ESTIMATING RATINGS IN FOOTBALL: BRAZILIAN CHAMPIONSHIP 2017

Luciano Ribeiro GALVÃO[1]
Júlio Sílvio de Sousa BUENO FILHO[2]

▪ ABSTRACT: Observing games from 2017 Brazilian Football Championship League series A, we adjusted models with linear predictors for the expected number of goals for home and away teams. We estimated values for attacking (γ) and defense (δ) strength parameters for each team. A common home advantage effect (μ) was used for all teams and conditionally estimated. The first leg of the double round-robin was used to produce initial estimates. Those were then re-estimated at each round of the second leg. We intend to present a model that could describe past performance and predict results (and scores) from each game. Additional features of the final classification of the tournament could also be predicted, such as the probability of playing South American Champions League or being demoted to series B of the competition. Proposed model can be considered flexible in the pure likelihood analysis version, but some desirable features of parametric Bayesian inference could enhance it's capabilities.

▪ KEYWORDS: Double round-robin; home advantage; Poisson distribution; prediction; relative strength

1 Introduction

Predicting events is of major interest in many fields in and outside academy. To assign probabilities of future events having little previous information (although many data) is one of the main goals of professionals from Statistics and Data Science. In the case of sports, some techniques and analytical tools have been developed to predict results extrapolating aspects of past performance from individuals or groups.

[1]Universidade do Estado de Minas Gerais - UEMG, Unidade Passos, CEP: 37902-407, Passos, Minas Gerais, Brasil. E-mail: *Luciano.galvao@uemg.br*
[2]Universidade Federal de Lavras - UFLA, Departamento de Estatística, Caixa Postal 3037, CEP: 37200-900, Lavras, Minas Gerais, Brasil. E-mail: *jssbueno@ufla.br*

Besides public interest there is also a great volume of financial contribution to sports and entertainment industry. Football is one of the most profitable sports in this sense, as it is the most popular in the world.

In the current literature it can be found methods for parametric estimation using football scores with different techniques. One defining difference is if realizations are actual game scores or aggregate results (victory, draw or defeat of the home team).

The most common choice for aggregate results are the binomial (ignoring ties) and the multinomial distributions. For the full scores many variations of Poisson distribution have been adjusted, in special to accommodate correlation between same game scores and extra-poisson variation. Examples would be the use of negative binomial and Skellam distributions (KARLIS and NTZOUFRAS, 2009; BAIO and BLANGIARDO, 2010). Poisson distribution was used to model scores based on goal differences (home minus away goals) using a Bayesian hierarchical model for the 2008 Italian Championship, with good predictive results (SUZUKI et al., 2010).

Sports ratings are in general a single real value associated to the relative strength of a player or team (ELO, 1978). In this paper we present sets of rating estimates for attack and defense parameters for each club from an actual championship (2017 Brazilian Football, series A) and use this estimates to predict future results and check with the actual realized scores to validate the model. Techniques used are general and could be replicated for any season of the same championship or similar sports.

The basic steps were to establish a linear predictor for expected score using attacking (γ_i) and defensive (δ_i) rating parameters for each club, and an additional modifier of club performance due to home advantage (μ) that can only be estimated conditioning on knowing initial ratings for all clubs.

This has a strong resemblance to state-space solutions implemented by Glickman and Stern (1998). Rating parameters are relative and have limited validity for this competition. Only observed results are used and no previous history or "known" nuisance factors has impact on estimation.

As noted by Andrade and Espirito Santo (2016), home advantage is an important factor, although its estimation can be much compromised. The model is not directly identifiable as the linear predictor using (μ) would be linearly dependent with previous parameters. A conditional estimation is, however, feasible. Our working hypothesis is that our estimation method model could elucidate the role of home advantage in analogous rating systems. In this way we establish a simple prediction model as a possible basis for a rating system in football.

2 Material e methods

For presentation purposes through this paper, the clubs are divided in three groups according to final classification in the championship. First four best positioned are called (G4) and would play South American Championship. Last four are called (Z4) and would be demoted, other 16 clubs are in the middle range (MR).

We used data from 2017 "series A" (main division) Brazilian (Association or Soccer) Football Championship. The tournament is a double round robin with 20 clubs in 38 rounds with 10 games per round. Data are scoring from home and away teams from each of the 380 games and can be obtained in UOL sports website (*https://www.uol.com.br/esporte/futebol/campeonatos/brasileirao/*).

Table 1 lists clubs enrolled for the Brazilian Championship series A, 2017 edition, in alphabetic ordering of popular names, and respective club initials for short.

Table 1 - Clubs enrolled for the Brazilian Championship series A - 2017

Club	Initials	Club	Initials
Atlético-GO	ACG	Bahia	BAH
Atlético-MG	CAM	Botafogo	BOT
Atlético-PR	CAP	Chapecoense	CHA
Avaí	AVA	Corinthians	COR
Grêmio	GRE	São Paulo	SPA
Palmeiras	PAL	Sport	SPT
Ponte Preta	PON	Vasco	VAS
Santos	SAN	Vitória	VIT
Fluminense	FLU	Coritiba	CTB
Cruzeiro	CRU	Flamengo	FLA

All programming was carried out in R (R CORE TEAM, 2019). Detailed programs can be obtained with authors.

2.1 Estimation

We initially fitted a static Poisson model for rating estimation that resembles Gamerman (2013). Unlike this paper we did not follow a dynamic Bayesian fitting, but carried out new estimates updating only the likelihood from games of each round.

Home advantage was estimated only for the second leg (from round 20 on). Home advantage parameter (μ) is not directly estimable (identifiable) but can be worked out conditioning on point estimates of the attack and defense parameters in a new restricted (or conditional) likelihood. Maximizing this likelihood we obtain $\hat{\mu}$.

Sequential estimation algorithm was:

1. optimize likelihood for attacking and defensive parameters (γ_i and δ_i) for the 190 games of first leg;
2. store $\hat{\gamma}_i$ and $\hat{\delta}_i$ for all clubs in this round;
3. using those point estimates we evaluate again the likelihood as a function of home advantage (μ) and maximize it again;
4. store $\hat{\mu}$ for this round
5. repeat the process for each new round

Direct sampling model is what follows, using (y_h) as the home scoring and (y_a) as the away scoring:

Definition 2.1.

$$y_h \sim Poisson(\lambda_h = e^{\eta_h})$$
$$\eta_h = \gamma_h - \delta_a$$
$$y_a \sim Poisson(\lambda_a = e^{\eta_a})$$
$$\eta_a = \gamma_a - \delta_h$$

in which h is a subscript for home parameters and a for away.

For the conditional likelihood, given estimates for $\hat{\gamma}$ and $\hat{\delta}$, we yield:

Definition 2.2.

$$y_h|\hat{\gamma}, \hat{\delta} \sim Poisson(\lambda_m = e^{\eta_m})$$
$$\eta_h = \hat{\gamma_m} - \hat{\delta_v} + \mu$$
$$y_a|\hat{\gamma}, \hat{\delta} \sim Poisson(\lambda_v = e^{\eta_v})$$
$$\eta_a = \hat{\gamma_v} - \hat{\delta_v} - \mu$$

Maximization of this conditional likelihood allows for estimation of $\hat{\mu}$.

Design coding and webscrapping are strictly operational tasks and details can be given by authors on demand. An example of data organization is presented in Table 2:

Table 2 - Example of data structure

Round	Game	Home Team	Away Team	Score	
				Y_h	Y_a
1	1	FLA	CAM	1	1
1	2	COR	CHA	1	1
1	3	FLU	SAN	3	2
⋮	⋮	⋮	⋮	⋮	⋮
38	379	ACG	FLU	1	1
38	380	CHA	CTB	2	1

All likelihood were maximized using NELDER and MEAD (1965) symplex algorithm as codded in *optim*() and *optimize*() functions form *R-stat*. Solutions (point estimates) were centralized to present effects as deviations from means. Some examples of summaries that illustrate properties and prediction potential of the method are given.

We worked out predictions before the rounds and compared with actual scoring. Comparison of expected and observed results are a reliable basis for quality checking of the fitting. It was also possible to estimate probabilities of each score as well as whole game result probabilities (home team victory, draw or lose). All those results were checked using model based credibility statistics.

3 Results

In this section we show estimation results for attacking and defense parameters for each club at the end of first and second legs. This helps comparing consistency of club strengths in both phases.

3.1 First leg

Rating patterns are shown in table 3. Higher values for both $\hat{\gamma}$ and $\hat{\delta}$ tend to be from teams with better classification in the first leg as well as at the tournament end.

Table 3 - First leg estimates of relative ratings for attack (γ), defense (δ) and provisional classification at the championship. Brazilian Football, series A, 2017

Order	Club	Attack		Defense	
		$\hat{\gamma}$	$s(\hat{\gamma})$	$\hat{\delta}$	$s(\hat{\delta})$
1	COR	0.09	0.02	0.27	0.03
2	GRE	0.21	0.02	0.08	0.03
3	SAN	0.03	0.02	0.05	0.03
4	PAL	0.25	0.02	0.18	0.03
5	FLA	0.07	0.02	0.04	0.02
6	SPT	0.07	0.02	-0.19	0.02
7	CRU	0.11	0.02	0.02	0.02
8	CAP	-0.07	0.03	-0.17	0.02
9	CTB	0.14	0.02	-0.14	0.02
10	FLU	0.16	0.02	-0.15	0.02
11	BOT	0.27	0.02	-0.03	0.02
12	VAS	0.05	0.02	-0.03	0.02
13	BAH	0.15	0.02	-0.06	0.02
14	CAM	0.21	0.02	-0.07	0.02
15	PON	0.05	0.02	-0.13	0.02
16	CHA	0.06	0.02	-0.07	0.02
17	SPA	0.13	0.02	-0.33	0.02
18	VIT	-0.10	0.03	-0.31	0.02
19	AVA	-0.06	0.03	0.03	0.02
20	ACG	-0.06	0.03	-0.33	0.02

COR, GRE, and PAL had high attacking rating estimates,while SAN had a good balance of positive $\hat{\gamma}$ and $\hat{\delta}$. This seems to be correlated with their good overall performance. Strong attacking in a system that rewards victory seems to be a good strategy, but it is interesting to note that the leader had higher defensive strength ($\hat{\gamma} < \hat{\delta}$). In general one can conclude that defensive parameters were more important that attack in the first leg. Better placed clubs had high values in both ratings. On the other hand, last positions are mainly due to low defensive strength. SPA is a less critical case among Z4 as the strong attacking parameter that could compensate this deficiency.

3.2 Second leg

There was a many changes in ordering except COR and ACG (leader and tailing clubs, respectively). It is a long term championship (almost seven months, from May 13^{th} to December 8^{th}). There were many manager changes and many athletes exchanged with European or Asian markets. Teams were not the same all along the road.

Ratings evaluated using 19 rounds of the first leg were updated at each round in the second leg. A single home advantage parameter was also estimated for each round. We could both predict results and check predictions with observed scores. Very high absolute values tended to have large impact on results, although general balance was the key for a good outcome for clubs in first positions.

In Table 4 are presented relative strength ($\hat{\gamma}_i$ and $\hat{\delta}_i$) and standard errors of estimates ($s(\hat{\gamma}_i)$ and $s(\hat{\delta}_i)$) for each $i^{th},\ i = 1, ..., 20$ club at the championship end. It can be seen that the general trends of the first round remains, but some emphasis on higher defensive parameters for the best clubs and low offensive parameter for the worse ones can also be seen. Two good examples of clubs that escape from Z4 due to good attacking performance are FLU and VIT, that had otherwise poor defense. The same type of counter intuitive rational can be seen in PAL that improved a lot his positioning in the table with a strong offensive team, but wit a mediocre defense has lost precious points. Again both COR and SAN had $\hat{\gamma} < \hat{\delta}$, being in G4 with strong defensive teams.

Table 4 - Estimates of relative ratings for attack (γ) and defense (δ). Ordering at the championship end. Brazilian Football, series A, 2017

Order	Clubs	Attack		Defense	
		$\hat{\gamma}$	$s(\hat{\gamma})$	$\hat{\delta}$	$s(\hat{\delta})$
1	COR	0.27	0.02	0.48	0.04
2	PAL	0.53	0.02	0.08	0.02
3	SAN	0.10	0.02	0.46	0.03
4	GRE	0.34	0.02	0.35	0.03
5	CRU	0.25	0.02	0.22	0.03
6	FLA	0.30	0.02	0.22	0.03
7	VAS	0.07	0.03	0.06	0.02
8	CHA	0.25	0.02	-0.07	0.02
9	CAM	0.28	0.02	-0.02	0.02
10	BOT	0.18	0.02	0.10	0.02
11	CAP	0.16	0.02	0.04	0.02
12	BAH	0.31	0.02	-0.02	0.02
13	SPA	0.28	0.02	-0.11	0.02
14	FLU	0.34	0.02	-0.11	0.02
15	SPT	0.24	0.02	-0.23	0.02
16	VIT	0.31	0.02	-0.20	0.02
17	CTB	0.20	0.02	-0.05	0.02
18	AVA	-0.31	0.04	0.01	0.02
19	PON	0.09	0.03	-0.06	0.02
20	ACG	0.04	0.03	-0.16	0.02

Note that stability was a distinctive mark of middle ground teams. Most of the clubs in MR had almost uniform ratings for both parameters. On the other hand, in the second phase of the tournament COR had also a strong offensive parameter, separating clearly from the chasing pack.

We also compared correlations of $\hat{\gamma}$ and $\hat{\delta}$ estimates with realized scoring (see Table 5). We assume that non-null correlation for those parameters are indication of a sensible model.

Table 5 - Correlation among rating parameters $\hat{\gamma}$, $\hat{\delta}$, pro and against scoring by game

	γ	δ
Pro scoring	$0.4627*$	0.2040
Against scoring	-0.2170	$-0.4086**$

Significance levels (α) for *Student's t* test at *5% and **1%

Using *Student's t* test we could see evidence of the association we suspected when devising the model, not only for pro scoring and δ but mostly for against scores and δ. As expected, other two correlations are not relevant.

Average attacking and defensive parameters along the second leg of the championship are depicted in Figure 1. Local regression *(Loess)* were estimated using automatic *(ggplot)* R package defaults. Note that greater variation are among demoted teams in Z4. A general trend on increasing scoring average along the second leg was also verified. This implies that as the competition reaches its final stages, most clubs start to face more risks and higher scores happened. To distinguish G4 from MR one should look primarily to defensive skills. On the other hand, to separate the demoted in Z4 from those in MR, attack is the most relevant factor.

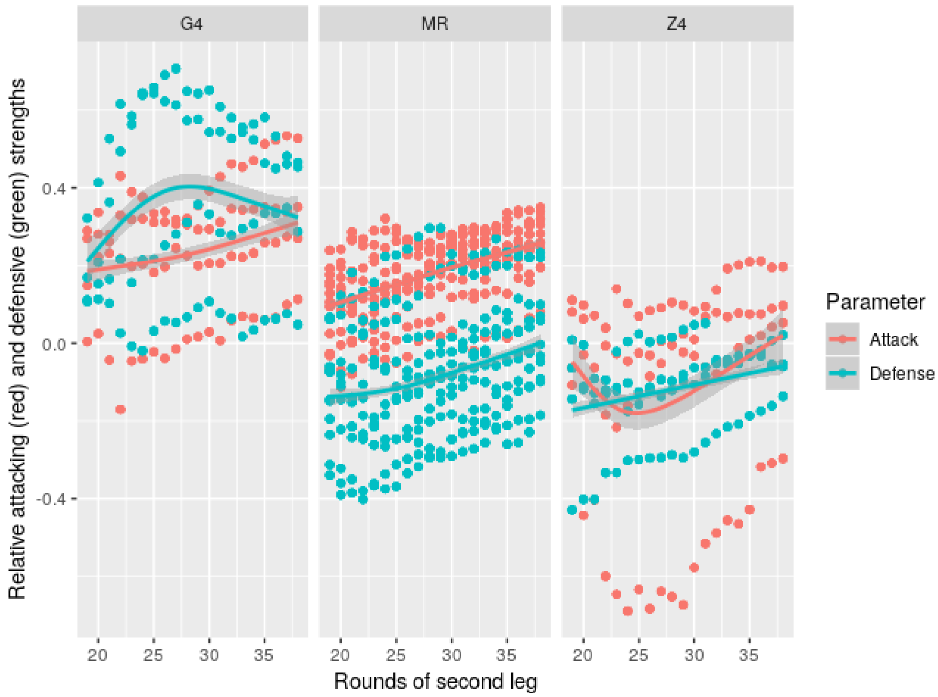

Figure 1 - Average estimates for atacking and defensive parameters for each round of the second leg. Brazilian Football, series A, 2017.

In Figures 3 and 2 are depicted relative attacking and defensive strengths, respectively. Unlike what is previously shown in Figure 1, results for each club are presented as a deviation from round average of all teams (more precisely, $\hat{\gamma}_i - \bar{\hat{\gamma}}$ and $\hat{\delta}_i - \bar{\hat{\delta}}$).

It can be seen that within groups the clubs are very much alike. For the attacking parameter in G4, COR is in the middle, GRE adn PAL are a bit higher and SAN a bit smaller. In Z4, AVA has the lowest attacking estimates, but the others are similar.

For the defensive parameter we can identify some greater differences among G4 and MR. SAN and COR are clearly better than the rest of their group. ACG is the negative highlight of the Z4 group.

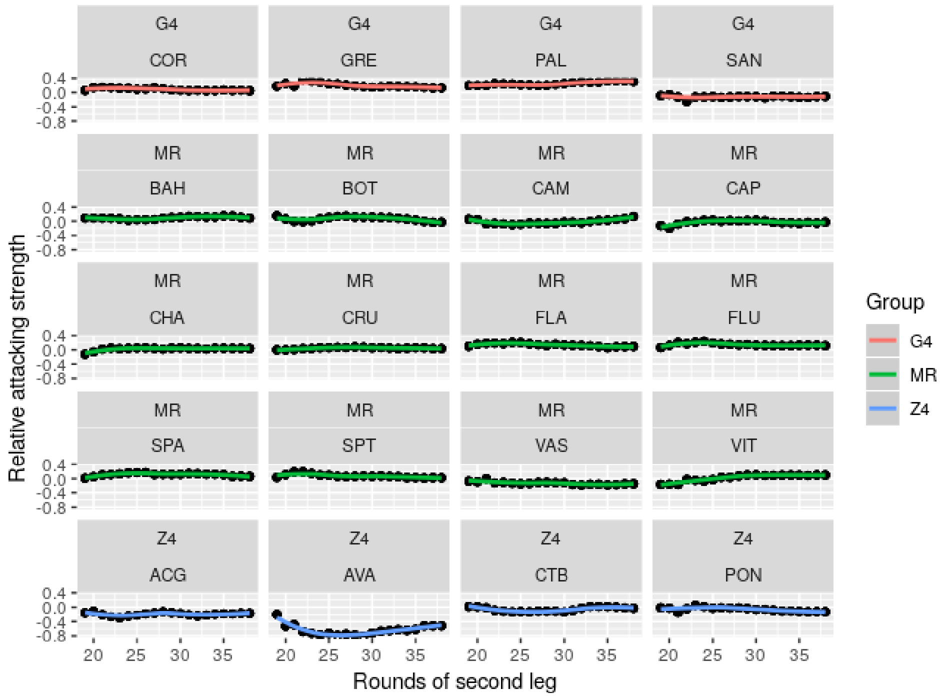

Figure 2 - Estimates of offensive rating ($\hat{\gamma}_i$) along the second leg. Brazilian Football, series A, 2017.

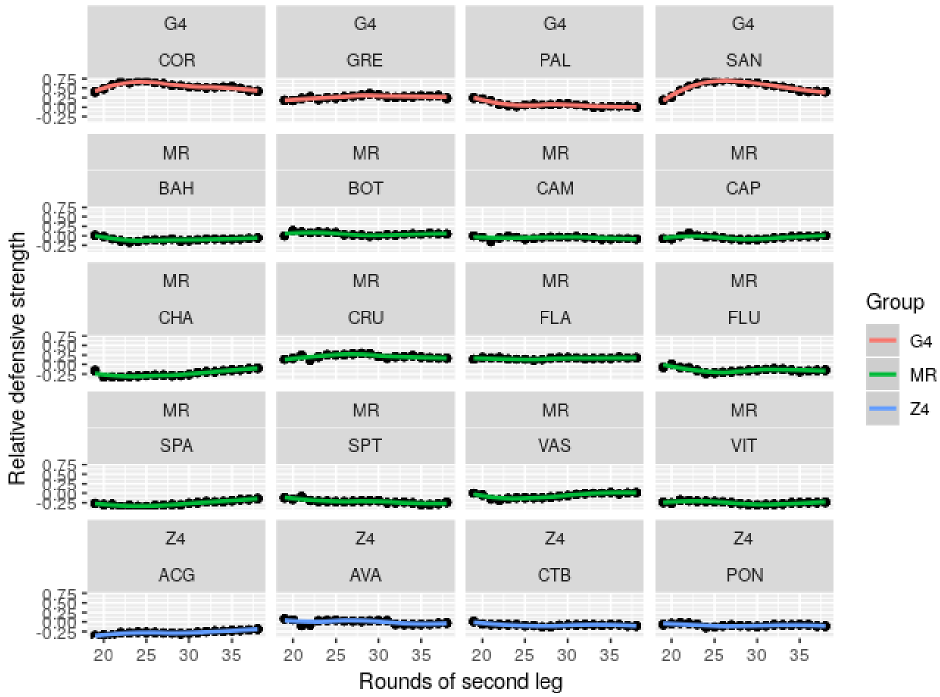

Figure 3 - Estimates of defensive rating ($\hat{\delta}_i$) along the second leg. Brazilian Football, series A, 2017.

Once evaluated attacking and defensive parameters for all clubs, we decided to depict both ratings in a single graph, trying to unravel any pattern of combined strength. Using numeric evaluation of standard errors we also depicted ellipses proportional to confidence regions, although we sacrificed the inference interpretation to depict a better graph. In the pictures ellipses were traced using a fraction (1/10) of their standard errors (Figures 4 and 5). Initials for i^{th} club are in the center of ellipse in the coordinates $(\hat{\gamma}_i, \hat{\delta}_i)$. Black was used for clubs in G4, red for those in MR and green for Z4 ones. It can be seen that ellipses for the second leg are more precise, reflecting larger information on model parameters. In the second leg it is also clear that better clubs somewhat clustered near or within first quadrant, being above average for attack (except Santos) and defense (despite Palmeiras mediocre estimates). Teams demoted in Z4 clustered within third quadrant, being worse in both criteria.

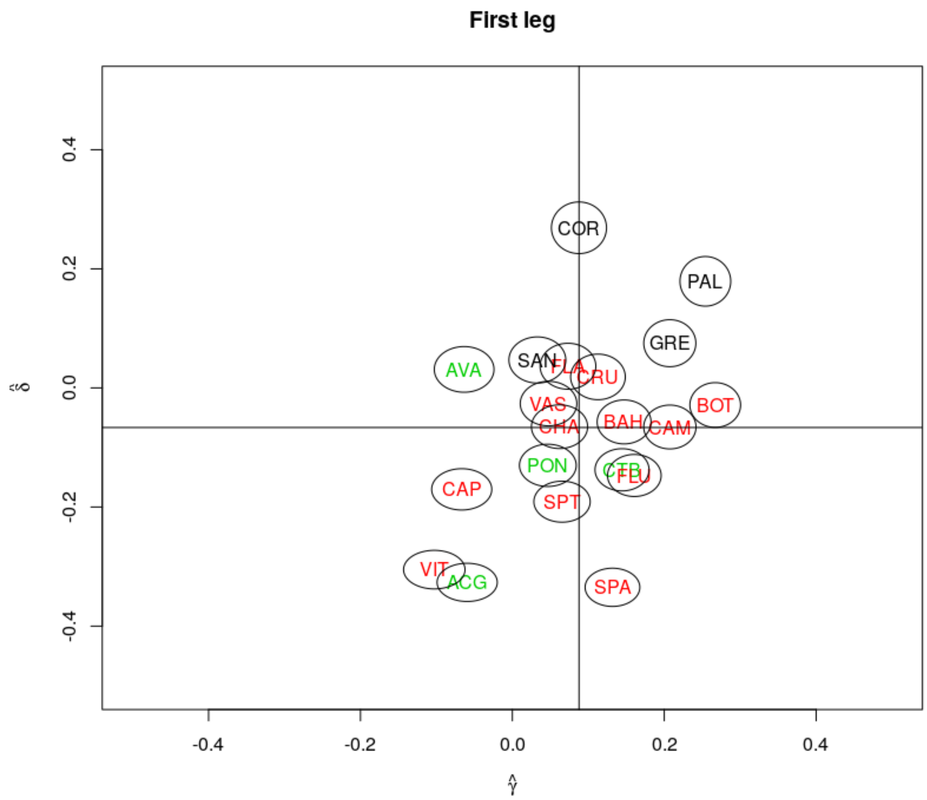

Figure 4 - Ellipses for *ratings* in the first leg of Brazilian Football, series A, 2017.

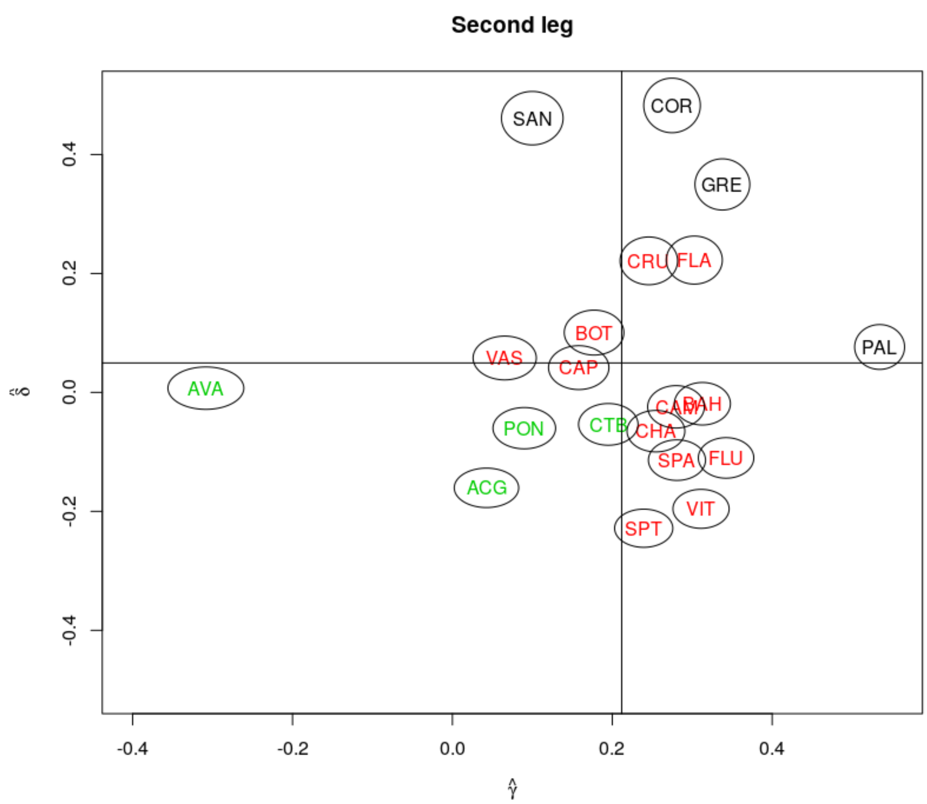

Figure 5 - Ellipses for *ratings* in the second leg of Brazilian Football, series A, 2017.

It is worth to note that Corinthians and Santos made their way through the competition based on defensive skills mainly. Only in the second leg they improved their attacking strength.

Note also that teams with median performance would be placed mainly in the second and fourth quadrants, compensating poor attack with good defense and vice versa.

3.3 Home advantage and probabilities of scores and results

After estimating main parameters for each round it was possible to estimate deviations in results due to an average parameter for home advantage μ. Estimate for this, in the whole tournament, are $\hat{\mu} = 0.31$, and this can have considerable influence in scoring probabilities (and affect game results). For each new set of estimates, a set of predictions $\hat{E[Y]}$ (expected scores) were worked out. It was also possible to work out complete tables of probabilities for each result and predict next rounds' scores. Table 6 brings examples on how $\hat{\mu}$ affects probabilities and expectations. This is done comparing ideal clubs ("A" home team and "B" away team). We considered their attack and defense parameters to be equal ($\gamma_A = \gamma_B = \delta_A = \delta_B = 1$) and three values of home advantage parameter: $\mu = -0.10$, $\mu = 0.31$ and $\mu = 0.50$, in which second value is our estimate for the whole tournament.

Scoring probabilities could also be translated in victory, draw or loss for the home team. Results in second leg were calculated considering probabilities estimated at each previous round for all 190 games of the second leg and their scores. This makes possible to find unexpected results (bad fitting of the model to realized values).

In Table 6 on can see the implications of home advantage parameter. To illustrate the application, for the COR vs SPA game of the second leg we evaluate probabilities for each scoring in Table 7. For presentation purposes only we limited our scores up to 4 goals each. Adding up values from Table 7, we evaluated Table 8, with expected probability of Corinthians winning, drawing or being defeated. Despite the ample favoritism for home team, the game was a draw.

Table 6 - Probabilities of victory (PV), draw (PE for equality) and defeat (PD) of home team

μ	PV	PE	PD
-0.10	31%	32.5%	36.5%
0.31	45%	28%	27%
0.50	52.5%	24.5%	23%
0	34.6%	30.8%	34.6%

Table 7 - Probabilities (%) for each possible score in COR vs SPA. Estimates used: $\hat{\gamma}_{COR} = 0.27$, $\hat{\gamma}_{SPA} = -0.28$, $\hat{\delta}_{COR} = 0.48$, $\hat{\delta}_{SPA} = -0.11$ and $\mu = 0.31$. Lower triangular represents victory of home team (COR), upper triangular for victory of the away team (SPA). Diagonal values add up to draw probability

Gols		Away					
		0	1	2	3	4	>4
Home	0	6.0	5.0	2.0	1.0	0	0
	1	12.0	10.0	4.0	1.0	0	0
	2	12.0	10.0	4.0	1.0	0	0
	3	8.0	7.0	3.0	1.0	0	0
	4	4.0	3.0	1.0	0	0	0
	>4	2.0	2.0	1.0	0	0	0

Table 8 - Probabilities of Victory, draw or defeat of the home team Corinthians playing São Paulo in the second leg. Used estimates: $\hat{\gamma}_{COR} = 0.27$, $\hat{\gamma}_{SPA} = -0.28$, $\hat{\delta}_{COR} = 0.48$, $\hat{\delta}_{SPA} = -0.11$ e $\mu = 0.31$

μ	PV	PE	PD
0.31	65%	21%	14%

The same exercise is repeated in Table 9, that bring probabilities of PAL vs FLU. Table 10 summarizes the results. Game score of 3 to 1 has a hefty 8% expected probability and can be considered quite normal.

Table 9 - Probabilities (%) for scores of home (PAL) and away (FLU) teams. Estimates used: $\hat{\gamma}_{PAL} = -0,53$, $\hat{\gamma}_{FLU} = 0.34$, $\hat{\delta}_{PAL} = 0.08$, $\hat{\delta}_{FLU} = -0.11$ e $\mu = 0.31$; Lower triangular for the home team victory and upper for the away team. In the diagonal are drawing scores

Gols		Away					
		0	1	2	3	4	>4
	0	2.0	3.0	2.0	1.0	0	0
	1	5.0	7.0	4.0	2.0	1.0	0
Home	2	7.0	9.0	6.0	3.0	1.0	0
	3	6.0	8.0	5.0	2.0	1.0	0
	4	4.0	5.0	3.0	1.0	1.0	0
	>4	3.0	4.0	3.0	1.0	0	0

Table 10 - Probabilities of home team winning, drawing or beying beaten. PAL vs FLU, second leg

μ	PV	PE	PD
0.31	66%	17%	17%

Table 11 brings some examples of unexpected results that are known as "zebras" in Brazilian football chronicle. A measure of the log-likelihood for this results can be compared to a critical quantile of χ^2 distribution considering $\nu = 2$ degrees of freedom, as a measure of evidence ($\chi^2_{\nu=2} = 5.99$). We used this method to detect surprising scores and results for each game. In that games, only game 378 can be considered a true "zebra" as Cruzeiro was having a dreadful tournament. Other games in this table had a rather unexpected score but cannot be considered "zebras".

Table 11 - Results of second leg and associated probabilities. Reference value for deciding on surprise effect: $\chi^2_{\nu=2} = 5.99$

J	$Y_m.$	M	$Y_v.$	V	PV	PE	PD	χ^2_c
214	2	SAN	0	COR	0.36	0.34	0.30	3.19
220	2	CRU	2	AVA	0.65	0.23	0.12	4.52
375	1	AVA	1	BOT	0.27	0.29	0.15	5.09
378	**1**	**ATG**	**2**	**CRU**	**0.66**	**0.20**	**0.14**	**6.03**

Conclusion

Maximum likelihood estimates of rating effects considering two parameters (attack and defense) for each club are a powerful tool to estimate ratings and predict both expected scores and results from games. Home advantage could be worked out conditioning on point estimates and works really well to enhance predictions.

It is possible to implement this in a straightforward fashion for a round robin tournament like Brazilian championship. Estimates could be updated in round basis and subsidize sports chronicle with more substantial material. Predictions were rather stable, although this is not too relevant for sports in general.

The model could be enhanced in many ways, specially by taking into account strategies to imput prior knowledge via Bayesian inference.

GALVÃO, L. R.; BUENO, J. S. S. Estimação de rating no futebol: Campeonato Brasileiro de 2017. *Rev. Bras. Biom.*, Lavras, v.38, n.1, p.1-17, 2020.

- *RESUMO: Com os jogos da série A do Campeonato Brasileiro de Futebol, 2017, ajustou-se modelos para preditores lineares da esperança do número de gols do time mandante e do time visitante. Estimou-se valores para parâmetros referentes à força de ataque (γ) e de defesa (δ) de cada time. Foi condicionalmente estimado um efeito para o mando de campo (μ) comum a todos os clubes. O primeiro turno da competição foi usado para obter estimativas iniciais, que por sua vez foram reestimadas no segundo turno, a cada rodada. Pretendeu-se com esse estudo apresentar um modelo que permite tanto descrever a performance passada quanto predizer as probabilidades associadas aos resultados de jogo e respectivos placares. Como aspectos adicionais pode-se prever a classificação final do campeonato e resultados derivados como as probabilidades de classificação para a Copa Libertadores da América ou de ser rebaixado para a segunda divisão da competição. O modelo proposto se mostrou flexível em sua versão inferência de verossimilhança, mas propriedades desejáveis de uma implementação paramétrica bayesiana podem expandir sua capacidade.*

- *PALAVRAS-CHAVE: Campeonato de turnos; distribuição Poisson; força relativa; futebol brasileiro; mando de campo; predição.*

References

ANDRADE, M. T. D.; ESPIRITO SANTO, L. C. Relação entre ações finais que resultaram em gol e o mando de campo no Campeonato Brasileiro de 2009. *Revista Brasileira de Ciências do Esporte*, v.38, n.4, p.363-369, 2016.

BAIO, G.; BLANGIARDO, M. Bayesian hierarchical model for the prediction of football results. *Journal of Applied Statistics*, v.37, n.2, p.253-264, 2010.

GAMERMAN, D; SANTOS, T. R. dos; FRANCO, G. C. A non-gaussian family of state-space models with exact marginal likelihood *Journal of Time Series*, v.34, n.6, p.625-645, 2013.

GAMERMAN, D.; MOREIRA, A. R. B. Multivariate spatial regression models. *Journal of Multivariate Analisys*, v.91, p.262-281, 2004.

GLICKMAN, M. E.; STERN, H. S. A state-space model for national football league scores *Journal of American Statistical Association*, v.93, n.441, p.25-35, 2012.

KARLIS, D.; NTZOUFRAS, I. Bayesian modelling of football outcomes: Using the Skellam's distribution for goal difference. *IMA Journal of Management Mathematics*, v.20, n.2, p.133-145, 2009.

NELDER, J.; MEAD, R. A simplex method for functions minimization. *The Computer Journal*, v.7, n.4, p.308-313, 1965.

R CORE TEAM. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org, 2019.

SUZUKI, A. K.; SALAZAR, L. E.; LEITE, J. G.; LOUZADA-NETO, F. A Bayesian approach for predicting match outcomes: The 2006 (Association) football world cup. *Journal of the Operational Research Society*, v.61, n.10, p.1530-1539, 2010.

Recebido em 11.09.2018.

Aprovado após revisão em 06.01.2020.