A New Performance Metric For Player Evaluation Based On Causality

Research Paper Track

Alessandro Cecchin
ceck1989@libero.it

1 Introduction

One of the most difficult tasks for the coaching staff in team sports, such as soccer, is to evaluate objectively the individual performance of the players. Most Possession Value models available attribute the value of an event to the player on the ball despite the change in game state being a result of the movement of several players, both attacking and defensive. This is a flaw that could be overcome through the appropriate use of the causal inference framework. The article presents a new performance metric based on the concept of causality whose goal is to decompose the possession value into the contribution of each player. The metric is consistent in measuring the effects of both offensive and defensive players; thus, it results useful for evaluating both on-the-ball and off-the-ball actions.

The article is structured as follows: firstly, in *section 2*, we present a literature review, while in *section 3* we introduce the concept of causality. In *section 4*, we define the new metric for individual players' evaluation. In *section 5*, we report a few examples in order to explain the main concepts about the new performance metric. In *section 6*, we show how to generalize the proposed approach. Finally, in *section 7*, we summarize the obtained results and we point out a relevant open question.

---

2 Literature review

In recent years, many books and papers have proposed methods for evaluating players' performances and scoring probabilities both in soccer ([1], [2], [3], [4], [5], [6], [7]) and other sports ([8], [9], [10], [11], [12]). With the advent of spatial-tracking data ([13]), it became possible to answer much deeper questions than before, when only statistics were recorded. In [10] the authors define the scoring probability in basketball as a function of shooting distance and few time measures, i.e. time from last event, total time played and time played in a given quarter. The works by Spearman ([5]) and Rios-Neto and al. ([6]) define a scoring probability in soccer for off-the-ball players considering their probability of receiving the ball, their probability of controlling the ball and their probability of scoring based on their position on the court. The M.Sc. thesis by Koren ([7]) proposes a way for considering also obstacles, i.e. defending players, other players in the range and the goalkeeper, as relevant factors for estimating scoring probabilities in soccer. Including the notion of obstacles is very important because it permits to link the effects of defensive players to the opposite offense. On the other hand, assigning a scoring probability to all players, even those off-the-ball, permits to have a factorization of the total scoring probability of a team. As proposed in [3], instead of considering only the scoring probability on the next action, it is also possible to consider the scoring probability in the next k≥1 actions.

While most of the works in the sport analytics cover only a specific task, in [1], [3] and [9] the authors presented new frameworks for a deeper understanding of the games. For example, in [1] and [9] the authors presented the concept of expected possession value in soccer and basketball respectively, while in [3] the authors proposed SPADL, that is a language for representing player actions, and HATTRICS, that is a framework for evaluating players' actions. These new metrics allow the analysts to evaluate all players' on-the-ball actions by measuring how they affect the overall team's probability of scoring. In particular, one of the proposed methods consists of measuring the impact of a player as the difference between "possession value at the end" and "possession value at the start" of an event.

In our opinion, the last defined metric has many limitations, even if it has the merit of tracing a new route regarding individual player analysis. It should be clear by looking at *figure 1* that at each event, what we are really measuring by considering the difference between "possession value at the end" and "possession value at the start" of the event is the global effect of a multi-agents (players) interactions and not an individual effect. *Figure 1* reports two consecutive states of a soccer game (we consider only three offensive players, three defensive players and the goalkeeper for simplicity). Players on offense are colored in blue and are numbered by 1, 2 and 3, while players on defense are colored in yellow and numbered again 1 to 3. The goalkeeper is colored in green. Blue or yellow dotted arrows represent players' movements, while black arrows represent passes. The two states represent the states at the beginning and at the end of the pass event reported. We can see offensive player number 1 making a pass to his teammate, i.e. offensive player number 2, who is cutting into the interior of the pitch; meanwhile offensive player number 3 is making a movement to create space. In this situation, it is not clear how many credit we have to give to each one of the players. Is it more relevant the pass, the cut for receiving the pass or the movement that creates space? At this point, what we want to highlight is that giving all the credit to the play on-the-ball leads to over/under-estimating the single player impact.

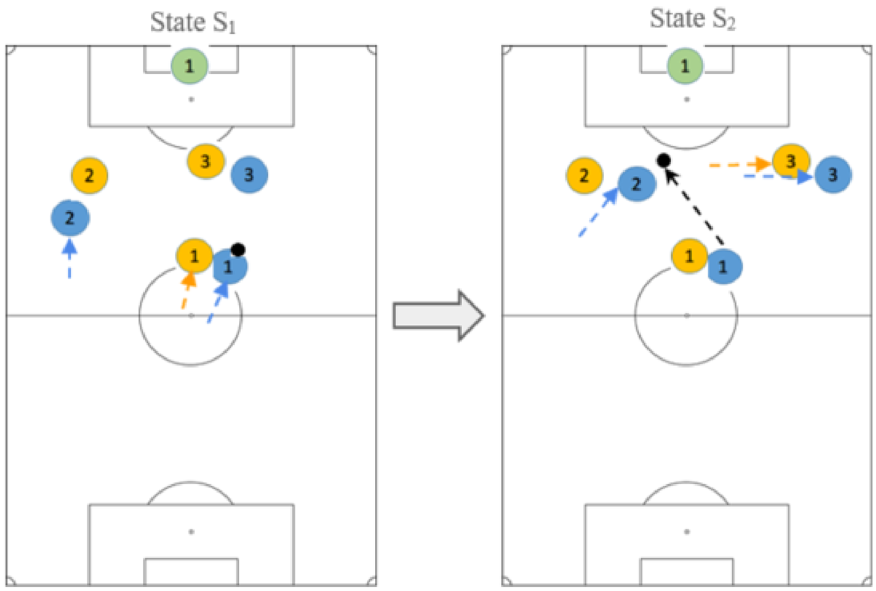

Figure 1: Two consecutive states S_1 and S_2 of a soccer game.

3 Causality

Following the ideas introduced by philosopher E. T. Knott in his work ([14]) on the value that theories of causation can bring to the sport, in this paper, we are going to define a new performance metric for players evaluation based on the concept of causality. The new metric must be able to capture the inherent multi-agent interactions of soccer by measuring the individual players' contribution to the team. Furthermore, the metric must be able to measure the effects of plays both on-the-ball and off-the-ball and thus, it have to be well defined both for offensive and defensive plays. Before going on with the metric definition, we need to do a brief introduction on causality.

Pearl ([15]) proposed a three-level hierarchy classification of causal information in terms of the kind of questions each class is capable of answering. The three levels are Association, Intervention, and Counterfactual, to match their usage. The level at the bottom is that of Association because it invokes purely statistical relationships, defined by the naked data. For instance, observing a customer who buys toothpaste makes it more likely that this customer will also buy floss; such associations can be inferred directly from the observed data using standard conditional probabilities and conditional expectation. Questions at this layer do not require causal information. The second level, Intervention, ranks higher than Association because it involves not just seeing what is, but changing what we see. A typical question at this level would be: “What will happen if we double the price?”. Such a question cannot be answered from sales data alone, as it involves a change in customers' choices in reaction to the new pricing. These choices may differ substantially from those taken in previous price-raising situations, unless we replicate

precisely the market conditions that existed when the price reached double its current value. Finally, the top level invokes Counterfactuals. A typical question in the counterfactual category is: “What if I had acted differently?” thus necessitating retrospective reasoning.
Considering that the literature about counterfactual reasoning in sports analysis is rather poor, to make the concepts more clear we borrow an example from Pearl ([16]): suppose that out of one million children, 99 percent are vaccinated, and 1 percent are not. If a child is vaccinated, he or she has one chance in one hundred of developing a reaction, and the reaction has one chance in one hundred of being fatal. On the other hand, he or she has no chance of developing smallpox. Meanwhile, if a child is not vaccinated, he or she obviously has zero chance of developing a reaction to the vaccine, but he or she has one chance in fifty of developing smallpox. Finally, let’s assume that smallpox is fatal in one out of five cases. Everyone should agree that vaccination looks like a good idea. The odds of having a reaction are lower than the odds of getting smallpox, and the reaction is much less dangerous than the disease. But now, let’s look at the data. Out of 1 million children, 990.000 get vaccinated, 9.900 have the reaction, and 99 die from it. Meanwhile, 10.000 don’t get vaccinated, 200 get smallpox, and 40 die from the disease. In summary, more children die from vaccination (99) than from the disease (40). Thus, the data seem to show that the vaccinations cause more deaths than smallpox itself. Should we ban vaccination after this evidence or should we take into account the deaths prevented? When we began, the vaccination rate was 99 percent. We now ask the counterfactual question “What if we had set the vaccination rate to zero?” Using the probabilities given above, we can conclude that out of one million children, 20.000 would have gotten smallpox, and 4.000 would have died. Comparing the counterfactual world with the real world, we see that not vaccinating would have cost the lives of 3.861 children (the difference between 4.000 and 139). We should thank the language of counterfactuals for helping us to avoid such costs.

4 The new metric

Before applying causality for inferring players’ performances, we must clarify what we are going to measure. Following our final consideration in section 2, it is evident that in team sports all players interact between each other and that each play by any given player has effects on teammates and opponents. Defenders must adapt to the offense, but also the offense must adapt to the defense and the ball carrier’s teammates must adapt to his/her plays. Therefore, the only way for assessing the performance of a player’s play is to measure how his/her play affects the whole team. For example, if player A makes a movement that creates space for player B so that he/she can have a better shot after player C’s pass, would player A’s movement be less important than player C’s assist? Alternatively, if a defender double the ball carrier, would his/her choice be worse than protecting the backcourt for preventing a deep pass? We are going to show how to answer these questions by making use of causality.

If the goal of the offense is to maximize its chance of scoring, we can define an offensive performance metric for a given player computing how he/she is affecting the team's total probability of scoring with his/her action, whatever it is a play on-the-ball or a movement off-the-ball. By duality, the goal of defense is to minimize opposite offense's chance of scoring. Because we can define the probability of scoring of each player as a function of defensive players' positions ([7]), we can define a defensive performance metric as an inverse measure of opponents offense's performance. In other words, the metric tries to capture the effect of each player by measuring how his/her play increase or decrease the probability of his/her team to score and how his/her play decrease or increase the probability of opponents to score.

We can do this in a causal way by measuring the effect of intervention: let's consider two timestamps T and $T + 1$. In the interval $\Delta T = (T + 1) - T$ each player of both teams has done something like a play on-the-ball, a movement off-the-ball or even stands still. Consider first the team on offense. To measure the effect that a given player P has on the total scoring probability of his/her team we can "block" the player P, i.e. we let all other players of both teams play as they did during the time interval ΔT except for the player P we are considering. Therefore, if p is the total scoring probability of the team on offense at time $T + 1$ and p' is the total scoring probability of the team on offense at time $T + 1$ while we "block" player P, we can measure the effect of the play by player P as the difference $p - p^{'}$. To be explicit, if p_0 is the scoring probability of the team on offense at time T, the measure of the effect of the play by player P is $(p - p_0) - (p^{'} - p_0)$, so that p_0 elides itself and we get exactly $p - p^{'}$. Thus, $p - p^{'} > 0$ means that player P's play has a positive effect in increasing his/her team chance of scoring, while $p - p^{'} < 0$ means that his/her play has a negative effect.

Following the previous reasoning we made on offense, we can measure the impact on defense of a player P as the inverse of opponents' total scoring probability change while we "block" the defender P. In other words, if p is the total scoring probability of the team on offense at time $T + 1$ and p' is the total scoring probability of the team on offense at time $T + 1$ while we "block" the defender P, we can measure the effect of the play by player P as $-(p - p')$. Hence, $-(p - p^{'}) > 0$ means that player P's play has a negative effect because it increase opponents' chance of scoring, conversely $-(p - p^{'}) < 0$ means that his/her play has a positive effect because it decrease opponents' chance of scoring.

It is easy to note that a given action by a player P can have an impact on the probability of scoring of both the teams. For example, if we consider the probability of scoring on the next action, then a steal or a turnover, changing the possession of the ball, have the effect of zeroing the probability of scoring of the team that was on offense while giving a positive probability of scoring to the team that was on defense. In this case, it is possible to observe that only the team in possession of the ball can score. Instead, if we consider the probability of scoring in the next $k > 1$ actions, then a given play by a player P have most of the time an impact on the probability of scoring of both the teams. It is because, in the span of the next k actions, the team in possession can lose the ball, hence also the probability of scoring of the team on defense can be assumed positive, even if infinitesimal in some situations. Thus, we can define the measure of the effects of a given play by a player P as the sum of the effects on offense and on defense of his/her play. Therefore, if P is a player on offense, we can measure the performance of his/her play as

$$perf(P) = \Delta\left(p_{off}\right) - \Delta\left(p_{def}\right),$$

where $\Delta\left(p_{off}\right)$ and $\Delta\left(p_{def}\right)$ are the effect of his/her play on the probability of scoring of the team on offense (his/her team) and on defense (the opponents), respectively.

Vice versa, if P is a player on defense, we can measure the performance of his/her play as

$$perf(P) = \Delta\left(p_{def}\right) - \Delta\left(p_{off}\right).$$

Note that our approach differs from that of [3] because at a given game state we are measuring each player effect one at the time while in [3] the authors awards only the play on-the-ball. At this point, it should be clear by our explanation that it is not possible to estimate correctly the performance of on-the-ball actions without also considering and estimating the performance of off-the-ball actions, otherwise we will finish to overestimate or underestimate the impact of the player on-the-ball.

A limitation of the proposed causal approach is that, to consider the entire effect of an offensive play, we must also take into account the related synchronous defensive plays like in an action-reaction relationship. The idea for overcoming this problem comes from the works [17], [18] and [19] on defensive metric definitions. The relevant innovation of the cited works is that of defensive matching: the authors propose ways for assigning to each defender a responsibility to guard an offensive player, letting more defenders to guard the same opponent. If we are able to assign defensive responsibilities at a given time T, i.e. we assign to each defender the matched opponent at that state of the game, then we can assume that the defender's action in the time interval $\Delta T = T - (T - 1)$ is related to the action of the opponent player P he/she is guarding. Therefore, to measure the effect of an offensive play by a player P in the time interval between T and $T + 1$ we have just to block player P and the matched defenders $D_1, \dots, D_k$ and consider the probabilities difference $\left(p_{off} - p_{off}'\right) - \left(p_{def} - p_{def}'\right)$, where p_{off} and p_{def} are the total scoring probability of the team on offense and on defense, respectively, at time $T + 1$ and p_{off}' and p_{def}' are the total scoring probability at time $T + 1$ of the team on offense and on defense, respectively, while we simultaneously "block" the player P and the matched defenders $D_1, \dots, D_k$. On the defensive side, there is no limitation in our approach because, in our view, defensive actions are always reactions to offensive plays or potential plays.

We can observe that this definition of the performance metric is coherent with the nature of soccer: each event is the effect of multiple plays, thus the plays on-the-ball cannot be the only actions awarded, but each player must be accounted for. An example is that of a player whose movement creates space for a teammate. His/her movement can be as valuable as an assist.

Another advantage of the proposed metric definition is that it is easy to generalizable. We just have to ask the counterfactuals question: "how would have changed the probability of scoring of both teams if player P would have acted differently?". Following our approach, it is sufficient to model an alternative scenario in which player P plays the action we want to analyze and look at the probability change with respect to the real game state. Moreover, by considering a set of feasible alternative plays, we can obtain a statistical measure of how good was the choice of player P in that game situation. This concept will be explain better in *section 6*.

However, some situations need to be dealt with separately. In soccer, the goalkeeper role is unique and therefore it has some peculiarities. Goalkeepers can be treated as all other players at every time except when a shot event occurs. After a shooting event, the ball can go out of the field, the goalkeeper can save it or a goal can occur. In the first case, whatever the goalkeeper does, he/she has no impact on the opponents' scoring probability, so his/her effect is 0. Clearly, the goalkeeper had an effect on the scoring probability of the shooter in the states before the shot, but when the shot goes out, in every counterfactual state regarding the goalkeeper, his/her action does not change the result of the shot. In the other two cases, instead, we have to adapt our definition of the performance metric. When the goalkeeper saves the ball he/she gains a score equal to the shooter probability of scoring, let's say p, while when he/she suffers a goal, he/she gains a negative score equal to $p - 1$. Eventually, we have also to sum up the change in scoring probability of both the teams in the next $k > 1$ actions if we are considering that measure of probability. To prove our need to treat the goalkeeper as a special case, if we do not change the performance metric definition, consider the following situation. The goalkeeper would gain a score equal to 1 everytime he/she saves the ball because if we would "block" the goalkeeper then there would be quite always a goal, thus the effect of a save is always 1, not distinguishing between easy and difficult saves. In reality, it can happen that a defender would save the ball even if the goalkeeper does not, but we can assume this case happens with a probability close to zero with respect to the whole set of shots, thus we can neglect it. On the other hand, the goalkeeper would gain a score equal to 0 everytime he/she suffers a goal because if we would "block" the goalkeeper there would be the goal anyway, thus the effect of the goalkeeper's play would be always 0. As before, we can assume that the probability that the goalkeeper saves the goal by not moving are close to zero. Thus, in our opinion, the alternative performance metric definition for the goalkeeper role is more suitable for these game situations.

Furthermore, we can see that foul events can be treated according to the original definition of the performance metric, even if the events of yellow and red cards should be addressed separately. In these cases, it is possible to define a penalty score to assign to the player who is booked or expelled because this will have an impact on the rest of the game. In fact, if the player is expelled his/her team will play with one less player, while if he/she is booked with a yellow card he/she will have to play more carefully to not be expelled.

5 Case studies

Because we do not have tracking data to build a real case study, we present three artificial examples in order to make the concepts more clear. We highlight that the scoring probabilities reported in *table 1* and *table 2* are invented estimates in order to run the

examples. It should be clear that, in a real application of the method, the scoring probabilities could be taken from any suitable model.
In all the case studies, the probabilities of scoring are considered over a span of $k = 10$ states. During this section we will use the following notation: we call each state of the action as S_i, $for\ i = 1, 2$; players on offense as O_j, $for\ j = 1, ..., 3$; players on defense as D_k, $for\ k = 1, ..., 3$; the goalkeeper as GK and the alternative states where we are blocking a player P as $S'_{i,P}$, $for\ i = 1, 2$. Thus, if O is a player on offense then $S'_{i,O}$ is the alternative state in which O and all the defenders associated with O have been blocked, while if D is a player on defense, then $S'_{i,D}$ is the alternative state in which only D has been blocked. We indicate the team on offense by T_{off} and the team on defense by T_{def}. With $P\left(S_i(P)\right)$ and $P\left(S_i(T)\right)$ we mean the probability of scoring of player P or the probability of scoring of team T at the state S_i, respectively. Finally, we call $PR_{S_i}(P)$ the performance achieved by a player P with his/her action at the state S_i.

5.1 Example 1

In the first example, we consider the game situation in *figure 1*, where we can see attacking player number 1 making a pass to the attacking player number 2, who is cutting into the interior of the pitch, meanwhile offensive player number 3 is making a movement in order to create space for his/her teammates. *Figure 2* represent the alternative state S'_2 where we are blocking, respectively, player O_1 and the associated defender D_1, player O_2 (in this scenario we don't block the associated defender D_2 because he/she has not moved), player O_3 and the associated defender D_3, and player D_3.

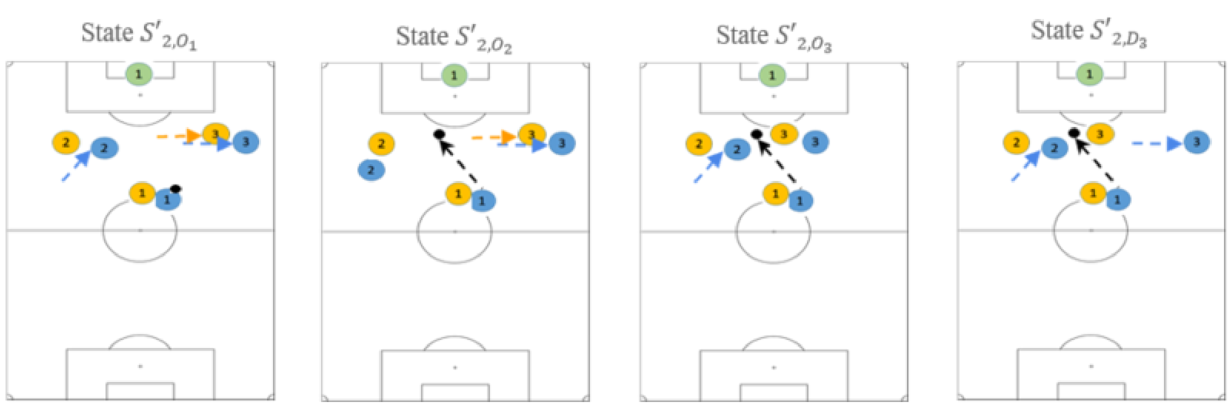

Figure 2: alternative states S'_2.

Table 1 reports the probabilities of scoring for each player and each team at each state and alternative state represented in *figure 1* and *figure 2*. Moreover, the last column of *table 1* reports the performance associated with each player during state S_2.

Player / Team	$P(S_1(*))$	$P(S_2(*))$	$P(S'_{2,O_1}(*))$	$P(S'_{2,O_2}(*))$	$P(S'_{2,O_3}(*))$	$P(S'_{2,D_3}(*))$	$PR_{S_2}(*)$
O_1	0.0085	0.0065	0.0083	0.0018	0.0048	0.0048	0.0082
O_2	0.0112	0.0300	0.0204	0.0042	0.0166	0.0166	0.0446
O_3	0.0123	0.0125	0.0128	0.0010	0.0154	0.0164	0.0138
D_1	0.0018	0.0012	0.0015	0.0021	0.0019	0.0020	0.0000
D_2	0.0012	0.0005	0.0008	0.0014	0.0012	0.0015	0.0000
D_3	0.0010	0.0008	0.0009	0.0016	0.0010	0.0010	-0.0132
GK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
T_{Off}	0.0320	0.0490	0.0415	0.0070	0.0368	0.0378	
T_{Def}	0.0040	0.0025	0.0032	0.0051	0.0041	0.0045	

Table 1: Scoring probability table for statesS_1, S_2 and alternative states S'_2.

As a useful exercise, we evaluate in detail the performances of players O_1 and D_3 at the game state S_2. For player O_1, we can see from *table 1* that in the alternative state S'_{2,O_1} the probabilities of scoring for the teams on offense and on defense are 0.0415 and 0.0032, respectively, while for player D_3 the probabilities of scoring for the teams on offense and on defense at the alternative state S'_{2,D_3} are 0.0378 and 0.0045, respectively. We can also observe that, at the real state S_2 the probabilities of scoring for the teams on offense and on defense are 0.049 and 0.0025, respectively. Thus, the performances achieved by players O_1 and D_3 at the state S_2 are:

$$PR_{S_2}(O_1) = [(0.049 - 0.0415) - (0.0025 - 0.0032)] = 0.0082;$$

$$PR_{S_2}(D_3) = [(0.0025 - 0.0045) - (0.049 - 0.378)] =- 0.0132.$$

If we do the same for the other active players, we get:

$$PR_{S_2}(O_2) = [(0.049 - 0.007) - (0.0025 - 0.0051)] = 0.0446;$$

$$PR_{S_2}(O_3) = [(0.049 - 0.0368) - (0.0025 - 0.0041)] = 0.0138.$$

The not active players D_1, D_2 and GK get an evaluation equal to 0 because they did not move during the action.

Looking at the probabilities of scoring of the real states $S_1$and S_2 we can note an increase in the chance of scoring for the team on offense and a decrease in the chance of scoring for the team on defense. This is coherent with the fact that the attacking team was able to get an open player in possession of ball much closer to the goal area. Moreover, all the offensive players got a positive score by the metric. Player O_2 receives the highest credit because he put himself/herself closer to score. Player O_3 receives the second highest score because if he/she had not created space for his/her teammate, defensive player D_3 would have been in a position to close on attacking player O_2 or even intercept the pass. Player O_1, instead, receives a lower score than his/her teammates because even if he/she had not made the pass, there would still be the time to make it without compromising too much the positional advantage of the team. Finally, defensive player D_3 receives a negative score because if he/she had not followed attacking player O_3, he/she would have been in a position to close on attacking player O_2 or to intercept the pass.

5.2 Example 2

In the second example, we consider the game situation in *figure 3*, where we can see attacking player number 1 winning a duel against his/her defender and attacking player number 3, followed by defender number 3, moving into the interior of the field in order to create space for offensive player number 1. Attacking player number 2 and the associated defender number 2 do not move during this frame.
Figure 4 represent the alternative states S_2' where we are blocking, respectively, player O_1 and the associated defender D_1, player O_3 and the associated defender D_3, player D_1 and finally player D_3. There are no alternative scenarios for player O_2, D_2 and goalkeeper GK because they did not move during this action.
Table 2 reports the probabilities of scoring for each player and each team at each state and alternative state represented in *figure 3* and *figure 4*. Again, the last column of *table 2* reports the performance associated with each player during state S_2.

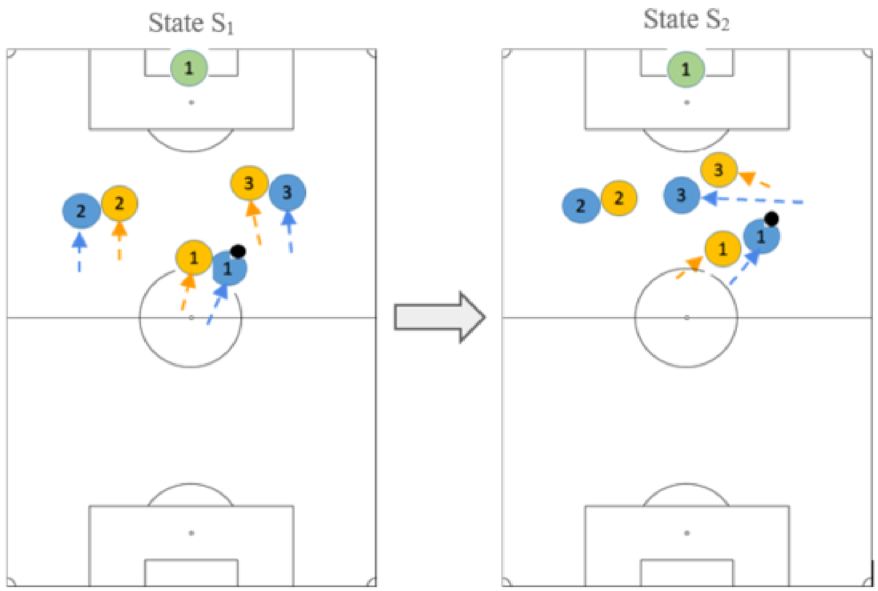

Figure 3: Two consecutive states S_1 and S_2 of a soccer game.

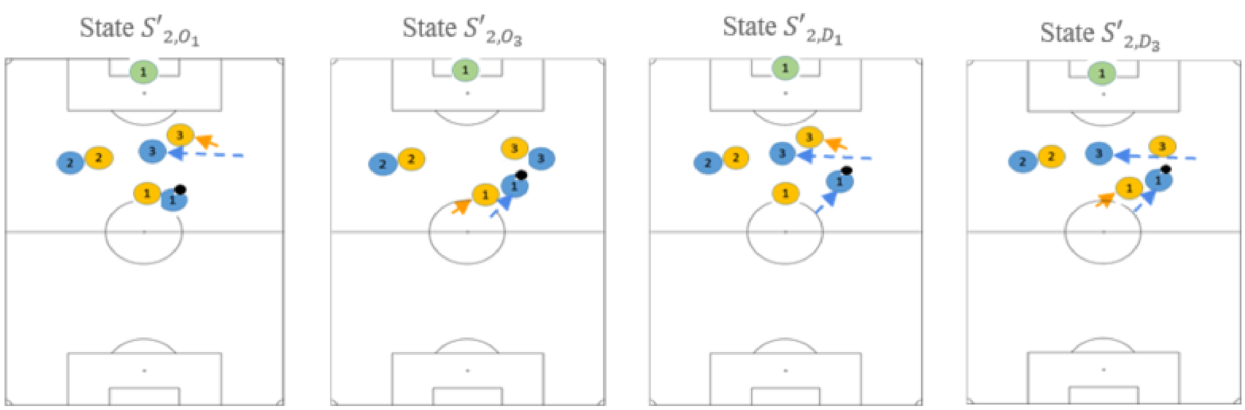

Figure 4: alternative states S'_2.

Looking at the probabilities of scoring of real states S_1 and S_2 we can note an increase in the chance of scoring for the team on offense and a decrease in the chance of scoring for the team on defense. This is coherent with the fact that the attacking team has improved its position thank to offensive player O_1 who wins the duel against his/her defender and find himself/herself in open space after his teammate O_3 cut into the interior of the field, followed by defender D_3. In this situation, all the active players, even the defenders, got a positive score by the metric. Player O_1 receives the highest credit after he/she has won the duel against defender D_1. Player D_3 receives the second highest score. It can appear unintuitive that a defender can get a positive score even if the offense is improving its scoring chances, but if we look at the counterfactual state, we can see that, had not player D_3 followed player O_3, than the probability of scoring of the offensive team would have been greater. In fact, player O_3 would have been free of receiving a pass and with no defender between him and the goalkeeper. Thus, the metric reflects the right choice of player D_3 rewarding him/her with a positive score. When this happen, it is possible to affirm that the merits of the offense are greater than the demerits of the defense. With the third highest score, but very close to the score of player D_3, there is Player O_3: he/she has the merits of deflecting defender D_3 towards the interior of the field, hence creating space for his teammate O_1. Finally, defensive player D_1 receives the lowest score because he/she lost the duel against attacking player O_1. Anyway, the score of player D_1 is positive because if he/she had not followed attacking player O_1, there would have been a two versus one situation where player D_3 have to defend both O_1 and O_3.

Player / Team	$P(S_1(*))$	$P(S_2(*))$	$P(S'_{2,O_1}(*))$	$P(S'_{2,O_3}(*))$	$P(S'_{2,D_1}(*))$	$P(S'_{2,D_3}(*))$	$PR_{S_2}(*)$
O_1	0.0140	0.0300	0.0200	0.0210	0.0350	0.0180	0.0204
O_2	0.0180	0.0200	0.0180	0.0200	0.0200	0.0520	0.000
O_3	0.0180	0.0240	0.0200	0.0190	0.0280	0.0200	0.0170
D_1	0.0030	0.0010	0.0028	0.0020	0.0004	0.0002	0.0113
D_2	0.0025	0.0010	0.0024	0.0022	0.0002	0.0001	0.0000
D_3	0.0025	0.0010	0.0022	0.0018	0.0001	0.0001	0.0186
GK	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
T_{Off}	0.0500	0.0740	0.0580	0.0600	0.0830	0.0900	
T_{Def}	0.0080	0.0030	0.0074	0.0060	0.0007	0.0004	

Table 1: Scoring probability table for states S_1, S_2 and alternative states S'_2.

5.3 Example 3

In the third example, we consider the game situation in *figure 5*. The real states are labeled by “State S_1” and “State S_2”. In this example we have drown only two offensive and two defensive players for simplicity.

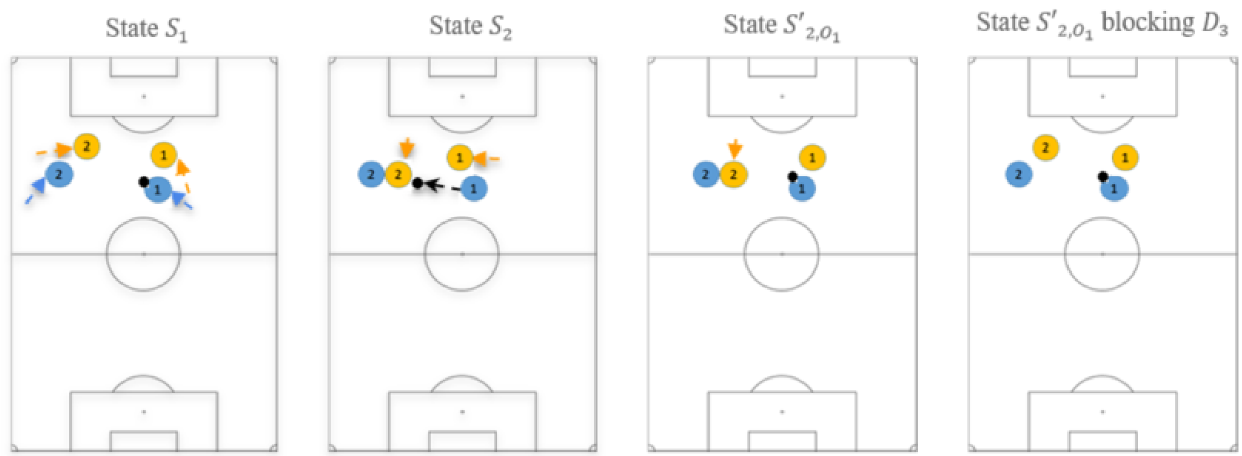

Figure 5: States S_1, S_2, S'_2 and the real S'_2 where we block also the indirect associated defender D_2 .

The game situation depicted is that of an intercepted pass: offensive player O_1 is passing the ball to his/her teammate, while defender D_2 intercept the ball. The sub-image of *figure 5* labeled "State S'_{2,O_1}" represents the counterfactual state where we have blocked player O_1 and the associated defender D_1. We can see that player D_2 is moving to intercept the pass even if there is no passage because player O_1 has been blocked. This creates an unlikely game situation because, having not considered the movement of player D_2 linked to the pass of player O_1, by letting D_2 move even if O_1 has not passed the ball, there would create a lot of space behind defender D_2 for a backdoor movement of player O_2, thus resulting in a much more dangerous game situation. If, instead, we block also player D_2 along with player O_1 (*figure 5*, sub-image labeled "State S'_{2,O_1} blocking D_3"), than we obtain a much more likely game situation where defender D_2 is guarding attacking player O_2 and is not reacting to a pass that is not happening.

6 A more general use case

The approach defined in *section 4* can be used also to measure how optimal a given play is with respect to a set of alternative plays. When we consider a given player for his/her evaluation, we can also consider the set of his/her feasible alternative plays, which result in a set of counterfactual states. For example, in *figure 6*, it is possible to observe the feasible area of movements and a set of alternative movements by player D_3 at state S_2 in the second example (*figure 3*).

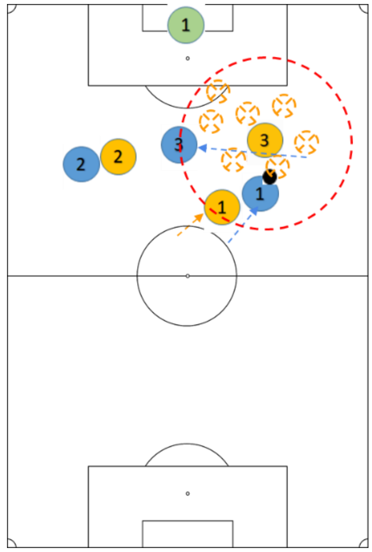

Figure 6: Feasible area of movements and a subset of alternative movements for player D_3.

Thus, we can evaluate a set of statistics regarding the goodness of a player's play by considering a finite subset of feasible plays and computing the performance for each alternative state. Then, we just need to compute the statistics of interest and to compare the performance with the actual play. For example, we can be interested in the performance ratio between the actual and the optimal play or in the comparison between the actual play and the optimal-worse plays performance interval. A final note is that the set of feasible alternative plays is infinite because we can consider every infinitesimal movement of a given player inside the feasible movement area, but we can focus on a finite set of alternative plays because to infinitesimal positional differences correspond infinitesimal differences in the scoring probabilities between the actual and the counterfactual states.

7 Conclusions

Most Possession Value models available attribute the value of an event to the player on the ball despite the change in game state being a result of the movement of several players, both attacking and defensive.
This is a flaw can could be overcome through the appropriate use of the causal inference framework.
By considering the mutual interactions between players, the new proposed metric results coherent with the teamwork nature of soccer. The metric allows evaluating a player's performance considering his/her all-around impact instead of attributing all the value of an event to the player on the ball.

In *section 4*, we have stressed that the link between offensive and defensive plays is a critical aspect of our approach, but that we can overcome it by extending the effect of an

offensive play by also taking into account the associated defensive plays of matched defenders.
In *section 5*, we have reported three examples. *Example* 1 and *example 2* show how the metric works. One of the main advantage of the metric is that a player can be evaluated positively even if his/her team is not doing well. In fact, in the second example, we have seen that defenders got positive evaluation scores even if the opponent team is improving its chance of scoring throughout the sequence of game states. This is possible because the merits of the offensive team are more than the demerits of the defensive team: if the players on defense had not acted as they did, than the probability of scoring of the team on offense would have increased much more. In *example 3*, by considering the case of an intercepted pass by a defender not matched with the player who makes the pass, we points out an important open question. The problem is that of indirect associations between defenders and attacking players. If we not consider that kind of associations, then the counterfactual state would become quite unlikely.

Moreover, it is possible to generalize our approach easily in order to measure the goodness of a play. As shown in *section 6*, given a game state, we can evaluate alternative scenarios and answer the counterfactual question "what if the player would have acted differently?". By analyzing alternative scenarios, it is possible to evaluate the range of probabilities between the best and worst actions playable by a given player and to measure how good the real play was.

A final note is that the proposed metric can be adapted to many other sports such as basketball or hockey.

References

[1] Fernández J., Bornn L. and Cervone D. Decomposing the immeasurable sport: A deep learning expected possession value framework for soccer. In: *MIT Sloan Sports Analytics Conference*, 2019.
[2] Llana S., Madrero P. and Fernández J. The right place at the right time: Advanced off-ball metrics for exploiting an opponent's spatial weaknesses in soccer. In: Londres: *MIT Sloan Sports Analytics Conference*, 2020.
[3] Decroos T., Bransen L., Van Haaren J., et al. *Actions Speak Louder Than Goals: Valuing Player Actions in Soccer*. Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019, pp.1851–1861.
[4] Pappalardo L., Cintia P., Ferragina P., et al. PlayeRank: Data-driven Performance Evaluation and Player Ranking in Soccer via a Machine Learning Approach. *ACM Transactions on Intelligent Systems and Technology*, 2019, 10(5): 1-27.
[5] Spearman W. Beyond Expected Goals. In: *Londres: MIT Sloan Sports Analytics Conference*, 2018.
[6] Rios-Neto H. M. R., Wagner Meira Jr. and Vaz-de-Melo P. O. S. A new look into Off-ball Scoring Opportunity: taking into account the continuous nature of the game. In: *Barcelona Analytics in Sports Tomorrow*, 2020.
[7] Koren A. *Estimating Goal-Scoring Probabilities in Soccer, Based on Physical and Geometric Factors.* M.Sc. Research Proposal, The Faculty of Industrial Engineering and Management Technion- Israel Institution of Technology, 2013.
[8] Oliver D. *Basketball on Paper: Rules and Tools for Performance Analysis*. Potomac Books Inc., 2004.
[9] Cervone D., D'Amour A., Bornn L., et al. A multiresolution stochastic process model for predicting basketball possession outcomes. *Journal of the American Statistical Association*, 2016, 111(514): 585-599.
[10] Zuccolotto P. and Manisera M. *Basketball Data Science*. Londres: Chapman and Hall, 2020.
[11] Fuentes C. S., Carlsson N. and Lambrix P. Player impact measures for scoring in ice hockey. In: *MathSport International 2019 Conference*, 2019.
[12] Zuccolotto P., Manisera M. and Sandri M. Big data analytics for modeling scoring probability in basketball: The effect of shooting under high-pressure conditions. *International Journal of Sports Science & Coaching*, 2018, 13(4): 569–589.
[13] Shea S. *Basketball Analytics: Spatial Tracking*. CreateSpace Independent Publishing Platform, 2014.
[14] Knott E. T. What Might a Theory of Causation Do for Sport?. *Philosophies*, 2019 4(2): 34-45.
[15] Pearl J. The seven tools of causal inference, with reflections on machine learning. *Commun. ACM*, 2019, 62(3): 54-60.
[16] Pearl J. and Mackenzie D. *The Book Of Why*. Basic Books, 2018.

[17] Franks A., Miller A., Bornn B., et al. Characterizing the spatial structure of defensive skill in professional basketball. *The Annals of Applied Statistics*, 2015, 9(1): 94–121.
[18] Keshri S. *Essays in Basketball Analytics*. Doctoral Thesis, Columbia University, 2019.
[19] Shaw L. and Glickman M. Dynamic analysis of team strategy in professional football. In: *Barça Sports Analytics Summit*, 2019.