The Bivariate Poisson Distribution and the Prediction of Football Matches

Gavin Whitaker

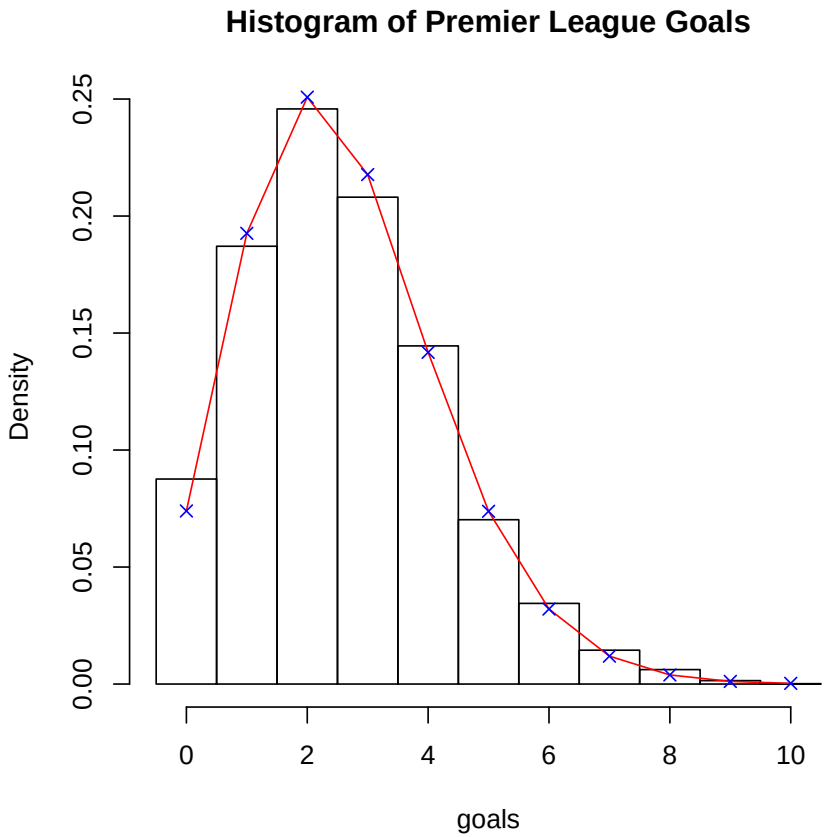
February 22, 2012

Supervisors: Phil Ansell and Dave Walshaw

Outline

- Motivation and the basics.
- Creating the Premier League.
- Teams attack and defence parameters.
- Results.
- A ‘real’ life example.

Beginnings



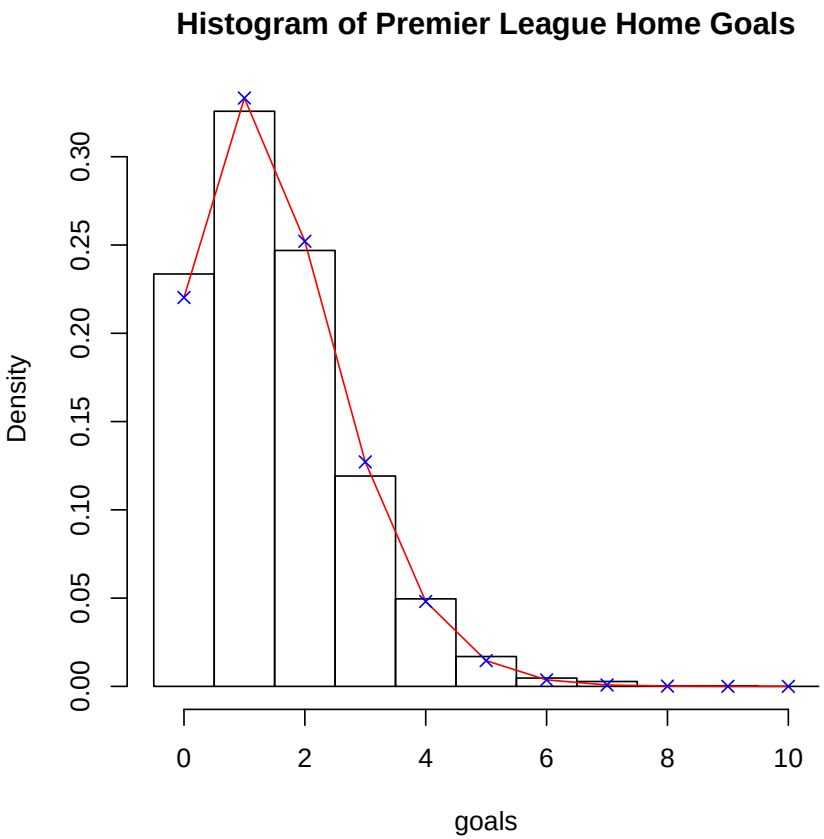
The mean number of goals scored in a game is 2.604.

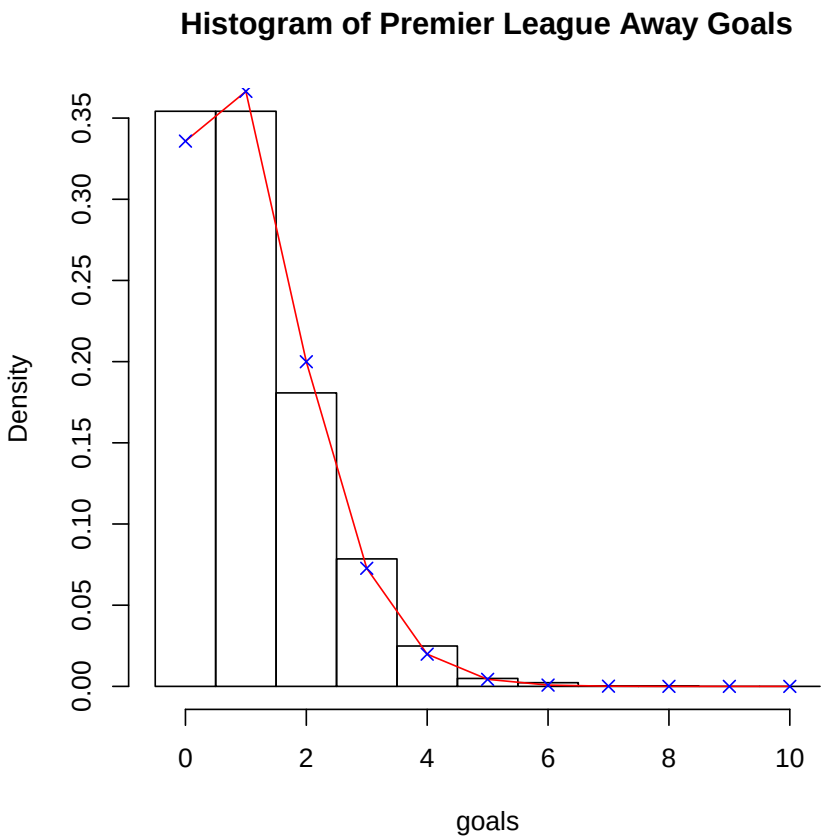
Beginnings

Sunderland, Liverpool game, (20^{th} March 2011), and compare it with the odds offered by bet365.

Total Goals	0	1	2	3	4
Pois(2.604)	0.074	0.193	0.251	0.218	0.142
bet365 odds	0.12	0.24	0.282	0.221	0.137

Beginnings





The home team mean is 1.513 and the away mean is 1.091.

Beginnings

Sunderland Goals (Home)	0	1	2	3
Pois(1.513)	0.220	0.333	0.252	0.127
bet365 odds	0.368	0.283	0.243	0.106

Liverpool Goals (Away)	0	1	2	3
Pois(1.091)	0.336	0.366	0.2	0.073
bet365 odds	0.312	0.365	0.222	0.101

The Basics

The Basics

- Bivariate Poisson Distribution.

The Basics

- Bivariate Poisson Distribution.
- Probability Generating Function

$$G(t_1, t_2) = exp\{\lambda_1(t_1 - 1) + \lambda_2(t_2 - 1) + \lambda_3(t_1 t_2 - 1)\}.$$

- Recall that the PGF of the univariate Poisson distribution is

$$exp\{\lambda(t - 1)\}.$$

The Basics

- Expanding this gives:

- Expanding this gives:

$$G(t_1, t_2) = e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{i=0}^{\infty} \frac{\lambda_1^i t_1^i}{i!} \sum_{j=0}^{\infty} \frac{\lambda_2^j t_2^j}{j!} \sum_{k=0}^{\infty} \frac{\lambda_3^k t_1^k t_2^k}{k!}.$$

- Expanding this gives:

$$G(t_1, t_2) = e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{i=0}^{\infty} \frac{\lambda_1^i t_1^i}{i!} \sum_{j=0}^{\infty} \frac{\lambda_2^j t_2^j}{j!} \sum_{k=0}^{\infty} \frac{\lambda_3^k t_1^k t_2^k}{k!}.$$

- On substituting t_1 and t_2 equal to zero we get the marginals

$$X \sim Po(\lambda_1 + \lambda_3)$$

and

$$Y \sim Po(\lambda_2 + \lambda_3).$$

Creating the Premier League

Creating the Premier League

- The Premier League season.

Creating the Premier League

- The Premier League season.
- Using R:

Creating the Premier League

- The Premier League season.
- Using R:
  - Factors, separating home and away teams.

Creating the Premier League

- The Premier League season.
- Using R:
  - Factors, separating home and away teams.
  - Points.

Creating the Premier League

- The Premier League season.
- Using R:
  - Factors, separating home and away teams.
  - Points.
  - Sorting.

Creating the Premier League

The final standings for The Premier League 09/10.

Team	Points	Goaldif	Goals	Conc
Chelsea	86	71	103	32
Man United	85	58	86	28
Arsenal	75	42	83	41
Tottenham	70	26	67	41
Man City	67	28	73	45
Aston Villa	64	13	52	39
Liverpool	63	26	61	35
Everton	61	11	60	49
Birmingham	50	-9	38	47
Blackburn	50	-14	41	55
⋮	⋮	⋮	⋮	⋮

Creating the Premier League

Home points and the question of home effect.

Team	Points at home	Team	Points at home
Chelsea	51	Aston Villa	24
Man United	48	Birmingham	24
Arsenal	45	Burnley	21
Tottenham	42	Stoke	21
Liverpool	39	West Ham	21
Man City	36	Bolton	18
Everton	33	Hull	18
Fulham	33	Wigan	18
Blackburn	30	Portsmouth	15
Sunderland	27	Wolves	15

Attack and Defence Parameters

Attack and Defence Parameters

- Team specifics.

Attack and Defence Parameters

- Team specifics.
- Regression using the R package Bivpois which has been used for:

Attack and Defence Parameters

- Team specifics.
- Regression using the R package Bivpois which has been used for:
  - General simulation.

Attack and Defence Parameters

- Team specifics.
- Regression using the R package Bivpois which has been used for:
  - General simulation.
  - Modelling football matches, specifically the 1991-1992 Italian serie A.

Attack and Defence Parameters

- Team specifics.
- Regression using the R package Bivpois which has been used for:
  - General simulation.
  - Modelling football matches, specifically the 1991-1992 Italian serie A.
  - Modelling water-polo games.

Attack and Defence Parameters

- Team specifics.
- Regression using the R package Bivpois which has been used for:
  - General simulation.
  - Modelling football matches, specifically the 1991-1992 Italian serie A.
  - Modelling water-polo games.
  - The demand for healthcare in Australia.

Attack and Defence Parameters

Attack and Defence Parameters

- Using this package for a specific home and away team we can fit the model:

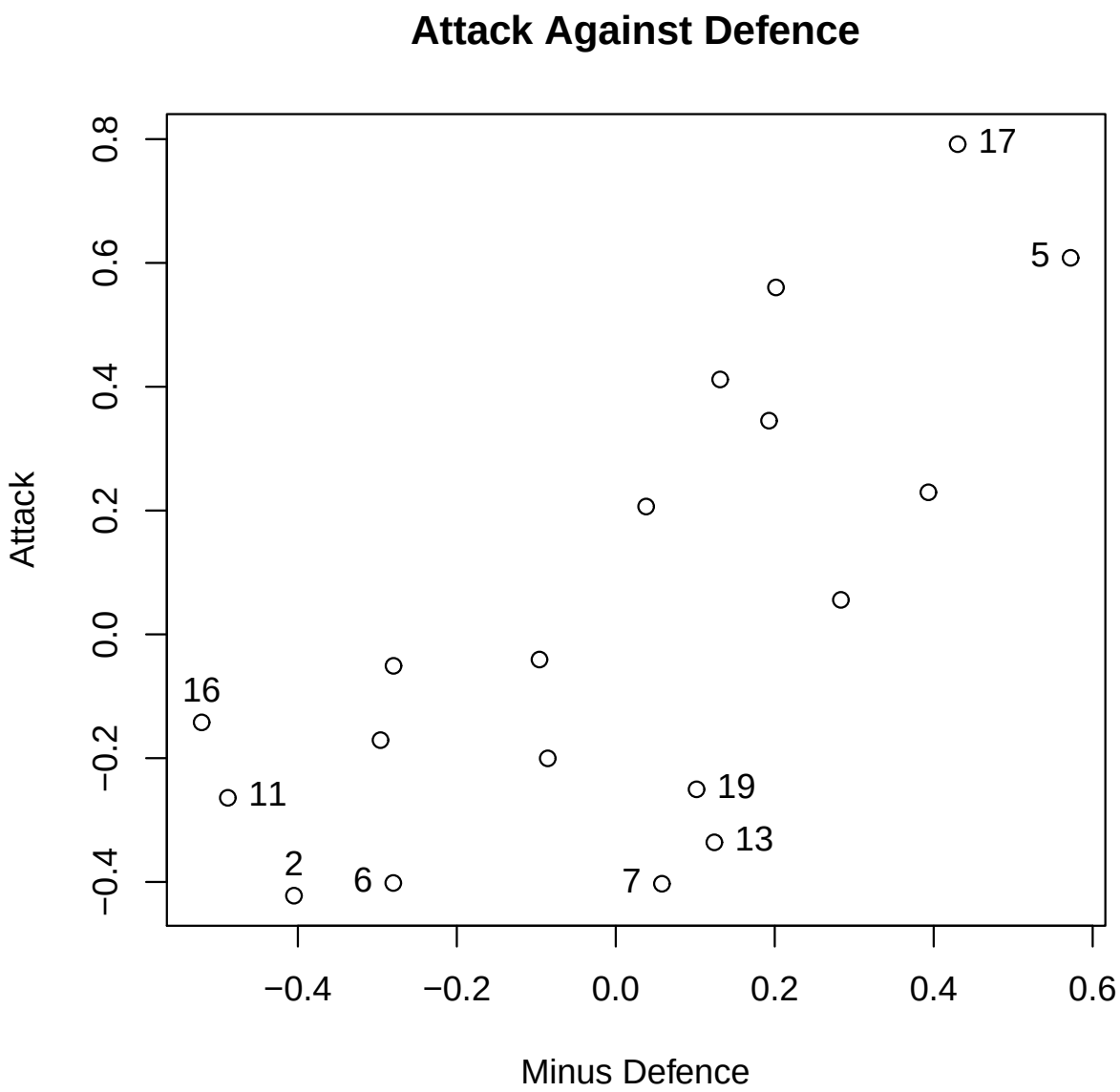
Attack and Defence Parameters

- Using this package for a specific home and away team we can fit the model:
  - $\log(\lambda_1)=\mu$ + home attack + away defence + home effect.
  - $\log(\lambda_2)=\mu$ + away attack + home defence.

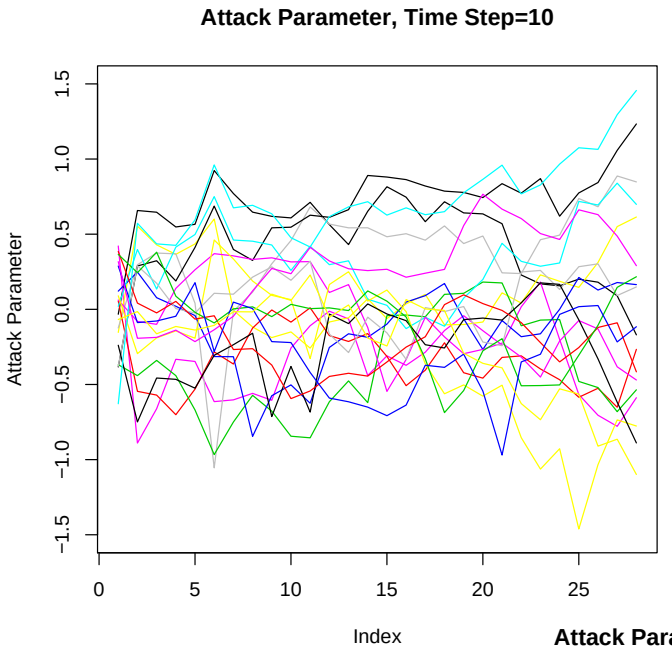
Attack and Defence Parameters

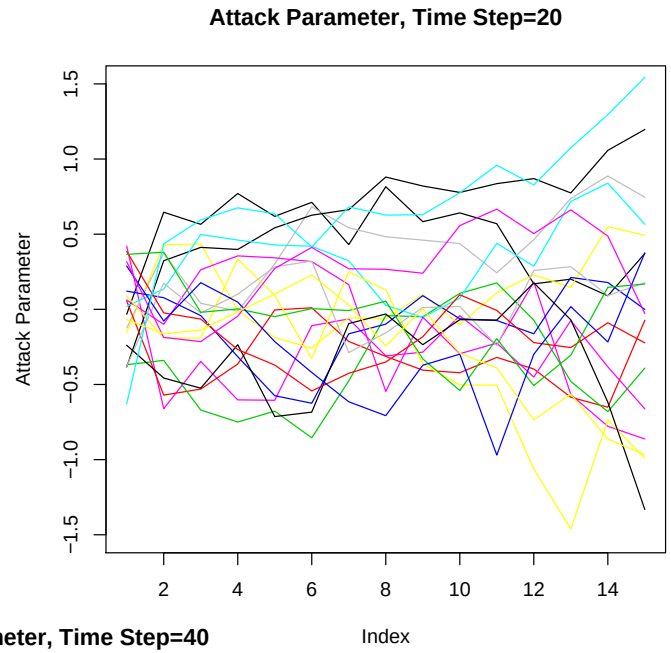
- Using this package for a specific home and away team we can fit the model:
  - $\log(\lambda_1)=\mu$ + home attack + away defence + home effect.
  - $\log(\lambda_2)=\mu$ + away attack + home defence.
- Time intervals and form.

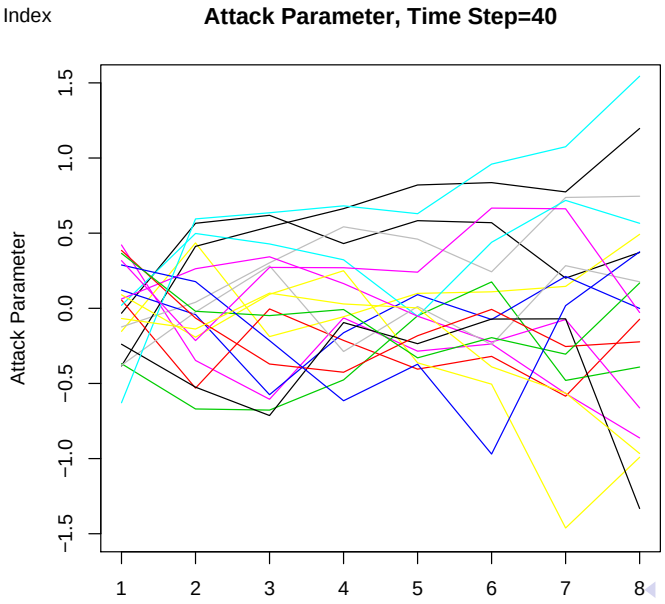
Attack and Defence Parameters



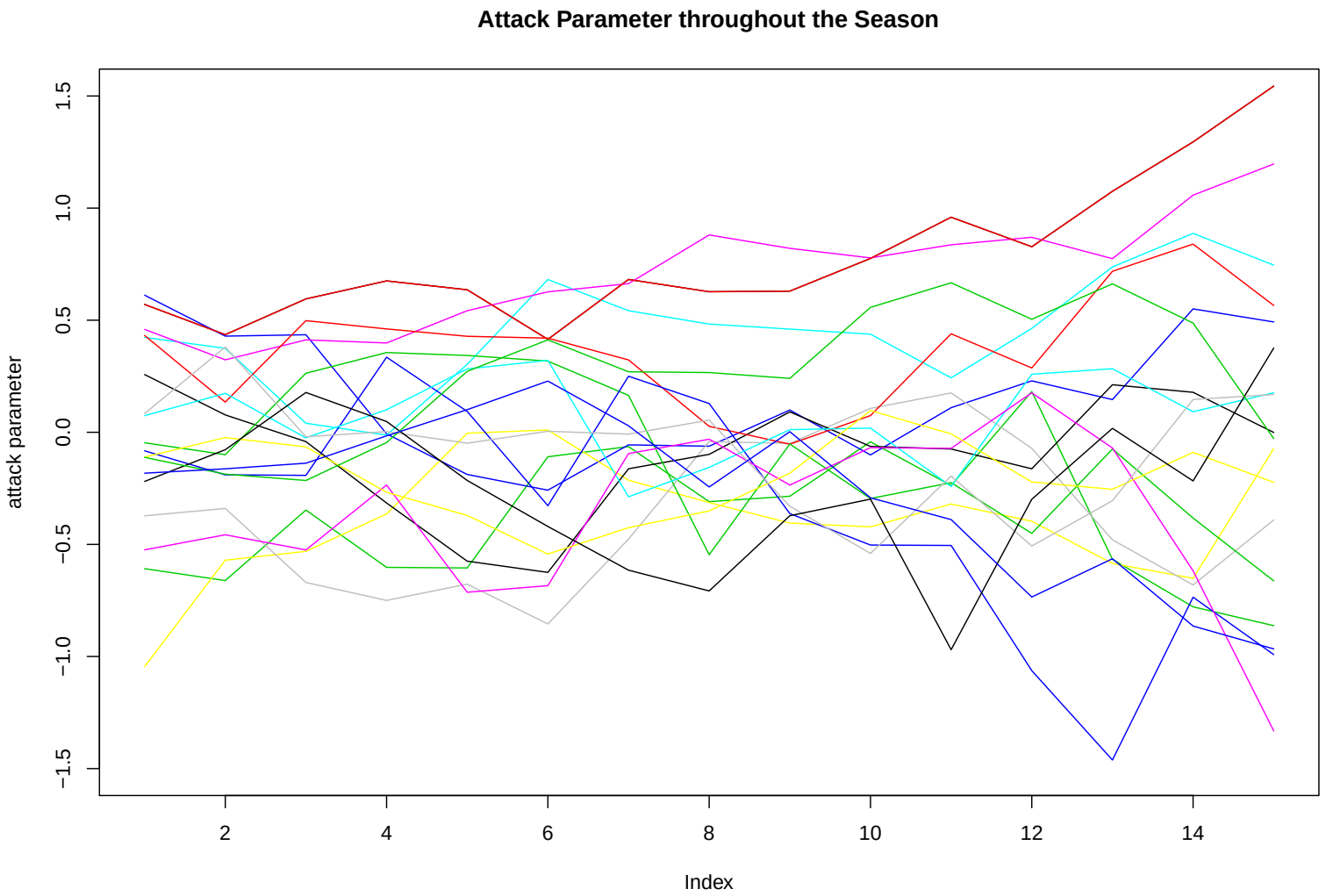
Attack and Defence Parameters



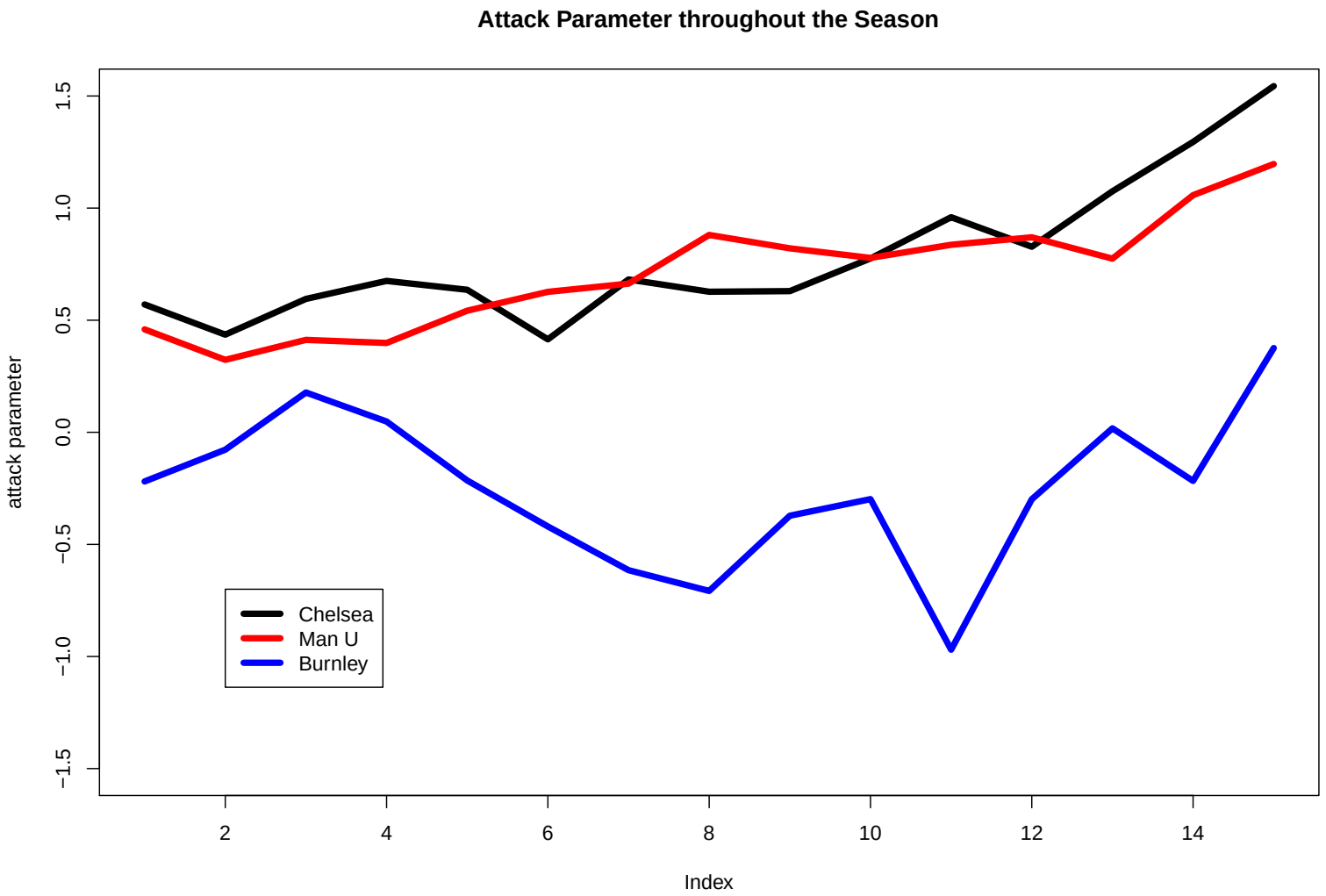




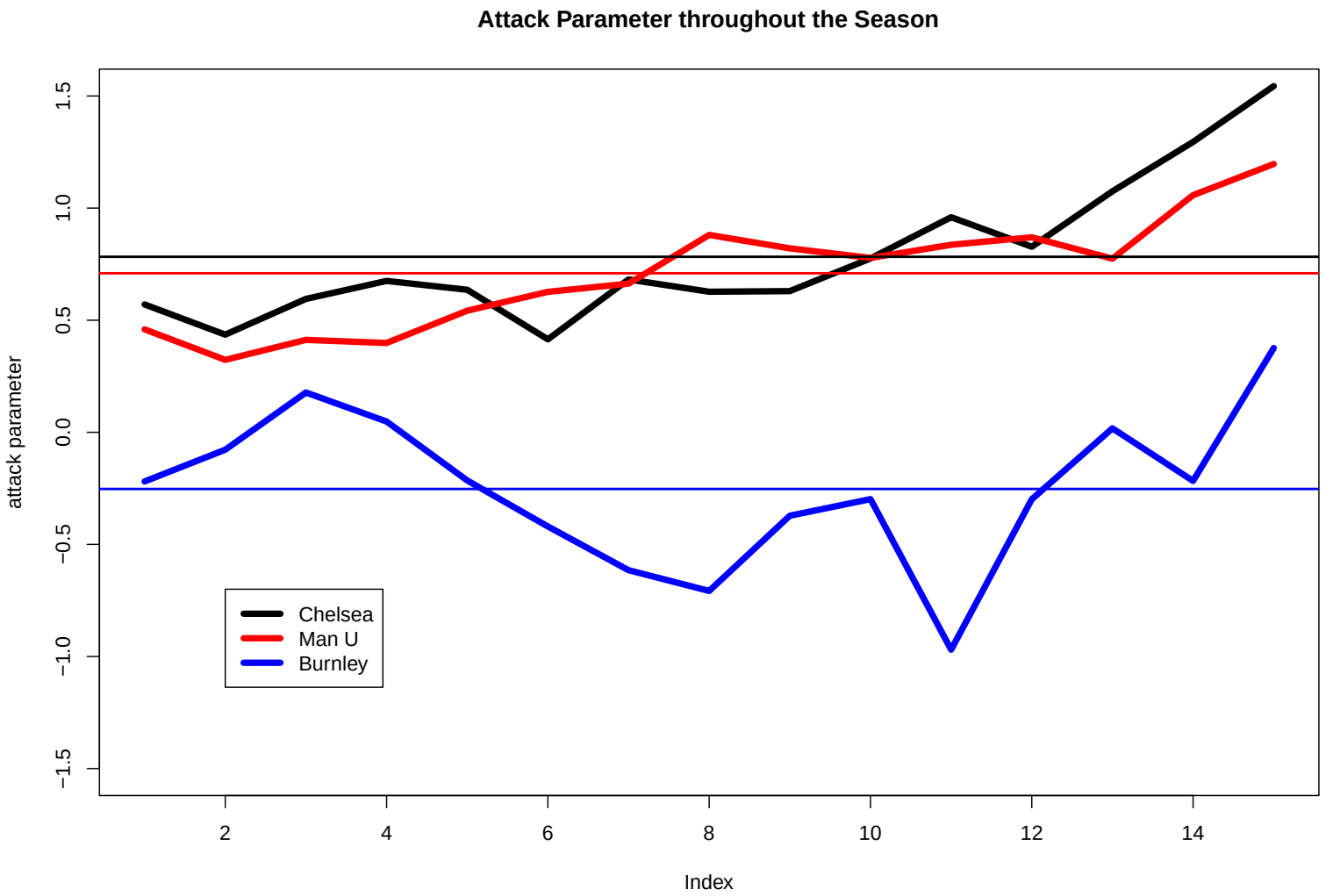
Attack Throughout the Season



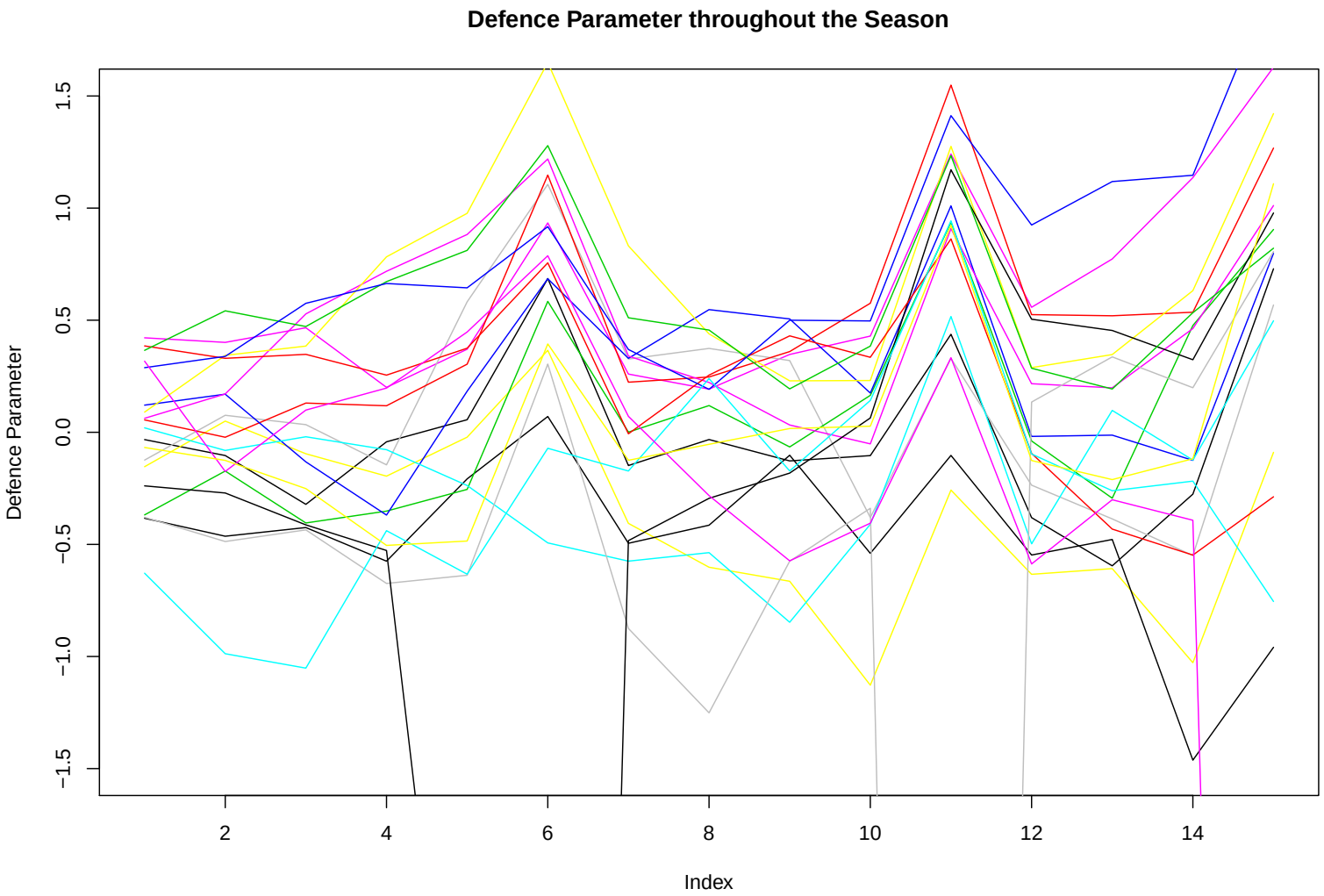
Attack for Specific Teams



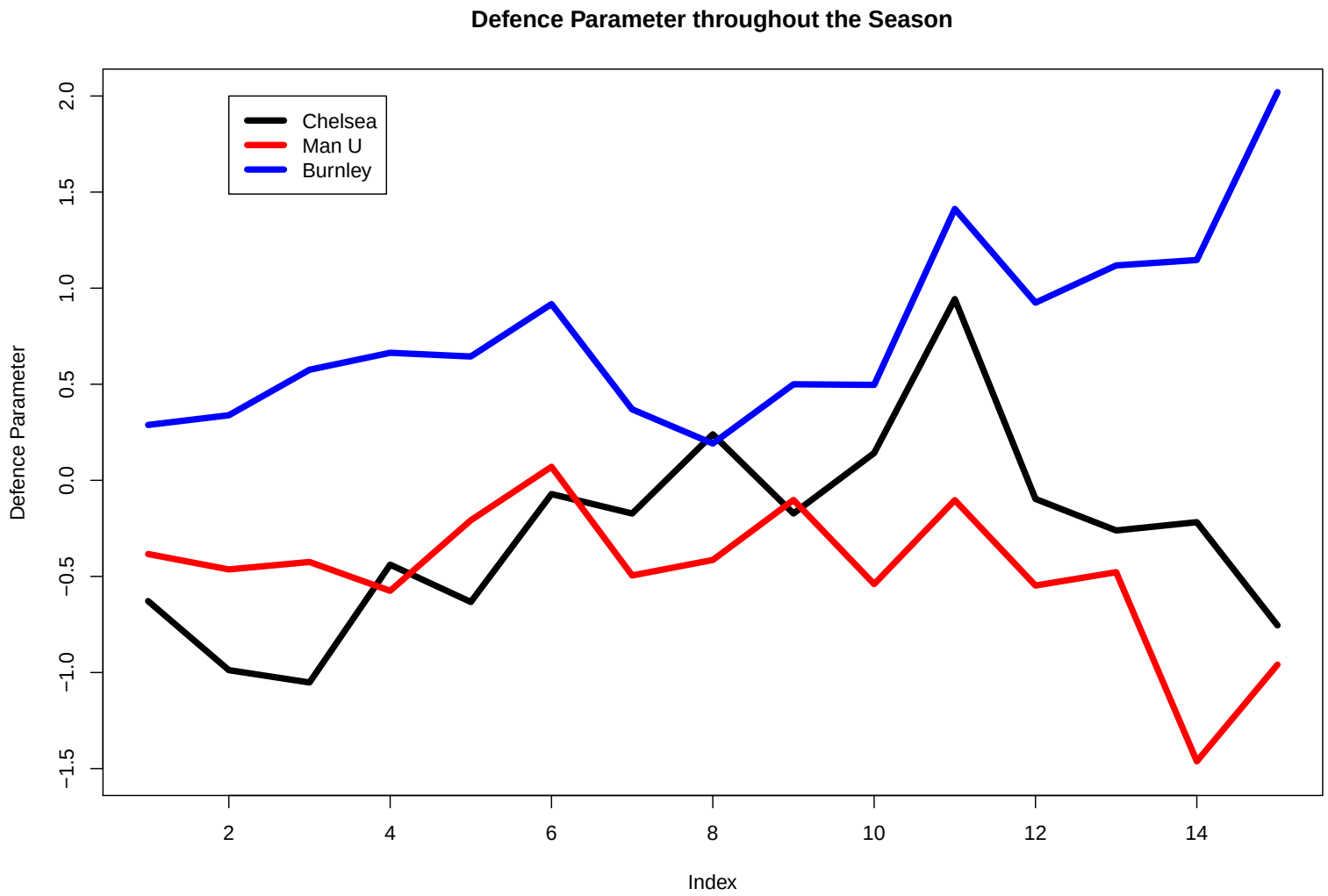
Attack Average



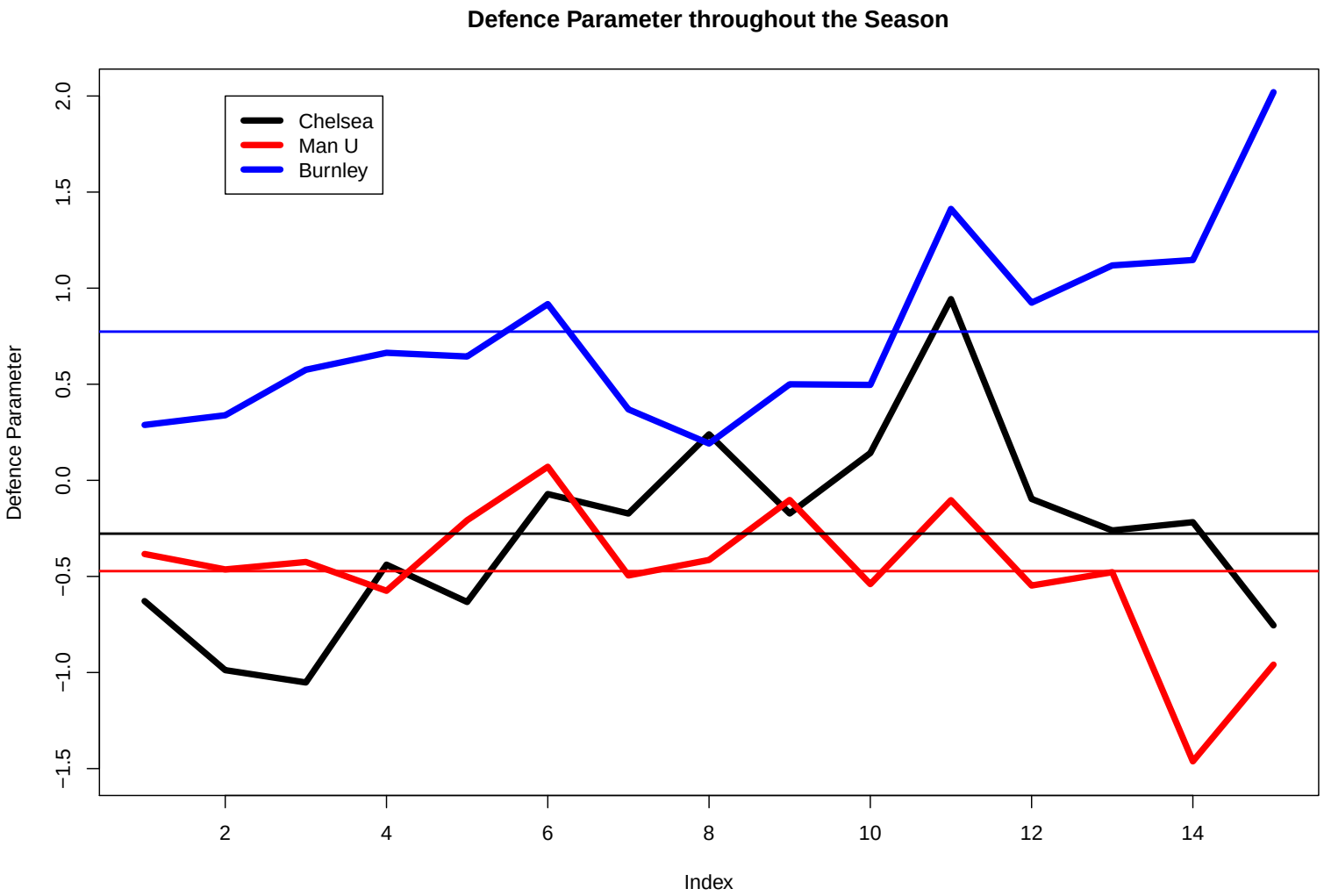
Defence Throughout the Season



Defence for Specific Teams



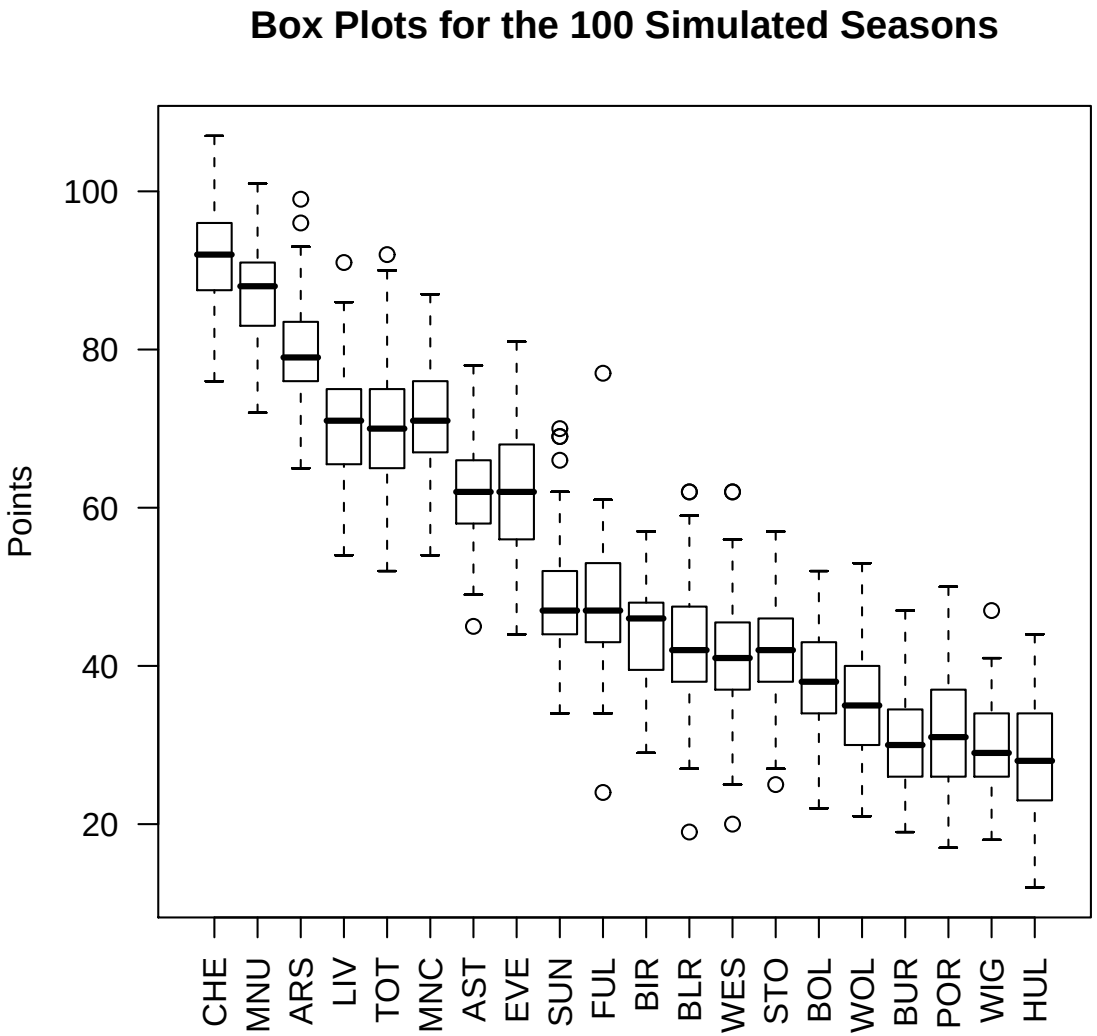
Defensive Average



Results-Season Estimate

Team	Points	Games Won	Games Drawn	Goals	Conc	Goaldif
Chelsea	90.5	28.44	5.18	111.54	36.86	74.68
Man United	88.4	27.37	6.29	95.82	31.34	64.48
Arsenal	78.35	23.97	6.44	91.29	44.72	46.57
Liverpool	71.23	20.89	8.56	67.32	39.29	28.03
Tottenham	71.1	20.99	8.13	75.63	45.96	29.67
Man City	70.14	20.78	7.8	77.08	48.16	28.92
Aston Villa	63.78	718.11	9.45	58.4	42.59	15.81
Everton	61.78	17.9	8.08	66.24	52.19	14.05
Sunderland	48.78	13.27	8.97	52.12	60.72	-8.6
Fulham	48.07	12.62	10.21	44.46	50.48	-6.02
⋮	⋮	⋮	⋮	⋮	⋮	⋮

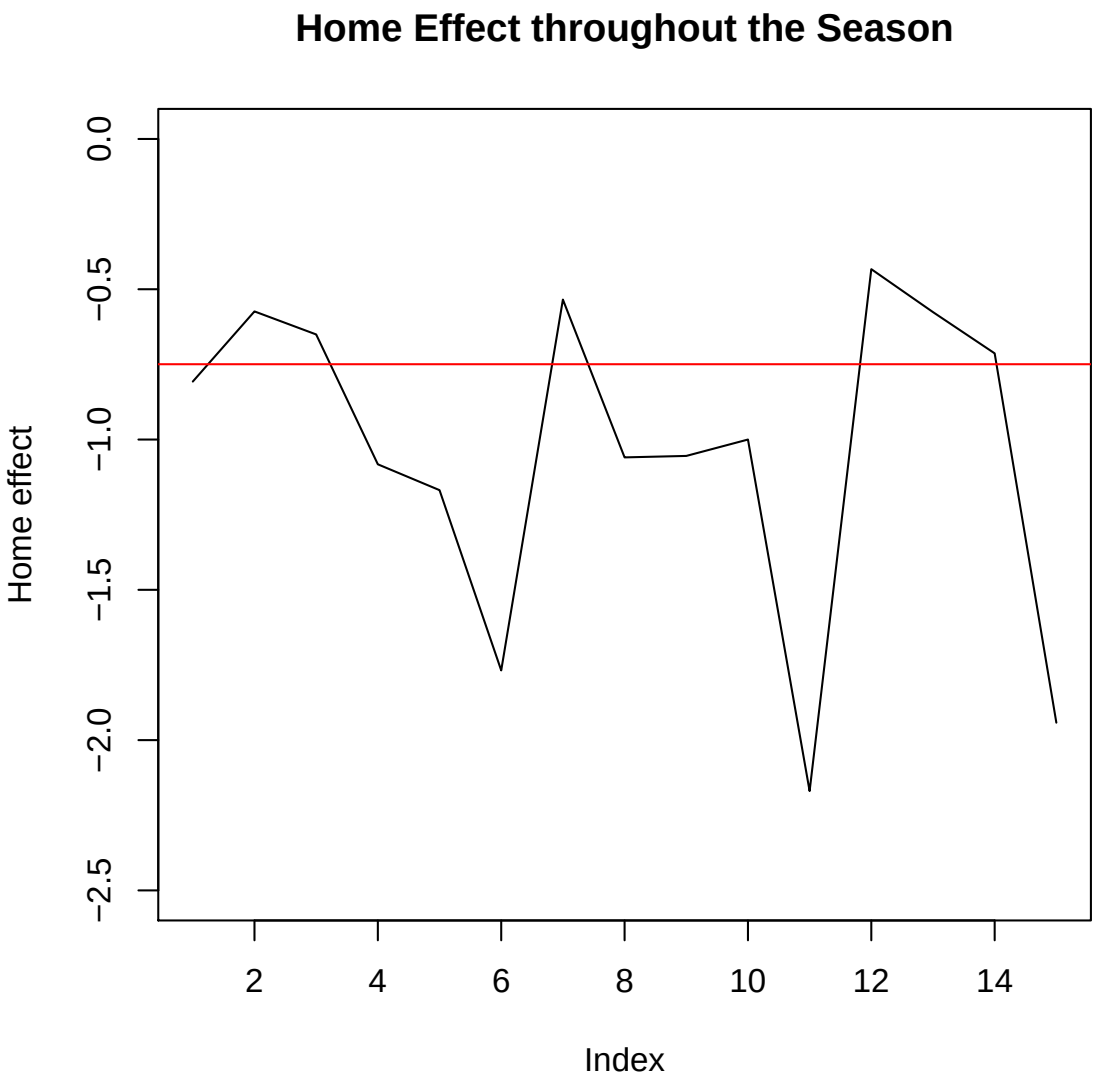
Results-Season Estimate



Results-Moving Average

Team	Points	Goals	Conc	Goaldif
Man United	70.96	105.82	53.71	52.11
Tottenham	69.13	97.84	57.77	40.07
Aston Villa	67.38	80.01	52.57	27.44
Chelsea	65.52	105.42	64.55	40.87
Arsenal	65.25	106.40	68.36	38.04
Man City	62.84	99.48	78.67	20.81
Everton	61.89	87.26	65.29	21.97
Liverpool	60.38	82.75	61.85	20.90
Blackburn	55.63	74.14	75.38	-1.24
Birmingham	54.24	63.10	68.24	-5.14
⋮	⋮	⋮	⋮	⋮

Results-Moving Average



A Real Life Example

A ‘Real’ Life Example

A ‘Real’ Life Example

- What is Football Manager.

A ‘Real’ Life Example

- What is Football Manager.
- Simulates depending on players not teams.

A ‘Real’ Life Example

- What is Football Manager.
- Simulates depending on players not teams.
- Simulating a Season.

A ‘Real’ Life Example

- What is Football Manager.
- Simulates depending on players not teams.
- Simulating a Season.
  - The Final Table.

A ‘Real’ Life Example

- What is Football Manager.
- Simulates depending on players not teams.
- Simulating a Season.
  - The Final Table.
  - The Results.

A 'Real' Life Example

The Premier League 09/10.

		Home					
		0	1	2	3	4	5
A	0	0.0842	0.0868	0.0895	0.0605	0.0263	0.0105
w	1	0.0684	0.1026	0.1105	0.0421	0.0132	0.0079
a	2	0.0421	0.05	0.0553	0.0211	0.0079	0.0053
y	3	0.0079	0.0211	0.0132	0.0105	0.0053	0.0026
	4	0.0105	0.0105	0.0026			
	5	0.0053	0.0026				

A 'Real' Life Example

The Premier League 09/10.

		Home					
		0	1	2	3	4	5
A	0	0.0842	0.0868	0.0895	0.0605	0.0263	0.0105
w	1	0.0684	0.1026	0.1105	0.0421	0.0132	0.0079
a	2	0.0421	0.05	0.0553	0.0211	0.0079	0.0053
y	3	0.0079	0.0211	0.0132	0.0105	0.0053	0.0026
	4	0.0105	0.0105	0.0026			
	5	0.0053	0.0026				

Results

home	draw	away
0.5079	0.2526	0.2395

A 'Real' Life Example

Football Manager simulation of the Premier League.

		Home					
		0	1	2	3	4	5
A	0	0.05	0.0868	0.0658	0.0447	0.0263	0.0026
w	1	0.0737	0.1263	0.0868	0.0421	0.0158	
a	2	0.0658	0.1053	0.0553	0.0316	0.0079	
y	3	0.0289	0.0421	0.0158	0.0053		
	4	0.0079	0	0	0.0079		
	5	0.0026					

A 'Real' Life Example

Football Manager simulation of the Premier League.

		Home					
		0	1	2	3	4	5
A	0	0.05	0.0868	0.0658	0.0447	0.0263	0.0026
w	1	0.0737	0.1263	0.0868	0.0421	0.0158	
a	2	0.0658	0.1053	0.0553	0.0316	0.0079	
y	3	0.0289	0.0421	0.0158	0.0053		
	4	0.0079	0	0	0.0079		
	5	0.0026					

Results

home	draw	away
0.4132	0.2368	0.35

Conclusion

- We've looked at some of the maths behind simulating football matches.
- We've looked at building The Premier League.
- We've looked at the R package Bivpois and considered a model for simulating football matches.
- We've looked at teams attack and defence parameters and talked about 'the home effect'.
- We've looked at a 'real' life example.

Conclusion

- We've looked at some of the maths behind simulating football matches.
- We've looked at building The Premier League.
- We've looked at the R package Bivpois and considered a model for simulating football matches.
- We've looked at teams attack and defence parameters and talked about 'the home effect'.
- We've looked at a 'real' life example.

THE END