Submitted for publication, 2012
draft version 2012.11.28

Determining the level of ability of football teams by dynamic ratings based on the relative discrepancies in scores between adversaries

ANTHONY C. CONSTANTINOU*, NORMAN E. FENTON

Risk and Information Management (RIM) Research Group,
Department of Electronic Engineering and Computer Science,
Queen Mary, University of London, UK, E1 4NS

THIS IS A PRE-PUBLICATION DRAFT OF THE FOLLOWING CITATION:

Constantinou, A. C. & Fenton, Norman E. (2013). Determining the level of ability of football teams by dynamic ratings based on the relative discrepancies in scores between adversaries. *Journal of Quantitative Analysis in Sports. Vol. 9, Iss. 1, 37–50.*

DOI: 10.1515/jqas-2012-0036
CONTACT: Dr. Anthony Constantinou, anthony@constantinou.info, a.constantinou@qmul.ac.uk

ABSTRACT

A rating system provides relative measures of superiority between adversaries. We propose a novel and simple approach, which we call pi-rating, for dynamically rating Association Football teams solely on the basis of the relative discrepancies in scores through relevant match instances. The pi-rating system is applicable to any other sport where the score is considered as a good indicator for prediction purposes, as well as determining the relative performances between adversaries. In an attempt to examine how well the ratings capture a team's performance, we have a) assessed them against two recently proposed football ELO rating variants and b) used them as the basis of a football betting strategy against published market odds. The results show that the pi-ratings outperform considerably the widely accepted ELO ratings and, perhaps more importantly, demonstrate profitability over a period of five English Premier League seasons (2007/08 to 2011/12), even allowing for the bookmakers' built-in profit margin. This is the first academic study to demonstrate profitability against market odds using such a relatively simple technique, and the resulting pi-ratings can be incorporated as parameters into other more sophisticated models in an attempt to further enhance forecasting capability.

Keywords: dynamic rating, ELO rating, football betting, football ranking, football rating, score prediction, soccer betting, soccer ranking, soccer rating, sports betting, sports rating

* Corresponding author. E-mail addresses: anthony@constantinou.info (Constantinou, A. C.), norman@eecs.qmul.ac.uk (Fenton, N. E.)

1 INTRODUCTION

A variety of Association Football (hereafter referred to simply as ‘football’) models formulated by diverse forecast methodologies have been introduced by researchers. This comes as no surprise given that football is the world’s most popular sport (Dunning & Joseph A. M., 1993; Mueller et al., 1996; Dunning E., 1999) and constitutes the fastest growing gambling market (Constantinou & Fenton, 2012).

Determining the relative ability between adversaries is probably the most important element prior to football match prediction, and the current league positions are widely assumed to be an accurate indication of this. However, league positions suffer from numerous drawbacks which makes them unreliable for prediction. For instance, a football league suffers from high variation at the beginning of the season, and from low variation by the end of the season. Additionally, competing teams during a season might not share the equivalent number of matches played due to postponements and thus, the league table will be erroneous for many weeks. In fact, the league table is inherently biased until the final match of the season is played, because for the ranking to be 'fair' each team has to play against residual teams on home and away grounds. Even at the end of the season, the ranking represents the overall performance over the period of a whole season, and fails to demonstrate how the ability of a team varied during that period. Further, it ignores Cup matches and matches from other competitions (e.g. Champions League), and fails to compare teams in different divisions/leagues. In summary, a league table will never be a true indicator of a team's current ability at any specific time. A rating system should provide relative measures of superiority between adversaries and overcomes all of the above complications.

In most of the football forecasting academic literature, the ability of a football team is dependent on the relevant probabilistic rates of historical match outcomes. Even though there have been numerous attempts in formulating more accurate football forecasting models (Maher, 1982; Kuonen, 1996; Buchner, et al., 1997; Dixon & Coles, 1997; Lee, 1997; Kuypers, 2000; Rue & Salvesen, 2000; Crowder et al., 2002; Tsakonas et al., 2002; Karlis & Ntzoufras, 2003; Koning et al., 2003; Dixon & Pope, 2004; Goddard & Asimakopoulos, 2004; Forrest et al., 2005; Goddard, 2005; Halicioglu, 2005a; Halicioglu, 2005b; Rotshtein et al., 2005; Joseph et. al., 2006; Min et al., 2008; Baio & Blangiardo, 2010; Constantinou et al., 2012a; Constantinou et al., 2012b), the use of pure rating systems has not been extensively evaluated. In fact, only three academic papers appear to have assessed the aid of such systems in football.

Knorr-Held (2000) was the first to propose a rating system that is primarily intended for rating football teams, even though it is also applicable to other sports. This proposed system was an extended version of the cumulative link model for ordered responses where latent parameters represent

the strength of each team. The system was tested according to four different measures and two of them disappointed in performance, whereas an assignment of a team-specific smoothing parameter turned out to be difficult for estimation. In (Hvattum & Arntzen, 2010) the authors suggested the use of the ELO rating for football match predictions. The ELO rating system was initially developed for assessing the strength of chess players (Elo, 1978) and is widely accepted and commonly used[†] as a measure of ability; notably in gaming and sports, but also in other disciplines such as recently in biometrics (Reid & Nixon, 2011). The authors (Hvattum & Arntzen, 2010) concluded that, even though the ratings appeared to be useful in encoding the information of past results for measuring the strength of a team, when used in terms of forecasts it appeared to be considerably less accurate compared to market odds. The ELO rating has also been assessed by (Leitner et al., 2010) along with the FIFA/Coca Cola World ratings (FIFA, 2012) for predicting tournament winners. However, both of these rating systems were said to be clearly inferior to bookmakers' odds, on the basis of EURO 2008 football data, which makes the study consistent with the former (Hvattum & Arntzen, 2010).

Harville (1977) stated that a team in American Football should be rewarded for winning *per* se and not for running up the score. Knorr-Held (2000) erroneously assumed that the same logic is applicable to association football on the basis of (Harville, 1977) when formulating performance ratings. In fact, Goddard (2005) demonstrated that no significant difference in forecasting capability is observed between goal-based and result-based regression models for match outcomes in football, and that some advantage is gained by using goal-based (rather than results-based) lagged performance covariates.

In a previous study (Constantinou et al., 2012a) we demonstrated how some of the disadvantages concerning team performances based on league tables can be overcome by introducing further model parameters that reflect a team's form and hence, adjust the ability of a team according to the inconsistencies between predicted and observed recent match performances. Furthermore, even though the model presented in (Constantinou et al., 2012a) appeared to be particularly successful at beating bookmakers' odds, its forecasts did not incorporate score-based information about the relevant football teams.

In this paper we propose a novel rating system that is computationally efficient with low complexity. The technique can be used to formulate both score-based and result-based match predictions, and the pi-ratings can be incorporated into other more sophisticated models in an attempt to further enhance forecasting capability. The model presented in (Constantinou et al., 2012a) is a good example of such a sophisticated model that can benefit from

[†] It might also worth mentioning that the ELO rating algorithm was featured prominently in the popular movie *The Social Network* (also known as *the Facebook movie*), whereby during a scene Eduardo Saverin writes the mathematical formula for the ELO rating system on Zuckerberg's dorm room window.

incorporating the pi-ratings for predictive inference, given that it completely ignores score based information for prediction.

The paper is organised as follows: Section 2 describes the rating system, Section 3 assesses the learning parameters used by the rating system, Section 4 evaluates the accuracy of the resulting ratings, and we provide our concluding remarks and future work in Section 5.

2 THE RATING SYSTEM

The rating system, which we call pi-rating assigns to every new team an initial performance rating of 0, and a rating of 0 represents the rating of the average team relative to the residual teams[‡]. This implies that no inflations or deflations of overall ratings occur over time and thus, if one of the teams gains rating *n* then the adversary loses rating *n.*

When it comes to football, to generate ratings that accurately capture a team's current ability, we have to at least consider:

a) the well known phenomenon of *home advantage* (Clarke & Norman, 1995; Hirotsu & Wright, 2003; Poulter, 2009);

b) the fact that most recent results are more important than less recent when estimating current ability (Constantinou et al., 2012b);

c) the fact that a win is more important for a team than increasing goal difference;

In view of the above 'rules', we propose the three following respective approaches:

a) different ratings for when a team is playing at home and away, but also a catch-up learning rate γ which determines to what extent the newly acquired information based on home performance influences a team's away rating and vice versa;

b) a learning rate λ which determines to what extent the newly acquired information of goal-based match results will override the old information in terms of rating;

[‡] If the rating is applied to a single league competition, the average team in that league will have a rating of 0. If the rating is applied to more than one league in which adversaries between the different leagues (or cup competitions) play against each other, the average team over all leagues will have a rating of 0.

c) high goal error differences, per match instance, are exponentially diminished prior to updating the pi-ratings.

Accordingly, the pi-rating system is built on the following hypothesis: Let us assume that team Y scored 240 and conceded 150 goals over the last 100 matches. Overall, team Y scored 90 goals more than those conceded; a rate of +0.9 goals in favour of team Y per match instance. If we were to predict Y's goal difference at match instance 101 against a random opponent, the best we could do on the basis of the above information is to predict +0.9 goals in favour of team Y and, in this paper, this is what we call Y's expected goal difference against the average opponent. What the pi-rating systems does is simply to revise this expected value on the basis of the rules (a), (b) and (c) specified above, and Sections 2.2 and 3 provide further information regarding the description of this revised value.

2.1. Defining the pi-rating

When a team is playing at home then their new home rating is dependent on (apart from the learning parameters):

- their current home rating;
- the opponent's current away rating;
- the outcome of the match.

In particular, the pi-rating is developed dynamically in cumulative updates whereby discrepancies between predicted and observed goal difference determine whether the rating will increase or decrease (i.e. a team's rating will increase if the score indicates a higher performance than that predicted by the pi-ratings). Accordingly, the overall rating of a team is the average rating between home and away performances, and this is simply defined as:

$$\mathrm{R}_{\tau} = \frac{\mathrm{R}_{\tau\mathrm{H}} + \mathrm{R}_{\tau\mathrm{A}}}{2}$$

where R_τ is the rating for team τ, $R_{\tau H}$ is the rating for team τ when playing at home, and $R_{\tau A}$ is the rating of team τ when playing away. Assuming a match between home team α and away team β, then the home and away ratings are respectively updated cumulatively as follows:

1. updating home team's home rating $\rightarrow \widehat{\mathrm{R}_{\alpha\mathrm{H}}} = \mathrm{R}_{\alpha\mathrm{H}} + \psi_H(e) \times \lambda$

2. updating home team's away rating $\rightarrow \widehat{\mathrm{R}_{\alpha\mathrm{A}}} = \mathrm{R}_{\alpha\mathrm{A}} + \left(\widehat{\mathrm{R}_{\alpha\mathrm{H}}} - \mathrm{R}_{\alpha\mathrm{H}}\right) \times \gamma$

3. updating away team's home rating $\rightarrow \widehat{\mathrm{R}_{\beta\mathrm{A}}} = \mathrm{R}_{\beta\mathrm{A}} + \psi_A(e) \times \lambda$

4. updating away team's away rating $\rightarrow \widehat{\mathrm{R}_{\beta\mathrm{H}}} = \mathrm{R}_{\beta\mathrm{H}} + \left(\widehat{\mathrm{R}_{\beta\mathrm{A}}} - \mathrm{R}_{\beta\mathrm{A}}\right) \times \gamma$

where $R_{\alpha H}$ and $R_{\alpha A}$ are the current home and away ratings for team α, $R_{\beta H}$ and $R_{\beta A}$ are the current home and away ratings of team β, $\hat{R}_{aH}$, $\hat{R}_{aA}$, $\hat{R}_{\beta H}$ and $\hat{R}_{\beta A}$ are the respective revised ratings, e is the error between predicted and observed goal difference (which we explain in detail in Section 2.3), $\psi(e)$ is a function of e (which we explain in detail in Section 2.2) and λ and γ are the learning rates (which we explain in detail in Section 3). Further, a step-by-step example of how the ratings are revised is presented in Section 2.4.

2.2. Weighting error (e)

The primary objective of this function is to diminish the importance of high score differences when updating the ratings. Figure 1 illustrates $\psi(\mathrm{e})$ against e. In particular, $\psi(\mathrm{e})$ is a function of e on the basis of the following equation:

$$\psi(\mathrm{e}) = \mathrm{c} \times \log_{10}(1 + \mathrm{e})$$

where c is a constant set to $c = 3$.

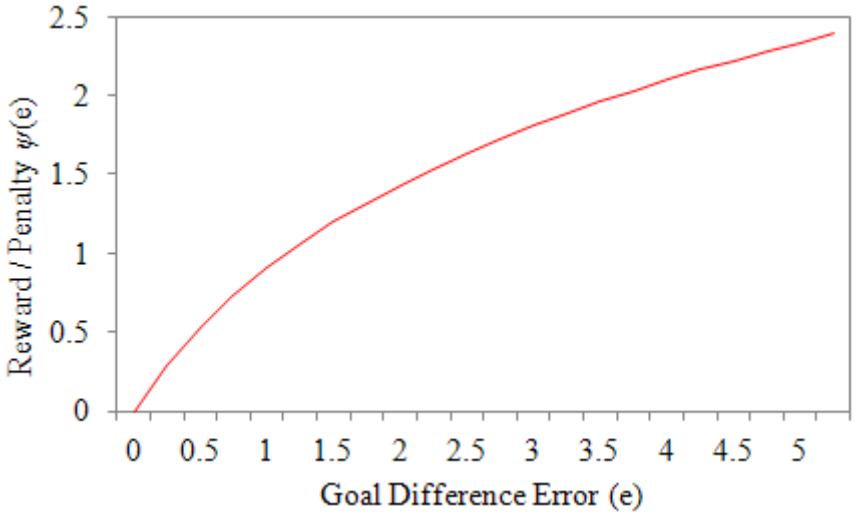

Figure 1. Weighted $\psi(e)$ reward/penalty relative to e, assuming $c = 3$.

Recognising that a win is more important than increasing goal difference is important in football. However, we do not know exactly how 'less' important each additional score difference becomes for individual match instances and hence, there are many possible ways to introduce diminishing returns. No relevant published paper appears to have sufficiently addressed this issue, and is an area of future investigation. In this paper we provide one possible approach for handling this. As a result, even though the deterministic function proposed in this section might appear to adequately capture (see Section 4) the importance of high goal differences, it should be noted that this approach is still a weakness. Perhaps a more traditional approach of dealing with this issue, is to consider a class of functions that can be optimised based on an appropriate data analysis.

2.3. Measuring error (e)

The observed goal difference g_D is simply the difference in goals scored, i.e. for the home team $g_D = g_H - g_A$ where g_H and g_A are the number of goals scored by the home and away team respectively. Accordingly, when g_D is positive implies that the home team wins and vice versa.

We describe $\widehat{g_D}$ as the expected goal difference, relative to the home team, against the average opponent (i.e. the revised expected value as discussed in section 2). There is no mathematical reason for the resulting computation to produce predicted score differences, but we accept that the resulting $\widehat{g_D}$ values are useful in earning such a description on the basis of the empirical evidence that we later provide in Section 3. Accordingly, the expected goal difference against the average opponent can then be measured as follows:

$$\widehat{g_{DG}} = b^{\frac{|R_{\tau G}|}{c}} - 1$$

where $\widehat{g_{DG}}$ is the expected goal difference against the average opponent for the team that plays at ground G (hence we have $\widehat{g_{DH}}$ and $\widehat{g_{DA}}$), b is equal to the base of the logarithm used $b = 10$, and $R_{\tau G}$ is the rating for team τ at ground G. When a team's rating is < 0 the outcome is simply $-\widehat{g_{DG}}$. The predicted goal difference between adversaries is then $\widehat{g_D} = \widehat{g_{DH}} - \widehat{g_{DA}}$. Accordingly, the error e between predicted and observed goal difference is[§]:

$$e = |g_D - \widehat{g_D}|$$

[§] If the prediction is +4 in favour of the home side then an actual result of $5 - 0$ will give you an error of approximately 1. But if the prediction is 0 in favour of the home side and the actual result is $1 - 0$, then this also gives you the same error as above.

2.4. Updating pi-ratings: An Example

Figure 2 illustrates a 6-step continuous cycle process for updating the pi-ratings. In this section, we follow this step-diagram to update the pi-ratings for two given teams with hypothetical home and away ratings.

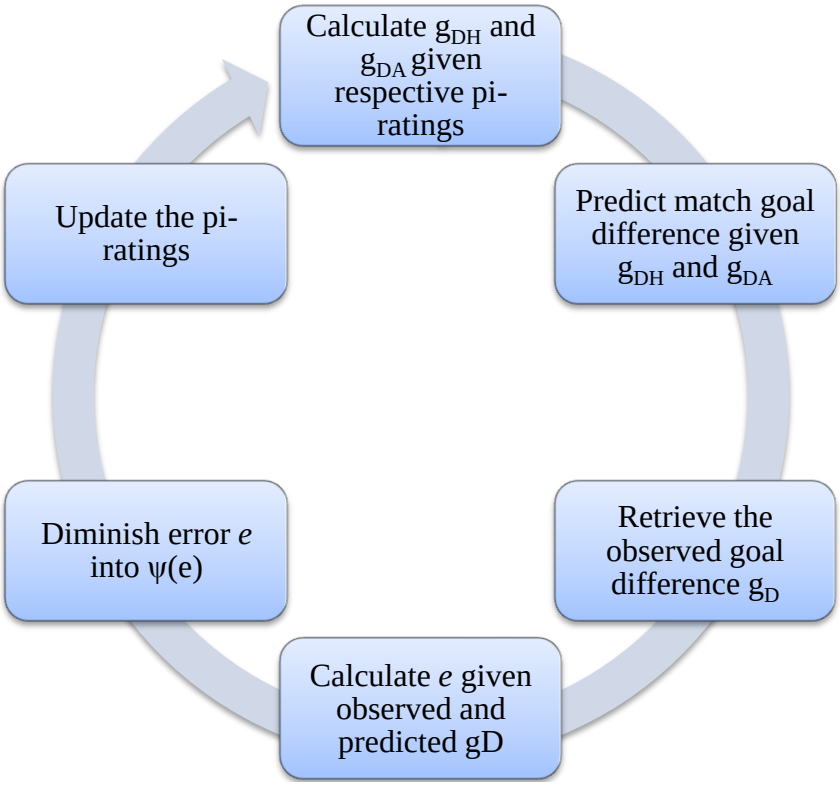

Figure 2. The process of updating the pi-ratings.

Step 1: Suppose that we have a match instance where team α (the home team) with ratings $\{R_{aH} = 1.6, R_{aA} = 0.4\}$ plays against team β (the away team) with ratings $\{R_{\beta H} = 0.3, R_{\beta A} = -1.2\}$. For this example, we have to consider team's α current home rating and team's β current away rating; R_{aH} and $R_{\beta A}$ respectively. Converting the pi-ratings into predicted goal differences against the average adversary $\widehat{g_{DH}}$ and $\widehat{g_{DA}}$, as demonstrated in Section 2.3, we retrieve the following information:

- team α is expected to win by:

$$\widehat{g_{DH}} = b^{\frac{|R_{\tau G}|}{c}} - 1 = 10^{\frac{1.6}{3}} - 1 = 2.4145$$

goals difference against the average opponent when playing at home;

- team β is expected to lose by:

$$\widehat{g_{DA}} = -\left(b^{\frac{|R_{\tau G}|}{c}} - 1\right) = \left(10^{\frac{1.2}{3}} - 1\right) = -1.5119$$

goals difference against the average opponent when playing away.

Step 2: Using the above information we can predict the match goal difference:

$$\widehat{g_D} = (\widehat{g_{DH}} - \widehat{g_{DA}}) = \left(2.4145 - (-1.5119)\right) = +3.9264$$

As defined in Section 2.3, the home team is expected to win by 3.9264 goals.

Step 3: We want to know the observed score. Suppose that we observe the score '4-1' $(+3)$. Therefore: $g_D = 3$.

Step 4: We can now compute the error e between predicted and observed goal difference. Based on the equation from Section 2.3 we get:

$$e = |g_D - \widehat{g_D}| = |(+3) - (+3.9264)| = 0.9264$$

Step 5: Prior to updating the respective pi-ratings, we want to first weight e. Accordingly, the diminishing equation will simply return the value of e:

$$\psi(e) = c \times \log_{10}(1 + e) = 3 \times 0.2847 = 0.8542$$

Therefore,

$$\psi_H(e) = \begin{cases} \psi(e), & \widehat{g_D} < g_D \\ -\psi(e), & \text{otherwise} \end{cases} \quad \text{and} \quad \psi_A(e) == \begin{cases} \psi(e), & \widehat{g_D} > g_D \\ -\psi(e), & \text{otherwise} \end{cases}$$

Step 6: We can now revise the pi-ratings. Assuming the learning rates of $\lambda = 0.1$ and $\gamma = 0.3$, the current ratings are revised as follows:

- $\mathbf{Team_a}$ **(home rating) :**

$\widehat{R}_{\alpha H} = R_{\alpha H} + \psi_H(e) \times \lambda = 1.6 + (-0.8542) \times 0.1 = 1.5145$ (down from 1.6);

- **$\mathbf{Team_a}$ (away rating) :**

  $\widehat{\mathrm{R}}_{\alpha\mathrm{A}} = \mathrm{R}_{\alpha\mathrm{A}} + \left(\widehat{\mathrm{R}}_{\alpha\mathrm{H}} - \mathrm{R}_{\alpha\mathrm{H}}\right) \times \gamma = 0.4 + (1.5145 - 1.6) \times 0.3 = 0.3743$ (down from 0.4);

- **$\mathbf{Team_\beta}$ (away rating) :**

  $\hat{R}_{\beta\mathrm{A}} = \mathrm{R}_{\beta\mathrm{A}} + \psi_A(e) \times \lambda = -1.2 + (+0.8542) \times 0.1 = -1.1145$ (up from -1.2);

- **$\mathbf{Team_\beta}$ (home rating)**:

  $\widehat{\mathrm{R}}_{\beta\mathrm{H}} = \mathrm{R}_{\beta\mathrm{H}} + \left(\widehat{\mathrm{R}}_{\beta\mathrm{A}} - \mathrm{R}_{\beta\mathrm{A}}\right) \times \gamma = 0.3 + \left(-1.1145 - (-1.2)\right) \times 0.3 = 0.3256$ (up from 0.3).

Even though team α beat team β '4-1', team's α ratings are decreased from $\{\mathrm{R}_{\alpha\mathrm{H}} = 1.6, \mathrm{R}_{\alpha\mathrm{A}} = 0.4\}$ to $\{\mathrm{R}_{\alpha\mathrm{H}} = 1.5145, \mathrm{R}_{\alpha\mathrm{A}} = 0.3743\}$, and team's β ratings are increased from $\{\mathrm{R}_{\beta\mathrm{H}} = 0.3, \mathrm{R}_{\beta\mathrm{A}} = -1.2\}$ to $\{\mathrm{R}_{\beta\mathrm{H}} = 0.3256, \mathrm{R}_{\beta\mathrm{A}} = -1.1145\}$. This happened because according to the ratings team α was expected to win by 3.7 goals against team β.

3 DETERMINING THE LEARNING RATES

In football, new observations are always more important than the former, and no matter how home and away performances differ for a team, we can still gain some information about a team's next away performance based on its previous home performance (and vice versa). Thus, determining optimal learning rates for parameters λ and γ is paramount for generating ratings that accurately capture the current level of performance of a team.

The learning parameters λ and γ can take values that go from 0 to 1. A higher learning rate λ determines to what extent the newly acquired information of match results will override the old information in terms of rating, and a higher learning rate γ determines the impact the home performances have on away ratings (and vice versa). For instance, when $\lambda = 0.1$ a team's rating will adjust with cumulative updates based on new match results with a weighing factor of 10%, and when $\gamma = 0.5$ a team's home performances will affect that team's away ratings with a weighting factor of 50% relative to the revised home rating.

In determining optimal learning rates we have assessed the ratings generated for different values of λ and γ by formulating score-based[**]

[**] The learning parameters could have been optimised based on predictions of type $\{\mathrm{H}, \mathrm{D}, \mathrm{A}\}$ (corresponding to home win, draw and away win), based on profitability, based on scoring

predictions, as demonstrated in Section 2, about the last five English Premier League (EPL) seasons; 2007/08 to 2011/12. For training the learning parameters[††] we have considered relevant historical data (Football-Data, 2012) beginning from season 1992/93 up to the end of season 2006/07. Accordingly, if a combination of learning rates λ and γ increase the forecast accuracy, then we assume that both λ and γ are a step closer to being optimal.

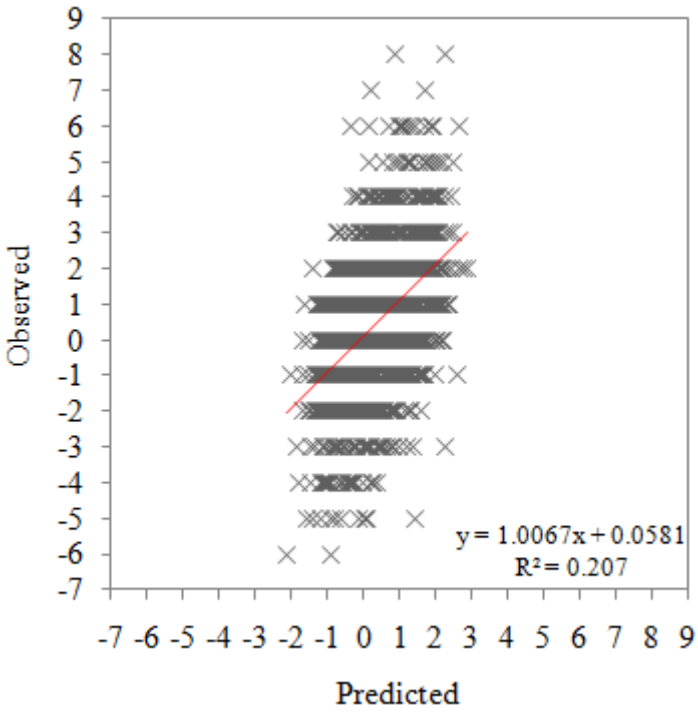

Figure 3. Predicted versus observed goal difference, with $Predicted = Observed$ superimposed, over the EPL seasons 2007/08 to 2011/12 (1900 match instances); where *Predicted* is $\widehat{g_D}$ *Observed* is the observed goal difference. The goal difference is illustrated relative to the home team as defined in Section 2.3; whereby a positive difference indicates a home win and vice versa.

Figure 4 illustrates how parameters λ and γ affect the squared error in predicted score difference over the EPL seasons 1997/98 to 2006/07 inclusive, where the error is simply the difference between predicted and observed goal difference (e.g. if a model predicts $+1$ goal for the home team and the observation is $+1$ goal for the away team then the absolute score error is 2

---

rules, or based on many other different accuracy measurements and metrics. We have chosen score difference for optimising the learning parameters since the pi-ratings themselves are exclusively determined by that information.

[††] The first five EPL seasons (1992/93 to 1996/97) are solely considered for generating the initial ratings for the competing teams. This is important because training the model on ignorant team ratings (i.e. starting from 0) will negatively affect the training procedure. Thus, learning parameters λ and γ are trained during the subsequent ten seasons; 1997/98 to 2006/07 inclusive.

goals). The generated values for each combination of learning rates are provided in Table B.1, Appendix B. Our results show that the combination of $\lambda = 0.035$ and $\gamma = 0.7$ generates the lowest prediction error.

Figure 3 provides empirical evidence, on the basis of a grid search over the error values presented in Table B.1, that the suggested combination of learning rates provides ratings that accurately capture a team's current performance. In particular, on the basis of *predicted* (effectively $\widehat{g_D}$) versus *observed* goal difference over the five EPL seasons, the identity line with $Predicted = Observed$ superimposed considers the two datasets to be significantly correlated. This information justifies treating $\widehat{g_D}$ as useful for predicting score differences. However, Figure 3 demonstrates the limitation in predicting fixed score differences on the basis of the large variability in observed scores. That is, even though we observe a relatively high number of score differences that are ≥ 4 (especially for the home team), the pi-rating system was never able to suggest such a high score difference as the most likely outcome (i.e. when a very strong team plays against a very weak team the most likely outcome in terms of expected score difference is normally approximately 3 goals in favour of the strong team, according to the pi-rating).

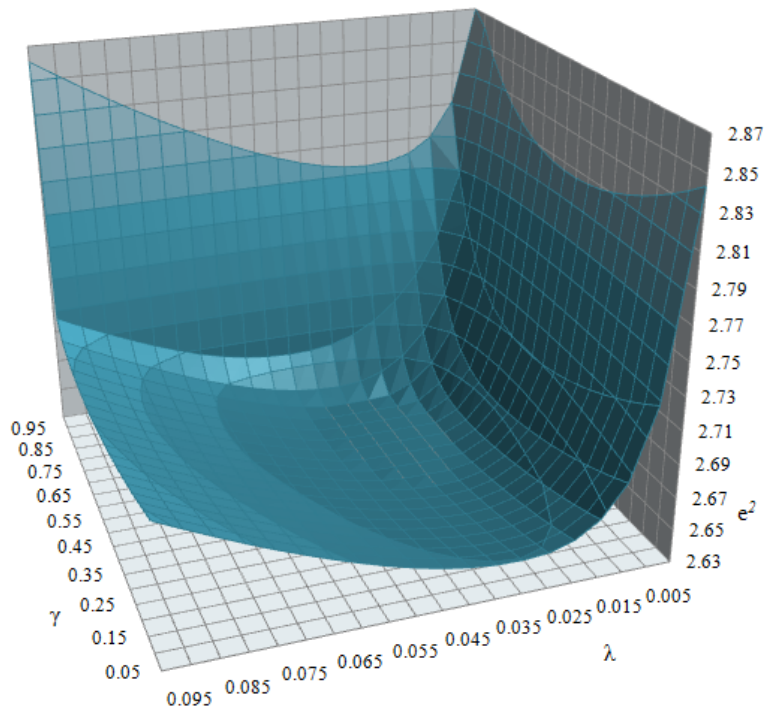

Figure 4. Estimating optimum λ and γ learning rates based on squared goal difference error e^2, for the EPL seasons 2007/08 to 2011/12. Minimum squared error of expected goal difference observed when $\lambda = 0.035$ and $\gamma = 0.7$, where $e^2 = 2.6247$

The section that follows demonstrates the performance of the pi-rating system, in terms of profitability, on the basis of the optimum learning rates of $\lambda = 0.035$ and $\gamma = 0.7$.

4 RATING DEVELOPMENT AND FORECASTING CAPABILITY

In an attempt to examine how well the pi-rating system captures a team's performance, we have compared it against the two ELO rating variants, the $\mathrm{ELO_b}$ and $\mathrm{ELO_g}$, which have recently been proposed for rating football teams (Hvattum & Arntzen, 2010). Appendix C provides the description of the two ELO ratings, as defined by the authors. For performance comparison, we have used the ratings as the basis of a football betting strategy against published market odds by considering all the match instances (1900) that had occurred during the EPL seasons 2007/08 to 2011/12 inclusive. For market odds data we have considered the *Betbrain* maximums (best available for the bettor) published by (Football-Data, 2012). The odds are available only in *home win - draw - away win* (HDA) form and hence, we have to formulate probabilities for each of those outcomes.

The predictive distribution $\{\mathrm{p(H)}, \mathrm{p(D)}, \mathrm{p(A)}\}$ is formulated directly from the historical database on the basis of predetermined levels of discrepancy between team ratings, and this method is similar to that proposed in (Constantinou et al., 2012a) on the basis of team strength. The granularity[‡‡] (of 28 levels of team rating discrepancy) has been chosen to ensure that for any match combination (i.e. a team of rating R_{aH} against a team of rating $R_{\beta A}$) there are sufficient data points for a reasonably well informed prior for $\{\mathrm{p(H)}, \mathrm{p(D)}, \mathrm{p(A)}\}$.

All of the rating systems consider identical datasets for formulating rating priors, training the ratings, and assessing profitability. When it comes to the ELO ratings, the home advantage is directly determined by the intervals (i.e. how a home team with an ELO rating x performs against an away team with an ELO rating of y). However, unlike the pi-rating system, the ELO ratings consider identical home advantage for all teams (that share identical rating discrepancies against the away team).

Further, we consider the parameters $[\mathrm{k}]$ and $[\mathrm{k_0}, \lambda]$ with the suitable values of $[\mathrm{k} = 20]$ and $[\mathrm{k_0} = 10,\ \lambda = 1]$, for $\mathrm{ELO_b}$ and $\mathrm{ELO_g}$ respectively, as suggested by the authors (Hvattum & Arntzen, 2010). However, we have also optimised the parameter values against our training data and assessed the difference between the two in terms of profitability. Accordingly, we found the optimum parameter values to be $[\mathrm{k} = 56]$ and $[\mathrm{k_0} = 2,\ \lambda = 2.8]$ for the $\mathrm{ELO_b}$ and $\mathrm{ELO_g}$ ratings respectively, assuming that the prior ELO rating for

[‡‡] For the pi-rating system the ratings are segregated into intervals of 0.10 (from ≤ -1.1 to > 1.6), for $\mathrm{ELO_b}$ the ratings are segregated into intervals of 25 (from ≤ -330 to > 345), and for $\mathrm{ELO_g}$ the ratings are segregated into intervals of 35 (from ≤ -475 to > 470).

each adversary is set to 1500. Appendix D illustrates how the ELO score error e converges over [k] and $[k_0, \lambda]$, where the minimum values of e observed are around 0.3514 and 0.3405 for ELO_b and ELO_g respectively.

4.1. Profitability Assessment

For betting simulation, we have followed a very simple strategy whereby for each match instance we place a £1 bet on the outcome with the highest discrepancy of which each rating system predicts with higher probability relative to published market odds. For example, assuming the predicted probabilities of $\{0.20, 0.36, 0.45\}$ against the published market probabilities of $\{0.30, 0.35, 0.40\}$[§§], then a bet is simulated against outcome A (which is the outcome with the highest discrepancy in favour of the rating system). If no discrepancy is observed in favour of the rating system, a bet will not be simulated.

Figures 5 and 6 demonstrate the distinct and overall cumulative profit/loss observed against published market odds during the five specified EPL seasons. Table 1 presents the summary statistics of the betting simulation. The simulation shows a rather consistent performance over the five seasons, whereby bets won vary between 28% to 37% at odds that vary between 2.79 and 3.27. Overall, the technique is profitable which implies that the pi-rating system properly captures the ability of a team at any time interval throughout the season. This implies that the pi-rating system was able to generate profit vial longshot bets. A behaviour that is similar to that demonstrated by the football forecast model of (Constantinou et al., 2012a), and this is interesting because the two models follow two completely different approaches to prediction.

Table 1. Betting simulation: outcomes and statistics.

EPL season	Match instances	Number of bets	Bets won	Winning odds (mean)	Total stakes	Total returns	Profit/ Loss
2007/08	380	372	121 (32.53%)	2.7959	£372	£338.31	-£33.69
2008/09	380	378	140 (37.04%)	3.1297	£378	£438.16	+£60.16
2009/10	380	380	109 (28.68%)	3.2603	£380	£355.38	-£24.62
2010/11	380	377	122 (32.36%)	3.2492	£377	£396.41	+£19.41
2011/12	380	380	127 (33.42%)	3.2784	£380	£416.36	+£36.36
TOTAL	1900	1887	619 (32.81%)	3.1415	£1887	£1944.62	+£57.62

[§§] Assumes a profit margin of 5%.

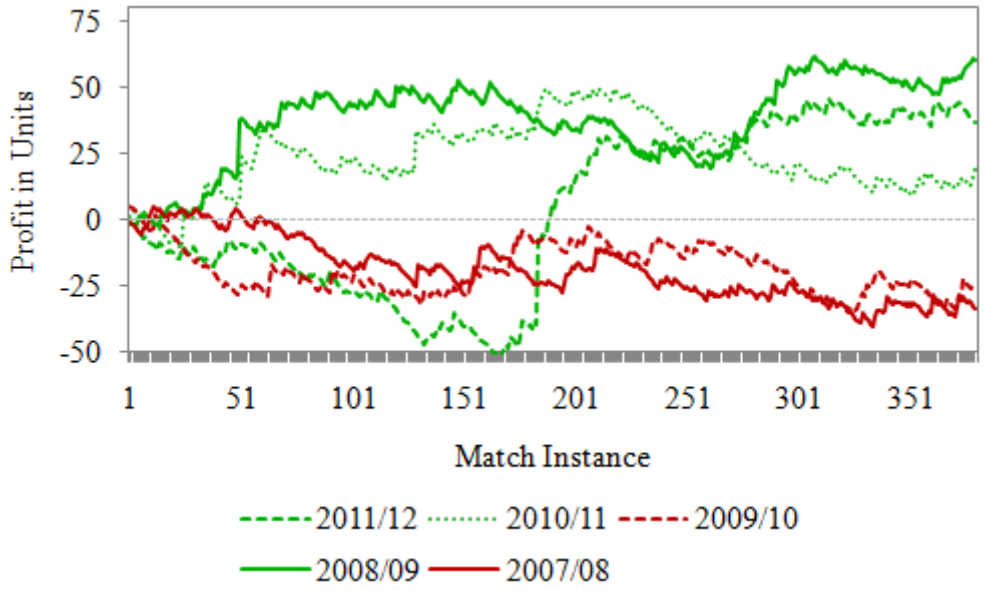

Figure 5. Distinct cumulative profit/loss observed against published market odds, based on pi-rating forecasts, during the EPL seasons 2007/08 to 2011/12 inclusive.

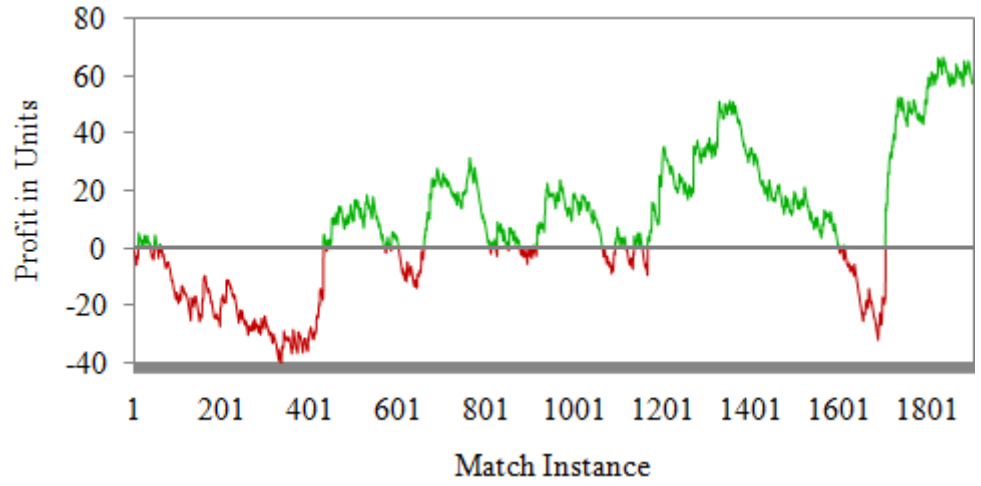

Figure 6. Overall cumulative profit/loss observed against published market odds, based on pi-rating forecasts, during the EPL seasons 2007/08 to 2011/12 inclusive.

In (Hvattum & Arntzen, 2010) the authors concluded that the ELO ratings appeared to be useful in encoding information of past results, but when used in terms of forecasts it appeared to be considerably less accurate when compared against published market odds. The authors recognised the popularity of the ELO ratings as a measure of team strength, but questioned their possibility of generating predictions that are on par with the market odds (Hvattum & Arntzen, 2010). In (Leitner et al., 2010) the authors provided

similar results based on another ELO rating variant called *The World Football Elo Rating*. In particular, the authors recognised the bookmakers' odds as a better performing model (on the basis of EURO 2008 tournament data) when compared to the ELO ratings, and suggested that various improvements are conceivable and deserve further study.

Figure 7 demonstrates the profitability of the pi-rating system against the ELO ratings. Our results are consistent with (Leitner et al., 2010; Hvattum & Arntzen, 2010). In particular, the ELO ratings perform considerably less accurately against the market, and it is clear that the pi-ratings are an improvement.

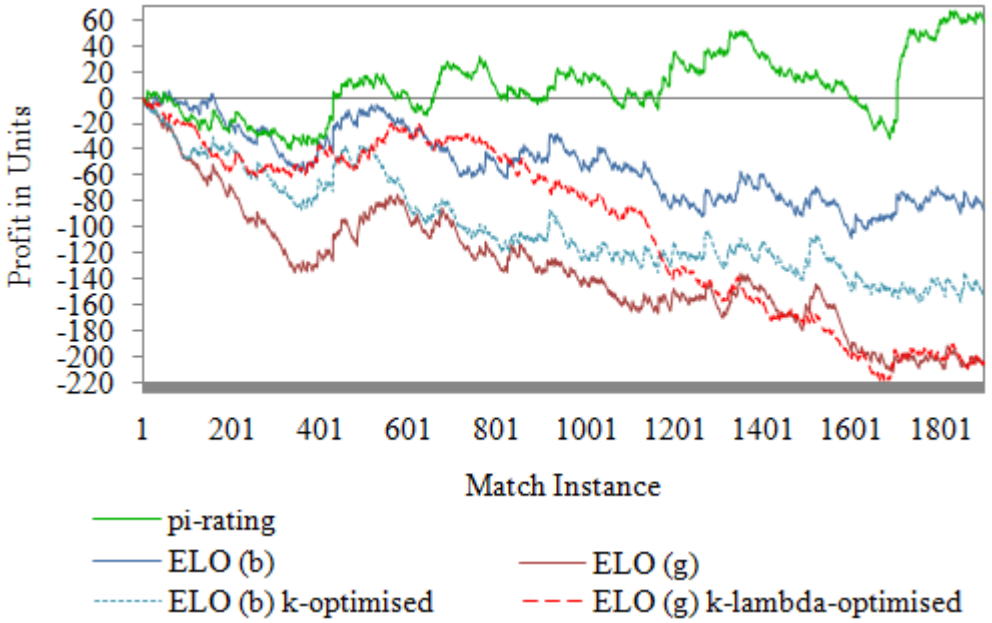

Figure 7. Overall cumulative profit/loss observed against published market odds, according to each of the rating procedures, during the EPL seasons 2007/08 to 2011/12 inclusive.

4.2. Rating development

Figure A.1 illustrates how the pi-ratings develop for the six most popular EPL teams over the course of the last 20 seasons, whereas Figure 8 illustrates how the pi-ratings develop for those identical teams during the last five seasons (1900 match instances) if we consider no previous relevant historical information. In particular, at match instance 1 (first match of season 2007/08) all six teams start at rating 0. The development of the rating shows that two seasons of relevant historical outcomes (76 match instances per team) might be enough for it to converge into acceptable estimates on the basis of

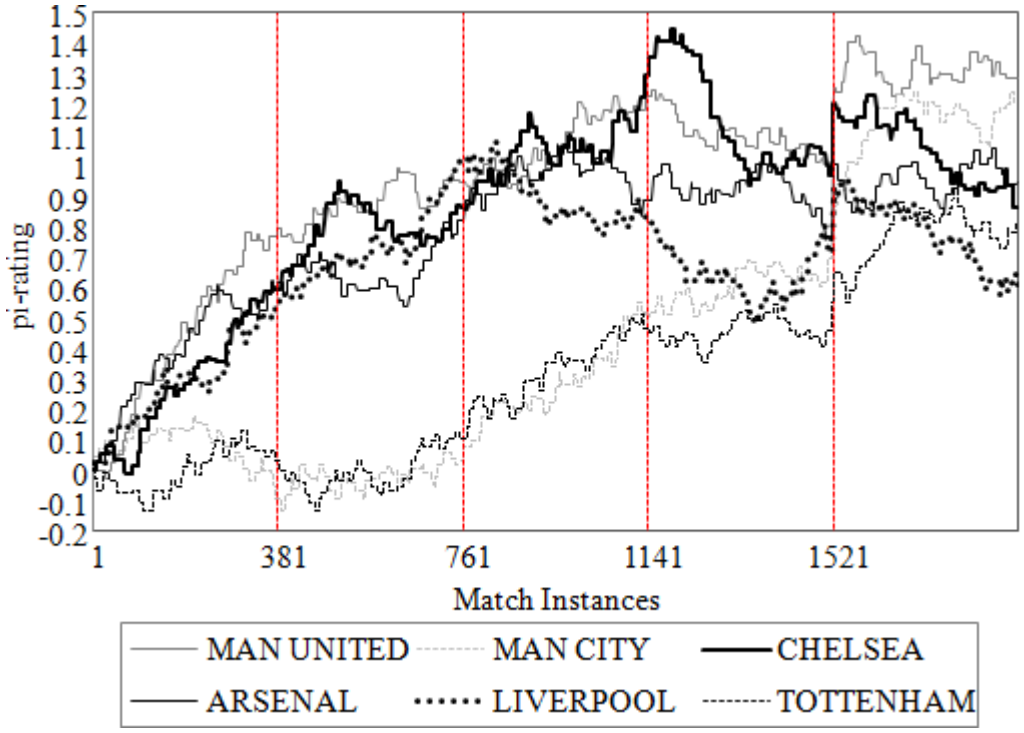

Figure 8. Development of the pi-ratings, assuming $\lambda = 0.035$ and $\gamma = 0.7$, for seasons 2007/08 and 2011/12.

the specified λ and γ rates. However, a further season of historical match outcomes might be required for teams with the uppermost difference from the average team (such as Chelsea and Manchester United).

In contrast to earlier studies that assumed or concluded that the home advantage factor is invariant between football teams and hence considered a single generalised model parameter for that matter (Knorr-Held 1997, 2000; Koning, 2000; Baio & Blangiardo, 2010; Hvattum & Arntzen, 2010; Leitner, 2010), our results show that this is not the case. Figure 9 illustrates how the ratings develop on the basis of home and away performances for Manchester United, Blackburn, Wolves and Arsenal during the same five EPL seasons. In particular, Manchester United and Blackburn demonstrate a high variation between home and away performances, whereas Wolves and Arsenal appear to perform almost indifferently between home and away. This outcome is consistent with (Clarke & Norman, 1995) who, in fact, reported that in many cases a team can even develop a negative home advantage.

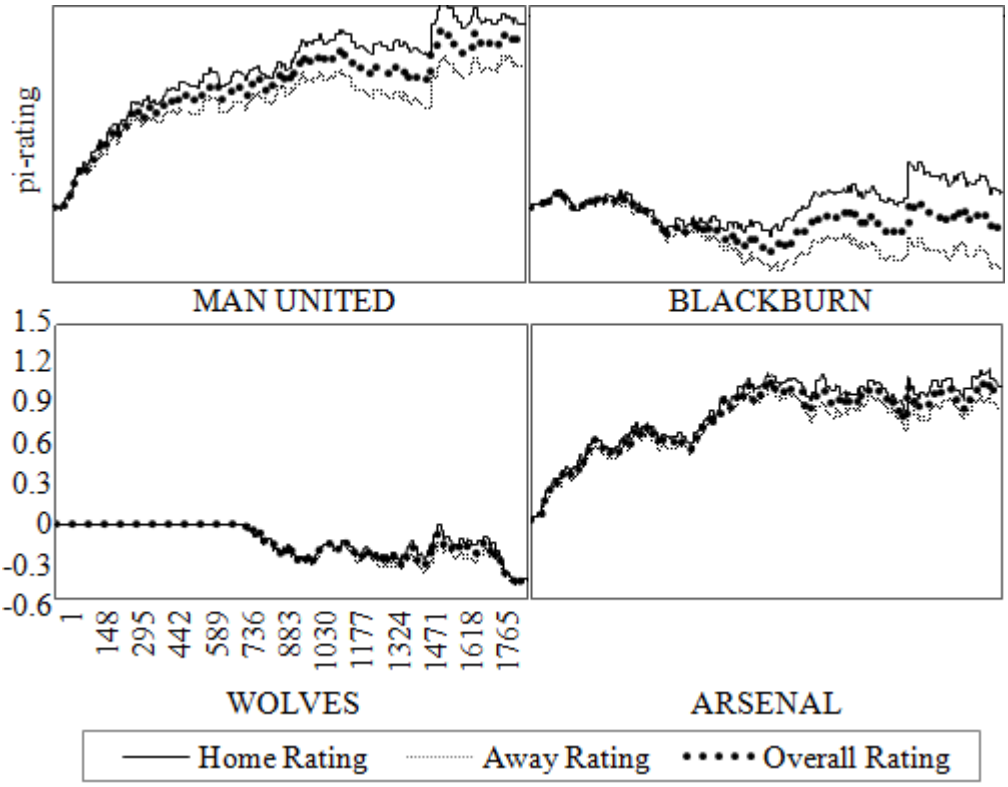

Figure 9. Development of the pi-ratings, assuming $\lambda = 0.035$ and $\gamma = 0.7$, based on individual home and away performances for the specified teams[***] and from season 2007/08 to 2011/12 inclusive.

5 CONCLUDING REMARKS AND FUTURE WORK

We have proposed a novel rating system, which we call pi-rating, for determining the level of ability of football teams on the basis of the relative discrepancies in scores between adversaries. The pi-rating is computationally efficient with low complexity and proceeds with dynamic modifications after every new match instance is observed. The pi-ratings can be used to formulate both score-based and result-based match predictions.

The pi-rating system considers different ratings for when a team is playing at home and away, considers the most recent results to be more important than the less recent, and diminishes the importance of high goal differences in predicted error when revising the pi-ratings. The learning parameters ensure that the newly acquired match results are more important than the former and that the newly acquired information based on a home ground performance influences a team's ratings when playing away and vice

[***] For the newly promoted team Wolves the development of the ratings start at match instance 760 since no performances have been recorded relative to the residual EPL teams during the two preceding seasons.

versa. More importantly, the learning parameters follow optimised rates which ensure that the pi-rating system proceeds with appropriate rating modifications.

In an attempt to examine how well the pi-ratings capture a team's performance, we have a) assessed it against two recently published football ELO rating variants and b) used it as the basis of a football betting strategy against published market odds. Over the period of five English Premier League seasons (2007/08 to 2011/12) the results show that the pi-ratings outperform the ELO ratings considerably and, perhaps more importantly, demonstrate profitability over a period of five English Premier League seasons (2007/08 to 2011/12), even allowing for the bookmakers' built-in profit margin. This implies that the pi-rating system generates performance values of higher accuracy when compared to the popular and widely accepted ELO rating system, while at the same time keeping the rating complexity and computational power required at roughly the same levels. Further, this is the first academic study to demonstrate profitability against published market odds on the basis of such a simple technique, and the resulting ratings can be incorporated as parameters into other more sophisticated models in an attempt to further enhance forecasting capability. In summary, the pi-ratings may:

a) simplify the process for a forecasting football model in the sense that the rating values will reflect a team's current performance and thus, further factors and techniques that are normally introduced for determining the *current form* of a team by weighting the importance of the more recent results may become redundant;

b) be incorporated into models that solely focus on results-based data and hence, enhance information considered on the basis that the pi-ratings are developed given score-based data;

Planned extensions of this research will determine:

a) the importance of the pi-ratings, by replacing relevant techniques of higher complexity for determining *current team form*, as inputs[†††] to the Bayesian network models that we have proposed in (Constantinou et al., 2012a; Constantinou et al., 2012b);

b) the value of pi-ratings in evaluating the relative ability of teams between different leagues, by considering relevant match occurrences between teams of those leagues (e.g. Uefa Champions League). If

[†††] Where the pi-ratings of the home and away team follow $\sim\text{Normal}(x, y)$ distributions for capturing rating uncertainty, where x is the pi-rating value (R_{aH} or $R_{\beta A}$) and y is the pi-rating variance, which can be measured over n preceding match instances.

successful, this will allow us to answer interesting questions such as *'which football league is best; the English Premier League or the Spanish La Liga?'*, and *'to what degree lower divisions differ from higher divisions in England'*, or even *'how much damage has the 2006 Italian football scandal, which was described as the biggest scandal in football history (Murali, 2011), caused to Serie A?'*.

ACKNOWLEDGEMENTS

We acknowledge the financial support by the Engineering and Physical Sciences Research Council (EPSRC) for funding this research, and the reviewers and Editors of this Journal whose comments have led to significant improvements in the paper.

Appendix A: Rating development over a period of 20 seasons

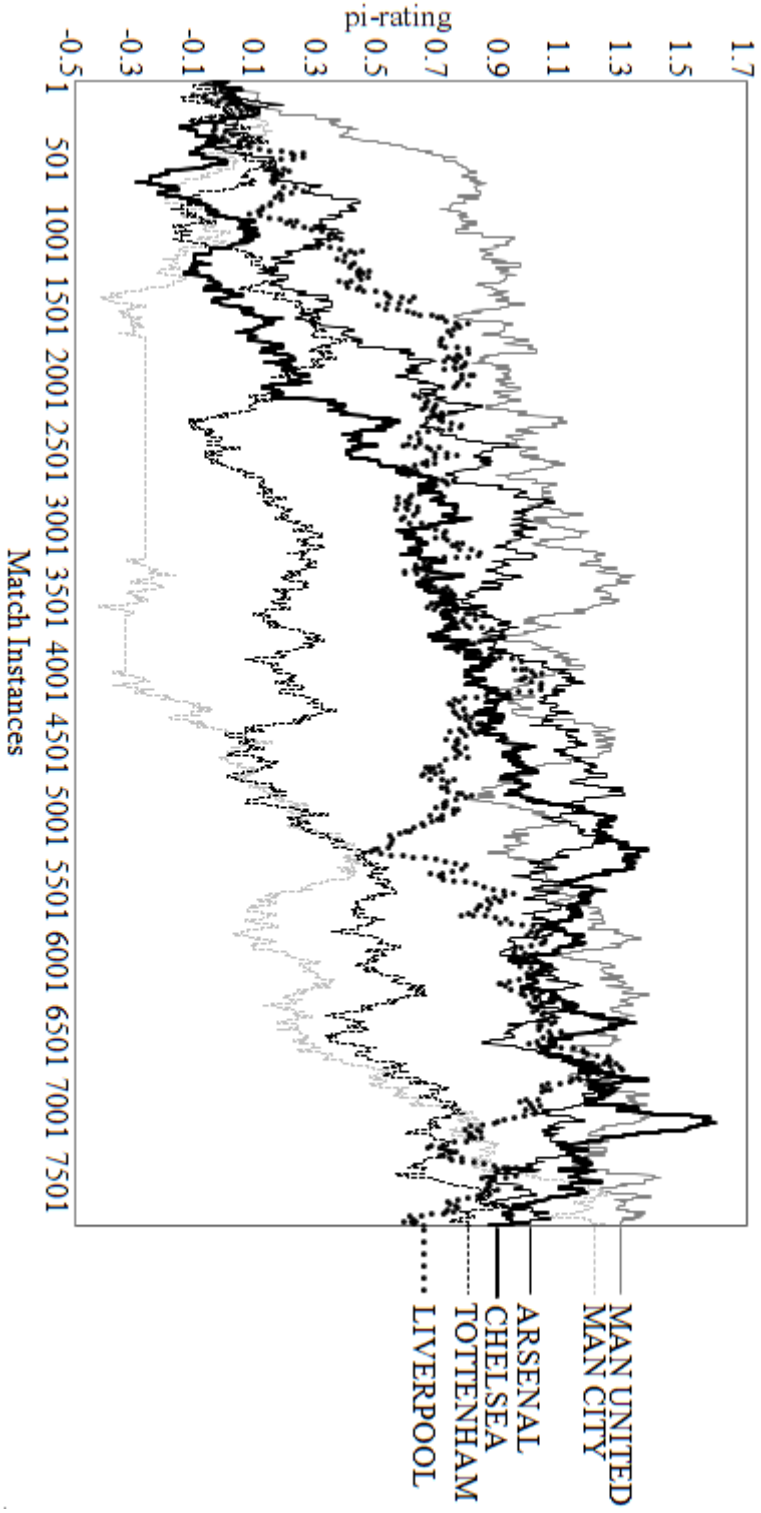

Figure A1. Rating development over a period of 20 seasons, assuming $\lambda = 0.035$ and $\gamma = 0.7$, for the six most popular EPL teams (from season 1992/93 to season 2011/12 inclusive).

Appendix B: Learning rates λ and γ

Table B.1. Squared error values generated based on learning rates λ and γ.

γ / λ	0.005	0.010	0.015	0.020	0.025	0.030	0.035	0.040	0.045	0.050	0.055	0.060	0.065	0.070	0.075	0.080	0.085	0.090	0.095	0.100
1.00	2.8809	2.8115	2.7896	2.7800	2.7760	2.7749	2.7758	2.7782	2.7818	2.7865	2.7918	2.7979	2.8044	2.8113	2.8187	2.8264	2.8345	2.8429	2.8518	2.8610
0.95	2.8616	2.7680	2.7261	2.7012	2.6842	2.6723	2.6644	2.6598	2.6577	2.6572	2.6579	2.6598	2.6628	2.6668	2.6718	2.6775	2.6838	2.6906	2.6977	2.7051
0.90	2.8461	2.7389	2.6918	2.6659	2.6506	2.6415	2.6369	2.6355	2.6359	2.6376	2.6405	2.6445	2.6492	2.6545	2.6605	2.6671	2.6738	2.6809	2.6884	2.6963
0.85	2.8340	2.7193	2.6730	2.6494	2.6371	2.6309	2.6286	2.6287	2.6300	2.6327	2.6362	2.6407	2.6456	2.6509	2.6569	2.6633	2.6701	2.6772	2.6847	2.6926
0.80	2.8246	2.7064	2.6624	2.6417	2.6315	2.6268	2.6258	2.6264	2.6282	2.6310	2.6348	2.6392	2.6441	2.6493	2.6550	2.6613	2.6680	2.6751	2.6826	2.6906
0.75	2.8173	2.6981	2.6567	2.6381	2.6293	2.6253	2.6249	2.6257	2.6277	2.6305	2.6343	2.6386	2.6435	2.6487	2.6542	2.6603	2.6670	2.6740	2.6815	2.6895
0.70	2.8119	2.6930	2.6538	2.6368	2.6286	2.6250	2.6247	2.6257	2.6279	2.6307	2.6343	2.6386	2.6433	2.6484	2.6540	2.6600	2.6666	2.6736	2.6811	2.6889
0.65	2.8082	2.6902	2.6527	2.6366	2.6289	2.6254	2.6251	2.6262	2.6284	2.6313	2.6348	2.6390	2.6436	2.6487	2.6542	2.6603	2.6667	2.6737	2.6811	2.6889
0.60	2.8059	2.6890	2.6527	2.6372	2.6297	2.6263	2.6258	2.6270	2.6291	2.6321	2.6356	2.6398	2.6444	2.6493	2.6549	2.6610	2.6674	2.6743	2.6816	2.6893
0.55	2.8049	2.6891	2.6535	2.6384	2.6310	2.6275	2.6269	2.6281	2.6302	2.6331	2.6367	2.6408	2.6454	2.6504	2.6559	2.6620	2.6684	2.6754	2.6826	2.6903
0.50	2.8050	2.6902	2.6549	2.6399	2.6325	2.6290	2.6283	2.6295	2.6315	2.6343	2.6379	2.6421	2.6468	2.6518	2.6572	2.6632	2.6698	2.6767	2.6839	2.6917
0.45	2.8061	2.6921	2.6568	2.6418	2.6342	2.6308	2.6300	2.6310	2.6330	2.6358	2.6394	2.6436	2.6485	2.6536	2.6589	2.6648	2.6713	2.6783	2.6856	2.6933
0.40	2.8082	2.6946	2.6591	2.6439	2.6362	2.6327	2.6318	2.6328	2.6348	2.6376	2.6412	2.6454	2.6503	2.6555	2.6609	2.6669	2.6734	2.6803	2.6875	2.6951
0.35	2.8111	2.6978	2.6617	2.6463	2.6384	2.6348	2.6339	2.6349	2.6368	2.6395	2.6431	2.6475	2.6524	2.6576	2.6631	2.6691	2.6757	2.6825	2.6896	2.6971
0.30	2.8148	2.7015	2.6647	2.6489	2.6408	2.6372	2.6362	2.6372	2.6389	2.6416	2.6453	2.6496	2.6546	2.6598	2.6655	2.6715	2.6779	2.6849	2.6920	2.6995
0.25	2.8192	2.7057	2.6680	2.6517	2.6434	2.6397	2.6387	2.6396	2.6414	2.6440	2.6476	2.6519	2.6568	2.6621	2.6679	2.6741	2.6805	2.6874	2.6946	2.7020
0.20	2.8243	2.7105	2.6716	2.6548	2.6462	2.6424	2.6414	2.6421	2.6440	2.6466	2.6502	2.6546	2.6594	2.6647	2.6706	2.6769	2.6833	2.6902	2.6974	2.7048
0.15	2.8301	2.7156	2.6756	2.6582	2.6493	2.6453	2.6442	2.6448	2.6467	2.6494	2.6530	2.6574	2.6622	2.6676	2.6735	2.6797	2.6864	2.6932	2.7005	2.7080
0.10	2.8365	2.7212	2.6800	2.6618	2.6526	2.6485	2.6473	2.6478	2.6497	2.6524	2.6559	2.6603	2.6653	2.6707	2.6766	2.6828	2.6895	2.6965	2.7037	2.7113
0.05	2.8436	2.7273	2.6847	2.6658	2.6561	2.6518	2.6506	2.6510	2.6527	2.6554	2.6591	2.6635	2.6685	2.6740	2.6799	2.6862	2.6929	2.6999	2.7072	2.7147

Appendix C: Description of the ratings ELO_b and ELO_g

In this section we provide a brief description of the ratings ELO_b and ELO_g as defined by the authors of the ratings (Hvattum & Arntzen, 2010).

C.1. Description of ELO_b

Let l_o^H and l_o^A be the ratings, at the start of a match, of the home and away teams respectively. The ELO ratings assume that the home and away teams should score γ^H and γ^A respectively where:

$$\gamma^H = \frac{1}{1+c^{\left(\frac{\left(l_o^A - l_o^H\right)}{d}\right)}} \qquad \text{and} \qquad \gamma^A = 1 - \gamma^H = \frac{1}{1+c^{\left(\frac{\left(l_o^H - l_o^A\right)}{d}\right)}}$$

and the parameters c and d serve only to set a scale of the ratings. The authors suggest that we use $c = 10$ and $d = 400$ (but alternative values of c and d give identical rating systems). Assuming that the score for the home team follows:

$$a^H = \begin{cases} 1, & if\ the\ home\ team\ won, \\ 0.5, & if\ the\ match\ was\ drawn, or \\ 0, & otherwise \end{cases}$$

Then the actual score for the away team is $a^A = 1 - a^H$. At the end of the match, the revised ELO rating for the home team is (the away team's l_1^A is calculated in the same way):

$$l_1^H = l_o^H + k(a^H - \gamma^H)$$

with $k = 20$ as a suitable parameter value.

C.2. Description of ELO_g

The ELO_g rating is a variant of ELO_b above, in an attempt to also consider score difference, with the difference that k is replaced by the expression:

$$k = k_0(1+\delta)^{\lambda}$$

where δ is the absolute goal difference, and assuming $k_0 > 0$ and $\lambda > 0$ as fixed parameters; suggesting $k_0 = 10$ and $\lambda = 1$ as suitable parameter values.

Appendix D: Optimised $[k]$ and $[k_0,\ \lambda]$ values for the ratings ELO_b and ELO_g

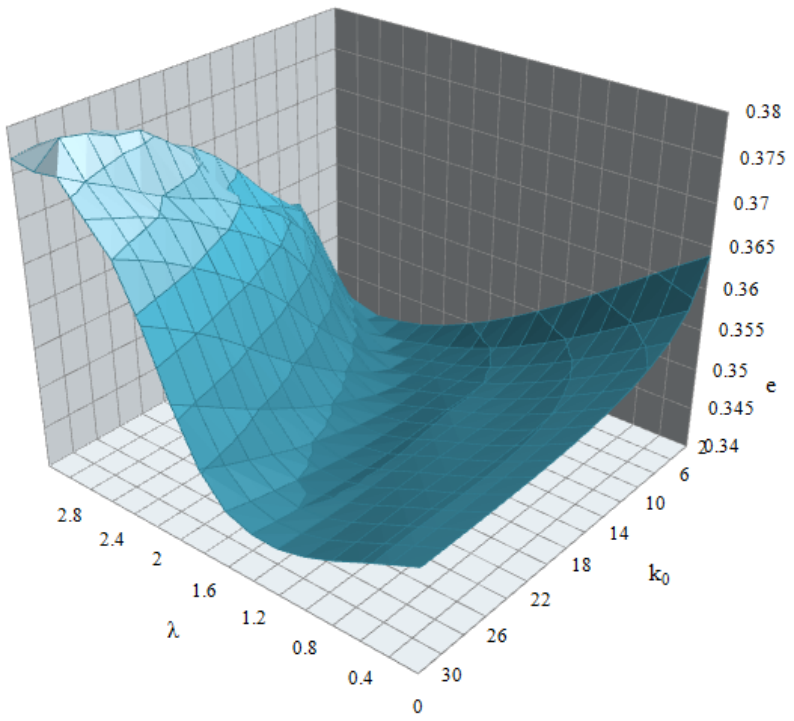

Figure D.1. Optimised k_0 and λ values for ELO_g. Minimum squared error of expected goal difference observed when $k_0 = 2$ and $\lambda = 2.8$, where $e = 0.3405$

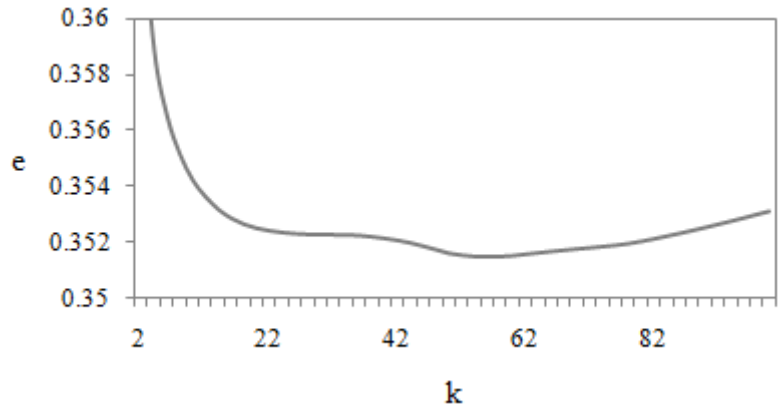

Figure D.2. Optimised k-value for ELO_b. Minimum squared error of expected goal difference observed when $k = 56$, where $e = 0.3514$.

REFERENCES

[1] Baio, G., & Blangiardo, M. (2010). Bayesian hierarchical model for the prediction of football results. Journal of Applied Statistics, 37:2, 253-264.

[2] Buchner, A., Dubitzky, W., Schuster, A., Lopes, P., O'Doneghue, P., Hughes, J., et al. (1997). Corporate evidential decision making in performance prediction domains. Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI'97). Brown University, Providence, Rhode Island, USA.

[3] Clarke, S. R., & Norman, J. M. (1995). Home ground advantage of individual clubs in English soccer. The Statistician, 44, 509–521.

[4] Crowder, M., Dixon, M., Ledford, A., & Robinson, M. (2002). Dynamic modelling and prediction of English football league matches for betting. The Statistician, 51, 157– 168.

[5] Constantinou, A. C., Fenton, N. E., & Neil, M. (2012a). pi-football: A Bayesian network model for forecasting Association Football match outcomes. Knowledge-Based Systems, 322-339. Draft available at: http://www.constantinou.info/downloads/papers/pi-model11.pdf

[6] Constantinou, A. C., Fenton, N. E., & Neil, M. (2012b). Profiting from an Inefficient Association Football Gambling Market: Prediction, Risk and Uncertainty Using Bayesian Networks. Under Review. Draft available at: http://www.constantinou.info/downloads/papers/pi-model12.pdf

[7] Constantinou, A. C., & Fenton, N. E. (2012). Evidence of an (intended) inefficient Association Football gambling market. Under Review. Draft available at: http://constantinou.info/downloads/papers/evidenceofinefficiency.pdf

[8] Dixon, M., & Coles, S. (1997). Modelling association football scores and inefficiencies in the football betting market. Applied Statistics, 46, 265-80.

[9] Dixon, M., & Pope, P. (2004). The value of statistical forecasts in the UK association football betting market. International Journal of Forecasting, 20, 697-711.

[10] Dunning, E. G., & Joseph A. Maguire, R. E. (1993). The Sports Process: A Comparative and Developmental Approach. Champaign: Human Kinetics, p. 129.

[11] Dunning, E. (1999). Sport Matters: Sociological Studies of Sport, Violence and Civilisation. London: Routledge.

[12] Elo, A. E. (1978). The rating of chess players, past and present. New York: Arco Publishing.

[13] Fenton, N. E. & Neil, M. (2012). Risk Assessment and Decision Analysis with Bayesian Networks. London: Chapman and Hall.

[14] FIFA. (2012). FIFA. Retrieved March 27, 2012, from FIFA/Coca-Cola World Ranking Procedure: http://www.fifa.com/worldranking/procedureandschedule/menprocedure/index.html

[15] Football-Data. (2012). Football-Data.co.uk. Retrieved August 2, 2012, from Football Results, Statistics & Soccer Betting Odds Data: http://www.football-data.co.uk/englandm.php

[16] Forrest, D., Goddard, J., & Simmons, R. (2005). Odds-setters as forecasters: The case of English football. International Journal of Forecasting, 21, 551-564.

[17] Goddard, J. (2005). Regression models for forecasting goals and match results in association football. International Journal of Forecasting, 21, 331-340.

[18] Goddard, J., & Asimakopoulos, I. (2004). Forecasting Football Results and the Efficiency of Fixed-odds Betting. Journal of Forecasting, 23, 51-66.

[19] Halicioglu, F. (2005a). Can we predict the outcome of the international football tournaments? : the case of Euro 2000. Doğuş Üniversitesi Dergisi, 6, 112-122.

[20] Halicioglu, F. (2005b). Forecasting the Professional Team Sporting Events: Evidence from Euro 2000 and 2004 Football Tournaments. 5th International Conference on Sports and Culture: Economic, Management and Marketing Aspects. Athens, Greece, pp. 30-31.

[21] Harville, D. A. (1977) The use of linear-model methodology to rate high school or college football teams. Journal of American Statistical Association, 72, 278-289.

[22] Hirotsu, N., & Wright, M. (2003). An evaluation of characteristics of teams in association football by using a Markov process model. The Statistician, 52: 4, 591-602.

[23] Hvattum, L. M., & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. International Journal of Forecasting, 26, 460-470.

[24] Joseph, A., Fenton, N., & Neil, M. (2006). Predicting football results using Bayesian nets and other machine learning techniques. Knowledge-Based Systems, 7, 544-553.

[25] Karlis, D., & Ntzoufras, I. (2000). On modelling soccer data. Student, 229–244.

[26] Karlis, D., & Ntzoufras, I. (2003). Analysis of sports data by using bivariate Poisson models. The Statistician, 52: 3, 381-393.

[27] Knorr-Held, L. (1997) Hierarchical Modelling of Discrete Longitudinal Data, Applications of Markov Chain Monte Carlo. Munich: Utz.

[28] Knorr-Held, L. (2000). Dynamic Rating of Sports Teams. The Statistician. 49(2), 261-276.

[29] Koning, R. (2000). Balance in competition in Dutch soccer. The Statistician, 49: 3, 419-431.

[30] Koning, R., H., Koolhaas, M., Renes, G. & Ridder, G. (2003). A simulation model for football championships. European Journal of Operational Research. 148: 268-276.

[31] Kuonen, D. (1996). Statistical Models for Knock-Out Soccer Tournaments. Technical Report, Department of Mathematics, Ècole Polytechnique Federale de Lausanne.

[32] Kuypers, T. (2000). Information and efficiency: an empirical study of a fixed odds betting market. Applied Economics, 32, 1353-1363.

[33] Lee, A. J. (1997) Modeling scores in the Premier League: is Manchester United really the best? Chance, 10, 15–19.

[34] Leitner, C., Zeileis, A., & Hornik, K. (2010). Forecasting sports tournaments by ratings of (prob)abilities: A comparison for the EURO 2008. International Journal of Forecasting, 26, 471-481.

[35] Maher, M. J. (1982). Modelling association football scores. Statististica Neerlandica, 36, 109–118.

[36] Min, B., Kim, J., Choe, C., Eom, H., & McKay, R. B. (2008). A compound framework for sports results prediction: A football case study. Knowledge-Based Systems, 21, 551- 562.

[37] Mueller, F. O., Cantu, R. C., & Camp, S. P. (1996). Catastrophic Injuries in High School and College Sports. Champaign: Human Kinetics, 57.

[38] Murali, V. (2011, October 28). Bleacher Report. Retrieved March 28, 2012, from World Football: 40 Biggest Scandals in Football History: http://bleacherreport.com/articles/909932-world-football-40-biggest-scandals-in-football-history

[39] Poulter, D. R. (2009). Home advantage and player nationality in international club football. Journal of Sports Sciences , 27(8): 797-805.

[40] Reid, D. A. & Nixon, M. S. (2011). Using Comparative Human Descriptions for Soft Biometrics. International Joint Conference on Biometrics (IJCB), 2011.

[41] Rotshtein, A., Posner, M., & Rakytyanska, A. (2005). Football predictions based on a fuzzy model with genetic and neural tuning. Cybernetics and Systems Analysis, 41: 4, 619- 630.

[42] Rue, H., & Salvesen, O. (2000). Prediction and retrospective analysis of soccer matches in a league. The Statistician, 3, 339-418.

[43] Tsakonas, A., Dounias, G., Shtovba, S., & Vivdyuk, V. (2002). Soft computing-based result prediction of football games. The First International Conference on Inductive Modelling (ICIM'2002). Lviv, Ukraine.