Formulating An Elo Rating for Major League Soccer

by Ryan Anderson

ryan@f2f2s.com

Last Edited: March 25th, 2011

Revision History

August 2nd, 2008 – initial version published to http://mlselo.f2f2s.com

July 22nd, 2009 – due to what I felt were unacceptably small variations due to recent results, the K-factors were changed from:

- **30** for an MLS Cup final
- **20** for a playoff game
- **10** for a regular season game

to the new values:

- **45** for an MLS Cup final
- **30** for a playoff game
- **15** for a regular season game

All previous results have been re-calculated to use the new factors.

March 25th, 2011 – After some conversations with other people who have developed ranking systems, I have decided to move to an iterative approach for calculating the Elo Ratings. The ratings are cycled through 15 times as that is the point that the all iterative changes showed up as less than 1% of 1% in testing.

In order to keep the ratings in the same 1350-1625 span, the K-factors were updated to:

- **1.50** for an MLS Cup final
- **1.25** for a playoff game
- **1.00** for a regular season game

For a full explanation of the updates see this blog post:
http://f2f2s.com/2011/03/21/serious-updates-to-the-elo-ratings/

For a comparison of the before and after rankings see this one:
http://f2f2s.com/2011/03/23/differences-in-the-two-elo-systems/

The source code for the python script is included in an appendix in case you're interested in copying it for something.

I. Introduction

The Elo rating system was developed by Arpad Elo in the early to mid 1900's as a way to measure the relative skill of chess players. It works under the assumption that the performance of each participant is a normally distributed random variable; meaning that while a player may perform better or worse from game-to-game, there is a baseline average he can expect to compete at. The ratings are developed so that the average player's rating is 1500.

Both the player's and his opponent's rating are used in the calculation of his next rating. Therefore, beating a player with a higher rating will cause a player to gain more points than beating a player with a lower rating.

The basic formula is below:

$$R'_A = R_A + K(S_A - E_A).$$

R'_A represents the new rating
R_A represents the initial rating before the most recent competition
K is a factor used to change the significance of the result based on the setting of the competition.
S_A is the number of points scored by the player
E_A is the number of points the player was expected to score

The formula for calculating E_A is as follows:

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}}.$$

R_B is the rating of the opponent and R_A is the player's rating

In the methodology section I'll cover how I altered these equations for use in the MLS Elo rankings.

Initially developed for the United States Chess Federation, a similar methodology has been used for rankings in Go, online computer games, college football, college basketball, Major League Baseball, and international soccer. It was the last one that helped spark my interest in this and led to this little experiment. The ultimate goal remains to develop some way to measure the quality of teams between leagues that play each other occasionally such as the English Premier League, Spain's La Liga, France's Ligue Un, and Italy's Serie A. This does not address that desire, but is a starting building block that may lead to that some time in the future.

II. Considerations

Soccer does not necessarily translate well into an Elo rating system. There are new factors that aren't originally factored in to the system that must be addressed as they offer information to the relative strength and weakness of teams. I'll list the potential issues I came up with here and I welcome anyone to bring up other issues by contacting me using the e-mail address provided.

Home Field Advantage: As in every sporting event there is an intrinsic advantage to being able to play at home. Whatever combination it may be of the support of the crowds, the ability to sleep in your own bed, familiarity with the playing surface, and other similar factors, this is clearly an issue that needed addressing. Nate Silver of Baseball Prospectus addresses this home field advantage by adding 25 points to the rating of the home team in his calculations. The soccer website http://www.eloratings.net adds 100 points to the home team involved in an international meeting. For the purposes of my calculations, I will be adding 85 points to the home team's rating. If you are interested in the math leading to this number you can read about it in the Methodology section.

Season-to-Season Carryover: A soccer team does not stay the same from year to year, so we cannot simply keep the teams' ratings the same to begin year 2 as they were at the end of year 1. However, discounting the previous year's work would be doing a grave disservice to teams who have demonstrated their superiority. In the end I followed Nate Silver's less than glamorous but extremely effective solution of halving the teams' difference against the baseline 1500 ranking. In other words, if DC United ends a season with a rating of 1550 they begin the next season with a rating of 1525. If New York Red Bulls end a season with a rating of 1420 they begin the next season with a rating of 1460. The Elo ratings are designed to be self-correcting, meaning that it is easier for a lower rated team to gain points than lose points and vice-versa for high rated teams, so potential bias should be squeezed out quickly.

Post-Season Play: This is an interesting discussion. I believe it is clear that post season results should factor into a team's rating, but the question is by how much. To my knowledge Nate Silver does not use any increased multiplier for post-season games. However, baseball teams also have a much longer season meaning there is much more data for an accurate measurement to be made. Eloratings.net utilizes the following system for multipliers based on the level of the match:

- **60** for World Cup finals;
- **50** for continental championship finals and major intercontinental tournaments;
- **40** for World Cup and continental qualifiers and major tournaments;
- **30** for all other tournaments;
- **20** for friendly matches.

Note that this would be considered the K-factor from the equation in the

Introduction. I'll be following that idea with by using different multipliers for regular season games, playoff games, and an MLS Cup final. I did some basic simulations with different multipliers to see what would put the velocity of movement at its most intuitively correct and this is what I came up with:

- **1.50** for an MLS Cup final
- **1.25** for a playoff game
- **1.00** for a regular season game

Please note than an MLS Cup final has no home team as it is played in a neutral site, and therefore the 85 point home field advantage is omitted.

Also, if these numbers seem small it's because they are. Instead of doing a single calculation on just each new game, I do 15 loops over the entire season up to that point. I find that reduces the noise in the rankings. See the blog posts linked in the revision sections for more information on this.

The Changes in MLS Game Structure: MLS has only been around since 1996 but has gone through significant rule changes in that time. For the first 4 years of its existence, there were no ties in MLS. Games that ended in a tie were decided by a hockey-style shootout where a player got the ball 35 yards out from goal and had 5 seconds to score on the keeper. That was abolished in the 2000 season and a tie game after regulation went into two 15-minute golden-goal overtimes, meaning that the first goal scored would end the game. If there were no goals scored then the game ended in a tie. Since the 2004 season the rules have basically mirrored the standard set of rules used for nearly every league and the World Cup, meaning to added extra time at all for regular season games.

In my attempt to deal with these changes I have developed the following set of rules to define what values to use for calculations.

1996-1999: All games that went to shootouts are counted as ties with the extra "goal" omitted.

2000-2003: Games ending with a golden goal in added extra time are considered a standard result where one team wins and another loses. No goals are subtracted or half goals awarded.

Playoffs: Any goal scored in added extra time is given full weight, so the scoring team would win with a goal differential of 1. Victories derived from a shootout are awarded to the winning team with a 0.5 goal differential.

Goal Differential: The goal differential between two teams is another way to determine their relative strength. If New England Revolution beats Chicago Fire by 5 goals in one game and Chicago beats New England by 1 goal in a second game we as observers can conclude that New England should be rated as the better team.

Because of this I will also be using a multiplier calculated by the goal differential. In practice this will be multiplied to the K-factor, and will now be referred to as the G-factor. In other words, my equation now looks more like:

$$R'_A = R_A + KG(S_A - E_A)$$

I did absolutely nothing to develop this formula and instead will use the same one employed by eloratings.net, with a minor addition for the shootout victories.

- 1 goal difference: G = 1
- 2 goal difference: G = 1.5
- 3 goal difference: G = 1.75
- 4+ goal difference: G = 1.75 + (N-3)/8 where N is the goal difference
- 0.5 goal difference: G = 0.75

Wins/Losses/Ties: Soccer does not really have a value of points scored in the sense required for these calculations. Yes, a win earns a team 3 points and a tie earns a team 1 point, but we need something with a value of 0 to 1 since the expected values we are working with will be percentages. If we used 3 and 1 as our values both teams would be able to gain in rating since they would have a positive number when subtracting their expected win percentage (E_A) from the actual result(S_A) of 1. Instead I will be following the lead of the standard chess Elo rating system and assigning a win a value of 1, a tie a value of 0.5 and a loss a value of 0.

Interleague Competitions: An MLS season does not consist solely of league matches. There is the US Open Cup which features teams from the USL-1, USL-2, PDL and USASA leagues in addition to MLS. There is also what was formerly called the CONCACAF Champions Cup and will now be transformed under a slightly different format to the CONCACAF Champions League starting in the fall of 2008. This competition features club teams from every country in the CONCACAF region. Also, there are invitational tournaments such as Superliga, where the MLS teams play against Mexican Football League teams.

There are two main ways to account for this. The obvious one is to continue with the same methodology used for MLS and calculate Elo ratings for every team in every league. Unfortunately that would require an enormous amount of data collection, and I don't believe there would be enough games between leagues to accurately measure them against each other. If we assume that MLS is better than USL-1 we would expect MLS to have a higher average rating. However, since not every USL-1 team has played every MLS team and the sample available is relatively minor, we would not be able to accurately portray the difference in skill level.

The second option is to calculate the total record for one league against another and then assume that every team from the league with lesser data has the same rating. We can calculate that rating using the leagues' historical winning

percentages against other leagues in the same way that I calculate the home field advantage. MLS would keep its base rating of 1500 and the others would be calculated accordingly. However, even with that I don't expect that we would have enough information to make an accurate decision. Also, assuming that every team in a certain league plays with equal quality brings a lot of room for error into the calculations.

In the end, I decided to exclude all games from US Open Cup and international club tournaments for the purpose of these ratings. I understand the limitations this places on the model, but I feel that the results will still be satisfactory for the purposes of rating MLS teams against each other. Perhaps in a situation like European soccer where the teams from top leagues play each other more often and results can be found relatively easily we can address this bias. However, I don't see any way to deal with these matches for MLS teams.

Expansion/Contraction: MLS started as a 10 team league, grew to 12 teams in 1998, contracted to 10 in 2002, expanded back to 12 in 2005, 13 in 2007, and 14 in 2008. Dealing with the added teams is easy, as I simply applied them ratings of 1500 and allowed the formula to do its work. It does have the unfortunate effect of giving inaccurate information for the first few iterations, but we just have to take that into account when looking at the numbers.

Contraction is a little trickier problem. Hopefully the MLS will remain a viable league and there will be no more contraction issues to deal with; however, since it has happened once, I had to come up with some way of dealing with it. What I did was, after the 2001 season I subtracted the Elo ratings for Miami Fusion and Tampa Bay Mutiny (1525.20 and 1409.01 respectively) from the total sum of Elo ratings (18,000). I then found what percentage of the total rating points left each team had and applied that percentage to the new total of 15,000 Elo ratings points. The new total is 15,000 so as to maintain a 1500 point average rating.

III. Methodology

This section will address how I collected the data used in these ratings, show the calculations for a home-field advantage of 85 ratings-points, and give the final formulas for calculation.

First, the data collection. All of the game results except for one come from the Historical Stats page on MLSnet.com. One game from the mid-2000s is missing from the page, but I found the result on ESPN's MLS page. The information was copied directly from the website and pasted into a Microsoft Excel page where I did all of the parsing using basic formulas and macros. After that, every individual entry was rechecked against the web page for accuracy. I believe that the results used in the calculations are 100% accurate but if you believe otherwise for any reason please contact me at the e-mail address provided.

From 1996 through 2007 MLS teams played 2,073 regular season games. In those games the home team won 1,028 times, the away team won 565 times, and there were 480 ties. These correspond to percentages of 49.59%, 27.26% and 23.15% respectively. From that we can calculate the average number of points per game using a standard weighted average as follows:

$$PPG = 2*(0.2315) + 3*(0.4959+0.2726)$$
$$PPG = 2.77$$

Now there is a similar calculation to determine the average points per game for the home team and away team respectively.

$$PPG_H = 1*(0.2315) + 3*(0.4959)$$
$$PPG_H = 1.72$$

$$PPG_A = 1*(0.2315) + 3*(0.2726)$$
$$PPG_A = 1.05$$

With simple division we can see that home team earns, on average, 62% of the points awarded per game, while the away team earns 38%. In other words, the home team has a 62/38 advantage.

Given that value of 0.62 we can plug that in for the E_A value in second equation in the Introduction section, and solve for $R_B - R_A$, which will be our home field advantage factor.

$$0.62 = \frac{1}{1 + 10^{(R_B - R_A)/400}}$$

$$10^{(R_B - R_A)/400} = \frac{1}{0.62} - 1 = 0.6129$$

Take the log(base 10) of both sides

$$(R_B - R_A)/400 = -0.21261$$

$$(R_B - R_A) = -85$$

In other words, in order for a team to earn 62% of the points over an infinite sample, it would need an 85 point advantage in Elo rating.

So now we can say that the full calculations for the Elo ratings are as follows:

$$R'_A = R_A + KG(S_A - E_A)$$

Where R_A is the team's previous Elo rating

K equals:

- **1.50** for an MLS Cup final
- **1.25** for a playoff game
- **1.00** for a regular season game

G equals:

- **0.75** for a 0.5 goal difference
- **1.00** for a 1 goal difference
- **1.50** for a 2 goal difference
- **1.75** for a 3 goal difference
- **1.75 + (N-3)/8** for an N goal difference where N >= 4

SA equals:

- **1** for a victory
- **0.5** for a tie
- **0** for a loss

$$E_A = \frac{1}{1 + 10^{(R_B - R_A)/400}}.$$

and E_A equals:

with an 85 point bonus added to the home team's rating.

IV. References/Thanks

This idea was entirely due to JaredL from the twoplustwo forums (http://forumserver.twoplustwo.com) . He is the one who introduced me to the idea of an Elo rating for soccer, and club teams in particular.

Thank you to Nate Silver (conveniently enough, also a poster on 2+2) for his excellent writeup (http://www.baseballprospectus.com/article.php?articleid=5247) explaining the steps he went through to formulate an Elo rating for Major League Baseball teams. When I saw the he had asked the same set of questions I had with regards to how to formulate the equations I felt like I had to be doing something right. Anyone interested in sports statistics would do well to sign up for BP, and anyone interested in election statistics should check out Nate's writing at http://www.fivethirtyeight.com where he discusses politics and polling from the same analytical perspective.

What turned out to be a big influence on both Jared and Nate (and therefore me) was the site http://www.eloratings.net, which does an Elo rating for international soccer teams. Thankfully, he was kind enough to post his methodology for everyone to see, which allowed me to piggyback on his ideas. I am very grateful for that.

Thanks also go to whoever keeps up the stats the MLS webpage (http://web.mlsnet.com/mls/history/index.jsp). I do think that there is a lot of improvement that can be done to the page from formatting the results the same year to year, to making everything easier to navigate, but for the purposes of this exercise I was able to find well over 99% of my data easily.

Again, if you have any questions or concerns, or would just like to talk about the information in here, please feel free to email me at ryan@f2f2s.com

A subset of the information gathered from this process will be posted at http://mlselo.f2f2s.com for you to look at. Feel free to check back from time-to-time as I will be adding to it.

V. Appendix A: Python Script

Below is the entire Python script that I use for updating the iterative Elo Ratings. I run this in Python 2.7, so you may have to change some syntax if you are on Python 3.

This script takes in 2 text files (one of the games, one of the starting Elo Ratings at the beginning of the season) and writes the final Elo Rating to a new file.

```
from sys import argv

teams = {}

gamefile = argv[1]

infile = open(gamefile, 'r')
games = [p.rstrip().split(',') for p in infile.readlines()]
infile.close()

elofile = argv[2]

infile = open(elofile, 'r')
ranks = [p.rstrip().split(',') for p in infile.readlines()]
infile.close()

for rank in ranks:
    teams[rank[0]] = float(rank[1])

def elo_weight(elo1, elo2, hfa=85):
    awgt = 1/(10**(((elo1+hfa)-elo2)/400.0)+1)
    hwgt = 1/(10**((elo2-(elo1+hfa))/400.0)+1)
    return hwgt, awgt

def g_fact(a, b):
    diff = abs(a-b)
    if diff == 0 or diff == 1:
        return 1
    elif diff == 0.5:
        return 0.75
    elif diff == 1.5:
        return 1.25
    elif diff == 2:
        return 1.5
    elif diff == 3:
        return 1.75
    return 1.75+(diff-3)/8
```

```
def k_fact(typ):
    if typ == 'r': return 1
    if typ == 'p': return 1.25
    return 1.5

def game_res(c, d):
    if c > d:
        return 1, 0
    elif d > c:
        return 0, 1
    return 0.5, 0.5

def calc_elo(h_elo, a_elo, h_s, a_s, setting):
    kfac = k_fact(setting)
    if setting == 'f':
        weight = elo_weight(h_elo, a_elo, 0)
    else:
        weight = elo_weight(h_elo, a_elo)
    gfac = g_fact(h_s, a_s)
    res = game_res(h_s, a_s)
    a2 = a_elo + (kfac * gfac * (res[1] - weight[1]))
    h2 = h_elo + (kfac * gfac * (res[0] - weight[0]))
    return h2, a2

for i in range(15):
    for game in games:
        output = calc_elo(teams[game[1]], teams[game[0]], float(game[3]),
float(game[2]), game[4])
        teams[game[0]] = output[1]
        teams[game[1]] = output[0]

outfile1 = open('endelo.txt', 'w')
outfile2 = open('newelo.txt', 'w')

for team in teams:
    print round(teams[team],4), "-", team
    outfile1.write("%s - %s\n" %(round(teams[team],4), team))
    outfile2.write("%s,%s\n" %(team, round((teams[team]-1500)/2 + 1500,4)))

outfile1.close()
outfile2.close()
```