Pass Impact Factor in Soccer Report

Nahid Aghababaeyan

October 2023

nahid.aut+soccer@gmail.com   LinkedIn

1 Introduction

Soccer, a sport beloved by millions worldwide, has long been a subject of fascination for enthusiasts, coaches, and researchers alike. From Charles Reep's pioneering efforts in the 1950s to the current era of advanced automated data collection, the study of soccer statistics has evolved dramatically. At the heart of this evolution lies the intricate art of passing, a skill that transcends individual brilliance to shape team success. Successful passes can unlock defenses, transition a team from defense to attack, and ultimately create those euphoric goal-scoring moments.

Going beyond numbers, this research explores the real impact of soccer passes, delving into the intricate fabric of the game. We introduce the idea of an "impact factor" for passes to measure their importance. By using the Metrica Sports dataset, we aim to provide valuable insights to the world of soccer analysis, offering coaches, analysts, and players a powerful tool for tactical improvement and informed decision-making.

2 Objective

The primary aim of this project is to establish a comprehensive and effective metric for evaluating the value of each pass made during a soccer game. In the world of soccer, every pass plays a crucial role in determining the outcome of a match, and having a quantifiable measure of the value of these passes can significantly enhance our understanding of player performance, team strategies, and overall game dynamics.
Implementation: Once the metric is devised, we will proceed to implement it in a practical and user-friendly manner. This includes the definition of the metric, data visualization, or any other necessary resources to facilitate the application of the metric based on the sample dataset.

3 About Dataset

In Game 2, there are three distinct datasets within the data folder, each providing valuable information related to the game. These datasets are:

- Tracking Data for Team Away: records the positions of Team Away players at 0.04-second intervals throughout the entire game.

- Tracking Data for Team Home: records the positions of Team Home players at 0.04-second intervals throughout the entire game

- Event Data: This dataset encompasses all the events that occurred during the game, including various types of events like player actions, fouls, and more.

Let's delve deeper into the structure and content of each of these datasets.

3.0.1 Tracking dataset

To facilitate the analysis and manipulation of the data, the first step involved merging the two tracking datasets into one cohesive dataset. This merged dataset contains information about both Team Away and Team Home players, making it easier to work with the tracking data. The resulting tracking dataset comprises 141,156 rows and 55 columns, with the following important variables:

Table 1: Tracking Dataset Snapshot

Index	Variable	Description
0	Period_x	Period or half of the game
1	Time [s]	Time in seconds
2	Away_25_x	Away Team Player 25 X-position
3	Away_25_y	Away Team Player 25 Y-position
4-25	Away 15 to Away 26	Away Team Player Positions (15 to 26)
26	ball_x_x	Ball X-position
27	ball_y_x	Ball Y-position
28	Home_11_x	Home Team Player 11 X-position
29	Home_11_y	Home Team Player 11 Y-position
30-55	Home 1 to Home 14	Home Team Player Positions (1 to 14)

	Period_x	Time [s]	Away_25_x	Away_25_y	ball_x_x	ball_y_x	Home_11_x	Home_11_y
50	1	2.040000	-52.666100	-0.199240	0.228960	-0.255680	47.478460	0.689520
80	1	3.240000	-52.802840	0.006800	-8.259520	-0.510680	47.065060	0.562360
200	1	8.040000	-48.668840	0.816000	-4.888720	19.216120	42.213440	1.462680
400	1	16.040000	-43.606280	0.264520	-25.529040	-13.872680	35.348880	-0.286280
500	1	20.040000	-42.908800	-1.392640	12.215440	-29.191720	36.740660	-3.355800

Figure 1: Snapshot of the Tracking Dataset: This image displays a sample of five rows from the tracking dataset, showcasing the positional data of one player from the Away team and one player from the Home team. For simplicity, some columns have been omitted in this representation.

3.0.2 Event dataset

The Event dataset is composed of 1,935 rows and 14 columns. It provides information about various events that occurred during the game. Here is an overview of the key columns in this dataset:

Table 2: Event Dataset Snapshot

Index	Variable	Description
0	Team	The team associated with the event
1	Type	The general type of the event
2	Subtype	The specific subtype of the event
3	Period	The period or half of the game in which the event occurred
4	Start Frame	The frame at which the event started
5	Start Time [s]	The time in seconds at which the event began
6	End Frame	The frame at which the event ended
7	End Time [s]	The time in seconds at which the event concluded
8	From	The entity or player initiating the event
9	To	The entity or player receiving the event or affected by it
10	Start X	The X-coordinate of the event's starting position
11	Start Y	The Y-coordinate of the event's starting position
12	End X	The X-coordinate of the event's ending position
13	End Y	The Y-coordinate of the event's ending position

	Team	Type	Subtype	Period	Start Frame	Start Time [s]	End Frame	End Time [s]	From	To	Start X	Start Y	End X	End Y
10	Home	BALL OUT	nan	1	504	20.160000	534	21.360000	Player3	nan	12.720000	-28.560000	4.240000	-34.680000
50	Away	BALL LOST	INTERCEPTION	1	2370	94.800000	2417	96.680000	Player19	nan	3.180000	-34.000000	15.900000	-19.720000
100	Home	PASS	nan	1	5707	228.280000	5745	229.800000	Player4	Player7	-19.080000	-25.840000	-31.800000	-33.320000
700	Away	PASS	nan	1	46534	1861.360000	46554	1862.160000	Player17	Player15	-39.220000	-4.760000	-38.160000	-15.640000
800	Away	CHALLENGE	TACKLE-WON	1	51550	2062.000000	51550	2062.000000	Player20	nan	-36.040000	-4.080000	nan	nan

Figure 2: Sample Event Data: In this snapshot, you can see five rows from the Event dataset, offering insights into various events during the game, including team, event type, time, coordinates, and involved entities.

4 Defining a Good Pass in Soccer

In order to establish a measurable metric for rating passes in a soccer game, it is essential to identify the key factors that contribute to a pass being considered 'good' or 'bad.' One of the fundamental elements to consider is the pressure under which a player receives a pass.

4.1 The Impact of Pressure

The level of pressure on a player when they receive a pass plays a pivotal role in determining the quality of the pass. This pressure is closely associated with the number of opponents near the player receiving the pass. When a player is surrounded by a higher number of opponents, there is an increased risk of losing possession of the ball, which can be detrimental to the team's performance.

In essence, a good pass should not only reach its intended target accurately but also be delivered in a manner that allows the recipient to maintain control of the ball, even under challenging circumstances created by the presence of multiple opponents. This emphasizes the need for precise passing and decision-making in soccer, as a well-executed pass should alleviate the pressure on the receiving player.

4.2 Pass Length and Attack Potential

Pass length directly influences a team's ability to create goal-scoring opportunities. Longer passes move the team closer to the opponent's goal, enhancing their offensive potential.

4.3 Passing to the Danger Zone

Passes that successfully reach a player in the opponent's penalty area offer a high chance of creating a shot and scoring. This is known as the danger zone, and passes into this area are of significant importance in soccer. Understanding the impact of these passes on goal-scoring opportunities is a key aspect of evaluating pass quality.

4.4 The Next Event After the Pass

Evaluating the next event immediately following a pass provides enhanced insights into the pass itself. When the next event is a shot or another pass within the team, it indicates effective ball possession. Conversely, if the next event is ball loss or recovery by the opposing team, it suggests that the pass may not have been successful in maintaining possession, which impacts its quality.

4.5 The Importance of Pass Types

The type of pass used in a soccer game, such as crosses, deep balls, headers, and more, carries varying levels of significance. Cross passes and deep balls, for instance, are often more valuable than standard passes. Understanding the influence of pass types on the Pass Impact Factor is essential in evaluating the quality of passes.

5 Pass Rating Metrics

Taking into account the five factors discussed earlier that influence the value of a pass, we have established six metrics to evaluate passes. These metrics consider factors all mentioned factors. To provide a comprehensive assessment, we combine these metrics with different weights, resulting in a single pass rating score. This combined rating helps us see

passes from a broader perspective and makes it easier to compare them.

5.1 Metric 1: Opponents-to-Players Ratio

The first pass rating metric focuses on the ratio of opponents to players in possession of the ball. This metric helps us assess the change in pressure during the pass. We calculate this ratio for both the passer and the receiver within a 5-meter radius of the player who passes the ball. A higher difference in this ratio indicates that the pass moves from a high-pressure area to a less pressured area, which is favorable as it reduces the risk of losing the ball. textbfOpponents-to-Players Ratio is calculated using the formula:

$$\text{Ratio} = \frac{O_p}{P_p} - \frac{O_r}{P_r}$$

O_p : Opponents in the 5-meter area around the passer

P_p : Players in possession around the passer

O_r : Opponents in the 5-meter area around the receiver

P_r : Players in possession around the receiver

why 5 meter? -¿ It takes about 0.8 seconds for a soccer player to reach a ball 5 meters away, assuming an average running speed of 22.5 kilometers per hour. The rating for the "ratio" metric is calculated as follows:

$$\text{rate}_{\text{ratio}} = \left(\frac{\text{diff\_ratio}}{2} \right) \times 0.2$$

This weight of 0.2 is applied to the "ratio" metric in our pass rating system, contributing to the overall assessment.

This metric allows us to quantify the pressure dynamics and contributes to our overall pass rating system.

5.2 Metric 2: Recipient's Defensive Pressure

This metric assesses the number of opponents near the player receiving the pass. Soccer often places high importance on minimizing defensive pressure during a pass. When multiple opponents surround the receiver, it increases the risk of losing possession, limits passing options, and adds pressure, increasing the likelihood of errors. The rating for the "Recipient's Defensive Pressure" metric is calculated as:

$$\text{rate}_{\text{Recipient's Defensive Pressure}} = \left(\frac{\text{number of opponents}}{2} \right) \times 0.1$$

The weight of 0.1 is assigned to this metric, indicating its lesser importance compared to the first metric, which carries a weight of 0.2.

5.3 Metric 3: Pass Length

This metric evaluates the length of a pass, with longer forward passes considered more valuable. The direction of the pass (forward or backward) is a significant factor in this assessment. The rating for the "Pass Length" metric is calculated as follows:

$$\text{rate}_{\text{Pass Length}} = \left(\frac{\text{pass\_vector}}{10} \right) \times 0.1$$

In this formula, for every 10 meters the ball passes, we add 0.1 point to the pass rating. This metric provides insights into the pass's direction and length, contributing to the overall assessment.

5.4 Metric 4: Penalty Area

This metric assesses whether the receiver obtains the ball within the opponent's penalty area, denoted as "penalty_area" (1 for inside, 0 for outside). The rating for the "Receiver Zone" metric is calculated as:

$$\text{rate}_{\text{Receiver Zone}} = \left(\frac{\text{penalty\_area}}{10} \right) \times 0.5$$

The weight assigned to this metric is 0.5, signifying its high importance. When a pass is received successfully in the opponent's penalty area, the team has a significant chance of scoring. This higher weight compensates for the earlier two metrics, given the heightened pressure in the receiver's area.

5.5   Metric 5: Next Event

This metric assesses the next event following a pass and assigns a rating based on the different next events that occur. Below we have shown 14 of these events from a larger list for simplicity(the full list is in the code more than 30 events). Each event is assigned a weight based on its importance, with corresponding weights for different outcomes. Here is a sample of these next events and their associated weights:

Table 3: Next Event after Away Pass

Next Event	Subtype Next Event	Next Event Team	Weight
SHOT	ON TARGET-GOAL	Away	1.0
SHOT	ON TARGET-SAVED	Away	0.9
SHOT	OFF TARGET-OUT	Away	0.8
SHOT	OFF TARGET-HEAD-OUT	Away	0.8
PASS	CROSS	Away	0.6
PASS	HEAD	Away	0.5
BALL OUT	CROSS	Away	0.4
CHALLENGE	TACKLE-WON	Away	0.3
CHALLENGE	GROUND-LOST	Home	0.2
RECOVERY	INTERCEPTION	Home	-0.5
BALL LOST	CROSS-INTERCEPTION	Away	-0.4
BALL LOST	INTERCEPTION	Away	-0.3
BALL LOST	0	Away	-0.2
BALL OUT	0	Away	-0.1

These weights reflect the significance of each event's outcome in determining the overall quality of the pass. This metric contributes to our comprehensive pass rating system, providing valuable insights into the subsequent actions following a pass.

5.6   Metric 6: Type of Pass

This metric evaluates the type of pass and assigns a rate based on each pass type. The following table illustrates the weights assigned to different pass types:

Table 4: Pass Types and Weights

Type	Weight
CROSS	0.7
DEEP BALL	0.2
HEAD	0.2
GOAL KICK	0.1
HEAD-INTERCEPTION-CLEARANCE	0.5

Each pass type is associated with a specific weight, indicating its importance in determining the pass rating. This metric allows us to account for the different values attributed to various pass types and contributes to our comprehensive pass rating system.

At the End we sum all the metrics and sacle the numbers to get a value for the pass rate at the end.

6   Final Pass Rating

At the end of our evaluation, we sum all the pass rating metrics and scale the numbers to obtain a final pass rating value. This value allows us to assess the overall quality of each pass in the game.

Now, let's visualize some key insights from our analysis:

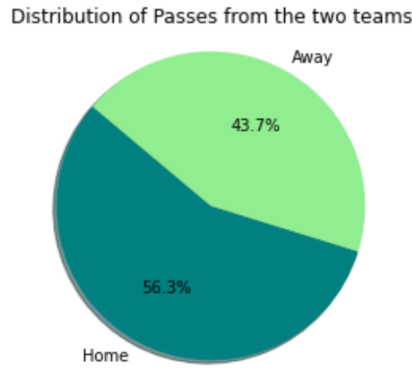Distribution of Passes from the two teams

Figure 3: Total Passes: This visualization shows that Team Home has a higher total number of passes.
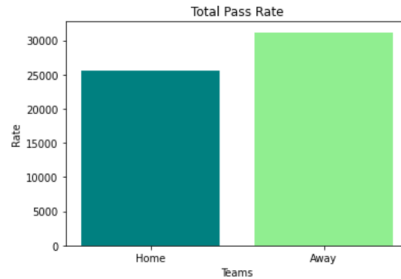

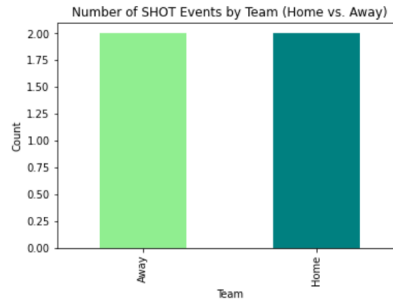Figure 4: Total Pass Rates: This visualization indicates that Team Home has higher pass rates in the game.

Figure 5: Number of Shots: This visualization reveals that both teams had a similar number of shots in total.

7 Sources

1. "Soccer Analytics: how Data Science is changing the 'Beautiful Game" by Luca Pappalardo. (February 2019)
https://github.com/metrica-sports/sample-data

2. "QPass: a Merit-based Evaluation of Soccer Passes" by Laszlo Gyarmati and Rade Stanojevic.
https://arxiv.org/abs/1608.03532

3. "Introducing Pass Value" by Delara Analytics.
http://www.delaraanalytics.com/introducing-pass-value/

4. ChatGPT (GPT-3.5)
As an AI language model developed by OpenAI, ChatGPT has been used to provide information and assistance in generating content for this report.