Evaluation of Football Forecasting Models: 2021 Brazilian Championship Case Study

Flavio Fontanella, Asla Medeiros e Sá, Moacyr Alvim H. B. da Silva

flavio.fontanella@gmail.com, asla.sa@fgv.br, moacyr@fgv.br

FGV/EMAp, Rio de Janeiro, Brazil

LACIAM Latin American Congress on Industrial and Applied Mathematics 2023

Introduction

Football forecasting models are computer programs/algorithms that use data such as previous results, match statistics, player performances and club economics to predict the probabilities for the results of future matches. In the present work, we have collected data (forecasts) from eight different forecasting models during the 2021 Brazilian football season. The forecasts have been analysed and evaluated.

Main definitions

Result for a football match in this context is either a *home win (H)*, a *draw (D)* or a *away win (A)*. A *forecast* for a match result is a triplet of probabilities (p_H, p_D, p_A), constrained to $p_H + p_D + p_A = 1$, where each p is the probability associated with the corresponding result. A match result forecast (p_H, p_D, p_A) may be visually represented by a point in $\mathbb{R}^3$, lying on the triangle with vertices $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, referred to as *2-simplex*. [1] A match result may be associated with the forecast that assigns probability 1 to that result (and 0 to the others), so being identified with a vertex of the 2-simplex. To plot the forecasts into 2-dimensional charts, we apply the linear mapping:

$$\begin{pmatrix} -\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{3} & -\frac{\sqrt{6}}{6} \end{pmatrix} \quad (1)$$

which preserves distances for points on the *2-simplex*.

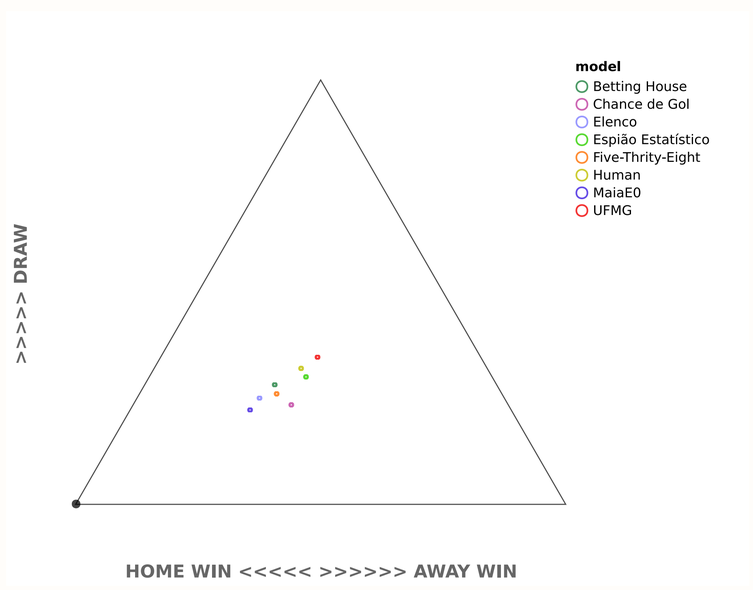

Figure 1: Forecast projections into the 2-simplex representation. Several forecasts and the actual result for a single match as the larger black dot.

Models' forecasts

Figure 2 shows the forecasts for all matches in the championship, each chart faceting one model. [3, 4] [5, 6, 7]

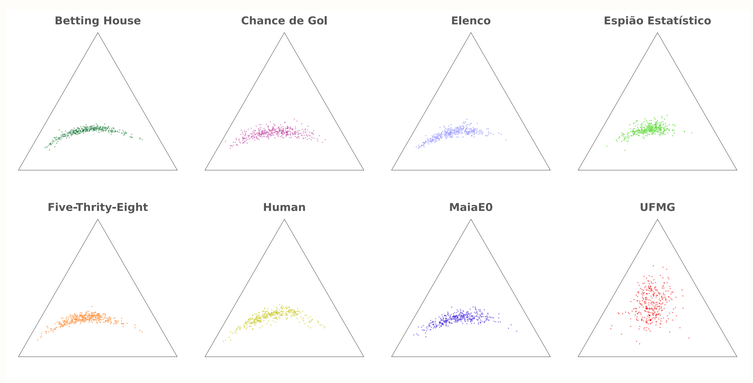

Figure 2: Forecast projections into the 2-simplex.

Observing the pattern formed by the cloud of points, we notice that most points lie close to an arc that starts near the home-win result, ends not so near the away-win result, and passes through the center of the triangle - that represents the naive $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ forecast. All but one model seem to follow this arc pattern, each with its own subpatterns: some more biased towards the home-win result, some displaying a shorter arc, some narrower, some more spread.

Table 1: Mean statistics for the models' forecasts. It is possible to confirm some impressions derived from the visual patterns observed in Figure 2.

Model	H	D	A	|H-A|
Betting House	0.432	0.279	0.289	0.231
Chance de Gol	0.434	0.267	0.298	0.233
Elenco	0.484	0.263	0.253	0.262
Espião Estatístico	0.408	0.294	0.298	0.150
Five-Thirty-Eight	0.445	0.268	0.285	0.225
Human	0.437	0.288	0.275	0.241
MaiaE0	0.443	0.273	0.284	0.224
UFMG	0.349	0.376	0.275	0.125

Forecasts evaluation

Evaluation of a forecast is an assessment of the quality of the forecast. A measure of closeness between the forecast and the actual result of the match can play the role of an *evaluator*. We adopt a scoring rule named *Ranked Probability Score (RPS)* as evaluator, defined as follows:

$$\mathrm{RPS}(forecast) = \frac{(p_H - e_H)^2 + (p_A - e_A)^2}{2} \quad (2)$$

where e_H is 1 if the actual match result is a home win and 0 for other results, while e_A is 1 for a away win and 0 otherwise. The RPS for a set of forecasts is obtained by averaging the RPS for each individual forecast and has been considered an appropriate score in football results forecasting context. [2]

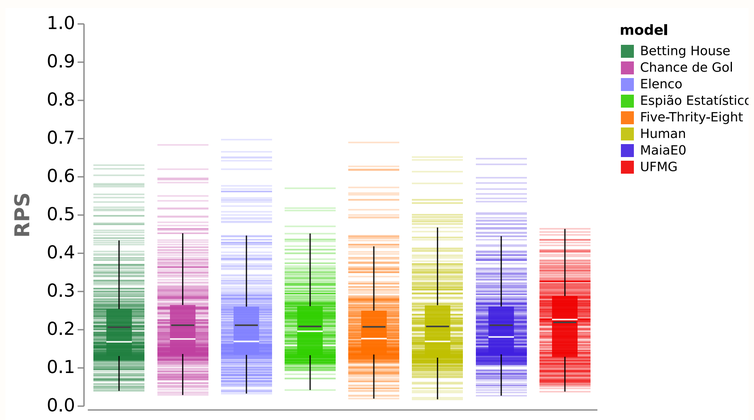

Figure 3: RPS evaluation for all models. Wide ticks represent evaluations for individual matches. The white tick over the box-plot marks the median and the black tick marks the mean evaluation.

Comparing to the previous figure, it's clear that models biased towards the middle of the 2-simplex get less spread evaluations. It doesn't lead necessarily to lower means or medians, as we can observe in the box-plot marks. It's arguable that all models performed relatively close, with ranking models' performance leading to millesimal scale comparison.

Table 2: Final mean RPS.

Model	RPS
Betting House	0.205
Chance de Gol	0.2097
Elenco	0.2096
Espião Estatístico	0.2074
Five-Thirty-Eight	0.2054
Human	0.2069
MaiaE0	0.2103
UFMG	0.2179

We also observed how the mean RPS evolved through time as the championship matches were played.

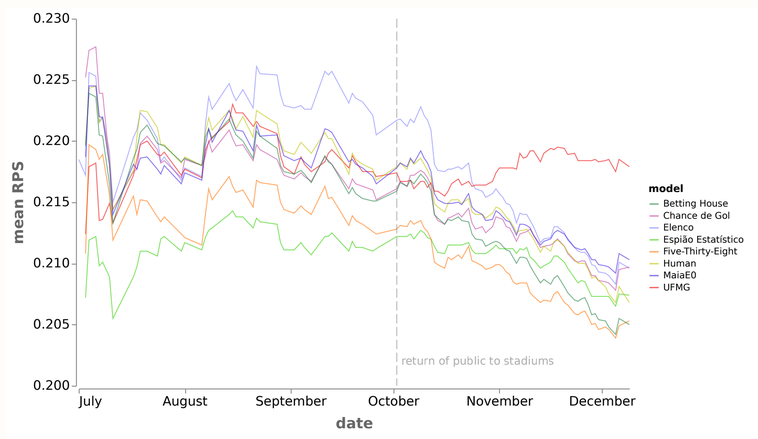

Figure 4: Time series of mean RPS for each model.

The chart presents an important event that may have caused great impact on forecasts performances: the return of the public to the stadiums on October 2nd, as the public presence seems to have increased home advantage, thus benefiting models with higher home-win bias.

Discussion and conclusion

Considering the arc pattern from the forecasts, we wonder if this arrangement is entirely due to the adopted mathematical modelling or if it could be intrinsic to the nature of the game. If it's the latter, football results forecasts could be reduced to one single variable, which would change mathematical basis of forecasting models and how we evaluate such forecasts. Since forecasts occur within a small region of the *2-simplex*, it's not surprising that the evaluations of models' performances were so close to one another. Figure 4 suggests though that models' ranking may become stable in time, with score variations likely to occur to all models similarly.

References

[1] E. S. Epstein and A. H. Murphy, A note on the attributes of probabilistic predictions and the probability score, *Journal of Applied Meteorology, 1965*

[2] A. Constantinou and N. Fenton, Solving the Problem of Inadequate Scoring Rules for Assessing Probabilistic Football Forecast Models, *JQAS, 2012*

[3] F. Fontanella, Métodos de avaliação de modelos de previsão de resultados de futebol, *Master's thesis, FGV/EMAP, 2021*

[4] L. F. G. N. Maia, Modelos in-play para partidas do campeonato brasileiro de futebol, *Master's thesis, FGV/EMAP, 2021*

[5] https://www.mat.ufmg.br/futebol/

[6] https://www.chancedegol.com.br/

[7] https://projects.fivethirtyeight.com/previsoes-de-futebol/brasileirao/