

- [Boxing](#)
- [Horse Racing](#)
- [Podcasts](#)
- [Local](#)
- [Running](#)
- [Water Polo](#)
- [Wrestling](#)
- [Outsports](#)
- [Video](#)
- [Odds](#)

• [All 308 blogs](#)



Follow



 1.9M

• [Search](#)



- [NHL](#)
- [Outnumbered](#)

## Quantifying the added importance of recent data

By [Eric T.](#)  [@BSH\\_EricT](#) on Jan 21 2014, 5:03p [11](#)



Do you think Phoenix looked back more than a year before deciding to give Smith \$34M for his age 31 to 36 seasons? - Marianne Helm

Past history shouldn't affect our projections as much as the most recent results, but it's not meaningless either. Let's build a model for how we think the importance drops off with time.

[Tweet Share on Twitter \(42\)](#) [Share Share on Facebook \(2\)](#) [Share Share on Google Plus 11 Comments](#) • [Rec Recommend this Post 2](#)

On Monday, I noted that one major divide in how people evaluate the NHL is [how much weight they give to recent results](#).

It almost certainly has to be true that the recent results matter more than the older ones, but how much more? If you're trying to guess how a goalie will do this year, is last year's performance 10 percent more important than the year before that, twice as important, or ten times as important?

One way to try to answer this is a direct analysis of how things have turned out for goalies in recent years, how their eventual performance compared to their most recently completed seasons.

More [Outnumbered](#)





[• Making the clutch save](#) • [Is Crosby hard to play with?](#) • [How does the NHL decide a suspension?](#) • [Should teams pay goalies big money?](#) • [Advanced stats proponents are not idiots](#)

In essence, this goes as follows: we imagine ourselves back in the summer of 2012 (or 2011 or 2010 or whatever), trying to predict each goalie's future save percentage using only his past numbers. We look at what system of weights on each of the previous years would have come up with the answer that most closely matches the actual outcomes.

## The model

There are a lot of details that go into this.

When you look at the past performance, do you look back three years or five or to the beginning of the goalie's career? When you predict the future, are you trying to guess next year or the next three years or to the end of the goalie's career? When you calculate how the predictions for a given system did, are you counting all goalies equally or paying more attention to the ones who had reasonably long histories?

Your exact answer to the weighting will depend on how you answer those questions. Here's what I did:

- I studied each goalie with NHL experience in each off-season from July 2005 to July 2013. That gave me 775 individual player projections to examine.
- I looked back to the goalies' previous four years for most of my work, but also tested three years or five.
- I looked ahead to the next three years for most of my work, but also tested just the coming season.
- In assessing a system of weights, I paid the most attention to projections where players had faced a lot of shots in their past *and* would go on to face a lot of shots in the future. (For those who might want to replicate this work, I weighted each player's outcome by the harmonic mean of the two numbers.)
- I normalized each year's performance to the league average, so I'm really asking "how far above average will his save percentage be" rather than "what will his save percentage be?"

So in my base case, I'm using years 1-4 to try to predict years 5-7. The best predictions came from weighting things like this:

- Each shot faced in year 3 counts 60 percent as much as shots in year 4
- Each shot faced in year 2 counts 50 percent as much as shots in year 4
- Each shot faced in year 1 counts 30 percent as much as shots in year 4

I think this passes the sniff test. I would've expected the drop-off to be less steep, personally -- before doing this, I normally would've just counted all four years equally. But I certainly don't look at this result and think "no way could that be true".

## Pressure testing

Since some of the decisions on how to construct this model were arbitrary, it's good practice to change those decisions a little and make sure it doesn't dramatically change the result.

If I look back three years instead of four, I get a reasonably similar answer. Then the result is that the previous year counts 50 percent as much as last year and the year before that counts 40 percent as much. If I look back five years, I get a 100-60-50-20-20 weighting.

If I just try to predict the coming season instead of the next three, I get a 100-70-50-10 weighting.

So everything's coming up more or less in the same ballpark -- last year is nearly as important to consider as the two previous years combined, and performance before that is a meaningful but modest adjustment.

### What does it mean?

The result is that recent performance matters a bit more than older data, but the older data is still quite significant.

Steve Mason has performed above his career average the last two years, so his projection comes in at .908 -- better than his career rate of .906, but less than the Flyers are hoping for. Conversely, Devan Dubnyk's off year puts his projection at .907, a little below his career .909 rate.

Ilya Bryzgalov's poor three-year run puts his projection at .909, significantly below his .913 career average. Jonathan Bernier's projection of .923 is considerably better than his career .918 rate.

The next step will be to combine the results of this weighting system with the Bayesian approach I [previously described](#). Then we'd be factoring in both the likelihood that recent results matter more than older ones and the likelihood that random chance played a role in the observed outcomes.

### More from [SB Nation NHL](#):

Follow @SBNationNHL

5,001 followers

Like

503

- [Canadiens struggles aren't going away without change](#)
- [Ryan Suter's incredible, record-setting workload](#)
- [Potty Mouths: The SB Nation hockey blog profanity rankings](#)
- [Photo: Here's Dodger Stadium with an ice rink](#)
- **Olympics:** [Injuries might alter Canada's roster](#) | [USA roster analysis](#)
- [Olympic rosters shouldn't be built with NHL restraints](#)

Please upgrade your browser to view HTML 5 content

[Tweet](#) [Share on Twitter \(42\)](#) [Share](#) [Share on Facebook \(2\)](#) [Share](#) [Share on Google Plus 11](#) [Comments](#) • [Rec](#) [Recommend this Post 2](#)

## Latest News

- [OutKast goes for 404 points in one game](#)
- [Spring ball is over. Now what?](#)
- [The athlete advocate](#)
- [CHART: LMA's scoring explosion in context](#)