Proceedings of the Annual General Donald R. Keith Memorial Conference
West Point, New York, USA
April 29, 2021
A Regional Conference of the Society for Industrial and Systems Engineering

Fenwick and Corsi in Army Hockey, Incorporating "Shot-Zones" and Rebounds into Weighted Shots

Mason Krueger and James Enos

Department of Systems Engineering
United States Military Academy
West Point, NY

Corresponding author's Email: mason.krueger@westpoint.edu

Author Note: Cadet Mason Krueger is a member of the Army hockey team as well as a senior student in the Department of Systems Engineering at the United States Military Academy. This paper is written as a component for the requirements to graduate with honors in Systems Engineering.

Abstract: Competitive mindsets and incentives for winning within our society generate an environment that forces athletes and organizations to continuously seek new opportunities to gain a competitive advantage. Applying analytics in sports to improve performance is a recent trend of this. The Moneyball strategy in baseball is a great example that displays successful implementation of sports analytics. For this, a player's value is gauged by using statistical data and metrics to predict which players will perform the best while also being minimally expensive to acquire. Due to its success, teams across sports today are attempting to use statistical data and metrics to analyze all facets of their game in hopes of acquiring this edge. However, this is challenging for sports containing less discrete data such as hockey because much of the statistical information is non-value added, providing no benefit towards real world application. Regardless, the growth of analytics in hockey shows that these strategies are finding their way into the hockey environment.

Keywords: Hockey, Analytics, Statistical Data, Non-Value Added

1. Introduction

The sport of hockey is quickly developing and is considered one of the big four professional sports in the United States. Despite being on the lower end of popularity in comparison to the NFL, MLB, and NBA, the sport has still rendered a large following. As a result, the NHL is currently utilizing many of the strategies originally deployed by these larger sports. However, there are numerous variables that impact the validity of analytics in a high-paced sport like hockey making the process different than a sport like baseball. Baseball is primarily a two-player interaction between the pitcher and the batter, whereas hockey is more complex with a greater amount of interactions between multiple players (Elitzur, 2018). The discrete nature of baseball allows organizations to utilize analytical concepts to help predict factors of the game such as the direction of the field that certain players are more likely to hit the ball. As a result, shifting the defense towards one side of the field or the other depending on which batter is now common in the MLB. Teams employed a shift in 34% of plate appearances that ended in a ball in play during the 2018 season (Sawchik, 2018). Ultimately, this strategy resulted in an overall improved game management in baseball. In 2017, Mike Petriello reported that against a shift the MLB had a 0.276 batting average on balls in play and a 0.306 batting average on balls in play without a shift (Petriello, 2017). These successes noticed in baseball through the study of analytics suggest that the NHL should attempt a similar strategy. However, it is the application of analytics in hockey that make the numbers more complicated. In hockey, players are constantly moving around, line-changes happen even when the game is not at a stoppage, and hockey players can switch positions regularly. These irregularities make predicting future occurrences much more complicated when using analytics.

Additionally, there is an increased demand in sports today for quantifying metrics that correlate to team success as well as an individual player's contributions that result in that team's success (Found, 2016). Hockey is no exception as noticed through the implementation of advanced hockey statistics, the spawning of statistical analysis websites for hockey, and most recently the development of analytics departments by NHL organizations (Sheps, 2020). Despite this, analytics has received an abundance of push-back from traditional thinkers in mainstream hockey. Part of the friction from mainstream hockey is that hockey analytics comes with the underlying issue that the data is noisy; quantifying player performance is complicated in hockey because of frequent line changes, continuity of play, and the infrequency of goals (Gramacy, 2013). Macdonald states that the low scoring rates compared to other sports makes analyzing team and player performance in hockey difficult

(Macdonald, 2012b). Despite this, experts in the hockey analytics world are constantly looking to improve upon the traditionally used box score metrics of goals, assists, and plus-minus rating.

2. Literature Review

There have been many profound jumps in the use of statistics throughout the hockey world as of recent. Many fans and researchers throughout the sport are challenging the box score metrics and are instead looking for new and improved methods to analyze games. The hopes of these new strategies are that they will lead to improved predictability of future events in hockey by finding ways to interpret noisy data from past games into reliable information for future games. Essentially, many fans and researchers are looking to follow-up the successes that sports like baseball and basketball have found through analytics.

2.1 Fenwick and Corsi Ratings

Fenwick rating and Corsi rating are two improved measures that are widely used in the hockey analytics world. The Fenwick rating is calculated by taking the sum of shots on goal for and missed shots for and subtracting off the sum of shots on goal against and missed shots against. Equation (1) below depicts the equation for Fenwick rating.

$$Fenwick = (Shots\ on\ goal\ FOR + missed\ shots\ FOR) - (Shots\ on\ gaol\ AGAINST + missed\ shots\ AGAINST) \quad (1)$$

The Corsi rating is calculated as a product of shots, blocks, and misses against a hockey team subtracted by the total shots, blocks, and misses generated for a team. Equation (2) below depicts the equation for Corsi rating.

$$Corsi\ (C) = CF - CA \quad (2)$$

In this equation CF represents the shots, blocks, and misses for whereas CA represents the shots, blocks, and misses against. Hockey analysts use these statistics to overcome the difficulty of analyzing team and player performance in hockey due to low scoring rates (Macdonald, 2012b). Essentially, there are substantially more data points for shots than there are for goals. Analysis conducted in, "Shots, Fenwick, and Corsi", shows that shot differential, Fenwick differential, and Corsi differential are all better at predicting goal differential than a team's current goal differential is when analyzing data from half of a season (JLikens, 2011; Macdonald, 2012b). Many hockey analysts favor this opinion that these ratings have value, however it is important to take note of the pro's and con's that these statistics entail. Additionally, Macdonald notes the importance of using these statistics alongside of goals as opposed to instead of the traditional goals statistic (Macdonald, 2012a).

2.2 Weighted Shots

Hockey analysts such as Macdonald and Krzywicki have looked at new ways to address the traditional shot metric in a method commonly referred to as weighted shots (Krzywicki, 2005, 2010; Macdonald, 2012d). The shot statistic is only able to tell the number of shots each team takes; however, utilizing new technology, analysts are now able to look at the type of shot taken as well as its location on the ice. Using this information, analysts can weight each shot upon its probability of being a goal (Macdonald, 2012a). Kryzwicki displays this in his equation below (Krzywicki, 2005).

$$P(\text{GOAL}) = \frac{1}{1+e^{-\sum \text{points}}}. \quad (3)$$

In this equation, 'points' is taken from the model scorecard Kryzwicki created as seen below in Figure 1 (Krzywicki, 2005). These points are a result of the weighted shot utilizing the variables for the shot of distance from the net, shot type, rebound, and the situation in terms of man-advantage or disadvantage. To assess NHL shot quality, Kryzwicki utilized a logistic regression model to predict the probability of a goal based upon distance, shot type, rebound, and situation (Krzywicki, 2005). As one might assume, Kryzwicki noticed that shots coming from straight on are the most dangerous and that taking an opponent wider than 30 degrees on either side of the net reduces the likelihood of the shooter scoring (Krzywicki, 2010). Figure 2 below depicts the hockey zone and the region of the ice that is considered to have the better quality of shots. In this picture, the diagonal black line protruding from the net represent the 30 degrees on either side of the net that Krzywicki considers. Additionally, Macdonald's work showed that the top five shots most likely to be a goal were all within 13 feet of the goal (Macdonald, 2012d). Aside from these somewhat intuitive results, there are some other factors that potentially play into the result of a goal. Two of these situations are the "Sniper" – an above average shooter – and the effect of efficient passing (Johns, 2004). For the latter, one might assume that a lateral pass or a pass from the side boards prior to a shot would make that shot more dangerous because it forces the goalie to move (Elomo, 2015). This is also depicted in Figure 2, where the dotted line perpendicular to the net represents the golden line. When passes are made across the line just prior to the shot, that shot is then considered to be more dangerous because it forces the goalie to move. One potential improvement upon the weighted shots statistic could be to address the location of the pass prior to the shot.

Variable	Range	Points
Intercept	Add to all records	-2.2369
Distance	Less than 10 ft	0.6884
	10 - 12 ft	0.6374
	13 - 14 ft	0.5564
	15 - 16 ft	0.5174
	17 - 22 ft	0.3654
	23 - 31 ft	0.0000
	32 - 36 ft	-0.3805
	37 - 38 ft	-0.4758
	39 - 44 ft	-0.8155
	45 - 57 ft	-1.0848
	58 ft or more	-1.3824
Shot Type	Wrap	-0.0742
	Slap	-0.0573
	Wrist	0.0093
	Snap	0.0130
	Backhand	0.0361
	Tip-In	0.1487
Rebound	Yes	1.3362
	No	0.0000
Situation	EV	-0.1244
	SH	0.0399
	PP	0.4007

Figure 1. Model Scorecard (Krzywicki, 2005).

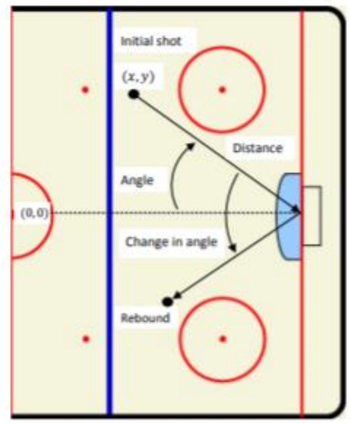

Figure 2. Quality Shot Location and Golden Line (Macdonald, 2012d)

3. Model Development

The Army hockey team at the United States Military Academy is regularly looking for ways to improve their likelihood of winning games. Analytics in sports are quickly gaining traction in the hockey community and the Army hockey team is following suit. The Army hockey team's recent adoption of the Iceberg Analytics software program is one example of this. Iceberg Analytics utilizes artificial intelligence to collect hockey data, and the coaches at West Point can use data from Iceberg Analytics to draw conclusions and aid in future decision making. For this project, raw data was pulled in from Iceberg Analytics for an Army versus Robert Morris game from the 2019-20 hockey season. This raw data set consists of every event that occurred throughout the game, the time of each event, and the location of each event in a x-y coordinate plane. The primary purpose of this model is to expound upon the concept of weighting shots in hockey as discussed previously in the literature review. The model utilizes the raw data from Iceberg Analytics to incorporate shots weighted by location into Corsi and Fenwick scores. Moreover, the model further weights each shot if they were a rebound. Building the model in Microsoft Excel allowed for easy manipulation of the data. Additionally, the model and dataset contain sufficient information for future hockey players and cadets to conduct follow-on research in hopes of improving its current state. The following steps were taken to develop the model. First, the x-y coordinates from the raw data were utilized to map out the hockey rink and decipher its correlation to the actual dimensions of a standard hockey rink. Second, this information was utilized to establish different "shot-zones" on the hockey rink. These "shot-zones" were created by utilizing previous research in hockey to determine which areas on the ice a shot is more likely to result in a goal. For example, a shot that occurs directly in front of the net is more likely to score than a shot along the perimeter boards of the hockey rink. Third, every shot in the data was grouped into either missed shots, blocked shots, rebounds, or shots-on-goal and classified by which team took the shot. Fourth, shots were grouped into their respective "shot-zone" by utilizing the location that they occurred in. Finally, the model was utilized to weight each shot and establish weighted Corsi and Fenwick scores for both Army and against Army. These statistics were notated consecutively as Corsi and Fenwick scores for and Corsi and Fenwick scores against.

3.1 Data Collection

One positive due to the recent surge of analytics in sports, is that data for collegiate and professional hockey is becoming more available. The raw data from Iceberg Analytics is extremely versatile because it encompasses every event throughout each game with its correlating time and location. For this project, data collection consisted of organizing the raw data into useful information for the model. To calculate Corsi and Fenwick scores a count of missed shots, blocked shots, and shots on goal is necessary. However, the raw data classifies shots into six different categories: shot attempt – blocked, shot attempt – deflected, shot attempt – on goal, shot attempt – over the net, shot attempt – wide left, and shot attempt – wide right. For this project, shot attempt – over the net, shot attempt – wide left, and shot attempt – wide right were simplified into one category, “missed shot”. Additionally, the classification of shot attempt – deflected was ignored because including these would result in data overlap. This is because a deflected shot is simply a deflection of a shot that previously occurred. Furthermore, the data was broken down into home and away categories to separate the shots by team. This avoids the potential for a shot to be counted towards the wrong team because of its location on the ice. For example, without separating shots by home and away, an Army shot that was taken from their defensive zone might be counted as a shot for Robert Morris because it also occurred in their offensive zone. Finally, rebound shots had to be classified for the purpose of this project because Iceberg Analytics raw data does not include rebounds. Therefore, rebounds were manually calculated utilizing goalie saves and the timestamp of events. This is discussed further in Section 3.2.

3.2 Establishing “Shot-Zones” and Classifying Rebounds

The standard ice hockey rink in North America is 200 ft. long and 85 ft. wide. Iceberg Analytics’ algorithm plots the rink as a coordinate plane utilizing centimeters where the x-direction represents length and the y-direction is width. Therefore, an ice rink in centimeters is approximately 6066 cm. long and 2600 cm. wide. In an x-y coordinate plane, center ice is depicted as (0,0), the end boards are represented by (3033,0) and (-3033,0), and the side boards are (0,1300) and (0, -1300). Additionally, the ice hockey rink is broken into three different zones: the offensive zone, neutral zone, and defensive zone. This model focuses on the offensive zone, which is where shots with the highest probability of scoring primarily occur. With this, “shot-zones” were established to represent different areas in the offensive zone with varying probabilities of scoring. The offensive zone was broken down into 7 different “shot-zones” denoted as: Zone 1A, Zone 1B, Zone 2, Zone 3A, Zone 3B, Zone 4, and Zone 5. Any shot that occurred outside of the offensive zone was denoted as No Zone. Figure 3 below displays a standard ice-hockey rink with the locations of each “shot-zone”. Of note, Zone 5 is located directly in front of the net; this is the highest probability area of scoring. In contrast, Zone 3A and 3B are in the vicinity of each corner of the offensive zone; these are the lowest probability areas of scoring. Once the zones were established, every shot from the game was placed into its respective zone using the raw data from Iceberg Analytics. Moreover, as shown in Table 1, each “shot-zone” was given a weight that was ranked in order based upon its likelihood of scoring.

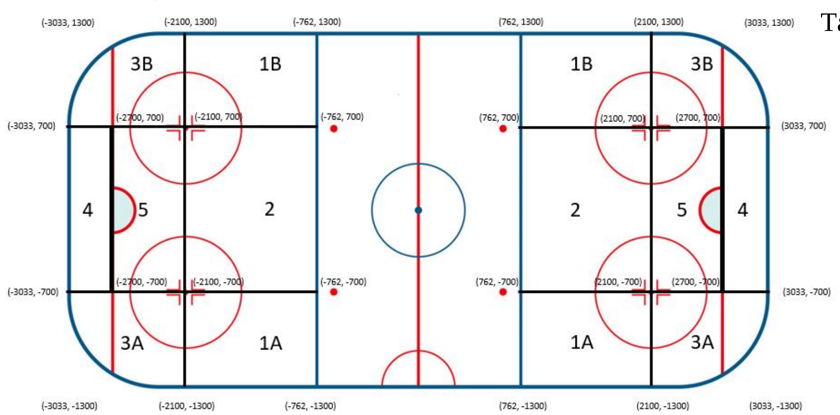

Figure 3. “Shot-Zones” overlayed onto ice rink with x-y coordinate values.

Table 1. Table of Zone and Rebound weights.

Location	Weight
Zone 1A	2
Zone 1B	2
Zone 2	3
Zone 3A	1
Zone 3B	1
Zone 4	4
Zone 5	5
No Zone	0.5
Rebound	1

A rebound in hockey is any shot that occurs directly after a save by the goaltender, without any stoppage in play between the save and shot. To incorporate rebounds into the model, a series of three steps were conducted. First, rebounds were determined by utilizing the raw data to find any shot that occurred within 0.5 seconds of a save by the goaltender. Second, these rebound shots were placed into their respective "shot-zone". Third, the rebound shots were given a weight of 1 as shown in Table 1 above. Rebounds were given this additional weight because the opposing goalie is typically less prepared to make a save on a rebound shot. This is because the goalie is still recovering from the previous shot and is often out of position.

4. Results and Analysis

In standard Fenwick and Corsi scores, simply increasing the total number of missed shots, blocked shots, and shots-on-goal would improve a team's score. However, in the weighted shots model, location and rebound shots are also important. Thus, increasing the number of missed shots, blocked shots, and shots-on-goal that are in higher priority "shot-zones" and increasing rebound shots will produce much higher Fenwick and Corsi scores. The weighted shots model provides a framework for discussion into the importance of shot location and a team's ability to generate rebounds throughout the course of a hockey game or season. The goal of this model is to provide players and coaches with added information when they reference the Fenwick and Corsi scores. Utilizing the weighted shots model, players and coaches can quickly reference these values to determine if they were generating shots from high-priority "shot-zones" and sufficiently creating rebound shots. Additionally, this weighted shots model for Fenwick and Corsi scores allows for easy iteration over entire seasons worth of Army hockey data as well as for other teams in other leagues. The current model also provides the flexibility to adjust the current "shot-zone" locations and their associated weights as deemed necessary. For example, further analysis into where in the offensive zone goals are primarily scored might prompt adjustments to the location or weights of the "shot-zones" in the future.

4.1 Weighted Shots Model

As previously stated, this model utilizes "shot-zones" with respective weights to add value to shots based upon their location. Furthermore, rebound shots are given an additional weight because of their increased likelihood of resulting in a goal. For this project, these weights are used to include the value of location on-ice to the missed shots, blocked shots, and shots-on-goal within the Fenwick and Corsi analytical scores. The weighted shots model of Zone 1 for both Fenwick and Corsi scores for are depicted below as Equations (4) and (5). The same model was conducted for Fenwick and Corsi scores against.

$$Fenwick = \sum \left[Zone1A_M(W_{1A}) + Zone1A_{MR}(W_{1A} + W_R)\right] + \left[Zone1A_{SoG}(W_{1A}) + Zone1A_{SoGR}(W_{1A} + W_R)\right] + \cdots \quad (4)$$

$$Corsi = \sum \left[Zone1A_M(W_{1A}) + Zone1A_{MR}(W_{1A} + W_R)\right] + \left[Zone1A_B(W_{1A}) + Zone1A_{BR}(W_{1A} + W_R)\right] + \left[Zone1A_{SoG}(W_{1A}) + Zone1A_{SoGR}(W_{1A} + W_R)\right] + \cdots \quad (5)$$

The Fenwick score for Army consists of the weighted sum for missed shots and shots-on-goal from each "shot-zone". The Corsi score is the weighted sum for missed shots, blocked shots, and shots-on-goal that occurred in each "shot-zone". To calculate the Fenwick and Corsi scores for Army, Equations (4) and (5) would be iterated for each "shot zone" including the "No Zone" classification. Again, the same method was conducted for Fenwick and Corsi scores against Army. For reference, W_{1A} represents the weight for location Zone 1A and W_R is the weight for a rebound shot. Furthermore, $Zone1A_M$, is the total number of missed shots that originated in Zone 1A, whereas $Zone1A_{MR}$ is the total number of missed rebound shots that occurred in Zone 1A. Similarly, $Zone1A_B$, is the total number of blocked shots in Zone 1A and $Zone1A_{BR}$ is the total number of blocked rebound shots in Zone 1A. Finally, $Zone1A_{SoG}$, is the total number of shots-on-goal in Zone 1A and $Zone1A_{SoGR}$ is the total number of rebound shots-on-goal originating in Zone 1A.

Missed shots, blocked shots, and shots on goal were then plotted onto a standard ice-rink, displayed as Figure 4, to visually depict the different "shot-zones". Army was the away team in this game, meaning the orange data points represent the shots for Army and the blue data points shots against Army, i.e., Robert Morris shots. Each shot is signified by MS, BS, or SoG representing its classification of missed shot, blocked shot, or shot on goal. Additionally, shots are denoted with the respective numerical value for the "shot-zone" they originated from. Figure 4 is valuable because it allows a quick analysis into which team generated a greater quantity of shots from higher value "shot-zones". In this game, for example, there were significantly more shots for Army generated from the regions around the goal and closer to center ice than the side boards. These regions contain Zone 2, Zone 4, and Zone 5 in the weighted shots model. These are also the "shot-zones" with the largest weights because goals typically occur more often in these areas of the ice. Furthermore, there were significantly more shots against Army that originated from regions closer to the side boards than the middle of the ice. The "shot-zones" closer to the side boards hold smaller weighted values because these shots are less likely to result in a goal. Therefore, plotting shots on the

ISBN: 97819384962-0-2

ice-rink provides the swift conclusion that the weighted shots model of Army's Fenwick and Corsi scores for should return a greater value than their Fenwick and Corsi scores against.

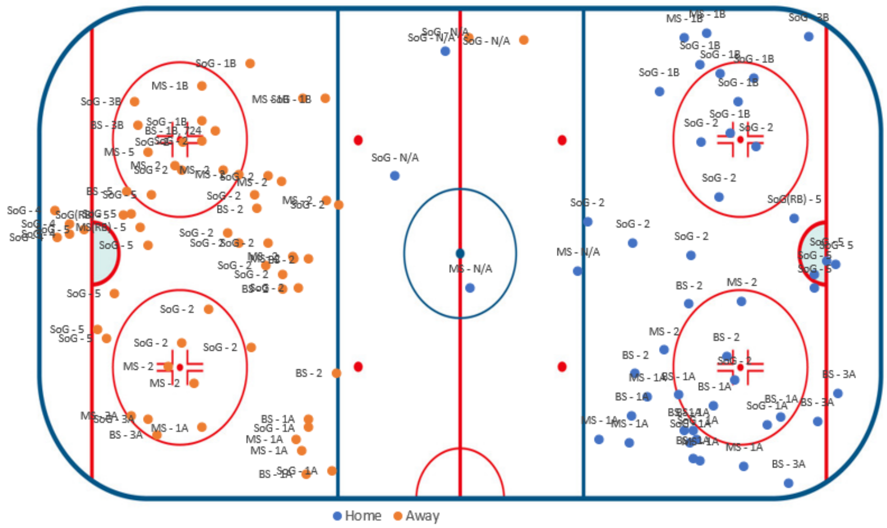

Figure 4. Plot of shots for Army (Orange) v. Robert Morris (Blue) 2019-20

4.2 Comparison of Weighted and Standard Fenwick and Corsi

Applying the weighted shots model to the Army versus Robert Morris game from the 2019-20 hockey season resulted in a Fenwick score for of 164 and a Corsi score for of 189 for Army. Conversely, Army earned a Fenwick score against of 88 and a Corsi score against of 114. Again, the Corsi and Fenwick scores for Army correlate to the respective shot types that Army produced, whereas the Corsi and Fenwick scores against Army relate to the respective shot types that Robert Morris created. These values provide minimal value initially, however when compared to the standard Fenwick and Corsi scores the benefits of weighting shots by location and rebounds are present. The standard Fenwick and Corsi scores for Army were 53 and 63 respectively. Furthermore, Army earned a standard Fenwick score against of 35 and a standard Corsi score against of 48.

Utilizing ratios between the for and against scores exhibits the effect that weighting shots holds over the Fenwick and Corsi ratings. To determine the ratio score for standard Corsi, the standard Corsi score for was divided from the standard Corsi score against. The same process was utilized to also determine the standard Fenwick ratio, the weighted Corsi ratio, and the weighted Fenwick ratio. Ultimately, the comparison between the standard and weighted ratios helps to reveal whether weighting the shots better explains the events that are occurring throughout the game or not. The goal is that weighting shots will improve the Corsi and Fenwick scores by using "shot- zones" to incorporate a shot quality aspect to the original Corsi and Fenwick scores that only provide shot quantity information. Essentially, the idea is to prove that including quality of shots information to the scores will provide more value to the statistic than simply utilizing the quantity of shots.

The ratio for standard Corsi scores between Army for and against yields a ratio score of 0.7619 however the ratio score between weighted Corsi scores is 0.6032. In addition, the ratio of standard Fenwick scores between for and against resulted in a ratio score of 0.6604 and a weighted ratio score of 0.5366. In both scenarios, the weighted ratio scores returned lower values showing that a significant portion of Army's shots originated in lower strength "shot-zones". Finally, the weighted Fenwick and Corsi ratings were calculated using Equation (1) and (2) from section 2.1. The new weighted ratings were a Fenwick rating of 76 and Corsi rating of 75. Each of these values were calculated and compiled into Table 2 below.

Analysis into the outcome of the game offers further insight towards what is taking place over the course of the game. This game resulted in a 4-0 win for Army, a strong reflection of the higher values held by Army in both standard and weighted Fenwick and Corsi scores for. Despite this, weighted Fenwick and Corsi scores portray a different story than both the final score and standard Fenwick and Corsi scores. The standard Fenwick ratio of 0.6604 shows a dominant performance by Army, however the weighted Fenwick ratio of 0.5366 reveals a much closer game. Similarly, the standard Corsi ratio of 0.7619 displays superiority by Army, yet the weighted Corsi ratio of 0.6032 is a more even game. This indicates that Army created more shots than Robert Morris throughout the game, but many of those shots were not taken from high priority "shot-zones" or rebound shots. In contrast, Robert Morris generated less shots over the course of the game but many of the shots were from higher priority "shot-zones" or rebounds.

Table 2. Summary of game totals for Army v. Robert Morris University

Home - Robert Morris University

Goals - Home	0

		Rebound?	Weighted
Missed Shot, Zone 1A - Home	5	0	10
Missed Shot, Zone 1B - Home	2	0	4
Missed Shot, Zone 2 - Home	2	0	6
Missed Shot, Zone 3A - Home	0	0	0
Missed Shot, Zone 3B - Home	0	0	0
Missed Shot, Zone 4 - Home	0	0	0
Missed Shot, Zone 5 - Home	0	0	0
Missed Shot, No Zone - Home	2	0	1
Total Mised Shot - Home	11		21
Blocked Shot, Zone 1A - Home	7	0	14
Blocked Shot, Zone 1B - Home	0	0	0
Blocked Shot, Zone 2 - Home	3	0	9
Blocked Shot, Zone 3A - Home	3	0	3
Blocked Shot, Zone 3B - Home	0	0	0
Blocked Shot, Zone 4 - Home	0	0	0
Blocked Shot, Zone 5 - Home	0	0	0
Blocked Shot, No Zone - Home	0	0	0
Total Blocked Shot - Home	13		26
Shot on Goal, Zone 1A - Home	3	0	6
Shot on Goal, Zone 1B - Home	6	0	12
Shot on Goal, Zone 2 - Home	7	0	21
Shot on Goal, Zone 3A - Home	0	0	0
Shot on Goal, Zone 3B - Home	1	0	1
Shot on Goal, Zone 4 - Home	0	0	0
Shot on Goal, Zone 5 - Home	4	1	26
Shot on Goal, No Zone - Home	2	0	1
Total Shot on Goal - Home	24		67

Away - Army

Goals - Away	4

		Rebound?	Weighted
Missed Shot, Zone 1A - Away	3	0	6
Missed Shot, Zone 1B - Away	2	0	4
Missed Shot, Zone 2 - Away	9	0	27
Missed Shot, Zone 3A - Away	1	0	1
Missed Shot, Zone 3B - Away	0	0	0
Missed Shot, Zone 4 - Away	0	0	0
Missed Shot, Zone 5 - Away	1	1	11
Missed Shot, No Zone - Away	0	0	0
Total Missed Shot - Away	17		49
Blocked Shot, Zone 1A - Away	2	0	4
Blocked Shot, Zone 1B - Away	1	0	2
Blocked Shot, Zone 2 - Away	4	0	12
Blocked Shot, Zone 3A - Away	1	0	1
Blocked Shot, Zone 3B - Away	1	0	1
Blocked Shot, Zone 4 - Away	0	0	0
Blocked Shot, Zone 5 - Away	1	0	5
Blocked Shot, No Zone - Away	0	0	0
Total Blocked Shot - Away	10		25
Shot on Goal, Zone 1A - Away	2	0	4
Shot on Goal, Zone 1B - Away	3	0	6
Shot on Goal, Zone 2 - Away	15	0	45
Shot on Goal, Zone 3A - Away	1	0	1
Shot on Goal, Zone 3B - Away	1	0	1
Shot on Goal, Zone 4 - Away	4	0	16
Shot on Goal, Zone 5 - Away	7	1	41
Shot on Goal, No Zone - Away	2	0	1
Total Shot on Goal - Away	36		115

	Standard	Weighted
Corsi Score Against	48	114
Fenwick Score Against	35	88
Corsi Score For	63	189
Fenwick Score For	53	164
Corsi Rating	15	75
Fenwick Rating	18	76

Standard Corsi Ratio	0.7619
Standard Fenwick Ratio	0.6032

Weighted Corsi Ratio	0.6604
Weighted Fenwick Ratio	0.5366

4.3 Recommendations

As discussed in the introduction, hockey data contains a lot of noise due to the high volume of activity at any time throughout a given game. Therefore, analytics in hockey should be viewed from a holistic perspective. Reviewing the Fenwick and Corsi ratio scores from the weighted shots model shows that including a location and rebound aspect will improve the information in each score. Despite this, the weighted shots model will not directly correlate to which team wins or outplays the other. Numerous factors outside of shot location and presence of rebounds can affect these outcomes. The benefit of this model is to improve the information analysts, coaches, or players receive from the Fenwick and Corsi ratings. Previously, these ratings only presented information containing which team generated a higher amount of missed shots, blocked shots, and shots on goal. Utilizing the weighted shot model, the Fenwick and Corsi scores will display the same information plus the added benefit of providing insight towards shot location and influence from rebound shots. Knowing this, coaches or players can utilize this number to determine if their team is producing shots or rebounds from higher quality “shot-zones”. In the Army versus Robert Morris example, Army’s coaches can utilize the standard Fenwick and Corsi scores to notice that they produced a high volume of shots. Additionally, they can reference the Fenwick and Corsi scores from the weighted shot model to realize they can improve by focusing these shots into higher weighted “shot-zones” or producing more rebounds.

5. Conclusion

This project proposes a new method for weighting shots in hockey. The proposed weighted shot model incorporates weights for rebound shots and varying weights for shots dependent upon the location they originated from on the ice. This weighted shot model allows the Fenwick and Corsi ratings to present more descriptive information to include strength for location of shots and rebounds. By adding shot weights, the ratio scores between Army for and against shifted from 0.6604 to 0.5366 for Fenwick and from 0.7619 to 0.6032 for Corsi. These simple additions show that weighting shots with the current model adjusts Fenwick and Corsi ratings to reflect "shot-zone" location and presence of rebound shots. This adjustment allows for quick analytical feedback to include a team's ability to produce shots from quality locations and generate shots from rebounds.

Furthermore, the weighted shot model can be used to analyze a larger sample of games or even seasons worth of data. Any game, utilizing Iceberg Analytics' raw data, can be easily placed into the Microsoft Excel spreadsheet to quickly generate weighted Fenwick and Corsi ratings. Due to this, the weighted shots model can be handed over to future hockey players, coaches, or cadets for further analysis in the future. The model is also very flexible. If different locations for the "shot-zones" are deemed necessary, the x-y coordinates for the zones can be altered and simulated in the future. Additionally, the weights for each "shot-zone" and for rebounds are easily manipulated if further research and analysis proves beneficial. In the future, it would be beneficial to expand research into NHL data for a larger data set. This is easily attainable with the NHL data Application Programming Interface (API), which provides access to analytical data for every team in the NHL. Teams in the NHL play significantly more games in a season than at the collegiate level, establishing a larger pool of data to simulate using this weighted shot model.

6. References

Elitzur, R. (2018). Data Analytics Effects in Major League Baseball. *Omega.* www.sciencedirect.com/science/article/pii/S0305048318300215.

Elomo, M., & Poikonen, T. (2015). *Analyzing reasons behind the goals in ice-hockey*. https://www.theseus.fi/bitstream/handle/10024/95439/Elomo%20Poikonen%20Final.pdf?sequence=1.

Found, R. (2016). Goal-based Metrics Better Than Shot-based Metrics at Predicting Hockey Success. *The Sport Journal*.

Gramacy, R., Taddy, M., & Jensen, S. (2013). Estimating Player Contribution in Hockey with Regularized Logistic Regression. *Journal of Quantitative Analysis in Sports.* 9. 10.1515/jqas-2012-0001.

JLikens. (2011, February 16). Shots, Fenwick and Corsi. http://objectivenhl.blogspot.com/2011/02/shots-fenwick-and-corsi.html.

Johns, G. (2004, October 17). *Statistical Shot Quality Weighing*. Hockey Analytics. http://hockeyanalytics.com/Research_files/Weighted_Shots.pdf.

Krzywicki, K. (2005, January). *Shot Quality Model*. http://www.hockeyanalytics.com/Research_files/Shot_Quality_Krzywicki.pdf.

Krzywicki, K. (2010, June). *NHL Shot Quality 2009-10*. http://hockeyanalytics.com/Research_files/SQ-RS0910-Krzywicki.pdf.

Macdonald, B. (2012a). Adjusted Plus-Minus for NHL Players using Ridge Regression with Goals,Shots, Fenwick, and Corsi. *Journal of Quantitative Analysis in Sports*. 8. 10.1515/1559-0410.1447.

Macdonald, B. (2012b). An Expected Goals Model for Evaluating NHL Teams and Players. In *Proceedings of the 2012 MIT Sloan Sports Analytics Conference*.

Macdonald, B. (2012c). Adjusted Plus-Minus for NHL Players using Ridge Regression with Goals,Shots, Fenwick, and Corsi. *Journal of Quantitative Analysis in Sports*. 8. 10.1515/1559-0410.1447.

Macdonald, B., Lennon, C., & Sturdivant, R. (2012d). *Evaluating NHL Goalies, Skaters, and Teams Using Weighted Shots.*

Petriello, M. (2018). 9 Things You Need to Know About the Shift. mlb.com/news/9-things-you-needto-know-about-the-shift-c276706888.

Sawchik, T. (2019). Don't Worry, MLB - Hitters Are Killing The Shift On Their Own. *FiveThirtyEight*. https://fivethirtyeight.com/features/dont-worry-mlb-hitters-are-killing-the-shift-on-their-own/.

Sheps, S. W. (2020). *Corsi, Fenwick and Gramsci: How bloggers and advanced analytics are changing the National Hockey League - Stephen W Sheps, 2020*. SAGE Journals. https://journals.sagepub.com/doi/10.1177/1012690219869192.