Stor 538 Championship
Rugby

Liam Hawkins, Tyler Youngberg, William Qin, Branson Pigg, Gabby Ruehle

Introduction

Rugby is a collective name for a couple of different team sports, mainly rugby union, rugby league, and rugby sevens. The most immediate difference between these sports is the number of players on each team during play with rugby union having 15 players, rugby league having 13 players, and rugby sevens having 7 players per team, all of which are typically split into forwards and backs. In all versions of rugby, the ball can only move in a forward direction if it is kicked or held by a running player. The ball can only be passed across the field or behind the advancing player, if the ball is passed forwards this results in a stoppage of play and the game restarts with a scrum which involves players on both teams packing together to attempt to gain possession of the ball. The objective of the game is to score the most points before time ends. There are a couple of different ways to score involving tries, conversion kicks, penalty kicks, and drop goals. A try is scored when a player grounds the ball in a downward motion in the other team's in-goal area. Conversion kicks occur after a try is scored; the ball is placed perpendicular to where the try was scored and the ball is kicked through the end posts. After a team is awarded a penalty, they have the option of taking a penalty kick, which is scored if the ball passes through the posts and above the crossbar. A drop goal is also scored by kicking the ball through the posts and above the crossbar but this occurs during game play. The scoring systems for rugby union and rugby sevens are the same with 5 points for a try, 2 points for a conversion kick, and 3 points for either a penalty kick or drop goal. In rugby league the scoring system is 4 points for a try, 2 points for a conversion kick, 2 points for a penalty goal, and 1 point for a drop goal.

The difference between rugby union and rugby sevens is minimal with the biggest difference being the number of players and the length of the halves (40 minutes for rugby union, 7 minutes for rugby sevens). Also, rugby sevens is typically the version played in the olympics and other international competitions. Rugby union and rugby league are very different versions of the game and often described as the difference between chess and checkers. Rugby union is often seen as more complicated, with more players on the field in more specialized positions. The biggest differentiator between the two styles is the possession of the ball after tackles. In rugby union once a player is tackled they must release their hold of the ball, and the ball can then be won by either side and play resumes. In rugby league the team has 6 attempts to move the ball down the field and an attempt ends when the player with the ball is tackled. Once the 6th tackle occurs, the ball switches possession and then the other team has 6 attempts to move the ball down the field and score (*Comparison of rugby league and rugby union*).

Rugby is currently the 9th most popular sport in the world with around 390 to 410 million fans. The sport is mainly played in England, France, Australia, New Zealand, and South Africa, but there are many other countries which compete in international competitions (*Top 10 Most Popular Sports in The World*). Rugby is typically played during the winter months, but there are games year-round due to the different hemispheres. In recent years rugby has experienced large growth with 9.6 million participants in 2018 (*Record number of people playing rugby globally - report*). Rugby has also recently had large growth in participation rates, with the U.S. seeing an 82.4% rise in participation from 2011-2016 (Soleimani). In the US there were an estimated 1.6 million rugby players in 2018 (Gough, 2020). There are a number of rugby leagues around the world including the Top 14 in France which sees 337.2 million in yearly revenue, the National Rugby League in Australia and New Zealand with 316 million in yearly revenue, Premiership Rugby in England with 251.4 million, Super Rugby with many countries with 94 million in revenue, and Super League in England and France with 60 million in revenue (*List of professional sports leagues by revenue*). Currently investment in rugby is rapidly changing with

private equity group CVC investing more than €600 million over the past 2 years into different leagues and tournaments. Currently, CVC is in talks to explore rugby streaming deals to digital players such as Amazon or create an internet subscription service for fans (Wiggins, 2020).

We decided to focus on rugby for this project because it is a quickly growing sport that is gaining dominance on the world stage, but still has had very little attention in terms of analytics. We believe that rugby is likely to continue to grow in popularity since rugby only entered the Olympics in 2016 and is likely to see large increases in participation after the next Olympics. Also with the new investment deals mentioned above, rugby is also likely to become more accessible to more viewers, increasing interest in the sport. So we thought that rugby has the interest, investment, and growing participation which makes it a great sport to evaluate for data analysis.

Literature Review

Over the years as improvements in both technology and hardware advanced, analytics too, has taken off in the sport of rugby. Across both rugby league and union, teams dedicate significant time and resources to find effective ways to utilize the vast amount of data collected both in training and gameday situations to improve their own team. Analytics have also helped coaches and scouts pick out the prospects in the field of rugby that go on to have successful careers. A popular system in which actions are given points or deductions based on their worth on the pitch, go on to measure the overall player's performances per game. Moreover, these points are weighted more in particular situations such as late game plays, especially if the game were tied, to award players that are least affected by mental stress and pressure. The metrics collected include number of tries, successful passes, drop kicks, successful tackles, and penalties then are assigned to positions. After this process, it is more clear which players are positively contributing to their team based on preset statistics on what players in a specific position should be good at. Finally there is even off ball data (when players are supporting the teammate carrying the ball) that is collected to give coaches a solid picture of their team. (Watts)

All major rugby union clubs today have used data to monitor fitness, prevent injuries and track player's positions through many different devices, most commonly wearable GPS trackers. The data gained from these devices has given coaches more insights on their teams as well as decreased player injuries. "GPS technology company Catapult - which develops wearable devices sewn into the back of players' shirts - recently aimed to deepen the use of data in rugby by launching a unique set of algorithms engineered to quantify key technical and physical demands in the sport." (Arastey) By automatically measuring actions of rugby players, these devices can capture and note every movement taken by players who wear the tracker, compile this data for staff to analyze, and share back with the players for them to improve their game. The data insights of the physical demands of rugby imposed on players also gives coaches important information on the physical levels of their players and helps guide them when they have to make tough decisions such as who to keep in the game and how he/her should train and recover after by monitoring the players' thresholds. After understanding this data, trainers can better design workout sessions to keep the players fresh for the games. Ultimately, controlling the levels of fitness of rugby athletics prevents injuries caused from physical overexertion. Injuries can also be caused by one singular play so the fitness monitors are also used to measure the time a player recovers in a game after a hard hit. Trainers therefore can measure each player's fatigue level and status and then decided if they should return to the game.

Another focal point in rugby that has been strongly impacted by analytics is concussions. Concussions are a growing issue in the sport, leading to players eventually suffering from chronic brain conditions with symptoms similar to Alzheimer's. There has been extensive research and development that aims to better prevent and monitor head injuries across all divisions of rugby, a heavy contact sport. In the past, when these technologies were not available yet, trainers had to rely on a player's self evaluation for their risk assessment when deciding whether the player should return to play after a hard hit. Now there have been personalised gum shields invented that are worn by players to both protect their teeth but also collect data from hits and tackles that impact the head. From the data collected by this device, doctors can better prevent injuries and treat players as the gum shields paint a more accurate picture of the forces involved in each impact to the head. The devices can also transmit impact data to a back to pitch-side doctors in real time so they can instantly assess if a player needs to come out of the game and receive further assessment. "This has proven particularly important in training sessions, where 20% of head injuries take place, although most of them go unseen. Thanks to this technology, coaches are now able to assess the forces exerted by players during drills and adjust the practice accordingly to avoid undetected head injuries." (Arastey)

The best and most successful team in rugby goes without a doubt to New Zealand's All Blacks. Since the start of rugby, the All Blacks have amounted a win percentage of almost 75% and after a coaching change in 2004, the win percentage has since grown to 85%. Some attribute their success to genetics as the gene pool of the country include the power, strength and speed found in Māori and Polynesian blood. However, there are many other factors that contribute to the All Black's success. In New Zealand, rugby is embedded in the culture so kids start young. They play Rippa Rugby, the non-contact, simplified version of rugby. But as they grow up in a country dominated by rugby with educational systems dedicated to it, they develop fast and are eager to be the next prospect at the next level of rugby. Although there is a high drop of rate after kids reach highschool, with coaches using specialized systems and software in training to develop and discover the best prospects of the kids that make it past highschool, New Zealand's national team is able to find athletes, improve their game, and keep them in the game. (Coates)

As rugby became more and more popular in the world, it also made its way into the world of gambling. Almost every professional rugby league and international match is available for betting. There are many different betting options but a few have emerged to become the most popular. There are head to head bets where gamblers simply bet on a winner of a match; however, there is also a more complicated style where a bookmaker sets a line and gamblers bet if the favored team will win by more than the set line.There are also halftime bets where gamblers can pick who is up at half. Finally, there are total match bets where gamblers attempt to predict the total points scored to other categories such as total match tries. Other miscellaneous categories include betting on single events in the game such as first try scorer. (Rugby League)

Future Work

With rugby sevens quickly becoming a favorite on the international sports scene comes much interest in the way the sport is analyzed and what these analyses say about the game and the skill levels of the people playing it. From our preceding literature review, we can see that rugby analysis is fairly new in terms of sports analytics, with major studies starting to show up in the 1990s. Additionally, the standard metrics studied for rugby are pretty basic when compared to the number of performance metrics we take into consideration for American sports, like football or basketball. This leaves us with a pretty broad spectrum of possibilities to consider for

how the future of rugby analysis will work and how it will build off of current knowledge and research.

One breakthrough we've seen in numerous sports, including rugby, over the past decade is the introduction of Global Positioning System (GPS) technology into the realm of sports analytics. In rugby, it is currently common for players to wear these GPS devices within their athleticwear. This allows rugby analysts to get an idea of each player's total distance traveled in a game or how fast a player was running on average throughout a match (Henderson et al). Using these GPS systems, rugby balls could be improved on by manufacturing the balls with miniature GPS devices installed within. Using these GPS-enabled rugby balls in conjunction with a dedicated analyst and specialized analysis software, we could collect information on how much distance a player covered while carrying the ball as well as how much time they spent carrying the ball. Utilizing this data alongside metrics on a player's number of successful ball passes and number of turnovers while in possession of the ball can give league coaches an insight into which team members are strong ball-movers and can work the ball down the field to the opponent zone with ease as opposed to who is better at defending the ball-movers. Furthermore, analysts could look into devising a ratio metric of the distance a player carried the ball per turnover during their possession. This ratio metric would allow analysts to compare players in terms of offensive skill across a team on an even playing field, including those who may not be heavy offensive players with much carry distance. This allows the coaches to make decisions on who should play in the offensive-heavy positions and who should typically man the defensive positions of the game, improving a team's physical and technical performance.

Existing as a prominent metric in sports like basketball, hockey, and soccer, the "plus-minus" (+/-) statistic can be applied in the scope of rugby quite similarly. Plus-minus for a player, in its simplest form, gives the difference in the total number of goals scored by a team when the player is on court and the total number of goals scored by the opposition when the player is on court. This metric, when considered with other selections of statistics for a specific sport of interest, can give insight into how valuable and impactful a player is to their team and their game performance in the sports it is tracked in, leading us to believe it could be just as applicable to rugby analysis (Schultze et al).

In the scope of rugby sevens, plus-minus could be an effective metric to keep in mind when analyzing a league's performance as well as their players' individual performances. Plus-minus for rugby would be defined similar to basketball's +/- statistic: +/- for player X = number of points scored by home league while X in play – number of points scored by visiting league while X in play. Studying this metric in rugby can help coaches and other decision makers recognize which players in a league may be lacking in their scoring abilities on the field and who is seemingly racking up the points. This insight would allow the coaches and athletic staff to adjust training regimens as needed to raise any struggling player's average +/- rating and to maintain players with higher averages. Beyond training, this statistic, not unlike the metrics gained from the hypothetical ball GPS data, could be used to determine who fits best in which position during the scrum, with those players averaging higher plus-minus ratings being placed into positions that tend to be try-heavy and have numerous scoring opportunities. Additionally, comparing plus-minus ratings among groups of players on the field at the same time could help coaches determine who to try to group together in any given match to maximize net scoring potential.

American football is a very similar game to rugby and thus some of the stats can transfer over to rugby as well. One particular advanced stat referred to as "Field Zone" is a relatively

unknown stat, however it can prove useful not just to American football, but to rugby as well. Field Zone is basically where exactly on the field a play is successful and which direction of the field the offense is heading. Typically, there are fifteen zones, (not including end-zones), a zone is 20 yards wide, and there are five zones on the left side of the hash marks, five in the middle of the hash marks, and five on the right side of the hash marks. Field Zone can be further broken down to passing plays, rushing plays, whether plays are more successful in the middle of the field, left side of the field, or right side. For rugby, the field itself is 100 meters long with two in-goal areas on opposite sides of the field. There will also be fifteen zones where each zone is 20 meters big on the left side, middle, and right side of the field. Zone success rates should track running as well as passing. Although some leagues don't allow forwards passing but for the ones that do the following scrum that results from a forward pass will be tracked as well. Running can be in any direction, so for the running to be determined as a "success" there'll need to be a certain amount of meters gained on the field before a tackle as a result of the running. Scoring a try as a result from running would also indicate success for the part of the field where the running began. This method can effectively show where and which plays are more effective and can lead to coaches making better decisions as to which strategy should be implemented at a certain spot on the field.

In American football, Markov Models can be used to determine win probability at a given point in the game. For American football, a drive is an absorbing Markov chain meaning that certain results in the game are absorbing states such as turnovers and scoring plays. This process can be implemented to rugby, however some changes are necessary. Football is easier because after each play there is a result that occurs and play is temporarily halted, even if the clock is still running. For rugby, play and especially running can theoretically last the entire half without pause or change of possession. So for our "states" in the Markov chain, it will be however long a particular team has possession of the ball and whether it results in a turnover, a try, a drop kick, a penalty, etc. These states are also independent of each other, meaning that the previous possession has no bearing on the next possession. Adding all the possessions together gives you the Markov Chain. Also, unlike football, there aren't stoppage in plays unless a team scores, so there are more results that can happen in a particular state, making the probability harder to predict. With this though, knowing the probability means that the individual outcomes of a possession are uncertain, but the probability of an event within the possession isn't. This knowledge can help teams and particularly coaches make more informed decisions on different strategies for when they have possession of the ball. Also, when they are playing defense they can determine what the other teams strengths are and their probability of scoring on a particular possession.

Conclusion

Rugby is one of the most popular sports in the world complete with several different variations on how to play. Today, rugby makes hundreds of millions of dollars of revenue per year in some European countries and Australia. Even now, rugby continues to grow quickly in popularity, especially in the US.

As it grows in popularity, rugby has also grown in interest to statisticians. In fact, most statistical work in rugby has been done within the last 30 years. Statisticians have compared players by attributing point-values to different actions and comparing player values across positions based on positional needs. In addition, more technically advanced analytics has been done using GPS trackers and sensors to better understand and track player fitness and injuries.

The current prevalence of these technologies in the game of rugby open up possibilities for future analytics. For example, GPS devices could be used to track the position of the ball as well as the players to give insights on which player is carrying the ball and which players are attempting to tackle them. We also expect that modern metrics in other sports, such as plus-minus, could be useful when applied to the game of rugby. In particular, we believe that analytical techniques used in American Football could be applied to rugby in the future because of the many similarities between the two sports.

All things considered, this analysis of rugby has given us a greater appreciation of the many interesting and unique characteristics of a sport that we previously knew as "like football." After taking the time to fully comprehend the rules of the game, understand the immense physical exertions that it takes to play it, and consider the many interesting analytical techniques that can be applied to this game, it is clear why rugby has drawn in so many dedicated fans around the world. Now, learning about its worldwide popularity and the growing interest in rugby year after year comes as little surprise to us as our analysis has brought rugby out from the shadow of American Football and put it in a whole new light.

Bibliography

Arastey, Guillermo Martinez. "Impact of Data Analysis And Technology in Rugby Union." Sport Performance Analysis, Sport Performance Analysis, 17 Jan. 2020, www.sportperformanceanalysis.com/article/data-analysis-and-technology-in-rugby-union

Coates, Joe. "Why New Zealand Dominates Global Rugby." Culture Trip, The Culture Trip, 21 Feb. 2018, theculturetrip.com/pacific/new-zealand/articles/why-new-zealand-dominates-global-rugby/.

Comparison of rugby league and rugby union. (2020, November 12). Retrieved November 18, 2020, from https://en.wikipedia.org/wiki/Comparison_of_rugby_league_and_rugby_union

Gough, C. (2020, July 28). Rugby: Number of participants U.S. 2018. Retrieved November 18, 2020, from https://www.statista.com/statistics/191684/participants-in-rugby-in-the-us-since-2006/

Henderson, Mitchell J, et al. "Rugby Sevens Match Demands and Measurement of Performance: A Review." *Supplement to International Journal of Fundamental and Applied Kinesiology*, vol. 50, no. 1, Mar. 2018, pp. 49–59. https://hrcak.srce.hr/file/284308

List of professional sports leagues by revenue. (2020, November 07). Retrieved November 18, 2020, from https://en.wikipedia.org/wiki/List_of_professional_sports_leagues_by_revenue

Record number of people playing rugby globally - report. (2019, July 10). Retrieved November 18, 2020, from https://www.rugbypass.com/news/record-number-of-people-playing-rugby-globally-report/

"Rugby League In 2018: History: Bet Types: Betting Tips." World Gambling List, 12 July 2020, worldgamblinglist.com/rugby-league.

Sharp, Warren. "NFL Offensive Success by Field Zone (OFF)." *Sharp Football Stats*, 2020, https://www.sharpfootballstats.com/nfl-offensive-success-by-field-zone--off-.html.

Schultze, Steven R., and Christian-Mathias Wellbrock. "A Weighted plus/Minus Metric for Individual Soccer Player Performance." *Journal of Sports Analytics*, vol. 4, no. 2, 2018, pp. 121–131., doi:10.3233/jsa-170225.

Soleimani, S. (n.d.). What is the Fastest Growing Sport in America? Retrieved November 18, 2020, from https://blog.sisuguard.com/what-is-the-fastest-growing-sport-in-america

Top 10 Most Popular Sports in The World. (n.d.). Retrieved November 18, 2020, from https://sportsshow.net/top-10-most-popular-sports-in-the-world/

Watts, Dan. "The Practical Applications Of Rugby Analytics: Articles: Sports." Articles | Sports | Innovation Enterprise, 10 Mar. 2017, channels.theinnovationenterprise.com/articles/the-practical-applications-of-rugby-analytics.

Wiggins, K. (2020, February 23). Buyout group CVC steps up plans to reshape global rugby. Retrieved November 18, 2020, from https://www.ft.com/content/e57f79b2-54b0-11ea-8841-482eed0038b1