Analyzing the Effect of Surface in Tennis Grand Slams

Kayla Frisoli, Shannon Gallagher, and Amanda Luby

September 07, 2018

Abstract

Tennis grand slams consist of the Australian Open, French Open, Wimbledon, and US Open, which are played on hard (Plexicushion), clay, grass, and hard (DecoTurf) courts, respectively. The surface type may substantially impact ball speed, height, and spin as well as player speed and agility. It is also believed that play style and practice habits may contribute to different results across surface types. For example, Rafael Nadal is thought to be the best clay court player of all time whereas Roger Federer is particularly known for dominance at Wimbledon. On the women's side, Serena Williams once struggled on clay courts but has seemingly transformed her style to perform better on clay courts, but has perhaps suffered on grass as a consequence. In this analysis, we examine the result of the top 100 players in grand slams from 2013-2017 across the four different surfaces. We create a hierarchical model with fixed and random effects to predict the number of points won in a match. We take into consideration player-specific effects, nationality (which is thought to have an effect on play style), sex, ranking, ELO, and game statistics. We assess the fit of our model using standard statistical techniques (e.g. MSE, AIC, BIC, residual diagnostics) in addition to 'common knowledge' factors (for instance, Rafael Nadal should be indicated as a superior clay court player by the model). We compare the results of top 100 players across grand slams to examine the effect of court surface. We also provide an in-depth analysis of Nadal, Federer, and S. Williams.

Introduction

Example reference(Centers for Disease Control and Prevention 2018)

Data

The data consists of 5080 matches split evenly over the four grand slams and the two leagues (ATP and WTA). Each match has 80 attributes, many of which are redundant. We focus on the following attributes for both the winner and loser of the match: games won, points won, retirement, break points faced, break points saved, aces, country of origin, and player attributes. Additionally, we take into account the number of sets in a match, the surface type, minutes played, and round of the tournament. A subset of the data is shown in Table 1.

Table 1: Example of the grand slam data. It includes winner and loser attributes, match attributes, and tournament attributes. Not all attributes are shown here.

Tournament	Year	W. Country	W. Points	Winner	W. Rank	L. Points	Loser
US Open	2014	USA	58	Serena Williams	1	31	Taylor Townsend
US Open	2013	ESP	136	Rafael Nadal	2	112	Philipp Kohlschreiber
Australian Open	2015	ESP	91	Rafael Nadal	3	51	Mikhail Youzhny
US Open	2013	USA	65	Serena Williams	1	41	Yaroslava Shvedova
Australian Open	2017	USA	73	Serena Williams	2	60	Lucie Safarova

Methods

Results

Discussion

References

Centers for Disease Control and Prevention. 2018. "Measles History." Available online at https://www.cdc.gov/measles/about.