13th International Conference on Applied Informatics
Eszterházy Károly Catholic University
Eger, Hungary, February 19–21, 2026

Expected Goals Integration in Form-Based Football Models: Extensions of the Veto and Balance Models

József Gergő Pál[a], **Csaba Bíró**[ab]

[a]Eötvös Loránd University
paljozsefgergo@gmail.com

[b]Eszterházy Károly Catholic University
biro.csaba@uni-eszterhazy.hu

Abstract

Form-based prediction models estimate football match outcomes by analyzing teams' recent performance histories. Our previous work [4] introduced the Veto and Balance models, which compute outcome probabilities from exponentially weighted win/draw/loss distributions derived from recent match results. While effective, these models rely exclusively on final scores, potentially missing crucial information about match quality and team performance. Expected Goals (xG) addresses this limitation by quantifying the quality of scoring opportunities based on shot characteristics such as distance, angle, and game context [3, 5]. Research demonstrates that xG provides a more stable and predictive measure of team strength than actual goals [1, 2, 6].

This paper systematically integrates xG metrics into form-based Veto and Balance models through three conceptually distinct approaches. The **Shift** method adjusts result-based probabilities using teams' underlying xG performance, tilting predictions toward teams with superior expected goal statistics. The **Blend** method constructs dual form distributions from both actual results and xG data, then combines them to leverage information from both sources. The **Luck** method reweights historical matches based on the divergence between actual outcomes and expected goal performance, emphasizing sustainable results while discounting anomalous performances. We investigate which xG integration strategy and model framework (asymmetric Veto or symmetric Balance) best captures the predictive

value of expected goal data.

For the Blend and Luck approaches, we implement multiple parameter configurations to explore different trade-offs between result-based and xG-based information. This yields 14 xG-enhanced variants, enabling a comprehensive comparison of how information weighting affects prediction quality across different modeling paradigms. We evaluate these models on international football leagues with complete xG statistics, measuring both probabilistic accuracy metrics and profitability in value betting scenarios. The study addresses whether expected goal integration consistently improves predictions across diverse league characteristics and competitive levels.

Keywords: expected goals, xG, form-based models, football prediction, sports analytics, Veto model, Balance model, probabilistic forecasting

References

[1] A. C. Constantinou, N. E. Fenton: *Solving the problem of inadequate scoring rules for assessing probabilistic football forecast models*, Journal of Quantitative Analysis in Sports 8.1 (2012), doi: 10.1515/1559-0410.1418.

[2] H. Eggels, R. van Elk, M. Pechenizkiy: *Expected Goals in Soccer: Explaining Match Results using Predictive Analytics*, in: Machine Learning and Data Mining for Sports Analytics Workshop, ECML/PKDD 2016, 2016.

[3] I. G. McHale, Ł. Szczepański: *Beyond Expected Goals*, Proceedings of the MIT Sloan Sports Analytics Conference (2014).

[4] J. G. Pál, C. Bíró: *Evaluating profitability in sports betting using probabilistic models and betting strategies*, Annales Mathematicae et Informaticae 61 (2025), pp. 202–214, issn: 1787-6117, doi: 10.33039/ami.2025.10.016, url: https://doi.org/10.33039/ami.2025.10.016.

[5] R. Pollard, J. Ensum, S. Taylor: *Estimating the probability of a shot resulting in a goal: The effects of distance, angle and space*, International Journal of Soccer Science 2 (2004), pp. 50–55.

[6] A. Rathke: *An examination of expected goals and shot efficiency in soccer*, Journal of Human Sport and Exercise 12.2proc (2017), S514–S529, doi: 10.14198/jhse.2017.12.Proc2.05.