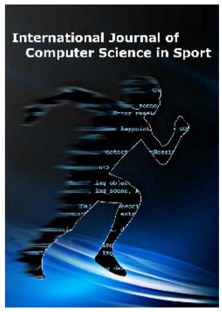
International Journal of Computer Science in Sport

Volume 17, Issue 2, 2018

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

sciendo

DOI: 10.2478/ijcss-2018-0012



Formation-based modelling and simulation of success in soccer

Perl, J.

University of Mainz, Germany

Abstract

The players' positions of tactical groups in soccer can be mapped to formation-patterns by means of artificial neural networks (Kohonen, 1995). This way, the hundreds of positional situations of one half of a match can be reduced to about 20 to 30 types of formations (Grunz, Perl & Memmert, 2012; Perl, 2015), the coincidences of which can be used for describing and simulating tactical processes of the teams (Memmert, Lemmink & Sampaio, 2017):

Developing and changing formations in the interaction with the opponent activities can be understood as a tactical game in the success context of ball control, space control and finally generating dangerous situations. As such it can be simulated using mathematical approaches like Monte Carlo-simulation and game theory in order to generate optimal strategic patterns. However, in accordance with results from game theory it turns out that in most cases the one optimal strategy does not exist (e.g. see Durlauf & Blume, 2010). Instead, a variety of partial strategies with different frequencies were necessary – an approach that is mathematically interesting but has nothing to do with soccer reality. An alternative approach, which is developed in the following, is to interrupt the strictness of a single strategic concept by creative elements, which improves flexible response to opponent activities as well as prevents from being analyzed by the opponent team.

The results of respective simulation reach from improving strategic behaviour to recognizing strategic patterns and in particular to analyzing role and meaning of creative elements.

KEYWORDS: FORMATION, SUCCESS, STRATEGY, CREATIVITY, ARTIFICIAL NEURAL NETWORK

Introduction

The aim of the presented approach is to model tactical activities in soccer in order to find optimal strategic concepts by means of simulation (also see Perl & Memmert, 2017). The used dynamic model is of type "state – transition", which means that the dynamic playing process is divided into discrete tactical states (here: one per second), where the changes between them are used as transitions.

Because of the fact that the dynamics of activities in soccer is extremely complex, the term state has to be defined on a rather high level of abstraction. One way to do this, and which has been successful in other studies, is to replace the original positions of the players of a tactical group by a location-independent pattern – what we call a formation type (see fig. 1) (Perl & Memmert, 2018).

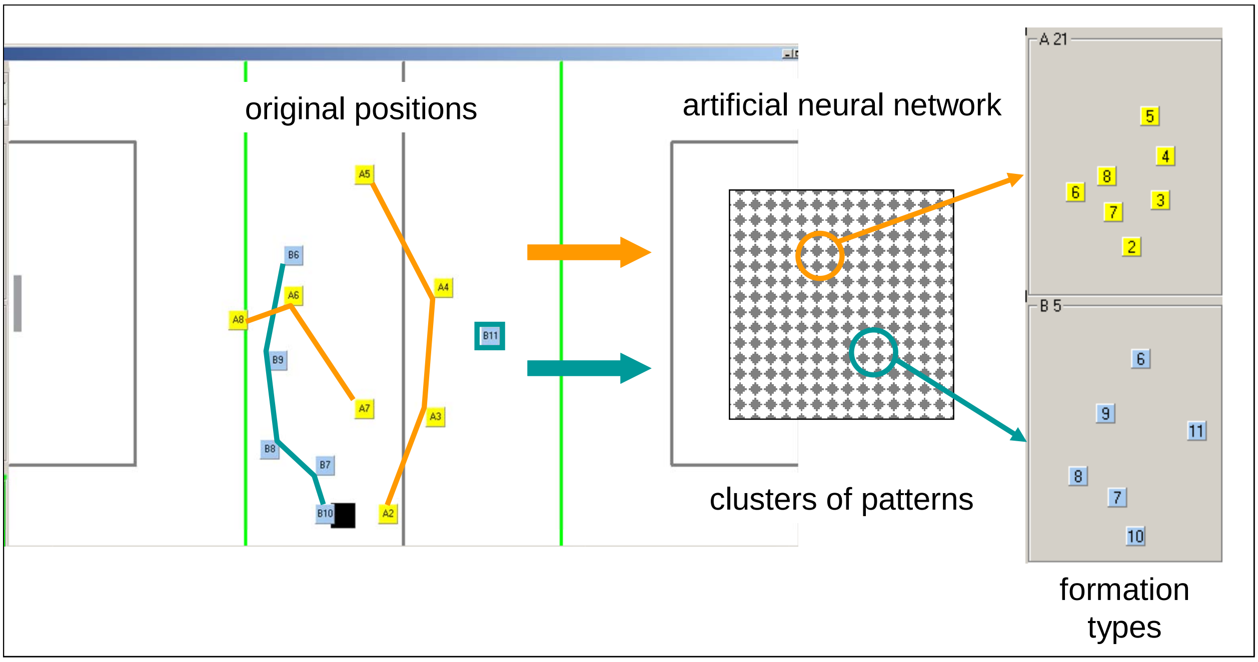


Fig. 1: The original positions of the offence group of team B (blue) and of the defence group of team A (yellow) are transferred to a trained network and identified to represent formation B 5 and A 21, respectively.

Those transitions build the basis for analyzing the dynamics of the tactical behaviour. They enable for simulating and optimizing processes by means of game theoretic approaches, which finally leads to the question of strictly vs. creatively following simulated strategic patterns.

Doing it this way, the complex interaction of players can be reduced to a comparably simple interaction of formation types:

A *state* can be defined as a pair of formation types of team A and team B at time t,

$$(AF[t]\ ,\ BF[t]).$$

A *transition* is the change of A- and B-formation types (independently of each other) to the state at time t+1 (also see fig. 2),

$$(AF[t]\ ,\ BF[t]) \rightarrow (AF[t+1]\ ,\ BF[t+1]).$$

Methods of modelling and simulation

Note: In the following the term *formation type* is abbreviated to *formation*.

Modelling: State-transition-approach

Modelling is based on the formations of the offensive and defensive tactical groups of teams A and B depending on time t, oriented in the values of success indicators like for instance ball control or space control (see Perl & Memmert, 2017):

Assume that at time t the state is given by the formation pair (AF[t] , BF[t]). Each time step t → t+1 defines a transition where both teams independently of each other change their formations to AF[t+1] and BF[t+1]. Those changes can be arbitrarily or oriented in kinds of success-optimization like e.g. improvement of ball control or space control. An example is given in fig. 2.

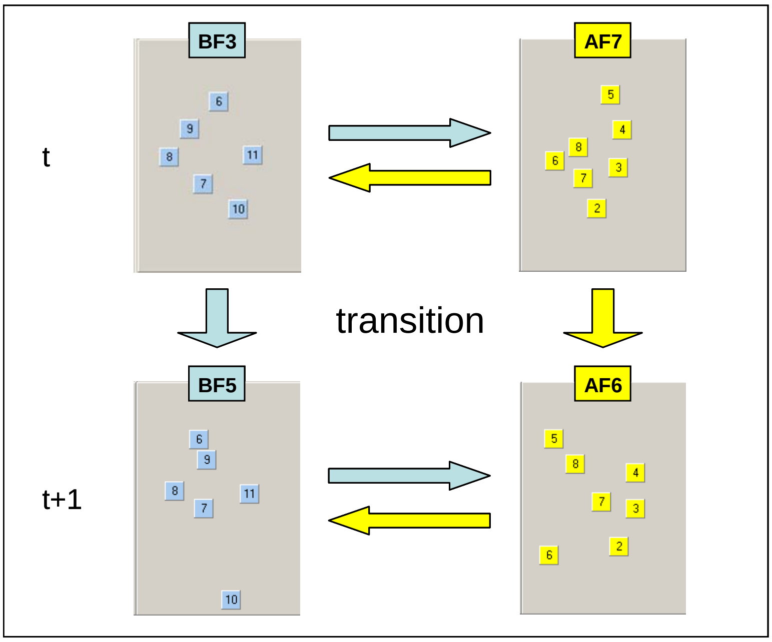


Fig. 2: The upper half shows the state (AF7 , BF3) of attacking team B and defending team A at time t. The transition to the subsequent state in the lower half consists in positioning B10 to a decentred position, while team A with A5 is opening to a more aggressive formation. Both changes could have been done in order to increase space control.

In the following, modelling, simulation, and analyses are done using specific components of the *SOCCER* software (© J. Perl, 2018) (Perl & Memmert, 2011).

Simulation: Monte-Carlo-method

Simulation of the time-depending state-transition-process can be done by means of the Monte-Carlo-method, where the original data provide frequency distributions that can be taken for generating simulated data. An example is given in fig. 3, where the left matrix shows the frequencies of states = pairs of formations (restricted to the first 10 formations, respectively), and the right tables in the first row show the distribution of subsequent formations of team A in case of the highlighted state (AF1,BF1): In most of the cases the subsequent formation is again AF1. But there are alternative changes to the formations AF2, AF3, AF4, AF6 (and 1 more AFn with n>10). Taking this distribution for MC-simulation would reproduce the original behaviour. Focussing the distribution on successful activities like ball contact, ball possession, or passes (rows 2, 3, and 4) could increase the success regarding these indicators. However, doing it this way the variety is reduced considerably, which would simplify the simulated behaviour of A a lot and therefore makes that behaviour much more predictable. Calculating and comparing such simulations can help to find optimal strategies and prevent strategic dead ends like the orientation in passes with its stereotype repetition of AF1 (see fig. 3, right matrices, last rows).

offence formations BF[t] / defence formations AF[t]

	1	2	3	4	5	6	7	8	9	10
1	30	3	37	11	1	1	2	11		
2	40	69	28	6	14	5	20	23		2
3	56	51	34	6	40		3	4		1
4	13	57	4	10		10	27	7		
5	7		2						9	
6	41	2	48	26	32	1				
7	42	35	27		44	17	3			2
8	4			2				9		18
9						1	1		9	
10	2		12	3		1	18			

distributions of A[t+1]-formations

	1	2	3	4	5	6	7	8	9	10
formations	23	2	1	2		1				
ball contacts	10	1	1							
ball possessions	7	1	1							
passes	2									

distributions of B[t+1]-formations

	1	2	3	4	5	6	7	8	9	10
formations	29			1						
ball contacts	3									
ball possessions	2									
passes	1									

Fig. 3: The left matrix shows the coincidences of defence formations of team A (left) and offence formations of team B (on top). The red mark highlights the state (AF1,BF1), which occurred 30 times. The right matrices in the first rows show the distributions of the transitions of the teams, where in case of team A 23 times the situation stayed stable in formation 1. (The reason that the numbers of continuing formations do not sum up to 30 is that the table is restricted to the first 10 formations.) The rows below show the numbers of successful activities in the contexts of those formations.

Technically, MC-simulation can be done as follows:

A random generator generates numbers x equally distributed over an interval I=[0,1), i.e. $0 \le x < 1$. To map a given distribution like the ball-contact (10, 1, 1, 0,..., 0) of team A from fig. 3 to the interval I, the numbers have to be transformed to corresponding interval-lengths: I1=[0,0.84), I2=[0.84,0.92), I3=[0.92,1.0), I4=[1.0,1.0),..., I10=[1.0,1.0). Doing it this way, the probability of randomly generating a "1" is 84% and that of generating a "2" or a "3" is 8%, respectively, corresponding to the respective interval-lengths; all other probabilities are 0%.

Before continuing with MC-simulation, a very simple but common strategic behaviour and method of simulation shall briefly be introduced and discussed: From the example of fig. 3 it seems to make sense, AF1 always to continue with AF1 – not only because AF1 is the most frequent formation but also because it is the most successful one. The corresponding formation of B in fig. 3 is BF1. Analogously, BF1 is the most frequent and/or successful continuation of BF1, and hence the simulation would reproduce the state (AF1,BF1) endless after once having reached it the first time.

The result is not restricted to bad simulation but can be found in practice as well. For example, such "stable" situations sometimes can be observed in tennis (base-line-duels or backhand-backhand-duels) or in soccer (concrete-defence, kick-and-rush).

Theory (2-persons-0-sum-theory, e.g. see Durlauf & Blume, 2010) as well as practical experience show that it is much more successful to not always play the "best" variant but to distribute it to a strategic mixture, which is much more difficult to anticipate by the opponent player or team. This is what MC-simulation with its orientation to distributions is helpful for – if the distribution is a substantial one and not degenerated like "passes" in fig. 3.

The problem is, however, that very often the distributions that can be taken as a basis for simulation are as degenerated as "passes" is in fig. 3. Note for example that "ball contacts" and "ball possessions" are not much better.

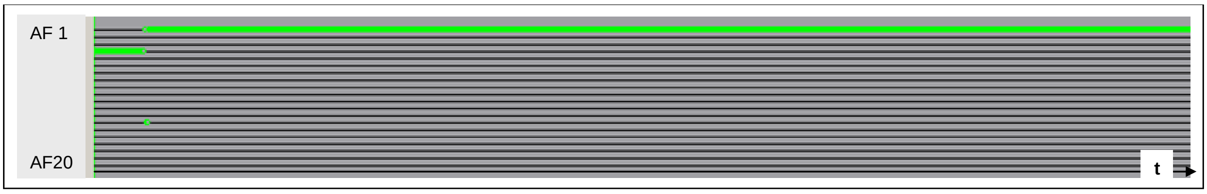


Fig. 4: Time-depending formations-diagram of team A as a result of MC-simulation oriented in optimizing the number of "ball contacts". As expected from fig. 3, the simulation is stabilizing with formation 1.

Fig. 4 shows the formations diagram of a strict "ball contacts"-oriented MC-simulation: It starts with the original start-formation AF4 and very fast stabilizes on formation AF1. Obviously, this result has nothing to do with playing soccer.

The reason is that strict MC-simulation does not take into account situation-oriented spontaneous actions that are not necessarily oriented in the optimization aim. Even small deviations from a 100%-simulation can produce results much closer to reality, as fig. 5 demonstrates: The upper diagram shows the original distribution of formations, while the lower diagram shows the MC-generated formations with a strictness of 95%, completed by 5% distribution-independent random values. Even though there is a clear priority on AF1, the original and the simulated formations diagrams become much more similar.

In the following, this difference between the strictness St and the 100%-simulation will be called flexibility Fl. In the case of fig. 5 this means: StA = 95, FlA = 100-95 = 5.

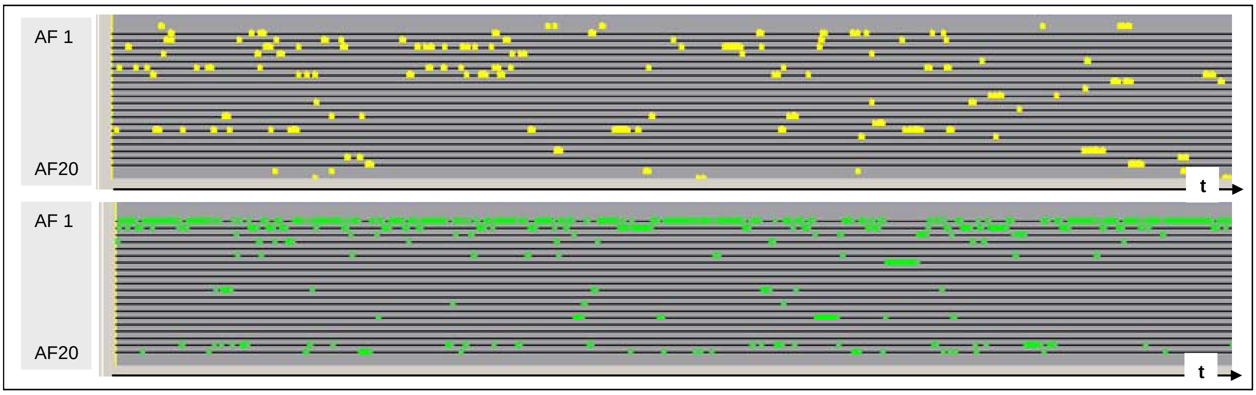


Fig. 5: Original formations diagram above (yellow dots). Below formations diagram (green dots) generated by "ball-contacts"-oriented MC-simulation with StA = 95.

Some more examples of the effect of strictness/flexibility in simulating formation distributions are given in fig. 6. As will be discussed later on under the aspect of creativity, in the case of flexibility = 0 the behaviour has a stereotype structure and seems to be easily to predict. In contrast, with flexibility ≥ 50 the behaviour becomes unmethodical, which is caused by the high rate of random generation.

The next step is to discuss, how simulation is usable under the aspect of analyzing soccer games.

Results

In the following, mainly two aspects of game simulation are dealt with. In the first part the success of strictness is analysed and discussed in the context of strategic interaction. In the second part it is presented how simulation can help to recognize the degrees of flexibility and creativity in the strategic behaviour of a team.

For a better comparability, all simulations are calculated using the same original data introduced in fig. 3.

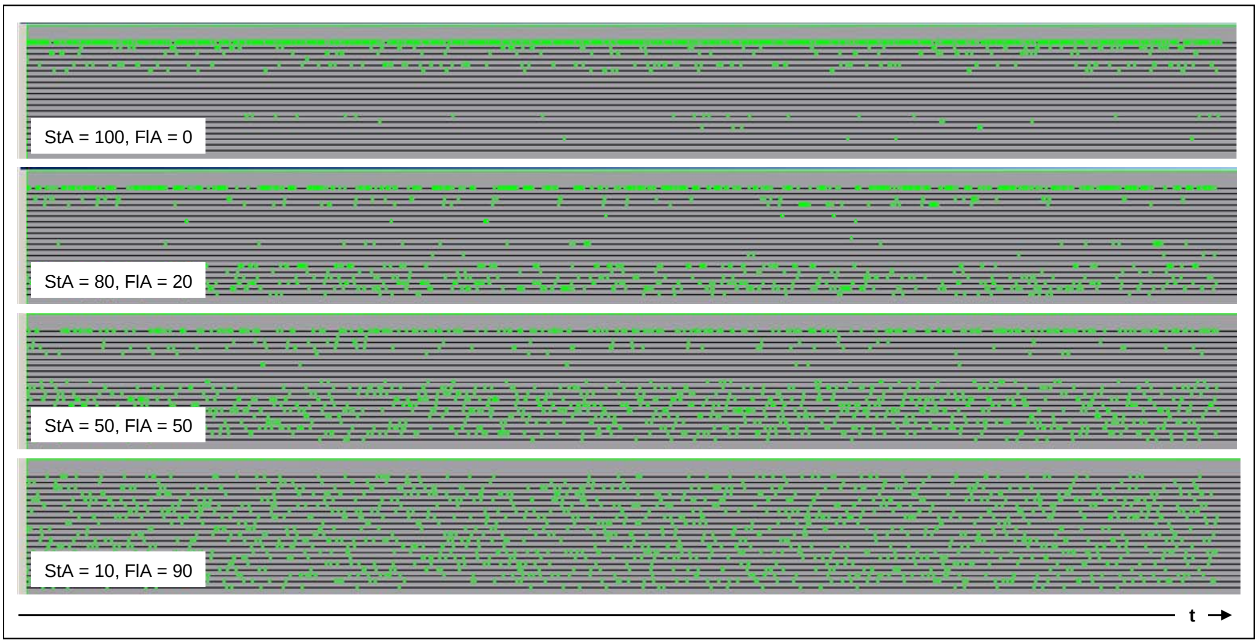


Fig. 6: Distributions of A-formations depending on values of strictness/flexibility.

Simulation-based analysis of soccer games

A first approach is to analyze how success can be improved, as is depicted in fig.7:

The yellow profile shows the increasing sum of the original ball contacts of team A (measured in seconds), while the small black circles mark the single contact events. The green profile in the same way shows the MC-generated values, simulated with *StA* = 95 and *StB* = 95.

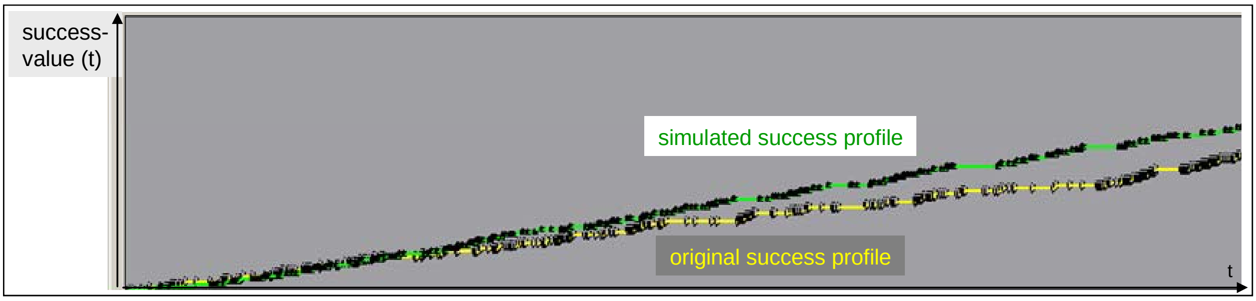


Fig. 7: Profiles of original (yellow) and simulated (green) contact events of team A.

The presented result – simulation improves success – of course is only one example. It depends on the definition of success (here: ball contacts) and the values of strictness (here: *StA* = 95, *StB* = 95). In fig. 8 four more scenarios are presented where different pairs (*StA*,*StB*) of strictness generate quite different profiles of success of team A: In the first example both teams act with the maximum strictness, (*StA*,*StB*) = (100,100), resulting in a simulated success profile that is much better than the original profile of team A. This is due to the fact that a 100%-strictness (theoretically) means an optimal behaviour – the success of which however depends strongly on the behaviour of the opponent team: The second example shows that reducing *StA* to 70 does not reduce the success of A that much. Instead reducing *StB* to 70 as in example 3, although reducing the success of team A significantly, the simulated success profile is still better than the original one. Increasing *StB* to 90, as in example 4, results in a simulated profile of team A that is worse than the original one.

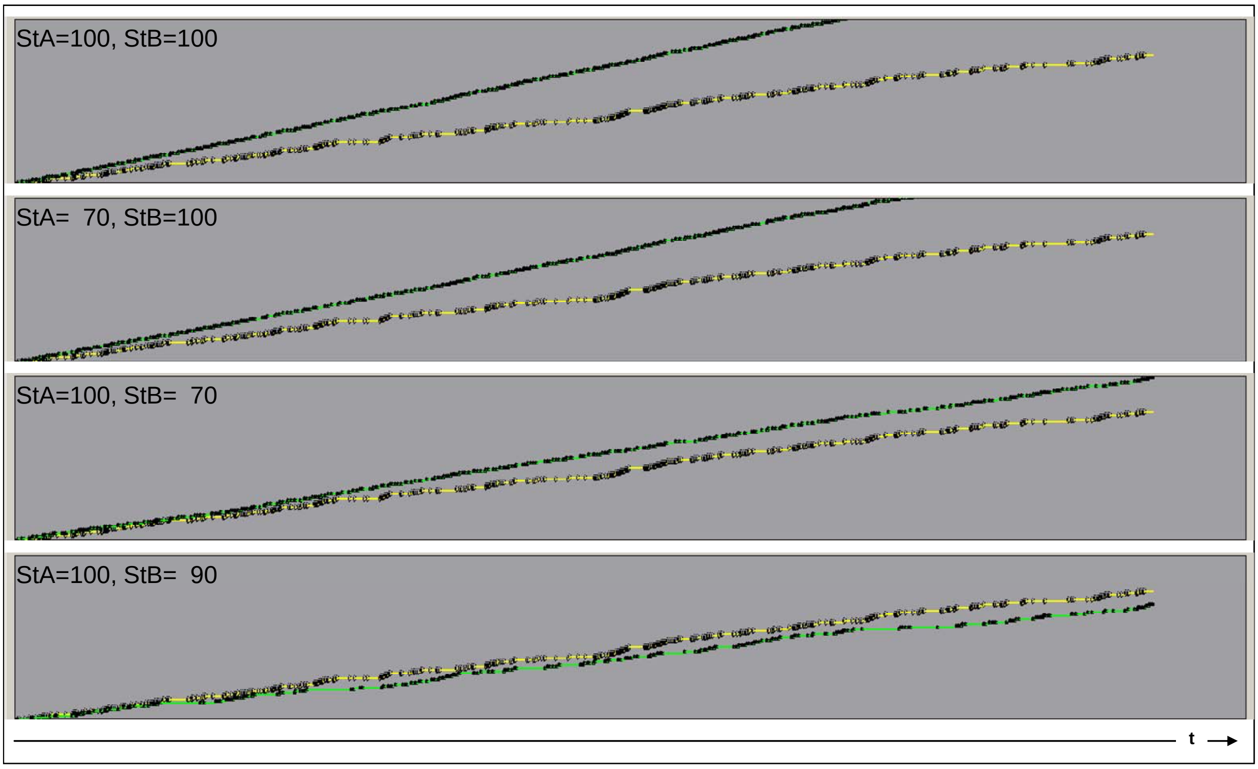


Fig. 8: Four examples of how strictness values affect the simulated success of A.

The examples in fig. 8 show how sensitive the success profiles and values react on the values of strictness. In order to better understand the connection between strictness and success it is helpful to read the "landscapes" of success, which are presented as matrices of success values of the teams A and B in fig. 9: The left matrix shows the final success differences of A and B. The strictness-pairs with advantage A (success difference > 0) are highlighted in orange with a focus on one maximum pair (*StA*,*StB*) = (72,98). In the same way the right matrix shows the final success differences of B and A, where the "clear" advantage of B (success difference > 1000) is highlighted in blue with the focus on the maximum pair (*StA*,*StB*) = (100,92).

A StA	StB FlA/B	80	82	84	86	88	90	92	94	96	98	100
		20	18	16	14	12	10	8	6	4	2	0
60	40	-59	-60	-48	-65	-71	-73	-65	-43	-13	4	7
62	38	-21	-26	-22	-35	-40	-39	-35	-16	7	34	37
64	36	-24	-35	-35	-49	-59	-59	-58	-21	5	32	39
66	34	-34	-51	-53	-68	-80	-72	-73	-37	-5	25	35
68	32	-31	-48	-56	-77	-87	-77	-77	-44	-7	27	36
70	30	-46	-60	-72	-91	-98	-88	-89	-45	-8	40	42
72	28	-55	-67	-79	-100	-109	-97	-101	-54	-1	55	55
74	26	-87	-94	-109	-130	-140	-130	-134	-81	-13	45	54
76	24	-108	-110	-124	-141	-148	-144	-143	-88	-26	44	52
78	22	-139	-146	-152	-165	-170	-180	-184	-125	-46	31	34
80	20	-222	-233	-270	-277	-307	-311	-312	-249	-142	-18	-19
82	18	-273	-295	-332	-356	-391	-390	-417	-324	-210	-76	-87
84	16	-313	-344	-392	-412	-445	-441	-484	-394	-251	-100	-109
86	14	-314	-349	-395	-423	-472	-483	-553	-446	-299	-98	-101
88	12	-347	-399	-450	-481	-543	-566	-636	-533	-360	-165	-133
90	10	-399	-448	-500	-525	-591	-658	-755	-624	-430	-203	-142
92	8	-527	-593	-669	-736	-774	-849	-981	-802	-513	-210	-152
94	6	-505	-563	-657	-760	-838	-899	-1052	-838	-520	-217	-135
96	4	-527	-599	-666	-775	-889	-951	-1159	-1045	-647	-267	-102
98	2	-544	-621	-661	-835	-1002	-1134	-1360	-1168	-718	-85	-96
100	0	-505	-595	-651	-820	-1034	-1234	-1821	-1587	-1072	-122	-63

B StA	StB FlA/B	80	82	84	86	88	90	92	94	96	98	100
		20	18	16	14	12	10	8	6	4	2	0
60	40	59	60	48	65	71	73	65	43	13	-4	-7
62	38	21	26	22	35	40	39	35	16	-7	-34	-37
64	36	24	35	35	49	59	59	58	21	-5	-32	-39
66	34	34	51	53	68	80	72	73	37	5	-25	-35
68	32	31	48	56	77	87	77	77	44	7	-27	-36
70	30	46	60	72	91	98	88	89	45	8	-40	-42
72	28	55	67	79	100	109	97	101	54	1	-55	-55
74	26	87	94	109	130	140	130	134	81	13	-45	-54
76	24	108	110	124	141	148	144	143	88	26	-44	-52
78	22	139	146	152	165	170	180	184	125	46	-31	-34
80	20	222	233	270	277	307	311	312	249	142	18	19
82	18	273	295	332	356	391	390	417	324	210	76	87
84	16	313	344	392	412	445	441	484	394	251	100	109
86	14	314	349	395	423	472	483	553	446	299	98	101
88	12	347	399	450	481	543	566	636	533	360	165	133
90	10	399	448	500	525	591	658	755	624	430	203	142
92	8	527	593	669	736	774	849	981	802	513	210	152
94	6	505	563	657	760	838	899	1052	838	520	217	135
96	4	527	599	666	775	889	951	1159	1045	647	267	102
98	2	544	621	661	835	1002	1134	1360	1168	718	85	96
100	0	505	595	651	820	1034	1234	1821	1587	1072	122	63

Fig. 9: Matrices of success value differences of A (left) and B (right) depending on pairs of *StA* and *StB*.

Before discussing these results regarding the meaning of strictness, in the following fig. 10 the success profiles corresponding to the maximum pairs are presented:

The first graphic shows the original success profiles of team A (yellow) and team B (light blue) compared to the simulated profiles of A (light green) and B (dark blue). As expected, the simulated profile of B is much better than the original one, and in contrast the simulated profile of A is much worse than the original one.

In the second graphic the situation is the other way round but, according to the smaller success difference, with smaller profile distances. The distance between team A and team B is marginal although A has improved a lot against its original profile.

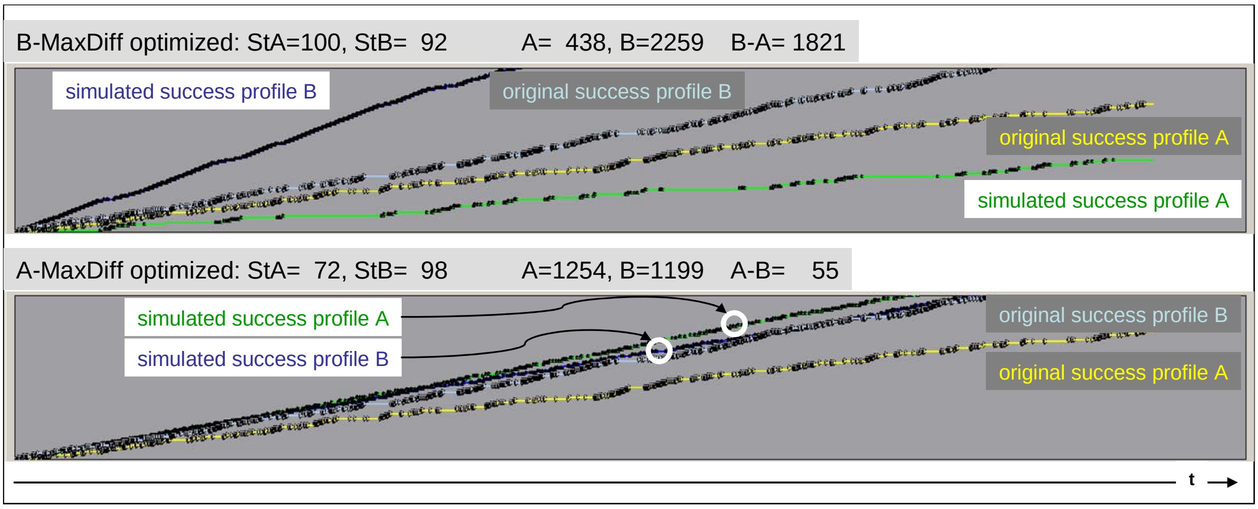


Fig. 10: Original and simulated success profiles of teams A and B in the cases of highest success differences.

The examples of fig. 10 emphasize the meaning of interaction in analysis and simulation of playing dynamics:

(1) It is obviously not sufficient to increase the number of successful events of a team if the opponent team is able to increase more.

(2) The degrees of strictness respectively flexibility in getting strategies to work seem to play an important role for being successful.

In fig.11 it is shown how much a strategy with an appropriate degree of flexibility can (theoretically) improve success:

In the cases of ball contacts and possessions there is a potential for improvement. Clearly, the numbers of ball possessions are always less than those of ball contacts, which is reproduced also in the simulation. In the case of passes the simulated improvement obviously is not a realistic one. The reason is that a pass is imbedded in a complex process context and cannot be simulated using only its quantitative correlation to formations. In contrast, space control is highly dependent on the interaction of formations, and therefore the simulation shows a realistic potential of improvement.

Quite similar is the situation with team B.

In the following, the relation between strictness and flexibility will be discussed in more detail:

An increasing degree of flexibility on the one hand decreases the orientation on the seemingly optimal strategy, i.e. the strategy with a strictness of 100%.

On the other hand, as is shown in fig. 6, an increasing degree of flexibility increases the variety of formations and therefore prevents from being analyzed by the opponent team.

Obviously, the optimal mixtures of strictness and flexibility as they are highlighted in fig. 9 depend on the strengths and weaknesses of A and B.

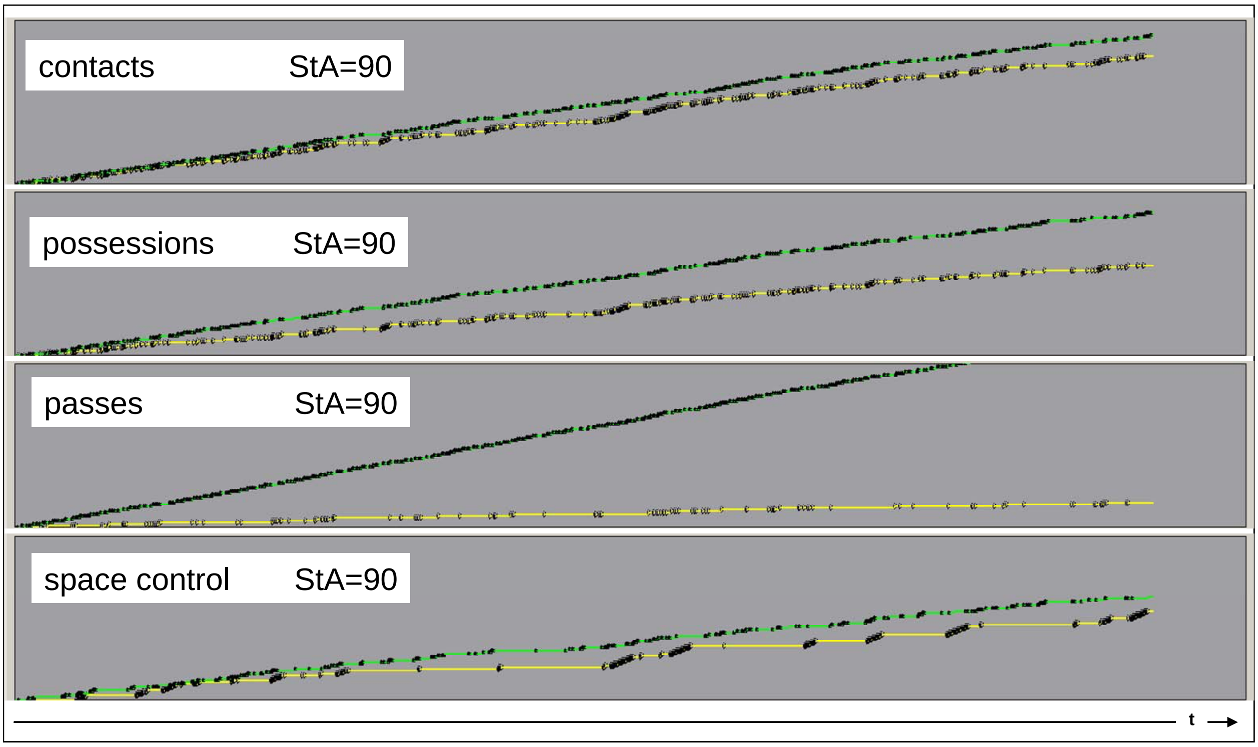


Fig. 11: Original (yellow) and simulated (light green) success profiles of team A.

On the first glance, the high degree of optimal flexibility of team A of up to 40% looks strange, because the strategy used for simulation is developed on the condition of maximizing success. And one could think that having success means to follow that strategy as strict as possible.

On a second glance, game theory demonstrates that a mixture of varying strategies often is superior to following one single strategy. The reason is that a single strategy can be recognized by the opponent and disproved by adapting its behaviour.

(Also see Durlauf & Blume, 2010 and compare the comments and examples regarding fig. 3 and the comments from above regarding fig. 6.)

Finally, the aspect of creativity has to be discussed:

Creativity

Activities are called *creative*, if they are *rare* as well as *successful* (Memmert, 2011). If a small x% of flexibility, i.e. of random events, improve the success like the reduction of strictness by about 8% in case of team B in fig. 9 does, this can be called a *creative* variation of the successful strategy. Transferred to the game this means that team B every 3 to 4 minutes continues its formation situation with a formation that is not expectable in the context of its recognizable strategic behaviour.

Different is the behaviour of team A: Fig. 9 shows a flexibility of about 30%, which does not mean *rare*, and transferred to the play means 1 unexpected transition every 1.5 minute. This could be said to be *unmethodical* or *out of strategy* rather than *creative* but could also represent a flexible mixture of strategies.

Note that all those discussions and interpretations are done on the basis of simulated behaviour. The question now is whether the original playing behaviour regarding ball contacts was strategic strict, creative or unmethodical. The answer can be found by simulating the playing behaviour under the condition of best fitting values of *StA* and *StB*, the surprising result of which is shown in fig. 12 in case of ball contacts:

The best fitting strictness value of team A is StA=100, which means that A in the original play follows strictly a single strategy.

The best fitting strictness value of team B is StB=30, which means that B plays more or less without any single strategic concept – or with a strategic concept consisting of a mixture of partial strategies as would be recommended by game theory.

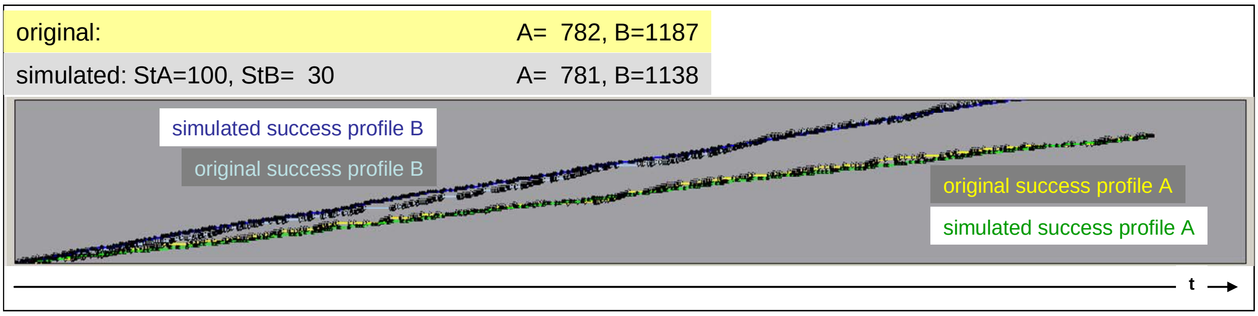


Fig. 12: Original success profiles of ball contacts compared to best fitting simulated profiles.

The result is that the strategic orientation of a team – e.g. unmethodical, flexible, creative, strict – could be detected by a best fitting simulation as is demonstrated exemplarily for team A for 3 success indicators in fig. 13.

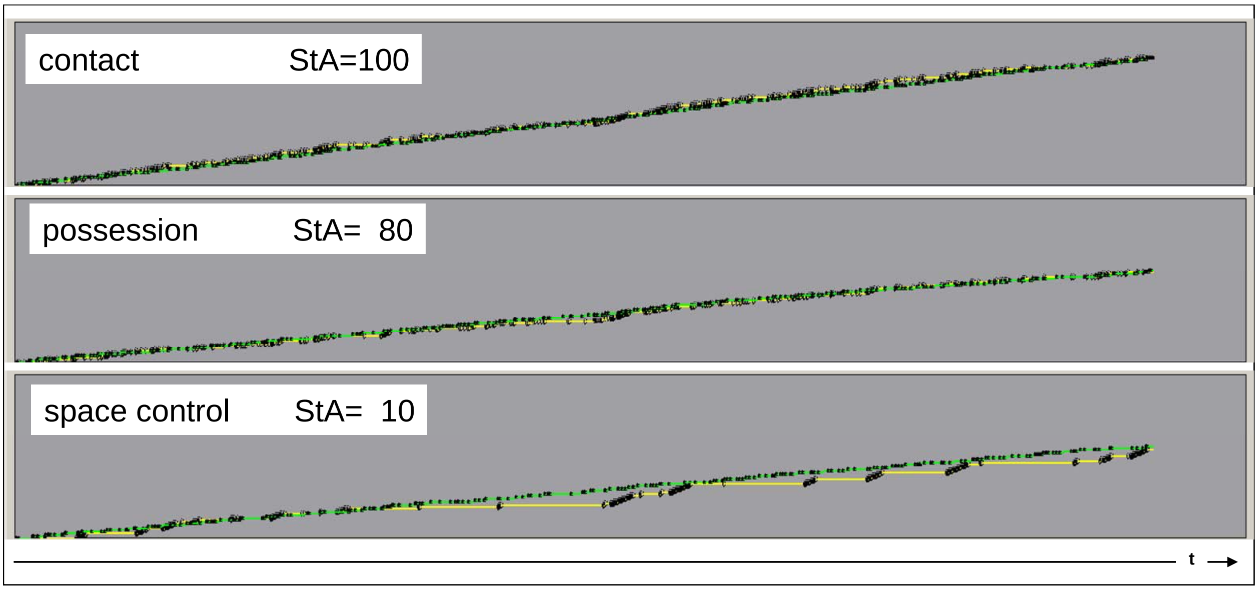


Fig. 13: Analyses of strategic strictness by means of best fitting simulation. The colours are the same as in fig. 12.

As can be seen, the original behaviour of team A (yellow profiles) can be simulated quite well (light green profiles) resulting in characteristic strictness values. Table 1 compares these characteristic values of team A with the respective values of team B. The result is that team A and B seem to act with quite different orientations on strategic planes.

Interpretation of the behaviour of team A (also compare fig. 6):

Regarding ball contacts, team A follows a strict plan where only few formations are involved in ball contacts.

In phases of ball possession, the variety of formations is greater and could be interpreted to be *strategic* with *creative* elements, although the degree of flexibility is rather high.

Finally, space control is not at all strategically planned but follows randomly the current situations.

Quite different is the situation with team B:

Ball contacts and possessions seem to follow randomly the current situations without orientation on strategic planes, while space control is generated much more purposeful than is done by team A, based on a flexible variety and a comparably great number of formations.

Table 1: Simulation-based interpretation of tactical concepts.

	A	B	A	B
	degree of strictness		interpretation	
ball contacts	100	30	strictly strategic	un-methodical
ball possessions	80	30	creative elements	un-methodical
space control	10	60	by chance	flexible

Discussion

As could be shown, simulation can help not only to improve strategic behaviour but also to better understand playing dynamics and (hidden) strategic concepts. Of course, such results have to be taken with care: There are broad gaps between strategic concepts, tactical activities and the way of technically put it to practice.

Moreover: While in "mathematical" games like for instance poker or chess all possible situations and activities are well-defined and (at least in principal) computable, the situation in a team play is much more difficult. The players do not have complete information about context situations, players' positions, or formations.

The quality of their activities is not perfect but depends on their technical abilities and attacking activities of opponent players. One example is that of passes: As is mentioned above, a pass is imbedded in a complex process context of time, space, positions and formations, the passing and the receiving player cannot recognize and evaluate completely every second as computers could. Insofar, simulation of course is not meant to optimize individual tactical behaviour.

Nevertheless, such simulative approaches could be helpful to compare and evaluate basic "one-dimensional" strategies like tiki taka, fast counter attack, concrete-defence or kick-and-rush regarding interaction and success. And, last but not least, they allow for better recognizing *how* teams get strategies to work: strictly, with creative elements, with flexibility, randomly, or not at all.

References

Durlauf, S. N. & Blume, L.E. (2010). *Game Theory*. Palgrave Macmillan.

Grunz, A., Perl, J., & Memmert, D. (2012). Tactical pattern recognition in soccer games by means of special Self-Organizing Maps. *Human Movement Science, 31*(2), 334-43.

Kohonen, T. (1995). *Self-organizing maps*. Berlin, Heidelberg, New York: Springer.

Memmert D. (2011). Sports and Creativity. M.A. Runco and S.R. Pritzker (Eds.) Encyclopedia of Creativity, Second Edition, *San Diego: Academic Press. 2,* 373-378.

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. doi:10.1007/s40279-016-0562-5.

Perl, J. (2015) Modelling and simulation, in A. Baca (ed) Computer Science in Sport, London, New York: Routledge.

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. & Memmert, D. (2017). A Pilot Study on Offensive Success in Soccer Based on Space and Ball Control – Key Performnace Indicators and Key to Understand Game Dynamics. *International Journal of Computer Science in Sport, 16* (1), 65-75.

Perl, J. & Memmert, D. (2018) Soccer: Process and Interaction. In A. Baca and J. Perl (eds.), *Modelling and Simulation in Sport and Exercise* (pp. 73-94). London, New York: Routledge.