Who Will (Most Likely) Win the 2018 FIFA World Cup?

Achim Zeileis

https://eeecon.uibk.ac.at/~zeileis/
Tournament forecast based on bookmakers odds.
Main results: Brazil and Germany are the top favorites with winning probabilities of 16.6% and 15.8%, respectively.
Top favorites are most likely to meet in the final (5.5%), then with odds very slightly in favor of Brazil (50.6% winning probability).
Bookmakers odds: Motivation

Forecasts of sports events:

- Increasing interest in forecasting of competitive sports events due to growing popularity of online sports betting.
- Forecasts often based on ratings or rankings of competitors’ ability/strength.

In football:

- Elo rating.
  - Aims to capture relative strength of competitors yielding probabilities for pairwise comparisons.
  - Originally developed for chess.
- FIFA rating.
  - Official ranking, used for seeding tournaments.
  - Often criticized for not capturing current strengths well.
  - June 2018: Decision to change calculation to be more similar to Elo.
Alternatively: Employ bookmakers odds for winning a competition.

- Bookmakers are “experts” with monetary incentives to rate competitors correctly. Setting odds too high or too low yields less profits.
- Prospective in nature: Bookmakers factor not only the competitors abilities into their odds but also tournament draws/seedings, home advantages, recent events such as injuries, etc.
- Statistical “post-processing” needed to derive winning probabilities and underlying abilities.
Bookmakers odds: Statistics

**Odds:** In statistics, the ratio of the probabilities for/against a certain event,

\[ \text{odds} = \frac{p}{1 - p} \].
Bookmakers odds: Statistics

**Odds**: In statistics, the ratio of the probabilities for/against a certain event,

\[ odds = \frac{p}{1 - p}. \]

**Illustrations:**
- Even odds are “50:50” (= 1).
- Odds of 4 correspond to probabilities 4/5 = 80% vs. 1/5 = 20%.
Bookmakers odds: Statistics

**Odds:** In statistics, the ratio of the probabilities for/against a certain event,

\[
\text{odds} = \frac{p}{1 - p}.
\]

**Illustrations:**
- Even odds are “50:50” (= 1).
- Odds of 4 correspond to probabilities 4/5 = 80% vs. 1/5 = 20%.

**Thus:** Odds can be converted to probabilities and vice versa.

\[
p = \frac{\text{odds}}{\text{odds} + 1}
\]

\[
1 - p = \frac{1}{\text{odds} + 1}
\]
Quoted odds: In sports betting, the payout for a stake of 1.
Quoted odds: In sports betting, the payout for a stake of 1.

Fair bookmaker: Given the probability $p$ for the event the bookmaker could set

$$quoted \ odds = \frac{1-p}{p} + 1.$$
Bookmakers odds: Quoted odds

**Quoted odds:** In sports betting, the payout for a stake of 1.

**Fair bookmaker:** Given the probability $p$ for the event the bookmaker could set

$$\text{quoted odds} = \frac{1 - p}{p} + 1.$$  

**Expected payout:** Wins and losses cancel out each other.

$$p \cdot \frac{1 - p}{p} - (1 - p) \cdot 1 = 0.$$
Bookmakers odds: Quoted odds

**Quoted odds:** In sports betting, the payout for a stake of 1.

**Fair bookmaker:** Given the probability $p$ for the event the bookmaker could set

$$quoted\ odds = \frac{1 - p}{p} + 1.$$ 

**Expected payout:** Wins and losses cancel out each other.

$$p \cdot \frac{1 - p}{p} - (1 - p) \cdot 1 = 0.$$ 

**Thus:** “Naive” computation of probability

$$p = \frac{1}{quoted\ odds}.$$
### Bookmakers odds: Quoted odds

**Illustration:** Quoted odds for bwin obtained on 2018-05-20.

<table>
<thead>
<tr>
<th>Team</th>
<th>Quoted odds</th>
<th>“Naive” probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>5.0</td>
<td>0.200</td>
</tr>
<tr>
<td>Germany</td>
<td>5.5</td>
<td>0.182</td>
</tr>
<tr>
<td>Spain</td>
<td>7.0</td>
<td>0.143</td>
</tr>
<tr>
<td>France</td>
<td>7.5</td>
<td>0.133</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>501.0</td>
<td>0.002</td>
</tr>
<tr>
<td>Panama</td>
<td>1001.0</td>
<td>0.001</td>
</tr>
</tbody>
</table>
**Bookmakers odds: Quoted odds**

**Illustration:** Quoted odds for bwin obtained on 2018-05-20.

<table>
<thead>
<tr>
<th>Team</th>
<th>Quoted odds</th>
<th>“Naive” probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>5.0</td>
<td>0.200</td>
</tr>
<tr>
<td>Germany</td>
<td>5.5</td>
<td>0.182</td>
</tr>
<tr>
<td>Spain</td>
<td>7.0</td>
<td>0.143</td>
</tr>
<tr>
<td>France</td>
<td>7.5</td>
<td>0.133</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>501.0</td>
<td>0.002</td>
</tr>
<tr>
<td>Panama</td>
<td>1001.0</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**Problem:** Probabilities of all 32 teams sum to $1.143 > 1$. 
Bookmakers odds: Adjustment

**Reason:** Bookmakers do not give honest judgment of winning chances but include a profit margin known as “overround”.

**Simple solution:** Adjust quoted odds by factor 1.143 so that probabilities sum to 1.

<table>
<thead>
<tr>
<th>Team</th>
<th>Adjusted odds</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>5.71</td>
<td>0.175</td>
</tr>
<tr>
<td>Germany</td>
<td>6.28</td>
<td>0.159</td>
</tr>
<tr>
<td>Spain</td>
<td>8.00</td>
<td>0.125</td>
</tr>
<tr>
<td>France</td>
<td>8.57</td>
<td>0.117</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Bookmakers odds: Adjustment

**Reason:** Bookmakers do not give honest judgment of winning chances but include a profit margin known as “overround”.

**Simple solution:** Adjust quoted odds by factor 1.143 so that probabilities sum to 1.

<table>
<thead>
<tr>
<th>Team</th>
<th>Adjusted odds</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>5.71</td>
<td>0.175</td>
</tr>
<tr>
<td>Germany</td>
<td>6.28</td>
<td>0.159</td>
</tr>
<tr>
<td>Spain</td>
<td>8.00</td>
<td>0.125</td>
</tr>
<tr>
<td>France</td>
<td>8.57</td>
<td>0.117</td>
</tr>
</tbody>
</table>
Refinement: Apply adjustment only to the odds, not the stake.

\[ \text{quoted odds}_i = \text{odds}_i \cdot \delta + 1, \]

- where \( \text{odds}_i \) is the bookmaker’s “true” judgment of the odds for competitor \( i \),
- \( \delta \) is the bookmaker’s payout proportion (overround: \( 1 - \delta \)),
- and \(+1\) is the stake.
Winning probabilities: The adjusted odds \( odds_i \) then corresponding to the odds of competitor \( i \) for losing the tournament. They can be easily transformed to the corresponding winning probability

\[
p_i = \frac{1}{odds_i + 1}.
\]

Determining the overround: Assuming that a bookmaker’s overround is constant across competitors, it can be determined by requiring that the winning probabilities of all competitors (here: all 32 teams) sum to 1:

\[
\sum_i p_i = 1.
\]
Data processing:

- Quoted odds from 26 online bookmakers.
- Computed overrounds $1 - \delta_b$ individually for each bookmaker $b = 1, \ldots, 26$ by unity sum restriction across teams $i = 1, \ldots, 32$.
- Median overround is 15.2%.
- Yields overround-adjusted and transformed winning probabilities $p_{i,b}$ for each team $i$ and bookmaker $b$. 
Modeling consensus and agreement

bwin
bet365
Sky Bet
Ladbrokes
William Hill
Marathon Bet
Betfair Sportsbook
SunBets
Paddy Power
Unibet
Coral
Betfred
Boylesports
Black Type
Betstars
Betway
BetBright
10Bet
Sportingbet
188Bet
888sport
Bet Victor
Sportpesa
Spreadex
Betdaq
Smarkets

BRA
GER
ESP
FRA
ARG
BEL
ENG
POR
URU
CRO
COL
RUS
POL
DEN
MEX
SUI
SWE
EGY
SRB
SEN
NGA
ISL
JPN
AUS
MAR
CRC
KOR
IRN
TUN
KSA
PAN
Modeling consensus and agreement

**Goal:** Get consensus probabilities by aggregation across bookmakers.

**Straightforward:** Compute average for team $i$ across bookmakers.

$$\bar{p}_i = \frac{1}{26} \sum_{b=1}^{26} p_{i,b}.$$
Modeling consensus and agreement

**Goal:** Get consensus probabilities by aggregation across bookmakers.

**Straightforward:** Compute average for team $i$ across bookmakers.

$$\bar{p}_i = \frac{1}{26} \sum_{b=1}^{26} p_{i,b}.$$

**Refinements:**

- Statistical model assuming for latent consensus probability $p_i$ for team $i$ along with deviations $\varepsilon_{i,b}$.
- Additive model is plausible on suitable scale, e.g.,

$$\text{logit}(p) = \log \left( \frac{p}{1 - p} \right).$$
Modeling consensus and agreement

**Model:** Bookmaker consensus model

\[
\text{logit}(p_{i,b}) = \text{logit}(p_i) + \varepsilon_{i,b},
\]

where further effects could be included, e.g., group effects in consensus logits or bookmaker-specific bias and variance in \(\varepsilon_{i,b}\).
Modeling consensus and agreement

**Model:** Bookmaker consensus model

\[ \text{logit}(p_{i,b}) = \text{logit}(p_i) + \varepsilon_{i,b}, \]

where further effects could be included, e.g., group effects in consensus logits or bookmaker-specific bias and variance in \( \varepsilon_{i,b} \).

**Analogously:** Methodology can also be used for consensus ratings of default probability in credit risk rating of bank \( b \) for firm \( i \).
Modeling consensus and agreement

Here:

• Simple fixed-effects model with zero-mean deviations.
• Consensus logits are simply team-specific means across bookmakers:

\[
\hat{\text{logit}}(p_i) = \frac{1}{26} \sum_{b=1}^{26} \text{logit}(p_{i,b}).
\]

• Consensus winning probabilities are obtained by transforming back to the probability scale:

\[
\hat{p}_i = \text{logit}^{-1} \left( \hat{\text{logit}}(p_i) \right).
\]

• Model captures 98.7% of the variance in \( \text{logit}(p_{i,b}) \) and the associated estimated standard error is 0.184.
Modeling consensus and agreement

<table>
<thead>
<tr>
<th>Team</th>
<th>FIFA code</th>
<th>Probability</th>
<th>Log-odds</th>
<th>Log-ability</th>
<th>Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brazil</td>
<td>BRA</td>
<td>16.6</td>
<td>−1.617</td>
<td>−1.778</td>
<td>E</td>
</tr>
<tr>
<td>Germany</td>
<td>GER</td>
<td>15.8</td>
<td>−1.673</td>
<td>−1.801</td>
<td>F</td>
</tr>
<tr>
<td>Spain</td>
<td>ESP</td>
<td>12.5</td>
<td>−1.942</td>
<td>−1.925</td>
<td>B</td>
</tr>
<tr>
<td>France</td>
<td>FRA</td>
<td>12.1</td>
<td>−1.987</td>
<td>−1.917</td>
<td>C</td>
</tr>
<tr>
<td>Argentina</td>
<td>ARG</td>
<td>8.4</td>
<td>−2.389</td>
<td>−2.088</td>
<td>D</td>
</tr>
<tr>
<td>Belgium</td>
<td>BEL</td>
<td>7.3</td>
<td>−2.546</td>
<td>−2.203</td>
<td>G</td>
</tr>
<tr>
<td>England</td>
<td>ENG</td>
<td>4.9</td>
<td>−2.957</td>
<td>−2.381</td>
<td>G</td>
</tr>
<tr>
<td>Portugal</td>
<td>POR</td>
<td>3.4</td>
<td>−3.353</td>
<td>−2.486</td>
<td>B</td>
</tr>
<tr>
<td>Uruguay</td>
<td>URU</td>
<td>2.7</td>
<td>−3.566</td>
<td>−2.566</td>
<td>A</td>
</tr>
<tr>
<td>Croatia</td>
<td>CRO</td>
<td>2.5</td>
<td>−3.648</td>
<td>−2.546</td>
<td>D</td>
</tr>
</tbody>
</table>
Pr($i$ beats $j$) = $\pi_{i,j}$

\[ \frac{\text{ability}_i}{\text{ability}_i + \text{ability}_j} \]

Abilities and tournament simulations

Further questions:

• What are the likely courses of the tournament that lead to these bookmaker consensus winning probabilities?
• Is the team with the highest probability also the strongest team?
• What are the winning probabilities for all possible matches?

Motivation:

• Tournament draw might favor some teams.
• Tournament schedule was known to bookmakers and hence factored into their quoted odds.
• Can abilities (or strengths) of the teams be obtained, adjusting for such tournament effects?
Abilities and tournament simulations

**Answer:** Yes, an approximate solution can be found by simulation when

- adopting a standard model for paired comparisons (i.e., matches),
- assuming that the abilities do not change over the tournament.

**Model:** Bradley-Terry model for winning/losing in a paired comparison of team $i$ and team $j$.

$$\text{Pr}(i \text{ beats } j) = \pi_{i,j} = \frac{\text{ability}_i}{\text{ability}_i + \text{ability}_j}.$$
Abilities and tournament simulations

“Reverse” simulation:

- If the team-specific ability $a_i$ were known, pairwise probabilities $\pi_{i,j}$ could be computed.
- Given $\pi_{i,j}$ the whole tournament can be simulated (assuming abilities do not change and ignoring possible draws during the group stage).
- Using “many” simulations (here: 1,000,000) of the tournament, the empirical relative frequencies $\tilde{p}_i$ of each team $i$ winning the tournament can be determined.
- Choose ability $a_i$ for $i = 1, \ldots, 32$ such that the simulated winning probabilities $\tilde{p}_i$ approximately match the consensus winning probabilities $\hat{p}_i$.
- Found by simple iterative local search starting from log-odds.
Abilities and paired comparisons

The image displays a matrix comparing the abilities of various teams. Each cell represents a paired comparison, with darker shades indicating higher abilities. The matrix includes teams such as BRA, GER, ESP, FRA, ARG, BEL, ENG, POR, URU, CRO, COL, COL, RUS, POL, DEN, MEX, SUI, SWE, EGY, SRB, SEN, PER, NGA, ISL, JPN, AUS, MAR, CRC, KOR, IRN, TUN, KSA, and PAN.
Tournament simulations: Survival curves

Group A
- URU
- RUS
- EGY
- KSA

Group B
- ESP
- POR
- MAR
- IRN

Probability (%)
- Round of 16
- Quarter
- Semi
- Final
- Winner

Graphs showing survival curves for each group.
Tournament simulations: Survival curves

Group C

- FRA
- DEN
- PER
- AUS

Group D

- ARG
- CRO
- NGA
- ISL

Probability (%)

Round of 16 | Quarter | Semi | Final | Winner
---|---|---|---|---
0 | 20 | 40 | 60 | 80 | 100

24/36
Tournament simulations: Survival curves

Group E

Group F

![Graph showing survival curves for Group E and Group F]
Tournament simulations: Survival curves

Group G

Probability (%)

Round
of 16 Quarter Semi Final Winner

BEL ENG TUN PAN

Group H

Probability (%)

Round
of 16 Quarter Semi Final Winner

COL POL SEN JPN
Outcome verification
Outcome verification

**Illustration:** Check results for UEFA Euro 2016.

**Question:** Was the bookmaker consensus model any good?
- Ex post the low predicted winning probability for Portugal (4.1%) seems wrong.
- However, they profited from Spain’s and England’s poor performances in the last group stage games.
- And they only won 1 out of 7 games in normal time.
- Even in the final Gignac might as well have scored a goal instead of hitting the post in minute 92... 

**Problems:**
- Just a single observation of the tournament and at most one observation of each paired comparison.
- Hard to distinguish between an unlikely outcome and systematic errors in the predicted (prob)abilities.
Outcome verification

Possible approaches:

- Compare forecasts with the observed tournament ranking (1 POR, 2 FRA, 3.5 WAL, 3.5 GER, ...).
- Benchmark against Elo and FIFA ratings.
- Note that the Elo rating also implies ability scores based on which pairwise probabilities and “forward” simulation of tournament can be computed:

  \[ \text{ability}_{Elo,i} = 10^{Elo_i/400}. \]

- Check whether pairwise probabilities roughly match empirical proportions from clusters of matches.
Outcome verification: Ranking

Spearman rank correlation of observed tournament ranking with bookmaker consensus model (BCM) as well as FIFA and Elo ranking:

- BCM (Probabilities) 0.523
- BCM (Abilities) 0.436
- Elo (Probabilities) 0.344
- Elo 0.339
- FIFA 0.310
Outcome verification: BCM pairwise prob.

Winning probability of stronger team (in %)

<table>
<thead>
<tr>
<th>Interval</th>
<th>Lose</th>
<th>Draw</th>
<th>Win</th>
</tr>
</thead>
<tbody>
<tr>
<td>[50,60]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(60,70)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(70,85]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Winning probability of stronger team (in %)
Outcome verification: Elo pairwise prob.

Winning probability of stronger team (in %)

<table>
<thead>
<tr>
<th></th>
<th>[50,65]</th>
<th>(65,80]</th>
<th>(80,95]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lose</td>
<td>0.0</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>Draw</td>
<td>0.6</td>
<td>0.8</td>
<td>1.0</td>
</tr>
<tr>
<td>Win</td>
<td>0.0</td>
<td>0.2</td>
<td>0.4</td>
</tr>
</tbody>
</table>

32/36
Outcome verification: BCM abilities
Outcome verification: Elo abilities

![Bar chart showing relative ability (Elo) for various countries, with France (FRA) and Germany (GER) having the highest median Elo values.](chart.png)
Discussion

Summary:
- Expert judgments of bookmakers are a useful information source for probabilistic forecasts of sports tournaments.
- Winning probabilities are obtained by adjustment for overround and averaging on log-odds scale.
- Competitor abilities can be inferred by post-processing based on pairwise-comparison model with “reverse” tournament simulations.
- Approach outperformed Elo and FIFA ratings for the last UEFA Euros and correctly predicted the final 2008 and winner 2012.

Limitations:
- Matches are only assessed in terms of winning/losing, i.e., no goals, draws, or even more details.
- Inherent chance is substantial and hard to verify.

