Competent Goalie Evaluation

Micah Blake McCurdy
hockeyviz.com

Seattle, Washington
Seattle Hockey Analytics Conference
March 9, 2019
Introduction

- Shooters shoot the puck.
- Goalies try to stop it.
Introduction

- Shooters shoot the puck.
- Goalies try to stop it.
  - How
Motivation

- Not all shots are equally likely to result in goals.
Motivation

- Not all shots are equally likely to result in goals.
  - Easy, everybody agrees
Motivation

- Not all shots are equally likely to result in goals.
  - Easy, everybody agrees
- How likely is it that a given shot will result in a goal?
Motivation

- Not all shots are equally likely to result in goals.
  - Easy, everybody agrees
- How likely is it that a given shot will result in a goal?
  - Hard, lots of yelling
First, the bad news

Input  Unblocked 5v5 and 5v4 shots, 2016-2018 regular seasons.
Testing  Unblocked 5v5 and 5v4 shots, 2018-2019 regular season so far.

What performs better: assuming that previous save percentages predict future goal likelihoods, or assuming that all goalies are identical?
Log-loss

A thing happens which you said would happen with probability $p$.

How right or wrong are you?
Log-loss

A thing happens which you said would happen with probability $p$.

- You score $-\log p$
Log-loss

A thing happens which you said would happen with probability $p$.

- You score $-\log p$
- The only fair way to evaluate predictions of probability.
Log-loss

A thing happens which you said would happen with probability $p$.

- You score $-\log p$
- The only fair way to evaluate predictions of probability.
- Lower is better, “less wrongness”
  - 0 is “not at all wrong”
Log-loss

A thing happens which you said would happen with probability $p$.

- You score $-\log p$
- The only fair way to evaluate predictions of probability.
- Lower is better, “less wrongness”
  - 0 is “not at all wrong”
- Unblocked goal % in 2016-2018, all situations: 4.6%.
Log-loss

A thing happens which you said would happen with probability $p$.

- You score $-\log p$
- The only fair way to evaluate predictions of probability.
- Lower is better, “less wrongness”
  - 0 is “not at all wrong”
- Unblocked goal % in 2016-2018, all situations: 4.6%.
  - Log-loss, 2018-2019 shots through Valentine’s: 0.1949
Log-loss

A thing happens which you said would happen with probability $p$.

- You score $\log p$.
- The only fair way to evaluate predictions of probability.
- Lower is better, “less wrongness”
  - 0 is “not at all wrong”
- Unblocked goal % in 2016-2018, all situations: 4.6%.
  - Log-loss, 2018-2019 shots through Valentine’s: 0.1949
- Individual save percentage for all goalies with at least 100 shots faced and 4.6% otherwise.
Log-loss

A thing happens which you said would happen with probability $p$.

▶ You score $-\log p$

▶ The only fair way to evaluate predictions of probability.
▶ Lower is better, “less wrongness”
   ▶ 0 is “not at all wrong”

▶ Unblocked goal % in 2016-2018, all situations: 4.6%.
   ▶ Log-loss, 2018-2019 shots through Valentine’s: 0.1949

▶ Individual save percentage for all goalies with at least 100 shots faced and 4.6% otherwise.
   ▶ Log-loss, 2018-2019 shots through Valentine’s: 0.1950
Log-loss

A thing happens which you said would happen with probability $p$.

- You score $-\log p$
- The only fair way to evaluate predictions of probability.
- Lower is better, “less wrongness”
  - 0 is “not at all wrong”
- Unblocked goal % in 2016-2018, all situations: 4.6%.
  - Log-loss, 2018-2019 shots through Valentine’s: 0.1949
- Individual save percentage for all goalies with at least 100 shots faced and 4.6% otherwise.
  - Log-loss, 2018-2019 shots through Valentine’s: 0.1950

Same story using only shots on goal, restricting to 5v5, both of those things, no matter.
Save percentage is completely useless.
Hear the Good News

We get something worth quoting by accounting for shot difficulty.
Hear the Good News

We get something worth quoting by accounting for shot difficulty.

- All goalies average: 0.1837
Hear the Good News

We get something worth quoting by accounting for shot difficulty.

- All goalies average: 0.1837
- About 6% better than guessing.
We get something worth quoting by accounting for shot difficulty.

- All goalies average: 0.1837
  - About 6% better than guessing.
- With goalie and shooter ability estimates: 0.1830
  - About 1% better still.
How
What affects goal likelihood?

- Distance
- Geometry ("Impossible angles")
- Vision of all concerned (screens)
- Sneaky passing and skating
- **Shooter quality**
- **Goalie quality**
- “Special plays”
  - Breakaways
  - 2-on-1s
  - Rebounds
What affects goal likelihood?

- Distance
- Geometry ("Impossible angles")
- Vision of all concerned (screens)
- Sneaky passing and skating
- **Shooter quality**
- **Goalie quality**
- “Special plays”
  - Breakaways
  - 2-on-1s
  - Rebounds
What affects goal likelihood?

- Distance
- Geometry ("Impossible angles")
- Vision of all concerned (screens)
- Sneaky passing and skating
  - 5v4 as partial proxy
- **Shooter quality**
- **Goalie quality**
- "Special plays"
  - Breakaways
  - 2-on-1s
    - "Rushes" as proxy
- Rebounds
  - Or at least rebound shots
Balance

How do we balance all these things to quantify how likely shots are to become goals?
How do we balance all these things to quantify how likely shots are to become goals?

- Attach electrodes to the skulls of fans to detect excitement
Balance

How do we balance all these things to quantify how likely shots are to become goals?

- Attach electrodes to the skulls of fans to detect excitement
- ♡ Regression ♡
How do we balance all these things to quantify how likely shots are to become goals?

- Attach electrodes to the skulls of fans to detect excitement
- Regression

Not just any regression: logistic ridge regression
If you pile up enough favourable things, a goal becomes likely.
Structure

The graph represents the "Logistic" function, defined as:

\[ x \mapsto \frac{e^x}{1 + e^x} \]

The x-axis represents the Goodness of Circumstances (in meaningless units), while the y-axis shows the Probability.
More $x$ means more “favourable things”. What is $x$, really?

$$p(x) = \frac{e^x}{1 + e^x}$$
Probability, Odds, and Log-Odds

More $x$ means more “favourable things”. What is $x$, really?

$$p(x) = \frac{e^x}{1 + e^x}$$

$$1 - p(x) = \frac{1}{1 + e^x}$$
Probability, Odds, and Log-Odds

More $x$ means more “favourable things”. What is $x$, really?

$$p(x) = \frac{e^x}{1 + e^x} \quad \quad 1 - p(x) = \frac{1}{1 + e^x}$$

Odds of a goal = \frac{\text{Probability of a goal}}{\text{Probability of no goal}} = \frac{p(x)}{1 - p(x)} = e^x
More $x$ means more “favourable things”. What is $x$, really?

$$p(x) = \frac{e^x}{1 + e^x} \quad \quad 1 - p(x) = \frac{1}{1 + e^x}$$

Odds of a goal = \frac{Probability of a goal}{Probability of no goal} = \frac{p(x)}{1 - p(x)} = e^x$

So $x$ for a shot is the logarithm of the odds of that shot becoming a goal: “log-odds”.

Probability, Odds, and Log-Odds
Logistic Regression

Trying to fit a regression:

- Targetting probabilities: Constrained to $[0, 1]$, awkward
- Targetting log-odds: can be any real number, smooth.
Included factors

- **Shooter ability (fixed)**
- **Goalie ability (fixed)**
- Distance to the net, normalized to 89 feet (the blue line).
  - Closer than 10 feet counts as 10 feet
- Visible net, normalized to 6 feet.
- **Shot type**
  - Slap / Wrist-Snap / Tip-Deflection / Wraparound / Backhand
- Rebound shot indicator (within 3s)
- Rush shot indicator (within 4s)
- Power-play indicator (5v4)
- Leading / Trailing
- Constant
Fitting

- Every set of values $\beta$ for each parameter gives a prediction for each shot $s$ becoming a goal, $p(s, \beta)$
Fitting

- Every set of values $\beta$ for each parameter gives a prediction for each shot $s$ becoming a goal, $p(s, \beta)$
- Multiplying all these probabilities gives the likelihood associated to a particular value of $\beta$. 
Fitting

- Every set of values $\beta$ for each parameter gives a prediction for each shot $s$ becoming a goal, $p(s, \beta)$
- Multiplying all these probabilities gives the likelihood associated to a particular value of $\beta$.
- The best fit is the value of $\beta$ that maximizes the likelihood.
Fitting

Every set of values $\beta$ for each parameter gives a prediction for each shot $s$ becoming a goal, $p(s, \beta)$.

Multiplying all these probabilities gives the likelihood associated to a particular value of $\beta$.

The best fit is the value of $\beta$ that maximizes the likelihood.

Or, the logarithm of the likelihood, because products are fidgety little things.
Fitting

- Every set of values $\beta$ for each parameter gives a prediction for each shot $s$ becoming a goal, $p(s, \beta)$
- Multiplying all these probabilities gives the likelihood associated to a particular value of $\beta$.
- The best fit is the value of $\beta$ that maximizes the likelihood
- Or, the logarithm of the likelihood, because products are fidgety little things.

$$L = \sum_{\text{goals}} \log p(s, \beta) + \sum_{\text{not goals}} \log(1 - p(s, \beta))$$
Overfitting and its discontents

- Fitting a model with ordinary least squares fitting is prone to matching the data too closely, being fooled by randomness.
Overfitting and its discontents

- Fitting a model with ordinary least squares fitting is prone to matching the data too closely, being fooled by randomness.
- In particular, ability estimates for players with small amounts of data can be outlandish.
Overfitting and its discontents

- Fitting a model with ordinary least squares fitting is prone to matching the data too closely, being fooled by randomness.
- In particular, ability estimates for players with small amounts of data can be outlandish.
- Disconnect arises because we know some things about these players that the model does not.
Fitting a model with ordinary least squares fitting is prone to matching the data too closely, being fooled by randomness.

In particular, ability estimates for players with small amounts of data can be outlandish.

Disconnect arises because we know some things about these players that the model does not.

Solution:
- Tell it.
Zero-biased regression

One thing we know about NHL players, prior to looking at any data about their play, is that they all play in the NHL.
Zero-biased regression

▶ One thing we know about NHL players, prior to looking at any data about their play, is that they all play in the NHL.

▶ Add a mild assumption that deviation from average is “bad”.
Zero-biased regression

- One thing we know about NHL players, prior to looking at any data about their play, is that they all play in the NHL.
- Add a mild assumption that deviation from average is “bad”.
- Instead of maximizing the likelihood $\mathcal{L}$, maximize

$$\mathcal{L} - \frac{1}{2} \beta^T \beta$$
Zero-biased regression

- One thing we know about NHL players, prior to looking at any data about their play, is that they all play in the NHL.
- Add a mild assumption that deviation from average is “bad”.
- Instead of maximizing the likelihood $L$, maximize

$$L - \frac{1}{2} \beta^T \Lambda \beta$$

where $\Lambda$ is a matrix that encodes how sure we are, before we start to look at data, that a given parameter “should” be close to zero.
Zero-biased regression

▶ One thing we know about NHL players, prior to looking at any data about their play, is that they all play in the NHL.
▶ Add a mild assumption that deviation from average is “bad”.
▶ Instead of maximizing the likelihood $\mathcal{L}$, maximize

\[
\mathcal{L} - \frac{1}{2} \beta^T \Lambda \beta
\]

where $\Lambda$ is a matrix that encodes how sure we are, before we start to look at data, that a given parameter “should” be close to zero.
▶ Starting from this prior, every shot updates our assumptions about the values of the factors involved, including the players.
Zero-biased regression

- For players, start with a prior of zero impact with a standard deviation of 0.1
Zero-biased regression

- For players, start with a prior of zero impact with a standard deviation of 0.1
  - An impact on goal odds of around 10-15%, for good or for ill
Zero-biased regression

- For players, start with a prior of zero impact with a standard deviation of 0.1
  - An impact on goal odds of around 10-15%, for good or for ill
- For other factors, start with a prior of zero impact with a standard deviation of 1,000,000
Zero-biased regression

- For players, start with a prior of zero impact with a standard deviation of 0.1
  - An impact on goal odds of around 10-15%, for good or for ill
- For other factors, start with a prior of zero impact with a standard deviation of 1,000,000
  - We have literally no idea and we’d like to be told, please.
Zero-biased regression

- For players, start with a prior of zero impact with a standard deviation of 0.1
  - An impact on goal odds of around 10-15%, for good or for ill
- For other factors, start with a prior of zero impact with a standard deviation of 1,000,000
  - We have literally no idea and we’d like to be told, please.
- We can choose how much zero-bias we mix in, and where.
Zero-biased regression

- For players, start with a prior of zero impact with a standard deviation of 0.1
  - An impact on goal odds of around 10-15%, for good or for ill
- For other factors, start with a prior of zero impact with a standard deviation of 1,000,000
  - We have literally no idea and we’d like to be told, please.
- We can choose how much zero-bias we mix in, and where.
- For every parameter, we get a “posterior” distribution, with its own mean and standard deviation, reflecting our new certainty about the impact of each factor.
Results!!

Input data: Every unblocked 5v5 and 5v4 shot in the 730 days up to and including January 1st, 2019.
Geometry

Unblocked, 5v5, non-rush, non-rebound wrist shots:
Historical

All unblocked shots, 5v5:
Shot Types

Odds Ratio of a Goal compared to a wrist shot from the same spot
Score Effects

Odds Ratio of a Goal compared to tied

Trailing

Leading
Score Effects

Leading teams take better shots, but trailing teams take more. Net effect favours trailing teams.
Other Factors

![Diagram showing odds ratio of a goal compared to baseline]

- Rush
- 5v4
- Rebound

Odds Ratio of a Goal compared to baseline

- 1.0
- 1.5
- 2.0
- 2.5
- 3.0
Shooter Distribution

Odds Ratio of a Goal
Shooter Distribution

≥ 200 shots
Shooter Distribution

Odds Ratio of a Goal

≥ 300 shots
Shooter Distribution

≥ 400 shots
Shooter Distribution

≥ 500 shots

Odds Ratio of a Goal

0.90 0.95 1.00 1.05 1.10 1.15 1.20 1.25 1.30
Shooter Distribution

Odds Ratio of a Goal

≥ 600 shots
Shooter Distribution

Odds Ratio of a Goal

≥ 700 shots
Shooter Distribution
Shooter Distribution

Odds Ratio of a Goal

≥ 900 shots
## Strongest Recent Shooters

<table>
<thead>
<tr>
<th>Player</th>
<th>Team</th>
<th>Impact on goal odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patrik Laine</td>
<td>WPG</td>
<td>+30%</td>
</tr>
<tr>
<td>Auston Matthews</td>
<td>TOR</td>
<td>+22%</td>
</tr>
<tr>
<td>Kyle Palmieri</td>
<td>N.J</td>
<td>+21%</td>
</tr>
<tr>
<td>Mikko Rantanen</td>
<td>COL</td>
<td>+21%</td>
</tr>
<tr>
<td>Nikita Kucherov</td>
<td>T.B</td>
<td>+20%</td>
</tr>
</tbody>
</table>
Auston Matthews
Alexander Ovechkin

![Graph showing the odds ratio of a goal before and after a certain event.](image)
## Weakest Recent Shooters

<table>
<thead>
<tr>
<th>Player</th>
<th>Team</th>
<th>Impact on goal odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brent Burns</td>
<td>S.J</td>
<td>-14%</td>
</tr>
<tr>
<td>Duncan Keith</td>
<td>CHI</td>
<td>-13%</td>
</tr>
<tr>
<td>Troy Stecher</td>
<td>VAN</td>
<td>-11%</td>
</tr>
<tr>
<td>Erik Karlsson</td>
<td>OTT / S.J</td>
<td>-11%</td>
</tr>
<tr>
<td>Oskar Klefbom</td>
<td>EDM</td>
<td>-11%</td>
</tr>
</tbody>
</table>
Weakest Recent Shooting Forwards

<table>
<thead>
<tr>
<th>Player</th>
<th>Team</th>
<th>Impact on goal odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kevin Labanc</td>
<td>S.J</td>
<td>-10%</td>
</tr>
<tr>
<td>Dmitri Jaskin</td>
<td>STL / WSH</td>
<td>-9%</td>
</tr>
<tr>
<td>Brock McGinn</td>
<td>CAR</td>
<td>-8%</td>
</tr>
<tr>
<td>Mikko Koivu</td>
<td>MIN</td>
<td>-7%</td>
</tr>
<tr>
<td>Carl Hagelin</td>
<td>PIT / L.A</td>
<td>-7%</td>
</tr>
</tbody>
</table>
Goalies
Goalies

Odds Ratio of a Goal

\[ \geq 2500 \text{ shots} \]
Goalies

Odds Ratio of a Goal

\( \geq 3000 \) shots
Goalies

Odds Ratio of a Goal

≥ 3500 shots
Goalies

Odds Ratio of a Goal

≥ 5000 shots
## Strongest Recent Goaltenders

<table>
<thead>
<tr>
<th>Player</th>
<th>Team</th>
<th>Impact on goal odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>John Gibson</td>
<td>ANA</td>
<td>-22%</td>
</tr>
<tr>
<td>Carter Hutton</td>
<td>STL / BUF</td>
<td>-12%</td>
</tr>
<tr>
<td>Antti Raanta</td>
<td>ARI</td>
<td>-12%</td>
</tr>
<tr>
<td>Jonathan Quick</td>
<td>L.A</td>
<td>-11%</td>
</tr>
<tr>
<td>Ben Bishop</td>
<td>DAL</td>
<td>-11%</td>
</tr>
</tbody>
</table>

Updated: hockeyviz.com/goalies
Quality of competition

Some players are facing better quality opposition systematically enough to be noticeable:

- Carter Hutton in 730 days up to Jan 30, 2019, faced nearly 2,000 shots with 1.3% higher odds of being goals purely because of shooter talent.
- Jacob Chychrun took 200 shots which had 2.9% lower odds of being goals purely because of the goalies he faced.
## Weakest Recent Goaltenders

<table>
<thead>
<tr>
<th>Player</th>
<th>Team</th>
<th>Impact on goal odds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calvin Pickard</td>
<td>COL / TOR / PHI / ARI</td>
<td>+14%</td>
</tr>
<tr>
<td>Chad Johnson</td>
<td>CGY / BUF / STL / ANA</td>
<td>+14%</td>
</tr>
<tr>
<td>Jared Coreau</td>
<td>DET</td>
<td>+13%</td>
</tr>
<tr>
<td>Cam Ward</td>
<td>CAR / CHI</td>
<td>+10%</td>
</tr>
<tr>
<td>Maxime Lagace</td>
<td>VGK</td>
<td>+10%</td>
</tr>
</tbody>
</table>
Conclusions

- No version of individual save percentage has any value.
- Taking shot quality into account gives us a framework for measuring shooting and goalie ability.
Historical Work

All up, a great deal of light but little heat.

- **Shot difficulty:**
  - Kryzwicki, 2005
    - Similar ideas sketched by Ryder the previous year.
  - Schuckers, 2011 (updated 2016)

- **Bayesian updating:**
  - MacDonald, 2013

- **Simultaneous treatment of shooters and goalies:**
  - Ventura and Thomas, 2015
Future Work

- Screens
- Above-ice geometry
- Pre-shot movement
- Granularity of rush chances
- Aging
- Chaos
Thanks!