

Eine systematische Analyse von Fußball-Vorhersagemodellen und Tipprunden

Malte Braband

Technische Universität Braunschweig, Deutschland
m.braband@tu-braunschweig.de

Zusammenfassung. In dieser Arbeit wird die Frage gestellt (und beantwortet), wie man systematisch in Fußball-Tipprunden gut abschneiden kann. Zur wissenschaftlichen Auswertung wurden zunächst Beziehungen zwischen verschiedenen Prognosemodellen ermittelt und alle Modelle in Beziehung zum Poisson-Modell gebracht. Teilweise wurden dabei neue Ergebnisse erzielt. Außerdem wurde eine Methodik entwickelt, mit der der Verlauf der WM berechnet werden kann, anstatt wie bisher in der Literatur zu simulieren. In der Auswertung haben die Systemtipps im Mittel deutlich besser abgeschnitten als die menschlichen Tipper, allerdings kann auch der beste Systemtipp realistischere Ergebnisse eines Tippwettbewerbs mit sehr vielen Tippern nicht gewinnen. Unter den Systemtipps gab es quantitativ und qualitativ relativ wenige Unterschiede, von zufälligen Schwankungen abgesehen. Verlässt man sich statt auf das eigene Expertenwissen also auf einen Systemtipp, dann wird man in den Tipprunden mit hoher Wahrscheinlichkeit zu den besten 20% der Teilnehmer gehören, aber nur selten gewinnen.

Abstract. This paper analyses the question how to systematically reach the top flight of soccer prediction leagues. In a first step several forecast models are compared and it is shown how most models can be related to the Poisson model. Some of the relations are new. Additionally a method has been developed which allows to numerically evaluate the outcome probabilities of soccer championships instead of simulation. The main practical result for the example of the 2014 soccer World Championship was that the forecast models were significantly better than the human participants of a large public prediction league. However the results between the forecast models were small, both qualitatively and quantitatively. But it is quite unlikely that a large prediction league will be won by a forecast model although the forecast models almost all belonged to the top flight of the prediction league.

Keywords. Soccer World Championship, prediction league, Poisson distribution, Skellam distribution, forecast models.

1 Einführung

Die Cäcilienhochschule Wilhelmshaven führt seit 2006 zu jeder Welt- und Europameisterschaft einen Tippwettbewerb durch, an dem sich zur WM 2014 über 2400 Teilnehmer

beteiligt haben [1]. In einem Jugend-Forscht-Projekt [2] wurde wissenschaftlich bewertet, wer die besten WM-Tipps abgab - und natürlich warum:

- Der (selbsternannte) „Fussballexperte“ mit seinem Fachwissen oder
- eine Tippgemeinschaft, die die „Weisheit der Vielen“ ausnutzt oder
- ein Systemtipp, der die Ergebnisse z.B. aus Ranglisten berechnet (Systemtipp) oder
- das Wettbüro mit seinen Quoten.

Damit die Systemtipps genau wie die menschlichen Teilnehmer am Tippwettbewerb teilnehmen konnten, galt es, aus den Modell-Ansätzen konkrete Tipps abzuleiten. Dabei stellte sich heraus, dass einige Zusammenhänge nicht in der Fachliteratur beschrieben waren. Daher habe ich als meinen Beitrag zur Jugend-Forscht-Arbeit teilweise neue mathematische Beziehungen zwischen den verschiedenen Prognosemodellen hergeleitet. Diese Arbeit beschreibt daher nur meine theoretischen Beiträge, für die kompletten Ergebnisse verweise ich auf [2].

2 Grundlagen

Der einfachste Ansatz zur Prognose von Fußballergebnissen wird im Buch von Tolan [3] beschrieben und beruht im Wesentlichen auf der Annahme, dass die Tore jeder Mannschaft in einem Spiel in guter Näherung Poisson-verteilt sind und zwar unabhängig von der gegnerischen Mannschaft. Der Erwartungswert der Poisson-Verteilung entspricht dabei der Anzahl von Toren pro Spiel für eine Mannschaft. Damit ist auch die Tor-Summe in einem Spiel wieder Poisson-verteilt, aber die Tor-differenz ist nicht wieder Poisson-verteilt, sondern es ergibt sich die sogenannte Skellam-Verteilung [4]. D.h. das Fußballspiel kann man statistisch als ein Würfelspiel mit einem Poisson-Würfel interpretieren. Damit lässt sich auch die große Spannung bei Fussballspielen bzw. die Unsicherheit über den Ausgang gut durch die hohe Streuung beider Verteilungen erklären.

2.1 Der Ansatz über Leistungsstärken nach Heuer

Tolan berücksichtigt allerdings nicht den Einfluss des Gegners. Heuer [5] wählt alternativ einen Ansatz mittels Tordifferenzen zur Prognose des Spiels von A gegen B, indem er die Schätzung der erwarteten Torzahlen auf sog. Leistungsstärken (LS) und die erzielten Torsummen (TS) der beiden Mannschaften zurückführt:

$$\lambda_A - \lambda_B = LS_A - LS_B \text{ (erwartete Tordifferenz)} \quad (1)$$

$$\lambda_A + \lambda_B = TS_A + TS_B - TS \text{ (erwartete Torsumme)} \quad (2)$$

Die Leistungsstärke ist bei Heuer definiert als die mittlere Tordifferenz beim Spiel gegen einen Gegner mittlerer Stärke auf neutralem Platz. Dieser Ansatz passt besonders gut für Meisterschaften von Vereinsmannschaften, wo quasi dieselben Mannschaften jedes Jahr gegeneinander spielen und es Statistiken von Vorsaisons gibt. Wenn man alle Parameter auf der rechten Seite schätzen kann, kann man die Gleichungen (1) und (2) zur konkreten Spielprognose auflösen und die konkreten erwarteten Torzahlen bestimmen. Zusätzlich können weitere Parameter wie der Heimvorteil, die aktuelle Fitness sowie der Marktwert der Mannschaften berücksichtigt werden [5].

2.2 Der Soccer Power Index nach Silver

Der Soccer Power Index (SPI) wurde von dem für Wahlprognosen bekannten amerikanischen Statistiker Nate Silver entwickelt [6] und bereits relativ erfolgreich für die Prognose der WM 2010 angewendet. Die genaue Herleitung ist proprietär, aber letztendlich wird für jede Mannschaft eine Offensiv- und Defensivstärke ermittelt, die der mittleren erzielten bzw. erhaltenen Torzahl gegen eine Mannschaft mittlerer Spielstärke auf neutralem Platz entspricht. Die Differenz entspricht der Leistungsstärke nach Heuer und man kann Spielprognosen nach (1) und (2) erstellen.

2.3 Besonderheiten bei Weltmeisterschaften

Weltmeisterschaften kommen im Vergleich zu Vereinsmeisterschaften selten vor und es werden auch nur relativ wenige Spiele gespielt. Da die Qualifikation über die verschiedenen Kontinentalverbände erfolgt, spielen Mannschaften unterschiedlicher Kontinentalverbände außerhalb der Weltmeisterschaft eigentlich nur Freundschaftsspiele. Dies bedeutet, dass nur wenige Statistiken vorliegen und dass diese teilweise auch noch recht alt und damit wenig aussagekräftig sind. Deswegen ist der Ansatz von Heuer nicht direkt übertragbar.

2.4 Ranglisten

In vielen Sportarten, auch für Fußball-Nationalmannschaften, werden Ranglisten geführt, um zu einer kontinuierlichen Einschätzung der Leistungsstärken zu kommen. Dabei wird angenommen, dass eine Mannschaft über einen gewissen Zeitraum eine etwa gleich große Leistungsstärke LS besitzt, die aus den Spielergebnissen geschätzt wird, und mit Hilfe derer die Mannschaften nach Leistungsstärke sortiert werden können, um z. B. Setzlisten für Turniere zu bilden. Für alle Ranglisten-Systeme sollte gelten:

1. Eine Mannschaft mit einer besseren Platzierung sollte gegen eine schlechter platzierte Mannschaft bei einer großen Anzahl von Spielen häufiger gewinnen
2. Die Gewinnwahrscheinlichkeit sollte umso größer sein, desto größer der Abstand in der Rangliste oder in der Leistungsstärke ist

Die FIFA-Rangliste [7] beruht auf einem Ansatz, bei dem für jedes Länderspiel eine Wertzahl ermittelt wird, die von den folgenden Parametern abhängt: dem Spielergebnis, der Wichtigkeit des Spiels, der Stärke des Gegners sowie dem Mittelwert der Stärken der Kontinentalverbände beider Mannschaften. Der Gewinner erhält diese Punktzahl. Die Ranglistenpunkte ergeben sich aus einem gewichteten Mittel aller Punkte aus den letzten 4 Jahren. Der Haupt-Kritikpunkt am FIFA-System besteht darin, dass sich auch die Punktzahl einer Mannschaft, die alle Spiele gewinnt, verringern kann, wenn sie z. B. als Gastgeber keine Pflichtspiele machen muss oder nur gegen schwächere Gegner spielt. Andererseits kann sich aber auch die Punktzahl ohne Spiele erhöhen, nur weil sich das Zeitfenster für die Berechnung geändert hat.

Daneben gibt es noch inoffizielle Elo-Ranglisten für Nationalmannschaften [8], die auf den im Schach bekannten System des ungarischen Physikers Elo beruhen. In diesem System bekommt das Team, das gewinnt, direkt Punkte vom unterlegenen Team. Der Haupt-Kritikpunkt ist hier, dass ein Team, das nicht spielt oder sich nur schwächere Gegner aussucht, keine Punkte verliert. Interessanterweise nutzt die FIFA dieses System für Frauen-, aber nicht für Männer-Nationalmannschaften.

Für die FIFA-Rangliste scheint es derzeit kein Verfahren zu geben, um aufgrund der Ranglisten-Position oder den -Punkten konkrete Spielprognosen abzuleiten. Die Elo-Rangliste dagegen beruht auf einem statistischen Modell [9]. Wenn r die Elo-Zahl einer Mannschaft ist, dann gilt für das Modell die Annahme, dass die wahre Spielstärke R normalverteilt mit dem Mittelwert r und einer Streuung von 200 (für jede Mannschaft gleich) ist. Spielt A gegen B, so gewinnt A, wenn deren tatsächlich realisierte Spielstärke größer ist als die von B. Die Differenz der Spielstärken ist dann wieder normalverteilt mit dem Mittelwert $\Delta r = r_A - r_B$ und der Streuung $200\sqrt{2}$. Damit kann man die Wahrscheinlichkeit berechnen, dass A gewinnt:

$$W(\Delta r) = P(\text{A gewinnt}) = P(R_A - R_B > 0) = 1 - \Phi_{\Delta r, 200\sqrt{2}}(0) \quad (3)$$

wobei Φ die Verteilungsfunktion der Normalverteilung bezeichnet.

Im Elo-Modell ist die Wahrscheinlichkeit für ein Unentschieden gleich Null, da die Normalverteilung eine stetige Verteilung ist. Da die Berechnung der Normalverteilung ohne Computer mühsam ist, hat Elo noch eine einfache Näherungsformel angegeben:

$$W(\Delta r) = \frac{1}{1 + 10^{-\Delta r/400}} \quad (4)$$

Außerdem kann der Heimvorteil berücksichtigt werden. Hier wird pauschal vorgeschlagen, die Punktzahl der Heimmannschaft um 100 zu erhöhen.

Schließlich könnte man auch noch die geschätzten Marktwerte der Mannschaften [22] als eine Art „Geld-Rangliste“ ansehen. Heuer [5] hatte hier für Bundesliga-

Mannschaften nachgewiesen, dass der statistische Einfluss des Marktwerts beträchtlich ist.

2.5 Prognose auf Grundlage von Wettquoten

Im Internet veröffentlichen die Wettanbieter die Quoten für Fussballwetten, manchmal gibt es bis zu 300 verschiedene Wetten für ein einziges Spiel. Interessant sind dabei vor allem die Quoten für eine gewisse Anzahl von Toren bzw. von Toto-Tipps. Leider ist die exakte Grundlage der Wettquoten ein Geschäftsgeheimnis der Wettanbieter. Im Prinzip sollte zwischen der Quote q für einen Tipp (Auszahlung bei Gewinn der Wette für 1 € Einsatz) und der Wahrscheinlichkeit p für diesen Tipp ein einfacher Zusammenhang bestehen, nämlich

$$p = \frac{1}{q}. \quad (5)$$

In diesem Fall würde sich eine faire Wette ergeben. Allerdings würde der Wettanbieter hier keinen Gewinn machen. Daher verringert der Wettanbieter die Quote um einen bestimmten Prozentsatz. In [10] wurde der Gewinnanteil vom Wettanbieter b_{win} mit ca. 8% abgeschätzt.

In [10] wurde außerdem anhand von Beispielen gezeigt, dass es plausibel ist, dass auch die Wettanbieter Poisson-Verteilungen zur Prognose benutzen. Wenn die Quote für die Wette „Die Mannschaft schießt null Tore“ bekannt ist, so weiß man für die Poisson-Verteilung, dass dann $p = e^{-\lambda} = 1/q$ gelten muss und der Mittelwert aus der Quote berechnet werden kann:

$$\lambda = \ln q \quad (6)$$

2.6 Verteilung des Toto-Tipps

Toto-Tipps können mit der Skellam-Verteilung berechnet werden. Die Berechnung ist sehr kompliziert, aber zum Glück hat sich jemand die Mühe gemacht, die Skellam-Verteilung für R zu programmieren [11]. Daraus ist dann relativ einfach die Wahrscheinlichkeit für den klassischen Toto-Tipp (Heimsieg, Unentschieden, Auswärtssieg) abzuleiten. Die Funktion `dske11am(0, x, y)` der Sprache R liefert auf einfache Weise die Wahrscheinlichkeit für ein Unentschieden (Tordifferenz 0) für die mittleren Torzahlen x und y . `pske11am(0, x, y)` entspricht der Verteilungsfunktion für Werte kleiner Null und damit der Wahrscheinlichkeit für einen Auswärtssieg.

Figur 1 zeigt die so berechneten Wahrscheinlichkeiten für Toto-Tipps. Dabei wurde von einer bei der WM im Mittel vorkommenden mittleren Torsumme von 2,5 Toren ausgegangen, die zwischen Heim- und Auswärtsmannschaft aufgeteilt wurde, d.h. für die Heimmannschaft läuft die mittlere Torzahl x von 0 bis 2,5 und für die Auswärtsmannschaft wird als mittlere Toranzahl $2,5 - x$ angesetzt. Die grüne Kurve in Figur 1 zeigt die Wahrscheinlichkeit für Heimsieg, die rote die für Auswärtssieg und die gelbe die für Unentschieden.

Man kann die Kurven in Figur 1 auch dazu nutzen, um aus den Wettquoten bzw. den Wahrscheinlichkeiten für den Toto-Tipp die mittleren Torzahlen der Mannschaften zurückzurechnen. Dabei muss man natürlich auch noch die Torsumme schätzen.

Um einen Zusammenhang mit dem Elo-Modell herzustellen, könnte man in (3) die Unentschieden als „halbe Siege“ mitzählen. Dann könnte man diesen Zusammenhang zwischen der Elo-Gewinnwahrscheinlichkeit und der Toto-Tipp-Wahrscheinlichkeit im Poisson-Modell bekommen

$$W(\Delta r) = P(A \text{ gewinnt}) + \frac{1}{2}P(\text{Unentschieden}) = F(\lambda_A, TS) \quad (7)$$

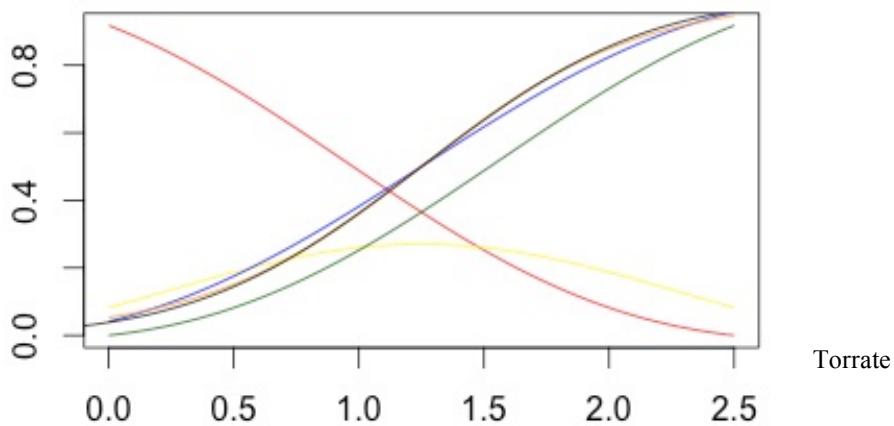


Fig. 1. Wahrscheinlichkeiten für Toto-Tipps mit Skellam-Verteilung und Vergleich mit Elo-Modell

Allerdings gilt $\Delta r = 0$ im Elo-Modell bei gleich starken Mannschaften, während dies im Poisson-Modell für $\Delta\lambda = \lambda_A - \lambda_B$ gilt. Um die beiden Modelle für eine vorgegebene Torsumme TS zu vergleichen, müssen wir im Elo-Modell den Ursprung noch um $TS/2$ verschieben. Dann können wir die Kurven in dasselbe Diagramm zeichnen, siehe Figur 1. Neu hinzugekommen ist die Gewinn-Wahrscheinlichkeit im Poisson-Modell (blaue Kurve, nach (7)) sowie die Gewinnwahrscheinlichkeit im Elo-Modell (die orange Kurve zeigt die Näherung nach (3), die schwarze Kurve die exakte Formel (4)).

Man erkennt vor allem im zentralen Bereich die sehr gute Näherung von (4) für (3) als auch die für praktische Auswertungen gute Übereinstimmung der beiden Seiten von (7). Die Elo-Formel hat im Vergleich zum Poisson-Modell eine geringere Streu-

ung, d. h. der Vorteil der stärkeren Mannschaft ist etwas höher als im Poisson-Modell.

Im Vergleich zu den anderen Ungenauigkeiten erscheint dieser Unterschied aber gering und man kann jetzt die Modelle ineinander umrechnen. Formal ergibt sich dann für ein vorher bestimmtes TS

$$\lambda_A = F^{-1}(W(\Delta r + TS/2), TS) \quad (8)$$

Mathematisch ist die Umkehrfunktion kompliziert. Man liest hier entweder die Werte aus Figur 1 ab oder berechnet sie direkt mit R.

2.7 Prognose von Fussballspielen aufgrund der „Weisheit der Vielen“

Surowiecki beschreibt das Prinzip in seinem Buch [14]. Das Prinzip ist vielen auch z. B. aus der Quiz-Show „Wer wird Millionär“ als Publikums-Joker bekannt. Wenn der Kandidat die Antwort nicht sicher weiß, lässt er das Publikum abstimmen. In der Regel wählt er dann die Antwort, die die meisten getippt haben und liegt damit in über 90% der Fälle richtig. Insbesondere ist die Quote deutlich besser als beim Experten- oder Telefon-Joker, bei dem nur eine Person befragt wird.

Dieses Prinzip kann man auch bei der Prognose von Fussballspielen anwenden. Hier gibt es ebenfalls nur eine richtige Antwort, die allerdings erst später feststeht. Nach Surowiecki muss aber eine ausreichende Meinungsvielfalt und Unabhängigkeit in der Gruppe gegeben sein, damit man gute Ergebnisse bekommt.

2.8 WM-Prognosen

Vor jeder WM werden zahlreiche Prognosen veröffentlicht, in der Regel Tipps auf das Abschneiden bestimmter Mannschaften oder „Wer wird Weltmeister?“. 2014 hat sich sogar der weltberühmte Physiker Stephen Hawking daran beteiligt [13], allerdings wohl nicht ganz ernsthaft. Dabei konnte man beobachten, dass die Weltmeister- oder Halbfinal-Tipps alle relativ ähnlich waren, nur Tolan wich mit seinem eindeutigen Favoriten Deutschland stark ab [14].

Allen Prognosen scheint gemein, dass sie entweder Formeln aufgrund von Regressionsmodellen aufstellen oder den WM-Verlauf simulieren. So wurden z. B. kurz vor der WM an der FU Berlin ein WM-Simulator im Rahmen einer Bachelorarbeit [15,16] und an der Universität Frankfurt am Main wurden neue Simulationsergebnisse [17] veröffentlicht.

3 Prognosemodelle für die WM

Im Tippwettbewerb der Cäcilienchule Wilhelmshaven [1] müssen die Tipps komplett vor dem ersten Spiel abgegeben werden. Dadurch werden einige Berechnungen

vereinfacht, da sonst die geschätzten Leistungsstärken der Mannschaften während des Turniers in Abhängigkeit der erzielten Ergebnisse angepasst werden müssten. Für korrekt vorhergesagte Platzierungen gibt es Zusatzpunkte und nur, wer die KO-Partien aufgrund seiner Vorrundentipps korrekt vorhergesagt hat, kann in der Hauptrunde noch weitere Punkte machen. Dies weicht z. B. von der Kicktipp-Plattform [18] ab, bei der man jedes Spiel bis zum Anpfiff tippen kann.

Außerdem wurden aufgrund der abgegebenen Einzel-Tipps nach dem Ansatz der „Weisheit der Vielen“ Tipps für „Tippgemeinschaften“ gebildet. Dabei kann man entweder den absolut häufigsten Tipp verwenden, oder man kann erst den häufigsten Toto-Tipp ermitteln und unter den zugehörigen Ergebnissen das am häufigsten getippte Ergebnis auswählen (sog. „Tendenz-Tipp“). Zusätzlich wurde noch zum Vergleich auf Tipps größerer Tippgemeinschaften wie Kicktipp zurückgegriffen.

Für die Systemtipps wurden erst einmal Ranglisten untersucht und versucht, die Informationen aus den Ranglisten auf das Poisson-Modell abzubilden. Außerdem wurde untersucht, wie aus den Quoten der Wettbüros Daten für das Poisson-Modell ermittelt werden können. Zusätzlich wurde noch angestrebt, aus den Daten der Marktwerte [22] der Nationalteams ein Poisson-Modell zu ermitteln. Schließlich wurden auch noch Tipps aus kombinierten Modellen, z. B. Rangliste und Marktwert, gebildet. Alle benötigten Daten wurden kurz vor Beginn der WM erhoben.

Figur 2 stellt in einer Übersicht die Zusammenhänge und Abhängigkeiten der verschiedenen WM-Spielprognosen dar. Die durchgezogenen Pfeile entsprechen dem oben dargestellten Forschungsstand. Die gestrichelt dargestellten Zusammenhänge wurden im Rahmen dieser Arbeit entwickelt und werden in den folgenden Unterabschnitten erläutert.

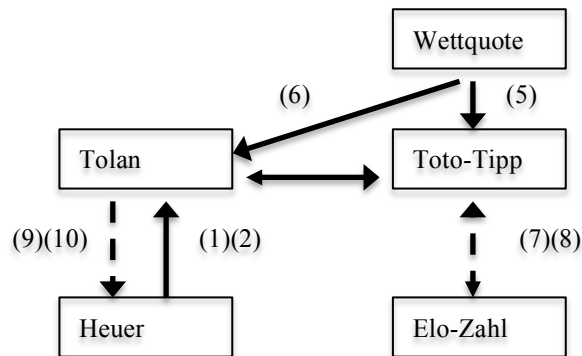


Fig. 2. Übersicht über Prognose-Modelle und Ergebnisse der Arbeit

3.1 Vorab-Prognose aufgrund von Wettquoten

Bei der Spielprognose habe ich die Formeln (1) und (2) benutzt, um aus den geschätzten Leistungsstärken (Statistiken) die erwarteten Torzahlen der beiden Mannschaften zu berechnen. Da ich keine verlässlichen Statistiken für Nationalmannschaften ähnlich der Bundesliga habe, versuche ich nun die Leistungsstärken aus den erwarteten Torzahlen für ein Spiel zu schätzen. Die erwarteten Torzahlen werden mit Hilfe von (6) berechnet. Allerdings stellen die Wettanbieter wie [20] vor WM-Beginn nur die kompletten Wetten für die ersten Spielpaarungen bereit.

Beispiel: Für das erste Spiel Deutschland gegen Portugal betragen die Quoten für Null Tore 4,50 bzw. 2,50. Um den Gewinn der Wettanbieter von etwa 8% auszugleichen, muss man diese Quoten noch entsprechend erhöhen. Dann ergeben sich mit (6) erwartete Torzahlen von 1,58 und 0,99, d.h. $LS_D - LS_P = 0,59$.

Dies bedeutet auch, dass man für ein Spiel nur eine Gleichung mit zwei Unbekannten erhält, denn in (2) kommen die Leistungsstärken nicht vor. Mit dem zweiten Gruppenspiel erhält man also zwei Gleichungen für vier Unbekannte. Für die anderen Spiele sind vor WM-Beginn nur die Quoten für Toto-Tipps bekannt. Aus den Wettquoten für die Toto-Tipps kann man die Wahrscheinlichkeiten für die Spielausgänge berechnen und mithilfe von Figur 1 die erwarteten Torzahlen schätzen. Dabei ist allerdings noch die Torsumme unbekannt, die bei Weltmeisterschaften mit 2,5 im Durchschnitt bisher geringer war als in der Bundesliga.

Beispiel (fortgesetzt): Für das zweite Gruppenspiel von Deutschland gegen Ghana betragen die Toto-Tipp-Quoten 1,33 bzw. 8,25. Nach Berücksichtigung des Gewinns ergeben sich erwartete Torzahlen von etwa 1,9 und 0,6, d.h. $LS_D - LS_G = 1,3$. Daraus folgt, dass man mit den ersten vier Spielen ein lineares Gleichungssystem mit vier Gleichungen und vier Unbekannten für die Leistungsstärken der vier Mannschaften erhält.

$$\begin{aligned}LS_D - LS_P &= \lambda_D - \lambda_P \\LS_G - LS_U &= \lambda_G - \lambda_U \\LS_D - LS_G &= \lambda_D - \lambda_G \\LS_G - LS_U &= \lambda_G - \lambda_U\end{aligned} \quad (9)$$

Leider erweisen sich die Gleichungen als nicht unabhängig, d. h. das Gleichungssystem ist nicht eindeutig lösbar. Zieht man von der ersten Gleichung die zweite und dann noch einmal die dritte Gleichung ab, so erhält man die vierte Gleichung usw.. Dies kann man auch leicht nachvollziehen, da alle Gleichungen nur Differenzen darstellen und irgendwie ein fester Punkt fehlt, auf den sich die Differenzen beziehen.

Man könnte sich aber diesen Fixpunkt beschaffen, indem man sich für eine Mannschaft die Leistungsstärke aus einer anderen Quelle besorgt. Z. B. kann man für die jeweils schwächste Mannschaft in der Gruppe die Wettquote im Internet nachschauen, dass diese Mannschaft Weltmeister wird. Daraus kann die Wahrscheinlichkeit mit (5) berechnet werden. Leider fehlte in der Literatur [5] ein Zusammenhang mit der Leis-

tungsstärke. Dieser wird im nächsten Abschnitt entwickelt und damit kann man aus der Wettquote die Leistungsstärke einer Mannschaft vor der WM bestimmen und dann das Gleichungssystem (9) lösen.

3.2 Weltmeister-Prognose

Um die Platzierung der Mannschaften in den Gruppen oder für die gesamte Weltmeisterschaft vorherzusagen, braucht man nur die Mannschaften nach ihrer Leistungsstärke oder die Position in der Rangliste zu ordnen. Dies ist die wahrscheinlichste Reihenfolge. Wenn man Aussagen über die Unsicherheit der Prognose machen will, kann man die Wahrscheinlichkeiten bestimmen, mit denen eine Mannschaft einen bestimmten Platz belegt, z. B. Gruppensieger wird. Stand der Technik ist hier die Simulation, siehe z. B. in [15,16,17,18] sowie [21]. Ich habe bei meinen Recherchen keine Quelle gefunden, die die Prognose berechnet anstatt zu simulieren, obwohl mit dem Poisson-Modell eigentlich der komplette Spielausgang statistisch beschrieben wird. Daher habe ich den folgenden Ansatz für eine exakte Berechnung entwickelt:

1. Bestimme für jedes Gruppenspiel die erwarteten Torzahlen beider Mannschaften
2. Bestimme für jedes (realistische) Ergebnis dessen Wahrscheinlichkeit
3. Bestimme für die Gruppe alle Ergebnis-Kombinationen und deren zugehörige Wahrscheinlichkeit
4. Ermittle die Abschlusstabelle und Platzierung der Mannschaften
5. Addiere die entsprechenden Platzierungswahrscheinlichkeiten

Dieser Ansatz entspricht der direkten Auswertung mehrstufiger Zufallsexperimente, allerdings ist die Auswertung komplizierter. Der Ansatz ist für alle Prognose-Modelle möglich, für die erwartete Torzahlen ermittelt werden können (vergleiche Figur 2). Es gibt in Schritt 3 sechs Stufen (die Gruppenspiele), aber in jeder Stufe (Schritt 2) gibt es unendlich viele Spielausgänge, deswegen muss man sich auf die typischerweise bei einer WM auftretenden Ergebnisse beschränken. Man kann statt der exakten Ergebnisse mit den Tordifferenzen arbeiten, und zwar mit Tordifferenzen zwischen -6 und +6, siehe Tabelle 1. Dort ist ein spezieller Ausgang rot hervorgehoben. In diesem Beispiel hätte Mannschaft A sieben Punkte und eine Tordifferenz von 3 erzielt. Sie wäre damit Gruppensieger vor D mit fünf Punkten geworden. Mit Hilfe der Skellam-Verteilung kann man nun die Wahrscheinlichkeit berechnen, mit der die sechs Spiele genauso ausgehen wie im Beispiel in Tabelle 1.

Es gibt 13 Möglichkeiten für jedes Spiel und insgesamt $13^6=4826809$ Kombinationen für jede Gruppe, für die das R-Programm auf meinem iMac etwa 10 Minuten benötigte. Manche Ausgänge sind dann zwar nicht eindeutig, z. B. wenn alle Spiele unentschieden ausgehen, aber dies ist in der Praxis ebenso und kommt nur mit einer Wahrscheinlichkeit kleiner als 0,0001 vor. In diesen Fällen wurde die Mannschaft höher platziert, die in der Rangliste besser steht. Auch die Anzahl der betrachteten Tordifferenzen wurde variiert und beobachtet, dass sich die Ergebnisse z. B. bei Tordifferenzen von -5 und +5 nur in der dritten Nachkommastelle unterscheiden.

Für Liga-Prognosen wie z. B. die Bundesliga mit 306 Spielen taugt der Ansatz allerdings selbst dann nicht, wenn man nur Toto-Tipps betrachten will, da man dann 3^{306} , also etwa 10^{145} Kombinationen berechnen müsste. Aber vielleicht lassen sich Qualifikationsgruppen für Weltmeisterschaften damit exakt prognostizieren.

Tordifferenz	>5	5	4	3	2	1	0	-1	-2	-3	-4	-5	<-5
A-B	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
C-D	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
A-C	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
B-D	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
A-D	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)
B-C	P(>5)	P(5)	P(4)	P(3)	P(2)	P(1)	P(0)	P(-1)	P(-2)	P(-3)	P(<-4)	P(-5)	P(<-5)

Tabelle 1. Exakte WM-Prognose mit mehrstufigem Zufallsexperiment

In den KO-Spielen wird die Spielprognose einfacher, da es keine Unentschieden mehr geben kann und das exakte Ergebnis egal ist. Hier kann man mit den Toto-Wahrscheinlichkeiten oder Gewinnwahrscheinlichkeiten rechnen. Da man für jede Gruppe für jede Mannschaft die Wahrscheinlichkeit kennt, dass sie Erster oder Zweiter wird, hat man für jedes Achtelfinale 16 verschiedene Paarungen zu berücksichtigen. Daraus kann man für jedes Achtelfinale für 8 Mannschaften die Wahrscheinlichkeit bestimmen, dass genau diese Mannschaft weiterkommt. Für jedes Viertelfinale hat man 64 verschiedene Möglichkeiten und bekommt die Wahrscheinlichkeiten für 16 Mannschaften. Dies gibt für jedes Halbfinale 256 Möglichkeiten bzw. für das Finale 1024.

Das Ergebnis ist überraschend: Trägt man die Leistungsstärke der Mannschaften z. B. aus dem SPI gegen die berechnete Wahrscheinlichkeit, Weltmeister zu werden, doppelt logarithmiert auf, so erhält man einen sehr guten linearen Zusammenhang, siehe Figur 3. Mit Hilfe einer Regressionsgeraden lässt sich mit einem Bestimmtheitsmaß $R^2 = 0,985$ folgender Zusammenhang bestimmen

$$\log P(\text{Weltmeister}) = 3,77 \log LS - 2,38 \text{ bzw. } P(\text{Weltmeister}) = \frac{LS^{3,77}}{238} \quad (10)$$

Der Vorteil dieses Ansatzes gegenüber Simulation besteht darin, dass die Ergebnisse numerisch exakt berechnet werden können und es keine statistischen Unsicherheiten gibt. Weiter besteht ein Vorteil bei der Rechenzeit, zumindest wenn eine hohe Genauigkeit angestrebt wird.

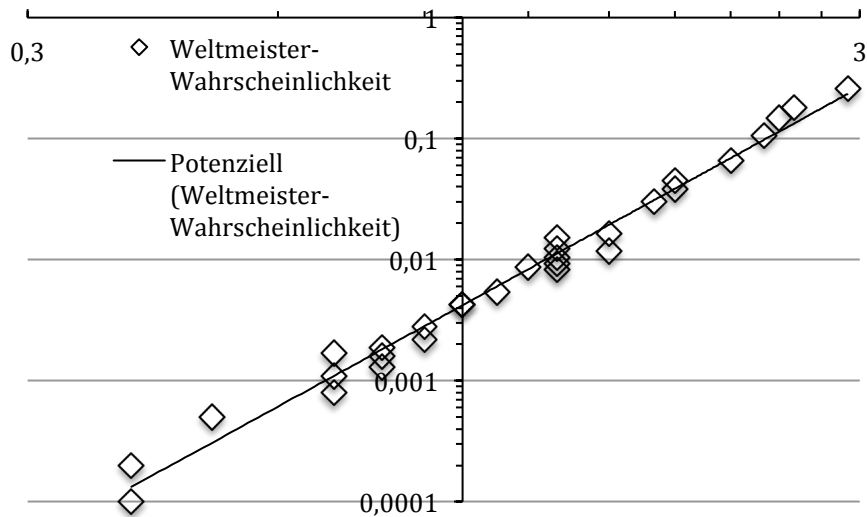


Fig. 3. Leistungsstärke und berechnete Weltmeister-Wahrscheinlichkeit für SPI

4 Auswertung des Tippspiels

Für die Teilnehmer des Tippspiels sowie für die Systemtipps habe ich getrennt Wahrscheinlichkeitsdichten mit R schätzen lassen, siehe die rote Kurve in Figur 4. Es fällt die deutlich steilere Form der Dichte auf, die einen wesentlich größeren Mittelwert und eine deutlich geringere Streuung aufweist als die andere Kurve.

Außerdem wurden die Systemtipps untereinander noch einmal qualitativ verglichen, indem die vorhergesagten Platzierungen der Mannschaften verglichen wurden, siehe Tabelle 2. Dabei konnte ich auch Tipps einbeziehen, für die es kein Modell gibt und die deswegen nicht an dem Tippwettbewerb teilnehmen konnten. Dort sieht man farblich, welche Mannschaften in Wirklichkeit besser (dunkelgrün) oder schlechter (rot) abgeschnitten haben als von den Tipps prognostiziert. Grüne Einträge stehen für exakt richtige Tipps und orange Einträge für etwas schlechtere Tipps. Ich habe ausgezählt, wie viel richtige Mannschaften die Tipps in den Stufen der KO-Runden vorhergesagt haben, die Unterschiede waren aber minimal. Fett gekennzeichnet sind die Tipps, die signifikant besser waren als der Zufall. Insgesamt am besten hat die „Weisheit der Vielen“ abgeschnitten, was aber daran liegt, dass die meisten Teilnehmer Deutsche waren, die aufgrund der „Lokalbrille“ häufig Deutschland ins Endspiel getippt haben. Nur dieser Systemtipp sowie der von Tolan haben Deutschland als Weltmeister vorhergesagt.

Zusätzlich wurde der Tipp der Plattform Kicktipp [18] nach der Methode „Weisheit der Vielen“ ausgewertet. Hier zeigten sich aber kaum Unterschiede zu den Systemtipps unseres Tippwettbewerbs. Die Kicktippdatei enthielt insgesamt 28864 Tipps, darunter aber nur 10785 Tipps vollständige. Aus diesen ermittelte ich für jedes Spiel sowohl den Roh-Tipp (häufigstes Ergebnis), aber auch den Tendenz-Tipp (d. h. häufigstes Ergebnis bei dem häufigsten Tototipp eines Spiels). Der Kicktipp-Tippwettbewerb unterscheidet sich insofern von unserem Wettbewerb, dass jedes Spiel erst getippt wird, wenn die Gegner bekannt sind. Für die Vorrunde macht dies keinen Unterschied, aber in der Endrunde waren die Kicktipp-Teilnehmer dadurch im Vorteil, so dass wir die Endrunde nicht vergleichen konnten. Der Roh-Tipp erreichte in der Vorrunde eine Punktzahl von 89 Punkten und der Tendenz-Tipp 92. Damit lag er nach der Vorrunde knapp hinter dem Tendenz-Tipp des Tippwettbewerbs, der nach der Vorrunde sogar auf Platz 39 der Gesamtwertung lag.

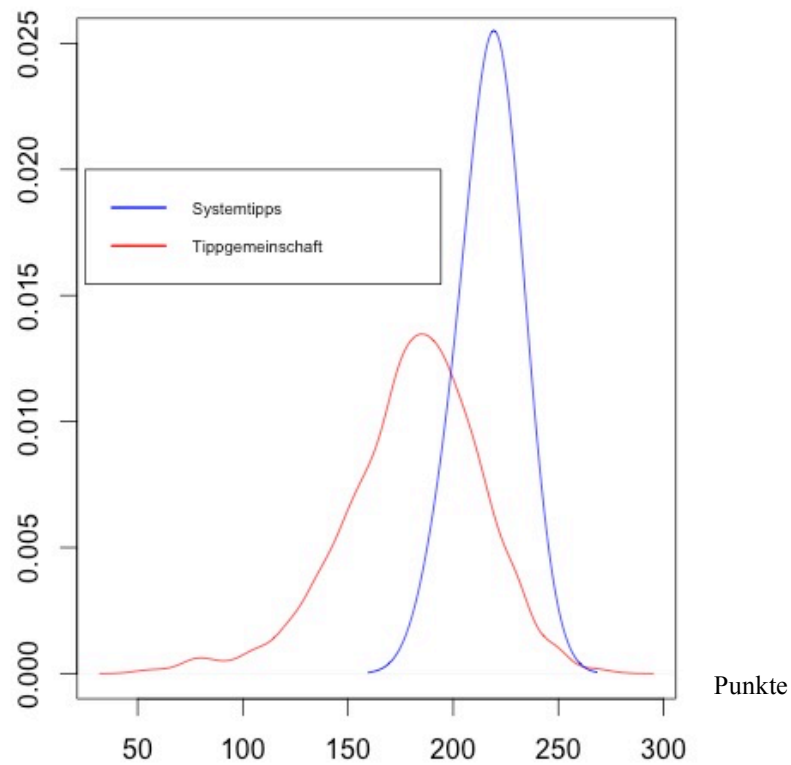


Fig. 4. Verteilung der Punkte für menschliche Teilnehmer und Systemtipps

5 Zusammenfassung und Diskussion

Zur Vorbereitung wurden die Zusammenhänge zwischen den verschiedenen Prognosemodellen ermittelt (s. Figur 2). Viele Zusammenhänge waren durch die Literatur bekannt, aber einige mussten neu entwickelt werden, wie z.B. die Berechnung der Leistungsstärke über die Wettquote oder den Zusammenhang zwischen Gewinnwahrscheinlichkeit und Toto-Tipp (Formeln (7), (8), (9), (10)). Damit wurde ein geschlossener Zusammenhang zwischen den Tipps ermittelt, der in dieser Form in der Literatur nicht vorkam.

	Tolan [14]	Ludwig [18]	Dormagen [16]	SPI [6]	bwin [20]	Elo [8]	FIFA [7]	Marktwert[22]	WdV [1]
WM	Deutschland	Brasilien	Brasilien	Brasilien	Brasilien	Brasilien	Spanien	Spanien	Deutschland
Vize	Niederlande	Argentinien	Spanien	Argentinien	Argentinien	Spanien	Deutschland	Deutschland	Brasilien
HF	England	Deutschland	Deutschland	Deutschland	Deutschland	Deutschland	Brasilien	Brasilien	Spanien
HF	Bosnien	Spanien	Argentinien	Spanien	Spanien	Argentinien	Portugal	Argentinien	Argentinien
VF	Ghana	Uruguay	Portugal	Chile	Belgien	Niederlande	Argentinien	Frankreich	Portugal
VF	Brasilien	Italien	Kolumbien	Frankreich	Italien	England	Schweiz	Belgien	Italien
VF	Elfenbeinküste	Frankreich	Frankreich	Kolumbien	Frankreich	Portugal	Uruguay	Portugal	Belgien
VF	Argentinien	Niederlande	Uruguay	Uruguay	England	Kolumbien	Kolumbien	Italien	England
AF	Algerien	Portugal	Chile	England	Portugal	Mexico	Italien	England	Kroatien
AF	Japan	Japan	England	Belgien	Niederlande	Uruguay	Belgien	Kolumbien	Niederlande
AF	Italien	Kolumbien	Mexico	Mexiko	Bosnien	Griechenland	Griechenland	Kroatien	Griechenland
AF	Russland	Russland	Belgien	Elfenbeinküste	Mexiko	Bosnien	Chile	Niederlande	Frankreich
AF	Frankreich	Belgien	Schweiz	Bosnien	Kolumbien	Ecuador	Bosnien	Schweiz	Schweiz
AF	Schweiz	Schweiz	Russland	Ecuador	Schweiz	Frankreich	Frankreich	Bosnien	Nigeria
AF	Chile	Nigeria	Elfenbeinküste	Portugal	Russland	Belgien	Kroatien	Elfenbeinküste	Russland
AF	Kroatien	Kroatien	Bosnien	Russland	Japan	Russland	Russland	Russland	Kolumbien
AF	8	10	10	9	9	10	10	8	10
VF	4	5	5	5	5	5	4	5	4
HF	2	3	3	3	3	3	2	3	3
F	1	1	0	1	1	0	1	1	1
WM	1	0	0	0	0	0	0	0	1

Tabelle 2. Qualitative Auswertung der WM-Systemtipps

Außerdem beschränkt sich die Literatur auf die Simulation von WM-Ergebnissen. Dies wurde durch persönliche Kommunikation mit den Wissenschaftlern der Universität Frankfurt am Main sowie der TU Berlin bestätigt [21]. Dagegen habe ich eine Möglichkeit gefunden, mit der der WM-Verlauf numerisch berechnet werden kann. Der Vorteil gegenüber einer Simulation liegt insbesondere darin, dass weniger Rechenzeit benötigt wird und die Ergebnisse bis auf Rundungsfehler exakt sind. Natürlich ergeben sich prinzipiell dieselben Resultate und man muss auch zugeben, dass dieser Ansatz nur für kleine Gruppengrößen und wenige KO-Spiele, wie das bei WM, EM oder Champions League der Fall ist, funktioniert. Insbesondere ist es mit diesem

Ansatz gelungen, einen direkten Zusammenhang zwischen der Leistungsstärke einer Mannschaft und der Wahrscheinlichkeit, Weltmeister zu werden, herzustellen (Formel (10)). Dies hat sich als nützlich für die Bestimmung der Leistungsstärken aus Wettquoten herausgestellt.

Die Systemtipps haben im Mittel über alle Teilnehmerklassen am besten abgeschlossen, aber den Tippwettbewerb nicht gewonnen. Die meisten Systemtipps sind auf den Plätzen 100 bis 200 gelandet und waren dutzende Punkte hinter den besten Tippfern. Dies lässt sich mit Figur 4 erklären. In dieser Abbildung sind die mit R geschätzten Dichtefunktionen der Punkte von menschlichen Teilnehmern und Systemtipps aufgetragen. Man erkennt, dass die Systemtipps im Schnitt deutlich besser sind. Auch ist die Streuung viel kleiner. Die Systemtipps sind zwar im Mittel besser, aber je größer die Runde ist, mit der getippt wird, desto wahrscheinlicher ist es, dass zufällig menschliche Tipper besser abschneiden als der beste Systemtipp. In unserem Fall waren dies 86 Tipper, die besser waren als jeder Systemtipp. Dies liegt an der großen Streuung der Punkteverteilung der menschlichen Tipper. Je mehr mitmachen, desto größer wird die Wahrscheinlichkeit, dass zufällig einer der menschlichen Tipper ein besonders gutes Ergebnis erzielt. Diesen Effekt kann man sich auch veranschaulichen, wenn man sich 10 Münzwürfe mit einer gezinkten Münze vorstellt, sagen wir mal Kopf kommt mit Wahrscheinlichkeit 0,8 vor. Der Systemtipp kennt das Modell und würde damit im Schnitt 8 von 10 Würfeln richtig vorhersagen. Aber z. B. bei 1000 Tippern, die zufällig tippen, hätte im Schnitt einer alle Würfeln richtig vorhergesagt und etwa 10 hätten neun Würfeln richtig vorhergesagt. Das heißt, wenn man sich auf Systemtipps verlässt, wird man in der Regel bei Tipprunden gut abschneiden, aber bei großen Tipprunden wird man nur selten gewinnen.

Die Unterschiede zwischen den Systemtipps lassen sich sowohl qualitativ als auch quantitativ gut durch zufällige Schwankungen erklären, d. h. wenn man in einer Tipprunde mit einem Systemtipp mitspielen will, kann man sich den Tipp aussuchen, der am einfachsten zu bestimmen ist. Dies ist aus meiner Sicht der Tipp, der auf Wettquoten beruht. Man kann entweder den Tipp nehmen, der die geringste Quote bietet, oder man schaut erst auf den Toto-Tipp, und sucht sich dann unter den Tipps mit der Tendenz des Toto-Tipps denjenigen aus, der die geringste Quote bietet.

Danksagung. Dieses Projekt wurde im Rahmen von "Jugend Forscht" am Gymnasium Neue Oberschule in Braunschweig sowie der Cäcilien- und Wilhelmshavener Schulen bearbeitet. Ich danke den Lehrern Fritz Maul und Thomas Eberhardt für die Betreuung und meinen Eltern für die Unterstützung. Bei meinen Mitautoren der Jugend-Forscht-Arbeit, Christoph Friedrich und Antonia Lücke, bedanke ich mich insbesondere für die Durchführung und Auswertung des Wilhelmshavener Tippspiels. Außerdem hat die Firma Kicktipp das Datenmaterial ihres öffentlichen WM-Tippspiels zur Auswertung zur Verfügung gestellt. Ganz besonders bedanke ich mich natürlich auch bei den vielen Teilnehmern der Tipp-Wettbewerbe.

Referenzen

1. Tippwettbewerb zur WM 2014 der Cäcilienchule Wilhelmshaven, www.tippwettbewerb.de (07.01.2015)
2. Braband, M., Friedrich, C. und Lücke, A.: Wer hat den besten WM-Tipp - die Experten, die große Masse, die Mathematik oder doch das Wettbüro?, Projekt Nr. 130423, Jugend Forscht (2015), <http://www.kicktipp.de/download/presse/2015-02-25-jugend-forscht-wm-tippspiel.pdf> (04.04.2015)
3. Tolan, M.: Manchmal gewinnt der Bessere - Die Physik des Fussballspiels, Piper, 2011
4. Wikipedia: http://en.wikipedia.org/wiki/Skellam_distribution (07.01.2015)
5. Heuer, A.: Der perfekte Tipp, Wiley-VCH, Weinheim, 2012
6. ESPN: Soccer Power Index, <http://www.espnfc.com/spi/rankings?cc=5739> (31.05.2014)
7. FIFA: FIFA-Weltrangliste, de.fifa.com/fifa-world-ranking/ranking-table/men/index.html (07.01.2015)
8. Wikipedia: World Football Elo Rankings, en.wikipedia.org/wiki/World_Football_Elo_Ratings (07.01.2015)
9. Wikipedia: Elo-Zahl, de.wikipedia.org/wiki/Elo-Zahl (07.01.2015)
10. Braband, M.: Fußball ist ein Würfelspiel oder: Mit welcher Wahrscheinlichkeit schafft Eintracht Braunschweig den Klassenerhalt?, Jugend Forscht Projekt Nr. 123374 (2014),
11. RStudio, www.rstudio.com
12. Surowiecki, J.: Die Weisheit der Vielen. Warum Gruppen klüger sind als der Einzelne („The wisdom of crowds“). Goldmann, München 2007
13. Spiegel Online, Physiker Hawking stellt Erfolgsformel für Englands Kicker vor, <http://www.spiegel.de/sport/fussball/fussball-wm-in-brasilien-stephen-hawking-und-die-erfolgsformel-a-972206.html>, (31.05.2014)
14. Brasilien, Spanien oder wir! DIE ZEIT Nr. 23/2014, 28.5.2014, S. 18
15. Rojas, R.: Wer gewinnt die Fußballweltmeisterschaft?, 28.5.2014 <http://www.heise.de/tp/artikel/41/41867/1.html> (31.05.2014)
16. Dormagen, D.: Development of a Simulator for the FIFA Worldcup 2014, Bachelorarbeit FU Berlin, 13.05.2014 http://www.inf.fu-berlin.de/inst/ag-ki/rojas_home/documents/Betreute_Arbeiten/Bachelor-Dormagen.pdf (31.05.2014)
17. Errechnet: Brasilien wird Weltmeister, DIE WELT, 27.5.2015, <http://www.welt.de/sport/article128441272/Errechnet-Brasilien-wird-Fussball-Weltmeister.html> (31.05.2014)
18. DIE ZEIT: Meisterchancen, Infografik Nr. 260 vom 12.6.2014, basierend auf Simulationen von Prof. Ludwig, Universität Frankfurt am Main
19. Kicktipp Tippenspiel, www.kicktipp.de (07.01.2015)
20. bwin Sportwetten, bwin.de (07.01.2015)
21. Private E-Mail-Kommunikation mit Prof. Ludwig, Prof. Rojas sowie Herr Dormagen
22. Transfermarkt, www.transfermarkt.de (07.01.2015)