

PREDICTING THE FOOTBALL WORLD CUP IN BRAZIL 2014

Magne Aldrin¹²

¹ Norwegian Computing Center, Oslo, Norway

² University of Oslo, Norway

The FIFA World Cup 2014 will be arranged in Brazil from 12. June to 13. July. During the tournament the Norwegian Computing Center (NR) will consider the chances of every team participating in the championship, based on a probability model. All calculations will be updated daily and published on <http://vm.nr.no/en>.

Model The prediction system is based on a Poisson regression model. Consider a game between team A and team B. In our model, the number of goals scored by country A are Poisson distributed with parameter $\lambda(A, B)$, given by

$$\lambda(A, B) = \gamma \cdot s_A / s_B,$$

where γ is the expected number of goals scored by one team in a match between two equally good teams, s_A expresses the goodness of team A and s_B expresses the goodness of team B.

Furthermore, the number of goals scored by team B are Poisson distributed with parameter $\lambda(B, A)$, independent of the number of goals scored by team A.

Estimation Before the start of the tournament, the model parameters are estimated from an artificial data set based on evaluations from several Norwegian football experts. The experts have guessed the results of several hundred hypothetical games between the 32 participating teams.

When the tournament starts, the real games are taken into account as well. The information value of the hypothetical games (the expert guesses) are weighted versus the real games.

The parameters are estimated by maximising a modified Poisson likelihood. The difference from an ordinary Poisson likelihood is that it is robustified by downweighting large victories and by adding a penalty term that shrinks the individual strength parameters towards a common mean.

Simulation The winner probabilities are found by simulating all remaining matches numerous times, or in other words, we “play” the remaining part of the tournament many times. From these simulations, we calculate all the teams’ chances to win the championship, to win their group, to reach the Second Round, etc.

Keywords: Football prediction, Poisson regression, Simulation.