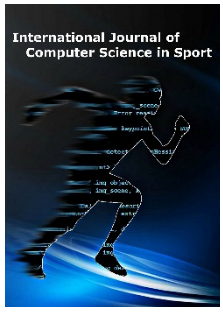
International Journal of Computer Science in Sport

Volume 16, Issue 1, 2017

Journal homepage: http://iacss.org/index.php?id=30

DE GRUYTER OPEN

DOI: 10.1515/ijcss-2017-0005



A Pilot Study on Offensive Success in Soccer Based on Space and Ball Control – Key Performance Indicators and Key to Understand Game Dynamics

Perl, J.[1]*, Memmert, D.*[2]

[1]*University of Mainz*

[2]*German University of Sport Science, Cologne*

Abstract

The intention of Key Performance Indicators (KPI) is to map complex system-behaviour to single values for scaling, rating and ranking systems or system components. Very often, however, this mapping only reduces important information about tactical behaviour or playing dynamics without replacing it by useful ones. The presented approach tries to bridge the gap between complex dynamics and numerical indicators in the case of offensive effectiveness in soccer in two steps. First, a model is developed which visualises offensive actions in a process-oriented way by using information units to represent offensive performance – i.e. Key Performance Indicators. Second, this model is organised in relation to time intervals, which enables to measure the effectiveness for a whole half-time as well as for arbitrary intervals of any desired lengths.

This contribution is meant as an introduction to a new modelling idea, where examples are calculated as case studies to demonstrate how it works. Therefore, only two games have been exemplarily analysed yet: The first one, which is used to demonstrate the method, is an example for similar quantitative indicators but different dynamic behaviour. The last one is used to demonstrate the results in the case of teams with extreme different strengths.

KEYWORDS: KPI, OFFENSIVE DYNAMICS, BALL CONTROL, SPACE CONTROL, CORRELATION, EFFICIENCY.

Introduction

Soccer is a complex game with a lot of technical and tactical facets (Rein & Memmert, 2016). Winning or losing a game often depends on small and often either hardly detectable details or complicated processes. Nevertheless, 0, 1, or 3 points for the result of the game assesses it simply by the result. The sum of these points is an important indicator for a team at the end of the season, which decides if it qualifies for higher challenges or not. This means that there is a basic and very simple visualising from complex quality to simple quantity. Insofar it is completely consistent to visualise qualitative aspects of the game like tactical behaviour as quantitative ratings like percentage of success.

There is a large number of such quite simple indicators (for a recent overview: Memmert, Lemmink, & Sampaio, 2017). Some examples are:

- Number of played passes;
- Mean number of passed players per pass;
- Mean length of passes;
- Ball wins / losses;
- Space wins / losses.

They all have in common that they do not say much about the playing process, the tactical quality, or the success of actions in defence or offence (Memmert et al., 2017). Different from these indicators, the presented approach tries to adapt "success" directly to the process dynamics, measuring it in terms of economics. In order to measure offensive success, there are – after shooting a goal, of course – two additional indicators of attacking which are extremely important. The two indicators are controlling the ball and controlling space in a specified critical area like the penalty-area and the 30-m-area in front of the opponent's goal (Memmert, Raabe, Knyazev, Franzen, Zekas, Rein, Perl, & Weber, 2016a,b).

The goal of this approach is to demonstrate how an appropriate combination of ball control and space control values leads to an indicator which provides quantitative information, like frequencies and distributions of events, without losing the qualitative information of the dynamic playing process. In the following method part, the term "success" which is usually used in an economy context is implemented into a model of offensive success in soccer (All presented tools are implemented in the soccer analysis software SOCCER©J.Perl, which contains more than 15 components for a wide range of statistical and dynamic analyses of processes in soccer; see Perl & Memmert, 2011; Perl, Grunz, & Memmert, 2013).

Method

In order to calculate offensive success depending on ball control and space control, the time-depending information is used as defined below.

The following parameters are used in the formulas and calculations:

- "t" is the point in time of calculation, in the following based on a grid of seconds.
- "IL" is the length of the time interval the calculation is applied to.
- "critical area" can be any area in front of the opponent's goal where offensive actions can generate danger, e.g. the penalty-area and the 30-m-area (all following examples are calculated using the 30-m-area as the critical area in order to guarantee full comparability).

- "LB" is a lower bound in order to specify relevant space control: A space control event is used only if the regarding control rate is greater than LB (for a good balance between noise and relevant values, the following examples use LB = 20%).

The term "controlled space at t" in (1 b) and (1 c) is defined as follows: At any point in time t, each player controls his own Voronoi cell (Fonseca, Milho, Travassos, & Araújo, 2012; Taki & Hasegawa, 2000), which is the set of all points of the pitch he can reach faster than any other player. This approach uses equal speed values and optimal player orientations and is, therefore, of course, an idealising model. Nevertheless, the combination of the Voronoi cells of all players, i.e. "controlled space at t", gives relevant information about the team's offensive effectiveness at time t.

*B*all *C*ontrol *E*vents

BCE(t) = if the team controls the ball at time t
then 1
else 0 (1 a)

Figure 1 shows the ball control events of team A (violet lines) in the 30-m-area in front of the goal of team B in a 300-sec-interval between 23:42 and 28:42 minutes. All following examples and figures refer to the 30-m-area and this particular time interval. The exemplary data are taken from the complete first half of a game that will be presented in the analysis section below in more detail.

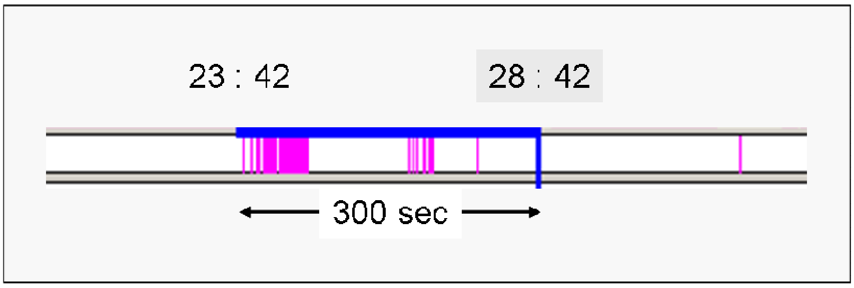


Figure 1: Ball control events presented as vertical violet lines.

*S*pace *C*ontrol *R*ates

SCR(t) = if %-rate of controlled space at t is > LB
then %-rate of controlled space at t
else 0 (1 b)

Figure 2 shows the profile of space control rates of team A (green line) in the 30-m-area in front of the goal of team B in the 300-sec-interval between 23:42 and 38:42 minutes. In figure 3, the corresponding space control events are added as gray lines, if the control-rates are greater than the lower bound LB (=20%).

*S*pace *C*ontrol *E*vents

SCE(t) = if %-rate of controlled space at t is >LB
then 1
else 0 (1 c)

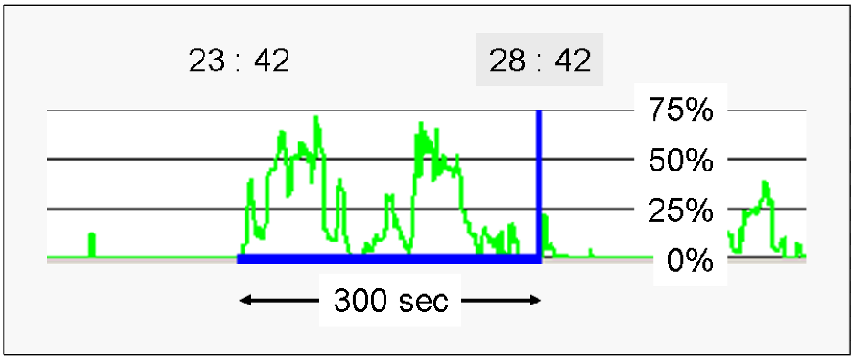


Figure 2: Space control rates presented as a green profile in a range of 0% to 75%.

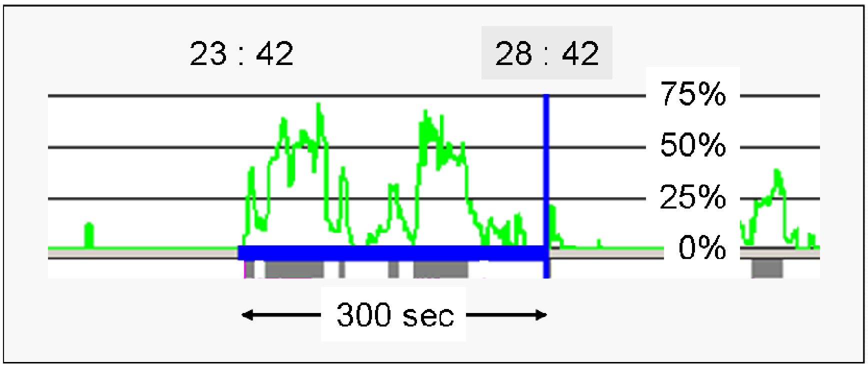


Figure 3: Space control events presented as vertical gray lines.

Of course, each of these indicators alone does not make a lot of sense. Controlling space without controlling the ball and vice versa is not very likely to generate dangerous situations. If, however, ball control meets space control in the critical area, this normally indicates a dangerous situation for the opponent. This means that not only the number of control events but also a kind of coincidence or correlation between them should play a role for indicating effectiveness or success of attacks.

A basic model from economy containing the term "success" is given by

$$\text{Efficiency} = \text{Success} / \text{Effort}, \quad (2)$$

which can easily be transformed to a model defining success depending on efficiency and effort:

$$\text{Success} = \text{Efficiency} \times \text{Effort}. \quad (3)$$

Projected to the situation of soccer, the terms Effort, Efficiency and Success can be defined as follows, using the formulas from above, regarding to an interval of length IL, ending in second t_0:

Effort is the number of all actions in order to generate dangerous situations in the opponent's critical area over a selected time interval – i.e.

$$\text{Effort}(IL,t_0) = \text{sum}(BCE(t) \oplus SCE(t)) / IL, \quad t=t_0-IL+1,\dots,t_0 \quad (4\text{ a})$$

(In order to avoid double counting active t-points, instead of "+" the operator "⊕" is used with a meaning similar to the logic "or", i.e. "1+1=1".)

Figure 4 shows the combination of space and ball control events, which defines the attacking effort in the analysed interval, while figure 5 shows the attacking efficiency as a correlation between ball control events and space control rates.

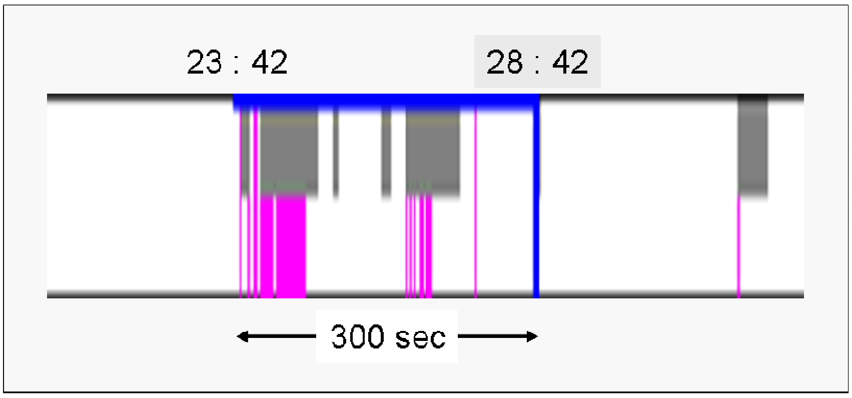


Figure 4: Space control events (gray lines) combined with ball control events (violet lines).

It was pointed out that neither ball control nor space control alone are normally successful in generating dangerous situations. Instead, a coincidence of ball and space control seems to be necessary:

Efficiency is therefore defined as a correlation between ball control events BCE(t) and space control rates SCR(rates), which leads to

$$\text{Efficiency}(IL,t_0) = \text{Corr}(\ BCE(t)\ ,\ SCR(t)\ ),\quad t=t_0-IL+1,\ldots,t_0. \qquad (4\text{ b})$$

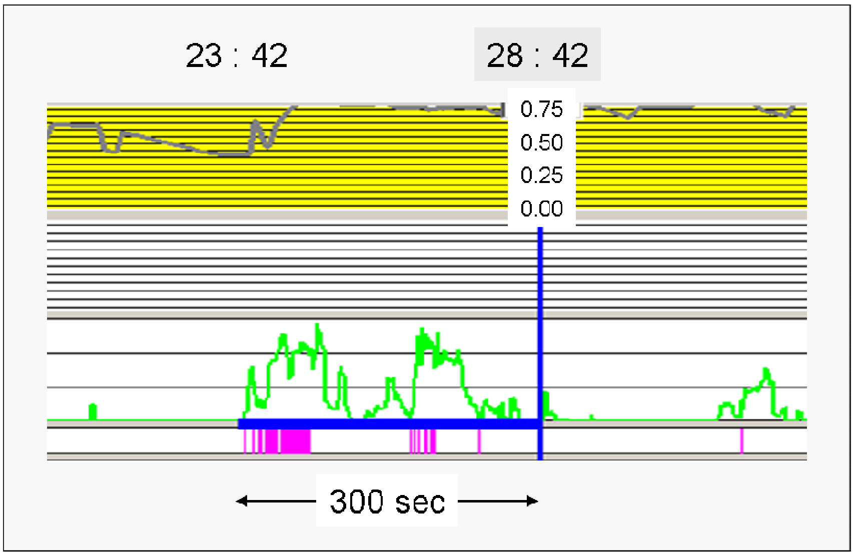


Figure 5: Correlation between ball control events and space control rates presented as a gray profile. In the selected interval, the correlation is on a comparably high level of about 0.75. Successful offensive activities normally show values between 0.50 and 0.70.

Offensive success can now be deduced from (4 a) and (4 b) as follows:

*O*ffensive *S*uccess

$$OS(IL,t_0) = \text{Efficiency}(IL,t_0) \times \text{Effort}(IL,t_0). \qquad (5)$$

Remark: To calculate the success indicator for the complete half-time, on the one hand, offers a simple key for categorising the game but loses a lot of information regarding game phases or specific attacks of the team. Therefore, it makes sense to orientate the calculation towards intervals of a given length IL which can be moved through the time grid, resulting in success indicators specific for a chosen interval.

The following figure 6 demonstrates that even with a rather high correlation the success value "0.42" seems to be comparably small (maximum value is 1). This is due to the fact that in the selected interval at only about half of the seconds (exactly: 56%) control events were generated. This clarifies that a reasonable indicator of success has to represent both, the correlation between control events and the number of control events.

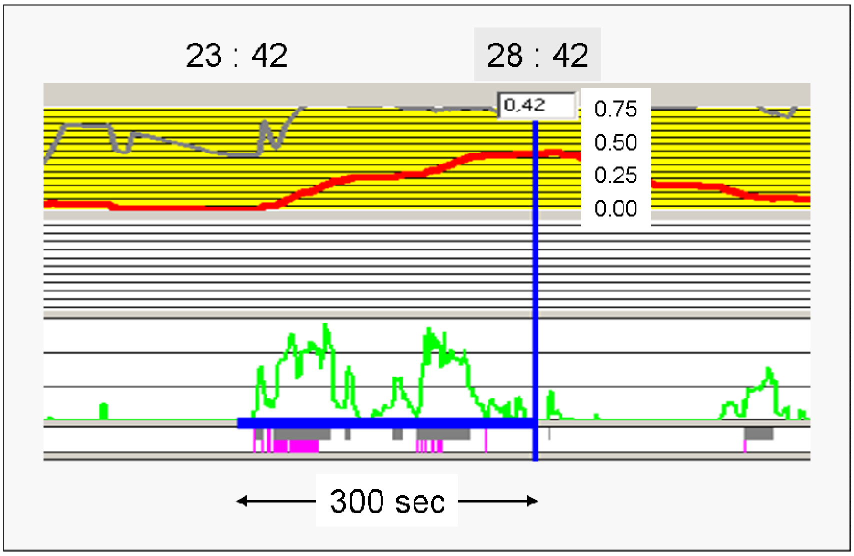


Figure 6: Completed presentation containing control rate and events as well as the gray correlation profile and the red success profile.

The following examples shall illustrate in more detail how the approach works.

Examples

OS as half-time KPI (Key Performance Indicator)

In case of calculating OS as a KPI for a complete half-time, only the values for $t_0 = t_{max}$ (= length of half-time) offer usable information, while the red profile over the half-time only represents the iterative process of approaching the OS-indicator to its final value.

In the following figures the upper parts with the yellow scales show information about team A while the lower parts with the blue scales show information about team B.

In figure 7, the interval is the half-time with a length of IL = t_{max} = 2712, and the point of calculation correspondingly is $t_0 = t_{max}$ = 2712. The red profiles show how the underlying iterative calculation process approaches the final OS-values "0.18" and "0.20". This calculation process strongly depends on iterative calculation of mean values.

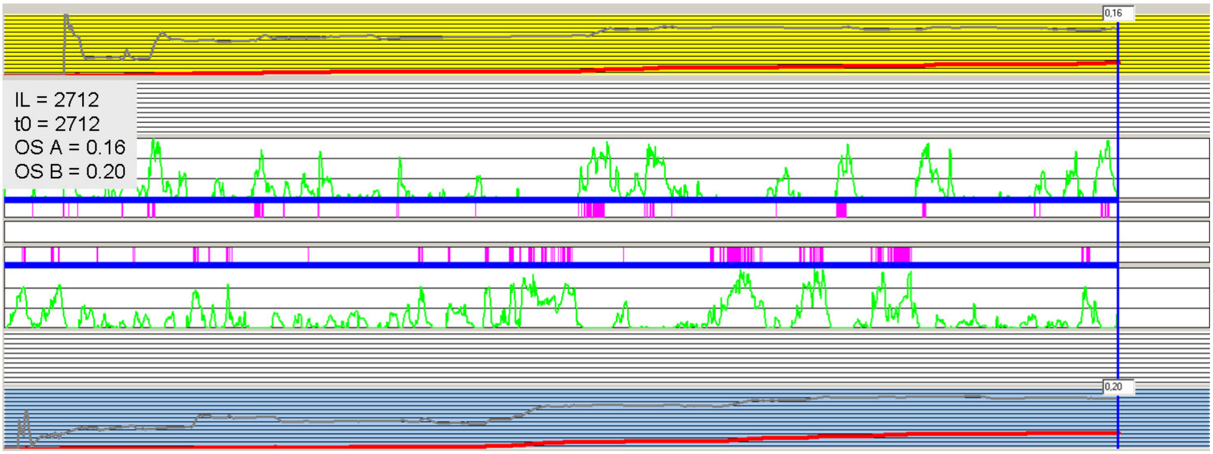


Figure 7: OS-values of team A and team B together with the red profiles of iterative approaches of the final values.

Different is the situation in case of OS as interval KPI.

OS as interval KPI

In the starting phase – i.e. t = 0, ..., IL-1 – the red profile represents the approaching calculation process.

For t = IL and then all t > IL, the red profile represents the values OS(IL,t).

In figure 8, the approaching zone regarding the interval length of IL = 900 is marked by a vertical red line. In the area right from this point, the red OS-profiles represent the dynamically changing OS-values depending on ball control events and space control rates.

Exemplarily, the blue lines mark the calculation phase of $OS(IL,t_0)$ = OS(900,2337). The results OSA(900,2337) = 0.25 and OSB(900,2337) = 0.28 say that – although team B was a bit better than team A – neither team A nor team B were really successful in the time interval [1438,2337].

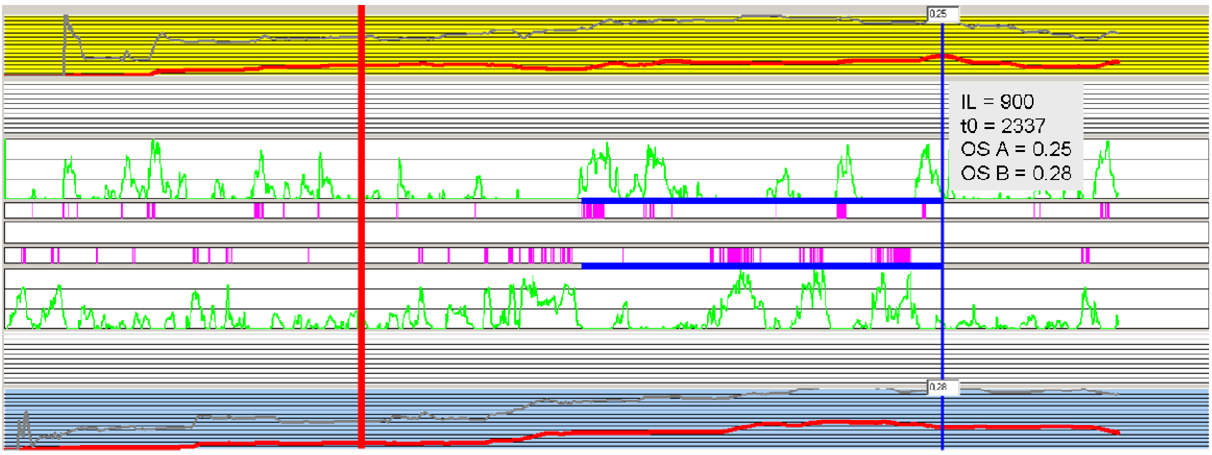


Figure 8: Approaching zone (red mark) and example of OS-calculation (blue marks) for IL = 900.

Obviously, interval-oriented OS-calculation opens a better view on the dynamic development of offensive success depending on control events. This effect becomes more impressive if the interval length becomes smaller, as demonstrated in figure 9:

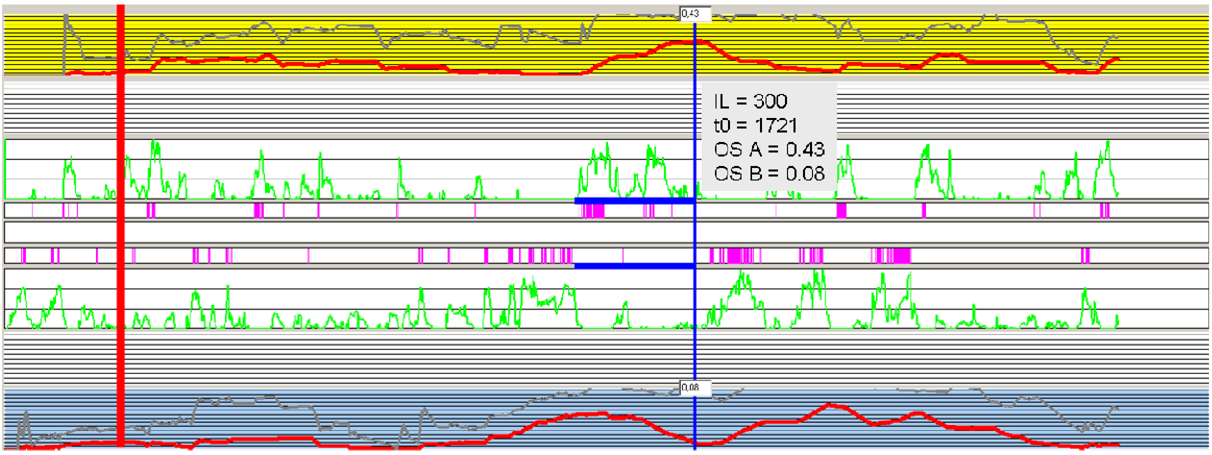


Figure 9: OS-profiles and example of OS-calculation for IL = 300.

From figure 9 it becomes clear how the offensive dynamics of team A (red profile in the upper half) and of team B (red profile in the lower half) change and interact with each other, which becomes even clearer if the interval length becomes smaller, as figure 10 demonstrates.

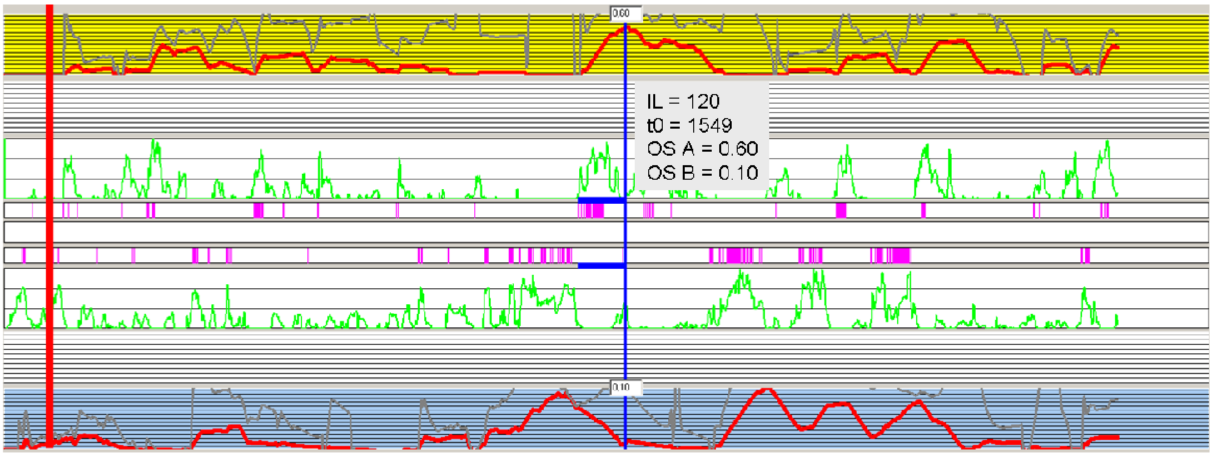


Figure 10: OS-profiles and example for IL = 120.

If, however, the interval length becomes too small, this positive effect disappears and the calculation mainly reflects dominating event constellations, as can be seen for IL = 60 in figure 11.

First analyses show that interval lengths between 600 and 100 fit best if the focus lays on dynamics and success of the offensive behaviour.

Analysing the first half-time under the aspect of offensive success and based on the results from above leads to the results listed in Table 1.

Team A has its strongest phase in the time interval [1420,1720], while team B has its strongest phase in [1720,2320] and an additional - but not quite as strong - phase in [1120,1720]. Combining these results makes understandable that over the complete half-time team A has a somewhat higher OS-KPI-value than team B. The comparison also shows that those rather low half-time values OSA = 0.18 and OSB = 0.20 do not reflect the comparably high offensive potential both teams have. In the second half-time, the values are even worse, OSA = 0.11 and OSB = 0.17. But at least team B has one rather high interval-value of OSB(200,1600) = 0.49.

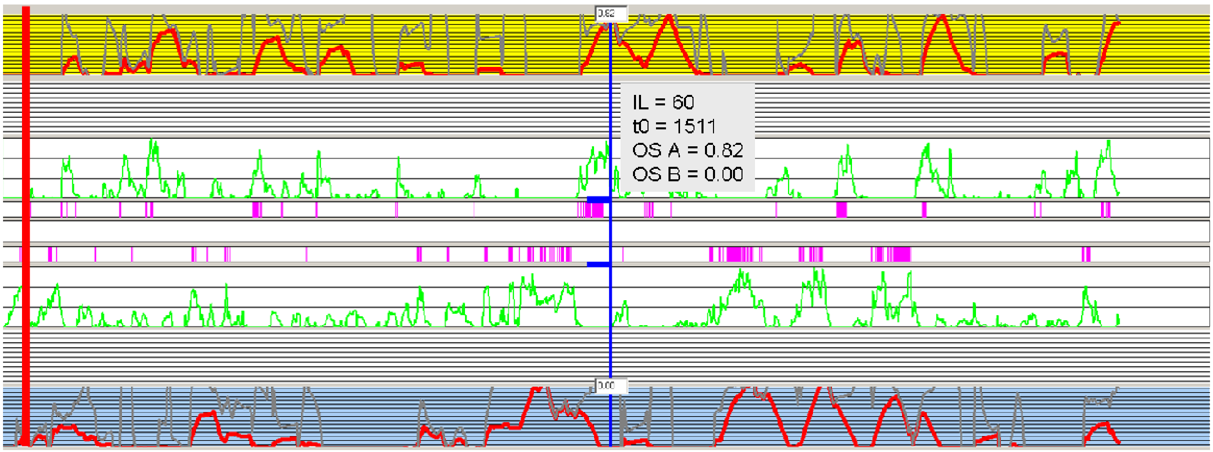


Figure 11: OS-profiles and example for IL = 60.

Table 1: Offensive success of team A and team B in the three phases of highest actions.

Interval	Length	OS A	OS B
half-time 1		0,18	0,20
[1020,1420]	400	0,02	0,40
[1450,1650]	200	0,55	0,07
[1760,2260]	500	0,10	0,46

A last example is taken from another game with a dominant team A against a weak team B, yielding the results listed in Table 2.

Table 2: Offensive success values of team A and team B.

Interval	Length	OS A	OS B
half-time 1		0,31	0,06
[1053,1253]	200	0,48	0,00
[1576,1776]	200	0,02	0,31
[1890,2190]	500	0,53	0,02

Even clearer becomes the difference between half-time KPI and interval values in that case by comparing the OS-profiles for IL = 200. As figure 12 shows, team B has not any chance against continuously running attacks of team A.

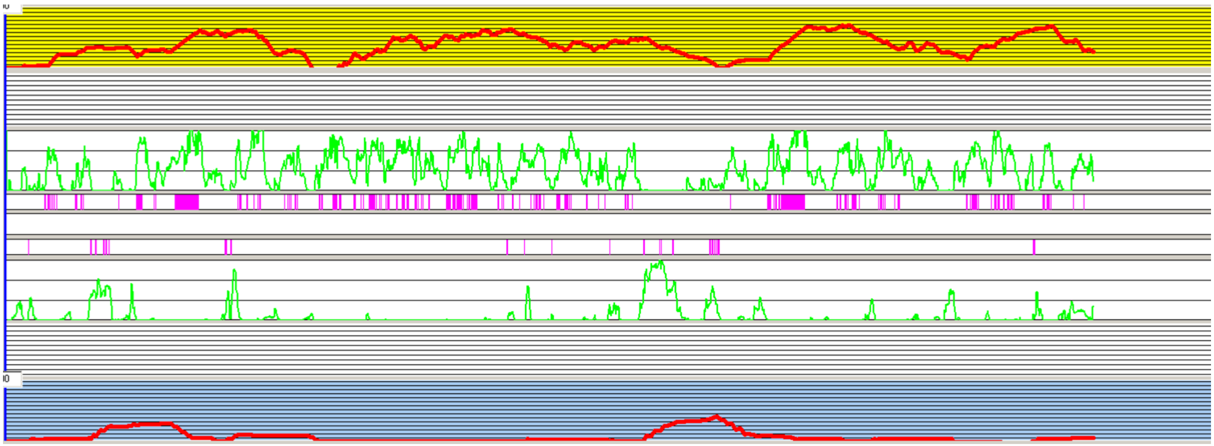


Figure 12: OS-profiles of a dominant team A against a weak team B, IL = 200.

Discussion

The presented approach demonstrates how a KPI can be deduced from simple modelling assumptions without neglecting the complex playing dynamics. In particular, the orientation to intervals opens access to a better understanding of offensive dynamics and offensive success. In turn, it becomes clear how much information gets lost if the analysis is reduced to calculating just one value representing the offensive quality of a complete half-time. Nevertheless, KPIs are thought to rank teams. Therefore, the OS-KPI can be helpful in order to compare teams regarding their offensive success indicators – in particular, in combination with interval-based success profiles.

References

Fonseca, S., Milho, J., Travassos, B., & Araújo, D. (2012). Spatial dynamics of team sports exposed by Voronoi diagrams. *Human Movement Science, 31*, 1652-1659.

Memmert, D., Lemmink, K. A. P. M., & Sampaio, J. (2017). Current Approaches to Tactical Performance Analyses in Soccer using Position Data. *Sports Medicine, 47*(1), 1-10.

Memmert, D., Raabe, D., Knyazev, A., Franzen, A., Zekas, L., Rein, R., Perl, J. & Weber, H. (2016a). Big Data im Profi-Fußball. Analyse von Positionsdaten der Fußball-Bundesliga mit neuen innovativen Key Performance Indikatoren [Big Data in professional football. Analysis of position data from the German Fußball-Bundesliga with new innovative Key Performance Indicators.]. *Leistungssport [High Performance Sport],* 1-13.

Memmert, D., Raabe, D., Knyazev, A., Franzen, A., Zekas, L., Rein, R., Perl, J. & Weber, H. (2016b). Innovative Leistungsindikatoren im Profifußball auf Basis von Positionsdaten [Innovative performance indicators in professional football based on position data.]. *Impulse [impulse], 2*, 14-21.

Perl, J., Grunz, A. & Memmert, D. (2013). Tactics in soccer: an advanced approach. *International Journal of Computer Science in Sport, 12*, 33-44.

Perl, J., & Memmert, D. (2011). Net-Based Game Analysis by Means of the Software Tool SOCCER. *International Journal of Computer Science in Sport, 10*, 77-84.

Perl, J., Grunz, A., & Memmert, D. (2013). Tactics in soccer: an advanced approach. *International Journal of Computer Science in Sport, 12*, 33-44.

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

Taki, T., & Hasegawa, J. I. (2000). Visualization of dominant region in team games and its application to teamwork analysis. In *Computer Graphics International, 2000. Proceedings* (pp. 227-235). IEEE.