Calvin Michael Floyd*, Matthew Hoffman and Ernest Fokoue

Shot-by-shot stochastic modeling of individual tennis points

https://doi.org/10.1515/jqas-2018-0036

Abstract: Individual tennis points evolve over time and space, as each of the two opposing players are constantly reacting and positioning themselves in response to strikes of the ball. However, these reactions are diminished into simple tally statistics such as the amount of winners or unforced errors a player has. In this paper, a new way is proposed to evaluate how an individual tennis point is evolving, by measuring how many points a player can expect from each shot, given who struck the shot and where both players are located. This measurement, named "Expected Shot Value" (ESV), derives from stochastically modeling each shot of individual tennis points. The modeling will take place on multiple resolutions, differentiating between the continuous player movement and discrete events such as strikes occurring and duration of shots ending. Multi-resolution stochastic modeling allows for the incorporation of information-rich spatiotemporal player-tracking data, while allowing for computational tractability on large amounts of data. In addition to estimating ESV, this methodology will be able to identify the strengths and weaknesses of specific players, which will have the ability to guide a player's in-match strategy.

Keywords: multiresolution; spatiotemporal data; stochastic modeling; sports; tennis.

1 Introduction

In tennis broadcasts and post-match commentary, the most common metrics shown are tally statistics – such as how many aces, unforced errors, or winners a player recorded – or simple percentages – such as what percentage of first serves were won by a player. While easy to compute, few of these common metrics really measure how a player is doing in terms of positioning or game strategy. Knowing that a player hit a winner gives important information about a result, but it hides information on what led to the won point. Did the player's opponent mis-hit the previous shot? Did the player set up the winner with a previous shot? Or did the player simply hit a brilliant shot? Distinguishing between these cases requires higher resolution spatiotemporal data. Professional tennis does have a system for collecting high-resolution spatiotemporal player and ball tracking data, but analysis of this data has lagged behind that of other sports.

Some of the most widely known analysis of player-tracking data is in basketball and the exploitation of camera-based tracking data for the National Basketball Association. The potential of this spatiotemporal data was explored by Kirk Goldsberry in 2012 with Goldsberry (2012), showing significant differences in players' shooting abilities. Later, the importance of players residing in certain regions on a basketball court was analyzed in Cervone, Bornn, and Goldsberry (2016a). New defensive metrics and evaluations based on spatial analysis were created in Franks et al. (2015) and Goldsberry and Weiss (2013) and have made their way into the play of the game. The strategy of teams to get players "open" to shoot the basketball was analyzed with the help of the spatiotemporal data in Lucey et al. (2014a). Also, Miller et al. were able to develop a machine learning approach to represent and analyze shot selection in Miller et al. (2014). Most closely related to this paper, stochastic modeling was applied to the NBA's player-tracking data to create an "Expected Possession Value" estimator in Cervone et al. (2016b).

Soccer is another sport which has embraced spatiotemporal player-tracking data, from an analysis perspective. Using soccer player-tracking data, classifications of passes in soccer matches were made in Horton et al. (2014), and shot prediction was improved in Lucey et al. (2014b). There has also been much analysis made on team strategy in Bialkowski et al. (2014); Lucey et al. (2013). In Brillinger et al. (2007) the single evolution of a play leading up to a goal by Argentina in the 2006 World Cup was analyzed, and the flow of the play was modeled by a stochastic process. Player-tracking data has recently become a staple of Major League Baseball broadcasts through their Statcast program. This includes the introduction of new metrics such as how optimal an outfielder's path was to a batted ball.

***Corresponding author: Calvin Michael Floyd,** Rochester Institute of Technology, Department of School of Mathematical Sciences, Rochester, NY 14623, USA, e-mail: calvin.michael.floyd@gmail.com
Matthew Hoffman and Ernest Fokoue: Rochester Institute of Technology, Department of School of Mathematical Sciences, Rochester, NY, USA, e-mail: mjhsma@rit.edu; epfeqa@rit.edu

As was mentioned earlier, professional tennis collects player-tracking data at many of its big tournaments. In Mora and Knottenbelt (2016), Mora and Knottenbelt, who used publicly-available 2012 Australian Open data, discussing the system used to produce the player-tracking data, stated: "*...the complexity and cost of this system make it only available to the main events at major tournaments. Therefore the data obtained with such systems is rich but limited in the number of matches and of difficult access.*" This sentiment is echoed in other papers, like Loukianov in Loukianov and Ejov (2012) who analyzes by-hand a 2009 Australian Open first set between Roger Federer and Rafael Nadal. He attempts to Markov model a tennis point, similar to this paper, but, ultimately, Loukianov simply comments on the "potential of Hawkeye TM technology," as the data is difficult to obtain. That being said, there has been interesting work done on the aforementioned 2012 Australian Open data. A Bayesian Network framework was used to "model player behavior using ball and player tracking information" in Wei et al. (2013). The same Australian Open data was used to evaluate fatigue and how it affected tennis players depending on variables such as how far they traveled during matches, how long each match lasted and how many days they had been playing for Reid and Duffield (2014). In Wei et al. (2015), Wei et al. were able to not only analyze the 2012 Australian Open data, but also tennis player-tracking data from the 2013 to 2014 Australian Open tournaments. They used their three tournaments worth of data to predict serve styles, based on the server's tendencies, the opponent's tendencies and the match context at the time of the serve. Also, in Spanias and Knottenbelt (2013), individual points were modelled as a Markov chain using historical serve data and improved on the predictive performance of existing models.

This paper proposes and develops an adaptation of the methodology of Cervone et al. (2016b) from basketball to tennis, utilizing tennis player-tracking data. Like basketball, tennis points can be broken down into discrete states and there is continuous player movement occurring at all times. We first develop the theory for stochastic modeling of the evolution of tennis points and then derive a measure for what we call the "expected shot value" (ESV) of each individual shot. This metric quantifies how many points a player can expect to score based on the circumstances at the time of a given shot, ideally correlating with win probability. The methodology behind the metric is tested using player-tracking data from the 2015 US Open men's tournament. This paper focuses on developing the theory and the US Open data is presented as validation for this proof-of-concept.

2 Theory

We first define a few terms which will be used throughout this paper. A *shot* will be a general term describing the process of making contact with the tennis ball and a *point* is a series of shots beginning with a serve and ending in one of several ways to be described later – with one player *winning* that point. For every tennis point, one player will be given a *server* label – the player who begins the point with a serve – and the other player will be labeled as the *receiver*. The server and receiver labels do not change between players during a point. For every shot during a given point, one player will be given a *striker* label and the other will be given a *returner* label. The striker will be the player who makes contact with the tennis ball (or the *strike*), and the returner will be the opposing player. Unlike the server and receiver labels, the striker and returner labels *will* change between players during the point, given that the point constitutes of more than one shot.

Let Ω represent the space of all possible tennis points in full detail, with $\omega \in \Omega$ describing the full path of a particular point. For any point path ω, we denote by $Z(\omega)$ the optical tracking time series generated by this point so that $Z_t(\omega) \in \mathcal{Z}$, with $t > 0$, is information contained in the player-tracking data exactly t seconds from the start of the point, beginning with the server setting up for his serve. $\mathcal{Z}$ is a high-dimensional space that includes the (x, y) coordinates for both players on the court, summary information such as which players are on the court, who struck the ball last, who is attempting to return that strike, how each shot is classified, and event annotations that are observable in real time, such as the ball being struck or the ball crossing the net. The possible point values of a particular point path ω for player i are either 0 or 1, denoted by $\Gamma^i(\omega) \in \{0, 1\}$. Taking the intuitive view of ω as a sample space of all individual tennis point paths, we define $Z_t(\omega)$ for each $t > 0$ as a random variable in $\mathcal{Z}$. Let $\mathcal{F}_t^{(Z)}$ represent the collection of all information from the player-tracking data up to time t of a particular point path ω such that $\mathcal{F}_t^{(Z)} = \{Z_s(\omega) : 0 \leq s \leq t\}$. We will define the *expected shot value*, or ESV, for player i at time $t \geq 0$ during a given point path ω, dependent on all available information up to time t, as follows:

$$v_t^i(\omega) = \mathbb{E}[\Gamma^i(\omega) | \mathcal{F}_t^{(Z)}] \tag{1}$$

Intuitively, (1) attempts to quantify how many points a specific player can expect to score based on all the available information about how the tennis point has played out up to a certain point t. In order to properly estimate the expected points, the player's tendencies as both a striker and returner will be taken into account.

Appendix A.1 defines ESV as a theoretical quantity, but we must develop a methodology to calculate it. In order to get the most out of the ESV estimator, we must require it to be stochastically consistent. Using a stochastically consistent ESV estimator guarantees that changes in the resulting ESV over the course of a point derive from players' on-court actions rather than artifacts or inefficiencies of the data analysis. In addition, we require that the ESV estimator be sensitive to the fine-grained details of the data without incurring undue variance or computational complexity. Applying a Markov chain-based estimation approach would require discretization of the data by mapping the observed spatial configuration $Z_t(\omega)$ into a simplified summary $C_t(\omega)$, which would potentially violate this criteria by trading potentially useful information in the player-tracking data for computational tractability.

In order to develop a methodology that meets both of the criteria above, we must understand that the information-computation trade-off results from choosing a single level of resolution at which to model the tennis point and compute all expectations. Adapting ideas from Cervone et al. (2016b), we estimate the ESV by combining models for the tennis point at two distinct levels of resolution. Specifically, we combine a fully continuous model of player movement and actions with a Markov chain model for a highly coarsened view of the point. This multi-resolution approach leverages the computational simplicity of a discrete Markov chain model while conditioning on exact spatial locations and high-resolution data features.

2.1 Coarsening of the spatiotemporal data

The Markov chain portion of our method requires a coarsened view of the data. For $t \geq 0$ during a point path ω, let $C(\cdot)$ be a coarsening that maps $\mathcal{Z}$ to a finite set $\mathcal{C}$, and call $C_t(\omega) = C(Z_t(\omega))$ the "state" of the point. To make the Markovian assumption plausible, we populate the coarsened state space $\mathcal{C}$ with summaries of the full resolution data so that transitions between these states represent meaningful events in a tennis point (Figure 1).

The coarsened state space is given by $\mathcal{C} = \mathcal{C}_{\text{ServerShot}} \cup \mathcal{C}_{\text{ReceiverShot}} \cup \mathcal{C}_{\text{end}}$. This makes it clear that we are breaking down tennis points shot-by-shot. In order to distinguish between states that continue the point and those that end it, we define $\mathcal{C}_{\text{Shot}} = \mathcal{C}_{\text{ServerShot}} \cup \mathcal{C}_{\text{ReceiverShot}}$. We formally define the "duration" of a shot as the interval of time that begins with the striker striking the tennis ball with his racket and ends with either the returner striking the ball or with the point ending in a state of $\mathcal{C}_{\text{end}}$.

A tennis point can end in three main ways: The bounce of a shot lands out-of-bounds, the shot goes directly into the net and lands on the striker's side of the court, or the ball is not returned by the returner and the ball bounces twice on the returner's side of the court. We note that there are additional ways for a point to end. For example, points can end if a shot strikes the returning player or if a shot bounces directly on the striker's

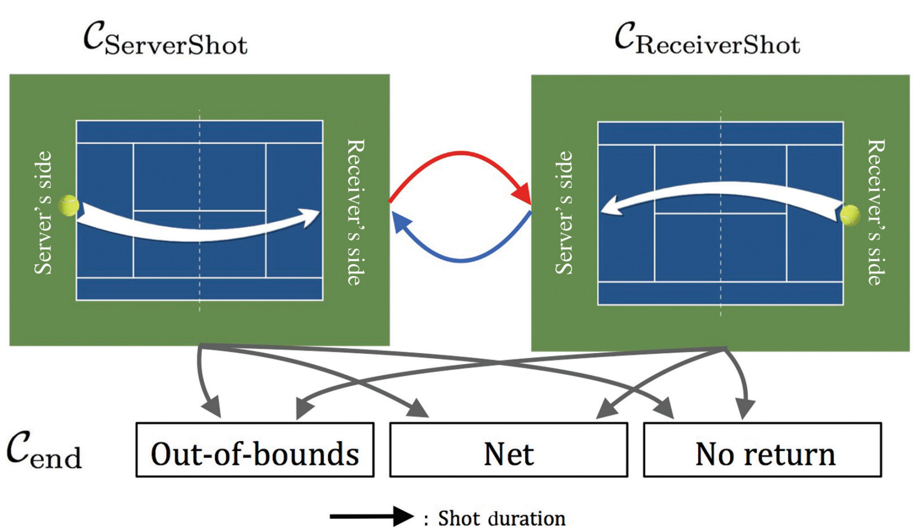

Figure 1: Visual representation of how the three coarsened states $\mathcal{C}_{\text{ServerShot}}$, $\mathcal{C}_{\text{ReceiverShot}}$ and $\mathcal{C}_{\text{end}}$ will interact with each other. The arrows between each state represent a "shot duration."

side of the court without going into the net. These cases will not be analyzed, as these are rare occurrences that will not significantly impact the results. Thus, $\mathcal{C}_{\text{end}} = \{\text{out-of-bounds, net, no return}\}$. Through the perspective of the striker and the returner, these end states will result in different point values. Let $X_S(c)$ and $X_R(c)$ denote the point value of the given end state $c \in \mathcal{C}_{\text{end}}$ through the perspective of the striker and returner, respectively. From the striking perspective, $X_S(\text{out-of-bounds}) = 0$, $X_S(\text{net}) = 0$ and $X_S(\text{no return}) = 1$. From the returning perspective, $X_R(\text{out-of-bounds}) = 1$, $X_R(\text{net}) = 1$ and $X_R(\text{no return}) = 0$.

At time $t \geq 0$ during a given point path ω, the most recent shot is given a *strike state*, $\beta_t^S(\omega)$, and a *return state*, $\beta_t^R(\omega)$. The different strike and return states and their corresponding point values, which make up the space $\mathcal{P}$, will be described in Section 3. Then when a strike by the server is the most recent shot at time t, we assume that $C_t(\omega)$ is equal to the strike state at that time, or $C_t(\omega) = \beta_t^S(\omega)$. The possible values of $C_t(\omega)$, if a strike by the server is the most recent shot at time t, thus live in $\mathcal{C}_{\text{ServerShot}} = \{\mathcal{P}\}$. Similarly, whenever a strike by the receiver is the most recent shot at time t, we again assume that $C_t(\omega)$ is equal to the strike state at that time, or $C_t(\omega) = \beta_t^S(\omega)$. The possible values of $C_t(\omega)$, if a strike by the receiver is the most previous shot at time t, live in $\mathcal{C}_{\text{ReceiverShot}} = \{\mathcal{P}\}$.

2.2 Model assumptions

We must make a few more assumptions about the processes $\mathcal{Z}$ and $\mathcal{C}$, which will allow them to be combined into a functional ESV estimator. First, we need to assume that $\mathcal{C}$ is marginally semi-Markov, since our underlying process of entering different states does not have a geometrically distributed duration, since the interval of time between strikes of the ball is not the same every time. This semi-Markov assumption guarantees that the embedded sequence of disjoint states $C^{(0)}, C^{(1)}, \ldots, C^{(K)}$ is a Markov chain.

In order to specify the relationship between coarsened and full-resolution conditioning, we will define two additional time points which mark changes in the future evolution of the point: Let the notation τ_t denote the first time increment after the beginning the shot duration occurring at time t, and δ_t is the endpoint of this shot duration, leaving the shot duration state into either the opposing player's shot duration state or into an end state, $\mathcal{C}_{\text{end}}$. The formal definitions of τ_t and δ_t can be found in Appendix A.2. We will assume that when a new shot duration state is passed into, this decouples the future of the point after time τ_t with its history up to time t. Let $\rho_{\tau_t}^S(\omega)$ denote the player ID of the striker of a shot occurring at τ_t during point path ω, and let $\rho_{\tau_t}^R(\omega)$ denote the player ID of the returner of a shot occurring at τ_t during point path ω.

Let us also assume for all $s > \tau_t$ and $c \in \mathcal{C}$:

$$\begin{aligned}\mathbb{P}(C_s(\omega) = c \,|\, C_{\tau_t}(\omega), \rho_{\tau_t}^S(\omega), \mathcal{F}_t^{(Z)}) \\ = \mathbb{P}(C_s(\omega) = c \,|\, C_{\tau_t}(\omega), \rho_{\tau_t}^S(\omega)).\end{aligned}$$

This assumption says that for predicting coarsened states beyond some point τ_t, all information in the point history up to time t can be summarized by the state of the point at τ_t, $C_{\tau_t}(\omega)$ – which most times includes the locations of each player at the beginning of the current shot duration–, and by who the striker is at τ_t, $\rho_{\tau_t}^S(\omega)$. The dynamics of tennis make this second assumption reasonable: Every time a player strikes the ball, this represents a structural transition in the tennis point to which both players (mostly the returner) react. Their actions prior to the most previous strike are not likely to influence their actions after it. Put simply, the information at the beginning of the shot duration, $C_{\tau_t}(\omega)$ and $\rho_{\tau_t}^S(\omega)$, will be used to predict the outcome of the shot duration, $C_{\delta_t}(\omega)$.

Using the two assumptions found in this Subsection we are able to simplify ESV from a theoretical quantity and combine aspects from the full-resolution and coarsened views of the process.

Theorem 1. *The full resolution ESV, v_t^i, for player i at time t of a point being played against player j can be rewritten:*

$$v_t^i = \mathbb{E}[\Gamma^i|\, \mathcal{F}_t^{(Z)}] = \sum_{c_v \in \mathcal{C}} \mathbb{E}[\Gamma^i|\, C_{\delta_t} = c_v]\, \mathbb{P}(C_{\delta_t} = c_v|\, \mathcal{F}_t^{(Z)})$$

$$= \begin{cases} \sum_{c_v \in \mathcal{C}} \mathbb{E}[\Gamma^i|\, C_{\delta_t} = c_v]\, \mathbb{P}(C_{\delta_t} = c_v|\, C_{\tau_t} = c_u, \rho_{\tau_t}^S = i), & \text{if } i = \rho_{\tau_t}^S \\ \sum_{c_v \in \mathcal{C}} \mathbb{E}[\Gamma^i|\, C_{\delta_t} = c_v]\, \mathbb{P}(C_{\delta_t} = c_v|\, C_{\tau_t} = c_u, \rho_{\tau_t}^S = j), & \text{if } i = \rho_{\tau_t}^R \end{cases} \quad (2)$$

Since this paper will focus on both players during a point, we will have to look at point values from two different perspectives: The striker's perspective, and the returner's perspective. This is reflected in (1). The proof of Theorem 1 follows directly from assumptions in this Subsection and is therefore omitted. Note that the dependence on ω was also omitted in (2) in order to save space. Heuristically, (2) expresses $v_t^i(\omega)$ as the expectation given by a homogeneous Markov Chain on $\mathcal{C}$ with a random starting point $C_{\delta_t}(\omega)$, where only the starting point depends on the full-resolution information $\mathcal{F}_t^{(Z)}$, which now can be summarized with $C_{\tau_t}(\omega)$ and $\rho^S_{\tau_t}(\omega)$.

2.3 Transition model specification

The representation of the ESV estimator in (2) shows that estimating ESV does not require a full-blown model for the entire tennis point at high resolution. Instead, the priority is to accurately predict how much the next state, denoted by $C_{\delta_t}(\omega)$, will contribute to a won point for both the striker and the returner. Let us define player i's *Markov transition probability matrix*, Kemeny and Snell (1976), for when he is the striker as $\mathbf{P}^i$, with

$$P^i_{uv} = \mathbb{P}(C^{(n+1)} = c_v | C^{(n)} = c_u, \rho^S = i) \tag{3}$$

for any $c_u,\ c_v \in \mathcal{C}$, and where ρ^S indicates the striker's player ID. Without any other probabilistic structure assumed for $\mathbb{P}(C^{(n+1)} = c_v | C^{(n)} = c_u, \rho^S = i)$ other than Markov, for all u, v, the maximum likelihood estimator of P^i_{uv} is the observed transition frequency,

$$P^i_{uv} = \frac{N^i_{uv}}{\sum_{v'} N^i_{uv'}}$$

where N^i_{uv} counts the number of transitions from c_u to c_v when player i is the striker. Intuitively, this states that the striker ultimately decides what the next state, c_v, will be based on the current state, c_u, whether that be the end of the point or the next strike state. Also, note that the results from this estimator will be undesirable if the number of visits to a particular state c_u is small. When it comes time to incorporate these probabilities in the ESV estimator, if the striker does not have any history of striking from a certain state, then we will use average probabilities of striking from that state amongst all of the players in the dataset.

3 Region system and location categorizations

There will be two distinct datasets used in this paper:

- **Dataset 1**: A five-match sample from the 2015 US Open, with all possible information available including bounce locations, shot speed and shot spin. Thus, any exact (x, y) coordinate, or any reference to speed or spin in this paper, will be from the five-match sample. This dataset, though only consisting of five matches, is still a significant amount of tennis player-tracking data, especially since the data is difficult to obtain.
- **Dataset 2**: A dataset with all available data from the 2015 US Open. The USTA was gracious enough to allow the ESV estimator and its related methodology to be run on the entirety of its 2015 US Open data, in order to give the ESV's results a dependable sample size. However, this dataset will not contain any specific information of the matches, such as coordinates and shot speed, and provides no way of linking to the five-match sample dataset. This dataset will mostly be used to validate the ESV estimator.

For any calculation in this paper, it will be specified which dataset was used. Also, the names of the players we are analyzing will not be used in this paper, in order to protect their privacy. Thus, there will be a focus on the mathematics and the potential of the ESV estimator and its methodology, rather than the performance of specific, well-known tennis players. Finally, speed and spin are not included in the current analysis. While these variables are almost certainly play a significant role in the outcome of a shot, they are withheld here to ensure that there are enough points in each bin. They will hopefully be included in future work with a larger data set.

3.1 Region system

In order to properly locate both players on the court for our model, we first define a set of regions for the tennis court. The court's in-bounds x-coordinates from baseline to baseline range from $x = -11.9135$ m to $x = 11.9135$ m (Figure 2). The center of the baselines are located at $x = \pm 11.887$ m, but if a ball touches any part of the line, it is considered inbounds. Therefore, we must take into account the width of the lines on the court, which are 0.053 m thick. This will be applied to the in-bounds sidelines on the court, as well. Only men's singles matches will be analyzed, thus the "doubles alleys" will not be in play, which are the long rectangular portions of the court whose y-coordinates range from $y \in [4.115,\ 5.487]$ m and $y \in [-5.487,\ -4.115]$ m. Therefore, the in-bounds y-coordinates from sideline to sideline range from $y = -4.1415$ m to $y = 4.1415$ m. Any shot bounce landing outside of these in-bounds x and y values will be deemed an out-of-bounds shot. This does not mean a player cannot go out-of-bounds to return a shot, however, so the region system must extend outside of the in-bounds section of the court. In order for the system to be as useful as possible, the regions should divide the court into meaningful subsets that represent normal play. A visualization of where all 2404 ground-strokes (non-serves) in the five-match sample dataset were struck

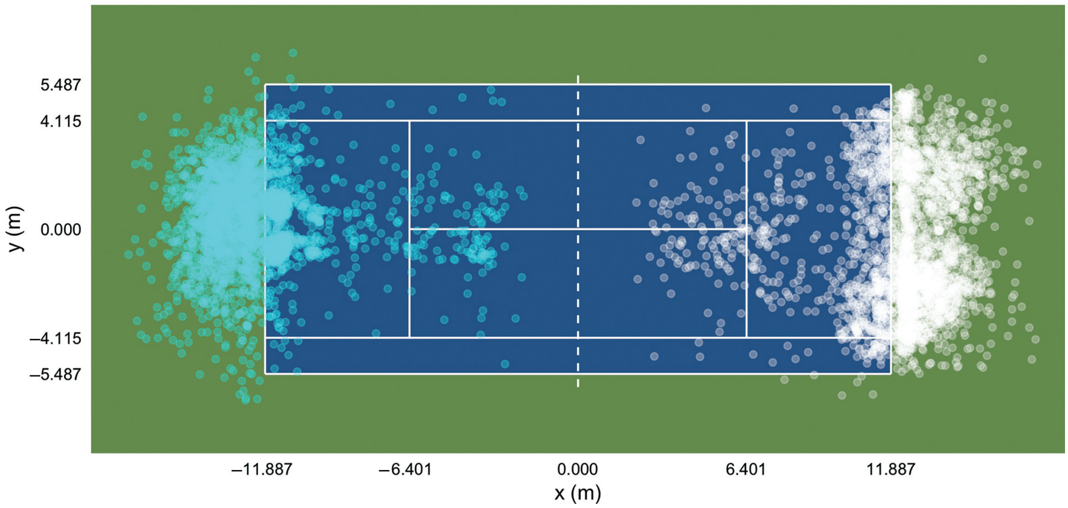

Figure 2: The layout of a tennis court with the locations of all strikes of ground-strokes in Dataset 1, represented by small white, transparent circles on the positive *x* side of the court. The blue areas indicate the in-bounds areas of the court, with the exception of the "double's alleys." The green areas represent the out-of-bounds areas of the court. The dashed white line at $x = 0$ m indicates the net. Locations of returners at the time of strikes of ground-strokes, represented by small blue, transparent circles on the negative *x* side of the court. Areas with a more solid color indicate a higher concentration of strike/return locations.

from – and where the returners were located at the time of those strikes – show that there are a high number of ground-strokes struck from the baseline area and, at the same time, a high number of returners who are setting up near the baseline area (Figure 2). Note that we did not map the locations of strikes which were returns of serves, as those will be treated separately.

Based on the distribution of the striker and returner locations, we define a region system with 15 cells on each side of the court and which has increased resolution near the baseline (Figure 3). The area around the baseline at $x = \pm 11.887$ m is treated differently from the rest of the region system, because of the high density of player locations around the baseline area: The inner 50% of the inbounds area will be encompassed by region 10, and the outer 25% portions of the inbounds area, extending to the out-of-bounds area, will be encompassed with regions 9 and 11. Regions 9 through 11 are centered on

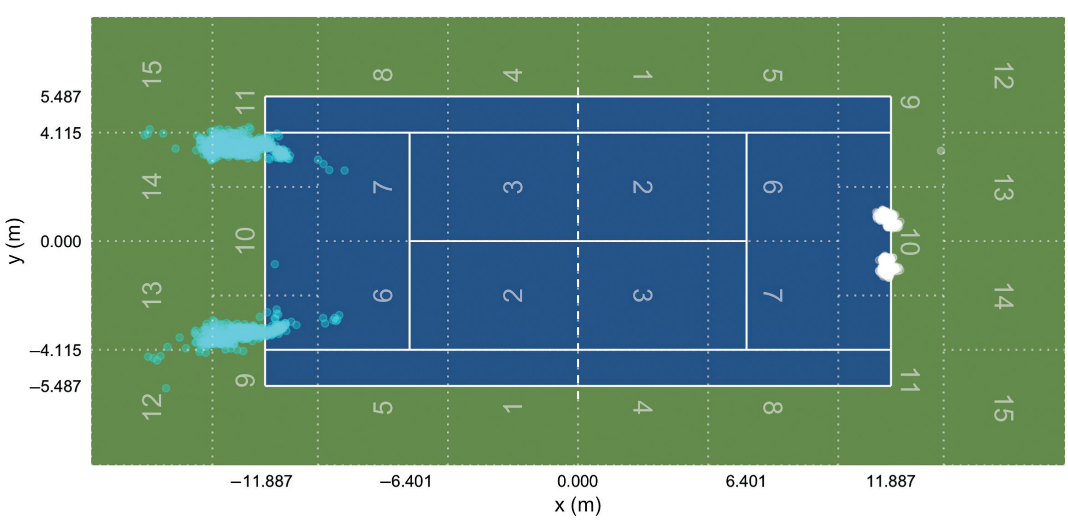

Figure 3: The region system which will be used for analysis, with each region given a number 1–15, along with the locations of strikes of first and second serves are displayed on the positive *x* side of the court, showing high concentrations in region 10. Locations of the returners at the time of first or second serve strikes are displayed on the negative *x* side of the court, showing high concentrations in regions 9 and 11.

the baseline and include two meters on either side: so $x \in [9.887, 13.887)$ m and $x \in (-13.887, -9.887]$ m. Continuing on, regions 5 through 8 represent player locations from an intermediate area of the court that is not quite at the baseline, but it is not close to the net, either. These regions' x-coordinates range from the end of the baseline regions to halfway to the net. That is, the x-coordinates for regions 5 through 8 range from $x \in [4.9435, 9.887)$ m and $x \in (-9.887, -4.9435]$ m.

Regions 1 through 4 encompass the area close to the net, ranging from $x \in (0, 4.9435)$ m and $x \in (-4.9435, 0)$ m. The region system also takes care of any player locations that would be considered the deep baseline area, as regions 12 through 15 encompass any x-coordinates where $x \geq 13.887$ m or $x \leq -13.887$ m. It should be noted that the the regions on the outside with the largest y magnitudes range from the middle of the sideline at $y \leq -4.115$ m and $y \geq 4.115$ m. This is to ensure any strikes or returns with large y magnitudes, while rare, are considered.

3.2 Location categorization

The driving belief behind this research is that the most important factor behind where a striker chooses to hit the ball next is where both the striker and the returner are located at the time of the strike. This region system will assist in grouping together similar shots, based on where the striker and returner are located. For example, if the striker strikes the ball from region 6 and, at the same moment the returner is located in region 10, the shot will be categorized as 6–10.

We will name this region combination of where the striker and returner are located as the *player location combination*. These player location combinations will be used to help categorize locations of players in order to assess their striking and return abilities. Since there are 15 distinct regions on either side of the court, there are 15^2 (or 225) possible player location combinations, which will be utilized to describe the players' locations. All possible *location categorizations* for a ground-stroke (GS) reside in the set

$$\mathcal{P}^{\mathrm{GS}} = \{a - b : a = \{1, 2, \ldots, 15\},\\ b = \{1, 2, \ldots, 15\}\}. \quad (4)$$

Most times, $\beta_t^S(\omega)$ and $\beta_t^R(\omega)$ will be the same, as they will simply be the player location combination at time t of point path ω, but there are special instances where this will not be the case, as explained in the Section 3.3.

Table 1: The most and least frequent ground-stroke location categorizations in Dataset 2, out of 38,025 total ground-strokes.

	Most frequent	Occurrences	Least frequent	Occurrences
1	10–10	4993	10–1	1
2	10–14	2757	3–1	1
3	14–10	2752	1–14	1
4	10–13	2634	5–15	1
5	13–10	2584	6–4	1

3.3 Location categorization of serve-related shots

One of the goals of this paper is to assess both a player's strike ability and return ability from each unique location categorization, in order to properly evaluate a player's skill set. Those familiar with the game of tennis know that ground-strokes and serves are very different shots. Every point begins with a serve, and, often times, it sets the tone for the rest of the point. On a serve, the server and the returner line up in approximately the same locations every time: The server must line up behind the baseline, on the opposite half of the court to the serve box he is attempting to land his serve in, and the returner usually lines up along the baseline behind the corresponding serve box. Thus, for a majority of first and second serve strikes, the player location combination will be either 10–9 or 10–11. This lack of variety in player locations during serves can be seen looking at all 1272 first or second serve strikes in Dataset 1 (Figure 3).

To lump serves and serves' returns in with 10–9 and 10–11 ground-strokes is unfair, as serves and serves' returns are much different than the average ground-stroke. Because of this we will treat the location categorizations of both serves and the returns of serves differently than ground-strokes.

In general, we assume a shot is being struck at time $t \geq 0$ of a specific point path ω. When a first serve is struck, $\beta_t^S(\omega)$ will be set to 1S (First Serve) and $\beta_t^R(\omega)$ will be set to 1SR (First Serve Return). For a second serve, $\beta_t^S(\omega)$ will be set to 2S (Second Serve) and $\beta_t^R(\omega)$ will be set to 2SR (Second Serve Return). When the return shot of a first serve is struck, the $\beta_t^S(\omega)$ will be set to 1SR, and $\beta_t^R(\omega)$ will be set to the player location combination at that time. Similarly, when the return of a second serve is struck, $\beta_t^S(\omega)$ will be set to 2SR, and $\beta_t^R(\omega)$ will be set to the player location combination at that time. Thus, for a strike or return which is serve-related (S-R), the possible location categorization's reside in the set

$$\mathcal{P}^{\text{S-R}} = \{1\text{S}, 2\text{S}, 1\text{SR}, 2\text{SR}\}.$$

All possible location categorization's will reside in the union, $\mathcal{P}$, of $\mathcal{P}^{\text{S-R}}$ and $\mathcal{P}^{\text{GS}}$:

$$\mathcal{P} = \mathcal{P}^{\text{S-R}} \cup \mathcal{P}^{\text{GS}}.$$

4 Strike and return categories

Through the perspective of a tennis player, every point either ends in a won point, given a point value of 1, or a lost point, given a point value of 0. There are shots in tennis which lead directly to a won point – such as pure winners – and shots which lead directly to a lost point – such as shots that land out-of-bounds or go directly into the net. Shots that lead directly to a point value of 0 or 1 – which we will refer to as *direct shots* – only make up approximately 29% of the total strikes in Dataset 2. In order to properly identify a given player's ability from different areas of the court, we need to take into account more than just these direct shots and look at *indirect shots*. Indirect shots are those that do not directly lead to the outcome of a point, but they contribute in some way to a won or lost point. These indirect shots will be given point values in-between 0 and 1. In this section, we will attempt to quantify just how much these indirect shots contribute to a won point, using what we will call *strike categories* and *return categories*.

4.1 Strike categories

All strikes during a point will fall into one of the eight categories described in Table 2. There are seven specific categories that cover most of the strikes that potentially contribute to the outcome of the point and one neutral category (λ_{NI}) that contains the remaining shots. It would be naive to claim this is a perfect approach. There may be further subdivisions possible and some strikes may be categorized as contributing in some way to a won or lost point when in reality they should be categorized as λ_{NI}. This approach to finding the most important strikes of a given point has the advantage of being simple and provides a jumping-off point for future approaches.

Table 2: The weights (w_{SC}) for each strike category (λ_{SC}), along with short definitions, and whether or not the category is considered a direct shot.

λ_{SC}	w_{SC}	Direct	Definition
λ_{PW}	1.00	Y	In-bounds strike which is not returned
λ_{SUPW}	0.75	N	Strike by winning player prior to λ_{PW}
λ_{FOB}	0.75	N	Strike by winning player prior to λ_{OB}
λ_{FN}	0.75	N	Strike by winning player prior to λ_{N}
λ_{NI}	0.50	N	Strike which does not fall into other categorizations
λ_{SUOPW}	0.25	N	Strike by losing player prior to λ_{PW}
λ_{OB}	0.00	Y	Strike which lands out-of-bounds
λ_{N}	0.00	Y	Strike which goes into net

Let all of the strike categories reside in the set Λ_{SC}:

$$\Lambda_{\text{SC}} = \{\lambda_{\text{PW}}, \lambda_{\text{SUPW}}, \lambda_{\text{SUOPW}}, \lambda_{\text{OB}}, \lambda_{\text{N}}, \lambda_{\text{FOB}}, \lambda_{\text{FN}}, \lambda_{\text{NI}}\}.$$

Some of these categories will contribute more to a winning point than others and the weights (w_{SC}) used in this paper for each strike category (λ_{SC}) are given in Table 2. Theoretically, these should correspond to how much each strike contributes to a winning point. The neutral w_{NI} (Non-Impactful) weight was set to 0.50 so that anything greater than 0.50 indicates a player tends to hits shot that lead to won points and anything below indicates a player tends to hits shot that lead to lost points. The weights w_{PW} (Pure Winner), w_{OB} (Out-of-Bounds) and w_{N} (Net) weights are self-explanatory since they are direct shots. Also, it should be noted there is one specific strike state which will never be penalized with w_{OB} or w_{N} weights: When the server is striking a first serve, he cannot directly lose the point on that specific strike. If he strikes the ball into the net or outside of the serve box he is aiming for, then the first serve is considered a "fault" and he continues onto the second serve. These types of points will be removed from any analysis done on players' striking ability.

The weight w_{SUPW} (Set-Up of Pure Winner) is 0.75 since it is a set-up to a λ_{PW}, which has a weight of 1.00, so the mean of 0.50 and 1.00 was chosen. The weight w_{SUOPW} (Set-Up of Opponent's Pure Winner) is 0.25 since it is not fully contributing to a lost point for the striker, but it is giving the returner the ability to strike a λ_{PW} the next shot, so the mean of 0.00 and 0.50 was chosen. The w_{FOB} (Forced Out-of-Bounds) and w_{FN} (Forced Net) weights were somewhat tougher to decide upon, but they both clearly deserved weights above 0.50. Ultimately, 0.75 was decided upon since if it is truly the belief that these strikes help cause the opponent to hit an λ_{OB} or λ_{N} strike, then the λ_{FOB} and λ_{FN} strikes deserve as much weight as a λ_{SUPW} strike.

Now, these categorizations of strikes may be somewhat counterintuitive. For example, if player i strikes a λ_{FOB}, and player j then strikes an λ_{OB}, does player j's strike truly deserve a 0.00 weight if player i hit a great strike which caused player j to hit an errant strike? There will be a lot of gray area and room for interpretation when it comes to sorting strikes into categories, but unless we were to watch each individual point and by-hand figure out which shots contribute more to a winning point than others, it is difficult to start sorting strikes into different categories and to weight strikes differently. It might be possible to better sort strikes by considering other available information such as speed and spin and this would be a good future project on a larger data set.

4.2 Return categories

A player's ability to return strikes is also a crucial skill as half of the game of tennis is a player putting himself into position to be able to return the ball and not letting the other player dictate how the point will be played. All strikes with the ability to be returned during a point will be categorized into one of the seven return categories described in Table 3.

"Strikes with the ability to be returned" includes all strikes that do not land out-of-bounds or go directly into net. Strikes that do not land in-bounds prevent the returner from having the opportunity to return them. All strikes like this as are labeled as "Insignificant" and were not included in any analysis of a player's return ability. Also, since they will have the same weights, returns that led to a λ_{OB} or λ_{N} strikes have been grouped together in the λ_{RLS} (Returned Losing Strike) return category, as have returns that led to λ_{FOB} or λ_{FN} strikes in the λ_{RFLS} (Returned Forced Losing Strike) return category. A λ_{NR} (No Return) return is when the returner is unable to get his racket on the ball, resulting in a pure winner for the opponent.

Let all of the return categories reside in the set Λ_{RC}. That is,

$$\Lambda_{\mathrm{RC}} = \{\lambda_{\mathrm{RPW}}, \lambda_{\mathrm{RSUPW}}, \lambda_{\mathrm{RSUOPW}}, \lambda_{\mathrm{RLS}}, \lambda_{\mathrm{RFLS}}, \lambda_{\mathrm{NR}}, \lambda_{\mathrm{R}}\}.$$

Table 3 contains weights (w_{RC}) for each return category (λ_{RC}) corresponding to how much each return category theoretically contributes to a winning point. Since, the λ_{NR} is a return where a player is unable to get his racket on the ball and thus a won point for the striker, $w_{\mathrm{NR}} = 0.00$. It was the intent of this research to keep the weights of the strike categories and the return categories symmetric,

Table 3: The weights (w_{RC}) for each return category (λ_{RC}), along with short definitions, and whether the category is deemed a direct shot.

λ_{RC}	w_{RC}	Direct	Definition
λ_{RPW}	1.00	Y	Returned a λ_{PW}
λ_{RSUPW}	0.75	N	Returned a λ_{SUPW}
λ_{RFLS}	0.75	N	Returned a λ_{FOB} or λ_{FN}
λ_{R}	0.50	N	Returned a λ_{NI}
$\lambda_{\mathrm{RSUOPW}}$	0.25	N	Returned a λ_{SUOPW}
λ_{RLS}	0.00	Y	Returned a λ_{OB} or λ_{N}
λ_{NR}	0.00	Y	Unable to return

therefore, the remaining return category weights, w_{RC}, correspond to the same strike category weights, which are explained in Section 4.1.

5 Calculation of ESV

Since we are incorporating transition probability matrices to predict how much the outcome state, or the *next* state, will contribute to a won point, we must consider all possibilities. Say player i is striking the ball from a given strike state. If that strike by player i does not result in the end of the point – which will give player i a defined point value of either 1 or 0 – then the point continues and we quantify how much the next potential return state will contribute to a won point for player i. Conversely, during a match between players i and j, if player i is attempting to return player j's strike and that strike does not end the point, then we must look ahead and quantify how much the next potential strike state will contribute to a won point for player i. To do this we will create two metrics: *Strike value* and *return value*.

5.1 Strike value and return value

The *strike value* (SV) for player k from strike state β^S will be defined as follows:

$$\mathrm{SV}(k, \beta^S) = \sum_{\lambda_{\mathrm{SC}} \in \Lambda_{\mathrm{SC}}} w_{\lambda_{\mathrm{SC}}} \frac{|S^k_{\beta^S, \lambda_{\mathrm{SC}}}|}{|S^k_{\beta^S}|}, \tag{5}$$

where $|x|$ is the cardinality of set x, counting the number of elements it contains. The *return value* (RV) for player k from return state β^R will be defined as follows:

$$\mathrm{RV}(k, \beta^R) = \sum_{\lambda_{\mathrm{RC}} \in \Lambda_{\mathrm{RC}}} w_{\lambda_{\mathrm{RC}}} \frac{|R^k_{\beta^R, \lambda_{\mathrm{RC}}}|}{|R^k_{\beta^R}|}. \tag{6}$$

Strike value and return value attempt to quantify how many points a specific player can expect based on how he has struck/returned from specific strike/return state in the past - whether that be a player location combination or a serve-related state. A strike/return value of 0.50 indicates a strike/return state that usually produces neutral strikes/returns which do not necessarily contribute to a winning point or a losing point. A strike/return value of 1.00 indicates a strike/return state that produces all pure winner strikes/returns, and a strike/return value of 0.00 indicates a strike/return state that produces all strikes/returns which directly lead to losing points. Definitions for strike/return subsets referenced in (5) and (6) can be found in Appendix A.3.

Similar to how we handle missing transition probabilities, if a player does not have a SV from a certain strike state or if he does not have a RV from a certain return state, then we will use the average SV/RV from that strike/return state amongst all the players in our dataset.

5.2 Defining the calculation of ESV

Recall our rewritten definition of the ESV estimator (2). First, let us look at the $\mathbb{P}(C_{\delta_t}(\omega) = c_v | \mathcal{F}_t^{(Z)})$ portion of our ESV estimator. To calculate this value, we create transition probability matrices $\mathbf{P}^i$ and $\mathbf{P}^j$ for a match between player i and player j. Then if we are attempting to calculate player i's ESV when he was the striker of the shot in question, the probability $\mathbb{P}(C_{\delta_t}(\omega) = c_v | \mathcal{F}_t^{(Z)})$ will be equivalent to P^i_{uv}. Conversely, if we are attempting to calculate player i's ESV when he was the returner of the shot in question, then the probability $\mathbb{P}(C_{\delta_t}(\omega) = c_v | \mathcal{F}_t^{(Z)})$ will be equivalent to P^j_{uv}. From a tennis perspective this makes sense, since the striker ultimately controls where the ball travels next, dictating the next player location combination or if the point ends as a result of this strike.

Now, we must look at the $\mathbb{E}[\Gamma^i(\omega) | C_{\delta_t}(\omega) = c_v]$ portion of our ESV estimator (2). As was discussed in Section 5.1, we must take into account not only the possibility that the point will end on a given strike, but also the possibility that the point continues. Thus, we will need to break down $\mathbb{E}[\Gamma^i(\omega) | C_{\delta_t}(\omega) = c_v]$ into two parts: One part which takes into account the potential end of the point, and the other which takes into account the potential next state and the continuance of the point. From player i's striking perspective, we will incorporate the return value metric, $\text{RV}(i,\beta^R \in \mathcal{P})$, and the point outcome from end states, $X_S(c \in \mathcal{C}_{\text{end}})$. From player i's returning perspective we will incorporate the strike value metric, $\text{SV}(i,\beta^S \in \mathcal{P})$, and the point outcome from end states, $X_R(c \in \mathcal{C}_{\text{end}})$. Thus, for a match being played between players i and j, the ESV for player i and point path ω will be calculated as:

$$
\begin{aligned}
v_t^i(\omega) &= \sum_{c_v \in \mathcal{C}} \mathbb{E}[\Gamma^i(\omega) | C_{\delta_t}(\omega) = c_v] \\
&\qquad \mathbb{P}(C_{\delta_t}(\omega) = c_v | C_{\tau_t}(\omega) = c_u, \mathcal{F}_t^{(Z)}) \\
&= \begin{cases} \displaystyle\sum_{c_w \in \mathcal{C}_{\text{Shot}}} \text{RV}(i, c_w) P^i_{uw} + \sum_{c_x \in \mathcal{C}_{\text{end}}} X_S(c_x) P^i_{ux}, \\ \hfill \text{if } i = \rho^S_{\tau_t}(\omega) \\ \displaystyle\sum_{c_w \in \mathcal{C}_{\text{Shot}}} \text{SV}(i, c_w) P^j_{uw} + \sum_{c_x \in \mathcal{C}_{\text{end}}} X_R(c_x) P^j_{ux}, \\ \hfill \text{if } i = \rho^R_{\tau_t}(\omega) \end{cases}
\end{aligned} \tag{7}
$$

6 Results

In Dataset 2, there are 70 distinct matches, and there are 10 matches where both players have at least four matches-worth of data. Contained in these ten matches were 2,020 points, 8,491 shots and 16,982 possible ESV values. To validate the ESV estimator, the points in these matches were analyzed.

6.1 ESV validation

A leave-one-out cross validation was implemented to these matches, fitting the transition probability matrices and the SV, RV metrics to every point in Dataset 2 except for the one in question. Below in Figure 4 shows the win rate versus the binned ESV value.

We note that points which had multiple serves were divided into two, and those sub-points which ended in a first-serve fault were not included since no player wins those sub-points. The best-fit line, produced using linear regression, outputs an R-squared value of 0.8438, which shows a linear trend between ESV and win rate. Generally, as the ESV values increase, the win rates increase, as well. In the future, the weightings used in the strike value and return value metrics can be tweaked to more appropriate settings, in order to increase the R-squared value.

It needs to be stated that ESV is not primarily a win probability, and is intended to be a measure of players' strengths and weaknesses. In order to validate ESV, it is logical that higher ESV values should result in higher winning rates and lower ESV values should result in lower winning rates – hence why we chose to validate ESV with win rate in this section. For now, this linear trend is reassuring and provides support that the ESV estimator is producing reliable results.

6.2 ESV applied to an individual point

The ESV estimator can also be used to visualize the evolution of an individual point (Figure 5). The point contained in both Dataset 1 and Dataset 2, so we have the ability to analyze player decision making and positioning afterwards. It should be noted that player 2 was the eventual winner of the point shown below.

When player 2 attempts to strike a second-serve return, player 1 achieves his highest ESV of the point at 0.5750. After the second-serve return, a series of shots occur next, with both players located near the baseline area, and the corresponding ESV's hovering around the

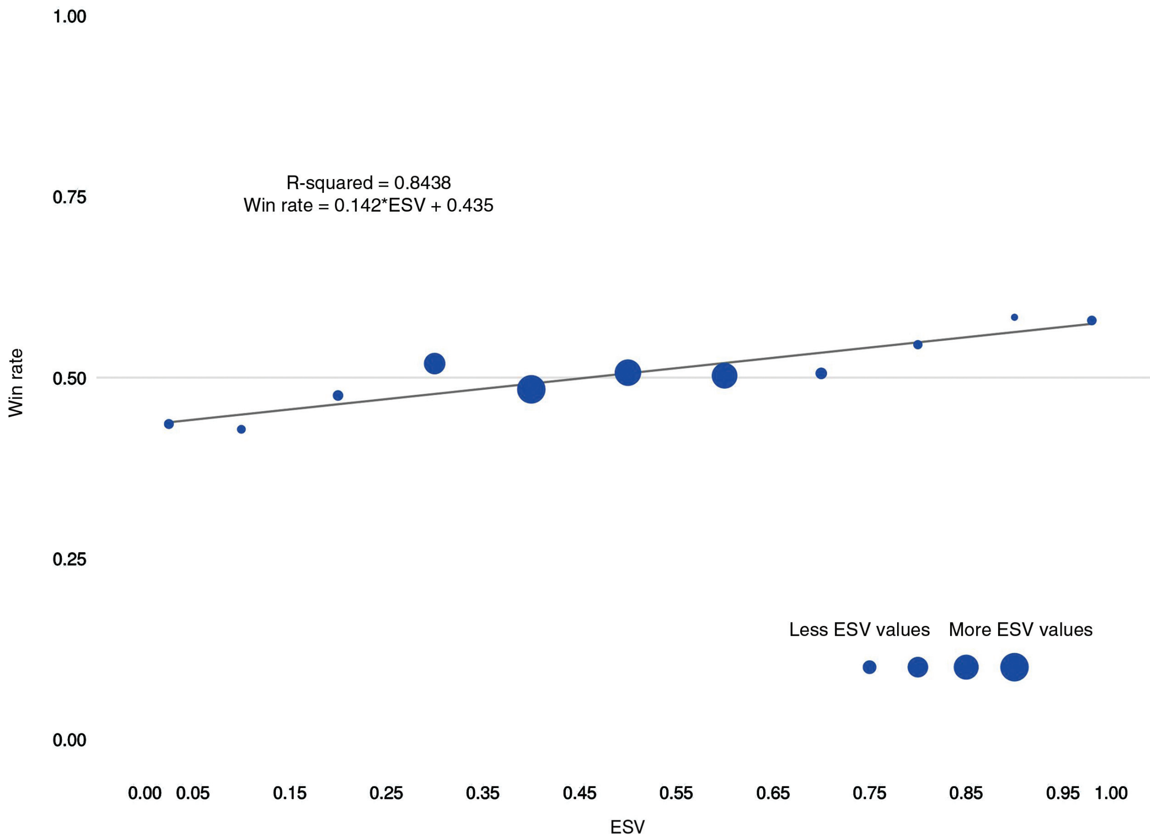

Figure 4: Win rate versus ESV value, for the calculated ESV values in the match involving the main two players in Dataset 2. The R-squared value of the best-fit line through the data points is displayed in the top left of the plot, along with equation of the best-fit line.

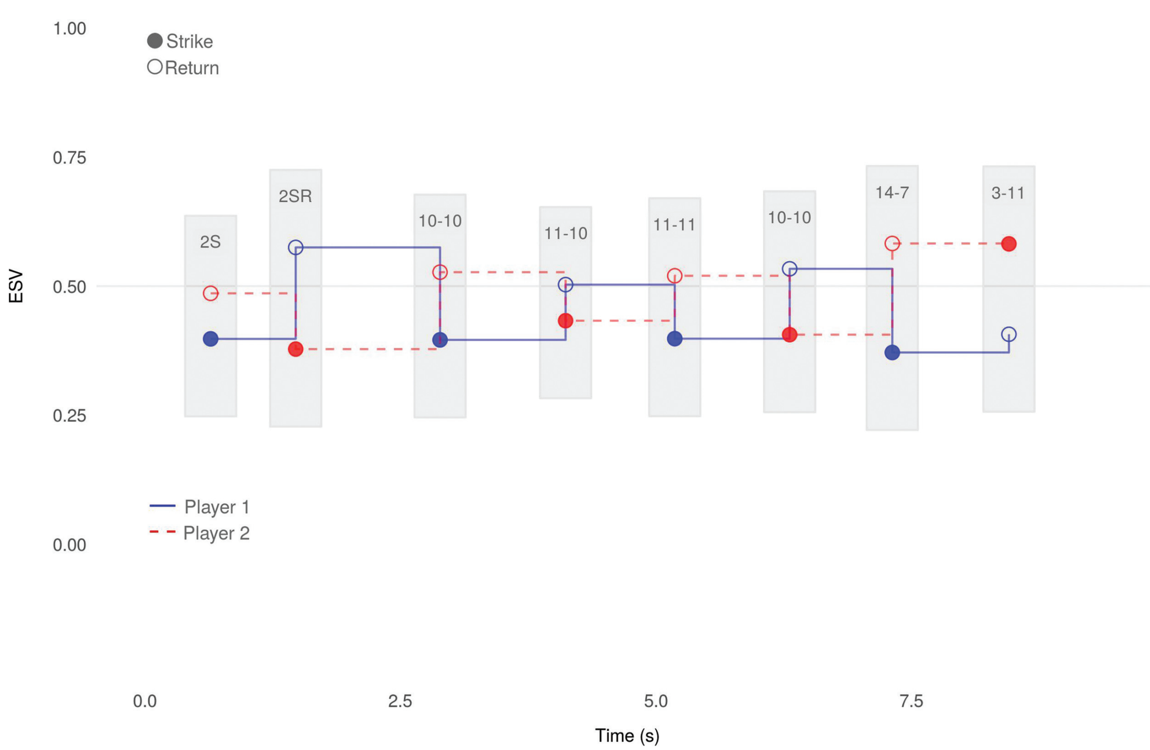

Figure 5: Shot-by-shot visualization of the ESV estimator applied to a point played between player 1 and player 2, based on data within Dataset 2. The labels above each of the points in the graph describe the corresponding strike state of each shot.

neutral 0.50. For most of the shots during this point, the returner has the advantage over the striker. On the third-to-last strike of the point, player 2 strikes an unexpected great strike from the strike state 10–10 (both players located at the middle of the baseline). This strike forced player 1 to back up and strike from region 14, and, at the same time, player 2 moved up towards the net to region 7. This results in a large difference player 1's ESV (0.5825) and player 2's ESV (0.3715). Player 1 ends up striking a weak shot, which player 2 is able to volley and strike a pure winner from the 3–11 strike state. Because of this advantageous strike state, the final shot of the point was the only one where the striker had a greater ESV (0.5818) than the returner (0.4069). Thus, player 2 forcing player 1 to strike from the strike state 14–7 effectively swung the point in player 2's favor – at least through the lens of the ESV estimator.

The approach taken in this paper to evaluating the evolution of individual tennis points was to go shot-to-shot and look at each shot individually, taking into account both direct shots and indirect shots. This specific point was included because it supports this approach, as the outcome of the point was swung on a single shot near the middle of the point. The five ground-strokes following the serve were all shots where both players were near the baseline, resulting in a lack of variation in the ESV since these are common types of shots. But, when player 2 is able to force player 1 to strike from the 14–7 strike state and force variation to where both players were located, player 2 is able to take control of the point.

Though the derivation of ESV may be somewhat complex, the calculation itself is not too involved, as transition probability matrices, SV and RV calculations mostly involve calculating sample averages. Although we acknowledge this could be problematic for small sample sizes, we argue the simplicity of our method should be considered a strength. Stepping through the calculation of player 2's ESV (0.5825) from the return state of 14–7 with player 1 as the striker:

1. Player 1 has struck from 14–7 previously 9 times in Dataset 2.
   - 22% of the strikes ended up as direct losing shots.
   - 11% of the strikes ended up as pure winners.
   - 67% of the strikes continued the point (once to each of the strike states 2–14, 6–11, 7–11, 7–14, 7–15 and 14–14).
2. Since we must consider the possibility the point continues, and player 2 will be the striker of the next shot, we calculate player 2's SV from each of the states 2–14 (SV of 0.750), 6–11 (0.438), 7–11 (0.433), 7–14 (0.550), 7–15 (0.595) and 14–14 (0.477). Thus, the point value for player 2 if the point continues is 0.540.
3. The calculation of ESV now comes down to weighting each of the probabilities with their respective point values (through player 2's perspective) and adding them together.
   - $0.22(1) + 0.11(0) + 0.67(0.540) \approx 0.5825$.

6.3 Guiding player strategy

In order to demonstrate how ESV and its methodology can guide player strategy, let us look at one of the players' decisions from the point analyzed in Section 6.2. Specifically, after player 2 struck from the 10–10 strike state at around 6.5 seconds, did he pick the best region to set up and return player 1's next strike from region 14? What is the worst region he could have set up in?

Assume that player 1 will still strike the ball from region 14. Player 2 traveled a distance of 3.04 m from when he struck from the 10 region to when he set up to return player 1's strike in region 7. This means that in the same distance or less, player 2 could have instead traveled to regions 6, 11, 14 or, of course, stayed put in region 12. Let us look specifically at the potential ESV's if the corresponding strike/return states were one of these other possibilities.

It should be noted that the ESV's of player 1 and player 2 are independent of each other, as ESV is more of a measure of the corresponding player's strengths and weaknesses than it is a win probability – so, it should not be expected that the ESV's of player 1 and player 2 add to one.

According to the potential ESV's (Table 4), player 2 actually picked the second-best area – region 7 – to set up to return player 1's strike. Player 2 could have traveled to virtually any of the other regions and still held the advantage over player 1 striking from region 14. However, if player 2 had chosen to set up on the other side of the court in region 6, he would have held a much bigger advantage in terms of difference in ESV. In Figure 6 we visualize player 2's potential ESV's if player 1 was striking from region 14. Then, we compare player 2 to the average player in Dataset 2 (Figure 7).

Table 4: Potential ESV's for both player 1 and player 2. The considered strike/return states consist of the striker residing in region 14, and the returner residing in several different regions.

β^S/β^R	P1 ESV	P2 ESV
14–6	0.1786	0.7723
14–7	0.3716	0.5825
14–10	0.4021	0.5044
14–11	0.3842	0.5164
14–14	0.3188	0.5558

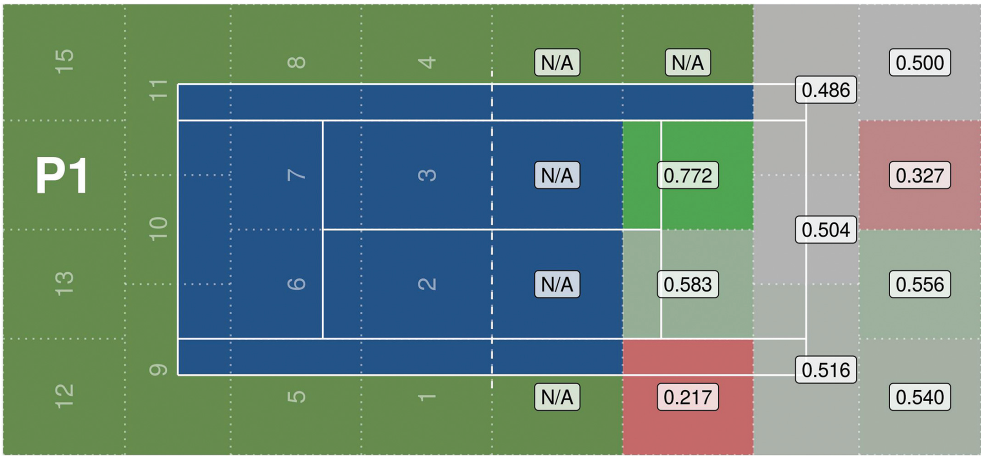

Figure 6: Potential ESV values for player 2 if player 1 was striking from region 14 and player 2 was setting up to return in any of the 15 regions on the opposite side of the court. Any region with "N/A" as its ESV value indicate that player 1 did not have sufficient prior history striking from the corresponding PLC. For example, player 1 did not strike from the 14–1 state in Dataset 2.

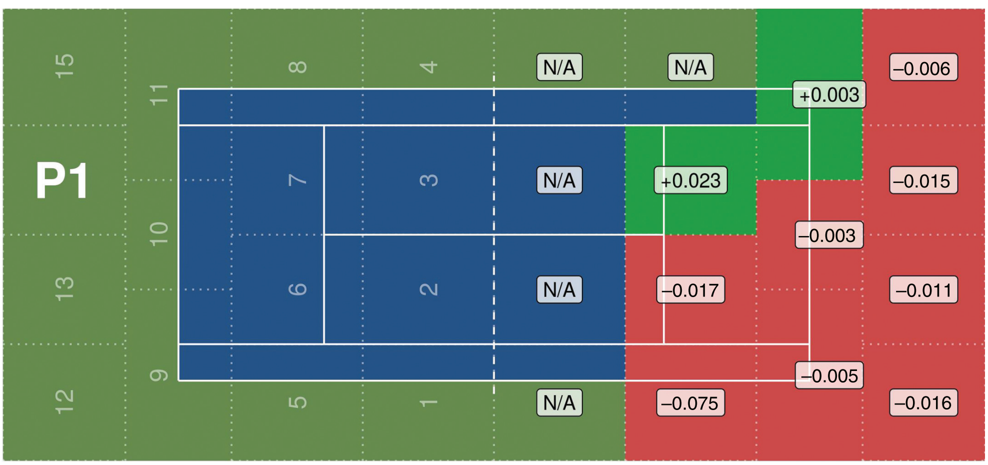

Figure 7: Player 2's ESV compared to the average player's in Dataset 2. For each potential PLC with player 1 striking from region 14, we calculated the ESV's for every player in Dataset 2 outside of player 1 and player 2 and took the mean of those values to define the average ESV from each PLC.

Compared to the average player, player 2 is better at returning strikes from region 14 on the right side of the court (above average ESV's from regions 6 and 9, and only slightly below average from region 12) than the left. This idea of comparing to the average player can be another way for players and coaches to identify strengths and weaknesses of a player's skill set.

7 Conclusions

In Section 6, the potential of the ESV estimator and its methodology was shown. Section 6.1 provided justification for the ESV estimator, as we showed a linear relationship between ESV and the win rate of the players who obtained them. Section 6.2 showed the ESV estimator applied to an individual point, and demonstrated how ESV can be used to describe the evolution of a point – *telling the story* of a point. The ESV estimator provides potential to give a new, shot-by-shot perspective to players and fans as to why and how a point or match was won. There is also potential for ESV to influence player strategy, as it is able to identify strong/weak strike or return states of specific players.

To be clear, we are not claiming to unearth any new findings of the game of tennis with this paper. ESV represents a new way to think about quantifying the skill of players from different areas of the court. As anyone who watches tennis knows, a majority of shots: Are taken when both players are located near the baseline area, are simply kept in play and do not necessarily affect the end of the point. From this perspective, it makes sense that most ESV values hover around 0.50, as seen in Sections 6.1 and 6.2. But what we also saw in Section 6.2 is that once there is variation to the where the players are located, the ESV values diverge from 0.50. We argue that their is value in being able to quantify both the baseline average and the escape from normalcy, no matter how predictable it may be.

We also emphasize that this this is a first effort to quantify tennis points in this manner and several assumptions were made facilitate the computation. In particular, somewhat arbitrary choices were made for each of the weights on the strike and return categories in Sections 4.1 and 4.2. In the future, these weights could be estimated with a larger dataset. It is likely that better weights along with more data would increase the R-squared value of the linear fit. Furthermore, this paper mainly considered ground-strokes, but it may be beneficial to give more attention to serves and their returns, as we somewhat lumped them all together during the calculation of ESV. The model also assumed that each point was of the same importance, but future work could differentiate between normal points, and point such as break points or match points. Another promising area for future work would be in including shot type or shot characteristics such as speed, spin, shot height, or landing location into the computation of the ESV. This would better distinguish between different shots struck from the same location of the court. Ultimately, more data and bigger sample sizes mean more possibilities for insight and understanding in the under-researched field of tennis analytics. Hopefully, in the future, the USTA and other tennis organizations will be able to provide researchers with more data and more information to work with.

A Appendix

A.1 Theoretical ESV

Letting $T(\omega)$ denote the time at which a point following the path ω ends, the point's outcome for player i then is a deterministic function of the full resolution data at this time, $\Gamma^i(\omega) = h^i(Z_{T(\omega)}(\omega))$. The expectation $\mathbb{E}[\Gamma^i(\omega)|\mathcal{F}_t^{(Z)}]$ is an integral over the distribution of future paths the current point can take. Recall our definition of $\Gamma^i(\omega)$ such that $\Gamma^i(\omega) = h^i(Z_{T(\omega)}(\omega))$. Thus, evaluating ESV amounts to integrating over the joint distribution of $[T(\omega), Z_{T(\omega)}]$:

$$\begin{aligned} v_t^i(\omega) &= \mathbb{E}[\Gamma^i(\omega)|\mathcal{F}_t^{(Z)}] \\ &= \int_t^\infty \int_{\mathcal{Z}} h^i(z)\,\mathbb{P}(Z_s(\omega) = z|T(\omega) = s, \mathcal{F}_t^{(Z)}) \\ &\qquad \mathbb{P}(T(\omega) = s|\mathcal{F}_t^{(Z)})\,dz\,ds. \end{aligned} \tag{8}$$

A.2 Additional time notations

The notations τ_t and δ_t will be defined as:

$$\tau_t = \begin{cases} \max\{s : s < t, C_s(\omega) \notin \mathcal{C}_{\text{ServerShot}}\} + \varepsilon, \\ \qquad\qquad \text{if } C_t(\omega) \in \mathcal{C}_{\text{ServerShot}} \\ \max\{s : s < t, C_s(\omega) \notin \mathcal{C}_{\text{ReceiverShot}}\} + \varepsilon, \\ \qquad\qquad \text{if } C_t(\omega) \in \mathcal{C}_{\text{ReceiverShot}} \end{cases} \tag{9}$$

$$\delta_t = \begin{cases} \min\{s : s > \tau_t, C_s(\omega) \notin \mathcal{C}_{\text{ServerShot}}\}, \\ \qquad\qquad \text{if } C_t(\omega) \in \mathcal{C}_{\text{ServerShot}} \\ \min\{s : s > \tau_t, C_s(\omega) \notin \mathcal{C}_{\text{ReceiverShot}}\}, \\ \qquad\qquad \text{if } C_t(\omega) \in \mathcal{C}_{\text{ReceiverShot}} \end{cases} \tag{10}$$

where ε is the temporal resolution of the player-tracking data. In this case, ε will be equal to 1/25 s.

A.3 Strike/return subset definitions

- Let $\mathcal{S}$ contain every strike contained in the dataset, which includes who struck the shot, its strike state and its corresponding strike category.
  - $\mathcal{S} = \{\text{player ID}\} \times \{\mathcal{P}\} \times \{\Lambda_{\text{SC}}\}$
- $S^i \subseteq \mathcal{S}$: The subset of all strikes in $\mathcal{S}$ struck by player *i*.
- $S^i_{\beta^S} \subseteq \mathcal{S}$: The subset of all strikes in $\mathcal{S}$ struck by player *i* from the strike state $\beta^S \in \mathcal{P}$.
- $S^i_{\beta^S,\lambda_{\text{SC}}} \subseteq \mathcal{S}$: The subset of all strikes in $\mathcal{S}$ struck by player *i* from the strike state β^S, with the strike category $\lambda_{\text{SC}} \in \Lambda_{\text{SC}}$.
- Let $\mathcal{R}$ contain every return contained in the dataset, which includes who is returning the shot, its return state and its corresponding return category.
  - $\mathcal{R} = \{\text{player ID}\} \times \{\mathcal{P}\} \times \{\Lambda_{\text{RC}}\}$
- $R^i \subseteq \mathcal{R}$: The subset of all returns in $\mathcal{R}$ returned by player *i*.
- $R^i_{\beta^R} \subseteq \mathcal{R}$: The subset of all returns in $\mathcal{R}$ returned by player *i* from the return state $\beta^R \in \mathcal{P}$.

– $R^i_{\beta^R,\lambda_{\mathrm{RC}}} \subseteq \mathcal{R}$: The subset of all returns in $\mathcal{R}$ returned by player i from the return state β^R, with the return category $\lambda_{\mathrm{RC}} \in \Lambda_{\mathrm{RC}}$.

References

Bialkowski, A., P. Lucey, P. Carr, Y. Yue, S. Sridharan, and I. Matthews. 2014. "Large-Scale Analysis of Soccer Matches using Spatiotemporal Tracking Data." in *Data Mining (ICDM), 2014 IEEE International Conference on*, Shenzhen, China: IEEE, 725–730.

Brillinger, D. R. 2007. "A Potential Function Approach to the Flow of Play in Soccer." *Journal of Quantitative Analysis in Sports* 3:3.

Cervone, D., L. Bornn, and K. Goldsberry. 2016a. "Nba Court Realty." in *10th MIT Sloan Sports Analytics Conference*, Boston, MA, USA.

Cervone, D., A. D'Amour, L. Bornn, and K. Goldsberry. 2016b. "A Multiresolution Stochastic Process Model for Predicting Basketball Possession Outcomes." *Journal of the American Statistical Association* 111:585–599.

Franks, A., A. Miller, L. Bornn, and K. Goldsberry. 2015. "Counterpoints: Advanced Defensive Metrics for nba Basketball." in *9th Annual MIT Sloan Sports Analytics Conference*, Boston, MA, USA.

Goldsberry, K. 2012. "Courtvision: New Visual and Spatial Analytics for the nba." in 2012 *MIT Sloan Sports Analytics Conference*.

Goldsberry, K. and E. Weiss. 2013. "The Dwight Effect: A New Ensemble of Interior Defense Analytics for the nba." *Sports Aptitude, LLC. Web*.

Horton, M., J. Gudmundsson, S. Chawla, and J. Estephan. 2014. "Classification of Passes in Football Matches using Spatiotemporal Data." *arXiv preprint arXiv:1407.5093*.

Kemeny, J. G. and J. L. Snell. 1976. *Markov Chains*. New York: Springer-Verlag.

Loukianov, A. and V. Ejov. 2012. "Markov Modeling for a Tennis Point Played." *International Journal of Sports Science and Sports Engineering* (in print).

Lucey, P., D. Oliver, P. Carr, J. Roth, and I. Matthews. 2013. "Assessing Team Strategy using Spatiotemporal Data." in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, Chicago, IL, USA: ACM, 1366–1374.

Lucey, P., A. Bialkowski, P. Carr, Y. Yue, and I. Matthews. 2014a. "How to get an Open Shot: Analyzing Team Movement in Basketball using Tracking Data." in *8th Annual MIT Sloan Sports Analytics Conference*, Boston, MA, USA.

Lucey, P., A. Bialkowski, M. Monfort, P. Carr, and I. Matthews. 2014b. "Quality vs Quantity: Improved Shot Prediction in Soccer using Strategic Features from Spatiotemporal Data." in *9th Annual MIT Sloan Sports Analytics Conference*, Boston, MA, USA, 1–9.

Miller, A., L. Bornn, R. P. Adams, and K. Goldsberry. 2014. "Factorized Point Process Intensities: A Spatial Analysis of Professional Basketball." in *International Conference on Machine Learning*, Beijing, China, 235–243.

Mora, S. V. and W. J. Knottenbelt. 2016. "Spatio-Temporal Analysis of Tennis Matches," in *KDD Workshop on Large-Scale Sports Analytics*, San Francisco, CA, USA.

Reid, M. and R. Duffield. 2014. "The Development of Fatigue During Match-Play Tennis." *British Journal of Sports Medicine* 48:i7–i11.

Spanias, D. and W. J. Knottenbelt. 2013. "Predicting the Outcomes of Tennis Matches using a Low-Level Point Model." *IMA Journal of Management Mathematics* 24:311–320.

Wei, X., P. Lucey, S. Morgan, and S. Sridharan. 2013. "Sweet-Spot: Using Spatiotemporal Data to Discover and Predict Shots in Tennis." in *7th Annual MIT Sloan Sports Analytics Conference*, Boston, MA, USA.

Wei, X., P. Lucey, S. Morgan, P. Carr, M. Reid, and S. Sridharan. 2015. "Predicting Serves in Tennis using Style Priors." in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2207–2215.