UNIVERSIDAD POLITÉCNICA DE MADRID

Facultad de Ciencias de la Actividad Física y del Deporte (INEF)

Exploring match performance of elite tennis players: the multifactorial game-related effects in Grand Slams

Tesis Doctoral

D. Yixiong Cui

崔一雄

Máster Universitario en Ciencias de la Actividad Física y del Deporte

Madrid, 2018

UNIVERSIDAD POLITÉCNICA DE MADRID

Facultad de Ciencias de la Actividad Física y del Deporte (INEF)

**EXPLORING MATCH PERFORMANCE OF ELITE
TENNIS PLAYERS: THE MULTIFACTORIAL GAME-
RELATED EFFECTS IN GRAND SLAMS**

TESIS DOCTORAL

D. Yixiong Cui

Madrid, 2018

DEPARTAMENTO DE CIENCIAS DE LA ACTIVIDAD FÍSICA,
DEL DEPORTE Y DEL OCIO

Facultad de Ciencias de la Actividad Física y del Deporte (INEF)

**EXPLORING MATCH PERFORMANCE OF
ELITE TENNIS PLAYERS: THE
MULTIFACTORIAL GAME-RELATED EFFECTS
IN GRAND SLAMS**

AUTOR:

D. Yixiong Cui

Máster Universitario en Ciencias de la Actividad Física y del Deporte
Universidad Politécnica de Madrid

DIRECTOR:

Dr. Miguel Ángel Gómez Ruano, Doctor en Ciencias de la Actividad
Física y del Deporte. Universidad Politécnica de Madrid

Madrid, 2018

Tribunal nombrado pro el Magfco. y Excemo. Sr. Rector de la Universidad Politécnica de Madrid, el día.....de.....de 2018.

Presidente: D.....

Vocal: D.....

Vocal: D.....

Vocal: D.....

Secretario: D.....

Realizando el acto de defensa y lectura de la Tesis el día.....de.....de 2018.

En.....

Calificación.....

EL PRESIDENTE

LOS VOCALES

EL SECRETARIO

Dedication

To my parents and all family members for the unselfish love, constant support, patience and comprehension they have for me.

To my dear departed grandparents whose sacrifices and nurture during the childhood illuminated my life path.

Dedication

A mis padres y los miembros de la familia por su amor incondicional, constante apoyo, paciencia y comprensión para me.

A mis queridos abuelos fallecidos cuyos sacrificios durante la infancia iluminaron el camino de mi vida.

谨献给

对我永远无私关爱、支持、耐心和包容的父母和家人

在孩童时期为用自己的无私奉献照亮我人生道路的外祖父母和祖母

Acknowledgement

Working as a Ph.D. student in INEF-UPM was a magnificent as well as challenging experience to me. Within all these years, many people were instrumental in various manners in shaping up my academic career. This thesis would not have become a reality without the kind support and help of them. I would hereby extend my sincere thanks to all of them.

First and foremost, I would like to express my genuine gratitude to my supervisor Dr. Miguel Ángel Gómez Ruano, for his support, mentorship and guidance throughout the conduct of my thesis. He guaranteed me the maximum freedom to explore my research interests during those years, and was always patient and dedicated whenever I need his help. His timely and scholarly advice, meticulous scrutiny, enthusiasm and scientific approach have enabled me to a great extent to accomplish this thesis.

I owe a deep sense of thanks to Dr. Luis Miguel Ruiz Pérez for his immensely valuable advice and wisdom. What he taught and offered me along those years of living and studying in Spain go beyond knowledge and will keep influencing me in the rest of my life.

I would extend sincere thanks to Dr. Jaime Sampaio for giving me opportunity of research exchange in his laboratory-Creativelab, where I got acquainted with a group of excellent “brains” in sport science. This experience was eye-opening and will always remind me how sport scientists should work, research and most importantly, think creatively every day.

I would also like to express my appreciation to Dr. José Antonio Aparicio Asenjo and coaches of Tennis Federation of Madrid during the early years when I arrived at Madrid. They helped me and offered me a place to work and to accumulate field experience of working with elite tennis players. This experience provided enormous practical knowledge of being a practitioner and a coach of tennis and gave me countless ideas for the completion of the thesis.

Thanks to all the academic staff, administrative staff, technical staff in INEF for their support and assistance. The studying environment in this faculty is friendly and fantastic, to name the least. Special thanks to my Spanish teachers in Salamanca for their efforts and dedication that helped me surmount the language obstacle at the first year of studying in Spain. Sincere appreciation to China Scholarship Council from Ministry of P.R. China, which funded my Ph.D. study.

I would like to express my gratitude to Dr. Hongyou Liu for his continuous help and friendship since the first time we got to know each during my master study. Big thanks to my Chinese and Spanish colleagues in INEF (Zhang, Zhou, Gai, Yi, Gong, Xiao, César) and friends in Spain and Portugal (Bruno, Diogo and Sara) for their company, conversation, and helps in this journey. Sincerely wish them best of luck in their research and work, and every success in their lives.

My privileged thanks to Yixue Lou for her constant support, understanding and love throughout all those years, and most importantly for believing in me even when I was doubtful.

Finally, my forever gratitude to my parents and other family members. The unwavering love, support and encouragement during every moment of my life make me who I am today and will always be the source of courage, faith and perseverance of my life.

Agradecimiento

Trabajando como un estudiante de doctorado en INEF-UPM fue una experiencia magnífica y desafiante. En estos años, muchas personas fueron instrumentales en diversas maneras en la configuración de mi carrera académica. Esta tesis no podría haberse convertido en una realidad sin el amable apoyo y ayuda de ellos. Por este medio, les agradezco sinceramente a todos ellos.

En primer lugar, me gustaría expresar mi genuino agradecimiento a mi supervisor, Dr. Miguel Ángel Gómez Ruano, por su apoyo inquebrantable, tutoría y orientación a lo largo de la realización de mi tesis. Él me garantizó la máxima libertad para explorar mis intereses de investigación durante esos años, y siempre fue paciente y dedicado cada vez que necesito su ayuda. Su oportuno y erudito consejo, meticuloso escrutinio, entusiasmo y enfoque científico me han permitido en gran medida lograr esta tesis.

Le debo un profundo sentido de agradecimiento al Dr. Luis Miguel Ruiz Pérez por su inmensamente valioso consejo y sabiduría. Lo que él me enseñó y proporcionó a lo largo de esos años viviendo y estudiando en España va más allá de conocimiento y seguirá influyéndome en el resto de mi vida.

Le agradezco sinceramente al Dr. Jaime Sampaio por darme la oportunidad de realizar un intercambio de investigación en su laboratorio, Creativelab, donde conocí a un grupo de excelentes "cerebros" en ciencias del deporte. Esta experiencia fue reveladora y siempre me recordará cómo los científicos del deporte deberían trabajar, investigar y, lo más importante, pensar creativamente todos los días.

También me gustaría expresar mi agradecimiento al Dr. José Antonio Aparicio Asenjo y entrenadores de la Federación de Tenis de Madrid durante los primeros años después de llegar a Madrid. Me ayudaron y me ofrecieron un lugar para trabajar y acumular experiencia en el campo de trabajo con los tenistas de élite. Esta experiencia me brindó un enorme conocimiento práctico de ser un practicante y entrenador de tenis y me dio innumerables ideas para completar la tesis.

Gracias a todo el personal académico, administrativo, técnico en INEF por su apoyo y asistencia. El ambiente de estudio en esta facultad es amigable y fantástico, por nombrar lo

menos. Un agradecimiento especial a mis profesores de español en Salamanca por su esfuerzo y dedicación que me ayudaron a superar el obstáculo del idioma en el primer año de estudio en España. Sincero agradecimiento a CSC del Ministerio de Educación de R.P. China por financiar el estudio doctorado.

Me gustaría expresar mi gratitud al Dr. Hongyou Liu por su continua ayuda y amistad desde la primera vez que conocimos a cada uno durante mi estudio de máster. Muchas gracias a mis colegas chinos y españoles en INEF (Zhang, Zhou, Gai, Yi, Gong, Xiao, César) y amigos en España y Portugal (Bruno, Diogo y Sara) por la compañía y ayuda en este viaje. Sinceramente les deseo mucha suerte en su investigación y trabajo, y mucho éxito en sus vidas.

Mis agradecimientos privilegiados a Yixue Lou por su constante apoyo, comprensión y amor a lo largo de todos estos años, y lo más importante por creer en mí incluso cuando tenía dudas.

Finalmente, mi gratitud para siempre a mis padres y otros miembros de mi familia. El amor, el apoyo y el aliento inquebrantables en cada momento de mi vida me convierten en lo que soy hoy y siempre serán la fuente del valor, la fe y la perseverancia de mi vida.

Abstract

Tennis is a complex and dynamic sport where coaches are required to prescribe accurate training programming and match planning, and players are expected to execute the strategy by constantly adjusting their tactics, in order to increase the winning probability. Match performance analysis provides them with accurate, valid and less subjective information that is beneficial for the improvement of player behaviors. Currently, the research of performance analysis in tennis is a well-established realm that addresses the technical-tactical, physical and physiological performance of players. However, there still exist some limitations that need to be further solved, such as the lack of consideration of contextual constraints, the exclusive focus on general performance instead of accounting for within-player variation in personal attributes and performance, and the failure to assess the non-linear relationships among different influencing components from macro-, meso-, micro- and nano-level and match performance. Therefore, in order to address these issues, five independent and correlated studies were undertaken in the thesis.

The first study analyzed the set-to-set variation in performance using match statistics of 146 completed main-draw matches in Australian Open and US Open 2016-2017 men's singles. Comparisons of technical-tactical and physical performance variables were done among different sets; and the coefficients of variation (CV) of these variables were contrasted between match winning and losing players. All comparisons were realized via standardized (Cohen's d) mean differences and uncertainty in the true differences were assessed using non-clinical magnitude-based inferences. Results showed that there was possibly to very likely decreases in serve, net and running related variables and an increase in return and winner related variables in the last sets when compared to the initial sets, indicating the influence of match fatigue and the player's adaption of match tactics and pacing strategy. Besides, winning players were revealed to have lower CV values in most of performance variables except for second serve, winner, and physical performance variables, which indicated that they would sacrifice the consistency to gain more aggressiveness and to dominate the match.

The second study was aimed to (i) analyze the match performance of professional female tennis players in different Grand Slams; (ii) model the relationships between match performance variables and ranking-based relative quality; and (iii) build typical performance profiles for female players in Grand Slams using regression-based approach. Data of a total of

1369 matches were collected within 2014-2017 four Grand Slams. Correlations between 37 performance variables and relative quality were determined and automatically classified into two groups of magnitudes via two-step cluster analysis. Higher-correlated variables were used to build players' typical performance profiles via regression-based technique to give percentage evaluation scores (%ES). Players had more service winners, double faults, return winners and return unforced errors in the Australian Open and US Open, implying a "fast-fast" serve strategy, and higher dominance ratio and better serving performance in Wimbledon. Physical performance became similar in all Grand Slams. All studied variables showed obvious correlation with RQ expect for those of physical performance. The findings indicate that female game in Grand Slams remained to be a contest over baseline; and demonstrate the influence of relative quality on performance.

The third and fourth study explored the differences in technical, tactical and physical performance of male tennis players in four Grand Slams 2015-2017, considering their anthropometric (height and weight) and individual characteristics (relative quality, experience, handedness, backhand style). These attributes were the used to classify players into different groups through cluster analysis. Afterwards, a descriptive discriminant analysis was performed to identify the most important variables that best predict player clusters in each Grand Slam. Results showed that players with higher experience, relative quality and height outperformed other players in serve and return, rally winners and breaking opportunities. While players with one-handed backhanded ran comparatively less distance. Left-handed players played more homogeneously and owned certain dominance on clay court, but their advantage is overrepresented at elite level. Results indicated that players' performance was clearly influenced by game location and personal features and currently male tennis is being tactically aggressive on all court surfaces.

The fifth study made a preliminary attempt to analyze the effects of contextual constraints on point-by-point performance. Data of 145 matches played by an elite male player within four Grand Slams over the consecutive six years (2011 to 2016) were used. Classification trees with the exhaustive CHAID (Chi-square Interaction Detection) algorithm were used to model the relationships between point outcome (win/lose) and various contextual variables (match location, match type, year, match status, quality of opposition, servers and serve number) and performance variables (serve speed and rally length). Game location was shown to have a major effect on the winning percentage of points. The winning chances decreased when

playing in the final rounds and against higher-ranked players. Furthermore, when drawing, losing or facing breakpoint in the game, the performance of player descended as well. While finishing points in short and long rally would contribute to higher winning percentage.

Findings from the current thesis add novel knowledge of how players' performance was influenced by contextual constraints and provide evidence-based insights for individualized training and performance optimization in player development and competition.

Keywords: tennis, performance analysis, performance variation, performance profiling, contextual variables, performance variables, Grand Slams, match statistics

Resumen

El tenis es un deporte complejo y dinámico que requiere por parte de los entrenadores una programación de entrenamiento y de partido precisa. Asimismo, en el tenis los jugadores deben ejecutar la estrategia ajustando constantemente sus tácticas para aumentar la probabilidad de ganar. El análisis de rendimiento de los partidos y la competición les proporciona información precisa, válida y menos subjetiva, que puede beneficiar la mejora de los comportamientos de los jugadores. Actualmente, la investigación del análisis del rendimiento en el tenis es un ámbito en auge que aborda el desempeño técnico-táctico, físico y fisiológico de los jugadores. Sin embargo, todavía existen algunas limitaciones que deben ser resueltas, como la falta de consideración de los factores contextuales, el enfoque exclusivo en el rendimiento general en lugar de tener en cuenta la variación dentro del jugador en los atributos personales y el rendimiento, y la falta de evaluación de las relaciones no lineales entre los diferentes componentes que influyen en el rendimiento: niveles macro-, meso-, micro- y nano en relación con el rendimiento del partido. Por lo tanto, para abordar estos problemas, se han planteado cinco estudios que tiene relación en los aspectos analizados sobre el análisis del rendimiento en el tenis.

El primer estudio analizó la variación en el rendimiento set-a-set utilizando las estadísticas de 146 partidos de individual masculino pertenecientes al cuadro principal en el Abierto de Australia y del Abierto de Estados Unidos 2016-2017. Se realizaron comparaciones de variables técnico-tácticas y de rendimiento físico entre los diferentes sets; y los coeficientes de variación (CV) de estas variables se contrastaron entre los jugadores ganadores y los que perdieron. Todas las comparaciones se realizaron a través de diferencias de medias estandarizadas (Cohen's d) y la incertidumbre en las diferencias verdaderas se evaluó mediante inferencias no clínicas basadas en la magnitud del tamaño del efecto. Los resultados mostraron que hubo disminuciones del rendimiento en las variables relacionadas con el servicio, juego de la red y la carrera, y un aumento del rendimiento en las variables relacionadas al resto y puntos ganadores en los últimos sets en comparación con los primeros sets. Estos resultados indican la influencia de la fatiga del partido y la adaptación del jugador en los aspectos tácticos y las estrategias en el ritmo de juego. Además, se reveló que los jugadores ganadores tienen valores de CV más bajos en la mayoría de las variables de rendimiento, excepto en el segundo servicio,

el ganador y las variables de rendimiento físico, lo que indica que sacrificarían la coherencia para ganar más agresividad y dominar el partido.

El segundo estudio tenía como objetivo: (i) analizar el rendimiento del partido de tenistas femeninas profesionales en diferentes Grand Slams; (ii) modelar las relaciones entre las variables de rendimiento del partido y la calidad relativa basada en el ranking; y (iii) construir perfiles de rendimiento típicos para jugadoras en Grand Slams usando un modelo estadístico de regresión lineal. Los datos de un total de 1369 partidos fueron recolectados dentro de los cuatro Grand Slams de 2014 a 2017. Las correlaciones entre las 37 variables de rendimiento y la calidad relativa se determinaron y se clasificaron automáticamente en dos grupos de magnitudes a través de un análisis clúster bietápico. Se usaron las variables con mayores coeficientes de correlación para construir los perfiles de rendimiento típicos de los jugadores a través de una técnica basada en la regresión lineal que permite una evaluación porcentual (%ES). Los jugadores tuvieron más ganadores de servicio, dobles faltas, ganadores y errores no forzados de resto en el Abierto de Australia y el Abierto de US, implicando una estrategia de servicio "rápido-rápido", y un mayor ratio de dominancia y mejor rendimiento de servicio en Wimbledon. El rendimiento físico resultó ser similar en todos los Grand Slams. Todas las variables del rendimiento estudiadas mostraron correlaciones claras con la calidad relativa excepto los de aspectos físicos. Los hallazgos indican que el juego femenino en Grand Slams seguía estando basado en el juego sobre la línea de fondo; y permiten demostrar la influencia de la calidad relativa en el rendimiento de las jugadoras.

El tercer y cuarto estudio exploraron las diferencias en el rendimiento técnico-táctico y físico de los tenistas masculinos en los cuatro Grand Slams de los años 2015-2017, considerando sus características antropométricas (altura y peso) e individuales (calidad relativa, experiencia, lateralidad, y estilo de revés). Estos atributos se usaron para clasificar jugadores en diferentes grupos a través del análisis de clúster. Posteriormente, se realizó un análisis discriminante para identificar las variables más importantes que mejor predicen los grupos de jugadores en cada Grand Slam. Los resultados mostraron que los jugadores con mayor experiencia, calidad relativa y altura superaron a otros jugadores en saque y resto, ganadores de peloteos y oportunidades de puntos de break. Mientras que los jugadores con el revés de una mano corrieron comparativamente menos distancia. Los jugadores zurdos jugaban de forma más homogénea y poseían cierto grado de dominancia en la cancha de arcilla, pero su ventaja está sobre-representada en el nivel de élite. Los resultados indicaron que el rendimiento de los

jugadores estaba claramente influenciado por la localización del partido y las características personales. Además, los resultados mostraron que el tenis masculino es tácticamente agresivo en todas las superficies de juego.

El quinto estudio realizó un intento preliminar para analizar los efectos de los factores contextuales en el rendimiento punto-a-punto de un jugador de élite. Para ello, se utilizaron datos de 145 partidos realizados por un jugador masculino de élite dentro de cuatro Grand Slams durante los seis años consecutivos (2011 a 2016). Se usaron los árboles de clasificación con el algoritmo exhaustivo CHAID (detección de interacción Chi-cuadrado) para modelar las relaciones entre el resultado del punto (ganar/perder) y las variables contextuales (lugar del partido, tipo de partido, año, estado del partido, calidad de oposición, saques y número de saque) y las variables de rendimiento (velocidad de servicio y la duración del peloteo). El lugar del partido demostró tener un efecto importante en el porcentaje de puntos ganadores. Las probabilidades de ganar disminuyeron cuando se juega en las rondas finales, y contra jugadores de mejores rankings. Además, al empatar, perder o enfrentar un punto de break en el juego, el rendimiento del jugador descendía significativamente. Mientras que los puntos de finalización en el peloteo corto y largo contribuirían a un mayor porcentaje de victorias.

Los hallazgos de la tesis actual incrementan el conocimiento actual sobre cómo el rendimiento de los jugadores está influenciado por los factores contextuales, y permiten proporcionar información basada en la evidencia para el entrenamiento individualizado y la optimización del rendimiento en el desarrollo de jugadores y en la preparación de la competición.

Palabras clave: tenis, análisis de rendimiento, variación de rendimiento, perfil de rendimiento, variables contextuales, variables de rendimiento, Grand Slams, estadísticas de juego

Table of Contents

Dedication	I
Acknowledgement	V
Abstract.....	IX
Table of Contents	XVII
List of Tables.....	XX
List of Figures.....	XXI
Chapter One – Introduction.....	1
1.1 Research Background	1
1.2 Limitation and issues of current investigation.....	4
1.3 Aims of the thesis.....	6
1.4 Thesis structure	7
Chapter Two – Set-to-set Match Performance Variation of Male Tennis Players in Grand Slams	9
2.1 Introduction.....	9
2.2 Method	11
2.2.1 <i>Samples</i>	11
2.2.2 <i>Variables and procedures</i>	11
2.2.3 <i>Statistical analysis</i>	12
2.3 Results.....	13
2.4 Discussion	18
2.5 Conclusion	23
Chapter Three - Performance Profiles of Professional Female Tennis Players in Grand Slams	25
3.1 Introduction.....	25
3.2 Method	27
3.2.1 <i>Sample and data Source</i>	27
3.2.2 <i>Variables and operational definitions</i>	27
3.2.3 <i>Procedures and statistical analysis</i>	28

3.3 Results.....	30
3.3.1 Descriptive statistics of all players.....	30
3.3.2 Results of correlation analysis and cluster analysis.....	34
3.3.3 Profiles of typical performance for player.....	37
3.4 Discussion.....	39
3.4.1 Serve and Return Performance.....	39
3.4.2 Technical and Physical Performance	41
3.4.3 Relationship between match variables and relative quality and cluster Analysis...	42
3.4.4 Performance profiles using percentage evaluation score (%ES)	43
3.5 Conclusion	44
3.6 Limitation.....	45
 Chapter Four – Effects of Experience and Relative Quality in Tennis Match	
Performance during Four Grand Slams.....	47
4.1 Introduction.....	47
4.2 Method	49
4.2.1 Subjects	49
4.2.2 Variables and Operational Definitions	49
4.2.3 Data Reliability.....	51
4.2.4 Statistical Analysis	52
4.3 Results.....	53
4.4 Discussion	61
4.4.1 Serve and Return Performance.....	61
4.4.2 Net Performance	62
4.4.3 Break Point Performance.....	63
4.4.4 Efficiency Performance.....	63
4.4.5 Movement Performance.....	64
4.5 Conclusion	65
 Chapter Five - Clustering tennis players' anthropometric and individual features helps to reveal performance fingerprints	
5.1 Introduction.....	67
5.2 Method	68
5.3 Results.....	71

5.4 Discussion	80
5.5 Conclusion	83
5.6 Practical Application.....	83
Chapter Six - Performance evaluation of point-by-point data for individual male tennis player in Grand Slams.....	85
6.1 Introduction.....	85
6.2 Methods.....	86
6.2.1 <i>Sample and subjects</i>	86
6.2.2 <i>Variables</i>	87
6.2.3 <i>Statistical Analysis</i>	89
6.3 Results.....	90
6.4 Discussion	97
6.5 Conclusion	99
Chapter Seven – Conclusion, Limitations and Future Research Lines	101
7.1 Main conclusions and outcomes of the thesis.....	101
7.2 Limitations	102
7.3 Perspectives for future research.....	103
Chapter Eight – Practical Application.....	105
References	107

List of Tables

Tables are numbered by prefixing the chapter number.

Table 2.1. Descriptive statistics of players' set-to-set match performance	14
Table 3.1. Serve performance of female players in four Grand Slams.	32
Table 3.2. Return, net, break point, efficiency and physical performance of female players in four Grand Slams.	32
Table 4.1. Description of variables and operational definitions	50
Table 4.2. Means (Standard deviations) and structure coefficients (SC) of technical and physical performance for four player clusters in Australian Open and Roland Garros	56
Table 4.3. Means (Standard deviations) and structure coefficients (SC) of technical and physical performance for four player clusters in Wimbledon and US Open	58
Table 5.1. Match performance variables and operational definitions	69
Table 5.2. Means (Standard deviations), discriminant function details and structure coefficients (SC) of technical-tactical and movement performance for four player clusters in Australian Open and Roland Garros	75
Table 5.3. Means (Standard deviations), discriminant function details and structure coefficients (SC) of technical-tactical and movement performance for four player clusters in Wimbledon and US Open	77
Table 6.1. Distribution of Points Played by Novak Djokovic in four Grand Slams	87
Table 6.2. Variables analyzed and correspondent sub-categories	88
Table 6.3. Frequency distribution (%) of point outcome related to situational variables (Crosstab Order: Pearson's Chi-square, degrees of freedom, significance value, expected frequency distribution, and effect size).....	91
Table 6.4. Frequency distribution (%) of point outcome related to situational variables (Crosstab Order: Pearson's Chi-square, degrees of freedom, significance value, expected frequency distribution, and effect size).....	92

List of Figures

Figures are numbered by prefixing the chapter number.

Figure 2.1. Standardized (Cohen's d) set-to-set differences of male players' match performance	15
Figure 2.2. Descriptive statistics the comparisons (standardized differences in Cohen's d, differences in means in percentage and inferences of the true magnitudes for) of CVs of technical-tactical and physical performance variables between winning and losing players.....	17
Figure 3.1. Match performance variables, operational definitions and abbreviations.....	30
Figure 3.2. Relationships between match variables and the relative quality (RQ) for female players in four Grand Slams	35
Figure 3.3. Performance profiles of percentage of evaluation score (%ES) for: all female players, Top-10 players and Serena Williams in four Grand Slams	38
Figure 4.1. Characteristics of four player categories classified by cluster analysis with playing experience (years) and relative quality (RQ) being the inputs	54
Figure 4.2. The distribution of four player clusters (with cluster centroids, best-fit ellipses and quarterfinalists) defined by two discriminant functions	60
Figure 5.1. Description of player groups by cluster analysis.....	73
Figure 5.2. The distribution of four player groups along with their convex-hull areas, complementary dispersion index areas and semifinalists featured by discriminant functions.....	79
Figure 6.1. Analysis by classification tree One of point-level performance for Novak Djokovic in Australian Open & US Open (including serve speed category)	94
Figure 6.2. Analysis by classification tree One of point-level performance for Novak Djokovic in Roland Garros & Wimbledon (including serve speed variable).....	95
Figure 6.3. Analysis by classification tree Two of point-level performance for Novak Djokovic in four Grand Slams (including rally length category)	96

Chapter One – Introduction

1.1 Research Background

Tennis is an open-skill sport where players continuously modify their tactics and techniques in response to particular match constraints. To obtain the success, coaches and players are expected to refine the training programming during match preparation, to optimize the strategy prior to the match, and finally players should apply the corresponding tactics efficiently within the match (Filipcic et al., 2015; O'Donoghue & Ingram, 2001). While in order to achieve those objectives, it is critical that they are supported by useful and precise evidence during the cycle of competing, reflecting, decision-making and preparing for further competition (O'Donoghue, 2014). However, getting that information is not easy as it revolves around a range of contextual constraints related to a complex and dynamic match situation (Carvalho et al., 2014; O'Donoghue & Mayes, 2013; Reid, McMurtrie, & Crespo, 2010). Meanwhile, the limited recall ability of coaches during subjective observation would potentially compromise the quality of evaluation and ultimately bias the data gathering of match events (Laird & Waters, 2008).

In light of this practical issue and the need for objective and unbiased feedback, sport performance analysis is applied to collect valid, accurate and reliable knowledge during competitive environments for later improving athletes' performance (Hughes & Franks, 2007; O'Donoghue, 2014). After more than thirty years of development, the realm of performance analysis has covered the analysis of techniques, tactics, effectiveness, movement patterns and decision-making happened in actual competition through qualitative and quantitative methods that involves little or no human judgement during data collection (Hughes & Franks, 2004; O'Donoghue, 2010, 2014). It overcome the limitation of human recall error and the knowledge gained from sport performance analysis could allow coach, fitness trainer, physiotherapist and players to maximize the possibility of winning via a more evidence-based and scientific planning (O'Donoghue, 2014; O'Donoghue & Mayes, 2013). And along with the evolution of sport technology and their application into the analysis of tennis, plenty of performance analysis research in sport were undertaken by investigators to address the following aspects.

Starting from the techniques analysis which is related with mechanical aspects of tennis techniques and how relevant skills are performed, it is regarded as one of the early research domains that attracted major research attention as techniques are fundamental aspects of sport (O'Donoghue, Girard, & Reid, 2013). Former studies that investigated this domain maintained an equilibrium between the need for experimental control and the need for ecological validity in order to cover relevant factors of tennis skills. The available research analyzed the biomechanical principles of different tennis forehand and backhand groundstrokes (Reid & Elliott, 2002; Reid, Elliott, & Crespo, 2013), serves (Girard, Micallef, & Millet, 2005; Reid, Elliott, & Alderson, 2008) and volleys (Chow, Knudson, Tillman, & Andrew, 2007). Results of those studies increased the knowledge of force generation, coordination of joints and muscular activity during the execution of techniques, providing insights into players expertise in specific skills.

Along with the technical analysis, the analysis of physical and physiological demands of tennis was conducted to quantitatively measure: i) the extrinsic game intensity represented by total playing time, effective playing time, work-to-rest ratio, rally count and duration, movement distance and speeds (Fernandez-Fernandez, Mendez-Villanueva, Fernandez-Garcia, & Terrados, 2007; Fernandez-Fernandez, Sanz-Rivas, & Mendez-Villanueva, 2009; Galé-Ansodi, Castellano, & Usabiaga, 2016; Hoppe et al., 2014; Ojala & Häkkinen, 2013); ii) the intrinsic and individual responses related to heart-rate, oxygen consumption (VO_2), blood lactate concentration, rating of perceived exertion (Fernandez-Fernandez, Sanz-Rivas, & Mendez-Villanueva, 2009; Fernandez-Fernandez, Sanz-Rivas, Sanchez-Muñoz, et al., 2009; Gomes, Coutts, Viveiros, & Aoki, 2011; Hornery, Farrow, Mujika, & Young, 2007). This is an important investigation branch of sport performance analysis in tennis since it directly describes the actual activity profiles of players and match intensity. In addition, findings obtained from the related studies help to deepen the understanding of players' physiological, psychological fatigue triggered by match load and how they modified the pacing strategy, considering the decrease in physical, perceptual and neuromuscular status. This type of studies provided more evidence for a match-oriented fitness-conditioning programming (Gescheit, Cormack, Reid, & Duffield, 2015; Reid & Duffield, 2014).

Further, the analysis of technical effectiveness in tennis was introduced to complement the analysis of techniques and to portray the use of different skills within actual match scenarios (O'Donoghue, 2014; O'Donoghue et al., 2013). Researchers examined the winner-to-error

ratios of different techniques during tennis match-play ranging from first and second serve and return efficiency (Gillet, Leroy, Thouwarecq, & Stein, 2009; Hizan, Whipp, & Reid, 2011), net point (O'Donoghue & Ingram, 2001) and ground stroke success (Martin-Lorente, Campos, & Crespo, 2017; Muhamad, Rashid, Razak, & Salamuddin, 2011). Instead of concerning about mechanical details of these skills, this kind of analysis focuses on their specific application into the match. Useful inferences were generated to indicate which techniques are more effective under certain conditions and to help distinguish novice and expert players (Hizan et al., 2011; Kovalchik & Reid, 2017). At the same time, the analysis of technical effectiveness is usually realized together with the following tactical analysis (O'Donoghue et al., 2013).

In tennis, tactical analysis is concerned with the evaluation of tactical decisions made and broader strategies used and it assesses whether correct technique is chosen and performed by players in determined match situation (O'Donoghue, 2014). Previously when moment-to-moment tactical decisions could not be observed, researchers made inferences from the patterns of play adopted by players during competitive tennis matches (O'Donoghue & Ingram, 2001). Such patterns consisted of the position of players, technique selections, shot types, shot direction and placement, and rally duration; and then researchers and coaches are allowed to infer the tactics from these factors and compare alternative ones in terms of effectiveness (Hizan, Whipp, & Reid, 2015; Hizan, Whipp, Reid, & Wheat, 2014; O'Donoghue et al., 2013). Within this scope, serve and return tactical effectiveness is one of the most studied topics. The serve and return direction, speed and distribution of professional male and female players (Cross & Pollard, 2009; Gillet et al., 2009), junior players (Hizan et al., 2011) and wheelchair players (Sanchez-Pay, Palao, Torres-Luque, & Sanz-Rivas, 2015) were analyzed across different match surfaces, showing different tactical preferences among distinct levels of players. Break point was also analyzed to inspect players' tactical effectiveness during this critical moment (O'Donoghue, 2012). Moreover, thanks to the advance in technology that permits the caption of ball movement (Bal & Dureja, 2012), a recent study took one step further to measure and cluster the serve and shot directions, speeds, landing locations and in- and out- bound angles performed by male and female players during Grand Slam tennis (Kovalchik & Reid, 2018). This research increased the knowledge of tactical performance in the highest level of tennis and provided a framework for characterizing playing style and designing a more representative practice.

In addition to the before-mentioned tactical analyses, the time-motion analysis of players has recently emerged as a “hot” topic in tennis performance analysis. It not only enables us to obtain players’ actual physical performance (movement techniques distance covered, running speed, acceleration, deceleration) during competition (Galé-Ansodi, Castellano, & Usabiaga, 2017; Hughes & Meyers, 2005), but most importantly, it fills the blank where only discrete description of players’ decisions was investigated. With the incorporation of new player-tracking innovations that are non-invasive and capable of capturing and processing player movement accurately during match, players’ spatiotemporal relationships in respect of the interactive coordination could be assessed and understood (Bal & Dureja, 2012; Pereira, van Emmerik, Misuta, Barros, & Moura, 2018). Some preliminary attempts haven been made in tennis to extract useful information from the massive tracking data and to identify the inter-player dynamics during baseline rally (Carvalho et al., 2013; Carvalho et al., 2014; Martinez-Gallego et al., 2013). From these studies, it is possible that shot-by-shot tactical behaviors could be analyzed in the future to unveil more dynamic complexity of competing players and balls; and to improve the understanding of how players regulate their tactical decisions on a point-basis.

1.2 Limitation and issues of current investigation

With all above-mentioned research in mind, it is undeniable that the analysis of tennis performance within actual competition setting has created ecological and meaningful information related to tennis behavior. However, there is a lack of depth in the information and limited the application of studies’ findings into practical scenes if we consider the following aspects.

First of all, thinking of the world-wide consensus that strategy and tactics are vital components of tennis match play performance (Crespo, Reid, & Miley, 2004; O'Donoghue & Ingram, 2001; Reid et al., 2013), it is notable that the contextual constraints involved in tennis match have not been given enough attention. This is a critical issue as the following constraints would influence the strategy planned before the match and tactics adopted during the match: (i) own strengths and weaknesses, and those of their opponents; (ii) the match venue, court surfaces; (iii) tournament type (main-draw format and rounds) and match type (best of three, best of five, super tiebreak, etc.); (iv) types of the ball (weight, coefficient of friction); (v) weather conditions (temperature, humidity, angle of the sunlight, etc.); (vi)

match status and match score (winning, losing or drawing) (Gómez, Lago, & Pollard, 2013); (vii) audience and umpire's influences (Carboch, Vejvodova, & Suss, 2016; Dube & Tatz, 1991). In order to maximize the chances of winning a match, the importance of these factors should be reinforced and logically, their influence have been investigated to give coaches and players more evidence while setting up strategies. But unfortunately, most existing studies considered the merely general match performance of players and tried to determine the key performance indicators without exploring the effects of contextual variables (Filipcic et al., 2015; Ma, Liu, Tan, & Ma, 2013; Reid, Morgan, & Whiteside, 2016). Therefore, the applicability and usefulness of future research will be enhanced if those multifactorial influences are considered.

In the second place, the current literature is limited in building performance profiles of tennis players without considering their performance variation. This statement could be argued via emphasizing that tennis is a highly individual but interacting sport that varied constantly either from match to match or from player to player (O'Donoghue, 2009). Hence, this characteristic determines that the building, interpreting and eventually predicting players' performance should be done acknowledging the variability of the game and different types of players (Butterworth, O'Donoghue, & Cropley, 2013). In the present, some studies attempted to explore the issue of variability by, for example, analyzing the players' physical and technical performance variation during consecutive days of simulated matches (Gescheit et al., 2015; Gescheit et al., 2016). However, the practical inferences were limited as they were not based on real matches. Meanwhile, in terms of player types, although couples of studies took into account the influence of players' handedness, height and ranking on match performance (Loffing, Hagemann, & Strauss, 2009; Reid & Morris, 2013; Vaverka & Cernosek, 2013), they just analyzed these characteristics in isolation and the number of sampling players was comparatively small. A holistic and complex approach that combines them and more players' attributes together is then expected in the future research.

Thirdly, at the elite level, the relationships among macro-level components (e.g. official organizing bodies and policies, long-term player development, annual tournament planning, etc.), meso-level components (e.g. match schedule during one tournaments, travels, injuries, etc.), micro-level components (match strategy, tactics, interpersonal dynamics between competing players, relative quality, match status, score-line, etc.) and nano-level components (physical, physiological, and mental status) are sometime non-linear and jointly determine

players' performance (Cowden, 2016; De Bosscher, De Knop, Van Bottenburg, & Shibli, 2006; Kovalchik, 2015; O'Donoghue, 2009; Palut & Zanone, 2005). Data obtained from these interacting factors is massive and could not be treated with traditional statistical methods. Whilst with the application of the recent popular and powerful data mining and machine learning techniques, this issue could be readily solved and useful and previously unknown information or patterns would be generated to benefit sport performance analysis and coaching process (Ofoghi, Zeleznikow, MacMahon, & Raab, 2013; Whiteside & Reid, 2016b). For the moment, few studies in tennis have already turned to these approaches to describe professional players' serve and shot pattern during Australian Open (Kovalchik & Reid, 2018; Whiteside & Reid, 2016b). These novel initiatives laid a foundation for applying relevant techniques to model the inter-relationships of performance measures and attributes and to contribute to a better ecological understanding of players' behaviours.

In summary, the future investigation of sport performance analysis in tennis is expected to: (i) account for multifactorial contextual constraints and explore their effects; (ii) recognize the variation in performance and evaluate holistically and complexly players' attributes; (iii) model the non-linear relationships among different levels of influencing components on match performance.

1.3 Aims of the thesis

In light of the above-mentioned introduction, the general aim of the current thesis is to evaluate the match performance of elite male and female tennis players during Grand Slams with the emphasis of different contextual constraints. Additionally, this aim is complemented by the following specific aims:

- (i) To analyze the set-based differences in technical-tactical and physical performance of professional male players during hard-court Grand Slams (Australian Open and US Open); and explore the within-player set-to-set variation in performance considering the match outcome
- (ii) To analyze and compare the match performance of professional female tennis players in recent four Grand Slams (Australian Open, Roland Garros, Wimbledon and US Open)

- (iii) To model the relationships between match performance variables and relative quality that is based on players' ranking and build typical performance profiles for female players in Grand Slams applying regression-based profiling technique.
- (iv) To explore different types of tennis players during Grand Slams according to their anthropometric and individual features using unsupervised learning approach and to encounter the differences in technical-tactical and physical performance among distinct player types.
- (v) To make a preliminary attempt of modelling the non-linear correlation between point outcomes and multiple contextual constraints (match location, opponent's quality, match type, match status, serve speed and rally numbers) using data mining technique.

1.4 Thesis structure

In total, the thesis includes seven chapters. Chapter one is the general introduction. Chapter two is elaborated to fulfill the aim (i), and Chapter three is done to complete the aim (ii) and (iii). Chapter four and five are undertaken to achieve the aim (iv) while Chapter six is realized to address the aim (v). Lastly, the Chapter eight concludes the findings from the previous chapters and illustrates the limitation and perspectives for future research.

Chapter two to six constitute the central part of the thesis and they are independent and correlated studies presented in manuscript formats of the scientific journals to which they were submitted. Therefore, they own distinct introduction, methods, results, discussion, conclusion, practical application and limitations according to the publication requirements of corresponding journals. The detailed information of the journals where they are published or under review is illustrated at the start of each chapter.

Chapter Two – Set-to-set Match Performance Variation of Male Tennis Players in Grand Slams

Cui, Y., Gómez, M.-Á., Liu, H. & Gonçalves, B. (In Review). Set-to-set performance variation in tennis Grand Slams: play with consistency and also risks. International Journal of Performance Analysis in Sport.

2.1 Introduction

The complexity and dynamics of tennis determine that player's game strategy and technical-tactical performance are critically influenced by a range of contextual-related constraints (Filipic, Leskosek, & Filipic, 2017; Ojala & Häkkinen, 2013; Pereira et al., 2018). The available research already evidenced that those constraints might be originated from the differences of match location (O'Donoghue & Ingram, 2001), players' qualities (Cui, Gómez, Gonçalves, Liu, & Sampaio, 2017) and weather condition of which players were conscious prior to a match (Morante & Brotherhood, 2007); and from the uncertainty of opponent's actual tactics (Ma et al., 2013), match score-line (O'Donoghue, 2012), spatial-temporal relationship among players and ball (Carvalho et al., 2013; Martinez-Gallego et al., 2013), physiological fatigue (Gomes et al., 2011; Reid & Duffield, 2014), noise of the audience (Dube & Tatz, 1991) and even the line-calling bias from referees that occurred within the match (Carboch et al., 2016). Therefore, adjusting to the constant change of those constraints and making correspondent tactical decisions should be an essential attribute reinforced in the development of professional tennis players (Murray & Hunfalvay, 2017; Triolet, Benguigui, Le Runigo, & Williams, 2013).

With this in mind, players are required to perceive, interpret and anticipate the before-mentioned factors efficiently in order to have enough time to respond and execute the next shot with good techniques (Cowden, 2016; Murray & Hunfalvay, 2017). However, it is difficult to maintain this ability with consistency under highly intensive and competitive match situation given that their physical and physiological status are significantly impaired by match fatigue and the technical performance underwent a decline in efficiency and accuracy (Davey, Thorpe, & Williams, 2002; Gomes et al., 2011; Hornery et al., 2007). Consequently, it would seem to be of interest for coaches, fitness trainers and players to customize the training routines and refine match strategies if more related facts are unveiled.

Contrary to the expectation, most existing literature exclusively focused on reporting the general match performance of players (Ma et al., 2013; Reid et al., 2016), failing to evaluate how different technical, tactical and physical variables were influenced within prolonged time or days as tennis is played as a tournament format. In light of this issue, several studies have attempted to assess how well-trained players' physiological and technical performance were influenced during simulated tournament that lasted 3 to 4 days with a match duration of 2 to 4 hours (Gescheit et al., 2015; Gescheit et al., 2016; Ojala & Häkkinen, 2013). Results of these studies were informative and showed that their technical, physical, physiological and perceptual status were compromised over consecutive days of the tournament due to heavy speed strength load, muscle damage, and neuromuscular and perceptual fatigue. Although being experimental, these studies owned the merit of mimicking what is actually taking place in elite tennis circuits in schedule and match duration. Afterwards, Whiteside and Reid (2016a) made the effort to measure the physical response of professional players during the first week of Australian Open, providing the first-handed feedback of the external hitting and movement workloads within the most onerous Grand Slam tournament and giving sex-specific and match-based inferences for training programming.

The former research inspected how performance changed within competitive tournaments, which would be regarded as a meso-level factor contributing to long-term development of players (De Bosscher et al., 2006). Nonetheless, the available research focused on the variation in performance at a micro-level is scarce. In tennis, this analysis level would denote for the set/game-based inter-relationship between two competing players in terms of relative quality (Klaassen & Magnus, 2001), actual match strategy and tactics (O'Donoghue & Ingram, 2001) and positioning dynamics during the point (Carvalho et al., 2013). Knowing the micro-level variation in performance would not only provide more evidence of how players are influencing each other during the match, but also facilitate to a deeper understanding about the physiological and mental state at an individual nano-level (Cowden, 2016; Gescheit et al., 2016). As far as the current study concerns, the only study that addresses this topic solely investigated the set-to-set difference of serve performance in Wimbledon male's singles (Maquirriain, Baglione, & Cardey, 2016) and found that male players maintained stable serve speeds and accuracy within five-set matches on grass court. Nonetheless, its inferences for practical application is limited if no other performance factors were included and the results should be further tested on other court surfaces.

Finally, when considering the issue of variation in performance from the perspective that athletes could benefit from this variation and eventually achieve a greater consistency in performance and better adaptation to complex match constraints (Seifert, Button, & Davids, 2013). In this vein, it is substantial to know what kind of performance aspects that are better if they are more consistent and are preferred if more varied along with the evolution of the match. Feedbacks produced from this process would be very useful to reinforce training programming and tactical execution.

Therefore, the aims of the current study were: (i) to analyze the set-based differences in technical-tactical and physical performance of professional male players during hard-court Grand Slams (Australian Open and US Open); and (ii) to explore the within-player set-to-set variation in performance considering the match outcome.

2.2 Method

2.2.1 Samples

Set-level statistics of 146 completed main draw matches from 2016 US Open (65 matches) and 2017 Australian Open men's singles (81 matches) were included. Data were extracted from the corresponding official website of each tournament. The inclusion and exclusion criteria for sampling matches were as follows: (i) matches were completed without occurring an early withdrawal of players; and (ii) matches were contested on the court where Doppler radar system (IBM: Armonk, NY, USA) and the Hawk-Eye camera system (Hawk-Eye Innovations, Southampton, UK) were mounted. As men's singles were played by the rule of "Best-of-Five", all the matches concluding in three, four and five sets were considered. In total, there were 543 individual sets collected and they were played on courts where ball speed and player's movement tracking systems were available. The objectivity and reliability of tracking technology have been previously verified, showing slight measurement error (ITF, 2005; Mecheri, Rioult, Mantel, Kauffmann, & Benguigui, 2016). Players involved in the research were ranked 1 to 631. The study was conducted under the approval of the local institutional research commission and all procedures were undertaken to accomplish all international standards.

2.2.2 Variables and procedures

Twenty-three match variables were grouped into serve, return and efficiency & physical performance related categories. Raw data of some match variables were standardized into their correspondent ratios to avoid bias arising from non-uniformity error. Serve performance group included: ace%, serve winner%, double fault%, 1st serve in%, 1st serve won%, 2nd serve won%, peak serve speed (km/h), 1st serve average speed (km/h) and second serve average speed (km/h); return performance group was comprised by: return winner%, return unforced error%, return 1st serve won%, return 2nd serve won%, break points per return game, break point won% and break point saved%; while efficiency and physical group consisted of: winner of total points won%, winner per unforced error ratio, dominance ratio, net points won%, net points won of total points won%, total distance covered in a set (m) and distance covered per point (m). In Grand Slams, the official notational match statistics were collected by professional coding personnel, and their inter- and intra-operator reliability were tested to have an intra-class correlation coefficient (ICC) higher than 0.90 (Hizan et al., 2011; Hizan, Whipp, & Reid, 2010). The inter- and intra-data collectors reliability in this study were also previously verified (Cui et al., 2017), results of intra-class correlation coefficient value ranged from 0.96 to 1 and standardized typical errors varied from 0.02 to 0.11. All tests showed very high level of reliability for the variables of interest (Hopkins, 2000).

2.2.3 Statistical analysis

Descriptive statistics for each of the performance variable was calculated and presented as means with standard deviations (SD). In the first place, a series of comparisons were realized in order to find the differences in performance among all the sets, which included: i) 1st set vs. 2nd, 3rd, 4th and 5th set; ii) 2nd set vs. 3rd, 4th and 5th set; iii) 3rd set vs. 4th and 5th set; and iv) 4th set vs. 5th set. Next, a comparison of set-to-set variation in performance was done between winners and losers in the matches. The coefficient of variations (CV) of all performance variables within all the sets of winner and loser players were compared. To avoid the shortcomings of research approaches using Null-Hypothesis Significance Testing (Batterham & Hopkins, 2006), the comparisons were done via standardized (Cohen's d) mean differences, computed with pooled variance and respective 90% confidence intervals (Hopkins, Marshall, Batterham, & Hanin, 2009). The difference within each pair of comparison was expressed as a factor of the smallest worthwhile difference, based on a small standardized effect of Cohen's d effect-size principle ($0.2 \times$ between-sets SD). Thresholds for effect size statistics were 0.2,

trivial; 0.6, small; 1.2, moderate; 2.0, large; and >2.0, very large. Uncertainty in the true differences of the comparisons was assessed using non-clinical magnitude-based inferences. Magnitudes of clear effects were described according to the following scale: 25–75%, possibly; 75–95%, likely; 95–99%, very likely; .99%, most likely (Hopkins et al., 2009). Differences were deemed clear if 90% CL for the difference in the means did not overlap substantial positive and negative values (Hopkins, 2007).

2.3 Results

Table 1 shows the descriptive statistics of all performance variables from first to fifth sets and Figure 1 show the results of inter-sets comparisons of these variables (standardized Cohen's d units). Compared to the first set: i) there were possibly to very likely decrease in serve winner%, 1st serve won%, 2nd serve won%, serve speeds, return unforced error%, break point saved% and dominance ratio when playing in the 3rd to 5th set, especially in the 4th and 5th set; ii) possibly increase in 1st serve in%, return 2nd serve won% and break point won% in the 5th set; iii) a possibly to likely increase in net success of total point won% in 2nd, 3rd and 5th set; iv) possibly increase in return winner%, return 1st won% and winner per unforced error ratio in the 4th and 5th set.

Trivial differences were shown in all match performance between the 2nd and 3rd set. However, in contrast with the 2nd set, it was found that: i) there were possibly to likely declines in serve winner%, serve speed, break point saved%, net success of total point won% and running distance in the 4th and 5th set; ii) the performance of 1st serve in%, return unforced error%, return 2nd serve won% and break point won% possibly ascended in the 4th and 5th set.

Possibly to likely decrease in double fault%, 1st serve won%, 2nd serve won%, serve speed, winner per unforced error ratio, dominance ratio and net success of total point won% were shown in the 4th and 5th set when in comparison with 3rd set. At the same time, 1st serve in%, return unforced error%, return 1st and 2nd serve won%, break point per return game, break point won% and running distance were possibly higher in the 4th and 5th set.

Finally, when contrasting with 4th set, the performance of 2nd serve won%, serve speed, return unforced error%, break point saved%, winner per unforced error ratio and dominance ratio possibly descended in the 5th set. But, there were possibly increases in 1st serve in%, 1st serve

won%, return 2nd serve won%, break point won%, net point won% and running distance in the last set. All the above results presented small effect size.

Table 2.1. Descriptive statistics of players' set-to-set match performance

Variables	Set No.	1st Set	2nd Set	3rd Set	4th Set	5th Set
		(n=292)	(n=294)	(n=294)	(n=150)	(n=56)
Ace%		9.0 (7.2)	8.3 (7.4)	8.6 (7.7)	8.6 (8.3)	8.3 (7.0)
Serve winner%		2.2 (3.2)	1.8 (2.8)	1.6 (2.3)	1.5 (2.4)	1.6 (2.5)
Double fault%		10.9 (9.7)	10.5 (10.0)	11.8 (10.7)	10.2 (9.7)	10.1 (10.7)
1 st serve in%		59.5 (10.6)	59.4 (11.1)	59.8 (11.1)	58.9 (10.1)	61.1 (9.9)
1 st serve won%		73.4 (12.8)	71.8 (14.2)	73.1 (13.3)	70.4 (13.4)	71.8 (11.7)
2 nd serve won%		51.7 (16.9)	50.7 (17.5)	50.5 (16.4)	49.9 (17.9)	47.9 (16.2)
Fastest Serve Speed (km/h)		204.3 (11.1)	203 (11.3)	202.6 (11.4)	201.9 (12.2)	199.2 (11.1)
1 st Average Speed (km/h)		185.1 (10.3)	184.0 (10.5)	183.0 (10.6)	181.3 (10.1)	180.7 (10.7)
2 nd Average Speed (km/h)		150.0 (11.4)	149.8 (11.3)	149.4 (11.5)	147.2 (10.0)	146.8 (10.6)
Return winner%		1.6 (2.6)	1.9 (3.0)	1.8 (2.6)	2.2 (3.2)	2.2 (2.9)
Return unforced error%		3.6 (4.1)	3.1 (3.8)	3.2 (3.9)	4.2 (5.2)	3.0 (3.6)
Return 1 st serve won%		26.6 (12.8)	28.2 (14.2)	26.9 (13.3)	29.6 (13.4)	28.2 (11.7)
Return 2 nd serve won%		48.3 (16.9)	49.3 (17.5)	49.5 (16.4)	50.1 (17.9)	52.1 (16.2)
Break points per return game		0.6 (0.5)	0.6 (0.5)	0.5 (0.5)	0.6 (0.5)	0.6 (0.5)
Break points won%		36.3 (36.1)	36 (34.3)	35.6 (36.4)	37.8 (33.3)	41.4 (37.1)
Break points saved%		63.7 (36.1)	64 (34.3)	64.4 (36.4)	62.2 (33.3)	58.6 (37.1)
Winner of total points won%		32.0 (11.7)	33.7 (11.5)	34.0 (11.4)	33.8 (12.0)	33.3 (13.0)
Winner per unforced error ratio		1.1 (0.9)	1.4 (1.6)	1.4 (1.5)	1.3 (1.1)	1.2 (0.9)
Dominance ratio		1.2 (0.6)	1.1 (0.7)	1.2 (0.9)	1.2 (0.9)	1.1 (0.5)
Net point won%		65.4 (23.9)	65.0 (22.5)	65.0 (23.1)	62.8 (23.1)	65.6 (25.1)
Net success of total points won%		14.2 (8.3)	16.8 (9.9)	16.7 (10.1)	15.2 (9.3)	15.3 (9.2)
Total distance covered (m)		640.4 (241.4)	658.5 (250)	630.7 (253.4)	623.2 (260)	653.3 (267.3)
Distance covered per point (m)		10.3 (2.7)	10.7 (2.9)	10.3 (2.9)	10.2 (2.8)	10.6 (3.2)

Figure 2.1. Standardized (Cohen's d) set-to-set differences of male players' match performance

Figure 2 illustrated the descriptive statistics of set-to-set CVs of performance variables between match winners and losers, and the results of comparisons between them. It was shown that winning players had possibly to very likely higher CVs in double fault%, 1st serve in%, 2nd average serve speed, return winner%, return UE%, winner per unforced error ratio, dominance ratio, and running distance than losers, with a small to moderate magnitude effect. While the match losers presented possibly to most likely higher CVs in ace%, serve winner%, 1st and 2nd serve won%, 1st and 2nd return point won%, winner of total point won% and especially in break-point related variables: break point per return game (88 ± 47), and break point won% (115 ± 66) than winners, with moderate to large magnitude effects. There was only trivial difference in relation to serve speed and net point performance related variables.

2.4 Discussion

This study was aimed to analyze the differences in performance on a set-to-set basis for Grand Slam tennis players, and to explore how winning players were differentiated from losing ones in performance stability when playing against heavy match load. Results demonstrated a decline in serve, net and running related performance and an increase in return and winner related performance of male players, as match proceeded into the later sets. Furthermore, match winners were revealed to have a better overall consistency in all performance than losing players except for serve speed, net point, and physical performance. The study provided evidence-based insight regarding the influence of fatigue within tennis matches. The findings allowed us to inspect tennis players' behavior throughout the match and helped to reconsider the evaluation of key performance indicators that differentiated winner and loser players.

The player's serve related performance in Wimbledon was previously analyzed and results showed that players could maintain constant serve performance over five-set match and were capable to overcome fatigue (Maquirriain et al., 2016). However, contrary to the above findings, the current study found that not only players' serve performance, but also return and winner related performance underwent a decrease. Among those variables, serve speed turned out to be the most obvious one. The peak serve speed, 1st and 2nd average serve speed all exhibited a gradual decrease from the first to the fifth set. Due to the influence of serve speed, players could not achieve the same number of serve winners like the first set as well, as returning players had more to react to a serve. These findings may imply that professional players' performance was still impaired by match fatigue.

The disagreement between the current study with the previous research could be attributed to the difference of court surfaces samples studied. As Wimbledon is a fast-pace tournament played on grass court, matches are finished in less time with fewer shot exchanges (Ma et al., 2013; O'Donoghue & Ingram, 2001). Therefore, players could get recovered from the last point and focus on their serving strategy, thus resulting in a constant serving behavior. But unlike the study of Maquirriain et al. (2016) that compared merely the performance between the first and the fifth set, our study evaluated players set-to-set performance in Australian Open and US open where court surface is slower. Consequently, it is reasonable to state that players' fatigue is gradually induced by lower shot rate and longer rallies, and then, the players could not maintain the same level of consistency in serve like in Wimbledon matches.

Further, player's return, breakpoint, efficiency and physical performances were also analyzed in order to comprehensively examine different performance aspects of tennis players within consecutive sets. Return players were able to win more first return points in the fourth set than in the first and third set, and more second return points in the fifth set than in the first set. In addition, they made more return unforced errors in the fourth set than in the second, third and fifth set. There are two possible reasons to these results. First of all, due to the decrease in serve speed and serve winner caused by match fatigue (Gescheit et al., 2015), returners could own more time to prepare for a good return in the last sets. Secondly, players developed perceptual adaptation to their opponents' playing styles and preferences as match proceeded, which would allow them to anticipate the server's tactics. Consequently, returning players could tend to use more aggressive returning tactics in order to obtain the dominance in return game through opening up more angles, fast return to server's feet, or return and volley like the recent famous Sneak-Attack-by-Roger. Notably, compared to the fourth set, the reduction of return unforced error in fifth set might imply that players chose a conservative returning tactic, given the fact that the serve speed has decreased obviously from the initial sets to the final sets, with no difference being found between the fourth and fifth set. Although being intuitive, we assume that under the high pressure of the last set, players tried to avoid losing easy points out of own mistakes while returning in order to break server's games in a safer way. However, as trivial differences were shown among all sets in breakpoint performance it would mean that this tactic did not lead to better performance in returning games. This finding was also reflected by the inter-set comparison of dominance ratio which exhibited a possible decrement in the fifth set. As dominance ratio represents the ratio of points won in return games against points lost in service games, a lower value in the final set indicates that the intensity and stress of the match was so elevated that it was hard for players to win "easy" points.

Possibly higher winner/total% and winner per unforced error ratio shown in the second to fifth set than the first set indicates that players built their game on active attacking either from baseline or go directly to the net to end the points. This was not only because they grew familiarity with opponents and the match surroundings, but also played with more risks to shorten point durations as they faced limitations in physical capacity under consecutive days and sets of high intensive matches (Gescheit et al., 2016). Additionally, this argument is supported by the finding that a possibly to likely higher net points won of total points won% was demonstrated in the second, third and fifth set than the first, which suggests that players

were inclined to finish more points by attacking in the net even though their net point success remained almost unchanged. Finally, in terms of distance covered, it was shown that only trivial to small differences were shown among all sets. From this perspective, it might seem that players running performance was not impaired by the match duration and fatigue and they could cover similar distance independently of the set. Whereas, when relating it to other performance variables discussed above (decreased serve speed, less serve winner and double faults, more first serve in% and net points played, etc.), it could be inferred that Grand Slam competitors still underwent heavy speed strength load and muscle damages. Besides, it is possible that the variation in technical and tactical performance was not only subject to situation variables such as location, surface, or opponents (Gómez, Lago, et al., 2013), but also greatly to the physical impairment (Gescheit et al., 2016). Furthermore, Grand Slam matches are played with one-day interval, meaning that player could have at least 24 hours of recovery if they get into the next round. But a previous study stated that one day of rest during the tournament was insufficient to recover explosive attributes of body muscles (Ojala & Häkkinen, 2013), which played major roles in the process of players' accelerating, decelerating, stretching, jumping and striking the ball. Therefore, in order to improve the understanding of how players' match performance is affected by the effects of consecutive sets and how they modify their match-play accordingly, it is expected that future research could inspect closely the following aspects the current study was unable to include: serve and return directions and techniques, running directions, speeds and strokes speeds, net clearance of balls.

Winning and losing performance is always one of the top research interests in performance analysis as it helps to identify key performance indicators (Sampaio & Leite, 2013) that better discriminate winner and loser as well as to tailor coaching process, setting concrete and practical goals in training and match preparation. Contrary to other major sports like basketball and football where numerous studies have shed light on how winning teams were differentiated with losing teams (Gómez, Lorenzo, Sampaio, Ibáñez, & Ortega, 2008; Liu, Hopkins, & Gomez, 2016), there are few studies that attempted to distinguish winners and losers' performance in tennis (Ma et al., 2013; Martinez-Gallego et al., 2013), not to mention assessing the within-match variability in performance. To the best of our knowledge, the current research is the first study that investigated tennis players' set-to-set variation of performance in such a complex manner. From the results of the comparisons, we could infer

that winners in Grand Slam matches had more consistency in performance than losers during the entire match.

The formers were more stable in almost all serve and return variables except for double fault%, 1st serve in%, 2nd serve speed, return winner% and return unforced error%, which represents the disparity between two groups of players in technical-tactical and physical capacities. On the one hand, more varied performance in these variables implies that winners maintained an aggressive strategy throughout the match. And instead of reducing serve speeds and aiming for safer locations, they probably ran more risks in both first and second serves and returns to chase accuracy and angles (Whiteside & Reid, 2016b) and establish similar level of dominance over opponents when serving within critical sets (Reid et al., 2016). On the other hand, it might be possible that they had clear arrangement of physical deposit and prioritized the serving and returning ability to avoid running with precipitated movements and striking excessive numbers of rally, which accelerate the appearance of match fatigue (Ojala & Häkkinen, 2013). To justify these findings, results demonstrated that they had a better stability in aces, 1st and 2nd serve and return points won%, while owning similar level of first serve speed. Therefore, based on the previous study that regarded the first and second serve as important performance indicators that differentiate player groups (Cui et al., 2017; Reid et al., 2010), it is suggested that taking risks while guaranteeing a relative lower variation in performance during service games is a critical predictor of winners and losers.

Furthermore, as winners own high efficiency and consistency in return games, they consequently obtained more breaking opportunities. But it is notable that losers had a higher variability in break point won% with a large magnitude and in break point saved % with a small magnitude than winners. Obviously, these would again highlight the different returning capacity between two groups. But it is also very likely that they were influenced by the match-status or score-line as there were many situations where winners had so great advantage that losing players might achieve none break point when returning, and similarly it might occur that losing players failed to save the first break point when serving (O'Donoghue, 2012). Whereas, to interpret the results from another perspective, we infer that winners could win or save stably more break points due to the following reasons: i) they know that were dominating the match and were motivated to win such critical points; and ii) they deliberately trained themselves under similarly stressful situations during training to be more prepared in real match scenarios. In other words, encountering a lower variability of breaking

performance in winners would mean that they were psychologically prepared and tactically resourceful. These findings are in agreement with previous studies (Cui et al., 2017; O'Donoghue, 2012), then future studies should consider the combined effect of match-scores, set scores and game-scores on player's performance and inspect their details behaviors under such circumstances.

In terms of physical performance, although male players covered similar total distance and distance per point on a set-wise, winners of the matches traversed with a higher variation in these variables when compared to losing players. It was reported previously that male professional winners ran more distance than losers during hard court tennis matches. Nonetheless, the identification of such perturbations in winning players is a factor that has never been reported before in tennis. A possible explanation that addresses this phenomenon would be related to players' positional status within matches (Carvalho et al., 2013). It was found that the spatial-temporal positioning stability between two competing tennis players within rally is defined by players' tactical arrangement and technical execution and players would either move in the same direction or the opposite ones (Carvalho et al., 2014), and any break in this inter-player coordination started by one player would promote an abrupt positional change of another. Based on this theory, it is inferred that winners would always try to take the initiative to move their opponents, bringing them more precipitations in movements and preventing them from adjusting to the stroke rhythm. But as in each set, losing players were also consistently adjusting themselves to the complex game status, and developing new tactical and perceptual adaptations, winning players have to modify the tactics in order to make opponents commit errors or to find better chances to terminate the points. However, due to the objective situation that massive spatial-temporal data relating to player's movement during official matches is currently restricted from public access, no further evidence could be generated from the limited results to support this argument. While facing this practical limitation, researchers and coaches is encouraged to find alternatives to address this recently popular research topic and to individualize player' training session.

While the current study offers more insights into professional players' set-based performance, there exist some pieces of consideration that should be further addressed. First of all, the study only used male players' performance in hard-court Grand Slams and cautions should be taken when generalizing the findings to other surfaces or to female counterparts. Secondly, detailed technical-tactical preferences of players within determined sets should be considered

and quantified along with contextual variations (match/set/game score-line, opponents' characteristics, etc.) to achieve a more practical and representative description of performance, and to provide first-handed illustration for practitioner during coaching process. Moreover, time-motion characteristics of two competing player in relation to ball features (location, speed and angle) would be the focus of future research in tennis and findings from these studies would be fundamental to help players full exploit the court space and optimize decision-making.

2.5 Conclusion

This study analyzed professional male tennis players' performance among different sets of match and compared the variability of performance between match winners and losers using the data from Australian Open and US Open. Results showed that: i) male players experienced a decrease in serve, net and running related performance and an increase in return and winner related performance within the match, indicating the influence of fatigue on technical performance, tactical adaptation and pacing strategy; ii) match winners played consistently in almost all performance related variables but they would sacrifice the stability in second serve, winner, and physical performance to be more aggressive so that they could dominate the match. The study provided evidence-based information regarding the influence of fatigue within Grand Slam tennis tournament. Findings of the current study advance the knowledge of elite players' match-play performance and game behavior that could be utilized for a fine-tuned training prescription.

Chapter Three - Performance Profiles of Professional Female Tennis Players in Grand Slams

Cui, Y., Gómez, M.-Á., Gonçalves, B., & Sampaio, J. (In Review). Performance profiles of professional female tennis players in Grand Slams. PLOS ONE.

3.1 Introduction

The competitiveness of four tennis Grand Slams (Australian Open, Roland Garros, Wimbledon and US Open) that are held every year represents the maximum level of competition of this sport and its progression (Cross & Pollard, 2009; Prieto-Bermejo & Gómez-Ruano, 2016). Each tournament requires the player to play with a draw of 128 competitors within a period of two weeks if they aspire to win the final trophy, higher ranking points or prizes (Corral, 2009). Therefore, the importance of those slams raised a great number of studies on how players performed in those events and provided valuable feedback for coaching and training in tennis (Hizan et al., 2015; Ma et al., 2013; Moss & O'Donoghue, 2015; O'Donoghue, 2012; O'Donoghue & Brown, 2009; O'Donoghue & Ingram, 2001; Reid et al., 2016; Whiteside & Reid, 2016b).

The difference of court surfaces in four Grand Slams determines that players adjust their strategies on a slam basis (O'Donoghue & Ingram, 2001) and it is necessary to understand how they respond accordingly in different events. Few studies have made thoroughly an examination of how professional female players performed in four Grand Slam events (O'Donoghue & Ingram, 2001; O'Donoghue & Brown, 2008), not to mention how they performed in different court surfaces (Gillet et al., 2009). O'Donoghue and Ingram (2001) reported the match time and rally patters in four Grand Slams by using notational data and O'Donoghue (2002) analyzed the game in the Australian Open based on match outcome. Recently Reid et al (2016) described the match-play characteristics in the Australian Open with data from Hawk-eye tracking system. Many of those studies just considered the performance of female players in certain Grand Slam like Australian Open (Hizan et al., 2014) and failed to compare that with other slams.

In general, most current studies are focused on the performance of male players from a physiological perspective. It was reported that players experienced a decrease in body mass and blood lactate while an increase in heart rate, perception of effort and salivary cortisol

concentrations on a set-to-set basis during the match, which showed the increasing fatigue and stress male players had to cope with under real match scenario (Gomes et al., 2011; Hornery et al., 2007). Meanwhile, the movement patterns (Gillet, Leroy, Thouwarecq, Megrot, & Stein, 2010; Martinez-Gallego et al., 2013) and the technical and tactical effectiveness of male players were also widely investigated (Cui et al., 2017; Gillet et al., 2009; O'Donoghue, 2012; Reid et al., 2010). The previous studies found that first and second serve capability (direction, speed and success), breakpoint opportunity, net point success and distance covered within the match were important predictors to their success and the differentiation of player types. Consequently, there is a need to update the understanding of female players' performance in Grand Slams, particularly by identifying their current profiles with recommended approaches (Butterworth et al., 2013; McGarry, O'Donoghue, & Sampaio, 2013).

The existing normative performance profiling technique (O'Donoghue, 2005) proposed to describe tennis performance uses the median, lower and upper quartiles of normalized performance indicators to represent the performance of a group of players or single player, and maps them onto a radar chart. This not only facilitate to an instant visualization of player's performance profile across various performance indicators which were originally in different units, but helps to inspect or compare the consistency and inconsistency in performance between determined players against certain percentile banding of population of interest (Butterworth et al., 2013; O'Donoghue & Cullinane, 2011). However, criticism of this approach is that there is a risk of misinterpreting individual performance if the quality of opponent is not included in the player profiles (McGarry et al., 2013). In response to the limitations, O'Donoghue and Cullinane (2011) brought out a new profiling technique that can properly evaluate and interpret the performance of individual tennis players while taking into account rivals of different strength. The approach considers primarily the relative quality (RQ) of two tennis players that compete in a match, utilizing their 52-week world ranking, and then it provides percentage evaluation scores (%ES) of each performance variable to interpret how good or bad players performed when they are against different opponents. The usefulness of the technique was evidenced in profiling individual performance, typical performance and performance trends in tennis. However, in the study of O'Donoghue and Cullinane (2011), there were few performance variables used to establish the profile of professional male tennis players in the Australian Open. The correlations between other performance variables and relative quality remain unknown not only for male players, but also for female players within four Grand Slams, considering the effects of different courts.

As match-related statistics can provide insightful information on players' perception and action. Former studies tended to analyze the performance variables in isolation: assessing solely serve and return performance variables like aces, first serve points in, first serve points won, second serve points won (Hizan et al., 2011); or break point performance variables (O'Donoghue, 2012); or physical performance variables (Pereira et al., 2017). In order to address this issue, it is necessary to bring together all performance variables of interest and to evaluate them comprehensively if we want to represent the performance of elite female players in different slams (Butterworth et al., 2013).

Therefore, the aim of the present study was to: (i) analyze and compare the match performance of professional female tennis players in recent four Grand Slams; (ii) model the relationships between match performance variables and relative quality; and (iii) build typical performance profiles for female players in Grand Slams.

3.2 Method

3.2.1 Sample and data Source

Data from 1369 matches played in four Grand Slams women's singles were obtained from separate official websites, which consisted of 499 matches from Australian Open (2014-2017, www.ausopen.com), 249 matches from Roland Garros (2015-2016, www.rolandgarros.com), 249 matches from Wimbledon (2015-2016, www.wimbledon.com) and 372 matches from US Open (2014-2016, www.usopen.org). Walkover matches or those with players who retired were excluded from the study. The data included the match statistics from 257 unique players (2738 observations) and their 52-week world rankings (ranking range: 1 to 992) in the corresponding slam. The study was approved by the local Institutional Review Board and all procedures are conducted following the European Data Protection Law in order to maintain the anonymity of sampled players.

3.2.2 Variables and operational definitions

The study included 37 variables that were selected from the raw data and Fig 1 shows their names and operational definitions.

It was impossible to assess the reliability of the tracking data (i.e. mean meters run per match/set/point) by Hawk-eye system, although being used in previous studies (Reid et al., 2016; Whiteside & Reid, 2016b). Therefore, two experienced performance analysts in tennis observed and collected the non-tracking data (all notational variables shown in Fig. 1 excluding physical performance-related variables) of two matches that were randomly selected. Afterwards, comparisons were made between inter-observers over all variables. The minimum Cohen's *kappa* value for all variables exceeded 0.90, while the intra-class coefficients (ICC) ranged from 0.96 to 1 and standardized typical errors varied from 0.02 to 0.11, demonstrating high inter-rater reliability (Hopkins, 2000; O'Donoghue, 2010).

3.2.3 Procedures and statistical analysis

Match statistics from the players and their 52-week world ranking in each slam were initially collected into spreadsheets for the calculation of relative quality. Afterwards, the raw data of each considered match performance variable was calculated into either percentage or ratio scale (shown in Fig 1) to avoid the misinterpretation caused by different units and the respective missing values were cleared. A one-way analysis of variance (ANOVA) was used to calculate the differences of all variables among four Grand Slams. Pairwise differences were assessed with Bonferroni *post hoc* test and partial eta squared (η^2) was used as the effect size estimate. The alpha level for significance was set at $p \leq 0.05$. The data package of IBM SPSS Statistics for Windows (Armonk, NY: IBM Corp.) was used for the ANOVA and the following cluster analysis.

Afterwards, according to the methods of O'Donoghue and Cullinane (O'Donoghue & Cullinane, 2011), the relative quality of for each player and their opponents was then calculated using their 52-week world ranking through the following two steps:

1. Using the equation to estimate the round of a slam that a player is expected to reach based on her ranking (Klaassen & Magnus, 2001):

$$R_X = 8 - \log_2(Rank_X)$$

where $Rank_X$ is the ranking of player X and R_X is the round she is expected to reach.

2. Determining the relative quality (RQ) for each match played by player X and player Y:

$$RQ = R_X - R_Y$$

Pearson correlation coefficient was calculated to determine the relationships between RQ and each of those 37 variables in four Grand Slams and the magnitudes of the associations was assessed by a spreadsheet developed by Hopkins (2006), applying non-clinical magnitude-based inferences with 99% confidence intervals. Strength of the correlation were interpreted as trivial (0.0–0.1), small (0.1–0.3), moderate (0.3–0.5), large (0.5–0.7), very large (0.7–0.9) and nearly perfect (0.9–1.0) (Hopkins et al., 2009). Magnitudes were defined as unclear when the confidence limits for the correlation coefficient overlapped both substantial positive and negative values, while clear correlation were described according to the following scale: 25–75%, possible; 75–95%, likely; 95–99%, very likely; >99%, most likely (Hopkins et al., 2009).

In order to select the variables that had clearly stronger correlations with RQ in different slams for the evaluation of player's performance, a two-step cluster analysis with Euclidean as the distance measure and Schwartz's Bayesian criterion was used to classify the magnitudes of correlations. These data were automatically clustered into weaker ($r = 0.089 \pm 0.092$, min to max: -0.117 to 0.221, $n = 90$ correlations) and stronger correlations ($r = 0.347 \pm 0.077$, min to max: 0.223 to 0.549, $n = 58$ correlations).

Finally, the typical performance profiles of all players were established by calculating the percentage evaluation score (%ES) as follows (O'Donoghue & Cullinane, 2011):

1. The simple linear regression equations of selected variables and players' RQ were built (one variable at a time) to determine the expected value of these variables.
2. Calculating the residual value of expected value and observed value, which stand for how much better or worse a player performed than she was expected to do when competing with a rival of certain ranking.
3. Dividing the residual by the standard deviation of the residuals in the data used in the regression model to give a Z-score.
4. Using the standard normal distribution to determine the area of the probability distribution for the residuals for Z-scores less than the one calculated the last step through the "NORMSDIST" function in Microsoft Excel and multiplying this probability by 100% to obtain the evaluation score (%ES). It means the percentage of matches where a performance variable value would be expected to be lower than the observed value considering the RQ of two competing players.

Means of %ES of those variables for all players, Top 10 ranked players and a determined individual were calculated and plotted into radar charts in the Excel.

Variable	Operational Definition	Abbreviation
Serve Performance		
Aces of all serves	% of aces in total number of serves (ace is a successful serve that opponent fails to touch by her racquet)	Ace%
Service winners of all serves	% of service winners in total services (return winner is a serve that is touched by opponent with her racquet but not returned into the court)	SW/TS%
First Serve		
First serve in	% of first serves that were in play	1stIn%
First serve points won	% of points won when first serves were in play	1stW%
in deuce side	% of points won when first serves were served into the deuce side	1stDW%
in advantage side	% of points won when first serves were served into the advantage side	1stAdW%
in Wide of the deuce side	% of points won when first serves landed in the Wide zone of the deuce side	1stDWide%
in T of the deuce side	% of points won when first serves landed in the T zone of the deuce side	1stDT%
in Body of the deuce side	% of points won when first serves landed in the Body zone of the deuce side	1stDBody%
in Wide of the advantage side	% of points won when first serves landed in the Wide zone of the advantage side	1stAdWide%
in T of the advantage side	% of points won when first serves landed in the T zone of the advantage side	1stAdT%
in Body of the advantage side	% of points won when first serves landed in the Body zone of the advantage side	1stAdBody%
Second Serve		
Second serve points won	% of points won when second serves were in play	2ndW%
Double Faults of second serves	% of double faults in total number of second serves	DF%
in deuce side	% of points won when second serves were served into the deuce side	2ndDW%
in advantage side	% of points won when second serves were served into the advantage side	2ndAdW%
in Wide of the deuce side	% of points won when second serves landed in the Wide zone of the deuce side	2ndDWide%
in T of the deuce side	% of points won when second serves landed in the T zone of the deuce side	2ndDT%
in Body of the deuce side	% of points won when second serves landed in the Body zone of the deuce side	2ndDBody%
in Wide of the advantage side	% of points won when second serves landed in the Wide zone of the advantage side	2ndAdWide%
in T of the advantage side	% of points won when second serves landed in the T zone of the advantage side	2ndAdT%
in Body of the advantage side	% of points won when second serves landed in the Body zone of the advantage side	2ndAdBody%
Return Performance		
Return points won of total returns	% of points won when returning serves	RW%
Returning first serves won	% of points won when returning first serves	R1stW%
Returning second serves won	% of points won when returning second serves	R2ndW%
Return winners of total returns	% of return winners in total serve returns (return winner is a serve that successful returned by returner into the court and the server does not hit the ball)	RWinner%
Return unforced errors of total returns	% of unforced errors in total serve returns (unforced error is when the player has time to prepare and position herself to get the ball back in play and makes an error)	RUE%
Net Performance		
Net points won	% of points won when player approached the net	NetW%
Net points won in total points won	% of net points won of total points won	Net/TW%
Break Performance		
Break points each return game	Number of break points had in every opponent's service game	Bpt/ReGame
Break points won	% of break points converted in opponent's service game	Bwon%
Break points saved	% of break points served in own service game	Bsaved%
Efficiency Performance		
Winner per unforced error ratio	Number of winners in every unforced error	Winner/UE
Dominance ratio	% of points won in opponent's serves / % of points lost in own serve	DR
Physical Performance		
Mean meters run per match	Mean distance covered in a match by individual player	Dist Total
Mean meters run per set	Mean distance covered in each set by individual player	Dist/Set
Mean meters run per point	Mean distance covered in each point by individual player	Dist/Pt

The illustration of the service side (Advantage and Deuce) and service placement (Wide, Body and T)

Figure 3.1. Match performance variables, operational definitions and abbreviations

3.3 Results

3.3.1 Descriptive statistics of all players

Table 3.1 presents the serve performance of the match variables in the four Grand Slams for all female players. Players won more first and second service points at an average of 66.78% (11.1%) and 46.44% (12.7%) respectively in Wimbledon (W), and served more aces with a mean of 4.84% (5%) and committed more double faults averaged at 13.40% (8.3%) in Australian Open (AO). Moreover, more service winners were obtained in slams of hard courts (AO and US Open). When considering the different serve directions in different serves, players won more points in all slams when the first serves landed in the T zone of deuce court

and wide zone of advantage court. But for second serves, higher percentage of wins was achieved when balls were served to the body zone in both of the deuce and advantage court.

Table 3.2. depicts the performance of return, net point, breakpoint, efficiency and distance covered. More points were won when returning first serves at an average of 38.08% (11.4%) in Roland Garros (RG), and AO was the slam where players had higher percentage of return winners, as well as the return unforced errors, with a mean of 5.19% (5.5%) and 8.69% (7.7%) respectively. The performance of returning the second serves is similar in four events.

Similar net performance was shown in all slams although players tended to approach the net more often in Wimbledon with 13.5% (7.9%) of the total points won being achieved in the net. In RG, players won slightly more break points at an average of 47.71% (23.8%) and in W, 55.56% (24.3%) of the break points was saved, higher than the rest of slams. Moreover, it is shown that players owned a higher average dominance ratio of 1.14 (1.0) in Wimbledon, achieving more winners against each unforced error at a ratio of 1.18 (0.8). Players ran more meters in RG with total distance covered, distance per set and distance per points being higher the other slams.

Table 3.1. Serve performance of female players in four Grand Slams.

Variables	Australian Open (n range: 764-998)	Roland Garros (n range: 350-498)	Wimbledon (n range: 392-498)	US Open (n range: 526-744)	F	p	partial eta squared	
	M (SD)	M (SD)	M (SD)	M (SD)				
Aces of all serves (%)	4.84 (5.0) #	3.37 (3.6) ¶†	4.79 (5.3)	4.43 (4.7)	8.42	<0.001	0.012	
Service winners of all serves (%)	1.05 (1.6) #‡	0.35 (0.8) ¶†	0.82 (1.3) §	1.19 (1.8)	34.57	<0.001	0.042	
First Serve								
First serve in (%)	61.32 (8.4) #‡	63.72 (8.4) †	62.76 (7.9) §	60.42 (8.4)	18.94	<0.001	0.02	
First serve points won (%)	64.67 (11.5) #‡	61.92 (11.4) ¶†	66.78 (11.1) §	64.40 (11.8)	14.98	<0.001	0.016	
in deuce court (%)	64.70 (14.0) ‡	62.42 (13.5) ¶†	67.71 (13.4)	65.65 (13.9)	9.62	<0.001	0.014	
in advantage court (%)	64.61 (14.3) #	61.91 (14.2) ¶	66.27 (14.2)	63.98 (14.2)	6.01	<0.001	0.009	
First serve points won	in Wide zone of the deuce court (%)	65.85 (20.6)	62.62 (23.9) ¶†	67.87 (17.8)	68.04 (23.3)	5.39	0.001	0.008
	in T zone of the deuce court (%)	67.55 (21.3)	63.93 (27.7) ¶†	69.68 (20.5)	68.80 (24.7)	4.41	0.004	0.006
	in Body zone of the deuce court (%)	52.60 (30.5) #	58.78 (21.6) ¶	49.68 (36.9) §	58.02 (24.2)	9.81	<0.001	0.014
	in Wide zone of the advantage court (%)	66.65 (24.6)	63.48 (24.5)	67.95 (21.6)	65.28 (26.1)	2.4	0.066	0.004
	in T zone of the advantage court (%)	64.35 (22.0) #	59.15 (26.9) ¶†	64.86 (20.8)	64.97 (24.8)	5.43	0.001	0.008
	in Body zone of the advantage court (%)	49.56 (33.6) #‡	57.78 (24.5) ¶	46.89 (38.0) §	56.92 (26.7)	13.11	<0.001	0.019
Second Serve								
Second serve points won (%)	45.09 (11.9)	45.54 (12.4)	46.44 (12.7)	45.40 (12.5)	1.36	0.252	0.001	
Double Faults of second serves (%)	13.40 (8.3) #	11.68 (8.7) †	12.40 (8.5)	13.30 (8.3)	4.25	0.005	0.006	
in deuce court (%)	52.40 (17.6)	50.10 (17.5)	52.13 (18.03)	52.56 (17.8)	1.66	0.173	0.002	
in advantage court (%)	52.20 (19.2)	52.84 (19.0)	54.29 (18.7)	51.81 (17.8)	1.52	0.207	0.002	
in Wide zone of the deuce court (%)	40.86 (40.0)	37.59 (39.8)	44.55 (36.3)	39.11 (41.8)	2.21	0.085	0.003	
Second serve points won	in T zone of the deuce court (%)	44.06 (34.8) #	36.67 (37.7) ¶	44.13 (34.7)	40.15 (39.0)	4.19	0.006	0.006
	in Body zone of the deuce court (%)	50.18 (27.5)	49.59 (23.3)	50.57 (32.5)	50.25 (23.7)	0.08	0.969	0
	in Wide zone of the advantage court (%)	44.92 (34.8)	49.30 (35.3) †	49.11 (31.8) §	40.71 (37.4)	6.1	<0.001	0.009
	in T zone of the advantage court (%)	41.79 (39.8) #‡	31.79 (40.0) ¶	41.89 (38.6)	35.40 (42.3)	6.9	<0.001	0.01
	in Body zone of the advantage court (%)	47.94 (29.5)	50.48 (27.7)	47.07 (35.8)	48.96 (24.2)	0.98	0.402	0.001

Note: the following legends denote significant difference between paired Grand Slams: # Australian Open (AO) vs. Roland Garros (RG); ‡ AO vs. Wimbledon (W); † AO vs. US Open (US); ¶ RG vs. W; † RG vs. US; § W vs. US.

Table 3.2. Return, net, break point, efficiency and physical performance of female players in four Grand Slams.

Variable	Australian Open (n range: 520-998) M (SD)	Roland Garros (n range: 60-498) M (SD)	Wimbledon (n range: 272-498) M (SD)	US Open (n range: 284-744) M (SD)	F	p	partial eta squared
Return Performance							
Return points won of total returns (%)	43.03 (9.8) ∫	44.13 (9.6) ¶	40.91 (9.8) §	43.27 (10.1)	9.81	<0.001	0.011
Returning first serves won (%)	35.33 (11.5) #∫	38.08 (11.4) ¶†	33.22 (11.1) §	35.60 (11.8)	14.98	<0.001	0.016
Returning second serves won (%)	54.93 (11.9)	54.56 (12.1)	53.56 (12.7)	54.60 (12.5)	1.40	0.241	0.002
Return winners of total returns (%)	5.19 (5.5) #‡	2.01 (2.8) †	2.51 (2.42)	2.69 (2.7)	95.93	<0.001	0.107
Return unforced errors of total returns (%)	8.69 (7.7) #‡	3.28 (3.6)	3.41 (2.9)	4.07 (3.3)	171.23	<0.001	0.177
Net Performance							
Net points won (%)	66.58 (20.1)	64.40 (19.2)	64.60 (17.1)	65.59 (18.6)	1.88	0.131	0.002
Net points won in total points won (%)	10.86 (6.7) ∫‡	11.19 (6.3) ¶†	13.5 (7.9)	12.78 (8.1)	18.33	<0.001	0.022
Break points Performance							
Break points per return game	0.77 (0.4) ∫	0.79 (0.4) ¶	0.69 (0.4)	0.74 (0.4)	4.42	0.004	0.006
Break points won (%)	45.73 (23.3)	47.71 (23.8)	44.29 (24.4)	46.17 (24.5)	1.76	0.153	0.002
Break points saved (%)	54.05 (23.1)	51.95 (23.8)	55.54 (24.2)	53.83 (24.5)	1.94	0.121	0.002
Efficiency Performance							
Winner per unforced error ratio	0.86 (0.6) #∫	1.01 (0.6) ¶†	1.18 (0.8) §	0.90 (0.6)	32.04	<0.001	0.036
Dominance ratio	1.09 (0.5)	1.08 (0.5)	1.14 (1.0)	1.09 (0.5)	1.03	0.38	0.001
Physical Performance							
Mean meters run per match (m)	1338.71 (571.7)	1452.19 (600.2)	1289.28 (567.9) §	1423.18 (589.1)	3.21	0.022	0.008
Mean meters run per set (m)	583.76 (198.1)	618.65 (221.0)	558.31 (188.2) §	608.84 (196.8)	3.65	0.012	0.010
Mean meters run per point (m)	9.71 (2.7) ∫	10.16 (2.6) ¶	9.14 (2.4) §	10.18 (2.9)	7.53	<0.001	0.020

Note: the following legends denote significant difference between paired Grand Slams: # Australian Open (AO) vs. Roland Garros (RG); ∫ AO vs. Wimbledon (W); ‡ AO vs. US Open (US); ¶ RG vs. W; † RG vs. US; § W vs. US.

3.3.2 Results of correlation analysis and cluster analysis

The strength of the relationships between match performance variables and relative quality (RQ) in each Grand Slam are presented in Figure 3.2 by the correlation magnitude, its 99% confidence interval and the likelihood for the magnitude of the true effect. Of all 37 variables, the majority of them showed clear positive or negative relationship with RQ in all Grand Slams (28 variables in AO, 26 in RG, 27 in W and US). Additionally, within the four slams, some variables differed in their correlation magnitudes with RQ.

Starting from serve performance, aces of all serves showed a small to moderate relationships with RQ in all the slams. And for the first serve performance, the first serve won (%), first serve won in deuce (%) and advantage court (%) had a moderate correlation with RQ in all events. Only in Wimbledon, there is a moderate correlation between first serve points won in T of the deuce court (%) and RQ. Among all variables of second serve performance, the second serve won (%) is the only one that had a moderate association with RQ in four slams while a moderate relation was only shown in RG between second serve won in deuce court (%) and RQ. Only double faults of the second serves (%) in AO was negatively associated with RQ with a possible small magnitude.

Considering the return performance, the moderate to large positive association between the total return points won (%) and RQ in all slams was stronger than that of the returning first and second serve points returns (%).

The net points won (%) performance in all slams were positively correlated with the RQ with a small magnitude while only trivial connection was found between the net points won in total points won (%) and RQ.

Break points opportunity in return game (%) demonstrated a moderate positive correlation with RQ in all Grand Slams. A small positive relation was shown between break points won (%), break points saved (%) and RQ. The winning efficiency was also positively associated with RQ in that the dominance ratio in AO, RG and US had a moderate to large positive correlation with it, and in Wimbledon they were moderately associated; and winner per unforced error ratio showed a small association with RQ in all slams. The total distance covered, distance covered per set and per point had only trivial correlations with RQ with the ones of RG being unclear.

Figure 3.2. Relationships between match variables and the relative quality (RQ) for female players in four Grand Slams

Asterisks represent the likelihood for the magnitude of true correlation as follow: *possible, **likely, ***very likely, ****most likely. Asterisks located in the trivial zone denote for trivial correlation. Legends: Ace% =Aces of all serves%; SW/TW% =Service winners of all serves%; 1stIn% =First serve in%; 1stW% =First serve points won%; 1stDW% and 1stAdW% =First serve points won in deuce and advantage

courts%; 1stDWide%, 1stDT% and 1stDBody% =First serve points won in Wide, T and Body locations of the deuce court; 1stAdWide%, 1stAdT% and 1stAdBody% =First serve points won in Wide, T and Body locations of the advantage court; 2ndW% =second serve points won%; 2ndDW% and 2ndAdW% =Second serve points won in deuce and advantage courts%; 2ndDWide%, 2ndDT% and 2ndDBody% =second serve points won in Wide, T and Body locations of the deuce court; 2ndAdWide%, 2ndAdT% and 2ndAdBody% =second serve points won in Wide, T and Body locations of the advantage court; DF% =double faults of the second serves%; RW% =Total Return points won%; R1stW% and R2ndW% =Return first and second serves won%; Rwinner% and RUE% =Return winner and unforced error of total returns%; NetW% =Net points won%; NetW/TW% =New won of total points won%; Bpt/Regame =Break points per return game ratio; Bwon% =Break points won%; Bsaved% =Break points saved%; Winner/UE =Winner per unforced error ratio; DR =Dominance ratio; Dist. Total =Mean meters run per match; Dist/Set =Mean meters run per set; Dist/Pt =Mean meters run per point. Note: Abbreviations can be applied to the whole text.

The total of 148 correlation magnitudes (37 in each of four slams) were classified into two groups by the two-step cluster analysis. In the group of stronger correlations, there are 14 variables respectively from AO and RG that belong to this group, and 15 variables respectively from W and US. Therefore, variables with stronger correlations with RQ in each slam were treated separately to calculate the percentage evaluation score (%ES). Table 3 presents the descriptive analysis of these variables and the relative qualities for Top 10 ranked players and Serena Williams in four slams. Meantime, those variables were combined together with those from all players shown in Table 1 and 2, to establish the typical performance profiles of players.

3.3.3 Profiles of typical performance for player

Figure 3.3 presents the means of the %ES scores for all variables (AO: 14, RG: 14, W: 15, US: 15) that were selected. The %ES of all players averaged round 50%. And when it comes to those of Top-10 players and Serena Williams, a clearly visual inspection detected the performance differences among them. In all slams, Serena Williams showed a dominant performance in serving more aces of total serves, considering of the quality of her opponents, whereas the Top-10 players only had a similar performance to that of all players. Better performance in her efficiency of winner per unforced error ratio was demonstrated in all slams except for Roland Garros, where Serena Williams performed inconsistent or similar like other player. Worse or equal %ES to that of general player was shown in her return and second serve performance. Moreover, results also showed that Serena Williams had more stable performance in saving more break points in fast surfaces events (AO, W and US), but not in slow surface(RG).

Figure 3.3. Performance profiles of percentage of evaluation score (%ES) for: all female players, Top-10 players and Serena Williams in four Grand Slams

3.4 Discussion

This study aimed to assess the female players' performance in four Grand Slams and build the typical match performance profiles based on relative quality (RQ). Performance variables that were more associated with RQ were identified from all variables. This is the first study to identify the correlation between match performance variables and RQ and to assess the performance of female players in four Grand Slams via the regression based method (Butterworth et al., 2013; O'Donoghue & Cullinane, 2011).

Effects of game location were observed across a range of performance variables all Grand Slams that implies the baseline oriented match pattern and the importance of serve and return efficiency in female tennis games. The majority of variables showed a positive correlation with relative quality, but not physical characteristic, which was revealed to be unaffected by quality of opponents. Different magnitudes of correlations for each variable in distinct Grand Slam suggested the joint effect of game location and quality of opposition. The analyses of typical performance underlined the usefulness of applying percentage evaluation scores to interpret player's performance.

The results of the current study confirm that match location had an influence on the performance of female players and their strategies in Grand Slams. Previous research showed that the difference of court surfaces, court dimension, court speed, ball types, temperature and humidity might cause players' distinct adaptation to the courts, thus resulting in their different performance (Cross & Pollard, 2009; Gillet et al., 2009; Gómez, Lago, et al., 2013; Hizan et al., 2011; Morante & Brotherhood, 2007; O'Donoghue, 2002; O'Donoghue & Ingram, 2001). Nonetheless, more factors such as surface abrasion, venue altitude and the distance between court and audience should be further looked into to determine and quantify the effects.

3.4.1 Serve and Return Performance

The importance of serve and return performance in four Grand Slams have been widely addressed in previous studies (Hizan et al., 2014; O'Donoghue & Brown, 2009; O'Donoghue & Brown, 2008; Reid & Morris, 2013; Whiteside & Reid, 2016b) and been treated as the two most important performance variables in tennis match (Gillet et al., 2009; Hizan et al., 2011; Unierzyski & Wieczorek, 2004). Consistent with former studies (Gillet et al., 2009; Hizan et al., 2011), the results indicated that in all Grand Slams, female players had an optimal first

serve in percentage of round 60% and still used a “fast-weak” serving strategies. This possibly due to that they were more aggressive in the first serve by hitting more flat and faster balls, and conservative in the second by using slower and more topspin serves (Gillet et al., 2009; King & Baker, 1979). Accordingly, receivers won less points when returning the first serves but more points when returning the second serves in all slams, which showed that they were more prepared to return aggressively the second serves and take control of the points (Hizan et al., 2011). However, it should be highlighted that female players had higher percentage of winning first serve returns in Roland Garros than the other slams probably because of the slower surface speed that permits returner to have more time to react (Gillet et al., 2009; ITF, 2016). This finding can also be due to the fact that players achieved less aces and serve winner in Roland Garros. Therefore, the strategy of the fast first serve would not be as effective as in other slams. Although slow surface also determined that returning players in RG could not hit more return winners neither, as shown in the results, they managed to reduce the return unforced errors and counteract the dominance of the servers over their first serves (Hizan et al., 2014).

Nevertheless, results demonstrate that in Australian Open and US Open, players committed more double faults and return unforced errors while having more service and return winners than other slams. On the one hand, this would indicate firstly that in these two slams, female players had a preference of “strong-strong” serve strategy that although enable them to be more aggressive at serves and win more points, but also made them to be too aggressive to make more double faults; and on the other hand, that they carefully prepared their positions in return with good anticipation, especially in Australian Open to win attain more return winners, which corroborates the finding of O’Donoghue and Brown (2008). Whilst in terms of similarly high percentage of return unforced errors, two possible explanations would be: (i). Serve returners tried to dictate the point at the onset of the second serve to put serving opponents under pressure so that they made more mistakes in returning; (ii). Players’ concentration and reaction in return was impaired by the high temperature and humidity in the period of celebrating Australian and US Open because the two slams are held in hot environmental conditions (Hornery et al., 2007; Kovacs, 2006) and the effect of physical fatigue, hypo-hydration, hyperthermia (increased body temperature), unpredicted match length with repeated movements and psychological stress players perceive would jointly impair their performance (Fernandez-Fernandez, Sanz-Rivas, & Mendez-Villanueva, 2009; Filaire, Alix, Ferrand, & Verger, 2009; Hornery et al., 2007; Morante & Brotherhood, 2007).

Further research could investigate the relationship between hitting a return winner or unforced error and player's return position, physiological and psychological responses, serve speed and angle of the serve (Whiteside & Reid, 2016b).

Additionally, results of the spatial distribution of the serves showed that in professional female matches, players generally won more points by placing the first serves to the T and wide zones of service box, independent of the deuce or advantage side. In that way, they could not only increase their possibility of serving aces (Whiteside & Reid, 2016b), but also made the receivers feel more difficult to return (Hizan et al., 2015). However, trying first serves into the body zones might not seem to be a bad option in female matches (Hizan et al., 2015), especially in Roland Garros and US Open where players could still win around 58% of total serve points in body zones. Concerning the second serves, the finding of the study agrees with Hizan et al. (2015) in that female players favored serving to body zone of the receivers and won more points. Because this strategy could avoid the receiver to use their double-handed backhand to attack the serve. But the results also showed that placing second serves to the backhand of the opponents was also as effective as serving to forehand side (T zone).

3.4.2 Technical and Physical Performance

Results have demonstrated that female players had merely around 10% of total points won that were won in the net, which suggests that professional female players still remained conservative in approaching the net and preferred baseline strategy in all Grand Slams to compete (O'Donoghue & Ingram, 2001) although they had good effectiveness in net performance, winning 65% of all net points. In terms of break point performance, female players obtained similar break opportunity per game, break points conversion and break point saved percentage to the ones of male players (Ma et al., 2013; O'Donoghue, 2012). When players had better efficiency in Wimbledon, they may win more points rather than making unforced errors and had higher dominance ratio of achieving more points in return game rather than losing points in own service game. A potential explanation is that in Wimbledon, the faster surface speed shortened the rally number and increased the rally rate, as a result it was easier for players to strike more direct winners (O'Donoghue & Ingram, 2001; O'Donoghue & Brown, 2008).

Similar to the finding of Reid et al. (2016), female players covered approximately 1289m to 1452m per match, 558m to 618m per set and 9.14m to 9.71m per point, with Wimbledon

being the one where players ran less meters and Roland Garros being the one players ran more distance. But the difference of distance covered among all slams was not that obvious because it was more dependent on the game and score-line of the match (Reid et al., 2016). Limited former studies have already considered the movement patterns of male professional players in real match situations and inter-player relationship (Carvalho et al., 2013; Martinez-Gallego et al., 2013; Pereira et al., 2017), but few has been done with female players (Reid et al., 2016). Therefore, it is necessary that future researches focus on the movement characteristics of female players based on different match situations and opponent's qualities.

3.4.3 Relationship between match variables and relative quality and cluster Analysis

The findings of the current study reinforced the O'Donoghue's (2009) theory and provided more information in that the difference between players' strength (ranking) also influenced female player's performance on many aspects.

First of all, results were in agreement with O'Donoghue and Cullinane (2011) on that the first serve in percentage is not correlated with relative quality (RQ) in any of the four slams; and percentage of serve and return points won are positively associated with RQ in all events. Meanwhile, other serve and return performance variables were also encountered to have positive correlation with RQ in all slams, such as percentage of first serve points won in deuce and advantage courts, percentage of second serve points won in deuce and advantage courts, percentage of aces and return winners. However, the percentage of return unforced error and the percentage of second serve points won in Wide zone of the deuce court had trivial correlation with RQ. Moreover, there were variables like percentage of service winners and double faults, which were revealed to have clear association with RQ in certain slams but not in all of them.

Secondly, in terms of the efficiency and physical variables, there were no clear relationships between the percentage of net points won in total points won, the distance covered in match, set and point and RQ in all Grand Slams. While net point won (%), break point per return game, break point won (%), break point saved (%), winner per unforced error ratio and dominance ratio were all positively associated with RQ.

The quality of opposition is one of the contextual variables that have been widely considered in team sports to evaluate and explain the performance of interested teams against different

level of opponents (Gómez, Silva, Lorenzo, Kreivyte, & Sampaio, 2017; Gómez, Lago, et al., 2013; Gonçalves, Marcelino, Torres-Ronda, Torrents, & Sampaio, 2016; Liu et al., 2016) but there is a paucity of investigation in tennis over this topic (Goossens, Kempeneers, Koning, & Spieksma, 2015; O'Donoghue & Cullinane, 2011; Reid et al., 2010). Gosseens et al. (2015) grouped the tennis players into 7 categories according to their world ranking and tried to predict the winning probability of male and female players in Grand Slams; Reid et al. (2010) used the match statistics of male tennis player tournament from Association of Tennis (ATP) to decide what performance variables best indicates of player's top-100 ranking. However, only O'Donoghue (2009) considered in the interacting performance theory that the strength and the type of opponents had major effects on the performance outcome of tennis players and O'Donoghue and Cullinane (2011) found the meaningful correlations between match statistics and relative quality in Grand Slams male games, which gave more evidence on how quality of opposition could vary the performance of players.

The results of correlation not only suggested that ranking of professional female players could be used to evaluate the player performance (Reid, Morgan, Churchill, & Bane, 2014) and the comparative strength between two players is a good predictor of their match behaviors (O'Donoghue & Cullinane, 2011); but also indicated the influence of different court surfaces on how players planned their strategies against players of various levels (O'Donoghue & Ingram, 2001).

The results of two-step cluster analysis supported the last statement as well in that not all variables in the stronger correlation group were the same ones in every slam, and some of their strengths of correlation varied from one to another. For example, the percentage of second serve points won in Body zone of the deuce court was the variable that had stronger relationship in Roland Garros as it is a preferred zone for female players to place their second serves (Hizan et al., 2015). Also, it could be implied that in further research, the selection of performance variables should could be altered by the necessity of coach and players, according to different match and opponents (Butterworth et al., 2013).

3.4.4 Performance profiles using percentage evaluation score (%ES)

In the final part of the current study, we built the typical performance profiles for Top-10 ranked players and Serena Williams, using the %ES of variables from stronger correlation group. Former profiling techniques (O'Donoghue, 2005) was questioned for not considering

the opposition effects when establishing performance profiles and individual performance may sometimes be underestimated (Butterworth et al., 2013). Therefore, the study applied the approach of O'Donoghue and Cullinane (2011) to describe the players' performance in relation to their strength. The approach is a more suitable profiling technique in that it gives an evaluation score to the player's performance based on her difference in ranking with opponents. Another merit of the approach is that: with only one unit of measurement percentage, the %ES of variables that are originally different in unit, could be readily visualized in coaching process. This avoids the normalization of performance variables, which sometimes omits the individual performance. Besides, there are certain variables the coach can easily understand without being normalized in the practical scene (Butterworth et al., 2013; McGarry et al., 2013). Therefore, from a practical perspective, the study suggests that visualized %ES can be provided along with actual match statistics in coaching process to have better understanding of player's performance.

3.5 Conclusion

The current study provides preliminary assessment on the match-play performance of female tennis players by: (i) analyzing the performance of professional female tennis players in four Grand Slams using match statistics; (ii) measuring the relationships between performance variables and relative quality; (iii) building the typical performance profiles of tennis players with the consideration of opposition effects. Results provide noble insights into the performance of professional tennis players and could improve the coaching process. Generally, female game in all Grand Slams remained to be a contest over baseline, although players were shown to have good efficiency at net. Relative quality was found to have an impact on various match performance variables related to serve and return, break point, net point and efficiency except for physical performance. These findings evidenced the joint effect of courts and opponents on players' match behaviors. And finally, using the percentage evaluation score (%ES) to establish typical performance profiles for players would facilitate a more reasonable assessment of match performance and help to set practical training and match targets. The profiles of professional players could serve as a benchmark for juniors who aspire to achieve a professional success. Moreover, in light of player's long-term development, it would be optimal if coaches and performance analysts could use this profiling technique to keep a longitudinal monitoring of their players' technical, tactical and physical

performance, exploring important information about individual variation within different match contexts.

3.6 Limitation

While the study has advanced the understanding about female players' performance in different Grand Slams and provided more information on the typical performance of players using ranking-based approach, there are some limitations that should be addressed. Firstly, the match statistics of just two seasons from Roland Garros and Wimbledon were used and this would lead to a possible results bias when comparing the performance among different Grand Slams. Besides, many players were multiply used as independent sample due to the that their performances were conditioned by various contextual factors (Gómez, Lago, et al., 2013), but sampling the same players more than once across different Grand Slams is a possible limitation. Moreover, considering only world ranking is insufficient to account for the quality of opposition, and it is preferred that future study could look into others player characteristics such as anthropometric features, handedness, backhand style, experience, etc. (Cui et al., 2017). In addition, relating to the last point, the current study failed to take into account the influence of handedness on players' serving strategy, so that caution should be taken when generalizing the findings to left-handed players. Lastly, rather than considering the absolute performance values, it is suggested that the relationships between differences in performance of two competing players and their relative quality should be explored in the future study.

Chapter Four – Effects of Experience and Relative Quality in Tennis Match Performance during Four Grand Slams

Cui, Y., Gómez, M.-Á., Liu, H., Gonçalves, B., & 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), 1-19.

4.1 Introduction

Research in professional tennis has been focused on describing game strategies (Gillet et al., 2009; O'Donoghue & Ingram, 2001), playing patterns (Hizan et al., 2014), technical and tactical efficiency (Filipic et al., 2015; Ma et al., 2013) and players' movement profiles (Pereira et al., 2017). The obtained results allowed us to understand the match characteristics and offered a grounding to coaches to develop accurate training plans. Currently, as performance analysis is continuously smoothing the game-to-training transition and shaping the specificity of individualized training and game preparation (Reid et al., 2010; Reid et al., 2016), it is required that this process should be constantly updated with the latest game information (O'Donoghue, 2013).

The four Grand Slams serve as gold-standard database for this purpose, due to their important influence (Maquirriain et al., 2016), higher competitiveness level, and high ranking points and prize money (Cross & Pollard, 2009). With advanced technology like the Hawk-eye system that provides accurate ball and player motion-related data (Whiteside, Bane, & Reid, 2016), recent studies have already reported the difference in technical, tactical and physical match-play of Grand Slams between female and male players. More specifically, movement characteristics, groundstroke speed, net clearance of shots (Reid et al., 2016) and ideal location for ace-serving (Whiteside & Reid, 2016b) were analyzed. These studies offered us useful insight on how top players performed. But unfortunately, they have somehow a limited applicability, for centering only on the sex-based performance difference or the synthesized analysis of certain techniques.

Admittedly, the complexity of professional sports like tennis is a multi-faceted issue that requires a holistic and comprehensive approach to analyze player's performance (Sampaio et al., 2015). Assessing performance after classifying players based on their quality can be an effective procedure to better understand the performance determinants. It is evidenced that the

quality of opposition provides great variance on player's performance in team (Gómez, Lago, et al., 2013; Gonçalves et al., 2016) and individual sport (O'Donoghue & Cullinane, 2011). Knowing the strength and weakness of opponents helps coaches to modify tactical decisions about how players perform in the match and, ultimately, to maximize situations where they can benefit from competitive advantages. The player's ranking seems to be an appropriate representation of quality (Reid et al., 2010; Reid et al., 2014), as it is characterized by his career achievement (Reid et al., 2014) and competitive strength (Reid et al., 2010). The final league standings was used as a discriminator of successful and non-successful basketball teams (Sampaio et al., 2015); and high, medium and low level football teams (Liu et al., 2016). In tennis, the ranking was already used earlier (Whiteside et al., 2016), but it is argued that the magnitude of a tennis player's quality can be better determined by relative quality (O'Donoghue & Cullinane, 2011), which is the relative strength or weakness of a player when competing against his opponent, considering their entry-rankings in Grand Slam.

Additionally, professional experience (Ma et al., 2013) is also a meaningful discriminator of player's quality. The existing literature in team sports has already examined the performance difference between expert/experienced and novice players (Didierjean & Marméche, 2005) and suggested that this comparison could facilitate a more exhaustive comprehension of how player's development is influenced by training factors (Leite & Sampaio, 2010). It has been found that experienced players outperformed the novices in detecting, remembering and making proper judgements (Vaz, Leite, João, Gonçalves, & Sampaio, 2012). Experienced basketball players were reported to have better passing performance when they were moderately or extremely fatigued (Lyons, Al-Nakeeb, & Nevill, 2006); experienced rugby players showed better performance than novices in game-related indicators (Vaz et al., 2012); and experienced football players were found to have cognitive advantage over inexperienced players for being faster to recognize and process the emergent features of play pattern so that they can anticipate better opponent's actions and make quicker reactions (Williams, 2000). In tennis, the only previous study that attempted to compare the visual search strategies of "expert and novice players" used players' ranking as the discriminator (Murray & Hunfalvay, 2017), but little is known about how tennis players' competitive performance is influenced by experience.

Therefore, the study was aimed to evaluate the match performance of different male player categories in four Grand Slams, based on their playing experience and relative quality. It was

hypothesized that players with longer playing experience and higher relative quality would outperform other players in all match-related statistics.

4.2 Method

4.2.1 Subjects

Data of 1188 players performing in 594 matches were collected from four 2015-2017 Grand Slams men's singles event, with two players' performances analysed for each match (Australian Open: range of rankings = 1 to 692; Roland Garros: 1 to 296; Wimbledon: 1 to 524; US Open: 1 to 581). Match statistics were included for players who played at least one match on the courts where both Doppler radar system (IBM: Armonk, NY, USA) and the Hawk-Eye camera system (Hawk-Eye Innovations, Southampton, UK) were installed. Matches with an early retirement of players were excluded. Data were gathered in each official Grand Slams website (Australian Open: www.ausopen.com; Roland Garros: www.rolandgarros.com; Wimbledon: www.wimbledon.com; US Open: www.usopen.org). The study was conducted under the approval of the local University Ethics Committee and all procedures were undertaken to accomplish all international standards.

4.2.2 Variables and Operational Definitions

The variables of present study can be divided into three categories: playing surfaces, player quality and match performance. The first one is intuitively explained by different Grand Slams because of their distinct court surfaces (Australian Open: acrylic hard surface; Roland Garros: clay surface; Wimbledon: grass surface; US Open: DecoTurf hard surface). In terms of player quality, relative quality and professional experience were determined to depict player characteristics. The former is considered as the relative strength or weakness of one player against his opponent using their entry-rankings in each Grand Slam (O'Donoghue & Cullinane, 2011); while the latter denotes the number of years that a player has competed in professional tournaments since he first achieved a professional ranking (Ma et al., 2013). There were 38 match performance variables that are presented in Table 4.1 along with their operational definitions.

Table 4.1. Description of variables and operational definitions

Variable	Operational Definition
<i>Serve Performance</i>	
Ace%	% of aces in total number of serves
Ace in Deuce%	% of aces in total serves to the deuce service box
Ace in Advantage%	% of aces in total serves to the advantage service box
Service winners%	% of service winners in total services
1 st serve in%	% of first serves that were in play
1 st serve points won%	% of points won when first serves were in play
1 st serve won in Deuce%	% of points won when first serves were served into the deuce service box
1 st serve won in Advantage%	% of points won when first serves were served into the advantage service box
2 nd serve points won%	% of points won when second serves were in play
2 nd serve won in Deuce%	% of points won when second serves were served into the deuce service box
2 nd serve won in advantage%	% of points won when second serves were served into the advantage service box
Double faults%	% of double faults in total number of second serves
Fastest serve speed (km/h)	The peak serve speed within the match
1 st Average speed Deuce (km/h)	Mean first serve speed when serving to the deuce service box
1 st Average speed Advantage (km/h)	Mean first serve speed when serving to the advantage service box
2 nd Average speed Deuce (km/h)	Mean second serve speed when serving to the deuce service box
2 nd Average speed Advantage (km/h)	Mean second serve speed when serving to the advantage service box
<i>Return Performance</i>	
Return points won%	% of points won when returning serves
Return winner%	% of return winners in total serve returns
Return unforced error%	% of unforced errors in total serve returns
Return 1 st serve won%	% of points won when returning first serves
Return 2 nd serve won%	% of points won when returning second serves
<i>Net Performance</i>	
Net points won%	% of points won when player approached the net
Net points won of total points won%	% of net points won of total points won
<i>Rally Performance</i>	
Winner in rally%	% of winners won in rally (ground strokes, approach, net, passing and lob shots)
Forced error in rally%	% of forced errors won in rally
Unforced error in rally%	% of unforced errors won in rally
<i>Break Performance</i>	
Break points per return game	Number of break points had in every opponent's service game
Break points won%	% of break points converted in opponent's service game

Break points saved%	% of break points served in own service game
<i>Efficiency Performance</i>	
Winners	Total number of winner shots in the match
Unforced errors	Total number of unforced errors in the match
Winner per unforced error ratio	Number of winners in every unforced error
Winner of total points won%	% of winner shots in total number of points won
Dominance ratio	% of points won in opponent's serves / % of points lost in own serves
<i>Physical Performance</i>	
Mean meters run per match (m)	Mean distance covered in a match by the player
Mean meters run per set (m)	Mean distance covered in each set by the player
Mean meters run per point (m)	Mean distance covered in each point by the player

4.2.3 Data Reliability

On Grand Slam courts equipped with the Hawk-Eye system, eight to ten high-speed video cameras are located around the arena and they are synchronized to accurately track the three dimensional movement of ball and player with a reported measurement mean error between 2 and 5 mm (Iwan, Colin, & Neil, 2005). The system can also automatically determine the stroke types (forehand or backhand) with validated high accuracy (Bal & Dureja, 2012; Whiteside & Reid, 2016b).

To ensure the data reliability, four broadcast matches that were randomly selected in each of the four Grand Slams, and two experienced performance analysts in tennis observed those matches and collected the non-tracking data. Afterwards, comparisons were made between inter-observers and between observers and the sequential match statistics (aces, aces in deuce service box, aces in advantage service box, double faults, first serves in, first serves won, first serves won in deuce, first serves won in advantage, second serves won, second serves won in deuce, second serves won in advantage, double faults, return points won, return first serves won, return second serves won, breakpoints per return game, breakpoints won, breakpoints saved) provided by IBM SlamTracker of official websites. The minimum Cohen's *kappa* value for all variables exceeded 0.90, demonstrating high inter-rater reliability (O'Donoghue, 2010). While for the continuous variables that are not included in the sequential data, intra-class coefficients (ICC) and standardized typical errors were calculated for inter-operators, intra-operators and operators vs. official data. Results of ICC ranged from 0.96 to 1 and standardized typical errors varied from 0.02 to 0.11, showing very high levels of reliability for

the following variables: service winner%, return winner%, return unforced error%, net points won%, net points won of total points won%, winner in rally%, forced error in rally%, unforced error in rally%, winners, unforced errors, winner per unforced error ratio, winner of total points won% and dominance ratio. Afterwards, comparisons were made between inter-observers and between observers and the sequential match statistics (aces, aces in deuce service box, aces in advantage service box, double faults, first serves in, first serves won, first serves won in deuce, first serves won in advantage, second serves won, second serves won in deuce, second serves won in advantage, double faults, return points won, return first serves won, return second serves won, breakpoints per return game, breakpoints won, breakpoints saved) provided by IBM SlamTracker of official websites. The minimum Cohen's *kappa* value for all variables exceeded 0.90, demonstrating high inter-rater reliability (O'Donoghue, 2010). While for the continuous variables that are not included in the sequential data, intra-class coefficients (ICC) and standardized typical errors were calculated for inter-operators, intra-operators and operators vs. official data. Results of ICC ranged from 0.96 to 1 and standardized typical errors varied from 0.02 to 0.11, showing very high levels of reliability for the following variables: service winner%, return winner%, return unforced error%, net points won%, net points won of total points won%, winner in rally%, forced error in rally%, unforced error in rally%, winners, unforced errors, winner per unforced error ratio, winner of total points won% and dominance ratio.

4.2.4 Statistical Analysis

Using the 52-week world ranking, the relative quality (RQ) for each player were firstly calculated according to the methods of O'Donoghue and Cullinane (2011), via the following equation:

$$\text{Equation 1. } RQ = (8 - \log_2(Rank_X)) - (8 - \log_2(Rank_Y))$$

where $Rank_X$ is the ranking of player X and $Rank_Y$ is the ranking of his opponent in determined match.

To classify the similar patterns of player type in all players from four Grand Slams, a *k*-means cluster analysis was performed using player's experience and RQ as input variables. Also, the Cubic Clustering Criterion (CCC) was applied together with Monte Carlo simulations to avoid subjective criteria in deciding cluster numbers. Based on the minimization of the

within-cluster sum of squares, the optimal number of clusters can be derived from this approach if the highest CCC value is given to certain number of clusters as the number of clusters is incremented by 1. Subsequently, a discriminant analysis was performed to determine which variables best predicts the player categories in each of the four Grand Slams. Structure coefficients greater than $|0.30|$ were interpreted as meaningful contributors for discriminant functions that differentiate player groups. Validation of discriminant models was conducted using the leave-one-out method of cross-validation, which takes subsets of data for training and testing and is necessary for understanding the usefulness of discriminant functions when classifying new data. The analyses were performed using JMP statistics software package (SAS Institute, Cary, NC, USA) and IBM SPSS software (Armonk, NY: IBM Corp.). Afterwards, a Matlab® (MathWorks, Inc., Massachusetts, USA) routine (Gal, 2003) was employed to generate the best fit for the ellipses of each player category in distinct Grand Slams (See Figure 4.2).

4.3 Results

Four clusters of players were identified by the cluster analysis based on the cubic clustering criterion (CCC). The maximal CCC value was obtained with four clusters ($CCC = -9.37$), indicating the optimal number of clusters. Figure 4.1 shows the labels and characteristics of four different player categories. There were 337 players with longest playing experience (mean $\pm$ SD: 13.7 ± 2.0) and highest RQ (3.1 ± 1.6) that were labeled as “High-experienced Higher RQ Players”; 338 players with longer playing experience (13.0 ± 2.4) and lower RQ (-2.6 ± 1.6) were labeled as “High-experienced Lower RQ Players”; 267 players with medium-long playing experience (8.4 ± 2.0) and higher RQ (1.8 ± 1.3) were labeled as “Experienced Higher RQ Players” and 246 players with shorter playing experience (5.9 ± 2.6) and lower RQ (-2.6 ± 1.6) were labeled as “Novice Lower RQ Players”.

Figure 4.1. Characteristics of four player categories classified by cluster analysis with playing experience (years) and relative quality (RQ) being the inputs

Table 4.2 depicts the means and standard deviations of the performance variables and the structure coefficients from discriminant analysis for four player clusters in Australian Open (AO) and Roland Garros (RG), and Table 4.3 presents the ones in Wimbledon (W) and US Open (US). Three functions were obtained in the discriminant analysis for each of Grand Slams, but only the first two were revealed to be statistically significant ($P < 0.01$). Meantime, the two functions could together explain 94.0% of the total variance in AO, 88.8% in RG, 92.4% in W, and 86.5% in US, with respective canonical correlation of 0.73 and 0.40 in AO ($\Lambda = 0.39$ and 0.78), 0.83 and 0.54 in RG ($\Lambda = 0.17$ and 0.53), 0.76 and 0.50 in W ($\Lambda = 0.28$ and 0.66) and 0.71 and 0.54 in US ($\Lambda = 0.29$ and 0.58). The obtained functions correctly classified 61.4% of the cases in AO, 72.7% of the cases in RG, 63.6% of the cases in W and 67.2% of the cases in US.

Figure 4.2 presents: i) four player categories and group centroids in Grand Slams; ii) best-fit ellipses with Least-Squares Criterion for every player category; iii) examples of players who played the quarterfinals of each tournament. They are all defined by the first and second discriminant functions. It is shown that HHRQ players had better performance in variables related to function 1. In contrast, EHRQ players exhibited overall better performance in variables relevant to function 2. HLRQ players had worse performance in variables related to both functions. NLRQ players showed a better performance than HHRQ and HLRQ players in variables related to function two.

Table 4.2. Means (Standard deviations) and structure coefficients (SC) of technical and physical performance for four player clusters in Australian Open and Roland Garros

Variable	Australian Open				Roland Garros							
Cluster	HHRQ	HLRQ	EHRQ	NLRQ	Function		HHRQ	HLRQ	EHRQ	NLRQ	Function	
	(n = 133)	(n = 146)	(n = 114)	(n = 99)	1	2	(n = 57)	(n = 49)	(n = 21)	(n = 27)	1	2
Ace%	11.0 (6.4)	7.7 (5.5)	13 (8.0)	8.1 (5.0)	0.28	0.37*	5.4 (3.5)	4.2 (3.1)	6.7 (4.1)	3.7 (3.3)	-0.14	0.29
Ace in Deuce%	11.8 (7.3)	9.2 (7.4)	13.8 (9.1)	9.6 (6.7)	0.19	0.28	6.4 (4.8)	4.5 (4.2)	7.0 (4.8)	4.1 (3.6)	-0.16	0.17
Ace in Advantage%	10.1 (7.7)	8.2 (6.5)	11.7 (8.6)	7.9 (6.2)	0.17	0.20	4.6 (3.8)	4.4 (3.8)	7.0 (5.2)	3.6 (4.5)	-0.06	0.32*
Serve winner%	2.0 (1.6)	1.3 (1.2)	2.1 (1.4)	1.5 (1.5)	0.20	0.23	0.7 (0.8)	0.6 (0.9)	0.9 (1.3)	0.9 (1.0)	0.02	0.17
1 st serve in%	62.0 (6.9)	61.4 (6.3)	63.4 (6.5)	61.1 (6.4)	0.08	0.18	63.8 (7.9)	60.6 (5.8)	60.5 (7.7)	61.6 (8.7)	-0.11	-0.14
1 st serve points won%	78.2 (8)	68.5 (8.6)	77.2 (7.1)	68.6 (7.6)	0.56*	0.13	74.3 (8.1)	65.1 (8.6)	74.3 (6.5)	60.6 (7.3)	-0.47*	0.25
1 st serve won in Deuce%	77.5 (10.5)	71.0 (10.4)	76.5 (9.4)	71.7 (9.4)	0.28	0.07	75.8 (10.3)	65.5 (10.5)	75.2 (7.4)	59.5 (8.0)	-0.44*	0.19
1 st serve won in Advantage%	75.0 (11.1)	69.9 (11.4)	75.1 (9.7)	68.4 (11.5)	0.25	0.06	73.9 (11.1)	65.9 (11.9)	73.0 (7.4)	61.8 (11.4)	-0.29	0.11
2 nd serve points won%	57.3 (10.1)	47.3 (9)	54.6 (10.4)	46.9 (9.9)	0.45*	-0.04	59.6 (8.1)	45.2 (9.7)	50.4 (8.3)	44.8 (7.1)	-0.52*	-0.20
2 nd serve won in Deuce%	61.8 (15.1)	54.0 (13.3)	57.4 (12.8)	52.5 (12.9)	0.24	-0.18	64.0 (14.3)	50.7 (14.4)	57.1 (12.9)	49.1 (13.2)	-0.31*	-0.05
2 nd serve won in Advantage%	60.6 (13.5)	53.5 (14.9)	60.3 (13.1)	54.2 (15.1)	0.22	0.09	66.1 (14.8)	49.2 (14.6)	54.9 (10.5)	50.5 (15.6)	-0.34*	-0.15
Double fault%	8.8 (5.2)	9.1 (5.4)	9.4 (5.2)	10.1 (5.3)	-0.05	0.14	5.9 (4.3)	8.4 (5.9)	10.1 (5.2)	8.6 (6.0)	0.14	0.31
Fastest Serve Speed (km/h)	206.2 (10)	202.9 (9.5)	212.5 (12)	207.6 (11.5)	0.15	0.66*	202.2 (10.0)	203.4 (10.2)	207.9 (8.0)	202.8 (11.1)	0.01	0.27
1 st Average Speed Deuce (km/h)	185.0 (8.6)	183.1 (9.4)	188.6 (10.5)	185.6 (10.6)	0.10	0.41*	185.2 (9.1)	185.9 (8.4)	186.4 (7.5)	182.9 (10.4)	-0.03	0.05
1 st Average Speed Ad (km/h)	186.3 (10.3)	183.6 (9.8)	189.3 (11.4)	186.0 (11.8)	0.12	0.34*	181.2 (8.6)	181.8 (9.2)	182.9 (9.4)	180.1 (11.7)	-0.01	0.08
2 nd Average Speed Deuce (km/h)	153.2 (9.1)	152.4 (9.6)	157.0 (13.3)	156.5 (11.3)	0.02	0.41*	150.3 (7.1)	153.3 (10.6)	151.8 (9.9)	153.9 (11.5)	0.11	0.02
2 nd Average Speed Ad (km/h)	149.6 (10.3)	148.5 (9.9)	151.5 (14.0)	151.5 (11.0)	0.02	0.23	144.8 (7.6)	145.5 (11.3)	143.6 (10)	148.1 (11.6)	0.07	-0.07
Return points won%	41.3 (6.8)	31.0 (7.3)	38.0 (6.5)	29.4 (6.4)	0.68*	-0.18	44.5 (7.2)	32.6 (7)	41.9 (6.2)	31.5 (6.9)	-0.56*	0.18
Return winner%	2.3 (1.9)	1.9 (1.9)	2.4 (1.8)	1.6 (1.3)	0.15	0.05	1.8 (1.6)	1.4 (1.3)	1.8 (1.7)	0.8 (1.1)	-0.15	0.08
Return unforced error%	3.3 (2.3)	3.3 (2.7)	4.1 (2.6)	3.8 (3)	0.01	0.33*	3.2 (2.3)	3.4 (2.1)	3.3 (1.9)	3.4 (2.4)	0.02	0.00
Return 1 st serve won%	33.1 (8)	22.8 (7.6)	29.8 (7.9)	21.1 (7.4)	0.60*	-0.16	37.1 (8.9)	25.3 (6.6)	34.9 (7.9)	25.6 (8.2)	-0.46*	0.19

Return 2 nd serve won%	54.2 (9.9)	44.9 (10.2)	51.0 (8.6)	42.4 (10.0)	0.45*	-0.16	55.5 (9.0)	43.4 (9.6)	52.9 (8.2)	41.5 (8.1)	-0.46*	0.15
Net point won%	71.9 (12.7)	64.2 (12.9)	68.4 (10.6)	62.4 (11.7)	0.28	-0.15	69.5 (11.6)	60.3 (13.9)	66.9 (10.4)	59.9 (11.3)	-0.24	0.06
Net success of total points won%	13.8 (6.1)	16.8 (8.3)	15.1 (6.9)	15.1 (7.7)	-0.12	-0.01	14.2 (6.5)	16.9 (9.5)	13.2 (4.8)	14.8 (5.0)	0.08	-0.16
Winner in rally%	34.7 (11.6)	29.4 (10.4)	33.0 (10.4)	27.0 (8.3)	0.26	-0.10	42.6 (9.5)	31.0 (7.9)	39.2 (6.1)	28.8 (8.2)	-0.48*	0.09
Forced error in rally%	31.3 (19.5)	32.9 (17.7)	31.4 (17.7)	35.0 (17.3)	-0.06	0.02	21.6 (7.2)	28.4 (7.3)	24.7 (7.1)	28.7 (8.6)	0.29	0.02
Unforced error in rally% [#]	34.1 (14.0)	37.7 (15.3)	35.6 (14.5)	38.1 (15.5)	-0.10	0.04	35.9 (9.3)	40.6 (8.4)	36.1 (6.8)	42.5 (9.2)	0.21	-0.10
Break points per return game	0.7 (0.3)	0.4 (0.3)	0.6 (0.3)	0.3 (0.2)	0.55*	-0.10	0.8 (0.3)	0.4 (0.2)	0.8 (0.2)	0.4 (0.3)	-0.53*	0.22
Break points won%	44.3 (17.1)	34.9 (25.2)	40.4 (17.6)	31.4 (30.0)	0.20	-0.11	46.0 (17.7)	37.5 (24.0)	42.3 (14.3)	27.4 (17.1)	-0.22	-0.01
Break points saved%	54.7 (33.3)	53.2 (20.2)	53.0 (29.2)	52.6 (16.8)	0.02	-0.05	68.3 (22.7)	56.2 (15.4)	60.4 (19.7)	51.6 (18.8)	-0.22	-0.08
Winners	42.8 (14.8)	34.6 (14.2)	45.2 (14.6)	34.8 (17.1)	0.28	0.24	37.8 (12.5)	31.4 (12.7)	41.6 (11.9)	27.6 (12.0)	-0.22	0.27
Unforced errors	34.0 (16.4)	38.9 (15.2)	38.1 (16.8)	40.9 (15.6)	-0.13	0.18	30.0 (12.0)	36.7 (11.5)	36 (14.2)	35.9 (11.1)	0.15	0.14
Winner per unforced error ratio	1.4 (0.7)	1.0 (0.6)	1.4 (0.8)	0.9 (0.5)	0.34*	0.05	1.6 (1.4)	0.9 (0.4)	1.3 (0.4)	0.8 (0.4)	-0.25	-0.02
Winner of total points won%	36.3 (9.3)	32.9 (8.3)	37.1 (9.8)	32.8 (9.0)	0.19	0.14	33.8 (7.2)	31.7 (8.1)	34.8 (7.5)	32.4 (8.9)	-0.08	0.13
Dominance ratio	1.5 (0.5)	0.8 (0.3)	1.3 (0.4)	0.8 (0.2)	0.78*	-0.06	1.5 (0.6)	0.8 (0.3)	1.2 (0.3)	0.7 (0.2)	-0.55*	0.01
Total distance covered (m)	2291.7 (901.1)	2361.2 (913.0)	2315.1 (784.8)	2190.5 (777.0)		0.01	2393.1 (893.9)	2567.4 (944.1)	2559.1 (806.4)	2219.8 (655.7)	0.00	0.05
Distance covered per set (m)	637.4 (182.5)	644.7 (199.0)	615.8 (166.7)	593.9 (136.8)		0.02	683.1 (186.6)	684.0 (180.2)	675.3 (165.0)	674.6 (161.9)	-0.01	-0.02
Distance covered per point (m)	10.5 (2.3)	10.4 (2.5)	9.9 (2.3)	9.7 (1.9)	0.03	-0.30*	11.8 (2.7)	11.6 (2.8)	11.2 (2.7)	11.6 (2.0)	-0.01	-0.11

Legend: [#] denotes the variable that were not included in the discriminant analysis; *Structure coefficient $\geq |0.30|$

HHRQ=High-experienced Higher RQ players; HLRQ=High-experienced Lower RQ players; EHRQ=Experienced Higher RQ players; NLRQ=Novice Lower RQ players

Table 4.3. Means (Standard deviations) and structure coefficients (SC) of technical and physical performance for four player clusters in Wimbledon and US Open

Grand Slam		Wimbledon					US Open					
Cluster	HHRQ	HLRQ	EHRQ	NLRQ	Function		HHRQ	HLRQ	EHRQ	NLRQ	Function	
Variable	(n = 74)	(n = 71)	(n = 70)	(n = 65)	1	2	(n = 73)	(n = 72)	(n = 62)	(n = 55)	1	2
Ace%	11.9 (6.7)	7.4 (5.5)	13.7 (8.5)	9.2 (5.7)	0.27	0.30*	9.2 (5.4)	5.8 (4.4)	11.7 (7.5)	6.8 (4.1)	0.32	-0.41*
Ace in Deuce%	13.0 (8.0)	7.6 (5.8)	16.0 (10.5)	10.4 (6.9)	0.28	0.38*	6.8 (5.0)	9.9 (8.0)	8.5 (6.9)	10.7 (6.7)	-0.22	-0.07
Ace in Advantage%	11.7 (8.0)	7.9 (7.1)	12.4 (8.5)	8.9 (6.6)	0.20	0.14	6.7 (4.8)	10.3 (8.6)	9.4 (6.8)	8.6 (6.3)	-0.15	-0.12
Serve winner%	1.2 (1.2)	1.1 (1.0)	1.6 (1.2)	1.3 (1.2)	0.06	0.22	2.4 (2.5)	1.5 (1.5)	2.7 (3.0)	2.1 (2.3)	0.16	-0.16
1 st serve in%	63.4 (5.9)	62.0 (6.2)	64.7 (6.6)	62.0 (6.3)	0.13	0.16	59.3 (7.3)	57.1 (7.1)	58.1 (5.9)	57.3 (6.7)	0.13	0.05
1 st serve points won%	80.4 (7.6)	69.6 (9.0)	79.5 (7.4)	70.9 (7.8)	0.53*	0.12	77.5 (7.7)	68.1 (8.6)	76.4 (7.5)	68.7 (8.0)	0.54*	-0.14
1 st serve won in Deuce%	80.8 (8.7)	70.7 (9.8)	80.1 (8.3)	72.3 (9.4)	0.43*	0.12	67.0 (11.8)	76.8 (9.2)	73.0 (9.5)	79.0 (9.5)	-0.44*	-0.20
1 st serve won in Advantage%	80.0 (10.0)	68.4 (12.3)	78.7 (9.6)	69.4 (9.2)	0.44*	0.08	68.2 (10.6)	76.0 (9.7)	70.6 (10.1)	74.9 (11.4)	-0.31*	0.01
2 nd serve points won%	57.0 (9.8)	48.1 (9.7)	56.1 (9.0)	48.9 (10.6)	0.36*	0.07	56.1 (10.5)	47.0 (8.0)	54.6 (10.5)	47.8 (10.6)	0.41*	-0.08
2 nd serve won in Deuce%	62.1 (13.2)	53.6 (15.3)	64.0 (14.3)	55.8 (15.8)	0.24	0.18	53.8 (13.1)	62.6 (14.5)	54.4 (12.3)	60.3 (12.6)	-0.28	0.11
2 nd serve won in Advantage%	61.3 (13.7)	55.1 (15.3)	60.2 (16.6)	56.6 (14.5)	0.14	0.02	54.7 (15.7)	65.4 (15.3)	54.7 (12.6)	61.3 (12.6)	-0.30*	0.16
Double fault%	6.8 (5.1)	9.8 (5.6)	9.5 (5.0)	10.2 (5.8)	-0.17	0.26	9.3 (5.7)	13.1 (5.8)	10.4 (6.1)	10.6 (6.0)	-0.20	0.04
Fastest Serve Speed (km/h)	206.1 (7.8)	204.2 (9.5)	211.5 (9.8)	207.6 (11.6)	0.10	0.44*	206.2 (9.5)	204.2 (8.8)	212.0 (11.6)	207.6 (11.3)	0.10	-0.41*
1 st Average Speed Deuce (km/h)	186.9 (6.8)	186.1 (9.2)	189.5 (9.9)	187.1 (10.3)	0.06	0.21	186.0 (9.5)	183.9 (9.8)	183.0 (11.1)	183.9 (9.1)	0.06	0.15
1 st Average Speed Ad (km/h)	190.0 (7.3)	186.6 (10.3)	192.1 (11.2)	188.3 (11.2)	0.14	0.20	185.2 (10.0)	184.1 (11.5)	181.3 (11.4)	182.8 (10.2)	0.03	0.21
2 nd Average Speed Deuce (km/h)	157.9 (10.8)	158.5 (11.9)	161.9 (13.5)	160.2 (11.4)	0.00	0.23	153.0 (11.4)	152.3 (10.9)	151.3 (11.8)	151.3 (9.8)	0.03	0.09
2 nd Average Speed Ad (km/h)	156.3 (10.1)	156.6 (11.5)	157.8 (14.5)	159.4 (13.4)	-0.04	0.11	148.9 (12.4)	148.0 (11.0)	147.9 (12.1)	146.9 (10.1)	0.05	0.04
Return points won%	39.5 (7.2)	29.3 (6.9)	35.9 (7.3)	27.7 (6.4)	0.59*	-0.18	42.3 (7.0)	31.3 (6.8)	38.4 (6.4)	31.7 (7.5)	0.68*	0.04
Return winner%	1.7 (1.5)	1.6 (1.5)	2.5 (1.9)	1.3 (1.1)	0.17	0.33*	1.6 (1.6)	1.7 (1.5)	1.9 (1.8)	2.3 (1.6)	-0.07	-0.14
Return unforced error%	1.4 (1.8)	1.3 (1.7)	1.4 (1.1)	1.3 (1.4)	0.04	0.02	3.1 (2.9)	2.9 (2.3)	3.3 (2.6)	3.2 (2.9)	0.02	-0.08
Return 1 st serve won%	31.1 (8.6)	20.9 (8.1)	27.0 (8.2)	19.3 (7.4)	0.50*	-0.20	33.6 (8.0)	22.9 (7.7)	29.2 (7.8)	22.0 (7.3)	0.62*	0.10
Return 2 nd serve won%	53.5 (10.6)	44.4 (10.2)	48.9 (8.9)	42.1 (8.7)	0.38*	-0.23	53.5 (9.3)	43.3 (9.8)	51.8 (8.9)	45.0 (11.0)	0.45*	-0.10
Net point won%	69.5 (12.4)	62.6 (11.9)	67.7 (9.6)	60.2 (11.8)	0.27	-0.06	69.4 (11.3)	62.5 (12.5)	70.1 (9.1)	60.6 (10.7)	0.35*	-0.14
Net success of total points won%	18 (6.8)	19.8 (9.5)	16.6 (5.9)	15.9 (8.2)	-0.03	-0.23	15.2 (6.0)	17.1 (7.8)	16.0 (5.7)	16.2 (7.8)	-0.09	-0.01
Winner in rally%	41.4 (9.6)	32.9 (9.2)	40.4 (9.0)	32.1 (8.6)	0.40*	0.03	37.8 (8.7)	30.0 (9.2)	35.8 (9.6)	32.2 (8.7)	0.34*	-0.05
Forced error in rally%	30.2 (7.9)	35.1 (8.7)	32.4 (7.7)	36.5 (9.0)	-0.24	0.13	24.8 (8.0)	26.5 (8.0)	22.6 (7.7)	24.4 (8.0)	-0.08	0.22

Unforced error in rally% [#]	28.4 (9.0)	32.1 (8.7)	27.2 (8.7)	31.4 (9.3)	-0.18	-0.15	37.4 (8.5)	43.5 (10.0)	41.5 (10.5)	43.5 (9.4)	-0.25	-0.14
Break points per return game	0.6 (0.3)	0.3 (0.3)	0.5 (0.3)	0.3 (0.2)	0.47*	-0.28	0.7 (0.3)	0.4 (0.2)	0.6 (0.2)	0.4 (0.2)	0.60*	0.02
Break points won%	42.1 (16.1)	29.5 (25.2)	42.6 (22.9)	28.3 (28.1)	0.25	0.07	46.7 (18.5)	30.7 (25.1)	38.8 (16.1)	36.5 (27.5)	0.25	0.06
Break points saved%	74.3 (26.2)	55.7 (18.0)	67.1 (27.1)	58.7 (19.5)	0.27	-0.08	66.6 (27.8)	56.1 (16.6)	66.0 (24.1)	58.0 (19.1)	0.21	-0.07
Winners	41.7 (15.6)	34.8 (18.4)	45.3 (14.1)	36.4 (18.4)	0.19	0.21	37.1 (12.9)	30.1 (13.2)	43.1 (13.9)	35.5 (14.6)	0.22	-0.39*
Unforced errors	22.4 (10.5)	27.9 (12.1)	23.1 (9.3)	27 (11.0)	-0.19	-0.04	30.2 (13.0)	37.3 (13.5)	38 (16.6)	38.8 (12.5)	-0.19	-0.26
Winner per unforced error ratio	2.2 (1.2)	1.3 (0.6)	2.2 (1.2)	1.5 (0.9)	0.35*	0.16	1.4 (0.6)	0.9 (0.4)	1.4 (0.9)	1.0 (0.5)	0.35*	-0.15
Winner of total points won%	35.1 (7.7)	32.2 (9.4)	36.8 (8.3)	33 (7.7)	0.16	0.19	32.8 (9)	29.7 (7.2)	35.7 (9.6)	33.4 (8.0)	0.13	-0.31*
Dominance ratio	1.5 (0.5)	0.8 (0.3)	1.3 (0.4)	0.8 (0.3)	0.71*	-0.05	1.5 (0.5)	0.8 (0.3)	1.2 (0.4)	0.8 (0.3)	0.73*	0.02
Total distance covered (m)	2108.8 (790.7)	2133.0 (786.1)	2011.8 (602.5)	1999.6 (623.2)	0.00	-0.12	2253.9 (812.8)	2240.3 (851.4)	2227.0 (753.4)	2233.5 (771.0)	0.01	0.02
Distance covered per set (m)	588.1 (150.8)	576.1 (155.6)	541.1 (131.4)	551.6 (122.8)	0.02	-0.23	643.1 (171.2)	609.9 (172.4)	579.6 (131.1)	610.8 (149.4)	0.05	0.20
Distance covered per point (m)	9.7 (2.2)	9.6 (2.5)	8.7 (2.1)	8.8 (1.8)	0.02	-0.34*	10.8 (2.3)	10.3 (2.4)	9.4 (1.9)	9.8 (2.0)	0.06	0.35*

Legend: [#] denotes the variable that were not included in the discriminant analysis; *Structure coefficient $\geq |0.30|$

HHRQ=High-experienced Higher RQ players; HLRQ=High-experienced Lower RQ players; EHRQ=Experienced Higher RQ players; NLRQ=Novice Lower RQ players

Figure 4.2. The distribution of four player clusters (with cluster centroids, best-fit ellipses and quarterfinalists) defined by two discriminant functions

4.4 Discussion

The aim of this study was to explore the match performance of different male tennis player categories and to compare their profiles in different Grand Slams. The findings could provide useful information about evaluating different types of players, given the fact that the available research has been exclusively focused on the general description of all player's performance (O'Donoghue & Ingram, 2001; Reid et al., 2016). Besides, in contrast to the previous studies (Reid et al., 2014) that empirically grouped players into different ranges as Top 10, top 20, top 50, etc. (which is also overemphasized by sport journalists and media), the present study highlighted that using cluster analysis techniques may avoid some subjectivity and prejudgment of player groups, and the results demonstrated its usefulness in differentiating players. Results of the discriminant functions showed high percentage of variance explained among all player groups and good ability of correctly classifying players in their respective categories, revealing the quality of the functions and the power of the structure coefficients of analyzed variables.

4.4.1 *Serve and Return Performance*

Overall, the serve-and-return related variables turned out to be the maximum contributors to these two discriminant functions in all Grand Slams, indicating the technical and tactical importance of serve and return in current professional male tennis (Hizan et al., 2014). In agreement with our hypothesis, High-experienced Higher RQ players (HHRQ) and Experienced Higher RQ players (EHRQ) players exceeded other players in those variables. It is possible that they had physical advantage of being taller and stronger, which could help them to serve 2—4 km/h faster than their opponents (Cross & Pollard, 2009). In contrast, High-experienced Lower Relative Quality Players (HLRQ) showed worse performance in serves, and it may be assumed that they were affected by the effect of aging on performance in that as players get older, they tend to be generally more experienced but are physically weaker because of the gradual loss of muscle function (Ma et al., 2013). However, it is speculated that being equally high-experienced players, HHRQ players were able to well maintain their physical and physiological competitiveness through better scientific and systematic training than HLRQ players.

From ace-related variables, ace in advantage% in Roland Garros and ace in deuce% in Wimbledon stood out as discriminant factors, which revealed the players' preference to serve into the wide zone to find more angles on clay court (Gillet et al., 2009), and to serve to all the corners of service boxes on grass court (Unierzyski & Wieczorek, 2004). HHRQ and EHRQ players hit more aces than the other players, suggesting the positive effect of faster serve speed on the probability of achieving aces. Correspondently, on slow court, the fast-serve strategy was employed like on fast courts (Gillet et al., 2009). Whiteside and Reid (2016b) argued that the fast serve speed (> 199 km/h) is beneficial to achieving an ace, but the most important factors are serve directionality in relation to the returner and ball's landing proximity to the nearest service box line, and that could suggest that experienced players in this study had better serving strategies. Concerning double fault%, all player categories were only clearly differentiated in RG. It was argued that best players could serve more aces but less double faults (Cross & Pollard, 2009). HHRQ players were able to commit less double faults than others in all Slams, but EHRQ players failed to meet this criterion despite aching more in their serves.

The higher winning percentage of return was achieved by HHRQ players, which would indicate their advanced visual search strategy when returning. Previous study showed that experienced and high ranked players had longer fixation duration for ritual, serve and action phases of the serve, compared to novice and lower ranked players. This gives them the perceptual advantage while processing information, anticipating accurately the strokes and making quicker reaction (Murray & Hunfalvay, 2017). Future study should examine the difference of return position, stroke techniques, net clearance and return locations (Reid et al., 2016) to better understand top player's returning behavior. After all, the serve and return performance profile of HHRQ players could be utilized as a model for the other players (especially novice and low-ranked players) who aspire to improve their performance through achieving specific goals and fostering cognitive strategies (Ma et al., 2013).

4.4.2 Net Performance

Net points won% was only able to discriminate the players groups in US Open, where HHRQ and EHRQ won more net points than the rest. But unexpectedly, all players had the similar performance in "net success of total points won%" in all Slams, which made us to infer that regardless of the absolute difference of court surface, the game strategy and tactics in elite

tennis are becoming similar. This is likely because, first of all, the improvement of technology in racket and ball enables players to hit faster serves and groundstrokes (Filipic et al., 2015); and secondly players are more physically competent through systematic and scientific training and nutrition prescription (Reid et al., 2016) to run and swing the racket faster (Pereira et al., 2017). Consequently, as the relative flying time of the ball and reaction time of players decrease, it is common to see players attacking in the net to finish the points. Thus, they could choose to end as many points as they would in the net, even on slower surface court. Nevertheless, the distinct winning percentages highlighted that there was a big difference in net skills (approach shot, volley and smash) among all players. Therefore, future investigation is suggested to closely inspect the players' execution of net points for the long-term perfection of their reservoir of techniques, as it constitutes one of the major tactical patterns in tennis.

4.4.3 Break Point Performance

Break points performance was one of the most contributing variable for the discriminant functions. HHRQ and EHRQ players were shown to own more break opportunities. Break point has a positive effect on receiving players' winning possibility in that they are more motivated to win the return game when a break point is created (Knight & O'Donoghue, 2012). Roland Garros was found to be the event where all players had more chances of break points because on slow surface, the advantage of serving players is neutralized by the slower speed and higher bounce of the ball (Filipic et al., 2015). Therefore, receiving players could have more time to react and try more tactical variability when returning the serves, thus winning more points. Nonetheless, players were not differentiated by break points won% and break points saved%, which probably suggests that professional players were mentally tough and more focused in crucial moments of breakpoints (Cowden, 2016), although they were under higher psychological stress. For this reason, it would be meaningful to investigate in the future the tactics players used in break points as well as their performance in the previous points.

4.4.4 Efficiency Performance

Our finding suggests that the winner per unforced error ratio is an important parameter to distinguish different player categories. Although it was not the case in RG, it is noteworthy

that in this clay event, players were discriminated by their ability to hit more winners in rally as well as in other Slams. The previous studies demonstrated that the ability of hitting more winners and committing less unforced errors is the benchmark of the technical and tactical prowess (Filipcic et al., 2015; Whiteside et al., 2016), physical potency (Murray & Hunfalvay, 2017; Reid et al., 2016) and psychological toughness for elite tennis player (Cowden, 2016). According to those findings, HHRQ and EHRQ players could be regarded as more outstanding groups in performance. But apart from their better skills or physical aptitudes, more studies on their point-level performance are needed to unravel their specific tactics.

4.4.5 Movement Performance

The physical movement performance has attracted substantial research interest in recent years (Pereira et al., 2017; Reid et al., 2016) thanks to the technological advances that facilitate the quantification of physical demands. But to the best of our knowledge, there is a limited research analysing movement characteristics for different tennis player groups. The current results increase the knowledge of tennis movement behavior and is useful for better training planning, considering the contextual constraints and recovery strategies (Pereira et al., 2017). Players ran approximately 2190.5 ± 777.0 — 2361.2 ± 913.0 m within the match in AO and 2233.5 ± 771.0 — 2253.9 ± 812.8 m in US, similar with the previous study (Reid et al., 2016) analyzing players in AO. While in RG, players traversed more distance, covering 2219.8 ± 655.7 — 2567.4 ± 944.1 m within the match, which was lower than previously reported 3160.0 ± 880.1 m in official ATP clay matches (Pereira et al., 2017). As male Grand Slams are played in best-of-five-set format (longer than ATP matches), this difference could only be attributed to the distinct measurement accuracy and criteria between the studies. Therefore, it is necessary that future study of movement performance apply the same or similar measurement standard to ensure the practical usefulness of the results. In Wimbledon, players covered less distance, around 1999.6 ± 623.2 — 2133.8 ± 786.1 m within the match as the points are generally ended quicker on grass surface court (O'Donoghue & Ingram, 2001).

Players were distinguished by the point-level movement performance in AO, W and US, where EHRQ players ran less distance per point than the other players (9.9 ± 2.3 m, 8.7 ± 2.1 m and 9.4 ± 1.9 m, respectively). These results could confirm the dominance of EHRQ players in serve-related variables (variables related to function 2) on fast court. There they had better chance to quickly win a point with high-velocity serve because it increased the

uncertainty for the opponents to return or to have good quality returns (Gillet et al., 2010). Therefore, they could run less distance to finish a point. Meanwhile, combining this finding with before-mentioned players' net performance in RG where all players traversed similar distance of 11.2 ± 2.7 — 11.8 ± 2.7 m, the study revealed that players were trying to play more aggressively on clay court, although their “big” serves would not work as efficient as the ones on fast surface courts. And remaining in the baseline in Roland Garros is no longer the best tactical option to win the game and attacking the net is also important on this surface. Future work is required to analyse the space-time dynamics of two players and its relationship with match scores, game tactics and stroke techniques.

4.5 Conclusion

In summary, this study analyzed the match performance of different tennis player groups in Grand Slams. Our findings suggest that playing experience and ranking-based relative quality (RQ) are good mediators of players' characteristics and match performance. Different player categories were clearly differentiated by their match performance, particularly related to serve and return, break points, net, rally and movement performance, facilitating a better interpretation of their performance profiles. High-experienced Higher RQ players performed overall better than other players, being especially more consistent in second serve and return, winner strokes and more breaking opportunities. Experienced Higher RQ players showed prominence in serving (faster serves and more aces) and aggressiveness in returning (more return winners and unforced errors), enabling them to run less distance. Additionally, our findings suggest that match tactics among court surfaces became less different and players approached similarly to the net in Roland Garros as they did in other Grand Slams. The findings provide insights for coach to prescribe individualized and tactics-oriented training programs for tennis players and to reconsider the match tactics in different Grand Slams against different opponents for the purpose of enhancing game performances.

Chapter Five - Clustering tennis players' anthropometric and individual features helps to reveal performance fingerprints

Cui, Y., Gómez, M.-Á., Gonçalves, B., & Sampaio, J. (In review). Clustering tennis players' anthropometric and individual features helps to reveal performance fingerprints. European Journal of Sport Science.

Abstract of the study published in Frontiers in Psychology (Special issue 5th, 2017).

5.1 Introduction

Collecting and analyzing match performance profiles has become an integral segment when coaching in professional sports (O'Donoghue & Mayes, 2013). In a complex and dynamic sport as tennis, where coaches and players need to make most precise decisions off and on the court to maximize the winning probability and to achieve longer sustainability of career. One of the keys to such success would lie in a comprehensive understanding of players' performances. Previous research have already focused intensively on how environmental factors such as court surface, temperature and consecutive competing days could influence the player's technical-tactical performance (Gillet et al., 2009; Smith, Reid, Kovalchik, Wood, & Duffield, 2017), movement performance (Ponzano & Gollin, 2017), physiological and perceptual response (Gescheit et al., 2015).

In such highly individual sport, tennis players have to constantly adjust their tactics according to their opponents. Hence, knowing well the opponent allows to better regulate behavior (Gómez, Lago, et al., 2013). But for an interactive sport like tennis, the quality of opposition has been overlooked or scarcely investigated for tennis until recently. Recent studies revealed the importance to evaluate the players' qualities in tennis (O'Donoghue & Cullinane, 2011) and identified that distinct relative quality and experience affected match performance (Cui et al., 2017).

Within this topic of assessing tennis player's performance, there are additional questions that seem worth exploring in further research. For instance, handedness has been deemed as an influencing factor to player's performance in many sports, with being left-handed considered as an advantage over right-handers (Hagemann, 2009). Possible explanations would be the innate superiority left-handers enjoy in terms of neuropsychological predispositions and strategic advantage while competing with right-handers (Loffing, Hagemann, & Strauss,

2010). Former research that investigated this polymorphism of handedness in tennis contrasted the year-end world ranking between left-handed and right-handed players (Loffing, Hagemann, & Strauss, 2012) and compared the serve performance between them (Loffing et al., 2009). These results implied that left-handers had certain advantage in the game but the effect was not very obvious in elite level.

Meanwhile, the influence of height on tennis serve performance was also examined but only clear positive correlations were found between body height and peak and first serve speed (Vaverka & Cernosek, 2013). Furthermore, given the importance of backhand style in tennis player's efficient stroke production, some studies analyzed the difference of swing velocity, (Reid & Elliott, 2002) shot accuracy (Muhamad et al., 2011) between two-handed backhand players and one-handed backhand players. No differences were found between them and it was suggested that more factors need to be considered such as player's kin-anthropometry and expertise level (Genevois, Reid, Rogowski, & Crespo, 2015). These attempts to discriminate players seem too much focused on the effects of one aspect, rather than integrating more performance determinants into a more holistic analysis.

Therefore, the current study was aimed to: i) explore and classify tennis players according to their anthropometric and individual features; ii) analyze the differences in technical-tactical and physical performance among distinct player types.

5.2 Method

Match dataset from four Grand Slams men's singles main draw (2015-2017) were obtained from the tournament organizers of Australian Open, Roland Garros, Wimbledon and US Open. The inclusion criteria of players were: i) having played on the court where both Doppler radar system (IBM: Armonk, NY, USA) and the Hawk-Eye camera system (Hawk-Eye Innovations, Southampton, UK) were mounted and ii) no early retirement occurred. The radar system comprises two radar units that are installed behind both baselines of the court and measures the serve speed with acceptable accuracy (average absolute measure difference: 0.8%) (ITF, 2005). The Hawk-Eye system consists of eight to ten high-speed video cameras that are located around the arena and the camera feeds are synchronized to track the movement of ball and player every 40 millisecond within point-play with a measurement mean error of 3.6 mm (Mecheri et al., 2016). Accordingly, a total of n=1188 players were

selected (Australian Open: range of ranking = 1 to 692; Roland Garros: 1 to 296; Wimbledon: 1 to 24; US Open: 1 to 581). Meanwhile, anthropometric and individual features for those players were collected from the official organizers of Association of Tennis Professional and International Tennis Federation. The study was conducted under the approval of the local University Ethics Committee and all procedures were undertaken to accomplish all international standards.

Two groups of variables were used in the study. The first included players' anthropometric and individual features: height (cm), weight (kg), handedness (left-handed or right-handed), backhand style (onehanded-backhand or two-handed-backhand) and playing experience (number of years a player has turned into the professional since the year he got his first professional ranking). The second group consisted of twenty-nine variables of interest related with serve and return, breakpoints, net, efficiency and movement performance in the matches (see Table 5.1). The reliability of intra-operators, inter-operators and operators vs. official data for match variables were previously tested (Cui et al., 2017). Results of ICC value ranged from 0.96 to 1 and standardized typical errors varied from 0.02 to 0.11, demonstrating very high level of reliability (Hopkins, 2000).

Table 5.1. Match performance variables and operational definitions

Variable	Operational Definition
<i>Serve Performance</i>	
Ace in Deuce%	% of aces in total serves to the deuce service box
Ace in Advantage%	% of aces in total serves to the advantage service box
Service winners%	% of service winners in total services
1 st serve points in%	% of first serves that were in play
1 st serve won in Deuce%	% of points won when first serves were served into the deuce service box
1 st serve won in Advantage%	% of points won when first serves were served into the advantage service box
2 nd serve won in Deuce%	% of points won when second serves were served into the deuce service box
2 nd serve won in advantage%	% of points won when second serves were served into the advantage service box
Double faults%	% of double faults in total number of second serves
Fastest serve speed (km/h)	The peak serve speed within the match
1 st Average speed Deuce (km/h)	Mean first serve speed when serving to the deuce service box
1 st Average speed Advantage (km/h)	Mean first serve speed when serving to the advantage service box
2 nd Average speed Deuce (km/h)	Mean second serve speed when serving to the deuce service box
2 nd Average speed Advantage (km/h)	Mean second serve speed when serving to the advantage service box
<i>Return Performance</i>	

Return points won%	% of points won when returning serves
Return winner%	% of return winners in total serve returns
Return unforced error%	% of unforced errors in total serve returns

Break Performance

Break points per return game	Number of break points had in every opponent's service game
Break points won%	% of break points converted in opponent's service game
Break points saved%	% of break points served in own service game

Net Performance

Net points won%	% of points won when player approached the net
Net points won of total points won%	% of net points won of total points won

Efficiency Performance

Winners	Total number of winner shots in the match
Winner per unforced error ratio	Number of winners in every unforced error
Winner of total points won%	% of winner shots in total number of points won
Dominance ratio	% of points won in opponent's serves / % of points lost in own serves

Physical Performance

Mean meters run per match (m)	Mean distance covered in a match by the player
Mean meters run per set (m)	Mean distance covered in each set by the player
Mean meters run per point (m)	Mean distance covered in each point by the player

Raw data were organized and match variables were standardized into their correspondent ratios. A two-step cluster analysis with log-likelihood as the distance measure was undertaken to classify players according to height, weight, handedness, backhand style and playing experience. The original data was grouped into pre-clusters by constructing a cluster features tree, then the standard hierarchical clustering algorithm on the pre-clusters was used and provided a range of solutions with different numbers of clusters. The Schwartz's Bayesian criterion (BIC), the BIC change value, the ratios of BIC changes and ratios of distance measures were calculated for each solution to find the optimal number of clusters, with the "better clustering solution" being expected to have a lower BIC value, larger value of BIC change, larger ratio of BIC changes and larger ratio of distance measures. Finally, the quality of the clustering model was measured by the average silhouette coefficient, which is a measure of both cohesion and separation. It was assessed as follows: -1.0 – 0.2, poor model, 0.2 – 0.5, moderate-to-fair model; > 0.5, very good model (Kaufman & Rousseeuw, 2009). Each input variable was ranked according to its predictor importance (PI), which shows its relative importance in estimating the cluster model. Interpretation of the cluster results was

described through calculating the mean values of the continuous variables and identifying the specific category of the nominal variables for each cluster (with according labels).

A descriptive discriminant analysis was used to identify which performance-related variables could best discriminate the clusters in every Grand Slam. The effect sizes of the discriminant functions were assessed by i) squared canonical correlation (r_c^2), which explains the variance associated with each function; and ii) partial η^2 , which describes the variance for the entire analysis and was computed by the following formula:

$$\text{partial } \eta^2 = 1 - \Lambda^{1/3}$$

where Λ represents Wilks' lambda (Leech, Barrett, & Morgan, 2014).

The interpretation of r_c^2 was as follows: small ($0.01 \leq r_c^2 < 0.09$), moderate ($0.09 \leq r_c^2 < 0.25$), large ($r_c^2 \geq 0.25$), while the strength of partial η^2 was interpreted by the following criteria: small ($0.01 \leq \text{partial } \eta^2 < 0.06$), moderate ($0.06 \leq \text{partial } \eta^2 < 0.14$) and large ($\text{partial } \eta^2 \geq 0.14$) according to the standards of Cohen (1988). Additionally, variables were considered as meaningful contributors to player group differences if their structure coefficients (SCs) in discriminant functions were higher than $|0.30|$. Validation of discriminant models was applied using the leave-one-out method of cross-validation, which takes subsets of data for training and testing, and is necessary for understanding the usefulness of discriminant functions when classifying new data. The analyses were performed using IBM SPSS software (Armonk, NY: IBM Corp.), and the significance level was set at $p \leq 0.05$. Finally, Matlab® (MathWorks, Inc., Massachusetts, USA) dedicated routines were used to generate a Convex hull area and a Cluster dispersion index area (see Figure 5.2), using discriminant scores from the obtained functions. The Convex hull area is the smallest polygonal region that marks the peripheral scores of certain clustering group. Complementarily, cluster dispersion index area demarcates the radial expansion of performance for each player cluster, and it was calculated using the mean distance from each player's discriminant score to the geometrical center of the corresponding cluster (Gonçalves et al., 2016; Sampaio, Lago, Gonçalves, Maças, & Leite, 2014). These complementary dispersion analyses offer an intuitive evaluation of the homogeneity of different groups' performance and the distribution of players within both cluster and functions.

5.3 Results

An optimal number of four player clusters were identified from a total of fifteen clustering solutions, with BIC = 2078.633, change of BIC = -544.021, ratio of BIC changes = 0.415 and ratio of distance measures = 3.181. The average silhouette coefficient was larger than 0.5, indicating a very good cluster model. Among all input variables of this cluster solution, the variable “backhand style” turned out to be the strongest predictor (PI = 1.00), followed by “handedness” (0.90), “weight” (0.65) and “height” (0.58); while “playing experience” was the weakest predictor (0.15). Figure 5.1 shows the labels and characteristics of four different player clusters. There were 387 righty two-handed backhand players with longest height (194.7 ± 4.8 cm), heaviest weight (89.5 ± 6.7 kg) and 9.3 ± 3.1 years of playing experience were labeled as Big-sized Righty Two-handed Players (BRT); 265 righty one-handed backhand players with medium height (185.9 ± 6.9 cm), medium weight (81.8 ± 7.3 kg) and longest playing experience (13.2 ± 3.7 years) were labeled as Medium-sized Righty One-handed Players (MRO); 414 righty two-handed backhand players with shortest height (183.3 ± 4.9 cm), lowest weight (75.3 ± 4.2 kg) and 10.2 ± 3.8 years of playing experience were labeled as Small-sized Righty Two-handed Players (SRT); and 122 lefty two-handed backhand players with medium height (186.3 ± 5.6 cm), medium weight (82.1 ± 6.5 kg) and 11.5 ± 4.0 years of playing experience were labeled as Lefty Two-handed Players (LT).

Figure 5.1. Description of player groups by cluster analysis

Result from the discriminant analysis showed that overall Wilks' lambda (Λ) for the first two discriminant functions (from the obtained three) were significant for Australian Open, Roland Garros and Wimbledon, while the Λ for all three functions in US Open were significant (detailed results are presented in Table 5.2 and 5.3). This showed that the discriminant models were able to significantly discriminate the player clusters in four Grand Slams. Serve and winner related variables were shown to be the major contributors to the separation of all player groups in all Slams. Besides, the following variables were identified as significant contributors to the discriminant functions: (i) Wimbledon, 1st serve won in advantage% (SC = 0.46), 2nd serve won in deuce% (0.30) and net success of total points won% (0.47); and (ii) US Open, serve winner% (0.31), return unforced error% (-0.32), total distance covered (0.33) and distance covered per point (0.36).

Using discriminant scores from first and second functions, the combined-plots (Figure 5.2, panel a), convex-hull area (panel b) and dispersion index area with group centroid (panel c)

were generated to depict the features of four player groups in each Grand Slam and to interpret their overall performance tendency. Compared with the rest of players, BRT players showed better performance in variables related to function 1 in all Slams. MRO players exhibited overall balanced performance in variables relevant to both functions during Australian Open and US Open, while having a better performance in variables related to function 2 in Roland Garros and Wimbledon. SRT players showed generally worse performance in variables related to both functions. In contrast with all groups, LT players showed better performance in Roland Garros. Considering convex-hull area and dispersion index, LT players were the most homogenous player group among all players, while MRO players demonstrated greater level of heterogeneity in all Slams as being identified with larger convex-hull area and dispersion radius.

Table 5.2. Means (Standard deviations), discriminant function details and structure coefficients (SC) of technical-tactical and movement performance for four player clusters in Australian Open and Roland Garros

Variables	Grand Slam	Australian Open					Roland Garros						
	Cluster Group	BRT	MRO	SRT	LT	Function		BRT	MRO	SRT	LT	Function	
		(n=154)	(n=104)	(n=179)	(n=55)	1	2	(n=45)	(n=37)	(n=56)	(n=16)	1	2
Ace in Deuce%		14.6 (8.4)	12.2 (8.3)	8.5 (6.2)	7.4 (5.8)	0.49 [#]	-0.10	7.7 (4.7)	6.2 (5.2)	4.1 (3.2)	2.2 (2.4)	0.36 [#]	-0.25
Ace in Advantage%		12.9 (8.2)	9.6 (7.5)	6.6 (5.2)	9.0 (7.2)	0.45 [#]	0.34 [#]	7.2 (5.0)	4 (3.9)	3.6 (2.9)	3.0 (3.6)	0.37 [#]	0.06
Serve winner%		2.2 (1.5)	2.0 (1.8)	1.2 (1.1)	1.1 (1.1)	0.44 [#]	-0.12	0.9 (1.2)	0.9 (0.9)	0.6 (0.9)	0.4 (0.6)	0.13	-0.16
1 st serve in%		62.8 (6.5)	61.9 (6.2)	60.6 (6.2)	64.1 (7.5)	0.11	0.43 [#]	62.2 (7.1)	59.8 (5.8)	62.3 (8.4)	65.0 (7.8)	-0.02	0.27
1 st serve won in Deuce%		77.2 (9.8)	74.6 (10.9)	72.5 (9.8)	70.3 (11)	0.28	-0.13	73.7 (10.3)	70.4 (12.0)	67.2 (10.8)	64.5 (13.6)	0.24	-0.12
1 st serve won in Advantage%		75.4 (9.5)	72.6 (11.8)	69.8 (11.2)	70.1 (13.1)	0.27	0.06	73.5 (11.4)	70.1 (11.8)	65.4 (12.1)	67.5 (9.5)	0.25	-0.05
2 nd serve won in Deuce%		56.6 (14.1)	58.7 (13.9)	55.7 (14.4)	55.3 (12.5)	0.06	-0.10	56.2 (15.5)	57.9 (15.7)	55.9 (12.7)	53.7 (21.5)	0.01	-0.10
2 nd serve won in Advantage%		58.8 (14.8)	56.2 (12.3)	55.9 (14.9)	58.4 (15.9)	0.08	0.19	58.1 (13.8)	56.4 (17.7)	53.2 (14.5)	63.4 (22.2)	0.07	0.11
Double fault%		10.1 (5.7)	9.1 (4.4)	8.5 (4.9)	9.7 (6.2)	0.13	0.22	8.5 (5.6)	6.3 (5.0)	8.1 (5.1)	7.4 (6.9)	0.04	0.16
Fastest Serve Speed (km/h)		214.6 (8.8)	209.3 (11.4)	200.8 (8.8)	201.3 (9.2)	0.82 [#]	0.07	212.6 (6.0)	204.2 (9.3)	197.8 (8.0)	195.8 (7.8)	0.74 [#]	-0.19
1 st Average Speed Deuce (km/h)		180.6 (25)	179.0 (22.8)	175.5 (19.7)	178.7 (17.8)	0.11	0.11	191.6 (6.2)	183.4 (17.6)	178.7 (13.0)	178.3 (8.5)	0.40 [#]	-0.04
1 st Average Speed Ad (km/h)		181.2 (27.2)	174.3 (26.4)	178.1 (16.9)	179.0 (14.8)	0.04	0.15	185.5 (11.7)	176.5 (15.6)	178.9 (9.4)	178.6 (8.2)	0.23	0.20
2 nd Average Speed Deuce (km/h)		152.5 (20.3)	149.9 (23.5)	143.5 (21.3)	137.5 (23.8)	0.28	-0.23	137.5 (33.3)	146.7 (24.7)	135.3 (25.7)	130.4 (26.3)	0.04	-0.25
2 nd Average Speed Ad (km/h)		145.3 (23.1)	141.5 (28.1)	137.8 (22.2)	144.1 (20.5)	0.14	0.24	131.2 (29.6)	137.5 (28.3)	133.1 (22.7)	121 (31.9)	0.00	-0.22
Return points won%		35.1 (8.5)	36.2 (8.3)	34.8 (8.4)	33.6 (8.1)	0.06	-0.18	38.7 (7.9)	37.2 (9.0)	38.1 (9.2)	38.4 (12.1)	0.02	0.08
Return winner%		2.1 (1.7)	1.8 (1.7)	2.1 (1.9)	1.9 (1.6)	0.01	-0.01	1.1 (1.1)	1.3 (1.3)	1.9 (1.7)	1.4 (1.6)	-0.18	0.05
Return unforced error%		3.6 (2.7)	3.9 (2.8)	3.4 (2.6)	3.4 (2.4)	0.07	-0.07	3.4 (2.0)	3.3 (2.3)	3.5 (2.2)	2.7 (2.2)	0.02	-0.07
Break points per return game		0.5 (0.3)	0.5 (0.3)	0.5 (0.3)	0.5 (0.3)	0.08	-0.09	0.6 (0.3)	0.6 (0.3)	0.6 (0.4)	0.7 (0.4)	0.01	0.08
Break points won%		36.0 (20.6)	39.8 (23.0)	40.1 (25.7)	33.2 (21.6)	-0.04	-0.29	37.3 (19.2)	36.4 (23.6)	44.2 (17.9)	36.5 (22.8)	-0.10	0.05
Break points saved%		55.4 (27.9)	51.3 (28.2)	53.7 (23.3)	51.1 (24)	0.03	-0.02	65.8 (20.1)	60.2 (22.0)	56.2 (19.3)	60.8 (19.4)	0.17	0.03

Net point won%	67.5 (12.7)	68.0 (11.7)	65.8 (12.9)	66.4 (13.2)	0.08	-0.01	63.7 (10.5)	65.8 (10.5)	64.3 (13.9)	64.6 (19.6)	-0.02	-0.06
Net success of total points won%	15.5 (6.9)	17.4 (7.5)	13.4 (6.6)	16.4 (9)	0.16	0.16	15.4 (4.6)	16.7 (9.8)	14.4 (7.1)	12.1 (6.1)	0.08	-0.24
Winners	44.5 (16.1)	42.4 (15.6)	34.1 (13.1)	36.1 (17.4)	0.39 [#]	0.06	38.1 (12.7)	38.6 (13.2)	32.3 (12.5)	23.1 (8.1)	0.23	-0.39 [#]
Winner per unforced error ratio	1.4 (0.7)	1.3 (0.7)	1.0 (0.6)	1.1 (0.7)	0.32 [#]	0.12	1.3 (1.2)	1.2 (0.4)	1.0 (0.5)	1.5 (1.9)	0.06	0.09
Winner of total points won%	38.4 (8.4)	37.5 (9.8)	30.9 (7.9)	31.9 (8.8)	0.51 [#]	0.01	33.4 (7.1)	36.1 (7.3)	32.2 (7.8)	27.6 (8.6)	0.10	-0.41 [#]
Dominance ratio	1.2 (0.5)	1.2 (0.5)	1.0 (0.5)	1.0 (0.5)	0.22	-0.06	1.2 (0.5)	1.1 (0.4)	1.0 (0.5)	1.2 (1.2)	0.06	0.09
Total distance covered (m)	2205.4 (711.5)	2225.4 (813.1)	2389.2 (978.2)	2391.9 (848.3)	-0.13	0.03	2594.8 (775.1)	2490.7 (899.5)	2426.3 (929.9)	1943.2 (632)	0.12	-0.16
Distance covered per set (m)	609.0 (156.4)	612.9 (180.5)	644.2 (200.5)	637.3 (130.1)	-0.12	-0.01	687.0 (141.1)	692.4 (176.3)	682.6 (200.1)	630.7 (181.3)	0.03	-0.10
Distance covered per point (m)	9.8 (2.1)	10.0 (2.3)	10.5 (2.5)	10.4 (2)	-0.16	0.00	11.6 (2.2)	11.5 (2.6)	11.8 (2.9)	11.3 (2.9)	-0.01	0.00
Wilks' lambda					0.51	0.82					0.24	0.54
Chi-square (χ^2)					318.74	92.44					194.87	84.09
Significance					<0.001	0.002					<0.001	0.009
Partial η^2					0.20**	0.06*					0.38**	0.19**
r_c^2					0.38**	0.11*					0.56**	0.34**
% of Variance					74.9	15.5					62.9	26.0
Reclassification (%)					54.9%						67.5%	

Legend: [#] Structure coefficient $\geq |0.30|$; * moderate effect; ** large effect. BRT=Big-sized Righty Two-handed players; MRO=Medium-sized Righty One-handed players; SRT=Small-sized Righty Two-handed players; LT=Lefty Two-handed players

Table 5.3. Means (Standard deviations), discriminant function details and structure coefficients (SC) of technical-tactical and movement performance for four player clusters in Wimbledon and US Open

Variables	Grand Slam	Wimbledon					US Open							
	Cluster Group	BRT	MRO	SRT	LT	Function		BRT	MRO	SRT	LT	Function		
		(n=103)	(n=58)	(n=95)	(n=24)	1	2	(n=85)	(n=66)	(n=84)	(n=27)	1	2	3
Ace in Deuce%		16.2 (9.2)	12.3 (8.1)	7.9 (5.9)	6.8 (5.1)	0.54 [#]	-0.07	13.2 (7.7)	8.4 (5.9)	6.2 (4.8)	4.7 (4.0)	0.50 [#]	0.02	0.11
Ace in Advantage%		13.8 (8.2)	11.3 (8.6)	6.0 (4.4)	9.8 (6.7)	0.50 [#]	0.15	12.9 (7.8)	8.8 (6.5)	5.4 (3.9)	5.9 (4.9)	0.48 [#]	0.12	-0.09
Serve winner%		1.7 (1.3)	1.2 (1.1)	1.0 (1.0)	1.2 (0.9)	0.28	-0.10	3.1 (2.7)	2.3 (2.8)	1.5 (1.5)	1.1 (1.1)	0.31 [#]	-0.06	-0.04
1 st serve in%		64.4 (6.0)	62.8 (6.1)	61.9 (6.5)	62.4 (6.8)	0.19	-0.06	57.7 (6.3)	58.7 (6.6)	56.8 (7.5)	61.0 (6.2)	-0.02	0.22	-0.31
1 st serve won in Deuce%		79.3 (7.8)	78.7 (10.5)	72.3 (10.5)	70.8 (10.0)	0.38 [#]	0.19	77.3 (10.0)	73.6 (10.9)	70.9 (11.7)	70.4 (9.3)	0.25	0.02	-0.02
1 st serve won in Advantage%		79.4 (10.9)	75.2 (10.8)	69.1 (11.1)	70.3 (8.3)	0.46 [#]	0.02	75.9 (10.8)	72.6 (10.3)	69.1 (10.9)	70.2 (9.1)	0.25	0.08	-0.09
2 nd serve won in Deuce%		59.1 (13.4)	63.1 (15.8)	56.4 (15.8)	57.6 (17.0)	0.09	0.30	59.1 (14.6)	59.1 (12.8)	55.8 (13.0)	56.2 (14.8)	0.09	-0.02	-0.13
2 nd serve won in Advantage%		59.1 (15.7)	60.5 (16.4)	56.0 (13.3)	59.5 (16.3)	0.09	0.19	60.3 (14.6)	59.1 (15.2)	57.3 (15.3)	60.1 (14.7)	0.06	0.10	-0.07
Double fault%		9.4 (5.0)	7.9 (4.5)	9.1 (6.5)	10.3 (5.6)	0.00	-0.18	11.0 (6.7)	11.2 (6.0)	10.2 (5.7)	11.7 (4.6)	0.02	0.09	-0.15
Fastest Serve Speed (km/h)		214.1 (8.0)	208.4 (10.4)	200.3 (6.6)	203.3 (7.8)	0.80 [#]	0.04	215.8 (7.6)	207 (9.3)	200.6 (9.1)	201.8 (6.9)	0.70 [#]	0.21	-0.10
1 st Average Speed Deuce (km/h)		174.4 (30.8)	174.0 (29.3)	165.3 (27.5)	176.5 (23.0)	0.13	0.15	190.6 (8.1)	185.3 (8.9)	178.9 (8.8)	178.5 (8.2)	0.56 [#]	0.01	-0.18
1 st Average Speed Ad (km/h)		168.8 (33.2)	174.1 (27.0)	172.5 (24.7)	166.6 (28.1)	-0.04	0.09	192.0 (9.0)	182.1 (9.8)	178.1 (9.0)	176.7 (6.5)	0.64 [#]	0.13	0.13
2 nd Average Speed Deuce (km/h)		160 (19.6)	156.2 (22.6)	142.5 (22.7)	138.3 (26.0)	0.41 [#]	0.11	158.1 (12.2)	155.1 (9.0)	146.2 (7.2)	143.5 (6.8)	0.54 [#]	-0.22	-0.32
2 nd Average Speed Ad (km/h)		151.7 (25.4)	152.7 (23.4)	144.3 (22.3)	142.3 (30.4)	0.16	0.13	153.5 (12.2)	149.5 (11.4)	141.8 (8.7)	146.6 (6.6)	0.39 [#]	0.20	-0.34
Return points won%		34.1 (8.3)	33.9 (7.8)	32.7 (8.9)	30.4 (8.6)	0.10	0.01	36.8 (8.3)	34.7 (7.9)	36.5 (8.5)	36.7 (9.1)	0.01	0.12	0.18
Return winner%		2.1 (1.7)	1.6 (1.4)	1.7 (1.6)	1.4 (1.4)	0.14	-0.22	2.1 (1.8)	1.5 (1.6)	1.9 (1.6)	1.6 (1.4)	0.06	0.04	0.25
Return unforced error%		1.5 (1.5)	1.5 (1.9)	1.1 (1.3)	1.3 (1.3)	0.12	0.14	3.8 (2.7)	3.4 (2.9)	2.1 (2.1)	3.6 (2.6)	0.20	0.24	-0.32 [‡]
Break points per return game		0.5 (0.3)	0.5 (0.3)	0.4 (0.3)	0.4 (0.3)	0.09	0.04	0.6 (0.3)	0.5 (0.3)	0.5 (0.3)	0.5 (0.3)	0.06	0.07	0.16
Break points won%		36.9 (23.4)	33.9 (20.8)	36.6 (26.4)	33 (26.3)	0.02	-0.12	35.5 (20.2)	36.3 (24.5)	42.2 (22.1)	39.9 (27.3)	-0.11	-0.01	0.14
Break points saved%		65.9 (25.2)	69.6 (26.2)	58.9 (22.8)	64.9 (15.7)	0.13	0.27	67.6 (24.0)	64.6 (23.5)	55.7 (20.4)	55.3 (19.8)	0.22	-0.04	-0.16

Net point won%	66.1 (12.0)	65.4 (9.9)	64.9 (12.4)	61.5 (15.2)	0.07	-0.05	65.8 (11.8)	65.0 (12.4)	65.2 (11.0)	69.8 (11.9)	-0.03	0.22	-0.05
Net success of total points won%	18.3 (8.4)	20.8 (7.8)	15.1 (6.8)	17.0 (6.0)	0.20	0.47 [#]	16.0 (6.2)	16.9 (8.1)	15.8 (6.4)	15.6 (7.3)	0.01	-0.08	-0.11
Winners	46.8 (16.3)	42.6 (16.3)	31.0 (14.0)	35.7 (18.6)	0.47 [#]	0.17	41.5 (14.9)	35.3 (14.7)	33.2 (12.5)	31.8 (12.2)	0.25	0.04	0.08
Winner per unforced error ratio	2.2 (1.2)	2.0 (1.1)	1.4 (0.8)	1.4 (0.6)	0.37 [#]	0.13	1.3 (0.6)	1.2 (0.7)	1.1 (0.8)	0.9 (0.3)	0.15	-0.09	-0.01
Winner of total points won%	37.9 (7.6)	37 (8.1)	29.4 (7.5)	31.8 (6.9)	0.54 [#]	0.34	36.8 (8.3)	33.2 (9.5)	29.9 (7.3)	28.1 (6.1)	0.37 [#]	-0.06	-0.03
Dominance ratio	1.2 (0.4)	1.2 (0.6)	1 (0.5)	0.9 (0.4)	0.28	0.17	1.2 (0.5)	1.1 (0.5)	1.0 (0.5)	1.0 (0.3)	0.16	0.02	-0.01
Total distance covered (m)	2071.9 (666)	2016.5 (589.2)	2078.5 (818.6)	2102.9 (708.3)	-0.01	-0.07	2110.9 (695.2)	2015.7 (812.3)	2469.6 (809.1)	2475.9 (815.4)	-0.19	0.14	0.33 [#]
Distance covered per set (m)	546.8 (126.1)	571.7 (128.6)	575.9 (166.9)	581.5 (127.3)	-0.10	0.08	587.0 (134.7)	574.4 (161.9)	651.6 (173.9)	661.1 (142.4)	-0.17	0.14	0.26
Distance covered per point (m)	8.8 (2.1)	9.2 (1.8)	9.7 (2.5)	9.6 (1.9)	-0.20	0.00	9.7 (2.1)	9.4 (1.9)	10.9 (2.4)	10.9 (2.2)	-0.22	0.15	0.36 [#]
Wilks' lambda					0.41	0.75					0.30	0.64	0.82
Chi-square (χ^2)					235.11	75.40					293.10	107.44	48.23
Significance					<0.001	0.04					<0.001	<0.001	0.007
Partial η^2					0.25**	0.09*					0.33**	0.14*	0.06*
r_c^2					0.46**	0.16*					0.53**	0.22*	0.18*
% of Variance					73.0	17.0					69.8	16.8	13.4
Reclassification (%)					65.7%						67.9%		

Legend: [#] Structure coefficient $\geq |0.30|$; *moderate effect; ** large effect. BRT=Big-sized Righty Two-handed players; MRO=Medium-sized Righty One-handed players; SRT=Small-sized Righty Two-handed players; LT=Lefty Two-handed players

Figure 5.2. The distribution of four player groups along with their convex-hull areas, complementary dispersion index areas and semifinals featured by discriminant functions

5.4 Discussion

This study examined match performances of different types of tennis players in Grand Slams according to their anthropometric and individual attributes. Results allowed us to build an unsupervised learning model to classify the players into four distinct groups. Furthermore, results of discriminant analysis identified serve and winner related variables as key performance indicators contributing to the differentiation of four player groups in all Grand Slam tournaments. In fast-pace court Slams like Wimbledon and US Open, net and running performance related variables also helped to discriminate these player groups.

Among all performance variables, fastest serve speed, ace in Deuce%, ace in Advantage% and winners of total points won% were identified as key indicators in all tournaments, with BRT and MRO players exhibiting a better performance. A strong correlation was addressed between body height and fastest serve speed in elite tennis due to the fact that increasing the height at which the ball is hit in 10-cm increments from 2.7 to 3.0 m would enable the landing position of the serves to move inside the service box by 25-30cm from the service line (Vaverka & Cernosek, 2013). This explains taller players' genetic advantage over shorter peers for being able to hit the serves at a greater height. Besides, as recent study argued that the most important factors for better serving ability are serve directionality in relation to the returner and ball's landing proximity to the nearest service box line rather than fast serve speed (Whiteside & Reid, 2016b), we infer that taller players also had greater control of serve directions in addition to their height advantage.

Moreover, in contrast with previous assumption that body height and second serve speed were weakly related (Vaverka & Cernosek, 2013), this study showed that taller players could also serve a significantly faster second serves in Wimbledon and US Open, which are fast-pace court tournaments. Meanwhile their second serves winning performance, especially second serves won in Deuce% was superior to SRT and LT players. This reigning serve capacity helped BRT and MRO players obtain higher winner of total points won% in all events but not in Roland Garros, where LT players overshadowed the others by hitting more winner strokes. In Wimbledon, BRT and MRO players won more net points than the rest of players, and finishing the point in the net was showed to be a major tactic for these taller players who has a convincing serve power (Hizan et al., 2014).

Our findings also imply that match circumstances affect the match-play of players to a great extent (Gómez, Lago, et al., 2013). Taller players could benefit from the variation of match

strategies and locations to maximize their predominance in serves, and serving the second serves like the first on fast-pace courts would give returners extra difficulty to return or to have a good return.

A number of studies highlighted the advantage of lefty-handed tennis players over their righty-handed counterparts in competition, especially in elite male category (Breznik, 2013; Loffing et al., 2009, 2012). However, the available findings suggested that their advantage is over-represented at elite level. There is a chance that in lower level tournaments like Challengers and Future tournaments this advantage persists. For higher-ranked players, the unfamiliarity with competing against left-handers is probably not the most serious issue. Therefore, the most obvious advantage of left-handers identified in this study resides in Roland Garros. In this clay court event, ball bounces and angles were higher for left-handed serves to the outside zone of the advantage court, which forced the returner to move a longer lateral distance and to hit the serve at a comparatively higher height (Gillet et al., 2009). This consequently helped lefty-handed players establish the advantage at the initial of points. Furthermore, when returning, the high friction of court diminished the prominence of righty-handed serve speed, giving LT players more time to react and prepare for a good return, particularly on advantage court, which is their favorite side of the court (Loffing et al., 2009).

Concerning the effect of player's backhand preference on match-play performance, the results showed that MRO players traversed less total distance and distance per point than the others in Wimbledon (2016.5 ± 589.2 m; 9.2 ± 1.8 m) and US Open (2015.7 ± 812.3 m; 9.4 ± 1.9 m), indicating a better running performance of MRO players in fast-pace tournaments. It has been verified that backhand strokes are less frequently played than forehand strokes even in professional circuit and no research has provided a clear answer to which backhand technique is better (Genevois et al., 2015). Nonetheless, two-handed backhand strokes rely more on trunk rotation for racquet velocity generation and one-handed backhand strokes rely more on segmental rotations of the upper limb. Thus, it implies that two-handed backhand players need to run more distance than one-handed backhanders to place themselves in a technically correct position for executing efficient backhand strokes. In fast court surface, this slight difference in distance would be critical in shot exchanges that happen mostly within less than one second. One-handed backhanders in this circumstance is, therefore, expected to have the advantage to cover more court space and guarantee good stroke quality within less distance. Furthermore, comparing the total distance of MRO to that of SRT and LT players in US Open, a difference

of 400 m would be substantial for maintaining the physical status and planning the recovery strategy at this top level.

Playing experience was another player characteristic that could not be overlooked, and former study demonstrated that it had great influence on professional player's match performance (Cui et al., 2017; Murray & Hunfalvay, 2017). Therefore, the better performance of MRO players may not be merely related to their backhand style, as they were the group with longest playing experience (13.2 ± 3.7 years). It is noteworthy that being highest-experienced group, they would benefit from a more advanced visual search strategy and perceptual advantage in processing game information, anticipating accurately the opponent's moves, reacting faster to strokes and making correct decisions (Murray & Hunfalvay, 2017). Finally, all these attributes could lead to their winning more points and committing less unforced errors.

With convex-hull area and complementary Cluster dispersion index area, it is possible to identify group homogeneity and recognize how well the discriminant functions separated them. To the best of our knowledge, this is the first study to apply such techniques in performance assessment and profiling. Of all groups, Lefty Two-handed Players were the most homogenous and this possibly indicates that they were able to exert a comparatively similar performance regardless of their opponents. It also could be inferred that fewer modifications in tactics were made throughout the course of tournaments. Whereas, taller players like Big-sized Righty Two-handed players and Medium-sized Righty One-handed players were in contrast more heterogeneous in their performance, showing that their match performance were varied match-to-match to a greater extent and the tactics could, therefore, be adjusted accordingly. Moreover, considering the fact that more players from these two groups made into the semi-finals in the four Grand Slams (Figure 2, panel c), it would be reasonable to conclude that adapting strategy and tactics is essential for players at the highest level. Additionally, these inter- and intra-groups variations in match performances deserve further assessment in research as well as practical attention in training. Based on the analysis, we would think that it is helpful to mix different types of player, which would add contextual variation to the daily training scenario, thus promoting player's better adaptation ability to distinct opponents.

5.5 Conclusion

Assessing player's match performance is always one of the most critical issues in coaching and training process. This study distinguished different types of tennis players utilizing a series of individual attributes and increased the knowledge of tennis match behavior by identifying key performance indicators that discriminate the player groups in each Grand Slam. The serve performance-related variables turned out to be the maximum contributors to the groups' separation. Big-sized Righty Two-handed players and Medium Righty One-handed players outperformed their counterparts within all tournaments but not Roland Garros, where Lefty Two-handed players established their advantage thanks to the slow court surface. The findings of the current study advance the knowledge of match-play performance and game behavior of distinct types of player, helping to identify the main features of different players in Grand Slams. Coaches and players could use this information to evaluate performance from a more complex perspective and design fine-tuned opponent-based training programs.

5.6 Practical Application

- Classifying players based on different features is useful for opponent assessment and match preparation
- Left-handed players showed most consistent performance but their advantage over their peers was no longer obvious at the highest level of tennis
- Mixing different types of player would add the contextual variation to the daily training scenario, thus promoting player's better adaptation ability to distinct opponents.

Chapter Six - Performance evaluation of point-by-point data for individual male tennis player in Grand Slams

Cui, Y., Liu, H. & Gómez, M.-Á. (In review). Performance evaluation of point-by-point data for individual male tennis player in Grand Slams. Kinesiology. Abstract of the study published in Journal of Human Sport and Exercise (Special issue, 2017).

6.1 Introduction

Match performance of tennis player is one of the topics that attract major research interest in sport performance analysis (McGarry et al., 2013). Competing in professional tennis requires player to win the game by making proper strategies and tactics according to different tournaments, rounds, sets, quality of opponents and consistently tailoring them based on the match periods and match status. With the application of performance profiling techniques, the technical, tactical, physical and psychological features of tennis players could be evaluated and compared (Butterworth et al., 2013). The available research has intended to describe the performance of tennis players of different sex (Hizan et al., 2011) and different levels (Galé-Ansodi et al., 2016; Sanchez-Pay et al., 2015), considering the match outcome or game location (McGarry et al., 2013; Reid et al., 2016). However, little was known about how the performance of a tennis player is influenced by other situational factors, such as the opponent's quality and match status, which have been investigated in major team sports like basketball and football (Gómez, Lorenzo, Ibañez, & Sampaio, 2013; Lago-Peñas, 2012). These situational factors takes into consideration the game location, game status, quality of opposition, period of play and type of competition and their effects on player's performance and game outcome (Gómez, Lago, et al., 2013).

Gillet et al. (2009) were ones of the first to assess the influence of game location on tennis player's serve and return performance, after which Galé-Ansodi et al. (2016) attempted to describe the physical performance of youth tennis player on hard and clay courts. More recently Reid et al. (2016) described the match performance of male and female tennis players in Australian Open, which was on hard court. However, the influence of other court surface like grass, strength of the opponents and game status were not considered. Hence, it is of interest to know how tennis player perform under these different situational circumstances.

As the application of new technology in sport fields is generating a great amount of data on match and training performance of athletes (Hughes & Franks, 2015; McGarry et al., 2013), it

is necessary to evaluate their performance in a more in-depth manner for advanced understanding of match behaviors. This ultimately can be used to help in-game coach decision-making with emphasis on those performance indicators which are of greatest relevance to achieving a winning outcome and also to optimizing individualized training for players (Reid et al., 2010; Reid et al., 2016).

Applying data mining and statistical modelling techniques, sport data could be interpreted to provide meaningful information and insights (Ofoghi et al., 2013). And being one of these techniques, classification tree analysis has been used lately by researchers of sport science, particularly in team sports (Gómez, Battaglia, et al., 2015) and proved to be useful in modeling non-linear phenomena. It establishes a hierarchical solution to classify complex problems, where a set of rules is derived from the interaction between attributes in a data set, with high predictive accuracy (Gómez, Battaglia, et al., 2015). Previous studies has applied the technique to model the ball screen effectiveness in basketball (Gómez, Lorenzo, et al., 2013), the match outcome in Australian football (Robertson, Back, & Bartlett, 2016), the ball possession effectiveness in futsal (Gómez, Moral, & Lago-Peñas, 2015) and the effect of scoring first on match outcome in football (Lago-Peñas, Gómez-Ruano, Megías-Navarro, & Pollard, 2016). This inspired us to use this model to analyse in detail how tennis player performed when competing against various level of opponents under distinct match status on different court surfaces and understand their interactions, instead of just analyzing and comparing the result of the performance.

Therefore, by applying the classification tree model, the present study was aimed to examine the point-by-point performance of individual tennis player in four Grand Slams, considering comprehensively various situational variables like court surfaces, opponent's quality and match status.

6.2 Methods

6.2.1 Sample and subjects

Point-level data of 145 Grand Slam matches played by Novak Djokovic within year 2011 to 2016 were collected from separate official tournament websites: Australian Open (www.ausopen.com), Roland Garros (www.rolandgarros.com), Wimbledon (<http://www.wimbledon.com>) and US Open (www.usopen.org). Novak Djokovic was stably

ranked top 3 in Association of Tennis Professionals (ATP, www.atpworldtour.com) through 2011-2016, being number 1 in ranking from 2014-2016. Besides, 11 of his 12 Grand Slam titles were obtained during this period. Therefore, the analysis of his performance could give us better understanding on how he behaved with these Slams and also provide a typical performance pattern of elite player for the rest who aspire to improve.

The data included the notational statistics such as first and second serve, rally numbers and game scores; and ball speed statistics collected by Doppler radar system (IBM: Armonk, NY, USA) that were installed in the courts. Matches that were not completed due to the retirement of the opponents or a walkover were excluded from the study. In total, there were 29675 points (See Table 6.1). The study was approved by the local University Ethics Committee.

Table 6.1. Distribution of Points Played by Novak Djokovic in four Grand Slams

	Australian Open	Roland Garros	Wimbledon	US Open	Total
Year	M—S—P	M—S—P	M—S—P	M—S—P	M—S—P
2011	6—19—1152	4—14—813	7—25—1386	5—19—1055	22—77—4406
2012	7—27—1504	7—26—1594	6—20—1087	7—24—1243	27—97—5428
2013	7—25—1505	5—18—1111	7—23—1498	7—25—1342	26—91—5456
2014	5—17—952	7—24—1401	7—27—1739	6—20—1118	25—77—5210
2015	7—24—1440	7—24—1388	7—24—1509	7—24—1329	28—96—5666
2016	7—24—1482	7—23—1331	3—10—696	N/A	17—57—3509
Total	39—136—8035	37—129—7638	37—129—7915	32—112—6087	145—506—29675

M—S—P denotes for total number of matches, sets and points played.

6.2.2 Variables

Table 6.2 presents eight situational variables and two performance-related variables that were included as predictor variables for point outcome (win/lose).

Based on the importance of Top-100 ranking that the previous research has continuously addressed (Prieto-Bermejo & Gómez-Ruano, 2016; Reid et al., 2010; Reid et al., 2014), we have established five categories to classify the quality of opposition, using Top-10, Top-20,

Top-50 and Top-100 world ranking as cut value. In addition, as there are few studies that have addressed the issue of game status in tennis, we decided to define 5 situations of game status, considering the sequence of scores within every game, as is suggested by O'Donoghue and Brown (2009).

A two-step cluster analysis with Euclidean as the distance measure and Schwartz's Bayesian criterion was used in the data package of IBM SPSS Statistics for Windows, Version 22.0 (Armonk, NY: IBM Corp.) to classify: (i). serve and return speed, and (ii) rally length (See Table 2). Former studies has discussed the tactical rationale of altering serve speed in first or second serve (Gillet et al., 2009; Hizan et al., 2011; O'Donoghue & Brown, 2008), and the result of the cluster analysis was consistent with that. The serve and return of serve was separately grouped into fast and slow speed. Meanwhile, rally length were pre-determined into short, medium and long length groups according to the routine procedure (O'Donoghue, 2013).

It was impossible to assess the reliability of the tracking data (serve speed) by Hawk-eye system, although being used in previous studies (Reid et al., 2016; Whiteside & Reid, 2016b). Therefore, two experienced performance analysts in tennis assessed the non-tracking notational (point sequence, count of first and second serve) and rally data (rally length) of two matches that were randomly selected. The minimum Cohen's *kappa* value for notational variables exceeded 0.90, demonstrating high inter-rater reliability (O'Donoghue, 2010). While for the continuous variables that are not included in the sequential data, intra-class coefficients (ICC) and standardized typical errors were calculated for inter-operators and operators vs. official data. Results of ICC ranged from 0.96 to 1 and standardized typical errors varied from 0.02 to 0.11, showing very high levels of reliability.

Table 6.2. Variables analyzed and correspondent sub-categories

Variable	Categories	M (SD)
<i>Situational Variables</i>		
Grand Slam	Australian Open, Roland Garros, Wimbledon and US Open	N/A
Year	2011, 2012, 2013, 2014, 2015, 2016	N/A
Game type	First round, second round, third round, fourth round, quarterfinals, semifinals and final	N/A
Set type	First set, second set, third set, fourth set and final set	N/A

Quality of opposition	Top10, top 11-20, top 21-50, top 51-100 and over 100.	N/A
Serving player	Novak and opponent	N/A
Serve number	First and second serve	N/A
Game status	Winning, losing, drawing, breaking opponent (when Novak returned), being broken (when Novak served)	N/A

Match Performance Variables

Serve and return categories for Novak Djokovic	First serve (km/h)	Fast speed (n = 7251)	188.8 (8.6)
		Slow speed (n = 1175)	154.7 (12.5)
	Second Serve (km/h)	Fast speed (n = 1879)	182.3 (13.9)
		Slow speed (n = 3016)	145.4 (10.0)
	Return first serve (km/h)	Fast speed (n = 6762)	190.9 (11.2)
		Slow speed (n = 2291)	155.0 (13.7)
	Return Second serve (km/h)	Fast speed (n = 2254)	178.8 (13.8)
		Slow speed (n = 3693)	145.0 (10.6)
	Double fault of Novak Djokovic (n = 263):		N/A
	Double fault of Opponent (n = 491):		N/A
Rally length	Short (n=4421)		2.2 (1.6)
	Medium (n=1568)		8.2 (1.9)
	Long (n=444)		16.7 (4.0)

N/A Denotes for the variable that is nominal.

6.2.3 Statistical Analysis

Crosstabs command were performed with the use of Pearson's Chi-square test to analyse the effects between the point outcome and situational variables related to Grand Slam, game type, set, quality of opposition, game status and point server. Effect sizes (ES) were calculated using the Cramer's V test and their interpretation was based on the following criteria: 0.10 = small effect, 0.30 = medium effect, and 0.50 = large effect (Volker, 2006).

To have a refined understanding of the previous analysis, two classification trees with the exhaustive CHAID (Chi-square Interaction Detection) algorithm were used to model the relationship among point outcome and situational variables. The first tree contained eight situational variables and serve performance while the second considered eight situational variables and rally performance. The algorithm is suitable for modelling nominal variables and has more precision because it merges categories until only two categories exist (Gómez, Battaglia, et al., 2015). The model was built under the following criteria: (i) the maximum number of iterations were 100; (ii) the minimum change in expected cell frequencies was

0.001; (iii) the significant values adjustment was done using the Bonferroni method; (iv) the tree based model has a maximum of 3 levels; (v) Grand Slam was forced to be the first variable as the influence of court surfaces on player's tactical and technical performance has already been highlighted (Gillet et al., 2009; O'Donoghue & Ingram, 2001). The risk of misclassification was calculated as a measure of the reliability of the model. The statistical analyses were done using the SPSS software (IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.). The level of significance was set at $p < 0.05$.

6.3 Results

Table 6.3 shows the distribution of Grand Slam, game type, quality of opposition, point server and serve number, which are related to point outcome (numbers and percentage). The variables related to game characteristics were shown to be significant and Novak won comparatively more points in Australian Open while the percentage of points won was decreasing gradually in every round as the game type changed from the first round (60.2%) to the final (51.1%). The quality of opposition was significantly related to point outcome ($p = 0.001$, $ES = 0.06$), greater success percentage was reached when playing against Top 51-100 players (60.2%), and lower success percentage occurred when competing with Top 10 players (52.6%). The results for variables related to serve exhibited that server and number of serve were also significantly related to point outcome ($p = 0.001$, $ES = 0.26$ and $p = 0.002$, $ES = 0.02$, respectively). Higher success was achieved when Novak was serving (69.4%) comparing to the situation when he returned the serve (43.1%). Whereas relatively more points were won when he or his opponent served a second serve (56.8%).

The sample distribution of game status, serve categories and rally length is presented in Table 4. The results demonstrated that game status had a significant relation with point outcome ($p = 0.001$, $ES = 0.12$), with higher point victory percentage when Novak was winning (62.5%) than when he was trying to break the opponent's service game (44.6%), and a relevantly greater success were shown when his service was being broken by the opponent (54.1%). Besides, the serve category was exhibited to be correlated significantly with the point outcome ($p = 0.001$, $ES = 0.36$). Serving fast in first and second serve enabled Novak to win more points (75.0% and 73.5% respectively) while he had greater success of winning points when returned slow second serve (52.7%) than retuning other serves. Finally, rally length was also significantly associated to point outcome ($p = 0.008$, $ES = 0.04$). A medium to long rally

helped him winning comparatively more points (58.8% and 58.1%, respectively) than short rally (54.5%).

Table 6.3. Frequency distribution (%) of point outcome related to situational variables (Crosstab Order: Pearson's Chi-square, degrees of freedom, significance value, expected frequency distribution, and effect size)

Situational Variables	Win		Lose		χ^2	df	P	EFD	ES
	N	%	N	%					
<i>Grand Slam</i>									
Australian Open	4542	56.53	3493	43.47	9.188	3	0.027*	2699.61	0.02
French Open	4201	55.00	3437	45.00					
Wimbledon	4326	54.66	3589	45.34					
US Open	3445	55.65	2642	44.35					
<i>Game Type</i>									
1st Round	2092	60.17	1385	39.83	103.481	6	0.001**	1542.07	0.06
2nd Round	2199	59.22	1514	40.78					
3rd Round	2334	56.89	1769	43.11					
4th Round	2640	56.06	2069	43.94					
Quarterfinal	2261	55.13	1840	44.87					
Semifinal	2748	53	2437	47					
Final	2240	51.06	2147	48.94					
<i>Quality of Opposition</i>									
Top 10	5816	52.58	5245	47.42	97.426	4	0.001**	984.14.	0.06
Top 11-20	2545	55.18	2067	44.82					
Top 21-50	3876	56.48	2987	43.52					
Top 51-100	2960	60.16	1960	39.84					
Over 100	1317	59.35	902	40.65					
<i>Point Server</i>									
Novak Djokovic	9843	69.4	4341	30.6	2080.056	1	0.001**	6290.67	0.26
Opponent	6671	43.06	8820	56.94					
<i>Serve Number</i>									
1st Serve	9964	54.93	8174	45.07	9.668	1	0.002**	5116.71	0.02
2nd Serve	6550	56.77	4987	43.23					

* P < 0.05; ** P < 0.01; EFD = expected frequency distribution.

Table 6.4. Frequency distribution (%) of point outcome related to situational variables (Crosstab Order: Pearson's Chi-square, degrees of freedom, significance value, expected frequency distribution, and effect size)

Situational Variables		Win		Lose		χ^2	df	P	EFD	ES
		N	%	N	%					
Game Status										
Winning		7416	62.50	4450	37.50	413.584	4	0.001**	280.29	0.12
Losing		4822	52.20	4415	47.80					
Drawing		3120	51.01	2996	48.99					
Being Broken		342	54.11	290	45.89					
Breaking Opponent		814	44.63	1010	55.37					
Serve Category										
First Serve	Fast	5435	74.96	1816	25.04	3844.983	9	0.001**	116.64	0.36
	Slow	1137	64.06	638	35.94					
Second Serve	Fast	1381	73.50	498	26.50					
	Slow	1890	62.67	1126	37.33					
Return First Serve	Fast	2256	33.36	4506	66.64					
	Slow	1091	47.62	1200	52.38					
Return Second Serve	Fast	888	39.40	1366	60.60					
	Slow	1945	52.67	1748	47.33					
Double Fault of Djokovic		0	0	263	100					
Double Fault of Opponent		491	100	0	0					
Rally Length										
Short		2409	54.49	2012	45.51	9.761	2	0.008**	196.29	0.04
Medium		922	58.80	646	41.20					
Long		258	58.11	186	41.89					

** P < 0.01; EFD = expected frequency distribution.

Figure 6.1 and 6.2 show the first exhaustive CHAID classification tree model for Novak Djokovic. It was demonstrated that game location, serve category contributed strongly to the tree while game type, quality of opponent, game status also had an influence.

In Australian Open (AO) and US Open (US), Novak won 56.6% of points, compared to 54.8% in Roland Garros (RG) and Wimbledon (W) (node 1 and 2). 75.8% and 73.0% of points were won when he served fast first and second serves in AO and US, while 74.3% and 73.9% were achieved for the ones in RG and W. In contrast, serving slow first and second serve made him win a lower 65.2% and 62.8% in AO and US, and 62.8% and 62.5% in RG

and W. In terms of return, similar results of approximately 33% and 40% of points could be obtained in all slams while Novak returned first and second fast serve. However, he could win 50.2% of points when returning slow second serve in AO and US (node 9), compared with 44.4% in RG and W (node 19). Lower winning percentage was obtained (node 25 and 37) when serving fast first serve in the final of all slams, compared to the previous rounds. In AO and US, serving slow second serves to Top 10 players achieved a lower 58% of points won (node 28), while in RG and W, this performance was influenced by game status in that Novak won just 45.2% of points when facing breakpoints (node 39). Similarly, under the drawing game status in AO and US, Novak could win only 46.2% of points while returning slow second serves (node 33).

Figure 6.3 presents the second tree model that considered the rally performance. The results showed that point server and rally length were two variables that strongly contributed to the model. In all slams, Novak won nearly 70% points in his service game (node 3 and 5) more than the 45% of return game (node 4 and 6). Finishing the points in short and long rally (node 12: 64.9% and node 11: 55.2%) during own service game helped him to win more points than medium rally, but it was not the case for Wimbledon, where only in medium rally could he win more points (node 13).

Figure 6.1. Analysis by classification tree One of point-level performance for Novak Djokovic in Australian Open & US Open (including serve speed category)

Figure 6.2. Analysis by classification tree One of point-level performance for Novak Djokovic in Roland Garros & Wimbledon (including serve speed variable)

Figure 6.3. Analysis by classification tree Two of point-level performance for Novak Djokovic in four Grand Slams (including rally length category)

6.4 Discussion

This study aimed to analyze the point-level performance of individual professional tennis player in different Grand Slams and used the classification tree model to achieve that. It provides a novel approach to assess the performance of player in relation to situational variables.

Results showed that the performance of Novak Djokovic was clearly influenced by game location, game type, quality of opposition, game status, point server, serve number, serve and return speed and rally length.

In tennis, the effect of game location may consist in the following “venue variations”: temperature, humidity (Morante & Brotherhood, 2007), court surfaces (O'Donoghue & Ingram, 2001), surface abrasion, ball types and the distance between court and the audience (Dube & Tatz, 1991; McGarry et al., 2013). They are factors that players and coaches need to consider when planning match strategies. Grand Slam tournaments symbolize the maximum level of professional tennis (Cross & Pollard, 2009) and each slam is played in different court surface, which was speculated as the major influencing factor on player's match-play (O'Donoghue & Ingram, 2001; Reid et al., 2016). Therefore, it is essential to have better understanding of its effect on player's performance. Partially in line with the study of Ma et al. (2013) that male players were more likely to win a match in the US Open and Wimbledon than in the French Open, our findings showed that Novak had more possibility to win Australian Open and US Open than Roland Garros and Wimbledon. On the one hand, this occurred probably due to the fact that in slow surface court tournament like Roland Garros, it was easier to break to the server's game and players tend to play more aggressively in the game. While in fast surface court of Wimbledon, serving players has greater advantage and it is hard to break their game, so that most players preferred to use safer playing style in fast court (Gillet et al., 2009; Tudor, Zecic, & Matkovic, 2014). On the other hand, it was also indicated that Novak Djokovic had a better familiarity with hard court surface. Further research is recommended to develop upon the current study to assess the inter-player performance variation in different court surfaces.

In the second place, the influence of game type on player's performance should be highlighted in Grand Slams due to the fact that male player has to compete with a draw of 128 players within two weeks. Each match is played in the best of five sets. These factors condition the performance of player in a way that they have to adjust the strategy and tactics on a match-to-

match basis to ensure the success as well as to avoid an early depletion of fitness and psychological state (Goossens et al., 2015; O'Donoghue, 2009). Results by Goossens et al. (2015) showed that in Grand Slam, the previous match could have an effect on the next round for players and the possibility to win the next round will decrease if he played more games in previous match. Consistent with that, our findings indicated that the point performance of Novak Djokovic was affected by game type. As he proceeded into the next round, the winning percentage of points decreased significantly. Additionally, results of the classification tree demonstrated that the efficiency of serve performance was conditioned by game type in all four slams, especially when Novak Djokovic served fast first and second serve. Serving fast in previous rounds could help him win more points while it was not the case in the last rounds, especially in the final.

Obviously, this phenomena should be examined together with the effect of quality of opposition because in eliminatory tournament, the opponents of the next round are expected to be higher-ranked rivals with better competitive skills (Reid et al., 2010; Reid & Morris, 2013). Previous research that addressed the effect of opponent's strength have identified that the serve and return performance of male players was influenced by quality of opposition (O'Donoghue, 2009; O'Donoghue & Cullinane, 2011). The present study provides more evidence on this issue. It was not surprising that higher winning percentage was achieved when Novak Djokovic competed with players ranked over Top 50 and then it descended when he was against Top 50, especially Top 10 players. However, the results from the classification tree offered more details on where and how the serve and return performance of Novak Djokovic was affected by different opponents. For instance, he turned out to be have more success when playing with lower ranked players over Top 20 when serving slow first serve in Roland Garros and Wimbledon, whereas a higher chance of winning was obtained when returning fast first serves by over Top 50 opponents Australian Open and US Open.

In elite tennis matches, winning the previous point particularly have a positive effect on winning the current point, and at crucial points the server has a disadvantage of winning because of the psychological pressure in stressful situations (Filipčič, Filipčič, & Berendijaš, 2008; Klaassen & Magnus, 2001), which suggested a score-line effect (Gómez, Lago, et al., 2013; O'Donoghue, 2012) within tennis game. In accordance with those previous findings, we used within-game status to assess the performance of Novak Djokovic. Results revealed that he had highest winning chance when he was winning, while he was losing, the percentage

decreased significantly. Furthermore, the first classification tree showed that Novak's slow second serve was influenced by game status in Roland and Wimbledon, where he won more points when he was winning, losing or drawing but less than half of all points when facing breakpoint. When he returned second slow serve in AO and UO, he had more success when winning, losing or breaking opponents than when he was drawing. This suggested that the positive effect of winning the previous point on winning the current one. However, future research should consider player's performance in critical points, such as set points, tie-break points and match points.

Furthermore, the results revealed the influence of rally length on performance. And in agreement with the previous research in that server has an advantage in points lasted to 4 shots (O'Donoghue & Brown, 2008), our findings that short rally and long rally in the service game of Novak Djokovic helped him to win more points than medium rally. In return game, he could win more points in medium to long rallies.

6.5 Conclusion

The results of the current study demonstrate the usefulness of the exhaustive CHAID tree model to interpret the match performance of tennis player and it allows to develop individual performance profiles. Although the court surfaces had a major effect on player's performance, our findings suggested that this importance should be considered together with others performance indicators such as game type, quality of opposition, match status. From a practical perspective, efforts could be made to improve Novak Djokovic's return performance, especially the second serve return when he was in critical game status like drawing, having breaking opportunity or being broken, and furthermore, to attempt playing aggressively to win return points in short rally. These results will give insightful ideas for coaches and players when planning the game strategy, helping to modify training contents to resemble real match situations where player learn to be more adaptable to play in different game status. Furthermore, it is possible that performance analysts in tennis use this method to model multi-player performance for specific match situations or in a time-wise manner.

Chapter Seven – Conclusion, Limitations and Future Research Lines

7.1 Main conclusions and outcomes of the thesis

(i) By analyzing set-level data gathered from 2017 Australian Open and 2016 US Open men's singles main-draw matches, it is shown that due to the influence of match fatigue, players experienced a decrease in serve, net and running related performance and an increase in return and winner related performance within the match, indicating an adaption of match tactics and pacing strategy. Besides, winning players performed consistently in almost all performance related variables but they would sacrifice the stability in second serve, winner, and physical performance to be more aggressive and to dominate the match.

(ii) Female tennis players during four Grand Slams (2014-2017) performed better in serve and return when competing on fast court surfaces and they preferred to stay in the baseline to finish the points, although owning good efficiency at net. Additionally, ranking-based relative quality was correlated with all technical-tactical performance variables related to serve and return, break point, net point and efficiency except for physical performance. These results evidenced the joint effect of match location and opponents on players' match behaviors. Finally, the use of the percentage evaluation score (%ES) to establish typical performance profiles for different groups of players demonstrated that it is a more reasonable profiling technique as it accounted for the influence of opponents' strength.

(iii) Playing experience and ranking-based relative quality are good mediators of players' characteristics and match performance. Different male player categories during four Grand Slams (2014-2017) were clearly differentiated by match performance related to serve and return, break points, net, rally and physical movement. High-experienced Higher RQ players performed overall better than other players, especially in second serve and return, winner strokes and breaking opportunities. Experienced Higher RQ players owned faster serve speed, more aces, return winners and unforced errors, which indicated that they played with more aggressiveness.

(iv) The anthropometric (height and weight) and individual features (handedness, backhand style and experience) of players were used to automatically classify players into four player groups. Big-sized Righty Two-handed players and Medium Righty One-handed players

outperformed their counterparts in Australian Open, Wimbledon and US Open. Although in Roland Garros, Lefty Two-handed players established their advantage thanks to the slow court surface, the findings suggested that their advantage is over-represented at elite level. Moreover, Lefty Two-handed Players were the most homogenous groups in performance. Big-sized Righty Two-handed players and Medium-sized Righty One-handed players were more heterogeneous, revealing how different types of players varied match strategy and tactics considering contextual constraints.

(v) The non-linear relationships between point outcome of studied player during 2011-2016 Grand Slams and contextual variables and performance variables were preliminarily modelled. Game location was shown to have a major effect on the winning percentage of points. Meanwhile results revealed that the winning chances decreased when playing in the final rounds and against higher-ranked players. Furthermore, when drawing, losing or facing breakpoint in the game, the performance of studied player descended as well. While finishing points in short and long rally would contribute to higher winning percentage.

7.2 Limitations

The limitations associated with the current thesis are summarized as follows:

- (i) Due to the availability of set-level data, the first study only considered male players' performance in hard-court Grand Slams and findings could not be generalized into other events played on clay and grass court, or into female players. Besides, detailed technical-tactical preferences of players (serve direction, groundstroke direction/depth, net point, movement pattern) in relation to specific match/set/game score-line were lacked.
- (ii) The study three to five used the match statistics of just two seasons of Roland Garros and Wimbledon which was an issue caused by the accessibility of archive data from official websites. The comparatively less data than the other two tournaments (Australian Open and US Open) might lead to possible bias in results when interpreting the match performance.
- (iii) The study four and five failed to consider female players and to include more personal characteristics of players, such as recent match result, record of injury or withdrawal, physical status, that help to understand the variation in performance.

(iv) Because of being a preliminary and case study, the study six only utilized the point-level statistics of one player despite that the sampled data was longitudinal. Therefore, the representativeness of results and findings were somehow limited.

(v) Overall, all studies focused majorly on technical-tactical performance although some physical variables such as serve speed and distance covered physical were included. While the following aspects have not been considered: physical variables such as movement speeds, distance covered at different speed zones, movements directions, numbers of acceleration and deceleration; physiological variables such as heart rate, rating of perceived effort; and most importantly, the variables related to inter-player as well as player-ball positional patterns.

7.3 Perspectives for future research

(i) Future studies are expected to go deeper to address the interactive effects of contextual variables studied in the thesis on match performance and explore more variables related to specific game situations. Meanwhile the current findings should be further extended to players of other categories to improve the understanding of within- and between group differences in performance.

(ii) Investigation of player's overall performance (technical-tactical, physical and physiological) in a longitudinal and consecutive manner (under different time intervals: weeks, months, years) is recommended so as to measure continuous variation in players' match behaviors, building a more complex performance profiles and providing more evidence-based insights for individualized training.

(iii) More research in spatiotemporal relationships between players and ball during the competition should be encouraged to dig out strategical and tactical patterns that are meaningful and representative.

(iv) It is vital that future study should emphasize the nano-level status of tennis players, namely the psychological status and mental toughness of players, in order to improve their decision-making from the ultimate level.

(v) Key performance indicators are expected to be refined in future investigations and this

process would be likely undertaken with the existence of massive data generated on and off the tennis courts. Therefore, researchers are encouraged to integrate different statistical approaches from machine learning and artificial intelligence that are more efficient in dealing with these data in a timely manner.

Chapter Eight – Practical Application

The findings of this thesis are novel and provide pertinent evidence-based information on technical-tactical and physical performance of elite tennis players during the highest level of the sport. Therefore, knowledge gained from the information could be used to inform training and match preparation, and this chapter is served to illustrate several practical applications that might be used in training. They are listed below according to the findings of all five studies.

1. Considering the decreases in serve and net performance for male players and higher variation of winning players in second serve, winners and running performance, it is suggested that: i) fitness coaches should prepare a more representative physical conditioning program that improve players' ability of fast recovery from a previous intensive work load. For example, some tennis-specific "high-intensity interval training" sessions could be prioritized and designed to improve the endurance and the heart rate recovery rate of players during prolonged training or match time; ii) technical coaches should help players to optimize their net approaches and defensive tactics (when the opponents approach the net) when experiencing match fatigue. They could use the end of their training sessions (where logically players are tired) to deliberately practice either net approaches (when serving, when returning, when playing under critical points, etc.) or defensive tactics (lobs, passing shots, drop-shots) and set up some punishments on players' failures in order to imitate the game-like pressure; iii) while contrary to the last point, coaches should encourage players during hitting drills or practice match to try to be more aggressive, and to take the initiative to move their opponents. For example, players could be awarded with "scores" every time their make offensive decisions, which are clearly confirmed by coaches. Afterwards, players could automatically win a point during once their "scores" reach certain sum. This will motivate players to always play positively and try to dominate the game.
2. Findings of the second study suggested that female players showed different match behaviors on distinct court surfaces and against different levels of opponents. As female players prefer to stay in the baseline, coaches are recommended to help female players to finish the point with shorter rally, which requires more angle-opening and variation of ball placements at the beginning of points. Therefore, specific drills could be designed

for the following three situations: i) rally, ii) serve + rally and iii) return + rally. In these drills, players are required to finish the points within less than 5 rallies, any extra rally could be accumulated for some “punishments” in order to help players dominate the point with the initial stokes. On the other hand, the profiling method used in this study allows for particularly robust monitoring of player’s performance during tournament. It could be, therefore, applicable for other player categories (e.g. junior) and for long-term tracking of performance.

3. In light of the findings from the third and fourth study, it is suggested that during the development of junior players, practitioners could use clustering approach to differentiate players based on their characteristics. Afterwards, mixing these player groups into training may help to improve their adaptability to different opponents. This could be realized by from two aspects: i) mixed-training of players with different styles; ii) deliberately imitating the playing styles of other players. Throughout this approach, it is expected that players could not only understand better how to compete against different types of opponents, but also develop a more stable performance, enhancing their weakness.
4. The last study provided a noble approach for coaches to inspect player’s performance. Analyzing point-level performance help them to prescribe evidence-based training doses that intend to improve the weak spots of players. Quoting the findings from Novak Djokovic, it would be beneficial that he makes efforts with his coaching teams on improving his return performance during critical point (drawing, break points, etc.). Stimulated match status is suggested to use to help develop representative returning tactics.
5. Finally, match preparation and recovery planning should be prescribed with the understanding of match workloads for players with different levels. Therefore, sex-specific and individual-specific conditioning is optimal for players’ progression.

References

- 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. doi:10.2478/v10237-012-0006-6
- Batterham, A. M., & Hopkins, W. G. (2006). Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance*, 1(1), 50-57.
- Breznik, K. (2013). On the Gender Effects of Handedness in Professional Tennis. *Journal of Sports Science & Medicine*, 12(2), 346-353.
- Butterworth, A., O'Donoghue, P., & Cropley, B. (2013). Performance profiling in sports coaching: a review. *International Journal of Performance Analysis in Sport*, 13(3), 572-593.
- Carboch, J., Vejvodova, K., & Suss, V. (2016). Analysis of errors made by line umpires on ATP tournaments. *International Journal of Performance Analysis in Sport*, 16(1), 264-275. doi:10.1080/24748668.2016.11868885
- Carvalho, J., Araujo, D., Travassos, B., Esteves, P., Pessanha, L., Pereira, F., & Davids, K. (2013). Dynamics of players' relative positioning during baseline rallies in tennis. *Journal of Sports Sciences*, 31(14), 1596-1605. doi:10.1080/02640414.2013.792944
- Carvalho, J., Araújo, D., Travassos, B., Fernandes, O., Pereira, F., & Davids, K. (2014). Interpersonal dynamics in baseline rallies in tennis. *International Journal of Sports Science and Coaching*, 9(5), 1043-1056. doi:10.1260/1747-9541.9.5.1043
- Chow, J. W., Knudson, D. V., Tillman, M. D., & Andrew, D. P. (2007). Pre-and post-impact muscle activation in the tennis volley: effects of ball speed, ball size and side of the body. *British Journal of Sports Medicine*, 41(11), 754-759.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences (2nd Ed.)*. . Hillsdale, NJ: Lawrence Earlbaum Associates.
- Corral, J. d. (2009). Competitive Balance and Match Uncertainty in Grand-Slam Tennis. *Journal of Sports Economics*, 10(6), 563-581. doi:10.1177/1527002509334650
- Cowden, R. G. (2016). Competitive Performance Correlates of Mental Toughness in Tennis. *Perceptual and motor skills*, 123(1), 341-360. doi:10.1177/0031512516659902
- Crespo, M., Reid, M. M., & Miley, D. (2004). Tennis: Applied Examples of a Game-based Teaching Approach. *Strategies*, 17(4), 27-30. doi:10.1080/08924562.2004.10591100
- Cross, R., & Pollard, G. (2009). Grand Slam men's singles tennis 1991-2009 Serve speeds and other related data. *COACHING & SPORT SCIENCE REVIEW*, 16(49), 8-10.

- 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), 1-19. doi:10.1080/24748668.2017.1399325
- Davey, P. R., Thorpe, R. D., & Williams, C. (2002). Fatigue decreases skilled tennis performance. *Journal of Sports Sciences*, 20(4), 311-318. doi:10.1080/026404102753576080
- De Bosscher, V., De Knop, P., Van Bottenburg, M., & Shibli, S. (2006). A Conceptual Framework for Analysing Sports Policy Factors Leading to International Sporting Success. *European Sport Management Quarterly*, 6(2), 185-215. doi:10.1080/16184740600955087
- Didierjean, A., & Marméche, E. (2005). Anticipatory representation of visual basketball scenes by novice and expert players. *Visual Cognition*, 12(2), 265-283.
- Dube, S. K., & Tatz, S. J. (1991). Audience Effects in Tennis Performance. *Perceptual and motor skills*, 73(3), 844-846. doi:10.2466/pms.1991.73.3.844
- Fernandez-Fernandez, J., Mendez-Villanueva, A., Fernandez-Garcia, B., & Terrados, N. (2007). Match activity and physiological responses during a junior female singles tennis tournament. *British Journal of Sports Medicine*, 41(11), 6. doi:10.1136/bjism.2007.036210
- Fernandez-Fernandez, J., Sanz-Rivas, D., & Mendez-Villanueva, A. (2009). A Review of the Activity Profile and Physiological Demands of Tennis Match Play. *Strength & Conditioning Journal*, 31(4), 15-26. doi:10.1519/SSC.0b013e3181ada1cb
- Fernandez-Fernandez, J., Sanz-Rivas, D., Sanchez-Muñoz, C., Pluim, B. M., Tiemessen, I., & Mendez-Villanueva, A. (2009). A Comparison of the Activity Profile and Physiological Demands Between Advanced and Recreational Veteran Tennis Players. *The Journal of Strength & Conditioning Research*, 23(2), 604-610. doi:10.1519/JSC.0b013e318194208a
- Filaire, E., Alix, D., Ferrand, C., & Verger, M. (2009). Psychophysiological stress in tennis players during the first single match of a tournament. *Psychoneuroendocrinology*, 34(1), 150-157. doi:<http://dx.doi.org/10.1016/j.psyneuen.2008.08.022>
- Filipic, A., Leskosek, B., & Filipic, T. (2017). Split-Step Timing of Professional and Junior Tennis Players. *Journal of Human Kinetics*, 55, 97-105. doi:10.1515/hukin-2017-0009

- 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. doi:10.7752/jpes.2015.04102
- Filipčič, T., Filipčič, A., & Berendijaš, T. (2008). Comparison of game characteristics of male and female tennis players at Roland Garros 2005. *Acta Universitatis Palackianae Olomucensis. Gymnica*, 38(3), 21-28.
- Gal, O. (2003). fit_ellipse. Retrieved from <https://www.mathworks.com/matlabcentral/fileexchange/3215-fit-ellipse>
- Galé-Ansodi, C., Castellano, J., & Usabiaga, O. (2016). Effects of different surfaces in time-motion characteristics in youth elite tennis players. *International Journal of Performance Analysis in Sport*, 16(3), 860-870.
- Galé-Ansodi, C., Castellano, J., & Usabiaga, O. (2017). Physical profile of young tennis players in the tennis match-play using global positioning systems. *Journal of Physical Education and Sport*, 17(2), 826.
- Genevois, C., Reid, M., Rogowski, I., & Crespo, M. (2015). Performance Factors Related to the Different Tennis Backhand Groundstrokes: A Review. *Journal of Sports Science & Medicine*, 14(1), 194-202.
- Gescheit, D. T., Cormack, S. J., Reid, M., & Duffield, R. (2015). Consecutive Days of Prolonged Tennis Match Play: Performance, Physical, and Perceptual Responses in Trained Players. *International Journal of Sports Physiology and Performance*, 10(7), 913-920. doi:10.1123/ijsp.2014-0329
- Gescheit, D. T., Duffield, R., Skein, M., Brydon, N., Cormack, S. J., & Reid, M. (2016). Effects of consecutive days of match play on technical performance in tennis. *Journal of Sports Sciences*, 1-7. doi:10.1080/02640414.2016.1244352
- Gillet, E., Leroy, D., Thouvarecq, R., Megrot, F., & Stein, J. F. (2010). Movement-Production Strategy in Tennis: A Case Study. *Journal of Strength and Conditioning Research*, 24(7), 1942-1947. doi:10.1519/JSC.0b013e3181dc4622
- Gillet, E., Leroy, D., Thouvarecq, R., & Stein, J. F. (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. doi:10.1519/JSC.0b013e31818efe29
- Girard, O., Micallef, J.-p., & Millet, G. P. (2005). Lower-limb activity during the power serve in tennis: effects of performance level. *Medicine and Science in Sports and Exercise*, 37(6), 1021-1029.

- Gomes, R. V., Coutts, A. J., Viveiros, L., & Aoki, M. S. (2011). Physiological demands of match-play in elite tennis: A case study. *European Journal of Sport Science*, 11(2), 105-109. doi:10.1080/17461391.2010.487118
- Gómez, M.-Á., Silva, R., Lorenzo, A., Kreivyte, R., & Sampaio, J. (2017). Exploring the effects of substituting basketball players in high-level teams. *Journal of Sports Sciences*, 35(3), 247-254. doi:10.1080/02640414.2016.1161217
- Gómez, M. Á., Battaglia, O., Lorenzo, A., Lorenzo, J., Jiménez, S., & Sampaio, J. (2015). Effectiveness during ball screens in elite basketball games. *Journal of Sports Sciences*, 33(17), 1844-1852.
- Gómez, M. Á., Lago, C., & Pollard, R. (2013). Situational variables. In T. McGarry, P. O'Donoghue, & J. Sampaio (Eds.), *Routledge handbook of sports performance analysis* (pp. 259-269). London: Routledge.
- Gómez, M. Á., Lorenzo, A., Ibañez, S.-J., & Sampaio, J. (2013). Ball possession effectiveness in men's and women's elite basketball according to situational variables in different game periods. *Journal of Sports Sciences*, 31(14), 1578-1587. doi:10.1080/02640414.2013.792942
- Gómez, M. Á., Lorenzo, A., Sampaio, J., Ibañez, S., & Ortega, E. (2008). Game-related statistics that discriminated winning and losing teams from the Spanish men's professional basketball teams. *Collegium Antropologicum*, 32(2), 451-456.
- Gómez, M. Á., Moral, J., & Lago-Peñas, C. (2015). Multivariate analysis of ball possessions effectiveness in elite futsal. *Journal of Sports Sciences*, 33(20), 2173-2181. doi:10.1080/02640414.2015.1075168
- Gonçalves, B., Marcelino, R., Torres-Ronda, L., Torrents, C., & Sampaio, J. (2016). Effects of emphasising opposition and cooperation on collective movement behaviour during football small-sided games. *J Sports Sci*, 34(14), 1346-1354. doi:10.1080/02640414.2016.1143111
- Goossens, D. R., Kempeneers, J., Koning, R. H., & Spieksma, F. C. R. (2015). Winning in straight sets helps in Grand Slam tennis. *International Journal of Performance Analysis in Sport*, 15(3), 1007-1021.
- Hagemann, N. (2009). The advantage of being left-handed in interactive sports. *Attention, Perception, & Psychophysics*, 71(7), 1641-1648.
- 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.
doi:10.1080/24748668.2011.11868556
- 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.
doi:10.1260/1747-9541.10.1.87
- Hizan, H., Whipp, P., Reid, M., & Wheat, J. (2014). A comparative analysis of the spatial distributions of the serve return. *International Journal of Performance Analysis in Sport*, 14(3), 884-893. doi:10.1080/24748668.2014.11868765
- Hizan, H., Whipp, P. R., & Reid, M. (2010). Validation of match notation (A coding system) in tennis. *Journal of Quantitative Analysis in Sports*, 6(3).
- Hopkins, W., Marshall, S., Batterham, A., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine+ Science in Sports+ Exercise*, 41(1), 3.
- Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, 30(1), 1-15.
- Hopkins, W. G. (2006). A Spreadsheet for Combining Outcomes from Several Subject Groups. Retrieved from sportsci.org/2006/wghcom.htm
- Hopkins, W. G. (2007). A spreadsheet for deriving a confidence interval, mechanistic inference and clinical inference from a P value. *Sportscience*, 11, 16-21.
- Hoppe, M. W., Baumgart, C., Bornefeld, J., Sperlich, B., Freiwald, J., & Holmberg, H.-C. (2014). Running Activity Profile of Adolescent Tennis Players during Match Play. *Pediatric Exercise Science*, 26(3), 281-290. doi:10.1123/pes.2013-0195
- Hornery, D. J., Farrow, D., Mujika, I., & Young, W. (2007). An integrated physiological and performance profile of professional tennis. *British Journal of Sports Medicine*, 41(8), 531-536. doi:10.1136/bjsm.2006.031351
- Hughes, M., & Franks, I. (2004). *Notational analysis of sport: Systems for better coaching and performance in sport*: Psychology Press.
- Hughes, M., & Franks, I. (2007). *The Essentials of Performance Analysis: An Introduction*. London: Taylor & Francis.
- Hughes, M., & Franks, I. (2015). *Essentials of performance analysis in sport*. London: Routledge.
- Hughes, M., & Meyers, R. (2005). Movement patterns in elite men's singles tennis. *International Journal of Performance Analysis in Sport*, 5(2), 110-134.

- ITF. (2005). Line Calling. Retrieved from <http://www.itftennis.com/technical/technical-centre/line-calling.aspx>
- ITF. (2016). ITF Approved tennis balls, classified surfaces & recognised courts 2016. *ITF Licensing (UK) Ltd.*
- Iwan, G., Colin, E., & Neil, G. (2005). Tracking the flight of a spinning football in three dimensions. *Measurement Science and Technology*, 16(10), 2056. doi:10.1088/0957-0233/16/10/022.
- Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis* (Vol. 344): John Wiley & Sons.
- King, H. A., & Baker, J. A. (1979). Statistical analysis of service and match-play strategies in tennis. *Canadian Journal of Applied Sport Sciences*.
- Klaassen, F. J., & Magnus, J. R. (2001). Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model. *Journal of the American Statistical Association*, 96(454), 500-509.
- 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. doi:10.1080/17461391.2011.577239
- Kovacs, M. S. (2006). Hydration and temperature in tennis—a practical review. *Journal of Sports Science and Medicine*, 5, 1-9.
- Kovalchik, S. (2015). Grand Slams are short-changing women's tennis. *Significance*, 12(5), 12-17. doi:10.1111/j.1740-9713.2015.00852.x
- Kovalchik, S., & Reid, M. (2018). A shot taxonomy in the era of tracking data in professional tennis. *Journal of Sports Sciences*, 1-9. doi:10.1080/02640414.2018.1438094
- 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 and Medicine*, 16, 489-497.
- Lago-Peñas, C. (2012). The Role of Situational Variables in Analysing Physical Performance in Soccer. *Journal of Human Kinetics*, 35, 89-95. doi:10.2478/v10078-012-0082-9
- Lago-Peñas, C., Gómez-Ruano, M., Megías-Navarro, D., & Pollard, R. (2016). Home advantage in football: Examining the effect of scoring first on match outcome in the five major European leagues. *International Journal of Performance Analysis in Sport*, 16(2), 411-421.

- Laird, P., & Waters, L. (2008). Eyewitness Recollection of Sport Coaches. *International Journal of Performance Analysis in Sport*, 8(1), 76-84.
- Leech, N. L., Barrett, K. C., & Morgan, G. A. (2014). Logistic Regression and Discriminant Analysis. In N. L. Leech, K. C. Barrett, & G. A. Morgan (Eds.), *IBM SPSS for intermediate statistics: Use and interpretation* (pp. 167-187): Routledge.
- Leite, N., & Sampaio, J. (2010). Early sport involvement in young Portuguese basketball players. *Perceptual and motor skills*, 111(3), 669-680.
- Liu, H., Hopkins, W. G., & Gomez, M. A. (2016). Modelling relationships between match events and match outcome in elite football. *Eur J Sport Sci*, 16(5), 516-525. doi:10.1080/17461391.2015.1042527
- 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., & Strauss, B. (2010). Automated processes in tennis: Do left-handed players benefit from the tactical preferences of their opponents? *Journal of Sports Sciences*, 28(4), 435-443. doi:10.1080/02640410903536459
- Loffing, F., Hagemann, N., & Strauss, B. (2012). Left-Handedness in Professional and Amateur Tennis. *PLOS ONE*, 7(11), e49325. doi:10.1371/journal.pone.0049325
- Lyons, M., Al-Nakeeb, Y., & Nevill, A. (2006). The impact of moderate and high intensity total body fatigue on passing accuracy in expert and novice basketball players. *Journal of Sports Science & Medicine*, 5(2), 215.
- Ma, S. M., Liu, C. C., Tan, Y., & Ma, S. C. (2013). Winning matches in Grand Slam men's singles: An analysis of player performance-related variables from 1991 to 2008. *Journal of Sports Sciences*, 31(11), 1147-1155. doi:10.1080/02640414.2013.775472
- Maquirriain, J., Baglione, R., & Cardey, M. (2016). Male professional tennis players maintain constant serve speed and accuracy over long matches on grass courts. *European Journal of Sport Science*, 16(7), 845-849. doi:10.1080/17461391.2016.1156163
- Martin-Lorente, E., Campos, J., & Crespo, M. (2017). The inside out forehand as a tactical pattern in men's professional tennis. *International Journal of Performance Analysis in Sport*, 1-13. doi:10.1080/24748668.2017.1349528
- Martinez-Gallego, R., Guzman, J. F., James, N., Pers, J., Ramon-Llin, J., & Vuckovic, G. (2013). Movement Characteristics of Elite Tennis Players on Hard Courts with

- Respect to the Direction of Ground Strokes. *Journal of Sports Science and Medicine*, 12(2), 275-281.
- McGarry, T., O'Donoghue, P., & Sampaio, J. (2013). *Routledge handbook of sports performance analysis*: Routledge.
- 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. doi:10.1002/sam.11316
- Morante, S. M., & Brotherhood, J. R. (2007). Air temperature and physiological and subjective responses during competitive singles tennis. *British Journal of Sports Medicine*, 41(11), 773-778.
- Moss, B., & O'Donoghue, P. (2015). Momentum in US Open men's singles tennis. *International Journal of Performance Analysis in Sport*, 15(3), 884-896.
- Muhamad, T. A., Rashid, A. A., Razak, M. R. A., & Salamuddin, N. (2011). A comparative study of backhand strokes in tennis among national tennis players in Malaysia. *Procedia-Social and Behavioral Sciences*, 15, 3495-3499.
- Murray, N. P., & Hunfalvay, M. (2017). A comparison of visual search strategies of elite and non-elite tennis players through cluster analysis. *Journal of Sports Sciences*, 35(3), 241-246. doi:10.1080/02640414.2016.1161215
- 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.
- O'Donoghue, P. (2005). Normative Profiles of Sports Performance. *International Journal of Performance Analysis in Sport*, 5(1), 104-119.
- O'Donoghue, P. (2009). Interacting Performances Theory. *International Journal of Performance Analysis in Sport*, 9(1), 26-46.
- O'Donoghue, P. (2010). *Research methods for sports performance analysis*. London: Routledge.
- O'Donoghue, P. (2012). Break points in Grand Slam men's singles tennis. *International Journal of Performance Analysis in Sport*, 12(1), 156-165.
- O'Donoghue, P. (2013). *Statistics for sport and exercise studies: an introduction*: Routledge.
- O'Donoghue, P. (2014). *An introduction to performance analysis of sport*: Routledge.
- O'Donoghue, P., & Brown, E. (2009). Sequences of service points and the misperception of momentum in elite tennis. *International Journal of Performance Analysis in Sport*, 9(1), 113-127.

- O'Donoghue, P., & Cullinane, A. (2011). A regression-based approach to interpreting sports performance. *International Journal of Performance Analysis in Sport*, 11(2), 295-307.
- O'Donoghue, P., & Ingram, B. (2001). A notational analysis of elite tennis strategy. *Journal of Sports Sciences*, 19(2), 107-115. doi:10.1080/026404101300036299
- O'Donoghue, P., & Mayes, A. (2013). Performance Analysis, Feedback and Communication in Coaching. In T. McGarry, P. O'Donoghue, & J. Sampaio (Eds.), *Routledge handbook of sports performance analysis* (pp. 155-164). London: Routledge.
- O'Donoghue, P. G., & Brown, E. (2008). The importance of service in Grand Slam singles tennis. *International Journal of Performance Analysis in Sport*, 8(3), 70-78.
- O'Donoghue, P. (2013). Sports performance profiling. In T. McGarry, P. O'Donoghue, & J. Sampaio (Eds.), *Routledge handbook of sports performance analysis* (pp. 127-139). London: Routledge.
- O'Donoghue, P., Girard, O., & Reid, M. (2013). Racket Sports. In T. McGarry, P. O'Donoghue, & J. Sampaio (Eds.), *Routledge handbook of sports performance analysis* (pp. 404-414). London: Routledge.
- Ofoghi, B., Zeleznikow, J., MacMahon, C., & Raab, M. (2013). Data mining in elite sports: a review and a framework. *Measurement in Physical Education and Exercise Science*, 17(3), 171-186.
- Ojala, T., & Häkkinen, K. (2013). Effects of the Tennis Tournament on Players' Physical Performance, Hormonal Responses, Muscle Damage and Recovery. *Journal of Sports Science & Medicine*, 12(2), 240-248.
- Palut, Y., & Zanone, P.-G. (2005). A dynamical analysis of tennis: Concepts and data. *Journal of Sports Sciences*, 23(10), 1021-1032.
- Pereira, T. J. C., Nakamura, F. Y., de Jesus, M. T., Vieira, C. L. R., Misuta, M. S., de Barros, R. M. L., & Moura, F. A. (2017). Analysis of the distances covered and technical actions performed by professional tennis players during official matches. *Journal of Sports Sciences*, 35(4), 361-368. doi:10.1080/02640414.2016.1165858
- Pereira, T. J. C., van Emmerik, R. E. A., Misuta, M. S., Barros, R. M. L., & Moura, F. A. (2018). Interpersonal coordination analysis of tennis players from different levels during official matches. *Journal of Biomechanics*, 67, 106-113. doi:<https://doi.org/10.1016/j.jbiomech.2017.11.036>

- Ponzano, M., & Gollin, M. (2017). Movement analysis and metabolic profile of tennis match play: comparison between hard courts and clay courts. *International Journal of Performance Analysis in Sport*, 17(3), 1-12. doi:10.1080/24748668.2017.1331572
- Prieto-Bermejo, J., & Gómez-Ruano, M. Á. (2016). Entering tennis men's Grand Slams within the top-10 and its relationship with the fact of winning the tournament.[Acceder a los Grand Slams de tenis masculino desde el top-10 y su relación con el hecho de ganar el torneo]. *RICYDE. Revista Internacional de Ciencias del Deporte*. doi: 10.5232/ricyde, 12(46), 410-422.
- Reid, M., & Duffield, R. (2014). The development of fatigue during match-play tennis. *British Journal of Sports Medicine*, 48(Suppl 1), i7-i11. doi:10.1136/bjsports-2013-093196
- Reid, M., & Elliott, B. (2002). Tennis: The one - and two - handed backhands in tennis. *Sports Biomechanics*, 1(1), 47-68. doi:10.1080/14763140208522786
- Reid, M., Elliott, B., & Alderson, J. (2008). Lower-limb coordination and shoulder joint mechanics in the tennis serve. *Medicine & Science in Sports & Exercise*, 40(2), 308-315.
- Reid, M., Elliott, B., & Crespo, M. (2013). Mechanics and learning practices associated with the tennis forehand: a review. *Journal of Sports Science & Medicine*, 12(2), 225.
- 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.
- Reid, M., Morgan, S., Churchill, T., & Bane, M. K. (2014). Rankings in professional men's tennis: a rich but underutilized source of information. *Journal of Sports Sciences*, 32(10), 986-992. doi:10.1080/02640414.2013.876086
- 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. doi:10.1080/02640414.2016.1139161
- Reid, M., & Morris, C. (2013). Ranking benchmarks of top 100 players in men's professional tennis. *European Journal of Sport Science*, 13(4), 350-355. doi:10.1080/17461391.2011.608812
- Robertson, S., Back, N., & Bartlett, J. D. (2016). Explaining match outcome in elite Australian Rules football using team performance indicators. *Journal of Sports Sciences*, 34(7), 637-644.

- Sampaio, J., & Leite, N. (2013). Performance Indicators in Game Sports. In T. McGarry, P. O'Donoghue, & J. Sampaio (Eds.), *Routledge handbook of sports performance analysis* (pp. 115-126). London: Routledge.
- Sampaio, J., McGarry, T., Calleja-González, J., Jiménez Sáiz, S., Schelling i del Alcázar, X., & Balciunas, M. (2015). Exploring Game Performance in the National Basketball Association Using Player Tracking Data. *PLOS ONE*, *10*(7), e0132894. doi:10.1371/journal.pone.0132894
- Sampaio, J. E., Lago, C., Gonçalves, B., Maças, V. M., & Leite, N. (2014). Effects of pacing, status and unbalance in time motion variables, heart rate and tactical behaviour when playing 5-a-side football small-sided games. *Journal of Science and Medicine in Sport*, *17*(2), 229-233.
- Sanchez-Pay, A., Palao, J. M., Torres-Luque, G., & Sanz-Rivas, D. (2015). Differences in set statistics between wheelchair and conventional tennis on different types of surfaces and by gender. *International Journal of Performance Analysis in Sport*, *15*(3), 1177-1188.
- Seifert, L., Button, C., & Davids, K. (2013). Key properties of expert movement systems in sport. *Sports Medicine*, *43*(3), 167-178.
- Smith, M. T., Reid, M., Kovalchik, S., Wood, T., & Duffield, R. (2017). Heat stress incidence and matchplay characteristics in women's Grand Slam tennis. *Journal of Science and Medicine in Sport*. doi:10.1016/j.jsams.2017.11.006
- Triolet, C., Benguigui, N., Le Runigo, C., & Williams, A. M. (2013). Quantifying the nature of anticipation in professional tennis. *Journal of Sports Sciences*, *31*(8), 820-830. doi:10.1080/02640414.2012.759658
- Tudor, P. B., Zecic, M., & Matkovic, B. (2014). Differences between 2010 and 2011 performance indicators of tennis play at the grand slam tournaments. *Kinesiology*, *46*, 101-106.
- Unierzyski, P., & Wieczorek, A. (2004). Comparison of tactical solutions and game patterns in the finals of two grand slam tournaments in tennis. *Science and racket sports III*, 200-205.
- Vaverka, F., & Cernosek, M. (2013). Association between body height and serve speed in elite tennis players. *Sports Biomechanics*, *12*(1), 30-37. doi:10.1080/14763141.2012.670664

- Vaz, L., Leite, N., João, P. V., Gonçalves, B., & Sampaio, J. (2012). Differences between Experienced and Novice Rugby Union Players during Small-Sided Games. *Perceptual and motor skills*, 115(2), 594-604. doi:10.2466/30.10.25.PMS.115.5.594-604
- Volker, M. A. (2006). Reporting effect size estimates in school psychology research. *Psychology in the Schools*, 43(6), 653-672. doi:10.1002/pits.20176
- Whiteside, D., Bane, M., & Reid, M. (2016). *Differentiating top-ranked male tennis players from lower ranked players using hawk-eye data: An investigation of the 2012–2014 Australian Open tournaments*. Paper presented at the ISBS-Conference Proceedings Archive.
- Whiteside, D., & Reid, M. (2016a). External Match Workloads During the First Week of Australian Open Tennis Competition. *International Journal of Sports Physiology and Performance*, 0(0), 1-20. doi:10.1123/ijsp.2016-0259
- Whiteside, D., & Reid, M. (2016b). Spatial characteristics of professional tennis serves with implications for serving aces: A machine learning approach. *Journal of Sports Sciences*, 35(7), 1-7. doi:10.1080/02640414.2016.1183805
- Williams, A. M. (2000). Perceptual skill in soccer: Implications for talent identification and development. *Journal of Sports Sciences*, 18(9), 737-750. doi:10.1080/02640410050120113