International Journal of Physical Sciences Research
Vol.1, No.2, pp.1- 13, August 2017
Published by European Centre for Research Training and Development UK (www.eajournals.org)

STOCHASTIC MODELING/GAME THEORY ANALYSIS OF SCORELINE

Etaga H.O, Umeokeke E.T, Nwosu C.R, Etaga N. C, Umeh E. U, Awopeju B, Eriobu N, Okoye V.C., Omoruyi F. A.,

Department of Statistics, Nnamdi Azikiwe University, Awka, Nigeria

ABSTRACT: *Prediction of a football match result arouses interest from different points of view; for different people, Hence the need for this work which aims at analysing the scores of the four top English clubs to enable prediction of future outcome of matches to be made in a more scientific manner. From the analysis of the scoreline of the top four EPL clubs; Manchester United (M.U), Chelsea (C), Arsenal (A), and Manchester City (M.C) from (2002-2015) using Game theory and Stochastic modelling, Chelsea emerged the best team with a selection probability of 0.41 while Manchester United also emerged second best with a selection probability of 0.37. From the steady state transition probability matrix, for all the six possible pairs of the four clubs shows that the probability of M.U wining C at home is 0.44 while C wins M.U at home with probability 0.67 depicting C as the stronger club. Similarly M.U is stronger than A, with a 0.71 winning probability as against 0.25 winning probability for A, while M.U and M.C appears to be equally matched with 0.48 and 0.49 probability of winning. C against A reveals a probability of 0.58 and 0.25 for A vs C. while C vs M.C showed C to have an upper hand with a 0.71 probability of winning and 0.44 for M.C vs C. Finally A vs M.C gives the two teams 0.53 and 0.42 winning probabilities. Thus, the two most viable clubs out of the four clubs are Manchester United and Chelsea. Using the four step TPM we also predicted the 2015/2016 matches to obtain their various probabilities given the previous game.*

KEYWORDS: Game theory, Stochastic modelling, English Premier League, Football, Operations Research, Stochastic Modelling, Prediction.

INTRODUCTION

Football is the most popular game all over the world; in Europe and South America it is the dominant spectator sport. People find interest in soccer for various reasons and at different levels, with a clear dominance for the males, **Reep and Benjamin (1968)** came to the conclusion that "chance does dominate the game" while **Hennessy (1969)** is of the opinion that only chance was involved. **Hill (1974**) argued that anyone who had ever watched a football match could reach the conclusion that the game was either all skill or all chance. He justified his opinion by calculating the correlation between the expert opinions and the final league tables result, concluding that even though chance was involved, there was also a significant amount of talent affecting the final outcome of the match. However the first real model to predict football scores was put forward by **Maher (1982)**.From his model, he obtained that the goals scored by two opposing teams in some particular match are drawn from independent Poisson distributions. Whilst introducing the home advantage factor, he assigned each team with a pair of fixed parameters (α and β) such that the model would simply consist in combining the respective attacking and defensive parameters of the opposing teams. **Nelson Mandela** (1992), avers that sports has the power to change the world. **Lee (1997)** relied on this model to simulate the English Premier League season 1995/1996 for around 1000 times, and investigated whether Manchester United really deserved to emerge victorious.

Steinmetz (2000) obtained a United States par tent for a statistical model (similar to a regression tree) that can be used for the prediction of future outcomes based on qualitative measures only, using historical parameters related to past performance, experience of team personnel, time of the season at which a game occurs, and the Las Vegas betting line. **Ferda (2009)** provides a statistical measurement to predict the possible winners of international football tournaments with specific reference to the Euro 2008 football tournaments. **Blundell (2009)** found that numerical models can facilitate the prediction of result in sporting events. The options within these models rely on data related to the competitors. He used a logistic regression model to predict the result of American football matches and incorporates data of the 2 teams' previous results, novel options like stadia size and the distance the away team has to travel.

In addition **Baio & Blangiardo (2010**) propose a Bayesian hierarchal model to address the prediction process by estimating the characteristics that bring a team to lose or win a particular game and predicting the score. They used the data of Italian serial A championship 2007/2008 to test the model adequacy. Also , almost all football clubs attracts a huge number of fans; emotional admires who are indirectly involved in the outcome of every game played by the club and often will want to predict the outcome of every game before its occurrence. Statistical modelling of the outcome of different game especially football has thus become a popular area of research. Out of the many football leagues available, the English Premier League (EPL), which is the world's most watched league with a TV audience of 4.7 billion people was chosen for this study, which aims at finding not only the best two EPL clubs but also finding the steady state probability of winning their home and away matches using their scoreline from 2002-2015 matches

The dataset used for this study is the results (scoreline) of 13 seasons (years) of the top four EPL clubs from the 2002-03 seasons to the 2014-15 seasons. Since each of the 4 teams play all other teams twice per season (home and away). This translates to 12 games. For each game, our dataset includes the home team, the away team, the score difference, the winner, and the number of goals for each team. By convention, club ("A vs. B") implies that A is the home team and B is the away team. The main aim of this research is to find the scoreline trend of the four clubs from 2002-2015, more specifically to obtain the best two clubs. This research aim would achieve by; Firstly, to obtain the best two clubs using the score of the game for each season; Secondly, to obtain the best two overall clubs for the 13 seasons; thirdly to obtain the estimate of the Transition probability matrix describing the game for the entire season; fourthly, to obtain the probability that a club wins and wins again after 4 plays and fifthly, to obtain the stationary matrix which describes the game.

DATA ANALYSIS AND RESULTS

Game theory Analysis of 2013/2014 season scoreline

In order to find the best club for each season, using the four clubs as strategies. A competitive situation where two individuals (MR A & MR B) select the best club using their scorelines is created so as to maximize profit and the other to minimize loss.

The top four EPL clubs are represented with the following abbreviations shown below

Manchester United (Man U)

Chelsea (C)

Arsenal (A) and

Manchester City (Man City)

Home Match 2013/2014 season

Table1, shows the difference in goals of the four clubs in their home matches. The value 1 in the diagonal matrix is for completeness in the season because a club cannot play themselves. The first row entry Man U had a draw with Chelsea, won Arsenal by 1 goal, lost 3 goals to Man City. In the second row entry Chelsea won Man U by 2 goals, won Arsenal by 6 goals, won Man City by 1 goal. In the third row entry, Arsenal had a draw with Man U, had a draw with Chelsea, and also had a draw with Man City. In the fourth row entry, Man City won Man U by 3 goals, lost 1 goal to Chelsea, won Arsenal with 3 goals.

Using the Minimax, & Maximin criteria, we observe that maximin = minimax = value of the game = 1. This implies the existence of a **saddle point** at (C,M.C), Thus the **optimal strategy** (the best clubs) is for MR A to select Chelsea and MR B to select Man City.

Table 1: Home Match 2013/2014 Payoff Matrix

		MR B				
		M	C	A	M.C	Minimum
MR A	M	1	0	1	-3	-3
	C	2	1	6	1	1
	A	0	0	1	0	0
	M.C	3	-1	3	1	-1
	Maximum	3	1	6	1	

Maximin

Minimax

Away Match 2013/2014 season

Using the same approach the payoff matrix for the away match 2013/2014 is obtained and shown in Table 2

Using the Minimax and Maximin Criteria, we observe that there is **no saddle point**; hence **linear programming method** is used to obtain the solution to the payoff matrix.

Table 2 Away Match 2013/2014 Payoff Matrix

		MR B				
		M	C	A	M.C	Minimum
MR A	M	1	-2	0	-3	-3
	C	0	1	0	1	0
	A	-1	-6	1	-3	-6

maximin

	M.C	3	-1	0	1	-1
	Maximum	3	1	1	$\boxed{1}$	

minimax

Table 2 is now converted to a payoff matrix with non negative entries by adding a constant number 6 to all the elements of the payoff matrix to give Table 3

Table 3: The modified payoff matrix with probabilities p_i & q_j

		MR B				
		M	C	A	M.C	Probability
MR A	M	7	4	6	3	p_1
	C	6	7	6	6	p_2
	A	5	0	7	3	p_3
	M.C	9	5	6	7	p_4
	Probability	q_1	q_2	q_3	q_4	

Where
p_i,i=1,2,3,4 and $q_{j,}$ j=1,2,3,4 are the strategy selection probabilities for both MR A and MR B respectively. Solving for MR B, the expected loss for MR B becomes system 1

$$7q_1 + 4q_2 + 6q_3 + 3q_4 \leq v$$

$$6q_1 + 7q_2 + 6q_3 + 6q_4 \leq v \qquad (1)$$

$$5q_1 \qquad + 7q_3 + 3q_4 \leq v$$

$$9q_1 + 5q_2 + 6q_3 + 7q_4 \leq v$$

$$q_1, q_2, q_3, q_4 \geq 0$$

Dividing system 1 by v gives system 2

Maximize $z_q \left(= {}^1/_v\right) = y_1 + y_2 + y_3 + y_4$

Subject to

$$7y_1 + 4y_2 + 6y_3 + 3y_4 \leq 1$$

$$6y_1 + 7y_2 + 6y_3 + 6y_4 \leq 1 \qquad (2)$$

$$5y_1 \qquad + 7y_3 + 3y_4 \leq 1$$

$$9y_1 + 5y_2 + 6y_3 + 7y_4 \leq 1$$

$$y_1, y_2, y_3, y_4 \geq 0$$

Where $y_1 = \frac{q_1}{v}, y_2 = \frac{q_2}{v}, y_3 = \frac{q_3}{v}, y_4 = \frac{q_4}{v}$

Converting system 2 to standard form gives system 3

Maximize $z_q = y_1 + y_2 + y_3 + y_4 + 0s_1 + 0s_2 + 0s_3 + 0s_4$

Subject to

$$7y_1 + 4y_2 + 6y_3 + 3y_4 + s_1 = 1$$

$$6y_1 + 7y_2 + 6y_3 + 7y_4 + s_2 = 1 \qquad 3$$

$$5y_1 \quad + 7y_3 + 3y_4 + s_3 = 1$$

$$9y_1 + 5y_2 + 6y_3 + 7y_4 + s_4 = 1$$

$$y_1, y_2, y_3, y_4\ s_1, s_2, s_3, s_4 \geq 0$$

See Table 4 for initial tableau for system 3

Table 4: The initial tableau

Basic	y_1	y_2	y_3	y_4	s_1	s_2	s_3	s_4	solution
Z	-1	-1	-1	-1	0	0	0	0	0
s_1	7	4	6	3	1	0	0	0	1
s_2	6	7	6	7	0	1	0	0	1
s_3	5	0	7	3	0	0	1	0	1
s_4	9	5	6	7	0	0	0	1	1

Using the statistical software TORA we solve the linear programming problem in Table 4 using simplex method to obtain the solution stated in Table 5

Table 5: Final iteration tableau

Basic	y_1	y_2	y_3	y_4	s_1	s_2	s_3	s_4	solution
Z	0	0	0	0.08	0	0.13	0.02	0.01	0.16
s_1	0	0	0	-2.73	1	-0.22	-0.25	-0.50	0.04
y_2	0	1	0	0.54	0	0.19	-0.11	-0.07	0.02
y_3	0	0	1	0.17	0	0.15	0.19	-0.20	0.13
y_1	1	0	0	0.36	0	-0.20	-0.07	-0.29	0.01

Table5 is the final tableau with the optimal solution for MR B given as

$$y_1 = 0.01; y_2 = 0.02;\ y_3 = 0.13, y_4 = 0, z = 0.16$$

Value of the game for the modified matrix is $v = \frac{1}{z} = 6.25$

These solution values are now converted back into the original variables:

$$y_1 = \frac{q_1}{v}, then\ q_1 = y_1 \times v = 0.01 \times 6.25 = 0.06$$

$$y_2 = \frac{q_2}{v}, then\ q_2 = y_2 \times v = 0.02 \times 6.25 = 0.12$$

$$y_3 = \frac{q_3}{v}, then\ q_3 = y_3 \times v = 0.13 \times 6.25 = 0.81$$

$$y_4 = \frac{q_4}{v}, then\ q_4 = y_4 \times v = 0.00 \times 6.25 = 0.00$$

The optimal strategies for MR A are obtained from the z row under the slack variables in Table 5

$$x_1 = 0.00;\ x_2 = 0.13, x_3 = 0.02, x_4 = 0.01, z = 0.16$$

$x_i = \frac{p_i}{v}$

$$x_1 = \frac{p_1}{v}, then\ p_1 = x_1 \times v = 0.00 \times 6.25 = 0.00$$

$$x_2 = \frac{p_2}{v}, then\ p_2 = x_2 \times v = 0.13 \times 6.25 = 0.81$$

$$x_3 = \frac{p_3}{v}, then\ p_3 = x_3 \times v = 0.02 \times 6.25 = 0.12$$

$$x_4 = \frac{p_4}{v}, then\ p_4 = x_4 \times v = 0.01 \times 6.25 = 0.06$$

Hence, the selection probabilities MR A and MR B respectively are:

(0.00, 0.81, 0.12. 0.06) & (0.06, 0.12, 0.81, 0.00) as shown in Table 6

While the expected value of the game for the problem is

$$v = {}^{1}/_{z} - k$$

$$v = 6.25 - 6$$

$$v = 0.25$$

Table 6: probabilities of using strategies by MR A and MR B

		MR B				
		M	C	A	M.C	Probability
MR A	M	7	4	6	3	0.00
	C	6	7	6	6	0.81
	A	5	0	7	3	0.12
	M.C	9	5	6	7	0.06
	Probability	0.06	0.12	0.81	0.00	

From Table 6, C(Chelsea) has the highest selection probability of 0.81 by MR A to maximize his profit. While A(Arsenal) has the highest selection probability of 0.81 of minimizing loss by MR B.

Using the same steps & methods, we obtain the best possible choices for MR A and MR B from 2002-2015 as shown in Table 7

International Journal of Physical Sciences Research
Vol.1, No.2, pp.1- 13, August 2017
Published by European Centre for Research Training and Development UK (www.eajournals.org)

Table 7: Choice selection of MR A and MR B

	MR A				MR B	
Season	Home	Away			Home	Away
2002/2003	M.U	M.U			M.U	M.U
2003/2004	A	M.C			M.U, C	C,A,M.C
2004/2005	C	M.C, C			M.C	C
2005/2006	C	C			C	M.U
2006/2007	C,A	C			M.U,C	M.U,A,M.C
2007/2008	C	M.U			M.U, A	C
2008/2009	M.U	M.U			M.U	M.U,A
2009/2010	M.U, M.C	M.U			M.U, C	M.U
2010/2011	M.U, C	C, A			M.U	C,A
2011/2012	M.C	M.U			M.U, M.C	M.C
2012/2013	M.U, M.C	M.C			M.U, M.C	M.U,C,M.C
2013/2014	C	C			C, M.C	A
2014/2015	C	C			C	A

Summarizing Table 7 to obtain the number of times a club is chosen by the competitors for the Home and Away matches gives Table 8

Table 8 Number of possible selection

	Number of occurrence of each club								
	M.U		C		A		M.C		
	Home	Away	Home	Away	Home	Away	Home	Away	Total
MR A	5	5	7	6	2	1	3	3	32
MR B	9	6	6	5	1	6	4	4	41

Thus the selection probabilities for the four clubs for the home & away matches are as shown in Table 9

Table 9 Probabilities selection for the four clubs by the two competitors

	M.U (H +A)	C (H+A)	A (H+A)	M.C(H+A)	Maximum
MR A	0.31	0.41	0.09	0.19	**0.41**
MR B	0.37	0.27	0.17	0.20	**0.37**

From Table 9, C(Chelsea) has the highest selection probability by MR A while M.U(Manchester United) has the highest selection probability by MR B.

Analysis using Transition Probability Matrix

The first step in the development of the transition probability matrix is to obtain the matrix of flow. From the combined scores for the 13 seasons, the matrix of flow was obtained for all possible pair of clubs for the entire season for "Home" and "Away" matches. Where

w represents a "win" L represents a "loss" D represents a "draw".

The matrix of flow, with the corresponding transition probabilities matrix (TPM) and its eigenvalues are given as follows

Matrices of flow, TPM and the eigenvalues for the Home and Away matches

HOME

Manchester United vs Chelsea

WDLW DWWLWWLDD

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 0 & 2 & 1 \\ 2 & 2 & 2 \\ 1 & 1 & 1 \end{pmatrix} \end{array}\begin{array}{c} \\ 3 \\ 6 \\ 3 \end{array} \quad P_{ij} = \begin{bmatrix} 0 & 0.7 & 0.3 \\ 0.3 & 0.3 & 0.3 \\ 0.3 & 0.3 & 0.3 \end{bmatrix}$$

Eigenvalues $=[0.9, -0.3, 0]$

AWAY

Chelsea vs Manchester United

DWWWDWDWWDLWW

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 4 & 3 \\ 1 & 3 & 0 \end{pmatrix} \end{array}\begin{array}{c} \\ 1 \\ 7 \\ 4 \end{array} \quad P_{ij} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0.6 & 0.4 \\ 0.3 & 0.7 & 0 \end{bmatrix}$$

Eigenvalues $= [0.8, -0.5, 0.3]$

Manchester United vs Arsenal

WDWWLWDWWWWWD

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 5 & 3 \\ 0 & 2 & 0 \end{pmatrix} \end{array}\begin{array}{c} \\ 1 \\ 9 \\ 2 \end{array} \quad P_{ij} = \begin{bmatrix} 0 & 1 & 0 \\ 0.1 & 0.6 & 0.3 \\ 0 & 1 & 0 \end{bmatrix}$$

Eigenvalues $=[1, -0.4, 0]$

Arsenal vs Manchester United

DDLDWDWLWLDDL

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 2 & 2 & 2 \end{pmatrix} \end{array}\begin{array}{c} \\ 3 \\ 3 \\ 6 \end{array} \quad P_{ij} = \begin{bmatrix} 0 & 0.3 & 0.7 \\ 0.7 & 0 & 0.3 \\ 0.3 & 0.3 & 0.3 \end{bmatrix}$$

Eigenvalues $= [0.9, -0.3, -0.3]$

Manchester United vs Manchester City

Manchester City vs Manchester United

DWDDWLWWWLLLW

WWLWLWLLDWLWW

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 1 \\ 0 & 2 & 1 \end{pmatrix} \end{array}\begin{array}{c} \\ 4 \\ 5 \\ 3 \end{array} \quad P_{ij} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.4 & 0.4 & 0.2 \\ 0 & 0.7 & 0.3 \end{bmatrix}$$

Eigenvalues $=[1, -0.2, 0.4]$

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 1 & 3 & 1 \\ 4 & 2 & 0 \\ 0 & 1 & 0 \end{pmatrix} \end{array}\begin{array}{c} \\ 5 \\ 6 \\ 1 \end{array} \quad P_{ij} = \begin{bmatrix} 0.2 & 0.6 & 0.2 \\ 0.7 & 0.3 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Eigenvalues $= [1, -0.25, -0.25]$

Chelsea vs Arsenal **Arsenal vs Chelsea**

DLDWDWLWWLWWW

WWDLDWLLWDLDD

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 0 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 2 & 0 \end{pmatrix} \end{array}\begin{array}{c} \\ 3 \\ 6 \\ 3 \end{array} \quad P_{ij} = \begin{bmatrix} 0 & 0.7 & 0.3 \\ 0.3 & 0.5 & 0.2 \\ 0.3 & 0.7 & 0 \end{bmatrix}$$

Eigenvalues $=[1, -0.3, -0.2]$

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix} \end{array}\begin{array}{c} \\ 4 \\ 4 \\ 4 \end{array} \quad P_{ij} = \begin{bmatrix} 0.25 & 0.25 & 0.5 \\ 0.25 & 0.25 & 0.5 \\ 0.5 & 0.25 & 0.25 \end{bmatrix}$$

Eigenvalues $= [1.04, -0.15, -0.15]$

Chelsea vs Manchester City **Manchester City vs Chelsea**

WWDWWWWLWWD **LLWLLLLWWWWLD**

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 5 & 3 \\ 0 & 2 & 0 \end{pmatrix} \end{array}\begin{array}{c} \\ 1 \\ 9 \\ 2 \end{array} \quad P_{ij} = \begin{bmatrix} 0 & 0 & 0 \\ 0.1 & 0.6 & 0.3 \\ 0 & 1 & 0 \end{bmatrix}$$

Eigenvalues $=[1, -0.4, 0]$

Matrix of flow **TPM**

$$\begin{array}{c} \\ L \\ W \\ D \end{array}\begin{array}{c} \begin{array}{ccc} L & W & D \end{array} \\ \begin{pmatrix} 4 & 2 & 1 \\ 2 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{array}\begin{array}{c} \\ 7 \\ 5 \\ 0 \end{array} \quad P_{ij} = \begin{bmatrix} 0.6 & 0.3 & 0.1 \\ 0.4 & 0.6 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Eigenvalues $= [0.9, 0, 0.25]$

Arsenal vs Manchester City **Manchester City vs Arsenal**

WWDWWWWDDWLDD **LLLLWLWWLWDWL**

Matrix of flow **TPM**

$$\begin{array}{c|ccc|c} & \boldsymbol{L} & \boldsymbol{W} & \boldsymbol{D} & \\ \hline \boldsymbol{L} & 0 & 0 & 1 & 1 \\ \boldsymbol{W} & 1 & 4 & 2 & 7 \\ \boldsymbol{D} & 0 & 2 & 2 & 4 \end{array} \quad \boldsymbol{P_{ij}} = \begin{bmatrix} 0 & 0 & 1 \\ 0.1 & 0.6 & 0.3 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

Eigenvalues $=[1, 0.05,\ 0.05]$

Matrix of flow **TPM**

$$\begin{array}{c|ccc|c} & \boldsymbol{L} & \boldsymbol{W} & \boldsymbol{D} & \\ \hline \boldsymbol{L} & 3 & 3 & 0 & 7 \\ \boldsymbol{W} & 3 & 1 & 1 & 5 \\ \boldsymbol{D} & 0 & 1 & 0 & 0 \end{array} \quad \boldsymbol{P_{ij}} = \begin{bmatrix} 0.5 & 0.5 & 0 \\ 0.6 & 0.2 & 0.2 \\ 0 & 1 & 0 \end{bmatrix}$$

Eigenvalues $= [1, -0.5, 0.2]$

Four Step TPM for Home &Away Match

The probability that a club will win his opponents given that he won the previous game (p_{ww}^4) are now calculated for the home & away matches for the six possible combinations as shown below;

Manchester United vs Chelsea

$$p^4 = \begin{bmatrix} 0.198 & 0.324 & 0.252 \\ 0.173 & 0.311 & 0.232 \\ 0.173 & 0.311 & 0.232 \end{bmatrix}$$

$p_{ww}^4 = 0.311$

Chelsea vs Manchester United

$$p^4 = \begin{bmatrix} 0.072 & 0.672 & 0.256 \\ 0.077 & 0.654 & 0.269 \\ 0.086 & 0.662 & 0.251 \end{bmatrix}$$

$p_{ww}^4 = 0.654$

Manchester United vs Arsenal

$$p^4 = \begin{bmatrix} 0.076 & 0.696 & 0.228 \\ 0.070 & 0.722 & 0.209 \\ 0.076 & 0.696 & 0.228 \end{bmatrix}$$

$p_{ww}^4 = 0.722$

Arsenal vs Manchester United

$$p^4 = \begin{bmatrix} 0.285 & 0.205 & 0.365 \\ 0.239 & 0.213 & 0.427 \\ 0.259 & 0.187 & 0.347 \end{bmatrix}$$

$p_{ww}^4 = 0.213$

Manchester United vs Man City

$$p^4 = \begin{bmatrix} 0.393 & 0.477 & 0.131 \\ 0.381 & 0.481 & 0.138 \\ 0.367 & 0.484 & 0.150 \end{bmatrix}$$

$p_{ww}^4 = 0.481$

Man City vs Manchester United

$$P^4 = \begin{bmatrix} 0.415 & 0.497 & 0.088 \\ 0.438 & 0.477 & 0.085 \\ 0.427 & 0.503 & 0.070 \end{bmatrix}$$

$p_{ww}^4 = 0.477$

Chelsea vs Arsenal

$$p^4 = \begin{bmatrix} 0.237 & 0.582 & 0.181 \\ 0.229 & 0.584 & 0.187 \\ 0.229 & 0.582 & 0.189 \end{bmatrix}$$

$p_{ww}^4 = 0.584$

Arsenal vs Chelsea

$$p^4 = \begin{bmatrix} 0.352 & 0.25 & 0.398 \\ 0.352 & 0.25 & 0.398 \\ 0.348 & 0.25 & 0.402 \end{bmatrix}$$

$p_{ww}^4 = 0.25$

Chelsea vs Manchester City

$$p^4 = \begin{bmatrix} 0.076 & 0.696 & 0.228 \\ 0.070 & 0.722 & 0.209 \\ 0.076 & 0.696 & 0.228 \end{bmatrix}$$

$p_{ww}^4 = 0.722$

Manchester City vs Chelsea

$$p^4 = \begin{bmatrix} 0.403 & 0.346 & 0.043 \\ 0.461 & 0.403 & 0.048 \\ 0 & 0 & 0 \end{bmatrix}$$

$p_{ww}^4 = 0.403$

Arsenal vs Manchester City

$$p^4 = \begin{bmatrix} 0.055 & 0.53 & 0.415 \\ 0.052 & 0.527 & 0.421 \\ 0.053 & 0.526 & 0.422 \end{bmatrix}$$

$p_{ww}^4 = 0.527$

Manchester City vs Arsenal

$$p^4 = \begin{bmatrix} 0.509 & 0.402 & 0.089 \\ 0.482 & 0.447 & 0.072 \\ 0.534 & 0.358 & 0.108 \end{bmatrix}$$

$p_{ww}^4 = 0.447$

Observe the following from the above 6 possible combinations

1. Man United has a probability of 0.31 of winning Chelsea considering home advantages while Chelsea has a higher probability 0.65 of winning Man U.
2. Man U has a high probability 0.72 of winning Arsenal and also Arsenal vs Man U has a low probability 0.21 of winning Man U.
3. Man U has a probability 0.48 of winning M.C and Man City has a high probability 0.48 of winning Man U.
4. Chelsea has a probability 0.58 of winning Arsenal and Arsenal has a low probability 0.25 of winning Chelsea.
5. Chelsea has a probability 0.72 of winning M.C and Man City has a low probability 0.40 of winning Chelsea.
6. Arsenal has a probability 0.53 of winning Man City and Man City has a low probability 0.45 of winning Arsenal.

Limiting Distributions of The Transition Probability Matrices

Since all the TPM are irreducible, recurrent Markov Chains and the eigenvalues satisfy the following equations $\lambda_1 = 1$ 1

$$|\lambda_1| < 1 \qquad j = 2,3$$

The limiting distributions of the transition probability matrices exist and are subsequently obtained for each team using the following equations

$$\sum_{i=0}^{\infty} P_{ij}\pi_i = \pi_j \qquad j \geq 0 \qquad 2$$

$$\sum_{i=0}^{\infty} \pi_i = 1$$

Where π_i is called the long run proportion of time spent at state i. Using equation 2, the limiting distributions of the TPM for each team are obtained as shown in Table 10

Table 10 Limiting Distributions of the TPM of each Team from 2002-2015

		L	W	D
Man U vs C		[illegible]63	[illegible]	[illegible]2
C vs Man U		[illegible]79	[illegible]	[illegible]3
Man U vs A		0.072	0.714	0.214
A vs Man U		0.323	0.247	0.446
Man U vs Man City		0.384	0.479	0.137
Man City vs Man U		0.427	0.488	0.085
C vs A		0.231	0.583	0.186
A vs C		0.35	0.25	0.4
C vs Man City		0.072	0.714	0.214
Man City vs C		0.506	0.443	0.114
A vs Man City		0.053	0.526	0.421
Man City vs A		0.5	0.417	0.083

From the above table we observe the following, Manchester United has a high probability 0.44 of winning Chelsea while Chelsea has a higher probability 0.66 of winning Manchester United considering home advantage, Secondly Man U has high probability 0.71 of winning Arsenal while Arsenal has a low probability of winning Man U,Thirdly; Man U and Man City has an equal probability 0.48 of winning in their respective homes,Fourthly;Chelsea has a high probability 0.71 of winning Man City while Man City has a probability of 0.44 of winning Chelsea and lastly Arsenal has a probability of 0.5 of winning Man City while Man City has a 0.4 probability of winning Arsenal.

CONCLUSION

Based on the finding of the analysis, one can say that the two most viable clubs out of the four clubs are Manchester United and Chelsea, considering their performance in the highly selected clubs and the scoreline of their games

REFERENCE

Baio,G., & Blangiardo, M. (2010). Bayesian hierarchical model for the prediction of Football results, Journal of Applied Statistics 32(7): 253-264.

Blundell J.D.(2009). Numerical Algorithms for predicting Sports Results. Leeds: University of Leeds.

Ferda,H.(2009). Research on the Prediction of the likely winners of the euro 2008 Football Tournament. Journal of Quatitative Analysis in sports, 1-13.

Goddard,J. and Asimakopulos,I.(2004). Forcasting Football results and the efficiency of fixed-odds betting. Journal of Forcasting 23,51-66.

Hennessy J (1969) The Times, March 21st 1969. 10

Hill I.D.(1974), Association Football and Statistical Inference. Applied Statistics,23,203-208.

Lee A.J.(1997), Modeling Scores in Premier League: Is Manchester United really the best. Chance,10,15-19.

Maher M.J (1982), Modeling Association Football Scores. Statistica Neerlandica,36,109-118.

Mandela,M.N.(1992),Sport has the power to change the Word

Nwosu, C.R. (2001). The Fundamentals of Operations Research. published by Afrika-Link Communications Ltd Owerri 88,104-107.

Reep C. and Benjamin B. (1968), Skill and Chance in Association Football Journal of the Royal Statistical Society, Series A, 131,581-585.

Steinmetz, J.G.(2000), System and method for predicting the outcome of college Football Games, United States patent 6112128,issued August 29.