2019 Continental Football Championship Forecasts via Nested Poisson Regression

Lorenz A. Gilch

Department of Informatics and Mathematics, University of Passau
Lorenz.Gilch@Uni-Passau.de

UNIVERSITÄT PASSAU
Fakultät für Informatik und Mathematik

Technical Report, Number MIP-1901
Department of Informatics and Mathematics
University of Passau, Germany
June 2019

2019 CONTINENTAL FOOTBALL CHAMPIONSHIP FORECASTS VIA NESTED POISSON REGRESSION

LORENZ A. GILCH

Abstract. This article is devoted to the forecast of the 2019 continental football championships in Africa and in the Americas, namely the Africa Cup of Nations, the Gold Cup and the Copa America. We present the simulation results for the three tournaments, where these simulations are based on a Poisson regression model that includes the Elo points of the participating teams as covariates and incorporates differences of team-specific skills. The proposed model allows predictions in terms of probabilities in order to quantify the chances for each team to reach a certain stage of the tournament. Monte Carlo simulations are used to estimate the outcome of each single match of the tournaments and hence to simulate the whole tournament itself. The model is fitted on all football games on neutral ground of the participating teams since 2010.

1. Introduction

1.1. **Problem formulation.** Football is a typical low-scoring game and games are frequently decided through single events in the game. While several factors like extraordinary individual performances, individual errors, injuries, refereeing errors or just lucky coincidences are hard to forecast, each team has its strengths and weaknesses (e.g., defense and attack) and most of the results reflect the qualities of the teams. We follow this idea in order to derive probabilities for the exact result of a single match between two teams, which involves the following three ingredients for both teams:

- Elo ranking
- attack strength
- defense strength

Since the complexity of the tournaments, with billions of different outcomes, making it very difficult to obtain accurate guesses of the probabilities of certain events, we do *not* aim on forecasting the exact outcome of tournaments, but we want to make the discrepancy between the participating teams *quantifiable* and to measure the chances of each team.

In this report we are interested in the continental championships of Africa (Africa Cup of Nations), of South America (Copa America) and North/Central America including the Caribbean (Gold Cup). Since the strengths of many of the participating teams are rather unknown it is still more unclear to estimate the strengths of the participating teams or

Date: June 14, 2019.
Key words and phrases. Africa Cup of Nations; Gold Cup; Copa America; football; Poisson regression; Elo.

even to determine the divergence of the teams' strengths, since many teams (and their players) are not so well-known as the teams from Europe or South America. This article summarises the results for the Africa Cup from (Gilch, 2019), but extends the results to the other tournaments.

Another note worth to mention is the fact that the groups are already drawn and so the tournament structure for each team (in particular, the way to the final) is set. Hence, it is the idea to measure whether a team has a rather simple or hard way to the final. Therefore once again, the aim is to quantify the difficulty for each team to proceed to the different stages of the tournament.

Since this is a technical report with the aim to present simulation results, we omit a detailed description of the state of the art and refer to (Gilch and Müller, 2018) for a discussion of related research articles and a comparison to related models and covariates under consideration. The results from the Africa Cup of Nations 2019 were presented in a detailed way in (Gilch, 2019)

Finally, let me say some words on the data available for feeding our regression model. In this article the Elo ranking (`http://en.wikipedia.org/wiki/World_Football_Elo_Ratings`) is preferably considered instead of the FIFA ranking (which is a simplified Elo ranking since July 2018), since the calculation of the FIFA ranking changed over time and the Elo ranking is more widely used in football forecast models. See also (Gásques and Royuela, 2016) for a discussion on this topic and a justification of the Elo ranking. The obtained results show that, despite the simplicity of the model, the model under consideration shows a good fit, the obtained forecasts are conclusive and give *quantitative insights* in each team's chances. In particular, we quantify the chances of each team to proceed to a specific phase of the tournaments, which allows also to compare the challenge for each team to proceed to the final.

1.2. **Questions under consideration.** The simulation in this article works as follows: each single match is modeled as G_A:G_B, where G_A (resp. G_B) is the number of goals scored by team A (resp. by team B). So much the worse not only a single match is forecasted but the course of the whole tournament. Even the most probable tournament outcome has a probability, very close to zero to be actually realized. Hence, deviations of the true tournament outcome from the model's most probable one are not only possible, but most likely. However, simulations of the tournament yield estimates of the probabilities for each team to reach different stages of the tournament and allow to make the different team's chances *quantifiable*. In particular, we are interested to give quantitative insights into the following questions:

(1) How are the probabilities that a team will win its group or will be eliminated in the group stage?
(2) Which team has the best chances to become new African champion, Gold Cup winner or Copa America winner?
(3) How big is the probability that a team will survive the preliminary round?

As we will see, the model under consideration in this article favors Senegal to win the Africa Cup of Nations 2019, Brazil to become new Copa America champion and Mexico to win the Gold Cup.

2. The model

2.1. **Involved data.** The model used in this article was proposed in (Gilch and Müller, 2018) (together with several similar bi-variate Poisson models) as *Nested Poisson Regression* and is based on the World Football Elo ratings of the teams. It is based on the Elo rating system, see (Elo, 1978), but includes modifications to take various football-specific variables (like home advantage, goal difference, etc.) into account. The Elo ranking is published by the website `eloratings.net`, from where also all historic match data was retrieved.

First, we present the formula for the Elo ratings, which uses the typical form as described in `http://en.wikipedia.org/wiki/World_Football_Elo_Ratings`: let $\mathrm{Elo}_{\mathrm{before}}$ the Elo points of a team before a match; then the Elo points $\mathrm{Elo}_{\mathrm{after}}$ after the match against an opponent with Elo points $\mathrm{Elo}_{\mathrm{Opp}}$ is calculated via the following formula:

$$\mathrm{Elo}_{\mathrm{after}} = \mathrm{Elo}_{\mathrm{before}} + K \cdot G \cdot (W - W_e),$$

where

- K is a weight index regarding the tournament of the match (World Cup matches have wight 60, while continental tournaments have weight 50)
- G is a number from the index of goal differences calculated as follows:

$$G = \begin{cases} 1, & \text{if the match is a draw or won by one goal} \\ \frac{3}{2}, & \text{if the match is won by two goals} \\ \frac{11+N}{8}, & \text{where } N \text{ is the goal difference otherwise} \end{cases}$$

- W is the result of the match: 1 for a win, 0.5 for a draw, and 0 for a defeat.
- W_e is the expected outcome of the match calculated as follows:

$$W_e = \frac{1}{10^{-\frac{D}{400}} + 1},$$

  where $D = \mathrm{Elo}_{\mathrm{before}} - \mathrm{Elo}_{\mathrm{Opp}}$ is the difference of the Elo points of both teams.

The Elo ratings as they were on 12 April 2019 for the top 5 participating nations in the Africa Cup of Nations (in this rating) are as follows:

Senegal	Nigeria	Morocco	Tunisia	Ghana
1764	1717	1706	1642	1634

The Elo ratings as they were on 18 May 2019 for the top 5 participating Gold Cup nations i are as follows:

Mexico	United States	Costa Rica	Honduras	Panama
1816	1766	1705	1596	1571

The Elo ratings as they were on 14 May 2019 for the top 5 participating Copa America nations are as follows:

Brazil	Colombia	Uruguay	Argentina	Chile
2132	1956	1933	1901	1834

The forecast of the outcome of a match between teams A and B is modelled as

$$G_A \ : \ G_B,$$

where G_A (resp. G_B) is the number of goals scored by team A (resp. B). The model is based on a Poisson regression model, where we assume (G_A, G_B) to be a bivariate Poisson distributed random variable; see (Gilch and Müller, 2018, Section 8) for a discussion on other underlying distributions for G_A and G_B. The distribution of (G_A, G_B) will depend on the current Elo ranking Elo_A of team A and Elo ranking Elo_B of team B. The model is fitted using all matches of the participating teams on *neutral* playground between 1.1.2010 and 12.04.2019. Matches, where one team plays at home, have usually a drift towards the home team's chances, which we want to eliminate. In average, we have for each African team 29 matches from the past and for the top teams even more; for the Copa America teams we have in average 38 matches (for Brazil even 56 matches) and for the Gold Cup teams we have in average 34 matches. In the following subsection we explain the model for forecasting a single match, which in turn is used for simulating the whole tournaments and determining the likelihood of the success for each participant.

2.2. **Nested Poisson regression.** We now present a *dependent* Poisson regression approach which will be the base for the whole simulation and which was already used in (Gilch and Müller, 2018) and (Gilch, 2019). The number of goals G_A, G_B respectively, shall be a Poisson-distributed random variable with rate $\lambda_{A|B}$, $\lambda_{B|A}$ respectively. As we will see one of the rates (that is, the rate of the weaker team) will depend on the concrete realisation of the other random variable (that is, the simulated number of scored goals of the stronger team).

In the following we will always assume that A has *higher* Elo score than B. This assumption can be justified, since usually the better team dominates the weaker team's tactics. Moreover the number of goals the stronger team scores has an impact on the number of goals of the weaker team. For example, if team A scores 5 goals it is more likely that B scores also 1 or 2 goals, because the defense of team A lacks in concentration due to the expected victory. If the stronger team A scores only 1 goal, it is more likely that B scores no or just one goal, since team A focusses more on the defence and secures the victory.

The Poisson rates $\lambda_{A|B}$ and $\lambda_{B|A}$ are now determined as follows:

(1) In the first step we model the number of goals $\tilde{G}_A$ scored by team A only in dependence of the opponent's Elo score $\mathrm{Elo} = \mathrm{Elo}_B$. The random variable $\tilde{G}_A$ is modeled as a Poisson distribution with parameter μ_A. The parameter μ_A as a function of the Elo rating Elo_B of the opponent B is given as

$$\log \mu_A(\mathrm{Elo}_B) = \alpha_0 + \alpha_1 \cdot \mathrm{Elo}_B, \tag{2.1}$$

where α_0 and α_1 are obtained via Poisson regression.

(2) Teams of similar Elo scores may have different strengths in attack and defense. To take this effect into account we model the number of goals team B receives against a team of Elo score $\mathrm{Elo} = \mathrm{Elo}_A$ using a Poisson distribution with parameter ν_B. The parameter ν_B as a function of the Elo rating Elo_B is given as

$$\log \nu_B(\mathrm{Elo}_B) = \beta_0 + \beta_1 \cdot \mathrm{Elo}_B, \tag{2.2}$$

where the parameters β_0 and β_1 are obtained via Poisson regression.

(3) Team A shall in average score $\mu_A\big(\mathrm{Elo}_B\big)$ goals against team B, but team B shall have $\nu_B\big(\mathrm{Elo}_A\big)$ goals against. As these two values rarely coincides we model the numbers of goals G_A as a Poisson distribution with parameter

$$\lambda_{A|B} = \frac{\mu_A\big(\mathrm{Elo}_B\big) + \nu_B\big(\mathrm{Elo}_A\big)}{2}.$$

(4) The number of goals G_B scored by B is assumed to depend on the Elo score $E_A = \mathrm{Elo}_A$ and additionally on the outcome of G_A. More precisely, G_B is modeled as a Poisson distribution with parameter $\lambda_B(E_A, G_A)$ satisfying

$$\log \lambda_B(E_A, G_A) = \gamma_0 + \gamma_1 \cdot E_A + \gamma_2 \cdot G_A. \tag{2.3}$$

The parameters $\gamma_0, \gamma_1, \gamma_2$ are obtained by Poisson regression. Hence,

$$\lambda_{B|A} = \lambda_B(E_A, G_A).$$

(5) The result of the match A versus B is simulated by realizing G_A first and then realizing G_B in dependence of the realization of G_A.

For a better understanding, we give an example and consider the match Senegal vs. Ivory Coast: Senegal has 1764 Elo points while Ivory Coast has 1612 points. Against a team of Elo score 1612 Senegal is assumed to score in average

$$\mu_{\text{Senegal}}(1612) = \exp(2.73 - 0.00145 \cdot 1612) = 1.48$$

goals, while Ivory Coast receives against a team of Elo score 1764 in average

$$\nu_{\text{Ivory Coast}}(1764) = \exp(-4.0158 + 0.00243 \cdot 1764) = 1.31$$

goals. Hence, the number of goals, which Senegal will score against Ivory Coast, will be modelled as a Poisson distributed random variable with rate

$$\lambda_{\text{Senegal}|\text{Ivory Coast}} = \frac{1.48 + 1.31}{2} = 1.395.$$

The average number of goals, which Ivory Coast scores against a team of Elo score 1764 provided that G_A goals against are received, is modelled by a Poisson random variable with rate

$$\lambda_{\text{Ivory Coast}|\text{Senegal}} = \exp(1.431 - 0.000728 \cdot 1764 + 0.137 \cdot G_A);$$

e.g., if $G_A = 1$ then $\lambda_{\text{Ivory Coast}|\text{Senegal}} = 1.33$.

As a final remark, let me mention that the presented dependent approach may also be justified through the definition of conditional probabilities:

$$\mathbb{P}[G_A = i, G_B = j] = \mathbb{P}[G_A = i] \cdot \mathbb{P}[G_B = j \mid G_A = i] \quad \forall i, j \in \mathbb{N}_0.$$

For comparision of this model in contrast to similar Poisson models, we refer once again to (Gilch and Müller, 2018). All calculations were performed with R (version 3.3.1). In the following subsections we present some regression plots and will test the goodness of fit.

2.3. **Regression plots.** As an example of interest, we sketch in Figure 1 the result of the regression in (2.1) for the number of goals scored by Costa Rica. The dots show the observed data (i.e, number of scored goals on the y-axis in dependence of the opponent's strength on the x-axis) and the line is the estimated mean μ_A depending on the opponent's Elo strength; the shaded grey area represents standard errors.

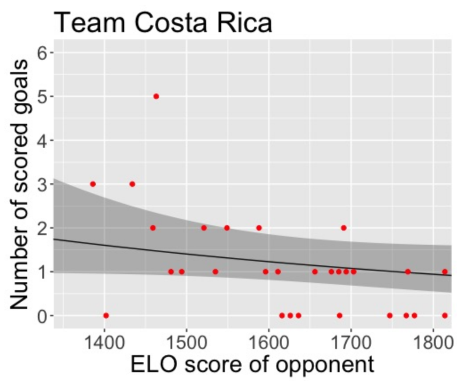

Figure 1. Plots for the number of goals scored by Costa Rica in regression (2.1).

Analogously, Figure 2 sketches the regression in (2.2) for the (unconditioned) number of goals against of Chile in dependence of the opponent's Elo ranking. The dots show the observed data (i.e., the number of goals against in the matches from the past) and the line is the estimated mean ν_B for the number of goals against. The most important countries

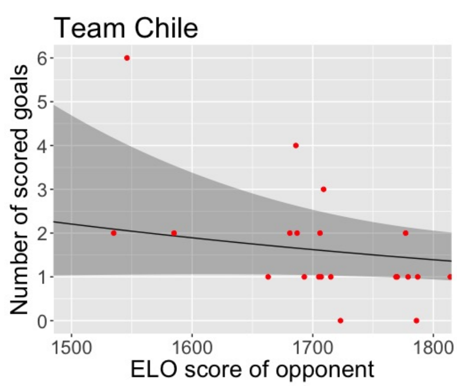

Figure 2. Plots for the number of goals scored by Chile in regression (2.2).

produce similar pictures, while the underdogs tend to fit less reasonable. Since the weakest teams won't have a big impact on the outcome of the tournaments, we still can accept this fitting.

2.4. **Goodness of fit tests.** We check goodness of fit of the Poisson regressions in (2.1) and (2.2) for all participating teams. For each team $\mathbf{T}$ we calculate the following χ^2-statistic from the list of matches from the past:

$$\chi_{\mathbf{T}} = \sum_{i=1}^{n_{\mathbf{T}}} \frac{(x_i - \hat{\mu}_i)^2}{\hat{\mu}_i},$$

where $n_{\mathbf{T}}$ is the number of matches of team $\mathbf{T}$, x_i is the number of scored goals of team $\mathbf{T}$ in match i and $\hat{\mu}_i$ is the estimated Poisson regression mean in dependence of the opponent's historical Elo points.

2.4.1. *Gold Cup.* We observe that all teams have a very good fit, except Trinidad & Tobago and Curaçao with a p-value of 0.01 and < 0.01. In average, we have a p-value of 0.397. In Table 1 the p-values for some of the top teams are given.

Team	Mexico	United States	Costa Rica	Honduras	Panama
p-value	0.43	0.67	0.31	0.15	0.59

TABLE 1. Goodness of fit test for the Poisson regression in (2.1) for some of the top teams.

We calculate a χ^2-statistic for each team which measures the goodness of fit for the regression in (2.2) which models the number of goals against. Here, we get an average p-value of 0.55; see Table 2. The rather less good fit of Mexico is not important because Mexico is the best-ranked team, so it is more important that the teams following Mexico in the ranking have a good fit!

Team	Mexico	United States	Costa Rica	Honduras	Panama
p-value	0.07	0.97	0.87	0.74	0.93

TABLE 2. Goodness of fit test for the Poisson regression in (2.2) for some of the top teams.

Finally, we test the goodness of fit for the regression in (2.3) which models the number of goals against of the weaker team in dependence of the number of goals which are scored by the stronger team; see Table 3. As a conclusion, the p-values suggest reasonable fits for the top teams.

Team	United States	Costa Rica	Trinidad & Tobago
p-value	0.63	0.15	0.44

TABLE 3. Goodness of fit test for the Poisson regression in (2.3) for some of the teams.

2.4.2. *Copa America.* In the case of the Copa America participating teams we also observe that all teams have a very good fit, except Chile with a p-value of 0.01. In average, we have a p-value of 0.46. In Table 4 the p-values for some of the top teams are given.

Team	Brazil	Colombia	Uruguay	Argentina
p-value	0.78	0.76	0.81	0.14

Table 4. Goodness of fit test for the Poisson regression in (2.1) for some of the top teams.

We calculate a χ^2-statistic for each team which measures the goodness of fit for the regression in (2.2) which models the number of goals against. Here, we get an average p-value of 0.38; see Table 5. Argentina has a non-satisfying fit which may cause due to contradictory results in the past few years.

Team	Brazil	Colombia	Uruguay	Argentina
p-value	0.51	0.54	0.69	0.01

Table 5. Goodness of fit test for the Poisson regression in (2.2) for some of the top teams.

Finally, we test the goodness of fit for the regression in (2.3) which models the number of goals against of the weaker team in dependence of the number of goals which are scored by the stronger team; see Table 6. As a conclusion, the p-values suggest reasonable fits for the top teams.

Team	Colombia	Ecuador	Qatar	Venezuela
p-value	0.32	0.29	0.15	0.33

Table 6. Goodness of fit test for the Poisson regression in (2.3) for some of the top teams.

2.4.3. *Africa Cup of Nations.* For the Africa Cup of Nations, we get similar results; see (Gilch, 2019), where also a detailed analysis of the null and residual deviances for each team for the regressions in (2.1), (2.2) and (2.3) is presented. Analogous deviance calculations with similar results are obtained for the Gold Cup and Copa America.

3. Gold Cup 2019

Finally, we come to the simulation of the Gold Cup 2019, which allows us to answer the questions formulated in Section 1.2. We simulate each single match of the Gold Cup 2019 according to the model presented in Section 2, which in turn allows us to simulate the whole Gold Cup tournament. After each simulated match we update the Elo ranking according to the simulation results. This honours teams, which are in a good shape during a tournament and perform maybe better than expected. Overall, we perform 100.000 simulations of the whole tournament, where we reset the Elo ranking at the beginning of each single tournament simulation.

3.1. **Group Forecast.** In the following tables 7-10 we present the probabilities obtained from our simulation, where we give the probabilities of winning the group, becoming runner-up, or to be eliminated in the group stage. In Group D, the toughest group of all, a head-to-head fight between Panama and Trinidad & Tobago is expected for the second place.

Team	1st	2nd	Preliminary Round
Mexico	85.20	12.40	2.40
Canada	11.00	52.40	36.70
Martinique	3.40	27.70	68.90
Cuba	0.40	7.50	92.10

Table 7. Probabilities for Group A

Team	1st	2nd	Preliminary Round
Costa Rica	66.80	25.80	7.40
Haiti	26.60	52.50	20.90
Nicaragua	6.60	21.70	71.70
Bermuda	0.00	0.00	100.00

Table 8. Probabilities for Group B

Team	1st	2nd	Preliminary Round
Honduras	35.70	30.80	33.60
Jamaica	35.30	30.80	33.80
El Salvador	26.50	30.90	42.60
Curacao	2.50	7.50	90.10

Table 9. Probabilities for Group C

Team	1st	2nd	Preliminary Round
United States	64.10	25.00	10.80
Panama	24.40	40.50	35.20
Trinidad & Tobago	11.50	33.90	54.60
Guyana	0.00	0.50	99.40

Table 10. Probabilities for Group D

3.2. **Playoff Round Forecast.** Our simulations yield the following probabilities for each team to win the tournament or to reach certain stages of the tournament. The result is presented in Table 11. The regression model favors Mexico, followed by the United States and Costa Rica. The remaining teams have only chances as underdogs to win the Gold Cup.

Team	Gold Cup	Final	Semifinal	Quarterfinal
Mexico	39.90	57.00	76.90	97.70
United States	26.20	51.80	70.20	89.10
Costa Rica	14.70	27.30	57.40	92.50
Panama	6.40	17.80	40.50	64.90
Jamaica	2.80	9.40	24.60	66.10
Haiti	2.40	7.40	26.30	79.10
Honduras	2.30	8.90	24.20	66.50
Canada	1.90	5.80	24.80	63.40
El Salvador	1.50	6.20	18.70	57.50
Trinidad & Tobago	1.30	5.90	20.70	45.40
Martinique	0.30	1.30	8.50	31.00
Nicaragua	0.20	1.00	5.10	28.40
Cuba	0.00	0.00	0.90	7.90
Bermuda	0.00	0.00	0.00	0.00
Curacao	0.00	0.10	1.10	9.90
Guyana	0.00	0.00	0.00	0.50

Table 11. Gold Cup 2019 simulation results for the teams' probabilities to proceed to a certain stage

The results are visualised in the heat map in Figure 3. The country's red tone colour represents the probability of winning the Africa Cup of Nations: the darker the colour tone, the higher the probability to win the Africa Cup of Nations.

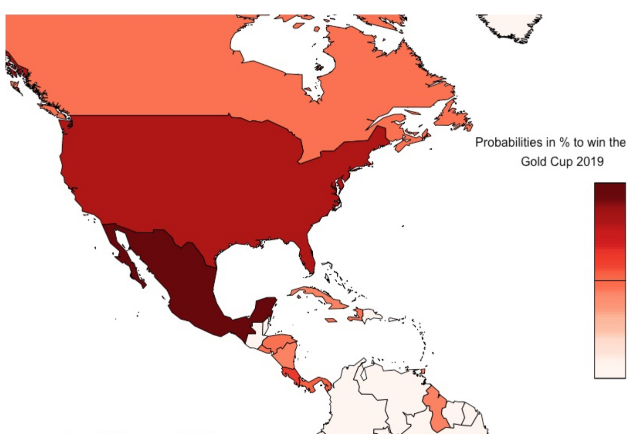

Figure 3. Heat map for the Gold Cup.

4. Copa America 2019

4.1. **Group Forecast.** In the following tables 12-14 we present the probabilities obtained from our simulation, where we give the probabilities of winning the group, becoming runner-up, qualifying for quarterfinals as a third-ranked team or to be eliminated in the group stage. As we can see, almost all teams have a good chance to qualify for the quarterfinals.

Team	1st	2nd	Qualified as Third	Preliminary Round
Brazil	55.90	29.40	12.80	1.90
Bolivia	0.30	3.30	5.80	90.70
Venezuela	18.00	30.20	26.30	25.50
Peru	25.80	37.20	22.10	14.90

Table 12. Probabilities for Group A

Team	1st	2nd	Qualified as Third	Preliminary Round
Argentina	43.40	29.70	13.80	13.00
Colombia	33.40	31.60	16.00	18.90
Paraguay	5.10	13.50	15.60	65.80
Qatar	18.00	25.20	19.10	37.60

Table 13. Probabilities for Group B

Team	1st	2nd	Qualified as Third	Preliminary Round
Uruguay	41.30	27.70	13.80	17.20
Ecuador	12.90	21.90	20.30	44.90
Japan	17.60	23.20	18.10	41.10
Chile	28.20	27.20	16.30	28.40

Table 14. Probabilities for Group C

4.2. **Playoff Round Forecast.** Our simulations yield the following probabilities for each team to win the tournament or to reach certain stages of the tournament. The result is presented in Table 15. The regression model clearly favors Brazil, while Argentina and Colombia seem to have only minor chances to win the Copa America.

Team	Copa Winner	Final	Semifinal	Quarterfinal
Brazil	33.00	48.30	73.10	98.00
Argentina	14.10	28.30	51.50	86.80
Colombia	12.40	24.20	45.20	80.90
Peru	10.40	21.00	45.60	85.00
Uruguay	8.20	21.30	41.00	82.80
Chile	8.10	17.70	35.10	71.60
Venezuela	5.30	12.80	32.40	74.50
Qatar	4.50	11.30	27.60	62.40
Japan	2.30	7.70	21.30	58.80
Ecuador	1.20	5.20	16.90	55.20
Paraguay	0.50	2.00	9.10	34.20
Bolivia	0.00	0.10	1.00	9.40

TABLE 15. Copa America 2019 simulation results for the teams' probabilities to proceed to a certain stage

The results are visualised in the heat map in Figure 4.

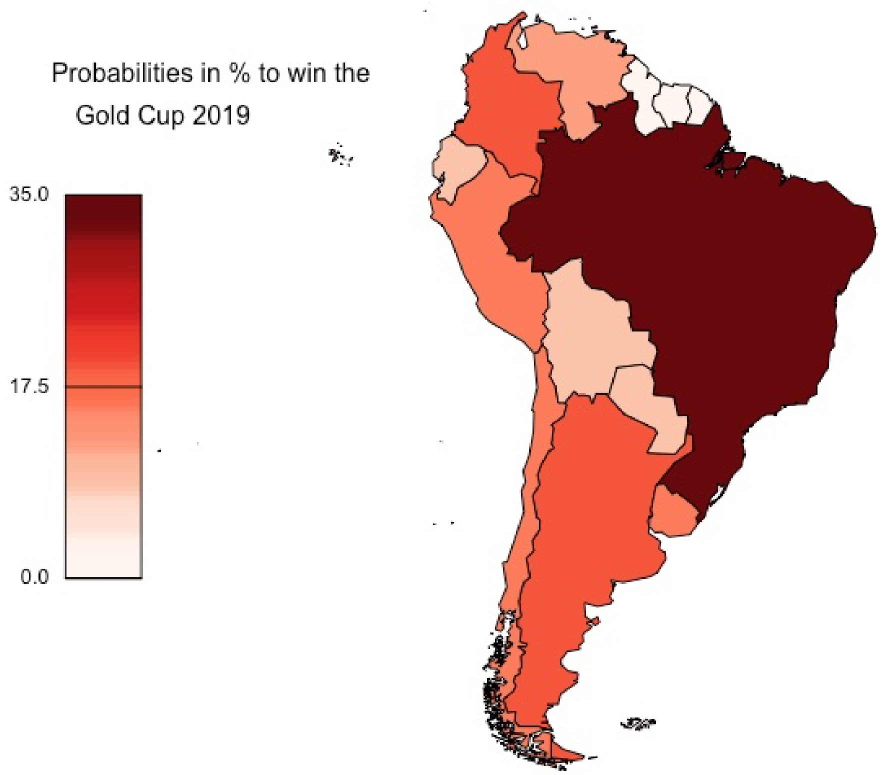

FIGURE 4. Heat map for the Copa America: the darker the red tone, the higher the probability to win the Copa America.

5. Africa Cup of Nations 2019 Simulations

A detailed analysis and presentation of the simulation results for the Africa Cup of Nations 2019 is presented in (Gilch, 2019). Thus, we just present here the forecast for the playoff round in Table 16. The regression model favors Senegal, followed by Nigeria, Ivory Coast and Egypt, to become new football champion of Africa.

Team	Champion	Final	Semifinal	Quarterfinal	Last16
Senegal	15.40	25.20	41.20	67.70	92.90
Nigeria	12.10	22.70	37.30	59.90	91.60
Ivory Coast	10.20	17.70	31.10	51.90	79.10
Egypt	10.10	19.20	34.60	56.60	90.60
Ghana	8.60	17.00	30.50	57.20	95.40
South Africa	8.40	15.50	28.50	48.80	76.50
Morocco	8.30	15.30	28.20	48.20	73.90
Tunisia	5.80	11.90	23.20	45.50	91.70
Algeria	5.10	10.30	21.40	43.30	77.80
Guinea	3.40	8.10	17.90	37.60	74.60
Cameroon	3.00	9.00	22.30	50.70	93.30
DR Congo	3.00	7.70	19.00	40.00	79.10
Mali	1.60	5.00	13.20	32.70	88.50
Madagascar	1.60	4.10	10.50	25.40	62.40
Kenya	1.10	3.10	9.10	23.90	58.40
Angola	1.00	2.80	8.00	22.10	66.10
Zimbabwe	0.40	1.80	7.40	22.80	59.50
Namibia	0.30	1.20	4.20	13.20	39.10
Uganda	0.10	0.50	2.60	10.30	34.90
Tanzania	0.10	0.50	2.60	10.10	36.90
Mauritania	0.10	0.40	1.50	5.90	24.40
Benin	0.10	0.60	3.40	15.10	58.00
Burundi	0.00	0.20	1.60	7.90	37.00
Guinea-Bissau	0.00	0.00	0.30	2.60	17.60

Table 16. Africa Cup of Nations 2019 simulation results for the teams' probabilities to proceed to a certain stage

The results can be visualised in a heat map in Figure 5.

Let me also mention that (Gilch, 2019) addresses also the question whether the current tournament structure, which allows third-placed teams in the preliminary round still to qualify for the round of 16, is reasonable or not. In particular, it is the question whether this structure is good or bad for the top teams and to quantify this factor. It turns out that a definitive elimination of third-ranked teams would increase the chances of top teams only slightly.

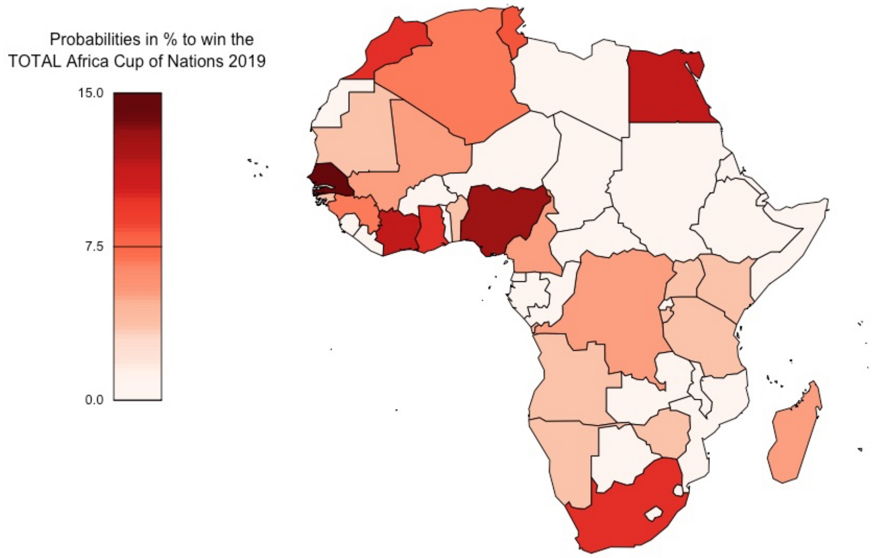

FIGURE 5. Heat map for the Africa Cup of Nations the darker the red tone, the higher the probability to win the Africa Cup of Nations.

6. Final remarks

For further discussion on adaptions and different models, we refer once again to the discussion section in (Gilch and Müller, 2018) and (Gilch, 2019). In particular, it turned out that more general models like Generalised Poisson models or Negative Binomial distribution did *not* lead to a remarkable better fit.

References

Elo, A. E. (1978). *The rating of chessplayers, past and present.* Arco Pub., New York.

Gásques, R. and Royuela, V. (2016). The determinants of international football success: A panel data analysis of the elo rating*. *Social Science Quarterly*, 97(2):125–141.

Gilch, L. A. (2019). Prediction Model for the Africa Cup of Nations 2019 via Nested Poisson Regression. *African Journal of Applied Statistics*, 6(1):599–615.

Gilch, L. A. and Müller, S. (2018). On Elo based prediction models for the FIFA Worldcup 2018. *Technical Report, Number MIP-1801, Department of Informatics and Mathematics, University of Passau, Germany.*

Lorenz A. Gilch: Universität Passau, Innstrasse 33, 94032 Passau, Germany

E-mail address: `Lorenz.Gilch@uni-passau.de`

URL: `http://www.math.tugraz.at/∼gilch/`