

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY
A
INFORMATIKY

Bakalárska práca

Bratislava 2011

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY
A
INFORMATIKY

FUTBALOVÝ ZÁPAS AKO DYNAMICKÝ SYSTÉM
II.

Bakalárska práca

Evidenčné číslo : 40f0fc2f-6e7e-46cd-9d56-9281b73a014c

Študijný program : Ekonomická a finančná matematika

Študijný odbor : 9.1.9 Aplikovaná matematika

Školiace pracovisko : Katedra aplikovanej matematiky a štatistiky

Školiteľ : Mgr. Richard Kollár, PhD.

Bratislava 2011

Jozef Iskra



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Jozef Iskra
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: bakalárska
Jazyk záverečnej práce: slovenský

Názov : Futbalový zápas ako dynamický systém II
Cieľ : Návrh dynamického systému, ktorý bude modelovať priebeh futbalového zápasu ako číselná funkciu.
Anotácia : Pred 20 rokmi zostavili psychológovia systém hodnotenia reakcií účastníkov diskusie, ktorý umožňuje zaznamenávať priebeh diskusie číselne ako funkciu času jednotlivcov. Cieľom je zostaviť podobný systém pre hráčov futbalového zápasu. Toto je jedna z dvojice súvisiacich tém, na ktoré hľadám dvojicu spolupracujúcich študentov.

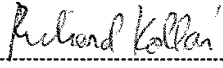
Vedúci : Mgr. Richard Kollár, PhD.

Dátum zadania: 27.10.2010

Dátum schválenia: 24.05.2011

doc. RNDr. Margaréta Halická, CSc.
garant študijného programu

.....
študent

.....

vedúci práce

Dátum potvrdenia finálnej verzie práce, súhlas s jej odovzdaním (vrátane spôsobu prístupnosti)

.....
vedúci práce

Čestné prehlásenie

Čestne prehlasujem, že som túto bakalársku prácu vypracoval samostatne za pomoci konzultácií a uvedených bibliografických odkazov.

V Bratislave, jún 2011

.....

Jozef Iskra

Pod'akovanie

Ďakujem vedúcemu svojej bakalárskej práce Mgr. Richardovi Kollárovi, PhD. za aktívne vedenie pri vypracovávaní témy a za ochotu a snahu naučiť a taktiež chcem poďakovať Borisovi Letkovi za výbornú spoluprácu.

Abstrakt

Jozef Iskra, Futbalový zápas ako dynamický systém II, Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky, vedúci práce Mgr. Richard Kollár, PhD., Bratislava, 2011, 40 strán.

Cieľom bakalárskej práce je vytvorenie matematického modelu futbalového zápasu, ktorý opisuje zotrvačnosť výkonu družstva a jeho reakciu na výkon súpera, podobne ako v existujúcom modeli manželskej interakcie. Parametre modelu založeného na diskretnom dynamickom systéme sme odhadovali lineárnou regresiou. Ich použitím by sme mohli lepšie pochopiť priebeh zápasu a vzájomnej interakcie tímov. V záverečnej časti sú zhrnuté výsledky všeobecného modelu spolu s čiastkovými výsledkami pre jednotlivé tímy.

Kľúčové slová :

matematické modely v športe, analýza futbalového zápasu, diskretný dynamický systém, lineárna regresia

Abstract

Jozef Iskra, Football match as the dynamical system II, Comenius University, Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics, supervisor Mgr. Richard Kollár, PhD., Bratislava 2011, 40 pages

The aim of the bachelor thesis is to create a mathematical model of football game, which describes inertia of the team performance and its reaction to the opponent's performance similarly as for the existing model of marital interactions. The parameters of the model based on a discrete dynamical system were established by linear regression. Using these parameters we could better understand the process of the match and the mutual interaction of the teams. In the final section the results of the general model are summarized along with the partial results for individual teams.

Key words :

mathematical models in sport, soccer game analysis, discrete dynamical systems, linear regression

Obsah

Zoznam ilustrácií a tabuliek.....	7
Úvod.....	8
1. Existujúci výskum.....	9
1.1 Futbalové výskumy.....	9
1.2 Modelling the dynamics of marital interaction.....	10
2. Použité dáta.....	11
3. Analýza modelov.....	14
3.1 Model s totožným vplyvom $B_{(t)}$ a $B_{(t-1)}$	14
3.2 Model s rôznym vplyvom $B_{(t)}$ a $B_{(t-1)}$	16
3.3 Model s rozličným sklonom $B_{(t)}$ a $B_{(t-1)}$	17
3.4 Model s rozličným sklonom $B_{(t-1)}$	21
3.5 Model s rozličným sklonom $B_{(t)}$	23
3.6 FC Barcelona.....	25
3.7 Manchester United.....	26
3.8 Tottenham Hotspur.....	28
4. Vyhodnotenie výsledkov.....	29
5. Riešenia systémov.....	34
Záver.....	39
Zoznam bibliografických odkazov.....	40

Zoznam ilustrácií a tabuliek

1. Existujúci výskum

Obr.1: Nabehaná vzdialenosť hráčov

Obr.2: Nabehaná vzdialenosť skupinou hráčov

Obr.3: Regulated

Obr.4: Nonregulated

Obr.5: Null clines

Obr.6: Pevné body

2. Použité dáta

Obr.7: Body vážené podľa čistého času

Obr.8: Kumulované body

3. Analýza modelov

Tabuľka 1: Výsledky modelu s totožným vplyvom $B_{(t)}$ a $B_{(t-1)}$

Tabuľka 2: Výsledky modelu s rôznym vplyvom $B_{(t)}$ a $B_{(t-1)}$

Tabuľka 3: Výsledky modelu s rozličným sklonom $B_{(t)}$ a $B_{(t-1)}$

Tabuľka 4: Výsledky modelu s rozličným sklonom $B_{(t-1)}$

Tabuľka 5: Výsledky modelu s rozličným sklonom $B_{(t)}$

Tabuľka 6: Výsledky modelu FC Barcelona

Tabuľka 7: Výsledky modelu Manchester United

Tabuľka 8: Výsledky modelu Tottenham Hotspur

5. Riešenia systémov

Obr.9: stabilita pevného bodu (0,0)

Obr.10: stabilita pevného bodu (0,0) pre FCB-TOT

Úvod

Futbal je obrovský fenomén a bezpochyby sa radí medzi najpopulárnejšie športy na svete. V každej krajine má svojich fanúšikov radiacich sa od laikov obdivujúcich presné prihrávky, rýchle kombinácie, či krásne góly, až po zarytých obdivovateľov, ktorí sa snažia z každého zápasu vydolovať čo najviac informácií.

Početné skupiny odborníkov sústavne pracujú na analyzovaní rôznych faktorov vplývajúcich na priebeh zápasu. Bolo vytvorených niekoľko výskumov zameraných napríklad na úspešnosť strelby v závislosti od početnosti prihrávok a od streleckej pozície, nabehanú vzdialenosť podľa hráčovho postavenia, a i. Ich cieľom je zvýšiť efektivitu tréningu zameraním sa na najdôležitejšie prvky, či vymyslieť najvhodnejšiu taktiku na zdolanie súpera. Všetky tieto výskumy sa však zameriavajú len na jeden tím, poprípade vplyv jednotlivca na výkon tímu, no nesledujú vplyv celého tímu na hru svojho súpera.

Cieľom tejto práce bolo skúmať futbal z matematického hľadiska, popísať ho vhodným modelom a pomocou toho určiť kvantitatívne parametre popisujúce hru jednotlivých družstiev. Nami hľadané parametre boli zotrvačnosť výkonu a taktiež reakcia na výkon súpera, vďaka ktorým by sme mohli predpovedať priebeh zápasu. Základ na testovanie matematického modelu tvoria dáta zozbierané z vybraných futbalových zápasov. Tieto dáta boli získané pomocou bodovacieho systému, ktorý zachytáva určité herné situácie v priebehu zápasu. Zaznamenáva kvalitu hry jednotlivých tímov a reakcie tímov na súperovu hru počas úvodných pätnástich minút.

V dosiahnutých výsledkoch sme pozorovali 3 rôzne formy reakcií na kvalitu súperovej hry a vlastnú zotrvačnosť výkonu.

Motivácia tejto práce je hlbšie pochopenie niektorých herných situácií a futbalových javov, za účelom ich lepšej interpretácie a následne možným zlepšením prípravy hráčov na stretnutie a reakcie na nepriaznivý priebeh zápasu.

V tejto práci sme sa opierali o existujúci výskum vedcov Johna M. Gottmana a Jamesa D. Murray-a – Modelling the dynamics of marital interaction, ktorí modelovali vývoj manželského vzťahu založeného na interakcii manželov počas ich vzájomného rozhovoru. Analógiou manželských rozhovorov budú v našom prípade futbalové zápasy.

1. Existujúci výskum

1.1 Futbalové výskumy

Ohľadom futbalu existuje veľké množstvo výskumov, ktoré sa snažia zaznamenávať a následne vyhodnocovať najdôležitejšie herné situácie. Vedome, či nevedome to začali praktizovať tréneri, skauti a manažéri, ktorí si počas zápasov robili poznámky o hráčoch a o priebehu zápasu, aby vedeli na čom treba popracovať. S postupom času sa tento systém vyvíjal a začali vznikať rôzne štúdie týkajúce sa futbalovej problematiky.

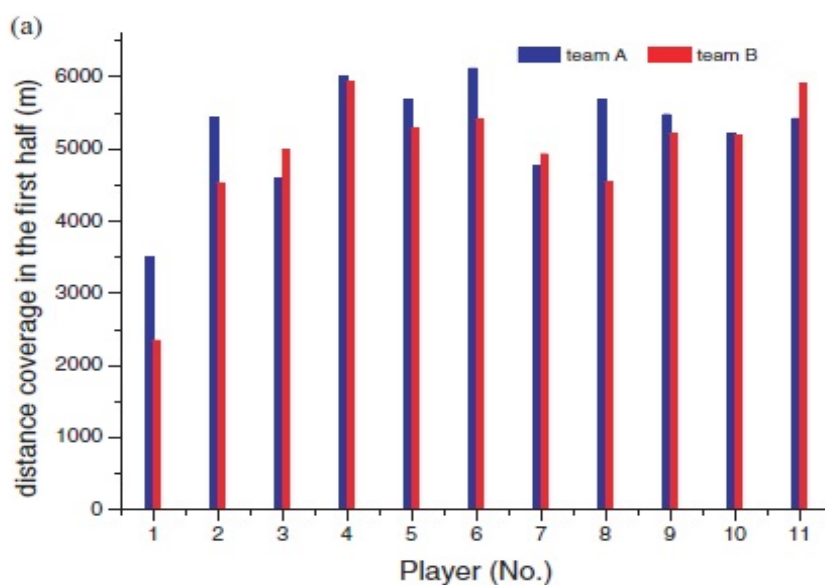
Jednou z nich je štúdia, ktorú vytvorili Reep a Benjamin (1968), ktorí zozbierali dáta z 3 213 zápasov medzi rokmi 1953 až 1968. Zamerali sa na prihrávky a strely. Zistili, že 80% gólov vyplynulo zo sekvencie troch alebo menej prihrávok, a že 50 percent všetkých gólov prišlo z pozícií získaných v poslednej útočnej štvrtine ihriska [1].

Neskôr Bate (1988) zistil, že 94% gólov, na všetkých stupňoch medzinárodnej úrovne, bolo strelených z akcií so štyrmi, alebo menej prihrávkami, a že 50-60% zo všetkých akcií vedúcich k strele pochádzalo z útočnej tretiny ihriska. Bate skúmal aspekty gólových šancí vo futbale a ich vzťah k taktike a stratégií so zreteľom na výsledky prezentované Keepom a Benjaminom (1968). Tvrdil, že tím nestrelí gól, pokiaľ nedostane loptu a jedného alebo viac útočníkov do útočnej tretiny ihriska. Čím viac držaní lopty tím má, tým väčšia je šanca, že prenikne do útočnej tretiny ihriska, a teda si vytvorí viac príležitostí na strelenie gólu. Čím je počet prihrávok počas držania lopty väčší, tým nižší je celkový počet držaní lopty počas zápasu tým aj celkový počet preniknutí do útočnej tretiny a teda aj počet šancí na strelenie gólu je menší. Tak Bate odmietol koncept hry s držaním lopty a uprednostňoval priamejšie stratégie. Uzavrel to tým, že pre zvýšenie množstva gólových príležitostí by tím mal:

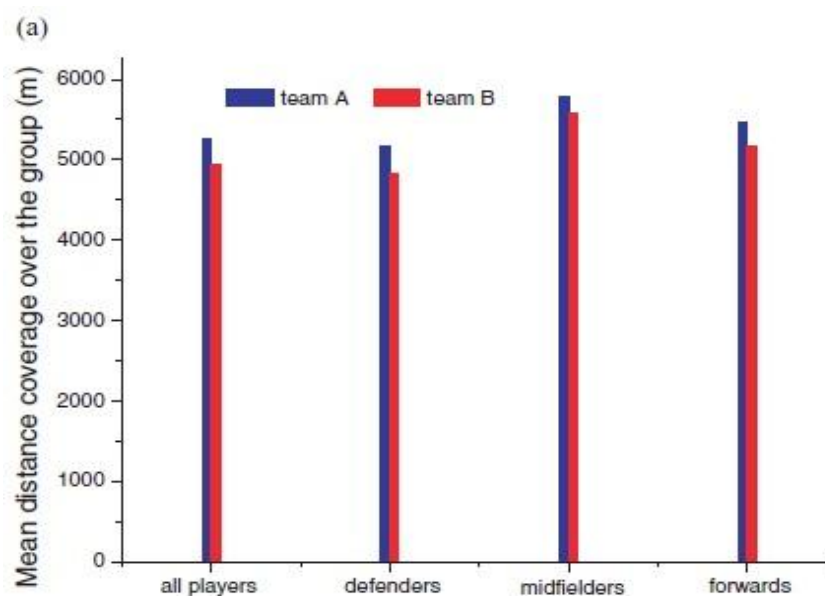
1. hrať loptu smerom dopredu tak často, ako je to možné
2. zredukovať počet spätných prihrávok na minimum
3. zvýšiť množstvo dlhých prihrávok smerom dopredu a útočníkov bežiacich s loptou
4. hrať loptu do voľného priestoru tak často, ako je to možné

Tieto odporúčania sú v súlade s tým, čo je známe ako ‘direct method’ (priama metóda) alebo ‘long-ball game’ (hra dlhými loptami). Tento prístup sa osvedčil u niektorých tímov v nižších divíziách anglickej ligy [1].

Účelom jednej z ďalších štúdií bolo uskutočniť podrobnú kvantitatívnu analýzu pre niekoľko aspektov z jedného zápasu. Individuálne a kolektívne správanie hráčov a tímov a pohyb lopty prezentovali ako rôzne časové rady. Jednou zo skúmaných vecí bola nabehaná vzdialenosť (obr. 1, 2).



Obr. 1 Nabehaná vzdialenosť hráčov



Obr. 2 Nabehaná vzdialenosť skupinou hráčov

Na obr. 1 Nabehaná vzdialenosť hráčov je graf nabehanej vzdialenosti v metroch počas prvého polčasu. Na osi x sú čísla všetkých hráčov oboch tímov (tím A – modrá, tím B – červená). Na obr. 2 Nabehaná vzdialenosť skupinou hráčov sú na x-ovej osi skupiny hráčov, v poradí: všetci hráči, obrancovia, stredopoliari, útočníci. Na y-ovej osi je priemerná vzdialenosť nabehaná skupinou v metroch, počas prvého polčasu. Modrá – tím A, červená – tím B.

Výsledky tejto štúdie by mohli byť použité na porovnávanie dvoch tímov, rozličných skupín hráčov, alebo jednotlivých hráčov [2].

O tom, že podobné štúdie majú význam, svedčí použitie výskumu na zistenie vplyvu zmien futbalových pravidiel. Aplikovali ho na zápasy Premier Division v Anglicku v sezóne 1992/93. Zistilo sa, že strata času spôsobená zdržiavaním brankára bola výrazne zredukovaná, takisto, že stred poľa bol viac vyťažovaný ako následok menšieho počtu spätných prihrávok, alebo, že napriek zvýšenému množstvu chýb v obrane sa nezvýšil počet gólov alebo gólových príležitostí.

V iných športoch, ako je napríklad basketbal existujú na popísanie zápasu určité kritické situácie, ktoré sa následne štatisticky vyhodnocujú. Jednou z takých štatistík je napríklad hodnotenie +/- bodov, ktoré vyhodnocuje prínos jednotlivých hráčov, ale aj kombinácie rôznych hráčov pre mužstvo. Vo futbale, však nie je jednoduché kritické situácie odhaliť a na tvorbu podobnej štatistiky ich je veľmi málo. Preto treba zvoliť iný prístup. Ideálne by bolo pristupovať k tomu objektívne, ale keďže je ťažké špecifikovať objektívne merateľné parametre činností vplyvujúcich na šancu na strelenie gólu (ako technika, napadanie a pod.) – v tejto práci sa budeme opierať o kombináciu objektívnych a subjektívnych kritérií, skombinovaných do komplexného hodnotenie výkonu tímu [3].

Veľkou výhodou je, že v tejto práci nie je hodnotenie výkonu zamerané len na útočnú aktivitu, ale aj na obrannú fázu, vrátane obrannej súčinnosti.

1.2 Modelling the dynamics of marital interaction

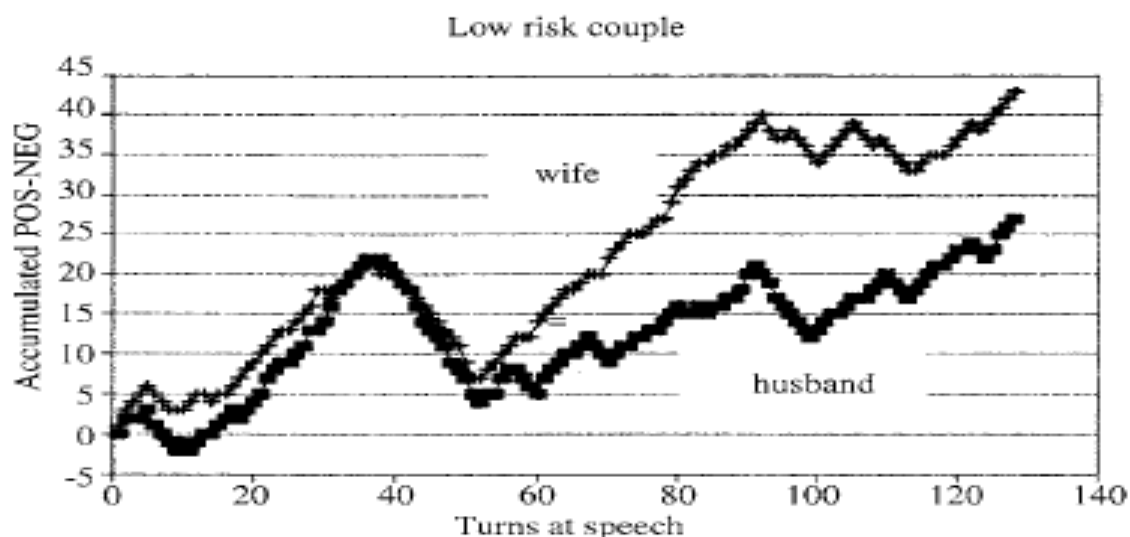
Hoci existuje dostatok futbalových štúdií, hlavný výskum, o ktorý sme sa pri písaní tejto práce opierali, je *Modelling the dynamics of marital interaction* [4]. V tejto sekcii priblížime obsah tejto práce, v ktorej sme sa zamerali hlavne na matematickú časť.

Budeme ilustrovať použitú matematickú metódu, keďže náš postup bude analogicky a táto metóda bude aplikovateľná aj v prípade modelovania futbalového zápasu.

Tento výskum bol pre nás kľúčový hlavne z toho dôvodu, že významnú časť tvorila interakcia medzi manželmi (manželský pár môžeme do istej miery vnímať ako futbalové tímy hrajúce spoločný zápas – jeden tím neustále reaguje na správanie toho druhého) a doteraz vo futbale neexistujú výskumy, ktoré by skúmali interakciu medzi dvomi tímami hrajúcimi spoločný zápas.

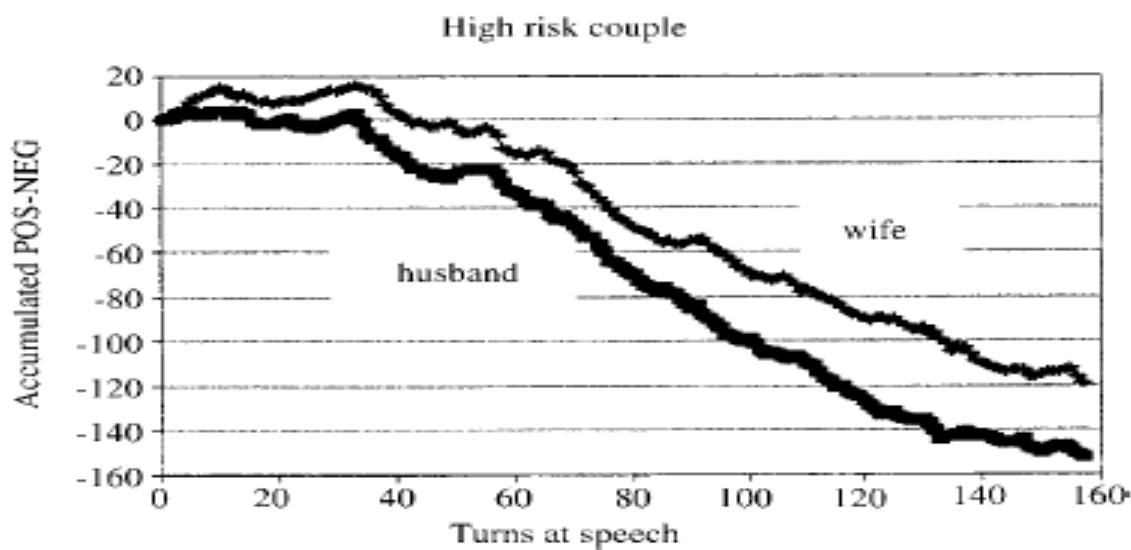
Jednou z príčin skúmania manželstva bola zvyšujúca sa miera rozvodovosti vo vyspelých krajinách. Napríklad, od 50 do 67% prvých manželstiev v USA skončí rozvodom v 40 ročnom období, pri druhých manželstvách je to približne o 10% viac. Pochopenie reakcií na partnera a predpovede pravdepodobnosti rozvodu by mohli zvýšiť úspešnosť manželských poradní a tým znížiť počet rozvodov.

Do štúdie bolo zahrnutých 73 manželských párov, ktorí boli sledovaní v období rokov 1983 až 1987. Na získanie dát bol použitý kódovací systém na zachytenie interakčného správania nazvaný *Rapid Couples Interaction Scoring System (RCISS)*. Na základe toho boli páry rozdelené do dvoch skupín (obr. 2, 3): *regulated* (nízko riziková) a *nonregulated* (vysoko riziková).



Obr.3: Regulated

Na obr.3: Regulated je nízko rizikový pár s prevahou pozitívnych reakcií.



Obr.4: Nonregulated

Na obr.4: Nonregulated je vysoko rizikový pár s prevahou negatívnych reakcií.

Následne bolo skóre získané z RCISS použité na modelovanie. Sled skóre je vyjadrený diferenčnými rovnicami:

$$\begin{aligned} W_{t+1} &= f(W_t, H_t), \\ H_{t+1} &= g(W_{t+1}, H_t), \end{aligned} \quad (5.1)$$

kde W_t a H_t je skóre manželky a manžela v čase t a predpokladáme, že je závislé od vlastného a partnerovho predchádzajúceho skóre, teda sa skladá z dvoch funkcií: *uninfluence function* (funkcia bez vplyvu partnera) a *influence function* (funkcia vplyvu na partnera). Asymetria v indexoch je spôsobená predpokladom, že manželka rozprávala prvá. Uninfluence function bolo modelované jednoduchou lineárnou rovnicou:

$$P_{t+1} = r_i P_t + a_i, \quad (5.2)$$

kde P_t je skóre v čase t , r_i je zotrvačnosť nálady a a je konštanta. Táto rovnica sa dá riešiť postupnými iteráciami. Ak P_0 je štartovací bod v čase 0, dostávame:

$$\begin{aligned} P_1 &= r_i P_0 + a \\ P_2 &= r_i P_1 + a = r_i^2 P_0 + a(1+r_i) \\ &\dots \\ P_t &= r_i^t P_0 + a(1+r_i^t) / (1-r_i) \end{aligned} \quad (5.3)$$

V takomto prípade je pevný bod stabilný, ak je splnená podmienka $|r_i| < 1$.

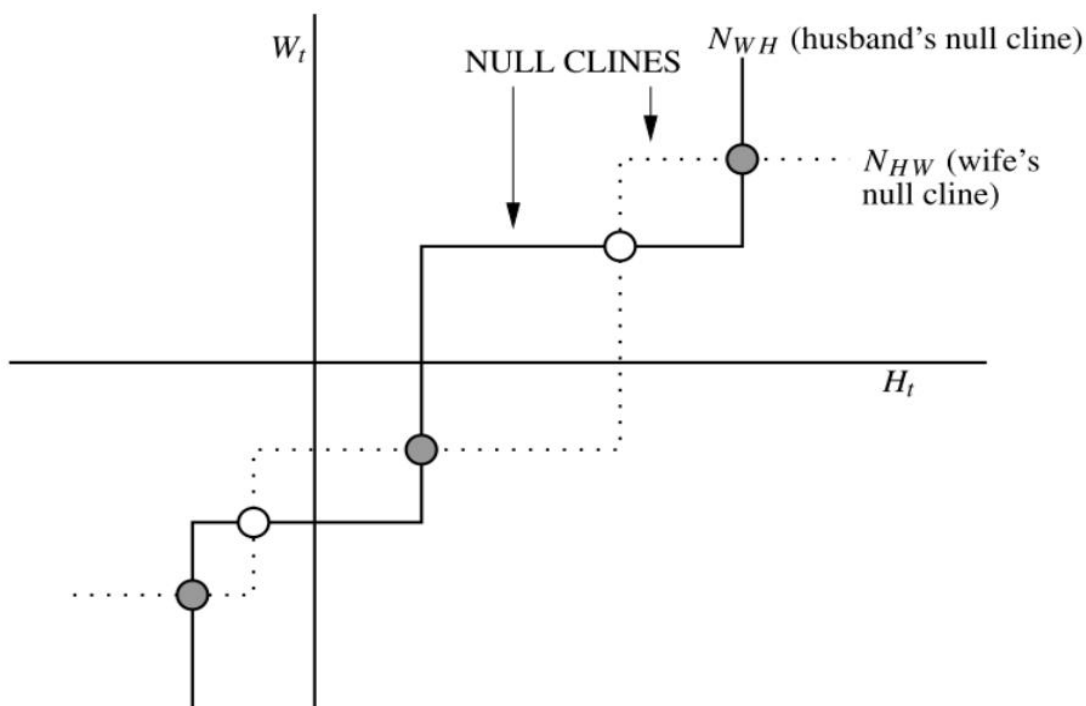
Na odhadnutie influence function bol použitý nasledujúci model:

$$W_{t+1} = I_{HW}(H_t) + r_1 W_t + a, \quad (5.4)$$

$$H_{t+1} = I_{WH}(W_{t+1}) + r_2 H_t + b, \quad (5.5)$$

kde $I_{AB}(A_t)$ je funkcia vyjadrujúca vplyv osoby A v čase t na osobu B , r_i je zotrvačnosť nálady a a, b sú konštanty. Za predpokladu, že influence function v bode nula má nulovú hodnotu, sa zoberali do úvahy len body, kde partnerovo skóre bolo nulové (približne 15%) a teda rovnica (5.4) sa stáva lineárnou. Pomocou metódou najmenších štvorcov potom boli z dát odhadnuté parametre r_1, r_2, a a b . Keď boli odhadnuté komponenty *uninfluence function* bolo možné vypočítať influence function $I_{HW}(H_t) = W_{t+1} - r_1 W_t - a$ pre každé H_t . Analogicky to isté pre $I_{WH}(W_t)$. Hľadaním stability a pevných bodov za podmienok $W_{t+1} = W_t$ a $H_{t+1} = H_t$ a dosadením do rovníc (5.4) a (5.5) dostaneme rovnice:

$$N_{HW}: W(H_t) = \frac{I_{HW}(H_t) + a}{1 - r_1}, \quad N_{WH}: H(W_t) = \frac{I_{WH}(W_t) + b}{1 - r_2}. \quad (5.6)$$



Obr.5: Null clines

Na obr.5: Null clines, kde x-ovú os tvoria body manžela H_t a y-ovú os tvoria body manželky W_t , sú zobrazené null clines z funkcií (5.6), ktorých priesečníky určujú pevné body. Ak je viacero pevných bodov, striedajú sa stabilné pevné body (plné kruhy) s nestabilnými pevnými bodmi (prázdne kruhy).

Riešením stability pevných bodov H_S, W_S , kde

$$W_t = W_S + w_t, \quad H_t = H_S + h_t, \quad (5.7)$$

pričom $|w_t|$ a $|h_t|$ sú malé, substitúciou (5.7) do (5.4) a (5.5) a upravením na Taylorov rozvoj s ponechaním iba lineárnej časti dostávame:

$$\begin{aligned} W_{t+1} &= W_S + w_{t+1} = I_{HW}(H_S + h_t) + r_1(W_S + w_t) + a \\ &\approx I_{HW}(H_S) + h_t I'_{HW}(H_S) + r_1(W_S + w_t) + a, \end{aligned} \quad (5.8)$$

kde $I'_{HW}(H_S)$ je $d I_{HW}(H_t)/dH$ vyčíslené v pevnom bode H_S , ale $I_{HW}(H_S) + r_1(W_S) + a = W_S$ podľa definície o pevnom bode W_S , teda (5.8) zjednodušíme na tvar:

$$w_{t+1} = r_1 w_t + I'_{HW}(H_S) h_t. \quad (5.9)$$

Podobný postup uplatníme aj na rovnicu (5.5) a dostaneme:

$$h_{t+1} = r_2 h_t + I'_{WH}(W_S) w_{t+1}, \quad (5.10)$$

čo spolu tvorí nasledujúci systém lineárnych diferenčných rovníc:

$$\begin{pmatrix} w_{t+1} \\ h_{t+1} \end{pmatrix} = \begin{pmatrix} r_1 & I'_{HW}(H_S) \\ r_1 I'_{WH}(W_S) & r_2 + I'_{WH}(W_S) I'_{HW}(H_S) \end{pmatrix} * \begin{pmatrix} w_t \\ h_t \end{pmatrix} = M * \begin{pmatrix} w_t \\ h_t \end{pmatrix}, \quad (5.11)$$

kde matica M , definovaná systémom (5.11) je matica stability a jej vlastné čísla sú

$$\lambda_1, \lambda_2 = [(r_1 + r_2 + I'_{WH} I'_{HW}) \pm \{(r_1 + r_2 + I'_{WH} I'_{HW})^2 - 4r_1 r_2\}^{1/2}] / 2. \quad (5.14)$$

Pre podmienky stability musí byť splnená podmienka $|\lambda_i| < 1$, v takomto prípade, w_t a h_t konverguje k nule. Túto podmienku aplikujeme na (5.14) a získame podmienky pre $r_1, r_2, I'_{HW}(H_S), I'_{WH}(W_S)$ pre stabilné pevné body H_S, W_S . Pre zachovanie stabilného pevného bodu z (5.3) uvažujeme iba r_i z intervalu (0,1). Podmienku stability $|\lambda_i| < 1$ teraz môžeme upraviť na tvar:

$$I'_{WH}(W_S) I'_{HW}(H_S) < (1-r_1)(1-r_2). \quad (5.17)$$

Teda ak chceme posúdiť stabilitu pevných bodov, musíme vyhodnotiť derivácie influence function a použiť (5.17). Ak teraz zderivujeme (5.6) podľa H a W dostaneme

$$\frac{dW}{dH} \Big|_{N_{HW}} = \frac{I'_{HW}(H)}{(1-r_1)}, \quad \frac{dH}{dW} \Big|_{N_{WH}} = \frac{I'_{WH}(W)}{(1-r_2)}$$

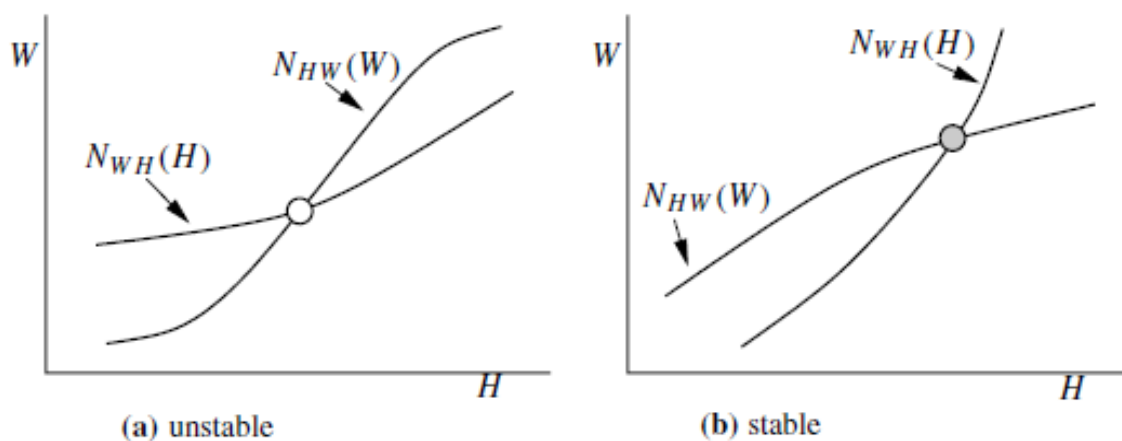
Podmienky stability (5.17) sa teda zjednodušia na tvar (v bodoch W_S, H_S)

$$\frac{dW}{dH} \Big|_{N_{HW}} * \frac{dH}{dW} \Big|_{N_{WH}} < 1.$$

Po zvolení osí W ako vertikálna os a H ako horizontálna os, sa dá posledný výraz zabezpečujúci stabilitu pevného bodu v bodoch W_S, H_S zapísať ako

$$\frac{dW}{dH} \Big|_{N_{HW}} < \frac{dH}{dW} \Big|_{N_{WH}}. \quad (5.19)$$

Teda všetko čo je teraz potrebné urobiť je vyhodnotiť gradienty a okamžite vieme rozhodnúť o stabilite, alebo nestabilite pevného bodu. Príklady použitia null clines a pevných bodov je vidieť na obrázku 6.



Obr.6: Pevné body

Na obr. 6: Pevné body sú príklady null clines a stability pevných bodov pre model manželstva. Pevné body korešpondujú bodom, v ktorých sa null clines pretínajú. Stabilita pevného bodu je určená vyšetrením gradientov. V a) je gradient manželkinho null cline väčší ako gradient manželovho null cline, teda podmienka (5.19) nie je splnená a dotyčný pevný bod je nestabilný. Keď sa pozrieme na b), vidíme, že podmienka (5.19) je splnená a pevný bod je stabilný. H = manžel, W = manželka.

Takto zostavený model poskytol bližší pohľad na jednotlivé manželstvá a vzťahy medzi manželmi. Umožňovalo to vyhodnocovať stabilitu a nestabilitu partnerov a tým do istej miery predpovedať úspešnosť vzťahov. Takisto bližšie poznanie vlastností partnerov má vplyv na zvýšenie úspešností manželských terapií.

Predpokladáme, že podobným spôsobom by sa dali vyhodnocovať aj vzťahy medzi futbalovými tímami. Zostavenie modelu a odhalenie interakcie by mohlo poskytnúť účinnú pomôcku pre manažérov a trénerov, aby včas zareagovali na nepriaznivý priebeh zápasu.

2. Použité dáta

Bližší rozbor použitých dát, ako napríklad systém vyhodnocovania jednotlivých kategórií, spôsoby váženia a ich výhody alebo nevýhody, rozbor grafy, či korelačná analýza dát, je popísaný v práci *Futbalový zápas ako dynamický systém I* [5]. V tejto časti sú v krátkosti popísané dáta s ktorými sme pracovali.

Pre začatie práce sme potrebovali získať dostatočnú vzorku futbalových zápasov. Pre väčšiu rozmanitosť sme vyberali zápasy z rôznych líg a súťaží. Menovite to boli zápasy španielskej a anglickej ligy a ligových pohárov, ligy majstrov a taktiež aj z vrcholného medzinárodného turnaja, z majstrovstiev sveta.

Z vybraných zápasov nás zaujímalo iba prvých 15 minút. Tento časový úsek sme ďalej delili na minútové úseky. Pre jeden zápas sme teda dostali 15 časových úsekov. V každej časti sme následne hodnotili výkony oboch zúčastnených mužstiev, na čo nám slúžili nasledujúce kategórie:

- Prihrávky
- Pohyb bez lopty
- Technika
- Získané územie
- Útočné súboje
- Obranné súboje
- Napádanie
- Pozičná hra

Prvých päť kategórií predstavuje útočnú fázu hry, zvyšné tri naopak obrannú fázu. Výberom týchto kategórií sme sa snažili zachytiť najdôležitejšie aspekty hry, tak aby dobre popisovali priebeh zápasu. Hoci sa môže zdať, že v zápase nesledujeme niektoré ďalšie významné ukazovatele, ako sú strely, alebo góly, no sú však zahrnuté v nami zvolených kategóriách.

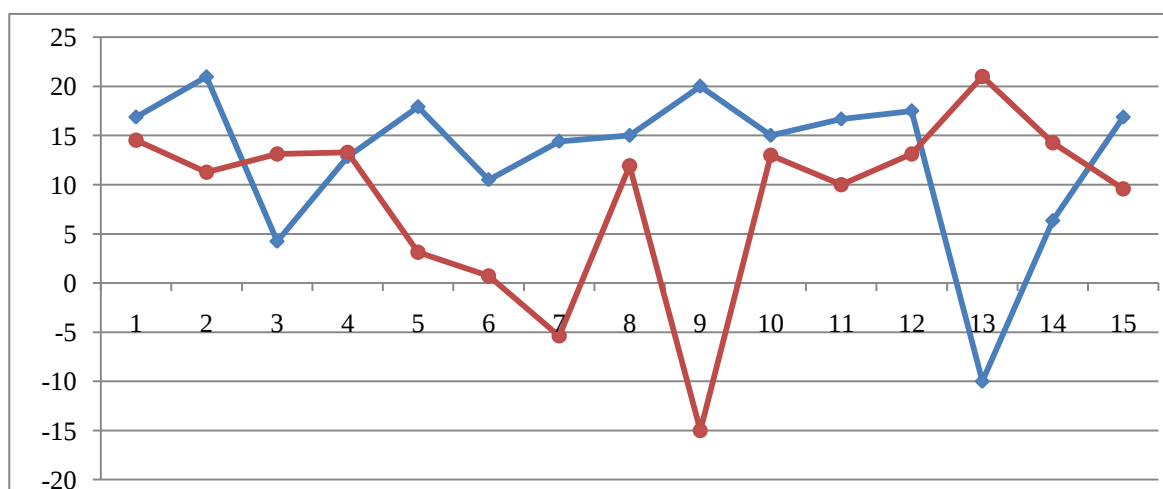
Každá z kategórií mala bodovaciú škálu od -5 po +5, pričom -5 vyjadruje, že pozorovaná činnosť bola najhoršie možné zvládnutá, +5 vyjadruje, že pozorovaná činnosť bola zvládnutá najlepšie ako bolo možné a 0 vyjadruje, že pozorovaná činnosť sa v danej minúte nevyskytla ani raz, poprípade bola neutrálna.

Vplyvom toho, že nami zvolené kategórie boli rozdelené na útočné a obranné, sme hru sme rozdelili na dve časti – útočná a obranná fáza. Útočná fáza hry nastáva vtedy, keď sa tím zmocní lopty, pokračuje jej držaním a končí sa stratou lopty (nepresnou prihrávkou, obranným zákrokom súpera) alebo zakončením akcie (strelou, gólom). Obranná fáza hry naopak začína stratou lopty. Počas tejto fázy má kontrolu nad loptou súper. Táto fáza končí získaním lopty, respektíve stratou lopty súpera, poprípade inkasovaným gólom.

Pomocou váženia sme získaným dátam dodali relevantnosť a komparatívnosť, vďaka čomu sme jednotlivé hodnoty mohli medzi sebou porovnávať. Hodnoty sme vážili tromi rôznymi spôsobmi: podľa celkového času (do výsledného času sa započítaval aj čas, kedy bola hra prerušená), podľa čistého času (hra bez prerušení) a podľa jednotlivých kategórií. Pri vážení podľa času sa útočné parametre vážili podľa času stráveného v útočnej fáze, pre obranné parametre to bolo podľa času, ktorý tím strávil v obrannej fáze.

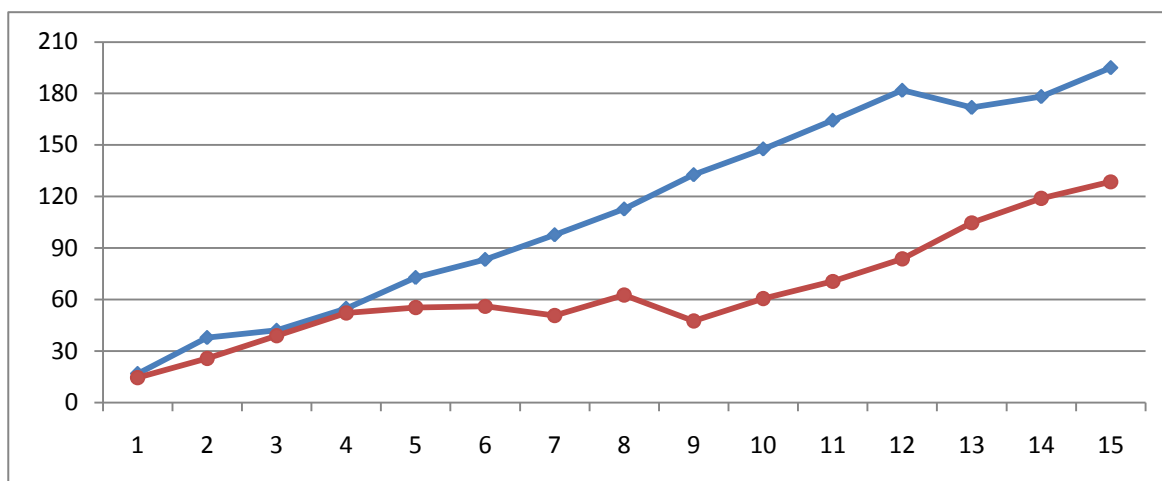
Váženie podľa čistého času má vyššiu výpovednú hodnotu ako podľa celkového času, keďže sa tam nezapočítava čas kedy bola hra prerušená. Na druhej strane je tento spôsob objektivnejší ako váženie podľa jednotlivých kategórií, kde výber ich váh je dosť subjektívny a väčšinou majú jednotlivé kategórie u rôznych tímov rôzne váhy. Pre univerzálnosť modelu sme teda volili váženie podľa čistého času.

Ako príklad zozbieraných dát môžeme uviesť dáta zo zápasu Arsenal – Tottenham (obr. 7, 8), odohraného dňa 20.11.2010.



Obr. 7: Body vážené podľa čistého času

Na *Obr. 7: Body vážené podľa čistého času* sú na osi x zobrazené minúty, na osi y sú hodnoty bodov v jednotlivých minútach, pričom modrá = Arsenal, červená = Tottenham. Pri bližšom pohľade na jednotlivé minúty, si môžeme všimnúť, že Arsenal mal v 11-tich prípadoch vyššie hodnoty ako Tottenham. Naopak Tottenham bol lepší iba v tretej, štvrtej, trinástej a štrnástej minúte.



Obr. 8: Kumulované body

Na *obr. 8: Kumulované body*, kde na x-ovej osi sú minúty je na y-ovej osi súčet bodov z obr. 7 do danej minúty, napr.: v tretej minúte je hodnota daná súčtom hodnôt v prvej, druhej a tretej minúte z obr. 7. Z tohto obrázku môžeme vidieť, že Arsenal (modrá) v úvodných 15-tich minútach nad Tottenhamom (červená) dominoval.

3. Analýza modelov

V tejto časti sa budeme venovať rozboru matematických modelov futbalového zápasu. Popíšeme dosiahnuté výsledky a pokúsime sa ich čo najlepšie interpretovať, aby poskytli hlbší pohľad na to, čo sa počas futbalového zápasu deje.

Na základe dát popísaných v predchádzajúcej kapitole sme sa snažili zostaviť adekvátny model, ktorý by spoľahlivo popisoval priebeh zápasu. Podobne ako tomu bolo v práci vedcov Johna M. Gottmana a Jamesa D. Murray-a – Modelling the dynamics of marital interaction [4], sme sa aj my zamerali hlavne na zotrvačnosť výkonu tímu (t.j. do akej miery si tím udržiava svoju výkonnosť) a interakciu medzi súpermi (t.j. akou mierou ovplyvňuje hru tímu výkon ich súpera).

Nasledujúce modely sme zostavili tak, aby v nich boli požadované okolnosti zachytené a hľadané koeficienty sme následne odhadovali pomocou lineárnej regresie [6].

3.1 Model s totožným vplyvom $B_{(t)}$ a $B_{(t-1)}$

Prvý model bol zostavený tak, aby zachytil zotrvačnosť výkonu a vplyv kvality hry súpera. Predpokladali sme, že hru tímu v aktuálnej minúte ovplyvňuje jeho predchádzajúca hra a to, že svojou hrou reagujú na predchádzajúcu hru súpera, ale taktiež musia reagovať na jeho aktuálnu hru. Odhadovali sme pri tom, že reakcia na aktuálnu a predchádzajúcu hru súpera je totožná. Model bol teda zostavený nasledovne:

$$A_{(t)} = \alpha A_{(t-1)} + \beta [B_{(t)} + B_{(t-1)}],$$

kde $A_{(t)}$, $B_{(t)}$ sú body tímu A , B v čase t , α je zotrvačnosť výkonu a β vyjadruje vplyv súpera. Výsledky regresie sa nachádzajú v tabuľke 1.

Tabuľka 1

	Estimate	Std. Error	t value	Pr(> t)
α	0.53117	0.05266	10.086	< 2e-16 ***
β	0.13244	0.02870	4.614	5.99e-06 ***

Pričom koeficient determinácie tejto regresie je 0.4461, upravený koeficient determinácie je 0.4422, F-štatistika = 114.4 a p-value < $2.2 \cdot 10^{-16}$.

Z p-value vidíme, že naša regresia je signifikantná a teda dáva zmysel, avšak koeficient determinácie nie je veľmi vysoký, čo znamená, že regresia nepopisuje dáta úplne presne [7], [8].

Čo sa týka koeficientov α a β , tie vyšli obidva signifikantné a ich hodnoty sú kladné. To znamená, že hra súpera B a takisto aj hra tímu A v predchádzajúcej minúte pôsobí pozitívne na aktuálnu kvalitu hry tímu A . Naš model teraz môžeme pomocou odhadnutých koeficientov prepísať na tvar

$$A_{(t)} = 0.53117 \cdot A_{(t-1)} + 0.13244 \cdot [B_{(t)} + B_{(t-1)}].$$

Z tohto sa dá usúdiť, že počas futbalového zápasu si tímy udržiavajú približne 53% -nú kvalitu svojej predošlej hry a hra súpera pôsobí na ich výkon približne 13-timi percentami z bodov súpera.

3.2 Model s rôznym vplyvom $B_{(t)}$ a $B_{(t-1)}$

Predchádzajúci model sme sa pokúsili mierne vylepšiť tým, že sme pridali nový predpoklad ohľadom reakcie na súperovu hru. Nové očakávania boli také, že tím A bude odlišne reagovať na hru súpera B v predchádzajúcej minúte, a inak bude reagovať na jeho aktuálnu hru. Model sme teda upravili na tvar

$$A_{(t)} = \alpha A_{(t-1)} + \beta B_{(t)} + \gamma B_{(t-1)},$$

kde α je zotrvačnosť výkonu, β je reakcia na aktuálnu hru súpera a γ je reakcia na hru súpera v predchádzajúcej minúte. Výsledky tohto modelu sú v tabuľke 2.

Tabuľka 2

	Estimate	Std. Error	t value	Pr(> t)
α	0.59132	0.05011	11.800	< 2e-16 ***
β	-0.20183	0.05822	-3.467	0.000609 ***
γ	0.44460	0.05521	8.053	2.27e-14 ***

Koeficient determinácie tejto regresie je 0.5175, upravený koeficient determinácie je 0.5124, F-štatistika = 101.2 a p-value < $2.2 \cdot 10^{-16}$.

Podarilo sa nám zvýšiť upravený koeficient determinácie, čo je znakom toho, že tento model je o niečo vhodnejší. Takisto koeficient determinácie sa zvýšil a presiahol hodnotu 0.5, čo znamená, že dáta sú takto popísané o niečo lepšie. P-value regresie je opäť menšia ako $2.2 \cdot 10^{-16}$, z čoho vyplýva signifikantnosť regresie.

Všetky koeficienty sú takisto signifikantné. Hodnoty α a γ sú kladné, čo znamená, že tak ako hra tímu A v predchádzajúcej minúte, tak aj hra súpera v predchádzajúcej minúte pozitívne pôsobí na kvalitu hry tímu A. Na druhej strane β vyšla záporná, teda čím lepšie tím B v aktuálnej minúte hrá, tým horšie hrá tím A.

Po dosadení odhadnutých koeficientov model vyzerá takto:

$$A_{(t)} = 0.59132 \cdot A_{(t-1)} - 0.20183 \cdot B_{(t)} + 0.44460 \cdot B_{(t-1)}.$$

Môžeme to teda zhrnúť tak, že zotrvačnosť výkonu počas zápasu je približne 59% z predošlého výkonu, predchádzajúci výkon súpera pôsobí na tím skoro 45-timi percentami bodov súpera a aktuálny výkon súpera 20-timi percentami z bodov súpera uberá na kvalite hry tímu.

3.3 Model s rozličným sklonom $B_{(t)}$ a $B_{(t-1)}$

Na základe práce o manželstve, kde vplyv manžela na partnerku keď mal manžel záporne reakcie bol iný ako keď bol manžel pozitívne naladený, sme sa snažili testovať podobné reakcie aj v našom prípade. Pokúsili sme sa odhadnúť, aký vplyv na výkon tímu má to, že

súper hrá danú minútu dobre (kladné hodnoty bodov) alebo zle (záporné hodnoty bodov), a takisto aký vplyv má dobrý/zlý výkon súpera v predchádzajúcej minúte. Nový model sme tak upravili na tvar:

$$A_{(t)} = \alpha A_{(t-1)} + \beta^+ * \max[B_{(t)}, 0] + \beta^- * \min[B_{(t)}, 0] + \gamma^+ * \max[B_{(t-1)}, 0] + \gamma^- * \min[B_{(t-1)}, 0],$$

kde α je zotrvačnosť výkonu, β^+ / β^- je reakcia na aktuálnu dobrú/zlú hru súpera a γ^+ / γ^- je reakcia na dobrú/zlú hru súpera v predchádzajúcej minúte. Výsledky modelu sa nachádzajú v tabuľke 3.

Tabuľka 3

	Estimate	Std. Error	t value	Pr(> t)
α	0.39165	0.05671	6.906	3.33e-11 ***
β^+	-0.14917	0.07108	-2.099	0.036750 *
β^-	-0.68627	0.18323	-3.745	0.000219 ***
γ^+	0.52841	0.06716	7.868	7.83e-14 ***
γ^-	-0.60000	0.19166	-3.131	0.001929 **

Koeficient determinácie tejto regresie je 0.5813, upravený koeficient determinácie je 0.5739, F-štatistika = 78.03 a p-value < $2.2 \cdot 10^{-16}$.

Celá regresia vyšla signifikantná a opäť sa nám podarilo zvýšiť upravený koeficient determinácie, z čoho môžeme usúdiť, že tento model je o niečo lepší ako predchádzajúci. Aj koeficient determinácie je vyšší a už sa blíži k hodnote 0.6, teda model už pomerne dobre popisuje naše dáta.

Všetky koeficienty sú na hladine významnosti 5% signifikantné, teda všetky parametre sú pre nás významné. Koeficient $\alpha = 0,39165$ znamená, že tím si zachováva približne 39% svojej predošlej výkonnosti. $\beta^+ = -0.14917$ nám hovorí, že aktuálny dobrý výkon súpera 15-timi percentami uberá na kvalite hry tímu A. Naopak $\beta^- = -0,68627$ znamená, že ak súper podáva v aktuálnej minúte zlý výkon, tímu A to pridáva na výkone takmer 69-imi percentami z výkonu súpera. Kladná hodnota γ^+ a záporná hodnota γ^- zase

vyjadrujú, že predchádzajúci výkon súpera tím A motivuje k lepšej hre, či už hral súper dobre, alebo zle. Rozdiel je ale v tom, že zlá hra pôsobí mierou 60% a dobrá hra mierou takmer 53% [9].

Z týchto výsledkov môžeme náš model upraviť do podoby

$$A_{(t)} = 0.39135 * A_{(t-1)} - 0.14917 * \max[B_{(t)}, 0] - 0.68627 * \min[B_{(t)}, 0] + \\ + 0.52841 * \max[B_{(t-1)}, 0] - 0.60000 * \min[B_{(t-1)}, 0].$$

3.4 Model s rozličným sklonom $B_{(t-1)}$

V tomto modeli sme testovali, či tím A nereaguje rozdielne iba na predchádzajúcu dobrú, respektíve zlú hru súpera. Všetky ostatné predpoklady ostali rovnaké ako v predchádzajúcom modeli. Tento model je teda určitým skĺbením predchádzajúcich dvoch:

$$A_{(t)} = \alpha A_{(t-1)} + \beta B_{(t)} + \gamma^+ * \max[B_{(t-1)}, 0] + \gamma^- * \min[B_{(t-1)}, 0].$$

Odhady koeficientov a výsledky regresie sa nachádzajú v tabuľke 4.

Tabuľka 4

	Estimate	Std. Error	t value	Pr(> t)
α	0.44124	0.05343	8.258	5.81e-15 ***
β	-0.25869	0.05571	-4.643	5.26e-06 ***
γ^+	0.60980	0.05886	10.360	< 2e-16 ***
γ^-	-0.66443	0.19152	-3.469	0.000603 ***

Koeficient determinácie tejto regresie je 0.5724, upravený koeficient determinácie je 0.5663, F-štatistika = 94.38 a p-value < $2.2 * 10^{-16}$.

Regresia je signifikantná, no upravený koeficient mierne klesol oproti predchádzajúcemu modelu, čo môže znamenať, že predchádzajúci model bol mierne lepší.

Koeficient determinácie je však stále nad hodnotou 0.5 a teda dáta sú takýmto spôsobom popísane celkom dobre.

Všetky koeficienty sú signifikantné s tým, že hodnoty α a γ^+ sú kladné a β^+ , γ^- sú záporné, čo potvrdilo naše predchádzajúce výsledky. Zmenila sa iba miera ich účinku. Zotrvačnosť výkonu je teraz 44%, aktuálny výkon súpera pôsobí na tím A mierou takmer 26%, pričom čím hrá súper lepšie, tým horší je výkon tímu A. Kladná γ^+ a záporná γ^- presne ako v predchádzajúcom modeli pôsobia pozitívne na výkon tímu A. To znamená, že ak hral tím B predchádzajúcu minútu dobre (kladné hodnotenie), hodnotenie tímu A bude lepšie o 61% ich bodov. Ak hral tím B predchádzajúcu minútu zle (záporné hodnotenie), hodnotenie tímu A bude lepšie o približne 66% bodov tímu B.

Po dosadení odhadnutých koeficientov vyzerá model nasledovne:

$$A_{(t)} = 0.44124 * A_{(t-1)} - 0.25869 * B_{(t)} + 0.60980 * \max[B_{(t-1)}, 0] - 0.66443 * \min[B_{(t-1)}, 0].$$

3.5 Model s rozličným sklonom $B_{(t)}$

Tento model je analógiou toho predchádzajúceho, no teraz sme však skúmali, iba rozdielnu reakciu tímu A na aktuálnu dobrú, respektíve zlú hru súpera. Model je teda veľmi podobný:

$$A_{(t)} = \alpha A_{(t-1)} + \beta^+ * \max[B_{(t)}, 0] + \beta^- * \min[B_{(t)}, 0] + \gamma B_{(t-1)},$$

kde α je zotrvačnosť výkonu, β^+ / β^- je reakcia na aktuálnu dobrú/zlú hru súpera a γ predstavuje reakcia na hru súpera v predchádzajúcej minúte. Výsledky modelu sa nachádzajú v tabuľke 5.

Tabuľka 5

	Estimate	Std. Error	t value	Pr(> t)
α	0.48602	0.05609	8.665	3.59e-16 ***
β^+	-0.03937	0.07087	-0.556	0.579
β^-	-0.88639	0.18710	-4.738	3.43e-06 ***
γ	0.34605	0.05971	5.796	1.82e-08 ***

Koeficient determinácie tejto regresie je 0.5415, upravený koeficient determinácie je 0.535, F-štatistika = 83.26 a p-value < $2.2 \cdot 10^{-16}$.

Celá regresia je signifikantná, o čom svedčí p-value, upravený koeficient determinácie je oproti modelom v kapitolách 5.3 a 5.4 nižší, čo znamená, že predchádzajúce modely boli mierne lepšie. Aj koeficient determinácie klesol, no stále je nad hodnotu 0.5.

Signifikancia koeficientov v tomto prípade nevyšla veľmi dobre. Keďže hodnota $\beta^+ = -0.03937$ je pomerne blízko nuly v našej regresii sa javí ako nevýznamná, čo potvrdzuje aj jej p-value = 0.579. V tomto prípade dobrý výkon tímu B ovplyvní negatívne výkon tímu A, avšak len necelými 4%, čo je takmer zanedbateľné. Ostatné koeficienty vyšli signifikantné a ich hodnoty prezrádzajú takmer 49% - nú zotrvačnosť výkonu, 89% - ný účinok zlej hry tímu B v aktuálnej minúte na zlepšení hry tímu A a 35% - ný vplyv predchádzajúcej hry tímu B na výkon tímu A.

Takto odhadnutý model teda vyzerá nasledovne:

$$A_{(t)} = 0.48602 \cdot A_{(t-1)} - 0.03937 \cdot \max[B_{(t)}, 0] - 0.88639 \cdot \min[B_{(t)}, 0] + 0.34605 \cdot B_{(t-1)}.$$

Ako najlepší sa javil model opísaný v kapitole 3.3, preto sme v ďalšej časti pracovali s ním. Použili sme ho na vyhodnotenie správania jednotlivých tímov počas zápasu. Aby sme však dokázali regresie dobre vyhodnotiť, zamerali sme sa iba na tímy, pre ktoré sme mali dostatočný počet dát. Najviac boli v našich dátach zastúpené tímy FC Barcelona (5 zápasov), Manchester United (3 zápasy) a Tottenham Hotspur (3 zápasy).

3.6 FC Barcelona

Použitím modelu z časti 3.3 výlučne na zápasy FC Barcelona, kde $A_{(t)}$ sú body Barcelony v čase t a $B_{(t)}$ sú body jej súperov v čase t , sme dostali nasledujúce výsledky (tabuľka 6):

Tabuľka 6

	Estimate	Std. Error	t value	Pr(> t)
α	0.45093	0.13205	3.415	0.00112 **
β^+	-0.08717	0.14753	-0.591	0.55671
β^-	-0.63595	0.28394	-2.240	0.02865 *
γ^+	0.57948	0.11474	5.050	4.03e-06 ***
γ^-	-0.54941	0.29230	-1.880	0.06479 .

Koeficient determinácie tejto regresie je 0.7835, upravený koeficient determinácie je 0.7663, F-štatistika = 45.6 a p-value < $2.2 \cdot 10^{-16}$.

Regresia vyšla signifikantná, s pomerne vysokým koeficientom determinácie a upraveným koeficientom determinácie, čo značí, že tento model dobre popisuje dáta získané zo zápasov FC Barcelona.

So signifikanciou jednotlivých koeficientov to už je o niečo horšie. Koeficient β^+ vyšiel nesignifikantný, s pomerne vysokou p-value. Koeficient γ^- je signifikantný na hladine významnosti 10%. Ostatné koeficienty sú signifikantné na hladine významnosti 5%.

Odhadnuté hodnoty koeficientov sú podobné hodnotám z tabuľky 3. Koeficient zotrvačnosti $\alpha = 0.45093$ hovorí, že Barcelona si udržiava 45% vlastného výkonu. Ak hrá súper v aktuálnej minúte dobre (kladné hodnotenia) Barcelone to znižuje výkon, no len necelými 9-timi percentami z výkonu ich súpera. Naopak, ak ich súper hrá v prebiehajúcej minúte zle (záporne hodnotenia), Barcelona sa to snaží využiť a povzbudzuje ich to k lepšiemu výkonu, takáto situácia pridáva ich hodnoteniu takmer 64% zo súperovho predošlého hodnotenia. Takisto výkon súpera z predchádzajúcej minúty im pridáva na výkone, či už hral súper dobre alebo zle. Ak hral súper minútu predtým dobre, prírastok na výkone Barcelony je 58% - ný, ak hral zle prírastok je 55% - ný zo súperovho výkonu.

Z tohto vyplýva, že Barcelona je mužstvo, ktorého výkon je vyvážené ovplyvnený vlastnou výkonnosťou, tým ako hral súper predchádzajúce minúty a takisto aj aktuálnou situáciou na ihrisku.

3.7 Manchester United

Rovnakým spôsobom sme vyhodnocovali zápasy Manchesteru United, pričom v tomto prípade sú $A_{(t)}$ body Manchesteru v čase t . Výsledky z regresie sú v tabuľke 7.

Tabuľka 7

	Estimate	Std. Error	t value	Pr(> t)
α	0.2165	0.1995	1.085	0.2883
β^+	0.1346	0.2107	0.639	0.5289
β^-	-0.8307	0.4997	-1.663	0.1089
γ^+	0.3850	0.2022	1.904	0.0685 .
γ^-	-0.2482	0.6006	-0.413	0.6830

Koeficient determinácie tejto regresie je 0.5602, upravený koeficient determinácie je 0.4722, F-štatistika = 6.369 a p-value = 0.0006084.

Táto regresia vyšla signifikantná, no nie tak výrazne ako predtým. Upravený koeficient determinácie je o dosť nižší ako v predchádzajúcich modeloch, takisto koeficient determinácie oproti zápasom Barcelony je dosť nižší, no stále nad hodnotou 0.5.

Koeficient γ^+ je signifikantný na hladine významnosti 10%, ostatné koeficienty nie sú signifikantné, čo môže byť spôsobené menším množstvom dát. Zotrvačnosť výkonu Manchesteru je značne nižšia ako u Barcelony a to necelých 22% čo znamená, že na ich výkon viac vplýva výkon súpera. Nízkou zotrvačnosťou vlastného výkonu kompenzujú pozitívnu reakciou na dobrý výkon súpera v prebiehajúcej minúte mierou 13% a prudko stúpa ich výkon, ak súper hrá zle (využívajú vo svoj prospech 83% jeho výkonu). Na

predchádzajúci výkon súpera reagujú podobne ako tomu bolo vo všeobecnom modeli (3.3), avšak nižšou mierou. Kvalitná súperova hra im pridáva 38.5% a nekvalitná hra im pridá na výkone 25%.

Výkon Manchesteru je menej ovplyvnený vlastným výkonom a výkonom súpera v predchádzajúcich minútach a väčší dôraz kladie na aktuálnu hru súpera.

3.8 Tottenham Hotspur

Podobný postup sme opakovali aj pre Tottenham. V tomto prípade $A_{(t)}$ predstavuje body Tottenhamu v čase t . Výsledky tejto regresie sú v tabuľke 8.

Tabuľka 8

	Estimate	Std. Error	t value	Pr(> t)
α	0.26827	0.15018	1.786	0.08163 .
β^+	-0.05133	0.17674	-0.290	0.77301
β^-	-0.32898	0.47364	-0.695	0.49133
γ^+	0.53911	0.18183	2.965	0.00508 **
γ^-	-1.01664	0.43949	-2.313	0.02593 *

Koeficient determinácie tejto regresie je 0.5652, upravený koeficient determinácie je 0.5108, F-štatistika = 10.4 a p-value = $1.930 \cdot 10^{-6}$.

Aj táto regresia je signifikantná s upraveným koeficientom determinácie mierne nad hodnotu 0.5, čo je o niečo lepšie ako pri Manchesteri. Koeficient determinácie je takisto mierne vyšší, tesne pod 0.6.

Koeficienty β^+ , β^- sú nesignifikantné, α je signifikantné na hladine významnosti 10% a koeficienty γ^+ , γ^- sú signifikantné na hladine významnosti 5%. Horšie výsledky signifikantnosti parametrov môžu byť opäť spôsobené menším počtom dát (3 zápasy, rovnako ako Manchester).

Tottenham má zotrvačnosť výkonu necelých 27%. Momentálny výkon súpera naň pôsobí dvojako. Ak hrá súper dobre, na hru Tottenhamu to pôsobí negatívne (uberá im 5% z výkonu súpera). Naopak, ak hrá súper zle, snažia sa to zužitkovať vo svoj prospech a ich výkon sa zlepší o 33% z výkonu súpera. Predchádzajúci výkon súpera pôsobí na Tottenham pozitívne. Ak súper hral dobre, zlepšujú sa s jeho výkonom mierou 54%. Ak súper hral zle, takisto to zlepšuje ich výkon a to viac ako o 101%.

Tottenham je mužstvo, ktoré je takisto ako Manchester menej ovplyvnené svojím predchádzajúcim výkonom. Odlišne však reaguje na hru súpera. Kým Manchester kládol väčší dôraz na aktuálny výkon súpera, Tottenham viac reaguje na súperovu predchádzajúcu hru, pričom jeho nekvalitnú hru dokáže využiť vo svoj prospech mierou 101%.

4. Vyhodnotenie výsledkov

V predchádzajúcej kapitole sme otestovali niekoľko modelov. Sústredili sme sa pri tom na schopnosť tímu udržať si svoju výkonnosť a na interakciu so súperom. Pri testovaní nám výsledky niektorých regresíí vyšli uspokojivo, pri iných to zas bolo horšie. V tejto časti sme sa zamerali na tie najlepšie modely a pokúsime sa ich previesť do praxe.

Použitím všetkých dát sme zostavili niekoľko modelov, ktorými sme sa snažili modelovať futbalový zápas. Snahou bolo opísať akýkoľvek zápas, vytvoriť všeobecný model, preto sme použili dáta zo zápasov rôznych tímov, súťaží, či úrovní. Model, ktorý najlepšie popisoval dáta je bližšie opísaný v kapitole 3.3 a výsledky regresie sa nachádzajú v tabuľke 3. Zostavený model zápasu, ktorý by mal do určitej miery popisovať ľubovoľný zápas, respektíve jeho prvý polčas (predpokladáme, že tréneri takisto nejakým spôsobom priebežne vyhodnocujú hru svojho tímu a počas polčasovej prestávky majú možnosť urobiť množstvo zmien, ktoré sledovaním prvých 15 minút zápasu nemáme možnosť odhadnúť) vyzerá nasledovne:

$$A_{(t)} = 0.39135 * A_{(t-1)} - 0.14917 * \max[B_{(t)}, 0] - 0.68627 * \min[B_{(t)}, 0] + \\ + 0.52841 * \max[B_{(t-1)}, 0] - 0.60000 * \min[B_{(t-1)}, 0].$$

Použitím tohto modelu by sa teda na základe niekoľkých úvodných minút zápasu dala odhadovať výkonnosť jednotlivých tímov v danom zápase. Toto dáva do rúk trénerov a manažérov silný nástroj, na základe ktorého potom môžu lepšie odhadnúť, akú zmenu je potrebné urobiť, aby to zlepšilo výkon tímu, alebo na ktoré situácie je potrebné sa zamerať počas tréningu.

Otestovaním tohto modelu na novo získaných dátach z dodatočne vybraných zápasov, pričom sme aktuálne hodnoty získaných bodov počítali pomocou vyššie uvedenej rovnice s odhadnutými koeficientmi z časti 3.3 sme dostali priemernú odchýlku od skutočnej hodnoty na úrovni 6.6. Táto odchýlka je pomerne vysoká, no nie je prekvapujúca vzhľadom na množstvo dát a aj preto, že sme popisovali správanie všetkých tímov naraz, pričom sme predpokladali, že sa správajú rovnako. Nasvedčuje to o určitých nedostatkoch, ktoré by sa mohli odstrániť dopĺňaním ďalších zápasov s ešte väčšou rozmanitosťou.

Ďalší spôsob ako zlepšiť relevantnosť výsledkov môže byť, nie tvorba všeobecného modelu, ale zameranie sa na jednotlivé tímy a sledovať ktorý tím akým spôsobom reaguje

na svojho súpera. Každý tím je jedinečný, má svoju mentalitu, zaužívaný štýl hry, trénera a mnohé iné faktory, ktoré ovplyvňujú napríklad ich vlastnú zotrvačnosť výkonu (ako rýchlo dokážu prekonať zlý vstup do zápasu, či ako dobre si budú udržiavať kvalitnú hru), alebo reakciu na hru súpera. Preto je lepšie odhadovať hodnoty koeficientov modelu z časti 3.3 pre jednotlivé tímy zvlášť ako sme to robili pre tímy FC Barcelona, Manchester United a Tottenham Hotspur v kapitolách 3.6 – 3.8.

Najlepšie výsledky regresie vyšli pre Barcelonu (kapitola 3.6, tabuľka 6), kde koeficient determinácie R^2 dosiahol hodnotu 0.7835. Dôvod, prečo sme u ostatných dvoch tímov nedosiahli podobne vysoký koeficient determinácie, alebo vyššiu signifikanciu parametrov môže byť menšie množstvo dát (5 zápasov Barcelony oproti 3 Manchesteru a 3 Tottenhamu). Výsledný model pre Barcelonu vyzerá takto:

$$A_{(t)} = 0.45093 * A_{(t-1)} - 0.08717 * \max[B_{(t)}, 0] - 0.63595 * \min[B_{(t)}, 0] + \\ + 0.57948 * \max[B_{(t-1)}, 0] - 0.54941 * \min[B_{(t-1)}, 0].$$

Takýmto spôsobom môžeme výkon Barcelony vyhodnocovať lepšie, než všeobecným modelom. To sa potvrdilo otestovaním dát pre ich zápasy použitím vyššie uvedenej rovnice. Priemerná odchýlka od skutočných hodnôt bola 4.17 bodov. Ani tento model teda nedáva presne hodnoty, no odchýlku oproti všeobecnému modelu sa nám podarilo znížiť o viac ako 2 body. Sledovaním väčšieho počtu zápasov by sme mohli koeficienty odhadnúť presnejšie a tým aj znížiť odchýlku od skutočných hodnôt.

Kedy sme takisto odhadli koeficienty daných parametrov pre ich súpera, dokázali by sme s určitou presnosťou zachytiť priebeh zápasu na základe jeho úvodných minút. Už po prvých pár minútach by teda tréneri oboch tímov vedeli či sa bude zápas vyvíjať v ich prospech, alebo nie a na základe toho by vedeli prispôbiť stratégiu k dosiahnutiu želaného výsledku.

5. Riešenia systémov

Použitím základného modelu z kapitoly 3.2 sme demonštrovali, ako sa dá zisťovať pomocou odhadnutých koeficientov priebeh zápasu a ku ktorému bodu bude konvergovať hodnotenie družstiev.

Za predpokladu, že odhadnuté koeficienty sú rovnaké pre oba tímy, sme zostavili rovnice:

$$A_{(t)} = \alpha A_{(t-1)} + \beta B_{(t)} + \gamma B_{(t-1)}$$

$$B_{(t)} = \alpha B_{(t-1)} + \beta A_{(t)} + \gamma A_{(t-1)}.$$

Po upravení na maticový tvar dostaneme:

$$u_t = \begin{pmatrix} \alpha & \gamma \\ \gamma & \alpha \end{pmatrix} * u_{t-1} + \begin{pmatrix} 0 & \beta \\ \beta & 0 \end{pmatrix} * u_t$$

kde $u_t = \begin{pmatrix} A_t \\ B_t \end{pmatrix}$ je hodnotenie tímu A, B v čase t. Po ďalšej úprave dostaneme tvar:

$$\begin{aligned} u_t &= \frac{1}{1-\beta^2} \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix} \begin{pmatrix} \alpha & \gamma \\ \gamma & \alpha \end{pmatrix} * u_{t-1} = \\ &= \frac{1}{1-\beta^2} \begin{pmatrix} \alpha + \beta\gamma & \gamma + \alpha\beta \\ \gamma + \alpha\beta & \alpha + \beta\gamma \end{pmatrix} * u_{t-1} \end{aligned}$$

Riešením pevného bodu, kde $A^* = A_t = A_{t-1}$ a $B^* = B_t = B_{t-1}$ sme dostali bod (0,0), respektíve bod (A, A), ak je splnená podmienka $\gamma^2 + \alpha^2 + 2 * \frac{\alpha + \beta\gamma}{(1-\beta)^2} = 1$. Riešením vlastných hodnôt matice zistíme stabilitu pevných bodov. Vlastné hodnoty danej matice sú:

$$\lambda_{1,2} = \frac{\alpha \pm \gamma}{1 \mp \beta}.$$

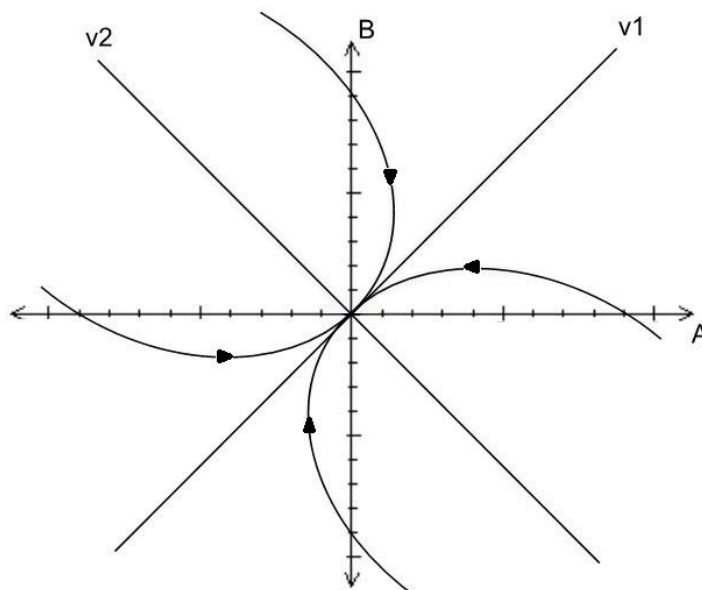
Ak sú vlastné hodnoty matice $|\lambda_i| < 1$, podmienka stability je splnená a body sú stabilné, ak sú vlastné hodnoty matice $|\lambda_i| > 1$, pevné body sú nestabilné.

Pre konkrétne hodnoty koeficientov z tabuľky 2, $\alpha = 0.59$, $\beta = -0.20$ a $\gamma = 0.44$, máme len jeden pevný bod a to bod (0,0), pričom vlastné hodnoty matice sú:

$$\lambda_1 = 0.858 \text{ a } \lambda_2 = 0.1875.$$

Obe vlastné čísla sú kladné a menšie ako 1, teda bod (0,0) je stabilný. Pre tieto vlastné hodnoty dostávame vlastné vektory v_1, v_2 :

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$



Obr. 9: stabilita pevného bodu (0,0)

Obr. 9: stabilita pevného bodu (0,0), kde na osi A sú hodnoty bodov tímu A a na osi B sú hodnoty bodov tímu B a v_1 , v_2 sú vlastné vektory vlastných hodnôt $\lambda_{1,2}$ opisuje, že nech je počiatočná hodnota bodov akákoľvek, postupom času sa dostaneme do pevného bodu (0,0).

Hodnoty bodov sa rýchlo ustália na hodnote (A, A), čiže dva rovnaké tímy budú hrať vyrovnané a postupne skonvergujú do bodu (0,0) z akejkoľvek štartovnej pozície. Znamená to, že ak bude zápas trvať príliš dlho, zápas stratí kvalitu, čo je vysvetliteľné tým, že výkon hráčov je limitovaný ich výdržou a každý tím má obmedzený počet striedaní.

Predchádzajúce výpočty boli ilustratívne a poskytujú spôsob, ako zisťovať priebeh zápasu pre konkrétne družstvá. Nasledujúce výpočty boli realizované na simulovaný zápas FC Barcelona – Tottenham Hotspur, pretože pre tieto tímy sme odhadovali ich koeficienty zotrvačnosti výkonu a reakcie na hru súperu, pričom ich regresie vyšli najviac signifikantné.

Pre oba tímy sme teda použili model z kapitoly 3.3 a platili nasledujúce rovnice:

$$A_{(t)} = \alpha_1 A_{(t-1)} + \beta_1^+ * \max[B_{(t)}, 0] + \beta_1^- * \min[B_{(t)}, 0] + \gamma_1^+ * \max[B_{(t-1)}, 0] + \gamma_1^- * \min[B_{(t-1)}, 0]$$

$$B_{(t)} = \alpha_2 B_{(t-1)} + \beta_2^+ * \max[A_{(t)}, 0] + \beta_2^- * \min[A_{(t)}, 0] + \gamma_2^+ * \max[A_{(t-1)}, 0] + \gamma_2^- * \min[A_{(t-1)}, 0]$$

kde A sú body Barcelony, B sú body Tottenhamu, pričom každý tím má rozdielne hodnoty koeficientov. Predchádzajúce rovnice sme upravili na tvar:

$$\begin{aligned}A_{(t)} &= \alpha_1 A_{(t-1)} + f_1(B_{(t)}) + g_1(B_{(t-1)}) \\B_{(t)} &= \alpha_2 B_{(t-1)} + f_2(A_{(t)}) + g_2(A_{(t-1)})\end{aligned}$$

Za podmienky pre riešenie pevných bodov $A^* = A_t = A_{t-1}$ a $B^* = B_t = B_{t-1}$ sme dostali

$$\begin{aligned}(1 - \alpha_1)A^* &= f_1(B^*) + g_1(B^*) \\(1 - \alpha_2)B^* &= f_2(A^*) + g_2(A^*),\end{aligned}$$

pričom $f_i(B^*) = \beta_i^+ B^* \{ \beta_i^- B^* \}$ pre $B^* > 0 \{ < 0 \}$ a $g_i(B^*) = \gamma_i^+ B^* \{ \gamma_i^- B^* \}$ pre $B^* > 0 \{ < 0 \}$. Dostali sme tak 4 prípady a ilustratívne budeme riešiť prípad kedy $A^*, B^* > 0$, pre ostatné prípady sú uvádzané iba riešenia.

1. $A^*, B^* > 0$:

Aplikovaní predchádzajúcich predpokladov sme dostali systém pre riešenie pevných bodov:

$$\begin{aligned}(1 - \alpha_1)A^* &= (\beta_1^+ + \gamma_1^+)B^* \\(1 - \alpha_2)B^* &= (\beta_2^+ + \gamma_2^+)A^*.\end{aligned}$$

Ak je splnená podmienka $\frac{\beta_1^+ + \gamma_1^+}{1 - \alpha_1} = \frac{1 - \alpha_2}{\beta_2^+ + \gamma_2^+}$, systém má nekonečne veľa riešení

$(\frac{\beta_1^+ + \gamma_1^+}{1 - \alpha_1} B^*, B^*)$, ak táto podmienka nie je splnená, systém má jediný pevný bod $(0,0)$.

Pre riešenie stability sme si systém diferenčných rovní prepísali na maticový tvar:

$$u_t = \begin{pmatrix} \alpha_1 & \gamma_1^+ \\ \gamma_2^+ & \alpha_2 \end{pmatrix} * u_{t-1} + \begin{pmatrix} 0 & \beta_1^+ \\ \beta_2^+ & 0 \end{pmatrix} * u_t,$$

kde $u_t = \begin{pmatrix} A_t \\ B_t \end{pmatrix}$ je hodnotenie tímov Barcelona a Tottenham v čase t . Po ďalšej úprave dostaneme tvar:

$$\begin{aligned}u_t &= \frac{1}{1 - \beta_1^+ \beta_2^+} \begin{pmatrix} 1 & \beta_1^+ \\ \beta_2^+ & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 & \gamma_1^+ \\ \gamma_2^+ & \alpha_2 \end{pmatrix} * u_{t-1} = \\&= \frac{1}{1 - \beta_1^+ \beta_2^+} \begin{pmatrix} \alpha_1 + \beta_1^+ \gamma_2^+ & \gamma_1^+ + \alpha_2 \beta_1^+ \\ \gamma_2^+ + \alpha_1 \beta_2^+ & \alpha_2 + \beta_2^+ \gamma_1^+ \end{pmatrix} * u_{t-1}\end{aligned}$$

Aplikovaním hodnôt z tabuľky 6 a tabuľky 8, teraz budeme systém riešiť pre konkrétne hodnoty koeficientov, z regresíí pre Barcelonu a Tottenham. Jediným pevným bodom je bod (0,0), ktorého stabilitu budeme teraz zisťovať riešením vlastných hodnôt matice.

Pre prvý kvartál sme dosiahli nasledujúce vlastné hodnoty a vlastné vektory:

$$\lambda_1 = 0.87 \text{ a } \lambda_2 = -0.22$$

$$v_1 = \begin{bmatrix} 0.77 \\ 0.64 \end{bmatrix}, v_2 = \begin{bmatrix} 0.67 \\ 0.75 \end{bmatrix}.$$

V tomto prípade je teda bod (0,0) stabilný.

2. $A^* > 0, B^* < 0$:

$$\lambda_1 = 0.21 + 0.62i \text{ a } \lambda_2 = 0.21 - 0.62i$$

$$v_1 = \begin{bmatrix} 0.76 \\ -0.1 - 0.64i \end{bmatrix}, v_2 = \begin{bmatrix} 0.76 \\ -0.1 + 0.64i \end{bmatrix}.$$

Aj v tomto prípade je bod (0,0) stabilný.

3. $A^* < 0, B^* > 0$:

$$\lambda_1 = 0.32 + 0.8i \text{ a } \lambda_2 = 0.32 - 0.8i$$

$$v_1 = \begin{bmatrix} -0.16 - 0.54i \\ 0.82 \end{bmatrix}, v_2 = \begin{bmatrix} -0.16 + 0.54i \\ 0.82 \end{bmatrix}$$

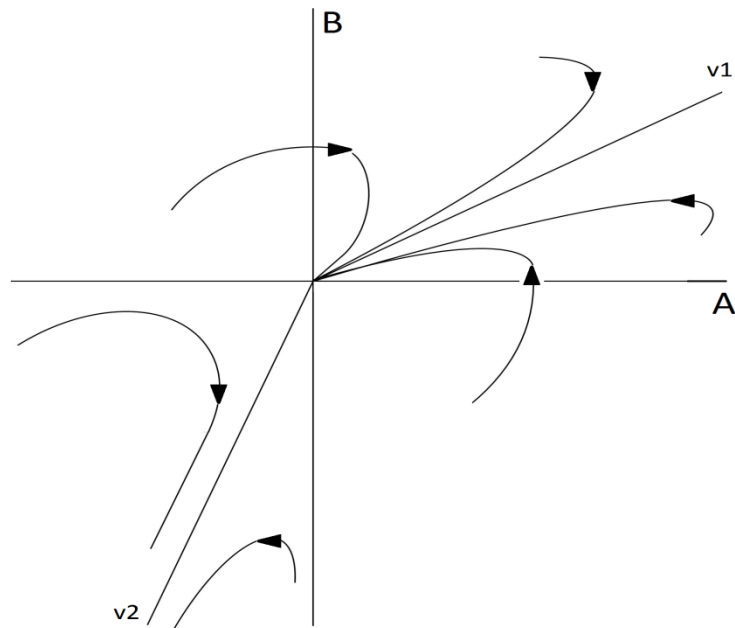
Aj v tomto prípade je teda bod (0,0) stabilný.

4. $A^*, B^* < 0$:

$$\lambda_1 = 2.22 \text{ a } \lambda_2 = -0.25.$$

$$v_1 = \begin{bmatrix} 0.74 \\ -0.67 \end{bmatrix}, v_2 = \begin{bmatrix} 0.49 \\ 0.87 \end{bmatrix}.$$

V tomto kvartáli je bod (0,0) nestabilný.



Obr. 10: stabilita pevného bodu $(0,0)$ pre FCB-TOT

Na Obr. 10: stabilita pevného bodu $(0,0)$ pre FCB-TOT, sú na osi A hodnoty bodov FC Barcelona a na osi B sú body tímu Tottenham, v_1 je prvý vlastný vektor pre prvý prípad a v_2 je druhý vlastný vektor pre štvrtý prípad, pričom malé záporné vlastné hodnoty spôsobia, že niektoré trajektórie budú skákať okolo v_1 a pomaly sa k nemu budú približovať. Z obrázku je pozorovateľné, že ak začnú oba tímy s kladnými hodnotami, tak prevahu bude mať Barcelona a pomaly výkon oboch tímov skonverguje k bodu $(0,0)$. Ak tímy odštartujú z pozície pre druhý prípad, t.j. Barcelona začne kvalitne a Tottenham zle, prevahu si bude udržiavať Barcelona, pričom Tottenham sa bude pomaly zlepšovať a keď sa dostane na kvalitnú hru, výkony oboch začnú pomaly konvergovať k pevnému bodu. V treťom prípade, t.j. Barcelona začne zle a Tottenham dobre, si prevahu bude udržiavať Tottenham, pričom oba tímy sa budú postupne zlepšovať. Keď Barcelona preskočí zo záporných hodnôt, výkon Tottenhamu začne klesať a hra sa vyrovná, až nakoniec skonverguje k $(0,0)$. Vo štvrtom prípade majú obidva tímy nekvalitný vstup do zápasu a ich hra sa môže mierne zlepšovať, no nedosiahne kladné hodnoty a následne sa výkon tímov začne rýchlejšie zhoršovať. Pri diskretnom dynamickom systéme sa však môže stať, že bod preskočí niektorú z osí a zmení sa priebeh zápasu. Spôsobiť to môže zmena herného štýlu, či striedanie alebo iný impulz, ktoré dokážu tímy vytrhnúť zo zlej vzájomnej hry. Ak nepríde externá zmena, tímy budú pokračovať v zhoršujúcom sa výkone.

Záver

Cieľom tejto práce bolo vytvoriť model futbalového zápasu pomocou ktorého by sme dokázali predpovedať jeho priebeh. Veľkou snahou bolo pri tom zachytiť interakciu medzi tímami a taktiež schopnosť udržiavať si vlastný výkon. Obe tieto veci sme pri testovaní modelov zachytili a potvrdilo sa, že počas zápasu vstupujú do výkonu oboch tímov vo veľkej miere.

Nosnou časťou odhaľovania dynamiky futbalového zápasu boli dáta. Tie zachytávali podstatné prvky opisujúce výkon tímov tak, aby zachytávali najdôležitejšie okamihy zápasu a vývoj celkovej hry. Ich následným vážením sme im dodali relevantnosť a komparatívnosť, na základe čoho sme mohli s nimi spoľahlivo pracovať.

Na základe určitých predpokladov sme následne vytvorili základný model, ktorý zahŕňal zotrvačnosť vlastného výkonu a reakciu na kvalitu súperovej hry. Následným testovaním a vylepšovaním predpokladov sme sa dopracovali k univerzálnemu modelu, ktorý podrobne zachytával udržateľnosť výkonu a reakcie na rôzne formy súperovej hry. Tento model, však mal vyššiu chybovosť, keďže nezachytával rôzne správanie jednotlivých tímov.

Aplikovaním získaného modelu na jednotlivé tímy sme chybovosť modelu znížili a tým sa aj zlepšila schopnosť predpovedať priebeh hry. Pre jednotlivé tímy sme odhadli ich vlastnú zotrvačnosť výkonu a takisto aj mieru, akou reagujú na svojho súpera. Odhalili sme tri rôzne formy správania sa tímov:

1. Vyrovnaná reakcia na vlastný výkon a výkon súpera.
2. Menší vplyv vlastného výkonu a výkonu súpera v predchádzajúcich minútach, no väčší dôraz na momentálnu hru.
3. Menší vplyv vlastného výkonu a aktuálneho výkonu súpera, no veľký vplyv predchádzajúceho priebehu.

Zoznam bibliografických odkazov

[1] T. REILLY, Science and Soccer, Liverpool : Spon Press 1996, p. 392, ISBN 0 419 18880 0

[2] Z. YUE, H. BROICH, F. SEIFRIZ, and J. MESTER, Mathematical Analysis of a Soccer Game, Studies in Applied Mathematics 121:223-243 (2008)

[3] +/- štatistika NBA

Dostupné na internete: <http://www.nba.com/statistics/plusminus/plusminus.jsp>, 28.5.2011

[4] J. D. MURRAY, Mathematical Biology, Berlin : Springer, 2003, p. 736, ISBN 9780387952284

[5] B.LETKO, Futbalový Zápas ako Dynamický Systém I., Bratislava, FMFI UK, 2011

[6] Regresná analýza

Dostupné na internete: http://en.wikipedia.org/wiki/Regression_analysis, 25.5.2011

[7] Koeficient determinácie

Dostupné na internete: http://en.wikipedia.org/wiki/Coefficient_of_determination, 25.5.2011

[8] P-value

Dostupné na internete: <http://en.wikipedia.org/wiki/P-value>, 25.5.2011

[9] Hladina významnosti

Dostupné na internete: http://en.wikipedia.org/wiki/Statistical_significance, 25.5.2011