Applied Artificial Intelligence
An International Journal

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/uaai20

Estimation of Market Values of Football Players through Artificial Neural Network: A Model Study from the Turkish Super League

Tugbay Inan & Levent Cavas

To cite this article: Tugbay Inan & Levent Cavas (2021) Estimation of Market Values of Football Players through Artificial Neural Network: A Model Study from the Turkish Super League, Applied Artificial Intelligence, 35:13, 1022-1042, DOI: 10.1080/08839514.2021.1966884

To link to this article: https://doi.org/10.1080/08839514.2021.1966884

APPLIED ARTIFICIAL INTELLIGENCE
2021, VOL. 35, NO. 13, 1022–1042
https://doi.org/10.1080/08839514.2021.1966884

Estimation of Market Values of Football Players through Artificial Neural Network: A Model Study from the Turkish Super League

Tugbay Inan[a] and Levent Cavas[b]

[a]Faculty of Sport Sciences, Department of Physical Education and Sport Teaching, Dokuz Eylül University, İzmir, Turkey; [b]Faculty of Science, Department of Chemistry, Dokuz Eylül University, İzmir, Turkey

ABSTRACT
Artificial intelligence (AI) has been widely affecting our lives in many ways for the last ten years. Potential applications of AI are currently being used in many sectors. However, the usage of AI has been very limited in sports science compared to the other sectors. In this study, we developed an artificial neural network model to estimate the market values of the players in the Turkish Super Football League. While the market value was selected as an output, the input values were: minutes played; goals scored; xG: assists; xA; defensive duels won %; tackle success %; shots on target %; Short-middle pass accuracy %; long pass accuracy %; and accuracy of passes to penalty area. After creating a neural network based on the input-output values with high training, validation and testing statistical values, input values were computed with the neural network created and then the output values were estimated. In conclusion, an artificial neural network is becoming one of the important modeling methods in all areas of life. Although the application of artificial neural networks is very limited in sports science, it is one of the suitable science disciplines where there is a lot of statistical data. The methodology proposed in this paper can also be used for talent selection in football. Moreover, it may help stop criticism by TV sports programmes and newspapers because market value estimation is based on fair performance parameters. Further studies are strongly recommended, not only for football but also for other sports disciplines.

ARTICLE HISTORY
Received 16 May 2021
Accepted 6 August 2021

Introduction

Artificial intelligence (AI) now affects all aspects of our daily lives. The tools developed under the area of artificial intelligence are commonly used in fields such as communication (Guzman and Lewis 2020), electronics (Bazrafkan and Corcoran 2018), chemistry (Cavas, Donut, and Mert 2016), medicine (Mamdani and Slutsky 2020) etc. The scientists in the field of artificial intelligence are highly interested in the areas in which there is big data. When they

CONTACT Tugbay Inan tugbay.inan@deu.edu.tr Faculty of Sport Sciences, Department of Physical Education and Sport Teaching, Dokuz Eylül University, Seferihisar, İzmir, Turkey
Supplemental data for this article can be accessed on the publisher's website

have data, they can easily model the system and make important predictions. One of the areas where big data is created is sports sciences. Although there is data available to use in the sports industry, the usage of artificial intelligence is still limited. Moreover, evaluation of some sports data, e.g. football data, is also connected with the sports industry. The number of studies related to predictions in sport are increasing. In football, artificial intelligence-based tools are being used to predict match results (Rahman 2020; Stübinger, Mangold, and Knoll 2020). Match results are strictly dependent on the quality of the players. Therefore, the clubs should include high quality players and so studies related to the selection of high-quality players are also important (Armatas et al. 2020; Pehlivan, Ünal, and Kahraman 2019).

Football player transfers are of great importance to the economic future of football clubs. Every year, the clubs spend time, energy and millions of dollars on player transfers. As such, it is very important for the clubs to make the right decisions when they transfer players but it is very difficult to predict whether the football player will contribute to her/his team or not.

To transfer a new player, generally, football clubs send their scout teams to watch the player live in football matches. This process should be carried out carefully because of the economics involved. Since it is based on human observations, sometimes subjective evaluations might occur (Bransen, Van Haaren, and Van De Velden 2019; Carmichael, Forrest, and Simmons 1999; Winemiller, Love, and Stamm 2020). On the other hand, a very long process is put in place to analyze match performance data and evaluate progress over the years. Football clubs must acquire the best players in order to compete with rival clubs and to win competitions. They spend a lot of money on football players to ensure that newly transferred footballers remain with the team for a long time. However, if they do not establish the right financial balance, or if they are not successful, the potential economic loss is expected. That is why UEFA has set some basic rules for clubs in European leagues regarding financial fair play(Union of European Football Associations UEFA 2015). In some cases, except for these basic rules, various developments may affect football finances, positively or negatively. As an example, the COVID-19 pandemic, which is currently affecting all countries, is also thought to affect the economy of football clubs and players on a large scale. Many clubs reduced the salaries of the footballers during the COVID-19 pandemic because of the decreased incomes of clubs. As such, it would be of little surprise if the market values of the players decreased because of the pandemic (Benchmark 2020).

Today, data collection and analysis in sport has become increasingly important, as it has in many other fields. In football, ranking is created according to the qualities of the players (speed, endurance, power, etc.) and in-competition performances (goals, assists, key passes, etc.). Team analysis also occurs before the competitions. Various analyses are made by using big data sources, such as Opta, Instat, Wyscout and Mediacoach.

Football is an important part in the clubs in terms of financial contribution compared to other sports disciplines. This contribution is mainly dependent on the fans of the football clubs. Therefore, this area can be named as "sport industry." Deloitte releases financial information about European Football. While the income of the top 20 clubs in the 1996–1997 season was 1.2 B €, it has now reached to 9.3 B € (DeloitteGroup 2020). This financial income generally comes from broadcasting, match day revenue and commercial activities.

Deloitte also reports that Barcelona is at the top of the money league for 2018–2019, with 800 M €. The Turkish Super League (TSL) is in 6th position in terms of broadcasting revenues within European Football. The commercial income of TSL is reported to be 200 M €. The total income also increased to 615 M € in 2018–2019 (Aktifbank 2019). One of the important criteria for the sporting success of football clubs is their ability to exercise strict financial discipline. In the past, Turkish Football clubs did not exercise this rigor for many reasons including pressure from fans, the structure of the clubs, governmental audits, absence of measurements, compensation claims, frequent changes of technical directors and payment of over-priced players and their representatives (2013; Akşar 2005; Akşar and Merih 2006). The failure of players with overestimated transfer values to adapt to their new clubs causes performance based problems in matches. In Turkey, four football clubs (Beşiktaş, Fenerbahçe, Galatasaray, Trabzonspor) are members of Borsa-İstanbul. Therefore, their financial structures are more transparent compared to the other clubs in the TSL. According to Public Disclosure Platform in Turkey (2020), the debts of Beşiktaş, Fenerbahçe, Galatasaray and Trabzonspor are 400 M €, 568 M €, 450 M € and 122 M €, respectively (Platform 2020). These debts are mainly due to the costs associated with players' transfers. These clubs are currently negotiating with UEFA for the structuring of their debts. From the data of transfermarkt, 500 M € has been spent for the transfer of the players within the TSL over the last 5 years (Transfermarkt 2020). From the market values of the players in TSL, it is very easy to see that these values have not been based upon performance parameters. Criticism of the market values of football players has become a frequent topic of conversation on Turkish TV. Moreover, there is a belief among sports audiences that payments to players in the Turkish Super Football League are unfair. This argument is mainly based on the fact that club budgets are spent in an uncontrolled manner (Akşar, 2013; Akşar and Merih 2006).

AI is playing a critically important role in all parts of our daily life. The evaluation of data created through daily actions is nearly impossible to analyze by traditional methods. The proper collection, storage and usage of human data (big data) are hot topics, not only within computer-related arenas but also within governments. It is very important to note that the countries which give special importance to AI technologies will be powerful countries of the future world. AI technologies have been successfully applied in many sectors, such as communication (Liu, Wang, and Zhou 2018), transportation (De Luca and Gallo 2020) and medicine (Kulkarni et al. 2020).

When the keywords *"artificial intelligence, sport"* were searched in the web of science (https://apps.webofknowledge.com/), only 240 records were found (the search date was February 25, 2020). From these statistics, it could be said that the application of AI technologies has been very limited in sports sciences compared to other industrial areas. The number of the records related to the keywords *"artificial intelligence, medicine," "artificial intelligence, communication"* and *"artificial intelligence, transportation"* was found to be 2204, 3802 and 735, respectively. No published data was found on the web of science related to the estimation of market values of football players via artificial intelligence. Therefore, we reviewed some of the published papers. We found related to the keywords "artificial intelligence, sport." (Şahin and Erol 2018) studied the estimation of the attendance demand in European football games by using an artificial neural network (ANN), adaptive neuro fuzzy inference system models and fuzzy rule-based system. They found that ANN is an effective predictor compared with the other models. Dynamical system theory and neural network modeling were studied by (Dutt-Mazumder et al. 2011) for the governing dynamics of association football players. The paperreveals basic theoretical backgrounds for dynamical system theory and neural network modeling which is necessary for sports scientists to enter this area of research. In the paper, ANN is also proposed to study the dynamical attributes of association football players. (Barron et al. 2020) studied the ANN estimation of player talent by using 340 performances, biographical and esteem variables. They found that passing, shooting, regaining possession and international appearances are important factors in the ANN prediction. They also suggest their methodology for identifying the correct players to transfer in football teams. (Barron et al. 2018) published a scientific report related to the identification of the key performance indicators in professional soccer influencing outfield players' league status through ANN. The authors reported that ANN efficiently estimated 78.8% of the players' league status with a test error of 8.3%. Moreover, the position-specific nature of performance in football may affect the current performance parameters used in scientific literature. Therefore, new methodologies such as those mentioned in the study of Barron et al. (2020) can be used.

Various applications of ANN in sports sciences – such as sport result prediction and analysis of tactical behaviors of handball players – can also found in the scientific literature (Bunker and Thabtah 2019).

The market values of the midfield players in TSL are modeled using ANN and these values are reevaluated in the present study. The contribution of the present study is to estimate the market values of TSL midfield players who have not been investigated via ANN before.

Data in the Study

In previous studies, it is seen that the variables used to estimate the market values of athletes are divided into three main categories: player characteristics, player performance, and player popularity (Müller, Simons, and Weinmann 2017). Although some studies have used player popularity to estimate the market values of football players, the most important variables that will determine the outcome of a football match are the performance parameters of the football player on the field rather than the popularity of the players outside the field.

Therefore, when the talent and performance measurements based literature were examined, it is clearly seen that the sports scientists have investigated the psychological characteristics (Christensen 2009; Miller, Cronin, and Baker 2015) anthropometric and physiological variables (Deprez et al. 2015; Mirkov et al. 2010) tactical skills (Christensen 2009; Kannekens, Elferink-Gemser, and Visscher 2011; Towlson et al. 2017) and technical skill variables (Huijgen et al. 2013; Olthof, Frencken, and Lemmink 2015; Waldron and Worsfold 2010) of football players. In classification studies related to different performance indicators in football, actions such as minutes played scoring, ball control, tackles, duels, shooting, assists, passing ability and decision-making have been the focus of recent papers. It is also reported in many studies that these parameters are used in estimation of market values (Frick 2011; Hughes and Bartlett 2002; Hughes et al. 2012; Lehmann and Schulze 2008; Müller, Simons, and Weinmann 2017; Ruijg and Van Ophem 2015; Serna Rodríguez, Ramírez Hassan, and Coad 2019; Wiemeyer 2003).

Midfield players are becoming increasingly popular due to their characteristics such as coordination between defense and forward players, blocking of attacks from opponent teams. Midfield players must be players who are resistant to high tempo, play an active role in defense and offense, and are equipped with football techniques. It is known that the majority of attacks during offense are initiated by midfielders. These players should have good ball control, accurate passing, shooting and crossing skills. In general, in today's football, it is seen that different types of players are selected while organizing the middle field. However, it should be taken into account that the offensive and defensive characteristics of the players are of equal weight and complement each other. From this point, midfield players have remarkable differences compared to the players in defense and forward positions. Therefore, the data were obtained from 358 midfield football players in the Turkish Super League and were based upon the 2018–2019 football season.

The market values of the players were taken from Transfermarkt (transfermarkt.com) and the performance parameters of the players were taken from Statsperform (OPTA, product of Statsperform). 306 matches were analyzed 11 performance parameters were selected for further analysis. Variable selections are based on the market value estimated related studies (Frick 2011; M. Hughes et al. 2012; Hughes and Bartlett 2002; Lehmann and Schulze 2008; Müller, Simons, and Weinmann 2017; Ruijg and Van Ophem 2015; Serna Rodríguez, Ramírez Hassan, and Coad 2019; Wiemeyer 2003). 83 of the players analyzed (358) did not play any matches in the 2018–2019 season. Therefore, these 83 players were omitted. 39 of the remaining 275 players played less than 90 minutes throughout the season, so they were also removed from the list. From the statistics of our manuscript, the maximum, minimum and mean minutes played per player during the season is 3294, 411 and 1394, respectively. The data of 236 players were analyzed in the study.

The data for these players was obtained from OPTA Sports Data (StatsPerform 2020). Data reliability and validity of OPTA Client system were studied by (Liu et al. 2013). Many papers based on OPTA data have been published (Gai et al. 2019; Konefał et al. 2018; Lago-Peñas et al. 2016; Oberstone 2011; Tunaru and Viney 2010). The definitions of the technical terms can be accessed from the web page for OPTA sports (StatsPerform 2020). Although there are some discussions about transparency, much research has been carried out by scientists on the validity and also reliability of TransferMarkt data. Transfermarkt data is based on player performance, age and possible injuries. The values of the players are continuously updated by the administrators, based on the discussions of registered users. Possible errors and their correction methodologies are discussed in the following papers (Gerhards and Mutz 2017; Gerhards, Mutz, and Wagner 2014; Herm, Callsen-Bracker, and Kreis 2014; Peeters 2018). Written permission was obtained from OPTA before the data was used. Only midfield players were analyzed in the study since there is a lot of variation if all positions are considered. Another reason is that there are more midfield players in the Turkish Super League than footballers playing other positions. Data used in the study is explained in Table 1.

Method

MATLAB 2019b was used to study ANN in this study (MATLAB 2020). The flowchart of the study is shown in Figure 1. The architecture of the ANN is given in Figure 2.

For the first part of the study, the data from OPTA was used to model by using ANN. We found 358 players' statistics from OPTA. However, since some of the data was missing, we reduced the data to 236. Since we could not create efficient ANN by using the data related to market value from OPTA,

Table 1. Data statistics of the study.

Variables Input layer	Data statistics Range	Mean ± S.D.
Minutes played	411–3294	1394.23 ± 716.795
Goals	0–14	2.22 ± 2.54
xG	0–15.27	2.25 ± 2.24
Assists	0–13	1.47 ± 2.05
xA	0–12.23	1.72 ± 1.79
Def duels won %	38.46–80.770	55.35 ± 7.31
Tackle success %	0–100	23.06 ± 24.64
Shots on target %	0–100	33.35 ± 15.66
Short, middle passes accurate %	69.79–95.08	86.05 ± 4.84
Long passes accurate %	0–88.89	55.7 ± 12.06
Passes to penalty area accurate %	0–100	51.14 ± 14.95
Output Layer		
Market value (€)	75000–17000000	1964195 ± 2917534

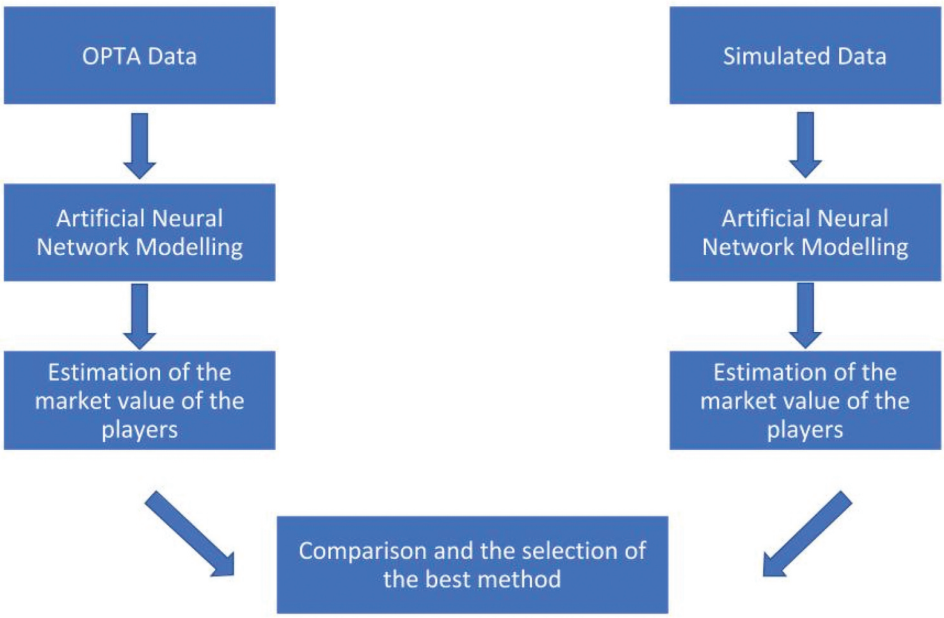

Figure 1. The flowchart of the study.

a new ANN model was created using simulated data. The simulated data was created based on the minimum and maximum values of the data from OPTA. Both models were compared by using performance parameters of ANN, such as training, testing and validation in MATLAB.

For the initial steps, the default settings of MATLAB Neural Network Fitting Tool were used. The percentages of training, validation and testing were 70%, 15% and 15%, respectively. The number of hidden layers was 10. The backpropagation algorithm was Levenberg-Marquardt (LM). Regression (R) values were used to estimate the accuracy and the precision of the network. R value if it is close to 1, shows the degree of relationship. Mean squared error

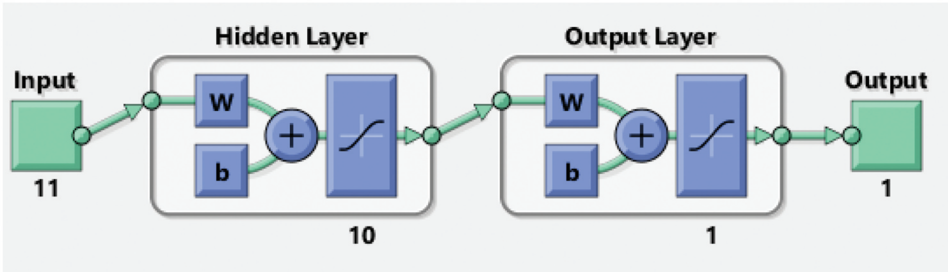

Figure 2. ANN structure of the study.

Table 2. The backpropagation used in the study.

Backpropagation algorithms	Function
BFGS Quasi-Newton backpropagation	*trainbfg*
Powell- Beale conjugate gradient backpropagation	*traincgb*
Fletcher- Reeves conjugate gradient backpropagation	*traincgf*
Polak-Ribiére conjugate gradient backpropagation	*traincgp*
Batch gradient descent	*traingd*
Batch gradient descent with momentum	*traingdm*
Variable Learning Rate backpropagation	*traingdx*
Levenberg-Marquardt backpropagation	*trainlm*
One Step Secant backpropagation	*trainoss*
Resilient backpropagation (Rprop)	*trainrp*
Scaled conjugate gradient backpropagation	*trainscg*

(MSE) is the average squared difference between outputs and targets. Lower MSE values show close relationships. Both R and MSE values were noted after each trial. Related to data analysis in MATLAB, raw and normalized versions of the data were used in the analysis. For the formation of the simulated data, the maximum and the minimum values from OPTA were taken and the random data between the maximum and the minimum values were created. The back propagation algorithms used in the study are listed in Table 2. For the final check, 60 inputs were randomly selected from OPTA and these were computed by using ANN to estimate ANN-predicted data. ANN-predicted data was also compared with the real market values.

Results

In this study, the market values of the midfield football players in Turkish Super League were modeled by using ANN. In the first step of the study, the data obtained from OPTA was used directly for the creation of ANN. The optimization of the number of hidden neurons for the simulated data is given in Table 3. The optimization of the number of hidden neurons for the data from OPTA (normalized and unnormalized) and the normalized simulated data are given in Supplementary Tables 1–3. In this analysis, the percentages of the training, validation and testing were 70, 15 and 15, respectively. While the

Table 3. The optimization of the number of hidden neurons for simulated data (data was not normalized).

Number of hidden neurons	Training percentage (n)	Validation percentage (n)	Testing percentage (n)	Epoch number	MSE			Regression coefficient		
					Training	Validation	Testing	Training	Validation	Testing
8	70(166)	15 (35)	15 (35)	5	0.033	0.064	0.097	0.962	0.953	0.972
9	70(166)	15 (35)	15 (35)	8	0.090	0.409	0.400	0.999	0.999	0.999
10	70(166)	15 (35)	15 (35)	6	0.097	0.003	0.847	0.999	0.999	0.999
11	70(166)	15 (35)	15 (35)	6	0.071	0.005	0.661	0.999	0.999	0.999
12	70(166)	15 (35)	15 (35)	5	0.069	0.065	0.016	0.983	0.977	0.983

performance parameters of the ANN related to data obtained from OPTA was extremely low in terms of R value, they were found significantly better than the simulated data. The optimum hidden neuron number was found to be 10 for the simulated data.

Normalized data was also used in the present study. The results from normalized data for both OPTA and the simulated data are given in Supplementary Tables 1–3. The performance parameters of the simulated data were significantly better than the data obtained from OPTA.

After studying the optimization of hidden neuron numbers, the optimization of the percentages of the training, validation and testing were studied under the optimized hidden neuron numbers. The same problem was also observed in optimization of the percentages of training, validation and testing. We found that the performance parameters (MSE, R) were lower than those of the simulated data (Table 4).

The normalized versions of the data for optimization of the percentages of the training, validation and testing were also explored in this study. Similar results were found and the performance parameters of the simulated data were much better compared to the data from OPTA (Supplementary Tables 4–6).

Different back-propagation algorithms for both simulated data and the data from OPTA were studied in the conditions of the optimized hidden neuron numbers and the optimized percentages of training, validation and testing. From the results, Levenberg-Marquardt backpropagation algorithm was selected as an optimum algorithm (Table 5 and Supplementary Tables 7–9).

After initial experiments, the market values of randomly selected data from OPTA were estimated by using ANN created under optimum conditions (simulated data (non-normalized), Levenberg-Marquardt Backpropagation Algorithm, Hidden neuron number to be 10 and the optimized percentages of the training, validation and testing to be 75, 15 and 15, respectively). The correlation coefficient was found to be −0.01994.

Discussion

This paper investigates, for the first time, the market values of the football players in the Turkish Super League by using ANN. There is always a discussion related to the market values of the players. Fans and also commenters criticize the market value of the players. They believe that, for some players, salaries are overestimated by club administrations (2013; Akşar 2005; Akşar and Merih 2006). Other players are thought to have market values that are lower than they deserve. Market values are affected by many different parameters. The clubs not only consider the performance parameters of players but also their advertising potential and other marketing opportunities they can bring. Since values are affected by more than just performance parameters, discussions regarding the fairness of market values occur.

Table 4. The optimization of percentages of training, validation and testing for simulated data (data was not normalized).

Number of hidden neurons	Training percentage (n)	Validation percentage (n)	Testing percentage (n)	Epoch number	MSE			Regression coefficient		
					Training	Validation	Testing	Training	Validation	Testing
10	90 (212)	5 (12)	5 (12)	33	0.018	0.034	0.059	0.999	0.999	0.999
10	80 (188)	10 (24)	10 (24)	41	0.071	0.037	0.061	0.999	0.999	0.999
10	70 (166)	15 (35)	15 (35)	7	0.011	0.015	0.019	0.999	0.999	0.999
10	60 (142)	20 (47)	20 (47)	35	0.010	0.025	0.034	0.999	0.999	0.999
10	50 (118)	25 (59)	25 (59)	14	0.021	0.032	0.057	0.998	0.993	0.996
10	40 (94)	30 (71)	30 (71)	13	0.032	0.012	0.077	0.998	0.094	0.995
10	30 (70)	35 (83)	35 (83)	6	0.093	0.094	0.089	0.991	0.988	0.990

Table 5. The comparative performance results of backpropagation algorithms for the simulated data (data was not normalized).

Backpropagation algorithms	Function	IN	MSE	R
BFGS Quasi-Newton backpropagation	*trainbfg*	44	0.011	0.999
Powell- Beale conjugate gradient backpropagation	*traincgb*	40	0.038	0.999
Fletcher- Reeves conjugate gradient backpropagation	*traincgf*	94	0.016	0.999
Polak-Ribiére conjugate gradient backpropagation	*traincgp*	6	0.194	0.999
Batch gradient descent	*traingd*	0	21.72	0.987
Batch gradient descent with momentum	*traingdm*	0	253.8	0.954
Variable Learning Rate backpropagation	*traingdx*	114	0.083	0.998
Levenberg-Marquardt backpropagation	*trainlm*	6	0.016	0.999
One Step Secant backpropagation	*trainoss*	19	0.062	0.998
Resilient backpropagation (Rprop)	*trainrp*	90	0.076	0.998
Scaled conjugate gradient backpropagation	*trainscg*	35	0.064	0.998

Table 6. The comparison of ANN estimated market values with the market values from OPTA (Values are given in Euro).

Sample Player	Market Value in OPTA	Estimated Market Value by ANN	Comment
Player 1	75000	6547700	Underestimated
Player 2	1000000	0	Overestimated
Player 3	17000000	3296100	Overestimated
Player 4	16000000	7729700	Overestimated

Football fans want to watch high quality football games and therefore the quality football games should be determined by performance parameters alone. As such, the aim of this paper is to propose a method of determining market values within the Turkish Super League that are only based on performance, thus bringing to an end a perceived lack of fairness. From our results, we found out that there is currently no correlation between performance parameters and the market values. Correlation coefficients were too low. This meant, an ANN to estimate market values could not be created by using the data from OPTA, so the authors of this paper wanted to create simulated data. Maximum and minimum market values from the Turkish Super League were used and simulated data was created ranging from 17 M € and 50 K €. Similarly, maximum and minimum performance values were also created. Since there was a strong correlation between the simulated performance parameters and the market values, an ANN was created efficiently. The ANN created by using simulated data was used to estimate the market values of 60 randomly selected players from the Turkish Super League. The correlation coefficient between the market values from OPTA and the estimated market values via ANN was −0.01994. This value showed that the current market values of midfield players in the Turkish Super League are not based on the performance parameters, which is one of the reasons for so much criticism on the subject. One of the important findings of this paper is that the ANN created by using simulated data efficiently estimated the market values of players based on the performance parameters. The samples estimated by the ANN created using simulated data are given in Table 6.

The market values and the performance parameters of the players are given in Table 6. Mean values are also given in Table 7.

Our results (Tables 6 and 7), revealed some very interesting and valuable outputs. In order to compare the market values, we selected some of the players from the Turkish Super League. According to Transfermarkt, the market value of player 1 is given as 75000 €. However, his value was estimated by the ANN created in this study was 6547700 €. When we compared his performance parameters and the market value with player 3 who has the highest market value, interesting results were seen. Player 3 has the highest market value and his total minutes played was 558. Player 1, however, has the lowest market value, yet he had played more minutes in total than player 3 (2151 min). Fans and TV commenters sometimes question how deserving players are of the money they receive. Certainly, from this perspective, it could be expected that a player with a high market value should also have played a high number of minutes. In the sports literature, played minutes is known as one of the important parameters in the transfer of players (Ruijg and Van Ophem 2015). The second interesting result is from player 2. The ANN created in this study found that the market value for this player had been overestimated. Although player 2 has a higher market value than player 1, his performance parameters are lower.

The most important differences between these two players are minutes played, asists, duels, tackle success, shots on target and pass during the 2018–2019 season. There are some proofs that some parameters such as shots on target, targeted passes percentages, duels, assists and played minutes are directly correlated with market values of the players (Frick 2011; M. Hughes et al. 2012; M. D. Hughes and Bartlett 2002; Lehmann and Schulze 2008; Müller, Simons, and Weinmann 2017; Ruijg and Van Ophem 2015; Serna Rodríguez, Ramírez Hassan, and Coad 2019; Wiemeyer 2003). The ANN created in this study also estimated a higher market value for player 4, since his performance parameters are higher than those of other players (Tables 6 and 7).

One disadvantage of the ANN created in the present study can be associated with the injuries. When a player experiences a sports injury, they can be away from the league for a long time and this will decrease her/his market value. In this case, market value estimation should not be created by using the ANN because it will decrease the market value as this is determined by performance parameters only. Instead, a coefficient constant may be created and the market value can be estimated by using the performance parameters of the previous years. Although injuries are seen as a disadvantage of the model, it is likely that players and managers will take greater care to avoid injuries if they understand that these will impact their market value.

Many efforts have been made to estimate players' market values and wages based on performance parameters. For example, (Barros and Leach 2006)

Table 7. The market values and the performance parameters of the players in the Table 6 with mean values.

Sample Player	Market value (€)	Minutes played	Goals	xG	Assists	xA	Def duels won %	Tackle succ. %	Shots on target %	Sh/m passes acc. %	Long passes acc. %	Passes to penalty area acc. %
Player 1	75000	2151	0	0.76	1	0.79	54.87	23.53	31.25	84.62	65.59	43.24
Player 2	1000000	505	0	0	0	0	44.12	0	0	84.08	66.67	0
Player 3	17000000	558	3	1.34	0	0.05	53.19	0	50	86.39	62.5	50
Player 4	16000000	1451	4	3.07	0	4.7	61.83	80	39.53	87.46	57.58	59.57
Mean Values	**4011667**	**1345**	**2.33**	**2.42**	**1.68**	**1.89**	**54.44**	**17.24**	**34.27**	**85.50**	**56.61**	**51.27**

studied performances in the English Premier Football League using data envelopment analysis (DEA). (Ribeiro and Lima 2012) estimated Portuguese football league players' wages by using the DEA. They showed that higher wages are related to increased efficiency. (He, Cachucho, and Knobbe 2015) investigated football players' performance from La Liga and their market values. They modeled market values (based on performance parameters) through extensive public data sources. (Carrieri, Principe, and Raitano 2018) studied the maximum earning parameters in Italian football players. They found popularity and the power exercised by players' agents to be more important than other parameters studied. (Caruso, Di Domizio, and Rossignoli 2017) also studied the relationship between Italian football players' wages and sports performance in Serie A. (Yaldo and Shamir 2017) investigated football players' wages by using computational methodology. Their method is based on pattern recognition algorithms. They applied their method on scoring, aggression and acceleration of football players. They compared the overpaid and underpaid football players in their article. We also present overpaid and underpaid football players in our study. However, it must be noted that it is very normal to obtain the results shown in Table 6, since no criteria are mentioned for the market values in the Turkish Super League. With the developments in the area of artificial intelligence, new algorithms will better estimate the market values of the players in the near future. This study can be considered as a starting point for future studies in Turkey.

There are many parameters that are investigated to estimate the market value of the football players. Although the correlations among players' market values and studied performance parameters were not found in the present study, some studies mention that there is a high correlation between market values and expert ratings. Some authors reported that market values can be estimated and the estimation could be highly correlated with actual transfer fees (Gerhards and Mutz 2017; Gerhards, Mutz, and Wagner 2014; Herm, Callsen-Bracker, and Kreis 2014). On the other hand, some papers have been published on the estimation of market values of the football players by using different parameters. (Singh and Lamba 2019) studied the influence of crowdsourcing, popularity and previous year statistics to estimate football players' market values. They found out that consistency, popularity, crowd estimation and performance parameters are effective factors to estimate the market values. Fantasy football based data also existed in the literature and it is also interesting to note that some authors have reported unusual factors such as birth month. Felipe et al. mentioned that the value of the players who were born in the first quarter of the year are higher than those of players who were born in the rest of the months of a year (Felipe et al. 2020). They also mentioned the importance of the parameters such as participating in international matches (UEFA) and being an attacking midfield player in terms of players' values. In our study, we could not create a model by using an ANN with data obtained from OPTA. We also checked the correlation among the variables. Since no correlation was found among the variables, other statistical methods were not applied. Other factors apart from performance parameters must have affected the market values of the players in Turkish Super League. It is suggested that market values of the players must be based on the performance parameters. The maximum and the minimum values can be determined by experts. The maximum and minimum values proposed by experts can be updated based on developments. This kind of dynamic information may contribute to the development of football itself. On the other hand, football is not only dependent on the performance parameters, but also on parameters outside the football such as advertising, jerseys, merchandising, popularity of players and social media influence. As it is explained in the paper, the modeling part of the study is based on the performance parameters alone. In some cases, some players whose transfer is delayed have to join the starting lineup without prior training. Since these kinds of players have to join matches without pre-season camps and enough preparation, they are more likely to become injured than their counterparts (Mallo and Alexandre 2012). One of the disadvantages of the model mentioned in the study could be long-term injuries, since injuries will decrease market value. On the other hand, as it was mentioned previously, players can be motivated not to be injured since they will know that their market values will decrease as a result of this. Pre-informing the football players of the details of how the models are formed is of great importance, if artificial intelligence based methods are applied in their leagues. The players' adaptation is also another important parameter for the application of the model.

In Turkey, the market value of a football player in the Turkish Super League depended on the players' reputation based on factors such as remarkable performance in his previous famous teams, high market value announcements from agents and media rumors. Other parameters such as rumors related to transfer of players to the top 4 teams, unwillingness of foreign players to transfer Turkish teams, missing infrastructure to educate and train young talents can also be listed (Akşar 2019, 2020). In a very new study carried out by Ćwiklinski, Giełczyk, and Choraś (2021), a machine learning approach is proposed to plan player transfers. They used Random Forest, Naive Bayes, and AdaBoost algorithms in their methodology and the authors reported to use neural network in their future studies (Ćwiklinski, Giełczyk, and Choraś 2021).

Professional football teams have to make a profit as they become businesses with multi-million dollar budgets. Therefore, the budgets must be managed carefully. Player transfers are one of the most important expenditure items of football teams. There are many football performance data providers. These data providers provide statistics that can help coaches to analyze and prepare matches. At the same time, these data providers allow scouts to identify promising future football player profiles. Although these platforms provide large and detailed data about a football player, they do not support market values since there is no correlation between the performance parameters and also market values. However, the results obtained from the study are promising in terms of guiding teams in the transfer planning processes.

In conclusion, use of ANNs has become one of the most important modeling methods in all areas of life. Although the application of ANN is very limited in sports science, it is one of the suitable science disciplines where there is a lot of statistical data. The present neural network may help to stop media criticism as it estimates the market value of players based on fair performance parameters. Moreover, talent selection can also be studied by using the proposed ANN in the study. Further studies are strongly suggested, not only for football but also for other sports disciplines.

Acknowledgments

The authors thank OPTA for providing data of the midfield players in the Turkish Super League (2018/2019) for this paper. Ms.Raines Walker is acknowledged for English proofreading.

ORCID

Tugbay Inan http://orcid.org/0000-0002-3573-8086

References

Akşar, T. 2005. *Endüstriyel Futbol*. 1st ed. İstanbul: Literatür Yayınları.

Akşar, T. 2013. *Krizdeki Futbol*. 1st ed. İstanbul: Literatür Yayınları.

Akşar, T. 2019. *Türk Futboluna Finansal Reçete*. Accessed December 30, 2020. http://www.futbolekonomi.com/index.php/haberler-makaleler/genel/122-tugrul-aksar/4494-futbol-ve-finansman-yapilandirma-futbol.html

Akşar, T. 2020. *Dört Kulübün Gelirleri, Zararlarını Karşılamaya Yetmiyor*. Accessed December 30, 2020. http://www.futbolekonomi.com/index.php/haberler-makaleler/mali/110-tugrul-aksar/4114-2017-04-26-14-35-22.html

Akşar, T., and K. Merih. 2006. *Futbol Ekonomisi*. 1st ed. İstanbul: Literatür Yayınları.

Aktifbank. 2019. *Ekolig*. Accessed December 30, 2020. https://www.aktifbank.com.tr/tr/aktifbank-hakkinda/BasinOdasi/bizden-haberler/Sayfalar/aktif-bank-2019-ekolig-raporu-aciklandi.aspx

Armatas, V., J. González-Ródenas, E. Varamenti, and D. Gkoumas. 2020. Effect of team ranking on defenders' offensive performance in greek top football league. *Journal of Physical Education and Sport* 20 (5). doi: 10.7752/jpes.2020.05361.

Barron, D., G. Ball, M. Robins, and C. Sunderland. 2018. Artificial neural networks and player recruitment in professional soccer. *PLoS ONE* 13 (10):e0205818. doi:10.1371/journal.pone.0205818.

Barron, D., G. Ball, M. Robins, and C. Sunderland. 2020. Identifying playing talent in professional football using artificial neural networks. *Journal of Sports Sciences* 38 (11–12):1211–20. doi:10.1080/02640414.2019.1708036.

Barros, C. P., and S. Leach. 2006. Performance evaluation of the English Premier Football League with data envelopment analysis. *Applied Economics* 38 (12):1449–58. doi:10.1080/00036840500396574.

Bazrafkan, S., and P. M. Corcoran. 2018. Pushing the AI envelope: Merging deep networks to accelerate edge artificial intelligence in consumer electronics devices and systems. *IEEE Consumer Electronics Magazine* 7 (2):55–61. doi:10.1109/MCE.2017.2775245.

Benchmark, K. F. 2020. *An analysis of the impact of the COVID-19 crisis on football players' market values*. Accessed December 30, 2020. https://footballbenchmark.com/library/player_value_not_immune_to_pandemic

Bransen, L., J. Van Haaren, and M. Van De Velden. 2019. Measuring soccer players' contributions to chance creation by valuing their passes. *Journal of Quantitative Analysis in Sports* 15 (2):97–116. doi:10.1515/jqas-2018-0020.

Bunker, R. P., and F. Thabtah. 2019. A machine learning framework for sport result prediction. *Applied Computing and Informatics* 15 (1):27–33. doi:10.1016/j.aci.2017.09.005.

Carmichael, F., D. Forrest, and R. Simmons. 1999. The labour market in association football: Who gets transferred and for how much? *Bulletin of Economic Research* 51 (2):125–50. doi:10.1111/1467-8586.00075.

Carrieri, V., F. Principe, and M. Raitano. 2018. What makes you "super-rich"? New evidence from an analysis of football players' wages. *Oxford Economic Papers* 70 (4):950–73. doi:10.1093/oep/gpy025.

Caruso, R., M. Di Domizio, and D. Rossignoli. 2017. Aggregate wages of players and performance in Italian Serie A. *Economia Politica* 34 (3):515–31. doi:10.1007/s40888-017-0062-6.

Cavas, L., N. Donut, and N. Mert. 2016. Artificial neural network modeling of diuron and irgarol-based HPLC data and their levels from the seawaters in Izmir, Turkey. *Journal of Liquid Chromatography and Related Technologies* 39 (2):87–95. doi:10.1080/10826076.2015.1128442.

Christensen, M. K. 2009. "An eye for talent": Talent identification and the "practical sense" of top-level soccer coaches. *Sociology of Sport Journal* 26 (3):365–82. doi:10.1123/ssj.26.3.365.

Ćwiklinski, B., A. Giełczyk, and M. Choraś. 2021. Who will score? A machine learning approach to supporting football team building and transfers. *Entropy* 23 (1):90. doi:10.3390/e23010090.

De Luca, G., and M. Gallo. 2020. The use of artificial neural networks for extending road traffic monitoring data spatially: An application to the neighbourhoods of Benevento. *Transportation Research Procedia* 45. doi:10.1016/j.trpro.2020.03.047.

DeloitteGroup, S. B. 2020. *Football Money League 2020*. Accessed December 30, 2020. https://www2.deloitte.com/bg/en/pages/finance/articles/football-money-league-2020.html

Deprez, D., J. Fransen, J. Boone, M. Lenoir, R. Philippaerts, and R. Vaeyens. 2015. Characteristics of high-level youth soccer players: Variation by playing position. *Journal of Sports Sciences* 33 (3):243–54. doi:10.1080/02640414.2014.934707.

Dutt-Mazumder, A., C. Button, A. Robins, and R. Bartlett. 2011. Neural network modelling and dynamical system theory. *Sports Medicine* 41 (12):1003–17. doi:10.2165/11593950-000000000-00000.

Felipe, J. L., A. Fernandez-Luna, P. Burillo, L. E. de la Riva, J. Sanchez-Sanchez, and J. Garcia-Unanue. 2020. Money talks: Team variables and player positions that most influence the market value of professional male footballers in Europe. *Sustainability (Switzerland)* 12 (9). doi: 10.3390/su12093709.

Frick, B. 2011. Performance, salaries, and contract length: Empirical evidence from German soccer. *International Journal of Sport Finance*, 6 (2): 87-118.

Gai, Y., A. Volossovitch, C. Lago, and M. Á. Gómez. 2019. Technical and tactical performance differences according to player's nationality and playing position in the Chinese Football Super League. *International Journal of Performance Analysis in Sport* 19 (4). doi: 10.1080/24748668.2019.1644804.

Gerhards, J., and M. Mutz. 2017. Who wins the championship? Market value and team composition as predictors of success in the top European football leagues. *European Societies* 19 (3):223–42. doi:10.1080/14616696.2016.1268704.

Gerhards, J., M. Mutz, and G. G. Wagner. 2014. Predictable winners. Market value, inequality, diversity, and routine as predictors of success in European soccer leagues. *Zeitschrift Für Soziologie* 43 (3). doi: 10.1515/zfsoz-2014-0305.

Guzman, A. L., and S. C. Lewis. 2020. Artificial intelligence and communication: A human–machine communication research agenda. *New Media and Society* 22 (1):70–86. doi:10.1177/1461444819858691.

He, M., R. Cachucho, and A. Knobbe 2015. Football player's performance and market value. *CEUR Workshop Proceedings, 1970, Porto - Portugal.*

Herm, S., H. M. Callsen-Bracker, and H. Kreis. 2014. When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community. *Sport Management Review* 17 (4):484–92. doi:10.1016/j.smr.2013.12.006.

Hughes, M., T. Caudrelier, N. James, A. Redwood-Brown, I. Donnelly, A. Kirkbride, and C. Duschesne. 2012. Moneyball and soccer - An analysis of the key performance indicators of elite male soccer players by position. *Journal of Human Sport and Exercise* 7 (SPECIALISSUE.2):402–12. doi:10.4100/jhse.2012.72.06.

Hughes, M. D., and R. M. Bartlett. 2002. The use of performance indicators in performance analysis. *Journal of Sports Sciences* 20 (10):739–54. doi:10.1080/026404102320675602.

Huijgen, B. C. H., M. T. Elferink-Gemser, A. Ali, and C. Visscher. 2013. Soccer skill development in talented players. *International Journal of Sports Medicine* 34 (8). doi: 10.1055/s-0032-1323781.

Kannekens, R., M. T. Elferink-Gemser, and C. Visscher. 2011. Positioning and deciding: Key factors for talent development in soccer. *Scandinavian Journal of Medicine and Science in Sports* 21 (6):846–52. doi:10.1111/j.1600-0838.2010.01104.x.

Konefał, M., P. Chmura, M. Zacharko, J. Chmura, A. Rokita, and M. Andrzejewski. 2018. Match outcome vs match status and frequency of selected technical activities of soccer players during UEFA Euro 2016. *International Journal of Performance Analysis in Sport* 18 (4):568–81. doi:10.1080/24748668.2018.1501991.

Kulkarni, S., N. Seneviratne, M. S. Baig, and A. H. A. Khan. 2020. Artificial Intelligence in Medicine: Where are we now? *Academic Radiology* 27 (1):62–70. doi:10.1016/j.acra.2019.10.001.

Lago-Peñas, C., M. Ã. Gómez-Ruano, A. L. Owen, and J. Sampaio. 2016. The effects of a player dismissal on competitive technical match performance. *International Journal of Performance Analysis in Sport* 16 (3):792–800. doi:10.1080/24748668.2016.11868928.

Lehmann, E. E., and G. G. Schulze. 2008. What does it take to be a star? – The role of performance and the media for German soccer players. *Applied Economics Quarterly* 54 (1):59–70. doi:10.3790/aeq.54.1.59.

Liu, H., W. Hopkins, M. A. Gómez, and J. S. Molinuevo. 2013. Inter-operator reliability of live football match statistics from OPTA Sportsdata. *International Journal of Performance Analysis in Sport* 13 (3):803–21. doi:10.1080/24748668.2013.11868690.

Liu, Y., Z. Wang, and D. Zhou. 2018. UKF-based remote state estimation for discrete artificial neural networks with communication bandwidth constraints. *Neural Networks* 108:393–98. doi:10.1016/j.neunet.2018.08.015.

Mallo, J., and D. Alexandre. 2012. Injury risk in professional football players with special reference to the playing position and training periodization. *Journal of Sports Medicine and Physical Fitness* 52 (6): 631-638.

Mamdani, M., and A. S. Slutsky. 2020. Artificial intelligence in intensive care medicine. *Intensive Care Medicine*. doi:10.1007/s00134-020-06203-2.

MATLAB. 2020. *Deep Learning*. Accessed December 30, 2020. https://www.mathworks.com/products/deep-learning.html

Miller, P. K., C. Cronin, and G. Baker. 2015. Nurture, nature and some very dubious social skills: An interpretative phenomenological analysis of talent identification practices in elite English youth soccer. *Qualitative Research in Sport, Exercise and Health* 7 (5):642–62. doi:10.1080/2159676X.2015.1012544.

Mirkov, D. M., M. Kukolj, D. Ugarkovic, V. J. Koprivica, and S. Jaric. 2010. Development of anthropometric and physical performance profiles of young elite male soccer players: Alongitudinal study. *Journal of Strength and Conditioning Research* 24 (10):2677–82. doi:10.1519/JSC.0b013e3181e27245.

Müller, O., A. Simons, and M. Weinmann. 2017. Beyond crowd judgments: Data-driven estimation of market value in association football. *European Journal of Operational Research* 263 (2):611–24. doi:10.1016/j.ejor.2017.05.005.

Oberstone, J. 2011. Comparing team performance of the English Premier League, Serie A, and la Liga for the 2008-2009 season. *Journal of Quantitative Analysis in Sports* 7 (1). doi: 10.2202/1559-0410.1280.

Olthof, S. B. H., W. G. P. Frencken, and K. A. P. M. Lemmink. 2015. The older, the wider: On-field tactical behavior of elite-standard youth soccer players in small-sided games. *Human Movement Science* 41:92–102. doi:10.1016/j.humov.2015.02.004.

Peeters, T. 2018. Testing the Wisdom of Crowds in the field: Transfermarkt valuations and international soccer results. *International Journal of Forecasting* 34 (1):17–29. doi:10.1016/j.ijforecast.2017.08.002.

Pehlivan, N., Y. Ünal, and C. Kahraman. 2019. Player selection for a national football team using fuzzy AHP and fuzzy TOPSIS. *Journal of Multiple-Value Logic and Soft Computing* 32 (5–6):369–405.

Platform, P. D. 2020. *BIST Companies*. Accessed December 30, 2020. https://www.kap.org.tr/tr/bist-sirketler

Rahman, M. A. 2020. A deep learning framework for football match prediction. *SN Applied Sciences* 2 (2). doi: 10.1007/s42452-019-1821-5.

Ribeiro, A. S., and F. Lima. 2012. Portuguese football league efficiency and players‘ wages. *Applied Economics Letters* 19 (6):599–602. doi:10.1080/13504851.2011.591719.

Ruijg, J., and H. Van Ophem. 2015. Determinants of football transfers. *Applied Economics Letters* 22 (1):12–19. doi:10.1080/13504851.2014.892192.

Şahin, M., and R. Erol 2018. Prediction of attendance demand in european Football games: Comparison of ANFIS, fuzzy logic, and ANN. *Computational Intelligence and Neuroscience Volume 2018,* Article ID 5714872, 14, https://doi.org/10.1155/2018/5714872 .

Serna Rodríguez, M., A. Ramírez Hassan, and A. Coad. 2019. Uncovering value drivers of high performance soccer players. *Journal of Sports Economics* 20 (6):819–49. doi:10.1177/1527002518808344.

Singh, P., and P. S. Lamba. 2019. Influence of crowdsourcing, popularity and previous year statistics in market value estimation of football players. *Journal of Discrete Mathematical Sciences and Cryptography* 22 (2):113–26. doi:10.1080/09720529.2019.1576333.

StatsPerform. 2020. *Detailed data to power your content and coverage*. Accessed December 30, 2020. https://www.statsperform.com/media-and-tech/digital-media/content-for-coverage/

Stübinger, J., B. Mangold, and J. Knoll. 2020. Machine learning in football betting: Prediction of match results based on player characteristics. *Applied Sciences (Switzerland)* 10 (1). doi: 10.3390/app10010046.

Towlson, C., S. Cobley, A. W. Midgley, A. Garrett, G. Parkin, and R. Lovell. 2017. Relative age, maturation and physical biases on position allocation in elite-youth soccer. *International Journal of Sports Medicine* 38 (3). doi: 10.1055/s-0042-119029.

Transfermarkt. 2020. *Transfers and rumours*. Accessed December 30, 2020. https://www.transfermarkt.com/

Tunaru, R. S., and H. P. Viney. 2010. Valuations of soccer players from statistical performance data. *Journal of Quantitative Analysis in Sports* 6 (2). doi: 10.2202/1559-0410.1238.

Union of European Football Associations UEFA. 2015. *Financial fair play: All you need to know*. Accessed December 30, 2020. https://www.uefa.com/news/0253-0d7f34cc6783-5ebf120a4764-1000–financial-fair-play-all-you-need-to-know/?referrer=%2Fcommunity%2Fnews%2Fnewsid%3D2064391

Waldron, M., and P. Worsfold. 2010. Differences in the game specific skills of elite and sub-elite youth football players: Implications for talent identification. *International Journal of Performance Analysis in Sport* 10 (1):9–24. doi:10.1080/24748668.2010.11868497.

Wiemeyer, J. 2003. Who should play in which position in soccer? Empirical evidence and unconventional modelling. *International Journal of Performance Analysis in Sport* 3 (1):1–18. doi:10.1080/24748668.2003.11868269.

Winemiller, S., A. Love, and J. Stamm. 2020. Recruiting reporters' perceptions of ethical issues. *Communication and Sport* 216747952094668. doi:10.1177/2167479520946683.

Yaldo, L., and L. Shamir. 2017. Computational estimation of football player wages. *International Journal of Computer Science in Sport* 16 (1):18–38. doi:10.1515/ijcss-2017-0002.