universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„A Data Science Approach for predicting soccer passes using positional data"

verfasst von / submitted by

Sebastian Eigenrauch, B.Sc.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna, 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 926
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Wirtschaftsinformatik
Betreut von / Supervisor:	ao. Univ.-Prof. tit. Univ.-Prof. Dipl.-Ing. Dr. Erich Schikuta

Abstract

Data-driven approaches for evaluating tactical team behavior in soccer are nowadays a widespread method in sport analytics. The large amount of data collections enables experts to generate a deep tactical understanding and extract valuable measurements out of team-performances. However, these approaches are often limited in their comprehensibility and applicability for domain experts. One of the most frequently occurring events during a soccer match are passes. The motivation of this research is the design, implementation and validation of data science algorithms, that predict tactical motion of defending players after an occurring event of a pass. The focus is the establishment and validation of different sets of rules, which simulate the movement behavior of the defending team, based on domain knowledge. The approach provides a high level of applicability for domain experts, in order to use and combine variable predefined rules for prediction, simulation and evaluation of different tactical approaches of defensive behavior.

Kurzfassung

Datengestützte Ansätze zur Bewertung des taktischen Mannschaftsverhaltens im Fußball sind heutzutage eine weit verbreitete Methodik in der Sportanalytik. Die große Anzahl von Datensammlungen ermöglicht es Experten, ein tiefgehendes taktisches Verständnis zu entwickeln und wertvolle Informationen aus den Leistungsdaten einer Mannschaft zu gewinnen. Allerdings sind diese Ansätze in ihrer Verständlichkeit und Anwendbarkeit für Domainexperten oft begrenzt. Eines der am häufigsten auftretenden Ereignisse während eines Fußballspiels sind Pässe. Die Motivation dieser Arbeit sind Design, Implementierung und Validierung von Data-Science-Algorithmen, die taktische Bewegungen von verteidigenden Spielern nach einem gespielten Pass vorhersagen. Der Schwerpunkt liegt dabei auf der Erstellung und Validierung verschiedener Regelsets, die das Bewegungsverhalten der verteidigenden Mannschaft auf der Basis von Domainwissen simulieren. Außerdem erhöht die wissenschaftliche Arbeit die Anwendbarkeit für Domainexperten, um variable vordefinierte Regeln für die Vorhersage, Simulation und Bewertung verschiedener taktischer Ansätze des Defensivverhaltens zu nutzen und zu kombinieren.

Contents

List of Tables

List of Figures

List of Algorithms

1. Introduction

The application of tracking technology, in order to measure the tactical match performance of a team, is nowadays a commonly used feature in professional soccer. A large amount of data, which provides the possibility to gain valuable insights into tactical team-performances, has been collected over the years. Data-driven approaches enable the extraction of representative measurements, which researchers, as well as domain experts can use, in order to gain a deeper understanding and insights into tactical soccer projections. In soccer, a team´s tactical behavior is an important component for a successful winning strategy, whereby the application of data-driven approaches for evaluation and analysis is a widespread method [Goes-Smit et al., 2018]. One of the most frequently occurring events during a soccer match are passes. Every pass of the attacking team results in a reaction of the defending team, in most cases through an adaptation of the positions of the defending players [Kempe et al., 2016].

In this research, we focus on the design, implementation and validation of data science algorithms, that predict tactical motion of defending players, based on the occurring event of a pass. Thus, we follow the approach of Rule-based Systems (RBS) [Liu et al., 2014]. We concentrate on the establishment and validation of different sets of rules, which simulate the movement behavior of the defending team. These sets of rules follow common defending tactics in modern soccer, which we use to predict and evaluate defending player motion after a played pass. We focus on the establishment of a framework, that is applicable for domain experts. Thus, we aim to enable soccer experts to use and combine variable predefined rules, in order to predict, simulate and evaluate different tactical approaches for defensive behavior.

At the beginning of this thesis, the current state of the art about movement prediction in soccer is described. Therefore, a general explanation about RBS with regard to predicting player motion is provided. Besides, we define what differs our approach from existing data-science researches. A description of occurring problems and the defined research questions as main focus of this thesis are explained in the subsequent section. In section 4, we describe the common structure of the implemented framework. Thus, different rules in order to define the tactical background of our prediction models are established. Besides, we demonstrate the general behavior of each model with regards to its underlying tactical nature. In section 5, the prediction result of each model is evaluated and presented. We compare accuracy and occurring defending patterns of our models with each other, in order to establish representative assumptions. In the last section, a discussion about the prediction results and assumptions is provided, where we discuss limitations and future work of our approach.

2. State of the Art

In the following section, we give a short introduction of the current state of the art about defending motion in soccer. First, we provide a general overview of different approaches for predicting player movement in team sport, based on data-driven methods. Therefore, we introduce approaches from Machine Learning (ML) and Rule-based Systems (RBS). In a next step, an explanation of how rules can be used, in order to extract knowledge from real-world scenarios and establish valuable predictions is provided. Subsequently, we describe how this method is currently applied for predicting player motion. Besides, we explain our approach, where we also take into consideration how our framework differs from previous researches.

Through large data sets, experts are capable of using the underlying data in order to extract information and establish valuable predictions. In soccer, positional data sets help researchers to simulate, how hypothetical actions or events influence the tactical team performance and player positioning of the participating actors [Cichy and Kaiser, 2019]. In general, we distinguish between the approaches of Rule-based Systems and Machine Learning, where the methods of Machine Learning can be further divided into Supervised Learning (SL), Unsupervised Learning (UL) and Reinforcement Learning (RL):

Unsupervised Learning

UL algorithms enable the extraction of rules from data without any human intervention. Occurring patterns and conclusions are found by the model through the provided unlabeled data. Thus, the model does not require any learning instructions, as it discovers solutions by itself [Fujii, 2021]. UL is applicable to large data sets and a commonly used approach for the evaluation of tactical team sport behavior. For instance, an Unsupervised Learning approach is applied by [Wang et al., 2015], where tactical soccer patterns based on match logs are automatically identified through the UL method of principal component analysis (PCA). The study applies PCA, in order to reduce redundant data, by mapping similar types of player to one role. These roles serve as representatives for a general tactical team evaluation. [Decroos et al., 2018] applies the UL method of hierarchical clustering to detect tactics out of soccer matches. Thus, different clusters of the phases of a match are established, where one team is in possession of the ball. Frequently occurring patterns are then identified and evaluated [Decroos et al., 2018].

Supervised Learning

Supervised Learning approaches are rather applicable for training algorithms to classify data or predict specific outcomes. Since these algorithms are defined by labeled data sets, they differ from an Unsupervised Learning approach [Jiang, 2021]. The method usually develops models from data sets, using previously known input and output values [BALADRAM et al., 2020]. Thus, models are fed with input data, where they adjust their weights until they have been fitted appropriately[Jiang, 2021]. An SL approach in team sport is applied in [Shah and Romijnders, 2016]. Therefore, neural networks are used, in order to classify offensive strategies in NBA basketball games, based on players tracking data. A similar approach is performed by [Thabtah et al., 2019]. The research focuses on artificial neural networks and decision trees, to forecast game results in basketball. Since basketball and soccer are based on a similar team sport nature, these approaches seem also applicable to predict player motion in soccer.

Reinforcement Learning

In Reinforcement Learning, an agent learns a policy independently through specific actions in order to maximize cumulative rewards. Thus, the model uses a trial and error approach, where a sequence of successful outputs is further reinforced to develop the best recommendation or policy for a given problem [Sutton, 2018]. A model´s reward function can be estimated from observed data through statistical learning techniques. Besides, explicit assumptions for an optimal criteria of player motion can be made with predefined reward functions such as a the score in soccer [Fujii, 2021]. A RL approach in team sport is used in [Wang et al., 2018]. This method establishes a mapping between player and the decision to double team. This mapping is compared to a cumulative reward of blocked shots. Thus, a policy can be established, where the actions that maximize the expected reward are selected [Wang et al., 2018]. Another projection to be mentioned is conducted by [Le et al., 2017]. This method follows an approach called 'Deep Imitation Learning', where a model automatically learns from expert behavior instead of a predefined reward function. The motion of defending soccer players can therefore be predicted, by comparing the movement of the league average or a specific team with the actual defending team [Le et al., 2017].

Rule-based Systems

Many different models exist, in order to forecast player motion in soccer, with a focus on rule-based approaches. These Rule-based Systems are usually defined as a set of rules, in order to apply and evaluate data. Thus, they use the knowledge of domain experts, to establish hypotheses and extract representative values from data [Liu et al., 2014]. RBS are a commonly used approach in different sports sectors, where experts define theories, based on the characteristical behavior of multi-agents, such as individual players in soccer[Fujii, 2021]. Representative values through RBS in soccer are applied in

[Lang et al., 2016], in order to analyze a team's goal scoring probability. The implemented rules define an increasing goal scoring probability, within a decreasing distance to the opponent's goal and an increasing centrality of the ball position. A rule based approach with a rather focus on predicting specific values, is applied in [Spearman, 2018]. Thus, the approach predicts the scoring opportunity of players, that are not in possession of the ball. Hypotheses are defined, with regard to an attacking team's ball control, the likelihood of a successful pass and the goal scoring probability. These prerequisites are used to provide a future indicator of a team's probability to score [Spearman, 2018].

With our framework, we follow the approach of establishing predefined rules, in order to forecast and evaluate the movement of defending players of different matches and game-constellations. We decide to establish prediction models, that use the knowledge of domain experts. While the previously described RBS approaches rather focus on the extraction of representative values in soccer, such as e.g. the evaluation of a team's goal scoring probability [Lang et al., 2016], our focal point is the establishment of different sets of rules, in order to predict and evaluate the movement of defending players. We thus use domain knowledge, in order to deploy different prediction models with a common tactical soccer approach as basement. We define rules for describing different projections of defending behavior in soccer, such as e.g. zone defense or man marking defending. Besides, the set up of our framework is designed, that it can easily be used by soccer experts, in order to combine different rules and build new models.

3. Problem Description and Research Question

In modern soccer, tactical defending behavior is an essential ability and decisive for the success of a team. Since various tactical approaches exist, the behavior of defending players differs, based on the underlying tactical inputs and leads to different movements of the defenders. As mentioned in the previous sections, literature provides different projections, in order to predict player motion in soccer. Even though, many of them enable a highly accurate prediction and a detailed match analysis such as [Le et al., 2017], an evaluation based on a team's underlying tactical background seems rather rare. Without this information, prediction models tend to be difficult to understand for domain experts and might complicate their applicability [Fujii, 2021]. In order to counteract these limitations, we use data-science algorithms for an establishment and validation of different sets of rules, that simulate the movement behavior of the defending team and follow some fundamental tactical team behavior approaches. Therefore, we use projections of Rule-based Systems (RBS), where we design models with different tactical backgrounds. We want to enable domain experts, to use and combine variable predefined rules, in order to predict, simulate and evaluate different soccer tactics. Thus, we aim to answer the following central questions:

1. Which tactical soccer approach seems most promising, in order to provide the highest prediction accuracy?
2. Are there recurring patterns within specific game constellations, where a prediction model seems particularly applicable?
3. Can general assumptions about movement behavior of defending players be established, based on different tactical backgrounds?

Our framework focuses on the design, implementation and validation of some principles of tactical team behavior in soccer. Through this projection, we analyze, which tactical approach is mostly applied in modern soccer, by evaluating our models and considering which forecast provides the highest accuracy. Beside a general evaluation of the prediction models, we also focus on the question, if some tactics are more applicable for specific game constellations or on specific parts of the field and if any recurring defending patterns can be identified. With the last question, we want to determine, if we can extract general assumptions about defending movement in soccer through our models. For instance, we analyze, if the evaluation results of a model differ for different teams and different games, or if they are consistent for overall matches and implicate a general indicator for defending player motion in soccer.

4. Prediction Framework

The following chapter explains the general set-up of the implemented prediction framework, that is used to forecast the movement action of the defending soccer team. At the beginning, a definition of the paper's scope and an overall explanation of the underlying coding structure is specified, with regard to the application of the framework for domain experts. Section 4.3 gives a short annotation about the used data set, relevant for the establishment and verification of the provided prediction models. To enhance understandability of the different prediction models, the tactical background and general structure for each model is described into detail, substantiated by several visualisations for each model. The code of our framework is provided in our GitHub repository [Bischofberger and Eigenrauch, 2023].

4.1. Scope and Objectives

As the prediction of real-world player motion in soccer might appear quite challenging and there exist various opportunities for a player's movement, we require a model that is capable of covering and predicting different in-game scenarios. We therefore apply the approach of conventional rule-based systems, where we use the knowledge of domain experts, in order to establish valuable prediction models [Liu et al., 2014]. Our main focus is the deployment of several models, that follow specific rules, based on real-world tactical soccer behavior. We therefore extract movement behavior out of soccer tactics and establish different rules. These rules are then combined to form our prediction models. Furthermore, our framework is built up, that it can be used by soccer experts, in order to forecast and evaluate different game constellations, based on the underlying tactical attitude. Due to the fast changing nature of modern soccer nowadays, we need to provide a high level of flexibility and dynamic to our models. Thus, our models require an easy adjustment, where they can be further enhanced by adding or removing rules to our models, which requires a flexible coding structure of the model framework.

In order to deploy a valid prediction framework, we build an interface, where we provide the relevant coding structure of our models (Fig. 4.1). In general, we distinguish between three classes: Rules, Rulesets and Metarulesets. A Rule is a single Python Class, that contains the business logic (tactical background for defending player motion) of a specific prediction Rule, such as e.g. a defender moves towards the current position of the ball. A Ruleset is a concatenation of several Rules, while a Metaruleset contains several Rulesets. The prediction logic, such as drawing the constellation plots or the evaluation of the occurring prediction error (chapter 5.1) is inherited by each Rule, Ruleset and Metaruleset. The framework follows a dynamical approach, as new Rules can easily be implemented to the existing structure. Besides, Rulesets and Metarulesets can be adjusted by adding or removing Rules/Rulesets to the particular set.

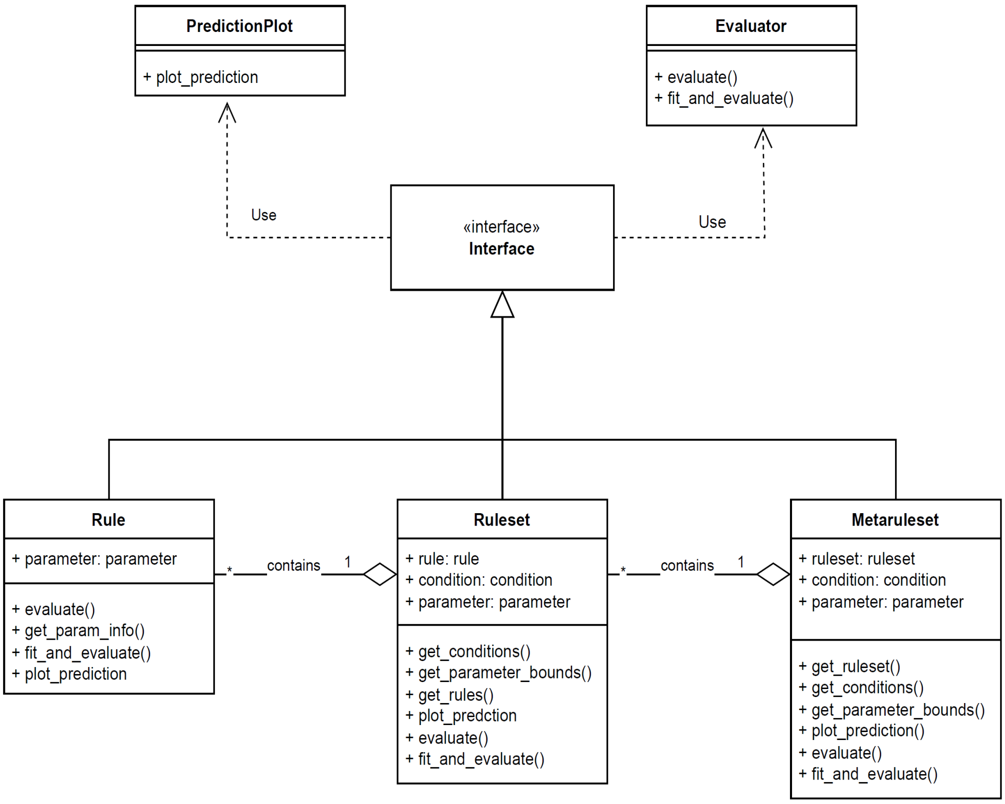
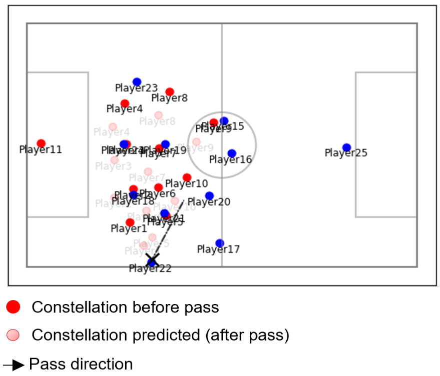
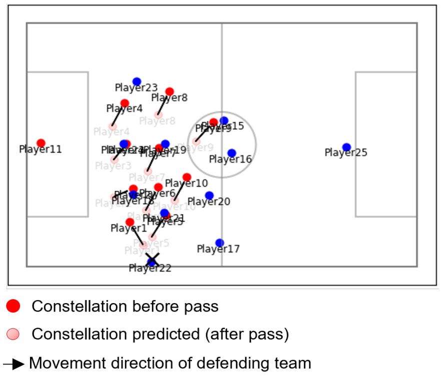


Figure 4.1.: UML model of our coding structure

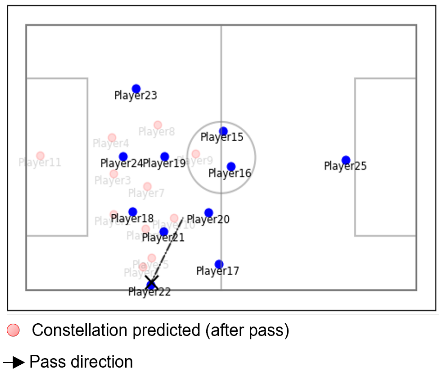
The framework for our prediction models is based on the approach of establishing rules, that describe some tactical player motion behavior in modern soccer. The objective of our models is to reduce a defined prediction error (chapter 5.1) and make the forecast of the movement of the defending team as accurate as possible. Besides, we evaluate different game constellations, to discover which of our models seems more applicable for specific scenarios. As our approach focuses mainly on the knowledge of domain experts, we require a high level of adaptability and abstraction, in order to ensure a continuous enhancement of our established models. In the following sections, we describe the structure of our different prediction models, with their underlying Rules and Rulesets.



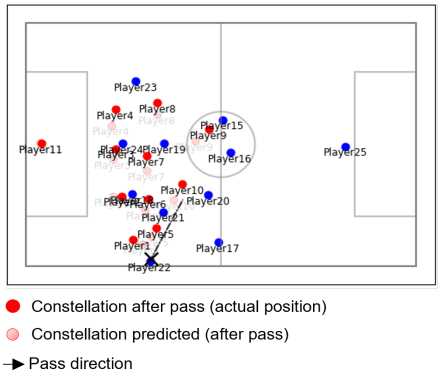
(a) Visualisation of the played pass (black arrow)



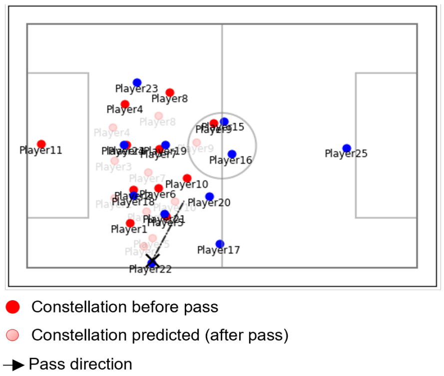
(b) Visualisation of the movement direction of each defender



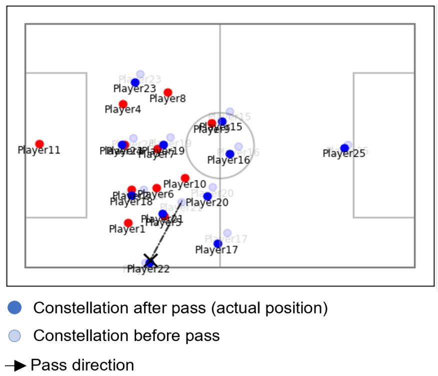
(c) Visualisation of only the predicted positions (transparent figures) of the defending team



(d) Comparison of predicted constellation (transparent) and actual position (filled) of the defending team



(e) Comparison of predicted constellation (transparent) and previous position (filled) of the defending team



(f) Visualisation of the previous (transparent) and the actual (filled) position of the attacking team

Figure 4.2.: Different game-visualisation features for the same game constellation

4.2. Graphical User Interface

As our prediction models is also usable by domain experts, in order to analyse and evaluate the prediction for specific game constellations, we design a Graphical User Interface (GUI). This GUI allows soccer professionals to establish and modify different Rulesets and Metarulesets, which can then be visualised and evaluated for specific game constellations. To enhance the tactical evaluation of different scenarios, we provide some game-visualisation features (Fig. 4.2) such as drawing arrows for the movement direction. The main reason for the GUI is the fast changing nature of modern soccer tactics and therefore the usability for soccer experts. We consider the prediction of defending motion in soccer as an on-going process, with a continuously changing tactical background. Thus adaptability and flexibility take an essential role. As an established model might be outdated after a few years, experts can use our approach to adjust or further enhance existing models and keep them up to date with the current tactical state of the art of modern soccer.

The design of the GUI is straightforward. Domain experts can select different implemented Rules, in order to build a prediction model. The quantity and order of the selected Rules can thus be adjusted individually. In a next step, different passes from the underlying data set are chosen, to predict specific game-constellations of the corresponding model. The evaluation of the established Ruleset, provides an analysis and visualization of the underlying tactical behavior, based on the concatenated Rules. In addition, several Rulesets can be combined to a Metaruleset, in order to analyze and evaluate a combination of different models. The framework offers an easy establishment of new prediction models for domain experts, which can be used for tactical analysis and game visualization. A description of the application through domain experts with example pictures is provided in Appendix.

4.3. Data Sets

For the evaluation and establishment of our prediction models, we use the public data set provided by Metrica Sports, which includes three independent and anonymous soccer games. The data has been collected through video tracking systems at a frequency of 25 Hz. To enhance data evaluation, the collected data is subdivided into different Game Events such as e.g. passes or dribblings. Passes are defined as the change of possession of the ball, between two players of the same team. As our prediction models are based on a pass event, we filtered out the events relevant for our models. This results in a total amount of 1074 passes for Game 1, 1191 passes for Game 2 and 1434 passes for Game 3 [Metrica-Sports, 2021].

4.4. Single-Rule Models

In order to provide an accurate prediction of the defending player constellation on the field, we implement three different single-rule models. These models consist of one specific rule and follow some trivial tactical approaches, to demonstrate some converse references for predicting player motion. They introduce some fundamental tactical team behavior and can later be combined and evaluated with further prediction models.

4.4.1. No Movement

As a first model we design a No Movement Model (NMM). In this scenario, the reaction of the defending team to a pass, is keeping their current position on the field, without movement in any direction. This inchoate approach rather serves as a baseline for comparing and evaluating further predictions. We define the rules of the model as the predicted constellation of the defending team is equal to the constellation before a pass (see Alg. 1).

Let *defense* be an array that contains all players of the defending team
Let *predicted constellation* be an empty array

foreach *defensePlayer* ϵ *defense* **do**
| **append** $defensePlayer_i$ to *predicted constellation*
end
return *predicted constellation*

Algorithm 1: Defined rules for No Movement Defending

In order to predict the constellation of the defending team after a pass, the NMM gets the current position (X-coordinate and Y-coordinate) of each player of the defending team. These positions are then set as the outcome of the prediction. Besides, the velocity (v) of each player on the field is set to zero. In Figure 4.3, the outcome of the prediction baseline after the pass is visualised. The transparent figures simulate the predicted outcome of the model, which is equal to the constellation before the pass. The filled figures demonstrate the real motion for each player to enhance visualisation and comparison.

The described model is applied as an important baseline, which we use in order to validate our implemented prediction models. It is used to demonstrate how accurate the movement prediction of another model is, compared to no movement of the players. For instance, if the NMM has a better prediction performance than other models, it indicates that the prediction from other models might be inaccurate.

4.4.2. Move Towards Ball

As a second single-rule model, we establish the Move Towards Ball Model (MTBM) model. It implicates a strict ball-oriented defending performance, where every defending player moves towards the position of the ball after a pass. This approach is chosen, as it might appear that in some specific game constellations, like e.g. after a corner kick, defenders orient themselves towards the ball position, in order to block a shot or to clear the ball

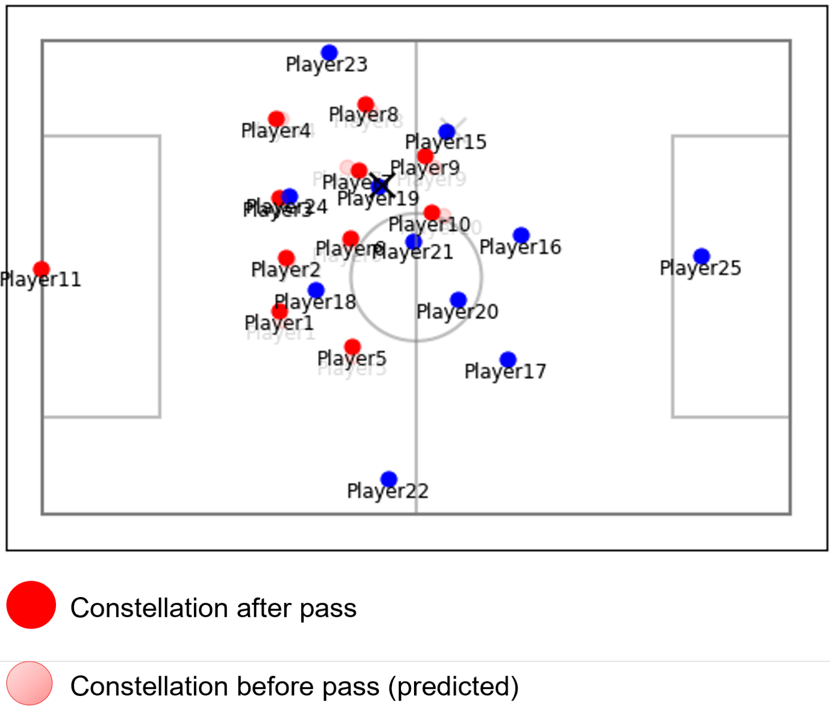


Figure 4.3.: Constellation plot of the 'No Movement Model' for the defending team (red figures). The transparent figures simulate the predicted constellation before the pass has been played. The filled figures demonstrate the real motion for each defender on the pitch. The transparent X marks the position of the ball before the pass, while the filled X simulates the ball position after the pass has been played.

out of the dangerous zone. Besides, it demonstrates some basic approaches for simulating pressing situations[e.V. (DFB), 2015]. The rule for the model is defined in algorithm 2.

The distance every defender covers on the pitch is computed by a given player velocity (v) and the time duration of a pass. In order to calculate the passing time, we get the game time at two different time frames. First, the game time at the moment the pass has been played (start time) and second, the match time when the ball reaches its target position (end time). The differential result is our passing time. For the velocity of each player we apply Bayesian Optimization, where the velocity is optimized for a current game constellation (see chapter 5.2). The mathematical formula for the frame distance of each player is described as followed:

Let *defense* be an array that contains all players of the defending team
Let *predicted constellation* be an empty array
Let *frameDistance* be the distance every defender covers within a frame
Let *movementAngle* be the direction every defender moves towards

foreach *defensePlayer* ϵ *defense* **do**
  get $frameDistance_i$
  get $movementAngle_i$
  calculate $playerPrediction_i$
  append $defensePlayer_i$ to *predicted constellation*
end
return *predicted constellation*

Algorithm 2: Defined rules for Move Towards Ball Defending

$$frameDistance_{player} = velocity_{player} * (time_{end} - time_{start}) \tag{4.1}$$

Besides the distance every defending player covers, we also need to calculate the movement direction for each individual defender. Therefore, we establish a right-angled triangle and used the angle between player position and target ball position as direction (see Figure 4.4 left-hand side). The distance between the defending player and the target ball position is defined as our Hypotenuse. We get our Adjacent and Opposite by subtracting the x-values and y-values of the particular defender from the x-values and y-values of the target ball position, declared as xDistance and yDistance:

$$xDistance_{playerToBall} = x_{targetBall} - x_{player} \tag{4.2}$$

$$yDistance_{playerToBall} = y_{targetBall} - y_{player} \tag{4.3}$$

With the Pythagorean theorem [Mathwarehouse.com,], we calculate the exact distance between each individual defending player and the target position of the ball, defined as our Hypotenuse.

$$distance_{playerToBall} = \sqrt{xDistance^2_{playerToBall} + yDistance^2_{playerToBall}} \tag{4.4}$$

We then apply the formula for right-angled triangles [Rudolph, 2018] in order to obtain the movement angle ($\boldsymbol{\alpha}$) towards the ball for each sole defending player.

$$\cos(\boldsymbol{\alpha}) = \frac{yDistance_{playerToBall}}{distance_{playerToBall}} \tag{4.5}$$

With the calculated movement angle ($\boldsymbol{\alpha}$) and the frame distance, we establish a new right-angled triangle, which defines distance and direction for the defending players movements (see Figure 4.4 right-hand side). The Adjacent and the Opposite are defined

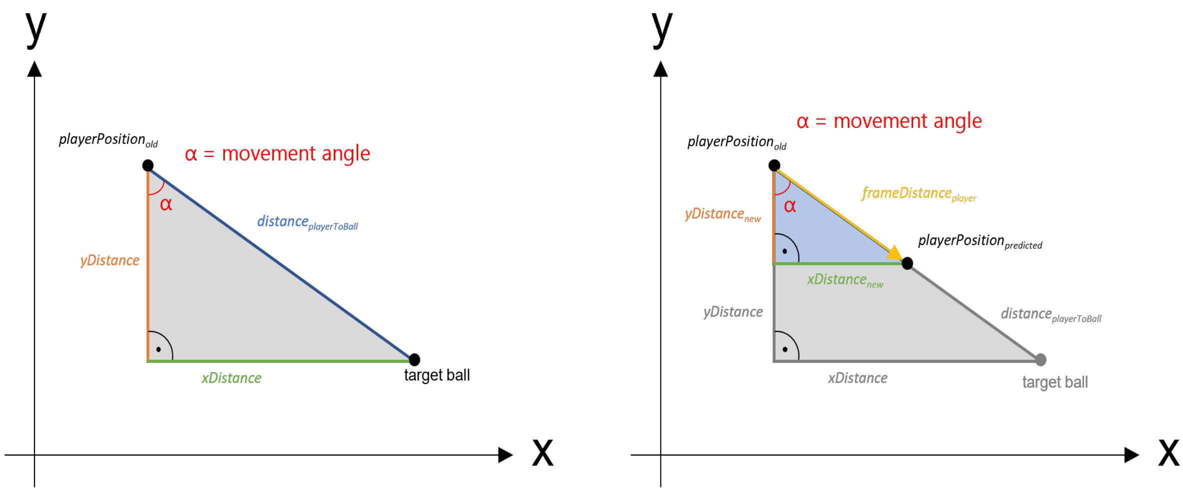


Figure 4.4.: Triangles in order to calculate the predicted position for each individual player. On the left-hand side we see the established right-angled triangle from each player position to the target ball position, which is used to calculate the appropriate movement angle. On the right-hand side we deviated a new triangle based on the distance every player covers within a passing frame (frameDistance) and the movement angle. The new xDistance and yDistance are then used in order to calculate the predicted position for each defender.

as the distances, a particular defender moves on the X-axis and Y-axis (see Figure 4.4) and added to the corresponding X-coordinates and Y-coordinates of each defender, in order to predict the new positions for the defending team [Rudolph, 2018].

Our Move Towards Ball Model is based on a simple tactical approach. Nonetheless, it allows the prediction of specific game constellations and serves as a first indicator of ball-oriented defending. Figure 4.5 shows a visualisation of the prediction outcome of the model at a given game constellation, where the prediction for every player of the defending team is towards the position of the ball after a pass has been played. As it is a single-rule model, the rule can easily be concatenated with other rules, to form a ruleset for a more detailed prediction model (see chapter 4.5).

4.4.3. Pass Angle

As a third single-rule model, we design a Pass Angle Model (PAM). This model has a rather ball-oriented focus with regard to zone defending behavior. It implicates, that every defending player on the field moves in the same direction, as the pass that has been played. This model has a similar approach as our introduced MTBM (see chapter 4.4.2), but slightly deviate regarding a team's nominal formation. While in our MTBM, every

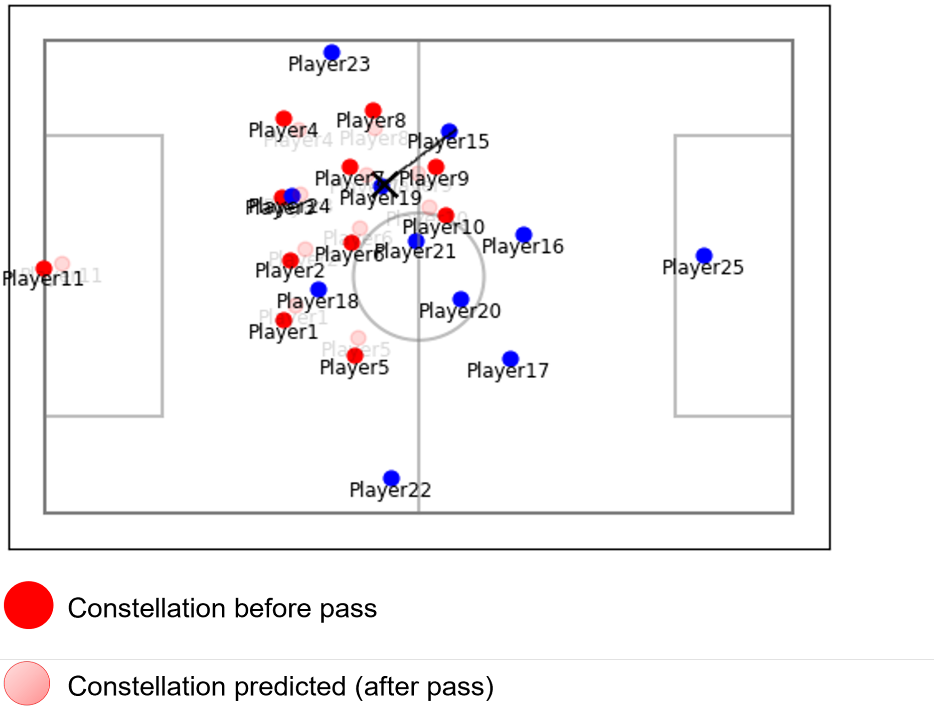


Figure 4.5.: Constellation plot of the 'Move Towards Ball Model' for the defending team (red figures). The transparent figures forecast, where the player of the defending team will move, after a pass has been played. The filled figures simulate the current game constellation and the actual position of the defenders. The black line visualises the pass that has been played, with the filled X as the ball position after the pass has been played.

player focuses on the current position of the ball, regardless of any team motion behavior, the PAM keeps the defending team's current formation. This makes it more applicable for game constellations, where the ball is in a neutral part of the pitch and where the defending team stays in its initial team formation[Spielverlagerung.de, 2013]. We define the rule for our PAM in alg. 3.

To calculate the distance each defender covers on the field, defined as frame distance, we use the same approach from the previous chapter (see formula 4.1). We apply Bayesian Optimization, in order to compute the best player velocity parameter applicable for our model (see chapter 5.2). To get the movement direction for every defender, which is in this case the same direction as the ball movement, we need to calculate the passing angle. Therefore, we follow the same structure from chapter 4.4.2, where we established a

Let *defense* be an array that contains all players of the defending team
Let *predicted constellation* be an empty array
Let *frameDistance* be the distance every defender covers within a frame
Let *passingAngle* be the direction the ball moves towards

foreach *defensePlayer* ϵ *defense* **do**
  get $frameDistance_i$
  get $passingAngle_i$
  calculate $playerPrediction_i$
  append $defensePlayer_i$ to *predicted constellation*
end
return *predicted constellation*

Algorithm 3: Defined rules for Passing Angle Defending

right-angled triangle. The direction is defined as our passing angle, which is the angle between the position of the ball after the pass (target ball) and the position of the ball before the pass has been played (previous ball position). Our Hypotenuse is defined as the distance between target ball and previous ball position. We deviate the corresponding Adjacent and Opposite by subtracting the x-values and y-values of the previous ball position from the x-values and y-values of the target ball position, defined as XDistance and yDistance [Rudolph, 2018]:

$$xDistance_{ballToBall} = x_{targetBall} - x_{previousBall} \tag{4.6}$$

$$yDistance_{ballToBall} = y_{targetBall} - y_{previousBall} \tag{4.7}$$

With the Pythagorean theorem, we deviate the distance between the previous position of the ball and the target position of the ball:

$$distance_{ballToBall} = \sqrt{xDistance^2_{ballToBall} + yDistance^2_{ballToBall}} \tag{4.8}$$

We define the angle of the passing direction as ($\boldsymbol{\alpha}$), which we calculate by applying the cosine formula for right-angled triangles.

$$\cos(\boldsymbol{\alpha}) = \frac{yDistance_{ballToBall}}{distance_{ballToBall}} \tag{4.9}$$

We apply the deviated passing angle ($\boldsymbol{\alpha}$) to every defender on the field, in order to get their movement direction. The distance, covered by each player of the defending team is calculated by the velocity of each player and the passing time (see formula 4.1). In Figure 4.6 (right-hand side) the new established right-angled triangle for the passing angle is visualised. Adjacent and Opposite are the distances, defenders move on the X-axis and Y-axis on the field (see figure 4.6). They are then added to the current position of each defender, in order to establish the new predicted positions of the defending team [Rudolph, 2018].

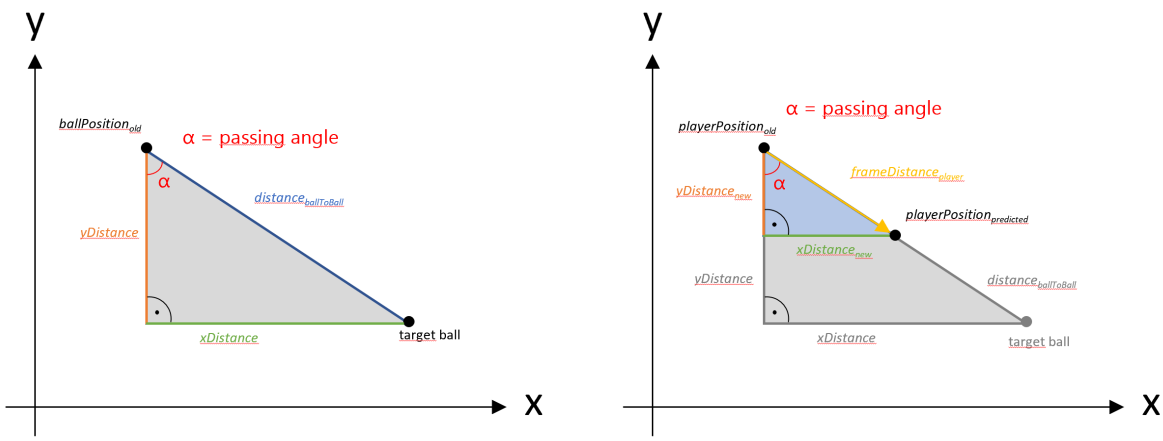


Figure 4.6.: Triangles in order to calculate the predicted position for each individual player. On the left-hand side we see the established right-angled triangle from the ball position before the pass to the target ball position, which is used to calculate the appropriate passing angle. On the right-hand side we deviated a new triangle based on the distance every player covers within a passing frame (frameDistance) and the passing angle. The new xDistance and yDistance are then used in order to calculate the predicted position for each defender.

The established PAM deals as a first indicator of ball-oriented defending with some regards to zone defending. Even though the model is based on a single-rule approach, it enables the prediction of some specific game constellations. In Figure 4.7 a visualisation of the prediction outcome is demonstrated. The defending players (red team) move in the same direction, as the pass that has been played (black line). Although the model is limited to only one rule, it provides some fundamental approaches of tactical behavior and can be combined with other rules, in order to form a tactical more sophisticated prediction model[Spielverlagerung.de, 2013].

4.5. Complex-Rule Models

Predicting the positions of soccer players might appear quite challenging, as there exist various factors that can influence a participant's motion, such as the physical condition or the complex individual decision-making of each player[Rein and Memmert, 2016]. Nonetheless, when defending an opponent's attack, professional soccer players follow some fundamental rules, in order to maximize the output of their defending actions. We therefore use some of these rules and established different rulesets to predict player

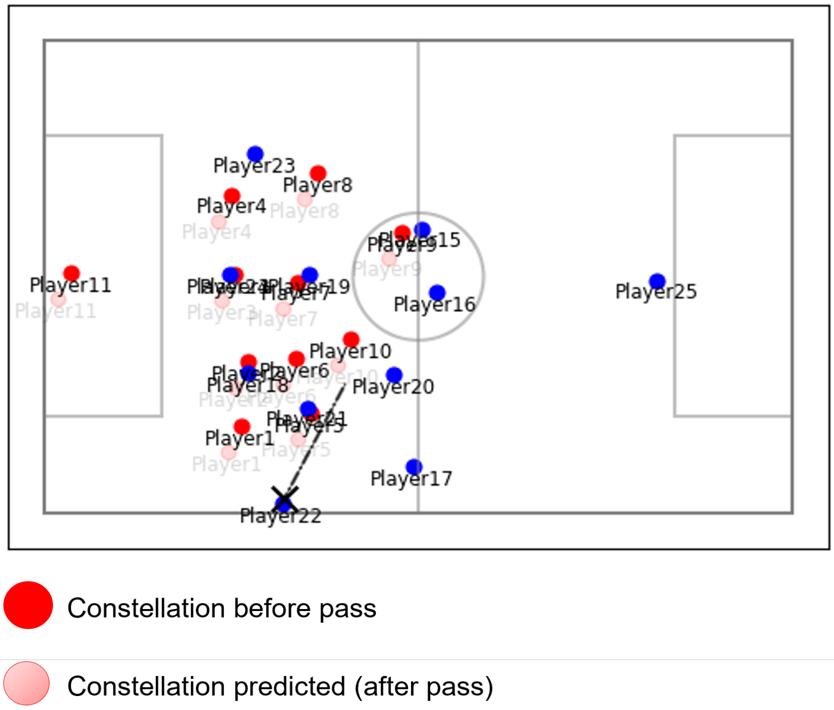


Figure 4.7.: Constellation plot of the 'Pass Angle Model' for the defending team (red figures). The transparent figures forecast, where the player of the defending team will move, after a pass has been played. The filled figures simulate the current game constellation and the actual position of the defenders. The black line visualises the pass that has been played, with the filled X as the ball position after the pass has been played.

positions, where we train our models with different principles of tactical behavior in soccer [Teoldo, 2021].

Complex-rule models are methods, that focus on domain experts' knowledge in order to establish hypothesis and extract representative values for a valuable evaluation [Liu et al., 2014]. The rules we use, follow the state of the art tactical behavior [Teoldo, 2021] and have been verified by two independent domain experts (see chapter 5.3). As real-world movement of players is rather complex and might deviate based on different in-game situations, our implemented models require a high level of flexibility and adaptability. To satisfy these requirements, every model consists of a set of rules (ruleset). A ruleset can easily be adjusted by adding new rules or removing old rules from the set.

4.5.1. Man Marking

Our first introduced prediction model follows the classic tactical man marking behavior. In general, this way of tactical defending assigns an opponent player to every player of the defending team [Teoldo, 2021]. While a 'strict man marking' approach assigns each player of the defending team to one specific attacker, whom they mark for the whole game duration, a 'loose man marking' approach is more flexible. In this play, the assignment of attackers to defenders can vary and they can change their opponent player e.g. due to change of positions of two different players of the attacking team [Spielverlagerung.de, 2012]. As our Man Marking Model (MMM) is designed for different in-game constellations during the game, we rather focus on the more flexible projection of a 'loose man marking', where the assignment of attackers to defenders is not static for the whole game duration and can vary in different game constellations. thus, we apply different rules to our model, defined as the Ruleset in algorithm 4.

Let *defense* be an array that contains all players of the defending team
Let *offense* be an array that contains all players of the attacking team
Let *distance* be the distance between defense and offense player
Let *ballDistance* be the distance between ball and defense players
Let MarkedPlayers be an empty array

sort *defense* by *mostDefensivePlayer*;
foreach *defensePlayer* ϵ *defense* **do**
  find *offensePlayer* ϵ *offense* **where** *distance* **is** *minimal*
  if *offensePlayer is not contained in MarkedPlayers* **then**
    defensePlayer **moves towards** *offensePlayer*;
    Append *offensePlayer* to *MarkedPlayers*;
  end
  if *ballDistance is minimal* ***AND*** *ballDistance* $\leq$ *15 m* **then**
    defensePlayer **moves towards** *ball*;
  end
  else if *distance* $\leq$ *2m* **then**
    defensePlayer **moves towards** *offensePlayer*;
  end
end

Algorithm 4: Defined rules for Man Marking Defending

In general the Ruleset of our Man Marking Model consists of a combination of four different rules. The first rule "Player Marking" establishes the general marking for each attacking player to a defender, for a given game constellation. Therefore, starting at the position of the most defensive player, we assign the closest player of the opponent team that has not yet been marked to this player. This implicates, that every defender marks his closest opponent, if this opponent is not yet marked by another defender. After

the pass has been played, defenders move towards their assigned opponents. A special role goes to the goalkeeper. As goalkeepers are no part of a man marking behavior and therefore not assigned to a any player, we apply the 'Don't Move' rule, where we assume, that goalkeepers do not move after a pass has been played. Besides, we limit the distance to the opponent player to a maximum of 25 meters. In case there is no available opponent player within this radius, our 'Player Marking' rule assumes that the defender does not move.
As a third rule, we establish the 'Go Towards Ball' rule for the defender, that is closest to the ball. As there might occur an interference between the marking of the player and the player that is closest to the ball, we assume that the 'Go Towards Ball' rule overrules the player mapping. Therefore, the player that is closest to the ball moves towards the ball, regardless his initial opponent marking. The last implemented rule is the 'Direct Opponent' rule. This rule implies, that if the distance of a defender's closest opponent is less that two meters, the player moves towards this attacking player, regardless of the previous player mapping. The concatenation of described rules, form our 'Man Marking Model'.

Figure 4.5.1 demonstrates a graphical visualisation of our MMM. Every defender moves towards the player they are assigned to. Players that do not move e.g. Player 8, have no corresponding player of the opponent team. This means that there is no unassigned player within the range of 25 meters. The closest player to the ball (Player 10) moves towards the current position of the ball.

Our MMM is considered as a more realistic prediction model that is applicable to rather many game constellations. Due to its complex-rule nature, it can easily be further enhanced through domain experts, by adding additional rules. Even though the tactical approach of man marking seems rather outdated in modern soccer, there still exist several game constellations, where a man marking approach seems more promising in defending actions [Spielverlagerung.de, 2012]. Besides, the MMM enables us to compare player movement with other prediction models and evaluate how applicable different tactical approaches for predicting defender motion are. Domain experts can use our model for simulating and predicting game constellations based on player marking defending.

4.5.2. Zone Defending

Our Zone Defending Model (ZDM) focuses on the tactical approach of zone defense. While our MMM assigns opponent players to every defender, the ZDM implicates a more ball- and zone-oriented defending measure, where it establishes several defending lines, to cover dangerous zones on the pitch and maximise the output of defending actions. In general, when a team defends with zone defense, the reference for each defender are his teammates. The defending team operates as a compact block, usually defined as an initial game formation. Within this block, the defending team is further divided into defending lines. Each line has a ball-oriented movement, with regard to the distance between each player in the line and the distance to other defending lines. Characteristics of a zone defense approach are noticeable defense lines, that shift all together as one defending block, based on the ball movement [Spielverlagerung.de, 2013]. For the implementation

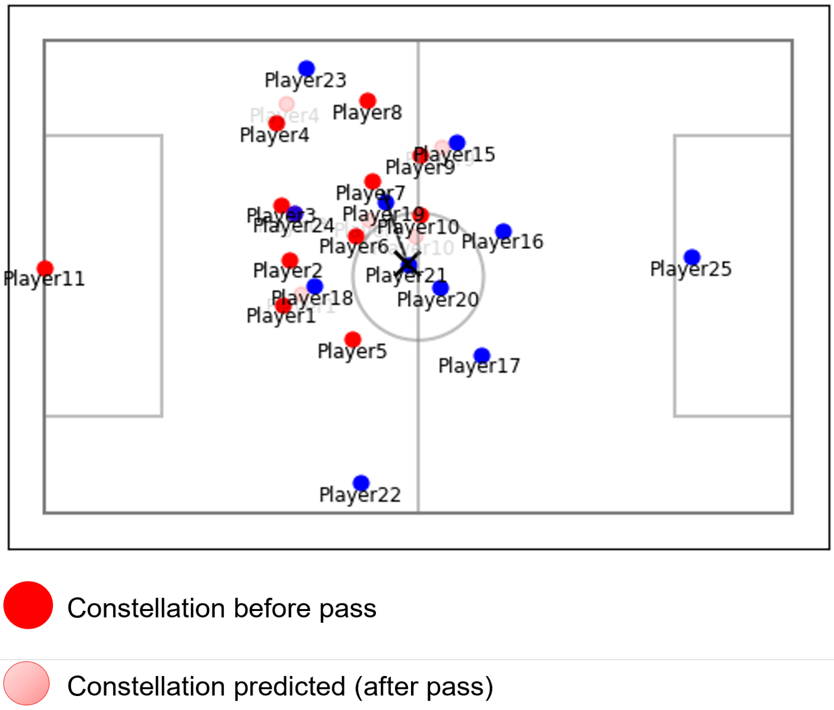


Figure 4.8.: Constellation plot of the 'Man Marking Model' for the defending team (red figures). The transparent figures forecast, where the player of the defending team will move, after a pass has been played. The filled figures simulate the current game constellation and the actual position of the defenders. The black line visualises the pass that has been played, with the filled X as the ball position after the pass has been played.

of our ZDM, we define different rules in alg. 5.

In order to provide a realistic prediction model based on the tactical approach of zone defense, we divide the defending team into different lines (see figure 4.9). The amount of players for each line can be set manually and is usually based on a team's initial game formation. For each defending line, we assign an orientation player, which is the most defensive player of each line. This orientation player, shifts his position in a ball-oriented movement, in the same angle as the pass that has been played. The other players of each chain adapt their positions, based on the distance to their teammates. Target position for each line-defender are the same y-coordinates as their corresponding orientation player, in order to form on straight defense line. The x-coordinates for the defenders are based on the gap to each teammate. For optimizing the distance between each player in a line, we apply Bayesian Optimization (chapter 5.2), where we set the distance for each defender

Let *defense* be an array that contains all players of the
defending team
Let *defenseBlock* be an array that contains all defense players of the
defending team
Let *midfieldBlock* be an array that contains all midfield players of the
defending team
Let *attackingBlock* be an array that contains all attacking players of the
defending team
Let *ballDistance* be the distance of the ball to all defending players
Let *lineDistance* be the distance between defenders within a lineBlock

foreach *lineBlock* **do**
    find $orientationPlayer_i$ **where** $orientationPlayer_i$ **is** *most defensive*
    get *newPosition* **from** *passingAngle*
    $orientationPlayer_i$ **moves towards** *newPosition*
    foreach *player* ϵ *lineBlock* **do**
        get *newPosition* **where** *lineDistance* is 10m
        $player_i$ **moves towards** *newPosition*
    end
end

Algorithm 5: Defined rules for Zone Defending

to a range from 5 to 15 meters.

Besides the establishment and shifting of different defense lines, we assume that the player closest to the current ball position, moves towards the ball, regardless his in-line position. This simulates a more ball-oriented approach and implicates direct pressure on the player in possession of the ball. As the goalkeeper is not part of a line of the defense block, we assume that he does not move and therefore apply the 'Don't Move' rule to the goalkeeper of the defending team.
Figure 4.10 demonstrates a graphical visualisation of our ZDM. Every player of the defending team is assigned to a line and moves towards his corresponding in-line position. The defense chains have been established manually (4-4-2 formation) but can easily be converted by domain experts to any formation. The distances between each player in a defending chain is optimized through Bayesian Optimization. We set the range from 5 meters to 15 meters. The player that is closest to the ball (Player 5) moves towards the ball, regardless his corresponding in-line position. This Model also assumes that goalkeepers do not move.

With our "Zone Defending Model" we provide a prediction model, that is based on a modern tactical soccer approach [Spielverlagerung.de, 2013]. It is applicable to any game constellation, and enables domain experts to easily add new rules, due to its complex-rule nature. The ZDM can be compared with other prediction models in order to evaluate

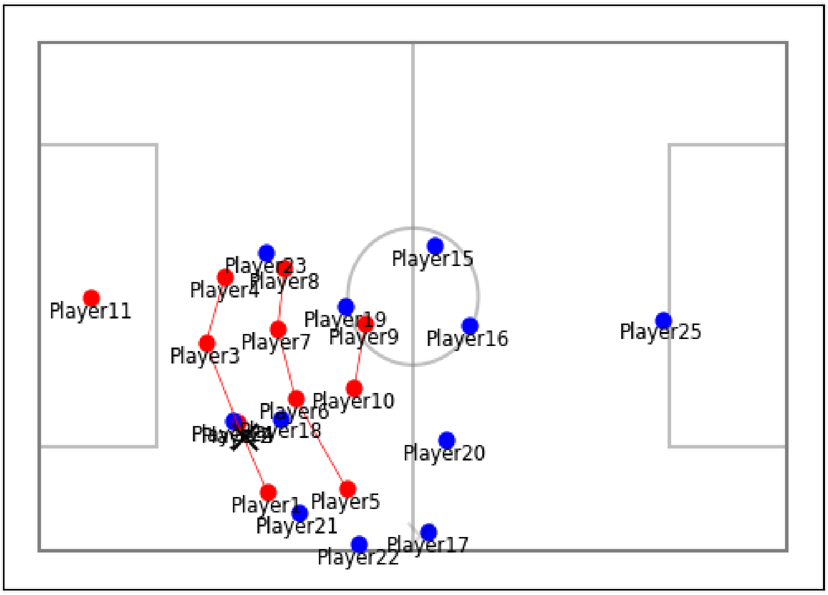


Figure 4.9.: Visualisation of the established lines for the defending team (red figures). The assigned team formation is a 4-4-2. The most defensive players of each line (Player 3, Player 7 and Player 10) are defined as orientation players and move in passing direction. The other players in each chain adjust their positions based on the movement of the orientation player. The red lines visualise the distances between the defenders of each defending line.

accuracy and also analyse in which game constellation the zone defending approach seems rather promising. It can be used by domain experts for predicting and simulating the defending behavior for specific game scenarios based on its underlying tactical conduct.

4.5.3. Combination Model

Soccer has evolved quite fast over the last decades. Beside of a better physical condition of each individual player, there has also been an improvement of tactical team behavior. Nowadays, professional soccer teams are capable of playing different tactical formations, which might change several times during a match. Furthermore, specific game constellations and attacking patterns require the defending team to apply different tactical measures. For instance, while the team in possession of the ball keeps the ball in their

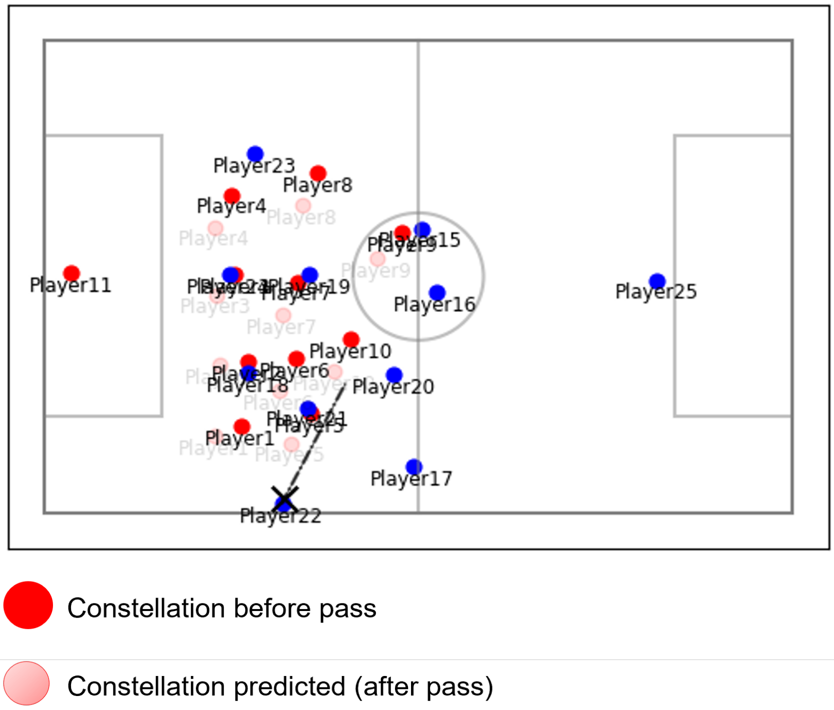


Figure 4.10.: Constellation plot of the 'Zone Defending Model' for the defending team (red figures). The transparent figures forecast, where the player of the defending team will move, after a pass has been played. The filled figures simulate the current game constellation and the actual position of the defenders. The black line visualises the pass that has been played, with the filled X as the ball position after the pass has been played.

own half, the approach of zone defense might seem more appropriate for the defending team. Nonetheless, with entering the 'Dangerous Zone' and decreasing goal distance, a defending team might rotate to a more man-oriented defending approach [Teoldo, 2021]. In order to provide a precise prediction, our models require a high amount of flexibility and adaptability. We therefore establish a 'Combination Model'. This approach provides a higher level of abstraction, which shifts our Rulesets to a meta level (Metaruleset). It enables the combination of different Rulesets with each other.

To provide a realistic prediction by combining different Rulesets with each other, we subdivide the pitch into three parts ('Attacking Zone', 'Midfield Zone' and 'Defending Zone'). To optimize the borders for each zone, we used Bayesian Optimization (chapter 5.2) and set the range for the distance to the goal between 15 to 35 meters. For each third, a different Ruleset can be applied (see figure 4.11). As we consider the zone defense approach more applicable when the ball is in a 'neutral area', we assign our ZDM to the 'Midfield Zone' and the 'Defending Zone'. When the distance of the ball to the goal of the defending team decreases, a more man-oriented defending measure seems rather promising. We therefore, allocate the MMM to the 'Attacking Zone'.

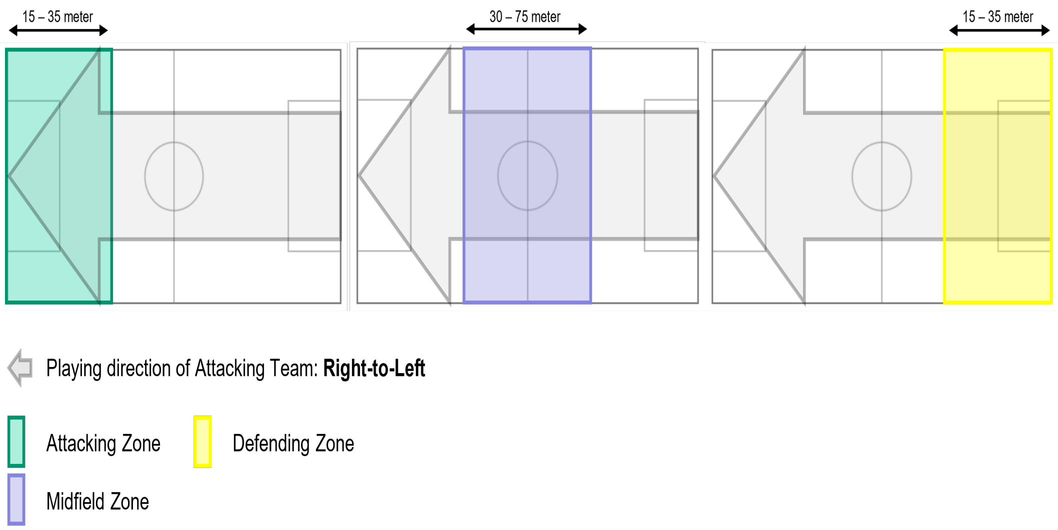


Figure 4.11.: Visualisation of the three different zones for our Metaruleset.

In figure 4.12 we visualise the approach of our Combination Model (CM). We contrast two different constellations with each other. On the left-hand side, the current ball position is in one of the 'neutral' parts of the pitch, which implicates the utilization of the ZDM. The display of the right-hand side line-up applies the MMM to the current in-game constellation, as the ball enters the 'Attacking Zone' and requires a different defending approach.

The intention of our Combination model is to satisfy the fast changing nature of modern soccer tactics. Through the establishment of a Metaruleset, we are able to combine several Rulesets with each other and apply them as one prediction model. This provides a more dynamic projection for predicting player motion in soccer. Besides, domain experts can use this approach, to establish models on a higher level of abstraction, where they can analyse and evaluate different Rulesets for different zones on the pitch.

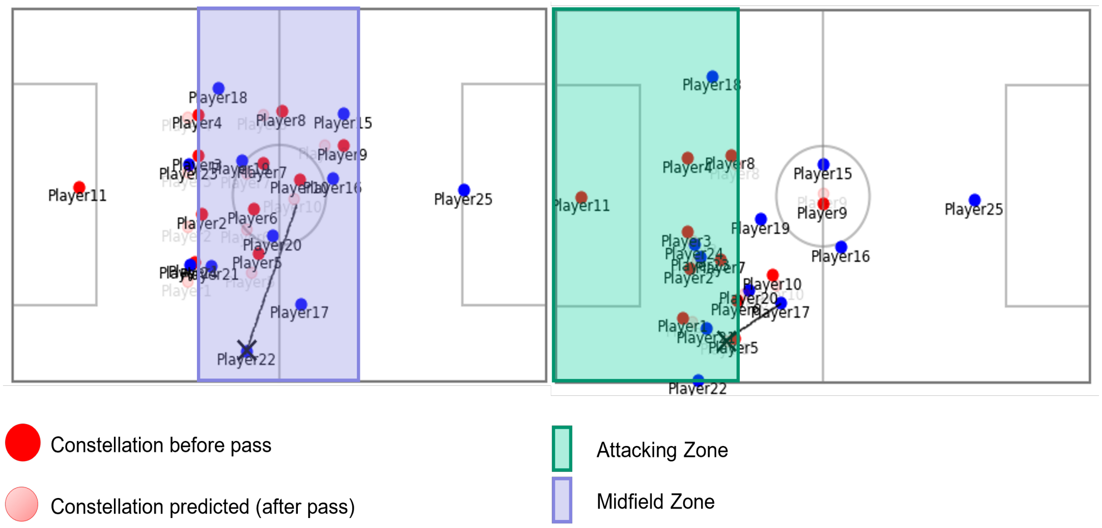


Figure 4.12.: Visualisation of our 'Combination Model'. On the left-hand side, the current ball position is located in the Midfield Zone, whereby the ZDM is applied. On the right-hand side, the ball enters the Attacking Zone, why defenders change their defending behavior according to the MMM.

5. Prediction Justifcation and Evaluation

In this section, we predict the defending player motion through our established models and evaluate the occurring result. Therefore, in order to provide a valuable benchmark for comparison, we define a prediction error ($\boldsymbol{\theta}$) for every model (section 5.1), which we use as prediction accuracy. In section 5.2 we explain the approach, how variable parameters of our models are optimized. A justification from different independent domain experts is provided in section 5.3 to ensure plausibility of the implemented models. In the subsequent sections, the evaluation of each prediction model is executed, where we analyze the prediction error and identify if there exist any recurring defending patterns. In section 5.4.4 we compare and analyze the results of our prediction models, where also assumptions about tactical defending approaches in soccer are generated.

5.1. Validation Model

In the following we describe our validation model, in order to provide a reliable measurement of the accuracy of the deployed prediction models. We establish a *prediction error* ($\boldsymbol{\theta}$), where we compare the actual motion of the defending players with the predicted movement provided by our models (alg. 5.1).

$$\boldsymbol{\theta}_{(\boldsymbol{c})} = \frac{\sum_d^{11} \sqrt{(x_{d(pr)} - x_{d(a)})^2 + (y_{d(pr)} - y_{d(a)})^2}}{\frac{\sum_d^{11} P_d}{\boldsymbol{\lambda_c}}} \tag{5.1}$$

Our defined *prediction error* ($\boldsymbol{\theta}$) is computed for a given game constellation *(c)*. To compare predicted with actual positions, we calculate the distance for each defender *(d)*, between his predicted and his actual position. thus, we apply the Pythagorean Theorem [Mathwarehouse.com,] with the x-coordinates (y-coordinates) of the predicted player positions $x_{d(pr)}$ respectively $(y_{d(pr)})$ and the actual positions $x_{d(a)}$ respectively $(y_{d(a)})$. In order to form the weighted average, we divide the quantum of the distances between predicted and actual position for each defender, through the total amount of defending players P_d, also taking into consideration the *time frame* ($\boldsymbol{\lambda}$) of the played pass. We involve the duration of a pass, as shorter passes might implicate lower prediction errors than a longer passes, due to a shorter movement time for each defender and therefore resultant a shorter deviation of the predicted and actual positions. In algorithm 5.2 we define the *passing time* $\boldsymbol{\lambda}$ at a given constellation *(c)* by subtracting the *game time* ($\boldsymbol{\sigma}$)

before a pass has been played ($\boldsymbol{\sigma_{(start)}}$) from the game time after the ball has reached its current position ($\boldsymbol{\sigma_{(end)}}$).

$$\boldsymbol{\lambda_{(c)}} = \boldsymbol{\sigma_{(end)}} - \boldsymbol{\sigma_{(start)}} (5.2)$$

The approach for validating accuracy of our prediction models is straight forward. We consider the deviation of the position of each individual defender and divide it through the total amount of defenders, in order to establish an average deviation for each constellation. To guarantee that the occurring error for each pass is weighted equal, we relativize each pass with its underlying time duration. A lower prediction error $\boldsymbol{\theta}$ results in a higher accuracy for a prediction. This approach allows us to provide a valuable comparison of different prediction models. Besides, we can evaluate the prediction error for a single game constellation or a total match. In case we evaluate a model for a whole game, we establish an average occurring prediction error, based on the error of each constellation. Due to the possibility of evaluating specific game constellations, we are also capable of estimating, which models seems rather promising for a specific game scenario.

5.2. Parameter Optimization

As mentioned in the previous chapters, our models are based on some Rules, that include variable parameters, such as e.g. a velocity (v) of each participating player. As it might appear quite difficult in real-world scenarios to determine these parameters within defined rules, so that they are applicable for every in-game scenario, we optimize the corresponding variable of each Rule for every predicted constellation. Therefore we apply Bayesian Optimization (BO), as it seems rather applicable for scenarios, where the sampling of the optimized function is an expensive aspiration [Nogueira, 2014].

In general, BO is used to optimize the output of a specific objective function [Nogueira, 2014]. This includes uncertainty of variable parameters that influence the outcome of the objective algorithm. Therefore, Gaussian Processes (GPs) are applied, through building a posterior distribution of functions, that maximize the output of the optimized algorithm [Garnett and Garnett, 2023]. The posterior distribution improves with a growing number of observations, as a Gaussian Process balances its underlying exploration and exploitation of the parameters, considering the knowledge of the function that is optimized. When executing GPs, new samples are fitted to the previously explored ones, helping to improve exploitation of the optimized function [Nogueira, 2014].

We apply BO in order to minimize our established *prediction error* ($\boldsymbol{\theta}$) for each constellation and thus enhance accuracy of our prediction models. We define our algorithm (alg. 5.1) as the objective function, where we want to optimize the corresponding output for every game scenario. We then set the bounds for the associated parameters that are fitted. For instance, for our prediction models, we define the boundaries for a player's velocity (v) from 0 to 4 m/s. The number of executions of BO (n_iter) is declared to 125 steps. As a higher amount of steps ensures a higher probability of finding a valuable maximum, we consider 125 steps as appropriate. Besides, we perform a number of 75 steps of random exploration (init_points) in order to diversify the exploration space. The bounds for each model and its corresponding parameters can be reviewed in table 5.1.

Parameter	**Units**	**Bounds**	**Model**
Rules & Rulesets			
Player velocity (v)	[m/s]	0 - 4	MMM, MTBM, PAM, ZDM
Distance to ball	[m]	0 - 25	MMM, ZDM,
Line distance	[m]	5 - 15	ZDM
Distance direct opponent	[m]	0 - 2	MMM
Distance marked player	[m]	0 - 25	MMM
Metarulesets			
Player velocity (v)	[m/s]	0 - 4	CM
Attacking Zone	[m]	15 - 35	CM
Midfield Zone	[m]	30 - 75	CM
Defending Zone	[m]	15 - 35	CM

Table 5.1.: Parameter bounds for optimization

5.3. Justification

In order to ensure, that our models follow some realistic player motion behavior and that the established rules are based on real-world tactical approaches, we consulted two independent domain experts. We conducted interviews with each expert, where we explained the behavior of our models and the tactical background they are based on. Both experts confirmed, that the tactical approach of our established models seems realistic and that they are applicable for real-world situations.

As a first domain expert, we considered Otto J. Band from the Otto-Friedrich University of Bamberg. Otto J. Band officiated for more than 20 years as Head-Coach of the university soccer team. With the selection of the University of Bamberg soccer team, Otto J. Band played several times at the 'Bavarian University Championship'. Besides, he traveled as Head-Coach to different countries all over the world, where he among other things played against the Student National Team of Argentina [inFranken.de, 2018] and several partner universities [Tim Kipphan, 2022].

As a second domain expert, we consulted Paul Ehmann. Paul Ehmann currently works as Head-Coach for the Under 15 youth team of the German Bundesliga club TSG 1899 Hoffenheim. Paul Ehmann started his career as player for the German teams 1.FC Kaiserslautern and TSG 1899 Hoffenheim. As a player, he played under Head-Coaches like David Wagner and Tayfun Korkut, which both later became coaches of German Bundesliga Clubs. After his career as midfield-player, he began working as assistant coach for FSV Frankfurt and later TSG 1899 Hoffenheim, where he became Head-Coach of the Under 15 youth team [tsg hoffenheim.de, 2021].

Both of our selected domain experts, have several years of coaching experience on a high-class level. We used this experience as proof, that the tactical base of models is consistent with the state of the art tactics in soccer. We provided valuable prediction models, that can be used to forecast and evaluate different game constellations, based on the principles of diverse tactical attitudes.

5.4. Evaluation

In the following, we evaluate the accuracy of each of our established models, based on our defined prediction error (chapter 5.1). We consider three different matches, where we analyse each game in particular, as well as a cumulative prediction error for the whole data set. As different teams in soccer might follow different tactical approaches, we evaluate both of the participating teams (home and away team). This allows us to analyse, if some of our models are more applicable to a specific team and the underlying tactical approach of this particular team. Besides we compute a prediction for both teams together, to provide general information about the deployed prediction model. In the end we compare the prediction results of the models with each other, to identify which approach contains the minimum prediction error.

Due to parameter optimization, there exists a slightly deviation in the outcome of our *prediction error* ($\boldsymbol{\theta}$). We cover this, by adjusting the steps within Bayesion Optimization (BO) (chapter 5.2) to 125 (n_iter) and 75 (init_points), which limits the deviation of $\boldsymbol{\theta}$ to a minimum for each run of our prediction models. For evaluation procedure, we set a fixed seed number generator, in order to ensure that we get the same samples for every run of our models and thus the prediction error stays the same for each individual model.

In order to analyze, in which area of the field a model seems most applicable, we evaluate the best prediction scores for each game. We define an individual threshold for the occurring prediction error, where we evaluate each pass, that is underneath that threshold. We demonstrate our identified 'best passes', through a visualization of the position of the ball after a pass. Since the visualized playing direction for the defending team is from right-to-left, we reverse the passes where the direction is left-to-right, in order to display them in the same figures. Through the distribution of the ball positions, we identify patterns on the pitch, where we evaluate, in which area of the field the corresponding model has the highest accuracy.

5.4.1. Single-Rule Models

In this section, we evaluate the prediction result of our single-rule models. We consider the No Movement Model (NMM) as our baseline, which we use as comparison to other prediction models. The Move Towards Ball Model (MTBM) and Pass Angle Model (PAM) are first indicators, for predicting player motion based on rules that follow some single-rule tactical behavior. The results are demonstrated in the following sections.

No Movement Model

As a fist step we analyse our baseline model, where defenders keep their current position and do not move after a pass. The results of the occurring *prediction error* ($\boldsymbol{\theta}$) can be reviewed in table 5.2. A graphical representation is provided in figure 5.1. The Total

	Game 1	Game 2	Game 3	Total
$\boldsymbol{\theta_{homeTeam}}$	2.830	2.626	2.642	2.699
$\boldsymbol{\theta_{awayTeam}}$	2.489	2.507	2.522	2.506
$\boldsymbol{\theta_{allTeams}}$	2.604	2.604	2.581	2.596

Table 5.2.: Prediction errors ($\boldsymbol{\theta}$) of our No Movement Model (NMM)

prediction error of the NMM results in ∼2.597 for all games. For the away team the Total $\boldsymbol{\theta_{awayTeam}}$ is slightly lower, while the home team has a higher prediction error. A similarity can be identified for Game 1, Game 2 and Game 3. This might implicate, that the away team follows a tactical approach, that is less running-intensive and therefore requires less adjustment of the positions. Considering $\boldsymbol{\theta_{allTeams}}$, we identify a similar prediction error for Game 1 and Game 2, that is above the Total $\boldsymbol{\theta_{allTeams}}$, while $\boldsymbol{\theta_{allTeams}}$ of Game 3 is slightly lower. We can assume, that the home team in general, and for Game 1 in particular, follows a rather running-intensive tactical background, that requires a lot of positioning adjustment of all defenders.

In figure 5.2 we visualized the distribution of the ball positions for the 'best passes' of the model, where the prediction error $\boldsymbol{\theta}$ of our NMM is less than the predefined threshold of 1.3 for each game. For Game 1, we identify 74 passes for the home team and 62 passes for the away team. As the visualized playing direction is from right to left, we reverse the passes where the direction is left-to-right, so that they can be displayed in the same figures. Analysing Game 1, passing patterns can be identified mainly on the wings and in the center for both teams. Markers, that are located outside the field implicate, that a bad pass has been played and that defender do not move their position as the pass has no receiver. The patterns implicate, that especially on the offense wings, defenders have a lower player motion, which might occur as the ball is considered to be in a non-dangerous area of the field for the defending team and therefore requires less defending movement. Considering Game 2 and Game 3, we identify a similar distribution of the markers on the pitch. While for the home team, the best passes seem to be mostly on the right respectively left wing, the away team appears additional in the center of the field.

Our No Movement Model (NMM) serves as an indicator for analysing and comparing further prediction models. Nonetheless, it enables some trivial evaluation of tactical approaches. No movement of defenders in soccer is considered as a rather unrealistic prediction for player motion, why we use this approach as a baseline, in order to compare the tactical background behavior of our established models, to a no movement approach. thus, we are capable of analyzing, if the underlying rules of the prediction models are reasonable and appropriate, when their occurring prediction error is lower than the baseline.

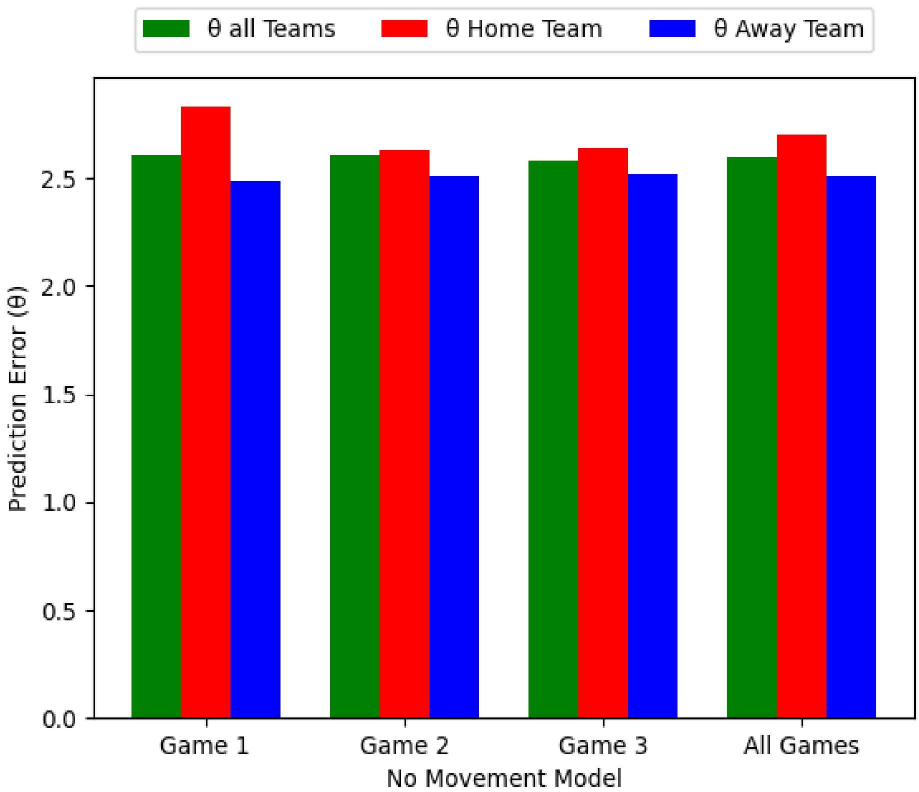
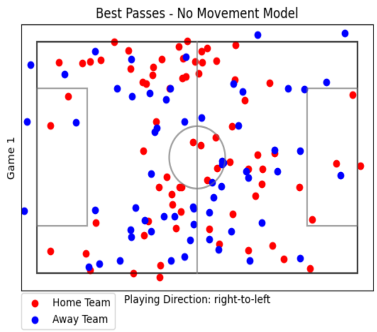
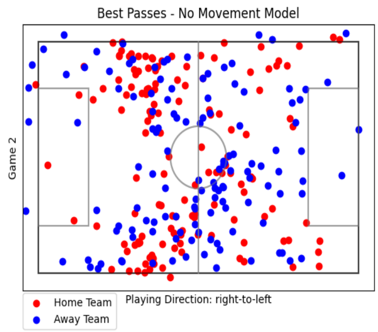


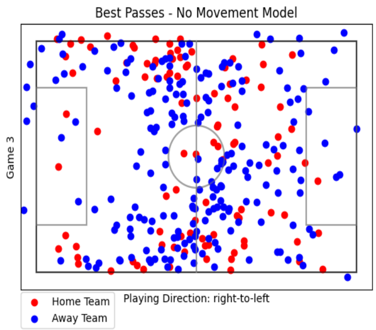
Figure 5.1.: Visualization of the prediction errors for No Movement Model (NMM)



(a) Distribution of the 'best passes' for Game 1



(b) Distribution of the 'best passes' for Game 2



(c) Distribution of the 'best passes' for Game 3

Figure 5.2.: Visualization of patterns for the best passes ($\boldsymbol{\theta} < 1.3$) of our No Movement Model (NMM) on the pitch. The markers represent the position of the ball, after a pass has been played. The playing direction for the defending team is set from right-to-left.

Move Towards Ball Model

When analyzing our MTBM, we consider a ball-oriented defending approach. Table 5.3 presents the prediction errors for all evaluated games, which is visualized in figure 5.3. Besides, the optimized parameters for each corresponding game can be reviewed. The observed Total error for both teams ($\boldsymbol{\theta_{allTeams}}$) is approximately $\sim$2.417, which indicates an overall higher accuracy than our baseline model. Similarities arise, considering $\boldsymbol{\theta_{homeTeam}}$ and $\boldsymbol{\theta_{awayTeam}}$, as the occurring error is higher for the home team for all of the three games. This implicates, that the tactical approach of the away team seems more applicable to our Move Towards Ball Model (MTBM). From the prediction results, we can assume, that in all three games the away team follows a more ball-oriented defending approach compared to the home team. Nonetheless, the overall prediction error is lower for both teams than the baseline. Considering the optimized parameters for the model, we identify a total average movement velocity for the defending players of $\sim$1.087 m/s, which is similar for all three games, without any peculiar outlier.

For identifying in which scenarios the Move Towards Ball Model (MTBM) is most applicable, we refine the 'best passes' for the model. In the previous section, the threshold was declared at an error of 1.3. This threshold was set, due to the high fluctuation and thus the high amount of outliers of the NMM, where some passes implicate a very high $\boldsymbol{\theta}$ and others a very low $\boldsymbol{\theta}$. As our Move Towards Ball Model contains less variation in the occurring prediction error and thus includes less outliers, we lift our threshold to $\boldsymbol{\theta} < 1.8$, in order to filter enough sample passes. This results in 41 passes of the home team and 26 passes of the away team for Game 1. In Game 2 the home team computes 22 passes and the away team 19 while Game 3 identifies 20 for the home respectively 38 passes for the away team. The distribution on the field for the 'best passes' are visualized in fig. 5.6. When analysing Game 1, patterns appear for both teams mostly in the central midfield zone of the pitch. This might implicate that there occur ball-pressing movements around the halfway line, occasionally including a lot of defense tackles. Considering Game 2, we identify that a ball-oriented defending approach is slightly more applied on the offense wings for the away team, while the distribution for the home team is more around the penalty area and halfway line. In Game 3, the patterns occur mostly on both wings for both of the teams.

Our Move Towards Ball Model (MTBM) follows a simple single-rule based approach. However, the underlying tactics serve as an indicator of defending motion in soccer, with a strong focus on ball-oriented defending. It results in a lower prediction error ($\boldsymbol{\theta}$) and thus provides a higher accuracy than our baseline model (NMM), where we can assume that the fundamental tactical approach is applicable to real-world scenarios. The identified prediction error is lower for the away team in all three games. thus we derive, that the tactical background of the away team indicates a more ball-oriented behavior and differs from the tactics of the home team.

	Game 1	Game 2	Game 3	Total
Prediction error ($\boldsymbol{\theta}$)				
$\boldsymbol{\theta_{homeTeam}}$	2.611	2.483	2.444	2.513
$\boldsymbol{\theta_{awayTeam}}$	2.357	2.262	2.339	2.319
$\boldsymbol{\theta_{allTeams}}$	2.432	2.415	2.404	2.417
Player velocity (v)				
$\boldsymbol{v_{allTeams}}$	1.089	1.014	1.158	1.087
$\boldsymbol{v_{homeTeam}}$	1.188	0.871	1.098	1.052
$\boldsymbol{v_{awayTeam}}$	0.959	1.198	1.216	1.124

Table 5.3.: Prediction errors ($\boldsymbol{\theta}$) and optimized parameters of our Move Towards Ball Model (MTBM)

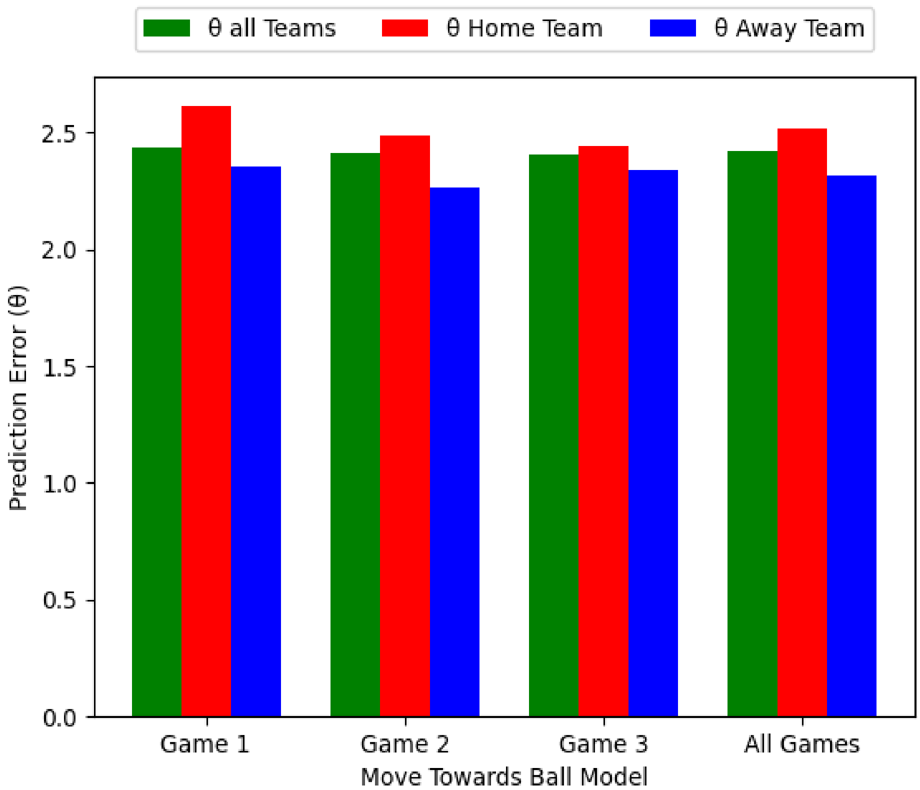
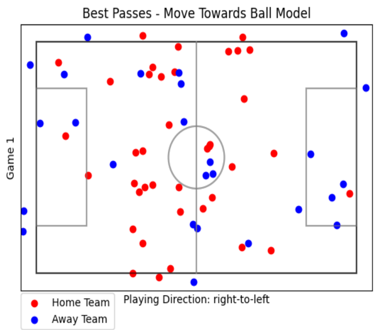
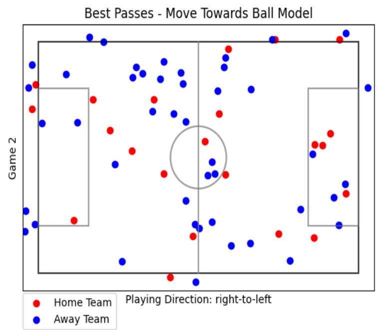


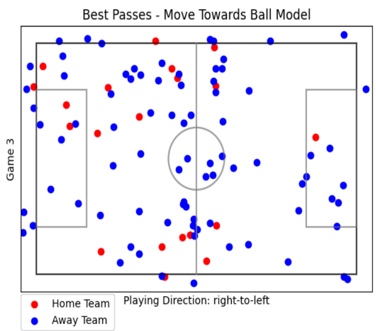
Figure 5.3.: Visualization of the prediction errors for Move Towards Ball Model (MTBM)



(a) Distribution of the 'best passes' for Game 1



(b) Distribution of the 'best passes' for Game 2



(c) Distribution of the 'best passes' for Game 3

Figure 5.4.: Visualization of patterns for the best passes ($\boldsymbol{\theta} < 1.8$) of our Move Towards Ball Model (MTBM) on the pitch. The markers represent the position of the ball, after a pass has been played. The playing direction for the defending team is set from right-to-left.

Pass Angle Model

As a second single-rule model, we evaluate our Pass Angle Model (PAM). This approach follows a ball-oriented tactical background with some regard to zone defense. The analysis of the prediction model identifies a Total prediction error $\boldsymbol{\theta_{allTeams}}$ for both teams of approximately ∼2.506. While the total error for the away team ($\boldsymbol{\theta_{awayTeam}}$) is slightly lower with ∼2.386, the $\boldsymbol{\theta_{homeTeam}}$ is higher with ∼2.625. This is reflected in all three games. While Game 1 has a higher deviation for $\boldsymbol{\theta_{awayTeam}}$ and $\boldsymbol{\theta_{homeTeam}}$, in Game 3 the gap of $\boldsymbol{\theta}$ between the two teams is minor, which appears to be similar to our No Movement Model. Nonetheless, the overall prediction error $\boldsymbol{\theta}$ of our Pass Angle Model is lower, compared to the NMM. The prediction results can be reviewed in table 5.4. As optimized parameters, we identify an average player velocity for all games of approximately ∼0.828, where the average $\boldsymbol{v_{homeTeam}}$ is slightly lower than $\boldsymbol{v_{awayTeam}}$. A graphical representation of the occurring prediction errors ($\boldsymbol{\theta}$) is provided in 5.5.

For the evaluation of the positions of the 'best passes', the threshold for the prediction errors is set to 1.8 ($\boldsymbol{\theta} < 1.8$). We define the threshold higher than our No Movement Model, as the prediction results include less outliers and a smaller fluctuation. Considering Game 1 (49 passes home team and 51 passes away team), we recognize a pattern, where most of the samples appear in the attacking half on the field. This might implicate, that in Game 1 when the ball is in a non-dangerous area for the defending team, defenders rather focus on a more zone oriented defending where they stay in their current formation and move within the same angle as the ball. Similarities are visible for Game 2 and 3 where the home team computed 37 passes and the away team 49 (Game 2) respectively 34 (home team) and 58 (away team) for Game 3. Most of the passes have been played in the attacking half with a higher distribution on the wings. Besides, a third pattern can be identified in between the halfway line and the penalty area, in a central position. This indicates, that the overall tactical approach of our PAM is mostly applicable for situations, where the ball is considered in a neutral area for the defending team.

Our Pass Angle Model is based on a single-rule tactical behavior, where the defending team follows a ball-oriented defending approach with slow regards to zone defending behavior. The overall prediction of our PAM offers a more accurate prediction than our baseline. As the underlying tactical approach follows a rather ball-oriented behavior, the identified error is lower in all three games for the away team, which is identical to our findings from the previous chapter. As the total errors ($\boldsymbol{\theta}$) for the PAM are higher than the MTBM, we assume that both teams have a rather focus on ball-oriented defending.

	Game 1	Game 2	Game 3	Total
Prediction error ($\boldsymbol{\theta}$)				
$\theta_{homeTeam}$	2.742	2.575	2.558	2.625
$\theta_{awayTeam}$	2.321	2.371	2.464	2.386
$\theta_{allTeams}$	2.497	2.518	2.500	2.505
Player velocity (v)				
$v_{allTeams}$	0.864	0.867	0.752	0.828
$v_{homeTeam}$	0.809	0.633	0.731	0.724
$v_{awayTeam}$	0.903	1.146	0.818	0.956

Table 5.4.: Prediction errors ($\boldsymbol{\theta}$) and optimized parameters of our Pass Angle Model (PAM)

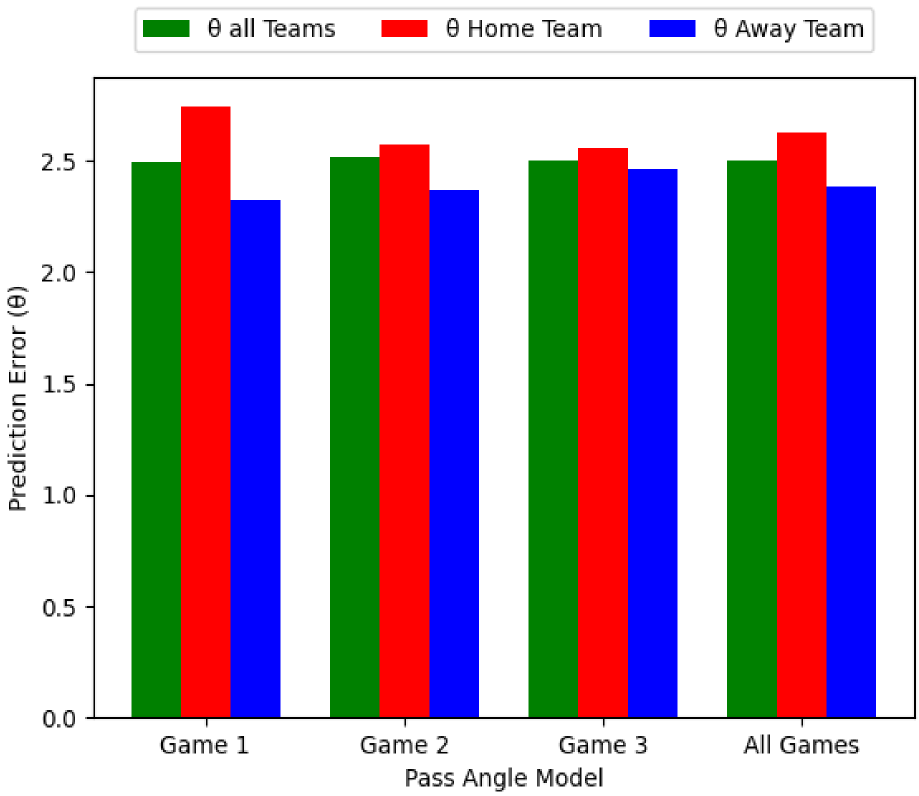
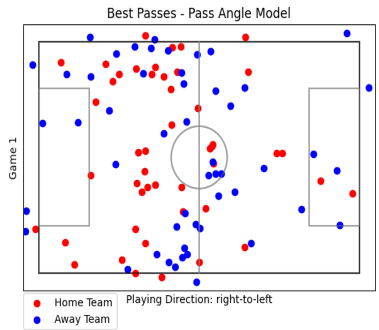
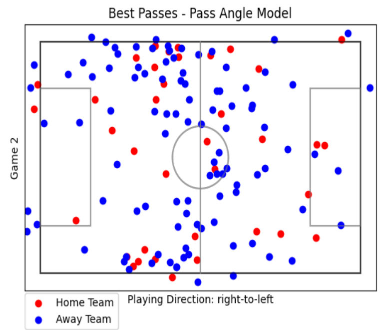


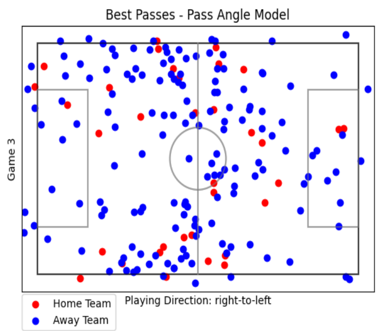
Figure 5.5.: Visualization of the prediction errors for Pass Angle Model (PAM)



(a) Distribution of the 'best passes' for Game 1



(b) Distribution of the 'best passes' for Game 2



(c) Distribution of the 'best passes' for Game 3

Figure 5.6.: Visualization of patterns for the best passes ($\boldsymbol{\theta} < 1.8$) of our Pass Angle Model (PAM) on the pitch. The markers represent the position of the ball, after a pass has been played. The playing direction for the defending team is set from right-to-left.

5.4.2. Complex-Rule Models

In the following, an evaluation of our established complex-rule models is demonstrated. These models are a concatenation of different Rules to a Ruleset, that follow some fundamental tactical approaches in soccer. Our first implemented model is our Man Marking Model (MMM), that is based on the tactical background of man marking behavior. The second model that is evaluated in this section is our Zone Defending Model (ZDM). In this section, we evaluate the prediction error for both models and analyze it with our baseline model (NMM). The results can be reviewed in the following.

Man Marking Model

Our Man Marking Model (MMM) is built on a tactical man marking approach, where basically every player of the attacking team is marked by a defender of the opposite team. It is considered as one of the principles of defending behavior in soccer. The analysis of the Total prediction error $\boldsymbol{\theta_{allTeams}}$ demonstrates an overall error of approximately ∼2.573, which is slightly lower than our identified error for the baseline (NMM). This might implicate, that both teams do not follow a man-oriented defending behaviour. Observing each individual game, we perceive that the model is more applicable to the away team, as the $\boldsymbol{\theta_{awayTeam}}$ is lower than $\boldsymbol{\theta_{homeTeam}}$ for every game. While $\boldsymbol{\theta_{awayTeam}}$ is rather constant for all three games, Game 1 for the $\boldsymbol{\theta_{homeTeam}}$ reveals a high inaccuracy. It indicates, that the home team in Game 1 plays different tactics, which is also confirmed by reviewing the player velocity (v) of the home team. The average speed for the home team in Game 1 and 2 demonstrates, that a movement below 1 m/s towards an opponent player indicates a higher accuracy of the model. Nonetheless, as the total accuracy for both teams is better than our established baseline, we assume that the approach is applicable even if the underlying tactical background of a team follows a different approach. The results of the identified prediction errors and the optimized parameters for the Man Marking Model (MMM) can be reviewed in table 5.6. Besides a graphical representation of the errors $\boldsymbol{\theta}$ are provided in fig. 5.7.

Even though we assume a different tactical approach for both teams of the evaluated games, there might occur some specific constellations, where a man marking approach seems rather promising. We therefore analyzed the 'best passes' of the prediction model, where the prediction error ($\boldsymbol{\theta}$) is set to 1.8, in order to highlight occurring patterns on the field. Due to a more constant emerging error than the NMM, the threshold for $\boldsymbol{\theta}$ is set to 1.8 ($\boldsymbol{\theta} < 1.8$), as our MMM contains less outliers and thus less variation within the occurring prediction error. The evaluation of identifiable patterns for Game 1 demonstrates an appearance for the best passes around the halfway line (44 passes home team and 33 passes away team). While in Game 1 the position of the ball is more located in the attacking half of the defending team, Game 2 implicates a higher appearance in the defending half. For both games, the MMM seems mostly applicable in the center circle of the pitch. A similarity can be analyzed for Game 3, where additionally patterns on both wings can be identified. The evaluation produces 33 passes for the home team and

	Game 1	**Game 2**	**Game 3**	**Total**
Prediction error ($\boldsymbol{\theta}$)				
$\boldsymbol{\theta_{homeTeam}}$	2.802	2.601	2.601	2.668
$\boldsymbol{\theta_{awayTeam}}$	2.467	2.471	2.495	2.478
$\boldsymbol{\theta_{allTeams}}$	2.570	2.583	2.567	2.573
Player velocity(v)				
$\boldsymbol{v_{allTeams}}$	0.810	1.976	0.857	1.215
$\boldsymbol{v_{homeTeam}}$	0.918	0.633	1.860	1.137
$\boldsymbol{v_{awayTeam}}$	1.582	1.714	0.910	1.402
Distance to Ball ($\boldsymbol{\mu}$)				
$\boldsymbol{\mu_{allTeams}}$	24.531	24.272	25.0	24.601
$\boldsymbol{\mu_{homeTeam}}$	24.517	24.640	24.703	24.620
$\boldsymbol{\mu_{awayTeam}}$	24.444	24.423	24.969	24.612
Distance Marked Player ($\boldsymbol{\rho}$)				
$\boldsymbol{\rho_{allTeams}}$	5.4394	9.967	14.271	9.894
$\boldsymbol{\rho_{homeTeam}}$	8.616	8.851	14.358	10.608
$\boldsymbol{\rho_{awayTeam}}$	10.077	9.275	13.593	10.982
Distance Direct Opponent ($\boldsymbol{\omega}$)				
$\boldsymbol{\omega_{allTeams}}$	1.994	0.324	2.0	1.439
$\boldsymbol{\omega_{homeTeam}}$	2.0	1.039	0.382	1.140
$\boldsymbol{\omega_{awayTeam}}$	0.834	0.834	1.938	1.202

Table 5.5.: Prediction errors ($\boldsymbol{\theta}$) and optimized parameters of our Man Marking Model (MMM)

29 passes of the away team (Game 2), respectively 40 passes home team and 51 passes away team for Game 3. Through our identified patterns, which are visualized in figure 5.8, we assume that especially around the halfway line, where a lot of defense tackles take place, a more man-oriented defending approach seems promising for specific constellations.

The Man Marking Model (MMM) serves as an indicator for player-oriented defending behavior. The overall prediction error is lower than the established baseline of no movement for defenders, which indicates, that in some in-game scenarios, defenders may perform a man-marking defending behavior. As the error $\boldsymbol{\theta}$ for all three games is lower for the away team, we suppose a different defending strategy of both teams. The model contains some basic rules, in order to perform a man-marking defense. It can easily be enhanced, by adding further Rules to the existing Ruleset, to provide a higher accuracy.

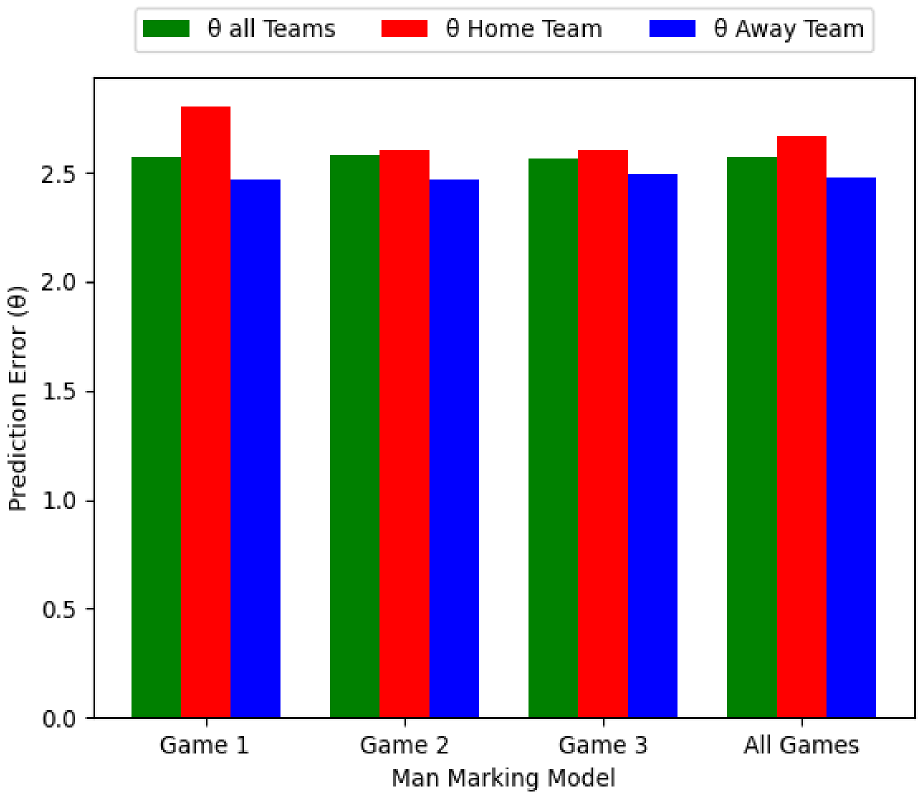
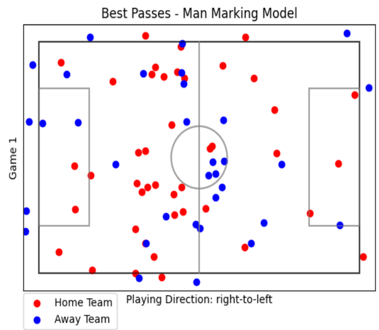


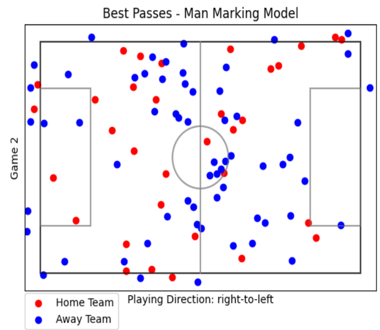
Figure 5.7.: Visualization of the prediction errors for Man Marking Model (MMM)

Zone Defending Model

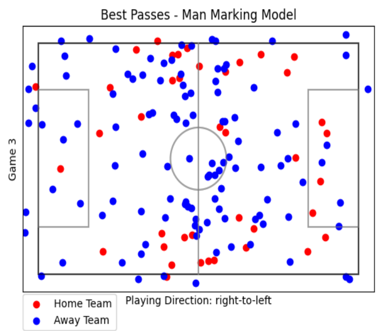
The Zone Defending Model (ZDM) serves as a second complex-rule based model, with a tactical background of zone defending behavior. It is considered as a modern tactical approach. Nonetheless, its underlying tactical behaviour is rather complex. When evaluating our ZDM, an overall prediction error for both teams ($\boldsymbol{\theta_{allTeams}}$) of approximately ∼2.556 can be identified. Compared to our baseline (No Movement Model (NMM)) and our other complex-rule model (Man Marking Model (MMM)), $\boldsymbol{\theta_{allTeams}}$ is slightly lower. thus, we can assume, that zone defense is more applicable for the analyzed games, as the overall prediction error is lower for every game. Similarities appear when considering the error ($\boldsymbol{\theta}$) for each team. For all of the three games, the prediction error is higher for the home team. While for Game 2 and 3 the gap between $\boldsymbol{\theta_{homeTeam}}$ and $\boldsymbol{\theta_{awayTeam}}$ is rather small, a higher deviation between both teams can be identified in Game 1. It implicates, that a zone defending approach is more applicable for predicting the defending behavior of the away team, as the occurring error ($\boldsymbol{\theta}$) is lower. The results of the prediction errors can be revwied in table **??**. A graphical representation is provided in figure 5.9.



(a) Distribution of the 'best passes' for Game 1



(b) Distribution of the 'best passes' for Game 2



(c) Distribution of the 'best passes' for Game 3

Figure 5.8.: Visualization of patterns for the best passes ($\boldsymbol{\theta} < 1.8$) of our Man Marking Model (MMM) on the pitch. The markers represent the position of the ball, after a pass has been played. The playing direction for the defending team is set from right-to-left.

Beside the evaluation of the occurring prediction errors ($\boldsymbol{\theta}$), an analysis of the 'best passes' for the Zone Defending Model (ZDM) is demonstrated in fig. 5.10. In order to refine the scenarios, where a prediction error is minimal, we set a predefined threshold of 1.6 ($\boldsymbol{\theta} < 1.6$) for our ZDM. As the threshold is above threshold of our baseline model (1.3) and at the same time our ZDM indicates an overall better accuracy than the NMM, we suppose less variability in the outcome of the calculated prediction error, which makes the ZDM more constant. When comparing ball positions of the 'best passes' for all three games, we identify the highest amount of played passes for Game 3 (44 passes home team and 50 passes away team), which implies that the approach of zone defending is performed by both of the teams. This is confirmed through the consideration of the prediction errors of both teams, where the gap between $\boldsymbol{\theta_{homeTeam}}$ and $\boldsymbol{\theta_{awayTeam}}$ is minimal. Analyzing the distribution of the passes on the field, patterns can be identified for Game 3 on both wings around the halfway line. A similar distribution appears for Game 2 (32 passes home team and 28 passes away team), with some additional allocation around the centre circle. Game 1, that has the lowest contingent of 'best passes' of all three games with 19 passes home team and 16 passes away team, demonstrates a passing pattern, that is mostly performed in the attacking half of the defending team. Through the established patterns, we assume, that a zone defending approach is mostly applicable, when passes are played in a 'neutral' area for the defending team.

Our Zone Defending Model (ZDM) represents a tactical defending approach, based on modern zone defense. As the overall prediction error is lower than our baseline, we indicate that the model is applicable for real-world scenarios. Besides, the approach implicates a high amount of 'best passes' around the halfway line, which demonstrates that the ZDM is especially appropriate of predicting player motion for passes played in that area. Our approach focuses on the fundamentals of zone defending behavior and can easily be enhanced through additional Rules. Nonetheless, it provides already some valuable predictions for specific in-game constellations.

	Game 1	**Game 2**	**Game 3**	**Total**
Prediction error ($\boldsymbol{\theta}$)				
$\boldsymbol{\theta_{homeTeam}}$	2.802	2.601	2.601	2.668
$\boldsymbol{\theta_{awayTeam}}$	2.467	2.471	2.495	2.478
$\boldsymbol{\theta_{allTeams}}$	2.570	2.583	2.567	2.573
Player velocity(v)				
$\boldsymbol{v_{allTeams}}$	0.810	1.976	0.857	1.215
$\boldsymbol{v_{homeTeam}}$	0.918	0.633	1.860	1.137
$\boldsymbol{v_{awayTeam}}$	1.582	1.714	0.910	1.402
Distance to Ball ($\boldsymbol{\mu}$)				
$\boldsymbol{\mu_{allTeams}}$	24.531	24.272	25.0	24.601
$\boldsymbol{\mu_{homeTeam}}$	24.517	24.640	24.703	24.620
$\boldsymbol{\mu_{awayTeam}}$	24.444	24.423	24.969	24.612
Distance Marked Player ($\boldsymbol{\rho}$)				
$\boldsymbol{\rho_{allTeams}}$	5.439	9.967	14.27	9.894
$\boldsymbol{\rho_{homeTeam}}$	8.616	8.851	14.35	10.60
$\boldsymbol{\rho_{awayTeam}}$	10.077	9.275	13.598	10.982
Distance Direct Opponent ($\boldsymbol{\omega}$)				
$\boldsymbol{\omega_{allTeams}}$	1.994	0.324	2.0	1.439
$\boldsymbol{\omega_{homeTeam}}$	2.0	1.039	0.382	1.140
$\boldsymbol{\omega_{awayTeam}}$	0.834	0.834	1.938	1.202

Table 5.6.: Prediction errors ($\boldsymbol{\theta}$) and optimized parameters of our Man Marking Model (MMM)

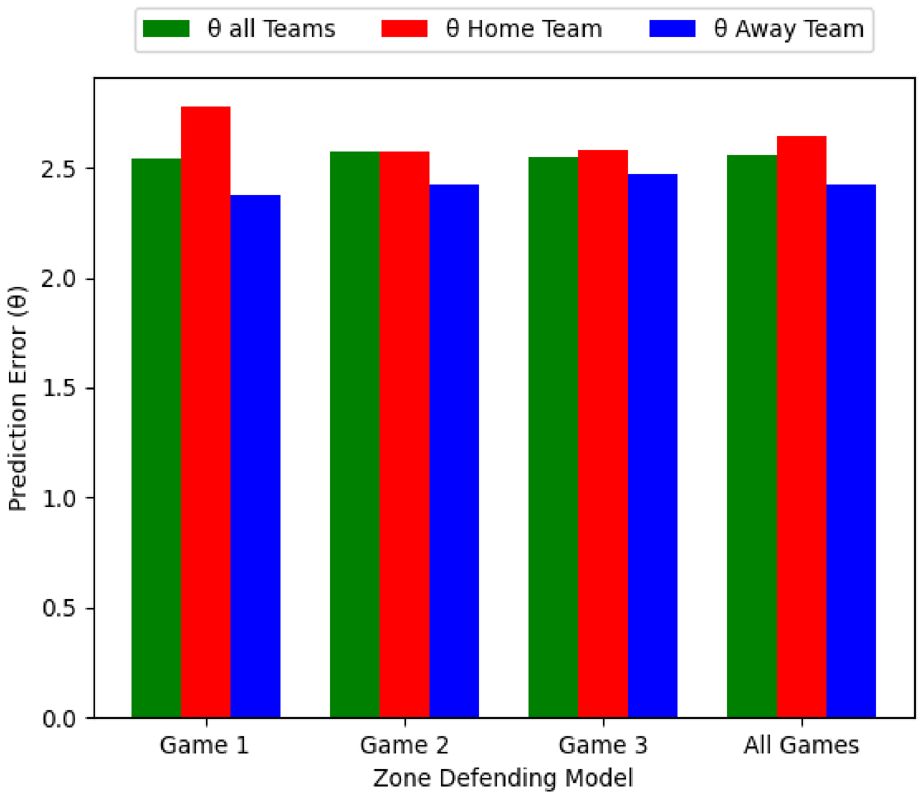
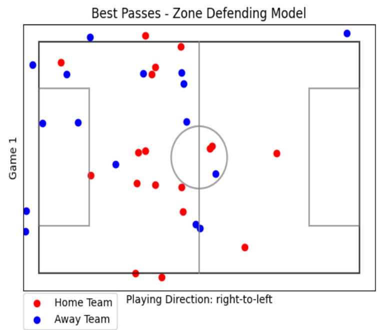


Figure 5.9.: Visualization of the prediction errors for Zone Defending Model (ZDM)

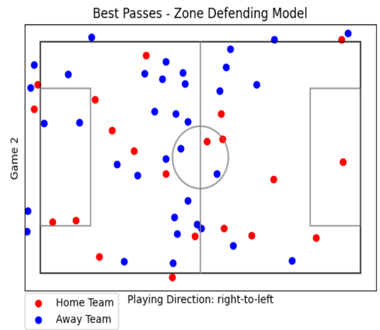
5.4.3. Combination Model

The following section provides an analysis and evaluation of our established Combination Model (CM). In principle, the framework of the model is defined as a Metaruleset, that enables the combination and application of several Rulesets. As a defending team's motion and defending behavior might differ, based on the ball position on the field, different defending strategies might be appropriate. In our CM, we combined the two complex-rule based models (Man Marking Model and Zone Defending Model) with each other. We assign the ZDM to the attacking and midfield zone in order to provide a tactical strategy, that is more based on zone defense, for the 'neutral' area on the pitch. For the defending zone, the MMM is applied, as we assume that a more man-oriented defending behavior seems rather reasonable the this area of the field. As we already assign several defending approaches to different parts of the field, we mainly focus on the evaluation of the occurring prediction and do not analyze the 'best passes', as done in the previous sections.

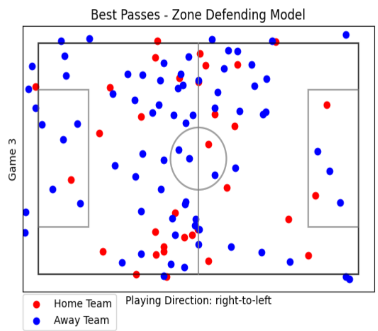
For the analysis of our Combination Model (CM), an overall Total prediction error for



(a) Distribution of the 'best passes' for Game 1



(b) Distribution of the 'best passes' for Game 2



(c) Distribution of the 'best passes' for Game 3

Figure 5.10.: Visualization of patterns for the best passes ($\boldsymbol{\theta} < 1.6$) of our Zone Defending Model (ZDM) on the pitch. The markers represent the position of the ball, after a pass has been played. The playing direction for the defending team is set from right-to-left.

	Game 1	**Game 2**	**Game 3**	**Total**
Prediction error ($\boldsymbol{\theta}$)				
$\boldsymbol{\theta_{homeTeam}}$	2.782	2.566	2.593	2.647
$\boldsymbol{\theta_{awayTeam}}$	2.411	2.450	2.484	2.448
$\boldsymbol{\theta_{allTeams}}$	2.553	2.570	2.540	2.554

Table 5.7.: Prediction errors ($\boldsymbol{\theta}$) of our Combination Model (CM)

both teams ($\boldsymbol{\theta_{allTeams}}$) of approximately $\sim$2.554 occurs. Compared to the underlying Rulesets, the Total $\boldsymbol{\theta_{allTeams}}$ for CM is lower than the error for our Man Marking Model and Zone Defending Model. It indicates, that a combination of both of our complex-rule models, might result in a higher prediction accuracy. Besides, similarities to the baseline model can be identified for the errors of each particular game of the home and away team, where the prediction for the away team implies a higher accuracy compared to the home team. Also, the Total error for the away team ($\boldsymbol{\theta_{awayTeam}}$) is slightly lower, compared to the Total error of the home team ($\boldsymbol{\theta_{homeTeam}}$). Evaluating the results of the CM for the home team, it can be identified that the error in every game in particular is similar to the error for the corresponding game of the ZDM. On the other hand, the $\boldsymbol{\theta_{awayTeam}}$ in every single game of our CM indicates a rather similarity to the errors of the MMM. It might demonstrate, that when combining both of our complex-rule models with each other, the away team follows a more man-oriented defending approach, while the tactics of the home team implies a rather zone-oriented defending. The prediction results can be reviewed in table 5.7. A graphical representation of the prediction error $\boldsymbol{\theta}$ for each game is demonstrated in figure 5.11. The optimized parameters for our Combination Model are provided in Appendix in the tables A.1 and A.2.

Our Combination Model combines the two tactical approaches of zone defense and man marking, in order to predict the movement of defending players. The evaluation of the model identifies an overall lower prediction error than the baseline model of no movement. Therefore, we assume that a combination of different prediction models, where they are applied to different parts of the field, result in a higher accuracy of the player motion. Besides, the combination of the models results in an overall lower prediction error than the particular error of each underlying model. This implicates, that a combination of different tactical approaches seems rather appropriate for predicting and evaluating real-world scenarios.

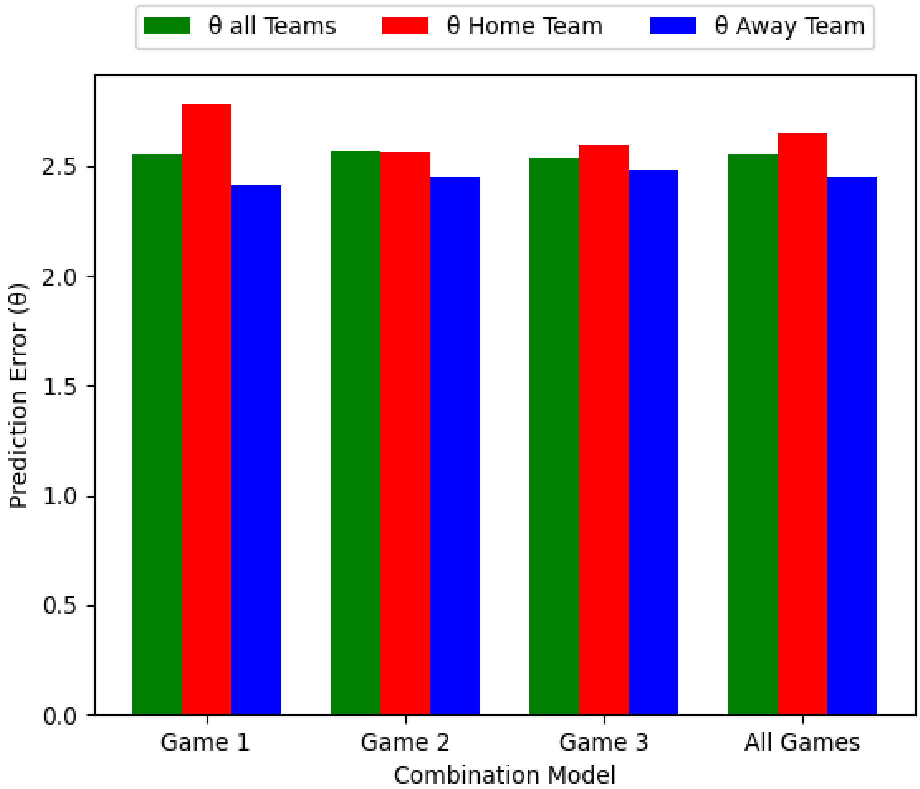


Figure 5.11.: Visualization of the prediction errors for Combination Model (CM)

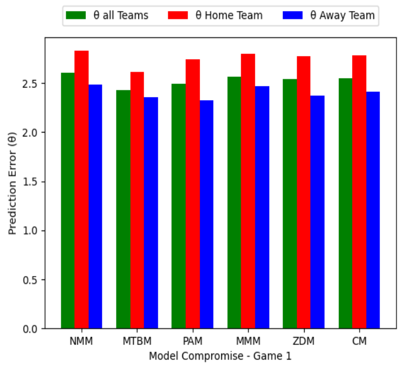
5.4.4. Model Compromise

In the following section, we evaluate and compare the introduced models from the previous sections. We analyze, which of the underlying tactical approach seems most promising for predicting defending player motion in soccer. Besides, an evaluation of occurring similarities regarding defending patterns and prediction errors of the established approaches are demonstrated. A comparison of the prediction results is provided in table 5.8. A graphical representation can be reviewed in fig. 5.12.

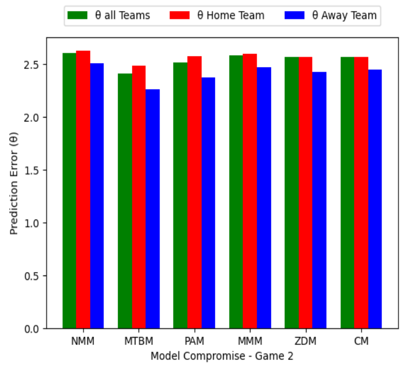
The evaluation of our models demonstrates the highest prediction error ($\boldsymbol{\theta}$) and thus the lowest accuracy for our baseline (No Movement Model). It implicates, that all of our established models provide a better overall prediction for defending movement, than the benchmark of no movement. Nonetheless, there might arise some specific game constellations e.g. when a prediction model forecasts a movement towards the opposite direction than the actual movement of a defender, where a no movement approach provides a higher accuracy. The lowest accuracy of the complex-rule models is identified for the MMM, which confirms the findings of the previous sections, that a man-marking approach seems rather outdated in modern soccer and is not applied by the teams of the analyzed games. The highest accuracy can be evaluated for our MTBM. Thus we assume, that a movement for the defenders of the evaluated games strongly depends on the ball-position on the field. A prediction with a underlying ball-oriented background seems therefore most promising. As our ZDM follows a straight zone-defending behavior, we identify a slightly higher inaccuracy. This might also result due to a rather rudimentary underlying zone-defending behavior, that can be further enhanced through establishing additional rules and shift the focus to a more ball-oriented approach. Considering the errors of the particular games, a variation between the accuracy of a model and the corresponding game can be identified. For instance, while our ZDM indicates the highest accuracy for both teams in Game 1, the lowest prediction error for both teams of the MMM is identified in Game 3. thus, we identify similarities for our PAM and ZDM, which implies that for Game 1 a more zone-oriented focus is conducted, while Game 3 identifies a rather ball-oriented defending manner. For the analysis of the different teams, we perceive similarities for all models, where the prediction for the away team provides a higher accuracy. thus, we can assume, that the defending motion of the home team follows either a more complicated, respectively different tactical background, or might vary more often during the evaluated games, than the tactics of the away team.

When evaluating different patterns of the 'best passes', similarities and deviations can be identified between the models regarding the position of the ball. Analysing the complex-rule models, an appearance of the best passes for both models occur mainly in the attacking part of the field. The best passes of our MMM mostly appear in the center around the centre circle, while the ZDM reveals a rather occurrence on the wings. We can thus suppose, that the analyzed teams follow a more zone-oriented defending approach, when passes are played towards the touchline. Passes played in the centre around halfway line, require a more man-oriented defense, according to the 'best passes' patterns of our MMM. These findings are also confirmed by the PAM, which implicates some regards to zone-defending behavior and thus indicates patterns of the 'best passes' on both wings. Some differences can be identified for the MTBM. Besides an appearance in the attacking zone on the pitch, the patterns also imply that a movement towards the ball is applicable around the penalty area of the defending team. We assume, that our MTBM has the highest prediction accuracy, within specific ball-pressure constellations, that require a strong movement towards the ball.

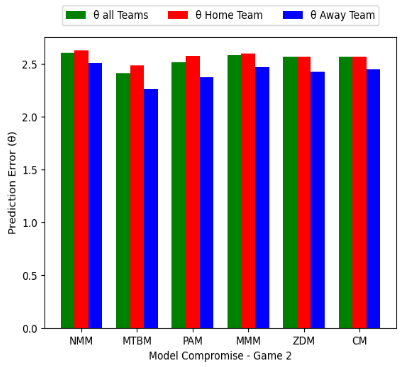
Through the evaluation of our established models we gain a deeper understanding of tactical defending behavior. Analyzing and comparing the prediction accuracy of each individual model allows us to identify, which approach seems most promising for different games and individual teams. Besides, it reveals specific constellations, where a prediction of a model provides the highest accuracy and is mostly applicable. Overall, our established models indicate a higher prediction accuracy than the considered benchmark, which makes them appropriate to forecast real-world player motion. The identified knowledge can be used, to establish additional Rules and thus enhance prediction accuracy of the existing



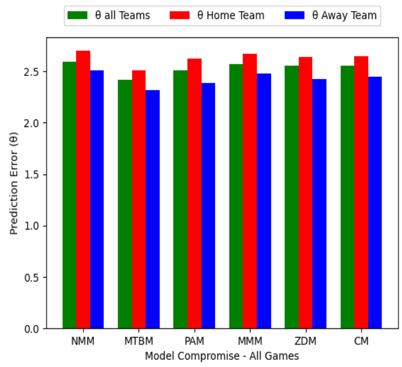
(a) Prediction error ($\boldsymbol{\theta}$) for Game 1



(b) Prediction error ($\boldsymbol{\theta}$) for Game 2



(c) Prediction error ($\boldsymbol{\theta}$) for Game 3



(d) Prediction error ($\boldsymbol{\theta}$) for all Games

Figure 5.12.: Comparison of the prediction errors ($\boldsymbol{\theta}$) of all prediction models

models.

	Game 1	**Game 2**	**Game 3**	**Total**
Prediction error ($\boldsymbol{\theta}$) NMM				
$\boldsymbol{\theta_{homeTeam}}$	2.830	2.626	2.641	2.699
$\boldsymbol{\theta_{awayTeam}}$	2.489	2.507	2.522	2.506
$\boldsymbol{\theta_{allTeams}}$	2.604	2.604	2.581	2.596
Prediction error ($\boldsymbol{\theta}$) MTBM				
$\boldsymbol{\theta_{homeTeam}}$	2.611	2.483	2.444	2.513
$\boldsymbol{\theta_{awayTeam}}$	2.357	2.262	2.339	2.319
$\boldsymbol{\theta_{allTeams}}$	2.432	2.415	2.404	2.417
Prediction error ($\boldsymbol{\theta}$) PAM				
$\boldsymbol{\theta_{homeTeam}}$	2.742	2.575	2.558	2.625
$\boldsymbol{\theta_{awayTeam}}$	2.321	2.371	2.464	2.386
$\boldsymbol{\theta_{allTeams}}$	2.497	2.518	2.500	2.505
Prediction error ($\boldsymbol{\theta}$) MMM				
$\boldsymbol{\theta_{homeTeam}}$	2.802	2.601	2.602	2.668
$\boldsymbol{\theta_{awayTeam}}$	2.467	2.471	2.495	2.478
$\boldsymbol{\theta_{allTeams}}$	2.578	2.583	2.567	2.573
Prediction error ($\boldsymbol{\theta}$) ZDM				
$\boldsymbol{\theta_{homeTeam}}$	2.776	2.569	2.581	2.642
$\boldsymbol{\theta_{awayTeam}}$	2.374	2.425	2.471	2.423
$\boldsymbol{\theta_{allTeams}}$	2.544	2.572	2.550	2.555
Prediction error ($\boldsymbol{\theta}$)CM				
$\boldsymbol{\theta_{homeTeam}}$	2.782	2.566	2.593	2.647
$\boldsymbol{\theta_{awayTeam}}$	2.411	2.453	2.484	2.448
$\boldsymbol{\theta_{allTeams}}$	2.553	2.570	2.540	2.554

Table 5.8.: Prediction errors ($\boldsymbol{\theta}$) of all established prediction models

6. Findings and Limitations

In this section, the presented results from the previous chapters are critically reviewed and discussed. We elaborate the findings through our prediction models, where we dispute strengths and weaknesses of the underlying approach. Additionally, we consider limitations and potential enhancements of our established framework, with regard to the application through domain experts.

We introduced different models, that follow some real-world tactical behavior. Based on these models, we extracted knowledge in the form of Rules out of fundamental soccer tactics, which we then used in order to establish our prediction models. These models are capable of predicting different game constellations, based on their underlying tactical approach. In order to evaluate if our approaches provide a valuable outcome, we established a benchmark for comparison, where players do not move their current positions after a pass has been played. Since all of our models provide a higher accuracy than this baseline, we can confirm that our approaches are capable of establishing validated predictions for real-world soccer scenarios.

We built a framework, for making highly accurate predictions about the movement of defending players after a pass. Through the evaluation of the models, we identified, that a more ball-oriented defending approach seems most promising in order to forecast defending behavior in soccer. Defense players follow some certain rules when adapting their positions on the field. These rules define the tactical approach, how players act in their defending motion. From our findings we inferred, that additionally to these rules, a strong dependency on the ball's movement on the field exists. This serves as a main indicator for defender's motion. We come to this conclusion, as our MTBM and PAM, which strongly depend on a ball movement, involve the highest prediction accuracy. Beside a general evaluation of our prediction models, we identified recurring defending patterns that describe the position on the field where a model seems most applicable. We perceived, that tactical approaches that are rather based on a ball- respectively zone-oriented defending, indicate the highest prediction accuracy around the halfway line of the pitch. Models with a rather man-oriented tactical background seem additionally applicable around the penalty area in the defending half of the field.

Comparing a man-marking defending approach with a zone defense projection, we come to the conclusion, that a zone-defending approach seems more applicable. The results of the two different prediction models indicate, that a zone defense approach seems more adaptable to in-game constellations than a man marking defending, as the ZDM outperforms the MMM in their prediction accuracy. Additionally, we infer, that our Zone Defending

Model could be further enhanced, with more regard to ball-oriented movement, in order to provide a lower prediction error. The evaluation results from a combination of both models indicate a higher accuracy for our CM. We assume, that teams do not follow one specific tactical approach during a whole game, and the defending behavior might vary due to specific in-game constellations on the field where the tactical behavior of a team changes.

Our findings of the evaluation indicate, that a ball-oriented defending strategy seems most promising, while a strictly man-marking defending seems rather outdated. Through our established framework, that contains Rules, Rulesets and Metarulesets, our models can easily be reconstructed and further adjusted by additional Rules. This implicates a high reproducibility of our approach. Thus, domain experts can rebuild our models with a rather focus on ball-oriented defending for future work. Furthermore, new Rules can be established, in order to adjust the existing Rulesets or to build new models with a different tactical background. These new models might serve as a new indicator of analyzing specific game constellations from a different tactical point of view.

Even though we implemented a valuable prediction framework, our models also face some limitations. The evaluation of the results demonstrates, that an accurate prediction for player motion in soccer is a high-complexity problem. Various factors might influence general tactical team behavior, as well as the movement of one particular player. Since we use the knowledge of domain experts, in order to establish the Rules for our prediction models, the approach is limited to the tactical background knowledge of the experts. Thus, the prediction depends on this knowledge and is limited to its underlying Rules. We counteract these limitations through the structure of our framework, where Rulesets can easily be adjusted and enhanced. Furthermore, the evaluation of models is limited to the three analyzed games. An analysis with more matches might implicate a better tactical understanding and also result into different findings. We do not know the participating teams of the evaluated games, which complicates the prediction analysis. For instance, a world-class team might follow a more offensive tactical background, and therefore a different defending approach, than a team from the third division. Furthermore, the prediction result does not consider any change of the current match-score, which also affects the tactical behavior of both teams, since they need to adjust their current tactics. Besides, our prediction models do not consider external factors such as e.g. physical conditions of each individual player, that might influence defender's motion.

We established a prediction framework, that is capable of making valuable forecasts about defending player motion in modern soccer. The approach implicates an easy reproducibility, as the underlying framework enables a trivial adjustment of existing Rulesets or the establishment of new models. Nonetheless, our research also confronts some limitations in its prediction and evaluation. Some of these limiting factors can be reduced by applying and evaluating additional data sets with knowledge of the participating teams. For future work, we can use our findings from the evaluation results, to establish models with a more complex tactical basement.

7. Summary and Future Work

We presented a novel approach for predicting and evaluating player motion in soccer. Our established framework is capable of making highly accurate predictions of defending players movement. Therefore, we applied the approach of Rule-based Systems (RBS), where we defined different Rules, that serve as basic module of our prediction models. Through the structure of our framework, domain experts can easily establish new models or adjust existing ones, which implicates a high applicability and flexibility of our approach.

The findings of the evaluation indicate, that a ball-oriented defending behavior provides the highest prediction accuracy and seems most applicable for analyzing specific game constellations. Nonetheless, we can confirm that all of our established models are capable of providing a valuable forecast for defending player motion, since all of our prediction models contain a better precision in their forecast than the established benchmark model of no player movement. Beside a general evaluation of our approach, an identification of any recurring defending patterns was conducted, that describe the position on the field where a model seems most applicable. We identified, that a man-oriented defending approach seems applicable around the penalty area in the defending half of the field and around the halfway line. Models with a rather zone- respectively ball-oriented defending approach indicate the highest prediction accuracy around the halfway line of the pitch. We considered, whether general assumptions about soccer tactics can be extracted, based on our models and the evaluated games. At this juncture, we analyzed the different games and models on any recurring similarities, which indicates that one approach is more applicable for overall matches. We identified, that a man-marking approach seems rather outdated and that a zone-defending approach, respectively ball-oriented defending measures are more employable for real-world scenarios.

For future work, we plan to use the knowledge through our evaluation findings and further enhance our existing models with additional Rules, that rather follow a ball-oriented defending behavior. Therefore, a deeper domain knowledge in order to establish valuable Rules for movement prediction is required, to enhance complexity of the underlying tactical background. Since our approach faces some limitations, a rather focus on an extension through Machine Learning algorithms, such as Imitation Learning, which follow the behavior of Reinforcement Learning seems promising in future work [Le et al., 2017]. Thus, Imitation Learning aims to extract knowledge from domain experts, in order to reproduce their behavior [Zheng et al., 2022], which we assume to be applicable to our approach.

Bibliography

[BALADRAM et al., 2020] BALADRAM, M., KOIKE, A., and YAMADA, K. (2020). Introduction to supervised machine learning for data science. *Interdisciplinary Information Sciences*, 26:87–121.

[Bischofberger and Eigenrauch, 2023] Bischofberger, J. and Eigenrauch, S. (2023). *football1234*. Available online at `https://github.com/jonas-bischofberger/football1234`, accessed on 28.06.2023, contact: jonas.bischofberger@univie.ac.at.

[Cichy and Kaiser, 2019] Cichy, R. and Kaiser, D. (2019). Deep neural networks as scientific models. *Trends in Cognitive Sciences*, 23.

[Decroos et al., 2018] Decroos, T., Van Haaren, J., and Davis, J. (2018). Automatic discovery of tactics in spatio-temporal soccer match data. pages 223–232.

[e.V. (DFB), 2015] e.V. (DFB), D. F. (2015). *Pressing: Wo und wann angreifen?* Available online at `https://www.dfb.de/trainer/artikel/pressing-wo-und-wann-angreifen-1638/`, accessed on 10.05.2023.

[Fujii, 2021] Fujii, K. (2021). Data-driven analysis for understanding team sports behaviors. *Journal of Robotics and Mechatronics*, 33:505–514.

[Garnett and Garnett, 2023] Garnett, R. and Garnett (2023). *Bayesian optimization*. Cambridge University Press, Cambridge, United Kingdom New York, NY.

[Goes-Smit et al., 2018] Goes-Smit, F., Kempe, M., Meerhoff, R., and Lemmink, K. A. (2018). Not every pass can be an assist: A data-driven model to measure pass effectiveness in professional soccer matches. *Big Data*, 7.

[inFranken.de, 2018] inFranken.de (2018). *Bamberger Uni-Auswahl kickt in Argentinien und Uruguay*. Available online at `https://www.infranken.de/lk/bamberg/sport/bamberger-uni-auswahl-kickt-in-argentinien-und-uruguay-art-3225650`, accessed on 17.05.2023.

[Jiang, 2021] Jiang, H. (2021). *Machine learning fundamentals : a concise introduction /*. Cambridge University Press,, Cambridge :.

[Kempe et al., 2016] Kempe, M., Vogelbein, M., and Nopp, S. (2016). The cream of the crop: Analysing fifa world cup 2014 and germany's title run. *Journal of Human Sport and Exercise*, 11.

[Lang et al., 2016] Lang, S., Link, D., and Seidenschwarz, P. (2016). Real time quantification of dangerousity in soccer using spatiotemporal tracking data.

[Le et al., 2017] Le, H., Carr, P., Yue, Y., and Lucey, P. (2017). Data-driven ghosting using deep imitation learning.

[Liu et al., 2014] Liu, H., Gegov, A., and Stahl, F. (2014). Categorization and construction of rule based systems. volume 459.

[Mathwarehouse.com,] Mathwarehouse.com. *Pythagorean Theorem.* Available online at `https://www.mathwarehouse.com/geometry/triangles/how-to-use-the-pythagorean-theorem.php`, accessed on 16.05.2023.

[Metrica-Sports, 2021] Metrica-Sports (2021). *Metrica sports sample data.* Available online at `https://github.com/metrica-sports/sample-data/commit/e706dd506b360d69d9d123d5b8026e7294b13996`, accessed on 18.05.2023.

[Nogueira, 2014] Nogueira, F. M. F. (2014). *Bayesian Optimization.* Available online at `https://github.com/bayesian-optimization/BayesianOptimization`, accessed on 25.05.2023.

[Rein and Memmert, 2016] Rein, R. and Memmert, D. (2016). Big data and tactical analysis in elite soccer: future challenges and opportunities for sports science. *SpringerPlus*, 5(1):1410–1410.

[Rudolph, 2018] Rudolph, D. (2018). *Winkelfunktionen: Sinus, Kosinus und Tangens.* Available online at `https://www.gut-erklaert.de/mathematik/winkelfunktionen-sinus-kosinus-tangens.html`, accessed on 16.05.2023.

[Shah and Romijnders, 2016] Shah, R. and Romijnders, R. (2016). Applying deep learning to basketball trajectories.

[Spearman, 2018] Spearman, W. (2018). Beyond expected goals.

[Spielverlagerung.de, 2012] Spielverlagerung.de (2012). *Die Manndeckung.* Available online at `https://spielverlagerung.de/2012/12/31/die-manndeckung/`, accessed on 13.05.2023.

[Spielverlagerung.de, 2013] Spielverlagerung.de (2013). *Die Raumdeckung.* Available online at `https://spielverlagerung.de/2013/01/02/die-raumdeckung/`, accessed on 11.05.2023.

[Sutton, 2018] Sutton, R. S. (2018). *Reinforcement learning : an introduction /.* Adaptive computation and machine learning. The MIT Press,, Cambridge, Massachusetts :, second edition. edition.

[Teoldo, 2021] Teoldo, I. (2021). *Football Intelligence : Training and Tactics for Soccer Success.* Routledge,, [Place of publication not identified] :, first edition.. edition.

[Thabtah et al., 2019] Thabtah, F., Zhang, L., and Abdelhamid, N. (2019). Nba game result prediction using feature analysis and machine learning. *Annals of Data Science*, 6.

[Tim Kipphan, 2022] Tim Kipphan, U. B. (2022). *Fußballer als Botschafter der Universität Bamberg*. Available online at `https://www.uni-bamberg.de/presse/pm/artikel/fussballtour-elfenbeinkueste/`, accessed on 17.05.2023.

[tsg hoffenheim.de, 2021] tsg hoffenheim.de (2021). *Paul Ehmann: Der Doktorand auf der Trainerbank*. Available online at `https://www.tsg-hoffenheim.de/aktuelles/news/2021/12/paul-ehmann-der-doktorand-auf-der-trainerbank/`, accessed on 17.05.2023.

[Wang et al., 2018] Wang, J., Fox, I., Skaza, J., Linck, N., Singh, S., and Wiens, J. (2018). The advantage of doubling: A deep reinforcement learning approach to studying the double team in the nba.

[Wang et al., 2015] Wang, Q., Zhu, H., Hu, W., Shen, Z., and Yao, Y. (2015). Discerning tactical patterns for professional soccer teams: An enhanced topic model with applications.

[Zheng et al., 2022] Zheng, B., Verma, S., Zhou, J., Tsang, I., and Chen, F. (2022). Imitation learning: Progress, taxonomies and challenges. *IEEE transactions on neural networks and learning systems*, PP.

Acronyms

BO Bayesion Optimization. 27, 28, 30

CM Combination Model. v, vi, 24, 28, 46, 48, 49, 52, 54, 67

GPs Gaussian Processes. 27

GUI Graphical User Interface. iii, vi, 9, 60, 61, 64, 65

ML Machine Learning. 2, 55

MMM Man Marking Model. v, vi, 18, 19, 24, 25, 28, 40–43, 45, 46, 48–50, 52, 53, 67

MTBM Move Towards Ball Model. v, vi, 10, 13, 28, 30, 34–37, 50, 52, 53

NMM No Movement Model. v, vi, 10, 30–34, 37, 40, 42, 44, 49, 52

PAM Pass Angle Model. v, vi, 13, 14, 16, 28, 30, 37–39, 50, 52, 53

RBS Rule-based Systems. 1–5, 55

RL Reinforcement Learning. 2, 3

SL Supervised Learning. 2, 3

UL Unsupervised Learning. 2, 3

ZDM Zone Defending Model. v, vi, 19–21, 24, 25, 28, 40, 42, 44, 46–48, 50, 52, 53, 66, 67

A. Appendix

In order to make our framework applicable to domain experts, we implemented a Graphical User Interface (GUI) of our approach. Through our GUI, experts are capable of establishing their own models by combining different predefined Rules with each other. Figure A.1 shows an overview of some of the Rules, that can be selected and further concatenated to a Ruleset. To enhance the analysis of a tactical approach, variable parameters such as e.g. a player´s velocity, can be set manually by the user. The result of the prediction is based on this selection. In fig. A.2, applied Rules with their underlying parameters are demonstrated. The order of the Rules can be changed manually. Through the confirmation button, the Rules of a model are applied to a Ruleset. In a next step, domain experts select different passes for the analysis, as demonstrated in fig. A.3. Furthermore, the underlying Rules of a model with their corresponding parameters can be reviewed. Through the 'draw prediction' button, the constellations of the selected passes are evaluated and visualized. Figure A.4 demonstrates the visualization of the prediction result. Besides, different models can be added to a Metaruleset. In figure A.5 the GUI of our Metaruleset is demonstrated. Thus, domain experts can review their previously established Rulesets and assign them to one of the predefined zones of the field. In a next step, passes from the underlying data sets can be selected, in order to analyze and visualize the corresponding game constellation.

Figure A.1.: Implemented Rules of our GUI

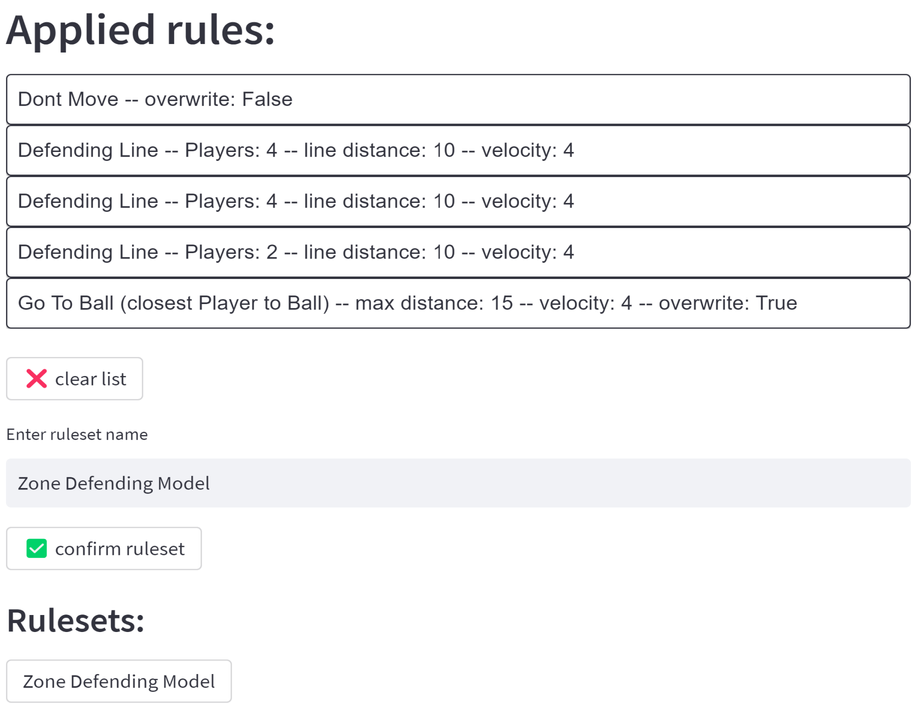


Figure A.2.: Applied Rules of a Model

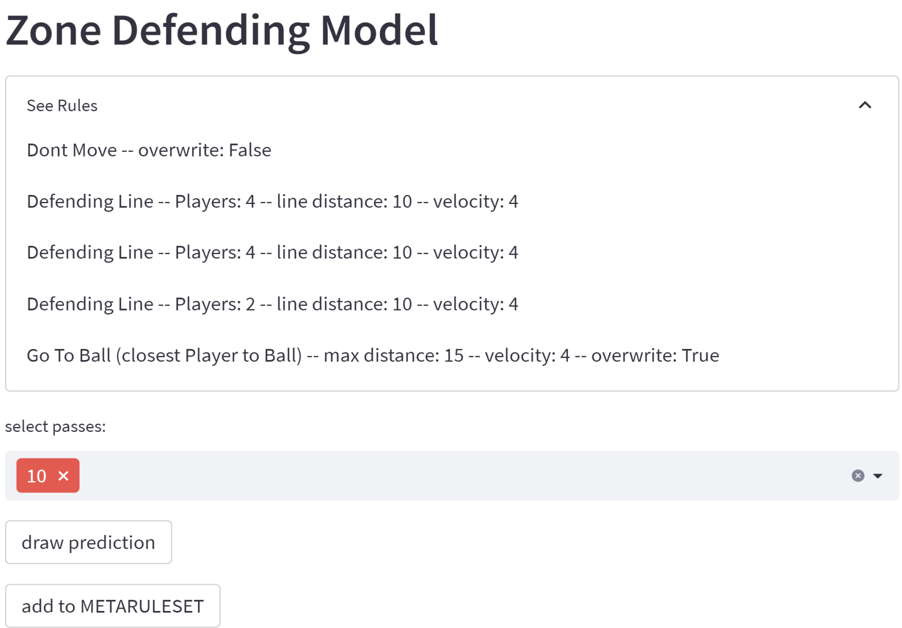


Figure A.3.: Selection of passes for the evaluation

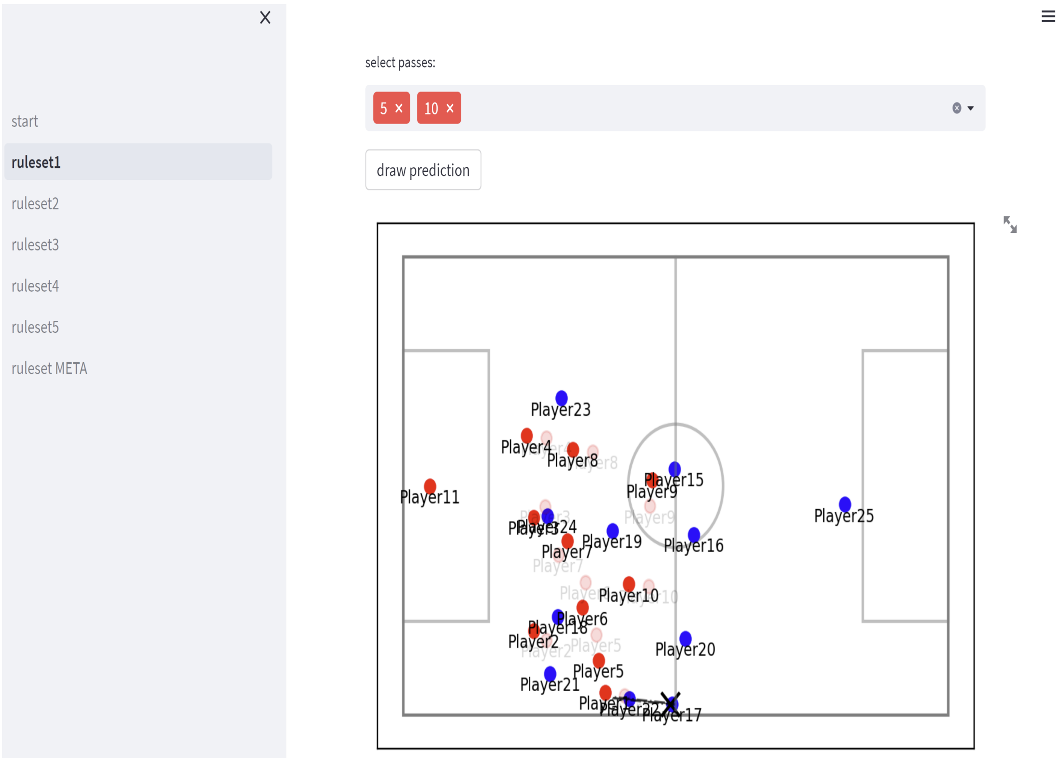


Figure A.4.: Visualization of the prediction plot through our GUI

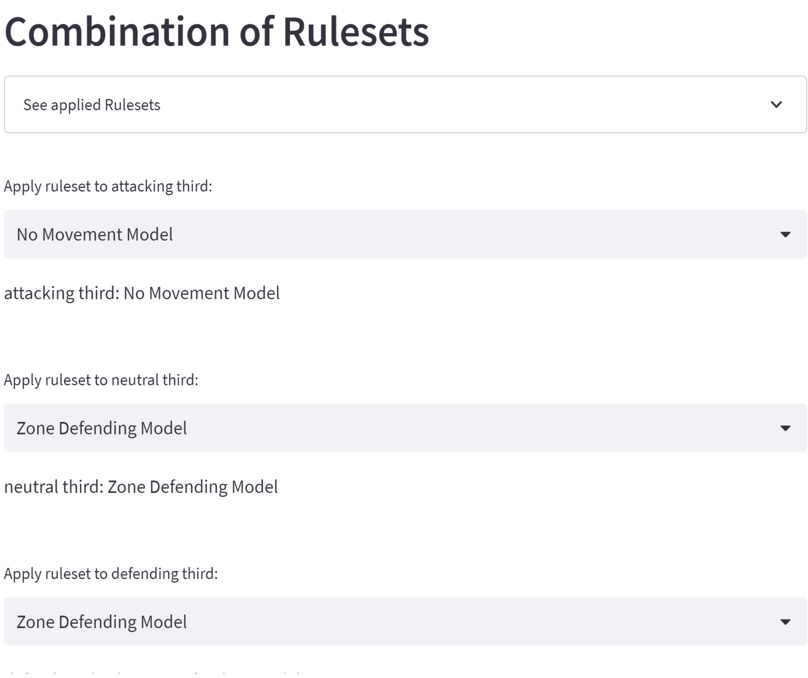


Figure A.5.: GUI of different Rulesets that can be applied to a Metarulest

	Game 1	Game 2	Game 3	Total
Distance Marked Player (ρ)				
$\boldsymbol{\rho_{homeTeam}}$	16.808	17.810	16.435	17.017
$\boldsymbol{\rho_{awayTeam}}$	0.213	16.691	17.810	11.571
$\boldsymbol{\rho_{allTeams}}$	17.343	17.810	16.433	17.195
Distance Direct Opponent (ω)				
$\boldsymbol{\omega_{allTeams}}$	0.145	0.451	0.270	0.289
$\boldsymbol{\omega_{homeTeam}}$	0.099	0.451	0.261	0.270
$\boldsymbol{\omega_{awayTeam}}$	0.807	1.215	0.451	0.824
Distance to Ball ($\boldsymbol{\mu}$)				
$\boldsymbol{\mu_{allTeams}}$	9.177	9.438	7.503	8.706
$\boldsymbol{\mu_{homeTeam}}$	8.688	9.438	7.412	8.512
$\boldsymbol{\mu_{awayTeam}}$	6.333	10.474	9.438	8.748
Distance Defending Line - 4 Players ($\boldsymbol{\delta}1$)				
$\boldsymbol{\delta 1_{allTeams}}$	11.202	7.846	10.372	9.807
$\boldsymbol{\delta 1_{homeTeam}}$	11.453	7.846	10.481	9.926
$\boldsymbol{\delta 1_{awayTeam}}$	13.533	7.641	7.846	9.673
Distance Midfield Line - 4 Players ($\boldsymbol{\delta}2$)				
$\boldsymbol{\delta 2_{allTeams}}$	11.293	13.905	12.462	12.553
$\boldsymbol{\delta 2_{homeTeam}}$	11.297	13.905	12.401	12.534
$\boldsymbol{\delta 2_{awayTeam}}$	6.012	15.0	13.905	11.639
Distance Attacking Line - 2 Players ($\boldsymbol{\delta}3$)				
$\boldsymbol{\delta 3_{allTeams}}$	7.171	6.561	6.429	6.723
$\boldsymbol{\delta 3_{homeTeam}}$	7.029	6.561	6.425	6.672
$\boldsymbol{\delta 3_{awayTeam}}$	5.961	14.559	6.561	9.025

Table A.1.: Optimized parameters of our Zone Defending Model (ZDM)

	Game 1	Game 2	Game 3	Total
Zone Localization Attacking Zone (AZ)				
$\boldsymbol{AZ_{total}}$	-35.355	-34.894	-28.178	-32.809
$\boldsymbol{AZ_{home}}$	-33.407	-34.894	-27.866	-32.056
$\boldsymbol{AZ_{away}}$	-29.454	-50.122	-34.894	-38.157
Zone Localization Defending Zone (DZ)				
$\boldsymbol{DZ_{total}}$	41.591	40.017	40.825	40.811
$\boldsymbol{DZ_{home}}$	42.306	40.017	40.855	41.059
$\boldsymbol{DZ_{away}}$	26.615	27.982	40.017	31.538
Velocity Attacking Zone ($\boldsymbol{V_{AZ}}$)				
$\boldsymbol{V_{AZtotal}}$	1.497	1.939	1.292	1.576
$\boldsymbol{V_{AZhome}}$	1.393	1.939	1.258	1.530
$\boldsymbol{V_{AZaway}}$	2.477	1.696	1.939	2.038
Velocity Defending Zone ($\boldsymbol{V_{DZ}}$)				
$\boldsymbol{V_{DZtotal}}$	2.625	2.920	2.805	2.783
$\boldsymbol{V_{DZhome}}$	2.628	2.920	2.802	2.783
$\boldsymbol{V_{DZaway}}$	2.005	0.881	2.920	1.935

Table A.2.: Optimized parameters of our Combination Model (CM). The Zone Localization values mark the x-coordinates, where the border for the corresponding zone is located on the field. $\boldsymbol{V_{AZ}}$ describes the best velocity within the Attacking Zone, where our MMM is applied. $\boldsymbol{V_{DZ}}$ defines the best velocity for the Defending and Midfield Zone, where our ZDM is applied.