

Please request reprints and cite the original paper!

This text is published in the following Reference:

Raab, M. & Gröschner, C. (2002). Wer wird Fußball Weltmeister 2002? In B. Strauß, M. Tietjens, N. Hagemann & A. Stachelhaus (Hrsg.), *Expertise im Sport: Lehren – lernen – leisten* (S. 223-224). Köln: bps-Verlag.

## **Wer wird Weltmeister im Fußball 2002?**

Markus Raab & Christian Gröschner

Max-Planck-Institut für Bildungsforschung, Center for Adaptive Behavior and Cognition, Berlin

Schlüsselwörter: Ergebnisvorhersage, Fußball, Heuristiken, Statistische Modelle

### **Einleitung**

Eine faszinierende Komponente des Sports ist die Unsicherheit wer einen Wettkampf gewinnen wird. Trotzdem wird Leistung aus unterschiedlichsten Gründen vorhergesagt. An einem aktuellen Beispiel, den Olympischen Winterspielen 2002 in Salt Lake City sei die Komplexität einer Vorhersage der erfolgreichsten Nation oder noch schwieriger der prognostizierten Anzahl von gewonnenen Medaillen erläutert. In einem Paarvergleich z.B. zwischen der USA und Deutschland soll die höhere Anzahl von (Gold)Medaillen vorhergesagt werden. Johnson & Ali (2002) benutzten ein ökonomisches Modell mit elf Parametern (z.B. Kapital des Landes, Regierungsform, Population, Anteil und Dauer des Landes unter Frost) von denen nur ein Sportparameter, der Heimvorteil benutzt wurde. Die Vorhersage von Johnson und Ali erreichte bei der Rangkorrelation mit der Gesamtmedaillenliste einen Wert von  $r = .67$  und mit den Rängen nach Anzahl der Goldmedaillen einen Wert von  $r = .71$ . Mit nur zwei Parametern (Heimvorteil, gewichtet mit 1.04 und Position aus der Medaillenrangliste in Nagano 1998) kann eine Korrelation von  $r = .79$  erreicht werden. Die Benutzung weniger Informationen als Bestandteil nichtkompensatorischer Heuristiken (vgl. TakeTheLast-Heuristic, vgl. Gigerenzer & Goldstein, 1999;) führt zu besseren Ergebnissen und soll im folgenden auf die Fußball WM 2002 angewendet werden.

### **Vorsagemodelle im Fußball**

Im Fußball gibt es wenige Versuche Ergebnisse von Ligen oder Meisterschaften vorherzusagen. Um den Fußball-Weltmeister 2002 (64 Spiele von 32 Mannschaften) vorherzusagen, werden existierende Modelle verglichen, die sich in Anzahl und Art der verwandten Informationen sowie der Verrechnung der Informationen unterscheiden. Kompensatorische Prädiktionsmodelle (vgl. Stefani & Clarke, 1992, für einen Überblick) werden mit Modellen nichtkompensatorischer Verrechnung (TakeTheLast-Heuristic) verglichen. Stefani (1983) sagte für vier europäische Ligen (4346 Spiele) den Gewinner mit einer 8% Fehlerrate mit einem Least-squares-Gaussian Prädiktionsmodell voraus. Es ist offen, welches Modell für Weltmeisterschaften bessere Vorhersagen macht. Die Vorhersagen beruhen auf verschiedenen Ranglisten und Hinweisreize. Die aktuelle FIFA-Rangliste (<http://www.fifa.com>; Stand Februar 2002) zeigt z.B. Deutschland auf dem zwölften bzw. dem elften Platz (FIFA „Big Count Survey

2000“), während die Benutzung der Ewigenliste der Weltmeisterschaften Deutschland auf dem zweiten Platz setzen würde. Neben den Rangpositionen aus den Ranglisten, wird die Anzahl gewonnener und verlorener Spiele sowie die Differenz erzielter Tore und Gegentore in der Vorbereitung, der Heimvorteil, die vorhergesagten Stärken und Schwächen von Mannschaften (vgl. Maher, 1982), länderspezifische Informationen (Population, Größe der Landfläche, Anzahl organisierten Fußballspieler, Anzahl der Vereine, Anzahl der Profis, etc.) benutzt.

### Vorsage zur Weltmeisterschaft 2002

Es ist offensichtlich, dass durch die Anzahl der verwandten Hinweisreize sowie der Art ihrer Gewichtung und Verrechnung unterschiedliche Vorhersagen getroffen werden. Tabelle 1 gibt die Art und Anzahl der benutzten Hinweisreize sowie ihre Gewichtung und Verrechnung an, die zur Vorhersage des Weltmeisters sowie der Gewinner der einzelnen Begegnungen benutzt werden.

*Tabelle 1: Vorhersage des Fußball-Weltmeisters 2002 nach Anzahl, Art, Gewichtung und Verrechnung der Hinweisreize. SL = (geschätzte Spielleistung in Angriff und Verteidigung), HV = Heimvorteil, Aktuelle Leistung (AL), SRS = Sport Rating System, RE = Random Error, LSG = Least-squares-Gaussian. HR = Hinweisreiz.*

| Modell                        | Anzahl HR | Art der HR    | Verrechnung                               | Vorhersage (1-4)                |                           |
|-------------------------------|-----------|---------------|---|---------------------------------|---------------------------|
| LSG (Stefani, 1983)           | 4         | SL, HV, RE    | Tore/Gegentore + HV + RE                  | 1. Kroatien<br>3. Nigeria       | 2. Senegal<br>4. Italien  |
| Maher-Poisson<br>Maher (1982) | 4         | SL, HV        | Tore/Gegentore + HV                       | 1. China<br>3. Nigeria          | 2. Kroatien<br>4. Italien |
| SRS Dixon & Coles (1997)      | 6         | SL, HV, AL    | FIFA Rangliste + HV (preliminary results) | 1. Frankreich<br>3. Argentinien | 2. Portugal<br>4. Italien |
| TakeTheLast                   | 2         | Random HR, AL | Letztplatzierung WM 1998 + HV             | 1. Frankreich<br>3. Argentinien | 2. Kroatien<br>4. Italien |

In die Berechnung der Vorhersage geht bei dem Modell LSG und Maher-Poisson für Frankreich, Japan und Korea (alle ohne Qualifikation) die Werte der FIFA (Final competition records) bei den anderen der FIFA (preliminary records der Qualifikation) ein.

### Diskussion

Die Überprüfung der Vorhersagen nach Abschluss der Weltmeisterschaften 2002 hat zwei Ziele. Erstens die Überprüfung der Vorhersagegenauigkeit existierender und auf dieses Problem angewandter Vorhersagemodelle. Dazu werden die anderen Hinweisreize (vgl. Vorhersagemodelle im Fußball) sowie die Platzierungen der anderen Mannschaften vorgestellt. Zweitens wird die Entwicklung von Strategien für die Vorhersage von Ergebnissen im Sport eruiert. Potenzielle Anwendung liegen im Wettverhalten von Zuschauern sowie in der Prognose von Sportverbänden oder Nationalen Olympischen Komitees.

### Literatur

Dixon, M. J. & Coles, S. G. (1997). Modelling association football scores and inefficiencies in the UK football betting market. *Journal of Royal Statistical Society Series C*, 46, 265-280.

- Gigerenzer, G. & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650-669.
- Johnson, D. K. N. & Ali, A. (2002). *A tale of two seasons: Participation and medal counts at the summer and winter Olympic games*. Working Paper 2002-02, January 2002. Wellesley, MA.
- Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica*, *36*, 109-118.
- Stefani, R. T. (1983). Observed betting tendencies and suggested betting strategies for european football pools. *The Statistician*, *32*, 319-329.
- Stefani, R. T. & Clarke, S. (1992). Predictions and home advantage for Australian rules Football. *Journal of Applied Statistics*, *19*, 251-259.