

# Evaluation of creating scoring opportunities for teammates in soccer via trajectory prediction

Masakiyo Teranishi<sup>1</sup>, Kazushi Tsutsui<sup>1</sup>, Kazuya Takeda<sup>1</sup>, and Keisuke Fujii<sup>1,2,3</sup>

<sup>1</sup> Graduate School of Informatics, Nagoya University, Nagoya, Japan.

<sup>2</sup> Center for Advanced Intelligence Project, RIKEN, Fukuoka, Japan.

<sup>3</sup> PRESTO, Japan Science and Technology Agency, Saitama, Japan.

**Abstract.** Evaluating the individual movements for teammates in soccer players is crucial for assessing teamwork, scouting, and fan engagement. It has been said that players in a 90-min game do not have the ball for about 87 minutes on average. However, it has remained difficult to evaluate an attacking player without receiving the ball, and to reveal how movement contributes to the creation of scoring opportunities for teammates. In this paper, we evaluate players who create off-ball scoring opportunities by comparing actual movements with the reference movements generated via trajectory prediction. First, we predict the trajectories of players using a graph variational recurrent neural network that can accurately model the relationship between players and predict the long-term trajectory. Next, based on the difference in the modified off-ball evaluation index between the actual and the predicted trajectory as a reference, we evaluate how the actual movement contributes to scoring opportunity compared to the predicted movement. For verification, we examined the relationship with the annual salary, the goals, and the rating in the game by experts for all games of a team in a professional soccer league in a year. The results show that the annual salary and the proposed indicator correlated significantly, which could not be explained by the existing indicators and goals. Our results suggest the effectiveness of the proposed method as an indicator for a player without the ball to create a scoring chance for teammates<sup>4</sup>.

**Keywords:** multi-agent · deep learning · trajectory · sports · football

## 1 Introduction

Assessing the movements of individual players for teammates in team sports is an important aspect of building teamwork, assessment of players' salaries, player recruitment, and scouting. In soccer, most analytics has focused on the outcomes of discrete events near the ball (on-ball) [30, 10, 35, 36, 3, 9, 29, 28] whereas much of the importance in player movements exist in the events without the ball (off-ball). For example, it is said that players in a 90-min game do not have the ball for about 87 minutes on average [14]. However, continuous off-ball movements are usually not discretized and difficult to understand except for core fans, experienced players, and coaches. Also for the media and building fan engagement, quantitative evaluation of off-ball players is an issue in demand, which provides

---

<sup>4</sup> Appendix: <https://arxiv.org/abs/2206.01899>, \*email: [fujii@i.nagoya-u.ac.jp](mailto:fujii@i.nagoya-u.ac.jp)

a common reference for beginners and experts in the sport e.g., when arguing a play of a favorite player.

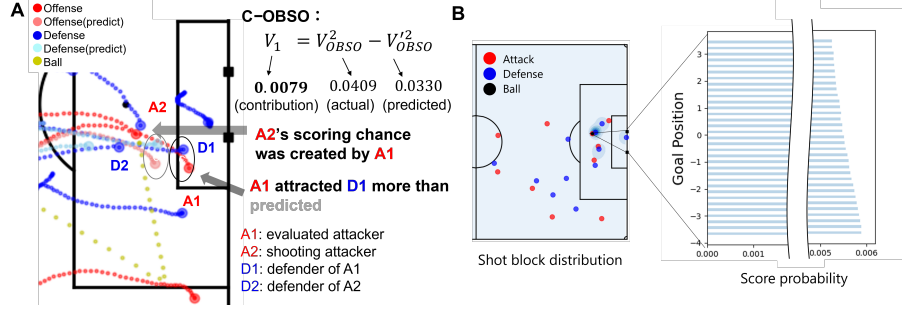
Regarding the off-ball player evaluation methods, the positioning itself related to the goal was evaluated from the location data of all players and the ball. For example, the method called off-ball scoring opportunity (OBSO) to evaluate the player who receives the ball [41] and the method to evaluate the movement to create space [14] have been proposed. However, it has been still difficult to clarify how movements contribute to the creation of scoring opportunities for teammates, to evaluate other attacking players who do not receive it (e.g., a player moving tactically for teammates), and often to evaluate a score prediction to reflect the position of the multiple defenders.

In this paper, we propose a new evaluation indicator, Creating Off-Ball Scoring Opportunity (C-OBSO in Fig. 1A), aiming for evaluating players who create scoring opportunities when the attacking player is without the ball. The overview of our method is as follows. (i) First, we modify the score model in the framework of OBSO [41] with the potential score model that reflects the positions of multiple defenders with a mixed Gaussian distribution (Fig. 1B). (ii) Next, we accurately model the relationship between athletes and perform long-term trajectory predictions (Fig. 1A) using the graph variational recurrent neural network (GVRNN) [46]. (iii) Finally, based on the difference in the modified off-ball evaluation index between the actual and the predicted trajectory (Fig. 1A), we evaluate how the actual movement contributes to scoring opportunity relative to the predicted movement as a reference.

In summary, our main contributions were as follows. (1) We proposed an evaluation method of how movements contributed to the creation of scoring opportunities compared to the predicted movements of off-ball players in team sports attacks. (2) As a score predictor, we proposed a potential score model that considers the positions of multiple defenders in a mixed Gaussian distribution. (3) In the experiment, we analyzed the relationship between the annual salary, the goals, and the game rating by experts, and show the effectiveness of the proposed method as an indicator for an off-ball player to create scoring opportunities for teammates. Our approach can evaluate continuous movements of players by comparing with the reference (here predicted) movements, which are difficult to be discretized or labeled but crucial for teamwork, scouting, and fan engagement. The structure of this paper is as follows. First, we overview the related works in Section 4 and present experimental results in Section 3. Next, we describe our methods in Section 2 and conclude this paper in Section 5.

## 2 Proposed framework

Here, we propose C-OBSO based on the motivation to evaluate players who create off-ball scoring opportunities for teammates. To this end, in Section 2.1, we first propose a potential score model that reflects the positions of multiple defenders with a mixed Gaussian distribution. Next, in Section 2.2, we predict multi-agent trajectory using GVRNN [46] and evaluate the difference between the actual value of the modified OBSO and the predicted value (as a reference) to evaluate how the movement contributed to the creation of scoring opportunities.



**Fig. 1.** Our C-OBSO example and potential score model. (A) Example of C-OBSO computation. A1 is the player to be finally evaluated, A2 is the shooting player, D1 and D2 are the defender of A1 and A2, respectively.  $V_1$  is the C-OBSO value of A1,  $V_{OBSO}^2$  is the actual OBSO, and  $V_{OBSO}'^2$  is the reference OBSO value of A2 using the predicted trajectory. (B) Potential score model. Left: shot-blocking distribution formed by defenders. Right: shot probability corresponding to each shot vector. The vertical axis is the goal position (m), and the horizontal axis is the shot probability corresponding to each shot vector.

## 2.1 Potential score model in modified OBSO

First, we describe the base model of our evaluation method called OBSO [41] and then propose the potential score model. OBSO evaluates off-ball players by computing the following joint probability

$$P(G|D) = \sum_{r \in R \times R} P(S_r \cap C_r \cap T_r | D) \quad (1)$$

$$= \sum_r P(S_r | C_r, T_r, D) P(C_r | T_r, D) P(T_r | D), \quad (2)$$

where  $D$  is the instantaneous state of the game (e.g., player positions and velocities).  $P(S_r)$  is the probability of scoring from an arbitrary point  $r \in R \times R$  on the pitch, assuming the next on-ball event occurs there.  $P(C_r)$  is the probability that the passing team will control a ball at point  $r$ .  $P(T_r)$  is the probability that the next on-ball event occurs at point  $r$ . Here, for simplicity, we can assume that  $P(S_r|D)$ ,  $P(T_r|D)$ ,  $P(C_r|D)$  are independent if the parameter  $\alpha = 0$  in the original work implementation (Eq. (6) in [41]). Then, the joint probability can be decomposed into a series of conditional probabilities as follows:

$$P(G|D) = \sum_{r \in R \times R} P(S_r | D) P(C_r | D) P(T_r | D). \quad (3)$$

$P(C_r|D)$  is the probability that the attacking team will control the ball at point  $r$  assuming the next on-ball event occurs there, which is called the potential pitch control field (PPCF).  $P(T_r|D)$  is defined as a two-dimensional Gaussian distribution with the current ball coordinates as the mean.  $P(S_r|D)$  is simply calculated as a value that decreases with the distance from the goal. We used the grid data and computed  $P(C_r|D)$  and  $P(T_r|D)$  based on the code at <https://github.com/Friends-of-Tracking-Data-FoTD/LaurieOnTracking>.

In the original OBSO [41], the scoring probability was calculated as the output  $P(S_r|D)$  of the score model as a function of the distance from the goal. However, the scoring probability may depend on the angle to the goal and the

defensive position of the opponent. Therefore, in this paper, we propose a score model that reflects the angle to the goal and the position of multiple defenders. Here, we consider the shot-blocking distribution of the defenders who can block shots in the field, and propose a potential model where the scoring probability decreases when defenders exist. The basic idea shown in Fig. 1B is to calculate the scoring probability from the angle to the goal at which the shot tends to be scored, considering the mixed distribution of the positions of multiple defenders. The proposed scoring probability  $P(S_r^p|D)$  at a certain point  $r$  is calculated as the sum of the shot value  $V_{shot}$  as follows:

$$P(S_r^p|D) = \sum_{i=1}^n V_{shot}(\vec{s}_i), \quad (4)$$

$$V_{shot}(\vec{s}) = C(c - V_{block}), \quad (5)$$

where  $n$  is determined by the angle from the shooting position to the goal, and  $\vec{s}$  is a shot vector per degree ( $n$  is larger when the shot from the center and smaller from the side). The shot value  $V_{shot}$  is calculated by subtracting the shot block value  $V_{block}$  from a certain constant  $c$  ( $c, C$  are parameters determined from data to be adjusted so that  $V_{shot} \geq 0$  and  $P(S_r^p) \in [0, 1]$ ). Let  $V_{block}$  be the sum of the shot block distribution values along the shot vector  $\vec{s}$ . The shot blocking distribution is the sum of the normal distributions (variance  $\sigma^2 = 0.5 + l_d$ ) assigned to each defender on the goal side of the shooting position (shot blockable players using legs), where  $l_d$  is the distance between the shooting position and the defender. We consider that goalkeepers have a shot blocking distribution with twice the value of normal defenders because of a higher shot-blocking ability. Here, we assume that the block distribution is not changed with the distance from the ball. A defender near to the ball may affect the ball, but far players use the flight time of the ball for their movement. This formulation is left for future work.

## 2.2 C-OBSO with trajectory prediction

Here, we describe the base model of our trajectory prediction method called GVRNN [46] and then describe our C-OBSO framework. Our contribution here is to evaluate how the actual “off-ball” movement contributes to scoring opportunity compared to the predicted movement (or trajectory) as a reference. In our method, we use GVRNN [46], which is a VRNN [8] combined with a graph neural network (GNN [22]). In GVRNN, the graph encoder-decoder network models the relationship between players as a graph, which is one of the best performing models for predicting player trajectories in team sports [46]. This is a probabilistic model which can sample multiple possible trajectories.

Based on the trajectory prediction, we propose an evaluation index C-OBSO of players who create scoring opportunities for teammates. The basic idea is to evaluate an off-ball player from the difference in the modified OBSO values between the predicted and actual movements of the players. The C-OBSO value of a player  $i$  without the ball can be expressed as follows.

$$V_i = V_{OBSO}^k - V_{OBSO}'^k, \quad (6)$$

where the player  $k$  is the ball carrier who performs a final action (e.g., shot),  $V_{OBSO}^k$  is the modified OBSO in the actual game situation, and  $V_{OBSO}^{k*}$  is the modified OBSO based on the predicted trajectory as a reference. For example, in Fig. 1A, C-OBSO is positive and the player to be evaluated (A1) contributes more to the shooter (A2) than the referenced (predicted) player. Specifically, A1 has created a more advantageous situation for A2 by attracting D1 more than expected. C-OBSO can evaluate a player in such situations with an interpretable value (i.e., the increase in scoring probability). Theoretically, if perfectly predicted, C-OBSO is always zero, but actually, if we apply this to a test data, the perfect prediction is impossible. In other words, we assume the imperfect trajectory prediction in this framework.

### 3 Experiments

In this section, we validate the proposed method of the potential score model, the trajectory prediction model (GVRNN), and the C-OBSO itself. For our implementation, the code is available at <https://github.com/keisuke198619/C-OBSO>.

#### 3.1 Dataset

In this study, we used all 34 games data of Yokohama F Marinos in the Meiji J1 League 2019 season to perform specific player-level evaluations in limited data. Note that the tracking data for all players and timesteps were not publicly shared in such amounts. The dataset includes event data (i.e., labels of actions, e.g., passing and shooting, recorded at 30 Hz and the simultaneous xy coordinates of the ball) and tracking data (i.e., xy coordinates of all players recorded at 25 Hz) provided by Data Stadium Inc. The company was licensed to acquire this data and sell it to third parties, and it was guaranteed that the use of the data would not infringe on any rights of the players or teams. For annual salaries, we used the salaries of the same team (Yokohama) in 2019 [40] because they were different valuation criteria for different teams and the transfer of the players took place during the season. The goals for each player in each match were collected from [20]. The rating by experts in each match [39] was also used for verification, which was scored in 0.5 point increments with a maximum of 10 points.

#### 3.2 Data processing for verification

We used the attacking data of Yokohama F Marinos for the test and those of the opponent teams for training the model or parameter fitting. Again, since the data was limited in this study, we split the data in such a way, and if we have more data, we can analyze all teams with the training data with the same team. Here we describe the processing of the potential score model, the trajectory prediction model, the C-OBSO, and their statistical analyses.

**Potential score model.** To validate the potential score model, the opponent’s shots (345 shots, 34 goals) were used for fitting the parameters  $c$  and  $C$ , and Yokohama F Marinos’ shots (494 shots, 59 goals) were used for verification. The parameters  $c, C$  of the potential score model were determined to be  $c = 1.1$ ,  $C = 1/150$  using the data of the opponents. The potential score model was verified by

the root mean square error (RMSE) between the actual score and the calculated scoring probability. We compared the RMSE with that of a simple score model as a function of distance from the goal for implementing the original OBSO [41] (see also Section 2.1). Although there have been more holistic score models such as [15, 1], to fairly compare with our potential model as a component of the modified OBSO, we consider the simple score model as an appropriate baseline.

**Trajectory prediction model.** For the test data of trajectory prediction and C-OBSO, we used 412 shot scenes of Yokohama F Marinos (we selected the sequences of consecutive events and excluded too short events such as a free kick). The trajectory prediction model was trained using the opponents’ data to generate “league average” trajectories. The tracking data were down-sampled to 10 Hz (after prediction, up-sampled at the original 25 Hz) based on [17]. To verify the accuracy of the long-term trajectory prediction, we set various time lengths (6, 8, 10, and 12 s) using mean trajectories in 10 samples. We divided into the opponent data for batch training (6 s: 94208 sequences, 8 s: 49152 sequences, 10 s: 33536 sequences, 12 s: 24320 sequences) and the validation (6 s: 10477 sequences, 8 s: 5479 sequences, 10 s: 3730 sequences, 12 s: 2721 sequences). Note that the end of all sequences was the moment of a shot. The input feature has 92 dimensions (the xy coordinates and the velocity of 22 players and the ball). During training, the model was trained based on the one-step prediction error of all combinations of the two attackers who invaded the attacking third. We simultaneously predicted the three players: one of the off-ball attackers and the defenders closest to each attacker. Note that we only consider the three players’ interactions and ignore others’ interactions, because the prediction error will increase if the numbers increase, and the increase of the predicted players is left for future work.

For the test data of the 412 sequences, the three relevant players and the attacker from the same criterion were predicted. At the inference, using 2 s sequences as burn-in period, we predicted the sequences for the subsequent time lengths (i.e., 4, 6, 8, and 10 s) by updating the estimated position and velocity (i.e., performed long-term prediction). For the training of the proposed and baseline models, we used the Adam optimizer [21] with a learning rate of 0.001 and 10 training epochs. We set the batchsize to 256. For the performance metrics, we used the endpoint error (mean absolute error: MAE) from the actual trajectory.

**C-OBSO.** To compute C-OBSO, predicted trajectories with 4 s (total 6 s) were used. This is because a longer prediction time will result in a larger prediction error, while a shorter prediction time will not make a difference in the evaluation of C-OBSO. Although the negative values of C-OBSO are also calculated by comparison with the reference, the negative values were calculated as 0, assuming that they may not have a negative effect on the behavioral players. This is because there were many situations with negative values in which the shooter’s defender did not take an appropriate defensive position in the predicted trajectory.

**Statistical analysis.** For the verification of C-OBSO, we examined the relationship with the annual salary, the goals, and the expert’s rating. Note that

there is no ground truth available for the verification. We also compared them with the existing OBSO [41]. Since some of the data often did not follow normal distributions, we used Spearman’s rank correlation coefficient  $\rho$  for these relationships. Regarding the RMSE in the potential score model and MAE in the trajectory prediction, for the same reason, we used nonparametric statistical tests to compare with the baselines. Regarding the potential score model, we used the Wilcoxon rank sum test. For all statistical calculations,  $p < 0.05$  was considered as significant.

### 3.3 Our model verification

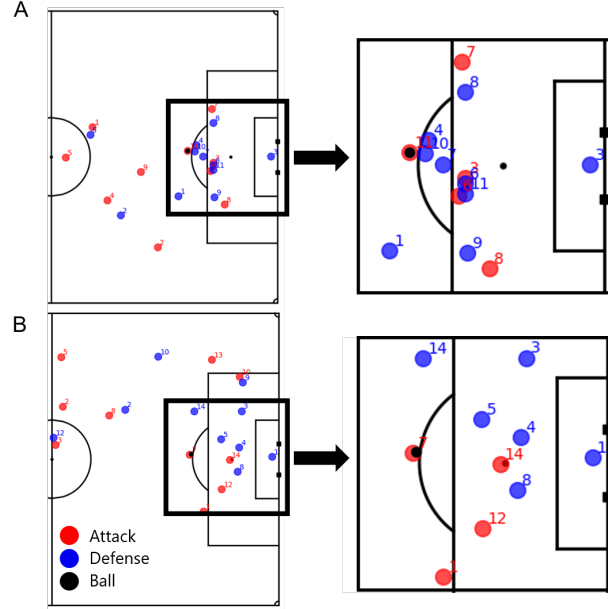
First, we validated the potential score model needed to calculate the C-OBSO. The RMSE with the actual scores was  $0.324 \pm 0.014$  for the conventional score model [41] without considering the defenders and goal angles, and  $0.309 \pm 0.0014$  for the potential score model ( $p < 10^{-10}$ ). This result suggests that the proposed method models the scores more accurately.

Figure 2 shows an example of the two methods in two actual situations where a shot is attempted from a similar distance. In the existing method, the probabilities were the same (both 0.1237) because the shots were taken from almost the same distance. The proposed method had a lower scoring probability with more defenders (upper: 0.0489, lower: 0.1202). We indicate that the proposed method reflects the position of multiple defenders and can model the score accurately.

Next, we show the results of the trajectory prediction model for computing C-OBSO. Endpoint errors (MAE and standard error, [m]) in GVRNN were  $0.608 \pm 0.014$ ,  $0.867 \pm 0.022$ ,  $1.701 \pm 0.045$ ,  $1.606 \pm 0.042$  in 4, 6, 8, 10 s prediction. In GVRNN, longer predictions show larger prediction errors except for the difference between 8 s and 10 s. Since the 4 s prediction of GVRNN achieved a low the MAE of less than 0.7 m, the GVRNN trajectory prediction of 4 s was used in the next C-OBSO.

### 3.4 C-OBSO results

Verification of C-OBSO is challenging because of no ground truth values or player ratings. Therefore, we analyzed the relationship with the annual salary, the goals, and the game rating by experts, whereas we admit that these variables include various confounding factors. The relationships between the average C-OBSO and OBSO values of each player of Yokohama F Marinos in 2019 and the annual salary of each player in 2019 are shown in Fig. 3 (note that the tracking data for all players and timesteps were not publicly shared). Here we analyzed 15 players with more than 10 sequences under evaluation. As a result, there was a significant positive correlation between annual salary and C-OBSO ( $\rho = 0.45, p = 0.046$ ). In addition, the two players with the higher evaluation values but lower salaries (in red in Fig. 3A) were highly evaluated players, who won the individual awards (the most valuable player and valuable player award). In fact, their annual salary for the following year (2020) was also increased (valuable player: increased from 11 million yen to 40 million yen; the most valuable player: increased from 20 million yen to 60 million yen).

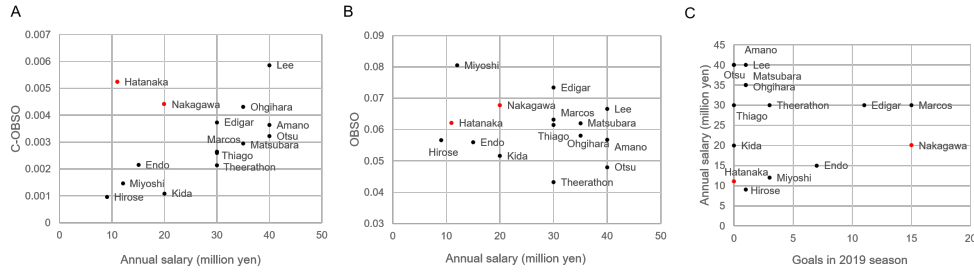


**Fig. 2.** Comparison between the score model of conventional and the proposed potential model. The scoring probability of our model is lower (A) when the defenders are crowded than (B), whereas that of the conventional score model in (A) was the same as (B).

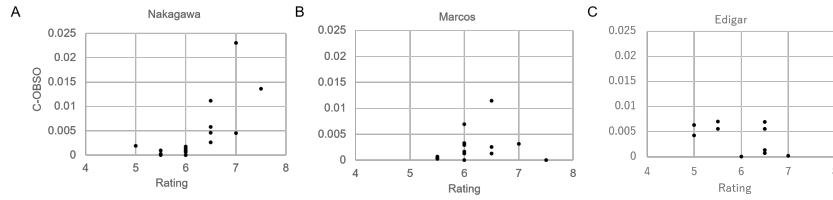
We found that these tendencies were similar to the C-OBSO and OBSO without the potential score model. On the other hand, there was no significant correlation ( $\rho = -0.28, p = 0.154$ ) for OBSO, which evaluates a player's own scoring opportunities (Fig. 3B). We also examined the relationship between annual salary and goals (Fig. 3C), and found no significant correlation ( $\rho = -0.23, p = 0.208$ ). Therefore, there was no relationship between annual salary and goals. There were many players with zero goals, and it is difficult to evaluate them only with the goals.

Next, in order to examine the relationship with player performance in more detail, we show the relationship between C-OBSO and the rating by experts of the top three scorers (Nakagawa with 15 goals, Marcos with 15 goals, and Edigar with 11 goals in this season) in Fig. 4. We analyzed the games in which there were two or more C-OBSO evaluations using the average of C-OBSO values on each game (17 games for Nakagawa, 14 games for Marcos, and 10 games for Edigar). A strong positive correlation was found only for Nakagawa ( $\rho = 0.75, p = 0.0003$ ) but not for Marcos ( $\rho = 0.27, p = 0.174$ ) and Edigar ( $\rho = -0.37, p = 0.145$ ). Similarly, there were no significant correlations between them for all players ( $\rho s < 0.190, p s > 0.05$ ). In addition to the number of scoring opportunities for teammates (17 times), the results found that Nakagawa would be subjectively and quantitatively an outstanding player.



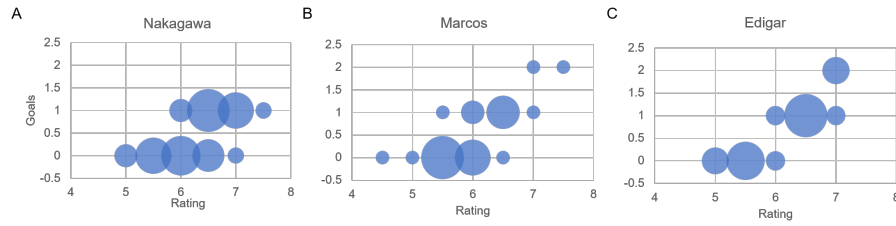


**Fig. 3.** Relationship between indicators, goal, and annual salary in a team. (A) Relationship between C-OBSO and the salary. (B) Relationship between OBSO [41] and the salary. (C) Relationship between each player’s goals and annual salary. Red players received individual awards (Hatanaka: valuable player Award, Nakagawa: the most valuable player award).



**Fig. 4.** Relationship between C-OBSO and the rating by experts of the top three scorers (A: Nakagawa, B: Marcos, C: Edigar) for each game.

For reference, we also show the relationship between the goals of the top three scorers and the ratings by experts in Fig. 5. We analyzed the games in which each player played (33 games for Nakagawa, 33 for Marcos, and 16 for Edigar). For each player, there were strong correlations between the goals and the rating (Nakagawa  $\rho = 0.63, p = 4.33 \times 10^{-5}$ , Marcos  $\rho = 0.71, p = 1.98 \times 10^{-6}$ , Edigar  $\rho = 0.91, p = 4.40 \times 10^{-7}$ ). We found that the rating of each game depends on a rare event (i.e., goals). Similarly, there were significant correlations between them for all players ( $\rho_s > 0.516, p_s < 0.018$ ). Recall that there was a stronger correlation between C-OBSO and Nakagawa’s rating than for the other two players. Nakagawa also had higher average ratings than the other players (6.26 for Nakagawa, 5.97 for Marcos, and 6.09 for Edigar), and he was the player who won the most valuable player award. The game rating by experts would depend on the goals, but it may also evaluate the creation of scoring opportunities only for Nakagawa. From these results, we speculate that Nakagawa was highly evaluated not only for his scoring but also for his contribution to other attacking players. Our method can also evaluate players difficult to be evaluated by conventional indicators, which is crucial for assessing teamwork and player salary, player recruitment, and scouting.



**Fig. 5.** Relationship between the goals and the rating by experts of the top three scorers (A: Nakagawa, B: Marcos, C: Edigar) for each game. The size of the circle represents the frequency because there are many combinations of the goals and the rating with the same value.

## 4 Related work

In the tactical behaviors of team sports, agents select an action that follows a policy (or strategy) in a state, receives a reward from the environment and others, and updates the state [16]. This is similar to a reinforcement learning framework (e.g., [2]). Due to the difficulty in modeling the entire framework from data for various reasons [44] (e.g., a sparse reward and difficulty in estimating intents), we can adopt two approaches: to estimate the related variables and functions from data (i.e., inverse approach) as a sub-problem, and to build a model (e.g., reinforcement learning model) to generate data in virtual space (i.e., forward approach, e.g., [23, 37]). Here, we focus on the former approach and introduce the research from the view of inverse approaches.

There have been many approaches to quantitatively evaluate the actions of attacking players about the scoring, such as based on the expected scores using tracking data [30, 10, 35, 36, 3], action data such as dribbling and passing [9, 13], and estimating state-action value function (Q-function) [45, 29, 28]. Some researchers have evaluated passes [32, 4, 12], and others evaluated actions to receive a ball by assigning a value to the location with the highest expected score [41, 27] and a rule-based manner [19]. In particular, Spearman [41] proposed an evaluation metric called OBSO to evaluate behavior based on location data and rule-based modeling. Defensive behaviors have also been evaluated based on data-driven [34, 43] and rule-based manners (e.g., [42]). However, these score evaluations do not often reflect the position of multiple defenders and goal angles in rule-based manner.

From the perspective of reinforcement learning, there have been many studies on inverse approaches. As for the study of state evaluation, there are several studies based on score expectation (e.g., [6, 7, 15]) and based on the value of space (e.g., [5, 14]). There is also research on estimating reward functions by inverse reinforcement learning [31, 33]. Researchers performed trajectory prediction sometimes in terms of the policy function estimation, as imitation learning [25, 24, 42, 17] and behavioral modeling [47, 46, 26, 18] to mimic (not optimize) the policy using neural network approaches. In this paper, we first propose a method to evaluate how the actual “off-ball” movement contributes to scoring

opportunity based on the difference between the state values generated from the actual and the reference policies.

## 5 Conclusion

In this paper, we evaluated players who create off-ball scoring opportunities by comparing actual movements with the reference movements generated by trajectory prediction. Our results suggest the effectiveness of the proposed method as an indicator for a player without the ball to create scoring opportunities for teammates. For future work, although the number of players to be evaluated was determined in the minimum setting, it is possible to evaluate the contribution to the scoring opportunities for teammates in a less limited way by predicting a larger number of players in both offense and defense. Furthermore, since our method evaluates off-ball players by comparing them with the referenced trajectory, the value becomes too small. Computing the evaluation value in a more interpretable way (e.g., in a score scale) would be future work. Finally, computing our indicators from broadcast videos (e.g., [11]) or other videos (e.g., top- or side-view [38]) would also be future work.

## Acknowledgments

This work was supported by JSPS KAKENHI (Grant Numbers 20H04075 and 21H05300) and JST Presto (Grant Number JPMJPR20CA).

## References

1. Anzer, G., Bauer, P.: A goal scoring probability model for shots based on synchronized positional and event data in football (soccer). *Frontiers in Sports and Active Living* **3**, 53 (2021)
2. Bernstein, D.S., Givan, R., Immerman, N., Zilberstein, S.: The complexity of decentralized control of markov decision processes. *Mathematics of Operations Research* **27**(4), 819–840 (2002)
3. Bransen, L., Van Haaren, J.: Measuring football players’ on-the-ball contributions from passes during games. In: *International workshop on machine learning and data mining for sports analytics*. pp. 3–15. Springer (2018)
4. Brooks, J., Kerr, M., Guttag, J.: Developing a data-driven player ranking in soccer using predictive model weights. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. pp. 49–55 (2016)
5. Cervone, D., Bornn, L., Goldsberry, K.: Nba court realty. In: *10th MIT Sloan Sports Analytics Conference* (2016)
6. Cervone, D., D’Amour, A., Bornn, L., Goldsberry, K.: Pointwise: Predicting points and valuing decisions in real time with nba optical tracking data. In: *Proceedings of the 8th MIT Sloan Sports Analytics Conference, Boston, MA, USA*. vol. 28, p. 3 (2014)
7. Cervone, D., D’Amour, A., Bornn, L., Goldsberry, K.: A multiresolution stochastic process model for predicting basketball possession outcomes. *Journal of the American Statistical Association* **111**(514), 585–599 (2016)

8. Chung, J., Kastner, K., Dinh, L., Goel, K., Courville, A.C., Bengio, Y.: A recurrent latent variable model for sequential data. *Advances in neural information processing systems* **28**, 2980–2988 (2015)
9. Decroos, T., Bransen, L., Van Haaren, J., Davis, J.: Actions speak louder than goals: Valuing player actions in soccer. In: *KDD*. pp. 1851–1861 (2019)
10. Decroos, T., Dzyuba, V., Van Haaren, J., Davis, J.: Predicting soccer highlights from spatio-temporal match event streams. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 31 (2017)
11. Deliege, A., Cioppa, A., Giancola, S., Seikavandi, M.J., Dueholm, J.V., Nasrollahi, K., Ghanem, B., Moeslund, T.B., Van Droogenbroeck, M.: Soccernet-v2: A dataset and benchmarks for holistic understanding of broadcast soccer videos. In: *7th International Workshop on Computer Vision in Sports (CVsports) at IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR’ 21)*. pp. 4508–4519 (2021)
12. Dick, U., Link, D., Brefeld, U.: Who can receive the pass? a computational model for quantifying availability in soccer. *Data Mining and Knowledge Discovery* **36**(3), 987–1014 (2022)
13. Dick, U., Tavakol, M., Brefeld, U.: Rating player actions in soccer. *Frontiers in Sports and Active Living* p. 174 (2021)
14. Fernández, J., Bornn, L.: Wide open spaces: A statistical technique for measuring space creation in professional soccer. In: *12th MIT Sloan Sports Analytics Conference* (2018)
15. Fernández, J., Bornn, L., Cervone, D.: Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer. In: *13th MIT Sloan Sports Analytics Conference* (2019)
16. Fujii, K.: Data-driven analysis for understanding team sports behaviors. *Journal of Robotics and Mechatronics* **33**(3), 505–514 (2021)
17. Fujii, K., Takeishi, N., Kawahara, Y., Takeda, K.: Policy learning with partial observation and mechanical constraints for multi-person modeling. *arXiv preprint arXiv:2007.03155* (2020)
18. Fujii, K., Takeuchi, K., Kuribayashi, A., Takeishi, N., Kawahara, Y., Takeda, K.: Estimating counterfactual treatment outcomes over time in complex multi-agent scenarios. *arXiv preprint arXiv:2206.01900* (2022)
19. Fujii, K., Yoshihara, Y., Matsumoto, Y., Tose, K., Takeuchi, H., Isobe, M., Mizuta, H., Maniwa, D., Okamura, T., Murai, T., et al.: Cognition and interpersonal coordination of patients with schizophrenia who have sports habits. *PLoS One* **15**(11), e0241863 (2020)
20. JLEAGUE: Jleague.jp 2019 data (2019), <https://www.jleague.jp/stats/2019/goal.html>
21. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: *International Conference on Learning Representations* (2015)
22. Kipf, T., Fetaya, E., Wang, K.C., Welling, M., Zemel, R.: Neural relational inference for interacting systems. In: *International Conference on Machine Learning*. pp. 2688–2697 (2018)
23. Kurach, K., Raichuk, A., Stańczyk, P., Zajac, M., Bachem, O., Espeholt, L., Riquelme, C., Vincent, D., Michalski, M., Bousquet, O., et al.: Google research football: A novel reinforcement learning environment. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. vol. 34, pp. 4501–4510 (2020)
24. Le, H.M., Carr, P., Yue, Y., Lucey, P.: Data-driven ghosting using deep imitation learning. In: *Proceedings of MIT Sloan Sports Analytics Conference* (2017)

25. Le, H.M., Yue, Y., Carr, P., Lucey, P.: Coordinated multi-agent imitation learning. In: Proceedings of the 34th International Conference on Machine Learning-Volume 70. pp. 1995–2003. JMLR. org (2017)
26. Li, L., Yao, J., Wenliang, L., He, T., Xiao, T., Yan, J., Wipf, D., Zhang, Z.: Grin: Generative relation and intention network for multi-agent trajectory prediction. *Advances in Neural Information Processing Systems* **34** (2021)
27. Link, D., Lang, S., Seidenschwarz, P.: Real time quantification of dangerousity in football using spatiotemporal tracking data. *PloS one* **11**(12), e0168768 (2016)
28. Liu, G., Luo, Y., Schulte, O., Kharrat, T.: Deep soccer analytics: learning an action-value function for evaluating soccer players. *Data Mining and Knowledge Discovery* **34**(5), 1531–1559 (2020)
29. Liu, G., Schulte, O.: Deep reinforcement learning in ice hockey for context-aware player evaluation. *arXiv preprint arXiv:1805.11088* (2018)
30. Lucey, P., Bialkowski, A., Monfort, M., Carr, P., Matthews, I.: quality vs quantity: Improved shot prediction in soccer using strategic features from spatiotemporal data. In: Proceedings of MIT Sloan Sports Analytics Conference. pp. 1–9 (2014)
31. Luo, Y., Schulte, O., Poupart, P.: Inverse reinforcement learning for team sports: Valuing actions and players. In: Bessiere, C. (ed.) Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20. pp. 3356–3363. International Joint Conferences on Artificial Intelligence Organization (7 2020)
32. Power, P., Ruiz, H., Wei, X., Lucey, P.: Not all passes are created equal: Objectively measuring the risk and reward of passes in soccer from tracking data. In: KDD. pp. 1605–1613 (2017)
33. Rahimian, P., Toka, L.: Inferring the strategy of offensive and defensive play in soccer with inverse reinforcement learning. In: Machine Learning and Data Mining for Sports Analytics (MLSA 2018) in ECML-PKDD Workshop (2020)
34. Robberechts, P.: Valuing the art of pressing. In: Proceedings of the StatsBomb Innovation In Football Conference. pp. 1–11. StatsBomb (2019)
35. Routley, K., Schulte, O.: A markov game model for valuing player actions in ice hockey. In: Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence. p. 782–791. UAI’15, AUAI Press, Arlington, Virginia, USA (2015)
36. Schulte, O., Khademi, M., Gholami, S., Zhao, Z., Javan, M., Desaulniers, P.: A markov game model for valuing actions, locations, and team performance in ice hockey. *Data Mining and Knowledge Discovery* **31**(6), 1735–1757 (2017)
37. Scott, A., Fujii, K., Onishi, M.: How does AI play football? An analysis of RL and real-world football strategies. In: 14th International Conference on Agents and Artificial Intelligence (ICAART’ 22). vol. 1, pp. 42–52 (2022)
38. Scott, A., Uchida, I., Onishi, M., Kameda, Y., Fukui, K., Fujii, K.: Soccertrack: A dataset and tracking algorithm for soccer with fish-eye and drone videos. In: 8th International Workshop on Computer Vision in Sports (CVsports) at IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR’ 22). pp. 3569–3579 (2022)
39. Soccer-digest: Soccer digest web j1 rating (2019), <https://www.soccerdigestweb.com>
40. Soccer-Money.net: Soccer-money.net (2019), <https://www.soccer-money.net>
41. Spearman, W.: Beyond expected goals. In: Proceedings of the 12th MIT sloan sports analytics conference. pp. 1–17 (2018)
42. Teranishi, M., Fujii, K., Takeda, K.: Trajectory prediction with imitation learning reflecting defensive evaluation in team sports. In: 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE). pp. 124–125. IEEE (2020)

43. Toda, K., Teranishi, M., Kushiro, K., Fujii, K.: Evaluation of soccer team defense based on prediction models of ball recovery and being attacked. *PLoS One* **17**(1), e0263051 (2022)
44. Van Roy, M., Robberechts, P., Yang, W.C., De Raedt, L., Davis, J.: Learning a markov model for evaluating soccer decision making. In: *Reinforcement Learning for Real Life (RL4RealLife) Workshop at ICML 2021* (2021)
45. Wang, J., Fox, I., Skaza, J., Linck, N., Singh, S., Wiens, J.: The advantage of doubling: a deep reinforcement learning approach to studying the double team in the nba. *arXiv preprint arXiv:1803.02940* (2018)
46. Yeh, R.A., Schwing, A.G., Huang, J., Murphy, K.: Diverse generation for multi-agent sports games. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2019)
47. Zhan, E., Zheng, S., Yue, Y., Sha, L., Lucey, P.: Generating multi-agent trajectories using programmatic weak supervision. In: *International Conference on Learning Representations* (2019)