A Bradley-Terry type model for forecasting tennis match results

Ian McHale[a], Alex Morton[b],*

[a] *The University of Salford – Centre for Operational Research and Applied Statistics, Salford, Greater Manchester M5 4WT, United Kingdom*
[b] *SANSTAT LTD, 483 River Valley Road, #10-07, Singapore 248368, Singapore*

Abstract

The paper introduces a model for forecasting match results for the top tier of men's professional tennis, the ATP tour. Employing a Bradley-Terry type model, and utilising the data available on players' past results and the surface of the contest, we predict match winners for the coming week's matches, having updated the model parameters to take the previous week's results into account. We compare the model to two logit models: one using official rankings and another using the official ranking points of the two competing players. Our model provides superior forecasts according to each of five criteria measuring the predictive performance, two of which relate to betting returns.

© 2010 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

Keywords: Bradley-Terry model; Logit; Ranking evaluation; Sport; Betting

1. Introduction

Sports forecasting models are often used to inform debate on some wider aspect of research, rather than being the subject of the research themselves. In their most common application, as a tool for assessing the efficiency of betting markets, the published results of sports forecasting models have rarely been found to be successful to the extent of enabling positive returns to be made. This could be a consequence of the proprietary nature of a successful forecasting model, where neither an odds-setter nor a bettor with the capacity to beat the bookmaker would be eager to release his or her formula.

Attempting to win money is not the only possible use for an objective forecasting model. Such models can be used by the media (see for example Finkelstein, Graham, Morton, & Stott, 2002, or Klaasen & Magnus, 2003) to analyse the psychology of betting markets (see for example Dixon & Pope, 2004 or Graham & Stott, 2008), uncover the dynamics of a sport (see for example Dixon & Robinson, 1998, or Holder & Nevill, 1997), construct ranking systems (Macmillan & Smith, 2007), or aid in the design of tournaments (see for example Szymanski, 2003).

In this paper we concentrate on forecasting tennis match results using a Bradley-Terry type model. Previous published papers on the subject

---

* Corresponding author.
  E-mail addresses: i.mchale@salford.ac.uk (I. McHale),
sanstat1@hotmail.com (A. Morton).

of forecasting tennis match results have employed official ATP (Association of Tennis Professionals) rankings or tournament seedings for producing match forecasts, rather than historical results. Klaasen and Magnus (2003) use a function of the ATP rankings of the two competing players to evaluate the probability of a single point being won, then use these to infer the probabilities of each player winning the match. The authors then demonstrate how they can track this probability over the course of the match. Boulier and Stekler (1999) show that the seedings of the top 16 players (with a dummy variable to represent non-seeds) are informative in predicting match outcomes in Grand Slam tennis tournaments. Clarke and Dyte (2000) also focus on Grand Slam tournaments, but use the difference in ATP ranking points between players, rather than the position, as a predictor. More recently, Corral and Prieto-Rodriguez (2010) compared the predictive powers of three types of variables: match characteristics (such as tournament and surface), player characteristics (for example, age and height) and past performance, as measured by the ATP rankings, in a probit model for the match winners in Grand Slam tournaments. The authors find, unsurprisingly, that past performance is the most important of the three types of variables in their series of fitted models.

However, the usefulness of official rankings or seedings as predictors is called into question in the study by Scheibehenne and Broder (2007), who predicted match winners to be the higher (better) ranked player, then compare these predictions with the predictions of randomly selected volunteers with no knowledge of tennis. The volunteers were able to pick winners with the same hit rate as the official ranking predictions, provided that they were familiar with at least one of the players. Meanwhile, the bookmaker's odds identified the winner almost 10% more often than the official rankings. Although this was based on the results of a single tournament, it is quite a damning indictment of the ATP rankings as useful predictors. Clarke and Dyte (2000) justify using the official rankings for prediction purposes, rather than the actual results, by saying that it is too difficult to update a results database continually. However, online resources such as www.tennis-data.co.uk now make the collection and manipulation of data straightforward.

Tennis is not the only sport to have had its officials' rankings called into question as useful predictors of future performance. McHale and Forrest (2007) find that for men's professional golf, additional forecasting power can be provided by adding recent results to a forecasting model which already uses world rankings as a predictor. In a similar study for soccer, McHale and Davies (2007) again find that there is additional information for forecasting match results in the recent results of international teams. Thus, the evidence from tennis, golf and soccer suggests that although official rankings of players and teams are useful as predictors, they do not contain all of the information which is relevant for forecasting results.

Unlike previous papers, we do not restrict our modelling approach to models which use information on official rankings. Our application of the Bradley-Terry model uses historical match results to obtain forecasts, and we show that these forecasts are more accurate, according to several criteria, than the forecasts obtained from standard models employed in the literature. In addition, our model is also used to assess the influence of the surface on match outcomes, with the results suggesting that tennis played on clay is very different to tennis on other surfaces. Lastly, unlike other previous studies, by updating the model as new results are recorded, our model provides out-of-sample forecasts which are good enough to enable positive returns from betting.

The paper is structured as follows. Section 2 presents the data and our model. Section 3 provides a comparison of our model with ranking-based models, and also disaggregates our model and identifies which information contributes to our model outperforming the official rankings based models. Section 4 considers using our model as the basis of a betting strategy, and some closing remarks are given in Section 5.

2. Data and model

We obtained match results on the top tier of men's professional tennis, the ATP tour, for nine seasons from 2000–2008, from www.tennis-data.co.uk. The details given are participants' names and ATP rankings, the match results in games and sets, the date of the match, tournament name, location, surface (hardcourt, carpet, clay or grass) and series. 'Series' relates to a tournament's importance in terms of