Introduction to Advanced Hockey Statistics

By Lars Skytte

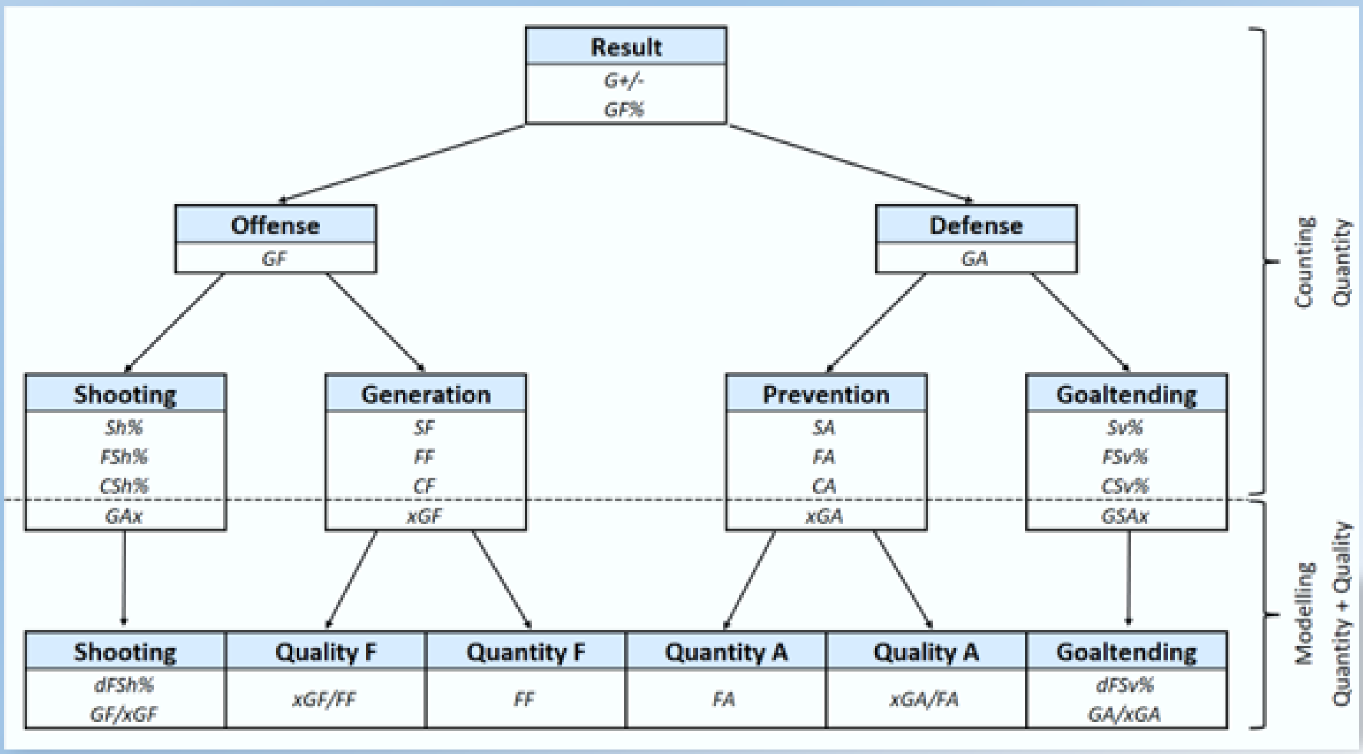


Table of contents

1. Introduction

Thank you so much for buying my book and supporting my continued work!

If you have comments, feedback or questions, please don't hesitate to contact me. You can write to: hockeystatistics.com@gmail.com or send a DM on Twitter.

The main purpose of this book is to give the reader an introduction to advanced hockey statistics. I try to write in an easy-to-understand manor. However, some of the content isn't trivial, so I hope the book will give the reader a few aha-moments along the way. This is why I think even experienced hockey analysts can learn a few things from the book.

The book will focus on 4 questions:

- ***What are advanced hockey statistics?***
- ***Why are advanced hockey statistics important?***
- ***How can I use hockey statistics in my own analysis?***
- ***Where can I find hockey statistics?***

Through reading this book you should:

- Learn about basic statistical metrics/concepts and how they are connected.
- Learn the difference between descriptive and predictive models.
- Understand the strengths and weaknesses of different models, so that you can add the proper context.
- Learn where to find hockey statistics.
- Gain the understanding and motivation to do your own data interpretation and research.

Hopefully, this book will teach you that hockey statistics doesn't have to be scary – Not everything requires a PHD in mathematics. Asking good questions and being curious will get you far.

"Curiosity is more important than knowledge"

- Albert Einstein

Website: Hockey-Statistics – A place for out-of-box thinking

Twitter: @HockeySkytte

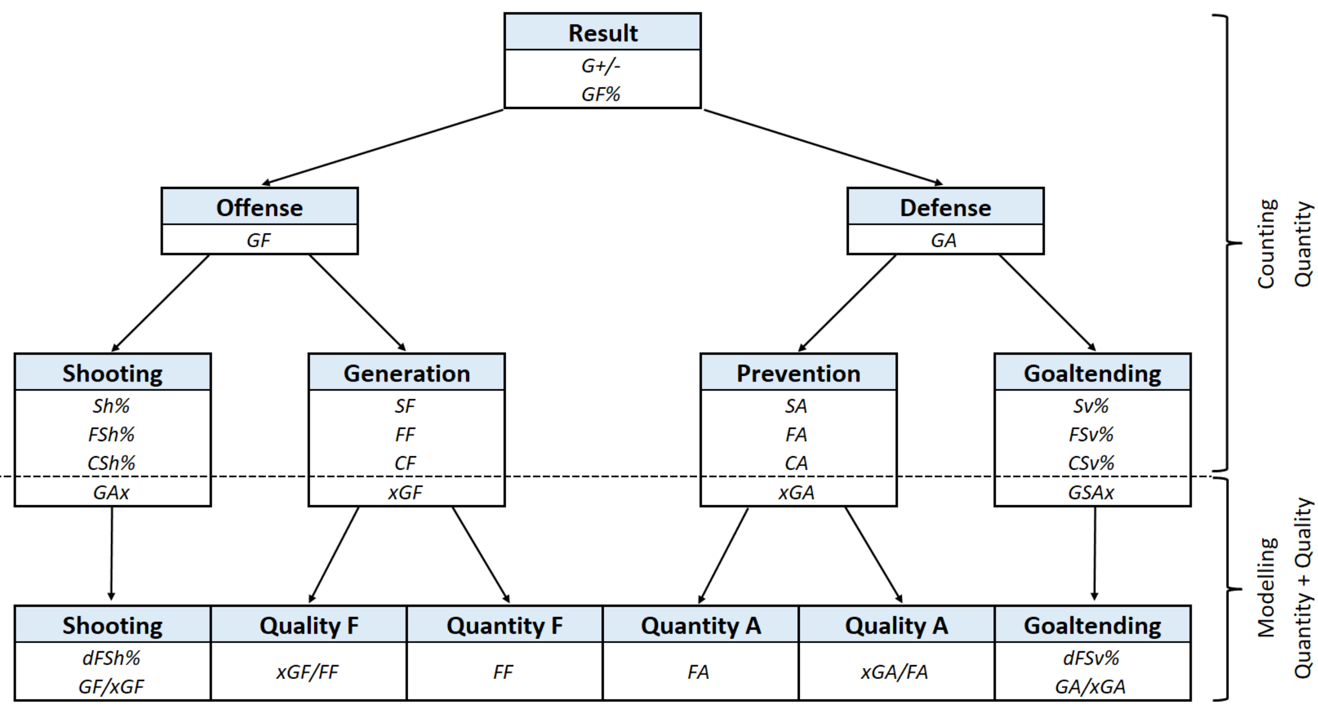
2. Shot statistics

Most statistics in hockey is based on shots. So, that will naturally be the focus in the first chapter. I think the best way to describe shot statistics is through a shot hierarchy. At the top you have the result, and then you can use shot data to describe how the result occurred.

2.1 – Shot Hierarchy

I will start off by showing the entire hierarchy in a single chart, and then I will walk you through each component step by step.

Here's the shot chart:



Abbreviations:

G+/-: Goal differential
GF%: Goals for percentage
GF: Goals for
GA: Goals against
Sh%: Shooting percentage
SF: Shots for
SA: Shots against
Sv%: Save percentage
FSh%: Fenwick Shooting percentage
FF: Fenwick for
FA: Fenwick against
FSv%: Fenwick Save percentage
CSh%: Corsi Shooting percentage
CF: Corsi for
CA: Corsi against
CSv%: Corsi Save percentage
GAx: Goals scored above expected
xGF: Expected Goals for
xGA: Expected Goals against
GSAx: Goals saved above expected
dFSh%: Delta Fenwick Shooting percentage: FSh% - xFSh%
GF/xGF: Goals for per Expected Goal for
xGF/FF: Expected Goals for per Fenwick for (Shot Quality)
xGA/FA: Expected Goals against per Fenwick against (Shot Quality)
dFSv%: Delta Fenwick Save percentage: FSv% - xFSv%
GA/xGA: Goals against per Expected Goal against

Result:

The ultimate goal in hockey is to score more goals than your opponent, so we can define the result as either goal differential (G+/-) or goals for percentage (GF%).

$$GF\% = \frac{GF}{GF + GA}$$

Clearly, the result is the end goal. All other analysis is either designed to describe how the results occurred or to predict future results.

Offense vs. Defense:

Statistically offense is defined as goals for and defense is defined as goals against. This makes sense statistically because it allows us to easily quantify offense and defense.

However, it's not how I (as a coach) would define offense and defense. In my definition, offense is the play when the team has possession of the puck and defense is the play when the opposition has possession of the puck. It's important to understand the difference!

Statistically, a team can play good defensive hockey by playing good possession hockey. If you possess the puck, you can't get scored on.

Say, a player in the offensive zone makes a high risk play which results in a breakaway and a goal against. Statistically speaking it's a bad defensive play, but intuitively it's a bad offensive play that leads to the goal against.

Shooting and Goaltending (Counting stats):

In the NHL data there are 4 types of shot events:

1) Goal
2) Shot (on goal)
3) Miss
4) Block

These shot types are then categorized in 4 statistical groups. It's a bit confusing that the event "shot" doesn't include goals, but the statistic "shots" does:

1) Goals
2) Shots: Goals + Shots
3) Fenwick: Goals + Shots + Misses
4) Corsi: Goals + Shots + Misses + Blocks

The naming here is somewhat stupid, but other that that it's pretty straight forward. Corsi simply means all shot attempts (SAT), fenwick means all unblocked shot attempts (USAT), shots mean shot attempts that hits the net (including goals) and goals just mean goals.

Goals, shots, fenwick and corsi are what we call simple counting stats. This means that you're simply counting the number of shot attempts. There's no modelling involved.

If we look at the shot hierarchy, we see that offense and defense is split into shooting/generation and goaltending/prevention respectively. At first, we're going to focus on shooting and goaltending. You would typically define shooting as Sh% and goaltending as Sv%:

$$Sh\% = \frac{GF}{SF}$$

$$Sv\% = 1 - \frac{GA}{SA}$$

However, you could also define shooting/goaltending based on fenwick or corsi:

$$FSh\% = \frac{GF}{FF}$$

$$FSv\% = 1 - \frac{GA}{FA}$$

$$CSh\% = \frac{GF}{CF}$$

$$CSv\% = 1 - \frac{GA}{CA}$$

If you sum shooting and goaltending, you get regular PDO:

$$PDO = Sh\% + Sv\%$$

This is the format you would see on stats sites, but you could also calculate PDO based on fenwick or corsi:

$$FPDO = FSh\% + FSv\%$$

$$CPDO = CSh\% + CSv\%$$

The reason you look at PDO, is because shooting and goaltending are less predictable than generation and prevention (More on this in chapter 3). This means that there's a higher variance in shooting and goaltending stats.

So, if the PDO is a lot higher or lower than 100%, there's a good chance the results are unsustainable. You often hear that PDO will regress to 100% over time, but this isn't necessarily true – Some teams have better than average goaltenders and shooters.

Over time the PDO should regress towards the true quality of the shooters/goaltenders.

Generation and Prevention (Counting stats):

Now, we will turn our attention to shot generation and shot prevention. This is simply the number of shots you generate and the number of shots you allow. Prevention means allowing as few shots as possible.

You often talk about shot differentials. The goal is to generate more shots than you allow. You can either look at normal differentials or you can look at percentages:

$$S\pm = SF - SA \qquad F\pm = FF - FA \qquad C\pm = CF - CA$$

$$SF\% = \frac{SF}{SF+SA} \qquad FF\% = \frac{FF}{FF+FA} \qquad CF\% = \frac{CF}{CF+CA}$$

The difference between shots generated and shots allowed is called play driving. So, when someone talks about play driving, they are likely referring to CF% or xGF% (more on that below).

Shooting and Goaltending (Modelling):

Until this point, we've only focused on counting stats. Let's now look at modelling stats (the stats below the dashed line in the shot hierarchy image).

What do I mean by a modelling stat? All the counting stats consists of just a single variable that you simply count. When we're building a model, we combine multiple variables and parameters.

All the modelling stats in the shot hierarchy are connected to expected goals. So, it's very important to understand what xG is.

Expected Goals:

The goal of an xG-model is to estimate the value of a shot – Clearly, a shot from the slot is worth more than a shot from the point.

So, xG-models combine different variables and parameters to best estimate the shot value. Exactly what goes into an xG-model differs slightly from model to model, but it's things like shot distance, shot angle, shot type, if it's a rebound shot, etc.

There are things that xG-models can't account for because the public NHL data is limited. As an example, you can't factor in cross ice passes since we don't have passing data. It's also difficult to properly estimate rush chances. This means that xG-models are most effective when we're looking at larger sample sizes – most of these problems will even out over time.

Of cause, certain players or certain teams may have a playing style that isn't measured particularly well with xG.

xG of a single shot is the estimated chance of a goal. If a shot has an xG-value of 0.091, then the shot has an estimated 9.1% chance of becoming a goal. xGF (expected goals for) is the sum of all shots for, so it combines shot quantity (number of shots) and shot quality (average xG-value). xGA (expected goals against) is the sum of all shots against.

All public xG-models that I know of are fenwick-based. This means that all blocked shots have an xG-value of 0. The primary reason for not including blocked shots in the model is that the NHL data doesn't include shot location for blocked shots.

Expected Goals summarized:

1) xG estimates the value of shots by combining a number of variables and parameters.
2) xG combines shot quality and shot quantity.
3) xG is fenwick-based (based on all unblocked shot attempts).
4) There are things xG-models can't account for, e.g. cross ice passes, rush chances.
5) xG is designed to equal actual goals when looking at the entire league.

Now that we have talked about expected goals, we can discuss the more advanced shooting and goaltender stats. In this paragraph we will go through 3 different ways of measuring shooting and goaltending:

1) GAx and GSAx
2) dFSh% and dFSv%
3) GF/xGF and GA/xGA

Let's go through them chronologically. GAx and GSAx is simply goals scored above expected and goals saved above expected respectively:

$$GAx = GF - xGF$$

$$GSAx = xGA - GA$$

This is a way to measure the total shooting/goaltending impact. How many goals have been scored/saved above expected?

Sometimes we're interested in measuring performance rather than impact. And what do I mean by that? If two goaltenders perform at the same level, but one of them is a starter and the other is a backup, then naturally the starter will have the greatest impact.

You can measure performance a few different ways. You could measure impact per 60 minutes, you could measure impact per shot, or you could measure impact per expected goal.

For goaltending and shooting it's common to measure performance as impact per shot. That's also what we do when we use Sh% and Sv%.

The advanced alternatives to Sh% and Sv% are dFSh% and dFSv%. These are defined as fenwick shooting percentage above expected and fenwick save percentage above expected:

$$dFSh\% = FSh\% - xFSh\%$$

$$dFSv\% = FSv\% - xFSv\%$$

It turns out that dFSh% is actually the same GAx per fenwick and dFSv% is the same as GSAx per fenwick:

$$dFSh\% = FSh\% - xFSh\% = \frac{GF}{FF} - \frac{xGF}{FF} = \frac{GF - xGF}{FF} = \frac{GAx}{FF}$$

$$dFSv\% = FSv\% - xFSv\% = \left(1 - \frac{GA}{FA}\right) - \left(1 - \frac{xGA}{FA}\right) = \frac{xGA}{FA} - \frac{GA}{FA} = \frac{xGA - GA}{FA} = \frac{GSAx}{FA}$$

So, dFSh% and dFSv% are ways to measure shooting and goaltending performance as impact per fenwick (unblocked shot attempt). This is often how you see goaltender performance being measured.

The 3rd option is to measure GF/xGF and GA/xGA. How many goals are scored/allowed per expected goal? Here we are measuring performance in regard to expected goals instead of fenwicks. If GF/xGF is above 1, you're scoring more than expected. If GA/xGA is above 1, your goaltending is below expected.

It's preferable to measure performance in regard to xG instead of fenwicks, because it's fairer. Say, two above average goaltenders have the same GSAx and they have faced the same number of shots (FA). One of them (goaltender A) has faced more difficult shots (higher xGA). The two goaltenders have performed equally well if we use dFSv% (they faced the same number of shots). But if we use GA/xGA then goaltender B has the best performance. He has the same impact (GSAx), but on a smaller workload (xGA).

So, I think GF/xGF and GA/xGA are the best ways to measure shooting/goaltender performance… But dFSh% and dFSv% are the consensus metrics. In other words, you will need to understand both metrics.

Earlier we discussed PDO. This was a way to combine goaltending and shooting into one metric. We can obviously do something similar with the more advanced goaltender/shooting stats. Here's how I would define delta-PDO or dPDO:

$$dPDO = dFSh\% + dFSv\% + 1$$

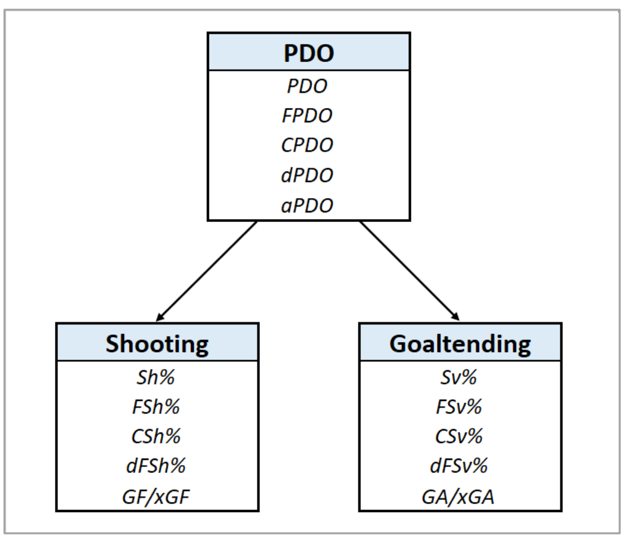
dPDO can be used the same way regular PDO, but it's a better metric since it includes shot quality.

Alternatively, we could define analytics-PDO or aPDO as:

$$aPDO = \frac{GAx}{xGF} + \frac{GSAx}{xGA} + 1$$

This would be my preferred metric for combined goaltending and shooting.

You won't find dPDO or aPDO anywhere. People are still using regular PDO, but I consider these metrics better. So, help spread the word.

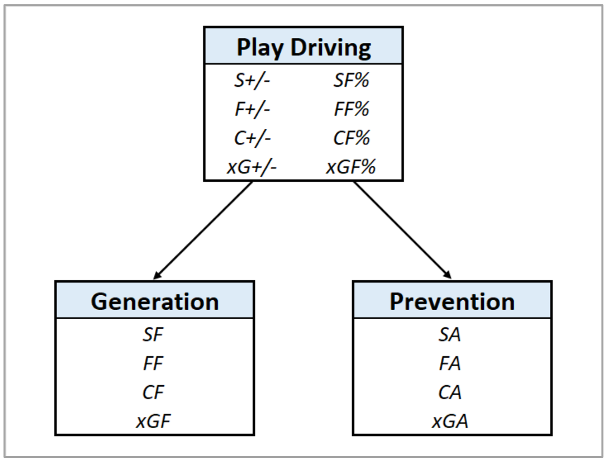


<u>Generation and Prevention (Modelling):</u>

The final part of the shot hierarchy is the advanced generation/prevention part. Generation is simply xGF, whereas prevention is xGA. Unlike the counting stats, these metrics combine shot quantity and shot quality. In fact, you can split up generation/prevention into quality and quantity.

Shot quality is defined as xG per fenwick – the average shot-value. Shot quantity is just the total number of shots. We're using fenwicks, because xG-models are fenwick-based.

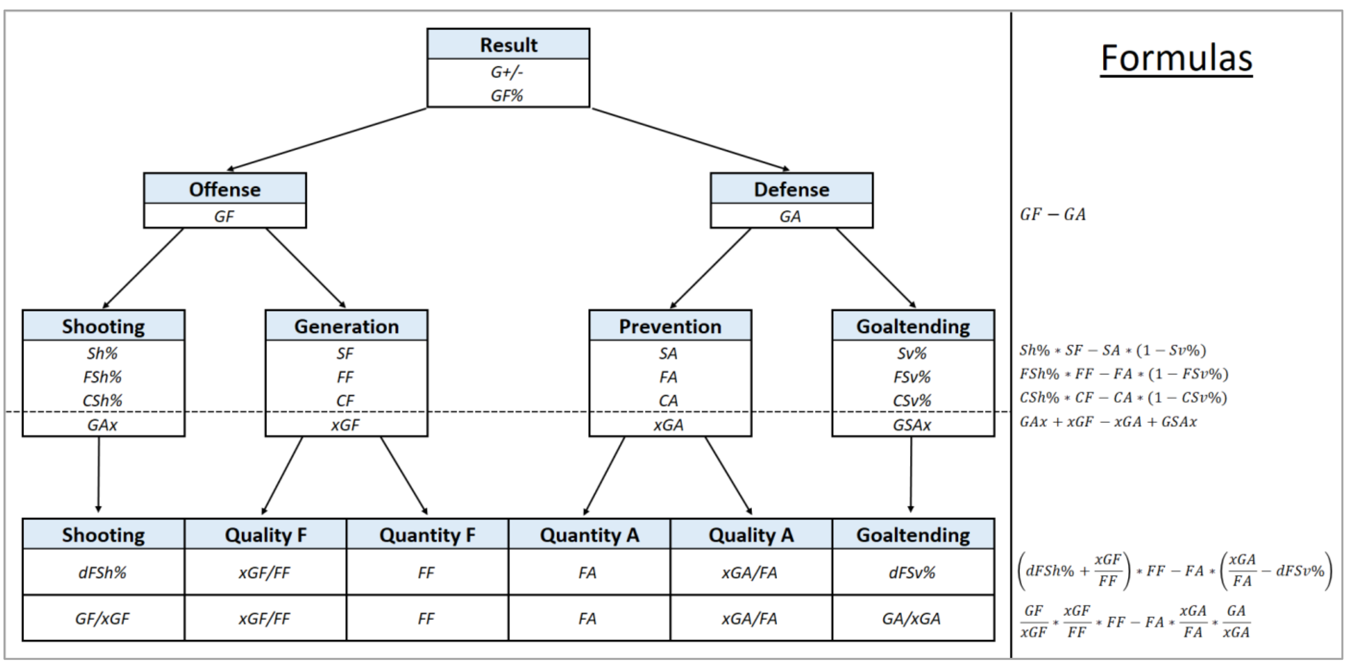
Like mentioned earlier, the combination of generation and prevention is called play driving:



2.2 – Shot Formulas

In this paragraph of the chapter, we will look at shot formulas. It's not necessary to understand all of this, but it should help give you an understanding of how the metrics are connected.

Here's the formulas for each step of the shot hierarchy. All the formulas are equal to the goal differential (G+/-), so it's a way to calculate the result.



On the next few pages, I will show all the calculations behind the formulas. If you're not interested in knowing how the formulas have come about, you can just skip this part and go directly to the "Shot Tracking" paragraph.

Shots:

For the rest of you, we will start off by looking at shots. I start by looking at goals for (GF), then goals against (GA) and lastly I combine the two.

The first step is to just divide by SF and multiply SF, and since GF/SF is the same Sh% we get:

$$GF = \frac{GF}{SF} * SF = Sh\% * SF$$

For GA it's a bit more complicated. We start by adding and subtracting SA. Then we divide and multiply by SA, and get:

$$GA = SA - SA + GA = SA - (SA - GA) = \left(\frac{SA}{SA} - \left(\frac{SA}{SA} - \frac{GA}{SA}\right)\right) * SA = \left(1 - (1 - \frac{GA}{SA}\right) * SA$$

Since: $Sv\% = 1 - \frac{GA}{SA}$, we must get:

$$GA = (1 - Sv\%) * SA$$

Finally, we can combine GF and GA:

$$G\pm = Sh\% * SF - SA * (1 - Sv\%)$$

Fenwick and Corsi:

The calculations for fenwick and corsi are exactly the same, so I will just show the results:

$$F\pm = FSh\% * FF - FA * (1 - FSv\%)$$

$$C\pm = CSh\% * CF - CA * (1 - CSv\%)$$

xG:

The next calculation is super simple. We just add and subtract xGF and xGA:

$$G\pm = GF - GA = (GF - xGF) + xGF - xGA + (xGA - GA)$$

And since: GAx = GF – xGF and GSAx = xGA – GA, we get:

$$G\pm = GAx + xGF - xGA + GSAx$$

This could also be written as:

$$G\pm = GAx + (xG\pm) + GSAx$$

So, shooting impact + play driving + goaltender impact equals result! This is a good thing to remember.

xG – (dFSh% and dFSv%):

In the next formula, we start by adding and subtracting xGF. Then we divide and multiply by fenwicks for (FF):

$$GF = GF - xGF + xGF = \left(\frac{GF - xGF}{FF} + \frac{xGF}{FF}\right) * FF$$

Since: GAx = GF – xGF and GAx/FF = dFSh%, we get:

$$GF = \left(\frac{GAx}{FF} + \frac{xGF}{FF}\right) * FF = \left(dFSh\% + \frac{xGF}{FF}\right) * FF$$

For GA the process is the same: We add and subtract xGA and then we divide and multiply by FA:

$$GA = -(xGA - GA) + xGA = -GSAx + xGA = \left(-\frac{GSAx}{FA} + \frac{xGA}{FA}\right) * FA = \left(-dFSv\% + \frac{xGA}{FA}\right) * FA$$

When we combine GF and GA, we get:

$$G\pm = \left(dFSh\% + \frac{xGF}{FF}\right) * FF - FA * \left(\frac{xGA}{FA} - dFSv\%\right)$$

xG – (GF/xGF and GA/xGA):

In the final formula we divide and multiply xGF and then we divide and multiply by FF:

$$GF = \frac{GF}{xGF} * xGF = \frac{GF}{xGF} * \frac{xGF}{FF} * FF$$

We do the same for GA:

$$GA = \frac{GA}{xGA} * xGA = \frac{GA}{xGA} * \frac{xGA}{FA} * FA$$

And combined it gives:

$$G\pm = \frac{GF}{xGF} * \frac{xGF}{FF} * FF - FA * \frac{xGA}{FA} * \frac{GA}{xGA}$$

So, results equal:

shooting*shot quality for*shot quantity for – shot quantity against*shot quality against*goaltending

That’s beautiful!

2.3 – Shot Tracking

Until now we have discussed the different shot metrics and how they are connected. In the final paragraph of this chapter, I would like to talk about the data itself. How is it tracked and are there any pitfalls?

How are shots tracked?:

All public NHL shot data is manually tracked. I expect some of the tracking to become automated in the very near future, but for now everything is manually tracked.

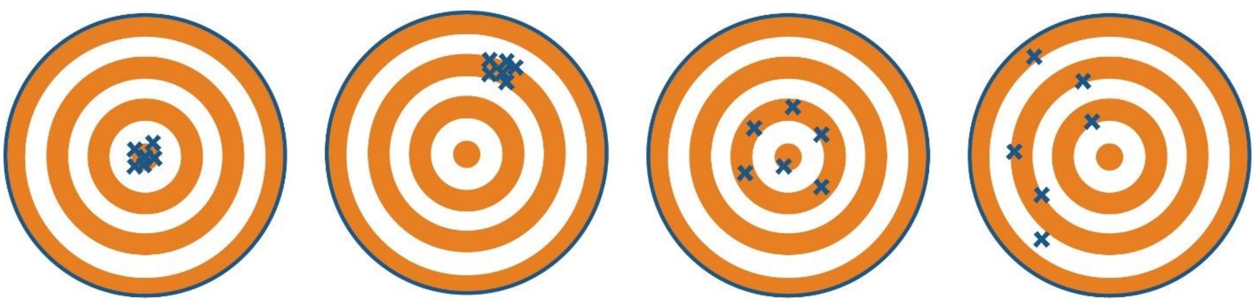
This means that the data isn't flawless, and we need to be aware of this. All data is tracked by the home team. When you track data in this fashion, it's very likely that you end up with data differences from arena to arena. This could skew the home data for some teams.

From a scientific standpoint you should randomize the trackers, so that potential differences would even out with a large enough sample size. As it is now, those differences will accumulate onto specific teams.

Accuracy vs. Precision:

This leads us to a discussion about accuracy versus precision. High accuracy means the average of the datapoints is correct. There might be mistakes, but those mistakes will even out, and the average will be correct. High precision means the datapoints are collected the exact same way, so they will always be close. However, if the accuracy is low, then all the datapoints will be wrong.

Accuracy and precision are illustrated in the following image:



High Accuracy High Precision | **Low Accuracy High Precision** | **High Accuracy Low Precision** | **Low Accuracy Low Precision**

How does this relate to NHL data? Since all data is tracked manually, we can't expect high precision. The same shot will be tracked differently – this is unavoidable. The hope is that on average the shot is tracked correctly (high accuracy). If the data has a high accuracy, then all our interpretations will be fine, for as long as we're looking at a sufficiently large sample size. However, if the accuracy is low, then we have problems.

I won't go into details in this book, but there are indications that shots in some arenas are tracked too close to the net, and in other arenas shots are tracked too far away from the net. This skews the xG data in those arenas. You can read about it in this article:

Indications that shot location data is flawed – Depends on where games are being played – Hockey-Statistics

When the data has a low accuracy, we need to calibrate the data, so that the average is correct. Some xG-models adjusts for rink bias (calibrates the model) and some xG-models don't. This is beyond the content of this book, but it's something to be aware of – especially when home numbers differ drastically from away numbers.

For more information about accuracy and precision I'd recommend reading this article by Garret Hohl:

Behind the Numbers: Pareto's Principle, Power Law Distribution, and when tracking data does not matter – Hockey Graphs (hockey-graphs.com)

2.4 – Summary

- Shot events can be: A Goal, a Shot, a Miss or a Block.

- Fenwick means all unblocked shot attempts: Goals + Shots + Misses.

- Corsi means all shot attempts: Goals + Shots + Misses + Blocks.

- Expected Goals is an estimate of shot value, and all xG-models are based on fenwicks.

- The shot hierarchy is designed to explain the results via shot metrics.

- Impact is measured in totals, whereas performance is measured in rates (Impact per 60, fenwick or xG)

- PDO metrics combine goaltending and shooting.

- Play driving combines generation and prevention.

- High accuracy means the average is correct, even if the individual datapoints aren't correct.

- High precision means the datapoints are always collected the same, even if it's incorrect.

3. Predictability

Hockey is all about winning – Scoring more goals than your opponent. So, why is it important to know how you win? A win is a win! It all comes down to predictability. Are the results you're getting sustainable or not?

In chapter 2 we looked at shot statistics. In this chapter we will show that play driving (xGF% or CF%) is a better predictor of future results than shooting/goaltending.

3.1 – Descriptive vs. Predictive

Before we get to the testing, it's worth discussing descriptive vs. predictive modelling. As the names indicate it's the difference between describing past events and predicting future events.

All the metrics in chapter 2 are designed to be descriptive. For instance, xG is designed to describe the shot value. In this chapter we will look at how predictive and repeatable certain metrics are. This will hopefully indicate why hockey analytics is important.

3.2 – Testing Predictability - Setup

Statistically this isn't the best way to test for predictability, but this book isn't about complex statistical modelling. This methodology is easy to understand, and it does illustrate the point quite well. If you're interested in a more scientific approach, I'd recommend reading this article:

Expected Goals are a better predictor of future scoring than Corsi, Goals | Hockey Graphs (hockey-graphs.com)

However, in our test we will just compare the first 41 games of the season to the last 41 games of the season. We are only looking at full seasons, so the following seasons are excluded: 12/13, 19/20 and 20/21.

Other than those 3 seasons we are looking at all seasons from 07/08 to 21/22. We are only using regular season data and it's all collected from www.Evolving-Hockey.com.

The approach is to simply compare data from the first half of the season to data from the second half of the season. This way we will get an indication of which metrics can be used to predict future results.

3.3 – Testing Predictability - Results

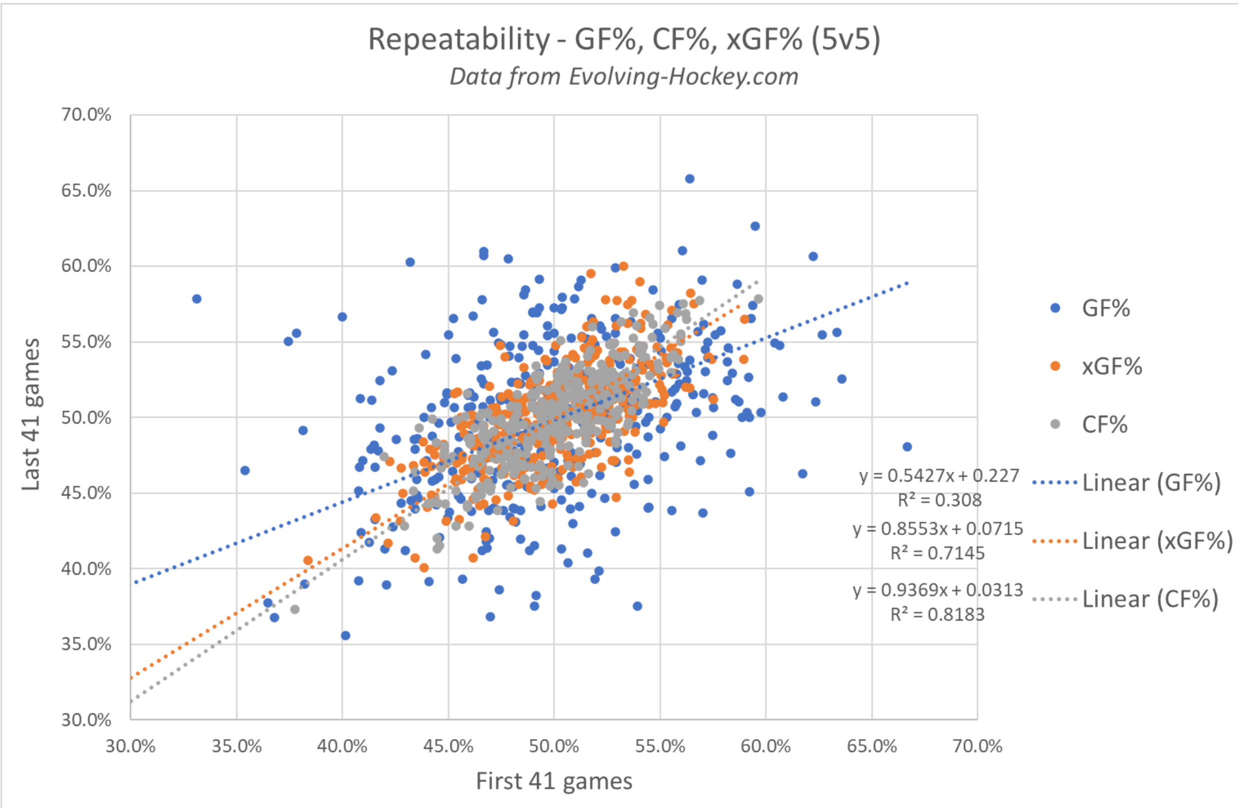
Repeatability:

We will start off by looking at five on five (5v5) repeatability. How well does a metric reproduce from the first half of the season to the second half of the season?

At first, we look at GF%, CF% and xGF%. The x-axis shows the metrics in the first 41 games of the season and the y-axis shows the metrics in the last 41 games of the season. If there were perfect correlation (repeatability), then the datapoints would be on a straight line.

Each datapoint represent a team from the 2007/2008 season to the 2021/2022 season (the shortened seasons are excluded). So, for instance one of the blue datapoints has the x-coordinate equal to Anaheim's GF% in the first 41 games of the 2014/2015 season and the y-coordinate equal to Anaheim's GF% in the last 41 games of the 2014/2015 season. Orange is xGF% and grey is CF%.

I'm using corsi instead of fenwick, because I already know corsi is the better predictor for future results.



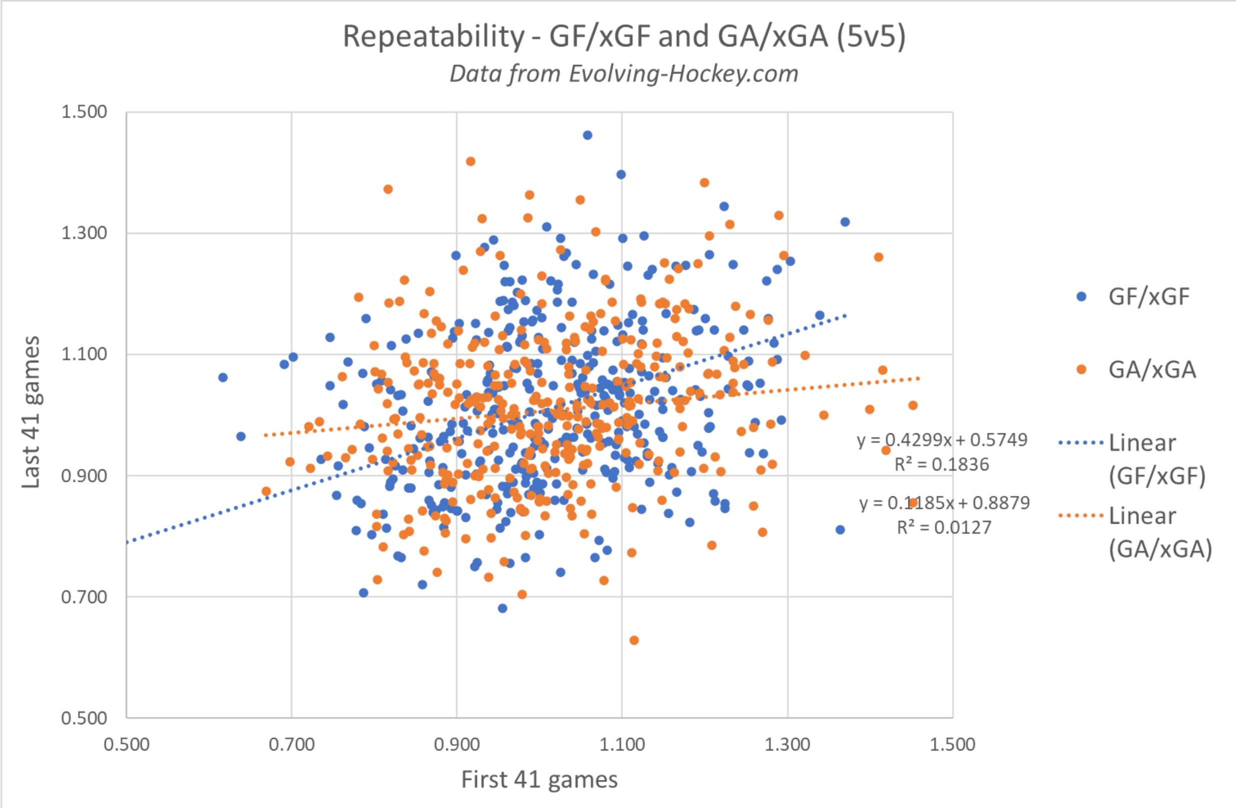
I have added trendlines to illustrate the repeatability (or lack thereof). I've put the trendline equations next to the correct legend entry.

The R^2 value is a measurement of correlation. If the value is 1, then there's perfect correlation (perfect repeatability) and if the value is 0 then there's no correlation.

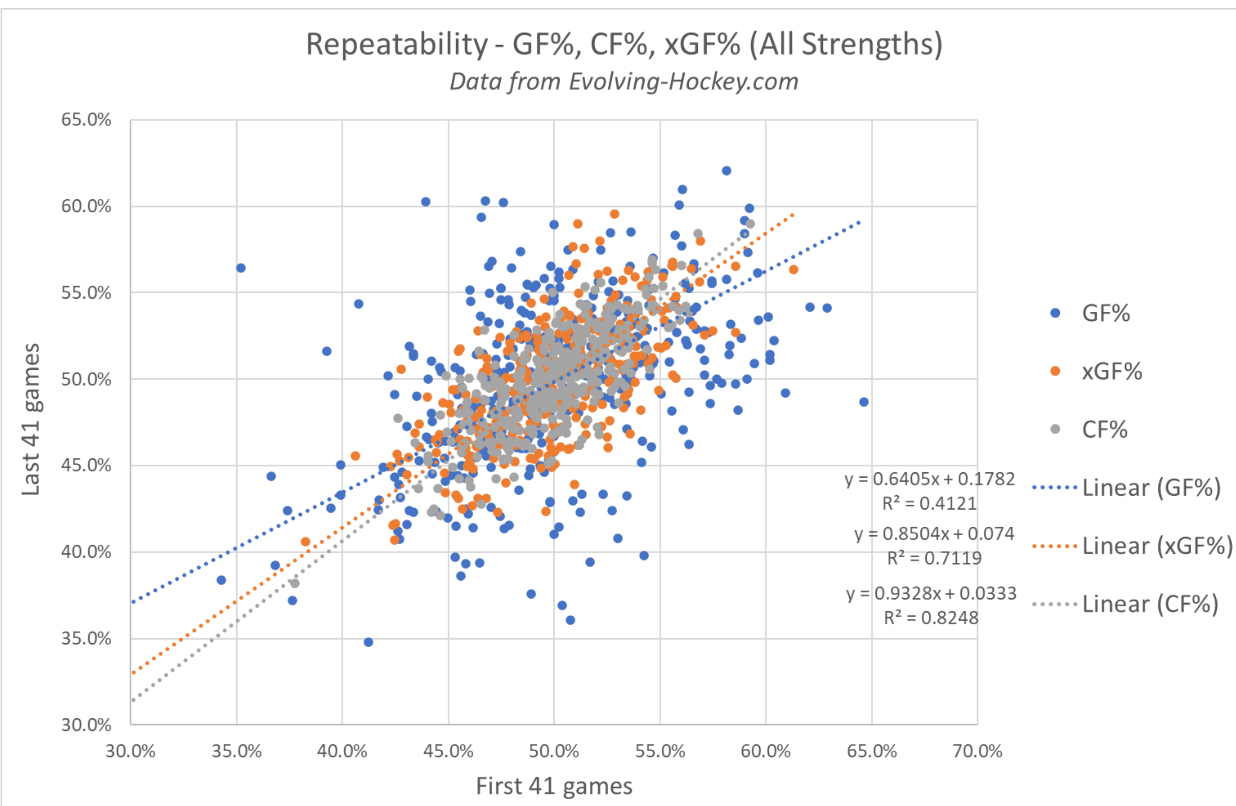
We clearly see xGF% and CF% are more repeatable than GF%. In other words, play driving is more repeatable than goal differential (results).

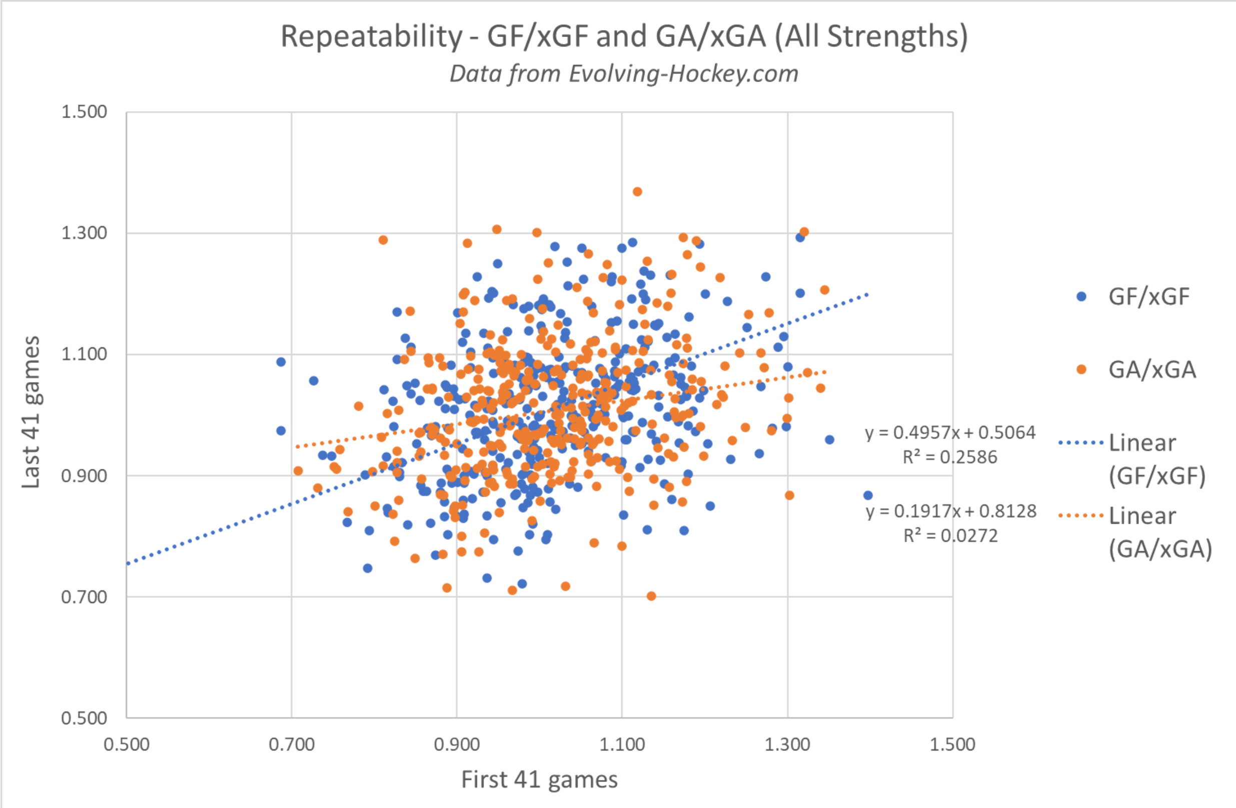
We saw in the Shot Hierarchy that results are the combination of play driving, shooting and goaltending. This means that the unrepeatability of GF% must primarily stem from shooting and goaltending.

The next graph shows the repeatability of shooting (defined as GF/xGF) and goaltending (defined as GA/xGA). We could have used classic Sh% and Sv% instead, and the results would have been similar. However, I think GF/xGF and GA/xGA are better metrics for shooting and goaltending.



We see that there's some repeatability for shooting, but almost none for goaltending. This is why analytics people say goaltending is impossible to predict. In the following graphs we look at all strengths data instead of 5v5:





At All Strengths the GF% is more repeatable, whereas the play driving metrics have similar repeatability. The increased repeatability of GF% comes from the shooting metric being more repeatable. This is probably because of increased sample size, but it could also stem from powerplay shooting being more repeatable than 5v5 shooting. This table shows all the R^2 values from the graphs:

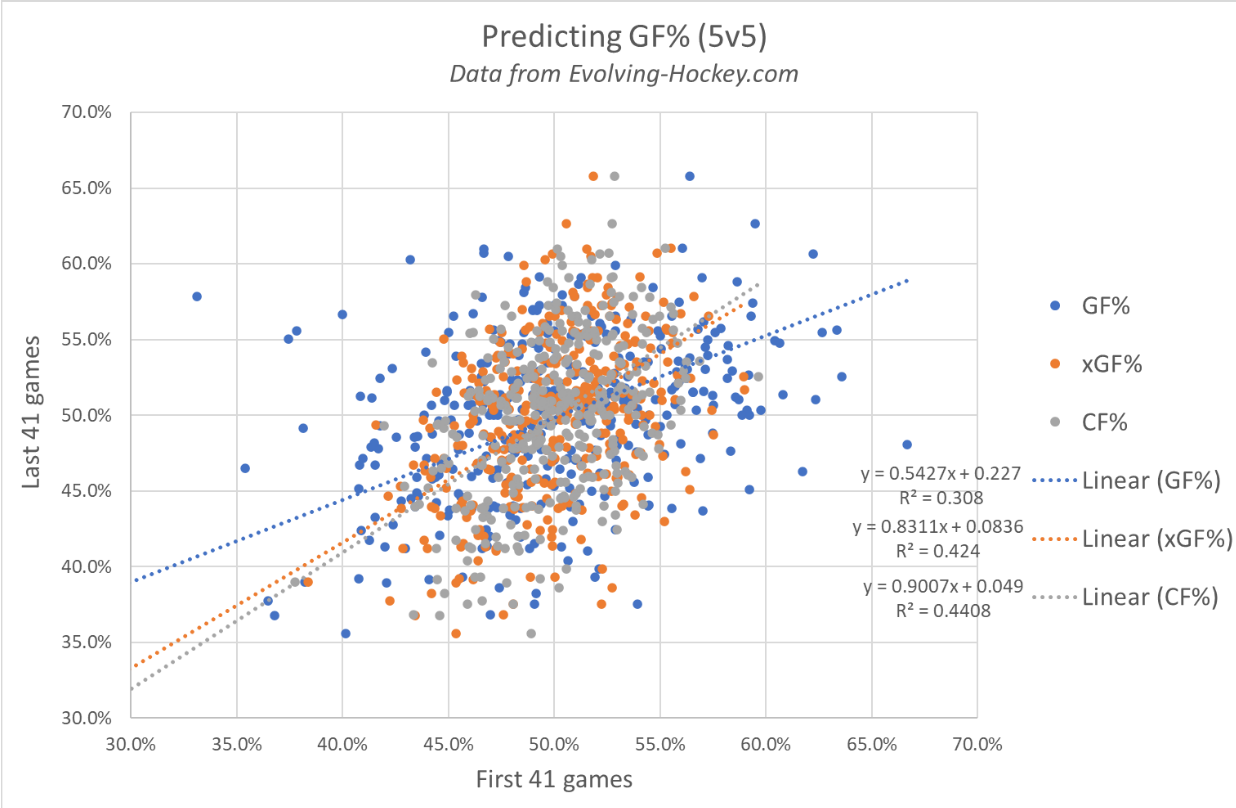
Metric	5v5	All Strengths
GF%	*0.3080*	*0.4121*
xGF%	*0.7145*	*0.7119*
CF%	*0.8183*	*0.8248*
GF/xGF	*0.1836*	*0.2586*
GA/xGA	*0.0127*	*0.0272*

So, just to summarize real quick: CF% and xGF% are more repeatable than GF%. This is because shooting and especially goaltending is less repeatable than play driving. This conclusion is exactly what I wanted to illustrate with the graphs!

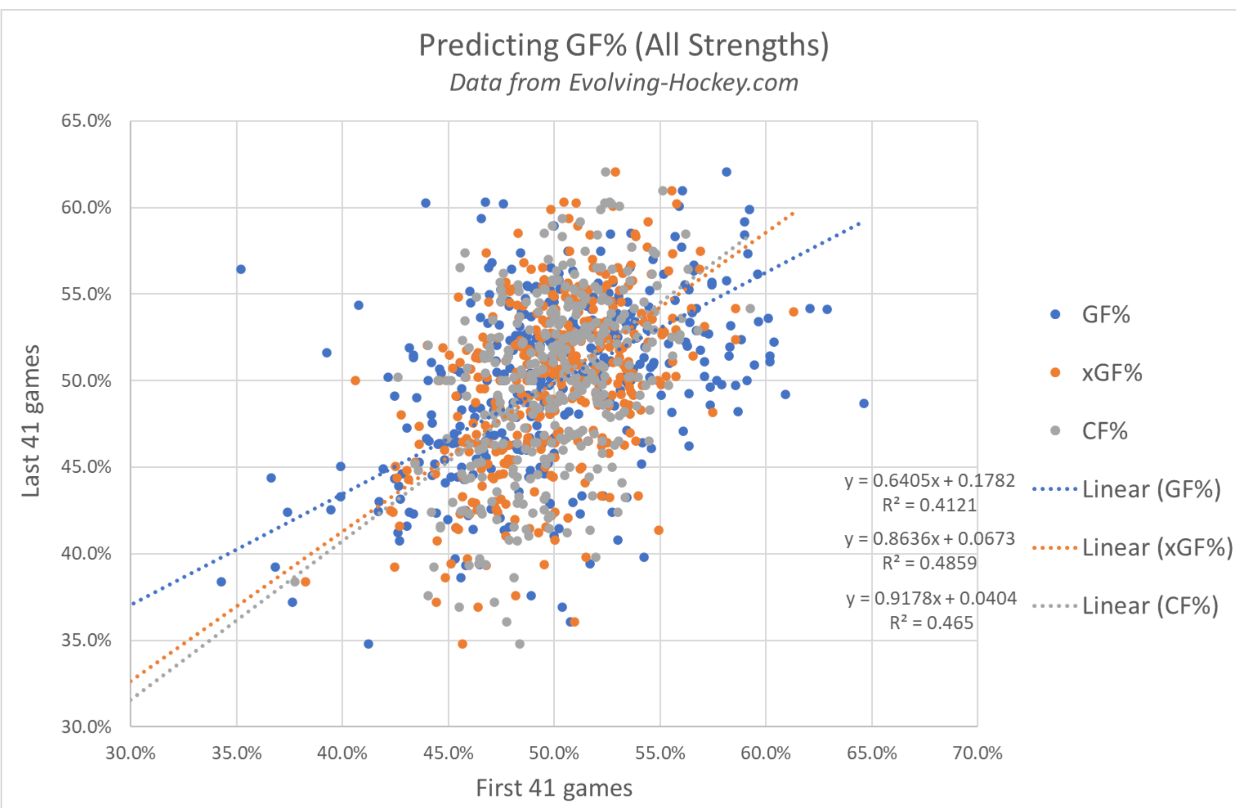
<u>Predictability:</u>

It's fine that play driving is more repeatable than results, but what we're really interested in, is predicting future results.

The next graph shows how well 5v5 GF%, CF% and xGF% in the first 41 games of the season correlate with GF% in the last 41 games of the season. Or put in another way, how well does GF%, CF% and xGF% predict future results (GF%).



At 5v5 we see here that xGF% and CF% are better predictors of future results than GF%. This is why it's interesting to know how you're winning, and not just if you're winning. If the results are driven by goaltending and shooting, they are less likely to be sustainable. The next graph shows the All Strengths predictability:



At all strengths the difference between GF%, CF% and xGF% is smaller than it was at 5v5. However, the play driving metrics are still better predictors of future results than GF%.

Metric	5v5	All Strengths
GF%	*0.3080*	*0.4121*
xGF%	*0.4240*	*0.4859*
CF%	*0.4408*	*0.4650*

3.4 – Testing Predictability - Conclusions

In chapter 2 we learned what Shot Statistics is. The point of this chapter is to illustrate the importance of shot statistics. Why is it relevant?

We can conclude that play driving is more repeatable than shooting and in particular goaltending performance. This is important to know when we try to foresee the future results.

If the goal is to build a predictive model, then we should put more weight on play driving metrics compared to shooting and goaltending metrics. This doesn't mean that goaltending and shooting aren't important – they are just difficult to predict.

It's fair to discuss whether the current goaltending/shooting metrics truly isolate goaltending and shooting performance. There are still many things the public xG-models can't fully describe – mainly pre-shot puck movement.

Are shooting and especially goaltending really this random? Or would improved models also lead to improved repeatability?

I certainly think goaltender performance varies a lot from season to season… but I also think our current goaltender metrics are flawed. If we could better isolate goaltender performance from team defense, we would likely be able to improve prediction models significantly.

3.5 – Sustainability vs. Luck

This leads me to another discussion. Because goaltending and shooting (PDO) is more random than play driving, it is often referred to as luck. Depending on your definition of luck, you could make this argument. However, when I hear the word luck, I think of something that requires no skill at all – that's clearly not what shooting and goaltending is!

So, in my mind when you call a team or a player lucky, you also miscredits the performance. I don't think that's fair. You can call the results unsustainable, though. It's perfectly reasonable to credit a team for its performance and still think it's unlikely to continue.

I think luck has become a "buzz-word" that creates an unnecessary division between the "analytics" crowd and the "eye-test" crowd. You can just as easily make your point by talking about sustainability and you wouldn't push people away.

The goal is to get more people interested in hockey analytics! Calling teams or players lucky has the opposite effect.

3.6 – Summary

- Descriptive modelling is describing past events – for instance by assigning value to shots, players etc.

- Predictive modelling is predicting future events, future performances.

- The play driving metrics: xGF% and CF% are more predictive of future results than GF% is.

- Goaltending is nearly impossible to predict.

- If the goal is to describe results, you would weigh Play driving, shooting and goaltending equally. If the goal is to predict future results, you would weigh Play driving more than shooting and goaltending.

- xG-models aren't perfect. They can't really account for pre-shot puck movement. This means that it's difficult to isolate goaltending from team defense.

4. Data Categories

Back in chapter 2 we talked about Shot Statistics in general terms. In this chapter we will try to categorize the stats and look at some possible applications.

There are 3 main categories of stats:

1) Team Statistics
2) Individual Statistics
3) On-ice Statistics

4.1 – Team-, Individual- and On-Ice Statistics

<u>Team Statistics</u>

There isn't much to say about team statistics because it's pretty obvious what it means. How well is the team performing?

The table below shows some common Team Statistics:

Team Statistics	Abbreviations
Results *GF%, Points, Wins* **Play driving** *SF%, FF%, CF%, xGF%* **Shooting Performance** *Sh%, dFSh%, GF/xGF, GAx* **Goaltending Performance** *Sv%, dFSv%, GA/xGA, GSAx* **Offense** *GF, SF, FF, CF, xGF* **Defense** *GA, SA, FA, CA, xGA*	***GF%*** *: Goals for percentage* ***SF%*** *: Shots for percentage* ***FF%*** *: Fenwick for percentage* ***CF%*** *: Corsi for percentage* ***xGF%*** *: Expected Goals for percentage* ***Sh%*** *: Shooting percentage* ***dFSh%*** *: Shooting percentage above expected* ***GF/xGF*** *: Goals scored per expected goal* ***GAx:*** *Goals scored above expected* ***Sv%:*** *Save percentage* ***dFSv%*** *: Save percentage above expected* ***GA/xGA*** *: Goals allowed per expected goal* ***GSAx:*** *Goals saved above expected* ***GF*** *: Goals for* ***SF*** *: Shots for* ***FF*** *: Fenwick for* ***CF*** *: Corsi for* ***xGF*** *: Expected Goals for* ***GA:*** *Goals against* ***SA:*** *Shots against* ***FA:*** *Fenwick against* ***CA:*** *Corsi against* ***xGA:*** *Expected Goals against*

There are other Team Statistics – E.g., hits, blocks, takeaways, faceoffs, penalties taken/drawn. The main focus here is on the Shot Statistics though.

Individual Statistics

That leads us to the Individual Statistics, which as the name indicates are statistics performed by the individual. The table shows some common Individual Statistics.

Individual Statistics	Abbreviations
SKATERS	***iSF:*** *Individual Shots for*
	iFF: *Individual Fenwick for*
Production	***iCF:*** *Individual Corsi for*
Goals, Assists, Points	***ixG:*** *Individual Expected Goals for*
	Sh% *: Shooting percentage*
Shot Creation	***dFSh%*** *: Shooting percentage above expected*
iSF, iFF, iCF, ixG	***GF/xGF*** *: Goals scored per expected goal*
	GAx: *Goals scored above expected*
Shooting	
Sh%, FSh%, dFSh%, GF/xGF, GAx	***Sv%:*** *Save percentage*
	FSv% *: Fenwick Save percentage*
	dFSv% *: Save percentage above expected*
GOALTENDERS	***GA/xGA*** *: Goals allowed per expected goal*
	GSAx: *Goals saved above expected*
Goaltending	***GSAA:*** *Goals saved above average*
Sv%, FSv%, dFSv%, GA/xGA, GSAx, GSAA	

Again, there are other Individual Statistics – E.g., Time on ice, hits, blocks, takeaways, faceoffs, penalties taken/drawn.

On-ice Statistics

The final category is On-ice Statistics. It's basically just Team Statistics, but from the perspective of an individual player. How is the team performing, when that specific player is on the ice?

The most famous On-ice Statistic is of course the classic +/-. What is the goal differential when the player is on the ice? There are some well known problems with +/-. You can get plusses while shorthanded and minuses while on the powerplay. This clearly helps PK-specialists and inhibits PP-players. The other problem is that +/- includes on-ice goaltending and shooting, which is much less repeatable than play driving (See Chapter 3).

I would recommend using 5v5 GF% or 5v5 G+/- instead of the classic +/- stat. It's almost the same thing, except the 5v5 stats don't give an advantage to PK-specialists and disadvantage to PP-specialists. However, I would still add on-ice goaltending and on-ice shooting for context.

The next table shows some of the common On-ice Statistics.

On-ice Statistics	Abbreviations
Results *GF%, +/-* **Play driving** *SF%, FF%, CF%, xGF%* **Shooting Performance** *oi-Sh%, oi-dFSh%, oi-(GF/xGF), oi-GAx* **Goaltending Performance** *oi-Sv%, oi-dFSv%, oi-(GA/xGA), oi-GSAx* **Offense** *GF, SF, FF, CF, xGF* **Defense** *GA, SA, FA, CA, xGA*	***GF%** : Goals for percentage* ***SF%** : Shots for percentage* ***FF%** : Fenwick for percentage* ***CF%** : Corsi for percentage* ***xGF%** : Expected Goals for percentage* ***oi-Sh%** : On-ice Shooting percentage* ***oi-dFSh%** : On-ice Shooting percentage above expected* ***oi-(GF/xGF)** : On-ice Goals scored per expected goal* ***oi-GAx:** On-ice Goals scored above expected* ***oi-Sv%:** On-ice Save percentage* ***oi-dFSv%** : On-ice Save percentage above expected* ***oi-(GA/xGA)** : On-ice Goals allowed per expected goal* ***oi-GSAx:** On-ice Goals saved above expected* ***GF** : Goals for* ***SF** : Shots for* ***FF** : Fenwick for* ***CF** : Corsi for* ***xGF** : Expected Goals for* ***GA:** Goals against* ***SA:** Shots against* ***FA:** Fenwick against* ***CA:** Corsi against* ***xGA:** Expected Goals against*

<u>WOWY Statistics</u>

This concludes the 3 main stat categories, but there is a 4th category – WOWY (With-Or-Without-You). It compares statistics when a player plays with you versus when the same player plays without you.

As an example, you could compare Evander Kane's stats when he's playing with Connor McDavid versus Evander Kane's statistics when he's playing without Connor McDavid. (Data from NaturalStatTrick)

Player 1	Player 2	GP	TOI	CF%	GF%	xGF%	oi-Sh%	oi-Sv%	PDO
Connor McDavid	Evander Kane	42	418	55.8%	63.8%	57.0%	10.6%	91.7%	1.023
Connor McDavid	w/o Evander Kane	80	909	57.2%	56.6%	60.0%	7.2%	92.8%	1.000
w/o Connor McDavid	Evander Kane	43	263	46.7%	64.3%	42.9%	7.3%	96.8%	1.041
w/o Connor McDavid	w/o Evander Kane	82	2394	50.5%	45.6%	47.9%	8.6%	90.4%	0.990

I would be very careful not to overestimate the importance of WOWY-numbers. However, it is a common type of analysis, so you will undoubtedly come across WOWY comparisons from time to time. Therefore, it's important to know what it means.

4.4 – Macro Stats and Micro Stats

All the metrics in the shot hierarchy are what we would typically call macro stats. They give you an overview of the results, but they don't really explain how the results came to be. Alternatively, there are micro stats: Zone entries, zone exits, passing, shot assists, puck battles etc.

The micro stats can add context to the macro stats – This is how the results occurred. If we generalize a bit, we could say that micro stats are what you see (eye test), and macro stats are the net result.

The biggest problem with micro stats is that the data isn't publicly available, so almost all micro stats data come from Corey Sznajder's amazing tracking project: Home (allthreezones.com)

Micro stats can add valuable context to your data, but using micro stats as the sole input for your analysis probably isn't a good approach. There are so many things you can't catch with micro stats.

You could say that micro stats are all the little things (problem is you can't measure every little thing), whereas macro stats are the sum of all the little things.

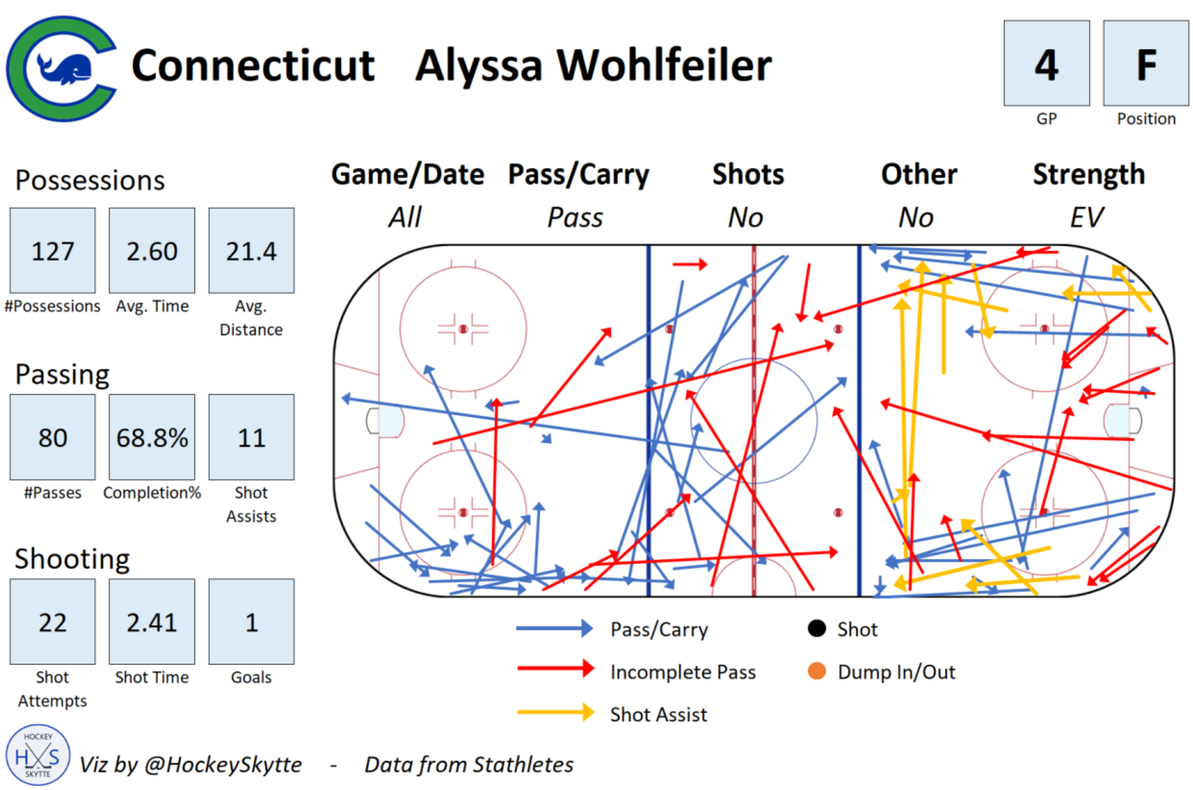
Big Data Cup Project 2022

In light of this discussion about micro stats, I think it's worth mentioning my Big Data Cup Project. The data from Stathletes includes some micro stats (passing, zone entries, zone exits).

I won't go into detail about the project, but you can find it here: BDC 2022 – Transforming Event Data Into Possession Data – Hockey-Statistics

Instead, I will just show a visualization of the micro stats. Here's the even strength passing map of Alyssa Wohlfeiler. Blue arrow means completed pass, red arrow means incomplete pass, and yellow arrow means shot assist.

So, Alyssa Wohlfeiler is near the top of the league in terms of shot assists per possession, but none of her shot assists are to the high danger area. This is one way you could use micro stats in your analysis.



Below is a short video presentation of the project. Here I go over the project really quickly and show where you can find all the cards and how to interpret them.

4.5 – Player Models

Now we've talked about the different data categories and macro stats versus micro stats. The final section of this chapter will be about player models – describing player impact with a single number.

The ultimate goal of hockey is to win hockey games and outscore your opponents. So, it's very easy to determine if a team is getting good or bad results. You simply look at the standings. However, in a team sport like hockey it's very difficult to assign player credit for those results. This is what player models attempt to do.

The problem is that we have no agreed upon definition of player quality. Some value goals, some value points, some value hits and some value something entirely different. The point is that player evaluations will always be subjective in nature. You can always argue for or against McDavid versus Matthews. It all depends on your preferred definition of player quality.

Player models are an attempt to take a more objective approach to player evaluations, but even within the models there are biases and preferences.

I will discuss some player models below, but first I want to talk about the models in more general terms.

- Player models (and RAPM models) attempt to isolate the player impact. They account for teammates and opponents in this process.
- Most player models use replacement level as the baseline.
- Player models are generally based on macro stats and not micro stats.
- Player models have difficulties with small sample sizes or when two players have played almost all their ice time together. Then it's difficult to isolate impact.
- Player models are designed to describe the results – Not to predict future results.

The Evolving-Hockey models:

Here are some very simple explanations of Evolving-Hockey's GAR and xGAR models:

GAR model: The GAR model is based on 3 things: Goals for (GF), Expected Goals against (xGA) and penalty differential. So, the model attempt to isolate each player's impact on these elements. Then the impact is compared to replacement level, hence the name Goals Above Replacement (GAR).

On-ice shooting plays an important role in GAR values, because one of the components is Goals for. This means that you should always look at the on-ice shooting for context. The GAR value is unlikely to be sustainable if the on-ice shooting is unsustainable.

xGAR model: The xGAR model is based on 4 things: Expected Goals for (xGF), Expected Goals against (xGA), Individual Shooting and penalty differential. Again, the model attempts to isolate the impact on these elements and compare the result to replacement level.

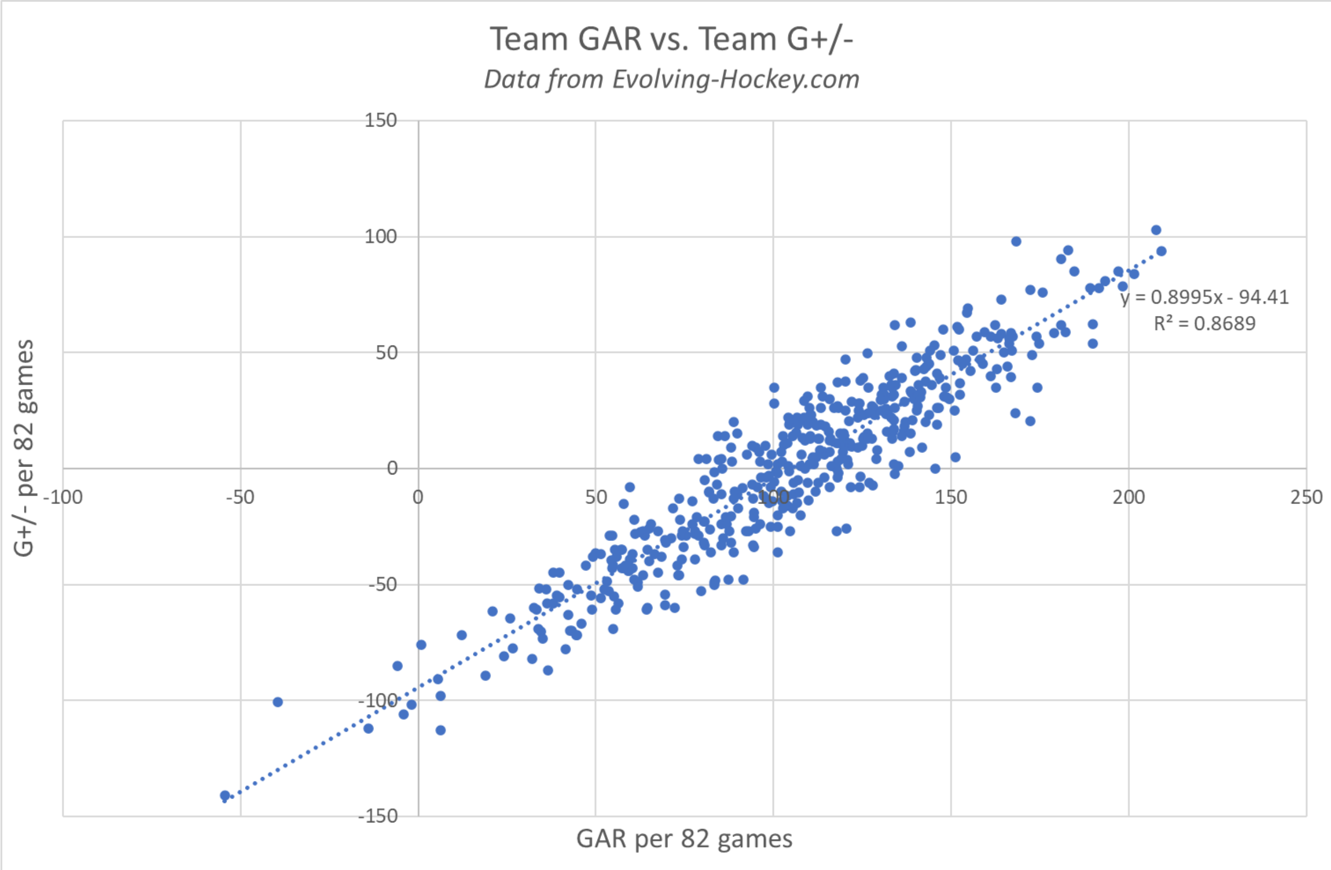
On-ice shooting is irrelevant in the xGAR model, because it's based xGF instead of GF. However, the Individual shooting is obviously very important. So, you need to look at individual shooting for context – is it sustainable or not?

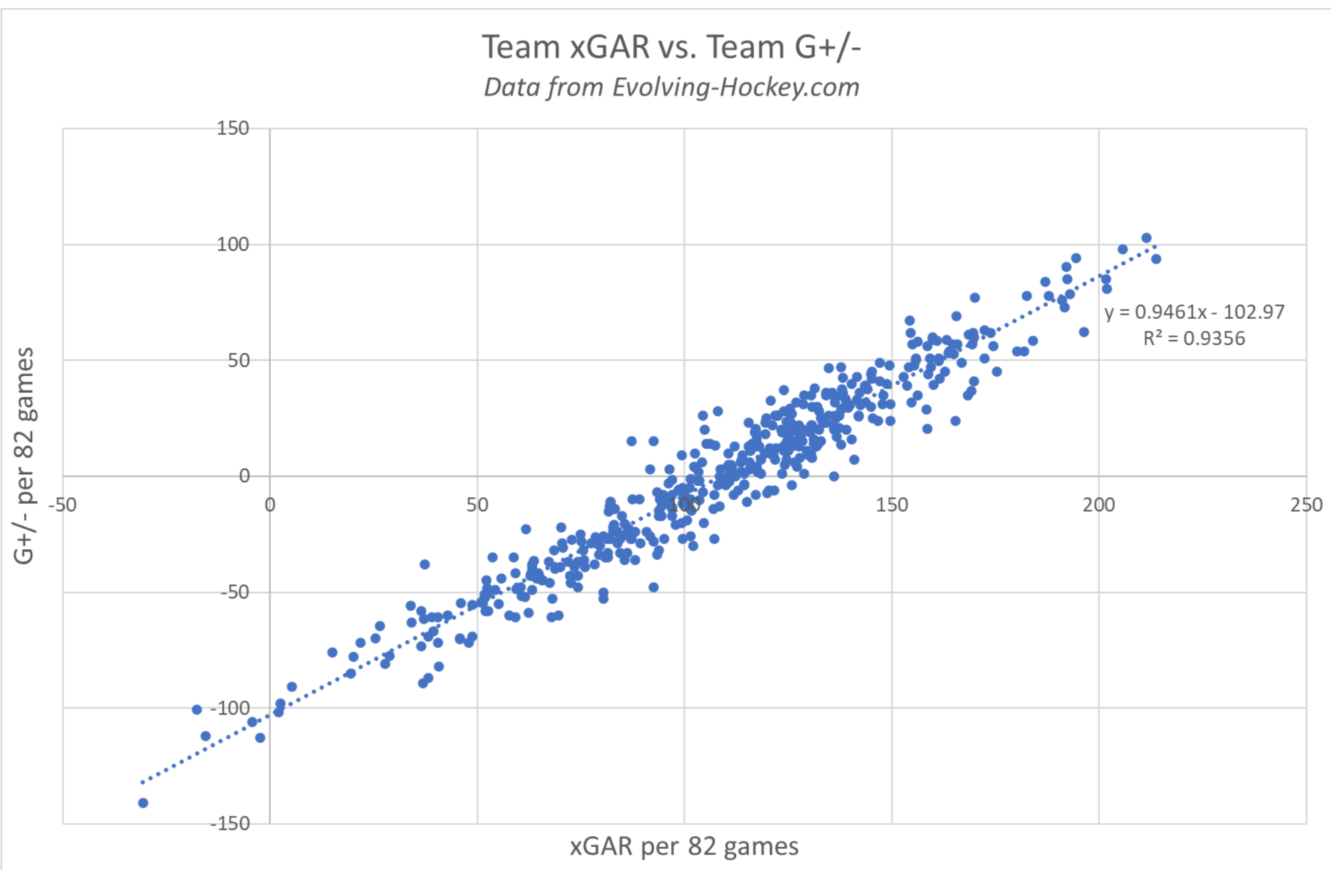
Generally speaking, I would say that the GAR model favors playmakers, whereas the xGAR model favors shooters. This is an oversimplification, but it's still something to be aware of.

Team GAR vs. Team Results

To illustrate that GAR models describe results, I will compare team GAR/xGAR and team results (team G+/-). I've simply plotted the Team GAR (x-axis) and Team Goal Differential (y-axis) of every season since 2007/2008. The data is prorated to 82 games to include shortened seasons.

The graphs can be found on the next page:


Team GAR vs. Team G+/-
Data from Evolving-Hockey.com
150
100
50
0
-50
-100
-150
-100
-50
0
50
100
150
200
250
G+/- per 82 games
GAR per 82 games
y = 0.8995x - 94.41
R² = 0.8689


Team xGAR vs. Team G+/-
Data from Evolving-Hockey.com
150
100
50
0
-50
-100
-150
-50
0
50
100
150
200
250
G+/- per 82 games
xGAR per 82 games
y = 0.9461x - 102.97
R² = 0.9356

These graphs clearly show that both models describe results quite well at the team level. There's a great correlation between Goal Differential and GAR/xGAR at the team level. This in itself doesn't mean that the GAR/xGAR models describe player impact well!

We now know that team GAR correlates well with team results, so if the models overestimate one player, they must also underestimate one player from the same team. We know that the sum of GAR/xGAR is correct for every team (correlates with team results)… But we don't know if the GAR value for each individual player is correct – And we can't really test for this.

It's difficult to isolate player impact in a team sport like Hockey!

Prediction Model

Single season player models aren't necessarily indicative of future impact – They are not designed to be predictive. So, if the goal is to predict the performance for the upcoming season, then you should increase the sample size and perhaps weigh play driving more than shooting. You could also include ageing curves – Expecting decline from older players.

4.6 – Summary

- Hockey Statistics can be split into three categories: Team Stats, Individual Stats and On-ice Stats.

- WOWY stands for With-Or-Without-You, and it compares the statistics of one player with or without another player.

- Micro stats are all the small plays (passes, zone entries, zone exits, puck battles etc.), whereas macro stats describe the overall results.

- Micro stats for the NHL aren't publicly available, but Corey Sznajder tracks an amazing amount of micro stats: [Home (allthreezones.com)](allthreezones.com)

- Player models like Evolving-Hockey's GAR and xGAR models attempt to isolate the impact of each player. They are designed to describe results – Not predict future results.

- Team GAR correlates with Team Results, so GAR models describe results well at the team level… But not necessarily at the individual player level.

5. Analyzing a Team Sport

This chapter is somewhat of a continuation of the last chapter, but this is something you're unlikely to find in any other statistical resource – It's more of a coaching perspective.

It should be clear that analyzing a team sport is different from analyzing an individual sport. The sum of all the parts doesn't always add up to the results!

5.1 – Productivity

In a team sport you have the following relationship:

$$Productivity = Potential +/\text{-}\ Processes$$

The productivity is the results/performance, the potential is the sum of all the individual players and the processes is everything else. In other words, the results are not just matter of the quality of your players, but also how well the players work together.

The processes could be hockey related things like role, line chemistry, strategy, on-ice communication etc., but it could also be more fluffy things like having fun, trusting your teammates/coaches, believing in the team, being happy etc.

The point is that the output/productivity of the team depends not only on the quality of the players, but also by culture and environment of the team.

General Managers and Coaches

The goal in pro sports is to increase the productivity to the point where you can win Championships. How can you do that?

The GM: As the General Manager you can increase the productivity in two ways: Getting better players (draft, trades, free agency) or getting a better culture/environment (finding players that fit together, hiring a coach that can optimize the potential).

$$\Delta Productivity = better\ players + better\ environment$$

The Coach: From the coaching perspective you can increase the productivity either via player development or via team development.

$$\Delta Productivity = player\ development + team\ development$$

Player development is difficult on a professional team, where most of the players are 20+ years old. So, the primary coaching job on a professional team is to best utilize the players at his disposal – giving players the correct roles, finding the best line combinations, matching lines, team strategy etc.

I'm a firm believer that you tactically can move the needle a lot… But it requires courage to revolutionize hockey tactics. You will sometimes need to take one step backwards in the hopes of eventually taking two steps forward. Hockey culture is very conservative.

Optimizing productivity

My last point in this section is about productivity optimization. If you aim to optimize the productivity for every single game, you will end up making some bad long-term decisions. We see teams in a rebuilding phase because they realize they can't win with the current players. So, they purposely decrease the productivity in the hopes of a long-term increase of the productivity. It makes perfect sense from a front office perspective.

It's less obvious from a player/coaching perspective. You always play the games to win… and that's fine. But you probably shouldn't play to win at all costs. That's when you see veterans play over prospects even though the difference is minimal. You see injured players take painkillers to play instead of resting. You see coaches apply the exact same strategy they've always done, because they are afraid to try something new.

Moving forward sometimes require a temporary step backwards.

5.2 – Output vs. Talent

I will try to bring back the discussion to hockey statistics in this section. We talked about player models in the previous chapter. They measure the impact/productivity of each player, but not necessarily the potential/talent. So, a player's output might change in a different environment (different role, line chemistry, strategy) even though the talent (natural ability) remains the same.

GAR models measure output – not talent!

I think this is important to know – putting together all the best GAR players won't necessarily lead to the highest productivity.

5.3 – Summary

- The results (productivity) aren't always equal to the quality of the players (the Potential). Other factors (Processes) will also affect the results.

- GAR models measure Productivity and not Potential. In other words, a player's GAR value might change in a different environment.

6. Data Interpretation

This chapter will focus on data interpretation. Interpreting data in a truly objective manor is extremely hard – There are so many pitfalls in hockey analytics.

6.1 – Biases

There are two types of biases we need to be aware of: Tracking biases and Interpretation biases.

Tracking Biases

All public data in the NHL is tracked manually. This means that the trackers themselves may have biases. The games are tracked by the home team, so the data may be imprecise if the shot trackers are biased (See Chapter 2). This is often referred to as Rink Bias because the biases differ from Rink to Rink.

I think there are 3 things we should be aware of when it comes to Rink Bias:

- Shot location: Shots are tracked too close to the net in some rinks and too far away from the net in other rinks. This leads to mistakes in the xG values.
- Shot count: Shots are overcounted in some rinks and undercounted in other rinks. Especially rebound shots are counted differently.
- Home team bias: The shot tracker may count more shots for the home team than the away team. I don't know if this bias actually exists today. I think the trackers aim to be neutral.

If you see big differences between the home data and away data, it may be connected to rink bias.

Interpretation Biases

Tracking biases lead to imprecise data, but the interpretation of the data could also be biased. This often happens when you already have a narrative, and you are trying to find data that confirms this narrative – Confirmation Bias. Then you will use the samples and metrics that confirm the narrative, while ignoring the samples and metrics that goes against the narrative. This can be either consciously or unconsciously.

Using data instead of the "eye-test" in itself doesn't mean your analysis is less biased. Making an unbiased and objective data analysis is difficult, and it requires a great understanding of the data.

A big problem is people claiming to know all the answers because they use facts and figures, when their analysis in reality is extremely biased and flawed. It makes hockey analytics look bad even though it's just the interpretation that's bad.

6.2 – Sample Size

Another classic pitfall is sample size. If the sample size isn't sufficient, then the uncertainty will be so large that you can't draw meaningful conclusions from the data.

Metrics like corsi and expected goals will stabilize much faster than goaltending and shooting… so a smaller sample size is needed. That's because goaltending and shooting include goals, so just a few lucky or unlucky bounces can completely skew the data. This aligns with what we saw back in Chapter 3 – Goaltending and shooting have a small repeatability when the sample size is just 41 games.

6.3 – Context

A single metric rarely tells the whole story, so you need to add context to the numbers. There's probably more, but these four elements are important to include your analysis:

- Teammates/Competition: Who is the most common line-mates? Does the player face 4th line competition or 1st line competition? Most player models (GAR/xGAR) already account for teammates and competition.
- Shooting: Is the on-ice shooting sustainable and is it impacting the numbers? On-ice shooting impact GAR, GF%, +/-, Assists… but it has no effect on CF%, xGF%
  Is the individual shooting sustainable? Individual shooting impact xGAR, Goals.
- Goaltending: Is the on-ice goaltending sustainable? On-ice goaltending impact GF%, +/-
- Role: How much ice-time is the player getting? Is he getting powerplay time? Is he playing shorthanded? These things will of course impact point totals, GAR/xGAR (if the player is above replacement level). The more ice-time a player gets, the higher the numbers.
  This is why we often look at things like: 5v5 points per 60 minutes, GAR_60, xGAR_60. It's a way to level the playing field.

6.4 – Knowledge and Creativity

One of the points of this chapter is to show that statistics don't necessarily tell the whole story. I like to view statistical models/knowledge as tools you can put in your toolbox. Having all the newest tools are great, but you still need carpentry skill to build a house.

Knowing when to use which tools and how to use the tools are essential for good data interpretation.

I even think, there's a more general discussion about knowledge versus creativity here. It's the difference between being able to play all of Mozart's compositions and composing your own music. One requires skill and knowledge, the other requires creativity.

It's not that knowledge isn't important, because clearly it is. It's just that being smart and creative in your usage of said knowledge is equally important. A player can have the hardest slap shot in the world, and it won't matter if he doesn't have the smartness and creativity to utilize it in game situations. It's the same thing with data. It's not enough to have all the newest data, you also need to know how to utilize it.

6.5 – Summary

- The NHL data may be imprecise due to tracking biases – Shot location tracked wrongly, Shots counted incorrectly, Home team shots counted more frequently.

- Confirmation bias is when you "make" the data fit your narrative.

- Insufficient sample size means the uncertainty is so large that you can't draw any meaningful conclusions.

- When finding outliers, you should always look at the context. Can we explain the results?

- Interpreting data requires more than just looking up the data. You need to understand the weaknesses and strengths of each model.

7. Data Sources

In this chapter I try to list some the most important data sources in advanced hockey statistics. I'm almost certainly missing a few great sources, but these links should get you a long way.

7.1 – Statistics Sources

NaturalStatTrick: Here you can find and download team stats, individual stats, on-ice stats and WOWY stats. There are also live game reports available.

Evolving-Hockey (Requires Subscription): Here you can find and download team stats, individual stats and on-ice stats. There are also live game reports available as well as player models (GAR, xGAR and RAPM). Evolving-Hockey also does projections: Point projections (Playoff probability), Game projections, Player projections and Contract projections. Finally, you can use Evolving-Hockey's scraper to download Play-by-Play data (the raw NHL data).

MoneyPuck: Here you can find and download team stats, individual stats, on-ice stats and Line stats. They also do live Game reports including the famous Deserve-To-Win-O-Meter. MoneyPuck also does Point projections (Playoff probability) and Game projections.

HockeyViz (Requires Subscription): Specializes in player visualizations where you can see the player impact as heat maps. HockeyViz also does Point projections (Playoff probability) and Game projections.

Dom Luszczyszyn (Requires Subscription): Dom makes player cards based on his GSVA player model. He also does Point projections (Playoff probability) and Game projections.

Corey Sznajder (Requires Subscription): Corey manually tracks Micro stats and this data is available with a subscription.

TopDownHockey: Here you can find and download team stats, individual stats, on-ice stats, GAR values and prospect data (NHLe). Patrick also does Point projections (Playoff probability) and Game projections.

JFresh (Requires Subscription): Specializes in team and player visualizations based on data from TopDownHockey and Corey Sznajder.

Byron Bader (Requires Subscription): Focuses on prospect projections (NHLe), and his prospect cards projects the probabilities of the player becoming a NHL player and a NHL star.

EliteProspects: An amazing resource for non-advanced hockey statistics from around the world.

WHKYHAC: Here you can find a list of links to great Women's Hockey data sources.

7.2 – Play-By-Play data

Most of the data you will come across is so-called summary data – Meaning it's the summed results over the course of a season or even a career. The alternative to this is to use the raw event data – the Play-by-Play data. Then you're looking at every single tracked event, so the data size is obviously much, much larger. There were 900,000 events last season if we include all events – There were around 150,000 shots.

If you work directly with the PBP data, then you will be less constrained by the data. As a simple example you could zoom in on a specific player's slap shots. This wouldn't be possible with the summary data, because shot type isn't a filter.

You can download PBP-data via Evolving-Hockey's PBP-quarry or you can download all shot data from MoneyPuck. For the more advanced users you can scrape the data directly from NHL API. I'm using Harry Shomer's Python scraper.

I don't expect the average reader to ever want to work with or interpret PBP-data, but I think it's important to know the difference between Summary data and PBP-data.

7.3 – Summary

- There are plenty of great NHL data sources. However, I'd recommend looking at the WHKYHAC links as well. There's some very interesting Women's Hockey data in there.

- Summary data is the summed data over the course of a season, a career or any other timeframe.

- Play-By-Play data is a table with every single event. It gives very large data-files.

8. Hockey-Statistics.com

In this final chapter, I will talk a bit about my statistical work and my website. Generally, you can categorize my work in 4 groups:

1) Modelling
2) Research
3) Model Tracking
4) Visualizations

I will talk about each category throughout the chapter.

I aim for transparency, originality and simplicity. I hope that most people can understand my writing and my visualizations – Or at least understand the thought process behind it.

My biggest strengths are by far my curiosity and creativity. I'm always trying to learn and I very rarely do it by taking the path most travelled by. You may not agree with everything I write… But that's in no way the goal anyway.

8.1 – Modelling

My primary model will be a projection model – A model that projects season results for the upcoming NHL season as well game projections for all NHL games. Exactly what the model will look like isn't determined yet, but I've written this article series and it will likely be very similar to that:

Game projection model – The Variables (Part I) – Hockey-Statistics

Game projection model – The Variables (Part IB) – Hockey-Statistics

Game projection model – In-season Model (Part II) – Hockey-Statistics

Game projection model – Pre-season Model (Part IIIA) – Hockey-Statistics

I never finished the series, but I plan to do a full writeup once the upcoming model is done.

I hope to do more model building in the near future and write about it in a way so that most people can follow along.

8.2 – Research

I've written some research articles along the way, so I will just link to some of those here.

Article about rink bias and how it impacts different metrics:

Indications that shot location data is flawed – Depends on where games are being played – Hockey-Statistics

Article about shot misses and whether or not goaltenders can directly impact the number of misses:

[Goaltenders have no apparent influence on shot misses! – Hockey-Statistics]

Article series about talent distribution in the NHL:

[Talent distribution – Percentiles (part I) – Hockey-Statistics]

[Talent Distribution – Predictability (Part II) – Hockey-Statistics]

[Talent distribution – Forwards vs. Defenders (Part III) – Hockey-Statistics]

[Talent distribution – Goaltending (Part IV) – Hockey-Statistics]

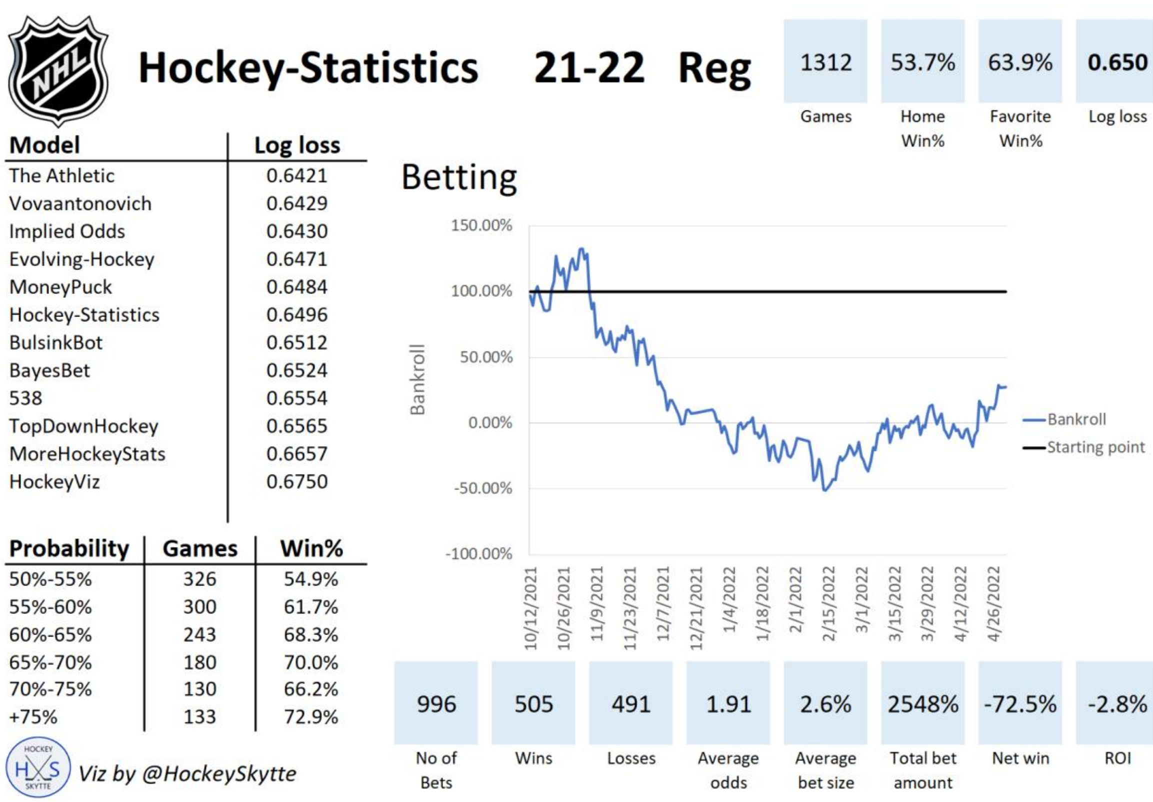
[Talent distribution – Contract value (Part V) – Hockey-Statistics]

I have some ideas for a couple of research articles, but it can be difficult to find the time and effort to get them done. Nonetheless, I hope to at least write a few research articles during the upcoming season.

8.3 – Tracking Game Projections

During the season I'm making daily game projections on Twitter (@HockeySkytte), but I'm also tracking other game projection models.

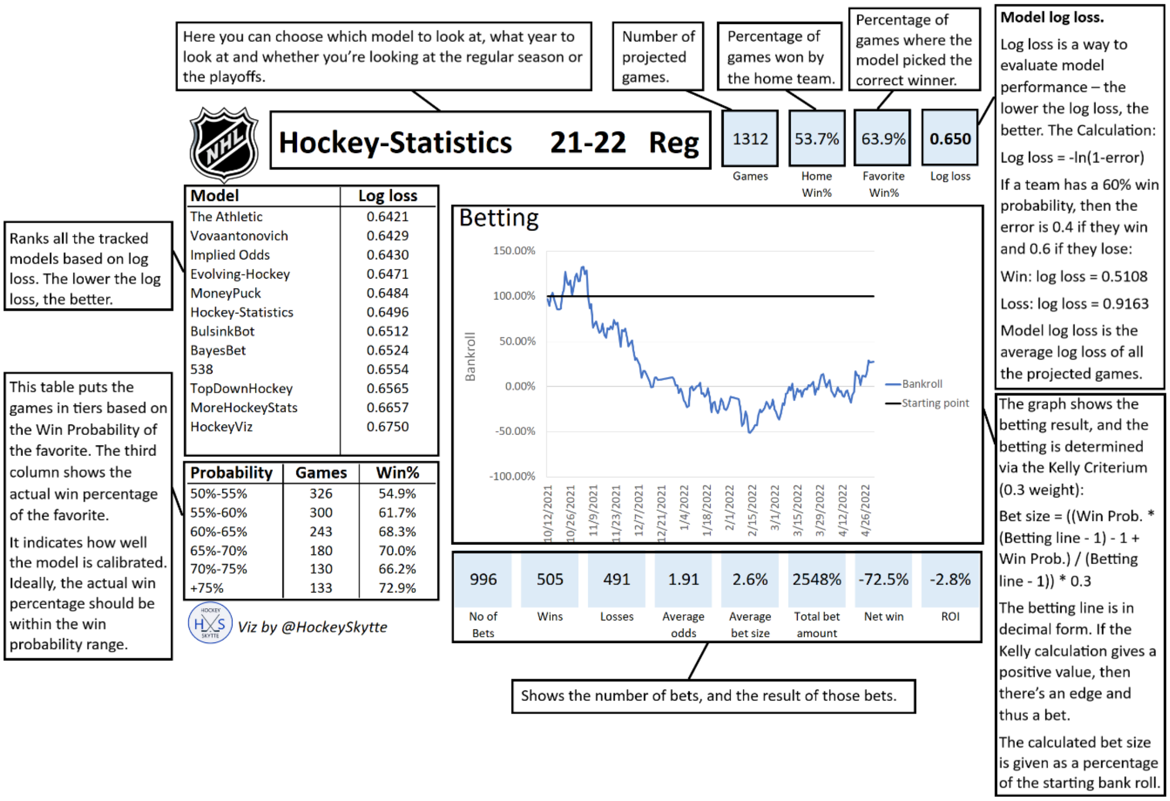
The model projections are then compared in my model cards. I try to update the cards on a daily basis, but sometimes it takes a bit longer. Here's the performance of my game projections last season:



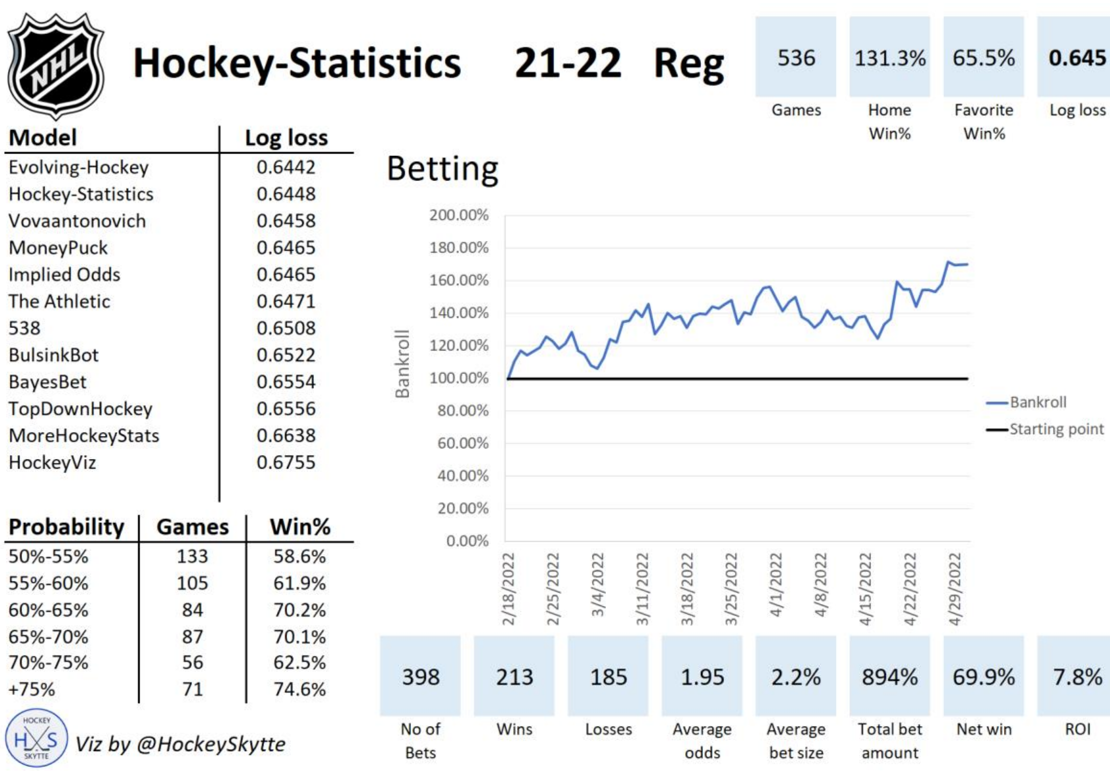
Model	Log loss
The Athletic	0.6421
Vovaantonovich	0.6429
Implied Odds	0.6430
Evolving-Hockey	0.6471
MoneyPuck	0.6484
Hockey-Statistics	0.6496
BulsinkBot	0.6512
BayesBet	0.6524
538	0.6554
TopDownHockey	0.6565
MoreHockeyStats	0.6657
HockeyViz	0.6750

Probability	Games	Win%
50%-55%	326	54.9%
55%-60%	300	61.7%
60%-65%	243	68.3%
65%-70%	180	70.0%
70%-75%	130	66.2%
+75%	133	72.9%

And here's an explainer card to help you understand the cards:

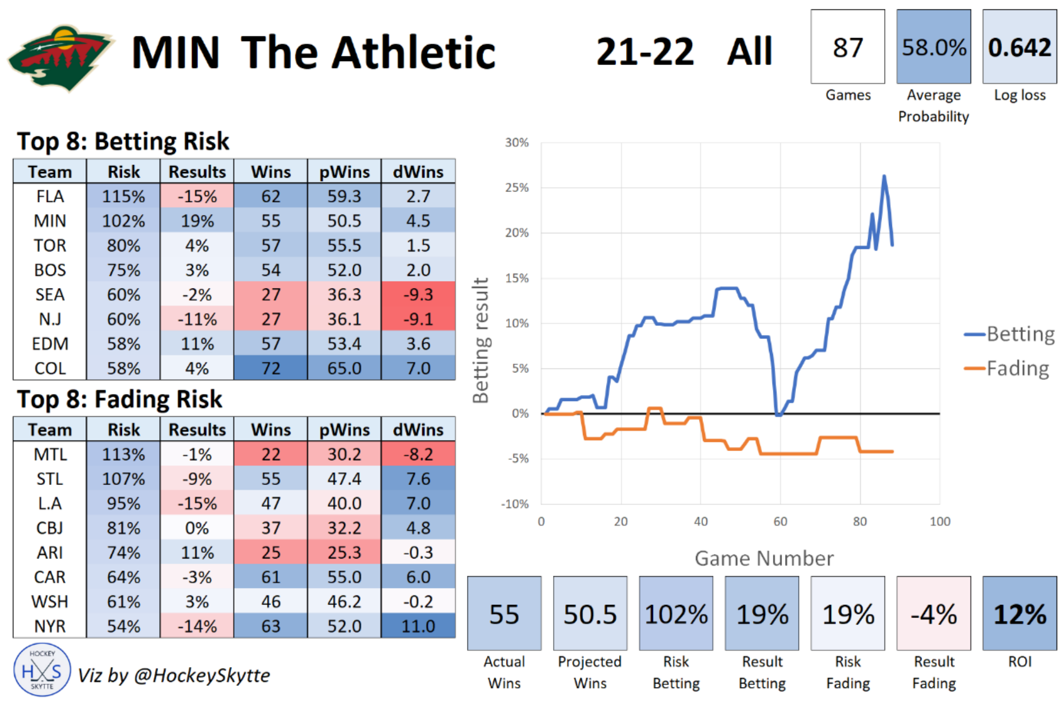


My model didn't perform particularly well last season, as you can see from the model cards, so back in February I switched to a completely different model. Here's the model performance since implementation of the new model (this gives me hope for the future):

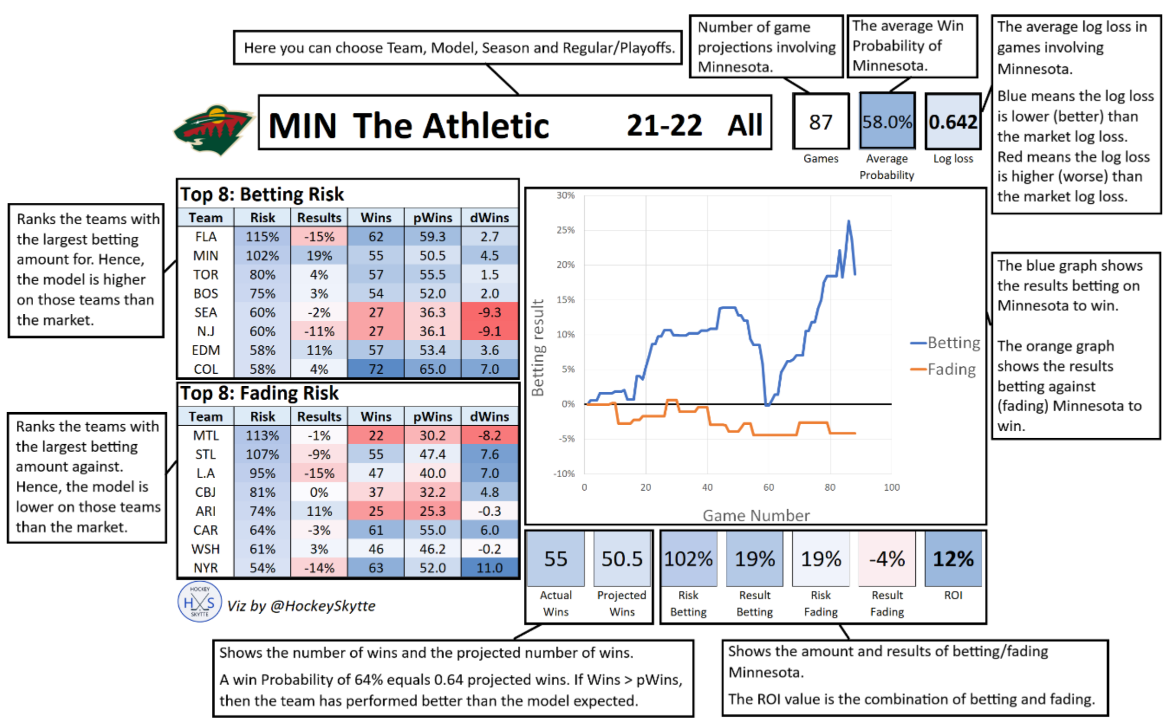


Building a model that can beat the closing betting lines is extremely difficult, but it's always interesting to compare the projections to the market even if you don't bet (I don't).

Some models will inevitably be consistently wrong about specific teams, so I also made a visualization that shows the betting results on or against the specific teams for each model. This way you can see if the model is higher or lower on a specific team than the market… and if the model beat the market on said team. Here's an example of how well Dom Luszczyszyn's model did betting on or against Minnesota in 2021/2022 season:



And here's the explainer card:



So, Dom's model was higher on Minnesota than the market… and the model had success betting on Minnesota.

My hope is that these Model cards can help add some transparency to the different game projections out there.

8.4 – Visualizations

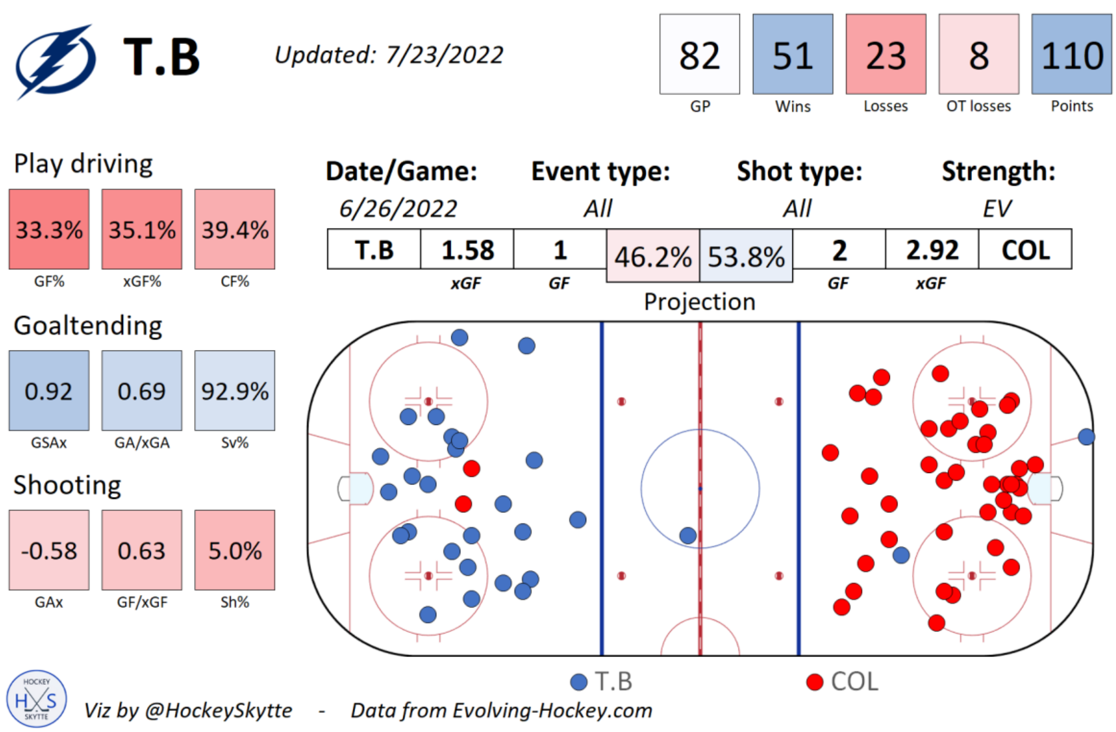
I've already talked about the Model cards, but you can find plenty of other visualizations in the cards section of my website.

Cards section:

I won't go through all the cards in this section, because I don't yet know exactly what cards/visualizations I will have for the upcoming season.

Game reports:

I plan to make some sort of Game Report card, where you choose a game and see the shot locations. It might look something like this:



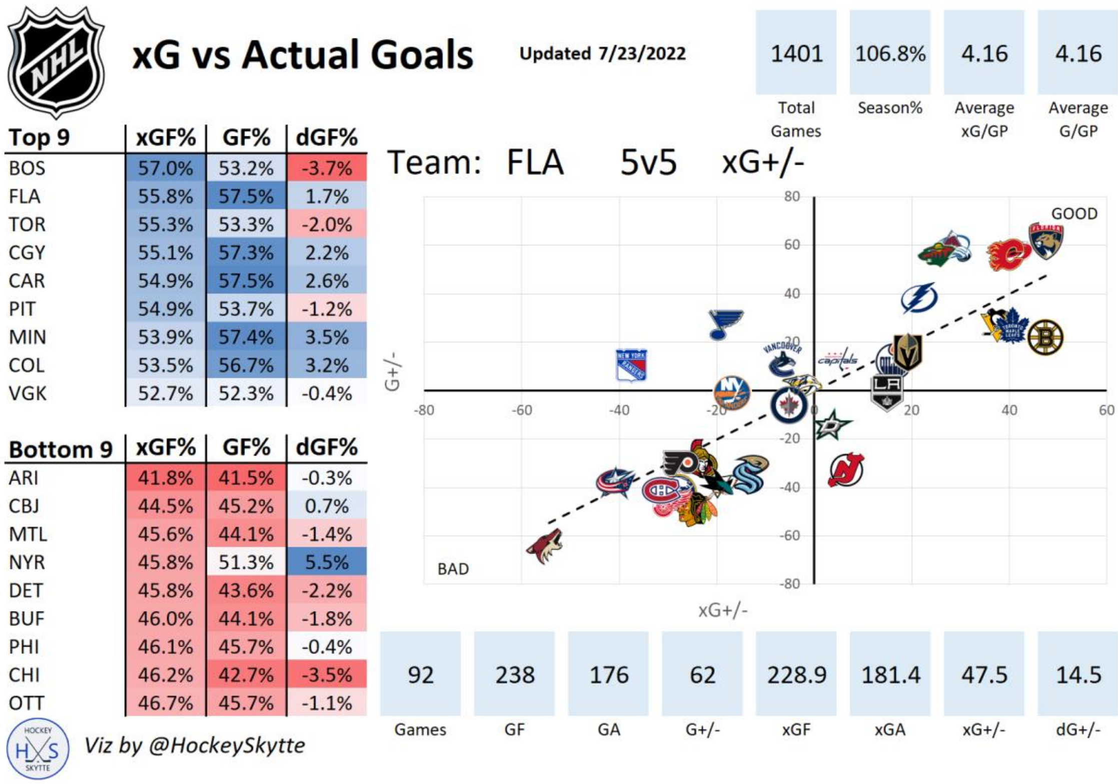
Player and Team Cards:

I will certainly have both player and team cards available, but I need to build the projection model before I know how they will look.

The player and team cards currently found on the site will likely change before the upcoming season.

xG vs. Actual Goals:

Last year I posted these cards comparing Expected goals to Actual goals. I really like the design here, because you can easily see if the results stem from Goaltending, Shooting or Play driving.



Top 9	xGF%	GF%	dGF%
BOS	57.0%	53.2%	-3.7%
FLA	55.8%	57.5%	1.7%
TOR	55.3%	53.3%	-2.0%
CGY	55.1%	57.3%	2.2%
CAR	54.9%	57.5%	2.6%
PIT	54.9%	53.7%	-1.2%
MIN	53.9%	57.4%	3.5%
COL	53.5%	56.7%	3.2%
VGK	52.7%	52.3%	-0.4%

Bottom 9	xGF%	GF%	dGF%
ARI	41.8%	41.5%	-0.3%
CBJ	44.5%	45.2%	0.7%
MTL	45.6%	44.1%	-1.4%
NYR	45.8%	51.3%	5.5%
DET	45.8%	43.6%	-2.2%
BUF	46.0%	44.1%	-1.8%
PHI	46.1%	45.7%	-0.4%
CHI	46.2%	42.7%	-3.5%
OTT	46.7%	45.7%	-1.1%

In this case we see that Florida created 47.5 goals from Play driving at 5v5, whereas 14.5 goals came from goaltending and shooting.

If the team is above the dotted line, it means that the team is outperforming their expected goals. If the team is under the dotted line, it means that the team is underperforming their expected goals. So, we would expect some regression towards the dotted line, since Goaltending/Shooting is less predictive of future results than Play driving (See chapter 3).

8.5 – Community

That was a short introduction to what you can expect to find on Hockey-Statistics.com, but ideally the site can become much more than this. The dream is to have Hockey-Statistics become more of a community rather than just my crazy ramblings. I wish the site would have much more life, but I neither have the time nor the skill to make that happen.

@ImpctSport has made a few contributions to the site:

Deep Learning Modeling of Hockey Game Contribution – Hockey-Statistics

Deep Dive into Offensive and Defensive Quality in the NHL – Hockey-Statistics

Hockey Deployment Management analysis using Machine Learning – Hockey-Statistics

…But it would very cool if others made contributions as well. When I started out a few years back, I could have really used a platform to write from. So, if you want to contribute to site, or you just have a good idea or a question, don't hesitate to reach out. You can either write to: hockeystatistics.com@gmail.com or send a DM on Twitter.

9. Conclusion

I hope you have enjoyed the book. Writing really isn't a strength of mine, so I always try to focus on structure and content. I'm sure you can find words missing, spelling errors or grammatical errors, but I promise you it's not because of laziness. I have a hard time detecting mistakes like that.

I can always use the excuse that English isn't my first language… but I would probably make the same number of mistakes in Danish.

I think we covered a lot of subjects in this relatively short book, and hopefully this gives you a good starting point to really get into hockey statistics. I don't necessarily consider this book a finished project, so I may add or elaborate on some things based on the feedback.

The goal is to get more people interested in hockey statistics, and hopefully this book can help in that matter… But we as a community also need to be inclusive and encouraging. I know Social Media is a tough place, but let's all try to be constructive in our feedback and let's try to spread content rather than clickbait and trolling.

"The true sign of intelligence is not knowledge but imagination"

- Albert Einstein

Perspectives:

I'm thinking about doing some video tutorials to help people get a more hands-on introduction to statistical analysis. It will be done in Excel, as all my visualizations are done in Excel. I think coding in Python or R is a lot more frightening to people (myself included), and the goal is to get more people interested in hockey statistics.

Obviously, Excel has its limitations when it comes to complex modelling, but in regard to data interpretation and visualizations Excel will do just fine.

I would also like to write more about model building. This will likely be done on the website at first, but eventually I might put in a book.

Lastly, I just want to thank you once more for buying my book. If you liked it, I hope you will help spread the word by sharing the link. It would mean a lot to my continued work and continued motivation.

9. References

Hockey-Graphs:

Behind the Numbers: Pareto's Principle, Power Law Distribution, and when tracking data does not matter | Hockey Graphs (hockey-graphs.com)

Expected Goals are a better predictor of future scoring than Corsi, Goals | Hockey Graphs (hockey-graphs.com)

Evolving-Hockey:

Evolving-Hockey

NaturalStatTrick:

Natural Stat Trick

Corey Sznajder:

Home (allthreezones.com)

Stathletes:

GitHub - bigdatacup/Big-Data-Cup-2021: Big Data Cup 2022: Powered by Stathletes

MoneyPuck:

MoneyPuck.com -NHL Analytics, Playoff Odds, Power Rankings, Player Stats

HockeyViz:

HockeyViz

Dom Luszczyszyn:

Dom Luszczyszyn - The Athletic

TopDownHockey:

Profile - topdownhockey | Tableau Public

JFresh:

JFreshHockey is creating NHL analytics-based visualization cards. | Patreon

Byron Bader:

Hockey Prospecting – Uncovering Tomorrow's Superstars

EliteProspects:

Elite Prospects - Hockey Players, Stats and Transactions

WHKYHAC:

Data (whkyhac.com)

Harry Shomer:

GitHub - HarryShomer/Hockey-Scraper: Python Package for scraping NHL Play-by-Play and Shift data