Politecnico di Torino

Campus Città Studi Biella

2nd level Master in Sports Engineering

Evaluation of referees' performance in international football games

Master Thesis

Author

Morgana Lalli

Supervisors

Professor Diego Tomalino Regruto

Mr. Daniele Berrone

Professor Gianluca Rosso

Professor Luca Malfatti

a.y. 2018/2019

Table of contents

Introduction

Football is one of the most loved and played game all over the world. According to the International Football Association Board (IFAB), that establishes the laws of the game, its strength lies in the way it is played. Indeed, the game rules are the same for all football matches throughout the world, from the FIFA World Cup™ Final to a game between children in a remote village. This unique mode of playing is a chance which must be taken advantages of for the *good football* everywhere [1].

Rules are fundamental to keep the game 'fair', to preserve the sporting spirit, to create a safe environment for the players and to produce a beautiful game for fans. As reported on IFAB website: "The best matches are those where the referee is rarely needed as the players play with respect for each other, the match officials and the Laws. The integrity of the Laws, and the referees who apply them, must always be protected and respected. All those in authority, especially coaches and team captains, have a clear responsibility to the game to respect the match officials and their decisions" [1].

Moreover, on IFAB website, it is possible to read: "Football's Laws are relatively simple, compared to other team sports, but as many situations are 'subjective' and referees are humans (and thus make mistakes) some decisions will inevitably cause debate and discussion. For some people, this discussion is part of the game's enjoyment and attraction but, whether decisions are right or wrong, the 'spirit' of the game requires that referees' decisions are always respected" [1]. To prevent discussions and make the decision less subjective as possible, over the Laws of the game which explain the behavior of officials in the match, IFAB has an entire specific and extra section for referees, a practical guideline for match officials.

Therefore, those statements underline how much the referee role is crucial – being the *judge of the game* – because on one hand he is exposed to criticisms for his choices and on the other hand he must be well-prepared, both psychology and physically, in order to take reliable and split-second decisions. Moreover, refereeing activities at highest levels must grow continuously, keep up with the demands of the modern-day football games [11]. Consequently, the evaluation of referees' performance by referee governing bodies is a very important point.

Love for the game and knowledge of the rules are not the only things necessary to become a referee at national or international level. For example, the Fédération Internationale de Football Association (FIFA) runs several specific tests on field to evaluate the official's fitness level; they must pass these tests every season, e.g. performing a $2 \times 50\ m$ and $2 \times 200\ m$ sprint tests, followed by a $12\ min$ run [8]. Also many aspects of officials' behavior tests are evaluated, throughout psychological tests. For the evaluation of referees' performance, as reported in Regulations on the Organization of Refereeing in FIFA Member Associations, "the Referees Committee shall establish a Referee Assessor panel to analyze the performance of the Referees in the matches in their category. [...] Consequently, the Referees Committee shall establish a Referee report template to record the relevant information and marks of the Referees." [2]. So, for each referee, a qualitative assessment of his performance must be done.

The criteria on which referees are evaluated, and on which talented officials are identified, are [2]:

- Age;
- Physical conditions;
- Technical skills;
- Psychological preparation;

- Health;
- Education.

Also considering another one of the greatest football organizations, the Union of European Football Association (UEFA), the same attention is put on refereeing activity. Indeed, due to the speed and fast movements of today's top-level football players [5] and due to the constant growth of media attention on field's actions, UEFA requires to its referees to [11]:

- Be well-prepared on the laws of the game;
- Have high level of physical fitness and tactical acumen;
- Have the mental strength to endure pressure;
- Be able to take split-second and reliable decisions.

As FIFA Referees Committee, also UEFA Referees Committee deals, promotes and supports refereeing activities. It manages two major referee gatherings each year, one of them before the new season, focused on fitness training and preparation. Fitness tests are used by national and international referee governing bodies as part of the match selection criteria for the following seasons [7]. Fitness level, analysis on match situations and reliability of decision-making are mainly considered during the gathering courses, promoting the exchange of knowledges among referees. These courses are based primarily on intensive and continuative training sessions over also on the analysis of video recordings for the decision-making process [11].

Referees' performance analysis

To a football referees it is required to keep up with play to ensure optimal positions for his key-decision making role [7]. FIFA, in Law of the Game, gives some general suggestions to referees for their positioning and conduct in different situations, as: "The best position is one from which the referee can make the right decision. All recommendations about positioning are based on probabilities and must be adjusted using specific information about the teams, the players and events in the match up to that point" [12]. On the other hand, it must be considered that is not possible to analyze performance without considering the rhythm of the match; a referee's lower physical performance could indeed be balanced out by official's experience and capacity to foresee the game.

Concerning this and the extreme pressure of thousands of fans and players under which referees must take spilt-second decisions, Pierluigi Collina, voted World's Best Referee of the Year six times, replays:

"*As a referee you have to be very self-confident. And you have to make sure you do everything you can to be perfectly prepared for the job. [...] For a referee, preparation means a lot of things, not just exercising, not just knowing the rules of the game. [...] Also, it is very important to improve your own performance, particularly if you're aiming to stay ahead of the pack."* He says. "*One of the things that have really changed in football is speed. In thirty years, it more or less tripled. [...] Tactics have changed completely. So as a referee, you need to know exactly where you have to be in order to make a good decision. And if you're going to be in the right place at the right time, you have to be able to anticipate what's going to happen. [...] During the match you mostly draw on your concentration and your experience*" [3].

With respect to the referee's physical preparation, researches do not agree on which are the key parameters to assess the performance, even if some elements are common in different studies; the most important and common key parameters are: speed (especially the amount of high intensity running (HIR)

[5, 8, 10]), total distance covered [5, 8, 10], referee distance from infringements [10], aerobic energy production and heart rate (HR). Collecting these data, it is possible to evaluate the performance looking at how they change in time, interpreting them as index of fatigue; obviously high match-to-match variability of the key features ad age-related physical decline must be considered when interpreting physical performances [7]. However other studies [9 – 10] investigate and suggest that clever positioning and anticipation of future events are fundamental elements to evaluate referee's performance. Moreover, the profiling of officials' match activity should be made in the context of the players' acts.

Referee governing bodies does not examine officials during regular matches but in specific test locations. This is correlated with the needs of the committees to evaluate referees' performances before the beginning of the seasons and choose the trained and more prepared officials for future games. Even if studies have shown that these tests used by FIFA and UEFA are not strictly correlated with match activities [5, 6, 8]. Some recent fitness tests introduced by FIFA, based on intensive and repetitive training session, are more appropriate for officials' preparation [5, 7]. A large amount of overall training time is spent by football officials developing their physical ability, since researches proved that high-intensity (>85% maximal heart rate) training protocols are effective for improving fitness and match running performance [7]. However, the performance standards of these tests have not been totally validated and the scientific basis behind them is not clear.

Besides, it must be considered that, even if some researches evaluate referees' performance in regular matches, this is generally done throughout quite invasive methodologies. The main ones of them are:

- Heart rate (HR) monitoring, expressing performance behavior as a percentage of maximal heart rate (HR_{max}) [4 – 6, 10]. This is measured with HR chest strap that referee has to wear during all the match;
- Finger-prick blood samples at the end of each time to analyzed lactate and glucose concentration [6, 10];
- Video recording of officials' performance from which distance run, speed and HIR trends are extracted [4, 5, 6, 8, 10]. Moreover, from videos and questionnaires the decisions made by the match officials are analyzed. [4 – 6, 8 – 10].

Results obtained in this way show than top-class referees have not constant performances and they experience fatigue during the game, in different phases and not only in the end [4 – 10]. It has been observed that the total distance covered during matches is not the best index of physical load imposed on the officials; better ones are the amount of high intensity work (related to HIR) and HR_{max} [4 – 10]. HR is strictly related to physical stress since it is an index of oxygen transport and muscles' activity.

Nonetheless these parameters are partially affected by the contest in which they are measured; Helsen and Bultynck [5] have shown that HR is conditionate by psychological component and 'stress hormones' as adrenaline. Indeed, pre-match HR is much higher (about 10% of HR_{max}) than HR before the begging of second half; this difference is not totally due to the warming-up because, both before the kick-off and during the break, referees have to go back in the dressing room for the same amount of time [5].

Even the amount of HIR could not be the best parameter to quantify official's activity all over the match. Truly, it is evaluated on the entire match and it is not automatically correlated to the official's positioning and speed at the moments of infringement. HIR also does not consider the circumstances in which referee whistles; for example, a skilled referee with strong ability to foresee the game, can jogging all the

time and starts running only when needed, whistling with an unobstructed view. An official is expected to be close to the game and to follow it without interfering [5]. In this case his amount of HIR can be lower than the one of other referees but this is not a sign of a worse performance. Kinematic activity profiles of top-class officials can be influenced by the distance covered by the ball [4]; then “slow matches” could affect the evaluation of referee’ performance if it is only based on the amount of HIR. Other studies [10] support this interpretation, underlining that the total distance covered could be affected by a clever position of the referee; Therefore, low distance traveled is not automatically a bad performance [10].

HIR Is generally obtained from video recording of the games, but also referee's positions, distance covered and speed are extracted from them [5, 8, 10]. In particular cases [8, 13] a bi-dimensional photogrammetric system with fixed camera is used to ensure high-resolution of data; the displacement and linear velocity of the referee is computed with spline polynomial functions, but more reliable data can be obtained using specific tracking algorithms [13]. Specific cameras placed to cover the whole field, allow indeed to obtain the positions of the referee, the players and the ball in a non-invasive way. This fact is very important because the data are extracted without being influenced by any psychological effect, since the referee is not aware that his performance is recorded.

All researches now are focused on the analysis of physical performance and fatigue management as well on the accuracy of the decision-making of the referees, but they do not generally consider the relative conditions of the referees. For example, they do not take into account the obstacles in officials’ view as well as the potential usage of VAR system, a specific support in refereeing.

VAR: Video Assistant Referee

In the modern-day football, since 2016, IFAB has approved the usage of the Video Assistant Referee system (VAR). This is a hawk-eye system that allow to review specific game actions during the match from which referee can modify his judgment in case of error. This system is already widely used in other sports, demonstrating to improve the reliability and accuracy of arbitration judgements [18]. Unlike other sports (e.g. tennis), football is subjected to a certain degree of uncertainty in establishing infringements; for this reason the introduction of VAR system has been widely criticized [14]. Truly, VAR has caused conflicting opinions [15]; on one hand there is someone who see in it a system to reduce the possibility of errors committed in the field while, on the other hand, there is someone else who see in VAR a killer of the emotion related to football games [15, 18].

The usage of VAR in football, as indicated by IFAB, was introduced to "correct clearly incorrect decisions or report serious unseen events to the referee" [16]. The system is based on a set of several cameras placed to cover the entire playing field; they record the game, broadcasting the data in streaming to a staff of expert football referees, composed by one VAR and three Assistant VARs (AVAR) [16]. After an event occurs, or if requested by the on-pitch referee, the VAR team checks the action and informs by earpiece the on-pitch official with the suggested measure. Then, stopping the game, the referee can either directly apply the suggestion or recheck the action through a specific monitor on the pitch's side line, confirming or changing his previous decision [16].

The infringements for which VAR is used are related to goals, penalties, direct expulsions and exchanges of people [16]. Another strong critique to this system, besides increasing the down time of the game, is its only partial 'automatization'. VAR is still based on 'human eye' and so the error possibilities is not totally avoided. [17].

Despite these criticisms, FIFA tested the system in the 2016 FIFA Club World Cup Japan and after some further examinations decided to adopt VAR in 2018 FIFA World Cup Russia, supported above all by its president Gianni Infantini. In fact, before the 2018 FIFA World Cup, he stated:

"I am, of course, very much in favour of the introduction of VAR: I was very sceptical two years ago, but we tried and saw that it works. The statistics say that we have gone from 93% of referees' decisions accuracy to around 99%. If we want the best football competition in the world to be arbitrated in the best possible way, we must help the referees and certainly with the VAR we can do it" [19].

Additionally, even the sport scientist and training expert for UEFA, Professor Werner Helsen, who was asked to investigate on VAR, underlined its benefits saying: "it's changing refereeing, it's changing the game" [14].

Despite the forecasts, given its recent introduction, there is no researches on how and how much the introduction of VAR modifies the football arbitration.

Research purpose

The performance evaluation and fatigue management of the referees in international games is a point of great interest, since it is where the maximum performances are required.

The aim of this work is to realize a simple method for the evaluation of the physical performance of referees in international games, analyzing the data collected in a non-invasive way of the matches of a given season. Total distance covered, speed, positioning and their changes during the game are described, as well as the conditions of the referee when he whistles. To *quantify* the performance assessments, the variability of the performances among referees and the "mean referee of the season" are extracted; the last one is obtained averaging the data of the all studied parameters. Thus the behaviour of individual referees is compared whit the one of the virtual official representing the "mean referee of the season" and an assessment is made. Finally, this work aims to esteem the impact of the introduction of VAR in football, evaluating if and how much it affects referee's behaviour and judgment.

This work was carried out at the Deltatre company during the curricular internship of the 2nd level Master in Sports Engineering of the Politecnico di Torino. It was conducted under the supervision of Mr. Daniele Berrone (Deltatre) and Professor Diego Regruto Tomalino (Politecnico di Torino), with the support of Professor Gianluca Rosso and Professor Luca Malfatti, of the Italian start-up Quant4Sport. Especially, the analyses were carried out thanks to a database of real data of international level football matches provided by Deltatre company, which have been properly anonymized.

Materials & methods

To evaluate the referees' performances, forty-six male international referees (40 ± 4 years old, 181.5 ± 4 cm height and 77 ± 5 kg weight), with different degree of experience at international level (10 ± 3 years) and hundred and forty-eight matches have been analyzed. From these data, both physical condition and technical skills of the referee have been examined [5, 7 – 10], looking at the total distance covered, the mean speed and the behavior of referee during the infringements [5, 8, 10]. Therefore, the trends of the speed and distance of the referee from the ball have been computed taking into account the number of players placed in the official's field of view. These parameters have been considered as indices of the "conditions" of the referee when he whistles an infringement, evaluating his behavior in the key moments

of the game and his fatigue management [7, 9 – 10]. The events studied are fouls, whistle-breaks and offsides, even if in offsides a key role is taken by assistant referees.

Tracking data

The information needed for this work has been extracted from data obtained recording the games through a tracking system, with the referees and the players not wearing any equipment and so without changing their behavior [8, 13]. This system is made up of two sets of cameras (each one consisting of three cameras), installed on the stands to have a clear view of the entire field. This system is semi-automatic and requires an initial calibration for matching the objects to track (ball, players and referees); then, it automatically maintains the tracking by itself. The pitch origin is set at the middle of the field: the X axis along the field edge lines and the Y axis parallel to the goal lines, as showed in the picture below (figure 1). Moreover, only for the ball, the Z component is determined as the over-elevation with respect to the pitch's height; the measure of Z is based on the stereoscopic effect or by algorithms, relating to the chosen provider. All the coordinates are provided in centimeters (e.g. the standard pitch size is 10500x6800 cm).

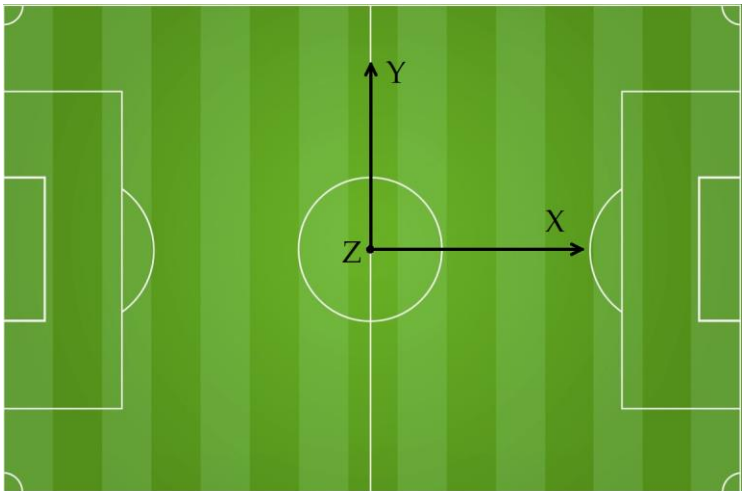


Figure 1. Tracking system: how it is placed the Cartesian reference system.

The raw data, returned by the tracking system, must be scrubbed because they can be affected by multiple errors. Indeed each object is compared with the same object of the previous frames and the position series are resampled, filtering the less significant measures. Then the objects are classified as reliable or unreliable. This operation considers also a possible mismatch between the tracking data of two players due, for example, to a collision. If the data of an object are considered unreliable or if there is a momentary loss of them, due to an error in the video recognition or to a delay of the operator in reacting to the changes, a gap in the positions of the objects is generated. This gap can be bridge only in some cases, reprocessing the measures collected backwards and assigning the data correctly to the players. Then the files with the reliable tracking data are produced. These are a set of tracking frame files (25 per second), consisting of lists of the tracked objects (ball, players and referees) and the timing in which these data were collected.

In parallel with the tracking of the positions, through gathering information, there is the creation of another file. In this file, all the events of the match and the main (summary) information are chronologically collected specifying the event's type, the position of the ball and the players involved as well as the event's timing. Knowing the timing of an event it is possible to match the tracking information

of all the objects during the infringement, comparing the event's timing with the one in tracking files. Therefore, all the necessary data are obtained and ready to be analyzed.

Matches analysis

The first step of the study consists in analyzing the games, individually. From these analyses the parameters for the arbitrator' evaluations have been extracted and the "virtual match official" (the mean referee of the season) computed. To get these elements all the data have been processed using the Python programming language and the associated software, through ad hoc algorithms created for this study.

Analyzing the single matches, the referees' behaviors have been strongly characterized, evaluating also the different types of official's race (walking, jogging, high-speed, etc.). Truly, to assess extensively the referee's behavior, not only the key moments positioning has been considered but also the position in the previous 15 seconds. The choice of a 15 seconds time range is experimentally motivated, as the average time interval to perform an action.

Therefore, the following information have been extracted for each match:

- Distance between the referee and the ball during the events and its trend;
- Initial referee-ball distance and its behavior;
- Average referee-ball distance for fouls-type events;
- Average referee-ball distance for whistle-breaks type events;
- Average number of players in the referee's field of view during the events;
- Average referee's speed at the event and its trend;
- Referee's speed characterization and average speed of the game;
- Referee's total distance covered;
- Percentage of change in the referee-ball distance and official's speed during the events. This parameter is measured at the end of each half of the game, with respect to the initial value.

To obtain these variables, the interest events have been recognized in the summary game, saving their timing and, fundamental for analyzing the official performance, the corresponding tracking data of the referee, the ball and the players have been identified. The events identification took place through specific labels, concerning all types of fouls (e.g. attacking foul, hand foul, etc. ...), whistle-break interruptions (e.g. player replacements), and offsides.

The tracking data directly provide the X and Y coordinates of the all objects. Therefore, simply calculating the Euclidean distance from a position to the next, it is possible to establish the referee's displacement. Then, adding all the consecutive displacements, the total distance covered by the official has been obtained. To reduce computational times, only one object's positions for second was considered, so the total displacement was determined as the sum of the successive distances travelled each second by the referee. From these displacements is also possible to extract instantaneous and mean speed in the game. Truly, it was firstly computed the three-seconds-average speed to check the reliability of the tracking data. This speed has been chosen as more informative and easier to check the raw data. In fact, if this was higher than 12 m/s, data were considered unreliable and so rejected. The threshold equal to 12 m/s has been selected because previous studies have shown that referee's maximal-speed sprint is slightly higher than 7 m/s [10] and, therefore, a slightly higher threshold has been chosen. On the reliable data the searched parameters have been extracted, including the total distance travelled.

After that, still using the Euclidean distance, the distance between the referee and the ball has been computed and the operation is repeated for all the 15 seconds preceding the event. Then, to determine the average speed of the referee when he *follows* an event, the total distance covered by the referee in the 15 seconds preceding the infringement is divided by the time range. Thus, considering fouls, whistle-breaks and offsides, the behaviours chronologically ordered of the initial referee-ball distance, the referee-ball distance evolution and the speed at the events, have been got.

Using linear regression, the trends of these parameters and their percentage of variations at the end of each game time (first and second half and any additional extra times[1]), with respect to the initial value, have been computed; this information quantifies the variations of their behaviours and so how the referee handles fatigue.

In addition, from the X and Y coordinates of the objects, the straight line passing by the referee and the ball is considered and is used as initial step to calculate the number of players in the referee's field of view. This line was determined as:

$$\frac{y - y_r}{y_r - y_b} = \frac{x - x_r}{x_r - x_b}$$

Where y_r and x_r represent the X and Y coordinates of the referee while y_b and x_b these of the ball. This line is can also be written as:

$$y = mx + q, \text{ where: } m = \frac{y_r - y_b}{x_r - x_b} \text{ and } q = \frac{y_r - y_b}{x_r - x_b} \cdot x_r + y_r \quad (1)$$

The *field of view* indicates the portion of space that can be perceived by an observer; it is approximated as an elliptical cone with variable angular opening, based on the peripheral vision considered. Generally, in the biocular vision, this cone is characterized by approximately 95° horizontal and approximately 80° vertical opening, but it can reach 130° – 135° vertical and 200° – 220° horizontal [20] (figure 2). Despite multiple studies have demonstrated the importance of peripheral vision in different sports including football [21], in this study the focus has been placed only on foveal vision; the reason is because foveal vision refers to vision in the *centr*e of the field of view, where visual acuity is at its highest.

Considering and analysing the entire biocular referee's field of view, distorted results could be obtained; indeed, in the 'entire field of view' of the official usually there are many players, but not all of them disturb the attention of the referee. The referee follows the game but he does not have to interfere with it; the aim of the research is to determine the real number of obstacles on the trajectory between the referee and an event when he whistles and makes a decision. Therefore, assuming that the referee *looks* at the action, his central view is considered and analysed. The foveal and parafoveal visual cones are very narrow, of about 10 ° [21] (figure 2), and for simplicity it was chosen to represent them as a rectangle centred on the referee-ball's line (1) and characterized to be included between the two lines (2)

$$y = mx + q \pm \delta\,, \quad (2)$$

where δ is equal to $2\ m$. In this range all the players present during the event were searched looking at the event's time and players' coordinates. Then, averaging the number of obstacles in all the events, the match average number of players in the official's field of view is computed.

[1] In all the games analysed only one presented extra time, so the analysis was focused only on first and second half.

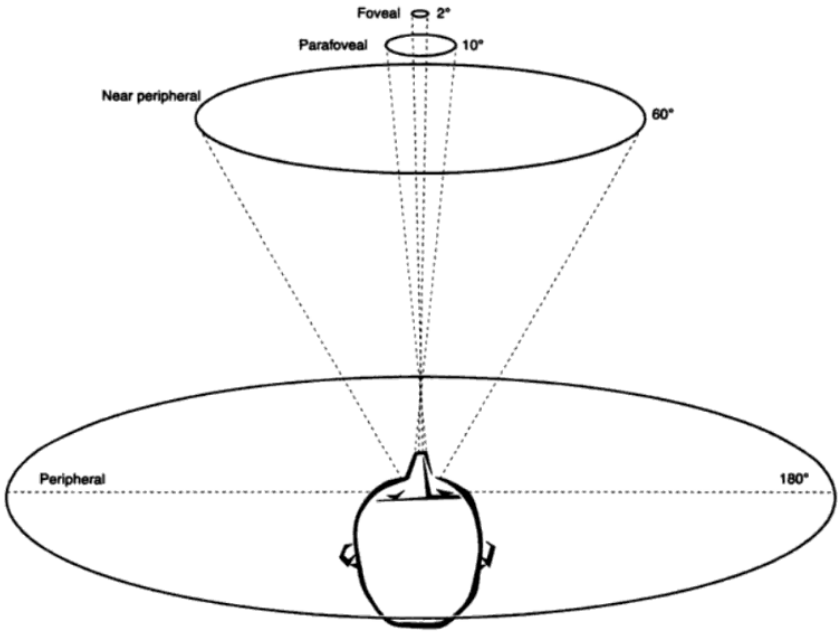


Figure 2. Cone of vision showing foveal, parafoveal, near peripherical and peripherical vision [22].

Instead, from the referee's speeds throughout the game, both the mean game's speed and the velocities characterization are extracted. The three-seconds-average speed has been considered and divided in five locomotive categories, that are:

- Standing & walking: 0.2-2.77 m/s (0.72-9.99 km/k);
- Low-speed jogging: 2.78-3.74 m/s (10-13.49 km/h);
- Medium-speed running: 3.75-4.99 m/s (13.5-17.99 km/h);
- High-speed running: 5-6.94 m/s (18.0-25 km/h);
- Maximal-speed sprinting: > 6.95 m/s (> 25 km/h) [10].

Finally, for each match, the rhythm of the game is evaluated to understand its correlation with the referee's performance. The total distance covered by the ball throughout the game has been chosen to characterize the rhythm of a game, because it is a summary way to represent this parameter but also because the official's performance is strictly correlated with the behaviour of the ball [4]. The ball's speed has not been chosen because very long or very short passages can alter the *real rhythm*, even if the total distance covered by the ball is not a parameter totally free from this problem.

VAR analysis

In thirty matches of the hundred and forty-eight ones analysed, the VAR system supported the referee's activity.

To analyse the events related to VAR, a methodology like the one used for study fouls, whistle-breaks and offsides events has been used. Always starting from the summary file of the game, through specific labels, the VAR calls and their related events have been identified. Then it was analysed how many times the VAR calls have modified the referee's decisions; without considering the outlines of the distribution, two ratios have been measured to quantify VAR importance in refereeing: the VAR calls over all arbitrated events and number of times that official has changed his mind over VAR calls.

Referees analysis

To give a simpler, immediate and more meaningful interpretation of the referee's activity, from each game, only some parameters of the ones extracted have been considered. The aim of the research was indeed to create an easy, informative and intuitive system that would allow to analyse the officials' performance. For an immediate interpretation, the system implements a 'traffic light' format of the data; comparing the single referee behaviour with the one of the virtual official, a judgment related to a colour scale is done. The scale colour is characterized such that in green is described a better parameter than the one of the virtual official, in white one with equal value and in red a pejorative parameter.

Accordance with previous studies [5 - 10], both elements closely related to the referee fitness level and parameters typical of his role have been considered evaluating the performance. These parameters are grouped into two clusters: a main set and a secondary one.

The main parameters chosen are:

- The average referee-ball distance at fouls;
- The average referee-ball distance at the whistle-breaks[2];
- The average number of players in the referee's field of view during the events;
- The average speed of the official throughout the game;
- The total distance covered.

The secondary parameters instead concern the behaviour of speed and distance between referee and ball at the events. The trend is represented through the initial values and the percentage changes at the end of each half time.

Both the elements of the main and the secondary set are put into a table. Then, for each field, the average and the standard deviation have been calculated. Averaging the data, the virtual referee has been computed while, through the standard deviation, an information on referees' performance variability is acquired. Then the coefficient of variation (CV) has also been calculated, to compare the variability of the different parameters and the maximum differences in the behaviour of the referees have been computed. Placing μ and σ as the mean and the standard deviation of a parameter, respectively, CV is defined as:

$$CV = \frac{\sigma}{\mu} \tag{3}$$

At a first look of the data, it seems that there is a relationship between referees' experience at international level and the referees' speed, the total distance covered and mean distance at the infringements; the correlation between the experience level (years of service) and the variables of the main set has been computed with the correlation coefficient, thus defined:

$$correl(X,Y) = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2\sum(y-\bar{y})^2}} \tag{4}$$

[2] The mean referee-ball distance at fouls and whistle-breaks events was extracted considering separately these two types of infringements. The offsides were not took into account because in their determinations a fundamental role is covered by the assistant referees.

Where x and y are the vectors of data implemented and $\bar{x}$ and $\bar{y}$ their averages. To verify the presence of general trends among all the referees, it has been calculated also the correlation coefficient between the parameters of the main set of variables and the rhythm.

After that, the traffic light format has been implemented with respect to the table containing the summary information of each game. Using the Microsoft Excel software, the formatting has set for each parameter (column of data), the red colour to the worst value, the white to the average value and the green to the best one; then it has been applied a colour scale on each cell of each column. Therefore, the virtual referee is characterized by the white cells.

Finally, a report for each referee is created containing the studied parameters for each game he has arbitrated for the evaluation of his performance. The report is scheduled to contain a box for personal data (such as birth, years of experience, etc.) and three other sections. The first section summarizes a radar chart of the main parameters, compared with the ones of the virtual referee; to simplify the interpretation, the data are normalized and represented as percentage differences, setting all the virtual referee's values to zero (0%, reference value). This is followed by an explanatory comment of the results. The second section is instead used to represents the trend of the secondary parameters using two dispersion graphs, always comparing the single referee and the virtual official. The third one reports instead all the tabular data of all the games arbitrated by the referee coloured with respect to the traffic light criterion, previously described; it also contains, for comparison, the virtual referee information and the maximum and minimum values of all the referees' team.

Results

Match and referee analysis

The first step of the analysis studies the arbitrated matches individually. Only one hundred and thirty-seven games have been analyzed of the one hundred and forty-eight ones considered, since games with problems in tracking data for at least half of the game have been discarded. No matches characterized by VAR usage have been discarded. For each game analyzed one report is created, extracting the parameters of interest and filling the summary table for evaluating the referees' performance (tables 1 and 2).

The recapitulatory table, containing also referees' personal data, is characterized by the traffic light formatting. From the data extracted from summary tables (tables 1 and 2), it is possible to observe both a certain variability in the officials' behaviors and during the different games that a single referee has arbitrated. To quantify these dispersions and *assess* the referees' performance, the virtual official, standard deviation, maximum and minimum values are collected (table 3). Then, also the coefficient of variation (CV) is computed for all the parameters, underlining where the maximum differences exist.

These differences are observed in the speed and distance events' trends, the elements that theoretically are related to the fatigue management and the level of experience; indeed it has been assumed that lower physical referee's performance is equalized by a clever position and a higher ability to foresee the game.

But these parameters can be also related to the pace of the game. Hypothetically in match with high pace the total distance covered by the referee should be higher than the one of games with lower rhythm, due to the greater movements of the ball.

Referee	Birth	International from	Height (cm)	Weight (kg)	Teams	Dist Fouls (m)	Dist WhistleBreaks (m)	n players
1	1979	2010	189	84	Team1 - Team2	16,0	18,4	2
1	1979	2010	189	84	Team3 - Team4	18,5	20,9	2
1	1979	2010	189	84	Team5 - Team6	22,1	21,0	2
2	1978	2011	197	85	Team7 - Team8	18,5	17,7	2
2	1978	2011	197	85	Team9 - Team5	13,8	18,5	2
3	1983	2011	173	62	Team10 - Team11	14,0	16,7	1
4	1975	2007	185	78	Team12 - Team13	18,7	19,5	1
4	1975	2007	185	78	Team14 - Team15	16,5	16,2	1
4	1975	2007	185	78	Team8 - Team7	17,2	18,3	1
4	1975	2007	185	78	Team5 - Team16	17,7	19,6	2
4	1975	2007	185	78	Team17 - Team8	14,9	17,9	2
4	1975	2007	185	78	Team16 - Team18	14,6	18,3	1
5	1977	2011	183	78	Team16 - Team13	18,4	17,9	1
6	1976	2006	176	68	Team19 - Team1	17,0	21,0	2
6	1976	2006	176	68	Team6 - Team20	16,0	15,3	1
6	1976	2006	176	68	Team18 - Team21	14,3	15,9	1
6	1976	2006	176	68	Team22 - Team17	12,9	15,7	1
6	1976	2006	176	68	Team23 - Team20	13,8	14,0	1
6	1976	2006	176	68	Team24 - Team2	13,9	16,9	1
6	1976	2006	176	68	Team8 - Team18	16,7	16,5	1
6	1976	2006	176	68	Team19 - Team17	16,6	16,6	1
7	1979	2006	174	70	Team3 - Team25	17,4	16,2	2
7	1979	2006	174	70	Team11 - Team10	14,3	14,0	1
7	1979	2006	174	70	Team15 - Team14	14,1	14,1	1
7	1979	2006	174	70	Team2 - Team1	13,9	15,4	1
8	1976	2013	177	70	Team16 - Team26	15,5	17,8	2
8	1976	2013	177	70	Team27 -Team28	14,5	15,3	1
8	1976	2013	177	70	Team11 -Team28	13,5	16,9	2
8	1976	2013	177	70	Team16 - Team15	13,6	13,7	1
9	1985	2013	187	83	Team29 - Team30	15,2	17,6	2
10	1987	2009	-	-	Team25 - Team3	14,7	15,8	2
11	1984	2014	179	73	Team11 - Team23	15,3	9,7	1
11	1984	2014	179	73	Team31 - Team2	16,7	19,5	2
12	1981	2012	-	-	Team26 - Team13	18,6	19,3	2
12	1981	2012	-	-	Team5 - Team9	15,8	16,1	1
13	1981	2008	181	74	Team6 - Team5	19,8	19,8	2
13	1981	2008	181	74	Team17 - Team21	15,9	17,7	1
13	1981	2008	181	74	Team15 - Team8	17,2	17,1	2
13	1981	2008	181	74	Team23 - Team11	17,2	18,4	2
14	1980	2007	184	84	Team32 - Team26	17,8	17,7	2
15	1986	2013	-	-	Team14 - Team8	17,5	18,1	1
16	1979	2010	-	-	Team15 - Team7	15,5	16,6	1
17	1975	2003	189	80	Team31 - Team19	15,8	16,8	1
17	1975	2003	189	80	Team29 - Team27	13,7	15,6	2
18	1980	2011	-	-	Team27 - Team30	15,6	19,4	2
19	1984	2010	-	82	Team14 - Team7	15,8	16,8	1
20	1977	2005	185	81	Team32 - Team16	15,4	17,7	2
21	1977	2005	185	81	Team33 - Team17	14,2	16,8	1
22	1984	2011	183	77	Team25 - Team4	18,0	20,3	2
22	1984	2011	183	77	Team18 - Team33	16,8	18,8	2

Table 1. Extraction of summary table of all the games analyzed (Part 1). The fields represented are: 'Referee', the ID of the official in this study; 'Birth', the officials' birth year; 'International from', the years in which the referee could arbitrate in international level matches; 'Height' and 'Weight' the physical characteristics of the referees; 'Teams', the teams that faced each other; 'Dist Foul' and 'Dist WhistleBreaks', the referee-ball distance in foul and whistle-breaks events, respectively; 'n players', number of players in referee's field of view at the events.

V mean (km/h)	V_in evt (km/h)	Dist_in evt (m)	V end 1 (%)	Dist end 1 (%)	V end 2 (%)	Dist end 2 (%)	Dist tot (km)	Rhythm	VAR
6,2	6,4	15,9	4,4%	-12,4%	9,3%	-26,2%	9,7	29,3	No
6,2	5,3	17,0	-5,9%	-18,3%	-11,0%	-33,9%	10,0	31,9	No
6,1	8,8	23,0	-11,3%	5,9%	-29,2%	15,2%	9,6	30,2	No
6,2	8,3	20,7	-18,9%	8,4%	-52,4%	23,2%	10,0	29,0	No
6,9	9,4	17,1	-23,6%	-0,2%	-46,5%	-0,4%	11,3	31,6	No
6,4	6,9	16,9	-10,1%	6,6%	-19,5%	12,8%	10,5	33,5	No
6,5	7,9	18,2	-18,8%	-4,6%	-35,8%	-8,8%	10,0	28,7	No
6,2	6,8	15,5	11,2%	-3,7%	25,3%	-8,4%	10,4	30,1	No
6,5	7,0	21,9	-1,5%	21,4%	-3,0%	42,8%	11,7	30,2	No
6,5	9,2	18,4	-20,4%	-4,5%	-38,6%	-8,5%	10,6	29,8	Sì
6,5	6,0	16,1	15,4%	-9,9%	28,1%	-18,0%	10,3	27,4	Sì
6,5	6,8	19,6	-8,4%	13,3%	-18,6%	29,4%	11,0	31,2	Sì
6,5	6,0	19,0	-0,1%	4,9%	-0,3%	10,1%	10,3	30,9	No
6,2	8,4	18,6	-18,2%	-4,2%	-38,8%	-9,0%	9,8	29,5	No
6,9	6,5	10,9	3,6%	-37,6%	8,0%	-82,4%	10,9	34,5	No
6,3	7,6	13,5	-15,8%	-11,4%	-36,1%	-26,0%	9,9	27,9	No
6,8	7,5	15,0	-10,0%	3,4%	-21,7%	7,5%	10,7	28,1	Sì
5,9	8,4	10,5	-23,9%	-28,1%	-46,6%	-54,6%	13,4	42,8	Sì
6,1	6,3	18,3	-3,1%	12,4%	-10,3%	40,9%	9,9	30,6	No
6,1	8,1	19,9	-8,4%	11,3%	-26,0%	35,1%	10,3	33,0	Sì
6,4	5,0	13,6	14,0%	-17,0%	33,5%	-44,5%	10,6	32,9	Sì
6,7	7,5	15,8	-13,5%	-5,4%	-24,3%	-9,7%	10,7	31,3	No
7,8	8,8	14,1	-12,8%	-3,9%	-26,9%	-8,2%	12,0	30,5	No
7,4	9,1	12,7	-21,0%	-8,5%	-51,0%	-20,7%	11,7	29,7	No
6,9	8,9	12,9	-22,9%	-13,7%	-53,8%	-32,1%	11,1	30,6	No
6,9	8,8	16,6	-10,0%	-1,7%	-22,0%	-3,7%	10,7	27,6	No
7,6	8,9	16,2	-4,3%	4,3%	-11,1%	11,3%	12,2	31,6	No
7,0	8,7	12,1	-32,4%	-39,3%	-48,6%	-58,9%	11,9	30,3	Sì
7,5	6,0	14,8	-7,9%	9,9%	-14,6%	18,3%	11,9	28,3	Sì
6,9	8,0	14,3	-10,4%	-23,9%	-18,6%	-42,7%	11,0	29,0	No
6,7	8,2	17,2	-8,0%	7,7%	-19,8%	19,0%	10,5	29,2	No
6,5	8,1	15,2	-10,0%	-10,9%	-26,6%	-29,0%	10,3	29,1	No
7,7	9,6	18,5	-25,4%	-1,4%	-66,0%	-3,7%	12,5	36,9	No
5,9	9,1	17,0	-24,9%	-8,8%	-51,3%	-18,1%	9,8	29,1	No
6,7	7,9	18,1	-14,9%	14,1%	-22,2%	20,9%	10,5	30,5	No
6,0	7,5	19,1	-14,3%	-5,3%	-20,2%	-7,5%	9,4	36,0	No
6,1	6,7	16,8	-13,4%	0,9%	-30,2%	2,0%	9,8	30,7	No
6,0	5,9	19,1	-8,8%	8,2%	-22,5%	20,3%	9,9	35,0	No
5,7	7,5	20,5	-19,6%	11,1%	-46,9%	26,5%	9,3	32,5	No
6,5	6,1	19,8	6,4%	10,1%	14,2%	22,5%	10,5	30,0	No
7,2	7,9	17,5	-9,7%	-1,4%	-23,7%	-3,5%	11,6	28,8	No
6,3	8,5	13,2	-18,0%	-24,9%	-36,6%	-50,8%	10,0	31,2	No
7,0	8,7	15,1	-22,4%	-10,1%	-43,4%	-19,6%	11,6	35,3	No
7,2	9,0	16,6	-7,4%	5,9%	-20,4%	16,1%	11,1	30,5	No
7,0	9,0	19,8	-4,4%	7,5%	-8,9%	14,9%	11,4	32,8	No
7,0	9,9	17,2	-17,8%	5,5%	-33,7%	10,5%	11,5	28,4	No
7,2	7,1	15,6	-12,4%	-5,3%	-42,1%	-18,1%	10,5	31,7	No
6,4	7,5	13,6	-23,6%	-16,2%	-48,0%	-33,0%	10,1	28,6	No
7,2	8,6	17,2	-10,5%	-7,8%	-20,0%	-14,8%	11,6	33,2	No
6,8	7,4	17,5	7,6%	-4,1%	14,2%	-7,7%	11,2	28,4	No

Table 3. Extraction of summary table of all the games analyzed (Part 2). The fields represented are: 'V mean', average speed during the match; 'V_in evt' and 'Dist_in_evt', initial value of the regression lines of the speed at the events and of the referee-ball distance to the infringements, respectively; 'V end 1' and 'Dist end 1', percentage changes at the end of first half; 'V end 2' and 'Dist end 2', percentage changes at the end of second half; 'Rhythm', total distance traveled by the ball; 'VAR', index that express if there was the possibility to use the VAR system.

Virtual Referee	Dist Fouls (m)	Dist WhistleBreaks (m)	n players	V mean (km/h)	V_in evt (km/h)	Dist_in evt (m)	V end1 (%)	Dist end1 (%)	V end2 (%)	Dist end2 (%)	Dist tot (km)
Mean	15,86	17,33	1,45	6,66	7,79	16,48	-0,13	-0,04	-0,28	-0,09	10,75
Standar deviation	2,26	2,08	0,25	0,46	1,25	3,00	0,12	0,16	0,23	0,33	0,76
Max	24,56	22,06	2,15	7,9	10,36	29	0,2071	0,2336	0,3778	0,4324	13,44
Min	9,85	9,7	0,88	5,7	4,97	10,31	-0,4123	-0,927	-0,7949	-1,4509	9,3
Coeff. of variation	0,14	0,12	0,17	0,07	0,16	0,18	0,86	3,82	0,84	3,85	0,07

Table 5. Mean virtual referee parameters, obtained averaging all the matches of the analyzed season. In addition, to mean values also the standard deviation, maximum and minimum values and coefficient of variation were reported.

Therefore, the correlation between the games' pace and the total distance traveled by the referee has been evaluated; but they were poorly correlated ($correl(pace, total\ distance) = 0{,}11$) so it is not possible to justify the performance variability with the rhythm thus defined. The correlation is then computed between the parameters of the secondary set and degree of experience; but also between these elements the correlation is very low (lower than 0.19).

To extract some general trends in referees' behavior the correlations between the variables of the main set and the years of experience have been calculated (table 4).

Coeff. of Correlation	Dist Foul	Dist WhistleBreak	n° players	V_mean	Dist_tot	Esperience
Dist Foul	1,0000	0,6797	0,2748	-0,1248	-0,1540	-0,0715
Dist WhistleBreak	0,6797	1,0000	0,3935	-0,1623	-0,2077	-0,0669
n° players	0,2748	0,3935	1,0000	-0,1196	-0,1405	0,0205
V_mean	-0,1248	-0,1623	-0,1196	1,0000	0,8322	-0,2409
Dist_tot	-0,1540	-0,2077	-0,1196	0,8322	1,0000	-0,1973
Esperience	-0,0715	-0,0669	0,0205	-0,2409	-0,1973	1,0000

Table 7. Coefficients of correlation between the parameters of the main set and the years arbitrated in international level.

In table 4 it can be observed that higher correlations exist between the average speed and the total distance travelled ($correl = 0.83$) as between distance at fouls and at whistle-breaks ($correl = 0.68$). The table also shows that more experienced referees tend to run less and less quickly ($corr = -0.24$ and $correl = -0.20$ respectively) and that the degree of experience is not related to the number of players in the referee's field of view.

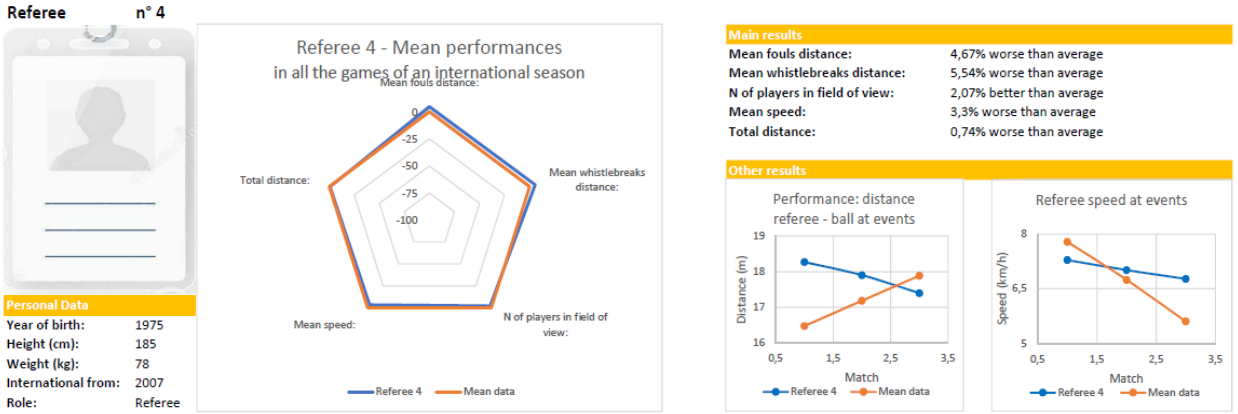
Finally, for each referee, a report is created, containing the studied parameters for each game. The report collects all the elements for evaluating referee's behavior. An example is shown in figure 3.

VAR analysis

To evaluate the contribution of VAR in refereeing, an analysis of its incidence has been done.

Considering only the types of events in which VAR is used, it has been studied both how many times it has been called and how many times its use modified referee's decisions (figure 4).

Then, rejecting outliers, its incidence has been computed as the ratio between the average number of VAR calls and the average number of total events and as the ratio between the average number of events modified by VAR and the average number of VAR calls, respectively. The results show that VAR has been called in 10% of the events and it has brought the referee to change his mind in 12% of these cases.



	Mean fouls (m)	Mean whistlebreak (m)	N of players	Mean speed (km/h)	Total distance (km)	Initial events speed (km/h)	Initial referee-ball distance (m)	Speed: end of 1 time (%)	Distance: end of 1 time (%)	Speed: end of 2 time (%)	Distance: end of 2 time (%)	Match
	18,67	19,48	1,3	6,48	10	7,89	18,17	-18,78	-4,59	-35,77	-8,75	Team12 - Team13
	16,54	16,15	1,3	6,23	10,42	6,79	15,46	11,17	-3,69	25,34	-8,38	Team14 - Team15
	17,17	18,28	1,4	6,48	11,66	7,04	21,93	-1,5	21,39	-3,01	42,78	Team8 - Team7
	17,71	19,63	1,6	6,52	10,64	9,19	18,4	-20,43	-4,49	-38,59	-8,49	Team5 - Team16
	14,89	17,93	1,5	6,45	10,32	6	16,09	15,38	-9,87	28,05	-17,99	Team17 - Team8
	14,59	18,29	1,4	6,47	10,99	6,82	19,55	-8,42	13,27	-18,64	29,38	Team16 - Team18
Referee means	16,6	18,29	1,42	6,44	10,67	7,29	18,27	-3,76	2	-7,1	4,76	
All referee means	15,86	17,33	1,45	6,66	10,75	7,79	16,48	-13,38	-4,29	-28,03	-8,56	
Normalized referee data*	4,67	5,54	-2,07	-3,3	-0,74	-6,42	10,86	-71,9	-146,62	-74,67	-155,61	
Mean norm. data	0	0	0	0	0	0	0	0	0	0	0	
Min value	9,85	9,7	0,88	5,7	9,3	4,97	10,31	-41,23	-92,7	-79,49	-145,09	
Max value	24,56	22,06	2,15	7,9	13,44	10,36	29	20,71	23,36	37,78	43,24	

* Values are normalized between [-100%,100%] where the mean is equal to zero

Figure 3. Example of referee's performance report.

Discussion and conclusions

The purpose of this work is to implement an easy and complete method for the evaluation of football referees' performance; the results, returned as reports, allow an intuitive and exhaustive interpretation of the analysis (figure 3).

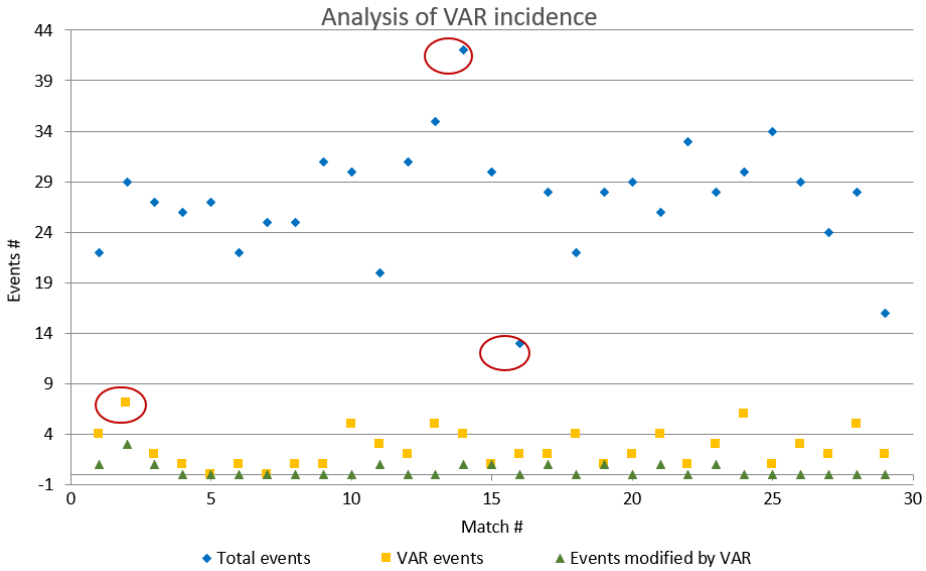


Figure 4. Analysis of VAR incidence: in blue the total numbers of events, in yellow the VAR calls and in green the number of events modified by the usage of VAR.

This method is based on tracking data obtained in a non-invasive way using video recording and they present both pros and cons. One of the main advantages is that performance measured on-pitch and during the season let to obtain more suitable results for referees' evaluation with respect to the usual tests made in specific locations by the referee governing bodies [5, 6, 8]. However, the tracking data can be affected by a problematic: the mismatch of the tracked objects (e.g. referee, ball, etc.); this problem can reduce the quality of the analysis of the referee's performance. Nonetheless, a research study [13] has proved the reliability of the overall methodology.

Another advantage of tracking data is, being a non-invasive way to collect information, that the referee is not conditionate by any psychological effect or by any equipment, then the real performance is recorded. However some researches [4 - 6, 10] consider the HR as a fundamental parameter to appraise the referees' performance, because it permits a deeper analysis in the mechanisms of the muscles that tracking data cannot do; but, on the other hand, HR is usually affected by stress hormones' effects, altering the results interpretation [5]. This misinterpretation is not present in the tracking data method and it is also possible to esteem the environment in which referees act.

According with researches [4 – 6, 8 – 10], a complete analysis of referees' behavior is given both by physical and psychological investigations. Tracking data let to obtain only the physical aspect, the one that referees train most [7]. In this work, like in studies [4, 8, 10], physical parameters have been linked with the conditions in which the referee takes a decision. The physical parameters considered are: total distance covered and average speed [5, 8 - 10]. Indeed, these parameters are strictly related to physical activity level that a referee has to sustain; the total distance covered by a sportsman is an indirect index of his energy consumption [8] and better physical performances ensure to arbitrate better [4]. Other parameters, taken into account to evaluate the refereeing conditions, are: the referee's distance from the ball during the infringements, referee's speed during the events and the number of players in his field of view when he whistles [4, 8]. Furthermore, the trends of these parameters have been extracted, evaluating the fatigue management of the match officials during the game and their different strategies to handle it. In fact, it has been observed that the referees feel fatigue not only in the end of the game and, to maintain the necessary playing level, they carry out different solutions [4, 8].

The development of the evaluating method consists in analyzing all the matches of a set of games (in this work an entire season has been analyzed). Therefore, all the previously described outcomes are extracted from each game and reported in the recapitulatory table (table 1 and 2). This table let to extract the variability in the referees' performance and the virtual mean referee for the comparison with other officials. The choice of a comparison with the virtual referee describes the performance of the officials with respect to an average behavior (figure 3). The report contains all the information of the games arbitrated by the referee considered, reporting both the 'raw data' and a graphical representation (figure 3); the graphs describes data in a brief and intuitive way and are followed by an explanatory comment.

The table illustrates the variability of the referees' performances and common trends. The maximum variability is observed in the values of the parameters characterizing the referee's behavior at the infringements (table 3). The correlation coefficient between the rhythm of the game and the total distance covered by the referee has been computed since it was supposed that the maximum variability depends on the game pace and the total distance is affected by a higher energy consumption [8] in games with higher pace. Nonetheless, no significant correlation value has been found; this probably depends on how the rhythm was defined (i.e. the total distance traveled by the ball) [4]; indeed this parameter can be

afflicted by long passage that increase 'the rhythm' even in low pace games. Then it has been searched if the different level of experience of the referees could justify this variability, computing the correlation between these elements and the years of experience of the subject. No relevant correlations have been found in the fatigue management, even if more experienced referees tend to run less, and less quickly, than the "beginner" ones (table 4). These results show that exist very different strategies to handle fatigue among the referees but also that more experienced ones tend to adopt a tactics concerning a clever position in the field. Studies [9 - 10] have proved that a smart positioning and the ability to foresee the game are fundamental elements of the match officials' behavior, explaining what has been observed in the speed results.

No other common behavior has been found, except for the correlation between the total distance covered and the average speed and the correlation between the referee-ball distance at fouls and whistle-breaks infringements (table 4). These correlation results were expected; even if referees adopt different strategies to handle fatigue, those ones with higher physical performance tend to run more and faster [4 - 6, 8, 10]. Evaluating instead the behavior at the infringements it makes sense, regardless to the fatigue or the ability to predict play, that the performances at fouls and whistle-breaks are linked; the referee has indeed to follow the whole game with the same level of attention.

Analyzing the incidence of VAR (figure 4), it can be observed that its use is quite the same in different games. The VAR system has been introduced in football because, as already proved in other sports, it improves the quality of the arbitration [14, 18 - 19]. Many criticisms have been moved against it because it has been accused of killing the game, increasing the down times and negatively affecting the referees' behavior; but the results obtained in this study show that there is not an abuse of the VAR system, having that only in 10% of the all arbitrated events it has been called. Moreover, of this 10%, it brought the referee to change his mind in 12% of these cases, so in less than 2% of all the events arbitrated.

In football, VAR has been introduced to "correct clearly incorrect decisions or report serious unseen events to the referee" [16] and this study supports the predictive analyses [14, 19], affirming that VAR is only a support in refereeing and does not condition negatively the game.

References

[1] IFAB, “The Laws of the game”. [https://www.theifab.com/laws], IFAB website, 2019.

[2] FIFA Executive Committee, “Regulations on the Organization of Refereeing in FIFA Member Associations”, 2010 Mar.

[3] Ceper C., Gavazzi N., and Mertens U., “Interview with star soccer referee Pierluigi Collina”. THE FOCUS. Vol X/1: 46- 49. 2017, Jan 1.

[4] Krustrup P et al., “Activity profile and physical demands of football referees and assistant referees in international games”. J Sports Sci. 2009 Sep; 27(11), 1167-1176, DOI: 10.1080/02640410903220310.

[5] Helse W. and Bultynck JB., “Physical and perceptual-cognitive demands of top-class refereeing in association football”. J Sport Sci. 2004 Feb; 22(2): 179-189. DOI: 10.1080/02640410310001641502.

[6] Castagna C., Abt G., D’ottavio S., “Physiological aspects of soccer refereeing performance and training.”. Sports Med. Jul 2007; 37(7): 625-46. DOI: 10.2165/00007256-200737070-00006.

[7] Weston M. at al., “Science and medicine applied to soccer refereeing: an update.” Sports Med. 2012 Jul; 42(7):615-31. DOI: 10.2165/11632360-000000000-00000.

[8] Mallo J., Navarro E., García-Aranda JM., Gilis B., Helsen W., “Activity profile of top-class association football referees in relation to performance in selected physical tests.” J Sports Sci. 2007 May; 25(7): 805-13. DOI: 10.1080/02640410600778602.

[9] Slack AL., Maynard IW., Butt J., Olusoga P., “Factors Underpinning Football Officiating Excellence: Perceptions of English Premier League Referees”. J App Sport Psych. 2013 Jul; 25(3): 298-315. DOI: 10.1080/10413200.2012.726935.

[10] Krustrup P. et al., “Activity profile and physical demands of football referees and assistant referees in international games”. J Sport Sci. 2009 Sept; 27(11): 1167–1176. DOI: 10.1080/02640410903220310.

[11] UEFA, “Refereeing”. [https://www.uefa.com/insideuefa/protecting-the-game/refereeing/index.html] UEFA website. 2019 Jun.

[12] FIFA, “Law of the Game 2012/2013”. FIFA website. 2012.

[13] Li Y., Dore A., Orwell J., “Evaluating the Performance of Systems for Tracking Football Players and Ball”. 2005 IEEE International Conference on Video and Signal Based Surveillance (AVSS'05), 15-16 Sept 2005, Como, Italy. DOI: 10.1109/AVSS.2005.1577342.

[14] Kidd R., “The Research That Changed Soccer: How Controversial VAR Joined The Beautiful Game”. Forbes. Jul 24, 2018. [https://www.forbes.com/sites/robertkidd/2018/07/24/the-research-that-changed-soccer-how-controversial-var-joined-the-beautiful-game/#33666eec35fb].

[15] Mirror Sport. “Big debate: Has VAR made a positive or negative impact at World Cup 2018?”. Mirror Sport. Jun 26, 2018. [https://www.mirror.co.uk/sport/football/news/big-debate-var-made-positive-12793420].

[16] FIFA. "VAR at the 2018 FIFA World Cup™". FIFA. 2018. [https://football-technology.fifa.com/en/innovations/var-at-the-world-cup/].

[17] Supereva. "Var : cos'è, cosa significa e come funziona" Supereva website. [https://www.supereva.it/var-cose-cosa-significa-e-come-funziona-40615].

[18] Bal.B et Dureja G., "Hawk Eye: A Logical Innovative Technology Use in Sports for Effective Decision Making". Sport Science Review. Jan 2012. 21(1). DOI: 10.2478/v10237-012-0006-6.

[19] LaPresse D., "Infantino: "L'introduzione della VAR è un momento storico, grazie anche all'Italia"". Eurosport. Mar 6, 2018. [https://it.eurosport.com/calcio/mondiali/2018/infantino-l-introduzione-della-var-e-un-momento-storico-grazie-anche-all-italia_sto6665243/story.shtml].

[20] Dagnelie G., "Visual Prosthetics: Physiology, Bioengineering, Rehabilitation". Springer Science & Business Media. 2011. ISBN 978-1-4419-0754-7.

[21] Williams AM., Devis K. et Williams JG., "Visual Perception and Action in Sport". Taylor & Francis Ltd. 1988. EAN: 9780419182900.

[22] Solso RL., "Cognition and the Visual Arts". MIT Press Ltd. 1966. EAN: 9780262691864.