E. Esme and M. S. Kiran, "Prediction of Football Match Outcomes Based on Bookmakers Odds by Using K-Nearest Neighbor Algorithm," International Journal of Machine Learning and Computing, vol. 8, no. 1, pp. 26-32, February 2018.
[8] J. M. Pérez-Sánchez, E. Góme-Déniz, and N. Dávila-Cárdenes, "A Comparative Study of Logistic Models Using an Asymmetric Link: Modelling the Away Victories in Football," Symmetry, vol. 10, no. 6, June 2018.

[9]   C. P. Igiri and E. O. Nwachukwu, "An Improved Prediction System for Football a Match Result," IOSR Journal of Engineering, vol. 4, no. 12, pp. 12-20, December 2014.

[10] H. Liu, M. Á. Gomez, C. Lago-Peñas, and J. Sampaio, "Match Statistics Related to Winning in the Group Stage of 2014 Brazil FIFA World Cup," Journal of Sports Sciences, vol. 33, no. 12, pp. 1205-1213, March 2015.

[11] "Korea Football Association," http://www.kfamatch.or.kr/svc/man/selectMainInfo.do, July 15, 2019.

[12] J. Goddard and I. Asimakopoulos, "Forecasting Football Results and the Efficiency of Fixed-Odds Betting," Journal of Forecasting, vol. 23, no. 1, pp. 51-66, January 2004.

[13] R. D. Baker and I. G. McHale, "Forecasting Exact Scores in National Football League Games," International Journal of Forecasting, vol. 29, no. 1, pp. 122-130, January-March 2013.

[14] G. Boshnakov, T. Kharrat, and I. G. McHale, "A Bivariate Weibull Count Model for Forecasting Association Football Scores," International Journal of Forecasting, vol. 33, no. 2, pp. 458-466, April-June 2017.

[15] M. J. Dixon and M. E. Robinson, "A Birth Process Model for Association Football Matches," Journal of the Royal Statistical Society Series D (The Statistician), vol. 47, no. 3, pp. 523-538, September 1998.

[16] J. H. Kim, G. T. Ro, J. S. Park, and W. H. Lee, "The Development of Soccer Game Win-Lost Prediction Model Using Neural Network Analysis," Korean Journal of Sport Science, vol. 18, no. 4, pp. 54-63, 2007.

[17] A. C. Constantinou, N. E. Fenton, and M. Neil, "Profiting from an Inefficient Association Football Gambling Market: Prediction, Risk and Uncertainty Using Bayesian Networks," Knowledge-Based Systems, vol. 50, pp. 60-86, September 2013.

[18] H. Rue and O. Salvesen, "Prediction and Retrospective Analysis of Soccer Matches in a League," Journal of the Royal Statistical Society: Series D (The Statistician), vol. 49, no. 3, pp. 399-418, September 2000.

[19] J. Sargent and A. Bedford, "Improving Australian Football League Player Performance Forecasts Using Optimized Nonlinear Smoothing," International Journal of Forecasting, vol. 26, no. 3, pp. 489-497, July-September 2010.

[20] S. H. Choi and J. H. Yoon, "General Fuzzy Regression Using Least Squares Method," International Journal of Systems Science, vol. 41, no. 5, pp. 477-485, March 2010.

[21] H. Y. Jung, J. H. Yoon, and S. H. Choi, "Fuzzy Linear Regression Using Rank Transform Method," Fuzzy Sets and Systems, vol. 274, pp. 97-108, September 2015.

[22] H. Tanaka, S. Uejima, and K. Asai, "Linear Regression Analysis with Fuzzy Model," IEEE Transactions on Systems, Man, and Cybernetics, vol. 12, no. 6, pp. 903-907, 1982.

[23] L. A. Zadeh, "Fuzzy Sets," Information and Control, vol. 8, no. 3, pp. 338-353, June 1965.

[24] S. H. Choi and J. J. Buckley, "Fuzzy Regression Using Least Absolute Deviation Estimators," Soft Computing, vol. 12, no. 3, pp. 257-263, May 2007.

[25] H. J. Jhang, H. Kwak, and S. H. Choi, "Analysis of the Outcome for the Korean Pro-Basketball Games Using Regression Model," Journal of The Korean Institute of Intelligent Systems, vol. 25, no. 5, pp. 489-494, October 2015.

[26] L. A. Zadeh, "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning—I," Information Sciences, vol. 8, no. 3, pp. 199-249, 1975.

[27] C. Mencar, M. Lucarelli, C. Castiello, and F. A. Maria, "Design of Strong Fuzzy Partitions from Cuts," 8th Conference of the European Society for Fuzzy Logic and Technology, September 2013, pp. 424-431.

[28] H. K. Kim and S. H. Choi, Statistical Analysis with Applications, Korea: Kyung Moon Sa, 2002. (In Korean)

[29] J. Lasek, Z. Szlávik, M. Gagolewski, and S. Bhulai, "How to Improve a Team's Position in the FIFA Ranking? A Simulation Study," Journal of Applied Statistics, vol. 43, no. 7, pp. 1349-1368, October 2016.

[30] K. Suzuki and K. Ohmori, "Effectiveness of FIFA/Coca-Cola World Ranking in Predicting the Results of FIFA World Cup Finals," Football Science, vol. 5, pp. 18-25, March 2008.

[31] L. M. Hvattum and H. Arntzen, "Using ELO Ratings for Match Result Prediction in Association Football," International Journal of Forecasting, vol. 26, no. 3, pp. 460-470, July-September 2010.

[32] "WFER: The World Football Elo Rating system," http://www.eloratings.net, June 10, 2019.

[33] "Sportstoto," http://www.sportstoto.co.kr/index.jsp, May 12, 2019.

[34] "Wisetoto," http://www.wisetoto.com/index.htm, May 12, 2019.

[35] "Betman," www.betman.co.kr, June 02, 2019.

[36] "William Hill," http://sports.williamhill.com/bet/en-gb, April 04, 2019.

[37] "Chosunilbo," http://premium.chosun.com/site/data/html_dir/2014/07/03/2014070300008.html, March 03, 2019.

[38] M. J. Dixon and P. F. Pope, "The Value of Statistical Forecasts in the UK Association Football Betting Market," International Journal of Forecasting, vol. 20, no. 4, pp. 697-711, October-December 2004.