Statistical Evaluation of Ice Hockey Goaltending

Michael E. Schuckers

May 3, 2016

Contents

1 Statistical Evaluation of Ice Hockey Goaltending **1**

1.1 Metrics and Notation 1

1.1.1 Wins 3

1.1.2 Goals Against Average 3

1.1.3 Save Proportion 4

1.2 Adjusted Save Proportion 8

1.2.1 Shot Probability Models 10

1.3 Rebounds 18

1.4 Discussion and Future Directions 20

1.5 Addendum: Pulling the Goalie 22

Chapter 1

Statistical Evaluation of Ice Hockey Goaltending

In this chapter we look at statistical methods for assessing the performance of ice hockey goaltenders. Since the time when ice hockey metrics began being recorded, several measures of goalie performance have been proposed. As we shall see, some of these incorporate factors that are beyond the control of goalies. We will assess metrics on goalie performance on how correlated they are with results from within the same season and how correlated their are with results from subsequent seasons. In particular, we look at a goalie's save proportion, their adjusted save percentage based upon shot difficulty and their ability to control rebounds from saved shots.

1.1 Metrics and Notation

Historically, there have been several metrics that have been proposed to assess and evaluate hockey goaltenders or goalies. These roughly are wins, goals against average (GAA) and save proportion (SVP). Each of these metrics has some merits, though SVP and its variants are generally considered the best of these for the evaluation of goalies. Below we discuss each of these in turn.

Notation

Throughout this chapter we will use a variety of different variables and metrics in order to evaluate goaltender performance. Here we introduce the basic notation that we will use. We will denote information about a shot using a vector of attributes $\mathbf{z}_i$ for the i^{th} shot from among n total shots with $i = 1, \ldots, n$. We will use an indicator function, Φ_i to denote whether or not a shot is saved using

$$\Phi_i = \begin{cases} 1 & \text{if shot } i \text{ is saved,} \\ 0 & \text{if shot } i \text{ is not saved.} \end{cases} \tag{1.1}$$

Since time on the ice is sometimes important in rate metrics, we will denote the time on the ice for a particular goalie over some interval, for example, a season by T. As needed we will use a subscript of j for the j^{th} goalie. We will use J to represent the total number of goalies considered in a particular analysis with $j = 1, \ldots, J$. For simplicity we will drop the goalie subscript if it is clear that we are considering a single goalie. Often we will make use of the additional measurements that are available about the i^{th} shot. The other additional notation that we will use is a result of information about a particular shot, $\mathbf{z}_i$. For such a shot, we will say the strength of the team taking the shot is denoted by s where s takes the values: EV (Even Strength), PP (Power Play), PK (Penalty Kill). As we will see below different types of shots will have different probabilities of becoming goals. The type of shot will be denoted w where the basic shot types are: Backhand, Deflection, Slap, Snap, Tip-In, Wrap(around) and Wrist. Additionally, for shots we may have location from which a shot was taken. From that location we will obtain the coordinates of a shot relative to center ice which will take the place of the origin $(0, 0)$. We will use x and y to denote the vertical and horizontal distances from center ice. The orientation here is that the x-axis goes through the center of both goals. Further we can also obtain the distance of a given shot from the center of the goal and the angle of the shot relative to the horizontal axis of the shot. We will use d to represent that distance and θ to represent the corresponding angle.

All of the information that we have regarding shots taken from the National Hockey League (NHL) comes from the NHL's Real Time Scoring System (RTSS). A description of the RTSS system can be found [Kasan, 2008]. Many authors including [Desjardins, 2010a], [Bruce McCurdy, 2010], [Fischer, 2010], [Zona, 2011] and [Awad, 2009] among others have noted that there are issues, specifically measurement error and bias, with the collection of these data. In particular, there seem to be some substantial differences in the recording of shot location from rink to rink in the NHL. We will discuss some of these issues below as well as some of the proposed remedies. Throughout this chapter we will use data from RTSS for the 2009-10 through the 2012-13 regular seasons.

1.1.1 Wins

The evaluation of goaltenders based upon wins is one manner to evaluate the contribution of a player to their team's performance. A win is awarded to a goalie if they are on the ice for the game winning goal. For many analysts, this suggests that winning is very team dependent. While it is common to have a goalie with a high save proportion win a large number of games, for example, during the 2011-12 season Jonathan Quick won 35 games with the fifth highest goalie save proportion, it sometimes happens that a goalie wins a very high number of games having had a save proportion that is not well above the league average. The latter can be exemplified by the noting that Marc-André Fleury won 37 games (out of 65 started and 67 total) during the 2009-10 season while having a 0.905 save proportion. The league average save proportion that season was 0.908. That winning tends to be associated with quality of a team rather than with quality of a goalie suggests that wins is not a strong measure of goalie performance.

1.1.2 Goals Against Average

Goals Against Average, or GAA, is the average number of goals that a goalie concedes per sixty minutes of ice time where 60 minutes is the regulation length of a hockey game, i.e. the length of a hockey game that does not go to overtime. The calculation of GAA is then number of goals conceded divided by time on ice (in minutes) multiplied by 60. Using our notation, this is

$$GAA = 60\left[\frac{n - \sum_{i=1}^{n} \Phi_i}{T}\right]. \tag{1.2}$$

We note that the GAA can be written as a function of a goalie's SVP in the following way:

$$GAA = 60\left[\frac{n}{T}(1 - SVP)\right]. \tag{1.3}$$

One result of writing GAA in this manner is that we can explicitly see that GAA is a function of time on ice and number of shots faced which are ultimately not under the direct control of a goalie. Consequently, it seems that SVP is a better choice for the evaluation of goalie performance than GAA.

1.1.3 Save Proportion

The most widely accepted traditional measure of goalie performance is the save proportion. One primary reason for this is that it is a rate over which a goalie has some measure of control, unlike wins and GAA which are dependent on the other things that the goalie's team is doing, scoring lots of goals or giving up lots of shots, respectively. A goalie's SVP is calculated as the number of shots saved divided by the number of shots faced. This is often done on a per game or per season basis. We can think of SVP as

$$SVP = \frac{\sum_{i=1}^{n} \Phi_i}{n}.$$

Figure 1.1 shows the average SVP in the NHL since the 1983-84 season. Any historical analysis needs to adjust for these changes which are likely due to changes in goalie athleticism and equipment. The increasing pattern here is discussed by [Paine, 2014]. Over this period goaltending styles improved and the size of equipment became larger. Both of these effects contributed to this trend.

More generally we can think of SVP as a function of the quality of goaltending and the

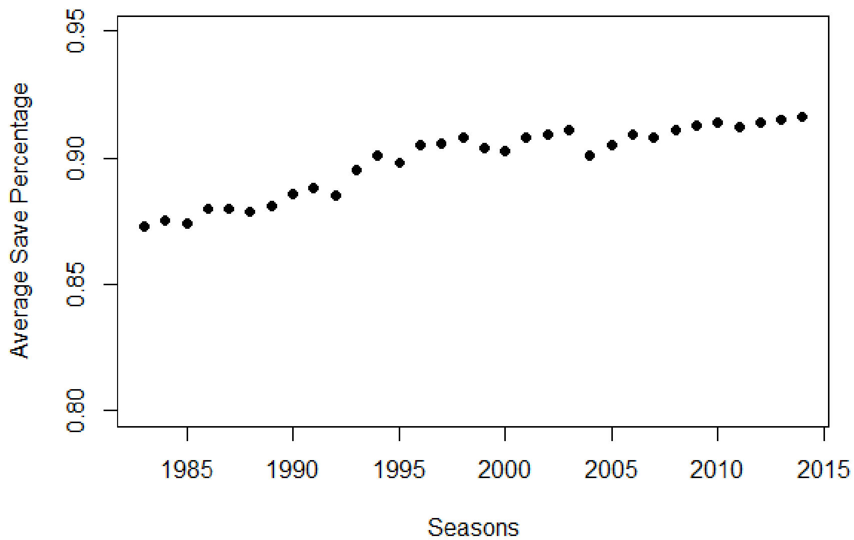


Figure 1.1: NHL League Average save proportion by Season 1983 to present

quality of shots faced. To that end, we can think of SVP as a weighted average with

$$SVP = \sum_{u=1}^{\mathcal{U}} G(u)S(u). \tag{1.4}$$

where $G(u)$ is a probability that a goalie saves a shot with attributes $\mathbf{z}_u$ and $S(u)$ is the probability of facing a shot with attributes $\mathbf{z}_u$. For this equation, $\mathbf{z}_u$ is a vector of descriptors for the attributes of a shot and $\mathbf{z}_u = (x, y, d, \theta, w, s, \ldots)^T$ where $u = 1, \ldots, \mathcal{U}$. These attributes might include things like the type of shot, the location of the shot and others. In this way, we can think of SVP as an expected save proportion taken with respect to some measure of shot distribution. This formulation will be especially useful in the next section where we consider adjustments to SVP. One reason that these adjustments are used is the possibility of differences between the distribution of shot types, i.e. different goalies could face different $S(u)$'s.

To understand our notation here, consider the following example using information from Table 1.1. Suppose that we have three goalies, A, B and C, who play in the same league and who faced four shots with $\mathcal{U} = 4$ different attribute types as listed in Table 1.1. Additionally,

Table 1.1: Example of notation

Shots Attributes, $\mathbf{z}_u$		Goalie A		Goalie B		Goalie C	
Shot Type	Distance	$G(u)$	n_u	$G(u)$	n_u	$G(u)$	n_u
Slap	Close	0.900	1200	0.880	100	0.880	500
Slap	Far	0.930	200	0.920	700	0.920	500
Wrist	Close	0.920	500	0.900	300	0.900	500
Wrist	Far	0.930	100	0.910	900	0.910	500
Overall		0.9095	2000	0.9105	2000	0.9025	2000

n_u is the number of shots with a given set of attributes that each goalie faced where we will let $n = \sum_{u=1}^{\mathcal{U}} n_u$ be the total number of shots faced of all attribute types. The overall SVP of each goalie is given in the last row of the aforementioned table. That is,

$$SVP = \sum_{u=1}^{\mathcal{U}} [G(u)] \frac{n_u}{n} = \sum_{u=1}^{\mathcal{U}} [G(u)] S(u). \tag{1.5}$$

So for goalie A, we have that

$$SVP_A = (0.900)\frac{1200}{2000} + (0.930)\frac{200}{2000} + (0.920)\frac{500}{2000} + (0.930)\frac{100}{2000} = 0.9095. \tag{1.6}$$

Similar calculations for goalies B and C yield overall save proportions of 0.9105 and 0.9025, respectively. Moving to our notation for the rest of the chapter, we will use $S(u) = \frac{n_u}{n}$. The more general notation of $G(u)$ and $S(u)$ will be used when we want to make adjustments to the SVP. Two things are also noteworthy from this example. First note that Goalie A has a lower overall save proportion than Goalie B because of the distribution of shot attributes despite being better than Goalie B at saving each type of shot. These examples are synthetic but they serve to illustrate that SVP is a function of $S(u)$. Similarly, Goalie C has a lower save proportion than Goalie B despite having the same save proportion for shots with the same attributes.

There has long been a discussion in 'hockey analytics' circles about whether or not there are differences between the $S(u)$ faced by different goalies over substantial periods of time, say more than half of an NHL season. This debate revolves around the notion that $S(u)$

for different goalies or different teams converges to a mean $S(u)$, say $\overline{S}(u)$ over time. Consequently, there is not an advantage in the prediction of future outcomes to be gained by adjusting for the differences in shot distributions. This debate is often referred to under the heading of shot quality [1]. Implicit in these discussions is that what is meant by shot quality is instead shot probability and average shot probability at that. All of these discussions are dependent upon the modelling of shot probability, a topic we take up later in this chapter.

Table 1.2: Intraseason correlation of Save Proportion (SVP) for even and odd shots

Season	NHL Average save proportion	More than 500 shots Correlation	J	More than 750 shots Correlation	J
2009/10	0.9198	0.203	48	0.072	33
2010/11	0.9221	0.282	46	0.385	32
2011/12	0.9215	0.180	45	0.165	35
2012/13	0.9203	0.145	24	0.252	9

The reliability of SVP has led some commentators and authors to suggest that goaltenders are difficult to assess[2]. Table 1.2 shows the within season correlation between a goalies even and odd shots faced in our data. Correlations are taken over goalies and J is the number of goalies involved in each correlation. Here we are limiting ourselves to looking at even strength 5-on-5 non-empty net shots as we will do throughout this chapter. We consider even versus odd shots in an attempt to balance the variety of factors that can possibly impact a goalie's SVP such as the players on the ice, the team they are playing against, the rink in which the game is played etc. By using even shots and odd shots we hope to split our data in such a way as to have two groups of data for which the impact of these factors is roughly the same. While the year to year correlations do fluctuate, there is some consistency in these correlations. These results mirror those found in, for example, [Macdonald et al., 2012]. In both the results for goalies who faced more than 500 shots and for those who faced more than 750 shots, the correlations average around 0.210 which is relatively weak. Note that there are fewer goalies in the results for the 2012/13 season because it was shortened due to the lockout. In Table 1.3 we look at the correlation in SVP from year to year among

[1] It is often said that shot quality in the NHL does not matter or it matters very little. See, for example, this blog post by Eric Tulsky, `http://nhlnumbers.com/2012/7/3/shot-quality-matters-but-how-much`.

[2] The phrase that is commonly used in hockey analytics is that *goaltending is voodoo* or *goalies are voodoo*. See, for example, `http://grantland.com/the-triangle/2015-nhl-awards-andrew-hammond-carey-price-mark-giordano/`

goalies facing 500 or more shots in each season. What is clear is that there is not a strong relationship between a goalie's save proportion in a given year and their save proportion in the subsequent year. The correlations range from -0.038 to 0.909 though the latter is based upon only 8 goalies from the lockout shortened 2012/13 season. Again, this uncertainty around the ability to predict future SVP confirms what some authors have suggested: it may be difficult to evaluate goaltenders on less than three years of data, [Desjardins, 2010b].

Table 1.3: Correlation between a goalie's SVP in one year and their SVP in subsequent years given a number of shots faced in each season.

Seasons	More than 500 shots		More than 750 shots	
	Correlation	J	Correlation	J
2009/10 v 2010/11	0.186	34	0.120	22
2010/11 v 2011/12	0.018	37	-0.038	28
2011/12 v 2012/13	0.060	20	0.909	8

While it is the case that save proportion is generally the best traditional metric to evaluate a goaltender, it has some issues. Foremost among these is that there is a good deal on variation in goaltender performance, even for a single goaltender, in the short term and that makes prediction of future performance difficult at best. We have seen this in the correlation results discussed above. Note that a goaltender who faces about 1500 shots and saves them at a rate of 0.925 would have a standard error of their save proportion of about 0.007 assuming a binomial distribution. Two standard errors from 0.925 yields an interval of 0.909 to 0.939. In terms of goalie performance this is a wide range of values. There are other factors as well including the distribution of shots faced. Below we will look at some attempts to deal with this inability to predict future save proportion.

1.2 Adjusted Save Proportion

There have been several attempts to create adjusted save proportion metrics. In this section, we describe some of these. The main motivations for these adjustments is to account for variation in shot difficulty, or the distribution of shot difficulty, and to improve upon our ability to predict a goalie's future SVP (or even strength, non empty net SVP). These approaches fall into two categories and use different information about shots that are part of $\mathbf{z}$. Both of these were first formulated by [Ryder, 2004]. The first type of adjusted save proportion essentially takes the league average for goaltending ability for a given set of shots. We will call this $\text{aSVP}_{\overline{G}}$. This can be calculated as

$$\text{aSVP}_{\overline{G}} = \sum_{u=1}^{\mathcal{U}} \overline{G}(u)S(u) \tag{1.7}$$

where $\overline{G}(u)$ is the league average probability that a shot with attributes $\mathbf{z}_u$ is saved. This metric has been traditionally referred to as the adjusted save proportion and is useful in comparison to the SVP. What this adjustment gives is the save proportion relative to what an average goalie's save proportion would have been, given the shots that an individual goalie faced. Comparisons of $\text{aSVP}_{\overline{G}}$ for different goalies will have different $S(u)$'s and so are still dependent upon the distribution of shots faced. This metric does allow for a comparison of how a goalie did relative to the league average on the shots they faced but each goalie faces a different distribution of shots which makes comparison across goalies problematic.

The second type of adjusted save proportion which we will refer to as shot quality neutral save proportion. We will denote this by $\text{aSVP}_{\overline{S}}$ and calculate it as the following:

$$\text{aSVP}_{\overline{S}} = \sum_{u=1}^{\mathcal{U}} G(u)\overline{S}(u) \tag{1.8}$$

where $\overline{S}(u)$ is the percentage of all shots in the league that are of attribute type u. This approach is useful because it allows direct comparison of individual goalies since they are compared based upon the same distribution of shot types. It is this general approach that has been adopted by `war-on-ice.com` for their adjusted save proportion, [War-On-Ice.com, 2014], and by the Defense Independent Goalie Rating of [Schuckers, 2011]. Below we will look at using this second type of aSVP, $\text{aSVP}_{\bar{S}}$, to predict future aSVP.

As mentioned above, all of the various methods that have been proposed for adjusting the save proportion fall into one of these two basic approaches. There is devil in the details of all of the methods that have been proposed by various authors to calculate these aSVP's. In some cases the authors use statistical smoothing or regression to estimate the goalie ability functions, the $G()$'s, and the shot distribution functions, the $S()$'s. Further, different authors have used different information about each shot to calculate their adjustments. Below we will go through some of the more prominent adjustments and place them into the framework we have established with aSVP$_{\overline{S}}$ and aSVP$_{\overline{G}}$.

1.2.1 Shot Probability Models

The goal of any shot probability model is to estimate the probability that a shot results in a goal given information about a particular shot. This is equivalent to the estimation of $G(u)$ in our earlier notation. As mentioned above, one of the earliest looks at aSVP was by [Ryder, 2004] who also developed the notation of shot neutral adjusted save proportion, our aSVP$_{\overline{S}}$. A logistic, or log odds, model for shot probability was introduced by [Krzywicki, 2005]. The final model that Kryzwicki reports in that paper includes "distance, shot type, rebound and situation" with each term additively effecting the log odds of the probability of a shot being a goal. Subsequently, Kryzwicki introduced an adjustment for rink biases in the recorded distances in [Krzywicki, 2009]. The adjusted distance which replaced distance in the previous model was a single quantity added (or subtracted) to each shot recorded in a particular rink. Again using a logistic linear model, Kryzwicki added an indicator for whether or not a shot was preceded by a giveaway. Again this model did not include any interaction terms. The following year in [Krzywicki, 2010], he introduced shot angles as an added addition to his logistic model. All of these models lead to predicted values for a shot. We will call the predicted probability that a given shot results in a goal $\hat{p}(\mathbf{z}_i)$. This can be thought of as one minus the predicted average save probability for this quantity. So that $\overline{G}(\mathbf{z}_i) = 1 - \hat{p}(\mathbf{z}_i)$ where $\overline{G}(\mathbf{z}_i)$ is a smoothed or averaged goalies ability to save a particular shot.

In many ways the above models and approaches are spatial in nature. They each try to model a probability surface where the x and y axes represent locations on a rink. (Note that in some cases x and y are replaced by polar coordinates, radius and angle.) [Schuckers, 2011] introduced the Defense Independent Goalie Rating, DIGR, which is a shot quality neutral adjusted save proportion using a general nonparametric spatial model. That is, their model did not assume a particular parametric form for the relationship between the x and y locations and the probability of a shot being a goal. Most previous models assumed a linear form for the probability surface on either the original or logistic scale. Further this model assumed that the impact of shot type, w, and strength, s were not consistent across locations. Therefore, DIGR effectively includes interactions for the factors of location, strength and shot type. Another innovation that [Schuckers, 2011] introduced was a form of shrinkage analysis by creating shot probability maps for each goalie that were weighted to the league average of all shots of a particular strength and shot type. Then the DIGR for the j^{th} goalie can be written as

$$\begin{aligned} DIGR_j &= \sum_{u=1}^{\mathcal{U}} \widetilde{G}_j(u)\overline{S}(u) \\ &= \sum_{u=1}^{\mathcal{U}} \left[\alpha(u)G_j(u) + (1-\alpha(u))\overline{G}(u)\right]\overline{S}(u) \end{aligned} \tag{1.9}$$

where $\alpha(u) = \frac{n_u}{n_u+n^\dagger}$, n_u is the number of shots faced by goalie $j, j = 1, \ldots, J$ of shottype w at strength s and $n^\dagger$ is a 'shrinkage' constant. Thus $\widetilde{G}_j(u)$ shrinks the save proportion of goalie j on shots with attributes u toward the league average, $\overline{G}(u)$, for those types of shots. $\alpha(u)$ determines the amount of shrinkage. If $n^\dagger = 0$ then $\alpha(u) = 1$ and DIGR is just the aSVP$_{\overline{S}}$ and as $n^\dagger$ gets large relative to n_u for each u then $\alpha(u)$ approaches zero and DIGR becomes close to the league average. The larger the value of $n^\dagger$ the more similar we are assuming is the performance of each goalie is to the league average for the characteristics, the $\mathbf{z}_u$'s, of shots. This makes for a broader, more flexible shot probability model than had been previously proposed.

The next improvement to the shot probability model was proposed by [Macdonald et al., 2012]. Their innovation was to incorporate the change in angle from the previous shot of a rebound shot to a logistic regression model of goal probability. This variable, change in angle, which we will denote by δ, is illustrated in Figure 1.2. The model that Macdonald *et al* use in that paper also have factors for length of time that the shooter was on the ice (shooter fatigue), length of time that the defense was on the ice (defense fatigue), length of time the offense was on the ice (offense fatigue), whether the shot was by

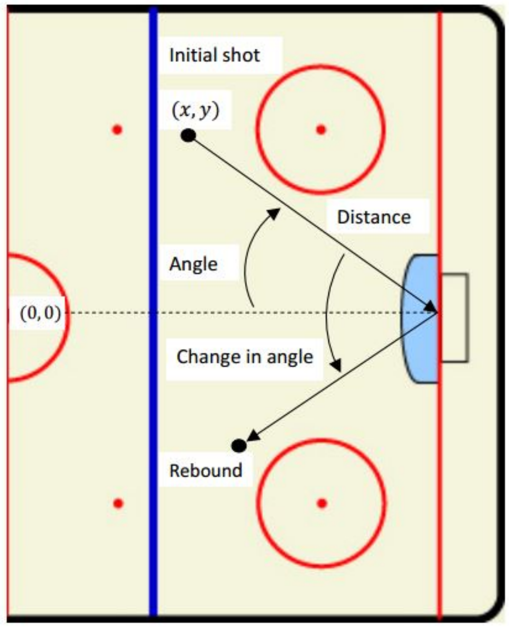


Figure 1.2: Visual explanation of distance, shot angle, and change in angle due to a rebound from [Macdonald et al., 2012]. Used by permission.

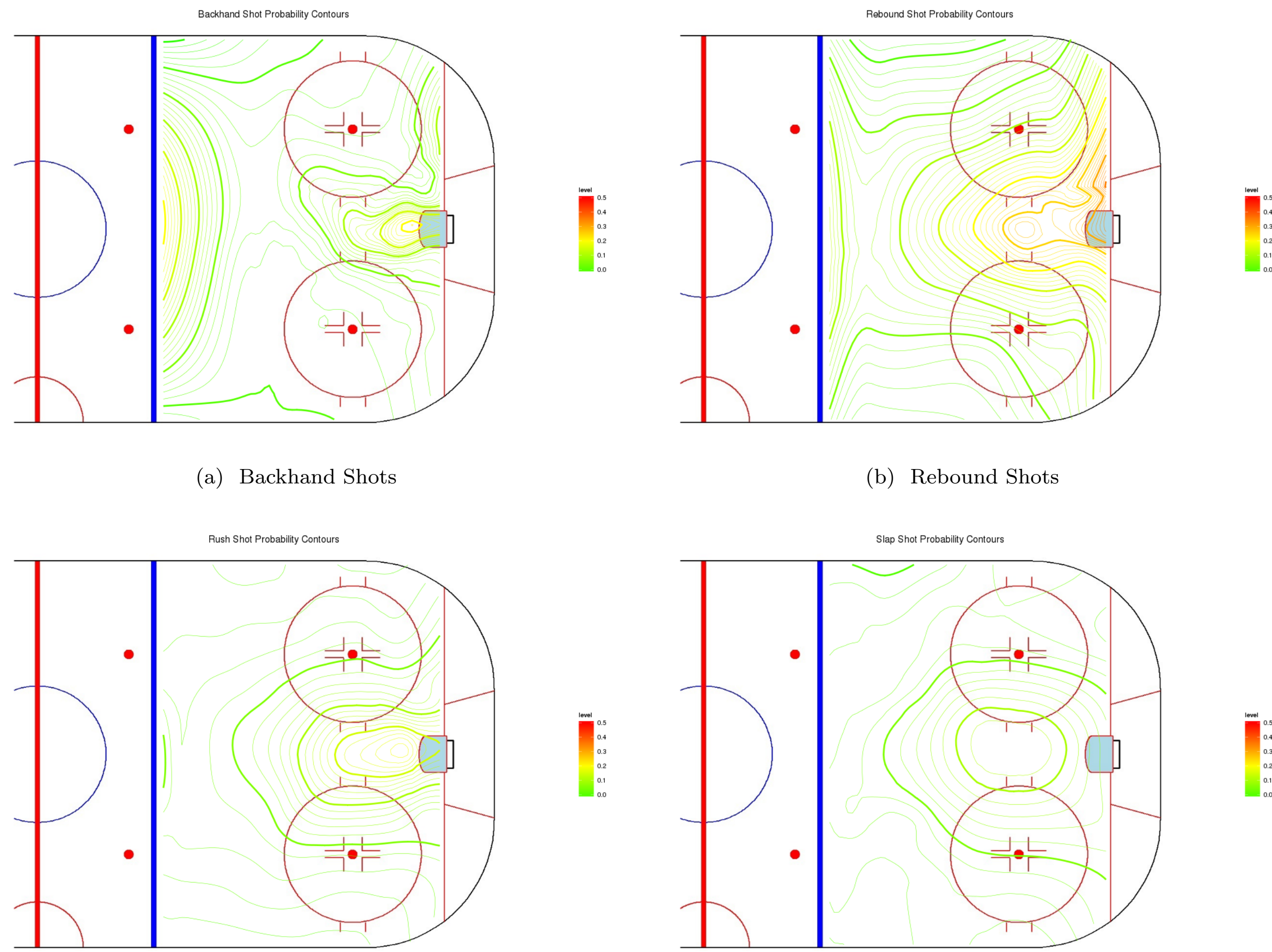
the home team (Byhome) and the score differential (Scorediff). We note here that Schuckers' original DIGR model did not include rebounds in that analysis.

Above we mentioned some of the issues with specific rinks and these are explored further in [Macdonald et al., 2012]. Specifically, the authors show that the adjusted save proportion is affected by a player's home rink. This is due to the aforementioned measurement errors in shot location by rink. The adjustment that they use which was previously suggested by others was to look at only shots faced when a goalie is on the road. In that way the biases of a particular rink are nearly averaged out. An alternative which has not been done would be to reweight all shots so that the weights for shots from a particular rink are equal across all rinks. For the analyses that follow we will use the location from which the rebound shot was taken not the x and y of the initial shot.

Although it was not the focus of their work, [Schuckers and Curro, 2013] created rink adjustments for shot location by using a discrete version of the probability integral transform to convert a shot coordinate, either x or y, from a particular rink to what the equivalent coordinate would have been by matching up the cumulative distributions from all rinks. Their adjustment, which we will use for plotting below, rounds to integer distances and makes the adjustments in a univariate rather than bivariate way. While this method is an improvement over most adjustments, a possible innovation would be to use an inherently bivariate adjustment and to allow for continuous scaling of the x and y (or r and θ) quantities. See, for example, [Pishchulin et al., 2012] for a discussion of how this might be accomplished. Using the adjusted shot locations from [Schuckers and Curro, 2013] we have plotted probability contours for each shot type. These can be found in Figures 1.3 and 1.4. The data we use below for shot locations has been adjusted following this methodology.

In 2014, the popular website `http://war-on-ice.com/` added an adjusted save proportion that was a simpler version of the above shot probability models just based upon shot location [Ventura, 2014]. Their adjusted save proportion is based upon a goalie's save proportion relative to the league average in each of three zones defined by the areas in Figure 1.5. This aSVP is the first type of adjusted save proportion given above, i.e. aSVP$_{\overline{G}}$. In that same year, a descendent of Ryder's and Kryzwicki's work on shot probability was introduced by [Johnson, 2014]. Kryzwicki had a term in his model for shots that followed a giveaway. Johnson modified and broadened this to include shots that followed shots by the other team as well as faceoffs, hits, takeaways, or giveaways in the neutral zone or the shooting team's defensive zone. Further, he provided some evidence that these shots have higher probabilities relative to other types of shots. Johnson referred to these types of shots as 'rush shots.'

Figures 1.3 and 1.4 show the contour plots for the probability of a goal from a given location for each of eight shot types. It is clear from those graphs that shot probability is a function of x and y locations as well as shot type s. In both of those figures, we are plotting the probability that shots from a given location result in goals. This is one minus the save proportion at those locations. The contours are given over the entire offensive zone but for most types of shots there are some locations where few shots of a given type were recorded. Reds and yellow contour lines correspond to higher probabilities than green or blue lines. Some of the contour plots for various shot types are clearly a function of few shots taken from certain locations, e.g. Wraparound shots and Tip-In shots. The volume of shots at a location are not accounted for in these graphs. Additionally, there are some combinations of shot type and shot location that are unlikely. For example, it is uncommon to take a backhand shot near the blue line or a slap shot within five feet of the goal.



(a) Backhand Shots

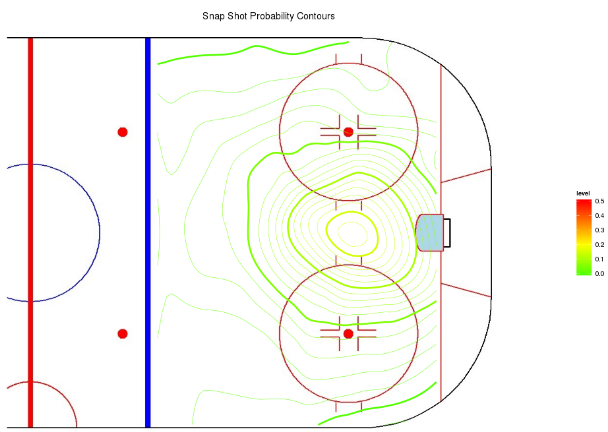
(b) Rebound Shots

(c) Rush Shots

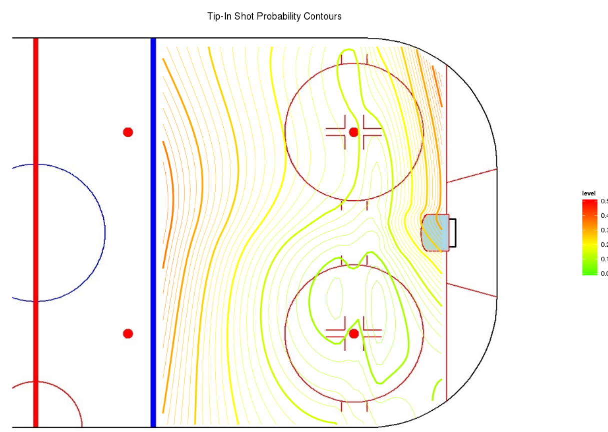
(d) Slap Shots

Figure 1.3: Using data from the NHL from the 2009 to 2013 seasons, these images are contour plots for shot probabilities of (a) Backhand shots, (b) Rebound shots, (c) Rush shots, and (d) Slap shots.

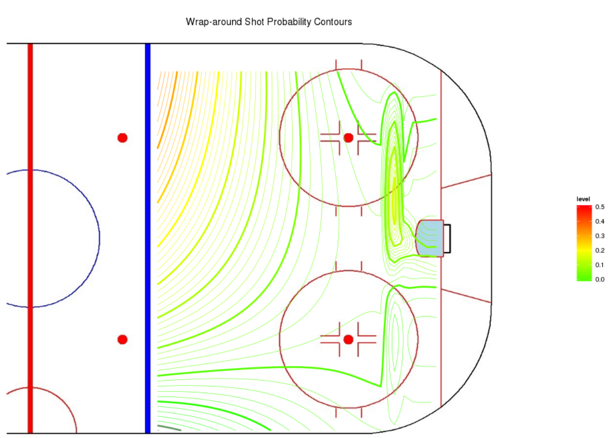
To assess the reliability of adjusted SVP methods, we will look at DIGR as representative of these approaches. The DIGR we use here is slightly different from the version in [Schuckers, 2011]. In this analysis we have added rebounds and rush shots to the shot types. We define a rebound as a shot that is the next event after another shot by the same team and that occurs within 3 seconds of the previous shot. For rush shots, we follow [Johnson, 2014] in calling a shot a rush shot if it is a shot taken within 10 seconds of a "shot attempt by



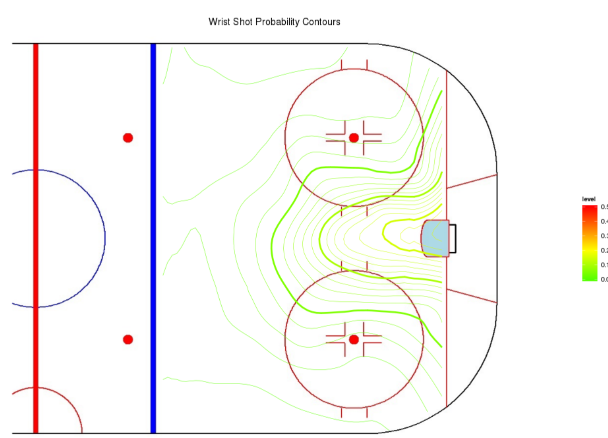
(a) Snap Shots



(b) Tip-In Shots

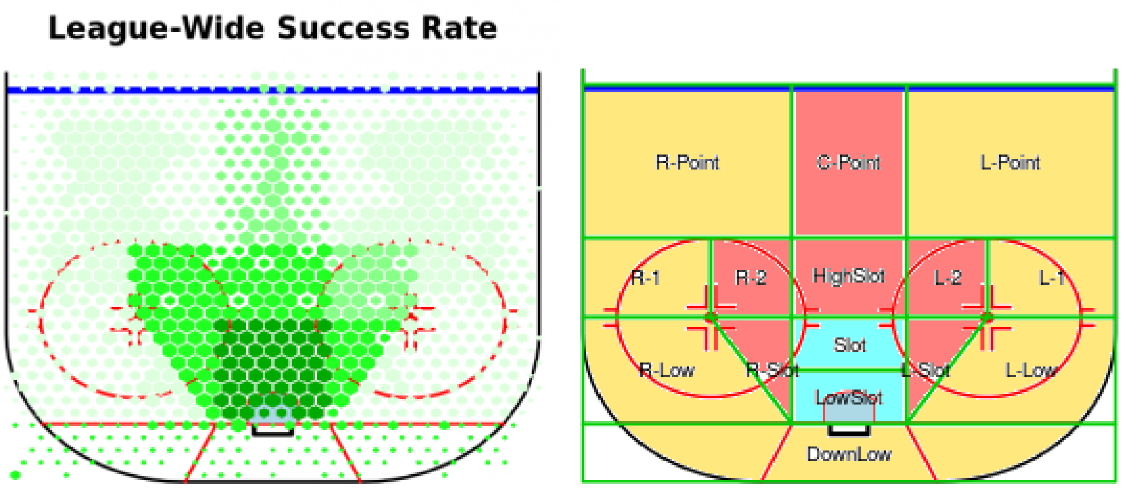


(c) Wraparound Shots



(d) Wrist Shots

Figure 1.4: Using data from the NHL from the 2009 to 2013 seasons, these images are contour plots for shot probabilities of (a) Snap shots, (b) Tip-In shots, (c) Wraparound shots, and (d) Wrist shots.



(a) Success Probabilities (b) Definition of Regions

Figure 1.5: (a) Mapping of shot success rate relative to unblocked shot attempts, and (b) represents the definition of the three (colored) regions of shot danger: high, medium and low from `war-on-ice.com`. Used by Permission.

the other team on the other net [or] . . . off a face off[sic] at the other end or in the neutral zone [or] . . . of a hit, giveaway or takeaway in the other end or the neutral zone.” Further, for consistency we will only analyze DIGR for even strength non-empty net shots.

Table 1.4: Intraseason correlation of DIGR for even and odd shots using $n^{\dagger} = 1000$

Season	More than 500 shots faced Correlation	J	More than 750 shots faced Correlation	J
2009/10	0.288	48	0.309	33
2010/11	0.644	45	0.751	36
2011/12	0.655	45	0.621	33
2012/13	0.251	24	0.178	9

We present results for a modified version of the original DIGR. First, we treated rebounds and rush shots as separate types of shots, s, as in Figures 1.3 and 1.4. Note that the probability contours found in those graphs are based upon the league averages for all shots not just those faced by a particular goalie. Second, we used the adjusted shot location from [Schuckers and Curro, 2013] as inputs to this version to account for the effect of different rinks. Finally, we altered the values for $n^{\dagger}$ to be 1000 which is much larger than the value

Table 1.5: Correlation between a goalie's DIGR in one year and their DIGR in subsequent years.

Seasons	More than 500 shots Correlation	J	More than 750 shots Correlation	J
2009/10 v 2010/11	0.237	34	0.393	22
2010/11 v 2011/12	0.647	37	0.665	28
2011/12 v 2012/13	-0.174	15	0.677	6

of $n^\dagger$ used in the original DIGR paper. This choice of $n^\dagger$ was driven by a trade-off in the ability to predict future aSVP while trying to minimize shrinkage. That is, we were trying to find the smallest value for $n^\dagger$ that yielded good correlations from one season to the next. We get similar results for those found here with $n^\dagger$ decreasing as n increases. Our results for the intra-season correlation between DIGR calculated on even shots and on odd shots faced by the same goalie can be found in Table 1.4 are stronger than those for just SVP that we saw in 1.2 when $n > 750$. However, there is one negative correlation (2011/12 v 2012/13 for goalies facing at least 500 shots) that stands out although for the same seasons with a larger number of shots faced, the correlation is positive and strongly so. These also mirror results for the aSVP approach in [Macdonald et al., 2012]. Those authors also found substantial improvement in reliability of aSVP (or DIGR) over SVP where reliability can be thought of as the correlation in a metric over time. These results are suggestive that adjusted SVP's may be moderately reliable for goalies who face a large number of shots. The top ten goalies as ranked by DIGR for the 2010-11 NHL regular season can be found in Table 1.6.

Table 1.6: Top 10 goalies ranked by DIGR for 2010-11 NHL Season using $n^\dagger = 1000$

Goalie	DIGR(aSVP)	SVP	n
T. Thomas	0.9314	0.9499	1416
S. Clemmensen	0.9308	0.9216	638
N. Backstrom	0.9305	0.9301	1202
J. Reimer	0.9301	0.9330	880
K. Lehtonen	0.9295	0.9290	1522
A. Niemi	0.9295	0.9310	1333
J. Quick	0.9295	0.9243	1255
S. Varlomov	0.9288	0.9370	571
M. Kiprusoff	0.9281	0.9172	1438
P. Rinne	0.9279	0.9364	1462

1.3 Rebounds

One area of goaltending that has received surprisingly little attention is a goalie's ability to control rebounds. [Myrland, 2009] wrote about the relationship between a teams rebound rate and their average shot probability against. Myrland argues that there is a relationship between the rate at which a teams gives up rebounds and the percentage of shots that are rebounds. In a similar vein, [Pettapiece, 2013] looked at the repeatability of a player's ratio of rebounds to unsaved shots. Pettapiece notes that "a goaltender who has the ability to cover up rebounds, or control them enough that a teammate can clear the zone, can significantly increase his value to his team." To assess this repeatability he compared the year 1 rebound rate to the year 2 rebound rate. Here we will define the rebound rate as the ratio of the number of rebounds shots allowed to the number of saved shots. As we did in the previous section we will use even strength, 5-on-5, non-empty net shot data from the 2009-10 through the 2012-13 regular seasons.

We begin by looking at the correlation between the rebound rate of even and odd shots. Table 1.7 shows the correlation between the rebound rate of even and odd shots. The counts in the table are the number of goalies involved in the correlations. As is perhaps not surprising there are higher correlations between those goalies that had more than 750 shots total in a season. In order for a goalie to be considered for this analysis, they had to have faced more than 500 shots in a season. Among goalies who faced at least 500 shots, the correlations here are moderate ranging from quite low, below 0.1, for 2009-10 to fairly strong, about 0.5, for the 2011-12 season. As always, the results lockout shortened 2012-13 season should be taken with a grain of salt since that season had only 48 regular season games relative to the usual 82. The correlations between even and odd shots in the same season are much stronger, all at least 0.3 for those goalies that faced more than 750 shots. These correlations suggest that the rebound rate is possibly something that would be considered repeatable. We will next look at the rebound rate from season to season.

The correlations for year to year rebound rate found in Table 1.7 are not great. For those that faced at least 500 shots the correlations have different signs which does not suggest consistency. For those that faced over 750 shots the correlations, at least, have the same

Table 1.7: Intraseason correlation of Rebound Rate for even and odd shots

Season	NHL Average Rebound Rate	More than 500 shots faced Correlation	J	More than 750 shots faced Correlation	J
2009/10	0.0823	0.066	46	0.327	33
2010/11	0.0825	0.329	48	0.395	32
2011/12	0.0842	0.512	45	0.642	35
2012/13	0.0838	0.268	24	0.336	9

sign, positive, but the values are diverse ranging from just under 0.1 to nearly 0.5. This suggests that past rebound rates are not strongly predictive of future rates though they are consistent within a given season.

We next investigated the relationship between rebound rates and SVP. Summaries of these relationships can be found in Table 1.8. It is clear from the correlations in that table that an increased rebound rate is associated with a lower current season save proportion. As might be expected a goalie's rebound rate is correlated with their save proportion although the relationship is not particularly strong. While not strong the relationship is consistently negative as would be expected. All of the goalies in the calculations given in Table 1.8 have faced at least 500 shots. Finally, we report the goalies with the lowest rebound rates for the 2010/11 season. This can be found in Table 1.10.

Table 1.8: Correlation between a goalie's rebound rate and their save proportion

Seasons	More than 500 shots Correlation	J	More than 750 shots Correlation	J
2009/10	-0.164	48	-0.114	33
2010/11	-0.440	46	-0.210	32
2011/12	-0.261	45	-0.252	35
2012/13	-0.066	24	-0.195	9

The numbers in parentheses correspond to the number of goaltenders included in the correlation calculations.

Table 1.9: Correlation between a goalie's rebound rate in one year and their rebound rate in subsequent years.

Seasons	More than 500 shots		More than 750 shots	
	Correlation	J	Correlation	J
2009/10 v 2010/11	-0.086	61	0.217	22
2010/11 v 2011/12	0.483	37	0.474	28
2011/12 v 2012/13	-0.261	20	0.067	8

Table 1.10: Top 10 goalies ranked by Rebound Rate for 2010-11 NHL Season among those goalies facing more than 500 shots

Goalie	Rebound Rate	SVP	n
J. Reimer	0.0572	0.9330	880
J. Quick	0.0586	0.9243	1255
S. Clemmensen	0.0663	0.9216	638
R. Luongo	0.0699	0.9333	1349
S. Mason	0.0714	0.9115	1152
P. Rinne	0.0723	0.9364	1462
C. Crawford	0.0726	0.9269	1218
O. Pavelec	0.0734	0.9295	1319
T. Vokoun	0.0735	0.9191	1347
B. Elliott	0.0748	0.9003	1144

1.4 Discussion and Future Directions

In this chapter we have looked at some of the ice hockey goalie metrics that have been used to evaluate goalies. It is clear that prediction of goalie performance in the National Hockey League is difficult. The primary traditional metric is the save proportion (SVP). Here we focused on even strength non-empty net five on five save proportion. Our analysis confirms what others have found which is that SVP is mildly correlated within a season using split data but not correlated with subsequent season's data. Adjusting for shot difficulty is one way that has been proposed to improve upon the prediction of subsequent season prediction. Methods for adjusting the SVP fall into one of two approaches either: adjusting for the performance of the goalie relative to an average goalie's shots faced or adjusting for the quality of shots relative to the performance of a particular goalie. Extending the work of [Schuckers, 2011], which is a method that falls into the latter category, we confirmed results found by [Macdonald et al., 2012] that adjusting the save proportion can improve upon the season to season reliability of SVP. Additionally we have added to the small literature on rebound rates among goalies to suggest that rebound rates are somewhat reliable within a season and between seasons. These relationships get stronger the more shots that a goalie has faced.

There are certainly other factors that impact how a goalie will perform in the future and one of these is their age. Certainly it is well known that athlete's performance decline as they get older. There has also been some work done on looking at player trajectories. This goes along with and could be used to improve the prediction of future goalie performance. Several authors including [Lundsford, 2012], [Tulsky, 2013] and [garik16, 2014] have added to our knowledge in this area. In general this work shows that goalies save proportion decreases with age and estimates the rates at which this occurs. Another metric that has received some interest is the Quality Starts metric, [Vollman, 2009]. A Quality Start is defined as one in which "the goaltender's save proportion is 0.913 or better, or at least 0.885 percent, but allowed fewer than 3 goals." A derivative of this is the percentage of a goalies starts that were quality.

One future direction for analysis would be to try to build a model, perhaps a generalized additive model, that might incorporate a variety of these metrics on a goalie to predict their save proportion or adjusted save proportion in future years. The variables in such a model might include the factors discussed in detail here: past save proportion, past rebound rate, adjusted save proportion as well as other variables such as age and performance in minor leagues such as the American Hockey League.

There are some ongoing efforts to improve the data in the NHL. As we (and many others) have noted, there are issues with data from the NHL's system for recording of data. See, for example, [Schuckers and Macdonald, 2014]. Some efforts have also been made to record NHL data by hand. In particular, [Boyle, 2013] has introduced a 'Shot Quality Project' that tried to track large numbers of NHL shots by hand to assess the impact of a variety of factors especially things like whether or not a goalie's view of the puck was obstructed. Having an obstructed view is something that likely has an impact on a goalie's ability to save a shot but is not in the current publicly available NHL data. Another project on which Boyle has collaborated is the so-called 'Royal Road' of [Valiquette, 2014]. The idea here is that an imaginary line extending from one goal mouth to the other exists — along the x axis of our notation — and passes that cross the royal road before a shot have an increased probability of being a goal. This project is in the early stages and the evidence they present is based upon goals rather than shots but this is an area that has potential for future work. Recently, the NHL has moved to add player tracking data for large amounts of additional automated information collection and for improved accuracy, [Masisak, 2015]. The amount and quality of these new data will no doubt provide new and important insights going forward which will impact our evaluation of goaltenders.

1.5 Addendum: Pulling the Goalie

Another area in which statistical analysis can have an impact is the altering of strategies or the development of new strategies. There are many strategies involved in any hockey game such as how to forecheck, how to carry the puck through the neutral zone, etc. The statistical evaluation of the impact of such strategies is difficult. As [Schuckers, 2014] wrote " with any strategic innovation, its success will lead to (being copied and eventually) its demise." The most common strategy advocated by hockey analysts is pulling the goalie early. The classic evaluation of pulling the goalie was done by [Beaudoin and Swartz, 2010]. They proposed an empirical Bayesian hierarchical Poisson model for goal scoring rates that suggests that being more aggressive about pulling the goalie can lead to an additional expected point or two over the course of a season. Recently, the idea of pulling the goalie has received additional attention in part because the aforementioned popular website, `war-on-ice.com`, created an interactive webpage that allowed users to compare the probability that the trailing team ties the game (if trailing by one or two goals) and the time at which there is the biggest increase in the probability of tying the game. These quantities are determined by comparing the scoring rates for the two teams at even strength and when one team has an extra attacker and their opponent is facing an extra attacker.

There is evidence that teams are potentially taking some of the information to heart.

[Davis and Lopez, 2015] recently looked at the times at which teams are pulling the goalie and they report that the length of time at the end of games when goalies are pulled for an extra skater has increased particularly for the 2013-14 and 2014-15 seasons. It seems that 'pulling the goalie' is becoming more acceptable to players, teams and coaches. As with any sort of strategy like pulling the goalie, the advantage gained will likely dissipate over time.

References

[Awad, 2009]Awad, T. (2009). Numbers on ice: A castle built on sand. `http://www.hockeyprospectus.com/puck/article.php?articleid=351`. Accessed September 19, 2013.

[Beaudoin and Swartz, 2010]Beaudoin, D. and Swartz, T. B. (2010). Strategies for pulling the goalie in hockey. *The American Statistician*, 64(3):197–204.

[Boyle, 2013]Boyle, C. (2013). Introducing the shot quality project. `http://www.sportsnet.ca/hockey/nhl/introducing-the-shot-quality-project/`. Accessed October 22,2015.

[Bruce McCurdy, 2010]Bruce McCurdy (2010). RTSS stats: Giveaways, takeaways, and throwaways. `http://www.coppernblue.com/2010/3/30/1396334/rtss-stats-giveaways-takeaways-and`. Accessed September 20, 2013.

[Davis and Lopez, 2015]Davis, N. and Lopez, M. (2015). NHL coaches are pulling goalies earlier than ever. `http://fivethirtyeight.com/features/nhl-coaches-are-pulling-goalies-earlier-than-ever/`. Accessed November 2, 2015.

[Desjardins, 2010a]Desjardins, G. (2010a). Clean up your act, Madison Square Garden. `http://www.arcticicehockey.com/2010/10/18/1756880/clean-up-your-act-madison-square-garden`. Accessed September 23, 2013.

[Desjardins, 2010b]Desjardins, G. (2010b). Fooled by randomness: Goaltender save percentage. `http://www.arcticicehockey.com/2010/4/20/1429278/fooled-by-randomness-goaltender`. Accessed October 30, 2015.

[Fischer, 2010]Fischer, J. (2010). Blocked shots, The New Jersey Devils' rink, and scorer bias - a follow up. `http://www.inlouwetrust.com/2010/7/8/1559914/`

blocked-shots-the-new-jersey. Accessed September 20, 2013.

[garik16, 2014]garik16 (2014). How well do goalies age? a look at a goalie aging curve. http://hockey-graphs.com/2014/03/21/how-well-do-goalies-age-a-look-at-a-goalie-aging-curve/. Accessed November 8, 2015.

[Johnson, 2014]Johnson, D. (2014). Introducing 'Rush' shots. http://hockeyanalysis.com/2014/07/09/introducing-rush-shots/. Accessed October 24, 2015.

[Kasan, 2008]Kasan, S. (2008). Off-ice officials are a fourth team at every game. http://www.nhl.com/ice/news.htm?id=388400. Accessed September 24,2013.

[Krzywicki, 2005]Krzywicki, K. (2005). Shot Quality Model:a logistic regression approach to assessing nhl shots on goal. http://hockeyanalytics.com/Research_files/Shot_Quality_Krzywicki.pdf. Accessed October 16, 2015.

[Krzywicki, 2009]Krzywicki, K. (2009). Removing Observer Bias from Shot Distance Shot Quality Model . http://hockeyanalytics.com/Research_files/SQ-DistAdj-RS0809-Krzywicki.pdf. Accessed October 17, 2015.

[Krzywicki, 2010]Krzywicki, K. (2010). NHL Shot Quality 2009-10 :a look at shot angles and rebounds. http://hockeyanalytics.com/2010/10/nhl-shot-quality-2010/. Accessed October 16, 2015.

[Lundsford, 2012]Lundsford, J. (2012). Aging goalies: What does the future hold for Luongo, Miller & Bryzgalov? http://www.coppernblue.com/2012/4/30/2987615/impact-age-goaltender-performance-save-percentage. Accessed November 8, 2015.

[Macdonald et al., 2012]Macdonald, B., Lennon, C., and Sturdivant, R. (2012). Evaluating nhl goalies, skaters, and teams using weighted shots. http://arxiv.org/abs/1205.1746. Accessed October 24, 2015.

[Masisak, 2015]Masisak, C. (2015). NHL, Sportvision test program to track players, puck. http://www.nhl.com/ice/news.htm?id=750201. Accessed October 31, 2013.

[Myrland, 2009]Myrland, P. (2009). The value of rebound control. http://brodeurisafraud.blogspot.com/2009/04/value-of-rebound-control.html. Accessed October 25, 2015.

[Paine, 2014]Paine, N. (2014). Why gretzky had it easy: The butterfly effect. http://fivethirtyeight.com/datalab/the-butterfly-effect/. Accessed January 13, 2016.

[Pettapiece, 2013]Pettapiece, R. (2013). Can goalies control the number of rebounds they allow? http://nhlnumbers.com/2013/7/15/can-goalies-control-the-number-of-rebounds-they-allow. Accessed October 25, 2015.

[Pishchulin et al., 2012]Pishchulin, L., Gass, T., Dreuw, P., and Ney, H. (2012). Image warping for face recognition: From local optimality towards global optimization. *Pattern Recognition*, 45(9):3131–3140.

[Ryder, 2004]Ryder, A. (2004). Goal Prevention 2004: a review of goaltending and team defense including a study of the quality of a hockey teams shots allowed. http://hockeyanalytics.com/Research_files/Goal_Prevention_2004.pdf. Accessed October 16, 2015.

[Schuckers, 2011]Schuckers, M. (2011). DIGR: A defense independent goalie rating based upon spatial probability maps. http://myslu.stlawu.edu/~msch/sports/Schuckers_DIGR_MIT_2011.pdf.

[Schuckers, 2014]Schuckers, M. (2014). Estimating the lower bound for the return on investment for hockey analytics. http://statsportsconsulting.com/2014/05/06/return-on-investment-for-hockey-analytics-1/. Accessed November 2, 2015.

[Schuckers and Curro, 2013]Schuckers, M. and Curro, J. (2013). Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events. Proceedings of the 2013 MIT Sloan Sports Analytics Conference http://www.statsportsconsulting.com/thor.

[Schuckers and Macdonald, 2014]Schuckers, M. and Macdonald, B. (2014). Accounting for rink effects in the National Hockey League's real time scoring system. http://arxiv.org/abs/1412.1035.

[Tulsky, 2013]Tulsky, E. (2013). Henrik Lundqvist's contract and estimating player value. http://www.sbnation.com/nhl/2013/12/10/5175204/henrik-lundqvist-contract-player-value-nhl. Accessed November 8, 2015.

[Valiquette, 2014]Valiquette, S. (2014). Red shots and green shots. http://www.msgnetworks.com/content/msgsite/en/videos/index.2.html?search=vally%20view. Accessed October 22,2015.

[Ventura, 2014]Ventura, S. (2014). Adjusted save percentage: Taking into account high, medium, and low probability shots. http://blog.war-on-ice.com/adjusted-save-percentage-taking-into-account-high-medium-and-low-probability-shots/. Accessed October 24, 2015.

[Vollman, 2009]Vollman, R. (2009). Howe and why: Quality starts. http://www.hockeyprospectus.com/puck/article.php?articleid=54. Accessed January 18, 2016.

[War-On-Ice.com, 2014]War-On-Ice.com (2014). Glossary. http://war-on-ice.com/glossary.html. Accessed January 4, 2016.

[Zona, 2011]Zona, D. (2011). Dear Oilers: Find new official scorers. http://www.coppernblue.com/2011/10/17/2495447/dear-oilers-find-new-official-scorers. Accessed September 20, 2013.