

## The case for xGoals

The advanced stat community has recently gotten fixated on Corsi and Fenwick as a measure of team performance and as a predictor of future success. The argument is that a team is much more likely to win if they control the puck more than the other team. This is theoretically a solid assumption, because after all you can't score if you don't control the puck; however this fails in a number of ways. The most obvious way is that possession stats do not take into account the quality of scoring chances.

If a team plays a collapsing, defensive style of hockey the large majority of the shots will come from the outside, where the defending goalie has a longer time to react to and generally has a better chance at saving the puck.

The New York rangers under Tortorella in 2011/2012 were notoriously defensive and collapsed back to the net and had a "shot-blocking" mentality. This style of hockey produced the franchise's third ever 50 win season for an Eastern conference best 51-24-7 for 109 points.

The Rangers also made it all the way to the conference finals.

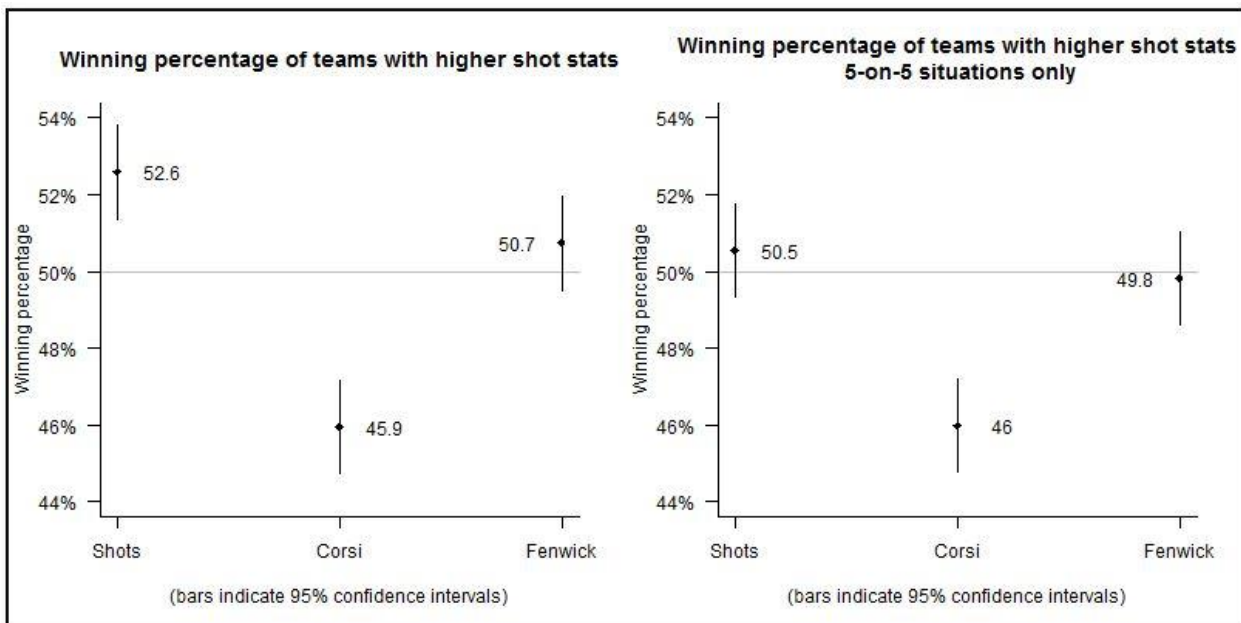
Their Corsi number? 47.8 during the regular season, 48.6 for the post season.

Fenwick? A slightly more palatable 49.4 for both regular season and post-season

### **The Rangers were 23<sup>rd</sup> in the entire league in Corsi in 2011-12**

This is where the Corsi and Fenwick argument get dicey. The common explanation from hockey analysts would be that this performance was due to "luck" and that it was "unsustainable" over time. I personally hate the use of the word "luck" in statistical analysis. Using the word "luck" is simply a shitty excuse for not being able to analyze the data properly, and as we'll see the Rangers had a lot more than just "luck".

### **As a matter of fact, Corsi and Fenwick seem to be a surprisingly bad indicator of win percentage.**



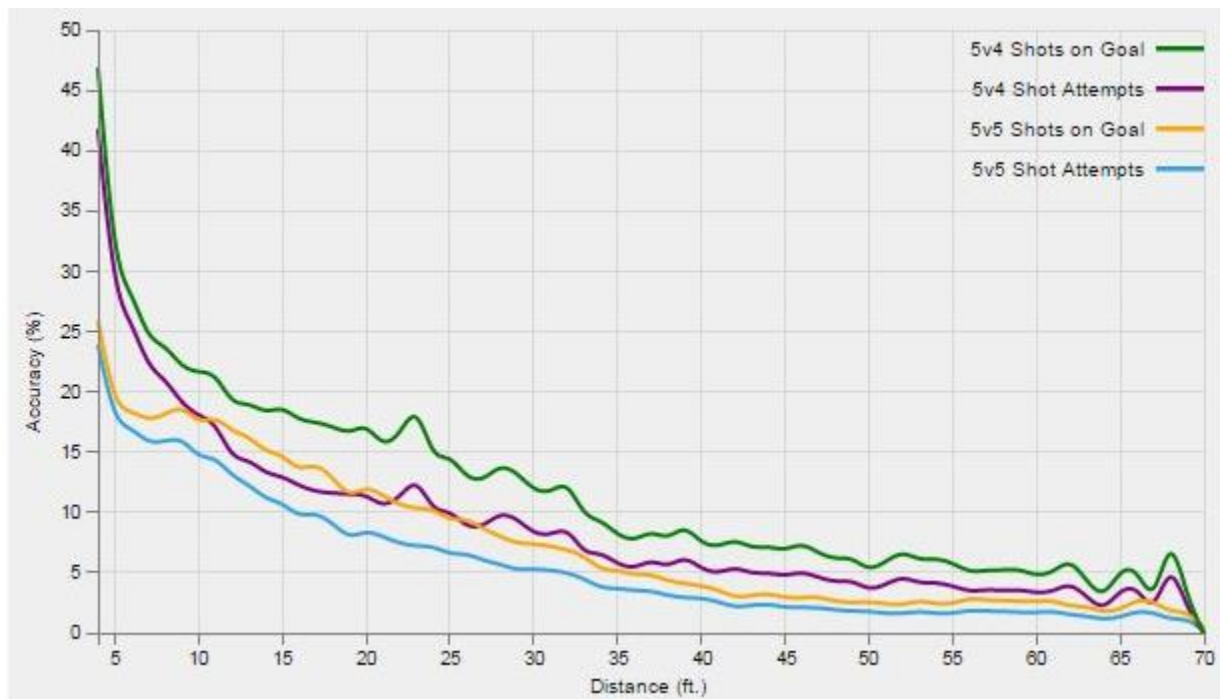
Scouring the internet, I found this very interesting graph, showing the correlation between win percentage and Fenwick and Corsi written by Stephen Pettigrew at <http://rinkstats.blogspot.ca/2013/12/why-popular-advanced-stats-are-bad-at.html>

It actually shows a **negative** correlation between Corsi and win percentage! While there are some flaws in Stephen's analysis (namely Score effect among others, read here for an explanation on score-effect: <http://nhlnumbers.com/2013/12/5/score-effects-and-you>) this makes a very interesting point: maybe these possession stats aren't the "money-balling of hockey" people seem to think they are.

A defensive, patient, counter-attacking style of hockey has existed for years and has always been thought as "outliers" in the possession stats obsessed hockey world, however there are still benefits to a counter attacking style like this. There is the old adage that good defense leads to offense, and in a lot of instances this is very true. When a team has a lot of offensive zone time, it naturally starts pressing and closing in on the opposing net to get to better scoring areas. Teams will activate their D men to try and penetrate the collapsing defense exposing lanes. When this happens, fast counter-attacking team are able to take advantage to create odd man rushes.

This essentially turns the game of hockey into: playing solid defense, limiting inside shots from the slot, and once the defensive team creates a turnover, it's a foot race the other way. These odd-man rushes also often lead to better quality scoring chances.

### **So how do we quantify "quality" scoring chances?**



Using data collected from the NHL since 2007-2008, xGoals measured the relationship between distance from the net and shooting percentage.

In the above graph, you can see that the data confirms what is intuitive: shooting closer to the net gives you a better chance to score. The difference in shooting percentage between shooting 10 ft away from the net and 40 ft away is an impressive 12.5% difference. **7 years of tangible NHL data tells us that you are 6 times more likely to score from 10 ft away than from 40ft.** This is the basis on which xGoals was created. By measuring the distance at which a player takes a shot, xGoals records the shooting percentage from the shot distance.

For example a shot from 40 ft gives around 2.5% chance of scoring a goal; this is equal to an xGoal score of 0.025. Statistics tell us that this particular player will have to shoot 40 times from 40ft out to score 1 goal.

So: if player x shoots 40 times from 40ft he gets 1 xGoal.

**From this basic idea, we are able to predict the amount of goals every single player on every single team will get based on historical shooting percentage.**

**So this includes an added factor to the shots for and against conversation, we also include shooting percentage**

From the amount of goals every player is predicted to get, we can get the **total** amount of goals a team is expected to get giving us a much, much more accurate metric of both performance and predictor of future success: xDiff.

xDiff is simply the amount of goals your team is expected to get **based on shot location (and subsequently shooting percentage)** over the amount of goals your team is expected to let in **based on shot location**. You can think of xDiff as Corsi or Fenwick but adding shooting percentage as an extra factor of accuracy.

So we can see that instead of measuring a team's performance by the amount of time they control the puck, we can measure it by the amount of goals they are predicted to score and the amount of goals they are predicted to let in.

This next part is the obligatory : **NOTHING IS PERFECT AND THERES NO WAY OF ACTUALLY PREDICTING OVECHKIN GETTING 43 GOALS IN 2013**

xGoals will not be able to say exactly how many goals a player will get, however it is a much better indicator of performance than how many shots are directed at the net over how many shots are directed at your own net. **xGoal factors in the actual amount of goals teams will score and not just controlling the puck** which coincidentally is how you win hockey games!

Oh and that Rangers 2011/12 team? **They were 8<sup>th</sup> overall in xDiff**: a massive improvement on their 23<sup>rd</sup> ranked Corsi rating, and a much more plausible explanation for their success that season.

To conclude:

**xDiff > Fenwick > Corsi**

***Check all of this out at : [www.fourthlineheroes.com](http://www.fourthlineheroes.com) and make sure to read the explanations for all the added goodies, including predicted team goals for in power play situations, predicted goals against in penalty killing situations, even predicted save percentage to give an extra measure of goaltender performance when faced with tough shots.***

**Summary of other goodies:**

**GAE: Goals above expected.**

-This measures the deviation from the goals a player actually scores to the amount of xGoals predicted. Think of this as the higher the deviation, the better the player is at scoring from far away. There are two explanations for this, one over a small sample size, one over a large sample size

Small sample: Overachievers. These are players that are getting timely goals from far away and are most likely not going to sustain their scoring pace without getting shots from closer in.

Large Sample: You are an elite sniper. The most interesting part of xGoals is seeing the difference between a player like Ryan Smyth who created an entire career of scoring from right in front of the net and a sniper like Stamkos. Smyth's expected xGoals was much higher than his actual goal total: this makes sense since most of his shots came in high shooting percentage areas. The guys who score more than expected are guys who are very good at shooting the puck. Since 2007, the guys with the highest GAE are:

-Steven Stamkos

-Alex Ovechkin

-Ilya Kovalchuk

-Jarome Iginla

-Phil Kessel

The trend? These guys are what is referred to as "elite snipers". So while xGoals is a very good predictor of goals in the general hockey player population, over a large enough sample size the GAE can actually tell you about the playing style of a player.

Low GAE? Brenden Gallagher, Ryan Smyth: Get the dirty goals type of player

High GAE? Steven Stamkos: One-Timer from thirty feet away type of player.

For the rest? xGoals is actually pretty good at predicting. Case in point? Corey Perry

Since 2007, xGoals for Perry = 218.084

Actual Goals for Correy Perry = 220

xSV%

Since we can predict goals for players, we can totally do it for goalies!

Expected save percentage tells us the how many goals are expected to be scored based again on shooting percentage from distance. A good goalie will make the tough saves in tight, while an overrated goalie will only have to make easy saves from a ways out. A very good case study in this is the recent tear of Devan Dubnyk.

Since 2007 for goalies with 1500 shots faced minimum, Devan Dubnyk ranks second in difference between expected SV% and actual SV%. Basically he is expected to have a 0.914 over that time span, however his actual SV% is a hot 0.922, meaning that he is saving a significant amount of shots he shouldn't be saving. His recent trade to Minnesota, (also a favourite of xGoal's with the top xDiff in the league this year) means that he no longer needs to be making the incredible saves he was asked to make in Edmonton. In short, Devan Dubnyk is seriously underrated and it's finally starting to show how good he actually is (for the record Cory Schneider is number 1 in Change%). Sorry Edmonton.