

Performance Evaluation of Women's Volleyball Players Using Bayesian Data Analysis

by

David Awosoga

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Data Science

Waterloo, Ontario, Canada, 2024

© David Awosoga 2024

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Understanding player contribution is an important component of lineup construction, advance scouting, and performance evaluation in volleyball. However, traditional methods utilize oversimplified percentages that fail to acknowledge latent variables and situational nuance. These shortcomings are addressed in this work via a holistic framework mirroring the structure of a story, with each sub-component representing an angle by which player contribution can be investigated. The emphasis here is on modelling player contribution via player presence, and a Bayesian logistic regularized regression with a horseshoe prior and custom model structure is employed. This approach incorporates player roles, lineup matchups, and additional context-specific information into its estimates. The resultant model outputs are tested against substantive knowledge and posterior predictive checks, with direct application to volleyball coaches, players, and the community at large. Extensions to this analysis consider other factors of player contribution such as player action, event sequences, and player intent, with applications arising from advancements in computer vision, sports science technology, and data acquisition also discussed.

Acknowledgments

I would like to thank everyone who has played a role in the completion of this thesis, starting with my supervisor Dr. Samuel Wong. Thank you for taking me on as a student and providing me with the flexibility and guidance necessary to create a special final project. I would also like to thank University of Waterloo Warriors Women's Volleyball Head Coach Richard Eddy for graciously giving me the opportunity to work with the team and learn firsthand the nuance and hard questions existing within this sport. Thank you to the other assistant coaches, and of course, the players, for intentionally making me feel welcome at practices, always being open to discuss volleyball strategy with me, and taking genuine interest in my work. Thank you to my colleagues at Canadian Sport Institute Ontario for being so flexible with their data solutions requests so as to accommodate the final sprint of this project. Finally, thank you to everyone who has been praying for and encouraging me over the past two years. This thesis completion is evidence of answered prayer and God's provision, and I cannot be more thankful for the support that I have received.

Dedication

All glory to the most high God.

Table of Contents

Author’s Declaration	ii
Abstract	iii
Acknowledgements	iv
Dedication	v
List of Figures	viii
List of Tables	ix
List of Abbreviations	x
1 Introduction	1
1.1 Background	2
1.2 Motivation	4
2 Literature Review	6
2.1 Performance Analysis and Skill Importance	7
2.2 Individual Player Evaluation	9

3	Data Exposition	10
3.1	Inferring Player Roles	12
3.2	Capturing Libero Presence	14
4	Modelling Contribution Via Player Presence	16
4.1	Plus/Minus	17
4.1.1	Discussion	21
4.2	Extensions to Plus/Minus	22
4.2.1	RAPM in Volleyball	25
4.2.2	Model Specification	26
4.2.3	Discussion	28
4.3	Posterior Predictive Checks	32
5	Conclusion	37
5.1	Future Extensions	38
5.1.1	What - Player Action	39
5.1.2	Where - Coordinate Data	40
5.1.3	How - Action Quality and Physical Attributes	41
5.1.4	When - Event Sequences	42
5.1.5	Why - Player Intent	42
	References	44
	APPENDICES	52
A	RAPM Player Estimates	53
	Glossary	61

List of Figures

1.1	Volleyball positions. The line that horizontally bisects the court divides the areas known as the front court and the back court.	3
3.1	Player roles at the beginning of a point when the setter starts in position 1	13
4.1	Distribution of positive pairwise plus/minus differentials for each player . .	23
4.2	Density plot of horseshoe prior shrinkage weights K_i and β estimates . . .	28
4.3	Density Plot of Player RAPM Estimates	29
4.4	Trace plots of top 4 OUA players, by RAPM	30
4.5	Simulated point-win probability versus observed data	34
4.6	Distribution of simulated maximum point spread versus observed data . . .	35
4.7	Binned Residual Plots	36
5.1	DataVolley Zone and Sub-Zone Definitions	40

List of Tables

1.1	The names, responsibilities, and typical skills performed by each player role in volleyball.	3
3.1	Sample information recorded in Data Volley 4 play-by-play data	11
3.2	Member Schools for Women’s Volleyball in the OUA.	11
4.1	Top and bottom 10 USPORTS plus/minus accumulators, 2021-2024 seasons	17
4.2	Top 10 USPORTS scaled plus/minus accumulators, 2021-2024 seasons (minimum 1000 points played)	18
4.3	Points played in one rotation and their ratios	20
4.4	Weighted ANOVA of Scaled Plus/Minus	21
4.5	RAPM player contribution estimates, separated by player role	31

List of Abbreviations

APM Adjusted Plus/Minus [22](#), [24](#), [26](#), [38](#)

BPM Box Plus/Minus [39](#), [42](#)

FIVB Fédération Internationale de volleyball [2](#)

NCAA National Collegiate Athletic Association [10](#), [14](#)

OUA Ontario University Athletics [11](#), [12](#), [14](#), [18–20](#), [28](#), [31](#), [37](#)

RAPM Regularized Adjusted Plus/Minus [5](#), [25](#), [26](#), [28](#), [31](#), [32](#), [38](#), [39](#)

USPORTS University Sports [11](#), [18](#), [19](#), [28](#), [31](#), [37](#)

Chapter 1

Introduction

Understanding an individual’s contribution towards overall group output is an important area of study in a wide domain of fields [1, 2, 3]. In sport, these types of problems garner much attention from important stakeholders such as fans, coaches, and players themselves. The quantitative analysis of sports, also known as **sports analytics**, is quite interested in this subset of performance analysis, investigating metrics that can be used as a proxy for quantifying a player’s skill, talent, and contribution. The field was pioneered in the 1960’s in baseball and has skyrocketed in popularity over the past twenty years, catalyzed by rapid methodological advances and the expanded accessibility of data [4]. Insights gained from sports analytics complement those that coaches bring via their intuition and experience, initiating systematic analyses that support, challenge, and provide further context to assumptions about the underlying processes that govern different game complexes. It is now commonplace to see professional sports teams with entire departments dedicated to research and development, with the Oakland Athletics of Major League Baseball and the Dallas Mavericks of the National Basketball Association (NBA) widely regarded as early adopters [5].

Data science has been integrated more broadly across a myriad of sport disciplines, including coaching, sports science, business, and player health and safety [6]. These integrations require three main points of consideration. The first and most important is the *practical significance* of the results, such that we can envision the kind of impact that they will have on how the game is played and coached. Tied closely to this impact is the readiness by which findings can be implemented within a daily training environment. The second most important point to consider is the *technical contribution* of a proposed integration. As sport garners increased attention from quantitative experts, methodological advancements readily follow, bringing clarity to previously opaque areas within sport [7]

while simultaneously inspiring further areas of investigation. There is also a large interdisciplinary aspect as researchers realize the transferability of notable approaches across domains. For example, “strain” - a concept in material science used to describe the deformation of a material from stress - has been applied to National Football League data to model the compression and collapse of the area of protection around the offensive team’s quarterback, which is known as the “pocket” [8]. The resultant metric is then used to measure a pass rusher’s effectiveness. Another example is the application of Voronoi diagrams from computational geometry to quantify regions of control for each player in sports such as soccer [9]. This helps model which player can reach the ball first based on their positional data.

Thirdly, both for implementation purposes and insights, *data constraints* must be taken into account for new research. Advancements in sports analytics motivate the adoption of technologies that attempt to capture previously understudied aspects of athlete performance. Some relevant examples for measuring internal load include using force plates for analyzing jumping profiles and heart rate monitors for athlete monitoring [10]. As both the quality and quantity of data proliferates, effective visualization, modeling, and communication of it become paramount to justify to the associated costs of acquisition. Therefore, careful consideration is still necessary when deciding which technological frameworks to adopt in sport and how well existing methods address pertinent challenges experienced by decision-makers [10]. Nevertheless, this continuous cycle of technological and methodological innovation inspired by the constant stream of new questions demonstrates that sports analytics will continue to have a major role in future operations within sport.

1.1 Background

The sport chosen for this analysis is volleyball, which originated in 1895 in the United States [11]. The governing body of international volleyball is the [Fédération Internationale de volleyball \(FIVB\)](#), and the sport is estimated to be played by over 800 million people worldwide [11]. Volleyball is played by two teams that each occupy a 9-meter by 9-meter side of a court separated by a net. The net height varies based on gender and level of competition, and at the senior level of the sport a men’s net is 243 centimeters tall, while the women’s net height is 224 centimeters. A volleyball match is divided into segments called *sets*, where the first team to win 3 sets wins the match. A point is initiated from a serve by the team that won the most recent point, and teams alternate possessions by returning the ball over the net after performing a maximum of three consecutive contacts on their side. A team can score a point in a set by having the ball touch the floor within the

boundary lines of their opponent's side of the court, having the ball touch the floor outside of boundary lines of either side of the court after last being contacted by the opposing team, or when the opponent commits a rule infraction.

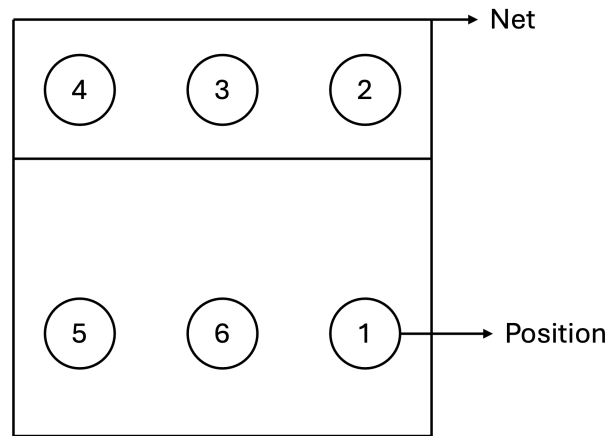


Figure 1.1: Volleyball positions. The line that horizontally bisects the court divides the areas known as the front court and the back court.

Table 1.1: The names, responsibilities, and typical skills performed by each player role in volleyball.

Role Title	Primary Responsibility	Typical Skills/Actions	Front Home	Back Home
Setter (S)	set distributor	Set, Block	2	1
Middle Blocker (MB, x2)	front court defender	Block, Attack	3	5
Outside Hitter (OH, x2)*	attacker and defender	Attack, Receive	4	6
Libero (L)	back court defender	Receive, Dig	NA	5
Right Side (RS)	attacker and defender	Block, Dig	2	1

*One outsider hitter plays in the front row, while the other plays in the back row.

A volleyball game is played with six players on each team on the court at a time - three in the front row (front court) and three in the back row (back court). Each player is assigned a **role** that they occupy for the duration of a set, and will start a point in one of six possible **positions** at the advent of an individual point. These positions are labelled from one through six in a counterclockwise manner starting from the back-right side of the court, as shown in Figure 1.1. The orientation of a team, called their **rotation**, is given relative to the position that the setter is in.

After the team that receives the serve successfully scores a point, the players on that side rotate one position clockwise. Although a player must begin a point in their assigned position, they typically move shortly after to a “home” position that is dictated by their role and whether they are in the front court or the back court. Table 1.1 describes the five distinct roles that players on each side can take, along with their primary responsibilities, typical skills, and home position when they are in the front or back row.

Although this introduction to the structure of volleyball is not intended to be comprehensive, the nuance associated with the role of the libero is worth mentioning in greater detail. The libero is a specialized role that was introduced in the late 90’s with the intention of leveraging their defensive skills to produce longer rallies and give undersized players greater opportunities to meaningfully contribute to team success [12]. Those who have watched volleyball can visually identify liberos as the player who wears a different-colored jersey from the rest of their teammates. A libero can substitute any player in the back row but is not permitted to play in the front row, which is why their front home position is marked by “NA” in Table 1.1. Liberos are traditionally very skilled passers and typically substitute out the middle blocker once they rotate into the back row. Different leagues employ various additional constraints around the allowed actions of liberos, including serving and front row setting.

1.2 Motivation

Mathematically, many aspects of volleyball can be framed as a constrained optimization problem. “Talent” is the scarce resource in question, and a coach’s objective is to construct a lineup that maximizes their team’s win probability. A number of factors must be considered when deciding on a lineup, and a major component of that is based on how a coach evaluates how each player contributes to team success. Utilizing a comprehensive yet digestible method to aid in this decision making becomes extremely advantageous. Once a lineup is chosen, it is assumed that a team’s cumulative talent is a fixed input that doesn’t significantly change within a season. External interventions such as player acquisition and long term player development are then the only other factors that can influence a team’s collective talent. Therefore, a coaches’ next point of business is advance scouting, the task of developing strategies and tactics that account for their team’s overall talent and maximize their win probability against upcoming opponents. This is another situation that relies heavily on an appropriate understanding of a player’s contribution to a team and how it is reflected in their performance.

These examples underscore the necessity of an extensive understanding of player con-

tribution, and the approach used in this work builds around the classical sub-components of a story, “Who”, “What”, “Where”, “When”, “Why”, and “How”, to drive the analysis. Each of these questions is addressed by varying granularities of data, and represent a different angle by which we can understand player contribution in volleyball. The rest of the paper is organized in the following manner. Chapter 2 gives a comprehensive review of existing attempts to evaluate player performance and quantify skill in volleyball, as well as how these attempts have progressed over time. Chapter 3 gives a detailed overview of the data utilized in this work, including commentary on necessary transformations and manipulations required to address missing and erroneous data. Chapter 4 addresses the “Who” of player contribution by considering player presence as the reference input. This is inspired by analogous approaches borrowed from other sports, centering around a metric called “plus/minus” and its related derivative [Regularized Adjusted Plus/Minus \(RAPM\)](#). Finally, Chapter 5 will summarize the findings and provide discussion on potential ways to extend this analysis to address the remaining “W’s” of player contribution.

Chapter 2

Literature Review

Across sports, research related to performance evaluation has received much attention due to the practical significance of the results. While match summary and event-level data are still the most accessible and utilized sources of data, object tracking systems such as global positioning systems (GPS), inertial measurement units (IMU), markerless motion capture, and radio frequency identification have quickly gained traction [13]. These technological advances have introduced a wealth of high resolution spatiotemporal information that has revolutionized analysis, particularly around the quantification of a player’s impact on defense [14]. Bayesian methods have also become increasingly popular in sports analytics [15] as they allow for the inclusion of prior information and the quantification of uncertainty. Recent examples include Bayesian hierarchical models in basketball [16, 17], Twenty20 (T20) Cricket [18], and effect size quantification in experimental training regimes [19]. To help decision makers filter out important metrics that provide the most unique and reliable information, a suite of **meta-analytics** have been introduced [20]. These have broad applications within sport, emphasizing *discrimination* - how well a metric can distinguish between players, *stability*, how the metric varies over time after removing chance variability, and *independence*, how different a proposed metric is from existing metrics.

The techniques utilized in modern volleyball analytics are much more primitive than those used in other sports, but there are noted opportunities for advancing the relative infancy of statistical research in volleyball. To start, there are evident similarities between volleyball and baseball in particular within the player-ball interactions. For example, observable similarities between serving a ball in volleyball and throwing a pitch in baseball motivate the potential for volleyball analogs of pitching metrics. There are also a handful of skill-based defensive metrics in baseball that can be translated to volleyball in a conceptually straightforward manner with appropriate data, such as a player’s “range” - how

much ground they can cover to make a play on a ball [21]. Volleyball can additionally borrow from methodologies developed in other team-based “invasion sports” such as gridiron football, soccer, basketball, and ice hockey. “Invasion sports” are loosely defined as sports where players can move freely within a playing area as the teams that they are on compete for possession of a ball or puck to score more points on their opponent than they concede [22]. Although volleyball is not by definition an invasion sport since non-adversarial contacts within a possession are conducted without interference from their opponents, it does share many game structures in common with these sports. This allows for research ideas from the better established sports to be adopted in volleyball, particularly within the areas of (static) game actions, dynamic game actions, movement patterns, collective team behaviours, social network analysis, and game styles [23]. Volleyball research has historically centered around performance and match analysis at a team level [24], with more recent works pivoting towards evaluating individual players. The quantity of studies of the former far outnumber those of the latter, and relevant developments across both categories will be reviewed in the following sections.

2.1 Performance Analysis and Skill Importance

Volleyball studies in performance analysis and skill importance are traditionally inferential and descriptive in nature, using common statistical techniques such as analysis of variance (ANOVA) and variations of discriminant analysis [25, 26, 27]. Some of the earliest published works investigated the playing characteristics of teams and the best predictors of success from different skill components [28]. Eom and Schutz [29] used multiple analysis of variance (MANOVA) and discriminant analysis and found that attacking and blocking were the most discriminative between a team winning and losing. More recently, Lopez et al. [30] performed a relational study and found a significant relationship between reception zone and the performance of the reception as well as an interaction between a team’s rotation and their reception performance. Most analyses take place within elite or international levels of play [31, 32], but others have run experimental studies with younger volleyball players to observe the outcomes from direct interventions towards the execution of various performance things [33]. For example, Oikonomopoulou et al. [34] assessed the spatial and temporal characteristics of the setting choices made by junior female volleyball setters and their performance concerning the origin of the ball and the game complex. They found demonstrated patterns in setting decisions depending on the tempo of the set, as well as in the lane of the court where the ball originated from.

Graph theoretic approaches such as social network analysis have also been employed to

evaluate different game phases, otherwise known as “game complexes”. These complexes are based on the state of a point as a rally progresses, and analysis is performed on the actions that a team conducts within that complex. There are 4-6 defined complexes in the literature [35], starting from a serve (**serve**, Complex 0) to the possession of the receiving team (**side-out**, Complex I), the possession of the serving team after receiving an attack (**sideout-transition**, Complex II), and all subsequent possessions by either team (**transition**, Complex III). Eigenvector centrality is a popular measure used to reflect the tendencies of different skills demonstrated by teams within each game complex. Loureiro et al. [36] found below average variability in attack organization and recommend higher training emphasis in “out-of-system (**transition**)” game complex, as that is where the majority of counter attack opportunities arise from. This is affirmed by Laporta [37], who focuses more on the functional differences between different game complexes. Although the majority of volleyball research papers focus on indoor volleyball, other variants, such as beach volleyball, have received recent attention [38]. The work in that space is less novel than in indoor volleyball, with interesting comparative analyses performed between genders and other investigations into the impact of rule changes in game play.

Machine learning and more advanced statistical models have also been employed to assess skill competencies. Miskin et al. [39] used a Markov chain with a multinomial likelihood and Dirichlet prior on the transition probabilities to quantify skill importance in Women’s volleyball. They give each skill an “importance score” that can be used to assist coaches in allocating practice time, developing new strategies, and optimizing team performance relative to player selection. Another skill evaluation framework was introduced by Bagley and Ware [40] with the objective of removing the bias that is inevitable with manual evaluations of skill execution. Both works agree that the traditional 3-and 4-point rating systems for skill do not appropriately consider impact on point outcome nor the difficulty of the action in question. Other examples include an unsupervised learning method to create player archetypes [41], an experimental study on volleyball spiking form using kinematic, dynamic, and surface electromyography parameters in a support vector machine [42], and an inductive logic programming (ILP) implementation used to discover strategies and understand frequency and efficiencies [43]. Computer vision and other deep learning models are becoming increasingly popular in volleyball analytics but have not quite replaced traditional machine learning methods in production [44].

2.2 Individual Player Evaluation

Of particular relevance to this work are efforts from Gill Fellingham [45] and Zachary Hass [46] to evaluate performances by individual players instead of aggregated for a team. Fellingham estimated the average contribution to team success of various positions using a Bayesian hierarchical mixed logistic model that evaluated the players who touched the ball in a team’s possession and what the outcome was. Touches per set are recorded for each player, and the position and player contribution to the log odds of a positive and negative outcome are separately modelled. From this he proposed a metric called “points scored above average”, with per-set and per-100 touches variants. This allows for comparisons to be made *within* a player’s role, such that players can be ranked and evaluated in a novel manner. The analysis does not separate a player’s performance from the performance of their team, nor can it be used to find the most valuable role, most valuable player, or the effect of teammates on the court. Nevertheless, the insights generated and the posterior distributions used for different player roles serve as starting points for possible future works.

In a 2017 PhD dissertation, Hass [46] developed a flexible and generalizable framework to divide credit between players. In it a comprehensive definition for division of credit models is detailed, highlighting the necessity for metrics to be relatable to an outcome, recordable or quantifiable, and attributable to specific players. Hass trials several regression models under a Bayesian framework to split credit and compare contribution between players and finds useful applications in optimal roster construction. A Markovian approach is then used to divide contribution by player action, using a Dirichlet model to estimate player strength. By modifying the model inputs to value action by their quality, a two-part model is developed to control for player role and game context. This culminates in a metric produced called Credit above Value Expected (CaVE), which is the difference in the observed credit given in the play and the value expected given the role and context. CaVE is validated via simulation studies and demonstrates promise in capturing underlying player strengths. However, robust inference of the metric is limited by the variability in its inputs, necessitating the implementation of a fully Bayesian model in future iterations.

Chapter 3

Data Exposition

After a volleyball game is played, the film is sent to be “coded” by a *scout*. The scout records game metadata, lineup and roster information, and most importantly, a wealth of detailed auxiliary information about the events that transpired. The quality of the information recorded is directly contingent on the thoroughness of the scout, and thus higher level leagues typically have higher quality coding performed than amateur ones. The data is collected in a variety of data structures and formats, the most popular being via game-level data (boxscores) and event-level data (play-by-play). Boxscores are compact tables that contain the summarized raw counts of player statistics, which are then aggregated to compose team performance. In volleyball specifically, an event is recorded each time a player contacts the ball, and play-by-play refers to a game report that gives the sequential listing of these events. This recounts contextual and identifiable information about each player at the time of any event occurrence. In terms of data storage, tabular formats such as comma-separated values (`.csv`) documents are the most common for boxscores, while volleyball play-by-play is stored in a variety of formats depending on the data vendor. Data Volley 4 [47] is the most popular software for collecting volleyball data, with over 10,000 users in over 65 leagues across the world, many in Europe. Data Volley 4 exports play-by-play in a custom `.dvw` format, with a component containing one row per event and over 80 columns of different variables describing it, some examples of which are given in Table 3.1. In spite of its popularity, Data Volley 4 is not without its shortcomings. The software has not been updated since 2007, prompting the adoption of successor programs such as VolleyStation [48] by many collegiate institutions and professional teams. Other notable rivals include Volleymetrics, which can export play-by-play data in `.json`, and StatCrew, a program previously used at many [National Collegiate Athletic Association \(NCAA\)](#) schools for collegiate volleyball, which exports the data as a raw text file.

Table 3.1: Sample information recorded in Data Volley 4 play-by-play data

Column Type	Example Variables
game metadata	match ID, point ID
event identifier	team, player number, player name, player ID
event description	skill, type, start zone, end zone
contextual data	set number, team scores, team rotations, serving team
player identities	player numbers and ID's in positions 1-6

The data utilized in this project comes from [Ontario University Athletics \(OUA\)](#) women's volleyball. The [OUA](#) is based in the Canadian province of Ontario and is the largest of the four conferences that make up [University Sports \(USPORTS\)](#), the governing body and leader of university sport in Canada. Very few analyses have been performed in this league, a notable exception being in basketball [49]. This provides a great opportunity to advance the analytics space and generate novel applications within an understudied domain. Information on the 14 women's volleyball teams in the [OUA](#) as of the 2023-24 season is listed in Table 3.2.

Table 3.2: Member Schools for Women's Volleyball in the OUA.

School	Team Name	Location
Brock University	Badgers	St. Catharines
Lakehead University	Thunderwolves	Thunder Bay
McMaster University	Marauders	Hamilton
Nipissing University	Lakers	North Bay
Queen's University	Gaels	Kingston
Royal Military College of Canada (RMC)	Paladins	Kingston
Toronto Metropolitan University* (TMU)	Bold	Toronto
Trent University	Excalibur	Peterborough
University of Guelph	Gryphons	Guelph
University of Toronto	Varsity Blues	Toronto
University of Western Ontario	Mustangs	London
University of Waterloo	Warriors	Waterloo
University of Windsor	Lancers	Windsor
York University	Lions	Toronto

*Formerly known as the Ryerson University Rams

All 362 regular season games played during the 2021-22, 2022-23, and 2023-24 were

extracted from Perfbook [50], the official data vendor for OUA volleyball. The OUA uses the aforementioned `.dvw` file format, which can be rendered in plain text and parsed using the open source **datavolley** [51] package in R. Extensive data cleaning is necessary to consolidate misspellings of player and team information from across seasons. Additional care is required for players who change jersey numbers or transfer schools between seasons, as these changes are not always caught by the scout.

There are three common types of errors that arise from coded games for which programmatic remedial measures could not be effectively implemented. They are listed for completeness but make up an estimated 5% of the data, a non-negligible amount but one that does not play a significant factor over the course of the entire analysis.

- **Missing actions** can occur when a scout misses a sequence of events in a rally. These can sometimes be detected by recognizing abnormal play numbering, but cannot be systematically recaptured without re-watching the game.
- **Mis-specified player contacts** can occur if a player is mistaken for another, or if their number is incorrectly recorded when an event occurs. These are very difficult to detect without a working knowledge of player roles.
- **Mis-specified player identities** occur via a missed substitution or when a player number is mis-coded at the beginning of a set. These can be detected by comparing recorded player actions with the list of recorded players on the court.

After combining the event information from each game, there were nearly 56,000 points played and $\sim 350,000$ recorded player contacts. With the data now cleaned and consolidated, several key transformations are necessary to prepare the data for analysis.

3.1 Inferring Player Roles

Player roles, as defined in Table 1.1, are not explicitly recorded in the `.dvw` file except for that of the setter. However, it is advantageous to directly extract these roles for all of the other players on a point-by-point basis from the contextual information. This is motivated by the necessity to compare the contribution of players with similar roles, as well as perform exploratory analysis on the implications of rotational match-ups in team success. We can deterministically infer player roles by taking advantage of two prominent spatial patterns typically observed in the lineup structure of teams. The first is that the

two outside hitters are always positioned diagonal from one another, as are the two middle blockers. The setter and the right side are the final pairing of players who stand diagonal from one another. This means that the players will have an offset of 3 positions from their diagonal pair, so if one outside hitter is in position 2, for example, the other outside hitter will be in position 5.

The second pattern is the order by which the players are aligned at the beginning of the set, which remains fixed throughout its duration. Teams overwhelmingly favour an “*SHM*” lineup orientation, an abbreviation for “Setter - Hitter - Middle”. This means that the setter is followed by an outside hitter, who in turn is followed by a middle blocker. To differentiate between the two outside hitters and the two middle blockers, the outside hitter who *follows* the setter is labelled as “Outside Hitter 1 (OH 1)” and the other as “Outside Hitter 2 (OH 2)”. Analogous to this, the middle blocker who follows the outside hitter in SHM is referred to as “Middle Blocker 2 (MB 2)”, and the other as “Middle Blocker 1 (MB 1)”. By these conventions the rest of the player roles can be inferred. Figure 3.1 illustrates this inference for a setter in position 1. By the diagonal pattern, we immediately know that the right side is in position 4. By the SHM convention, we know that that OH 1 is in position 2, meaning that the OH 2 is in position 5. Similarly by SHM we know that MB 2 (most likely replaced by the libero) is in position 6, so MB 1 will be in position 3.

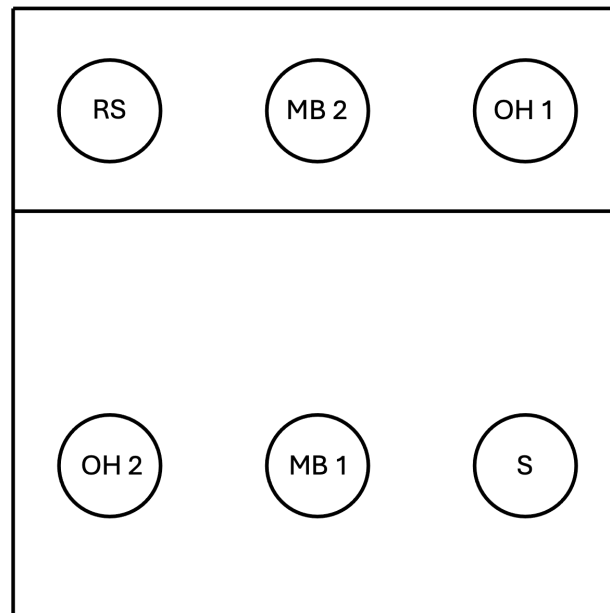


Figure 3.1: Player roles at the beginning of a point when the setter starts in position 1

3.2 Capturing Libero Presence

While we have shown that most player roles and positions can be reconstructed based on the provided information of the setter, this does not include the position of the libero. As stated in Chapter 1, liberos can substitute any player in the back row and move freely in and out of the game without recording a substitution. Since these substitutions happen so frequently, only libero skills are recorded in Data Volley 4 but not their positions on the court at the start of the point. Unlike the other methods used to clean and transform the data, which for any volleyball data format, this item has some nuance based on league-specific libero rules. In the [NCAA](#), for example, only one libero can be declared by a team during a set. That means that ignoring misspecified player contacts, a player who is recorded as making a contact without being registered in the list of known player identities can be presumed to be the libero.

From here two popular approaches have been employed to infer libero presence. Powers et al. [52] used subject matter knowledge that liberos replace the middle blocker after they have served in the back row to make a deterministic replacement that persisted throughout the entire set. Hass [46] used a probabilistic approach with a Bayesian multinomial model that leveraged the rotation and serve information of the game state. Instead of predicting where the libero would be, the model predicts which of the four possible back row players would be off of the court. This removes the necessity for the assumption of replacement based on subject matter knowledge, at the detriment of being much more complicated to model. These are both reasonable approaches, but neither are as effective with the current dataset because two liberos can be declared in a set in the [OUA](#) and these declarations can change in between sets, another unique characteristic of [OUA](#) rules. This motivates the implementation of a specialized approach that accounts for these differences.

First, we filter roster metadata to create a set of possible liberos. This set may contain more than two liberos due to constraints from Data Volley 4 software. For example, if player A and player B are designated as liberos at the beginning of a match, but players B and C are designated as liberos at for set 2, there are 3 possible liberos to choose from, and we have to consider all of them as possibilities. After getting this set of possible liberos and filling in matches where a libero was not declared, we borrow from the approach of Powers and note all players who recorded a contact but do not appear in the player identity list. We then filter this list of unidentified players by our set of possible liberos, and tabulate the proportion of touches that each libero has in each set during the course of a match. There are then three possible scenarios by which libero presence can be assigned:

1. **Confirmed Libero Presence:** If a registered libero makes a contact during a

possession, then we trivially know that they are present and mark them accordingly

2. **Confirmed Libero Absence:** For rotations where the the middle is serving, we follow the trends of convention and infer that no libero is present on the court.
3. **Unknown Libero Presence:** For plays where we don't know which libero is present, we use the proportion of individual touches throughout the set by each potential libero as a probabilistic indicator of their presence.

With this final pre-processing step complete, the data is now ready to be used in both exploratory and more rigorous analysis of player contribution.

Chapter 4

Modelling Contribution Via Player Presence

We focus our investigation on the “**who**” of player contribution. How can we quantify the magnitude of contribution via knowledge of player identity? This question is addressed by considering **player presence**. Player presence models decompose team success and divide credit from these outcomes between the individual players on the court. These methods aim to capture player impact without explicitly specifying the forms that it takes. This can account for contributions that cannot be tracked by current technology, such as teamwork effects and other player intangibles [53]. Player presence methods also help quantify the influence that players have on the decision making of their opponents and teammates. Typical measures for team success used as outcomes in player presence models are wins, points, and their derivative probabilities and scaled differentials [46].

Understanding individual contribution through this lens is the first level considered because it requires the coarsest granularity of data. In its simplest formulation, one only needs to record the outcome of a point and the identities of the players on each team, without accounting for any of the additional action that takes place during the point. This ease of data collection makes player presence models very accessible and achievable to implement across various levels of sport, from recreational to amateur to professional. In this chapter we perform exploratory analysis on methods derived using player presence, and create a more robust model that accounts for some of the contextual covariates that affect its serviceability. We then introduce adjustments to these metrics based on observed challenges with interpreting their raw form, introducing a Bayesian framework to quantify uncertainty in the player estimates. Finally, we perform posterior predictive checks on the model outputs via simulation to ensure that they provide meaningful information.

4.1 Plus/Minus

We begin with **plus/minus**, which is a simple counting statistic that for each player takes the difference of the points scored and points conceded by their team while they are on the court. Plus/minus originated in the National Hockey League (NHL) in the 1950’s and 60’s [54], and has since been adopted extensively in basketball, soccer, and other leagues around the world [55]. Plus/minus is widely used by players and fans alike but has also faced criticisms due to its lack of contextual awareness. Nevertheless, it serves as a springboard for inquiry into the data and its utility. The segment of time that a player is active for their team is known as a “shift” in hockey, and these combinations of active players on at one time is what is used to calculate plus/minus. We refer to them here as “stints”, consistent with definitions applied in better established basketball research in plus/minus. Unlike hockey, basketball, and soccer, in which player stints are recorded on a continuous domain of minutes and seconds, volleyball player stints are recorded over a discrete domain of points. This is because the aforementioned sports are constrained by a time limit, whereas a volleyball game only ends when the winning team scores enough points. Additionally, substitutions in volleyball can only occur during a stoppage in play, whereas in sports like hockey they can take place during live action. These attributes, along with the fact that points are scored at a greater rate in volleyball than in hockey and soccer, provide potential for interesting insights to be gleaned even from its raw output.

Table 4.1: Top and bottom 10 USPORTS plus/minus accumulators, 2021-2024 seasons

Rank	Player	Team	Plus/Minus	Role
Top 10				
1	Aleiah Torres	Brock Badgers	+1,006	Libero
2	Madison Chimienti	Brock Badgers	+864	Outside Hitter
2	Sara Rohr	Brock Badgers	+864	Setter
4	Emily Foest	Brock Badgers	+797	Outside Hitter
5	Sadie Dick	Brock Badgers	+651	Right Side
6	Shea Baker	Queen’s Gaels	+445	Setter
7	Mary Stewart	Queen’s Gaels	+437	Outside Hitter
8	Aimee Lemire	Queen’s Gaels	+408	Outside Hitter
9	Hannah Duchesneau	Queen’s Gaels	+400	Right Side
10	Laurin Ainsworth	Brock Badgers	+398	Outside Hitter
Bottom 10				

1	Isabelle Pellerine	Trent Excalibur	−1,027	Setter
2	Ashley Raymond	RMC Paladins	−966	Outside Hitter
3	Emma Titley	RMC Paladins	−914	Outside Hitter
4	Natalie Fournier	Trent Excalibur	−869	Right Side
5	Taryn Bender	RMC Paladins	−734	Right Side
6	Emilie L’Abbe	RMC Paladins	−679	Setter
7	Emilie Boileau	RMC Paladins	−668	Libero
8	Ashley Chartier	Trent Excalibur	−613	Outside Hitter
9	Justine Colambot	Trent Excalibur	−591	Outside Hitter
10	Jamie Spencer	Trent Excalibur	−534	Middle Blocker

Table 4.1 displays the output of the raw plus/minus totals of [USPORTS](#) athletes from the entire dataset. What immediately stands out is that both lists are dominated by two teams. It comes at no surprise that these teams were the ones with the best and worst win-loss records in the [OUA](#) during the time frame from which the data is taken from. The Brock Badgers won three consecutive [OUA](#) championships between 2021-2024, boasting an impressive 49 wins compared to only 4 losses in the regular season. The Queen’s Gaels had a stellar run during this period as well, posting 39 wins and 13 losses. Conversely, the Trent Excalibur posted a record of 8 wins and 45 losses, while the RMC Paladins won just one game over those three seasons. Therefore, we can see that raw/plus minus matches intuition and from a bottom-up angle reflects accurately on the players.

Table 4.2: Top 10 USPORTS scaled plus/minus accumulators, 2021-2024 seasons (minimum 1000 points played)

Rank	Player	Team	Role	Scaled Plus/Minus Per Set
1	Sophia Currier	Toronto Varsity Blues	Libero	+8.16
2	Aleiah Torres	Brock Badgers	Libero	+7.50
3	Madison Tuck	Brock Badgers	Right Side	+7.06
4	Laurin Ainsworth	Brock Badgers	Outside Hitter	+6.89
5	Hayley Goodwin	Toronto Varsity Blues	Setter	+6.56
6	Madison Chimienti	Brock Badgers	Outside Hitter	+6.17
7	Sara Rohr	Brock Badgers	Setter	+5.98
8	Holly Biasi	Queen’s Gaels	Libero	+5.93

9	Emily Foest	Brock Badgers	Outside Hitter	+5.59
10	Julie Gordon	TMU Bold	Outside Hitter	+5.54

Table 4.2 presents the top ten [USPORTS](#) scaled plus/minus accumulators, where the scaling divides plus/minus by total points played to discount for relative opportunity. These scaled differentials can be more intuitively viewed on a per-set basis instead of per-point by multiplying by 46, the average number of total points in a set. With this definition, we can credit Sophia Currier, the [OUA](#) leader in scaled plus/minus, with a +8.16 scaled points per set differential. We can see a different prominent pattern arise from this table, with another interesting area of application. While the list remains dominated by outside hitters and members of the Brock Badgers, three of the players on are liberos, compared to only one libero appearing on the top 10 list in Table 4.1. These liberos all play for teams with successful win-loss records, and their abnormally high standing in this list can be explained by rules surrounding their usage, understood under the umbrella of substitutional patterns in volleyball.

In [USPORTS](#), liberos are not permitted to serve, meaning that in typical settings where they substitute the back row middle blocker, they do so *after* the middle blocker has served. This means that for a complete rotation of a team, a libero will be on the court for 4 out of the 6 serves that their team performs. This is significant because serving teams are inherently at a disadvantage in terms of scoring a point, since the receiving team has the first opportunity to formulate an attack. In fact, from the data we find that receiving teams have a 55% point-win probability for a given point. Therefore, players such as liberos and other defensive specialists who appear more frequently when their team is receiving a serve are in an inherently better position to have inflated raw and scaled plus/minus totals compared to other player roles. The opposite is true for middle blockers and serving specialists, who not only play fewer overall possessions compared to their teammates, but actually play more possessions while their team is serving than when their team is receiving, placing a ceiling on their plus/minus potential. It comes as no surprise then that the highest ranked middle blocker in our data set (Angelina Markotic, unsurprisingly from Brock) doesn't appear until 21st on the raw plus/minus list, while 5 out of the 10 players ranked 11-20 on that list are also liberos. The full points played distributions are displayed in Table 4.3, where we see that outside hitters, setters, and right sides do not experience such nuanced usage rates. We see incredible agreement between the theoretical and observed serve-to-receive ratios, with the only discrepancy coming from middle blocker proportions. We theorize that middle blockers serve less often than expected because they are often substituted for by designated servers, especially late in sets. Therefore, steps must

be taken in future modelling iterations to account for these role effects and substitution patterns.

Table 4.3: Points played in one rotation and their ratios

Role	on Serve	on Receive	Theoretical Ratio	Observed Ratio
OH (x2)	6	6	1:1	50:50
S	6	6	1:1	50:50
RS	6	6	1:1	51:49
MB (x2)	4	3	4:3	54:46
L	4	6	2:3	42:58

For now, we use the scaled plus/minus values to tackle some other exploratory questions. The goal is to decompose the data into sub-levels corresponding to different contextual considerations that are taken strategically by coaches when constructing a lineup, to test whether or not they have a noticeable effect on scaled plus/minus. The effects of home court advantage have been studied in several works across volleyball [56, 57], and it is of interest to understand its magnitude in the OUA as well. Having player identities as a level is intended to test how many statistically significant differences can be detected between players. Distinguishing scaled plus/minus between whether a team is on serve or receiving comes as a byproduct of the aforementioned advantage that receiving teams have in terms of scoring a point. Most literature considers volleyball without matchup-specific adjustments, either due to lack of knowledge or lack of actionable insights provided [52]. However, rotational offsets continue to be employed by coaches with the intention of increasing their win probability, which justifies continued analysis to bring more clarity to this presupposition. Finally, role-specific effects are of interest because most literature performs analysis by labeling roles in their general form - “middle”, “outside hitter”, “setter”, “right side”, and “libero”. No distinction is typically made between the MB 1 and MB 2 as well as the OH 1 and OH 2. In practice, however, coaches have been known to switch their OH 1 and OH 2 to achieve a different lineup configuration, instead of replacing a player entirely. Similar to rotations, this is done to achieve a more favourable spatial matchup for offensive and/or defensive assignments. Therefore, investigating whether or not there is a substantial effect in scaled plus/minus with these lineup changes is of interest for strategic considerations. The interaction between player and role was also tested to see if any signal could be found for player-specific differences in quality of performance when playing in different roles.

4.1.1 Discussion

The statistical methods used to investigate the differences in scaled plus/minus at each level were **analysis of variance (ANOVA)** and **Tukey’s honest significant differences (Tukey’s HSD)**. ANOVA is used to test the difference in group mean values, while Tukey’s HSD performs pairwise group comparisons within a level. Both ANOVA and Tukey’s HSD assume independent observations and normally distributed data with equal variance among each group. The ANOVA was weighted by the number of points played, to capture the same effect of relative opportunity as in the scaled plus/minus. The ANOVA output is displayed in Table 4.4.

Table 4.4: Weighted ANOVA of Scaled Plus/Minus

	df	Sum Sq	Mean Sq	F value	Pr(>F)
at_home	1	280.87	280.87	287.00	0.0000
on_serve	1	7666.47	7666.47	7833.62	0.0000
player_on	368	5518.53	15.00	15.32	0.0000
role	6	9.14	1.52	1.56	0.1556
player_on x role	725	1152.19	1.59	1.62	0.0000
matchup	25	34.99	1.40	1.43	0.0755
Residuals	55077	53901.81	0.98		

We see that **at_home** and **on_serve** have unsurprisingly high F-values, matching generally accepted anecdotal evidence and volleyball research. The significance of **player_on** is also comforting as it provides evidence for scaled plus/minus to be able to distinguish players. We see interestingly enough that there aren’t statistically significant differences in scaled plus/minus averages between different **roles**, a pattern that we had identified earlier. This implies that this effect is more context dependent than originally perceived and requires a more granular analysis than what is provided by the current formulation of scaled plus/minus. The interaction between **player_on** and **role** was intended to test whether a systematic difference in scaled plus/minus could be identified when individual players play different roles. No such pattern was found, indicating that these strategic changes are more reasonably reasonably attributed to team-level effects. Finally, the fact that **matchup** is only statistically significant at the 10% level suggests slight but non-convincing support for our intuition on rotation effects.

We use the Tukey pairwise comparisons to analyze the direction and magnitude of the differences in means. We find from the differences that the home teams have an average advantage of +0.041 scaled points over visiting teams, or just under 2 extra points per

set. The positive coefficient is a good sanity check that the ANOVA output is correctly specified, while the interpretation holds weight as it implies that sets played between two evenly-matched teams will on average be decided by which team is playing on their home court. We use the results of the Tukey test to see how many pairwise differences between players were statistically significant with a p-value of 0.0001 and filter accordingly. We then rank players based on how many positive scaled plus/minus differentials they have over other players, which in other words reveals how many other players someone is “better than”, according to this metric. The summary statistics of these rankings reveal some interesting results. Firstly, the mean number of players that someone is “better than” is 5, though the median value is only 1. With nearly 370 players in the data set, this would be an absurd conclusion in practice. In fact, Figure 4.1 illustrates that nearly 90% of players have positive scaled point differentials over less than 15 other players, after which there is a steep increase for a small subset of players. Outsider hitter Laurin Ainsworth, the leader of the list, has a statistically significant positive scaled plus/minus differential over 37 other players, followed by fellow Brock Badger teammates Madison Chimienti (33) and Emily Foest (30). However, we find that of these players, and more generally of any player with positive scaled differentials, that these statistically significant differences only come against players from bottom-dwelling teams such as RMC and Trent. Therefore, we have further evidence that scaled plus/minus cannot appropriately differentiate between teammates *nor* opponents, as the correlation between player estimates renders rankings uninformative.

4.2 Extensions to Plus/Minus

Although we are able to gain preliminary insights into player contribution from raw and scaled plus/minus, there is an obvious need to extend our methodology. We notice immediately that plus/minus does not account for the quality of teammates and ignores the effects of opponents, inflating the totals for poor players on good teams, and suppressing the contribution of good players on poor teams [55]. Fortunately, these shortcomings have been studied in the literature, and a series of advancements have followed. The first improvement was introduced in 2004 by Dan Rosenbaum [58]. Denoted as [Adjusted Plus/Minus \(APM\)](#), Rosenbaum fit a multiple linear regression model by regressing plus/minus on player stints and contextual game variables in the following manner:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_M X_M + \epsilon \quad (4.1)$$

such that

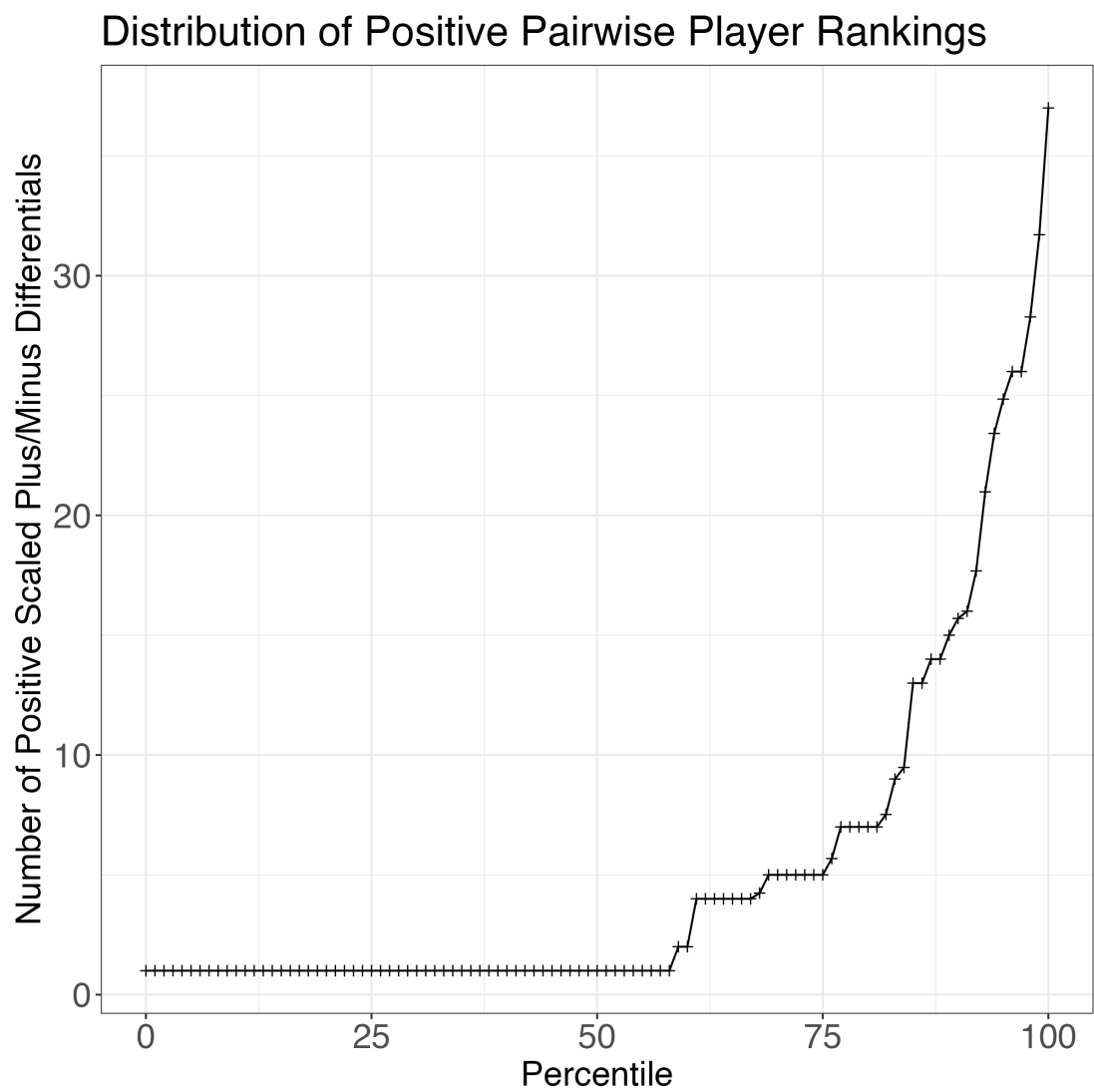


Figure 4.1: Distribution of positive pairwise plus/minus differentials for each player

$$E(Y|X_1, \dots, X_M) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_M X_M \quad (4.2)$$

ϵ is the error term of the model and is assumed to have constant variance such that $Var(\epsilon_i|X) = \sigma^2$, as well as zero covariance between different error terms ϵ_i and ϵ_j when conditioned on $\mathbf{X}$. Finally, the model is assumed to have strict exogeneity such that $E(\epsilon_i|X) = 0$. Each observation has the score differential during stint $s \in S$ as the dependent variable Y_s and indicators for each player as the set of independent variables $\mathbf{X} = \{X_m, m \in 0, 1, 2, \dots M\}$. Under this construction, an active player X_i on the home team is given a +1, while active players on the visiting team are given a -1. Players who are inactive during the stint are given a 0. The score differential is then given from the perspective of the home team, such the intercept term β_0 in the regression serves to capture home court advantage. Additional contextual game variables serving as proxies for leverage were also used in the model, improving APM's ability to distinguish player contribution from situational factors. To ensure that the matrix necessary for the ordinary least squares estimate is invertible, Rosenbaum removes players who have extremely low usage rates from the data set.

While using the coefficients β_i for $i = 1, \dots, M$ as estimates for player contribution improved the interpretability and utility of plus/minus methodology, the standard deviation of the estimates were massive, making them extremely noisy. Another glaring issue is that high correlation of values between the players who are on the same team was observed. This happens because players, especially good players, will play together very often, making it difficult to differentiate single contributions from individuals within frequently observed lineups. Otherwise known as *multicollinearity*, this stifles our ability to make statistically significant inferences on differences in the quality of players. Rosenbaum recognized this and made several other amendments to his original method, but the next biggest breakthrough in the multicollinearity and estimation issues was proposed by Joseph Sill in 2010 [59].

Recall that in traditional multiple linear regression, the parameters $\beta_0, \beta_1, \dots, \beta_M$ of the model are estimated by minimizing the following expression, which is known as *Ordinary Least Squares*:

$$\sum_{i=1}^N [y_i - E(y_i)]^2 \quad (4.3)$$

Here the sum of squared errors for each parameter are minimized in the regression to produce a linear unbiased estimator for each model coefficient. Sill proposed the ap-

plication of a Bayesian technique called **Tikhonov regularization**, which is a form of penalized regression where bias is introduced into the model in a controlled manner to combat overfitting and improve accuracy. More commonly known as *ridge regression*, the new estimates of the parameters β_i are found by minimizing

$$\sum_{i=1}^N [y_i - E(y_i)]^2 + \lambda \sum_{i=0}^N [E(y_i)]^2 \quad (4.4)$$

where the added λ term corresponds to the penalty introduced to control the variance parameter estimates. The optimal value of the λ hyperparameter is found in Sill's implementation using a common statistical technique called *cross-validation*, and is equivalent to using a Gaussian prior distribution to constrain the bounds of player contribution estimates within a specified window. This approach came to be known as Regularized Adjusted Plus Minus (**RAPM**), which was demonstrated to be much less sensitive to multicollinearity and no longer required players to be removed from the regression. Since its inception, several other more advanced plus/minus techniques have been introduced which include more detailed sources of data, such real plus/minus [60] and estimated plus/minus [61].

4.2.1 RAPM in Volleyball

The most comprehensive analysis to-date of plus/minus in volleyball was performed by Hass in 2018 [62]. Hass recovered libero court presence from the provided play-by-play data via a Bayesian multinomial model using the maximum *a posteriori* (MAP) estimate. The posterior distribution can then be used to quantify uncertainty that arises due to the incomplete presence information in the computed plus/minus metric. Instead of combining sequences of points by player stints as done in basketball, Hass considered each point played as its own independent entity. Therefore we are now predicting point-win probability, and at a high level the model can be understood as

$$\text{Win probability} \sim \text{contextual variables} + \text{player variables} \quad (4.5)$$

This converts the linear model formulation to a logistic regression, which has a binary outcome of +1 if the team of interest wins a point and a -1 if they lose it. The set of explanatory variables $\mathbf{X}$ remain unchanged from the linear regression definition, except for the fact that it only considers players from a single team and not their opponents due to data constraints. The model now is expressed as:

$$Pr(Y = 1|\mathbf{X}) = F(\beta_0 + \beta_1 X_1 + \dots + \beta_m X_m) \quad (4.6)$$

where F is the cumulative logistic distribution function:

$$F(\beta_0 + \beta_1 X_1 + \dots + \beta_m X_m) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \dots + \beta_m X_m)}} \quad (4.7)$$

Several penalized regression models and their Bayesian analogs are run to compute variants of [APM](#). Hass considers two classes of penalized models - the aforementioned ridge penalty, and a *LASSO* penalty, which sums the absolute value of the regression coefficients instead of squaring them [\[63\]](#). Bayesian and frequentist formulations for each model class are also constructed, and an *elastic net*, a hybrid between the ridge and lasso penalties, is also implemented. The model estimates were compared and contrasted, but generalizable results were rudimentary due to the fact that the data is taken only from a single season’s worth of data from one team. However, their findings were generally in agreement with coach ratings and are purported scale to an entire conference or league.

4.2.2 Model Specification

The rendition of [RAPM](#) used in this work generally follows that of Hass [\[62\]](#) by retaining the Bayesian formulation to enable uncertainty quantification and probabilistic estimates, but with the following key differences:

- Since we have an entire league’s worth of data at our disposal, opposing players are also included in the logistic regression. The model is taken from the perspective of the home team, with outcomes of 1 if they win the point and 0 if they lose it.
- A different set of context estimates are chosen based on the exploratory analysis performed in Chapter [3](#). In particular, variables representing leverage are included.
- Instead of the traditional Gaussian priors used in ridge regression and Laplacian priors used in LASSO, a special shrinkage penalty called the **horseshoe prior** [\[64\]](#) is used on player coefficients.

The sparse matrix $\mathbf{X}$ of player presence indicators maintain values of -1, 0, and 1 for most roles, but use the probabilistic presence estimates for potential liberos as defined in Chapter [3](#). Context variables such as an indicator for the serving team and the rotational

matchup are chosen based on the exploratory analysis performed in Section 4.1. The *set number* and the *score differential*, which proxy for the leverage of the game situation, are also included because it has been demonstrated in sport that teams with a lead tend to play more conservatively than teams who are trailing [65]. These effects are uncertain in volleyball because the sport is not constrained by a time limit - the leading team has to actively score to win. Therefore, even though substitute players are more frequently used in situations where the game is out of reach - also known as “garbage time”, the inclusion of leverage variables would hopefully capture this and discount the points played by their quality. Finally, a binary indicator for whether a point came following a time out or not is included in the model, consistent with considerations made in other literature [66].

The horseshoe prior has flat tails that prevent the overshrinkage of coefficients that are actually meaningful, and an infinitely tall spike at its origin to severely shrink those that aren’t. This flexibility in handling different sparsity patterns provides the prior with robustness that serves well in regressions with a large number of parameters where many of them are not significantly different from 0. There are two types of hyperpriors specified as part of the horseshoe, *local* shrinkage parameters λ_i and a *global* shrinkage parameter τ . τ estimates overall sparsity levels and “pulls” coefficient estimates to 0, while λ_i ’s have a half-Cauchy ($C^+(0, 1)$) distribution that allow for non-zero elements to be recognized and escape the strict shrinkage imposed by τ . The horseshoe prior derives its name from the density plot of the shrinkage weights that it applies, illustrated in Figure 4.2 alongside an image of the density of the prior itself with its recognizable spike at zero.

The model is constructed by applying separate priors to the contextual variables and the player variables, represented as a composite formula with two non-linear components. The horseshoe prior is applied solely to the player variables, while a $Normal(0, 10)$ prior is placed on the contextual variables. This is a relatively flat prior that avoids infinite spike at 0 that the horseshoe imposes due to these variables being chosen based on substantive knowledge. Model predictions for win probability estimates can then be performed for lineups that have not been observed in the data by using the posterior estimates for players that are present in model construction and sampling from the horseshoe prior for players who are not. The model was implemented using the **brms** [67] package in R, with 6000 total iterations, 3000 set aside for burn in. A seed was imposed to ensure consistent initialization and four chains were run in parallel to reduce total run time.

It has been well documented that models in **brms** with horseshoe priors experience divergent transitions, so in this implementation the **adapt_delta** parameter, which controls the target average proposal acceptance rate of samples, is set to 0.98, far above the default value of 0.80. This helps ensure model convergence, at the cost of greatly increasing the computation time. Another technique used to catalyze convergence is performed by

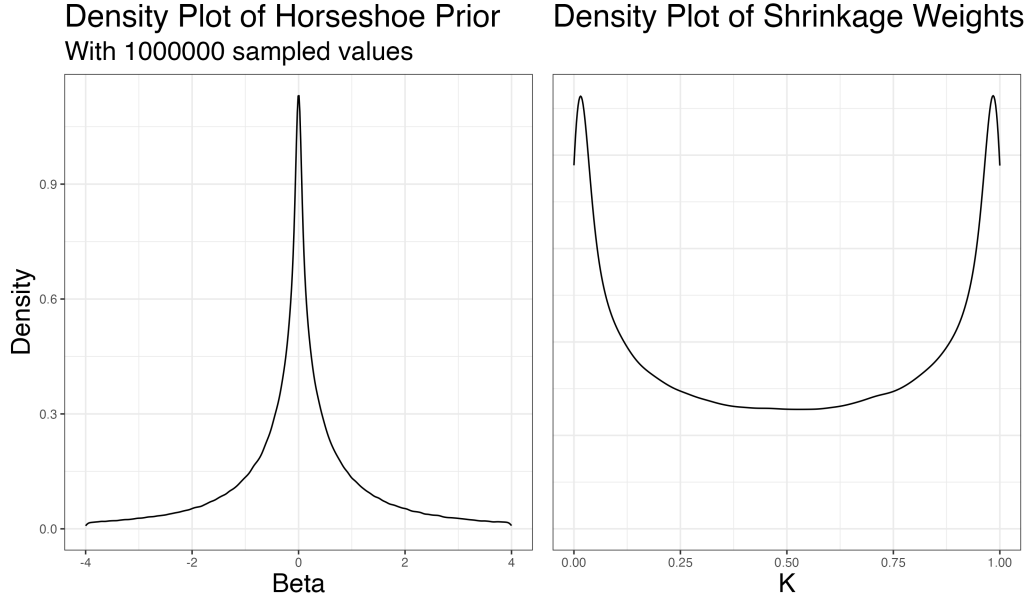


Figure 4.2: Density plot of horseshoe prior shrinkage weights K_i and β estimates

modifying the `par_ratio` hyperparameter, which is the ratio of expected non-zero to zero coefficients [68]. This is interpreted as our educated guess on the number of players who are statistically better or worse than an “average” or “replacement-level” player. Based on the intuition provided in Section 4.1.1, the selected ratio is 0.25.

4.2.3 Discussion

Even with the remedial measures employed to improve convergence, the final model experienced a handful of divergent transitions. Nevertheless, the supplementary information provided on the algorithm’s ability to estimate the posterior distribution of each parameter had good results, so we proceeded with additional analysis. Figure 4.3 displays a density plot of the player RAPM estimates. Estimates are filtered for players who appear in at least 8% of their team’s games across the data (~ 600 points played). Both the mean and median are very close to zero, with a longer tail towards the negative values.

To build confidence about the validity of our model output, we are interested in assessing how well our model fits both the data and our domain knowledge. We start by comparing the players with the largest RAPM estimates to our knowledge of USPORTS and OUA

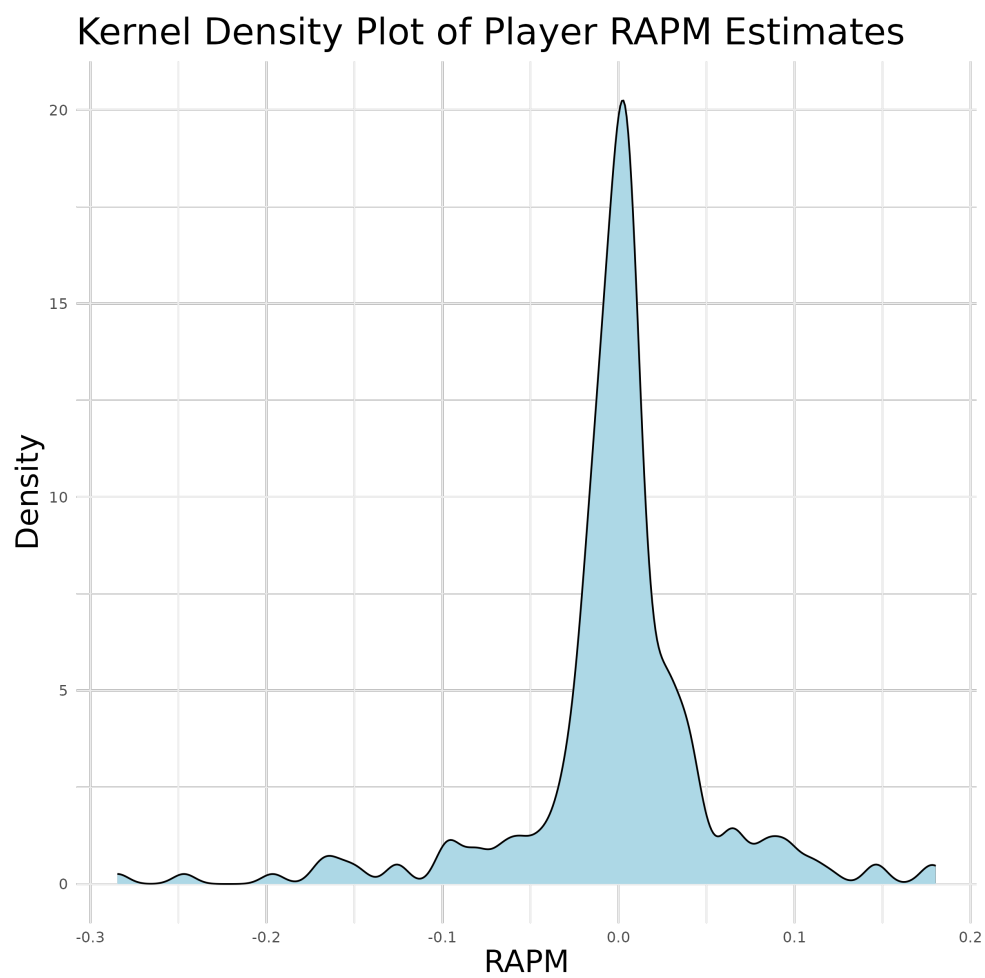


Figure 4.3: Density Plot of Player RAPM Estimates

Trace Plots of RAPM for Top 4 Players

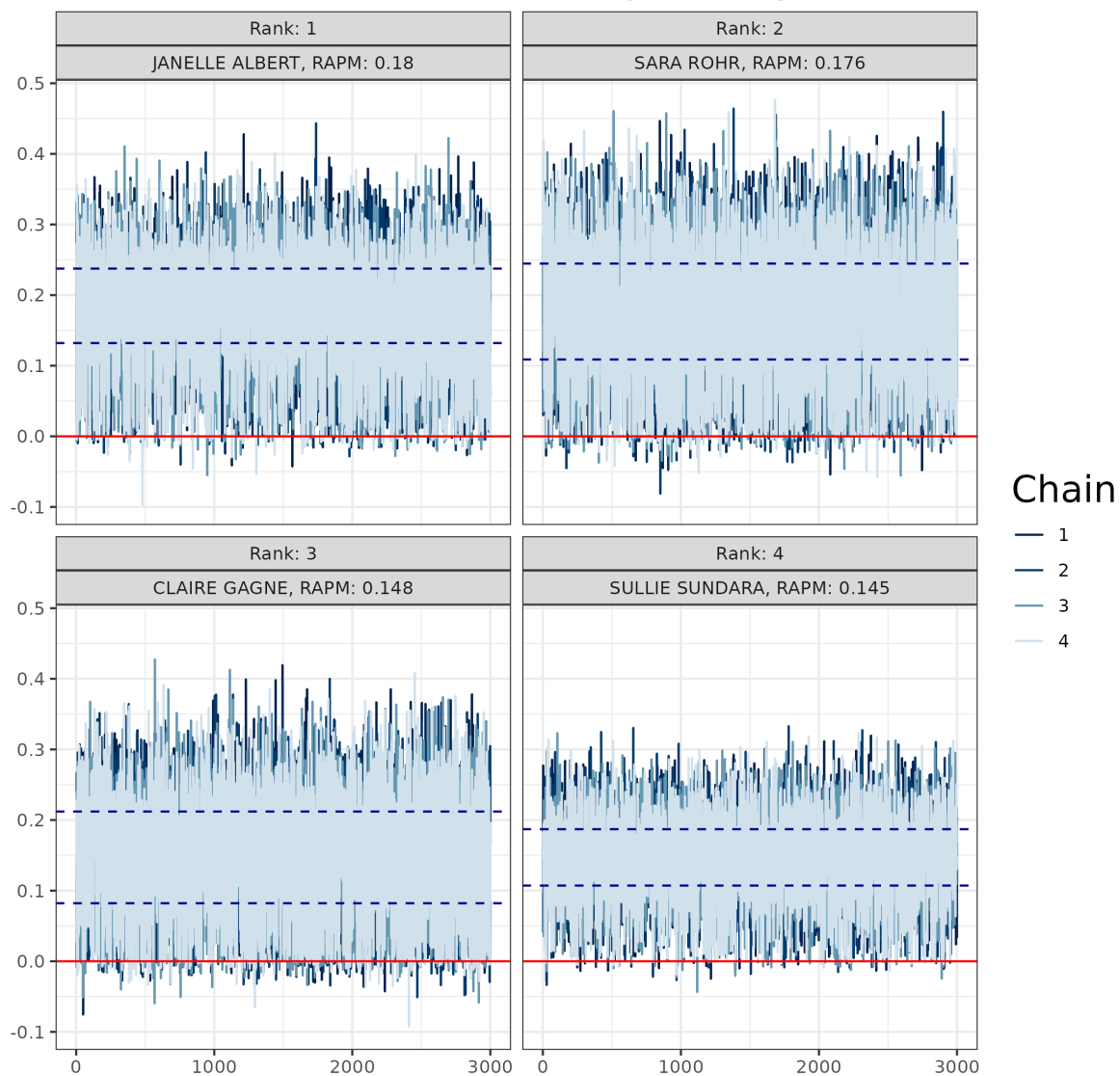


Figure 4.4: Trace plots of top 4 OUA players, by RAPM

award winners. Figure 4.4 shows the trace plots for the top four overall players in the OUA by RAPM. The blue dashed lines show the 50% credible interval for the player estimates, and they are all comfortably different from 0, which is shown as the solid red line in each figure. For those familiar with the league, these names come as no surprise. Janelle Albert, an outside hitter from the University of Western Ontario and back-to-back winner of the OUA Most Valuable Player in 2021-22 and 2022-23, leads all players in RAPM. She is followed closely by setter Sarah Rohr, the 2024 OUA Female Athlete of the Year and 2024 USPORTS First-Team All-Canadian. Rohr played an integral role in the dominance of the Brock Badgers over the conference during her illustrious university career. Claire Gagne, also a setter, led the 2022-23 Waterloo Warriors to a 16-win and 4-loss regular season record, being named an OUA First-Team All-Star in the process. Sullie Sundara, an outside hitter with the McMaster Marauders, was the 2021-22 OUA Rookie of the Year and was named to 3 consecutive OUA First-Team All-Star teams as well.

Table 4.5: RAPM player contribution estimates, separated by player role

Rank	RAPM	Athlete Name	Team	50% CI
Outside Hitter				
1	0.18	Janelle Albert	Western Mustangs	[0.13, 0.24]
2	0.14	Sullie Sundara	McMaster Marauders	[0.11, 0.19]
3	0.11	Madison Chimienti	Brock Badgers	[0.04, 0.17]
4	0.10	Olivia Kim	York Lions	[0.03, 0.15]
5	0.10	Julia Murmann	Toronto Varsity Blues	[0.05, 0.14]
Setter				
1	0.18	Sara Rohr	Brock Badgers	[0.11, 0.24]
2	0.15	Claire Gagne	Waterloo Warriors	[0.08, 0.21]
3	0.11	Hayley Goodwin	Toronto Varsity Blues	[0.01, 0.20]
4	0.08	Jenna Peters	Toronto Varsity Blues	[0.01, 0.14]
5	0.07	Shea Baker	Queen's Gaels	[0.01, 0.11]
Libero				
1	0.12	Vivian Zanutel	Queen's Gaels	[0.06, 0.17]
2	0.09	Aleiah Torres	Brock Badgers	[0.03, 0.15]
3	0.09	Jenna Gravina	Waterloo Warriors	[0.01, 0.16]
4	0.09	Sophia Currier	Toronto Varsity Blues	[0.00, 0.15]
5	0.07	Holly Biasi	Queen's Gaels	[0.01, 0.12]

Middle Blocker				
1	0.08	Christina Piccinin	York Lions	[0.02, 0.13]
2	0.04	Ellie Hatishita	McMaster Marauders	[0.00, 0.07]
3	0.04	Jordyn Brearley	Windsor Lancers	[0.00, 0.06]
4	0.03	Ana Strbac	McMaster Marauders	[0.00, 0.05]
5	0.03	Alexandra Iggulden	Western Mustangs	[0.00, 0.05]
Right Side				
1	0.07	Julia Typa	York Lions	[0.01, 0.11]
2	0.05	Madi Lohmann	Trent Excalibur	[0.00, 0.08]
3	0.04	Maddy Lutes	McMaster Marauders	[0.00, 0.07]
4	0.03	Karen Fan	Waterloo Warriors	[0.00, 0.06]
5	0.02	Jessica Chung	Guelph Gryphons	[0.00, 0.04]

We now evaluate [RAPM](#) by player role to see if the contextual variables included to account for serving patterns and rotational matchups are able to normalize the estimates for each role. These results for the top five players are displayed in Figure 4.5, alongside the 50% credible intervals. Full [RAPM](#) results can be found in Appendix A. We see that outside hitters and setters have the largest estimates, with non-zero credible intervals as well. The estimates for middle blockers are still quite low, which implies that we are still not able to completely capture the patterns present in lineup construction, even with the covariates included to account for situational decision-making. We also find that estimates with a mean of less than 0.04 rarely have non-zero credible intervals unless the sample size is very large. The estimates for right side players were also surprisingly low compared to other positions, which could be evidence of another pattern not captured in model results. In conclusion, we see that although there are some nuances within volleyball not captured in the model, [RAPM](#) is able to do a much better job at differentiating players and giving results in line with substantive knowledge than simply the raw and scaled plus/minus.

4.3 Posterior Predictive Checks

To further test model output, we perform a series of graphical and probabilistic validation tests, known as **posterior predictive checks** [69]. While conceptually straightforward and relatively intuitive to implement, such checking is imperative for mitigating the adverse affects on inference that poorly specified models may contain. We start by drawing

simulated values from the joint posterior predictive distribution of replicated data and comparing these samples to observed data, looking for systematic differences between the simulations and data. The general idea is that a well-fitted model should be able to generate data similar to the observed data. We define *average point-win probability* as the test statistic against which model output will be initially checked. We also look at the posterior predictive p-values for this check, which are defined as the probability that the replicated data could be more extreme than the observed data.

The distribution of mean win-probabilities from 500 samples of the posterior predictive distribution is displayed in Figure 4.5. Here we see that the observed mean point-win probability of 0.510 is right in the middle of the posterior predictive distribution, with a predictive p-value of 0.49. The 95% credible interval is [0.504, 0.515]. However, since `mean` is a sufficient statistic of this quantity there are likely better alternatives to evaluate model efficacy. One such example is comparing the predicted *point-spread*, the difference in points between two teams as a set progresses, to observed values. This helps us understand how well our model can capture individual point fluctuations in win-probability estimates. Since we use contextual data directly from the observed data, we manually cap simulated sets in accordance with volleyball rules to ensure that the point bounds are realistic.

After simulating point spread from the same 500 samples used to compute mean win-probability, we consider three test statistics - maximum point spread, its standard deviation, and mean point spread. The distribution of maximum point spread is displayed in Figure 4.6, where the mode of this spread is 22 points, one more than the observed maximum point spread of 21 points. The Bayesian p-value of this quantity is 0.404. While this is also an encouraging sign, we find that the standard deviation and mean point spreads of the simulated values are consistently wider and larger than those from the observed data, with Bayesian p-values close to 0. Put another way, sets are predicted to be more lopsided than what is observed, and this occurs across simulations. This could be due in part to the simulation construction, which requires special care and numerous assumptions to faithfully reproduce situational context such as substitutions.

Finally, we perform graphical model checking, where we are interested looking at a direct display of the data with summaries of the data/parameter inferences, and plotting the residuals and other measures of discrepancies between the model and data. Here, graphical displays are generated using the `bayesplot` [70] package in R. Figure 4.7 shows the binned residual plots for 9 samples of replicated data. Each dot is a predictive error, and there seem to be no systematic patterns in over-estimates or under-estimates. Overall, our posterior predictive checks demonstrate that while not perfect, our model does a reasonable job of representing the data.

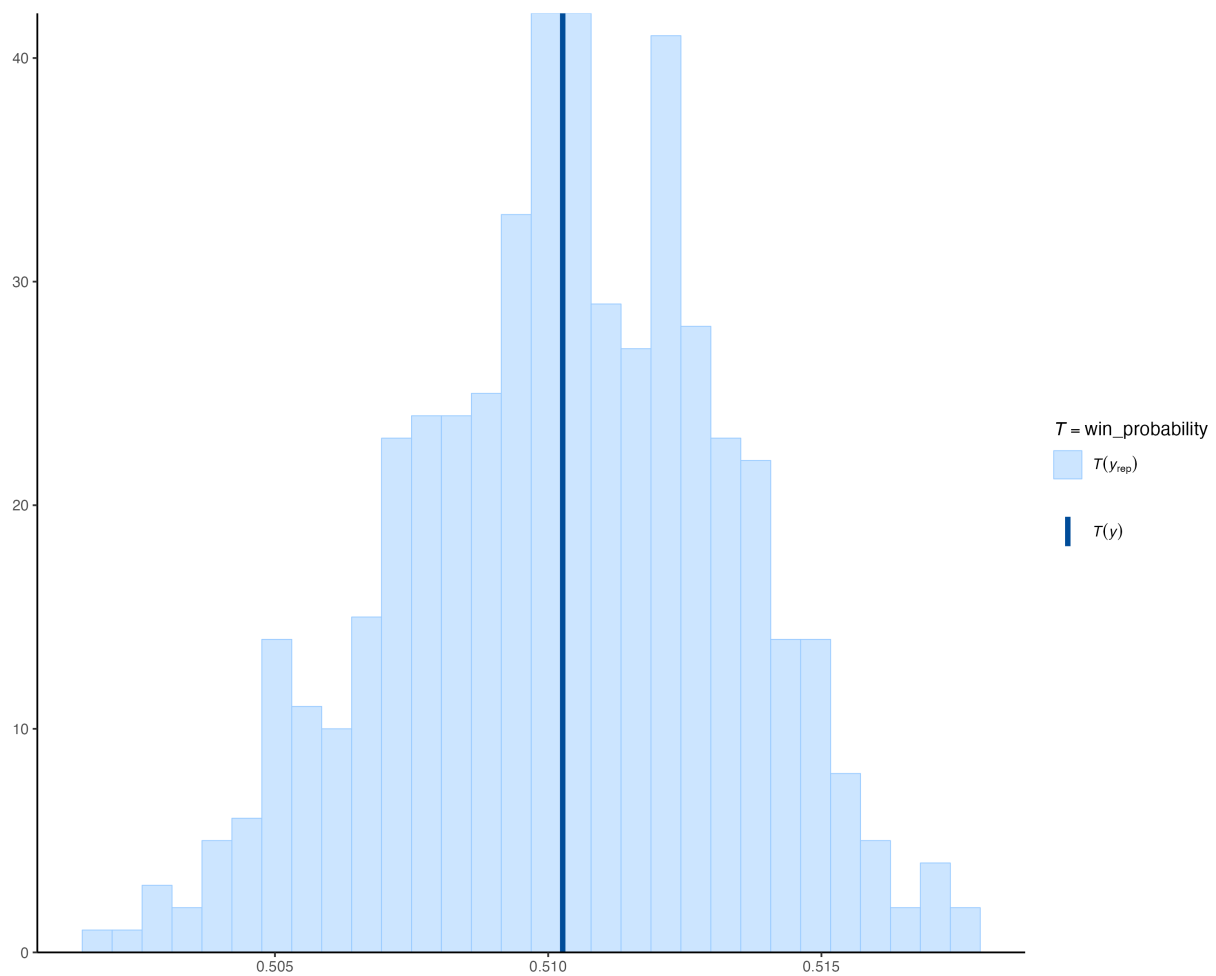


Figure 4.5: Simulated point-win probability versus observed data

Maximum Point-Spread

500 simulations, dark line is from observed data

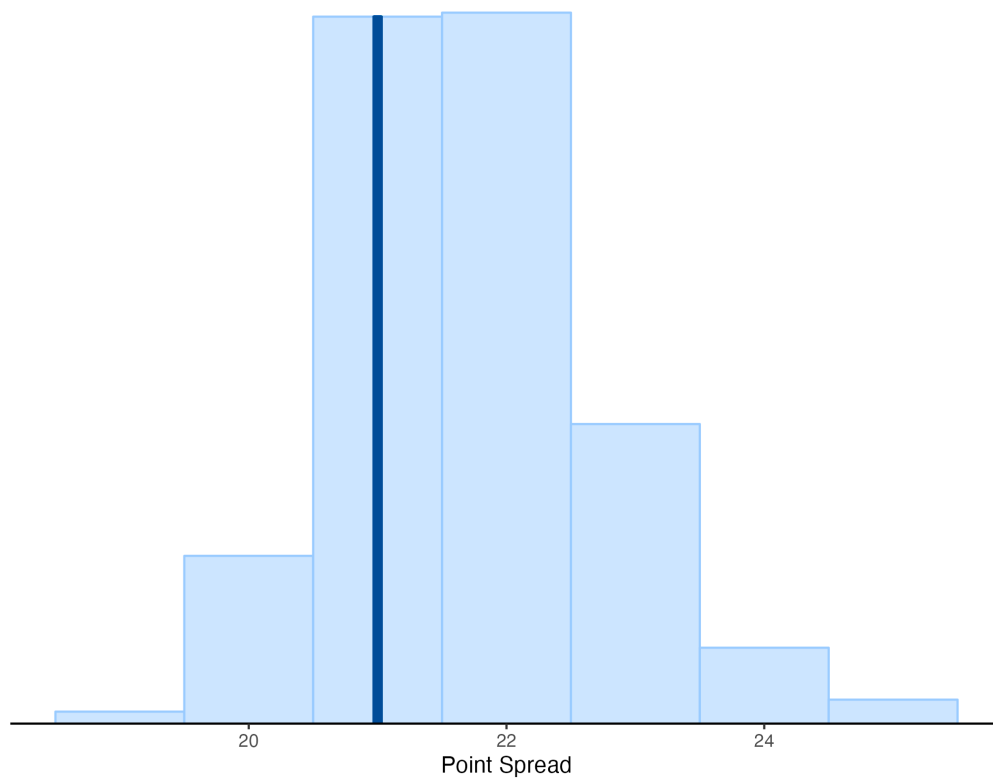


Figure 4.6: Distribution of simulated maximum point spread versus observed data

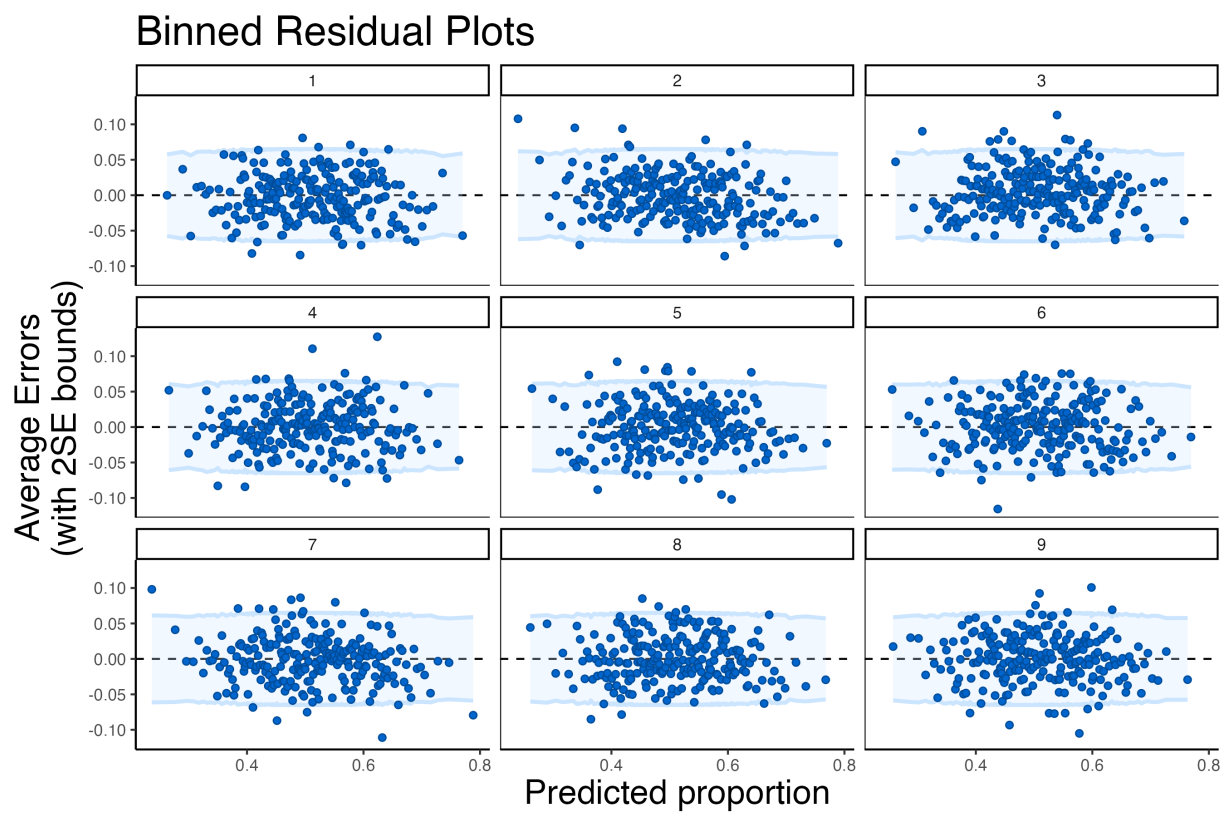


Figure 4.7: Binned Residual Plots

Chapter 5

Conclusion

Understanding contribution and appropriately dividing credit between players is a very interesting and relevant problem in sports analytics. Insights generated from such analyses can be applied to a broad set of disciplines, including salary allocation, roster construction, advance scouting, and role assignment. The methods developed to address these types of questions must consider the practical significance of the findings, the technical contribution of the statistical methodology, and relevant data constraints. Volleyball is one of the most popular sports in the world, and its intricacies and underlying structure motivate a wealth of intriguing analytical questions. The literature surrounding player contribution models in volleyball focus primarily on performance analysis and skill importance. Early works analyzed winning styles of play within teams for match discrimination, while other approaches have leveraged graph theory and machine learning to investigate specific in-game components. Recent research advances have navigated away from team performance and honed in on player evaluation, developing novel metrics that attempt to isolate individual contribution in a way that is comparable within positions and independent of player role.

The league of interest was Ontario University Athletics ([OUA](#)), the largest of the four conferences in University Sports ([USPORTS](#)), the leader of university sport within Canada. Event-data from every regular season game played in the 2021-22, 2022-23, and 2023-24 seasons was extracted and extensive data cleaning was performed to consolidate player identities and produce a dataset with nearly 56,000 points played and $\sim 350,000$ player contacts. Remedial measures were performed in the acquired data to infer player roles and probabilistically capture libero presence on the court. These measures leveraged common patterns in volleyball gameplay as well as league-specific rules surrounding allowable libero actions. In this work, we investigated player contribution from the lens of player presence, one of six identified components of contribution within a unified and comprehensive

framework. Player presence models attempt to capture individual contribution without specifying the means by which these contributions take place, enabling them to capture qualitative pieces of information as well. The foundation of those models used here is plus/minus, a statistic calculated by taking the difference in points for and points against for a team while a particular player is active. Exploratory analysis of raw and scaled plus/minus provides insight into the impact of different game scenarios on point differential. An example of this is the advantage of receiving teams in point-win probability and the effect that this has on player roles such as liberos and middles, who have unequal distributions of points played when their team is serving and receiving and therefore inherently different distributions in plus/minus.

Analysis of variance (ANOVA) and Tukey’s honest significant differences (Tukey’s HSD) provide quantification to the marginal differences in these levels and motivate extensions to plus/minus models due to the unclear ability to differentiate between individual players and account for context. This inspired the implementation of the regression-based Adjusted Plus/Minus (APM), which accounts for situational context and opponent effects but suffers from noisy predictions and multicollinearity in coefficient estimates. These issues are addressed by Regularized Adjusted Plus/Minus (RAPM), a family of penalized regression models much less sensitive to correlated estimates and small sample sizes. This is where plus/minus literature in volleyball currently stands [71]. Here we develop a Bayesian RAPM model using a horseshoe prior [72] instead of the traditional Laplacian and Gaussian priors utilized in other literature. The horseshoe is able to flexibly shrink coefficients without meaningful impact on estimates while simultaneously maintaining the magnitudes of those that do matter. From the model output we find that the output for players at the top match substantive knowledge, while still providing enough new information to differentiate it from existing statistics. However, we are still unable to fully account for substitution patterns that introduce bias into libero and middle estimates, and also find unexpectedly low right side player estimates as well. Finally, we performed posterior predictive checks and assessed output from other graphical approaches to validate the model specification and identify the independent information that it provides.

5.1 Future Extensions

We now turn our attention towards how we can quantify player contribution using the other 5 “W” questions. For each component, we walk through relevant research and potential extensions that could be applied, and postulate on applications to volleyball analysis.

5.1.1 What - Player Action

Although player presence models bring an elegant simplicity due to their relative ease of calculation and data collection, they are unable to discern *what* makes a player valuable. For this we need to upgrade our data collection to account for player actions by acquiring match summary data. Skill importance has been researched in previous works, but has focused more so on team tactics than how they constitute an individual player’s contribution. Isolating and aggregating individual player actions allows us to define the notion of *efficiency* in different contexts as the proportion of effective actions over ineffective ones. Similar to player presence models, some notion of win probability is typically chosen as the outcome of interest, and in volleyball this is evaluated on a per-point or per-set basis. At a high level, Fellingham’s Points Above Average Per Set (PAAPS) [73] can be understood as an [RAPM](#) regressed on player contacts instead of simply who was present. This adds more context to contribution in terms of the quantity of contacts and how those influence a team’s possession outcome, but still fails to discriminate between different contact types and how their credit is distributed between players.

Therefore, an alternative approach that could have potential utility in volleyball is an extension of [RAPM](#) called [Box Plus/Minus \(BPM\)](#) [74]. [BPM](#) was created to combine the benefits of player presence and player action frameworks, while remaining incredibly accessible to many levels of play due to its reliance on match summary statistics instead of sequential event-data. This reliance on means that it is far from comprehensive, but it does a solid job. [BPM](#) estimates player contribution in points above league average, where the base model regresses a player’s [RAPM](#) on their match summary statistics. Adjustments to player output are made to account for overall team quality, the usage rate of a player, and their contribution relative to the role that they have. The quality of [BPM](#) is directly contingent on how well the chosen set of match summary statistics capture the skills necessary for success in the sport it’s applied to. In basketball, where defensive proficiency cannot be recorded as effectively as offensive statistics in boxscores, [BPM](#) does well at measuring offense but does not give definitive defensive numbers. Applied to volleyball, finer grain statistics with demonstrated value such as block deflections and pass quality are not documented in traditional boxscores and may need to be collected through other means for defensive numbers generated by [BPM](#) to be more insightful. Additional modifications are necessary to translate the basketball-specific implementation of [BPM](#) to one that accounts for the associated nuances of volleyball, but [BPM](#) nevertheless provides an exciting opportunity to extend the plus/minus framework into the domain of player action and improve our notion of player contribution.

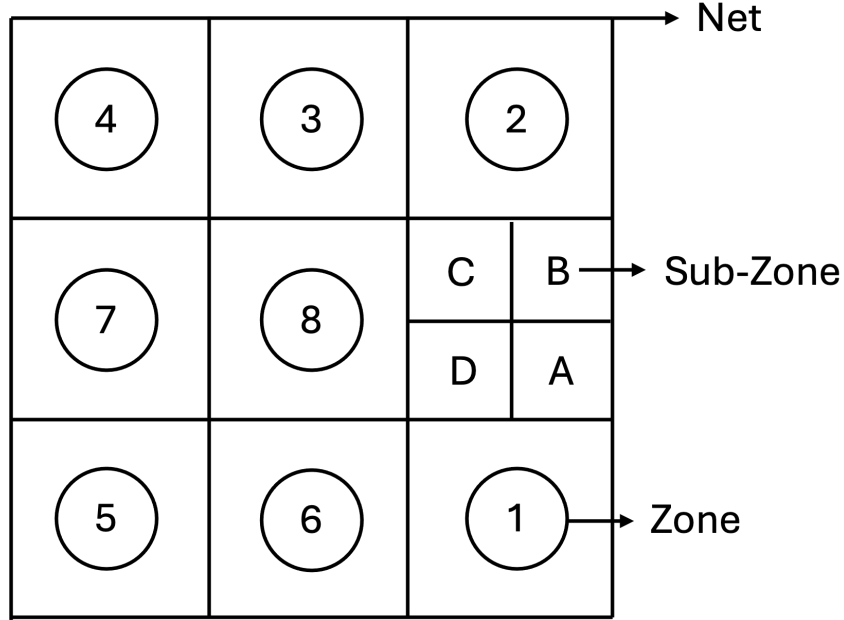


Figure 5.1: DataVolley Zone and Sub-Zone Definitions

5.1.2 Where - Coordinate Data

Spatial analysis in sport has seen many advances with the increasing availability of coordinate data. Such information provides contextual awareness to model predictions and helps reveal spatial patterns that are not apparent from match summary data alone. An example of this is an algorithm developed by Cantú-González et al. [75], where kinematic equations of motion for the ball and three-dimensional coordinates of attackers and blockers are modeled to provide coaches the ability to simulate attack outcomes based on customized team defensive structures. Unfortunately, with event-level data such as that used in this work, the coordinates are relatively coarse. Scouts discretize a volleyball court into 9 one-meter by one-meter squares, with some going a step further and dividing these squares into four equally sized quadrants. This discretization is illustrated in Figure 5.1.

Each contact is then given a zone and potentially a sub-zone identifier for its start and end location, from which visualizations such as heatmaps can be generated. There is inherent variability with this method as balls that are blocked or go into the net will often either be given a missing value for their end location or have it inferred based on the scout’s intuition of where the ball *would have* landed. Such analysis would be greatly improved by higher resolution coordinate data such as that provided by **tracking data**. Tracking data

uses computer vision or motion capture software to ascertain the x, y, and z coordinates for objects of interest at regular time intervals. This has revolutionized analyses in other sports [16], and approaches leveraging computer vision from Gordon [76] and Raymond [77] have shown tremendous potential for volleyball applications. Integration of tracking data within volleyball analytics can take inspiration from analogous metrics used in baseball such as “jump” [78], which measures a defender’s reaction time and route to intercept a ball. This would also allow us to model adversarial contacts in a much better way since we would know the locations of all players on the court at once and not just that of the player making the contact.

5.1.3 How - Action Quality and Physical Attributes

How individual players contribute to team success can be considered via action quality and the physical characteristics of these actions. Action quality is a hotly contested area within the volleyball community due to the inherently subjective assignment of ratings and their non-mathematical foundation. Actions are typically evaluated on a 0-4 point scale based on the situation that they put their team in, where a higher numerical value corresponds to a higher-quality action. One major flaw of the current approach is that these values are determined at the discretion of the scout in charge of coding the game, which may differ from a coach’s personal definitions of quality. Additionally, a numeric scale assumes linearity between values, implying for example that an action with an assigned quality of 4 is twice as valuable as one with a quality of 2. However, numerous studies have demonstrated that this is not the case, which has motivated implementations like that of Bagley [79] to provide more objective methodologies of grading action quality that scale appropriately as well.

Another way to understand the “how” of player contribution is by making anthropometric, biomechanical, and physiological assessments of physical attributes. With this we can consider their physical attributes to understand which components play a substantial role in their demonstrated skill. For example, if a player is a very good blocker, is this because they are naturally very tall or have an unusually long wingspan? Is it because they have above-average jumping ability through force generation? Or is it due to superb eye-hand coordination and proprioception for the player to be able to time a block perfectly and position their hands accordingly? These questions can be addressed using specialized techniques and instruments commonly used in sports science disciplines, such as inertial measurement units (IMU’s) and joint and pose estimation. The high costs associated with adopting these technologies, however, inhibits their widespread adoption, limiting the feasibility of such analyses by most programs.

5.1.4 When - Event Sequences

By taking player actions and evaluating them as they occur sequentially during a point, we are able to incorporate leverage into analysis of player contribution. Analysis of sequential actions also allows us to understand team strategy and recognize trends in specific player actions. Instead of aggregating player actions, as performed in [BPM](#) and spatial visualizations, we analyze the sequence of actions that lead to a particular outcome. For example, set distribution - how a setter distributes attack opportunities to their teammates during different phases of game play - can be analyzed to inform blocking and defensive strategies. The volume of repetition, as well as any patterns in setting tendencies from a team's first attack opportunity in a point versus in their subsequent possessions can be gleaned by breaking down setter decisions temporally. This is an ideal use case for event-level data, and similar studies using Markov chain implementations by Miskin [\[80\]](#) and most recently by Powers [\[81\]](#) have introduced major improvement in individual skill analysis.

Powers [\[81\]](#) modeled the progression of a point as a Markov chain with a state space defined by the sequence of contacts in the current possession. Changes in win probability calculated at every contact in the chain are attributed to each player, with adjustments made based on the “strength of schedule” - the difficulty of executing the contact based on the quality of competition faced by the player. These include a hierarchical linear mixed effects model for adversarial contacts such as serves and attacks, with random effects for the players involved. The resulting base metric, called “Points Gained” (PG), provides a way by which players from different positions can be accurately compared in terms of points. This lends itself to interesting case studies in player recruitment and strategic interventions, such as the benefits of using a defensive specialist in place of an outside hitter in the back row. While demonstrating a major advancement in individual player performance evaluation, this approach does not account for explicit spatial information that would undoubtedly improve player estimates, nor does it completely eliminate bias found in evaluation codes in certain situations, such as when a setter makes a dig. Additionally, expanding the state space of the Markovian transition matrix could also potentially refine model output.

5.1.5 Why - Player Intent

Understanding player intent is a fascinating concept that is by far the most difficult attribute to capture of the 6 “W’s”. By creating a metric to capture player intent, we can then define *skill* as a player’s ability to translate concerted intent into an executed outcome. We

can also devise a metric of “luck” calculated as the difference in win probability, for example, between the predicted outcome and the realized outcome. A “lucky” player is then one who makes low-efficiency decisions but experiences better-than-expected outcomes, while an “unlucky” player is one who makes high-quality decisions but has lower-than-expected outcomes. The function used to model this translation from intent to execution can be conceptually difficult to formulate, but some game actions in volleyball follow repeatable patterns that can potentially be modelled under this framework. For example, in an ideal scenario for a team that receives a serve or an attack, the second contact in their possession should be taken by the setter near between zone 2 and 3. Therefore, the ability of a player who receives a serve to accurately pass the ball to the setter is one way that we can measure their passing skill, because we know (or can infer with high degrees of confidence based on typical game patterns) the intended result of their action.

Another potential application of player intent models in volleyball could take a format similar to what is now possible in baseball via the invention of PitchComTM. PitchComTM is a portable device worn by baseball players that communicates which pitch a pitcher should throw. The ability to export the audio files from these devices and compare the instructed location to that which was performed throughout the course of a game can provide a quantitative notion of pitcher skill. In volleyball, an associated representation can be implemented for a player’s serving target. By recording the accuracy and other intended characteristics of a serve before it is initiated, analysis can be performed to differentiate between poor execution and poor intention.

References

- [1] Shouvik Sanyal and Mohammed Hisam. The Impact of Teamwork on Work Performance of Employees: A Study of Faculty Members in Dhofar University. 20, March 2018.
- [2] Zachary J Hass. Division of Credit Modeling for Team Sports with an Emphasis on NCAA Volleyball. *Theses and Dissertations Available from ProQuest*, pages 1–161, January 2017.
- [3] Thuy Ngoc Nguyen, Chase McDonald, and Cleotilde Gonzalez. Credit Assignment: Challenges and Opportunities in Developing Human-like AI Agents, July 2023. Comment: 11 figures; 3 tables.
- [4] Saiem Gilani. Sportsdataverse: The Set of R Packages for Sports Data. <https://r.sportsdataverse.org>. R package version 0.2.0.
- [5] James Cochran. The Emergence of Sports Analytics | Analytics Magazine. <https://pubsonline.informs.org/doi/10.1287/LYTX.2010.01.06/full/>. (accessed 2024-08-04).
- [6] Johann Windt, David Taylor, Dean Little, and Ben C Sporer. Making everyone’s job easier. How do data scientists fit as a critical member of integrated support teams? *British Journal of Sports Medicine*, 55(2):73–75, January 2021.
- [7] Edgar Santos-Fernandez, Kerrie L. Mengersen, and Paul Wu. Bayesian Methods in Sport Statistics. In N. Balakrishnan, Theodore Colton, Brian Everitt, Walter Piegorsch, Fabrizio Ruggeri, and Jozef L. Teugels, editors, *Wiley StatsRef: Statistics Reference Online*, pages 1–8. Wiley, 1 edition, February 2019.
- [8] Quang Nguyen, Ronald Yurko, and Gregory J. Matthews. Here Comes the STRAIN: Analyzing Defensive Pass Rush in American Football with Player Tracking Data. *The American Statistician*, 78(2):199–208, April 2024. Comment: 12 figures, 7 tables.

- [9] Sarthak Batra. The Role of Data Analytics in Modern Day Sports. 2017.
- [10] Johann Windt, Kerry MacDonald, David Taylor, Bruno D. Zumbo, Ben C. Sporer, and David T Martin. “To Tech or Not to Tech?” A Critical Decision-Making Framework for Implementing Technology in Sport. *Journal of Athletic Training*, 55(9):902–910, September 2020.
- [11] Naveen Peter. History of volleyball: From origins to the Olympics. <https://olympics.com/en/news/what-history-volleyball-game-origin-mintonette-ymca-fivb-olympics>, February 2023. (accessed 2024-08-04).
- [12] Mesquita Isabel, F Manso, and José Palao. Defensive participation and efficacy of the libero in volleyball. *Journal of Human Movement Studies*, 52:95–108, January 2007.
- [13] Lorena Torres-Ronda, Emma Beanland, Sarah Whitehead, Alice Sweeting, and Jo Clubb. Tracking Systems in Team Sports: A Narrative Review of Applications of the Data and Sport Specific Analysis. *Sports Medicine - Open*, 8(1):15, December 2022.
- [14] Alexander Franks, Andrew Miller, Luke Bornn, and Kirk Goldsberry. Characterizing the spatial structure of defensive skill in professional basketball. *The Annals of Applied Statistics*, 9(1), March 2015.
- [15] Jonah Gabry, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. Visualization in Bayesian Workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 182(2):389–402, February 2019.
- [16] Zachary Turner and Alexander Franks. Modeling Player and Team Performance in Basketball. *Annual Review of Statistics and Its Application*, 8(1):1–23, March 2021.
- [17] Daniel Cervone, Alex D’Amour, Luke Bornn, and Kirk Goldsberry. A Multiresolution Stochastic Process Model for Predicting Basketball Possession Outcomes. *Journal of the American Statistical Association*, 111(514):585–599, April 2016.
- [18] Jack Davis, Harsha Perera, and Tim B. Swartz. Player evaluation in Twenty20 cricket. *Journal of Sports Analytics*, 1(1):19–31, September 2015.
- [19] Kerrie L. Mengersen, Christopher C. Drovandi, Christian P. Robert, David B. Pyne, and Christopher J. Gore. Bayesian Estimation of Small Effects in Exercise and Sports Science. *PLOS ONE*, 11(4):e0147311, April 2016.

- [20] Alexander M. Franks, Alexander D’Amour, Daniel Cervone, and Luke Bornn. Meta-analytics: Tools for understanding the statistical properties of sports metrics. *Journal of Quantitative Analysis in Sports*, 12(4):151–165, December 2016.
- [21] Baseball Almanac Inc. Range Baseball Dictionary | Baseball Almanac. <https://www.baseball-almanac.com/dictionary-term.php?term=range>. (accessed 2024-08-04).
- [22] Joachim Gudmundsson and Michael Horton. Spatio-Temporal Analysis of Team Sports. *ACM Computing Surveys*, 50(2):1–34, March 2018.
- [23] Felicity Lord, David B Pyne, Marijke Welvaert, and Jocelyn K Mara. Methods of performance analysis in team invasion sports: A systematic review. *Journal of Sports Sciences*, 38(20):2338–2349, October 2020.
- [24] Miguel Silva, Rui Marcelino, Daniel Lacerda, and Paulo Vicente João. Match Analysis in Volleyball: A systematic review. 2016.
- [25] Julie Anna Street. *APPLYING SPORTS ANALYTICS TO DIVISION 1 WOMEN’S VOLLEYBALL: SERVE RECEIVE PASS RATING RELATIVE TO SERVER LOCATION, SPEED OF THE SERVE, LOCATION OF THE PASSER, AND PASSING ANGLE*. PhD thesis, Miami University, 2018.
- [26] Miguel Silva, Daniel Lacerda, and Paulo Vicente João. Match analysis of discrimination skills according to the setter attack zone position in high level volleyball. *International Journal of Performance Analysis in Sport*, 13(2):452–460, August 2013.
- [27] G. Giatsis, S. Drikos, and A. Lola. Analysis of match report indicators in men’s volleyball Olympics and world championships (2014–2021) depending on the type of final score. *International Journal of Sports Science and Coaching*, 18(3):874–882, 2023. Cited By :6.
- [28] J.M. Palao, J.A. Santos, and A. Ureña. Effect of team level on skill performance in volleyball. *International Journal of Performance Analysis in Sport*, 4(2):50–60, December 2004.
- [29] Han Joo Eom and Robert W. Schutz. Statistical Analyses of Volleyball Team Performance. *Research Quarterly for Exercise and Sport*, 63(1):11–18, March 1992.
- [30] Eduardo López, Ignacio Díez-Vega, and Juan José Molina. Effect of serve on reception and side-out performance in relation to ball’s contact with the net and type of serve, in high level male volleyball. *RICYDE. Revista internacional de Ciencias del Deporte / The International Journal of Sport Science*, 17(63):56–68, 2021.

- [31] Sotirios Drikos and George Vagenas. Multivariate assessment of selected performance indicators in relation to the type and result of a typical set in Men’s Elite Volleyball. *International Journal of Performance Analysis in Sport*, 11, January 2011.
- [32] Jara González-Silva, Carmen Fernández-Echeverría, Manuel Conejero, and M. Perla Moreno. Characteristics of Serve, Reception and Set That Determine the Setting Efficacy in Men’s Volleyball. *Frontiers in Psychology*, 11:222, February 2020.
- [33] V. Luis-Del Campo, E. Cabeza, and J.S. Damas. KNOWLEDGE OF OPPOSITION ATTACKING TENDENCIES ON VOLLEYBALL BLOCKING. *Revista Internacional de Medicina y Ciencias de la Actividad Fisica y del Deporte*, 22(88):845–861, 2022.
- [34] Aikaterini Oikonomopoulou, Karolina Barzouka, Konstantinos Sotiropoulos, Sotirios Drikos, and Konstantinos Noutsos. Spatiotemporal analysis of setting per game complex and origin of the ball in junior female volleyball players. *Journal of Physical Education and Sport*, 22(3), March 2022.
- [35] Marta Hurst, Manuel Loureiro, Beatriz Valongo, Lorenzo Laporta, Pantelis Nikolaidis, and José Afonso. Systemic Mapping of High-Level Women’s Volleyball using Social Network Analysis: The Case of Serve (K0), Side-out (KI), Side-out Transition (KII) and Transition (KIII). *International Journal of Performance Analysis in Sport*, 16:696–710, July 2016.
- [36] Manuel Loureiro, Marta Hurst, Beatriz Valongo, Pantelis Nikolaidis, Lorenzo Laporta, and José Afonso. A Comprehensive Mapping of High-Level Men’s Volleyball Gameplay through Social Network Analysis: Analysing Serve, Side-Out, Side-Out Transition and Transition. *Montenegrin Journal of Sports Science & Medicine*, 6:35–41, July 2017.
- [37] Lorenzo Laporta, José Afonso, and Mesquita Isabel. Interaction network analysis of the six game complexes in high-level volleyball through the use of Eigenvector Centrality. *PLOS ONE*, 13:e0203348, September 2018.
- [38] Alexandre Igor Araripe Medeiros, José Manuel Palao, Rui Marcelino, and Isabel Mesquita. Systematic review on sports performance in beach volleyball from match analysis. *Revista Brasileira de Cineantropometria & Desempenho Humano*, 16:698–708, September 2014. Publisher: Universidade Federal de Santa Catarina.
- [39] Michelle A. Miskin, Gilbert W. Fellingham, and Lindsay W. Florence. Skill Importance in Women’s Volleyball. *Journal of Quantitative Analysis in Sports*, 6(2), April 2010. Publisher: De Gruyter.

- [40] Corley Bagley and Brett Ware. Bump, Set, Spike: Using Analytics to Rate Volleyball Teams and Players. *MIT Sloan Sports Analytics Conference 2017*, 2017.
- [41] M.Á. Casimiro-Artés, R. Hilenó, and A. García-de Alcaraz. Applying Unsupervised Machine Learning Models to Identify Serve Performance Related Indicators in Women’s Volleyball. *Research Quarterly for Exercise and Sport*, 2023.
- [42] Xianyan Dai and Shangbin Li. Volleyball Data Analysis System and Method Based on Machine Learning. *Wireless Communications and Mobile Computing*, 2021:1–11, May 2021.
- [43] Jan Van Haaren, Horesh Ben Shitrit, Jesse Davis, and Pascal Fua. Analyzing Volleyball Match Data from the 2014 World Championships Using Machine Learning Techniques. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 627–634, San Francisco California USA, August 2016. ACM.
- [44] Rabiú Muazu Musa, Anwar P. P. Abdul Majeed, Muhammad Zuhaili Suhaimi, Mohd Azraai Mohd Razman, Mohamad Razali Abdullah, and Noor Azuan Abu Osman. Nature of Volleyball Sport, Performance Analysis in Volleyball, and the Recent Advances of Machine Learning Application in Sports. In Rabiú Muazu Musa, Anwar P. P. Abdul Majeed, Muhammad Zuhaili Suhaimi, Mohd Azraai Mohd Razman, Mohamad Razali Abdullah, and Noor Azuan Abu Osman, editors, *Machine Learning in Elite Volleyball: Integrating Performance Analysis, Competition and Training Strategies*, SpringerBriefs in Applied Sciences and Technology, pages 1–11. Springer, Singapore, 2021.
- [45] Gilbert W. Fellingham. Evaluating the performance of elite level volleyball players. *Journal of Quantitative Analysis in Sports*, 18(1):15–34, March 2022. Publisher: De Gruyter.
- [46] Zachary J Hass. Division of Credit Modeling for Team Sports with an Emphasis on NCAA Volleyball. *Theses and Dissertations Available from ProQuest*, pages 1–161, January 2017.
- [47] Data Project Sport Software. Data Volley Media. <https://www.dataproject.com/Products/EN/en/Volleyball/DataVolley4>, 2007. (accessed 2024-08-24).
- [48] VolleyStation. VolleyStation. <https://volleystation.com>. (accessed 2024-08-03).

- [49] Peter L’Oiseau. *Process-Oriented Player Evaluation Metrics for USports Basketball*. Master of Science, Carleton University, Ottawa, Ontario, 2021.
- [50] Perfbook. Perfbook. <https://perfbook.fr/#perfbook>.
- [51] Ben Raymond, Adrien Ickowicz, and Tyler Widdison. Datavolley: Reading and Analyzing DataVolley Scout Files. <https://datavolley.openvolley.org>, 2024.
- [52] Scott Powers, Luke Stancil, and Naomi Consiglio. Estimating individual contributions to team success in women’s college volleyball, February 2024. arXiv:2402.01083 [stat].
- [53] Rob Antle. Regularized Adjusted Plus/Minus (RAPM). <https://basketballstat.home.blog/2019/08/14/regularized-adjusted-plus-minus-rapm/>, August 2019. (accessed 2024-08-04).
- [54] A. C. Thomas, Samuel L. Ventura, Shane Jensen, and Stephen Ma. Competing Process Hazard Function Models for Player Ratings in Ice Hockey, February 2013. arXiv:1208.0799 [stat].
- [55] Lars Magnus Hvattum. A comprehensive review of plus-minus ratings for evaluating individual players in team sports. *International Journal of Computer Science in Sport*, 18(1):1–23, July 2019.
- [56] Rui Marcelino, Isabel Mesquita, José Manuel Palao Andrés, and Jaime Sampaio. Home Advantage in High-Level Volleyball Varies According to Set Number. *Journal of Sports Science & Medicine*, 8(3):352–356, September 2009.
- [57] Fabio Angioluci Diniz Campos, Leandra Cristina Benetti Campos, Luiz Cláudio Reeberg Stanganelli, and Miguel Angel Gómez. The home advantage effect in Volleyball: The coaches’perspective. *IOSR Journal of Sports and Physical Education*, 2023.
- [58] Dan Rosenbaum. Picking the Difference Makers for the All-NBA Teams. <http://www.82games.com/comm30.htm>, April 2004. (accessed 2024-07-26).
- [59] Joseph Sill. Improved NBA Adjusted +/- Using Regularization and Out-of-Sample Testing. *MIT Sloan Sports Analytics Conference 2010*, 2010.
- [60] Söhnke M. Bartram, Kewei Hou, and Sehoon Kim. Real effects of climate policy: Financial constraints and spillovers. *Journal of Financial Economics*, 143(2):668–696, February 2022.

- [61] Taylor Snarr. Dunks & Threes: Pro Basketball Analysis. <https://dunksandthrees.com/>. (accessed 2024-08-04).
- [62] Zachary Hass and Bruce A. Craig. Exploring the potential of the plus/minus in NCAA women’s volleyball via the recovery of court presence information. *Journal of Sports Analytics*, 4(4):285–295, November 2018.
- [63] Robert Tibshirani. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996. Publisher: [Royal Statistical Society, Oxford University Press].
- [64] Carlos M. Carvalho, Nicholas G. Polson, and James G. Scott. Handling Sparsity via the Horseshoe. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, pages 73–80. PMLR, April 2009. ISSN: 1938-7228.
- [65] Tianyu Guan, Jiguo Cao, and Tim B. Swartz. Parking the bus. *Journal of Quantitative Analysis in Sports*, 19(4):263–272, December 2023.
- [66] Carmen Fernández-Echeverría, Jara González-Silva, Inma T. Castro, and M. Perla Moreno. The Timeout in Sports: A Study of Its Effect on Volleyball. *Frontiers in Psychology*, 10, October 2019. Publisher: Frontiers.
- [67] Paul-Christian Bürkner. brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80:1–28, August 2017.
- [68] Juho Piironen and Aki Vehtari. Sparsity information and regularization in the horse-shoe and other shrinkage priors. *Electronic Journal of Statistics*, 11(2), January 2017. arXiv:1707.01694 [stat].
- [69] Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. *Bayesian Data Analysis Third edition (with errors fixed as of 15 February 2021)*. 2021.
- [70] Jonah Gabry, Daniel Simpson, Aki Vehtari, Michael Betancourt, and Andrew Gelman. Visualization in Bayesian Workflow. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 182(2):389–402, February 2019.
- [71] Zachary Hass and Bruce A. Craig. Exploring the potential of the plus/minus in NCAA women’s volleyball via the recovery of court presence information. *Journal of Sports Analytics*, 4(4):285–295, November 2018.

- [72] Carlos M. Carvalho, Nicholas G. Polson, and James G. Scott. Handling Sparsity via the Horseshoe. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*, pages 73–80. PMLR, April 2009.
- [73] Gilbert W. Fellingham. Evaluating the performance of elite level volleyball players. *Journal of Quantitative Analysis in Sports*, 18(1):15–34, March 2022.
- [74] Daniel Myers. About Box Plus/Minus (BPM) | Basketball-Reference.com. <https://www.basketball-reference.com/about/bpm2.html>, February 2020. (accessed 2024-08-03).
- [75] José Roberto Cantú-González, Filiberto Hueyotl-Zahuantitla, Jesús Abraham Castorena-Peña, and Mario A. Aguirre-López. The Attack-Block-Court Defense Algorithm: A New Volleyball Index Supported by Data Science. *Symmetry*, 14(8):1499, July 2022.
- [76] Chad Gordon and Andrew Tao. Computer Vision Introduction. <https://www.volleydork.com/post/computer-vision-introduction>, May 2023. (accessed 2024-08-04).
- [77] Adrien Ickowicz and Ben Raymond. Multi-camera volleyball tracking. <https://untan.gl/multicamera-volleyball-tracking.html>, March 2022. (accessed 2024-08-04).
- [78] Major League Baseball. Jump. <https://www.mlb.com/glossary/statcast/jump>. (accessed 2024-08-04).
- [79] Corley Bagley and Brett Ware. Bump, Set, Spike: Using Analytics to Rate Volleyball Teams and Players. 2017.
- [80] Michelle A. Miskin, Gilbert W. Fellingham, and Lindsay W. Florence. Skill Importance in Women’s Volleyball. *Journal of Quantitative Analysis in Sports*, 6(2), April 2010.
- [81] Scott Powers, Luke Stancil, and Naomi Consiglio. Estimating individual contributions to team success in women’s college volleyball, February 2024. Comment: 23 pages, 10 figures, 2023 NESSIS.

APPENDICES

Appendix A

Full Table of RAPM Player Estimates

Rank	RAPM	Athlete Name	Team	50% CI
Outside Hitter				
1	0.18	Janelle Albert	Western Mustangs	[0.13, 0.24]
2	0.14	Sullie Sundara	McMaster Marauders	[0.11, 0.19]
3	0.11	Madison Chimienti	Brock Badgers	[0.04, 0.17]
4	0.10	Olivia Kim	York Lions	[0.03, 0.15]
5	0.10	Julia Murmann	Toronto Varsity Blues	[0.05, 0.14]
6	0.08	Julie Gordon	TMU Bold	[0.01, 0.13]
7	0.06	Scarlett Gingera	TMU Bold	[0.02, 0.10]
8	0.06	Olivia Julien	McMaster Marauders	[0.00, 0.10]
9	0.04	Mary Stewart	Queen's Gaels	[0.00, 0.07]
10	0.04	Sarah Glynn	Waterloo Warriors	[0.00, 0.07]
11	0.04	Aimee Lemire	Queen's Gaels	[0.00, 0.07]
12	0.04	Lauren Veltman	TMU Bold	[0.00, 0.07]
13	0.04	Emma Mckinnon	McMaster Marauders	[0.00, 0.06]
14	0.04	Jaeya Brach	Western Mustangs	[0.00, 0.06]
15	0.03	Laurin Ainsworth	Brock Badgers	[0.00, 0.06]
16	0.03	Oceane Raymond-Leduc	Nipissing Lakers	[0.00, 0.05]
17	0.03	Mackenzie Valle	Waterloo Warriors	[0.00, 0.04]
18	0.03	Devin Gerth	Western Mustangs	[0.00, 0.05]

19	0.03	Mya Newton	Western Mustangs	[0.00, 0.05]
20	0.03	Noa Pijl	Guelph Gryphons	[0.00, 0.04]
21	0.03	Genna Saunders	Trent Excalibur	[−0.01, 0.04]
22	0.02	Julia Hapke	Trent Excalibur	[0.00, 0.04]
23	0.02	Julie Moore	TMU Bold	[0.00, 0.03]
24	0.02	Hope Kotun	Toronto Varsity Blues	[0.00, 0.03]
25	0.01	Avery Kelly	Waterloo Warriors	[0.00, 0.02]
26	0.01	Yaila Sealy	Windsor Lancers	[0.00, 0.02]
27	0.01	Arielle Palermo	Queen's Gaels	[−0.01, 0.02]
28	0.01	Sophie Fernback	Guelph Gryphons	[−0.01, 0.02]
29	0.01	Madison Shanks	York Lions	[−0.01, 0.02]
30	0.01	Emily Foest	Brock Badgers	[−0.01, 0.01]
31	0.00	Jessica Reynolds	York Lions	[−0.01, 0.01]
32	0.00	Lauryn Langley	Lakehead Thunderwolves	[−0.01, 0.01]
33	0.00	Rachael Meilikhan	Waterloo Warriors	[−0.01, 0.01]
34	0.00	Leah Stewart	Guelph Gryphons	[−0.01, 0.01]
35	0.00	Emma Titley	RMC Paladins	[−0.01, 0.01]
36	0.00	Tori Dauvin	Lakehead Thunderwolves	[−0.01, 0.01]
37	0.00	Briar Crerar	Guelph Gryphons	[−0.01, 0.01]
38	0.00	Olivia Zhu	Toronto Varsity Blues	[−0.01, 0.01]
39	0.00	Lael Harrison	Western Mustangs	[−0.01, 0.01]
40	0.00	Paige Entwistle	McMaster Marauders	[−0.01, 0.01]
41	0.00	Emily Tsiandoulas	Western Mustangs	[−0.01, 0.01]
42	0.00	Katelyn Grasman	TMU Bold	[−0.01, 0.01]
43	0.00	Maia Dan	Guelph Gryphons	[−0.01, 0.01]
44	0.00	Aliah Admans	Windsor Lancers	[−0.01, 0.01]
45	−0.01	Paige Owen	Nipissing Lakers	[−0.01, 0.01]
46	−0.01	Jennifer Zhu	Waterloo Warriors	[−0.02, 0.00]
47	−0.01	Victoria Herrington	York Lions	[−0.02, 0.01]
48	−0.01	Sophia Kalbfleisch	Guelph Gryphons	[−0.02, 0.00]
49	−0.02	Alysha Goundrey	Lakehead Thunderwolves	[−0.03, 0.00]
50	−0.02	Zoe Kuck	TMU Bold	[−0.03, 0.00]
51	−0.02	Shayla Kessler	Nipissing Lakers	[−0.03, 0.00]
52	−0.02	Aryn Markett	Windsor Lancers	[−0.04, 0.00]
53	−0.02	Brooklyn Parliament	York Lions	[−0.04, 0.00]
54	−0.02	Nicole Muzyka	Trent Excalibur	[−0.04, 0.00]
55	−0.02	Megan Nahachewsky	Lakehead Thunderwolves	[−0.04, 0.00]
56	−0.03	Kianna Drewniak	York Lions	[−0.05, 0.00]

57	-0.04	Shaelyn Francis	Nipissing Lakers	$[-0.06, 0.00]$
58	-0.04	Sydney Maclellan	Trent Excalibur	$[-0.07, 0.00]$
59	-0.05	Justine Colambot	Trent Excalibur	$[-0.09, 0.00]$
60	-0.06	Maya Marshall	RMC Paladins	$[-0.10, 0.00]$
61	-0.07	Cayleigh Fossum	Lakehead Thunderwolves	$[-0.11, -0.01]$
62	-0.10	Ashley Chartier	Trent Excalibur	$[-0.15, -0.03]$
63	-0.15	Elisa Marketin	RMC Paladins	$[-0.22, -0.07]$
64	-0.16	Adyson Wilson-Hands	TMU Bold	$[-0.23, -0.08]$
65	-0.20	Ashley Raymond	RMC Paladins	$[-0.26, -0.14]$
Setter				
1	0.18	Sara Rohr	Brock Badgers	$[0.11, 0.24]$
2	0.15	Claire Gagne	Waterloo Warriors	$[0.08, 0.21]$
3	0.11	Hayley Goodwin	Toronto Varsity Blues	$[0.01, 0.20]$
4	0.08	Jenna Peters	Toronto Varsity Blues	$[0.01, 0.14]$
5	0.07	Shea Baker	Queen's Gaels	$[0.01, 0.11]$
6	0.04	Christina Stratford	McMaster Marauders	$[0.00, 0.07]$
7	0.03	Lauryn Colpitts	McMaster Marauders/ Queen's Gaels	$[0.00, 0.04]$
8	0.02	Alex Curran	Guelph Gryphons	$[0.00, 0.03]$
9	0.02	Chiara Da Silva	York Lions	$[0.00, 0.03]$
10	0.02	Hannah Joo	Western Mustangs	$[0.00, 0.03]$
11	0.01	Lauren Pastor	Western Mustangs	$[-0.01, 0.02]$
12	0.01	Madison Seo	Waterloo Warriors	$[0.00, 0.02]$
13	0.01	Mackenna Knox	Brock Badgers	$[-0.01, 0.02]$
14	0.01	Sarah Zonneveld	TMU Bold	$[0.00, 0.02]$
15	0.01	Alexa Shields	Guelph Gryphons	$[-0.01, 0.02]$
16	0.00	Brooke Davis	Windsor Lancers	$[-0.01, 0.01]$
17	0.00	Katrine Eistrat	Waterloo Warriors	$[-0.01, 0.02]$
18	0.00	Rebecca Illingworth	Nipissing Lakers	$[-0.01, 0.01]$
19	0.00	Natalie Lefaive	York Lions	$[-0.01, 0.01]$
20	0.00	Thea Ishida	Western Mustangs	$[-0.01, 0.01]$
21	0.00	Danika Scharff	Windsor Lancers	$[-0.01, 0.01]$
22	0.00	Payton Entwistle	McMaster Marauders	$[-0.01, 0.01]$
23	-0.01	Kennedy Millin	Queen's Gaels	$[-0.02, 0.01]$
24	-0.01	Alicia Lam	TMU Bold	$[-0.02, 0.01]$
25	-0.01	Anna Arellano	Toronto Varsity Blues	$[-0.02, 0.01]$
26	-0.01	Jasmine Safar	Toronto Varsity Blues	$[-0.02, 0.00]$

27	-0.01	Robin Melnick	Toronto Varsity Blues	$[-0.02, 0.00]$
28	-0.01	Nicole Woelk	Windsor Lancers	$[-0.02, 0.00]$
29	-0.01	Emilie L'abbe	RMC Paladins	$[-0.02, 0.01]$
30	-0.02	Sarah Hewer	Waterloo Warriors	$[-0.03, 0.00]$
31	-0.02	Cassidy Chang	RMC Paladins	$[-0.03, 0.00]$
32	-0.03	Megan Mcdonald	York Lions	$[-0.05, 0.00]$
33	-0.03	Logan Garvey	Nipissing Lakers	$[-0.05, 0.00]$
34	-0.03	Chayse Victoria	McMaster Marauders	$[-0.05, 0.00]$
35	-0.06	Ella Carter	Guelph Gryphons	$[-0.09, -0.01]$
36	-0.08	Alexis Cancade	Lakehead Thunderwolves	$[-0.14, 0.00]$
37	-0.10	Payton Leflar	Lakehead Thunderwolves	$[-0.18, -0.01]$
38	-0.12	Jayda Phillipson	RMC Paladins	$[-0.18, -0.06]$
39	-0.25	Heidi Hayes	Trent Excalibur	$[-0.36, -0.12]$
40	-0.28	Isabelle Pellerine	Trent Excalibur	$[-0.39, -0.20]$
Libero				
1	0.12	Vivian Zanutel	Queen's Gaels	$[0.06, 0.17]$
2	0.09	Aleiah Torres	Brock Badgers	$[0.03, 0.15]$
3	0.09	Jenna Gravina	Waterloo Warriors	$[0.01, 0.16]$
4	0.09	Sophia Currier	Toronto Varsity Blues	$[0.00, 0.15]$
5	0.07	Holly Biasi	Queen's Gaels	$[0.01, 0.12]$
6	0.06	Sam Loshusan	Windsor Lancers	$[0.01, 0.11]$
7	0.05	Tyneille Swaby	York Lions	$[0.00, 0.08]$
8	0.05	Mary Ann Riofloridio	TMU Bold	$[0.01, 0.08]$
9	0.04	Kassandra Pavlic	Nipissing Lakers	$[0.00, 0.07]$
10	0.04	Maddie Bowles	Western Mustangs	$[0.00, 0.06]$
11	0.03	Ashley Vang	Guelph Gryphons	$[0.00, 0.04]$
12	0.03	Nora Kotun	Waterloo Warriors	$[0.00, 0.05]$
13	0.02	Ashlynn Freiburger	Waterloo Warriors	$[0.00, 0.03]$
14	0.01	Mackenzie Proulx	Guelph Gryphons	$[0.00, 0.02]$
15	0.01	Jada Weir	Nipissing Lakers	$[-0.01, 0.02]$
16	0.01	Payton King	Guelph Gryphons/ Toronto Varsity Blues	$[-0.01, 0.02]$
17	0.01	Medeleine Pelosi	Western Mustangs	$[-0.01, 0.02]$
18	0.00	Natalia Souza	Toronto Varsity Blues	$[-0.01, 0.01]$
19	0.00	Alex Sinik	Waterloo Warriors	$[-0.01, 0.01]$
20	0.00	Kylie Maxwell	Western Mustangs	$[-0.01, 0.01]$
21	0.00	Brooke Driscoll	Trent Excalibur	$[-0.01, 0.02]$

22	0.00	Delaney Watson	Toronto Varsity Blues	$[-0.01, 0.01]$
23	0.00	Lauren Boddington	York Lions	$[-0.01, 0.01]$
24	0.00	Hayley Brookes	McMaster Marauders	$[-0.01, 0.01]$
25	0.00	Joey Hewick	Windsor Lancers	$[-0.01, 0.01]$
26	0.00	Anna Scarfe	Windsor Lancers	$[-0.01, 0.01]$
27	0.00	Maddie Cavanaugh	Trent Excalibur	$[-0.01, 0.01]$
28	-0.01	Leah Van Zeumeren	Waterloo Warriors	$[-0.02, 0.00]$
29	-0.01	Lilie Demelo	RMC Paladins	$[-0.02, 0.01]$
30	-0.01	Jade Butterworth	Trent Excalibur	$[-0.02, 0.01]$
31	-0.02	Kristine Boileau	Queen's Gaels	$[-0.03, 0.00]$
32	-0.05	Sydney Simonson	Guelph Gryphons	$[-0.08, 0.00]$
33	-0.08	Alexie Chevrier	Nipissing Lakers	$[-0.13, -0.02]$
34	-0.08	Rebekah Hara	Lakehead Thunderwolves	$[-0.14, -0.02]$
35	-0.09	Emilie Boileau	RMC Paladins	$[-0.14, -0.04]$
36	-0.15	Annie Whelen	Trent Excalibur	$[-0.22, -0.09]$

Middle Blocker

1	0.08	Christina Piccinin	York Lions	$[0.02, 0.13]$
2	0.04	Ellie Hatishita	McMaster Marauders	$[0.00, 0.07]$
3	0.04	Jordyn Brearley	Windsor Lancers	$[0.00, 0.06]$
4	0.03	Ana Strbac	McMaster Marauders	$[0.00, 0.05]$
5	0.03	Alexandra Iggulden	Western Mustangs	$[0.00, 0.05]$
6	0.02	Jenna Woock	McMaster Marauders	$[0.00, 0.04]$
7	0.02	Jenna Woock	Toronto Varsity Blues	$[0.00, 0.03]$
8	0.02	Ramanda Bell	Nipissing Lakers	$[0.00, 0.03]$
9	0.02	Jaden Germann	Nipissing Lakers	$[0.00, 0.03]$
10	0.02	Elizabeth Bilawey	Toronto Varsity Blues	$[0.00, 0.03]$
11	0.01	Ava Ebert	Western Mustangs	$[0.00, 0.02]$
12	0.01	Taya Knier	Trent Excalibur	$[-0.01, 0.02]$
13	0.01	Grace Pyatt	Brock Badgers	$[0.00, 0.02]$
14	0.01	Grace Coulter	Western Mustangs	$[-0.01, 0.02]$
15	0.01	Mia Miskovic	Guelph Gryphons	$[-0.01, 0.02]$
15	0.01	Mia Miskovic	Western Mustangs	$[-0.01, 0.02]$
16	0.01	Elina Zvigelsky	York Lions	$[-0.01, 0.02]$
17	0.01	Britney Veltman	TMU Bold	$[-0.01, 0.02]$
18	0.01	Bethany Smith	TMU Bold	$[-0.01, 0.02]$
19	0.01	Anna Maria Gadomski	Toronto Varsity Blues	$[-0.01, 0.02]$
20	0.01	Samantha Ingleton	Guelph Gryphons	$[-0.01, 0.01]$

21	0.00	Christina Jovetic	Brock Badgers	$[-0.01, 0.01]$
22	0.00	Christine Hachokake	McMaster Marauders	$[-0.01, 0.01]$
23	0.00	Ella Mackenzie	Guelph Gryphons	$[-0.01, 0.01]$
24	0.00	Amanda Siksna	Queen's Gaels	$[-0.01, 0.01]$
25	0.00	Angelina Markotic	Brock Badgers	$[-0.01, 0.01]$
26	0.00	Sydney Degraauw	Waterloo Warriors	$[-0.01, 0.01]$
27	0.00	Avery Isaacs	Waterloo Warriors	$[-0.01, 0.01]$
28	0.00	Grace Crooks	Waterloo Warriors	$[-0.01, 0.01]$
29	0.00	Lauren Mellon	Queen's Gaels	$[-0.01, 0.01]$
30	0.00	Jamie Spencer	Trent Excalibur	$[-0.01, 0.01]$
31	0.00	Josephine Calverley	RMC Paladins	$[-0.01, 0.01]$
32	-0.01	Lisa Wedel	Guelph Gryphons/ Toronto Varsity Blues	$[-0.01, 0.01]$
33	-0.01	Megan Winter	Nipissing Lakers	$[-0.02, 0.01]$
34	-0.01	Brooke Elliot	Trent Excalibur	$[-0.02, 0.01]$
35	-0.01	Megan Gahima	Windsor Lancers	$[-0.02, 0.00]$
36	-0.01	Eve Smith	RMC Paladins	$[-0.02, 0.01]$
37	-0.01	Lauren Hebert	Waterloo Warriors	$[-0.02, 0.01]$
38	-0.01	Amarai Endes	York Lions	$[-0.02, 0.00]$
39	-0.01	Sophie Hotte	RMC Paladins	$[-0.02, 0.01]$
40	-0.01	Christine Abu	Queen's Gaels	$[-0.02, 0.01]$
41	-0.01	Madison Murphy	Nipissing Lakers	$[-0.02, 0.00]$
42	-0.01	Kristine Ledger	Nipissing Lakers	$[-0.02, 0.00]$
43	-0.01	Elizabeth Turner	Trent Excalibur	$[-0.02, 0.00]$
44	-0.02	Mia Chartrand	Nipissing Lakers	$[-0.03, 0.00]$
45	-0.02	Meera Sayal	Brock Badgers	$[-0.03, 0.00]$
46	-0.02	Cassidy Collins	Guelph Gryphons	$[-0.03, 0.00]$
47	-0.02	Ema Baruchelli	Brock Badgers	$[-0.03, 0.00]$
48	-0.02	Claire Stewart	Western Mustangs	$[-0.03, 0.00]$
49	-0.02	Jaime Couvrette	Lakehead Thunderwolves	$[-0.03, 0.00]$
50	-0.02	Ashley Ditchfield	TMU Bold	$[-0.04, 0.00]$
51	-0.02	Milena Ljubicic	Lakehead Thunderwolves	$[-0.04, 0.00]$
52	-0.03	Hannah Gingera	Lakehead Thunderwolves	$[-0.04, 0.00]$
53	-0.03	Serena Otshudi	Waterloo Warriors	$[-0.05, 0.00]$
54	-0.04	Katie Murray	Windsor Lancers	$[-0.07, 0.00]$
55	-0.05	Emily Heil	Lakehead Thunderwolves	$[-0.08, 0.00]$
56	-0.06	Thea Ziesmann	Toronto Varsity Blues	$[-0.10, 0.00]$
57	-0.07	Telessa Salmikivi	RMC Paladins	$[-0.11, -0.01]$

58	-0.09	Chloe Hallett	Trent Excalibur	$[-0.15, -0.01]$
59	-0.10	Ella Rohani	York Lions	$[-0.16, -0.03]$
60	-0.17	Sara Grosse	RMC Paladins	$[-0.24, -0.10]$
Right Side				
1	0.07	Julia Typa	York Lions	$[0.01, 0.11]$
2	0.05	Madi Lohmann	Trent Excalibur	$[0.00, 0.08]$
3	0.04	Maddy Lutes	McMaster Marauders	$[0.00, 0.07]$
4	0.03	Karen Fan	Waterloo Warriors	$[0.00, 0.06]$
5	0.02	Jessica Chung	Guelph Gryphons	$[0.00, 0.04]$
6	0.02	Asya Dogruer	Toronto Varsity Blues	$[0.00, 0.03]$
7	0.01	Hannah Duchesneau	Queen's Gaels	$[0.00, 0.02]$
8	0.01	Jessie Nairn	McMaster Marauders	$[0.00, 0.02]$
9	0.01	Mikayla Sherriffs	TMU Bold	$[0.00, 0.02]$
10	0.01	Madison Tuck	Brock Badgers	$[-0.01, 0.02]$
11	0.01	Kate Vancamp	Western Mustangs	$[-0.01, 0.02]$
12	0.01	Sadie Dick	Brock Badgers	$[-0.01, 0.02]$
13	0.01	Elizabeth Junger	Nipissing Lakers	$[-0.01, 0.02]$
14	0.01	Jade Longlad	Western Mustangs	$[-0.01, 0.01]$
15	0.01	Alisa Trubyuk	York Lions	$[-0.01, 0.01]$
16	0.01	Julia Wiercigroch	Toronto Varsity Blues	$[-0.01, 0.01]$
17	0.00	Raiya Millin	Windsor Lancers	$[-0.01, 0.01]$
18	0.00	Rachael Meilikhan	Queen's Gaels	$[-0.01, 0.01]$
19	0.00	Ana-Maria Horvat	Nipissing Lakers	$[-0.01, 0.01]$
20	0.00	Abbey Neufeld	Waterloo Warriors	$[-0.01, 0.01]$
21	0.00	Sarah Mcnames	Guelph Gryphons	$[-0.01, 0.01]$
22	0.00	Hailey Holland	Guelph Gryphons	$[-0.01, 0.01]$
23	0.00	Teah Cook	Queen's Gaels	$[-0.01, 0.01]$
24	0.00	Natalie Fournier	Trent Excalibur	$[-0.01, 0.01]$
25	-0.01	Amy Connelly	Nipissing Lakers	$[-0.01, 0.01]$
26	-0.01	Rachel Gray	Western Mustangs	$[-0.01, 0.01]$
27	-0.01	Robin Henshaw	Western Mustangs	$[-0.02, 0.01]$
28	-0.01	Giulia Chimilovski	Toronto Varsity Blues	$[-0.02, 0.01]$
29	-0.01	Anjallia Lonardi	TMU Bold	$[-0.02, 0.00]$
30	-0.01	Emma Shepherd	Windsor Lancers	$[-0.02, 0.00]$
31	-0.01	Kylie Ferguson	TMU Bold	$[-0.02, 0.00]$
32	-0.01	Jaymie Stephenson	Guelph Gryphons	$[-0.02, 0.00]$
33	-0.01	Virginie Franks	Lakehead Thunderwolves	$[-0.02, 0.00]$

34	−0.01	Quincy Johnson	York Lions	[−0.02, 0.00]
35	−0.02	Jaide Cummings	Toronto Varsity Blues	[−0.03, 0.00]
36	−0.03	Kathryn Beresford	Trent Excalibur	[−0.04, 0.00]
37	−0.03	Abby Delamere	McMaster Marauders	[−0.06, 0.00]
38	−0.13	Mackenzie Lockett-Hiemstra	RMC Paladins	[−0.21, −0.03]
39	−0.17	Taryn Bender	RMC Paladins	[−0.24, −0.09]

Glossary

Attack The final contact in a team's possession, intended to score a point on the opposition's side [3](#)

Block An attempt by the defensive team to intercept the attack of an offensive team at the net and send it back to side of the offensive team [3](#)

Dig The first contact in a team's possession after the opposing team *attacks* the ball over the net [3](#)

Receive The first contact in a team's possession after the opposing team *serves* the ball over the net [3](#)

Set The second contact in a team's possession, intended to facilitate a successful attack [3](#)