# Supplementary material for “Predicting match outcomes in association football using team ratings and player ratings”

**Halvard Arntzen [1] and Lars Magnus Hvattum [1]**

[1] Faculty of Logistics, Molde University College, Norway

**Address for correspondence:** Lars Magnus Hvattum, Faculty of Logistics, Molde University College, P. O. Box 2110, N-6402 Molde, Norway.

**E-mail:** `hvattum@himolde.no`.

**Phone:** (+47) 71 21 42 23.

**Fax:** NA.

**Abstract:** This online supplement presents details of calculations for team ratings and for individual player ratings in association football.

**Key words:** Elo rating; plus-minus rating; ranking method; soccer; sports

# A Ratings

The main manuscript describes a comparison of team ratings and player ratings in association football, testing their performance as covariates in models used to derive predictions for match outcomes. Section  describes the how the team ratings used are calculated. Then, Section  outlines how individual player ratings are calculated.

## A.1 Team ratings

The team ratings described here are based on the rating system of Elo (1978) for chess players, with adjustments for association football by Hvattum and Arntzen (2010). The ratings are dynamic, and updated after each match. Let $r_i$ and $r_j$ be the ratings of team $i$ and team $j$ before a match. The outcome of the match is denoted by $m_{ij}$, with $m_{ij} = 1$ if team $i$ wins, $m_{ij} = 0.5$ if the game is drawn, and $m_{ij} = 0$ if team $j$ wins. Let $g_i$ and $g_j$ be the number of goals scored by team $i$ and team $j$, respectively. Four parameters, $b = 1$, $c = 10$, $d = 400$, and $k = 10$, control the update of ratings after games, and their values were tuned by Hvattum and Arntzen (2010) to obtain the best possible predictions for match outcomes when using the resulting ratings as the only covariate in an OLR model. The updated rating $r_i'$ for team $i$ after the match can now be written as

$$r_i' = r_i + k\left(1 + |g_i - g_j|\right)^b \left(m_{ij} - \frac{1}{1 + c^{(r_j - r_i)/d}}\right).$$

The rating update formula above requires that both teams involved in the match have a prior rating. This is not the case for the first matches in the data set, when the teams do not have established ratings. Therefore, initial ratings are calculated as the performance ratings for the first two seasons of data. Two or more seasons are used to be enable the ratings to reflect differences in strength between the different divisions of the league system. Performance ratings are calculated by first setting an initial guess $r_i = r_i^{INIT} = 0$ for the rating of each team $i$. Then, the rating update formula is used for all the matches of the two first seasons, yielding an estimated performance rating $r_i^{END} = r_i$ for each team. If, for all teams, $r_i^{INIT} \approx r_i^{END}$, then $r_i = r_i^{END}$ is the rating performance of team $i$ over the two seasons of matches, which is then taken as the established rating at the beginning of the third season. While $r_i^{INIT} \not\approx r_i^{END}$ for any $i$, the initial guess is updated by setting $r_i^{INIT} = r_i^{END}$ and the procedure is repeated.

Another case where a team does not have an established rating occurs when a team is promoted into the lowest division covered by the data. In this situation, results for the team are recorded until a given number of matches has been played. Then, the first established rating for the team is calculated similarly as for the teams in the first two seasons: an initial guess $r_i = r_i^{INIT} = 0$ is made, then the rating update formula is applied for each match, but only for the new team $i$ and not for the opposing teams. This results in a performance rating $r_i^{END} = r_i$. As long as $r_i^{END} \not\approx r_i^{INIT}$, the initial guess is updated by setting $r_i^{INIT} = r_i^{END}$ and the process is repeated.

When using the Elo-ratings to make predictions of match outcomes, a covariate is calculated as the difference in rating between the home team $i$ and the away team $j$, $x_{Elo} = r_i - r_j$. The effect of home advantage is taken into account by the prediction model, and not in the Elo calculations. As most teams play an equal number of home matches and away matches, in an alternating fashion, the effect of compensating for home field advantage in the Elo calculations is negligible.

## A.2 Player ratings

The player ratings examined in this paper are based on the plus-minus ratings developed for association football by Sæbø and Hvattum (2015, 2019) and later refined by Pantuso and Hvattum (2019) and analyzed by Gelade and Hvattum (2020). The starting point is to split all matches into segments where the set of players on the pitch is unchanged. In other words, segments are created whenever a substitution is made or a player leaves the pitch following a red card. An unconstrained quadratic program is then formulated, based on the idea of using ridge regression to estimate a multiple linear regression model, to determine player ratings, while controlling for league differences, the home field advantage, the effect of red cards, and the effect of age on player performances.

To describe the model, let $M$ be a set of matches, with each match $m \in M$ divided into segments $s \in S_m$. The duration of segment $s$ is denoted by $d(m,s)$ . Let $t^{MATCH}(m)$ denote the time that match $m$ is played, and let $T$ be the time at which ratings are calculated, as illustrated in Figure 1. Each match $m$ belongs to a given competition type, denoted by $c(m)$, which in this study corresponds to either the Premier League, the Championship, League One, or League Two.

Let $h = h(m)$ and $a = a(m)$ be the two teams involved in match $m$, where $h$ is the home team if the match is not played on neutral ground. Define $g^S(m,s)$ as the goal difference in favour of $h$ at the beginning of the segment, and $g^E(m,s)$ as the goal difference at the end of the segment. The change of the goal difference in favour of $h$ in segment $s$ of match $m$ then becomes $g(m,s) = g^E(m,s) - g^S(m,s)$.

The sets of players of the home team and away team that appear on the pitch for segment $s$ are denoted by $P_{msh}$ and $P_{msa}$, respectively. For $n = 1, \ldots, 4$, define $r(m, s, n) = 1$ if team $h$ has received $n$ or more red cards before segment $s$ and team $a$ has not, $r(m, s, n) = -1$ if team $a$ has received $n$ or more red cards and team $h$ has not, and $r(m, s, n) = 0$ otherwise. In the case that a team has made all its available substitutions and a player on the team becomes injured and must leave the pitch, the situation is handled in the same way as a red card.

We let $C$ be the set of competition types and define $C_p$ as the set of competition types in which player $p$ has participated. Each player $p$ is associated to a set $P_p^{SIM}$ of players that are considered to be similar. This set is based on which players have appeared together on the same team for the most minutes of playing time. The time of the last match where players $p$ and $p'$ appeared together is denoted by $t^{SIM}(p, p')$ (see Figure 1).

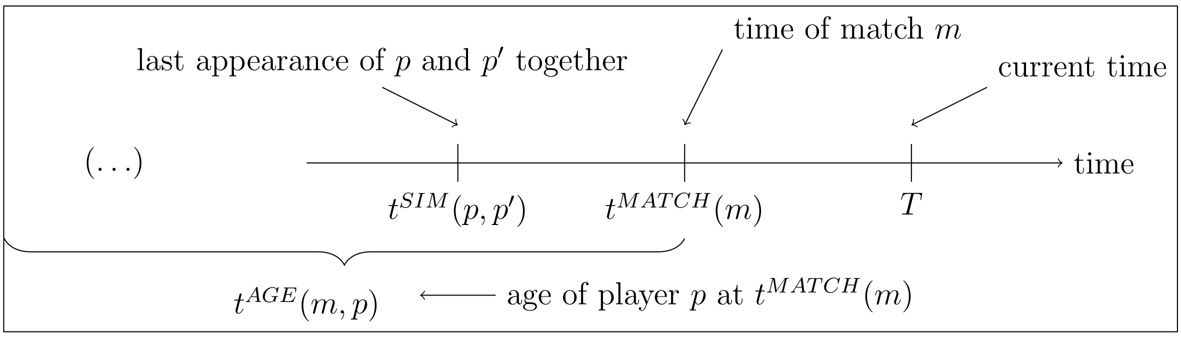


Figure 1: Illustration of the parameters referring to the time of events.

The quality of players is assumed to depend on their age, allowing the model to capture their typical improvement in early years as well as their decline when getting older. Define $t = t^{AGE}(m, p)$ as the age of player $p$ at the time of match $m$, as illustrated in Figure 1. A set of integer age values $Y = \{y^{MIN}, \ldots, y^{MAX}\}$ is defined, and for a given match and player, the exact age of the player is expressed as a convex combination of the nearest two ages in $Y$. Thus, we use weights $u_y(t)$ defined as

$$u_y(t) = 1 \text{ if } t < y^{MIN} = y \text{ or } t > y^{MAX} = y \ .$$

Moreover, if $y \leq t \leq y+1$ we let

$$u_y(t) = (y+1) - t, \quad u_{y+1}(t) = t - y \ .$$

In all other instances we let $u_y(t) = 0$. For example, if $t = 19.25$, we get $u_{19}(t) = 0.75, u_{20}(t) = 0.25$. This allows a precise representation of player ages, while limiting the number of age effect parameters to be estimated.

The following parameters are defined to control the behavior of the model: $\lambda$ is a regularization factor, $\rho_1$ is a discount factor for older observations, $\rho_2$ and $\rho_3$ are parameters regarding the importance of the duration of a segment, and $\rho_4$ is a factor for the importance of a segment depending on the goal difference at the start of the segment. The parameter $w^{AGE}$ is a weight to balance the importance of the age factors when considering similarity of players. Finally, $w^{SIM}$ is a weight that controls the extent to which ratings of players with few minutes played are shrunk towards 0 or towards the ratings of similar players. The values of the parameters in this work are set as used by Pantuso and Hvattum (2019), resulting in $\lambda = 16.0$, $\rho_1 = 0.1$, $\rho_2 = 300.0$, $\rho_3 = 300.0$, $\rho_4 = 2.5$, $w^{AGE} = 0.35$, and $w^{SIM} = 0.85$. In addition, the set of similar players is limited in size to $|P_p^{SIM}| \leq 35$, and players' ages are confined by $y^{MIN} = 16$ and $y^{MAX} = 42$.

The variables used in the quadratic program can be stated as follows: The base rating of player $p$ is denoted by $\beta_p$. The value of the home advantage in competition $c(m)$

is represented by $\beta^{H}_{c(m)}$. For a given age $y \in Y$, the age effect is denoted $\beta^{AGE}_{y}$. For a player of age $t$ the age effect is

$$u_y(t)\beta_y^{AGE} + u_{y+1}(t)\beta_{y+1}^{AGE} \ ,$$

provided $y \leq t \leq y+1$. The influence of red cards are captured by the variables $\beta_n^{HOMERED}$ and $\beta_n^{AWAYRED}$, for $n = 1, \ldots, 4$. An adjustment for each competition $c \in C$ is given by the variables $\beta_c^{COMP}$. Let $\beta$ be the vector of all the decision variables of the model and let $V$ be the set of indices of $\beta$.

The model to calculate plus-minus ratings can now be stated as

$$\min_{\beta} Z = \sum_{m \in M} \sum_{s \in S_m} \left( w(m,s) \left( f^{LHS}(m,s) - f^{RHS}(m,s) \right) \right)^2 + \lambda \sum_{j \in V} \left( f^{REG}(\beta_j) \right)^2$$

where

$$\begin{aligned}
w(m,s) =& w^{TIME}(m,s) w^{DURATION}(m,s) w^{GOALS}(m,s), \\
w^{TIME}(m,s) =& \exp\left(\rho_1 \left(T - t^{MATCH}(m)\right)\right), \\
w^{DURATION}(m,s) =& \frac{d(m,s) + \rho_2}{\rho_3}, \\
w^{GOALS}(m,s) =& \begin{cases} \rho_4 & \text{if } |g^S(m,s)| \geq 2 \text{ and } |g^E(m,s)| \geq 2 \\ 1 & \text{otherwise,} \end{cases} \\
f^{RHS}(m,s) =& g(m,s), \\
f^{LHS}(m,s) =& \frac{d(m,s)}{90} \left( \frac{11}{|P_{msh}|} \sum_{p \in P_{msh}} f^{PLAYER}(m,s,p) \right. , \\
& \left. - \frac{11}{|P_{msa}|} \sum_{p \in P_{msa}} f^{PLAYER}(m,s,p) + f^{SEGMENT}(m,s) + f^{MATCH}(m) \right),
\end{aligned}$$

$$f^{SEGMENT}(m,s) = \begin{cases} \sum_{n=1}^{4} r(m,s,n)\beta_n^{HOMERED} & \text{if } \sum_{n=1}^{4} r(m,s,n) \geq 0 \\ \sum_{n=1}^{4} r(m,s,n)\beta_n^{AWAYRED} & \text{if } \sum_{n=1}^{4} r(m,s,n) < 0, \end{cases}$$

$$f^{MATCH}(m) = \begin{cases} \beta_{c(m)}^{H} & \text{if team } h(m) \text{ has home advantage} \\ 0 & \text{otherwise,} \end{cases}$$

$$f^{PLAYER}(m,s,p) = \beta_p + \sum_{y \in Y} u_y(t^{AGE}(m,p))\beta_y^{AGE} + \frac{1}{|C_p|}\sum_{c \in C_p} \beta_c^{COMP},$$

$$f^{REG}(\beta_y^{AGE}) = \begin{cases} \left(\beta_y^{AGE} - (\beta_{y-1}^{AGE} + \beta_{y+1}^{AGE})/2\right) & \text{if } y \in Y \setminus \{y^{MIN}, y^{MAX}\} \\ 0 & \text{if } y \in \{y^{MIN}, y^{MAX}\}, \end{cases}$$

$$f^{REG}(\beta_c^{COMP}) = \beta_c^{COMP},$$

$$f^{REG}(\beta_c^{H}) = \beta_c^{H},$$

$$f^{REG}(\beta_p) = \left( f^{AUX}(p,T,1) - \frac{w^{SIM}}{|P_p^{SIM}|} \sum_{p' \in P_p^{SIM}} f^{AUX}\left(p', t^{SIM}(p,p'), w^{AGE}\right) \right),$$

$$f^{AUX}(p,t,w) = \beta_p + w\sum_{y \in Y} u_y(t)\beta_y^{AGE} + \frac{1}{|C_p|}\sum_{c \in C_p} \beta_c^{COMP}.$$

When this program has been solved, the estimated rating for player $p$ at time $T$ is $f^{AUX}(p,T,1)$. To make predictions of match outcomes, a covariate $x_{PM}$ is calculated based on the plus-minus ratings before a match by considering the average rating of players in the starting line-up of the home team minus the average rating of players in the starting line-up of the away team. Players with no prior match appearances are removed from this calculation, as they have no estimated plus-minus rating.

## References

Elo, A. (1978). *The Rating of Chessplayers, Past and Present.* Arco Publishing, New

York.

Gelade, G. and Hvattum, L. (2020). On the relationship between +/− ratings and event-level performance statistics. *Journal of Sports Analytics*. Forthcoming.

Hvattum, L. and Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting*, **26**, 460–470.

Pantuso, G. and Hvattum, L. (2019). Maximizing performance with an eye on the finances: a chance-constrained model for football transfer market decisions. arXiv:1911.04689.

Sæbø, O. and Hvattum, L. (2015). Evaluating the efficiency of the association football transfer market using regression based player ratings. In *NIK: Norsk Informatikkonferanse*. Bibsys Open Journal Systems. 12 pages.

Sæbø, O. and Hvattum, L. (2019). Modelling the financial contribution of soccer players to their clubs. *Journal of Sports Analytics*, **5**, 23–34.