TESIS DOCTORAL

ANÁLISIS TÉCNICO-TÁCTICO DEL SERVICIO EN LA DISCIPLINA DE DOBLES DE TENIS PROFESIONAL A TRAVÉS DE LA TECNOLOGÍA DE OJO DE HALCÓN Y DEL ANÁLISIS NOTACIONAL

Doctorando:

Fernando Vives Albiol

Directores:

Dr. José Francisco Guzmán Luján

Dr. Rafael Martínez Gallego

Dr. Miguel Crespo Celda

2024

Esta tesis ha sido depositada en la Facultad de Ciencias de la Actividad Física y del Deporte de la Universidad de Valencia, por D. Fernando Vives Albiol con DNI 29197695-T.

Valencia, septiembre de 2024

A mis dos mejores golpes,

Luca y Maël

"It's no accident, I think, that tennis uses the language of life. Advantage, service, fault, break, love, the basic elements of tennis are those of everyday existence, because every match is a life in miniature."

— Andre Agassi, Open

AGRADECIMIENTOS

Recuerdo con claridad aquellas tardes en el chalet de mis abuelos, cuando, antes de jugar al tenis, preparábamos la pista como si de un ritual sagrado se tratase. El aroma inconfundible de la tierra batida, húmeda y recién mojada, impregnaba el ambiente. En esos momentos, el mundo parecía detenerse y una calma indescriptible lo envolvía todo. Desde la preparación de la pista hasta el primer servicio, todo adquiría un estado de paz y armonía que, aún hoy, evoca en mí una profunda nostalgia.

Mis padres me abrieron las puertas a este maravilloso deporte que ha marcado mi vida desde la infancia y que, con el paso del tiempo, se ha convertido en mi pasión y profesión. No solo me enseñaron a coger una raqueta, sino que también me transmitieron valores fundamentales que han trascendido más allá del propio deporte. Para mi hermana Cristina y para mí, son un ejemplo de vida, tanto en su dedicación al deporte como en su integridad como personas. Me han enseñado que el éxito en el tenis y en la vida se construye a base de esfuerzo, trabajo, respeto y humildad.

Gracias al tenis, he tenido la inmensa suerte de conocer a gente maravillosa que han sido fundamentales en mi vida, tanto dentro como fuera de la pista. A lo largo de los años, he compartido experiencias enriquecedoras no solo con jugadores, sino también con árbitros, entrenadores o padres, quienes han contribuido de manera significativa a mi crecimiento, tanto a nivel profesional como personal. En especial, me gustaría dar las gracias a Juanmi Mirálles, director deportivo del Club Español de Tenis, y a mis dos «maestros» deportivos Juan Giner y Josevi Pérez.

Quisiera expresar mi más sincero agradecimiento a mis directores de tesis, Miguel y José, por su infinita paciencia y dedicación. Su guía ha sido un ejemplo constante de excelencia, no solo en el ámbito profesional, sino también en el personal, demostrando que el rigor académico puede ir de la mano de una gran humanidad y cercanía. Y, por supuesto, a Rafa, quien me introdujo en este maravilloso mundo de la investigación y ha estado a mi lado en cada paso del proceso. No encuentro palabras suficientes para expresar mi profunda gratitud.

Me gustaría darle las gracias a Javi L., que no dudó en acompañarme en esta travesía y me abrió las puertas al fascinante universo del *Machine Learning.* Su predisposición y dedicación han sido fundamentales para llevar a buen puerto este proyecto. También quisiera agradecer a Dr. Roig, referente en el ámbito académico, y a Javi F., el más

optimista de los pesimistas, cuya ayuda ha sido siempre tan valiosa como desinteresada en los momentos que más la he necesitado.

Por último, quiero dedicar un agradecimiento muy especial a Sandra, quien ha tenido la inmensa paciencia de acompañarme durante todos estos años. Su generosidad, apoyo y comprensión han sido fundamentales para mí. Al igual que le ocurre al personaje de Jack Nicholson en Mejor imposible, tú me inspiras a ser una mejor versión de mí mismo.

ÍNDICE

INDICE DE TABLAS

INDICE DE FIGURAS

RESUMEN

El saque se ha convertido en una herramienta clave en el tenis moderno, con una incidencia directa en el juego, especialmente en el contexto de los partidos de dobles. Los estudios realizados se centran en el servicio en los partidos de dobles de la Copa Davis y la Copa Billie Jean King, y emplean datos obtenidos a través del sistema Ojo de Halcón (*Hawk-Eye*) y técnicas avanzadas de aprendizaje automático (*Machine Learning*). El primer estudio tuvo como objeto analizar las diferencias en las características espacio-temporales y la efectividad de los servicios en el tenis profesional de dobles masculino. Para ello se analizaron un total de 20.419 saques en 96 partidos de la Copa Davis disputados entre 2010 y 2019. Los resultados mostraron que los primeros saques dirigidos a las zonas Abierta y T fueron los más efectivos, mientras quela posición del sacador variaba en función de la dirección del servicio, independientemente del lado donde se producía el servicio. El segundo estudio utilizó técnicas de inteligencia artificial (*Deep Learning*) para analizar un conjunto de datos compuesto por 14.146 saques en 97 partidos de la Copa Davis disputados entre 2010 y 2019, con el fin de identificar las variables clave que maximizan la efectividad del servicio. Los resultados mostraron que las variables del ángulo de servicio y la distancia del bote a las líneas laterales eran las que tenían una mayor influencia a la hora de lograr un *ace* o saque directo. En cambio, la velocidad del servicio mostró una influencia menor en la efectividad en dobles que en el tenis individual, ya que velocidades superiores a 187 km/h solo aumentaban un 10 % la probabilidad de convertir un saque directo o *ace*. El objetivo del tercer estudio fue identificar las variables que influyen en la efectividad del primer saque en partidos de dobles femenino profesional. Se analizaron 4.939 primeros saques de 145 jugadoras en partidos de la copa Billie Jean King. Los resultados identificaron que el ángulo del saque era la variable de mayor relevancia, seguida por la velocidad y la distancia del bote respecto a las líneas laterales. Estos hallazgos proporcionan información valiosa para las estrategias de entrenamiento y preparación táctica en dobles profesional, en concordancia con investigaciones previas sobre tenis individual masculino. Además, ofrece a jugadores y entrenadores datos muy concretos que pueden utilizarse durante la competición para mejorar la efectividad del saque o para ayudar a mejorar la anticipación en el resto.

RESUM

El servei s’ha convertit en una eina clau en el tennis modern, amb una incidència directa en el joc, especialment en el context dels partits de dobles. Els estudis realitzats es centren en el servei dels partits de dobles de la Copa Davis i la Copa Billie Jean King, i utilitzen dades obtingudes a través del sistema de Ull de Falcó (Hawk-Eye) i tècniques avançades d'aprenentatge automàtic (Machine Learning) . El primer estudi va tenir com a objecte analitzar les diferències en les característiques espai-temporals i l'efectivitat dels serveis en el tennis professional de dobles masculí. Per això, es van analitzar un total de 20.419 serveis en 96 partits de la Copa Davis disputats entre 2010 i 2019. Els resultats van mostrar que els primers serveis dirigits a les zones Oberta i T van ser els més efectius, mentre que la posició del servidor variava en funció de la direcció del servei, independentment del costat on es produïa el servei. El segon estudi va utilitzar tècniques d'intel·ligència artificial (Deep Learning) per analitzar un conjunt de dades compostes per 14.146 serveis en 97 partits de la Copa Davis disputats entre el 2010 i el 2019, per identificar les variables clau que maximitzaven l'efectivitat del servei. Els resultats van mostrar que les variables de l'angle de servei i la distància del pot a les línies laterals eren les que tenien una influència més gran a l'hora d'aconseguir un *ace* o servei directe. En canvi, la velocitat del servei va mostrar una influència menor en l'efectivitat en dobles que en el tennis individual, ja que velocitats superiors a 187 km/h només augmentaven un 10% la probabilitat de convertir un servei directe o *ace*. L'objectiu del tercer estudi va ser identificar les variables que influeixen en l'efectivitat del primer servei en partits de dobles femení professional. Es van analitzar 4.939 primers serveis de 145 jugadores en partits de la copa Billie Jean King. Els resultats van identificar que l'angle de la treta era la variable de més rellevància, seguida per la velocitat i la distància del pot respecte a les línies laterals. Aquestes troballes proporcionen valuosa informació per a les estratègies d'entrenament i preparació tàctica en dobles professional, en concordança amb investigacions prèvies sobre tennis individual masculí. A més, ofereix a jugadors i entrenadors dades molt concretes que es poden utilitzar durant la competició per millorar l'efectivitat del servei o per ajudar a millorar l'anticipació a la resta.

ABSTRACT

The serve has become a key tool in modern tennis, with a direct impact on the game, especially in the context of doubles matches. The studies conducted focus on the serve in Davis Cup and Billie Jean King Cup doubles matches, using data from the Hawk-Eye system and advanced Machine Learning techniques. The first study aimed to analyse the differences in the spatio-temporal characteristics and effectiveness of serves in professional men's doubles tennis. A total of 20,419 serves were analysed in 96 Davis Cup matches played between 2010 and 2019. The results showed that first serves aimed at the open and T zones were the most effective, while the position of the server varied according to the direction of the serve, regardless of the side from which the serve was made. The second study used artificial intelligence (Deep Learning) techniques to analyse a dataset of 14,146 serves from 97 Davis Cup matches played between 2010 and 2019 to identify the key variables that maximise serve effectiveness. The results showed that the variables of service angle and distance from the bounce to the sideline had the greatest influence on achieving an ace or direct serve. In contrast, service speed had less of an impact on effectiveness in doubles than in singles, with speeds above 187km/h only increasing the probability of a direct serve or ace by 10%. The aim of the third study was to identify the variables that influence the effectiveness of the first serve in professional women's doubles matches. A total of 4,939 first serves from 145 players in Billie Jean King Cup matches were analysed. The results showed that the angle of the serve was the most important variable, followed by the speed and distance of the sideline bounce. These findings provide valuable information for training strategies and tactical preparation in professional doubles, in line with previous research in men's singles tennis. It also provides players and coaches with very concrete data that can be used during competition to improve the effectiveness of the serve or the anticipation of the return.

1. PRESENTACIÓN DE LOS ESTUDIOS

Este trabajo de investigación se ha desarrollado a partir de una recopilación de artículos publicados durante el transcurso de los últimos años, coincidiendo con la realización del programa de doctorado en Actividad Física y Deporte por parte del autor.

Estudio 1.
Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology (2022).

Revista: *Journal of Performance Analysis in Sport*
DOI: https://doi.org/10.1080/24748668.2022.2110555

Año	2021
JCR (Journal of Citations Report)	**Impact Factor:** 1.9
Scopus Metrics	**CiteScore**: 3.7 **CiteScore Best Quartile**: Q2
SJR (SCImago Journal Rank)	0.71

Estudio 2.
Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve (2023).

Revista: *Applied Sciences*
DOI: https://doi.org/10.3390/app132413213

Año	2022
JCR (Journal of Citations Report)	**Impact Factor:** 2.7
Scopus Metrics	**CiteScore**: 4.5 **CiteScore Best Quartile**: Q2
SJR (SCImago Journal Rank)	0.49

Estudio 3.

Serve Strategy in Elite Women's Doubles Tennis - A Data-Driven Perspective (2024).

Revista: *RETOS*

DOI: https://doi.org/10.47197/retos.v60.107972

Año	2023
JCR (Journal of Citations Report)	**Impact Factor:** 1.99
Scopus Metrics	**CiteScore**: 3.2 **CiteScore Best Quartile**: Q2
SJR (SCImago Journal Rank)	0.38

2. INTRODUCCIÓN

2.1. LA DISCIPLINA DE DOBLES EN TENIS

Desde sus albores, el tenis moderno ha integrado la modalidad de dobles, como así evidencia la primera pareja ganadora en esta disciplina en Wimbledon en el año 1884, apenas siete años después de la primera final del torneo individual masculino. A lo largo de la historia, han destacado diversos especialistas en esta disciplina, quienes han recibido reconocimiento tanto por sus logros deportivos como a nivel económico (Breznik, 2015). Tanto la Federación Internacional de Tenis (ITF), la Asociación de Tenis Femenino (WTA) y la Asociación de Tenistas Profesionales (ATP) otorgan una gran importancia a esta modalidad, estando presente en todos sus torneos.

En la modalidad de dobles, en contraste con el juego individual, se experimenta un cambio significativo en cuanto al número de jugadores, pasando de dos a cuatro participantes en la pista. Asimismo, se modifican las dimensiones de la pista, con la inclusión de pasillos laterales. Este ajuste, que amplía el espacio de juego casi 3 metros, altera la dinámica táctica del encuentro. Una de las consecuencias más notables radica en el posicionamiento estratégico de los jugadores durante el desarrollo del juego. Esta variación en la táctica se manifiesta en la disposición de los competidores en la pista, lo que incide directamente en la toma de decisiones y en la ejecución de las jugadas durante el transcurso del partido (Crespo y Miley, 1998).

En cuanto al formato de juego, los dobles masculinos y femeninos son la disciplina del tenis que más modificaciones han sufrido durante los últimos años, tratando de darles un mayor atractivo para los potenciales espectadores. Los torneos del circuito ATP y WTA desde 2012 se disputan con punto decisivo en caso de *Deuce* (40 – 40), y un *match tiebreak* a 10 puntos en el lugar de tercer set. En cambio, en los *Grand Slam* si que se sigue jugando con ventajas, pero hay disparidad de criterios en el desenlace del tercer set, mientras el Open de Australia y el de Estados Unidos utilizan el *match tiebreak set,* Roland Garros y Wimbledon siguen utilizando un *set* de *tiebreak* (formato tradicional) con diferencia de dos juegos. Por último, en las dos grandes competiciones por equipos nacionales, la Copa Davis y la *Billie Jean King* los partidos se disputan al mejor de tres sets de *tie break* con ventajas (ITF, 2024a).

A pesar de la reconocida importancia y tradición de la modalidad de dobles en el tenis profesional, existe una notable carencia de investigación específica en este contexto. Aunque podemos encontrar estudios como el de Breznik (2015) el cual analizó el rendimiento y los logros de los jugadores de dobles para determinar los equipos y jugadores más exitosos de la historia del tenis. También han sido analizadas las diferentes formaciones que utilizan los diferentes equipos en competición a nivel táctico, la estructura temporal y la finalización de los puntos en pista rápida (Kocib et al., 2020; Martínez-Gallego et al., 2020). La comunicación entre los jugadores de los equipos ha sido motivo de estudio, mostrando diferencias entre las parejas ganadoras y perdedoras (Lausic et al. 2009; 2014; 2015).

Aunque el servicio, como veremos más adelante, ha sido el golpe que ha recibido un mayor interés por parte de los investigadores, otros golpes específicos han sido evaluados como la volea y sus diferentes patrones de posicionamiento durante los partidos (Martínez-Gallego, 2021c). La experiencia de los equipos también ha sido analizada para ver su incidencia en el juego durante la disputa de la competición (Martínez-Gallego, 2021d). Por último, se ha comparado en dobles femenino la actividad de las jugadoras, la estructura de juego y la finalización de los puntos en torneos de Grand Slam (Borderías et al., 2021; 2023).

2.2. LA COPA DAVIS Y LA COPA BILLIE JEAN KING

En el verano de 1899, un grupo de estudiantes de Harvard, encabezados por Dwight F. Davis, Holcombe Ward y Malcolm Whitman, junto con su compañero Beals Wright y su padre presentaron al presidente de la Asociación Nacional de Tenis de Estados Unidos, James Dwight, la propuesta de una competición internacional. Este evento marcó el inicio de la Copa Davis, aunque inicialmente se conociera como International Lawn Tennis Challenge (Gillmeister, 1997).

El primer enfrentamiento entre Estados Unidos y las Islas Británicas tuvo lugar en 1900 en el Longwood Cricket Club de Boston. La Federación Internacional de Tenis (ITF) conmemoró el Centenario de la competición en 1999, destacando entre los numerosos eventos conmemorativos el enfrentamiento de segunda ronda entre Estados Unidos y Australia, celebrado en el mismo escenario que la primera edición del torneo. El palmarés de la Copa Davis incluye a 16 países diferentes que han logrado alzarse con el trofeo de campeón. Encabezando la lista se encuentran los Estados Unidos, con un total de 32 títulos (*Davis Cup*, s. f.).

En 2018, la Asamblea General Anual de la Federación Internacional de Tenis (ITF) aprobó un cambio de formato histórico para la competición, propuesto por la empresa Kosmos y la propia ITF. Esta modificación implicó la disputa de la fase final de la Copa Davis durante una única semana y en una sede única, con la participación de 18 equipos nacionales distribuidos en 6 grupos. En la primera edición en 2019, bajo este nuevo formato, España se proclamó campeona al vencer a Canadá por 2-0 en la final celebrada en la Caja Mágica de Madrid (RFET, Real Federación Española de Tenis, s. f.).

El concepto inicial de la Copa Federación se remonta a 1919 cuando Hazel Wightman propuso una competición de equipos femeninos. Aunque la competición no se materializó hasta 1963, coincidiendo con el 50 aniversario de la ITF y corono a los Estados Unidos como campeonas en esta primera edición, un logro que marcó el inicio de una larga historia de éxito para el equipo americano. En 1994, participaron más de 70 países, lo que evidenció la creciente importancia y popularidad del torneo, que se convirtió en un pilar para el desarrollo del tenis femenino en todo el mundo.

La evolución del formato de la Copa Federación incluyó la introducción de competiciones regionales en 1992 y un nuevo formato en 1995, inspirado en la Copa Davis masculina. A lo largo de los años, se han realizado ajustes en el formato, culminando en el actual sistema introducido en 2020, que incluye las Finales de la Billie Jean King Cup (BJKC) con 12 equipos clasificados. Además, se establecieron grupos regionales donde las naciones compiten por ascensos y descensos anuales, con el objetivo de fortalecer la competición a nivel regional y global.

El cambio de nombre de la competición en 2020 representa un homenaje a la figura icónica de la jugadora Billie Jean King. Este cambio refleja los valores y principios de King, que han sido fundamentales en la promoción del tenis femenino y en la búsqueda de la igualdad de género en el deporte. Con su legado como inspiración, la BJKC continúa siendo el principal evento del tenis femenino, equiparando su prestigio y premios con los de la Copa Davis masculina (BJKC, s. f.).

2.3. EL SERVICIO

Según la regla del tenis nº16 de la ITF (2024b): "El servidor se pondrá con ambos pies situados detrás de la línea de fondo (es decir, lo más alejados de la red) y dentro de las prolongaciones imaginarias de la marca central y la línea lateral. Entonces, el servidor lanzará la pelota al aire con la mano en cualquier dirección y la golpeará con la raqueta antes de que toque el suelo. Se considerará que se ha ejecutado el servicio en el momento en que el jugador da un raquetazo, lo mismo si toca la pelota como si falla al intentarlo. Un jugador que solamente usa un brazo puede utilizar la raqueta para lanzar la pelota. La pelota servida deberá pasar por encima de la red y aterrizará dentro del recuadro de servicio que esté diagonalmente opuesto, antes de que el restador la devuelva".

El servicio se ha estudiado en diversos contextos, desde las diferencias según el nivel y la superficie de juego (Varverka et al., 2018; Fernández-García et al., 2019) hasta las estrategias de ritmo y velocidad de los sacadores en partidos a cinco sets (Martin et al., 2019). Mecheri et al. (2016) examinaron diferentes características del servicio para determinar su incidencia en el juego. Mientras que Whiteside y Reid (2017) identificaron las variables más importantes para lograr un saque directo o *ace*.

La velocidad del servicio ha sido posiblemente la variable que más interés ha recibido por parte de los investigadores entre las diferentes relacionadas con el saque. Los estudios han demostrado que los jugadores tienden a utilizar un primer saque más rápido para tomar ventaja en el punto, mientras que prefieren un segundo saque más lento para asegurar la consistencia (Cui et al., 2017; Brown, 2021; Martínez-Gallego et al., 2021a). Las investigaciones han concluido que los jugadores utilizan saques planos para maximizar la velocidad en el primer saque y saques liftados para reducir el riesgo en el segundo saque (Barnett, 2009; Gilet et al., 2009; Sakurai et al., 2013).

En cuanto a las diferencias de género en el saque, se han encontrado variaciones en la velocidad media, el número de saques directos y el porcentaje de puntos de servicio ganados (Elliot et al., 2013; Torres-Luque et al., 2019; Reid et al., 2016). Las elecciones de dirección del primer saque también difieren entre hombres y mujeres (Hizan et al., 2015; Sánchez-Pay et al., 2019). Estas diferencias se han estudiado en relación con los diferentes tipos de superficie de juego y su impacto en los resultados de los partidos (Carboch, 2017; Fitzpatrick et al., 2019). Además, se han investigado las diferencias de rendimiento entre jugadores profesionales y juveniles, con importantes implicaciones para la planificación de los entrenamientos y las competiciones (Hizan et al., 2011; Krause et al., 2019; Kovalchik y Reid, 2017).

En relación a la disciplina de dobles, Carboch y Kočíb (2016) compararon la eficiencia del servicio entre parejas de dobles masculinas y femeninas, encontrando que los hombres ganan más puntos al servicio y las mujeres al resto. También se ha analizado las diferencias en la dirección y efectividad del saque equipos en hombres y mujeres (Martínez- Gallego et al., 2021a) o los diferentes tipos de formación de los equipos al iniciar el punto al resto y al servicio (Kocib et al., 2020). Martínez-Gallego et al. (2021b) examinaron el impacto del saque en el tenis masculino de dobles, desafiando las creencias comunes sobre su importancia y Avkhimovich, (2023) analizó el orden del servicio de la pareja en función de las características del servicio de los jugadores.

2.4. LA TECNOLOGÍA DE OJO DE HALCÓN

El estudio del rendimiento deportivo ha sido objeto de numerosas investigaciones, tanto en las sesiones de entrenamiento como en la competición oficial. Esto se ha convertido en una herramienta muy útil para que los entrenadores puedan optimizar su trabajo, para que los jugadores logren un mejor rendimiento, o para una mejor comprensión y análisis por parte de los investigadores (Hughes & Bartlett, 2008; O'Donoghue, 2014).

En el tenis, como ocurre en otros deportes, se ha producido una gran evolución en la obtención y recogida de la información. Dichos registros pueden utilizarse para optimizar el proceso de entrenamiento o prepararse de forma más eficaz y específica para la competición (Morgulev et al., 2018). En la actualidad, se utiliza una amplia gama de dispositivos, como son las cámaras de alta velocidad, los sistemas de seguimiento y etiquetado, o el sistema GPS, para la recopilación de los datos (Barris & Butto, 2008; Jindo et al., 2022).

Un ejemplo destacado de esta tecnología es el sistema de Ojo de Halcón, utilizado por primera vez en el críquet y posteriormente, en el tenis profesional durante la Copa Hopman en la edición de 2006. Este sistema de seguimiento emplea cámaras de alta velocidad y algoritmos computacionales para rastrear la trayectoria de la pelota con un margen de error mínimo (3,6 mm). Su implementación ha revolucionado la forma en que se juzgan las jugadas en eventos deportivos, ofreciendo una herramienta precisa y confiable para la toma de decisiones (Shigh Bal & Dureja, 2012; Baodong, 2014; Mecheri et al., 2016).

El uso de la tecnología de Ojo de Halcón en el tenis ha dado acceso a una amplia gama de datos que han sido objeto de estudio en diversos proyectos de investigación. El servicio dado su papel fundamental en el juego, ha sido analizado en profundidad en función de su trayectoria y eficacia en trabajos como el de Rioult et al. (2015) o Mecheri et al. (2016). Estos estudios han abarcado las estrategias y distribución del servicio en jugadores profesionales o un propio sistema de análisis notacional para la evaluación del saque y el resto (Loffing et al., 2009; Hizan et al., 2010; Kolbinger y Lames, 2013). Otros estudios en cambio, han intentado predecir el servicio en un contexto y rival determinado (Wei et al., 2015) o han identificado los factores más importantes para lograr un saque directo (Whiteside y Reid, 2017). También ha sido tratado y analizado el servicio en la disciplina de dobles por Martínez-Gallego et al. (2021a) y Vives et al. (2022). Y estudios más recientes han propuesto mejorar la efectividad del servicio en el tenis de élite (Vives et al., 2023; Fizpatrick et al., 2023).

Otros aspectos del juego han sido explorados, así Wei et al. (2013) analizaron la predicción de golpes o Loffing y Hagemann (2014) que examinaron la posición en la pista de los jugadores en relación con la anticipación del golpe de fondo del adversario. También se han analizado las características de los partidos de Grand Slam se ha comparado a tenistas juniors y profesionales o se ha evaluado un modelo teórico de la trayectoria de la pelota (Reid et al.,2016; Kovalchik y Reid, 2017; Cant et al., 2020). En su estudio de 2018, Kovalchik y Reid clasificaron los tipos de golpes en el tenis profesional, asociando los tiros ganadores con una mayor velocidad. Giles et al. (2020) analizaron los cambios de dirección entre jugadores profesionales y juniors. Por último, Meurs et al. (2021) investigaron el Índice de Ventaja Posicional como una herramienta para identificar la ventaja de un jugador sobre su oponente.

2.5. APRENDIZAJE AUTOMÁTICO EN EL TENIS.

En los últimos años ha habido una proliferación en el uso del aprendizaje automático (Machine Learning, ML) en el tenis, principalmente debido a los avances de la tecnología en la recopilación de datos a través de diferentes sistemas, siendo los sistemas de seguimiento y etiquetado los utilizados. Esta gran cantidad de información ha requerido un tratamiento más sofisticado y complejo que los anteriores modelos de análisis, dando lugar a una mayor profundidad y complejidad en los resultados obtenidos.

La literatura científica ha dado lugar a diversos estudios sobre la predictividad del golpeo en el tenis profesional basándose en determinados parámetros del jugador y de las zonas de golpeo (We et al., 2016, Shimizu et al., 2019; Makino et al., 2020; Zhou y Liu, 2024). Por el contrario, otros autores han intentado clasificar o determinar patrones específicos en relación a diferentes situaciones del juego, golpes de fondo, voleas y restos (Kovalchik y Reid, 2018; Martínez-Gallego et al., 2021c).

Varios estudios han explorado diversos aspectos del rendimiento en el tenis. Cui et al. (2019) identificaron los factores clave para el éxito en tenistas masculinos de Grand Slam. Giles et al. (2020, 2021, 2023) desarrollaron una metodología para detectar y clasificar los movimientos de cambio de dirección en jugadores profesionales. En cambio, otras investigaciones se han centrado en entender las características morfológicas de los deportistas en edad juvenil para proyectar su futuro rendimiento en la élite (Panjan, 2010; Siener et al., 2021).

Como hemos visto anteriormente, el servicio debido a su relevancia en el tenis moderno ha sido motivo de interés en una gran variedad de estudios. Así Mecheri et al. en 2016, determinaron que el saque tiene un efecto significativo en el resultado del punto. La predicción de la dirección del servicio también ha sido abordada en otros estudios como el de Wei et al. (2015). También se ha analizado los principales parámetros de los primeros servicios para lograr un saque ganador, tanto en individuales (Whiteside y Reid, 2017) como en dobles (Vives et al., 2023). O el último ejemplo lo podemos encontrar en el trabajo de Gao y Kowalcyk (2021) identificaron el saque como un predictor crucial para determinar el vencedor del partido.

2.6. REFERENCIAS

Avkhimovich, N. (2023). Should the strongest serve first? Strategic ordering of servers in doubles tennis. *Applied Economics*, 1-10.

Baodong, Y. (2014). Hawkeye technology using tennis match. *Computer Modelling & New Technologies*, *18*(12C), 400-402.

Barnett, T. (2009). Applying tennis match statistics to increase serving performance during a match in progress. *Journal of Medicine and Science in Tennis*, *14*(3), 16-19.

Barris, S., & Button, C. (2008). A review of vision-based motion analysis in sport. *Sports medicine*, *38*, 1025-1043.

BJKC (s.f.). Billie Jean King Cup History. https://www.billiejeankingcup.com/es/history

Breznik, K. (2015). Revealing the best doubles teams and players in tennis history. *International Journal of Performance Analysis in Sport, Taylor & Francis Journals*, *15*(3), 1213-1226.

Brown, E. G. (2021). A faster serve has more impact on success for female elite tennis players than males. *International Journal of Performance Analysis in Sport*, *21*(4), 600-610.

Borderias, M., Crespo, M., Martínez-Gallego, R., & Baiget, E. (2023). Comparison of the game structure and point ending during Grand Slam women's doubles tennis. *International Journal of Sports Science & Coaching*, *18*(4), 1248-1255.

Borderias, M., Fernández-Fernández, J., Martínez-Gallego, R., & Baiget, E. (2021). Comparison of the activity profile during Grand Slam women's doubles tennis. *International Journal of Performance Analysis in Sport*, *21*(6), 1066–1080.

Cant, O., Kovalchik, S., Cross, R. & Reid, M. (2020). Validation of ball spin estimates in tennis from multi-camera tracking data. *Journal of Sports Sciences*, *38*(3), 296-303.

Carboch, J. (2017). Comparison of game characteristics of male and female tennis players at grand-slam tournaments in 2016. *TRENDS in Sports Sciences,* (4), 151-154.

Carboch, J., & Kočíb, T. (2016). A comparison of service efficiency between players of male and female doubles at professional tennis tournaments. *Auc Kinanthropologica*, *51*(2), 56-62.

Cayer, L. (2018). El entrenamiento en dobles en el circuito ATP. En Crespo, M., Braam, M. y Probert, A. (Ed.). *Lecturas esenciales para entrenadores del circuito (Vol.I),* 31-38. International Tennis Federation.

Crespo, M., & Miley, D. (1998). *Advanced coaches manual*. International Tennis Federation.

Cui, Y., Gómez, M. Á., Gonçalves, B., Liu, H. & Sampaio, J. (2017). Effects of experience and relative quality in tennis match performance during four Grand Slams. *International Journal of Performance Analysis in Sport*, 17(5), 783-801.

Cui, Y., Liu, H., Liu, H., & Gómez, M. Á. (2019). Data-driven analysis of point-by-point performance for male tennis player in Grand Slams. *Motricidade*, *15*(1), 49-61.

Davis Cup (s.f.). Copa Davis - Historia. https://www.daviscup.com/es/organizaci%C3%B3n/historia.aspx

Gao, Z., & Kowalczyk, A. (2021). Random forest model identifies serve strength as a key predictor of tennis match outcome. *Journal of Sports Analytics*, *7*(4), 255-262.

Giles, B., Kovalchik, S. & Reid, M. (2020). A machine learning approach for automatic detection and classification of changes of direction fromplayer tracking data in professional tennis. *Journal of Sports Sciences*, *38*(1), 106-113.

Gillet, E., Leroy, D., Thouvarecq, R. & Stein, J.F. (2009). A notational analysis of elite tennis serve and serve-return strategies on slow Surface. *J Strength Cond Res, 23*(2), 532–539.

Gillmeister, H. 1997 . *Tennis: A Cultural History*, London: Leicester University Press.

Fernández-García, A. I., Torres-Luque, G., Hernández-García, R. & Blanca-Torres, J. C., (2019). Análisis de las variables estadísticas relacionadas con el servicio en tenis masculino de alto rendimiento en categoría junior y absoluto. *Cultura Ciencia Deporte*, 14(42), 285-295.

Fitzpatrick, A., Stone, J.A., Choppin, S. & Kelly, J. (2019). Important performance characteristics in elite clay and grass court tennis match-play. *International Journal of Performance Analysis in Sport*, 19(6), 942-952.

Fitzpatrick, A., Joseph A. Stone, Joseph A., Choppin, S. & Kelley, J. (2023) Analysing Hawk-Eye ball-tracking data to explore successful serving and returning strategies at Wimbledon, International Journal of Performance Analysis in Sport, DOI: 10.1080/24748668.2023.2291238

Hawk-Eye Innovations. (2015). Hawk-Eye's accuracy & reliability: Electronic line calling. Basingstoke.

Hizan, h., Whipp, P. & Reid, M. (2010). Validation of Match Notation (A Coding System) in Tennis. *Journal of Quantitative Analysis in Sports*, *6*(3), 11.

Hizan, h., Whipp, P. & Reid, M. (2011). Comparison of serve and serve return statistics of high performance male and female tennis players from different age-groups. *International Journal of Performance Analysis in Sport*, *11*(2), 365-375.

Hizan, h., Whipp, P. & Reid, M. (2015). Gender Differences In The Spatial Distributions Of The Tennis Serve. *International Journal of Sports Science & Coaching*, *10*(1), 87-96.

Hughes, M. & Bartlett, R. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences*, *20*(19), 739-754.

ITF (2024a). Davis Cup Rules and Regulations. https://www.daviscup.com/media/330736/330736.pdf

ITF (2024b). Rules of Tennis. Rules and Regulations. ITF LTD.

Kocib, T., Carboch, J., Cabela M. & Kresta, J. (2020). Tactics in tennis doubles: analysis of the formations used by the serving and receiving teams. *International Journal of Physical Education, Fitness and Sport*, *9*(2).

Kolbinger, O., & Lames, M. (2013). Ball trajectories in tennis-Lateral and vertical placement of right handed men's singles serves. *International Journal of Performance Analysis in Sport*, *13*(3), 750-758.

Kovalchik, S. & Reid, M. (2017). Comparing Matchplay Characteristics and Physical Demands of Junior and Professional Tennis Athletes in the Era of Big Data. *Journal of Sports Science and Medicine*, *16*, 489-497.

Kovalchik, S. & Reid, M. (2018). A shot taxonomy in the era of tracking data in professional tennis. *Journal of Sports Sciences*, *36*(18), 2096-2104.

Lausic, D., Tennebaum, G., Eccles, D., Jeong, A., y Johnson, T. (2009). Intrateam Communication and Performance in Doubles Tennis. *Research Quarterly for Exercise and Sport*, *80*(2), 281–290.

Lausic, D., Razon, S., & Tenenbaum, G. (2014). Verbal Communication in Doubles Tennis: An Analysis of Close Games. *The Sport Psychologist*, *28*(4), 361-374.

Lausic, D., Razon, S., & Tenenbaum, G. (2015). Nonverbal sensitivity, verbal communication, and team coordination in tennis doubles. *International Journal of Sport and Exercise Psychology*, *13*(4), 398–414.

Loffing, F., Hagemann, N., & Strauss, B. (2009). The serve in professional men's tennis: effects of players' handedness. *International Journal of Performance Analysis in Sport*, *9*(2), 255–274.

Loffing, F. & Hagemann, N. (2014). On-court position influences skilled tennis players' anticipation of shot outcome. *J Sport Exerc Psychol*, *36*(1), 14–26.

Makino, M., Odaka, T., Kuroiwa, J., Suwa, I. & Shirai, H. (2020). Feature selection to win the point of ATP tennis players using rally information. *International Journal of Computer Science in Sport*, *19*(1), 37-50.

Martínez-Gallego, R., Crespo, M. & Jiménez, J. (2021a). Analysis of the differences in serve effectiveness between Billie Jean King Cup (former Fed Cup) and Davis Cup doubles tennis matches. I*nternational Journal of Sports Science & Coaching*.

Martínez-Gallego, R., Crespo, M., Ramón-Llin, J., Micó, S. & Guzmán, J.F. (2020). Men's doubles profesional tennis on hard courts: Game structure and point ending characteristics. *Journal of Human Sport and Exercise*, *15*(3), 633-642.

Martínez-Gallego, R., Ramón-Llin, J. & Crespo, M. (2021c). "A Cluster Analysis Approach to Profile Men and Women's Volley Positions in Professional Tennis Matches (Doubles)" *Sustainability* 13, no. 11: 6370.

Martínez-Gallego, R., Salvador, S. M., Luján, J. F. G., Reid, M., Ramón-Llin, J., & Crespo, M. (2021b). Challenging serve myths in doubles tennis. *International Journal of Sports Science & Coaching*, *16*(6), 1305-1311.

Martínez-Gallego, R., Vives, F., Guzmán, J.F., Ramón-Llin, J. & Crespo, M. (2021d). Time structure in men's professional doubles tennis: does team experience allow finishing the point faster. *International Journal of Performance Analysis in Sport*, *21*(2), 215-225.

Mecheri, S., Rioult, F., Mantel, B., Kauffmann, F. & Benguigui, N. (2016). The serve impact in tennis: First large-scale study of big Hawk-Eye data. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *9*(5), 310-325.

Morgulev, E., Azar, O. H., & Lidor, R. (2018). Sports analytics and the big-data era. *International Journal of Data Science and Analytics*, *5*, 213-222.

O'Donoghue, P. y Brown, M. (2008). The Importance of Service in Grand Slam Singles Tennis. *International Journal of Performance Analysis in Sport, 8*(3), 70-78.

Sakurai, S., Reid, M. & Elliot, B. (2013). Ball spin in the tennis serve: spin rate and axis of rotation. *Sports Biomechanics*. *12*(1), 23-29.

Sánchez-Pay, A., Giménez-Cárceles, J. M., & Sánchez-Alcaraz, B. J. (2019). Diferencias en la distribución espacial del saque en función del rendimiento y tipo de superficie. *Kronos*, *18*(2), 1-13.

Shimizu, T., Hachiuma, R., Saito, H., Yoshikawa, T., & Lee, C. (2019, October). Prediction of future shot direction using pose and position of tennis player. In *Proceedings Proceedings of the 2nd International Workshop on Multimedia Content Analysis in Sports* (pp. 59-66).

Singh Bal, B. & Dureja, G. (2012). Hawk-Eye: A logical innovative technology use in sports for effective decision making. *Sport Science Review*, *21*(1-2), 107-119.

RFET, Real Federación Española de Tenis. (s. f.). *Historia*. RFET, Real Federación Española de Tenis. https://www.rfet.es/es/copa-davis-historia.html

Reid, M., Morgan, S. & Whiteside, D. (2016). Matchplay chracteristics of Grand Slam tennis: implications for training and conditioning. *Journal of Sports Sciencies*, *34*(19), 1791-1798.

Torres-Luque, G., Sánchez-Pay, A., Fernández-García, A. & Palao, J. (2014). Características de la estructura temporal en tenis. *Journal of Sport and Health Research*; 6(2), 117-128.

Vaverka, F., Nykodym, J., Hendl, J., Zhanel, J. & Zahradnik, D. (2018). Association between serve speed and court surface in tennis. *Internacional Journal Performance Analysis in Sport*, *18*(2), 262-272.

Vives, F., Crespo, M., Guzmán, J. F., & Martínez-Gallego, R. (2022). Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology. *International Journal of Performance Analysis in Sport*, *22*(4), 638-648.

Vives, F., Lázaro, J., Guzmán, J. F., Martínez-Gallego, R., & Crespo, M. (2023). Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve. *Applied Sciences*, *13*(24), 13213.

Wei, X., Lucey, P., Morgan, S., & Sridharan, S. (2013). Predicting shot locations in tennis using spatiotemporal data. *(DICTA) International Conference on*, 1–8.

Wei, X., Lucey, P., Morgan, S., Carr, P., Reid, M & Sridharan, S. (2015). Predicting Serves in Tennis using Style Priors. *KDD '15: The 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* 2207-2215.

Whiteside, D. & Reid. M. (2017). Spatial characteristics of profesional tennis serves with implications for serving aces: A machine learning approach. *Journal of Sports Sciences*, *35*(7), 648-65.

3. OBJETIVOS

Los objetivos de este trabajo fueron los siguientes:

3.1. OBJETIVOS DEL ESTUDIO 1

Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology.

Objetivo general nª1:

Describir el primer servicio en dobles profesionales masculino.

Objetivo específico nº1.1:

Analizar las características espacio-temporales del primer servicio.

Objetivo específico nº1.2:

Analizar la efectividad del primer servicio en función de su dirección.

Objetivo general nº2:

Describir el segundo servicio en dobles profesionales masculino.

Objetivo específico nº2.1:

Analizar las características espacio-temporales del segundo servicio.

Objetivo específico nº 2.2:

Analizar la efectividad del segundo servicio en función de su dirección.

3.2. OBJETIVOS DEL ESTUDIO 2

Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve.

Objetivo general nª1:

Descubrir las variables más relevantes del servicio en dobles profesionales masculino.

Objetivo específico nº1.1:

Cálculo de las variables más relevantes para lograr un *ace.*

Objetivo específico nº1.2:

Generación de un conjunto de datos sustituyendo los valores originales por los valores recomendados de la variable en cuestión y predicción de la efectividad.

3.3. OBJETIVOS DEL ESTUDIO 3

Serve Strategy in Professional Women's Doubles Tennis - A Data-Driven Perspective.

Objetivo general nª1:

Descubrir las variables más relevantes del servicio en dobles profesionales femenino.

Objetivo específico nº1.1:

Cálculo de las variables más relevantes para lograr un *ace*.

Objetivo específico nº1.2:

Generación de un conjunto de datos sustituyendo los valores originales por los valores recomendados de la variable en cuestión y predicción de la efectividad.

4. ESTUDIO 1: *EFFECTIVE SERVING STRATEGIES IN MEN'S* DOUBLES *DAVIS CUP MATCHES: AN ANALYSIS USING TRACKING TECHNOLOGY.*

Vives, F., Crespo, M., Guzmán, J. F., & Martínez-Gallego, R. (2022). Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology. *International Journal of Performance Analysis in Sport*, *22*(4), 638-648.

4.1. ABSTRACT AND KEYWORDS

Despite the importance of the serve in doubles tennis, there is very little information on the tactical characteristics of this game situation. The aim of this study was to analyse the differences in the spatial-temporal characteristics and the effectiveness of the serves in professional doubles tennis. Using the Hawk-Eye system, 20419 serves from 96 Davis Cup doubles matches played between 2010 and 2019 were analysed. First serves to the Wide and T zones showed the highest effectiveness. The fastest serves were those played to the T, while the slowest were those played to the Wide zone, except for the second serves on the Deuce side, where the opposite occurred. The serves played to the T were hit from positions closer to the centre of the court, and the time from serve contact to ball bounce was shorter than in the rest of the serves played to other directions. The data provided in this study have clear practical implications, both for the design of specific training sessions and for the tactical preparation of matches.

Keywords: Service; tennis; Hawk-Eye; tactics; racquet sports

4.2. INTRODUCTION

The serve is the stroke that starts the point in tennis and has always been considered one of the most important strokes in the game. This has made it the most studied stroke in tennis, with many studies showing its relationship with performance and ranking (Klaassen & Magnus, 2014; O'Donoghue & Ingram, 2001; Reid et al., 2010). In addition, from a strategic and tactics perspective, the serve has been related to different variables that affect the tactics used by players when hitting this stroke, such as the surface (Vaverka & Cernosek, 2016), the importance of the point (Knight & O'Donoghue, 2012; Meffert et al., 2018), and the gender (Reid et al., 2016) or the playing level of the players (Kovalchik & Reid, 2017). Despite the considerable influence of all these factors on the tactics used by the players, previous studies have shown that, in general, the most used serve tactics by players is to take more risks with the first serve and to play more consistently with the second serve (Cross & Pollard, 2011; Martínez-Gallego et al., 2021).

Regarding the doubles discipline, it is worth highlighting its relevance in professional tennis, since it is played in the Davis Cup and Billie Jean King Cup, in all four Grand Slams, as well as in all ATP and WTA circuits tournaments, and even in exhibition events such as the Laver Cup. In fact, with the change of format of the Davis Cup, it has now become the decisive final point of the tie in case of a draw in the singles matches (Martínez-Gallego et al., 2020). Despite this, as compared to the singles game, substantially fewer studies have focused on this discipline. Previous studies that have analysed doubles matches have mainly focused on the analysis of match structure, such as Martínez-Gallego et al. (2020) who studied the game structure and the point completion in men's doubles, Borderias et al. (2021) who compared the activity profile of women's doubles matches on different surfaces, Martínez-Gallego et al. (2021) who evaluated the influence of partner experience and match type on the time structure of matches, or Dell'Edera et al. (2018) who compared certain performance indicators between junior and professional doubles. In relation to the role of the serve in doubles matches, only a few studies have been found. Two of them compared the effectiveness of the serve between male and female teams (Carboch & Kocǐb, 2016; Martínez-Gallego et al., 2021). Carboch and Kocǐb (2016) found that male teams won 10% more serves than female teams and that they needed more points to the serve to win the match. In addition, they showed that, on hard courts the serve was a greater advantage for both male and female teams, while the effectiveness of the serve on clay court was lower for both genders. Martínez-Gallego et al. (2021) indicated that with the first serve both male and female teams tried to move the receiver serving to the sides to take the initiative of the point, while with the second serve they took lower risks and mainly served to the body. Furthermore, they concluded that male partners were more effective than male partners, especially with the first serve. On the other hand, Kocib et al. (2020) analysed the formations used by the teams. As per the serving team, the most used formations were the classic and the "I" formation. The two most used formations in the other teams were: both players at the back of the court or the back/back combination, highlighting the use of the back/back formation in the second serves compared to the formation of both players at the back of the court.

Although the information provided by these studies is very relevant, it seems necessary to further investigate the characteristics of the serve to understand the aspects that determine its effectiveness in the doubles game. In this context, advances in sports technology are allowing scientists as well as player support teams to record large amounts of match data that provide sound data-based insights allowing deeper understanding of the key performance indicators of the different disciplines. In tennis, one of the most important advances in the evaluation of the technical and tactical performance of tennis players has been the use of tracking systems and, specifically, the analysis of data provided by the Hawk-Eye system (Martínez- Gallego et al., 2021). The high level of precision and the large amount of data recorded through this system offer a wide range of opportunities for research and practice (Mecheri et al., 2016). Specifically, this system allows the tracking of coordi- nates of the ball and the players during the points, so that variables that are of interest for analysing the performance of players in specific actions can be studied with greater depth and precision.

In singles tennis, some studies have used the Hawk-Eye system to analyse and demonstrate the importance of some spatiotemporal variables related to the serve, such as serve speed and serve location (Mecheri et al., 2016; Whiteside & Reid, 2017). However, other variables such as the hitting position, the ball trajectory time, or the difference in speed before and after the bounce have not been studied in depth. As mentioned above, these variables have been even less studied in doubles, in which only one study use this system to analyse the serve location. Therefore, the aim of this study was to analyse the differences in spatio-temporal characteristics such as speed, server position, time from hit to bounce, loss of speed after bounce and efficiency of first and second serves, depending on the side and its direction in professional men's doubles tennis using data provided by the Hawk-Eye system.

4.3. METHODS

4.3.1. Data collection

Data in the present work were obtained through the Hawk-Eye system (Hawk-Eye Innovations Ltd., 2015) used during the competition. The Hawk-Eye system is calibrated before the competition by means of the actual coordinates of the court where it is installed. It uses a total of 10 cameras recording at a frequency of 50–60 Hz that track the movement of the ball and players in four dimensions during points (Cartesian coordinates and time). Independent tests have validated the accuracy and reliability of the system, reporting an average error of 2.6 mm (Reid et al., 2016). Spatio-temporal variables such as the hitting position, the ball bounce position, the service speeds, and velocities before and after the bounce were calculated from the data provided by the system. Player consent for the use of the data during the competition was obtained as part of the tournament agreement.

4.3.2. Data processing

A total of 29,419 services were analysed from 96 full men's doubles matches played during Davis Cup qualifier ties between 2010 and 2019. The match data were imported into R Statistical Software (R Foundation for Statistical Computing, Vienna, Austria), where the data were processed to obtain the variables of interest and subsequently filtered, and outliers were removed. Means were calculated for each variable by match, obtaining a total of 96 records for each variable, each corresponding to one of the matches.

4.3.3. Data analysis

The variables analysed were the following:

Independent variables

- Type: first or second serve.
- Court side: side of the court from which the service occurs. Deuce (right) or Advantage (left) side.
- Direction: direction of the service (Figure 1).

Dependent variables

- Speed: average speed of the service.
- Position: position of the server when serving.

- Time: time from the impact to the bounce of the ball.
- Loss of speed: loss of speed of the ball after the bounce.
- Effectiveness: point won by the serving team without the serve having been returned (Martínez-Gallego et al., 2021).

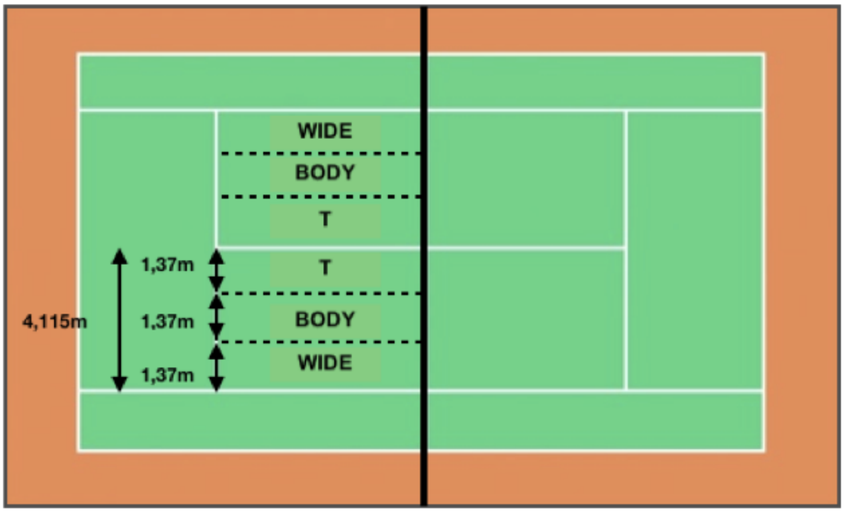

Figure 1. Serve directions.

Firstly, a descriptive analysis was carried out and mean and standard deviation were calculated. Assumptions of normality and homogeneity of variances were verified using the Kolmogorov-Smirnov test and the Levene test. The Kruskal – Wallis test and post-hoc pairwise comparison was used to identify differences in the dependent variables, setting the significance level at $p < 0.05$. Effect size was calculated using eta-squared, considering small (0.01–0.06), moderate (0.061–0.14) and large (>0.141) effects (Cohen, 2013).

4.4. RESULTS

4.4.1. Spatio-temporal characteristics of the first serve

Table 1 shows the descriptive statistics of the kinematic variables as a function of the direction of the serves. Regarding the first serve on the deuce side, it could be observed that there were significant differences in the speed between the three directions ($H(2) = 119.14$; $p < 0.01$; $\eta^2 = 0.41$), being the serves played to the T the ones that presented a higher speed (mean = 188.35 km-h^{-1}), followed by those played to the Body zone (mean = 182.17 km-h^{-1}) and, finally, those played to the Wide zone were the slowest serves (mean = 176.41 km-h^{-1}). Regarding the Position, serves played to the T were hit significantly closer to the centre of the court (mean = 1.56 m; H(2) = 13.77; $p < 0.01$; $\eta^2 = 0.04$) than those played to the Body (mean = 1.70 m) and to the Wide zone (mean = 1.70 m). Regarding Time, significant differences were found between the three directions of the serves (H(2) = 178.31; $p < 0.01$; $\eta^2 = 0.62$), with the serves played to the T showing a lower value of Time (mean = 0.38 s), followed by the serves played to the Body zone (mean = 0.40 s) and, finally, those played to the Wide zone were the serves that showed a lower Time (mean = 0.43 s). In the variable Loss of Speed, no differences were found between the serve directions (F = 0.68; $p = 0.71$; $\eta^2 = 0.01$).

Table 1.

Descriptive statistics of the serve kinematic variables in function of the direction

			Wide Mean (SD)	Body Mean (SD)	T Mean (SD)	F	p	η^2
First serve	Deuce side	Speed (km h^{-1})	175. 76 (12.70)$	181. 77 (13.11)$	188. 22 (14.64)$	555.42	<0.01	0.13
		Position (m)	1. 73 (0.62)*	1. 72 (0.61)†	1. 56 (0.61)*†	60.55	<0.01	0.02
		Time (s)	0. 43 (0.04)$	0. 40 (0.03)$	0. 38 (0.03)$	1123.10	<0.01	0.24
		Loss of speed (km h^{-1})	31. 15 (3.17)*	31. 33 (3.21)	31. 50 (3.52)*	7.28	<0.01	0.00
	Advantage side	Speed (km h^{-1})	181. 42 (17.72)$	182. 64 (14.82)$	185. 67 (12.81)$	48.06	<0.01	0.02
		Position (m)	1. 87 (0.72)*	1. 82 (0.70)†	1. 67 (0.73)*†	47.91	<0.01	0.02
		Time (s)	0. 42 (0.04)$	0. 40 (0.04)$	0. 39 (0.03)$	353.53	<0.01	0.10
		Loss of speed (km h^{-1})	31. 33 (3.33)	31. 50 (3.27)	31. 34 (3.50)	1.59	0.2	0.00
Second serve	Deuce side	Speed (km h^{-1})	152. 75 (12.57)$	154. 81 (12.13)$	155. 73 (13.37)$	12.29	<0.01	0.01
		Position (m)	1. 94 (0.68)*	1. 92 (0.68) †	1. 72 (0.60)*†	34.38	<0.01	0.02
		Time (s)	0. 50 (0.04)$	0. 47 (0.04)$	0. 46 (0.04)$	143.54	<0.01	0.08
		Loss of speed (km h^{-1})	29. 88 (3.13)	30. 18 (3.15)	30. 14 (3.15)	2.48	0.08	0.00
	Advantage side	Speed (km h^{-1})	153. 22 (13.44)$	155. 19 (12.87)$	156. 71 (13.49)$	15.142	< 0.01	0.01
		Position (m)	1. 97 (0.64)$	1. 89 (0.66)$	1. 67 (0.63)$	54.278	< 0.01	0.03
		Time (s)	0. 49 (0.04)$	0. 47 (0.04)$	0. 46 (0.04)$	144.49	< 0.01	0.08
		Loss of speed (km h^{-1})	29. 87 (3.19)#*	30. 24 (3.16)#	30. 42 (3.38)*	6.141	<0.01	0.00

SD = Standard deviation; $ Significant differences (p < 0.05) between W, B and T; * Significant differences (p < 0.05) between W and T; † Significant differences (p < 0.05) between B and T; # Significant differences (p < 0.05) between W and B.

Regarding the first serve on the advantage side, it could be observed that there were significant differences in speed between the three directions (H(2) = 21.25; $p < 0.01$; $\eta^2 = 0.07$), being the serves played to the T the ones with the highest speed (mean = 186.25 km- h^{-1}), followed by those played to the Body zone (mean = 183.67 km-h^{-1}) and, finally, those played to the Wide zone were the slowest serves (mean=181.65km-h^{-1}). Regarding Position, the serves played to the T were hit significantly closer to the centre of the court (mean = 1.67 m; H(2) = 12.35; $p < 0.01$; $\eta^2 = 0.04$) than those played to the

Body (mean = 1.82 m) and to the Wide zone (mean = 1.85 m). Regarding Time, significant differences were found between the three directions of the serves ($H(2) = 111.73$; $p < 0.01$; $\eta^2 = 0.39$), with the serves played to the T the ones showing a lower value of Time (mean = 0.39 s), followed by the serves played to the Body zone (mean = 0.40 s) and, finally, those played to the Wide zone were the serves that showed a lower Time (mean = 0.42 s). In the Loss of Speed variable, no significant differences were observed between the three directions of the serves.

4.4.2. Spatio-temporal characteristics of the second serve

As per the second serve on the deuce side, the serves played to the Wide showed a higher speed (mean = 158.98 km-h^{-1}; $H(2) = 20.12$; $p < 0.01$; $\eta^2 = 0.06$) than those played to the Body (mean = 156.16 km-h^{-1}) and to the T (mean = 154.70 km-h^{-1}). Regarding Position, no significant differences were observed between the three directions of the serves. Regarding Time, significant differences were found between the serves played to the Wide and T zones ($H(2) = 9.66$; $p < 0.01$; $\eta^2 = 0.03$), being the serves played to the T the ones that presented a lower value of Time (mean = 0.47 s) as compared to the serves played to the T (mean = 0.48). In the Loss of Speed variable, no significant differences were observed between the three directions.

Finally, significant differences in Speed between the three directions ($H(2) = 78.43$; $p < 0.01$; $\eta^2 = 0.27$) can be observed in the second serve on the advantage side, with the serves played to the T showing the highest speed (mean = 162.15 km-h^{-1}), followed by those played to the Body zone (mean = 154.35 km-h^{-1}) and, finally, those played to the Wide zone were the slowest serves (mean = 151.40 km-h^{-1}). Regarding the Position, serves played to the T, were hit significantly closer to the centre of the court (mean = 1. 90 m; $H(2) = 54.28$; $p < 0.01$; $\eta^2 = 0.03$) than those played Wide (mean = 2.07 m) and to the Body (mean = 2.10 m). As per Time, significant differences were found between the three directions of the serves ($H(2) = 145.08$; $p < 0.01$; $\eta^2 = 0.51$), with the serves played to the T showing a lower value of Time (mean = 0.45 s), followed by the serves played to the Body zone (mean = 0.48 s) and, finally, those played to the Wide zone were the serves that presented a higher Time (mean = 0.50 s). In the variable Loss of Speed, serves played to the T lost more speed (mean = 31.06; $H(2) = 13.41$; $p < 0.01$; $\eta^2 = 0.04$)

than serves played to the Body zone (mean = 30.22) and serves played to the Wide zone (mean = 29.82).

4.4.3. Effectiveness of the first serve depending on direction

Figure 2 shows the effectiveness of the first serve on the deuce and the advantage sides, for each of the directions. Regarding the first serves on the Deuce side, there was a lower effectiveness when the serves were played to the Body (mean = 29.95 %; H(2) = 34.72; $p < 0.01$; $\eta^2 = 0.12$), compared to the serves played to the T (mean = 38.45 %) and those played to the Wide zone (mean = 40.71 %).

Similarly, on the advantage side, there was a lower effectiveness of serves played to the Body (mean = 28.27 %; H(2) = 40.16; $p < 0.01$; $\eta^2 = 0.13$), with respect to those played to the T (39.28 %) and to the Wide zones (37.54 %).

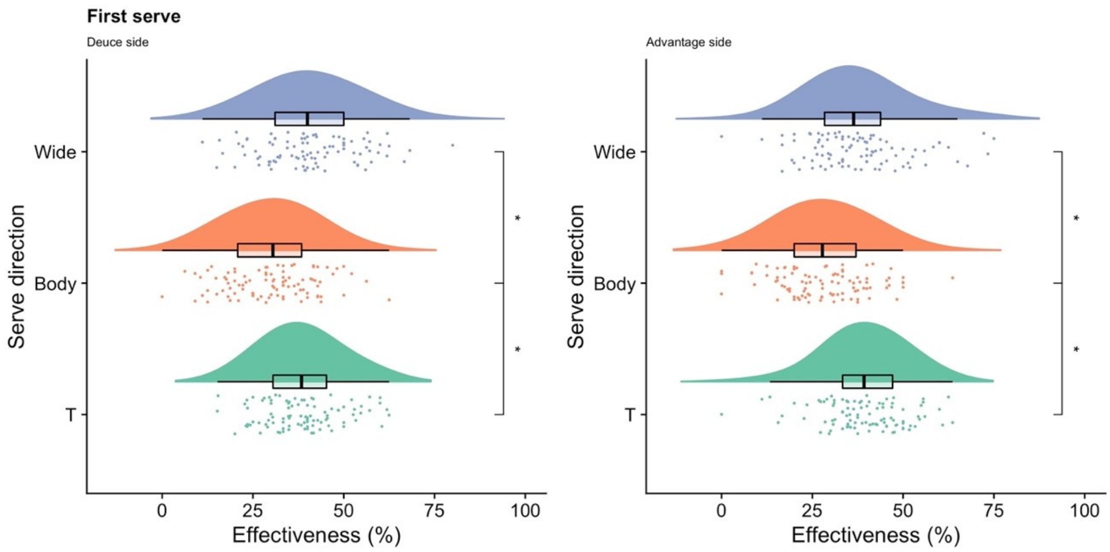

Figure 2. Differences in effectiveness of the first serve.

4.4.4. Effectiveness of the second serve depending on direction

Figure 3 shows the effectiveness of the second serve on the deuce and the advantage sides, for each of the directions. As per the second serves on the deuce side, there was a higher effectiveness of the serves played to the Wide zone (mean = 30.74 %; H(2) = 8.24; $p < 0.05$; $\eta^2 = 0.02$), with respect to those played to the T (21.14 %). On the advantage side, no significant differences were found (H(2) = 4.28; $p = 0.12$; $\eta^2 = 0.01$).

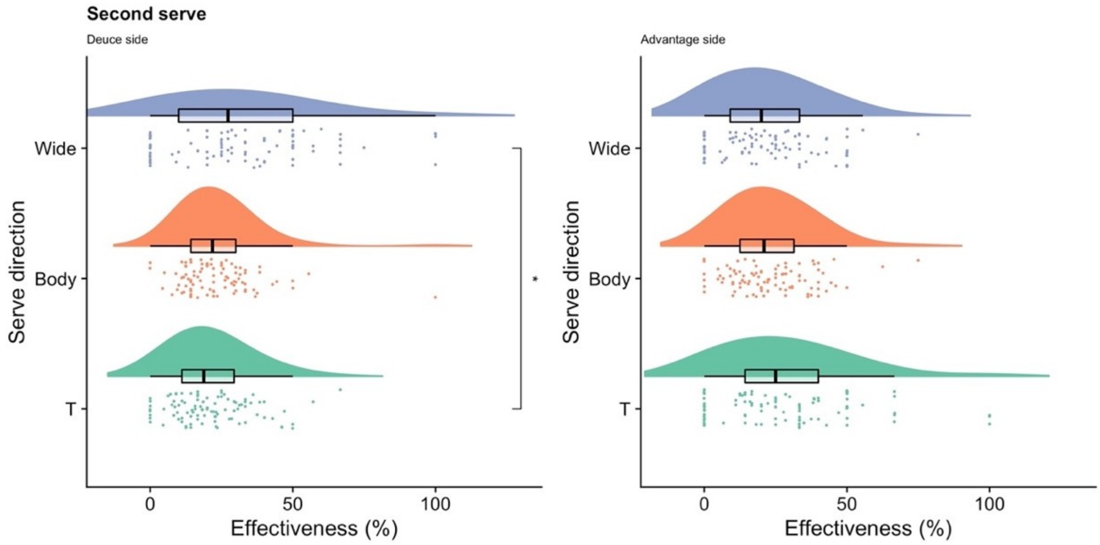

Figure 4.3. Differences in effectiveness of the second serve.

4.5. DISCUSSION

The results found have allowed to observe how the speed of the serve varied for each of the directions, both on the first and second serves on both sides of the court. In all cases, the fastest serves were those played to the T, while the slowest ones were those played to the Wide zone, except for the second serves on the Deuce side, where the opposite occurred. The serves to these two zones have a clear tactical intention, which is to try to place the ball as far as possible from the receiver's zone to make the return more difficult, or to prevent the receiver from taking the initiative of the point (Martínez-Gallego et al., 2021). However, the fact that the speed of the serves to the Wide zone is much lower is directly related to the type of spin used to hit the serve. To serve to this zone, players tend to use a slice serve, which opens up a greater angle but reduces the speed of the ball (Gillet et al., 2009). In the case of second serves next to the Deuce side, it seems that players increase the speed of the serve in order to surprise the opponent. Previous studies have shown that this is the least used direction for second serves in doubles (Martínez-Gallego et al., 2021), so serving to this area with a higher speed can be a good strategy that, as seen in the results, increase the effectiveness of second serves. As per the position of the servers, it could be observed that their position tended to be closer to the centre when they played the serves to the T zone. Previous studies have shown the importance of the angle of the trajectory of the ball with respect to the receiver's position (Whiteside & Reid, 2017), therefore, it seemed that players opted for these positions closer to the centre of the court to obtain angles that placed the ball away from the receiver's position when serving to this area. However, previous studies that have analysed the position of servers in the singles game have shown values considerably closer to the centre of the court (Whiteside & Reid, 2017). This may be because the size of the court in doubles is larger than in singles, which forces players to stand further away from the centre of the court to be able to correctly cover the furthest areas of the court for the next shot. This fact may also explain the reason why when servers play second serves to the advantage side, they stand significantly further away from the centre than when they serve to the body and the T. On the one hand, to hit the serve with a greater angle, and on the other hand, to have a better position to cover the court in case the receiver can attack the serve.

As for the time from ball hit to bounce, in all situations, serves played to the T zone showed the shortest time, and those played to the Wide zone showed the longest time. These results seem quite logical considering that the distance travelled by the ball is greater as the bounce occurs further away from the centre of the court (Mecheri et al., 2016). In addition, as noted above, the speed of the serves played to the *T*-zone is higher, which also contributes to this reduction of time.

With respect to the loss of speed of the ball after the bounce, differences have only been found in the second serves on the advantage side, where the serves played to the T zone were the ones that lost the most speed. These differences in speed loss are substantially smaller than the differences in service speed, so this variable might have a low influence on service performance compared to service speed, which has been shown to be a determinant variable (Vaverka et al., 2018).

Regarding the effectiveness of the first serves, on both sides, serves played to the Wide and T zones were the most effective. These data coincide with those obtained in studies that have analysed service effectiveness, both in singles (Mecheri et al., 2016; Whiteside & Reid, 2017) and doubles (Martínez-Gallego et al., 2021). Therefore, as in the singles game, playing the serves to the areas furthest away from the receiver's position seems to be the best strategy to increase the effectiveness of the serve. With respect to the second serves, differences were only found on the deuce side, where serves played to the Wide zone were more effective than those played to the T zone. As indicated above, this could be due to the fact that players serve a lower percentage of serves to the Wide zone (Martínez- Gallego et al., 2021) and with a higher speed, so that when they do so, they surprise the opponent and win a greater number of points.

As it can be observed, the information provided in this study has clear practical implications, both for the training and for the strategical and tactical preparation of matches. When relating the results on serve effectiveness with the rest of the variables studied in the doubles game, it can be noted that when players serve Wide, they try to obtain an advantage by moving the receiver further away from the court, serving at a greater angle and using the slice and topspin effects to a greater extent, except on second serves to the deuce side, where they try to surprise the opponent by increasing the speed.

On the other hand, when they serve to the T zone, they try to reduce the receiver's response time by positioning themselves closer to the centre of the court, serving at a higher speed and mainly using a flat serve. Knowing this data, coaches can be more specific when designing and conducting the tasks and drills when working the service in doubles training sessions. In addition, the application of this information to the characteristics of the players themselves and of the opposing teams can help to better prepare the tactics to be used with the service in doubles matches. Despite the new findings provided in our research, this study has also certain limitations. Firstly, the analysis of the serve has been carried out in isolation, without considering its relationship with subsequent actions. This could be an important aspect to consider for further studies, as effective serves are not only those in which the serving team wins the point directly with the serve, but also those in which the serve provides an advantage that allows the point to be won in the following shots. In addition, it could also be interesting to know how the serve influences the characteristics and performance of the return. Finally, it could be of interest to apply this type of analysis not only to professional men's doubles, but also to professional women's doubles, as well as to different categories such as collegiate, junior, seniors, and recreational players.

4.6. CONCLUSIONS

This article analyses the differences in the spatial-temporal characteristics of serves in men's professional doubles tennis according to their direction. The results found showed that the fastest serves were those played to the T, while the slowest were those played to the Wide zone. The serves played to the T were hit from positions closer to the centre of the court, and the time from the impact to the bounce of the ball was shorter than in the rest of the directions. The first serves played to the Wide and T zones were the most effective, with no differences being found in the second serves. These data suggest that when the players serve Wide, they try to obtain an advantage by moving the receiver to positions further away from the court. On the other hand, when they serve to the T zone, they try to reduce the receiver’s response time, positioning themselves closer to the centre of the court and serving at a higher speed.

4.7. REFERENCES

Borderias, M., Fernandez-Fernandez, J., Martinez-Gallego, R., & Baiget, E. (2021). Comparison of the activity profile during grand slam women's doubles tennis. *International Journal of Performance Analysis in Sport*, *21*(6), 1066–1080. https://doi.org/10.1080/24748668.2021. 1972611

Carboch, J., & K o cˇi b , T. (2016). A comparison of service efficiency between players of male and female doubles at professional tennis tournaments. *Auc Kinanthropologica*, *51*(2), 56–62. https://doi.org/10.14712/23366052.2015.32

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Cross, R., & Pollard, G. (2011). Grand slam men's singles tennis 1995-2009 part 2: Points, games and sets. *ITF Coaching and Sport Science Review*, *2011*(53), 3–6.

Dell'Edera, M., Bertino, L., & Campagnoli, D. (2018). Junior doubles development: An on-court progressive approach based on facts and stats. *ITF Coaching & Sport Science Review*, *26*(75), 4–6. https://doi.org/10.52383/itfcoaching.v26i75.244

Gillet, E., Leroy, D., Thouvarecq, R., & Stein, J. (2009). A notational analysis of elite tennis serve and serve-return strategies on slow surface. *Journal of Strength and Conditioning Research*, *23* (2), 532–539. https://doi.org/10.1519/JSC.0b013e31818efe29

Hawk-eye innovations. (2015). *Hawk-Eye's accuracy & reliability: Electronic line calling*. Basingstoke.

Klaassen, F., & Magnus, J. R. (2014). *Analyzing wimbledon: The power of statistics*. Oxford University Press.

Knight, G., & O'Donoghue, P. (2012). The probability of winning break points in grand slam men's singles tennis. *European Journal of Sport Science*, *12*(6), 462–468. https://doi.org/10.1080/ 17461391.2011.577239

Kocib, T., Carboch, J., Cabela, M., & Kresta, J. (2020). Tactics in tennis doubles: Analysis of the formations used by the serving and receiving teams. *International Journal of Physical Education, Fitness and Sports*, *9*(2), 45–50. https://doi.org/10.34256/ijpefs2026

Kovalchik, S. A., & Reid, M. (2017). Comparing matchplay characteristics and physical demands of junior and professional tennis athletes in the era of big data. *Journal of Sports Science & Medicine*, *16*(4), 489–497.

Martínez-Gallego, R., Crespo, M., & Jiménez, J. (2021). Analysis of the differences in serve effectiveness between Billie Jean King cup (former fed cup) and Davis cup doubles tennis matches. *International Journal of Sports Science & Coaching*, *16*(3), 777–783. https://doi.org/ 10.1177/1747954120988545

Martínez-Gallego, R., Crespo, M., Ramón-Llin, J., Micó, S., & Guzmán, J. F. (2020). Men's doubles professional tennis on hard courts: Game structure and point ending characteristics. *Journal of Human Sport and Exercise*, *15*(3), 633–642. https://doi.org/10.14198/jhse.2020.153.13

Martínez-Gallego, R., Ramón-Llin, J., & Crespo, M. (2021). A cluster analysis approach to profile men and women's volley positions in professional tennis matches (doubles). *Sustainability*, *13* (11), 6370. https://doi.org/10.3390/su13116370

Martínez-Gallego, R., Vives, F., Guzmán, J. F., Ramón-Llin, J., & Crespo, M. (2021). Time structure in men's professional doubles tennis: Does team experience allow finishing the points faster? *International Journal of Performance Analysis in Sport*, *21*(2), 215–225. https://doi.org/ 10.1080/24748668.2021.1872218

Mecheri, S., Rioult, F., Mantel, B., Kauffmann, F., & Benguigui, N. (2016). The serve impact in tennis: First large-scale study of big hawk-eye data. *Statistical Analysis and Data Mining*, *9*(5), 310–325. https://doi.org/10.1002/sam.11316

Meffert, D., O'Shannessy, C., Born, P., Grambow, R., & Vogt, T. (2018). Tennis serve performances at break points: Approaching practice patterns for coaching. *European Journal of Sport Science*, *18*(8), 1151–1157. https://doi.org/10.1080/17461391.2018.1490821

O'Donoghue, P., & Ingram, B. (2001). A notational analysis of elite tennis strategy. *Journal of Sports Sciences*, *19*(2), 107–115. https://doi.org/10.1080/026404101300036299

Reid, M., McMurtrie, D., & Crespo, M. (2010). Title: The relationship between match statistics and top 100 ranking in professional men's tennis. *International Journal of Performance Analysis in Sport*, *10*(2), 131–138. https://doi.org/10.1080/24748668.2010.11868509

Reid, M., Morgan, S., & Whiteside, D. (2016). Matchplay characteristics of grand slam tennis: Implications for training and conditioning. *Journal of Sports Sciences*, *34*(19), 1791–1798. https://doi.org/10.1080/02640414.2016.1139161

Vaverka, F., & Cernosek, M. (2016). Quantitative assessment of the serve speed in tennis. *Sports Biomechanics*, *15*(1), 48–60. https://doi.org/10.1080/14763141.2015.1123763

Vaverka, F., Nykodym, J., Hendl, J., Zhanel, J., & Zahradnik, D. (2018). Association between serve speed and court surface in tennis. *International Journal of Performance Analysis in Sport*, *18*(2), 262–272. https://doi.org/10.1080/24748668.2018.1467995

Whiteside, D., & Reid, M. (2017). Spatial characteristics of professional tennis serves with implications for serving aces: A machine learning approach. *Journal of Sports Sciences*,*35*(7), 648–654. https://doi.org/10.1080/02640414.2016.1183805

5. ESTUDIO 2: *OPTIMIZING SPORTING ACTIONS EFFECTIVENESS: A MACHINE LEARNING APPROACH TO UNCOVER KEY VARIABLES IN THE MEN'S PROFESSIONAL DOUBLES TENNIS SERVE.*

Vives, F., Lázaro, J., Guzmán, J. F., Martínez-Gallego, R., & Crespo, M. (2023). Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve. *Applied Sciences*, *13*(24), 13213.

5.1. ABSTRACT & KEYWORDS

This study used a novel machine learning approach to uncover key serve variables that maximize effectiveness in men's professional doubles tennis. A large dataset of 14,146 serves from 97 Davis Cup doubles matches played between 2010 and 2019 was analyzed using explainable AI techniques. The angle and distance from the bounce to the sidelines of the serves were found to best distinguish the points won with aces from rallies lasting more than three strokes. Optimal serve angle ranges of 5.7–8.7° substantially increased the probability of serving an ace by over 80%, compared to around 30% when serving used more central angles. Lateral bounce distances of 0–28 cm from the sidelines also boosted the ace probability by over 50%. The serve speed was shown to have less influence on serve effectiveness as compared to singles tennis, with velocities above 187 km h^{-1} only increasing the probability of serving an ace by 10%. These findings have important practical implications for the tactical decision-making and technical training of serves in men's professional doubles tennis. The data highlight that the angle and placement of serves are more important than velocity for attaining effective serves in doubles. Coaches and players can use this knowledge to pay special attention to the most important variables in the effectiveness of serves, such as the line distance and angle, in order to maximize the performance of the doubles serve. The novel methodology used in this study provides a valid and reliable way to calculate the efficiency of actions in various sport disciplines using tracking data and machine learning approaches.

Keywords: racquet sports; tactics; performance analysis; sport analytics; tracking technology

5.2. INTRODUCTION

The serve has become the most decisive stroke in modern professional tennis, particularly in the doubles game (Martínez-Gallego et al., 2020). It has been shown to be a key shot for controlling the game and even winning points directly through aces, thus being a critical technical and tactical skill for success at the professional levels of the game (Kostoglou et al., 2017; Browning, 2009).

In the men's singles game, servers have demonstrated a significant advantage over returners, especially in points lasting fewer than 4 strokes, in which they win approximately 60% of them (O'Donoghue & Brown, 2008). Notably, Reid et al. (2017) suggested that, among the top 100 ATP-ranked players, the most important predictors of success are points won with the second serve and points won when returning the second serve. Furthermore, research has shown that games won when serving and points scored with the first and second serves become particularly relevant for the outcome of the match (Mecheri et al., 2016).

In the doubles game, the characteristics of the discipline highlight a greater impact of serve as compared to singles matches, especially when points are shorter than 4 strokes (Martínez-Gallego et al., 2020). This has been explained by the position of the serving player at the net (Kocib et al., 2020). As it is the case for the singles game, research has concluded that serving teams tend to take more risks with their first serve while seeking greater reliability with their second serve (Martínez-Gallego et al., 2021b). Moreover, studies have shown that factors such as the playing surface and the experience of the doubles team provide further advantages for serving teams at professional level (Carboch & Kocib, 2016; Martínez-Gallego et al., 2021c).

Apart from the advantage of the serve in doubles tennis, serve tactics are influenced by a variety of contextual and situational informations. Variables such as the score and the returner's position play a crucial role in the players' decision-making when serving (Whiteside & Reid, 2017). Effectively managing this contextual information allows players to maximize the performance of their serves, enabling them to increase the variability of serve types, exploit the returner's weaknesses, surprise their opponents, and maintain a high level of consistency (Giampolo & Levey, 2013).

The introduction of Hawk-Eye technology in tennis has provided access to a wealth of data which has been used in various studies to predict game indicators, particularly parameters related to the serve (Wei et al., 2015; Mecheri, 2016). In the singles game, Whiteside and Reid (2017) analyzed direct serves and reported that accuracy-related variables played a more significant role than speed-related ones. In doubles play, Hawk-Eye data has been used to analyze the characteristics of serves based on their direction. The position of the server has been identified as a crucial factor in determining the direction of the first or second serve, making it a variable to consider for the effectiveness of the serve and the anticipation in the return (Vives et al., 2021). Additionally, differences between men’s and women’s doubles tennis in the effectiveness and the directions of the serve have found that men’s teams are more effective, especially when serving to areas far from the center of the serve box than their women counterparts (Martínez-Gallego et al., 2021a).

The abundance of information collected through digital systems has led to a proliferation of scientific articles in tennis that utilize machine learning (ML) for data analysis. ML refers to the scientific study of algorithms and statistical models that enable computer systems to perform specific tasks without explicit programming (Mahesh, 2020). In tennis, from a performance analysis perspective, Mlakar and Luštrek (2017) developed a method to classify different types of strokes, as well as the active and passive playing time. Giles et al. (2020; 2021) conducted two machine learning studies, where the first aimed to automate the identification and classification of change of direction (COD) movements in professional tennis using tracking data. The second study analyzed the temporal and movement characteristics of COD in professional tennis, suggesting that players should develop a variety of COD styles to improve their performance. Specific players have also been subject to studies, such as that of Liu (2022) who examined Roger Federer's technical-tactical performance, and Perri et al. (2022) who classified the strokes and movements of eight high-performance tennis players, with the serve being the most precise stroke, allowing for detailed player tracking throughout training (Perri et al., 2022). Most of the studies analyzing serve through ML have focused on singles play. For instance, Kovalchik et al. (2017) analyzed 25,680 serves to determine the characteristics of the 'perfect' serve in professional men's tennis, identifying up to 13 distinct types of serves. Other studies have developed different prediction methods for various serve directions, considering both the server and the returner actions and positions. These studies emphasized the importance of contextual information in predicting the directions used by the servers (Zhu & Naika, 2022; Brouwer, 2021).

Prediction of match results is an area that has seen numerous investigations through Machine Learning. Sipko and Knottenbelt (2015) used historical player performance to predict match outcomes. Kovalchik (2016) compared three match outcome prediction models: regression, point-based, and pair-wise comparison, with regression models performing best. Yang et al. (2017) adopted a hybrid model combining historical match data and real-time statistics, while Cornman et al. (2017) also used historical match data to predict future match results. Other authors attempted predictions using different approaches, such as SVM, Naive Bayes, Logistic regression, and Random Forest (Sekar, 2019), artificial neural networks (Candila & Palazzo, 2020), or by comparing different models like Support Vector Machine, Random Forest, and Extreme Gradient Boosting for classification (Van Rooij, 2021). Wilkens (2021) concluded that the most efficient option was to use model ensembles that combine predictions from multiple approaches. Moreover, other authors have explored the combination of various approaches, with De Serrano (2021) using open-source data with historical player information or employing a Glicko mathematical classification model, commonly used in professional chess (Yue et al., 2022). Finally, two studies combined individual player statistics and game parameters such as scores to predict match outcomes (Solanki et al., 2022; Gupta et al., 2022). Therefore, it can be concluded that this area of the game has attracted considerable attention from researchers and that the various studies conducted have explored different approaches to the challenge of match outcome prediction in tennis.

Despite the importance of serve in doubles tennis as shown above, to the knowledge of the authors no previous study has utilized Big Data and ML to analyze its effectiveness in men's professional doubles. Therefore, the aim of this study was to explore this topic and create a model which could describe the variables associated with a greater serve effectiveness in Davis Cup doubles matches.

5.3. METHODS

5.3.1. Sample

A total of 14146 serves was analysed from 97 full men's doubles matches played during Davis Cup qualifying ties between 2010 and 2019. Our study featured the participation of 123 teams and 160 players from 34 different countries, with an average age of 30.03 ± 4.73 years.

5.3.2. Instruments

The data utilized in this study were obtained from the Hawk-Eye system (Hawk-Eye Innovations Ltd., 2015) deployed during the Davis Cup events. It consists of a group of ten cameras placed around the circumference of the court, which captures and records the trajectory of the ball and the players' movement on the court (Vives et al., 2022). The Hawk-Eye System is well-known for its high accuracy and reliability in tracking and analyzing tennis matches. It is widely used in professional tennis and has provided the data for numerous research papers (Mecheri et al., 2016; Whiteside, 2017; Martinez-Gallego, 2021a; Vives, 2022).

5.3.3. Procedure

A complete E2E AI pipeline was built to experiment with data processing and Explanatory Artificial Intelligence (XAI):

1. Experimental setup
2. Data processing
3. Training the deep learning model
4. Feature importance algorithms
5. Probabilistic and statistical analysis. Values selection based on maximizing desired effectiveness and minimizing the undesired one.
6. Synthetic data-set generation, simulation based on the predictions on this data-set and analysis of effectiveness improvements.

1. Experimental setup

The XAI process can be summarized in the steps 2 to 6. A pipeline that included all xAI processes from Data Processing to training the Deep Learning Model was built to calculate the feature, select values, and test the results on a synthetic data-set.

The whole process is summarized in Figure 1.

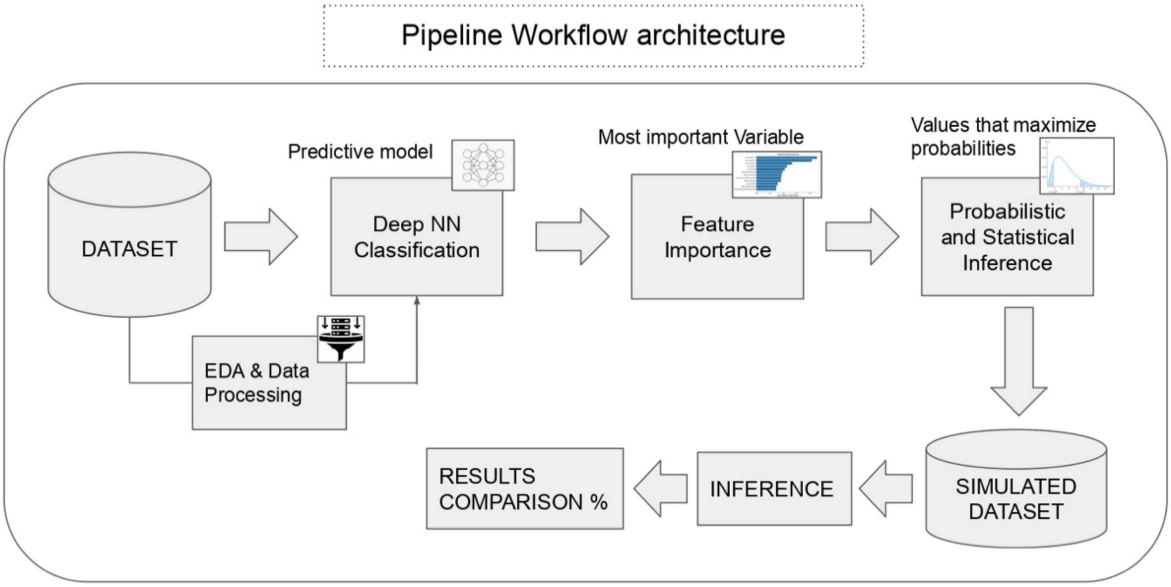

Figure 1. Pipeline workflow architecture

2. Data processing

2.1. Data-set

The data set consisted of the following variables:

<u>Independent variables</u>

- Court side:
  Side of the court from where the serve is played. This variable was used to divide the data-set and perform the analysis for each side of the court.
- Effectiveness:
  - Type 1: The point finishes with one shot.
  - Type 2: The point finishes with two shots.
  - Type 3: The point finishes with three shots.
  - Type 4: The point finishes with four shots.

Dependent variables

- Speed: Mean speed of the serve.
- Position: Position of the server when hitting the serve (A in Figure 2).
- Time: Time between ball impact and ball bounce.
- Speed loss: Loss of speed of the ball after its bounce.
- Impact Z: Height of the ball at impact.
- Net clearance: Height of the ball when passing over the net.
- δ: Serve angle (Figure 2).
- β: Vertical projection angle.
- dL: Distance from ball bounce to the sideline of the service box.

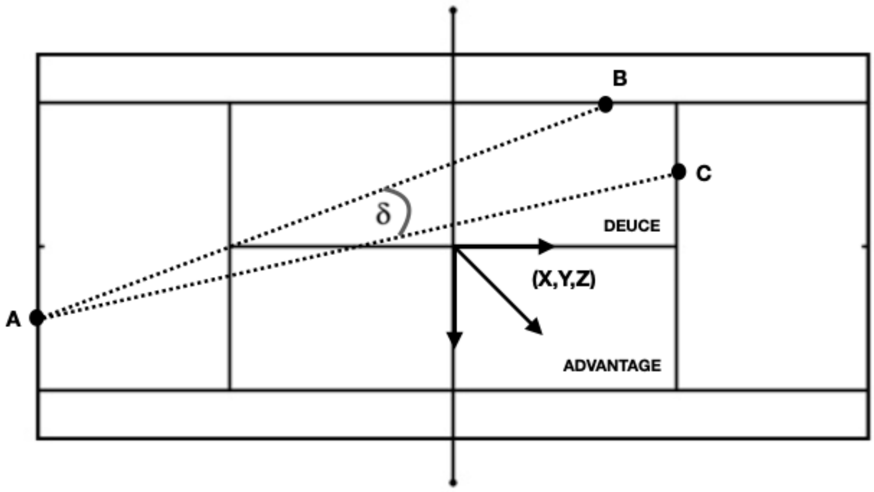

Figure 2. Angle of the serve.

2.2. EDA and Data Processing

After some Exploratory Data Analysis (correlation matrices, boxplots, histograms, Kernel Density Estimation plots - KDE, etc.), the data was prepared to apply machine learning classification algorithms. Some variables were not informative as their variance was equal or close to 0. The main data processing steps were the following:

1. Oversampling the weak target samples to solve the data-set "class unbalancing" problems.

2. Feature selection: This was done using a combination of tennis knowledge criteria together with correlation analysis and feature selection algorithms.

3. Outliers deletion and filtering: Outliers were detected based on statistic methods and were filtered afterwards.

Furthermore, some features were not following a simple normal distribution but a mixture of gaussians. The collinearity issues between variables, which didn't affect the model performance, but affected the interpretability of the results, were also solved. Thus, Feature importance was applied only with uncorrelated features. This process was done separately for δ and dL.

3. Training the Deep Learning Model

AutoGluon and FastAI were utilized in this study for this purpose. AutoGluon presents commendable tools for explanation and automated feature enhancement. However, it exhibited a lack of the flexibility required for this research as certain outcomes were inconsistent. Consequently, it was decided to employ FastAI for the analysis of tabular data.

The implementation of AutoGluon was straightforward, as this tool undertakes the preprocessing of the dataset, trains multiple models, selects the optimal one, assesses feature importance, and employs various XAI algorithms to deliver global explanations for the models.

On the other hand, FastAI is a tool that offers functionalities complimentary to those of PyTorch and AutoGluon. It provides certain data processing functionalities while facilitating the fine-tuning training process. Although it demands a greater degree of software development, it offers advantageous high-level flexibility options. These include the ability to select the number of hidden layers, determine the loss function, and automate the training process, commonly known as "babysitting", through the application of beneficial callbacks such as a state-of-the-art learning rate finder, early stopping, and model checkpointing.

Deep Learning model backbone Architecture and loss function included the following steps are the following:

Backbone Network: A tabular model with 5 layers of feature size [256, 128, 128, 128, 64]. The basic layer block (named LinBnDrop) is formed by the next transformations:

1. Linear Layer (torch.nn.linear)
2. Reactified Linear Unit – ReLU (torch.nn.relu)
3. Batch Normalization - BatchNorm1d (torch.nn. BatchNorm1d)

Loss Function: Focal Loss Flat (Lin et al., 2018). The focal loss works specially well with imbalanced data as it adapts its weights to focus learning on hard misclassified examples. Optimization Algorithm: Adam (Kingma & Ba, 2015).

4. *Feature Importance algorithms for feature explanation*

Feature importance algorithms were used to select the variables that have a higher impact on the outcome of the models. Feature Permutation Importance and SHAP Summary Plots are the two algorithms commonly accepted in the scientific community (Altmann et al., 2010; Lundberg & Lee, 2017).

Feature permutation importance (Altmann et al., 2010) is based on repeated permutations of the outcome vector to estimate the distribution of measured importance for each variable in a non-informative setting.

SHAP assigns each feature an importance value for a particular prediction (Lundberg & Lee, 2017). SHAP values attribute to each feature the change in the expected model prediction when conditioning on that feature.

5. *Probabilistic and Statistical Analysis. Values Selection based on Maximizing desired Effectiveness and Minimizing the undesired one*

Shapley values and its variants are adequate “Local Model-Agnostic methods”, which provide interesting insights of a single prediction. Global Explanation methods can help understand the model through visualizations and explanations of the final model weights (Molnar, 2022). As the aim of the study was to understand how to maximize the serve effectiveness (how and why specific values of the selected features affect the predicted target variables), the goal was to calculate which values threshold of the selected features maximize the likelihood of Effectiveness type 1 and minimize the likelihood of the Effectiveness type 4. For this purpose, a novel semiautomated algorithm based on classical statistics was developed. The first steps of this process were the following:

1. First, a smooth probability distribution of the data using the Kernel Density Estimation algorithm (KDE), together with statistical estimators was calculated for each selected feature. The estimators were mainly the Mean, the MLE, and percentiles 5 and 95, for each type of Effectiveness.
2. Then, the values were mostly identified by selecting the area of the plot which combined a higher density of the points with the desired result (type1) and a lower density of the undesired result (type4). The statistical estimators also helped to understand the probability distribution.

6. ***Synthetic Data-set Generation, simulation based on the Predictions on the data-set and Analysis of Effectiveness Improvements.***

In the final step a synthetic dataset using the selected values was generated, and the prediction was simulated. Since the model reached 93% f1-score, the predictions were reliable, at least for this specific data-set. By doing this, the approach was validated, and values could be adjusted to improve the Effectiveness Rate. The next steps in the process were the following:

3. The values outside the desired selected threshold were substituted by random values inside it and, therefore, a synthetic dataset was generated.
4. Then, predictions on the whole dataset were performed using the Deep Learning Model previously trained. Then the final rate of Ef1 and Ef4 was calculated.

It should be noted that, the most relevant part is the one described in step 2, where the threshold values for each variable were selected. Steps 3 and 4 also helped with validation and fine tuning of the selection, but were mainly an evaluation of the final results to quantify the success of the process.

To efficiently perform multiple experimentation scenarios, all the processes were automated, except one aspect of the second step, in which the areas that maximize and minimize the target variable in the desired way were selected.

5.4. RESULTS

5.4.1. Training of a Deep Neural Network

Table 1 displays the testing data metrics of the dataset for different types of effectiveness. As it can be observed, the model achieved high performance metrics for the dataset of type 1 and of type 4, being f1-score 0.95 and 0.94 for each class. The overall classification accuracy on the evaluation test set was 0.94. Therefore, this model, which serves as the basis for the other algorithms, has been shown to be reliable for defining the characteristics of the serve when points end in an ace and those that last 4 or more shots.

Table 1.

Sklearn Classification Report Type1 vs Type4 - Deuce & Advantage

Effectiveness	Precision	Recall	F1-score	Macro average	Weighted average
Type 1	0.91	0.99	0.95	0.95	0.94
Type 4	0.99	0.89	0.94	0.95	0.94

5.4.2. Calculation of the most relevant variables

Once the reliability of the model in distinguishing between points that end with an ace and points that last for 4 or more strokes was established, the next step was to calculate the serve variables that had the greatest relevance for each type of effectiveness. Figures 3,4 and 5 depict the mean SHAP values, which indicate the average impact on the model outputs for each variable on the deuce and advantage sides, respectively. As it can be observed, the results of the most important variables were very similar for both sides. Regardless of the serving side, the variables that had the most relevance in the model were δ and dL. Additionally, "Speed" was the other variable that had significant importance in the developed model.

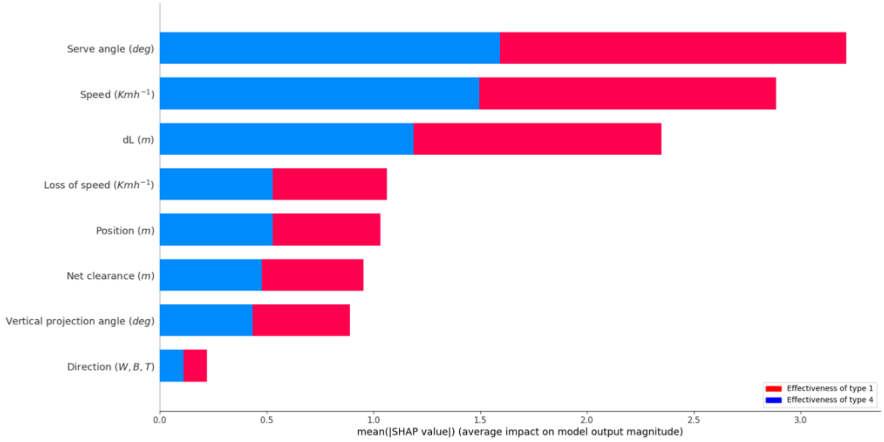

Figure 3. Effectiveness of type1 vs type4 in the DEUCE side.

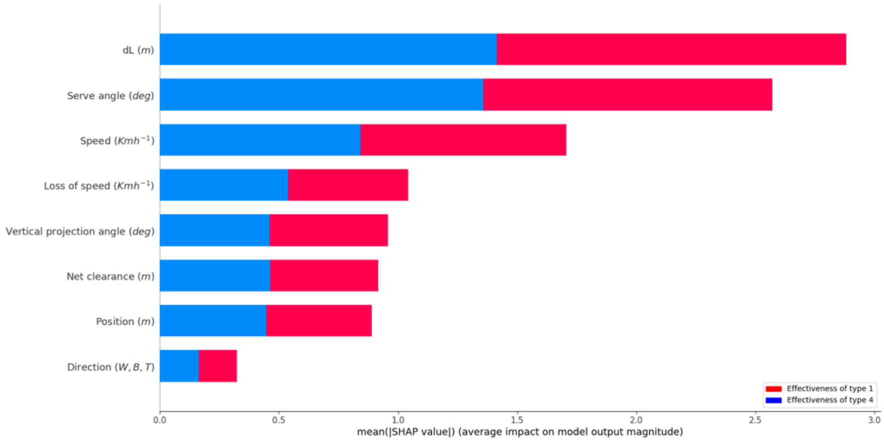

Figure 4. Effectiveness of type1 vs type4 in the ADVENTAGE side.

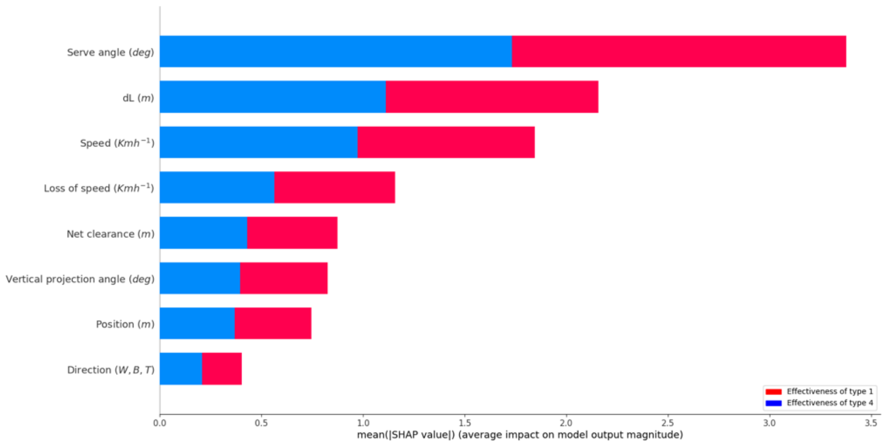

Figure 5. Effectiveness of type1 vs type4 in both sides.

The mean SHAP values show the average impact on the model outputs for each variable.

The results of the most important variables were very similar for the deuce and advantage sides.

After separating highly correlated variables (δ and dL) the SHAP values summary showed that in both cases, δ and dL were the most important variables.

The algorithm showed that the Speed is the next most important variable for the model.

5.4.3. Values that maximize the effectiveness type 1 (and the effectiveness minimize type4). That is, recommended values.

Table 2, show the distributions of the variables dL, δ, and Speed of the serves for points ending with an ace or with more than 3 shots for both sides of the court. Regarding dL, aces showed a similar distribution on both sides of the court. However, the serves for points with 4 or more shots showed significant differences in the distribution of their population. It can be observed that on the advantage side, a greater number of serves with higher values of dL were found.

Regarding δ, the distribution of serves where the point lasted more than three strokes was similar on both sides of the court. However, it can be observed that aces have a wider range on the deuce side compared to the advantage side which indicates that on the deuce side, larger angles would maximize the probability of achieving an ace to a greater extent than on the advantage side.

Regarding Speed, again, the distribution of serves where the point lasted more than 3 strokes was similar on both sides of the court. However, the recommended values to maximize the probability of serving an ace would be slightly lower on the deuce side than on the advantage side.

The recommended values to maximize the probability of serving an ace would be somewhat lower for the deuce than for the advance side. For the other efficiencies, no significant differences were found.

Table 2. Descriptive values of the variables

			dL				Serve angle				Speed			
	Effectiveness	**Population (%)**	**MaxLike (m)**	**Percentils (m)**	**Mean**	**SD**	**MaxLike (deg)**	**Percentils (deg)**	**Mean**	**SD**	**MaxLike (km h-1)**	**Percentils (deg)**	**Mean**	**SD**
	Type 1	9.68	0.19	0.04 - 1.41	0.41	0.41	5.78	2.04 - 7.78	5.76	1.52	179.86	159.01 - 210.86	185.72	16.85
DEUCE	Type 2	25.15	0.63	0.21 - 1.93	0.95	0.54	4.92	0.87 - 6.31	3.71	1.76	186.24	160.89 - 205.72	184.20	14.15
SIDE	Type 3	19.58	0.59	0.22 - 1.97	1.02	0.57	4.37	0.76 - 6.19	3.44	1.77	184.82	159.01 - 204.54	183.13	14.31
	Type 4	35.54	0.77	0.28 - 1.93	1.05	0.52	4.67	0.86 - 6.13	3.45	1.70	180.57	156.56 - 202.56	180.39	14.32
	Type 1	9.17	0.18	0.03 - 1.4	0.40	0.42	5.85	2.11 - 7.51	5.62	1.43	193.91	162.58 - 211.39	189.23	16.44
ADVANTAGE	Type 2	24.34	0.52	0.18 - 1.95	0.99	0.57	5.06	0.84 - 6.32	3.66	1.78	189.45	155.56 - 206.65	185.25	14.67
SIDE	Type 3	19.03	0.49	0.19 - 1.97	1.00	0.59	4.79	0.77 - 6.42	3.69	1.85	186.48	155.56 - 204.71	183.03	15.26
	Type 4	37.33	0.78	0.27 - 1.95	1.07	0.54	4.56	0.86 - 6.04	3.46	1.67	187.22	153.23 - 203.38	181.26	15.37
	Type 1	9.46	0.18	0.03 - 1.41	0.40	0.41	5.79	2.08 - 7.69	5.70	1.47	193.91	160.65 - 211.20	187.32	16.75
BOTH	Type 2	24.78	0.61	0.19 - 1.94	0.97	0.58	5.05	0.84 - 6.32	3.69	1.77	187.22	160.26 - 206.36	184.68	14.40
SIDE	Type 3	19.34	0.52	0.2 - 1.97	1.01	0.58	4.59	0.77 - 6.25	3.56	1.81	185.73	157.61 - 204.66	183.08	14.76
	Type 4	36.38	0.76	0.28 - 1.95	1.06	0.53	4.63	0.86 - 6.08	3.46	1.69	187.22	155.33 - 202.94	180.81	14.76

5.4.4. *Generation of a dataset substituting original values for recommended values of the variable in question and prediction of efficiencies (the model trained in point 1 was used to make the predictions).*

Using the variables associated with the population density graphs, a statistical estimation was performed on the values that would, a priori, maximize the probability of serving an ace and minimize the probability of the point lasting four or more strokes. Table 3 shows the results of the predictions using the trained model described earlier and the synthetic dataset, which resulted from replacing values that fell outside the recommended range with random values within the recommended range. As can be observed, dL values between 0 and 0.28 m increase the probability of achieving an ace to values close to 90%. Similarly, δ values between 5.7º and 8.7º elevate the possibility of achieving an ace to values above 80%. On the other hand, Speed values between 187 and 220 km h^{-1} only increase the probability of achieving an ace to approximately 40-45%.

Table 3. Results of the predictions using the trained model.

Variable	Side	Values (min-max)	Effectiveness Before (%)	Effectiveness After (%)
dL (m)	DEUCE	0 - 0.28	33.05	83.40
	AD		30.64	81.26
	BOTH		33.49	88.48
Serve angle (deg)	DEUCE	5.7 - 8.7	33.70	88.68
	AD		29.77	88.21
	BOTH		33.25	89.04
Speed (km h^{-1})	DEUCE	187 - 220	35.17	45.75
	AD		34.36	44.13
	BOTH		30.38	39.91

5.5. DISCUSSION

Using a novel methodological proposal this study sought to identify which characteristics of serves determine their efficiency in men's professional doubles tennis. The results showed that the characteristics of the serves in which the point ended in two, three or more shots were very similar. However, it was possible to identify different characteristics for the aces, compared to those in which the point lasted more than three shots. Therefore, these results allow to establish the variables and values which are associated with a higher probability of a serving an ace in men's professional doubles tennis. In addition, the methodology used has been shown to be a novel, valid and reliable way that can be replicated for the calculation of the efficiency of actions in different sports disciplines.

The results have shown that the service angle is one of the variables that best discriminates the probability of serving an ace in men's professional doubles tennis. These results are consistent with a previous study of the serve in singles tennis, which indicated that if the serve angle was less than 5.88º the serve will be returned with a probability of 92.52% (Whiteside & Reid, 2017). In doubles tennis, according to the data obtained in the present study, an angle of between 5.7º and 8.7º theoretically increases the probability of serving an ace from 29.77% to 88.21% on the advantage side and from 33.70% to 88.68% on the deuce side. These results support those obtained by previous studies where Whiteside and Reid (2017), reported that there is a higher probability of getting an ace when the serve is played wide than to the T zone, with probabilities of 42% on the advantage side and 52% on the deuce side. Wide serves generate a greater angle than those played to the T, so it is logical, according to the data obtained, that they are also more effective. Furthermore, from a technical point of view, and specifically related to the different types of serve spin used, it has been shown that players use the slice to generate a greater serve angle on the deuce side, whereas they opt for flat or topspin for a greater angle when serving to the T on the advantage side (Matsuzaki, 2004; Gillet et al., 2009; Mecheri et al., 2016).

Another of the variables found to have a greater relevance in determining the probability of serving an ace has been the distance from the bounce of the ball to the sideline of the service box. This reinforces the idea that one of the main objectives of the server is to gain an advantage over the receiver, trying to get the ball as far away from the receiver as possible (Mecheri, 2016; Martínez-Gallego et al., 2021; Vives et al., 2022). Whiteside and Reid (2017) determined that in 77.53% of cases, if the distance to the line was less than 15.27 cm and, in addition, the serve angle was greater than or equal to 5.88º the serve would be an ace. In the case of men's professional doubles tennis, in this study it has been shown that, if the serve distance is between 0 and 28 cm, the probability of serve being an ace would go from 30.64% to 81.26% on the advantage side and from 33.05% to 83.40% on the deuce side.

The results on the importance of the angle of the serve and the distance of the bounce of the ball to the line also support what has been found in previous studies on the professional doubles game. According to Vives et al. (2022) players try to take the initiative of the point by moving the receiver out of the court, both on the advantage and deuce sides. This study reported a high efficiency of the serve when players serve wide on both sides of the court, with values of 40.71% on the deuce side and 37.54% on the advantage side. Similarly, the percentages of effectiveness of serves to the T were considerably high, with values of 38.45% and 39.28% respectively. The information provided in this study regarding the most important variables and the values that maximize the probabilities of direct serves could be very valuable for coaches and players in order to further increase the effectiveness of this type of service.

In singles, serve speed has been shown to be one of the most important variables when serving an ace. Previous studies have reported that the probability of serving and ace if the serve speed is greater than 198 km h^{-1} is 58.68% (Whiteside and Reid, 2017). However, the data obtained in our study show that the serve speed is less important in doubles tennis. Thus, it was found that once 187 km h^{-1} is exceeded, the probability of serving an ace would only increase by about 10% regardless of the side of the serve.

Therefore, considering the results from this study and the values found for the most important variables that determine the execution of serving an ace, it can be stated that, although speed is an important variable, very high speeds do not considerably increase the probability of winning the point directly with the serve. However, the angle of the serve and the bounce of the ball close to the sidelines do significantly increase the probability of serving an ace.

These results have very relevant practical implications related to the training regime for serve practices in men's professional doubles. From a tactical perspective, knowledge of the insight found in this study can assist tennis players in their decision making to maximize their serve advantage. In addition, technically, physically, and mentally, it is also important to prepare doubles tennis players for efficient service execution with respect to the values of the indicated variables. Therefore, the effective application of the values found for variables such as service angle, distance to the line, or speed will be determinant for the design of specific training tasks and for the definition of specific patterns of play (Whiteside & Reid, 2017; Vives et al., 2022).

5.6. CONCLUSIONS

In this study, key characteristics for a direct serve in men's professional doubles tennis were identified. The angle of the serve was found to be a crucial variable, while the distance from the ball bounce to the sideline was also relevant. These findings are consistent with previous research in doubles tennis. It was observed that an optimal service angle between 5.7° and 8.7° significantly increased the probability of a direct serve. In addition, maintaining a distance from the ball bounce to the sideline between 0 and 28 cm increased the probability of success. The speed of the serve in doubles showed less influence compared to singles tennis. These results have practical implications for training and tactical decision-making in doubles tennis to maximise the effectiveness of serves.

5.7. REFERENCES

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: a corrected feature importance measure. *Bioinformatics*, *26*(10), 1340-1347.

Brouwer, M. J. (2021). *Classification of tennis serve direction using supervised machine learning* (Doctoral dissertation).

Browning, R. (2009). The Serve: The Most Important Shot in Tennis. *USTA Magazine.*

Candila, V., & Palazzo, L. (2020). Neural networks and betting strategies for tennis. *Risks*, 8(3), 68.

Carboch, J., & Kočíb, T. (2016). A comparison of service efficiency between players of male and female doubles at professional tennis tournaments. *Auc Kinanthropologica*, 51(2), 56-62.

Cornman, A., Spellman, G., & Wright, D. (2017). Machine learning for professional tennis match prediction and betting. *Working Paper, Stanford University*..

De Seranno, A. (2020). Predicting Tennis Matches Using Machine Learning.

Giampolo, F., & Levey, J. (2013). Championship tennis (pp. 59). Champaign, IL: *Human Kinetics.*

Giles, B., Kovalchik, S., & Reid, M. (2020). A machine learning approach for automatic detection and classification of changes of direction from player tracking data in professional tennis. *Journal of sports sciences*, 38(1), 106-113.

Giles, B., Peeling, P., Kovalchik, S., & Reid, M. (2021). Differentiating movement styles in professional tennis: A machine learning and hierarchical clustering approach. *European Journal of Sport Science*, 1-10.

Gillet, E., Leroy, D., Thouvarecq, R., & Stein, J.-F. (2009). A notational analysis of elite tennis serve and serve-return strategies on slow sur- face. *Journal of Strength and Conditioning Research*, 23, 532–539.

Gupta, M., Gupta, R., Ansari, S., & Tambe, D. N. (2022). Tennis Match Prediction using Machine Learning. *International Journal of Research Publication and Review*, 3(4), 693-696.

Hawk-eye innovations. (2015). *Hawk-Eye's accuracy & reliability: Electronic line calling*. Basingstoke.

Kingma, D. P., & Ba, J. (2015). Adam: a method for stochastic optimization. Conference Paper at ICLR 2015. *arXiv preprint arXiv:1412.6980*.

Kocib, T., Carboch, J., Cabela, M., & Kresta, J. (2020). Tactics in tennis doubles: analysis of the formations used by the serving and receiving teams. *Int J Phys Educ Fit Sport*, 9(2), 45-50.

Kostoglou, G. T., Tsitas, K., Kostoglou, V., & Kostoglou, P. (2017). The importance of the serve in tennis: A systematic review. *International Journal of Performance Analysis in Sport*, 17(2), 505-518.

Kovalchik, S. A. (2016). Searching for the GOAT of tennis win prediction. *Journal of Quantitative Analysis in Sports*, 12(3), 127-138.

Learning, M. (2017). Final Project Report: Real Time Tennis Match Prediction Using Machine Learning

Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980-2988).

Liu, L. (2022). Model for Evaluating the Technical and Tactical Effectiveness of Tennis Matches Based on Machine Learning. Mobile Information Systems, 2022.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, *30*.

Mahesh, B. (2020). Machine learning algorithms-a review. *International Journal of Science and Research (IJSR*), 9, 381-386.

Mlakar, M., & Luštrek, M. (2017). Analyzing tennis game through sensor data with machine learning and multi-objective optimization. In Proceedings of the 2017 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2017 ACM international symposium on wearable computers (pp. 153-156).

Martínez-Gallego, R., Crespo, M., & Jiménez, J. (2021a). Analysis of the differences in serve effectiveness between Billie Jean King Cup (former Fed Cup) and Davis Cup doubles tennis matches. *International Journal of Sports Science & Coaching, 16*(3), 777-783.

Martínez-Gallego, R., Crespo, M., Ramón-Llin, J., Micó, S., & Guzmán, J.F. (2020). Men's doubles professional tennis on hard courts: Game structure and point ending characteristics. *Journal of Human Sport and Exercise, 15*(3), 633-642. doi: https://doi.org/10.14198/jhse.2020.153.13

Martínez-Gallego, R., Salvador, S. M., Luján, J. F. G., Reid, M., Ramón-Llin, J., & Crespo, M. (2021b). Challenging serve myths in doubles tennis. *International Journal of Sports Science & Coaching*, 16(6), 1305–1311.

Martínez-Gallego, R., Vives, F., Guzmán, J. F., Ramón-Llin, J., & Crespo, M. (2021c). Time structure in men's professional doubles tennis: does team experience allow finishing the points faster?. *International Journal of Performance Analysis in Sport*, 21(2), 215-225.

Matsuzaki, C. (2004). Tennis fundamentals (pp. 91). Champaign, IL: *Human Kinetics*.

Molnar, C. (2020). Model-Agnostic Methods. *Interpretable machine learning*. Lulu. com. 143-221.

O'Donoghue, G. P., & Brown, E. (2008). The importance of service in Grand Slam singles tennis. International *Journal of Performance Analysis in Sport*, 8(3), 70-78.

Perri, T., Reid, M., Murphy, A., Howle, K., & Duffield, R. (2022). Prototype Machine Learning Algorithms from Wearable Technology to Detect Tennis Stroke and Movement Actions. Sensors, 22(22), 8868.

Reid, M., McMurtrie, D., & Crespo, M. (2010). The relationship between match statistics and top 100 ranking in professional men's tennis. *International Journal of Performance Analysis in Sport*, 10(2), 131-138.

Reid, M., Whiteside, D., & Elliott, B. (2011). Serving to different locations: set-up, toss, and racket kinematics of the professional tennis serve. *Sports Biomechanics*, 10(4), 407-414.

Sekar, A. (2019). Predicting the Winner of a Tennis Match Using Machine Learning Techniques (Doctoral dissertation, Dublin, National College of Ireland).

Sipko, M., & Knottenbelt, W. (2015). Machine learning for the prediction of professional tennis matches. MEng computing-final year project, Imperial College London, 2.

Solanki, S., Jakir, V., Jatav, A., & Sharma, D. (2022) Prediction of tennis match using machine learning. *International journal of progressive research in engineering management and science*, 2 (6), 5-7.

Vives, F., Crespo, M., Guzmán, J. F., & Martínez-Gallego, R. (2022). Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology. *International Journal of Performance Analysis in Sport*, 22(4), 638-648.

Yue, J. C., Chou, E. P., Hsieh, M. H., & Hsiao, L. C. (2022). A study of forecasting tennis matches via the Glicko model. *Plos one*, *17*(4), e0266838.

Wilkens, S. (2021). Sports prediction and betting models in the machine learning age: The case of tennis. *Journal of Sports Analytics*, *7*(2), 99-117.

Wei, X., Lucey, P., Morgan, S., Carr, P., Reid, M., & Sridharan, S. (2015). Predicting serves in tennis using style priors. In Proceedings of 21st ACM SIGKDD international conference on knowledge discovery and data mining (pp. 2207–2215), Sydney. New York, NY: Association of Computing Machinery.

Zhu, Y., & Naikar, R. (2022). Predicting Tennis Serve Directions with Machine Learning. In *International Workshop on Machine Learning and Data Mining for Sports Analytics* (pp. 89-100). Cham: Springer Nature Switzerland

6. ESTUDIO 3: *SERVE STRATEGY IN ELITE WOMEN'S DOUBLES TENNIS - A DATA-DRIVEN PERSPECTIVE.*

Vives, F., Lázaro, J., Guzmán, J. F.., Crespo, M., & Martínez-Gallego, R. (2024). Estrategia de servicio en el tenis profesional de dobles femenino - una perspectiva basada en datos (Serve strategy in professional women’s doubles tennis - a data-driven perspective). *Retos*, *60*, 955–962.

6.1. ABSTRACT & KEYWORDS

This study aimed to identify the key variables that influence the effectiveness of the first serve in women's doubles matches using a deep learning training model. A total of 4939 first serves by 145 players from Billie King Cup matches were classified according to their effectiveness. Data from the Hawk-Eye system were filtered and processed to obtain descriptive statistics of the spatio-temporal predictor variables and then to calculate the most important variables for first serve effectiveness. The angle of the service was the feature that gained more relevance, along with the speed and distance from the bounce to the sidelines, regardless of the side from which the service was made. The training model showed novel results, identifying that the values of the service angle between 5.5º-8.7º increased the effectiveness from 22.2% to 48.5% on the Deuce side and from 17.3% to 45.7% on the Advantage side. These results not only align with existing research on men's professional tennis across both singles but also offer valuable tactical and training guidance. The insights gathered from this study are poised to enhance training methodologies and match-day strategies, providing a strategic edge in serve selection during competitive play.

Keywords: racquet sports, machine learning, strategy, sport analytics, serve

6.2. INTRODUCTION

Tennis is an open skill sport in which, from a strategy perspective, players make tactical decisions based on their strengths and weaknesses, as well as those of their opponents. In the quest to understand the critical factors that influence success in tennis, numerous studies have analysed players' match statistics (Crespo et al., 2015; O'Donoghue, 2002).

The serve is one of the key performance aspects in tennis, as it is a fundamental skill that plays a crucial role in the match outcome. In fact, points won with the second serve and points won with the return of the second serve have been identified by authors such as Reid, McMurtrie and Crespo (2010) as the most relevant variables of the game. Furthermore, it has been shown that players improve their serving statistics over the course of a tournament (Grambow et al., 2020), and that those with a better serve win more points with the second serve and in break point situations and win more points and games when serving (Infantes et al., 2019), which highlights the importance of this stroke in the performance of the player.

Despite being the most decisive stroke in tennis today, the scarcity of studies dedicated to the serve in the women's game is evident (Sánchez-Pay et al., 2022). Both the serve and the return have been shown to play a crucial role in competition (Stănescu, 2016; Abidin & Ruslan, 2020) and can have an impact on subsequent strokes (Born, 2021). Authors such as Grambow et al. (2021) have suggested that players should focus on increasing the percentages of games and points won with the serve, as well as perfecting aces to improve serving effectiveness. Other authors have explored biomechanical aspects (Whiteside et al., 2015) and the kinematics of the first serve at different stages of training (Whiteside et al., 2013). Furthermore, studies have also analyzed how players' ranking can influence their performance (Sánchez-Alcaraz, 2018; Kovalchik et al., 2017)).

As per the serve in the doubles game, it has received considerably less attention from researchers than that of the singles. When comparing the service efficiency between male and female professional players Carboch & Kočíb (2016) concluded that men won around 10% service points more than women, and that to win a match they also needed to win more service points than women. Kocib et al. (2020) tactically analysed the formations of serving and returning teams, and Avkhimovich (2023) investigated the order of servers in doubles and how this affects who should serve first. Martinez-Gallego et al. (2021c) demystified popular serving beliefs in doubles, concluding that the serving team's advantage is most notable on points of up to four shots, especially with new balls on clay, although this advantage diminishes on break points. Lastly, a study of Davis Cup doubles matches revealed that the effectiveness of the serve in doubles depends more on the angle and placement than on the speed of the serve (Vives et al., 2023).

Similarly, an obvious lack of research can be observed in women's doubles, despite its relevance in team competitions such as the Billie Jean King Cup (BJKC). The BJKC, formerly known as the Fed Cup until 2020, is the most important global competition played by 110 national teams annually, which is the tennis World Cup for women. Created in 1963 by the ITF, it is the equivalent for women of the Davis Cup (BJKC, 2023). As aforementioned, research on this competition is substantially less common than in the men's events.

In this context, other studies have focused on gender differences in the volley game in doubles, highlighting certain court coverage differences between men and women (Martínez-Gallego et al, 2021b), or the activity profile of players on different surfaces in doubles Grand Slam matches, resulting in short and intense points with a predominance of net play (Borderías et al., 2021). Finally, Borderías et al. (2023) analysed structural variables and the completion of points in women's doubles, concluding that taking the initiative and being aggressive in the game is beneficial for the point outcome.

From a technological perspective in terms of the tools used for match analysis, the use of tracking technologies in tennis, such as the Hawk-Eye (HE), has allowed different studies to use a large amount of data for the analysis of the singles game performance indicators. The serve has been analyzed in a myriad of studies (Loffing, 2009; Rioult et al., 2015; Wei et al., 2015; Mecheri, 2016; Whiteside and Reid (2017). For instance, Fitzpatrick et al. (2023) explored the different serve and return strategies used by male and female players at Wimbledon. Other studies have also investigated the different movement parameters, changes of direction and hitting dynamics of male and female players during matches (Wei et al., 2013; Reid and Duffield, 2014; Reid et al., 2016; Giles et al., 2020; Brandon et al., 2021; Wei et al., 2013; Meurs et al., 2021).

In the case of the doubles game, HE data have been used to analyze the characteristics of the services as a function of their direction. Vives et al. (2021) found that the position of the server was shown to be a determining factor in the direction of the first or second serve and, therefore, a variable to be considered for the effectiveness of the serve itself and the anticipation in the return. The differences in the characteristics of the serve in men's and women's doubles tennis have also been analysed in terms of its effectiveness and the directions used, and Martínez-Gallego et al. (2021a) concluded that men's teams are more effective than women's, especially in serves directed to areas far from the center of the service box.

Despite the acknowledged importance of the service in tennis and the numerous studies that have analyzed its different aspects in the singles and the men's game, there is a notable lack of specific research focused on women's doubles, especially in high-level competitions such as the Billie Jean King Cup. This gap in the scientific literature justifies the need for this research that, by using currently available technologies and analyses, explores service dynamics and strategies in this tennis discipline. Therefore, this study aims to fill this research gap by identifying and analyzing the key variables that influence the effectiveness of the first serve in Billie Jean King Cup women's doubles matches.

6.3. METHODS

6.3.1. *Sample*

A total of 4939 first serves was analysed from 65 women's doubles matches played during BJKC qualifying ties between 2011 and 2019. Our study featured the participation of 107 teams and 145 players from 23 different countries, with an average age of 25.81 ± 4.74 years.

6.3.2. Process

The data used in this study was obtained using the HE tracking system (Hawk-Eye Innovations Ltd., 2015). The HE tracking technology has passed rigorous ITF tests that have validated the accuracy and reliability of the system. A system of synchronized cameras is used to track the movement of the ball in real time. Each camera captures the position of the ball in its own plane. By knowing the position and orientation of each camera, the 3D position of the ball can be estimated by intersecting the trajectories defined by the cameras. This allows the three-dimensional coordinates of the ball to be determined accurately in space (Singh Bal & Dureja, 2012; Reid et al., 2016).

After acquiring data from the Hawk-Eye system, a template was developed using Microsoft Excel version 16.16.7 (Microsoft, Redmond, USA) to facilitate data processing and the calculation of targeted variables. The derived variables included:

Target variables

The variables derived and analyzed in this study were:

- Court side: Side of the court from where the serve is played. This variable was used to divide the dataset and perform the analysis for each side of the court.
- Effectiveness:
    - Type 1: The point finishes with one shot.
    - Type 2: The point finishes with two shots.
    - Type 3: The point finishes with three shots.
    - Type 4: The point finishes with four or more shots.

Input features

- Speed: Mean speed of the serve.
- Position: Position of the server when hitting the serve.
- Time: Time between ball impact and ball bounce.
- Speed loss: Loss of speed of the ball after its bounce.
- Impact Z: Height of the ball at impact.
- Net clearance: Height of the ball when passing over the net.
- δ: Serve angle.
- β: Vertical projection angle.
- dL: Distance from ball bounce to the sideline of the serve box.
- Direction: Wide (W), Body (B) and T (Figure 1).

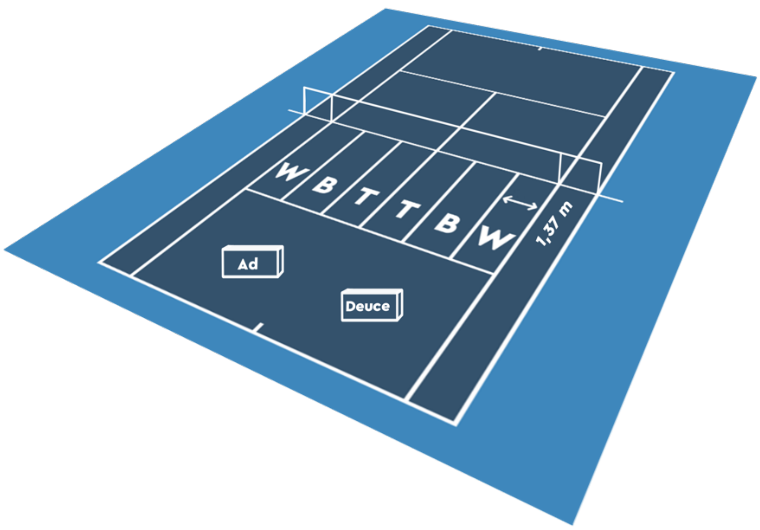

Figure 1. Serve direction.

6.3.3. Data analysis

The Machine Learning methodology used in this study was used and validated in a previous study by Vives et al. (2023), using the following steps:

6.3.3.1. Predictive model: Training the deep learning model

This study compared AutoGluon and FastAI for tabular data analysis. Although AutoGluon offers outstanding tools, its lack of flexibility leading to the choice of FastAI. The implementation of AutoGluon was straightforward, while FastAI required more development, but provided flexibility.

A tabular model with 5 layers of feature size [256, 128, 128, 128, 128, 64] and LinBnDrop basic blocks were used. The loss function was Focal Loss Flat, and the Adam optimization algorithm (Kingma & Ba, 2015) was used. This work contributes to the understanding of tools for tabular data analysis.

6.3.3.2. Feature importance: Most important variable

The Feature Permutation Importance methodology emerges as a robust tool, facilitating the identification of features with predictive capability, irrespective of the employed model. This approach provides a valuable perspective for assessing the individual contribution of variables within the broader analytical framework (Kuhn & Johnson, 2013).

Feature importance algorithms were employed to identify variables with a substantial impact a predictive model, indicating the relative significance of each feature in making predictions. It is based on repeated permutations of the outcome vector to estimate the distribution of measured importance for each variable in a non-informative setting (Rajbahadur et al., 2022).

6.3.3.3. Probabilistic and statistical analysis: Values that maximize the probabilities

In the context of a study focused on maximizing serve effectiveness and understanding the impact of selected features on predicted target variables, a novel semi-automated algorithm based on classical statistics was developed.

The initial steps involved calculating a smooth probability distribution using the Kernel Density Estimation algorithm for each selected feature, along with statistical estimators such as Mean, Maximum Likelihood Estimation (MLE), and percentiles 5 and

95 for each type of effectiveness. Subsequently, values were identified by selecting areas on the plot with a higher density of points corresponding to the desired result (type 1) and a lower density of points for the undesired result (type 4). Statistical estimators played a crucial role in interpreting the probability distribution during this process.

6.3.3.4. Synthetic data-set generation: simulation based on the predictions on this dataset and analysis of effectiveness improvements.

The final step was to create a synthetic dataset by applying the identified thresholds to the original dataset Through validation procedures, we made adjustments to improve the effectiveness rate. To further refine the dataset, any values outside the specified thresholds were replaced with random values within the acceptable range, creating a new synthetic dataset. We then used the previously trained deep learning model to make predictions on the synthetic dataset, calculating the final rates for effectiveness type 1 and effectiveness type 4.

The Effectiveness results using the original dataset and the new synthetic dataset with the recommended threshold values are shown in table 2. They show how the model trained on the original dataset impacts the effectiveness on the second dataset, increasing the overall probability on the synthetic dataset by 47.1%

6.4. RESULTS

6.4.1. Descriptive statistics of the spatiotemporal predictor variables

Table 1 shows the descriptive statistics of the spatio-temporal predictor variables. As it can be seen in the results, the mean of the variable speed is higher in aces than in the rest of the serves, both on the Deuce side with 160.19 km/h and on the Advantage side with 165.90 km/h. This difference is also reflected in the δ, being the angle greater in aces with an average of 5.99º on the Deuce side and 5.74º on the Advantage side. On the other hand, the dL was another of the variables that showed significant differences between the aces and the rest of the serves, being the shortest distance in aces regardless of the side. Regarding the rest of the variables, as can be seen, their means were quite similar on the Deuce and Advantage sides, both for the serves that were aces and for the rest of the first serves.

Table 1. Descriptive statistics of the spatio-temporal predictor variables

	DEUCE				ADVANTAGE			
	Total (N=2615)		Aces (N=156)		Total (N=2324)		Aces (N=113)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Speed (km/h)	154.14	13.63	160.19	15.69	155.88	13.72	165.90	13.64
Position (m)	2.03	0.73	1.96	0.79	2.03	0.82	1.82	0.89
Impact Z (m)	2.67	0.14	2.70	0.13	2.68	0.15	2.72	0.14
Net Clearance (m)	1.14	0.11	1.13	0.11	1.14	0.11	1.13	0.11
Time (s)	0.48	0.05	0.47	0.05	0.48	0.05	0.45	0.04
Speed loss (km/h)	7.80	1.20	7.76	3.28	7.85	1.14	7.50	2.83
Serve angle (deg.)	3.59	1.91	5.99	1.64	3.63	1.74	5.74	1.56
Vertical projection (deg.)	9.05	0.61	9.07	0.56	9.02	0.64	9.05	0.55
dL (m)	1.01	0.54	0.39	0.37	0.99	0.56	0.40	0.47

6.4.2. Calculation of the most relevant variables

The next step was the calculation of the serve variables that were most important for type 1 effectiveness. Figure 2 shows the mean values indicating the average impact on the model outputs for each variable on the Deuce side, the Advantage side and both sides respectively. As it can be seen, the most important variables were δ, speed and dL to the line irrespective of the side where the serve was played.

Figure 2. Importance of the variables for the aces.

6.4.3. Results of the predictions using the trained model.

Table 2 shows the results of the predictions using the trained model and the synthetic dataset, which resulted from replacing values outside the recommended range with random values within the recommended range. Serve angle values between 5º and 8.7º double the initial probability of making an ace on both the deuce and the advantage side. Similarly, although to a lesser extent, both the variable of serve speed (between 157 – 200 km/h) and the variable of distance to the line < 0.42m, also improve the probability of making an ace, regardless of the side of the court.

Table 2.

Results of the predictions using the trained model.

Variable	Side	Values (min-max)	Effectiveness Before (%)	Effectiveness After (%)
Serve angle (deg)	DEUCE	5.5 – 8.7	22.2	48.5
	AD		17.3	45.7
	BOTH		18.7	47.1
Speed (km/h)	DEUCE	157 – 200	18.8	24.3
	AD		16.7	23.9
	BOTH		16.8	22.8
dL (m)	DEUCE	0 – 0.50	20.3	21.7
	AD		16.4	20.3
	BOTH		17.5	20.7

6.5. DISCUSSION

This study addressed serve effectiveness in women's doubles tennis, using advanced machine learning techniques to analyze a large dataset collected from BJKC matches. From the analysis of 4939 serves in 65 matches, it was found that serve angle, serve speed and distance to the serve line are critical variables in increasing the likelihood of hitting an ace.

The results shown in Table 2, have determined that serve angles between 5.5º-8.7º have a direct impact on the serving team achievement of an ace in women's professional doubles tennis, increasing the overall probability by 47.1 % (48.5% on the deuce side and 45.7% on the advantage side). These results are in the lines of previous studies in both singles and men's doubles tennis, where it was reported that angles from 5.88º and 5.7º, respectively, significantly increased the probability of hitting an ace (Whiteside & Reid, 2017; Vives et al., 2022). These results confirm that, in women's doubles tennis, as in the men's game, employing strategies in which the direction of the serve seeks to place the ball away from the receiver increases the efficiency of the serve partner and, therefore, their chances of winning the point, either through an ace or by taking the initiative of the point (Rioult et al., 2015; Mecheri et al., 2016; Martínez-Gallego et al., 2021; Vives et al., 2022). However, it is important to note that these improvements in efficiency in women's doubles (≈30 %) are considerably lower than those reported in the men's game (≈50-60 %). This may be due, on the one hand, to the lower overall efficiency of the serve in women's doubles compared to men's doubles (Martínez-Gallego et al., 2021a) and, on the other hand, to the position adopted by the players in the return of serve. In this regard, although there are no data in doubles tennis, in the singles game it has been reported that female players adopt positions closer to the net than male players (Reid et al., 2016). This could allow them to shorten the trajectory of the ball and, therefore, move a shorter distance, thus increasing the likelihood of returning the serve.

The variable of serve speed has been explored due to its impact on the game in several studies, which have found a significant relationship with the probability of winning the point (Rioult et al., 2015). In contrast, no direct relationship has been found between aces and speed in both male and female players (Brown, 2021). However, the results obtained in this study show that, in women's doubles tennis, first serve speed, although limited, does improve the likelihood of hitting an ace. Specifically, it has been found that serves with speeds of 157 km/h and above improve the probability of hitting an ace on both sides of the court by around 5%. These results are in line with those found in men's doubles, where once 187 km/h was exceeded, the probability of an ace was increased by approximately 10% (Vives et al., 2023). Thus, although serve speed is an important variable that has received much attention and interest (Rioult et al., 2015; Grambow et al., 2021; Mecheri et al., 2016), its contribution to improving effectiveness is limited compared to other variables such as angle. Currently, probably due to improved return performance, returners are able to return a very high percentage of serves, even at high speeds (Born et al., 2021; Fitzpatrick et al. 2023), which means that, in isolation, speed is not a determining variable in serve effectiveness. Therefore, it seems reasonable that, although serve speed is an aspect to work on with the players, it would be advisable to integrate it with variables related to direction and angle to maximize effectiveness.

Similar to speed, the distance from the ball bounce to the serve box sideline, although it improves the effectiveness of serves, does so to a limited extent, increasing the probability of hitting an ace by approximately 5% when the distance is less than 0.50 m. These results differ from those observed in men's doubles tennis, where bouncing the ball within 0.28 m of the line was reported to increase the probability of hitting an ace by 50% (Vives et al., 2023). This also contrasts with a study of men's singles tennis, which found that bouncing within 0.15 m of the line significantly increased the probability of hitting an ace (Whiteside and Reid, 2017). These differences are likely due to the interaction of this variable with speed. The fact that male players serve at a higher speed than female players (Brown, 2021) means that, when the ball bounces close to the lines, effectiveness increases to a greater extent. On the other hand, in the case of the female players, although the ball bounces close to the lines, serving at a slower speed allows the receiver to reach the ball in better conditions and, therefore, the serve is less effective.

Even though this study is comprehensive, there are certain limitations that should be taken into account. These limitations also provide some ideas for future research directions in this particular area of study.

One of the key limitations of this study is the isolated analysis of the serve, conducted without considering its impact on subsequent shots or shot patterns that develop later in the game. This approach omits how the characteristics of the serve can affect the dynamics of the point. Future research could analyze the effectiveness of the serve using a more comprehensive approach, by examining how the variables of the serve contribute to the overall development of the game, thus allowing for a more holistic understanding of the various nuances in the strategy of women's doubles tennis.

Another important limitation is the independent analysis of the variables conducted, without exploring the interaction between them. This simplified approach may not fully capture the complexity and multifactorial nature of tennis performance. In this scenario, future research should consider the interactions between variables, which would offer deeper insights into how specific combinations of factors influence serve success in women’s doubles.

Finally, the exclusive use of data accessed from the BJKC may limit the absolute applicability of the results to other competitive contexts in women’s doubles tennis. The specificity of the sample suggests the need for caution in absolute applying the results. Future research could also extend the analysis to different competitions and levels of play, thus enriching the understanding of serving in women’s doubles tennis and its impact on competitive performance. These studies could reveal tactical and strategic variations adapted to different playing contexts, offering a broader and globally applicable perspective.

From a practical application perspective, the results of this study reveal critical aspects that can be considered by coaches and trainers to improve serve effectiveness in women's doubles on court training sessions. The identification of speed, bounce distance to the line, and particularly, serve angle as key variables that increase the likelihood of hitting aces suggests that serve training should prioritize these elements. This aligns with the findings of Filipcic et al. (2021), who highlighted the importance of shot angles in their analysis of matchplay characteristics and performance indicators. Therefore, specific technical training that incorporates exercises aimed at improving mechanical variables, such as serve angle and speed, could be particularly beneficial. Furthermore, when planning doubles game strategies, and considering the observed strengths and weaknesses in the opponents, it would be relevant to understand how the variables identified in this study can help not only to improve serving performance, but also in shaping the post-serve game patterns, allowing players to take the initiative from the start of the point.

6.6. CONCLUSIONS

This study explored the effectiveness of the serve in women's doubles tennis, using advanced machine learning techniques to analyse data collected from Billie Jean King Cup matches. The results revealed that serve angle is the variable that determines the most the likelihood of hitting an ace. Speed and distance to the serve sidelines, although to a lesser extent, were also related to a greater likelihood of hitting an ace. These findings underline the importance of integrating technical, tactical, and biomechanical aspects to improve women's doubles serving performance through the appropriate use of variables. Therefore, it is suggested that coaches and physical trainers focus their serve training sessions on improving the angle and speed of the shot, as well as on improving the strategic selection of serve direction, to enhance the chances of success in women's doubles tennis.

6.7. REFERENCES

Abidin, A. & Ruslan, N. (2020). Exploring the importance of players' characteristics and performance on serve and return of serve in winning the women's singles Grand Slam tennis tournaments. *Journal of Physics. Conference Series*, *1496*(1), 012008.

Avkhimovich, N. (2023). Should the strongest serve first? Strategic ordering of servers in doubles tennis. *Applied Economics*, 1-10.

Baiget, E., Corbi, F., & López, J. L. (2023). Influence of anthropometric, ball impact and landing location parameters on serve velocity in elite tennis competition. *Biology of Sport*, *40*(1), 273.

Bal, B. S. & Dureja, G. (2012). Hawk Eye: A Logical Innovative Technology Use in Sports for Effective Decision Making. *Sport Science Review*, *21*.

BJKC (2023). Billie Jean King Cup History. Available at: https://www.billiejeankingcup.com/en/about/history. *Accessed 8 March.*

Borderias, M., Crespo, M., Martínez-Gallego, R. & Baiget, E. (2023). Comparison of the game structure and point ending during Grand Slam women's doubles tennis. *International Journal of Sports Science & Coaching*, *18(4),* 1248-1255. https://doi.org/10.1177/17479541221098226

Borderias, M., Fernández-Fernández, J., Martínez-Gallego, R. & Baiget, E. (2021). Comparison of the activity profile during Grand Slam women's doubles tennis. *International Journal of Performance Analysis in Sport, 21, 1066 – 1080.*

Born, P., Malejka, L., Behrens, M., Grambow, R., Meffert, D., Breuer, J., & Vogt, T. (2021). Stroke placement in women's professional tennis: What's after the serve?.

Brown, E. G. (2021) A faster serve has more impact on success for female elite tennis players than males, *International Journal of Performance Analysis in Sport*, 21:4, 600-610.

Carboch, J., & Kočíb, T. (2016). A comparison of service efficiency between players of male and female doubles at professional tennis tournaments. *Auc Kinanthropologica*, *51*(2), 56-62.

Crespo, M., Filipcic, A., Panjan, A., Reid, M., Šarabon, N., & Zečić, M. (2015). Differences in performance indicators of elite tennis players in the period 1991-2010. *Journal of physical education and sport, 15, 671-677.*

Durán, A., Martínez-Gallego, R., & Gimeno, M. (2022). Diferencias de género en la aproximación al cambio de dirección en tenistas profesionales (Gender differences in the approach to change of direction in professional tennis players). *Retos*, *43*, 938–943. https://doi.org/10.47197/retos.v43i0.88803

Filipcic, A., Leskosek, B., Crespo, M., & Filipcic, T. (2021). Matchplay characteristics and performance indicators of male junior and entry professional tennis players. *International Journal of Sports Science & Coaching*, *16*(3), 768-776.

Filipcic, A., Zecic, M., Reid, M., Crespo, M., Panjan, A., & Nejc, S. (2015). Differences in performance indicators of elite tennis players in the period 1991-2010. *Journal of Physical Education and Sport*, *15*(4), 671.

Fitzpatrick, A., Stone, J. A., Choppin, S., & Kelley, J. (2024). Analysing Hawk-Eye ball-tracking data to explore successful serving and returning strategies at Wimbledon. *International Journal of Performance Analysis in Sport*, *24*(3), 251-268.

Fitzpatrick, A., Stone, J. A., Choppin, S., & Kelley, J. (2019). Important performance characteristics in elite clay and grass court tennis match-play. *International Journal of Performance Analysis in Sport*, *19*(6), 942–952. https://doi.org/10.1080/24748668.2019.1685804

Giles, B., Kovalchik, S. & Reid, M. (2020) A machine learning approach for automatic detection and classification of changes of direction from player tracking data in professional tennis, Journal of Sports Sciences, 38:1, 106-113, doi: 10.1080/02640414.2019.1684132

Grambow, R., Born, P., O'Shannessy, C., Breuer, J., Meffert, D., & Vogt, T. (2022). Serve efficiency development in women's vs. men's professional tennis. *Human movement*, *23*(2), 128-137.

Grambow, R., Oshannessy, C., Born, P., Meffert, D. & Vogt, T. (2021). Serve efficiency development indicates an extended women's tennis world class cohort: Analysing 14 years of Ladies Wimbledon Championships – implications for coaching. *Human Movement.*

Hawk-eye innovations. (2015). *Hawk-Eye's accuracy & reliability: Electronic line calling*. Basingstoke.

Infantes, P., Hidalgo, C. & Sánchez-Álcaraz Martínez, B. J. (2020). Perfiles de juego en función de las estadísiticas del saque del top 25 masculino de tenis. Education, Sport, Health and Physical Activity (ESHPA*): International Journal*, 4(3), 313-323.

Jindo, T., Mitsuhashi, D., & Kubota, T. (2022). Accuracy of subjective stats of key performance indicators in tennis. International Journal of Racket Sports Science, 4(2), 32-39.

Kingma, D. P., & Ba, J. (2015). Adam: a method for stochastic optimization. Conference Paper at ICLR 2015. *arXiv preprint arXiv:1412.6980*.

Kocib, T., Carboch, J., Cabela, M., & Kresta, J. (2020). Tactics in tennis doubles: analysis of the formations used by the serving and receiving teams. *Int J Phys Educ Fit Sport*, *9*(2), 45-50.

Kovalchik, S. A., Bane, M. K. & Reid, M. (2017). Getting to the top: an analysis of 25 years of career rankings trajectories for professional women's tennis. *Journal of sports sciences*, *35*(19), 1904-1910.

Krishnan Rajbahadur, G., Wang, S., Kamei, Y., & Hassan, A. E. (2022). The impact of feature importance methods on the interpretation of defect classifiers. *arXiv e-prints*, arXiv-2202.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26, p. 13). New York: Springer.

Martínez-Gallego, R., Crespo, M. & Jiménez, J. (2021a). Analysis of the differences in serve effectiveness between Billie Jean King Cup (former Fed Cup) and Davis Cup doubles tennis matches. *International Journal of Sports Science & Coaching, 16*(3), 777-783.

Martínez-Gallego, R., Ramón-Llín, J. & Crespo, M. (2021b). A Cluster Analysis Approach to Profile Men and Women's Volley Positions in Professional Tennis Matches (Doubles). Sustainability, 13*, 6370.*

Martinez-Gallego, R., Salvador, S. M., Luján, J. F. G., Reid, M., Ramón-Llin, J., & Crespo, M. (2021c). Challenging serve myths in doubles tennis. *International Journal of Sports Science & Coaching*, *16*(6), 1305-1311.

Mecheri, S., Rioult, F., Mantel, B., Kauffmann, F. & Benguigui, N. (2016). The serve impact in tennis: First large-scale study of big Hawk-Eye data. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *9*(5), 310-325. https://doi.org/10.1002/sam.11316

Reid, M., & Duffield, R. (2014). The development of fatigue during match-play tennis. *British journal of sports medicine*, *48*(Suppl 1), i7.

Reid, M., McMurtrie, D. & Crespo, M. (2010). The relationship between match statistics and top 100 ranking in professional men's tennis. *International Journal of Performance Analysis in Sport*, 10(2), 131-138.

Reid, M., Morgan, S. & Whiteside, D. (2016). Match play characteristics of Grand Slam tennis: implications for training and conditioning. *Journal of sports sciences*, 34(19), 1791-1798.

Rioult, F., Mecheri, S., Mantel, B., Kauffmann, F., & Benguigui, N. (2015, September). What can hawk-eye data reveal about serve performance in tennis?. In *MLSA15-Machine Learning and Data Mining for Sports Analytics, workshop (ECML/PKDD 2015)* (Vol. 1970, pp. 36-45).

Sánchez-Pay, A., Ortega-Soto, J.A., Courel-Ibáñez, J. & Sánchez-Alcaraz, B.J. (2022). El servicio en tenis profesional femenino: influencia del ranking y la lateralidad. *Revista internacional de medicina y ciencias de la actividad física y el deporte*, 22(86), 255-267.

Sánchez-Alcaraz Martínez, B. J., Perona Arce, F., & Courel-Ibáñez, J. (2018). Parámetros de Rendimiento en el Servicio y Resto del Top-100 Femenino en Tenis. *Revista Kronos*, *17*(1).

Stănescu, R. (2016). Comparative Study on the Efficiency of Women Tennis Players' First Serve at Australian Open 2014 and 2015. In V. Grigore, M. Stanescu, & M. Paunescu (Eds.), *Physical Education, Sport and Kinetotherapy* - ICPESK 2015, vol 11. European Pcroceedings of Social and Behavioural Sciences (pp. 202-209). Future Academy. https://doi.org/10.15405/epsbs.2016.06.28

van Meurs, E., Buszard, T., Kovalchik, S., Farrow, D., & Reid, M. (2021). Interpersonal coordination in tennis: assessing the positional advantage index with Australian Open Hawkeye data. *International Journal of Performance Analysis in Sport*, *21*(1), 22-32.

Vaverka, F., Nykodym, J., Hendl, J., Zhanel, J., & Zahradnik, D. (2018). Association between serve speed and court surface in tennis. *International Journal of Performance Analysis in Sport*, *18*(2), 262–272. https://doi.org/10.1080/24748668.2018.1467995

Vives, F., Crespo, M., Guzmán, J. F., & Martínez-Gallego, R. (2022). Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology. *International Journal of Performance Analysis in Sport*, *22*(4), 638-648.

Vives, F., Lázaro, J., Guzmán, J. F., Martínez-Gallego, R., & Crespo, M. (2023). Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve. *Applied Sciences*, *13*(24), 13213.

Wei, X., Lucey, P., S. Morgan, S. & Sridharan, S. "Predicting Shot Locations in Tennis Using Spatiotemporal Data," *2013 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Hobart, TAS, Australia, 2013, pp. 1-8, doi: 10.1109/DICTA.2013.6691516.

Wei, X., Lucey, P., Stuart Morgan, Carr, P., Reid, M. & Sridharan, S. (2015). Predicting Serves in Tennis using Style Priors. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15). Association for Computing Machinery, New York, NY, USA, 2207–2215. https://doi.org/10.1145/2783258.2788598

Whiteside, D., & Reid, M. (2017). Spatial characteristics of professional tennis serves with implications for serving aces: A machine learning approach. *Journal of sports sciences*, *35*(7), 648-654.

Whiteside, D., Elliott B.C., Lay B.S. & Reid, M. (2013). The effect of age on discrete kinematics of the elite female tennis serve. *J Appl Biomech*;29(5):573-82

Whiteside, D., Elliott, B.C., Lay, B.S., & Reid, M. (2015). Coordination and variability in the elite female tennis serve. *Journal of Sports Sciences*, 33*, 675 - 686.*

7. CONCLUSIONES

Las principales conclusiones que se obtuvieron por estudio fueron las siguientes:

Estudio 1: *Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology.*

- Se observó que los primeros saques más rápidos fueron los dirigidos a la zona de la T, por contra los más lentos fueron los dirigidos a la zona Abierta.
- La velocidad media del primer servicio en dobles es cercana a los 180km/h y varía en función de la dirección del servicio.
- La posición del sacador varía según la dirección del servicio y si se trata de un primer o segundo servicio, con un claro componente táctico.
- Los primeros saques más efectivos son los dirigidos a las zonas Abierta y T, mientras que no se encontraron diferencias en los segundos saques independientemente de la dirección y del lado.

Estudio 2: *Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve*.

- Se identificaron las características más importantes para lograr un *ace* en la modalidad de dobles profesionales masculino, resultando el ángulo del servicio y la distancia del bote de la pelota a la línea lateral como las variables más relevantes.
- Un ángulo de servicio óptimo entre 5.7º y 8.7º se asoció significativamente con una mayor probabilidad de un saque directo, coincidiendo con investigaciones previas en la disciplina individual.
- Mantener una distancia del bote de la pelota a la línea lateral entre 0 y 28 cm también aumentó la probabilidad de éxito en la consecución de un saque directo.
- La velocidad del servicio mostró una influencia menor en comparación con la disciplina individual, sugiriendo que otros factores son más determinantes en el dobles masculino.

Estudio 3: *Serve Strategy in Professional Women's Doubles Tennis - A Data-Driven Perspective.*

- Se identificaron el ángulo de servicio, la velocidad y la distancia del bote de la pelota a la línea lateral las características más importantes para lograr un *ace* en la modalidad de dobles profesionales femenino.
- Se encontró que el ángulo de servicio es la variable más determinante para lograr un saque directo en el tenis de dobles femenino.
- La velocidad del servicio y la distancia a las líneas laterales también se relacionaron con una mayor probabilidad de realizar saques directos, aunque en menor medida.
- Los valores recomendados del ángulo de servicio son muy similares en la categoría femenina y masculina, mientras que la velocidad y la distancia del bote de la pelota a la línea son significativamente diferentes.

8. LIMITACIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN.

A pesar de que se intentó abarcar en amplitud y profundidad las características principales del servicio en el tenis profesional, tanto en la categoría femenina como en la masculina, existen ciertas limitaciones que deberían tenerse en cuenta. Estas limitaciones también podrían convertirse en futuras líneas de investigación en esta área de estudio en particular.

Una de las principales limitaciones de estos estudios es la escasez de investigaciones previas en la disciplina de dobles. Existen numerosos estudios de todo tipo sobre el tenis individual, pero hay muy poca literatura sobre el tenis de dobles profesional. Es por este motivo por el que en muchas ocasiones tuvimos que basarnos en estudios previos de tenis individual.

Otra limitación del estudio es el uso exclusivo de datos procedentes de la copa Davis y la copa Billie Jean King, lo que podría restringir la generalización de los resultados a otros contextos competitivos en el tenis profesional de dobles. En futuras investigaciones, se podría ampliar el análisis a otras competiciones y niveles de juego, lo que permitiría comprender mejor el saque en la modalidad de dobles y su influencia en el rendimiento de los jugadores en la competición.

En varios de estos estudios (2 y 3) se ha tratado de determinar la incidencia del servicio en el juego, pero se ha llevado a cabo de forma aislada, ya que no se encontraron diferencias significativas entre los diferentes tipos de servicio que se analizaron. En el futuro, se podrían llevar a cabo investigaciones que abordaran el análisis del saque desde una perspectiva más integral, explorando cómo las distintas variables asociadas al saque influyen en el desarrollo global del juego. Esto permitiría obtener una comprensión más completa de los diversos matices técnico-tácticos presentes en el tenis de dobles y proporcionaría un enfoque más profundo y enriquecedor para el estudio de esta fase del juego.

Por último, otra futura línea de investigación debería ser la interacción entre el servicio y el resto, y cómo esto afecta al resto del juego. Además, en la disciplina de dobles, existe la figura del compañero del sacador, que influye de forma directa tanto en el saque como en el resto. Dadas las características específicas de la disciplina de dobles, estos futuros estudios podrían proporcionar a entrenadores y jugadores herramientas muy útiles para diseñar sesiones de entrenamiento concretas y para tomar decisiones tácticas durante la competición.

9. APLICACIONES PRÁCTICAS

9.1. DISEÑO DE EJERCICIOS PARA LAS SESIONES DE ENTRENAMIENTO

A continuación, se describen unas propuestas de ejercicios basadas en la información obtenida en los estudios previos:

- Propuesta 1

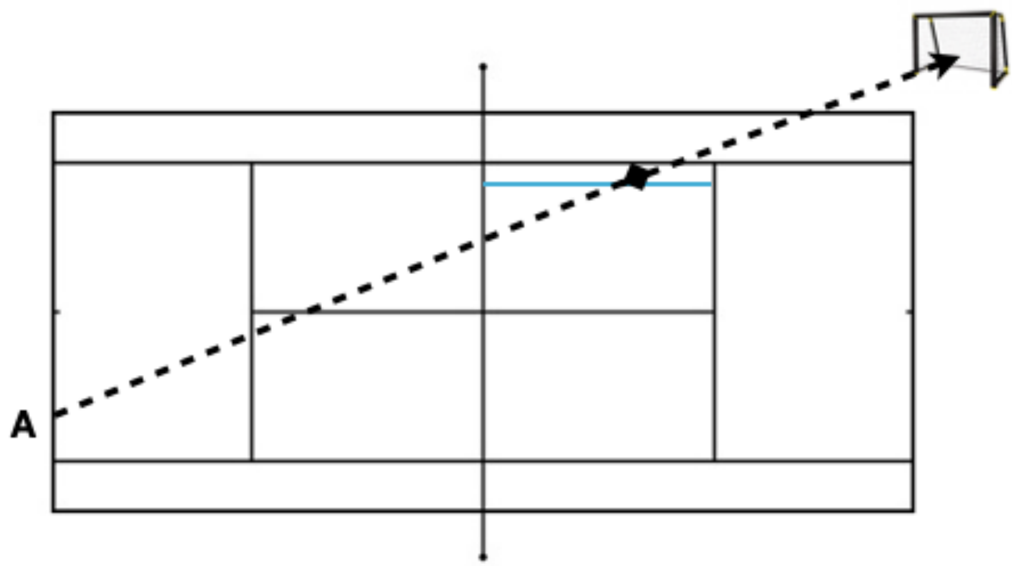

Figure 1. Servicio Abierto.

Objetivo:

Primer Servicio a la zona Abierta del lado del Deuce.

Tipo de Dinámica:

Cerrado

Descripción:

El jugador situado en la posición A realizará primeros servicios a la zona Abierta del lado del *Deuce*. Para facilitar el ángulo de tiro al sacador se colocará una portería (Figura 1) y en el cuadrado de saque de recepción se señalizará con la distancia obtenida en los resultados previos, , siendo 30 cm para los hombres y 50 cm para las mujeres (línea azul). Para potenciar el ejercicio se podría jugar en formato *match tie-break* a 10 puntos, siendo un punto a favor del sacador si entra la pelota en la zona delimitada y en contra si no entra. Este ejercicio se podrá replicar también en el lado de la Ventaja.

- Propuesta 2

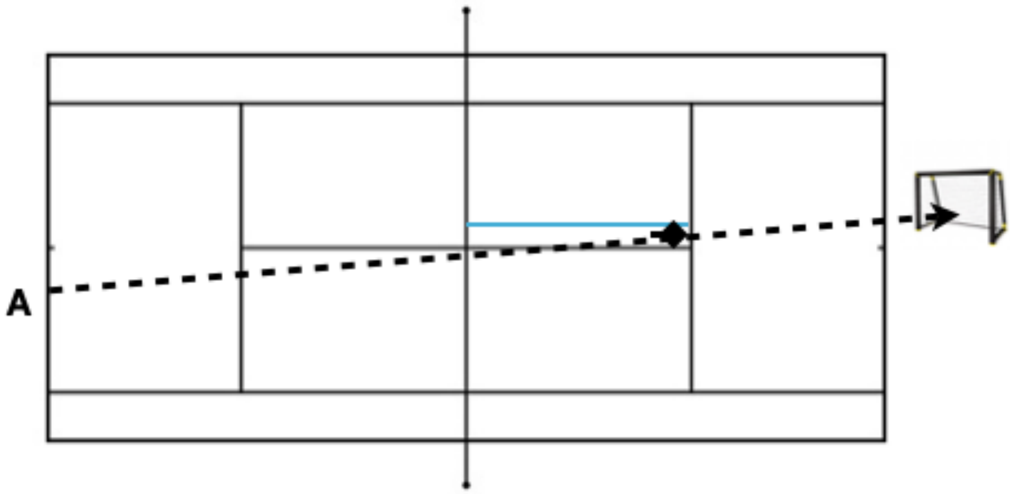

Figure 2. Servicio a la T.

Objetivo:

Primer Servicio a la zona del T del lado del Deuce.

Tipo de Dinámica:

Cerrado

Descripción:

El jugador situado en la posición A realizará primeros servicios a la zona de la T del lado del *Deuce*. Para facilitar el ángulo de tiro al sacador, al igual que se realizaba en el ejercicio anterior, se podrá colocar una portería (Figura 2), y en el cuadrado de saque de recepción se señalizará con la distancia obtenida en los resultados previos, siendo 30 cm para los hombres y 50 cm para las mujeres (línea azul). Para potenciar el ejercicio se podría jugar en formato juego, siendo 15-0 a favor del sacador si entra la pelota en la zona delimitada y en contra si no entra, así sucesivamente hasta completar el tanteo. Este ejercicio se podrá replicar también en el lado de la Ventaja.

- Propuesta 3

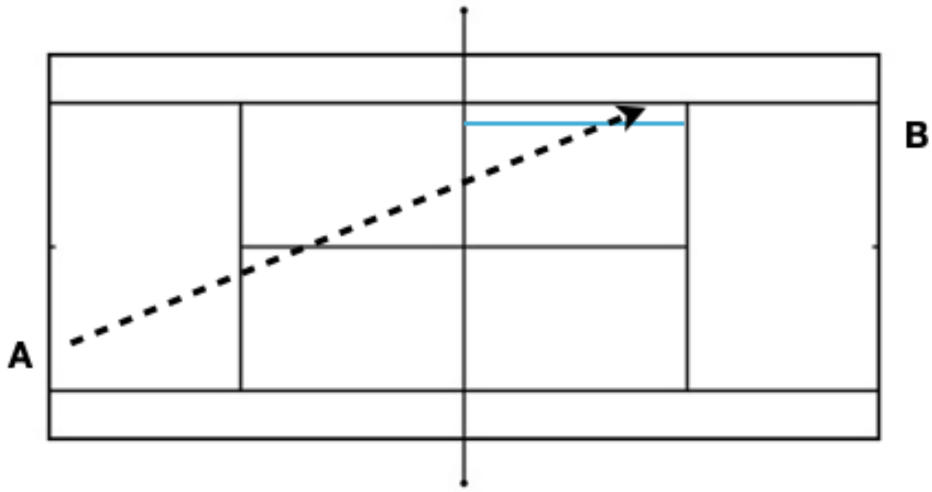

Figure 3. Servicio con Oposición.

Objetivo:

Primer servicio con Oposición

Tipo de Dinámica:

Cerrado

Descripción:

El jugador situado en la posición A realizará primeros servicios a la zona Abierta del lado del *Deuce*, mientras que el jugador situado en la posición B intentará devolver dichos servicios. Se podrá jugar con puntuación a 10 puntos, el sacador si logra un *ace* o un error del restador conseguirá un punto o, por el contrario, si el restador logra devolver el servicio lo sumará él. Este ejercicio se podrá replicar sacando a la zona de la T en el lado del *Deuce*, o del mismo modo, variando la dirección (zona Abierta o a la T) en el lado de la Ventaja.

- Propuesta 4

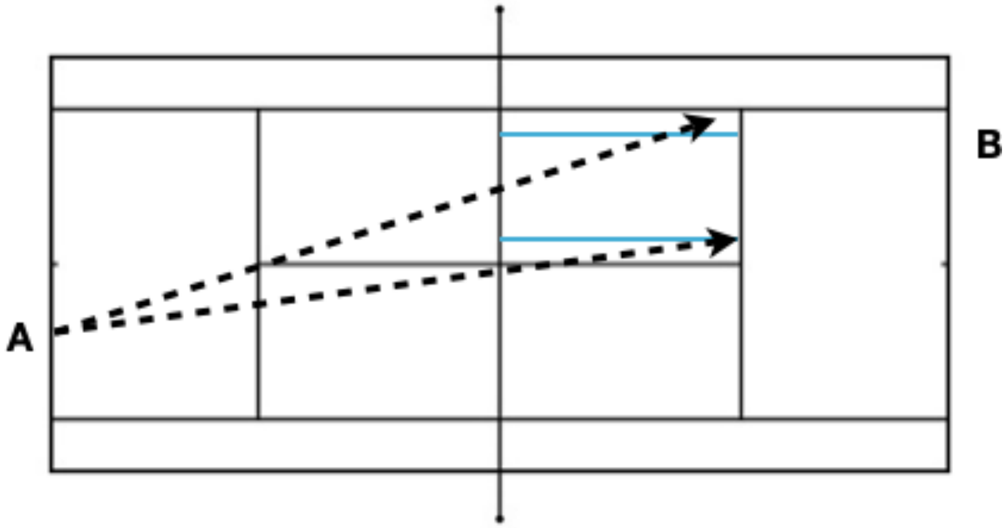

Figure 4. Servicio con Oposición e Incertidumbre

Objetivo:

Primer servicio con Oposición + Incertidumbre

Tipo de Dinámica:

Semi-Abierto

Descripción:

El jugador situado en la posición A podrá realizar primeros servicios a la zona Abierta o a la zona de la T del lado del *Deuce*, el jugador situado en la posición B intentará devolver el servicio. Una de las características más importantes que han mostrado los resultados, son las diferencias significativas en la posición del sacador en función de la dirección del servicio, por tanto, entrenar una posición 'híbrida' donde el sacador pudiera realizar servicios tanto a la zona Abierta como a la zona de la T le proporcionaría una ventaja sobre el restador que podría intuir la zona del servicio en función de la posición del sacador. También podría funcionar esta propuesta en el sentido contrario, entrenando la anticipación del restador en función de la posición del sacador si este ultimo ocupa una determinada colocación al realizar el servicio.

- Propuesta 5

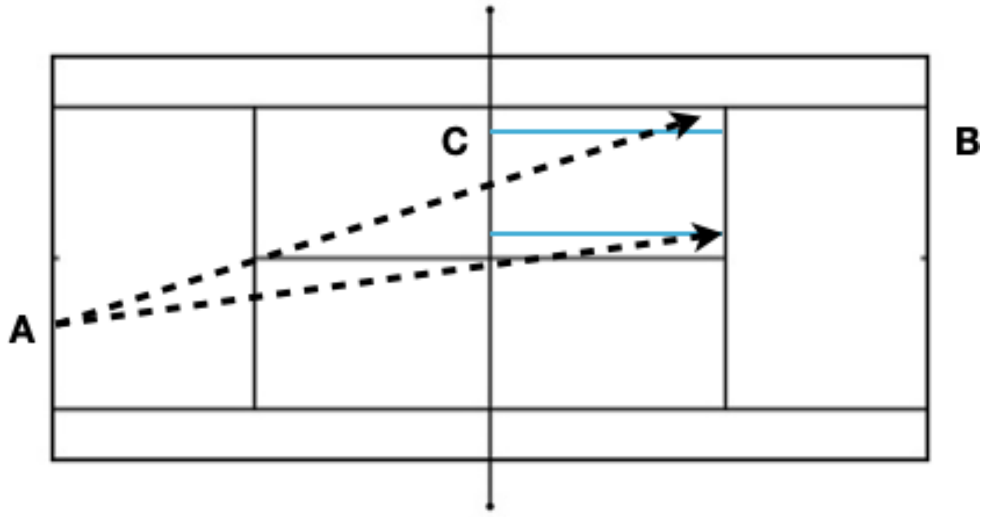

Figure 5. Servicio con Oposición y Jugador en la Red

Objetivo:

Primer servicio con Oposición + Jugador en la Red

Tipo de Dinámica:

Semi-Abierto

Descripción:

El jugador situado en la posición A podrá realizar primeros servicios a la zona Abierta o a la zona de la T del lado del *Deuce*, el jugador situado en la posición B intentará devolver el servicio y además se colocará un tercer jugador situado en la posición C. En este ejercicio se añade la figura del compañero en la red para asemejar lo más posible la acción de juego modificada a la situación real de juego. En este caso también se podría jugar con marcador hasta los 21 puntos, otorgando 3 puntos al saque directo, 2 puntos al error del restador y 1 punto si el jugador de la red consigue una volea, así mismo se podría puntuar de forma inversa si el restador consigue restar superando al jugador de la red con 3 puntos, 2 puntos si el jugador de la red falla la volea y 1 punto si no se logra meter el primer servicio.

- <u>Propuesta 6</u>

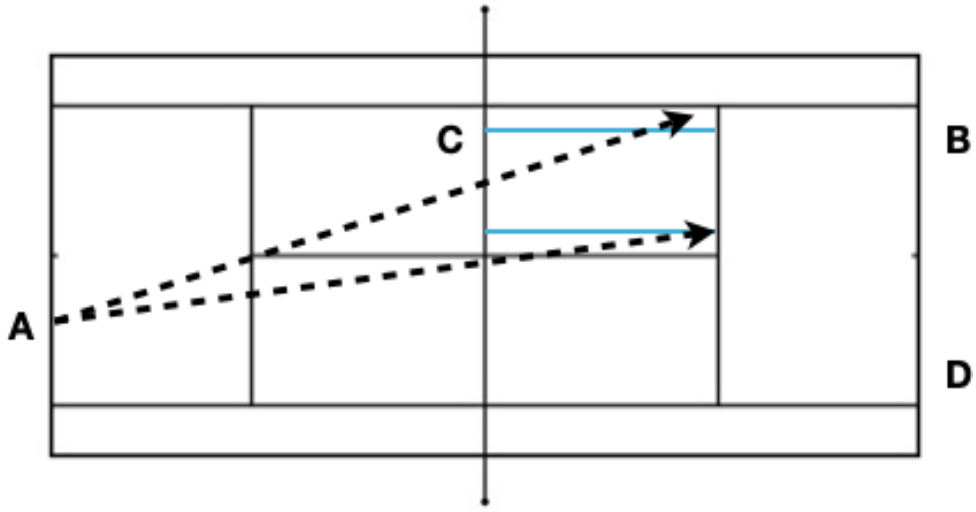

Figure 6. Servicio en situación de Partido

Objetivo:

Partido Modificado

Tipo de Dinámica:

Abierto

Descripción:

Se jugará un partido de dobles con los 4 jugadores que vemos en la figura X en las posiciones A, B, C y D. El formato de juego será un *match tie-break* a 10 puntos, pero por ejemplo el saque directo valdrá 3 puntos, el error del restador siempre que la pelota bote en la zona delimitada 2 puntos y el resto de los puntos 1.

10. ANEXOS

Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology

Fernando Vives, Miguel Crespo, José Francisco Guzmán & Rafael Martínez-Gallego

INTERNATIONAL JOURNAL OF PERFORMANCE ANALYSIS IN SPORT
2022, VOL. 22, NO. 4, 638-648
https://doi.org/10.1080/24748668.2022.2110555

Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology

Fernando Vives [a], Miguel Crespo [b], José Francisco Guzmán [a] and Rafael Martínez-Gallego [a]

[a]Department of Sport and Physical Education, University of Valencia, Valencia, Spain; [b]Head Participation & Education, International Tennis Federation, London, UK

ABSTRACT

Despite the importance of the serve in doubles tennis, there is very little information on the tactical characteristics of this game situation. The aim of this study was to analyse the differences in the spatial-temporal characteristics and the effectiveness of the serves depending on their direction in professional doubles tennis. Using the Hawk-Eye system, 20,419 serves from 96 Davis Cup doubles matches played between 2010 and 2019 were analysed. First serves to the Wide and T zones showed the highest effectiveness. The fastest serves were those played to the T, while the slowest were those played to the Wide zone, except for the second serves on the Deuce side, where the opposite occurred. The serves played to the T were hit from positions closer to the centre of the court, and the time from serve contact to ball bounce was shorter than in the rest of the serves played to other directions. Generally, compared to singles tennis, players served from positions further away from the centre of the court. The data provided in this study have clear practical implications, both for the design of specific training sessions and for the tactical preparation of matches.

ARTICLE HISTORY
Received 4 April 2022
Accepted 1 August 2022

KEYWORDS
Service; tennis; Hawk-Eye; tactics; racquet sports

1. Introduction

The serve is the stroke that starts the point in tennis and has always been considered one of the most important strokes in the game. This has made it the most studied stroke in tennis, with many studies showing its relationship with performance and ranking (Klaassen & Magnus, 2014; O'Donoghue & Ingram, 2001; Reid et al., 2010). In addition, from a strategic and tactics perspective, the serve has been related to different variables that affect the tactics used by players when hitting this stroke, such as the surface (Vaverka & Cernosek, 2016), the importance of the point (Knight & O'Donoghue, 2012; Meffert et al., 2018), and the gender (Reid et al., 2016) or the playing level of the players (Kovalchik & Reid, 2017). Despite the considerable influence of all these factors on the tactics used by the players, previous studies have shown that, in general, the most used serve tactics by players is to take more risks with the first serve and to play more consistently with the second serve (Cross & Pollard, 2011; Martínez-Gallego et al., 2021).

CONTACT Fernando Vives fervial@alumni.uv.es Department of Sport and Physical Education, University of Valencia, Valencia, 46010, Spain

Regarding the doubles discipline, it is worth highlighting its relevance in professional tennis, since it is played in the Davis Cup and Billie Jean King Cup, in all four Grand Slams, as well as in all ATP and WTA circuits tournaments, and even in exhibition events such as the Laver Cup. In fact, with the change of format of the Davis Cup, it has now become the decisive final point of the tie in case of a draw in the singles matches (Martínez-Gallego et al., 2020). Despite this, as compared to the singles game, substantially fewer studies have focused on this discipline. Previous studies that have analysed doubles matches have mainly focused on the analysis of match structure, such as Martínez-Gallego et al. (2020) who studied the game structure and the point completion in men's doubles, Borderias et al. (2021) who compared the activity profile of women's doubles matches on different surfaces, Martínez-Gallego et al. (2021) who evaluated the influence of partner experience and match type on the time structure of matches, or Dell'Edera et al. (2018) who compared certain performance indicators between junior and professional doubles. In relation to the role of the serve in doubles matches, only a few studies have been found. Two of them compared the effectiveness of the serve between male and female teams (Carboch & Kočíb, 2016; Martínez-Gallego et al., 2021). Carboch and Kočíb (2016) found that male teams won 10% more serves than female teams and that they needed more points to the serve to win the match. In addition, they showed that, on hard courts the serve was a greater advantage for both male and female teams, while the effectiveness of the serve on clay court was lower for both genders. Martínez-Gallego et al. (2021) indicated that with the first serve both male and female teams tried to move the receiver serving to the sides to take the initiative of the point, while with the second serve they took lower risks and mainly served to the body. Furthermore, they concluded that male partners were more effective than male partners, especially with the first serve. On the other hand, Kocib et al. (2020) analysed the formations used by the teams. As per the serving team, the most used formations were the classic and the "I" formation. The two most used formations in the other teams were: both players at the back of the court or the back/back combination, highlighting the use of the back/back formation in the second serves compared to the formation of both players at the back of the court.

Although the information provided by these studies is very relevant, it seems necessary to further investigate the characteristics of the serve to understand the aspects that determine its effectiveness in the doubles game. In this context, advances in sports technology are allowing scientists as well as player support teams to record large amounts of match data that provide sound data-based insights allowing deeper understanding of the key performance indicators of the different disciplines. In tennis, one of the most important advances in the evaluation of the technical and tactical performance of tennis players has been the use of tracking systems and, specifically, the analysis of data provided by the Hawk-Eye system (Martínez-Gallego et al., 2021). The high level of precision and the large amount of data recorded through this system offer a wide range of opportunities for research and practice (Mecheri et al., 2016). Specifically, this system allows the tracking of coordinates of the ball and the players during the points, so that variables that are of interest for analysing the performance of players in specific actions can be studied with greater depth and precision.

In singles tennis, some studies have used the Hawk-Eye system to analyse and demonstrate the importance of some spatiotemporal variables related to the serve, such as serve speed and serve location (Mecheri et al., 2016; Whiteside & Reid, 2017). However, other variables such as the hitting position, the ball trajectory time, or the difference in speed before and after the bounce have not been studied in depth. As mentioned above, these variables have been even less studied in doubles, in which only one study use this system to analyse the serve location. Therefore, the aim of this study was to analyse the differences in spatio-temporal characteristics such as speed, server position, time from hit to bounce, loss of speed after bounce and efficiency of first and second serves, depending on the side and its direction in professional men's doubles tennis using data provided by the Hawk-Eye system.

2. Methods

2.1. Data collection

Data in the present work were obtained through the Hawk-Eye system (Hawk-Eye Innovations Ltd., , 2015) used during the competition. The Hawk-Eye system is calibrated before the competition by means of the actual coordinates of the court where it is installed. It uses a total of 10 cameras recording at a frequency of 50–60 Hz that track the movement of the ball and players in four dimensions during points (Cartesian coordinates and time). Independent tests have validated the accuracy and reliability of the system, reporting an average error of 2.6 mm (Reid et al., 2016). Spatio-temporal variables such as the hitting position, the ball bounce position, the service speeds, and velocities before and after the bounce were calculated from the data provided by the system. Player consent for the use of the data during the competition was obtained as part of the tournament agreement.

2.2. Data processing

A total of 29,419 services were analysed from 96 full men's doubles matches played during Davis Cup qualifier ties between 2010 and 2019. The match data were imported into R Statistical Software (R Foundation for Statistical Computing, Vienna, Austria), where the data were processed to obtain the variables of interest and subsequently filtered, and outliers were removed. Means were calculated for each variable by match, obtaining a total of 96 records for each variable, each corresponding to one of the matches.

2.3. Data analysis

The variables analysed were the following:

2.3.1. Independent variables

- Type: first or second serve.
- Court side: side of the court from which the service occurs. Deuce (right) or Advantage (left) side.

- Direction: direction of the service (Figure 1).

2.3.2. Dependent variables

- Speed: average speed of the service.
- Position: position of the server when serving.
- Time: time from the impact to the bounce of the ball.
- Loss of speed: loss of speed of the ball after the bounce.
- Effectiveness: point won by the serving team without the serve having been returned (Martínez-Gallego et al., 2021).

Firstly, a descriptive analysis was carried out and mean and standard deviation were calculated. Assumptions of normality and homogeneity of variances were verified using the Kolmogorov-Smirnov test and the Levene test. The Kruskal – Wallis test and post-hoc pairwise comparison was used to identify differences in the dependent variables, setting the significance level at $p < 0.05$. Effect size was calculated using eta-squared, considering small (0.01–0.06), moderate (0.061–0.14) and large (>0.141) effects (Cohen, 2013).

3. Results

3.1. Spatio-temporal characteristics of the first serve

Table 1 shows the descriptive statistics of the kinematic variables as a function of the direction of the serves. Regarding the first serve on the deuce side, it could be observed that there were significant differences in the speed between the three directions ($H(2) = 119.14$; $p < 0.01$; ц$^2 = 0.41$), being the serves played to the T the ones that presented a higher speed (mean = 188.35 km-h^{-1}), followed by those played to the Body zone (mean = 182.17 km-h^{-1}) and, finally, those played to the Wide zone were the slowest serves (mean = 176.41 km-h^{-1}). Regarding the Position, serves played to the T were hit

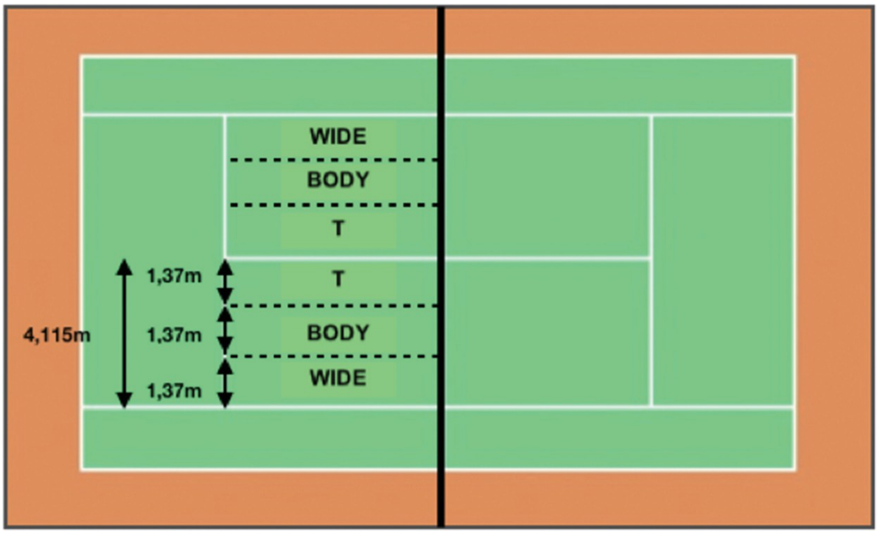

Figure 1. Serve directions.

Table 1. Descriptive statistics of the serve kinematic variables in function of the direction.

			Wide Mean (SD)	Body Mean (SD)	T Mean (SD)	H	*p*	η^2
First serve	Deuce side	Speed (km h^{-1})	176.41 (6.67)$	182.17 (7.09)$	188.35 (7.49)$	119.144	<0.01	0.41
		Position (m)	1.70 (0.32)*	1.70 (0.32)†	1.56 (0.30)*†	13.77	<0.01	0.04
		Time (s)	0.43 (0.01)$	0.40 (0.01)$	0.38 (0.01)$	178.31	<0.01	0.62
		Loss of speed (km h^{-1})	31.51 (2.34)	31.85 (2.76)	31.90 (2.67)	0.68	0.71	0.01
	Advantage side	Speed (km h^{-1})	181.65 (7.96)$	183.67 (8.15)$	186.25 (6.73)$	21.25	<0.01	0.07
		Position (m)	1.85 (0.36)*	1.82 (0.34)†	1.67 (0.40)*†	12.35	<0.01	0.04
		Time (s)	0.42 (0.02)$	0.40 (0.02)$	0.39 (0.01)$	111.73	<0.01	0.39
		Loss of speed (km h^{-1})	31.78 (2.80)	32.19 (2.64)	31.91 (2.81)	1.19	0.55	0.00
Second serve	Deuce side	Speed (km h^{-1})	158.98 (8.62)*#	156.16 (6.51)#	154.70 (7.03)*	20.12	<0.01	0.06
		Position (m)	1.70 (0.40)	1.69 (0.40)	1.62 (0.31)	3.49	0.18	0.01
		Time (s)	0.48 (0.03)*	0.47 (0.02)	0.47 (0.02)*	9.66	<0.01	0.03
		Loss of speed (km h^{-1})	30.36 (2.61)	30.37 (2.27)	30.16 (2.43)	2.48	0.68	0.00
	Advantage side	Speed (km h^{-1})	151.40 (8.46)$	154.35 (6.87)$	162.15 (8.65)$	0.78	<0.01	0.00
		Position (m)	2.07 (0.38)*	2.10 (0.38)†	1.90 (0.46)*†	12.71	<0.01	0.04
		Time (s)	0.50 (0.03)$	0.48 (0.02)$	0.45 (0.02)$	145.08	<0.01	0.51
		Loss of speed (km h^{-1})	29.82 (2.85)*	30.22 (2.43)†	31.06 (2.92)*†	13.41	<0.01	0.04

SD = Standard deviation; $ Significant differences ($p < 0.05$) between W, B and T; * Significant differences ($p < 0.05$) between W and T; † Significant differences ($p < 0.05$) between B and T; # Significant differences ($p < 0.05$) between W and B.

significantly closer to the centre of the court (mean = 1.56 m; H(2) = 13.77; $p < 0.01$; $ц^2$ = 0.04) than those played to the Body (mean = 1.70 m) and to the Wide zone (mean = 1.70 m). Regarding Time, significant differences were found between the three directions of the serves (H(2) = 178.31; $p < 0.01$; $ц^2$ = 0.62), with the serves played to the T showing a lower value of Time (mean = 0.38 s), followed by the serves played to the Body zone (mean = 0.40 s) and, finally, those played to the Wide zone were the serves that showed a lower Time (mean = 0.43 s). In the variable Loss of Speed, no differences were found between the serve directions (F = 0.68; $p = 0.71$; $ц^2$ = 0.01).

Regarding the first serve on the advantage side, it could be observed that there were significant differences in speed between the three directions (H(2) = 21.25; $p < 0.01$; $ц^2$ = 0.07), being the serves played to the T the ones with the highest speed (mean = 186.25 km-h^{-1}), followed by those played to the Body zone (mean = 183.67 km-h^{-1}) and, finally, those played to the Wide zone were the slowest serves (mean = 181.65 km-h^{-1}). Regarding Position, the serves played to the T were hit significantly closer to the centre of the court (mean = 1.67 m; H(2) = 12.35; $p < 0.01$; $ц^2$ = 0.04) than those played to the Body (mean = 1.82 m) and to the Wide zone (mean = 1.85 m). Regarding Time, significant differences were found between the three directions of the serves (H(2) = 111.73; $p < 0.01$; $ц^2$ = 0.39), with the serves played to the T the ones showing a lower value of Time (mean = 0.39 s), followed by the serves played to the Body zone (mean = 0.40 s) and, finally, those played to the Wide zone were the serves that showed a lower Time (mean = 0.42 s). In the Loss of Speed variable, no significant differences were observed between the three directions of the serves.

3.2. Spatio-temporal characteristics of the second serve

As per the second serve on the deuce side, the serves played to the Wide showed a higher speed (mean = 158.98 km-h^{-1}; H(2) = 20.12; $p < 0.01$; $ц^2$ = 0.06) than those played to the Body (mean = 156.16 km-h^{-1}) and to the T (mean = 154.70 km-h^{-1}). Regarding Position, no significant differences were observed between the three directions of the serves. Regarding Time, significant differences were found between the serves played to the Wide and T zones (H(2) = 9.66; $p < 0.01$; $ц^2$ = 0.03), being the serves played to the T the ones that presented a lower value of Time (mean = 0.47 s) as compared to the serves played to the T (mean = 0.48). In the Loss of Speed variable, no significant differences were observed between the three directions.

Finally, significant differences in Speed between the three directions (H(2) = 78.43; $p < 0.01$; $ц^2$ = 0.27) can be observed in the second serve on the advantage side, with the serves played to the T showing the highest speed (mean = 162.15 km-h^{-1}), followed by those played to the Body zone (mean = 154.35 km-h^{-1}) and, finally, those played to the Wide zone were the slowest serves (mean = 151.40 km-h^{-1}). Regarding the Position, serves played to the T, were hit significantly closer to the centre of the court (mean = 1. 90 m; H (2) = 54.28; $p < 0.01$; $ц^2$ = 0.03) than those played Wide (mean = 2.07 m) and to the Body (mean = 2.10 m). As per Time, significant differences were found between the three directions of the serves (H(2) = 145.08; $p < 0.01$; $ц^2$ = 0.51), with the serves played to the T showing a lower value of Time (mean = 0.45 s), followed by the serves played to the Body zone (mean = 0.48 s) and, finally, those played to the Wide zone were the serves that presented a higher Time (mean = 0.50 s). In the variable Loss of Speed, serves played to the T lost more speed (mean = 31.06; H(2) = 13.41; $p < 0.01$; $ц^{2} = 0.04$) than serves played to the Body zone (mean = 30.22) and serves played to the Wide zone (mean = 29.82).

3.3. *Effectiveness of the first serve depending on direction*

Figure 2 shows the effectiveness of the first serve on the deuce and the advantage sides, for each of the directions. Regarding the first serves on the Deuce side, there was a lower effectiveness when the serves were played to the Body (mean = 29.95 %; H(2) = 34.72; $p < 0.01$; $ц^{2} = 0.12$), compared to the serves played to the T (mean = 38.45 %) and those played to the Wide zone (mean = 40.71 %).

Similarly, on the advantage side, there was a lower effectiveness of serves played to the Body (mean = 28.27 %; H(2) = 40.16; $p < 0.01$; $ц^{2} = 0.13$), with respect to those played to the T (39.28 %) and to the Wide zones (37.54 %).

3.4. *Effectiveness of the second serve depending on direction*

Figure 3 shows the effectiveness of the second serve on the deuce and the advantage sides, for each of the directions. As per the second serves on the deuce side, there was a higher effectiveness of the serves played to the Wide zone (mean = 30.74 %; H(2) = 8.24; $p < 0.05$; $ц^{2} = 0.02$), with respect to those played to the T (21.14 %). On the advantage side, no significant differences were found (H(2) = 4.28; $p = 0.12$; $ц^{2} = 0.01$).

4. Discussion

The results found have allowed to observe how the speed of the serve varied for each of the directions, both on the first and second serves on both sides of the court. In all cases, the fastest serves were those played to the T, while the slowest ones were those played to

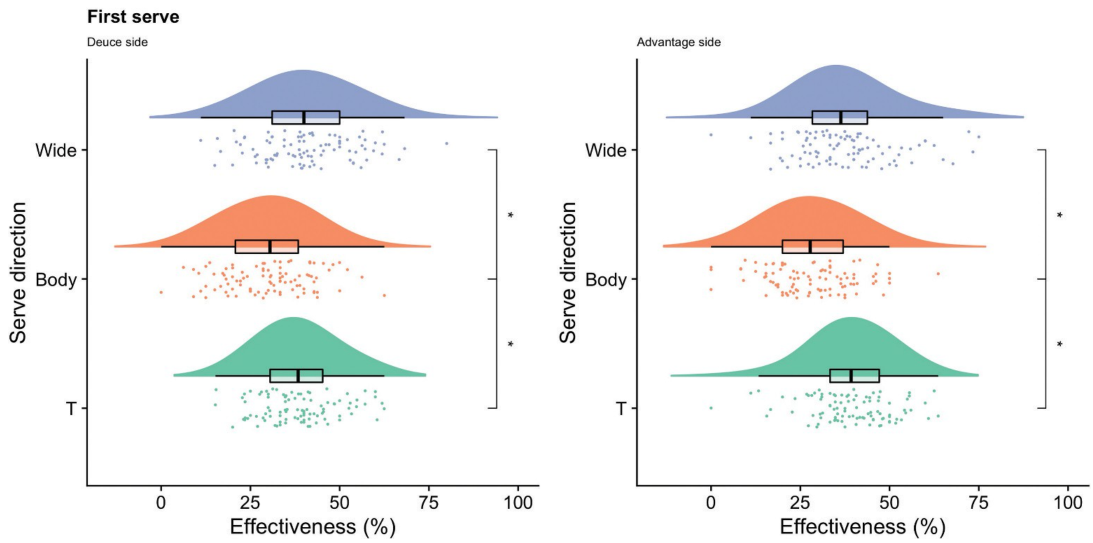

Figure 2. Differences in effectiveness of the first serve.

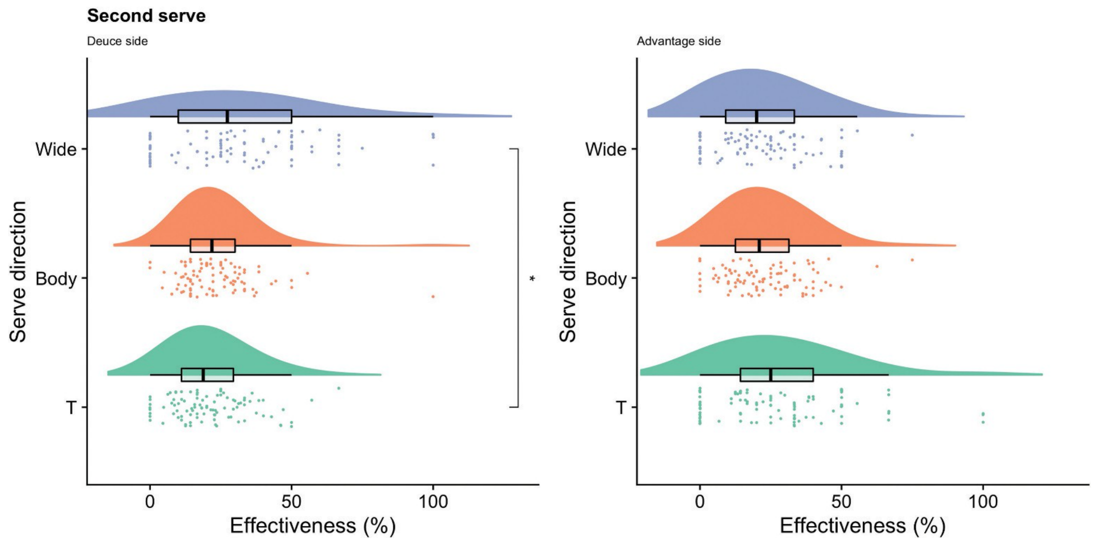

Figure 3. Differences in effectiveness of the second serve.

the Wide zone, except for the second serves on the Deuce side, where the opposite occurred. The serves to these two zones have a clear tactical intention, which is to try to place the ball as far as possible from the receiver's zone to make the return more difficult, or to prevent the receiver from taking the initiative of the point (Martínez-Gallego et al., 2021). However, the fact that the speed of the serves to the Wide zone is much lower is directly related to the type of spin used to hit the serve. To serve to this zone, players tend to use a slice serve, which opens up a greater angle but reduces the speed of the ball (Gillet et al., 2009). In the case of second serves next to the Deuce side, it seems that players increase the speed of the serve in order to surprise the opponent. Previous studies have shown that this is the least used direction for second serves in doubles (Martínez-Gallego et al., 2021), so serving to this area with a higher speed can be a good strategy that, as seen in the results, increase the effectiveness of second serves.

As per the position of the servers, it could be observed that their position tended to be closer to the centre when they played the serves to the T zone. Previous studies have shown the importance of the angle of the trajectory of the ball with respect to the receiver's position (Whiteside & Reid, 2017), therefore, it seemed that players opted for these positions closer to the centre of the court to obtain angles that placed the ball away from the receiver's position when serving to this area. However, previous studies that have analysed the position of servers in the singles game have shown values considerably closer to the centre of the court (Whiteside & Reid, 2017). This may be because the size of the court in doubles is larger than in singles, which forces players to stand further away from the centre of the court to be able to correctly cover the furthest areas of the court for the next shot. This fact may also explain the reason why when servers play second serves to the advantage side, they stand significantly further away from the centre than when they serve to the body and the T. On the one hand, to hit the serve with a greater angle, and on the other hand, to have a better position to cover the court in case the receiver can attack the serve.

As for the time from ball hit to bounce, in all situations, serves played to the T zone showed the shortest time, and those played to the Wide zone showed the longest time. These results seem quite logical considering that the distance travelled by the ball is greater as the bounce occurs further away from the centre of the court (Mecheri et al., 2016). In addition, as noted above, the speed of the serves played to the *T*-zone is higher, which also contributes to this reduction of time.

With respect to the loss of speed of the ball after the bounce, differences have only been found in the second serves on the advantage side, where the serves played to the T zone were the ones that lost the most speed. These differences in speed loss are substantially smaller than the differences in service speed, so this variable might have a low influence on service performance compared to service speed, which has been shown to be a determinant variable (Vaverka et al., 2018).

Regarding the effectiveness of the first serves, on both sides, serves played to the Wide and T zones were the most effective. These data coincide with those obtained in studies that have analysed service effectiveness, both in singles (Mecheri et al., 2016; Whiteside & Reid, 2017) and doubles (Martínez-Gallego et al., 2021). Therefore, as in the singles game, playing the serves to the areas furthest away from the receiver's position seems to be the best strategy to increase the effectiveness of the serve. With respect to the second serves, differences were only found on the deuce side, where serves played to the Wide zone were more effective than those played to the T zone. As indicated above, this could be due to the fact that players serve a lower percentage of serves to the Wide zone (Martínez-Gallego et al., 2021) and with a higher speed, so that when they do so, they surprise the opponent and win a greater number of points.

As it can be observed, the information provided in this study has clear practical implications, both for the training and for the strategical and tactical preparation of matches. When relating the results on serve effectiveness with the rest of the variables studied in the doubles game, it can be noted that when players serve Wide, they try to obtain an advantage by moving the receiver further away from the court, serving at a greater angle and using the slice and topspin effects to a greater extent, except on second serves to the deuce side, where they try to surprise the opponent by increasing the speed. On the other hand, when they serve to the T zone, they try to reduce the receiver's response time by positioning themselves closer to the centre of the court, serving at a higher speed and mainly using a flat serve. Knowing this data, coaches can be more specific when designing and conducting the tasks and drills when working the service in doubles training sessions. In addition, the application of this information to the characteristics of the players themselves and of the opposing teams can help to better prepare the tactics to be used with the service in doubles matches. Despite the new findings provided in our research, this study has also certain limitations. Firstly, the analysis of the serve has been carried out in isolation, without considering its relationship with subsequent actions. This could be an important aspect to consider for further studies, as effective serves are not only those in which the serving team wins the point directly with the serve, but also those in which the serve provides an advantage that allows the point to be won in the following shots. In addition, it could also be interesting to know how the serve influences the characteristics and performance of the return. Finally, it could be of interest to apply this type of analysis not only to professional men's doubles, but also to professional women's doubles, as well as to different categories such as collegiate, junior, seniors, and recreational players.

5. Conclusions

This article analyses the differences in the spatial-temporal characteristics of serves in men's professional doubles tennis according to their direction. The results found showed that the fastest serves were those played to the T, while the slowest were those played to the Wide zone. The serves played to the T were hit from positions closer to the centre of the court, and the time from the impact to the bounce of the ball was shorter than in the rest of the directions. The first serves played to the Wide and T zones were the most effective, with no differences being found in the second serves. These data suggest that when the players serve Wide, they try to obtain an advantage by moving the receiver to positions further away from the court. On the other hand, when they serve to the T zone, they try to reduce the receiver's response time, positioning themselves closer to the centre of the court and serving at a higher speed.

Acknowledgement

The authors wish to acknowledge the contributions of the International Tennis Federation Coaching Department.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Fernando Vives http://orcid.org/0000-0002-4516-2391

References

Borderias, M., Fernandez-Fernandez, J., Martinez-Gallego, R., & Baiget, E. (2021). Comparison of the activity profile during grand slam women's doubles tennis. *International Journal of Performance Analysis in Sport*, *21*(6), 1066–1080. https://doi.org/10.1080/24748668.2021.1972611

Carboch, J., & Kočíb, T. (2016). A comparison of service efficiency between players of male and female doubles at professional tennis tournaments. *Auc Kinanthropologica*, *51*(2), 56–62. https://doi.org/10.14712/23366052.2015.32

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Cross, R., & Pollard, G. (2011). Grand slam men's singles tennis 1995-2009 part 2: Points, games and sets. *ITF Coaching and Sport Science Review*, *2011*(53), 3–6.

Dell'Edera, M., Bertino, L., & Campagnoli, D. (2018). Junior doubles development: An on-court progressive approach based on facts and stats. *ITF Coaching & Sport Science Review*, *26*(75), 4–6. https://doi.org/10.52383/itfcoaching.v26i75.244

Gillet, E., Leroy, D., Thouvarecq, R., & Stein, J. (2009). A notational analysis of elite tennis serve and serve-return strategies on slow surface. *Journal of Strength and Conditioning Research*, *23*(2), 532–539. https://doi.org/10.1519/JSC.0b013e31818efe29

Hawk-eye innovations. (2015) . *Hawk-Eye's accuracy & reliability: Electronic line calling*. Basingstoke.

Klaassen, F., & Magnus, J. R. (2014). *Analyzing wimbledon: The power of statistics*. Oxford University Press.

Knight, G., & O'Donoghue, P. (2012). The probability of winning break points in grand slam men's singles tennis. *European Journal of Sport Science*, *12*(6), 462–468. https://doi.org/10.1080/17461391.2011.577239

Kocib, T., Carboch, J., Cabela, M., & Kresta, J. (2020). Tactics in tennis doubles: Analysis of the formations used by the serving and receiving teams. *International Journal of Physical Education, Fitness and Sports*, *9*(2), 45–50. https://doi.org/10.34256/ijpefs2026

Kovalchik, S. A., & Reid, M. (2017). Comparing matchplay characteristics and physical demands of junior and professional tennis athletes in the era of big data. *Journal of Sports Science & Medicine*, *16*(4), 489–497.

Martínez-Gallego, R., Crespo, M., & Jiménez, J. (2021). Analysis of the differences in serve effectiveness between Billie Jean King cup (former fed cup) and Davis cup doubles tennis matches. *International Journal of Sports Science & Coaching*, *16*(3), 777–783. https://doi.org/10.1177/1747954120988545

Martínez-Gallego, R., Crespo, M., Ramón-Llin, J., Micó, S., & Guzmán, J. F. (2020). Men's doubles professional tennis on hard courts: Game structure and point ending characteristics. *Journal of Human Sport and Exercise*, *15*(3), 633–642. https://doi.org/10.14198/jhse.2020.153.13

Martínez-Gallego, R., Ramón-Llin, J., & Crespo, M. (2021). A cluster analysis approach to profile men and women's volley positions in professional tennis matches (doubles). *Sustainability*, *13*(11), 6370. https://doi.org/10.3390/su13116370

Martínez-Gallego, R., Vives, F., Guzmán, J. F., Ramón-Llin, J., & Crespo, M. (2021). Time structure in men's professional doubles tennis: Does team experience allow finishing the points faster? *International Journal of Performance Analysis in Sport*, *21*(2), 215–225. https://doi.org/10.1080/24748668.2021.1872218

Mecheri, S., Rioult, F., Mantel, B., Kauffmann, F., & Benguigui, N. (2016). The serve impact in tennis: First large-scale study of big hawk-eye data. *Statistical Analysis and Data Mining*, *9*(5), 310–325. https://doi.org/10.1002/sam.11316

Meffert, D., O'Shannessy, C., Born, P., Grambow, R., & Vogt, T. (2018). Tennis serve performances at break points: Approaching practice patterns for coaching. *European Journal of Sport Science*, *18*(8), 1151–1157. https://doi.org/10.1080/17461391.2018.1490821

O'Donoghue, P., & Ingram, B. (2001). A notational analysis of elite tennis strategy. *Journal of Sports Sciences*, *19*(2), 107–115. https://doi.org/10.1080/026404101300036299

Reid, M., McMurtrie, D., & Crespo, M. (2010). Title: The relationship between match statistics and top 100 ranking in professional men's tennis. *International Journal of Performance Analysis in Sport*, *10*(2), 131–138. https://doi.org/10.1080/24748668.2010.11868509

Reid, M., Morgan, S., & Whiteside, D. (2016). Matchplay characteristics of grand slam tennis: Implications for training and conditioning. *Journal of Sports Sciences*, *34*(19), 1791–1798. https://doi.org/10.1080/02640414.2016.1139161

Vaverka, F., & Cernosek, M. (2016). Quantitative assessment of the serve speed in tennis. *Sports Biomechanics*, *15*(1), 48–60. https://doi.org/10.1080/14763141.2015.1123763

Vaverka, F., Nykodym, J., Hendl, J., Zhanel, J., & Zahradnik, D. (2018). Association between serve speed and court surface in tennis. *International Journal of Performance Analysis in Sport*, *18*(2), 262–272. https://doi.org/10.1080/24748668.2018.1467995

Whiteside, D., & Reid, M. (2017). Spatial characteristics of professional tennis serves with implications for serving aces: A machine learning approach. *Journal of Sports Sciences*, *35*(7), 648–654. https://doi.org/10.1080/02640414.2016.1183805

applied sciences

MDPI

Article

Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve

Fernando Vives [1], Javier Lázaro [2], José Francisco Guzmán [1], Rafael Martínez-Gallego [1,*] and Miguel Crespo [3]

[1] Department of Sport and Physical Education, University of Valencia, 46010 Valencia, Spain; fervial@alumni.uv.es (F.V.); jose.f.guzman@uv.es (J.F.G.)
[2] Independent Researcher, 46100 Valencia, Spain; javilazarogarcia@gmail.com
[3] Development Department, International Tennis Federation, London SW15 5XZ, UK; miguel.crespo@itftennis.com
* Correspondence: rafael.martinez-gallego@uv.es

Abstract: This study used a novel machine learning approach to uncover key serve variables that maximize effectiveness in men's professional doubles tennis. A large dataset of 14,146 serves from 97 Davis Cup doubles matches played between 2010 and 2019 was analyzed using explainable AI techniques. The angle and distance from the bounce to the sidelines of the serves were found to best distinguish the points won with aces from rallies lasting more than three strokes. Optimal serve angle ranges of 5.7–8.7° substantially increased the probability of serving an ace by over 80%, compared to around 30% when serving used more central angles. Lateral bounce distances of 0–28 cm from the sidelines also boosted the ace probability by over 50%. The serve speed was shown to have less influence on serve effectiveness as compared to singles tennis, with velocities above 187 km h^{-1} only increasing the probability of serving an ace by 10%. These findings have important practical implications for the tactical decision-making and technical training of serves in men's professional doubles tennis. The data highlight that the angle and placement of serves are more important than velocity for attaining effective serves in doubles. Coaches and players can use this knowledge to pay special attention to the most important variables in the effectiveness of serves, such as the line distance and angle, in order to maximize the performance of the doubles serve. The novel methodology used in this study provides a valid and reliable way to calculate the efficiency of actions in various sport disciplines using tracking data and machine learning approaches.

Keywords: racquet sports; tactics; performance analysis; sport analytics; tracking technology

check for updates

Citation: Vives, F.; Lázaro, J.; Guzmán, J.F.; Martínez-Gallego, R.; Crespo, M. Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve. *Appl. Sci.* **2023**, *13*, 13213. https://doi.org/10.3390/app132413213

Academic Editors: Peter Dabnichki and Juliana Exel

Received: 5 November 2023
Revised: 3 December 2023
Accepted: 8 December 2023
Published: 13 December 2023

1. Introduction

The serve has become the most decisive stroke in modern professional tennis, particularly in the doubles game [1]. It has been shown to be a key shot for controlling the game and even winning points directly through aces, thus being a critical technical and tactical skill for success at the professional levels of the game [2,3].

Previous research has established that the serve provides a key advantage in professional tennis, and this effect is amplified in doubles matches. Servers demonstrate clear dominance over returners in singles tennis, winning approximately 60% of points lasting fewer than four strokes [4]. Moreover, serve metrics like the second serve points won are among the strongest predictors of overall success [5,6]. The positional dynamics and rules of doubles tennis further accentuate the serve's impact, with data showing an even greater serving advantage for doubles compared to singles, especially in short points [1,7,8]. Additional factors related to the surface speed [9] and team experience [10] also influence professional doubles servers' effectiveness. Apart from the heightened serve advantage in doubles, service tactics are also shaped by contextual information like the score and the returner position, which influence players' decision-making [11]. By leveraging such situational data, servers can increase serve variety, target specific opponent weaknesses, maintain consistency, and increase unpredictability [12]. Collectively, these findings underscore the serve as a critically decisive stroke with major tactical and strategic implications in high-level doubles tennis.

The introduction of tracking technology like Hawk-Eye in tennis has enabled access to rich datasets, which researchers have utilized to predict game indicators and parameters especially related to the serve [6,13]. For singles tennis, an analysis of serving patterns shows accuracy measures are more impactful than speed [11]. For doubles, tracking data reveals that the serving placement and direction strongly influence the effectiveness of and anticipation by returners [14,15].

Machine learning techniques have also been applied in recent tennis literature leveraging these large datasets [16]. Broad applications include the stroke classification [17], an automated recognition of movement patterns [18,19], evaluating player performance [20,21], and modeling shot variety and effectiveness [22]. Significant attention has focused on predicting serve characteristics and overall match outcomes. The different modeling approaches explored include regression, point-based comparisons, neural networks, random forests, and ensemble methods [23–34].

In summary, ample research has investigated match prediction in tennis using diverse machine learning techniques on increasing volumes of tracking data. However, doubles tennis has received less specific focus, representing an open area for deeper investigation, particularly around service analytics. Therefore, the aim of this study was to uncover the variables associated with a greater serve effectiveness in men's professional doubles matches using a large dataset from Davis Cup doubles matches and machine learning techniques. The results obtained highlight the angle and the distance of the bounce of the ball from the line as the most important and decisive variables related to the effectiveness of serves.

2. Materials and Methods

2.1. Sample

A total of 14,146 serves were analysed from 97 full men's doubles matches played during the Davis Cup (qualifying ties) between 2010 and 2019. Our study featured the participation of 123 teams and 160 players from 34 different countries, with an average age of 30.03 ± 4.73 years.

2.2. Instruments

The data utilized in this study were obtained from the Hawk-Eye system [35] deployed during Davis Cup events. It consists of a group of ten cameras which are placed around the circumference of the court and capture and record the trajectory of the ball and the players' movement on the court [14]. The Hawk-Eye System is well-known for its high accuracy and reliability in tracking and analyzing tennis matches. It is widely used in professional tennis and has provided data for numerous research papers [6,11,14,15].

2.3. Procedure

A complete E2E AI pipeline was built to experiment with data processing and explainable artificial intelligence (XAI):

1. Experimental setup.
2. Data processing.
3. Training the deep learning model.
4. Feature importance algorithms.
5. Probabilistic and statistical analysis. Values selection based on maximizing the desired effectiveness and minimizing the undesired one.

6. Synthetic dataset generation, and simulation based on the predictions of this dataset and an analysis of improvements in effectiveness.

2.3.1. Experimental Setup

The XAI process can be summarized in steps 2 to 6. A pipeline that included all xAI processes from data processing to training the deep learning model was built to calculate the feature, to select values, and to test the results on a synthetic dataset. The whole process is summarized in Figure 1.

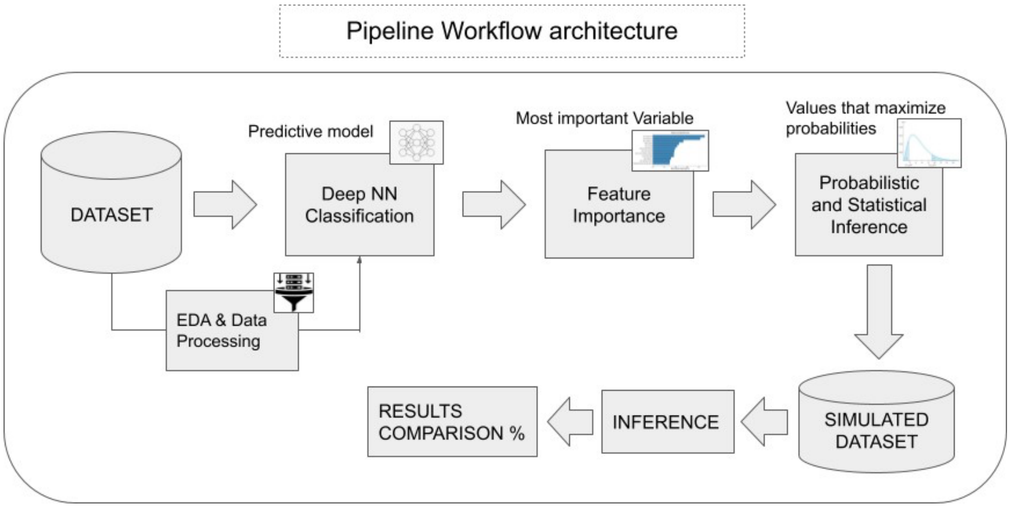

Figure 1. Pipeline workflow architecture.

2.3.2. Data Processing

Dataset

The dataset consisted of the following variables:

Independent variables:

- Court side: Side of the court from where the serve is played. This variable was used to divide the dataset and perform the analysis for each side of the court.
- Effectiveness:
    - Type 1: The point finishes with one shot.
    - Type 2: The point finishes with two shots.
    - Type 3: The point finishes with three shots.
    - Type 4: The point finishes with four shots.

Dependent variables:

- Speed: Mean speed of the serve.
- Position: Position of the server when hitting the serve (A in Figure 2).
- Time: Time between ball impact and ball bounce.
- Speed loss: Loss of speed of the ball after its bounce.
- Impact Z: Height of the ball at impact.
- Net clearance: Height of the ball when passing over the net.
- δ: Serve angle (Figure 2).
- β: Vertical projection angle.
- dL: Distance from ball bounce to the sideline of the service box

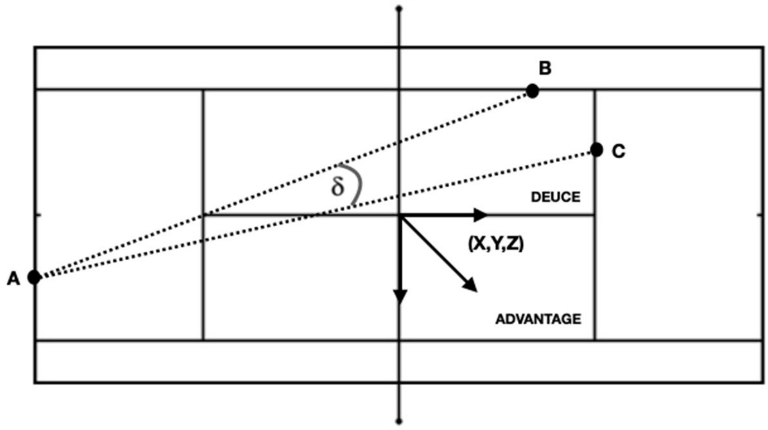

Figure 2. Angle of the serve. δ: Serve angle; A: Position of the server when hitting the serve; B: ball direction; C: central point of the service box line.

EDA and Data Processing

After some exploratory data analysis (correlation matrices, boxplots, histograms, kernel density estimation plots (KDE), etc.), the data were prepared to apply machine learning classification algorithms. Some variables were not informative as their variance was equal or close to 0. The main data processing steps were the following:

1. Oversample the weak target samples to solve the dataset "class unbalancing" problems.
2. Feature selection was conducted using a combination of tennis knowledge criteria together with correlation analysis and feature selection algorithms.
3. Outliers were detected based on statistic methods and were filtered afterwards.

Furthermore, some features did not follow a simple normal distribution but a mixture of gaussians. The collinearity issues between variables, which did not affect the model performance but affected the interpretability of the results, were also solved. Thus, feature importance was applied only with uncorrelated features. This process was conducted separately for δ and dL.

2.3.3. Training the Deep Learning Model

The dataset was divided into train (70%), validation (15%) and test (15%) before applying any data processing, keeping the test and validation subsets completely independent from data leaks from the training process. The validation dataset was used to apply callbacks such as early stopping, the model check pointer, etc. Cross-validation was used to validate the model. For validation, the balanced accuracy score and the loss of each p were monitored. For the final evaluation of the model, the accuracy and balanced accuracy (as well as F1 score, precision and recall) on the test subset were calculated.

AutoGluon v0.5.2 and FastAI v2.5.6 were utilized in this study for this purpose. AutoGluon presents commendable tools for explanation and automated feature enhancement. However, it exhibited a lack of the flexibility required for this research, as certain outcomes were inconsistent. Consequently, it was decided to employ FastAI for the analysis of tabular data.

The implementation of AutoGluon was straightforward, as this tool undertakes the preprocessing of the dataset, trains multiple models, selects the optimal one, assesses feature importance, and employs various XAI algorithms to deliver global explanations for the models.

On the other hand, FastAI is a tool that offers functionalities complimentary to those of PyTorch and AutoGluon. It provides certain data processing functionalities while facilitating the fine-tuning training process. Although it demands a greater degree of software development, it offers advantageous high-level flexibility options. These include the ability to select the number of hidden layers, determine the loss function, and automate the training process, commonly known as "babysitting", through the application of beneficial callbacks such as a state-of-the-art learning rate finder, early stopping, and model checkpointing.

The backbone architecture for the deep learning model and the loss function included in the following steps are the following:

Backbone Network: A tabular model with 5 layers of feature size: 256, 128, 128, 128, 64. The basic layer block (named LinBnDrop) is formed by the following transformations:

- Linear Layer (torch.nn.linear);
- Rectified Linear Unit—ReLU (torch.nn.relu);
- Batch Normalization—BatchNorm1d (torch.nn.BatchNorm1d).

The determination of the optimal number of layers and feature size were driven by experimentation. A standard backbone for tabular data was initially employed in the investigation. For the dataset under consideration, it was observed that a shallow deep neural network (DNN) possessed an adequate capacity to extract relevant patterns. The introduction of additional layers resulted in overfitting the model, whereas a reduction in the number of layers led to underfitting. Similar considerations were applied to feature sizes, commencing with a standard size and subsequently adjusting them iteratively until the size yielding optimal performance on the dataset was identified.

Loss function: Focal loss flat [36]. The focal loss works especially well with imbalanced data as it adapts its weights to focus learning on hard misclassified examples. The primary rationale behind its utilization stems from the presence of imbalanced classes in the dataset. Additionally, an exploration of the flat cross entropy loss was conducted, revealing a comparable performance to the focal loss. However, it was observed that the flat cross entropy loss exhibited a slightly diminished robustness in addressing class imbalances when compared to the focal loss on our dataset.

Optimization algorithm: The Adam optimization method [37] was employed in this study. Adam is characterized as an adaptive learning rate optimization algorithm, with its principal hyperparameters being the learning rate and momentum. The rationale behind incorporating momentum in conjunction with re-scaling lacks clear theoretical motivation. Nevertheless, Adam is widely acknowledged for its robustness across a range of hyperparameter choices.

Callbacks: Several callbacks were implemented to enhance training performance in this study. Early stopping was employed to halt training when the model ceased to exhibit improvement over the last epochs. The model checkpoint callback was also utilized to create copies of the model at specific intervals during training. This allowed for the subsequent selection of the best-performing model among the saved copies.

2.3.4. Feature Importance Algorithms for Feature Explanation

Feature importance algorithms were used to select the variables that have a higher impact on the outcome of the models. Feature permutation importance and SHAP summary plots are the two algorithms commonly accepted in the scientific community [38,39].

Feature permutation importance [39] is based on repeated permutations of the outcome vector to estimate the distribution of measured importance for each variable in a non-informative setting.

SHAP assigns each feature an importance value for a particular prediction [40]. SHAP values attribute to each feature the change in the expected model prediction when conditioning on that feature.

2.3.5. Probabilistic and Statistical Analysis: Values Selection Based on Maximizing the Desired Effectiveness and Minimizing the Undesired One

Shapley values and its variants are adequate "local model-agnostic methods", which provide interesting insights of a single prediction. Global explanation methods can help understand the model through visualizations and explanations of the final model weights [40]. As the aim of the study was to understand how to maximize the serve effectiveness (how and why specific values of the selected features affect the predicted target variables), the goal was to calculate which values threshold of the selected features maximize the likelihood of effectiveness type 1 and minimize the likelihood of effectiveness type 4. For this purpose, a novel semiautomated algorithm based on classical statistics was developed. The first steps of this process were the following:

- First, a smooth probability distribution of the data using the kernel density estimation algorithm (KDE), together with statistical estimators, was calculated for each selected feature. The estimators were mainly the mean, the MLE, and percentiles 5 and 95 for each type of effectiveness.
- Then, the values were mostly identified by selecting the area of the plot which combined a higher density of the points with the desired result (type 1) and a lower density of the undesired result (type 4). The statistical estimators also helped to understand the probability distribution.

2.3.6. Synthetic Dataset Generation and Simulation Based on the Predictions on the Dataset and Analysis of Improvements in Effectiveness

In the final step, a synthetic dataset using the selected values was generated, and the prediction was simulated. Since the model reached 93% f1-score, the predictions were reliable, at least for this specific dataset. By doing this, the approach was validated, and values could be adjusted to improve the rate of effectiveness. The next steps in the process were the following:

- The values outside the desired selected threshold were substituted by random values inside it and, therefore, a synthetic dataset was generated.
- Then, predictions on the whole dataset were performed using the deep learning model previously trained. Then, the final rates of Ef1 and Ef4 were calculated.

It is important to comment that the synthetic dataset assumes a role in the validation of the explanatory algorithm. Following the application of "feature importance" algorithms to identify the most significant variables, an assessment is conducted to determine values that contribute to increased efficiency, relying on insights derived from probability density plots. The values selected are oriented towards maximizing the probability of efficiency 1 within the confines of the real dataset.

To validate this optimization, the original dataset undergoes modification, wherein the values of each variable are adjusted in accordance with the proposed recommendations. Specifically, the dataset is 'clipped', entailing the substitution of values outside the established range limits with randomly generated values falling within the recommended limits. Subsequently, the trained model is employed to generate new predictions on this adjusted dataset. This methodology serves to validate the potential impact on efficiency, providing insights into how the metric might evolve when utilizing the recommended values for the selected variables.

Also, it should be noted that the most relevant part is the one described in step 2, where the threshold values for each variable were selected. Steps 3 and 4 also helped with the validation and fine-tuning of the selection but were mainly an evaluation of the final results to quantify the success of the process.

To efficiently perform multiple experimentation scenarios, all the processes were automated, except one aspect of the second step, in which the areas that maximize and minimize the target variable in the desired way were selected.

3. Results

3.1. Training of a Deep Neural Network

Table 1 displays the testing data metrics of the dataset for different types of effectiveness. As it can be observed, the model achieved high performance metrics for the dataset of type 1 and of type 4, which were f1-score 0.95 and 0.94 for each class. The overall classification accuracy of the evaluation test set was 0.94. Therefore, this model, which serves as the basis for the other algorithms, has been shown to be reliable for defining the characteristics of the serve when points end in an ace and those that last four or more shots.

Table 1. Sklearn classification report type 1 vs. type 4—deuce and advantage.

Effectiveness	Precision	Recall	F1-Score	Macro Average	Weighted Average
Type 1	0.91	0.99	0.95	0.95	0.94
Type 4	0.99	0.89	0.94	0.95	0.94

3.2. Calculation of the Most Relevant Variables

Once the reliability of the model in distinguishing between points that end with an ace and points that last for four or more strokes was established, the next step was to calculate the serve variables that had the greatest relevance for each type of effectiveness. Figures 3–5 depict the mean SHAP values, which indicate the average impact on the model outputs for each variable on the deuce and advantage sides, respectively. As it can be observed, the results of the most important variables were very similar for both sides. Regardless of the serving side, the variables that had the most relevance in the model were δ and dL. Additionally, "Speed" was the other variable that had significant importance in the developed model.

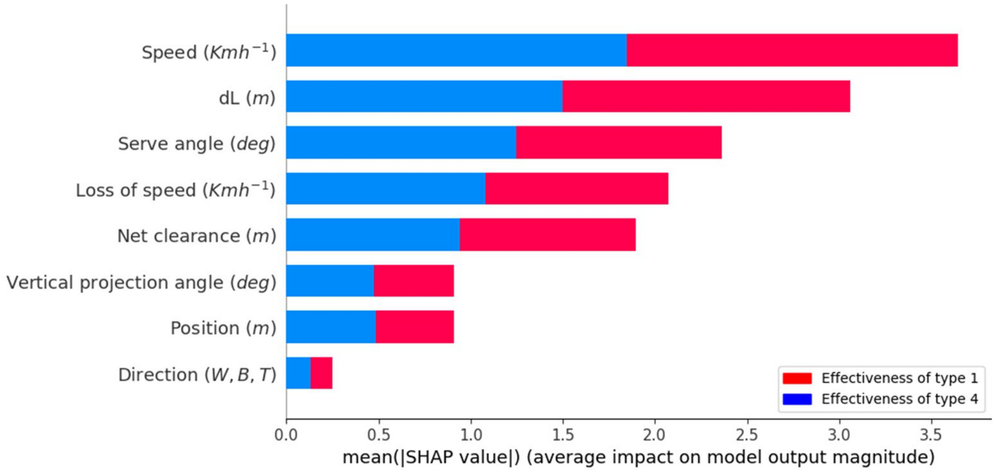

Figure 3. Effectiveness of type 1 vs. type 4 in the DEUCE side.

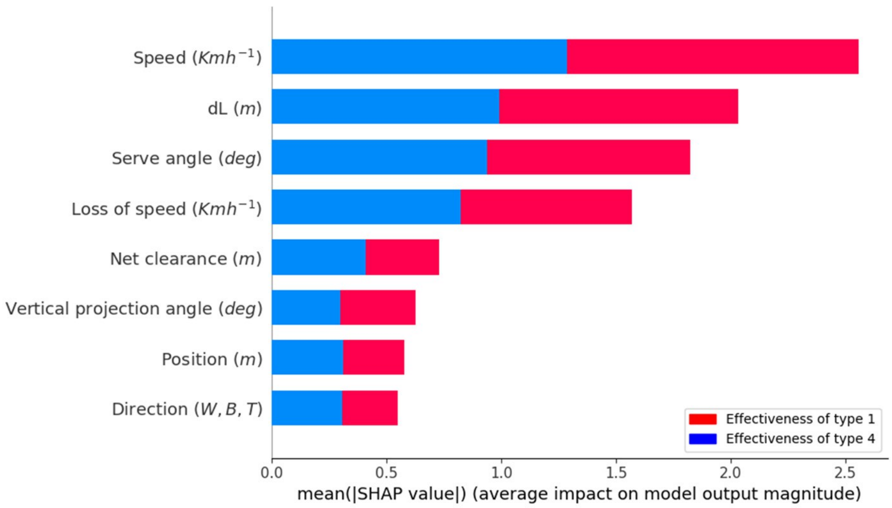

Figure 4. Effectiveness of type 1 vs. type 4 in the ADVANTAGE side.

The mean SHAP values show the average impact on the model outputs for each variable. The results of the most important variables were very similar for the deuce and advantage sides. After separating highly correlated variables (δ and dL) the SHAP values summary, it was shown that in both cases, δ and dL were the most important variables. The algorithm showed that the Speed is the next most important variable for the model.

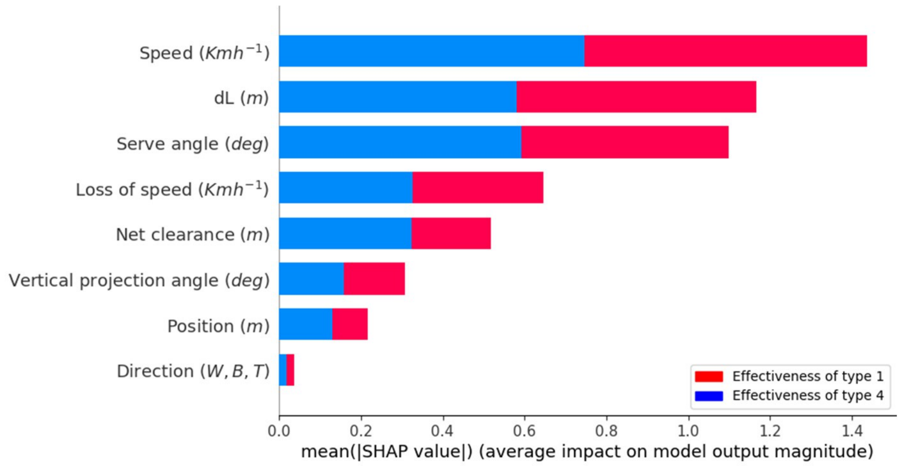

Figure 5. Effectiveness of type 1 vs. type 4 in BOTH sides.

3.3. Values That Maximize the Effectiveness of Type 1 (and Minimize the Effectiveness of Type 4): That Is, Recommended Values

Table 2 shows the distributions of the variables dL, δ, and Speed of the serves for points ending with an ace or with more than three shots for both sides of the court. Regarding dL, aces showed a similar distribution on both sides of the court. However, the serves for points with four or more shots showed significant differences in the distribution of their population. It can be observed that on the advantage side, a greater number of serves with higher values of dL were found.

Table 2. Descriptive values of the variables.

			dL				Serve Angle				Speed			
	Effectiveness	Population (%)	MaxLike (m)	Percentils (m)	Mean	SD	MaxLike (deg)	Percentils (deg)	Mean	SD	MaxLike ($km\ h^{-1}$)	Percentils (deg)	Mean	SD
DEUCE SIDE	Type 1	9.68	0.19	0.04–1.41	0.41	0.41	5.78	2.04–7.78	5.76	1.52	179.86	159.01–210.86	185.72	16.85
	Type 2	25.15	0.63	0.21–1.93	0.95	0.54	4.92	0.87–6.31	3.71	1.76	186.24	160.89–205.72	184.20	14.15
	Type 3	19.58	0.59	0.22–1.97	1.02	0.57	4.37	0.76–6.19	3.44	1.77	184.82	159.01–204.54	183.13	14.31
	Type 4	35.54	0.77	0.28–1.93	1.05	0.52	4.67	0.86–6.13	3.45	1.70	180.57	156.56–202.56	180.39	14.32
ADVANTAGE SIDE	Type 1	9.17	0.18	0.03–1.4	0.40	0.42	5.85	2.11–7.51	5.62	1.43	193.91	162.58–211.39	189.23	16.44
	Type 2	24.34	0.52	0.18–1.95	0.99	0.57	5.06	0.84–6.32	3.66	1.78	189.45	155.56–206.65	185.25	14.67
	Type 3	19.03	0.49	0.19–1.97	1.00	0.59	4.79	0.77–6.42	3.69	1.85	186.48	155.56–204.71	183.03	15.26
	Type 4	37.33	0.78	0.27–1.95	1.07	0.54	4.56	0.86–6.04	3.46	1.67	187.22	153.23–203.38	181.26	15.37
BOTH SIDES	Type 1	9.46	0.18	0.03–1.41	0.40	0.41	5.79	2.08–7.69	5.70	1.47	193.91	160.65–211.20	187.32	16.75
	Type 2	24.78	0.61	0.19–1.94	0.97	0.58	5.05	0.84–6.32	3.69	1.77	187.22	160.26–206.36	184.68	14.40
	Type 3	19.34	0.52	0.2–1.97	1.01	0.58	4.59	0.77–6.25	3.56	1.81	185.73	157.61–204.66	183.08	14.76
	Type 4	36.38	0.76	0.28–1.95	1.06	0.53	4.63	0.86–6.08	3.46	1.69	187.22	155.33–202.94	180.81	14.76

Regarding δ, the distribution of serves where the point lasted more than three strokes was similar on both sides of the court. However, it can be observed that aces have a wider range on the deuce side compared to the advantage side which indicates that, on the deuce side, larger angles would maximize the probability of achieving an ace to a greater extent than on the advantage side.

Regarding Speed, again, the distribution of serves where the point lasted more than three strokes was similar on both sides of the court. However, the recommended values to maximize the probability of serving an ace were slightly lower on the deuce side than on the advantage side.

The recommended values to maximize the probability of serving an ace would be somewhat lower for the deuce than for the advance side. For the other efficiencies, no significant differences were found.

3.4. Generation of a Dataset Substituting Original Values for Recommended Values of the Variable in Question and Prediction of Efficiencies (the Model Trained in Point 1 Was Used to Make the Predictions)

Using the variables associated with the population density graphs, a statistical estimation was performed on the values that would, a priori, maximize the probability of serving an ace and minimize the probability of the point lasting four or more strokes. Table 3 shows the results of the predictions using the trained model described earlier and the synthetic dataset, which resulted from replacing values that fell outside the recommended range with random values within the recommended range. As can be observed, dL values between 0 and 0.28 m increase the probability of achieving an ace to values close to 90%. Similarly, δ values between 5.7° and 8.7° increase the possibility of achieving an ace to values above 80%. On the other hand, Speed values between 187 and 220 km h^{-1} only increase the probability of achieving an ace by approximately 40–45%.

Table 3. Results of the predictions using the trained model.

Variable	Side	Values (Min–Max)	Effectiveness before (%)	Effectiveness after (%)
dL (m)	DEUCE	0–0.28	33.05	83.40
	AD		30.64	81.26
	BOTH		33.49	88.48
Serve angle (deg)	DEUCE	5.7–8.7	33.70	88.68
	AD		29.77	88.21
	BOTH		33.25	89.04
Speed (km h^{-1})	DEUCE	187–220	35.17	45.75
	AD		34.36	44.13
	BOTH		30.38	39.91

4. Discussion

Using a novel methodological proposal, this study sought to identify which characteristics of serves determine their efficiency in men's professional doubles tennis. The results showed that the characteristics of the serves in which the point ended in two, three or more shots were very similar. However, it was possible to identify different characteristics for the aces, compared to those in which the point lasted more than three shots. Therefore, these results allow us to establish the variables and values which are associated with a higher probability of serving an ace in men's professional doubles tennis. In addition, the methodology used has been shown to be a novel, valid and reliable way that can be replicated for the calculation of the efficiency of actions in different sports disciplines.

The results have shown that the service angle is one of the variables that best determines the probability of serving an ace in men's professional doubles tennis. These results are consistent with a previous study of the serve in singles tennis, which indicated that if the serve angle was less than 5.88°, the serve will be returned with a probability of 92.52% [11]. In doubles tennis, according to the data obtained in the present study, an angle of between 5.7° and 8.7° theoretically increases the probability of serving an ace from 29.77% to 88.21% on the advantage side and from 33.70% to 88.68% on the deuce side. These results support those obtained by previous studies, where Whiteside and Reid [11] reported that there is a higher probability of getting an ace when the serve is played wide than to the T zone, with probabilities of 42% on the advantage side and 52% on the deuce side. Wide serves generate a greater angle than those played to the T, so it is logical, according to the data obtained, that they are also more effective. Furthermore, from a technical point of view, and specifically related to the different types of serve spin used, it has been shown that players use the slice serve to generate a greater serve angle on the deuce side, whereas they opt for flat or topspin serves for a greater angle when serving to the T on the advantage side [6,19,41].

Another of the variables found to have a greater relevance in determining the probability of serving an ace has been the distance from the bounce of the ball to the sideline of the service box. This reinforces the idea that one of the main objectives of the server is to gain an advantage over the receiver, trying to get the ball as far away from the receiver as possible [6,14,15]. Whiteside and Reid [11] determined that in 77.53% of cases, if the distance to the line was less than 15.27 cm and, in addition, the serve angle was greater than or equal to 5.88°, the serve would be an ace. In the case of men's professional doubles tennis, in this study, it has been shown that, if the serve distance is between 0 and 28 cm, the probability of a serve being an ace would increase from 30.64% to 81.26% on the advantage side and from 33.05% to 83.40% on the deuce side.

The results on the importance of the angle of the serve and the distance of the bounce of the ball to the line also support what has been found in previous studies on the professional doubles game. According to Vives et al. [14], players try to take the initiative of the point by moving the receiver out of the court, both on the advantage and deuce sides. This study reported a high efficiency of the serve when players serve wide on both sides of the court, with values of 40.71% on the deuce side and 37.54% on the advantage side. Similarly, the percentages of the effectiveness of serves to the T were considerably high, with values of 38.45% and 39.28%, respectively. The information provided in this study regarding the most important variables and the values that maximize the probabilities of direct serves could be very valuable for coaches and players in order to further increase the effectiveness of this type of service.

In singles, the serve speed has been shown to be one of the most important variables when serving an ace. Previous studies have reported that the probability of serving an ace if the serve speed is greater than 198 km h^{-1} is 58.68% [11]. However, the data obtained in our study show that the serve speed is less important in doubles tennis. Thus, it was found that once 187 km h^{-1} is exceeded, the probability of serving an ace would only increase by about 10%, regardless of the side of the serve.

Therefore, considering the results from this study and the values found for the most important variables that determine the execution of serving an ace, it can be stated that, although speed is an important variable, very high speeds do not considerably increase the probability of winning the point directly with the serve. However, the angle of the serve and the bounce of the ball close to the sidelines do significantly increase the probability of serving an ace.

The primary advantages of this method lie in its capacity for variable explainability. Feature importance scores are computed not only to discern the most pivotal variables, but also to propose optimal value ranges for these variables to maximize efficiency. It is crucial to note, however, that this process is not readily automated and necessitates human intervention for analysis. A future direction is proposed, aiming to establish mathematical foundations to formulate the selection of value ranges for selected variables that maximize the probability of the desired variable. Furthermore, noteworthy correlations between variables, indicative of collinearity, were identified in the study. Although these correlations did not adversely impact model prediction performance, they significantly influenced feature importance algorithms. Moving forward, addressing such collinearity could be pivotal for more refined analyses.

On the other hand, the exclusivity of the dataset to Davis Cup hard court matches introduces a limitation to the generalizability of the findings, precluding straightforward extrapolations for all tennis surfaces, levels and genders. While augmenting the dataset with diverse surfaces, levels and genders could enhance the dataset variance, it concurrently poses challenges in deriving specific and universally applicable conclusions. Furthermore, the study underscores the prominence of a player's playing style over physical attributes concerning the maximization of serve efficiency. For future endeavors, an exploration of player grouping based on similar characteristics, facilitated by clustering algorithms, could be pursued to train distinct models for each group.

These results have very relevant practical implications related to the training regime for serve practices in men's professional doubles. From a tactical perspective, the insights found in this study can assist tennis players in their decision-making to maximize their serve advantage. In addition, technically, physically, and mentally, it is also important to prepare doubles tennis players for efficient serve executions with respect to the values of the indicated variables. Therefore, the effective application of the values found for variables such as the serve angle, distance to the line, or speed will determine the design of specific training tasks and help define specific patterns of play [11,14].

5. Conclusions

In this study, key characteristics for a direct serve in men's professional doubles tennis were identified. The angle of the serve was found to be a crucial variable, while the distance from the ball bounce to the sideline was also relevant. These findings are consistent with previous research on doubles tennis. It was observed that an optimal serve angle between 5.7° and 8.7° significantly increased the probability of a direct serve. In addition, maintaining a distance from the ball bounce to the sideline between 0 and 28 cm increased the probability of success. The speed of the serve in doubles tennis showed less influence compared to singles tennis.

The novel methodology employed here utilizing tracking data and explainable AI techniques provides a valid and reliable way to model service efficiency in tennis. The major contribution of this work is the extraction of specific parameter values to optimize the doubles serve through data-driven analytics. From a practical perspective, the serve angle and placement recommendations offer directly implementable guidelines for players and coaches. Technically, physically and mentally targeting these service metrics should become integral in training programs and match tactics.

This research focused exclusively on elite men's hard court doubles matches, so an area for future investigation is expanding dataset diversity to include different surfaces, levels and genders. Furthermore, grouping players by style and developing individualized models could provide more tailored strategic recommendations. The advanced analysis framework presented can be extended to other facets of tennis performance or different sports relying on spatiotemporal tracking data. As machine learning approaches, wearable sensors and tracking tools continue to proliferate in sports, ample opportunities exist to generate actionable intelligence maximizing success.

Author Contributions: Conceptualization, F.V., J.F.G., M.C. and R.M.-G.; methodology, F.V., J.F.G., M.C. and R.M.-G.; formal analysis, J.L., F.V. and R.M.-G.; investigation, F.V., J.F.G., M.C. and R.M.-G.; data curation, J.L. and F.V.; writing—original draft preparation, F.V., J.F.G., M.C. and R.M.-G.; writing—review and editing, F.V., J.F.G., M.C. and R.M.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from the International Tennis Federation and are available from the authors with the permission of the International Tennis Federation.

Acknowledgments: The authors wish to acknowledge the contributions of the International Tennis Federation Participation and Education Department.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Martínez-Gallego, R.; Crespo, M.; Ramón-Llin, J.; Micó, S.; Guzmán, J.F. Men's doubles professional tennis on hard courts: Game structure and point ending characteristics. *J. Hum. Sport Exerc.* **2020**, *15*, 633–642. [CrossRef]
2. Kostoglou, G.T.; Tsitas, K.; Kostoglou, V.; Kostoglou, P. The importance of the serve in tennis: A systematic review. *Int. J. Perform. Anal. Sport* **2017**, *17*, 505–518.
3. Filipcic, A.; Zecic, M.; Reid, M.; Crespo, M.; Panjan, A.; Nejc, S. Differences in performance indicators of elite tennis players in the period 1991–2010. *J. Phys. Educ. Sport* **2015**, *15*, 671–677. [CrossRef]

4. O'Donoghue, G.P.; Brown, E. The importance of service in Grand Slam singles tennis. *Int. J. Perform. Anal. Sport* **2008**, *8*, 70–78. [CrossRef]
5. Reid, M.; Whiteside, D.; Elliott, B. Serving to different locations: Set-up, toss, and racket kinematics of the professional tennis serve. *Sports Biomech.* **2011**, *10*, 407–414. [CrossRef] [PubMed]
6. Mecheri, S.; Rioult, F.; Mantel, B.; Kauffmann, F.; Benguigui, N. The serve impact in tennis: First large-scale study of big Hawk-Eye data. *Statistical Anal. Data Min. ASA Data Sci. J.* **2016**, *9*, 310–325. [CrossRef]
7. Kocib, T.; Carboch, J.; Cabela, M.; Kresta, J. Tactics in tennis doubles: Analysis of the formations used by the serving and receiving teams. *Int. J. Phys. Educ. Fit. Sport* **2020**, *9*, 45–50. [CrossRef]
8. Martínez-Gallego, R.; Salvador, S.M.; Luján, J.F.G.; Reid, M.; Ramón-Llin, J.; Crespo, M. Challenging serve myths in doubles tennis. *Int. J. Sports Sci. Coach.* **2021**, *16*, 1305–1311. [CrossRef]
9. Carboch, J.; Koc˘íb, T. A comparison of service efficiency between players of male and female doubles at professional tennis tournaments. *Auc Kinanthropol.* **2016**, *51*, 56–62. [CrossRef]
10. Martínez-Gallego, R.; Vives, F.; Guzmán, J.F.; Ramón-Llin, J.; Crespo, M. Time structure in men's professional doubles tennis: Does team experience allow finishing the points faster? *Int. J. Perform. Anal. Sport* **2021**, *21*, 215–225. [CrossRef]
11. Whiteside, D.; Reid, M. Spatial characteristics of professional tennis serves with implications for serving aces: A machine learning approach. *J. Sports Sci.* **2017**, *35*, 648–654. [CrossRef] [PubMed]
12. Giampolo, F.; Levey, J. Serves. In *Championship Tennis*; Human Kinetics: Champaign, IL, USA, 2013; p. 59.
13. Wei, X.; Lucey, P.; Morgan, S.; Carr, P.; Reid, M.; Sridharan, S. Predicting serves in tennis using style priors. In Proceedings of the KDD '15: The 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, 10–13 August 2015; pp. 2207–2215. [CrossRef]
14. Vives, F.; Crespo, M.; Guzmán, J.F.; Martínez-Gallego, R. Effective serving strategies in men's doubles Davis cup matches: An analysis using tracking technology. *Int. J. Perform. Anal. Sport* **2022**, *22*, 638–648. [CrossRef]
15. Martínez-Gallego, R.; Crespo, M.; Jiménez, J. Analysis of the differences in serve effectiveness between Billie Jean King Cup (former Fed Cup) and Davis Cup doubles tennis matches. *Int. J. Sports Sci. Coach.* **2021**, *16*, 777–783. [CrossRef]
16. Mahesh, B. Machine learning algorithms—A review. *Int. J. Sci. Res.* **2020**, *9*, 381–386.
17. Mlakar, M.; Luštrek, M. Analyzing tennis game through sensor data with machine learning and multi-objective optimization. In Proceedings of the UbiComp '17: The 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Maui, HI, USA, 11–15 September 2017; pp. 153–156. [CrossRef]
18. Giles, B.; Kovalchik, S.; Reid, M. A machine learning approach for automatic detection and classification of changes of direction from player tracking data in professional tennis. *J. Sports Sci.* **2020**, *38*, 106–113. [CrossRef] [PubMed]
19. Giles, B.; Peeling, P.; Kovalchik, S.; Reid, M. Differentiating movement styles in professional tennis: A machine learning and hierarchical clustering approach. *Eur. J. Sport Sci.* **2021**, *23*, 44–53. [CrossRef] [PubMed]
20. Bayram, F.; Garbarino, D.; Barla, A. Predicting Tennis Match Outcomes with Network Analysis and Machine Learning. In *SOFSEM 2021: Theory and Practice of Computer Science*; Bureš, T., Ed.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2021; Volume 12607, pp. 525–536. [CrossRef]
21. Perri, T.; Reid, M.; Murphy, A.; Howle, K.; Duffield, R. Prototype Machine Learning Algorithms from Wearable Technology to Detect Tennis Stroke and Movement Actions. *Sensors* **2022**, *22*, 8868. [CrossRef]
22. Kovalchik, S.; Reid, M. A shot taxonomy in the era of tracking data in professional tennis. *J. Sports Sci.* **2018**, *36*, 2096–2104. [CrossRef]
23. Sipko, M.; Knottenbelt, W. *Machine Learning for the Prediction of Professional Tennis Matches*; MEng Computing-Final Year Project; Imperial College London: London, UK, 2015; Volume 2.
24. Kovalchik, S.A. Searching for the GOAT of tennis win prediction. *J. Quant. Anal. Sports* **2016**, *12*, 127–138. [CrossRef]
25. Yang, M.; Qu, Q.; Shen, Y.; Zhao, Z.; Chen, X.; Li, C. An effective hybrid learning model for real-time event summarization. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4419–4431. [CrossRef]
26. Cornman, A.; Spellman, G.; Wright, D. *Machine Learning for Professional Tennis Match Prediction and Betting*; Working Paper; Stanford University: Stanford, CA, USA, 2017.
27. Sekar, A. Predicting the Winner of a Tennis Match Using Machine Learning Techniques. Ph.D. Thesis, College of Ireland, Dublin, Ireland, 2019.
28. Candila, V.; Palazzo, L. Neural networks and betting strategies for tennis. *Risks* **2020**, *8*, 68. [CrossRef]
29. Van Rooij, C. Machine Learning in Tennis: Predicting the Outcome of a Tennis Match Based on Match Statistics and Player Characteristics. Master's Thesis, School of Humanities and Digital Sciences of Tilburg University, Tilburg, The Netherlands, 2021.
30. Wilkens, S. Sports prediction and betting models in the machine learning age: The case of tennis. *J. Sports Anal.* **2021**, *7*, 99–117. [CrossRef]
31. De Seranno, A. Predicting Tennis Matches Using Machine Learning. Master's Thesis, Ghent University, Ghent, Belgium, 2020.
32. Yue, J.C.; Chou, E.P.; Hsieh, M.H.; Hsiao, L.C. A study of forecasting tennis matches via the Glicko model. *PLoS ONE* **2022**, *17*, e0266838. [CrossRef] [PubMed]
33. Solanki, S.; Jakir, V.; Jatav, A.; Sharma, D. Prediction of tennis match using machine learning. *Int. J. Progress. Res. Eng. Manag. Sci.* **2022**, *2*, 5–7.

34. Gupta, M.; Gupta, R.; Ansari, S.; Tambe, D.N. Tennis Match Prediction using Machine Learning. *Int. J. Res. Publ. Rev.* **2022**, *3*, 693–696.
35. Hawk-Eye Innovations. *Hawk-Eye's Accuracy & Reliability: Electronic Line Calling*; Hawk-Eye Innovations: Basingstoke, UK, 2015.
36. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2980–2988.
37. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**. [CrossRef]
38. Altmann, A.; Tolos¸i, L.; Sander, O.; Lengauer, T. Permutation importance: A corrected feature importance measure. *Bioinfmatics* **2010**, *26*, 1340–1347. [CrossRef]
39. Lundberg, S.M.; Lee, S.I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 4765–4774.
40. Molnar, C. Model-Agnostic Methods. In *Interpretable Machine Learning*; Leanpub: Victoria, BC, Canada, 2020; pp. 143–221.
41. Matsuzaki, C. *Tennis Fundamentals*; Human Kinetics: Champaign, IL, USA, 2004; p. 91.

2024, Retos, 60, 955-962

Estrategia de servicio en el tenis profesional de dobles femenino - una perspectiva basada en datos
Serve strategy in professional women's doubles tennis - a data-driven perspective

*Fernando Vives, **Javier Lázaro, *José Francisco Guzmán, ***Miguel Crespo, *Rafael Martínez-Gallego
*Universidad de Valencia (España), ** Independent researcher, ***International Tennis Federation (UK)

Abstract. This study aimed to identify the key variables that influence the effectiveness of the first serve in women's doubles matches using a deep learning training model. A total of 4939 first serves by 145 players from Billie Jean King Cup matches were classified according to their effectiveness. Data from the Hawk-Eye system were filtered and processed to obtain descriptive statistics of the spatio-temporal predictor variables and then to calculate the most important variables for first serve effectiveness. The angle of the serve as the feature that gained more relevance, along with the speed and distance from the bounce to the sidelines, regardless of the side from which the serve was made. The training model showed novel results, identifying that the values of the serve angle between 5.5º-8.7º increased the effectiveness from 22.2% to 48.5% on the Deuce side and from 17.3% to 45.7% on the Advantage side. These findings align with existing research on men's professional tennis singles and offer valuable insights into strategic and coaching approaches. The insights gathered from this study are poised to enhance training methodologies and match-day strategies, providing a strategic edge in serve selection during competitive play.
Keywords: racquet sports, machine learning, strategy, sport analytics, serve

Resumen. Este estudio tuvo como objetivo identificar las variables más determinantes que influyen en la efectividad del primer servicio en los partidos de dobles femenino utilizando un modelo de entrenamiento de aprendizaje profundo. Se clasificaron 4939 primeros servicios de 145 jugadoras según su efectividad, pertenecientes a partidos de la Copa *Billie Jean King*. Los datos del sistema de Ojo de Halcón fueron filtrados y procesados para obtener estadísticas descriptivas de las variables predictoras espacio-temporales y, posteriormente, se calcularon las variables más importantes para la efectividad del primer servicio. El ángulo del servicio fue la característica que obtuvo una mayor relevancia, junto con la velocidad y la distancia desde el bote de la pelota hasta las líneas laterales, independientemente del lado donde se ejecutó el servicio. El modelo de entrenamiento mostró que los valores del ángulo de servicio entre 5.5º-8.7º incrementaban la efectividad, pasando del 22.2% al 48.5% en el lado del Iguales, y del 17.3% al 45.7% en el lado de la Ventaja. Estos resultados no sólo se alinean con la investigación existente sobre el tenis profesional masculino tanto en la disciplina individual como la de dobles, sino que también ofrecen una valiosa orientación táctica y de entrenamiento. La información obtenida en este estudio está destinada a mejorar las metodologías de entrenamiento y las estrategias en el propio partido, proporcionando una ventaja estratégica en la selección del servicio durante la competición.
Palabras clave: deportes de raqueta, aprendizaje automático, estrategia, análisis deportivo, servicio

Fecha recepción: 23-06-24. Fecha de aceptación: 30-08-24
Fernando Vives
fervial@alumni.uv.es

Introduction

Tennis is an open skill sport in which, from a strategy perspective, players make tactical decisions based on their strengths and weaknesses, as well as those of their opponents. Additionally, other contextual variables, such as playing surface, type of balls, temperature, humidity or altitude, also significantly influence player performance (Fitzpatrick et al., 2019; Vaverka et al., 2018). In the quest to understand the critical factors that influence success in tennis, numerous studies have analysed players' match statistics (Crespo et al., 2015; Durán et al., 2022), as well as key performance indicators in junior and professional tennis players in different game situations (Jindo et al., 2022; Filipcic et al., 2021) and different types of playing surfaces (Fitzpatrick et al., 2019; Vaverka et al., 2018).

Similarly, an obvious lack of research can be observed in women's doubles, despite its relevance in team competitions such as the Billie Jean King Cup (BJKC). The BJKC, formerly known as the Fed Cup until 2020, is the most important global competition played by 110 national teams annually, which is the tennis World Cup for women. Created in 1963 by the ITF, it is the equivalent for women of the Davis Cup (BJKC, 2023). As aforementioned, research on this competition is substantially less common than in the men's events.

The serve is one of the key performance aspects in tennis, as it is a fundamental skill that plays a crucial role in the match outcome (Filipcic *et al. 2015; Grambow et al. 2022)*. In fact, points won with the second serve and points won with the return of the second serve have been identified by authors such as Reid, McMurtrie and Crespo (2010) as the most relevant variables of the game. In addition, it has been shown that players tend to improve their serving statistics as a tournament progress (Grambow et al., 2020). Also, players with a more effective serve achieve a higher percentage of points won on the second serve, show better performance in break point situations, and win more points and games when serving (Infantes et al., 2019). These findings highlight the importance of the serve in overall player performance.

Despite being the most decisive stroke in tennis today, the scarcity of studies dedicated to the serve in the women's game is evident (Sánchez-Pay et al., 2022). It has been demonstrated that both serving and returning are crucial elements in competition (Stănescu, 2016; Abidin & Ruslan, 2020). Furthermore, these elements significantly influence successive shots (Born, 2021). Authors such as Grambow et al. (2021), it is recommended that players prioritize the increase in the percentage of games and points won on serve, as well as perfecting aces to improve serving effectiveness. Other authors have explored biomechanical aspects (Whiteside et al., 2015) the kinematics of the first serve at different stages of training (Whiteside et al., 2013) or the relationship between serve speed and anthropometric parameters, ball impact and landing location (Baiget et al., 2023). Furthermore, studies have also analyzed how players' ranking can influence their performance (Sánchez-Alcaraz, 2018; Kovalchik et al., 2017)).

As per the serve in the doubles game, it has received considerably less attention from researchers than that of the singles. When comparing the serve efficiency between male and female professional players Carboch & Kočíb (2016) concluded that men won around 10% serve points more than women, and that to win a match they also needed to win more serve points than women. Kocib et al. (2020) tactically analysed the formations of serving and returning teams, and Avkhimovich (2023) investigated the order of servers in doubles and how this affects who should serve first. Martinez-Gallego et al. (2021c) demystified popular serving beliefs in doubles, concluding that the serving team's advantage is most notable on points of up to four shots, especially with new balls on clay, although this advantage diminishes on break points. Lastly, a study of Davis Cup doubles matches revealed that the effectiveness of the serve in doubles depends more on the angle and placement than on the speed of the serve (Vives et al., 2023).

In this context, other studies have focused on gender differences in the volley game in doubles, highlighting certain court coverage differences between men and women (Martínez-Gallego et al, 2021b), or the activity profile of players on different surfaces in doubles Grand Slam matches, resulting in short and intense points with a predominance of net play (Borderías et al., 2021). Finally, Borderías et al. (2023) analysed structural variables and the completion of points in women's doubles, concluding that taking the initiative and being aggressive in the game is beneficial for the point outcome.

From a technological perspective in terms of the tools used for match analysis, the use of tracking technologies in tennis, such as the Hawk-Eye (HE), has allowed different studies to use a large amount of data for the analysis of the singles game performance indicators. The serve has been analyzed in a myriad of studies (Loffing, 2009; Rioult et al., 2015; Wei et al., 2015; Mecheri, 2016; Whiteside and Reid (2017). For instance, Fitzpatrick et al. (2023) explored the different serve and return strategies used by male and female players at Wimbledon. Other studies have also investigated the different movement parameters, changes of direction and hitting dynamics of male and female players during matches (Wei et al., 2013; Reid and Duffield, 2014; Reid et al., 2016; Giles et al., 2020; Brandon et al., 2021; Wei et al., 2013; Meurs et al., 2021).

In doubles, HE data has been used to analyse how serve direction varies with server position. Vives et al. (2021) found that server position was a key factor in the direction of the first and second serves, affecting both serve efficiency and return anticipation. The differences in the characteristics of the serve in men's and women's doubles tennis have also been analysed in terms of its effectiveness and the directions used, and Martínez-Gallego et al. (2021a) concluded that men's teams are more effective than women's, especially in serves directed to areas far from the center of the serve box.

Despite the recognized importance of the serve in tennis and the numerous studies that have examined its various aspects in singles and the men's game, there is a significant lack of research specifically focused on women's doubles, particularly in high-level competitions such as the Billie Jean King Cup (BJKC). This gap in the scientific literature highlights the necessity for research that utilizes current technologies and analytical methods to explore serve dynamics and strategies in this discipline. Accordingly, this study aims to address this gap by identifying and analyzing the key variables that influence the effectiveness of the first serve in BJKC women's doubles matches.

Methods

Sample

A total of 4939 first serves was analysed from 65 women's doubles matches played during BJKC qualifying ties between 2011 and 2019. Our study featured the participation of 107 teams and 145 players from 23 different countries, with an average age of 25.81 ± 4.74 years.

Process

The data used in this study was obtained using the HE tracking system (Hawk-Eye Innovations Ltd., 2015). The HE tracking technology has passed rigorous ITF tests that have validated the accuracy and reliability of the system. A system of synchronized cameras is used to track the movement of the ball in real time. Each camera captures the position of the ball in its own plane. By knowing the position and orientation of each camera, the 3D position of the ball can be estimated by intersecting the trajectories defined by the cameras. This allows the three-dimensional coordinates of the ball to be determined accurately in space (Singh Bal & Dureja, 2012; Reid et al., 2016).

After acquiring data from the HE system, a template was developed using Microsoft Excel version 16.16.7 (Microsoft, Redmond, USA) to facilitate data processing and the calculation of targeted variables.

Target variables

The variables derived and analyzed in this study were:

- Court side: Side of the court from where the serve is played. This variable was used to divide the dataset and perform the analysis for each side of the court.
- Effectiveness:
  - Type 1: The point finishes with one shot.
  - Type 2: The point finishes with two shots.

2024, Retos, 60, 955-962

- Type 3: The point finishes with three shots.
- Type 4: The point finishes with four or more shots.

Input features

- Speed: Mean speed of the serve.
- Position: Position of the server when hitting the serve.
- Time: Time between ball impact and ball bounce.
- Speed loss: Loss of speed of the ball after its bounce.
- Impact Z: Height of the ball at impact.
- Net clearance: Height of the ball when passing over the net.
- d: Serve angle.
- b: Vertical projection angle.
- dL: Distance from ball bounce to the sideline of the serve box.
- Direction: Wide (W), Body (B) and T (Figure 1).

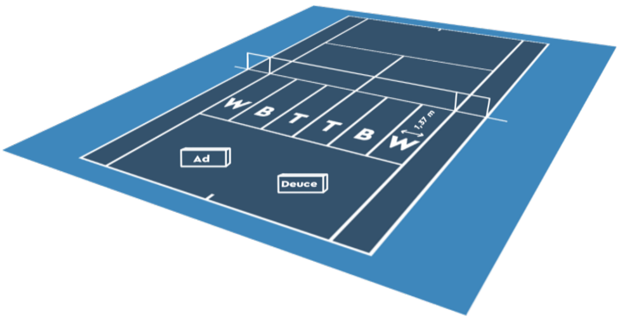

Figure 1. Serve direction

Data analysis

The Machine Learning methodology used in this study was used and validated in a previous study by Vives et al. (2023), using the following steps:

Predictive model: Training the deep learning model

This study compared AutoGluon and FastAI for tabular data analysis. Although AutoGluon offers outstanding tools, its lack of flexibility leading to the choice of FastAI. The implementation of AutoGluon was straightforward, while FastAI required more development, but provided flexibility.

A tabular model with 5 layers of feature size [256, 128, 128, 128, 128, 64] and LinBnDrop basic blocks were used. The loss function was Focal Loss Flat, and the Adam optimization algorithm (Kingma & Ba, 2015) was used. This work contributes to the understanding of tools for tabular data analysis.

Feature importance: Most important variable

The Feature Permutation Importance methodology emerges as a robust tool, facilitating the identification of features with predictive capability, irrespective of the employed model. This approach provides a valuable perspective for assessing the individual contribution of variables within the broader analytical framework (Kuhn & Johnson, 2013).

Feature importance algorithms were employed to identify variables with a substantial impact a predictive model, indicating the relative significance of each feature in making predictions. It is based on repeated permutations of the outcome vector to estimate the distribution of measured importance for each variable in a non-informative setting (Rajbahadur et al., 2022).

Probabilistic and statistical analysis: Values that maximize the probabilities

In the context of a study focused on maximizing serve effectiveness and understanding the impact of selected features on predicted target variables, a novel semi-automated algorithm based on classical statistics was developed.

The initial steps involved calculating a smooth probability distribution using the Kernel Density Estimation algorithm for each selected feature, along with statistical estimators such as Mean, Maximum Likelihood Estimation (MLE), and percentiles 5 and 95 for each type of effectiveness. Subsequently, values were identified by selecting areas on the plot with a higher density of points corresponding to the desired result (type 1) and a lower density of points for the undesired result (type 4). Statistical estimators played a crucial role in interpreting the probability distribution during this process.

Synthetic data-set generation: simulation based on the predictions on this dataset and analysis of effectiveness improvements

The final step was to create a synthetic dataset by applying the identified thresholds to the original dataset Through validation procedures, we made adjustments to improve the effectiveness rate. To further refine the dataset, any values outside the specified thresholds were replaced with random values within the acceptable range, creating a new synthetic dataset. We then used the previously trained deep learning model to make predictions on the synthetic dataset, calculating the final rates for effectiveness type 1 and effectiveness type 4.

The Effectiveness results using the original dataset and the new synthetic dataset with the recommended threshold values are shown in table 2. They show how the model trained on the original dataset impacts the effectiveness on the second dataset, increasing the overall probability on the synthetic dataset by 47.1%

Results

Descriptive statistics of the spatio-teEporal predictor variables

Table 1 shows the descriptive statistics of the spatio-

2024, Retos, 60, 955-962

temporal predictor variables. As it can be seen in the results, the mean of the variable speed is higher in aces than in the rest of the serves, both on the Deuce side with 160.19 km/h and on the Advantage side with 165.90 km/h. This difference is also reflected in the d, being the angle greater in aces with an average of 5.99º on the Deuce side and 5.74º on the Advantage side. On the other hand, the dL was another of the variables that showed significant differences between the aces and the rest of the serves, being the shortest distance in aces regardless of the side. Regarding the rest of the variables, as can be seen, their means were quite similar on the Deuce and Advantage sides, both for the serves that were aces and for the rest of the first serves.

Table 1.
Descriptive statistics of the spatio-temporal predictor variables

	DEUCE				ADVANTAGE			
	Total (N=2615)		Aces (N=156)		Total (N=2324)		Aces (N=113)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Speed (km/h)	154.14	13.63	160.19	15.69	155.88	13.72	165.90	13.64
Position (m)	2.03	0.73	1.96	0.79	2.03	0.82	1.82	0.89
Impact Z (m)	2.67	0.14	2.70	0.13	2.68	0.15	2.72	0.14
Net Clearance (m)	1.14	0.11	1.13	0.11	1.14	0.11	1.13	0.11
Time (s)	0.48	0.05	0.47	0.05	0.48	0.05	0.45	0.04
Speed loss (km/h)	7.80	1.20	7.76	3.28	7.85	1.14	7.50	2.83
Serve angle (deg.)	3.59	1.91	5.99	1.64	3.63	1.74	5.74	1.56
Vertical projection angle (deg.)	9.05	0.61	9.07	0.56	9.02	0.64	9.05	0.55
dL (m)	1.01	0.54	0.39	0.37	0.99	0.56	0.40	0.47

Calculation of the Eost relevant variables

The next step was the calculation of the serve variables that were most important for type 1 effectiveness. Figure 2 shows the mean values indicating the average impact on the model outputs for each variable on the Deuce side, the Advantage side and both sides respectively. As it can be seen, the most important variables were d, speed and dL to the line irrespective of the side where the serve was played.

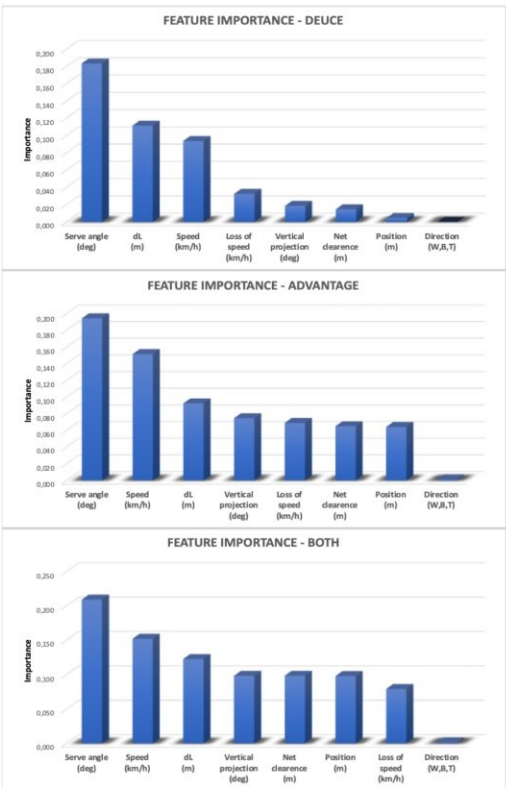

Figure 2. Importance of the variables for the aces in the Deuce, Advantage and Both sides

Results of the predictions using the trained Eodel

Table 2 shows the results of the predictions using the trained model and the synthetic dataset, which resulted from replacing values outside the recommended range with random values within the recommended range. Serve angle values between 5º and 8.7º double the initial probability of making an ace on both the deuce and the advantage side. Similarly, although to a lesser extent, both the variable of serve speed (between 157 – 200 km/h) and the variable of distance to the line < 0.42m, also improve the probability of making an ace, regardless of the side of the court.

Table 2.
Results of the predictions using the trained model

Variable	Side	Values (min-max)	Effectiveness Before (%)	Effectiveness After (%)
	DEUCE		22.2	48.5
Serve angle (deg)	AD	5.5 – 8.7	17.3	45.7
	BOTH		18.7	47.1
	DEUCE		18.8	24.3
Speed (km/h)	AD	157 – 200	16.7	23.9
	BOTH		16.8	22.8
	DEUCE		20.3	21.7
dL (m)	AD	0 – 0.50	16.4	20.3
	BOTH		17.5	20.7

Discussion

This study addressed serve effectiveness in women's doubles tennis, using advanced machine learning techniques to analyze a large dataset collected from BJKC matches. From the analysis of 4939 serves in 65 matches, it was found that serve angle, serve speed and distance to the serve line are critical variables in increasing the likelihood of hitting an ace.

The results shown in Table 2, have determined that serve angles between 5.5º-8.7º have a direct impact on the serving team achievement of an ace in women's professional doubles tennis, increasing the overall probability by 47.1 % (48.5% on the deuce side and 45.7% on the advantage side). These results are in the lines of previous studies in both singles and men's doubles tennis, where it was reported that angles from 5.88º and 5.7º, respectively, significantly increased the probability of hitting an ace (Whiteside & Reid, 2017; Vives et al., 2022). These results confirm that, in women's doubles tennis, as in the men's game, employing strategies in which the direction of the serve seeks to place the ball away from the receiver increases the efficiency of the serve partner and, therefore, their chances of winning the point, either through an ace or by taking the initiative of the point (Rioult et al., 2015; Mecheri et al., 2016; Martínez-Gallego et al., 2021a; Vives et al., 2022). However, it is important to note that these improvements in efficiency in women's doubles (≈30 %) are considerably lower than those reported in the men's game (≈50-60 %). This may be due, on the one hand, to the lower overall efficiency of the serve in women's doubles compared to men's doubles (Martínez-Gallego et al., 2021a) and, on the other hand, to the position adopted by the players in the return of serve. In this regard, although there are no data in doubles tennis, in the singles game it has been reported that female players adopt positions closer to the net than male players (Reid et al., 2016). This could allow them to shorten the trajectory of the ball and, therefore, move a shorter distance, thus increasing the likelihood of returning the serve.

The variable of serve speed has been explored due to its impact on the game in several studies, which have found a significant relationship with the probability of winning the point (Rioult et al., 2015). In contrast, no direct relationship has been found between aces and speed in both male and female players (Brown, 2021). However, the results obtained in this study show that, in women's doubles tennis, first serve speed, although limited, does improve the likelihood of hitting an ace. Specifically, it has been found that serves with speeds of 157 km/h and above improve the probability of hitting an ace on both sides of the court by around 5%. These results are in line with those found in men's doubles, where once 187 km/h was exceeded, the probability of an ace was increased by approximately 10% (Vives et al., 2023). Thus, although serve speed is an important variable that has received much attention and interest (Rioult et al., 2015; Grambow et al., 2021; Mecheri et al., 2016), its contribution to improving effectiveness is limited compared to other variables such as angle. Currently, probably due to improved return performance, returners are able to return a very high percentage of serves, even at high speeds (Born et al., 2021; Fitzpatrick et al. 2023), which means that, in isolation, speed is not a determining variable in serve effectiveness. Therefore, it seems reasonable that, although serve speed is an aspect to work on with the players, it would be advisable to integrate it with variables related to direction and angle to maximize effectiveness.

Similar to speed, the distance from the ball bounce to the serve box sideline, although it improves the effectiveness of serves, does so to a limited extent, increasing the probability of hitting an ace by approximately 5% when the distance is less than 0.50 m. These results differ from those observed in men's doubles tennis, where bouncing the ball within 0.28 m of the line was reported to increase the probability of hitting an ace by 50% (Vives et al., 2023). This also contrasts with a study of men's singles tennis, which found that bouncing within 0.15 m of the line significantly increased the probability of hitting an ace (Whiteside and Reid, 2017). These differences are likely due to the interaction of this variable with speed. The fact that male players serve at a higher speed than female players (Brown, 2021) means that, when the ball bounces close to the lines, effectiveness increases to a greater extent. On the other hand, in the case of the female players, although the ball bounces close to the lines, serving at a slower speed allows the receiver to reach the ball in better conditions and, therefore, the serve is less effective.

Even though this study is comprehensive, there are certain limitations that should be taken into account. These limitations also provide some ideas for future research directions in this particular area of study.

One of the key limitations of this study is the isolated analysis of the serve, conducted without considering its impact on subsequent shots or shot patterns that develop later in the game. This approach omits how the characteristics of the serve can affect the dynamics of the point. Future research could analyze the effectiveness of the serve using a more comprehensive approach, by examining how the variables of the serve contribute to the overall development of the game, thus allowing for a more holistic understanding of the various nuances in the strategy of women's doubles tennis.

Another important limitation is the independent analysis of the variables conducted, without exploring the interaction between them. This simplified approach may not fully capture the complexity and multifactorial nature of tennis performance. In this scenario, future research should consider the interactions between variables, which would offer deeper insights into how specific combinations of factors influence serve success in women's doubles.

Finally, the exclusive use of data accessed from the BJKC may limit the absolute applicability of the results to other competitive contexts in women's doubles tennis. The specificity of the sample suggests the need for caution in absolute applying the results. Future research could also extend the analysis to different competitions and levels of play, thus enriching the understanding of serving in women's doubles tennis and its impact on competitive performance. These studies could reveal tactical and strategic variations adapted to different playing contexts, offering a broader and globally applicable perspective.

From a practical application perspective, the results of this study reveal critical aspects that can be considered by coaches and trainers to improve serve effectiveness in women's doubles on court training sessions. The identification of speed, bounce distance to the line, and particularly, serve angle as key variables that increase the likelihood of hitting aces suggests that serve training should prioritize these elements. This aligns with the findings of Filipcic et al. (2021), who highlighted the importance of shot angles in their analysis of matchplay characteristics and performance indicators. Therefore, specific technical training that incorporates exercises aimed at improving mechanical variables, such as serve angle and speed, could be particularly beneficial. Furthermore, when planning doubles game strategies, and considering the observed strengths and weaknesses in the opponents, it would be relevant to understand how the variables identified in this study can help not only to improve serving performance, but also in shaping the post-serve game patterns, allowing players to take the initiative from the start of the point.

Conclusions

This study explored the effectiveness of the serve in women's doubles tennis, using advanced machine learning techniques to analyse data collected from Billie Jean King Cup matches. The results revealed that serve angle is the variable that determines the most the likelihood of hitting an ace. Speed and distance to the serve sidelines, although to a lesser extent, were also related to a greater likelihood of hitting an ace. These findings underline the importance of integrating technical, tactical, and biomechanical aspects to improve women's doubles serving performance through the appropriate use of variables. Therefore, it is suggested that coaches and physical trainers focus their serve training sessions on improving the angle and speed of the shot, as well as on improving the strategic selection of serve direction, to enhance the chances of success in women's doubles tennis.

Acknowledgements

The authors wish to acknowledge the contributions of the Education Department, International Tennis Federation.

Disclosure statement

The authors have no disclosures or conflicts of interest

References

Abidin, A. & Ruslan, N. (2020). Exploring the importance of players' characteristics and performance on serve and return of serve in winning the women's singles Grand Slam tennis tournaments. *Journal of Physics. Conference Series*, *1496*(1), 012008.

Avkhimovich, N. (2023). Should the strongest serve first? Strategic ordering of servers in doubles tennis. *Applied Economics*, 1-10.

Baiget, E., Corbi, F., & López, J. L. (2023). Influence of anthropometric, ball impact and landing location parameters on serve velocity in elite tennis competition. *Biology of Sport, 40(1*), 273.

Bal, B. S. & Dureja, G. (2012). Hawk Eye: A Logical Innovative Technology Use in Sports for Effective Decision Making. *Sport Science Review*, *21*.

BJKC (2023). Billie Jean King Cup History. Available at: https://www.billiejeankingcup.com/en/about/history. *Accessed 8 March.*

Borderias, M., Crespo, M., Martínez-Gallego, R. & Baiget, E. (2023). Comparison of the game structure and point ending during Grand Slam women's doubles tennis. *International Journal of Sports Science & Coaching*, *18(4),* 1248-1255. https://doi.org/10.1177/17479541221098226

Borderias, M., Fernández-Fernández, J., Martínez-Gallego, R. & Baiget, E. (2021). Comparison of the activity profile during Grand Slam women's doubles tennis. *International Journal of Performance Analysis in Sport, 21, 1066 – 1080.*

Born, P., Malejka, L., Behrens, M., Grambow, R., Meffert, D., Breuer, J., & Vogt, T. (2021). Stroke placement in women's professional tennis: What's after the serve?. *Int. j. racket sports sci. vol. 3(1),* 2021, 37-44.

Brown, E. G. (2021) A faster serve has more impact on success for female elite tennis players than males, *International Journal of Performance Analysis in Sport*, *21(4)*, 600-610.

Carboch, J., & Kočíb, T. (2016). A comparison of service efficiency between players of male and female doubles at professional tennis tournaments. *Auc Kinanthropologica*, *51(2),* 56-62.

Crespo, M., Filipcic, A., Panjan, A., Reid, M., Šarabon, N., & Zečić, M. (2015). Differences in performance indicators of elite tennis players in the period 1991- 2010. *Journal of physical education and sport*, *15*, 671-677.

Durán, A., Martínez-Gallego, R., & Gimeno, M. (2022). Diferencias de género en la aproximación al cambio de dirección en tenistas profesionales (Gender differences in the approach to change of direction in professional tennis players). *Retos*, *43*, 938–943. https://doi.org/10.47197/retos.v43i0.88803

Filipcic, A., Leskosek, B., Crespo, M., & Filipcic, T. (2021). Matchplay characteristics and performance indicators of male junior and entry professional tennis players. *International Journal of Sports Science & Coaching*, *16(3)*, 768-776.

Filipcic, A., Zecic, M., Reid, M., Crespo, M., Panjan, A., & Nejc, S. (2015). Differences in performance indicators of elite tennis players in the period 1991-2010. *Journal of Physical Education and Sport*, *15(4)*, 671.

Fitzpatrick, A., Stone, J. A., Choppin, S., & Kelley, J. (2024). Analysing Hawk-Eye ball-tracking data to explore successful serving and returning strategies at Wimbledon. *International Journal of Performance Analysis in Sport*, 24(3), 251-268.

Fitzpatrick, A., Stone, J. A., Choppin, S., & Kelley, J. (2019). Important performance characteristics in elite clay and grass court tennis match-play. *International Journal of Performance Analysis in Sport*, *19(6)*, 942–952.

https://doi.org/10.1080/24748668.2019.1685804

Giles, B., Kovalchik, S. & Reid, M. (2020) A machine learning approach for automatic detection and classification of changes of direction from player tracking data in professional tennis. *Journal of Sports Sciences, 38(1)*, 106-113, doi: 10.1080/02640414.2019.1684132

Grambow, R., Born, P., O'Shannessy, C., Breuer, J., Meffert, D., & Vogt, T. (2022). Serve efficiency development in women's vs. men's professional tennis. *Human movement, 23(2)*, 128-137.

Grambow, R., Oshannessy, C., Born, P., Meffert, D. & Vogt, T. (2021). Serve efficiency development indicates an extended women's tennis world class cohort: Analysing 14 years of Ladies Wimbledon Championships – implications for coaching. *Human Movement.*

Hawk-eye innovations. (2015). *Hawk-Eye's accuracy & reliability: Electronic line calling.* Basingstoke.

Infantes, P., Hidalgo, C. & Sánchez-Álcaraz Martínez, B. J. (2020). Perfiles de juego en función de las estadísiticas del saque del top 25 masculino de tenis. Education, Sport, Health and Physical Activity (ESHPA*). International Journal*, *4(3),* 313-323.

Jindo, T., Mitsuhashi, D., & Kubota, T. (2022). Accuracy of subjective stats of key performance indicators in tennis. *International Journal of Racket Sports Science, 4(2*), 32-39.

Kingma, D. P., & Ba, J. (2015). Adam: a method for stochastic optimization. Conference Paper at ICLR 2015. *arxiv preprint arxiv:1412.6980*.

Kocib, T., Carboch, J., Cabela, M., & Kresta, J. (2020). Tactics in tennis doubles: analysis of the formations used by the serving and receiving teams. *Int J Phys Educ Fit Sport*, *9*(2), 45-50.

Kovalchik, S. A., Bane, M. K. & Reid, M. (2017). Getting to the top: an analysis of 25 years of career rankings trajectories for professional women's tennis. *Journal of sports sciences*, *35*(19), 1904-1910.

Krishnan Rajbahadur, G., Wang, S., Kamei, Y., & Hassan, A. E. (2022). The impact of feature importance methods on the interpretation of defect classifiers. *arxiv e-prints*, arXiv-2202.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26, p. 13). New York: Springer.

Martínez-Gallego, R., Crespo, M. & Jiménez, J. (2021a). Analysis of the differences in serve effectiveness between Billie Jean King Cup (former Fed Cup) and Davis Cup doubles tennis matches. *International Journal of Sports Science & Coaching, 16*(3), 777-783.

Martínez-Gallego, R., Ramón-Llín, J. & Crespo, M. (2021b). A Cluster Analysis Approach to Profile Men and Women's Volley Positions in Professional Tennis Matches (Doubles*). Sustainability, 13, 6370.*

Martinez-Gallego, R., Salvador, S. M., Luján, J. F. G., Reid, M., Ramón-Llin, J., & Crespo, M. (2021c). Challenging serve myths in doubles tennis. *International Journal of Sports Science & Coaching*, *16(6),* 1305-1311.

Mecheri, S., Rioult, F., Mantel, B., Kauffmann, F. & Benguigui, N. (2016). The serve impact in tennis: First large-scale study of big Hawk-Eye data. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *9(5*), 310-325. https://doi.org/10.1002/sam.11316

O'Donoghue, P. (2002). Performance models of ladies' and men's singles tennis at the Australian Open. *International Journal of Performance Analysis in Sport, 2(1),* 73-84.

Reid, M., & Duffield, R. (2014). The development of fatigue during match-play tennis. *British journal of sports medicine*, *48*(Suppl 1), i7.

Reid, M., McMurtrie, D. & Crespo, M. (2010). The relationship between match statistics and top 100 ranking in professional men's tennis. *International Journal of Performance Analysis in Sport, 10(2)*, 131-138.

Reid, M., Morgan, S. & Whiteside, D. (2016). Match play characteristics of Grand Slam tennis: implications for training and conditioning. *Journal of sports sciences*, *34(19*), 1791-1798.

Rioult, F., Mecheri, S., Mantel, B., Kauffmann, F., & Benguigui, N. (2015, September). What can hawk-eye data reveal about serve performance in tennis?. In *MLSA15-Machine Learning and Data Mining for Sports Analytics, workshop (ECML/PKDD 2015)* (Vol. 1970, pp. 36-45).

Sánchez-Pay, A., Ortega-Soto, J.A., Courel-Ibáñez, J. & Sánchez-Alcaraz, B.J. (2022). El servicio en tenis profesional femenino: influencia del ranking y la lateralidad. *Revista internacional de medicina y ciencias de la actividad física y el deporte*, *22(86)*, 255-267.

Sánchez-Alcaraz Martínez, B. J., Perona Arce, F., & Courel-Ibáñez, J. (2018). Parámetros de Rendimiento en el Servicio y Resto del Top-100 Femenino en Tenis. *Revista Kronos*, *17(1).*

Stănescu, R. (2016). Comparative Study on the Efficiency of Women Tennis Players' First Serve at Australian Open 2014 and 2015. In V. Grigore, M. Stanescu, & M. Paunescu (Eds.), *Physical Education, Sport and Kinetotherapy - ICPESK 2015, (11)*. European Pcroceedings of Social and Behavioural Sciences (pp. 202-209). Future Academy. https://doi.org/10.15405/epsbs.2016.06.28

van Meurs, E., Buszard, T., Kovalchik, S., Farrow, D., & Reid, M. (2021). Interpersonal coordination in tennis: assessing the positional advantage index with Australian Open Hawkeye data. *International Journal of Performance Analysis in Sport*, *21(1),* 22-32.

Vaverka, F., Nykodym, J., Hendl, J., Zhanel, J., & Zahradnik, D. (2018). Association between serve speed and court surface in tennis. *International Journal of Performance Analysis in Sport*, *18(2),* 262–272. https://doi.org/10.1080/24748668.2018.1467995

Vives, F., Crespo, M., Guzmán, J. F., & Martínez-Gallego, R. (2022). Effective serving strategies in men's doubles Davis cup matches: an analysis using tracking technology. *International Journal of Performance Analysis in Sport*, *22*(4), 638-648.

Vives, F., Lázaro, J., Guzmán, J. F., Martínez-Gallego, R., & Crespo, M. (2023). Optimizing Sporting Actions Effectiveness: A Machine Learning Approach to Uncover Key Variables in the Men's Professional Doubles Tennis Serve. *Applied Sciences*, *13(24),* 13213.

Wei, X., Lucey, P., S. Morgan, S. & Sridharan, S. "Predicting Shot Locations in Tennis Using Spatiotemporal Data," *2013 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Hobart, TAS, Australia, 2013, pp. 1-8, doi: 10.1109/DICTA.2013.6691516.

Wei, X., Lucey, P., Stuart Morgan, Carr, P., Reid, M. & Sridharan, S. (2015). Predicting Serves in Tennis using Style Priors. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15). Association for Computing Machinery, New York, NY, USA, 2207–2215. https://doi.org/10.1145/2783258.2788598

Whiteside, D., & Reid, M. (2017). Spatial characteristics of professional tennis serves with implications for serving aces: A machine learning approach. *Journal of sports sciences, 35(7),* 648-654.

Whiteside, D., Elliott B.C., Lay B.S. & Reid, M. (2013). The effect of age on discrete kinematics of the elite female tennis serve. *J Appl Biomech*;*29(5)*:573-82

Whiteside, D., Elliott, B.C., Lay, B.S., & Reid, M. (2015). Coordination and variability in the elite female tennis serve. *Journal of Sports Sciences, 33, 675 - 686.*

Datos de los/as autores/as:

Fernando Vives	fervial@alumni.uv.es	Autor/a
Javier Lázaro	javilazarogarcia@gmail.com	Autor/a
José Francisco Guzmán	jose.f.guzman@uv.es	Autor/a
Miguel Crespo	miguel.crespo@itftennis.com	Autor/a
Rafael Martínez-Gallego	Rafael.Martinez-Gallego@uv.es	Autor/a