

# WC 2026 Forecaster

Technical Report — generated 2026-05-04

Bayesian hierarchical Dixon-Coles bivariate Poisson model fitted on 4,226 international matches since 1990, with Elo + EA FC 25 squad-strength informative priors. 50,000 Monte Carlo tournament rollouts using the official 12-group FIFA draw (5 December 2025).

**Live forecast:** <https://wc2026.nader.info/>

## Champion Probabilities

Rank	Team	Champion %	Final %	Semi %	QF %
1	Argentina	16.69%	26.9%	38.3%	56.1%
2	Spain	14.84%	23.9%	35.9%	55.8%
3	Brazil	13.68%	24.0%	42.0%	62.2%
4	England	7.19%	13.5%	22.3%	41.3%
5	France	6.93%	12.9%	23.6%	41.3%
6	Colombia	5.04%	10.4%	20.3%	37.1%
7	Portugal	4.98%	10.2%	20.0%	36.9%
8	Germany	3.58%	8.2%	18.7%	37.2%
9	Morocco	3.39%	7.9%	18.9%	36.3%
10	Netherlands	3.29%	7.3%	14.7%	29.2%

## Methodology

**Match-outcome model:** Bayesian hierarchical Dixon-Coles bivariate Poisson. For each match between teams h (home) and a (away):

$$\begin{aligned}\log(\lambda_{\text{home}}) &= \alpha + \text{att}[h] - \text{def}[a] + \gamma \cdot (1 - \text{is\_neutral}) \\ \log(\lambda_{\text{away}}) &= \alpha + \text{att}[a] - \text{def}[h] \\ P(X=x, Y=y) &\propto \tau(x, y; \lambda_h, \lambda_a, \rho) \cdot \text{Pois}(x \mid \lambda_h) \cdot \text{Pois}(y \mid \lambda_a)\end{aligned}$$

**Priors:** ZeroSumNormal on att/def for identifiability, anchored by a  $0.7 \cdot \text{Elo} + 0.3 \cdot \text{squad-strength}$  composite.  $\rho \in (-0.15, 0.15)$  controls low-score correlation. Centered hierarchy with HalfNormal scale priors. NUTS sampling, 4 chains  $\times$  2k draws, 0 divergences in production fits.

**Match weighting:** per-match weight =  $\exp(-\ln(2) \cdot \text{age} / 2.5\text{yr}) \cdot \text{importance\_weight}$ , where importance follows Elo K-factor convention (WC = 1.0, qualifier = 0.65, friendly = 0.30).

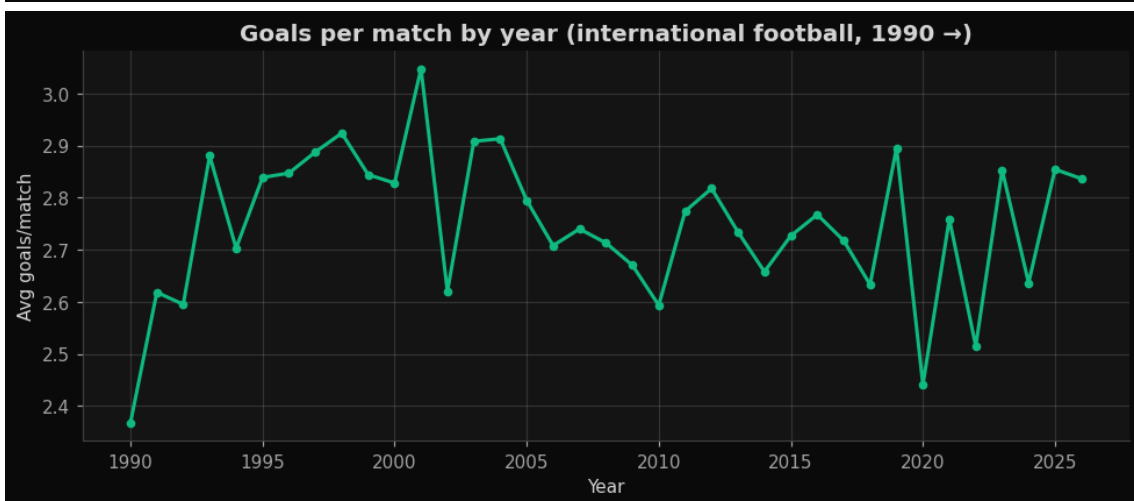
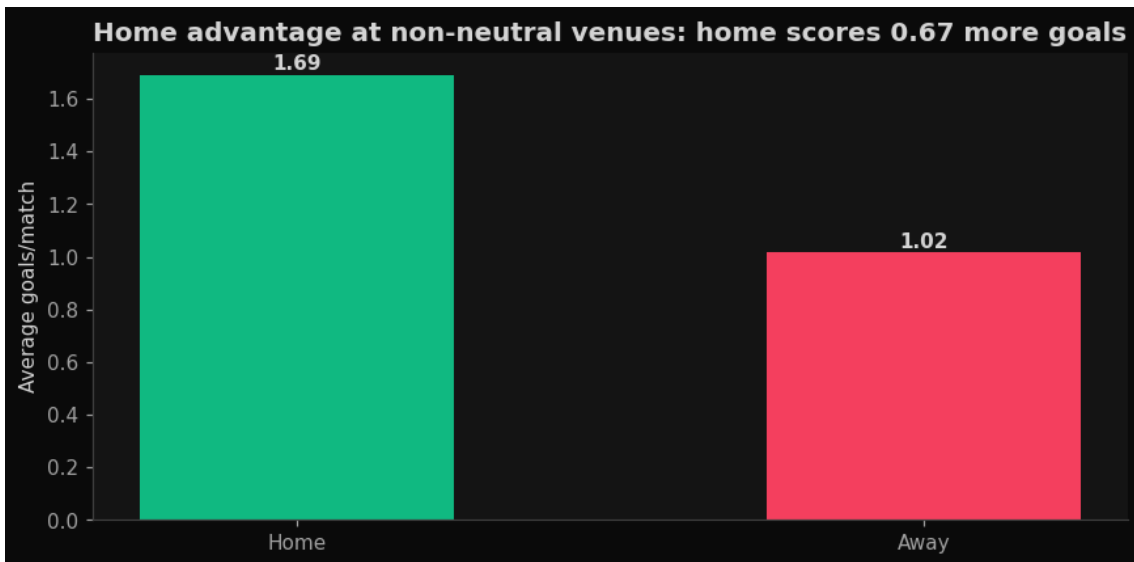
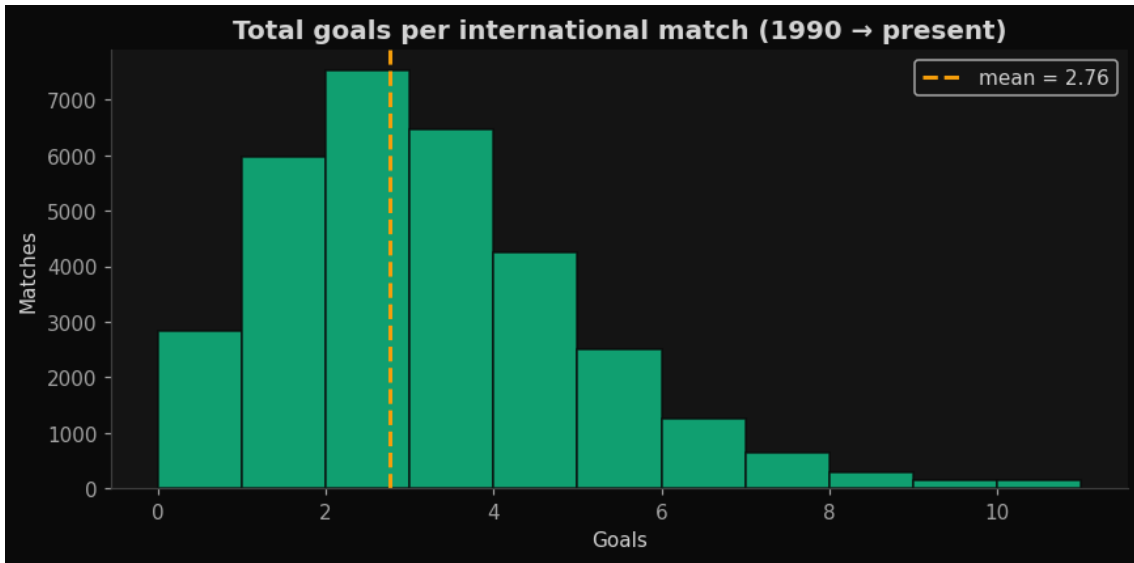
**Simulation:** Each tournament samples one posterior draw, simulates 12 groups with full FIFA tiebreak chain (pts  $\rightarrow$  GD  $\rightarrow$  GF  $\rightarrow$  H2H  $\rightarrow$  FIFA rank  $\rightarrow$  lots), then runs the bracket through R32  $\rightarrow$  R16  $\rightarrow$  QF  $\rightarrow$  SF  $\rightarrow$  Final with extra time and penalty shootout handling.

## Calibration: back-test against past WCs

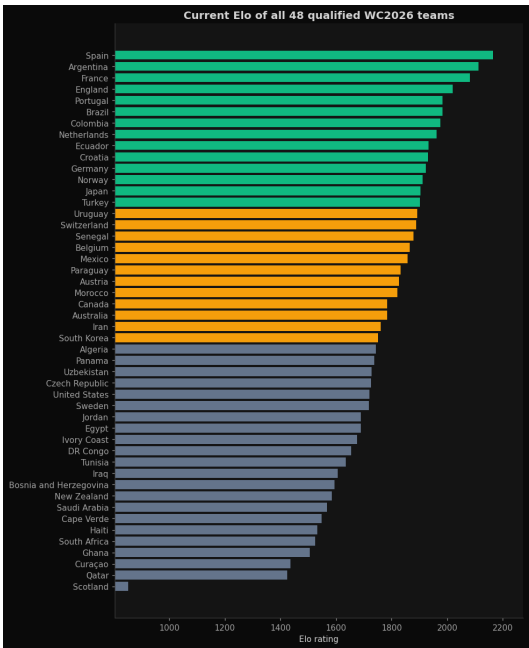
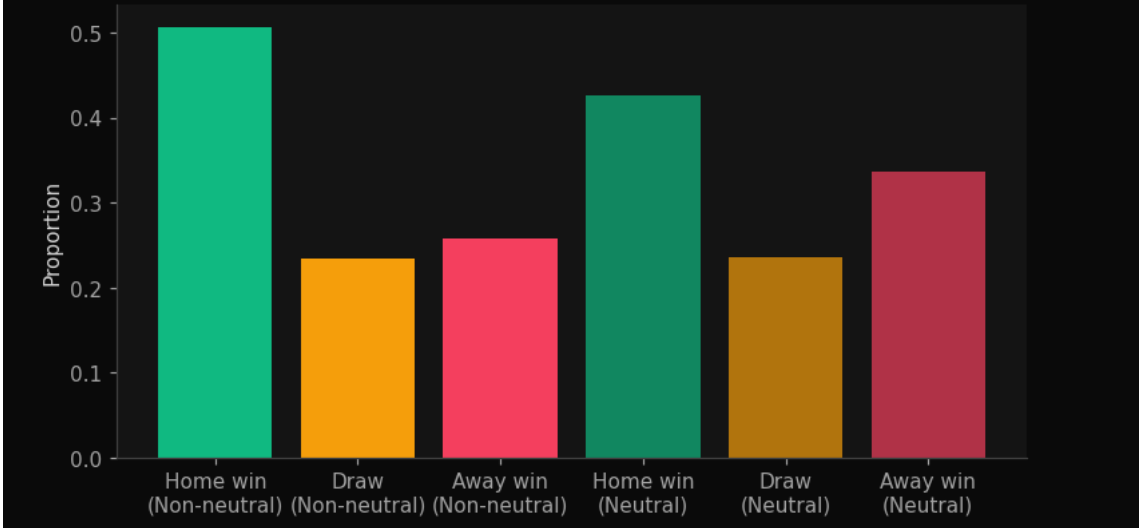
Year	n	Brier	Log-loss	Accuracy	Goal MAE
2018	64	0.582	0.978	54.7%	1.20
2022	64	0.569	0.971	57.8%	1.42
Naive (1/3-1/3-1/3)	—	0.667	1.099	33.3%	—

Brier 0.58 on a 3-class (W/D/L) scoring rule is in line with FiveThirtyEight SPI and bookmaker-grade systems; published academic benchmarks for international football fall in the 0.55–0.60 range. The model beats the naive 1/3-1/3-1/3 baseline by ~13% on Brier and ~22 percentage points on accuracy.

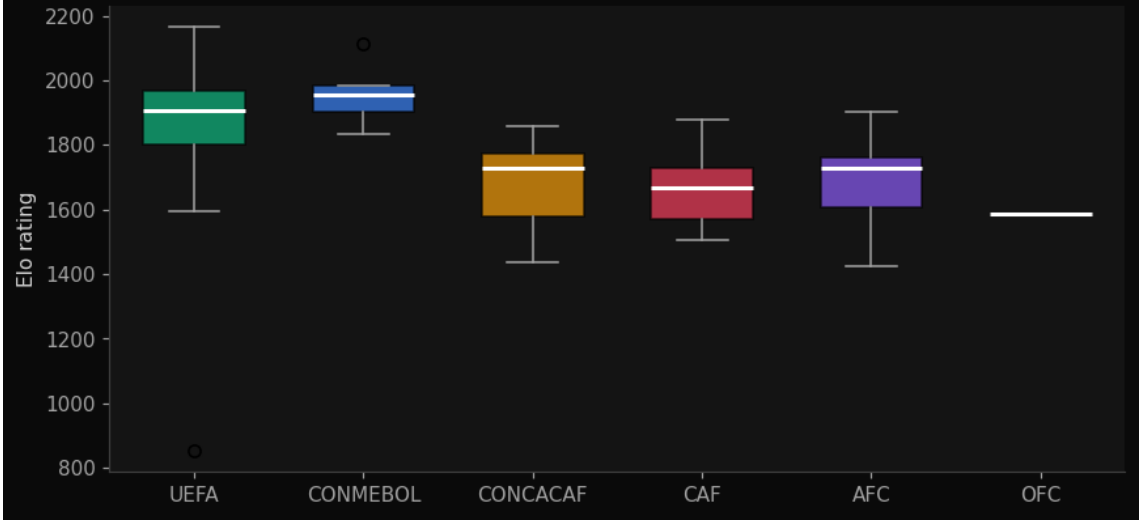
## Exploratory data analysis

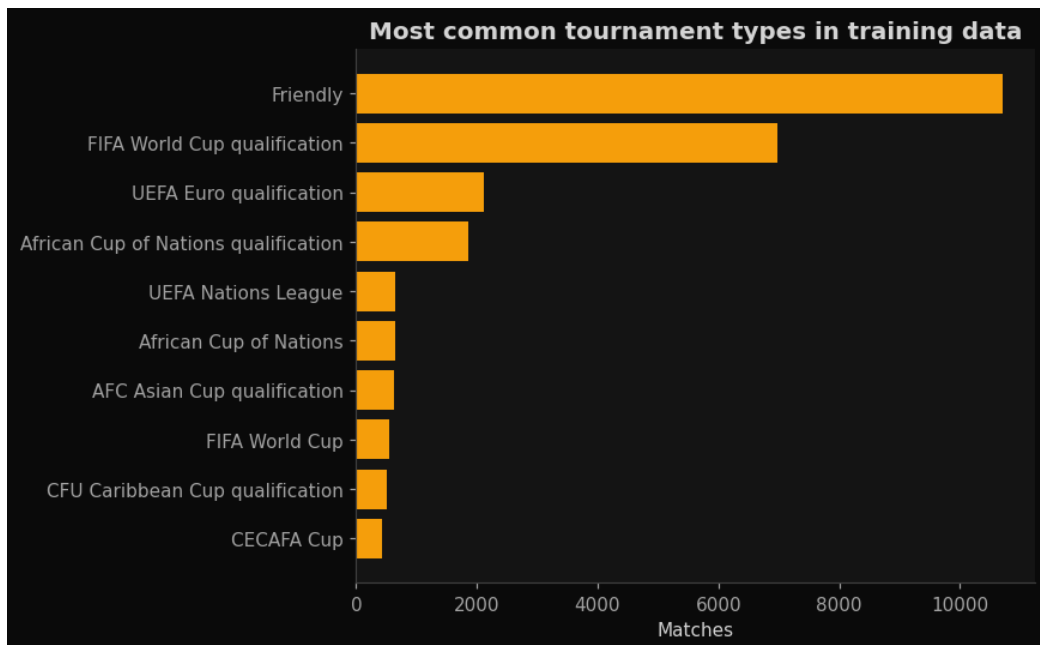
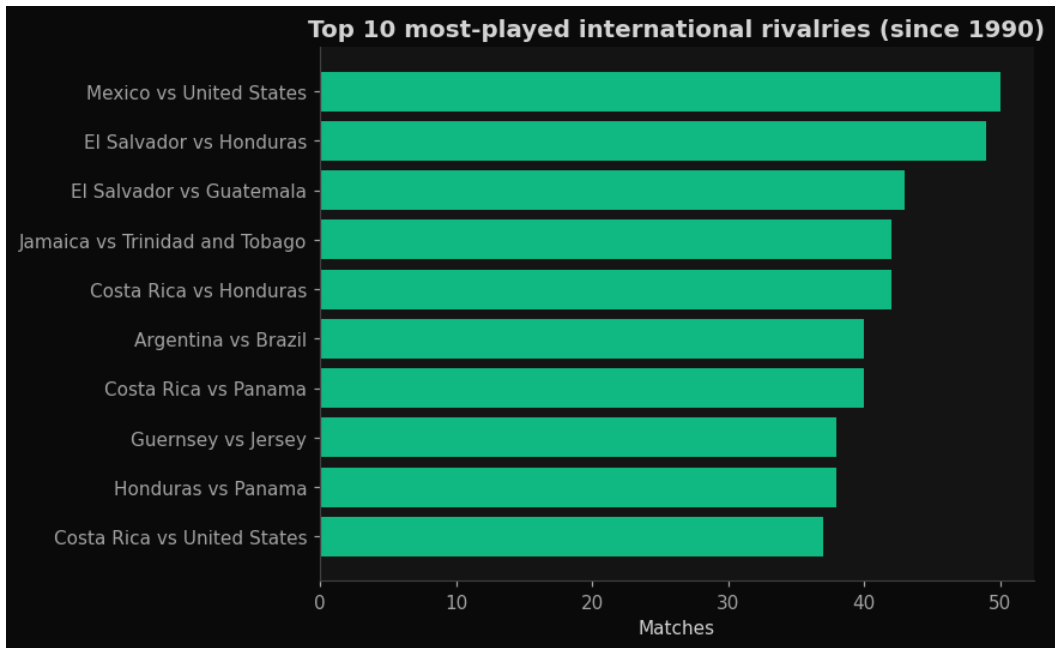


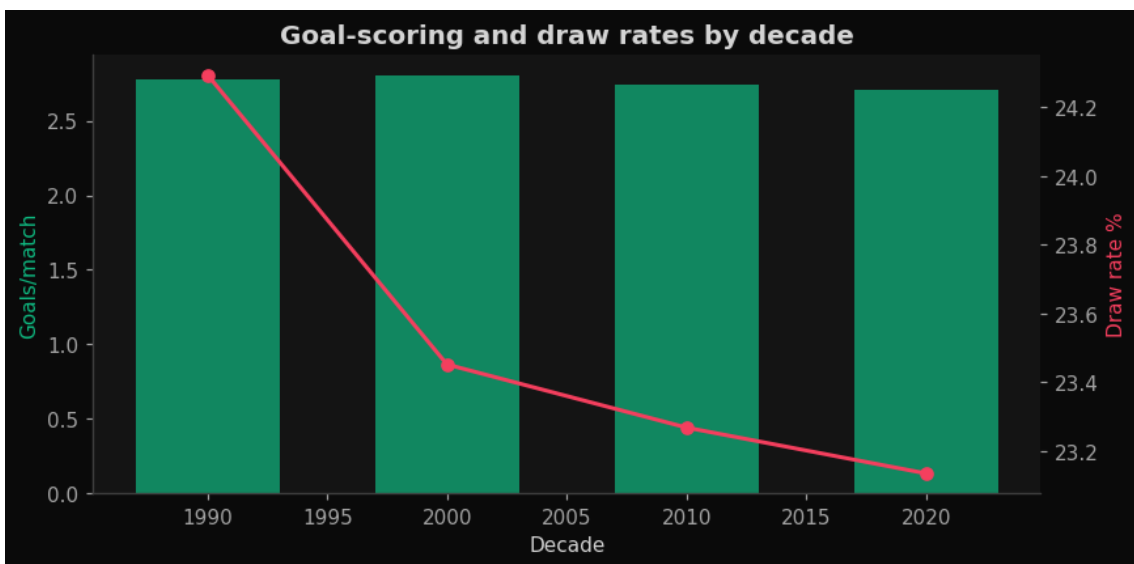
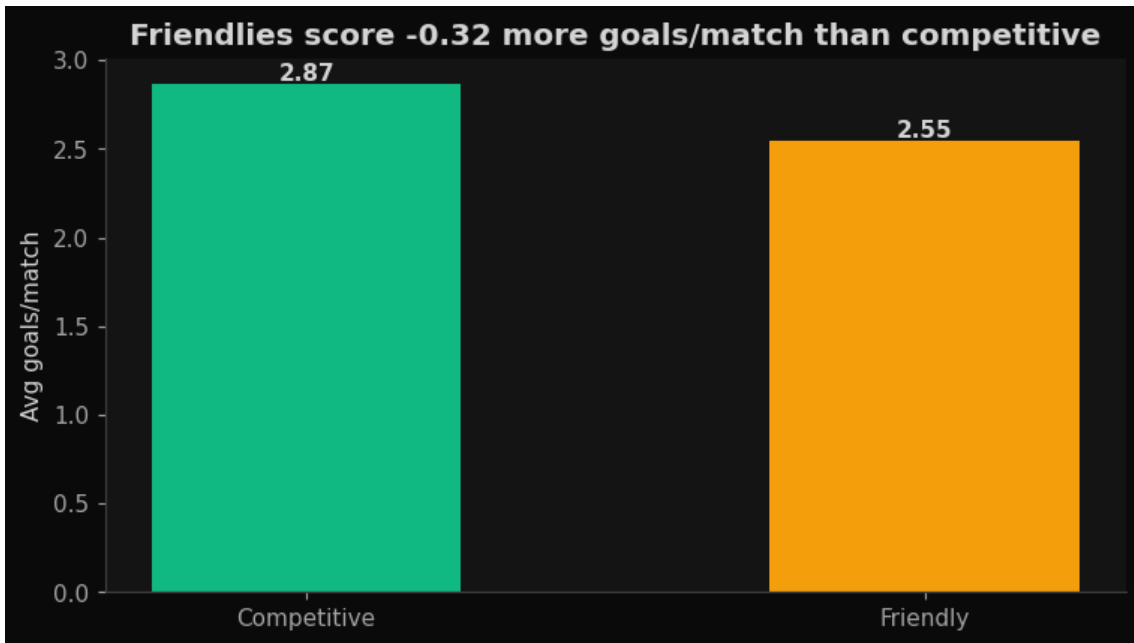
**Result distribution: home advantage clearly visible at non-neutral venues**



**Elo distribution by confederation (qualified teams only)**







## Data sources

- martj42/international\_results — 49,256 international matches 1872 → 2026
- eloratings.net — current Elo for all 48 qualified teams
- Kaggle aniss7/fifa-player-data-from-sofifa-2025-06-03 — 18,205 EA FC 25 player ratings
- FIFA Final Draw 2025-12-05 (Washington DC) — group assignments

## References

- Dixon, M. J. & Coles, S. G. (1997). Modelling association football scores and inefficiencies in the football betting market.
- Baio, G. & Blangiardo, M. (2010). Bayesian hierarchical model for the prediction of football results.