DIGR: A Defense Independent Rating of NHL Goaltenders using Spatially Smoothed Save Percentage Maps

Michael E. Schuckers
St. Lawrence University and Statistical Sports Consulting,
Canton, NY 13617 USA
Email: schuckers@stlawu.edu, michael.schuckers@statsportsconsulting.com

Abstract

Evaluation of NHL goalies is often done by comparing their save percentage. These save percentages depend highly upon the defense in front of each goalie and the difficulty of shots that each goalie faces. In this paper we introduce a new methodology for evaluating NHL goalies that does **not** depend upon the distribution of shots that any individual goalie faced. To achieve this new metric we create smoothed nonlinear spatial maps of goalie performance based upon the shots they did face and then evaluate these goalies on the league average distribution of shots. These maps show the probability of a goalie giving up a goal from across the playing surface. We derive a general mathematical framework for the evaluation of a goalie's save percentage. Using data from the 2009-10 NHL regular season, we apply this new methodology and calculate our new defense independent goalie rating (DIGR) for each goalie that face more than 600 shots. Results of this evaluation are given and possible extensions of the methodology are discussed.

1 Introduction

Currently, the most commonly used metric for evaluation of goalies is the save percentage. However, this metric is dependent upon the distribution of shots that each goalie has faced and, therefore, does not allow for the direct comparison of goalie performance. For example, during the 2009-10 regular season Johan Hedberg of the Atlanta Thrashers had a save percentage of 0.915 and Tim Thomas had a save percentage of 0.915. But, as [1] points out the shots faced by Hedberg were, on average, much more difficult than the shots faced by Thomas. To overcome this drawback, we propose in this paper the Defense Independent Goalie Rating (DIGR) which provides two innovations. First, the DIGR is based not on the shots that an individual goalie faces but on a single distribution of shots for comparison across goalies. Second, to derive the DIGR we develop a generalized smooth shot probability mapping of the shots faced by a given goalie. Thus, the DIGR is a metric that allows the direct comparison of goalies since under the DIGR their rating is based upon the exact same distribution of shots.

Several authors including [2], [3] and most recently [1] have proposed methodology for comparing how an average goalie would have done with the distribution of shots that each goalie faced. They do this by comparing a given goalie's performance against certain shot types to how the league average faired against those same shot types. For the DIGR, we develop a spatial approach for generating the league distribution of shots that takes into account shot location, opponents strength and type of shot. We apply our methodology to all of the shots taken in the 2009-10 NHL regular season. The structure of the remainder of this paper is as follows. Section 2 contains a discussion of the mathematical notation that we will be using, introduces the framework for the analyses and proposes our defense independent goalie rating. The next section, Section 3, describes our approach for deriving a nonparametric spatial mapping of shot probabilities. We then apply our methodology and calculate the DIGR, our goalie rankings, in Section 4. A discussion of these results and this new metric is given in Section 5.

2 Mathematical Notation and Framework

The primary metric for evaluating goalie performance is the save percentage. For the i^{th} goalie, let G_i be the save percentage. That percentage is calculated by taking the number of shots saved by the i^{th} goalie, E_i, divided by the total number of shots faced by the i^{th} goalie, N_i. We can then write G_i as

$$G_i = \frac{E_i}{N_i}.$$

We can generalize G_i by letting $X_i(s)$ be the number of shots of type *s* that goalie *i* saves out of $T_i(s)$ shots where $T_i(s)$ is the total number of shots of type *s* that goalie *i* faced. N_i is then the total number of shots that goalie *i* faced of all types. Then we can write the save percentage for goalie *i*, G_i, as

$$G_i = \frac{\sum_s X_i(s)}{N_i}$$

which we can then rewrite as the following

$$G_i = \sum_s \frac{X_i(s)}{T_i(s)} \frac{T_i(s)}{N_i}$$

which is the sum of a product of two terms $\delta_i(s) = X_i(s)/T_i(s)$ and $\Gamma_i(s) = T_i(s)/N_i$. The first term, $\delta_i(s)$, is the percent of saves that goalie *i* makes for a particular type of shot *s*. The second term, $\Gamma_i(s)$, is the percent of the total shots faced by goalie *i* that are of type *s*. We can think of $\delta_i(s)$ as the performance term as it reflects how a goalie performs on shots of type *s*. The $\Gamma_i(s)$'s define the distribution of shots that goalie *i* faced. Then, the i^{th} goalie's save percentage, G_i, is the average or expected save percentage against the distribution of shots defined by the $\Gamma_i(s)$'s. Decomposing a save percentage in this way allows us to generate metrics for goalie performance.

A shot quality adjusted (SQA) save percentage which has been proposed by several authors including [1] and can be written using the notation we have introduced above. This metric finds the average league save percentage for each shot type *s*, $\bar{\delta}(s)$, and substitutes it for the $\delta_i(s)$. Thus we get

$$G_i^\diamond = \sum_s \bar{\delta}(s)\Gamma_i(s).$$

which represents the save percentage that an average goalie would have had for the same distribution of shots. $G_i^\diamond$, the SQA for the i^{th} goalie, is then usually compared to G_i to give an idea of how a goalie compares to the league average for the shots that they faced. If $G_i > G_i^\diamond$ this suggests that goalie *i* outperformed the league average save percentage for the shots that goalie *i* faced, while $G_i < G_i^\diamond$ suggests that goalie *i* underperformed the league average save percentage for the shots that goalie *i* faced. What is unclear from this analysis is how to compare performance between goalies. For example, using results from [1], we find that both Kiprusoff (Calgary) and Niittymaki (Tampa Bay) had a difference, $G_i - G_i^\diamond$, of 0.007 so that they both outperformed the league average for the shots that they faced by the same amount. But Kiprusoff had a $G_i^\diamond$ of 0.913 while Niittymaki faced considerably harder shots on average and had a $G_i^{\Diamond}$ of 0.902. We can conjecture that Niittymaki is a better goalie since he outperformed the league having faced more difficult shots on average but the evidence provided by this measure makes it difficult to conclude this definitively. This is because of the dependence of this performance metric on the shots faced. Next we consider a metric that does not depend upon the shots faced by an individual goalie.

To address the dependence of the i^{th} goalie's performance on the distribution of shots that they faced, $\Gamma_i(s)$, we introduce a new metric based upon a reformation of G_i. Instead of replacing $\delta_i(s)$ with $\bar{\delta}(s)$, we use the league distribution of shots of type *s*, $\bar{\Gamma}(s)$, to replace $\Gamma_i(s)$. In this way we are standardizing the shots that each goalie faces since we are using a single distribution of shots of type *s*. Our new metric is then

$$G_i^* = \sum_s \delta_i(s)\bar{\Gamma}(s).$$

G_i^* is then the save percentage that the i^{th} goalie would have had had they faced the league distribution of shots of type *s*. The difficulty with this measure is to ensure that there is sufficient information about each shot type *s* to ensure that it is possible to estimate $\delta_i(s)$ for each goalie *i* and for each shot type *s*. Since this metric, G_i^*, does not depend on the distribution of shots that the i^{th} goalie faced but on an average distribution of shots that is the same across all goalies, we will call this metric the defense independent goalie rating (DIGR). This metric allows for a direct comparison of goalies since the shots being considered are for the same distribution of shots, $\bar{\Gamma}(s)$. Below we will use spatial maps to estimate $\delta_i(s)$ so that we can get an estimated G_i^* for each goalie during the 2009-10 NHL season.

3 Spatially Smoothed Goalie Performance

The basis for the metrics defined above is the shot types, s. In this section we further define our choice for *s*. For the remainder of this paper, we will define $\boldsymbol{s}$ as a vector of values. [1] uses a logistic regression model. We extend this work to allow for alternative forms for the relationship between the probability of a goal and $\boldsymbol{s}$. Since our goal is to create a spatial map of goalie performance, part of the vector $\boldsymbol{s}$ will be the *x*- and *y*-coordinates for each shot. Additionally, we will incorporate the type of shot (*w*=backhand, deflection, slap, snap, tip-in, wrap and wrist) as well as the strength of the team (*v*=shorthanded, even or power play) taking the shot. Thus, we have that $\boldsymbol{s}=(x, y, w, v)$. For this analysis we have eliminated empty net shots, penalty shots, and shootout shots.

It is not always possible to define $\delta_i(\mathbf{s})$ for each shot type and each goalie since not every goalie faces each shot combination $\boldsymbol{s}$. Because of this we will use a nonparametric spatial smoothing weighted estimation for $\delta_i(\mathbf{s})$. That is, for each goalie *i* for each type of shot *w* and strength *v*, we will create a smoothed spatial smoothed map of the save percentage at each location *x* and *y*. Specifically, for estimation of $\delta_i(x, y, w, v)$, we add all additional shots of type *w* to our estimation but with total weight of those shots equivalent to 30 shots which is approximately the average number of shots taken per game in the 2009-10 regular season. We use the loess (or LOWESS for locally weighted scatterplot smoothing) function in the statistical software R [5] assuming a locally linear polynomial fit (degree=1). We will refer to these spatially smoothed versions of $\delta_i(x, y, w, v)$ as $\tilde{\delta}_i(x, y, w, v)$. Figure 1 shows several example mappings of $\tilde{\delta}_i(x, y, w, v)$ for different goalies. In that figure red indicates a higher probability of a goal from that location and blue indicates a lower probability probabilities of a Since there are seven different shot types and three different strengths, we derive 21 different mappings for each goalie.

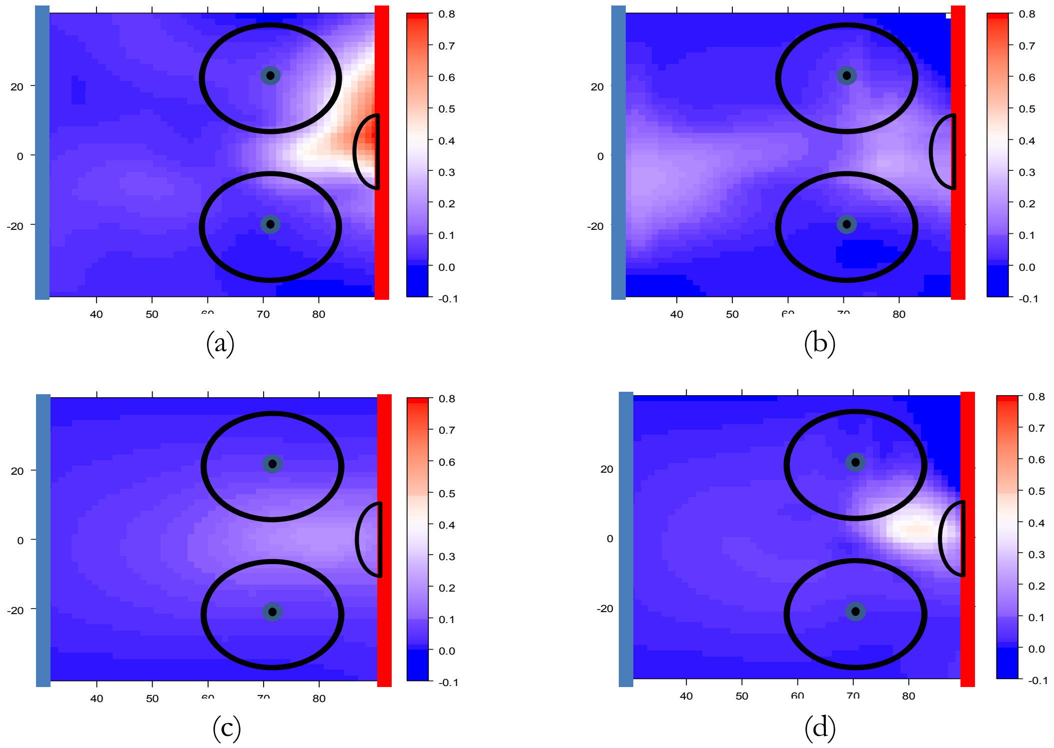

Figure 1: Spatial mappings, $\tilde{\delta}_i(x, y, w, v)$'s, for selected goalies, selected shot types and selected strengths. Figure 1(a) is for Martin Brodeur of the New Jersey Devils for slap shots faced at even strength; (b) is for Tim Thomas of the Boston Bruins for wrist shots faced during opponents power play; (c) is for Marc-André Fleury of the Pittsburgh Penguins for snap shots at even strength; (d) is for the Phoenix Coyotes' Ilya Bryzgalov for slap shots taken during opponent's power plays.

4 Application and Results

The data that we will use for this analysis is every NHL shot from the 2009-10 regular season. The data was downloaded from ESPN.com's GameCast of each regular season game and processed into an appropriate format[1]. As mentioned above we excluded several types of shots (penalty shots, shootout shots and empty net shots) from this analysis. All other shots were included in the calculation of $\overline{\Gamma}(\mathbf{s})$ and were mapped to a single offensive zone. There were n=74300 shots in our dataset for which there was enough information to analyze a given shot. Following [1] and [2], we adjusted shots taken at the New York Rangers home ice because of a observer bias in both the x and y positions for those shots. Here we used a probability integral transform to adjust those shots

[1] Special thanks to Dan Downs for his assistance in downloading and processing the data and to Ken Krzywicki for information about ESPN's GameCast.

locations. We evaluated all 49 of the NHL goalies that faced more than 600 shots or approximately 20 games worth of shots. Table 1 contains an ordered listing of the goalies with the 40 highest DIGR's as well as their raw save percentage from the NHL.

Table 1: DIGR for NHL 2009-10 Regular Season

Rank	Player (Team)	DIGR Rating (G_i^*)	Save Percentage (G_i)
1	Ryan Miller (BUF)	0.9285	0.9285
2	Ty Conklin (STL)	0.9280	0.9215
3	Jaroslav Halak (MTL)	0.9269	0.9242
4	Jonas Hiller (ANA)	0.9243	0.9183
5	Henrik Lundqvist (NYR)	0.9237	0.9208
6	Evgeni Nabokov (SJS)	0.9227	0.9216
7	Ilya Bryzgalov (PHX)	0.9226	0.9204
8	Tuukka Rask (BOS)	0.9218	0.9312
9	Antti Niemi (CHI)	0.9215	0.9124
10	Tomas Vokoun (FLA)	0.9191	0.9246
11	Johan Hedberg (ATL)	0.9190	0.9151
12	Roberto Luongo (VAN)	0.9186	0.9128
13	Jose Theodore (WSH)	0.9185	0.9105
14	Cam Ward (CAR)	0.9185	0.9155
15	Dwayne Roloson (NYI)	0.9182	0.9068
16	Miikka Kiprusoff (CGY)	0.9178	0.9199
17	Semyon Varlamov (WSH)	0.9163	0.9095
18	Ondrej Pavelec (ATL)	0.9159	0.9061
19	Chris Mason (STL)	0.9158	0.9129
20	Manny Legace (CAR)	0.9155	0.9129
21	Craig Anderson (COL)	0.9136	0.9167
22	Scott Clemmensen (FLA)	0.9135	0.9117
23	Ray Emery (PHI)	0.9132	0.9055
24	Antero Niittymaki (TBL)	0.9132	0.9085
25	Mike Smith (TBL)	0.9130	0.8996
26	Jonathan Quick (LAK)	0.9129	0.9066
27	Mathieu Garon (CBJ)	0.9128	0.9033
28	Pekka Rinne (NSH)	0.9125	0.9111
29	Martin Brodeur (NJD)	0.9122	0.9162
30	Jimmy Howard (DET)	0.9117	0.9237
31	Dan Ellis (NSH)	0.9113	0.9092
32	Jean-Sebastien Giguere (TOR, ANA)	0.9110	0.9069
33	Carey Price (MTL)	0.9108	0.9124
34	Marty Turco (DAL)	0.9101	0.9128
35	Jonas Gustavsson (TOR)	0.9087	0.9023
36	Martin Biron (NYI)	0.9082	0.8964
37	Michael Leighton (PHI, CAR)	0.9076	0.9055
38	Brian Elliot (OTT)	0.9073	0.9087
39	Marc-Andre Fleury (PIT)	0.9069	0.9052
40	Tim Thomas (BOS)	0.9064	0.9148

We first note that there is a wide range of values for the DIGR in our data. The goalie with the highest DIGR was Ryan Miller of the Buffalo Sabres with G_i^*=0.9285. Miller would be predicted to have a save percentage of 92.85%, if he faced the distribution of shots taken by the entire NHL. Note that league average save percentage was 91.15%. The top five goalies were Ryan Miller (BUF), Ty Conklin (STL), Jaroslav Halak (MTL), Jonas Hiller (ANA), and Henrik Lundqvist (NYR). Some of the goalies in Table 1 had lower save percentages G_i's than their performance warranted suggesting that their save percentages were likely hurt by the difficulty of the shots they faced. These goalies include Mike Smith (TBL) , Martin Biron (NYI) and Dwayne Roloson (NYI), in particular. The goalie whose save percentage, G_i, most benefitted from the distribution of shots that they faced as Jimmy Howard. The distribution of the G_i^*'s is approximately Normal for the 49 goalies with at least 600 shots faced. Table 2 of the Appendix contains the remaining DIGR results for the nine goalies not given here.

5 Discussion

In this paper we have presented two innovations. The first of these is the defense independent goalie rating (DIGR) and the second is a methodology for mapping shot probabilities. The DIGR, G_i^*, is a new goalie performance metric that allows for comparison across goalies by evaluating them on the same distribution of shots. For the DIGR, we have chosen to evaluate predicted performance based upon the league average distribution of shots. From this evaluation we found that Ryan Miller, Ty Conklin and Jaroslav Halak were the best performing NHL goalies for the 2009-10 regular season. This metric should be a useful tool for valuing and evaluation of NHL goalies in future seasons. Our generalized non-linear spatially smooth shot probability mappings allow for interactions between the effects of shot type, opponents strength and the location from which the shot was taken. This extends the previous work of [1]. The mathematical framework that we have developed here can be extended to allow for other metrics based upon specific versions of $\Gamma(\mathbf{s})$. For example, it is possible to predict how Marty Turco, our 34th rated goalie, would have been predicted to perform if he had faced the shots taken against Antti Niemi, our 9th rated goalie. (Note that Turco replaced Niemi as the primary goaltender for Chicago for the 2010-11 season.) This could be written as

$$G_{ij} = \sum_s \bar{\delta}_i(\boldsymbol{s})\Gamma_j(\boldsymbol{s}).$$

We have previously considered a simplified version of this based solely on shot location in [4]. There are other refinements that are possible within this framework. It is possible to incorporate other information into our shot probability model $\delta_i(\mathbf{s})$. [6] has suggested that score differential has a relationship with the probability of a given shot being a goal. [1] has proposed using whether or not a shot is a rebound. Additionally, looking at the shot target, where on the goalie the shot was aimed, could be worth incorporating into our shot model but that variable was not available for this analysis. Finally, we note some limitations of this methodology. First, we are averaging each goalie's performance over the course of the season. Goalie performance is likely to fluctuate within and between seasons. Consider the performance of Tim Thomas of the Boston Bruins, the 40th rated goalie, during the 2010-11 regular season. His current save percentage is 0.940 (as of 2/14/11) which is first in the NHL. The DIGR only evaluates performance not potential. Second, we are predicting performance, G_i^*, rather than observing performance and, consequently, there are standard errors associated with our predictions. Though we have not included those standard errors here, it is important to keep those in mind when comparing DIGR performance.

6 References

[1] Ken Krzywicki, "NHL Shot Quality 2009-10." Hockey Analytics.com, 2010. Web. 10 November 2010 http://hockeyanalytics.com/Research_files/SQ-RS0910-Krzywicki.pdf.

[2] Alan Ryder, "Product Recall Notice for 'Shot Quality'." Hockey Analytics.com, 2007. Web. 10 November 2010 http://hockeyanalytics.com/Research_files/Product_Recall_for_Shot_Quality.pdf.

[3] Ken Krzywicki, "Removing Observer Bias from Shot Distance." Hockey Analytics.com, 2009. Web. 10 November 2010 http://hockeyanalytics.com/Research_files/SQ-DistAdj-RS0809-Krzywicki.pdf.

[4] Michael Schuckers, "Turco for Niemi (4 goals for $1.45 Million)." Statistical Sports Consulting.com, 2010. Web. 9 August 2010 http://blog.statsportsconsulting.com/2010/08/09/turco-for-niemi-4-goals-for-145-million.aspx

[5] R Development Core Team, R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2010.

[6] Tom Awad, "Numbers on Ice: Does shot quality exist?" Puck Prospecus, 2010. Web. 10 November 2010 http://www.puckprospectus.com/article.php?articleid=540

Appendix

Table 2: Remaining DIGR results for the 2009-10 NHL regular season

Rank	Player (Team)	DIGR Rating (G_i^*)	Save Percentage (G_i)
41	Cristobal Huet (CHI)	0.9045	0.8947
42	Jeff Deslauriers (EDMj)	0.9037	0.9006
43	Steve Mason (CBJ)	0.9033	0.9014
44	Brian Boucher (PHI)	0.9023	0.8995
45	Josh Harding (MIN)	0.9016	0.9046
46	Alex Auld (DAL, NYR)	0.9011	0.8951
47	Pascal Leclaire (OTT)	0.9006	0.8869
48	Niklas Backstrom (MIN)	0.8988	0.9032
49	Vesa Toskala (TOR, CGY)	0.8969	0.8797