1. Overview

The ECF grading team have introduced changes to the grading system. There has been a step change in all grades, which is not discussed here and a change in method for grading juniors.

This note is an investigation into whether the method for grading juniors is

- inflationary or deflationary
- stretches or shrinks the grading range

This paper attempts to investigate the complexities of the system with some heavy maths which remain difficult to interpret, but can give an underpin to further theoretical work.

2. Executive Summary

This version is work in progress. However the emerging results appear to confirm comments made on the 2009 grades. The ECF grading system is summarised on the ECF website and further information and comment has been available on the grading forum linked to that site.

The system is inherently unstable with both deflationary and inflationary features. The new treatment of juniors makes the system less stable. It appears unreliable in ranking players with few or no games against rated players.

Tentative conclusions are that the system is not inherently inflationary or deflationary. It stretches junior and new entrants grades.

Any weaknesses in the approach to unrated players feeds back into the rating of regular players.

Regrettably there is a lot of mathematics first beginning with some definitions

3. Definitions

The mathematics ignore the 40 point rule and complications arising from treating categories below A as carrying forward older ratings. Although these are discussed. The final junior adjustment (+5 or +10) is ignored.

Define r as the number of players being rated. currently, r is of the order of 10,000.

Each player r plays m(r)rated games

P^1 is an rx1 matrix of the final grades of those r players

N is an rxr matrix where element n(i,j) is the proportion of the graded games in the period by player i played against player j. This entry will invariably be 0 or 1/m(i).

S is a rxr matrix is the grading result adjustment, where element s(i,j)is [(100*Σ scores by i against j)-(50*games between i and j)] all divided by number of graded games by player i.

1 is a variable length column vector which in effect sums each row of the prior matrix.

Let p^* be the reference grades (explained below)

Then the general formula for ECF rating is:

$$P^1 = N\ p^* + S1$$

The general idea is that gradings are assumed not to change fast so that the reference grade is that of the previous year. However there are new entrants who have no prior grade and juniors who are being treated as new entrants, because many of them appear to be improving .

4. Model Development

To investigate further the matrix is partitioned between regulars, R - non juniors with a prior grade and others, U, who are treated differently.

Hence the general formula is partitioned and sorted in a natural indexation as:

$$\begin{bmatrix} P^R \\ P^U \end{bmatrix} = \begin{bmatrix} N^{RR} & N^{RU} \\ N^{UR} & N^{UU} \end{bmatrix} \begin{bmatrix} p^0 \\ p^* \end{bmatrix} + \begin{bmatrix} S^{RR} & S^{RU} \\ S^{UR} & S^{NU} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

where p^0 is the available prior year grade and p^* can be interpreted from a description of the process used by the graders as the solution to the following equation.

$$[P^*] = [N^{UR}\ N^{UU}] \begin{bmatrix} p^0 \\ p^* \end{bmatrix} + S^{UR}1 + S^{UU}1$$

The natural solution (it is not necessarily unique, but it is a solution that makes sense)for p^* is:

$$P^* = [I- N^{UU}]^{-1} \{ N^{UR} p^0 + S^{UR} 1 + S^{UU} 1\} \qquad (1$$

I is the identity matrix. Equation (1) is far to large to solve by matrix manipulation, so, loosely, the grader's method is to take a first guess by first calculating ignoring the last two entries in the final bracket, then add the second item and solve, then add the third item and carry out iterations until the result is stable.

This vector can be plugged into the partitioned matrix giving two matrix equations for the separate methods.

$$P^R = N^{RR} p^0 + N^{RU} p^* + S^{RR} 1 + S^{RU} 1 \qquad (2A$$

$$P^U = N^{UR} p^0 + N^{UU} p^* + S^{UR} 1 + S^{UU} 1 \qquad (2B$$

Equation 2A can be re-arranged as follows:

$$P^R = \{N^{RR} p^0 + S^{RR} 1 \}+ \{N^{RU} p^* + S^{RU} 1\}$$

The first brackets encompasses the general equation in a closed system. The second brackets cover method adjustments.

Substituting Equation 1 into equation 2A gives:

$$P^R = N^{RR} p^0 + S^{RR} 1 + N^{RU} [I- N^{UU}]^{-1} N^{UR} p^0 + N^{RU} [I- N^{UU}]^{-1} S^{UR} 1 + N^{RU} [I- N^{UU}]^{-1} S^{UU} 1 + S^{RU} 1$$

This equation is close to unfathomable in this state.

5. Situation with no games between ungraded players

To grapple with the issues, I'm going to start by assuming there are no games between ungraded players. Hence N^{UU} & S^{UU} are null, so $[I- N^{UU}]^{-1}$ becomes the identity matrix. Next plug a simplified equation 1 into 2A:

$$P^R = N^{RR} p^0 + S^{RR} 1 + N^{RU} N^{UR} p^0 + N^{RU} S^{UR} 1 + S^{RU} 1 \qquad (3$$

What do these matrices look like?

- Normally players will play at each other at most once in each grading period so elements $n^{RR}(i.)$ are generally of the form 0 with m elements 1/m, where m is the number of rated games played by player i. The matrix will be closely symmetric, where this term means if element e(i,j) is non zero then so is e(j,i). All diagonal elements of N^{RR} must be zero.

- Again diagonal elements of the matrix S^{RR} must be zero and it is "closely symmetric", but with negative elements. Elements s^{RR} will be more sparsely non zero, since draws and matched wins and losses will score zero along with entries for unplayed opponents
- Comparing N^{RU} and N^{UR} : if n(i,j) >0 then n(j,i) >0 but will differ if the players have played a different number of games.
- When considering elements s(i,j) similar considerations apply, but if s(i,j) >0, then s(j,i) <0.

($N^{RU}.N^{UR}$) and ($N^{RU}.S^{UR}$) are closely symmetric rxr matrices. There effects are to add back grades/scores as compensation for those lost when deflating N^{RR} by using all games played rather than games played against rated players as a denominator.

6. Implications

I now move onto some assumptions.

1. I assume generally that rated players play more than unrated players
2. Generally both rated and unrated players will play a higher proportion of their games against rated players. The exception would be some juniors (remember these are now treated as "unrated" in this method); juniors would also play a significant proportion of games amongst fellow youngsters. Juniors will be considered later.
3. Evidence suggests rated players are overall stronger than unrated players. This would need reconsideration if the influx of new overseas players into the system were to change. Overseas players tend to be equivalent to strong rated players rather than beginners.
4. Rated players are sorted broadly in decreasing order of strength. For example based on prior year's grading, although the exact method is not important.

If 2A is re-arranged as follows:

$P^R = (N^{RR} + N^{RU}.N^{UR})\ p^0 + (S^{RR} + N^{RU}.S^{UR})1 + S^{RU}1,$

which I am going to describe in a loose fashion as player's grade is a weighted average of opponent's strength +/- a performance adjustment.

To simplify, assume nobody plays anyone else more than once. If multiple games are played between a rated and an unrated player then denominator of weightings increases by the square of the multiple, whereas the numerator increases by the excess games. Take a look at each matrix.

Consider the elements of row i.
N^{RR} : The general formula in a closed state, but the sum of the row will now only add up to the proportion of games played against rated players, not 1.

$N^{RU}.N^{UR}$: This product is a bit of a black box. The diagonal elements are (typically) the highest elements. The diagonal element is the proportion of games against unrated players by player i, multiplied by the sum over all those unrated players k, of the inverse of all games played by k. Non diagonal elements (i,j) are non zero where common unrated opponents are played by both i and j. Non diagonal elements will generally be zero, but clusters will occur where there are relationships, for instance clubs or leagues; even then the elements should be relatively small. The sum of each row of $N^{RR} + N^{RU}.N^{UR}$ is one.

S^{RR} : Similarly to N^{RR}, elements have the denominator based on all games rather than just those between the rated players.

$N^{RU}.S^{UR}$: This product is a more complex version of the matrix $N^{RU}.N^{UR}$ discussed above. Since S^{UR} is generally negative this item tends to work in the opposite direction to the expected benefits a rated player will get from S^{RU}.

S^{RU} :Typically more positive the lower the value of i. The effects are more diluted if the proportion of games against rated players is high.

So what is this weighted average of opponent's strength? Ignoring clustering and repeat matches. The weighted average is:

The proportion of games against rated opponents multiplied by the average grade of those opponents plus the proportion of games played against unrated players multiplied by a weighted average of all rated players. The biggest influence on the weighted average is the player's own prior grade.

The performance adjustment is constructed in a similar way.

7. The effect of games between unrated players.

We need to return to equation 2A and consider games between unrated players. With no such games the matrix $[I- N^{UU}]^{-1}$ has no influence.

Re-introducing this matrix inflates the system. This offsets the effect of using all games as a denominator rather than only those where opponents have a grade.

8. Investigating the grading of previous ungraded players

For completeness one should review the remaining grades. The implications of this section are important in any critical appraisal of the treatment of juniors. Returning to equation 2B:

$$P^U = N^{UR} p^0 + N^{UU} p^* + S^{UR} 1 + S^{UU} 1 \qquad \text{(2B}$$

Plugging in equation 1 and re-arranging gives a solution:

$$P^U = \{I + N^{UU} [I- N^{UU}]^{-1}\}\{N^{UR} p^0 + S^{UR} 1 + S^{UU} 1\}$$

The second bracket can be described as a proportion of the performance against rated players plus relative performance against unrated players. The proportion in each element i is the ratio of games against rated to total games played by i. This ratio is then grossed up by the effect of the first bracket.

At present I have failed to find a generalisation of the net effects of the whole equation. Investigations suggest that:

- There is no net effect if each player has the same proportion of unrated games (including nil)
- The method cannot be expected to get unrated players in the right ranking where there is only a thin link to games with rated players. Essentially the formula depends on the central limit theorem working on the numbers of games by unrated players to rated players. A different proposition to total games played being controlled by the theorem.

- Any weaknesses in the ratings for unrated players will feedback into those of rated players.

9. Other issues

The introduction of the 40 point rule would introduce stretch into the system since when grades are recalculated after p^* is fixed. For instance, points will be added where a win has reduced a grade score and subtracted where a lose has increased a grade, because of the grading difference.

The introduction of categories will tend to lag the method effects for those players not category A. calculations.

The change in average grades between years will be affected by the relative grade of those leaving the system. Measures of inflation or deflation need careful treatment of leavers. In this note these issues are considered from the point of view of survivors. An implicit assumption in the system are in some universal sense their performance does not change.