Journal of Quantitative Analysis in Sports

Volume 7, *Issue* 3 | 2011 | *Article* 24

Estimating NHL Scoring Rates

Samuel E. Buttrey, *Naval Postgraduate School*
Alan R. Washburn, *Naval Postgraduate School*
Wilson L. Price, *Universite Laval*

Recommended Citation:
Buttrey, Samuel E.; Washburn, Alan R.; and Price, Wilson L. (2011) "Estimating NHL Scoring Rates," *Journal of Quantitative Analysis in Sports*: Vol. 7: Iss. 3, Article 24.
Available at: http://www.bepress.com/jqas/vol7/iss3/24
DOI: 10.2202/1559-0410.1334

Estimating NHL Scoring Rates

Samuel E. Buttrey, Alan R. Washburn, and Wilson L. Price

Abstract

We propose a model to estimate the rates at which NHL teams score and yield goals. In the model, goals occur as if from a Poisson process whose rate depends on the two teams playing, the home-ice advantage, and the manpower (power-play, short-handed) situation. Data on all the games from the 2008-2009 season was downloaded and processed into a form suitable for the analysis. The model seems to perform adequately in prediction and should be useful for handicapping and for informing the decision as to when to pull the goalie.

KEYWORDS: Poisson process, GLM, hockey

1. INTRODUCTION

In the final game of the 2009 Stanley Cup playoff series, with just over a minute remaining, Detroit Red Wings' head coach Babcock called a timeout and pulled goalie Osgood in favor of an extra attacker. The Red Wings hurled their powerful man-advantage attack at the Pittsburgh Penguins, pressing into their zone, cycling and looking for an opening. A whistle set up a face-off near the Penguins' goal with 6.5 seconds left. The Wings won the draw and Zetterberg got a solid shot away, but Fleury made a save for the Pens. The rebound went to Lidstrom, who let go a likely winner, but Fleury made another save. The Wings battered the defensive box as the puck bounded to the right of the net, but time expired before the Wings could make another shot. If there had been a few more seconds on the clock, would the Wings have scored? Did Detroit pull its goalie soon enough? Could Operations Research have brought the cup back to the Motor City?

The question of when to pull the goalie has been looked at before. We have nothing new to contribute to the methodology for computing the best time to pull the goalie, and will use dynamic programming method of Washburn (1991) to do so when one is needed. Our main contribution here is to introduce a statistical method that will tailor the needed parameters to the particular teams that are playing. This same methodology will also enable the construction of a scale on which hockey teams can be ranked. We will do so, again using statistics from the 2008-2009 season (the Red Wings will end up at the top).

Hockey is a fluid game where puck possession changes frequently, and where scoring can happen at any time. We take the point of view that scoring for each team is well modeled as a Poisson process (Breiman, 1986), and our first object will be to realistically estimate the rate parameters (the "scoring rates") for those processes in a given game. Statistically speaking, the winner of a hockey game will be the team whose Poisson random variable (number of goals scored) is larger than the other's.

In considering questions related to goalie-pulling, the scoring rates will have subscripts that correspond to various game states that the National Hockey League (NHL) records as each game progresses. Principal among these is the number of players on the ice, which varies with the imposition of penalties. The identities of the participating players are also recorded, but we will not use them. Puck possession is not part of the recorded state, mainly because puck possession is not as clearly defined in hockey as it is in (say) American football.

The usual statistical dilemma is present here: we want as much data as possible, but on the other hand we don't want to go back so far in time that the nature of a team essentially changes. Our solution will be to use all data from the current season, but none from previous seasons. Our justification for this is that a team's roster can fluctuate significantly between seasons, but not within a season.

This will make our estimates statistically unreliable at the beginning of a hockey season, but at their best at the end of the season, as in the introductory vignette.

The rest of the paper is organized as follows: section 2 describes the basic model and describes the different manpower situations in hockey. Accounting for these differing situations is important for estimating scoring rates and, eventually, for addressing the question of the optimal time to pull the goalie. Section 3 describes the data acquisition, processing, and analysis task. Section 4 gives the details of the model, together with the results and evaluation of the estimation. Finally, section 5 offers some conclusions and describes our directions for future work.

2. The Basic Model and Team Ranking

2.1: Scoring Rates

The most direct method of estimating scoring rates would be to simply divide the number of goals scored by the amount of time available for scoring them. Since there are 30 teams in the NHL, this method would estimate 30×29=870 parameters λ_{ij} from a single season's data, each one being the rate at which team *i* scores goals on team *j*. There would, for example, be a rate at which the Wings scores goals on the Pens. There is not enough data in a single season to reliably estimate that many parameters — many pairs of teams play only one game per season. We therefore make the reasonable structural assumption that λ_{ij} is a product of factors that include an offensive factor for team *i* and a defensive factor for team *j*, in addition to a factor for which team is at home. (A tiny number of NHL games – two of the 1,230 in our data – are played at neutral sites, in which cases we use the NHL's designation of the "home" team.) Specifically, we assume that the scoring rate of team *i* against team *j*, for $i \neq j$, is

$$\lambda_{ij} = \lambda_0 A_i B_j D\,, \tag{0.1}$$

where λ_0 is the base scoring rate, A_i is the offensive factor, B_j is the defensive factor and the factor D depends on whether team *i* is at home or away. Putting aside the home team advantage, in any matchup the team more likely to win is the one with the higher scoring rate, so the crucial question is whether $A_i B_j$ is larger than $A_j B_i$. This is equivalent to whether A_i / B_i is larger than A_j / B_j, so the ratio of each team's offensive factor to its defensive factor is the desired method of ranking teams.

We choose to deal with (1.1) in its logarithmic form, which is a sum or four terms instead of a product of four factors. In that form we have

$$L_{ij} = \mu + \alpha_i + \beta_j + \gamma\,, \tag{0.2}$$

DOI: 10.2202/1559-0410.1334

where $\mu \equiv \ln(\lambda_0)$, etc. Since equation (1.2) is a sum, we can arbitrarily set $\alpha_1 = \beta_1 = 0$, and $\gamma = 0$ when team i is away. This leaves 60 parameters to be estimated. In this form the teams can be ranked by the difference $\alpha_i - \beta_i$.

2.2: Adjusting For "Power Plays"

We downloaded and analyzed every National Hockey League (NHL) box score from the 2008-2009 season in order to compute team-specific scoring rates. (Section 3, below, gives some details on acquiring and processing the data.) The ordinary estimate (goals divided by games) feels naïve because no adjustment is made for the "power play," which takes place after some penalties when a player is temporarily ejected and his team plays with one fewer player for the duration of the penalty. These power-play situations obviously make it easier for the team at full strength to score. By computing the lengths of all power-play situations we were able to compute improved scoring rates for each team for each manpower situation. Under a suitable model these rates can be employed to compute the probability of one team beating another. Our model performs reasonably well when using twenty days' worth of data to predict the next day's games but the real contribution of this work may be the construction of this goals-by-manpower data set, which we have not seen elsewhere. This procedure also serves as an instructive example of data collection in the internet age.

2.2. The Types and Effects of Penalties in Hockey

When a hockey game begins, each team has six players on the ice. One is the goalie; the other five are collectively referred to as "skaters." So in the usual terminology the game starts in the "five-on-five" manpower situation. There are five sorts of penalties in hockey. "Minor" penalties are the most common sort, and are assessed for two minutes, during which the offender (or in some cases a proxy) is removed from the ice. The penalized team then plays with four skaters until the penalty's two-minute duration expires, or until the other team scores a goal. (We will use "expire" to refer to either of these two outcomes.) However, if two penalties are called simultaneously, both players are removed but the manpower situation remains unchanged (unless other penalties are also declared) because new players are substituted. In general, when the two teams have different numbers of skaters, the team with more is said to be on the "power play," and the team with fewer, to be "short-handed." Otherwise the two teams are said to be "at even strength," even if the manpower situation is, for example, four-on-four. If a goal is scored on a team with an existing (non-simultaneous) minor penalty, the penalty (the earliest minor penalty, if more than one is in effect) is ended.

A "major" penalty lasts five minutes and is not ended when a goal is scored. "Misconduct," "game misconduct" and "match" penalties do not result in a change in the manpower situation and so these are ignored here.

Complications in keeping track of the manpower situation ensue in two situations. A team may not have fewer than three skaters, so if a team with three skaters commits a penalty the penalty is "stacked" – that is, it is not imposed until an existing penalty expires. A second complication occurs during a "double minor." This is a pair of minor penalties imposed on the same player. The first of these is imposed immediately (or, if necessary stacked); the second is imposed when the first expires. So the second part of a double minor acts like a stacked penalty, even when the penalized team has four skaters, except that it is only imposed when the first part expires, not when an earlier minor expires.

A hockey game consists of three twenty-minute periods. A penalty imposed late in one period will extend into the next if necessary. In the NHL's regular season, if a game is tied at the end of sixty minutes, an overtime period lasting up to five minutes is played. In the case where no penalties are in effect, the overtime period starts in the with four-on-four condition. (A team with penalties in effect can have three skaters; if two penalties are in effect for a team, then the other team starts the overtime with five skaters.) Such a period is ended immediately by a goal. We treat overtime play in the same way as play in regulation time. A game that is still tied after the five-minute overtime goes to a shootout; we essentially ignore shootouts for the purposes of this paper. In the playoffs, there are no shootouts: overtime periods in the playoffs are played at full strength, last twenty minutes each, and continue until the first goal is scored.

3. Data Analysis

3.1: Acquiring and Processing The Data

3.1.1. Events and Intervals

Our goal in this research was to compute the number of minutes that each team played against each opponent in each manpower situation. We define an "event" as an action in the game which changes (or has the potential to change) the manpower situation. We also include the beginning and the end of the game as events. A full-strength goal, or a goal scored against a team that is short-handed as a result of only major penalties, does not change the manpower situation, but we will call these "events" as well. Events that can affect manpower include

DOI: 10.2202/1559-0410.1334

penalties, the expiration of penalties, and goals scored on a team that is short-handed as a result of at least one minor penalty.

Box scores list goals and penalties, but not expirations, so we had to deduce these ourselves. Our parser, then, converted a partial list of events into a list of "intervals." An interval is a set of playing time during which there is no change in manpower, together with the number of skaters on each team and the number of goals scored by each team during that interval (of course, these could both be zero). By computing every interval for every game we were able to model goal-scoring (and goal-yielding) rates for each team for each manpower situation.

3.1.2. The Data And the Parser

Data for this effort came from the website `sports.yahoo.com`. We were able to acquire all the game identifiers by automatically examining the "scores and schedule" pages, and then read and dump all the box scores using a text-based web browser. (Some care had to be taken to separate regular-season from pre-season games, and the all-star game also needed to be excluded.) As an example, the boxscore from the Vancouver/San Jose game of January 20, 2009 is found at `http://sports.yahoo.com/nhl/boxscore? gid= 2009012018`. (The final two digits of the address identify the home team, mostly alphabetically by city name, so that Boston is 01, Buffalo 02 and so on, down to Washington, which is 23; then fairly recent additions Phoenix, Anaheim, Florida, Nashville, Atlanta, Columbus, and Minnesota have numbers 24 through 30.) In a small number of cases the box scores were internally inconsistent and we verified their contents by comparing them to the corresponding ones at `CBS Sports.com`. The boxscore for the Vancouver/San Jose game of January 20, 2009 can be found at`http://www.cbssports.com/nhl/gamecenter/boxscore/NHL_20090120_VAN@SJ`. Other researchers will find useful similar documents at `nhl.com` as well.

In text form each of the Yahoo box scores is about 25,000 bytes, and, because each of thirty teams plays 82 games, there are 1,230 regular-season box scores, all in a consistent format.

We parse the box scores by means of a number of functions in S-Plus (Insightful Corp., 2005). One function converts the text of the box score to a set of lists giving times and teams responsible for penalties and goals, in ascending time order. (Although it is conventional to refer to times within periods, we convert times to global ones in decimal minutes, so that "5:30 of the third period" becomes 45.5). When a goal and penalty take place at the same time, the goal must have come first.

The current status of the game is maintained in a list with five components. One is the complete list of goals and penalties, in time order, produced by the parser. The set of penalties also includes an indicator of whether each is a major, minor, or part of a double-minor. At this stage misconduct and paired major penalties have been discarded.

A second component is the list of intervals generated up to the current time. The set of penalties currently in effect, if any, makes up a third component. This list includes the offender, the starting time of the penalty, its scheduled ending time, and an indicator as to whether it is stacked (in which case the starting time shows the time at which it was called, not the time at which it is imposed). Other parts of this list are the names of the teams, with an indication of which is the home team, and the way the game ended: regulation, overtime, or shootout.

3.1.3 The Parsing Algorithm

Once the game's text has been parsed, a second function runs through the set of goals and penalties implementing this general algorithm:

1.) Read an event from the list of penalties and goals.
2.) Check for expiration. That is, before processing this event, find out whether any penalties have expired since the most recent event before this one was processed. (If this event is the first of the game excluding the "game start" event, of course, no penalty can have expired between the start of the game and now.) If a penalty has expired between the time of the most recent event and the time of the current one, adjust the manpower situation at the time of the penalty's expiration. This step requires a loop, because there could have been stacked penalties (or the second half of a double minor) that needed to be imposed between then and now, and those newly imposed penalties may themselves have expired by the current time. A "start new interval" function creates new intervals and adds them to the end of the existing list of intervals.
3a.) If the current event is a goal, score the goal by incrementing the scoring team's goal counter in the current interval. Now check for expiration again, since if the team scored on is short-handed it may (or may not) get an additional skater. The check for expiration step may therefore bring about the creation of a new interval.
3b.) If the current event is a penalty, impose it. That is, we add this penalty to the list of penalties currently in effect. If the penalized team already has only three skaters, the penalty is added to the list of penalties, together with an indication that it is to be stacked. Otherwise we will decrement the number of skaters on the offender's team and start a new interval.

DOI: 10.2202/1559-0410.1334

3c.) The final event is the end of the game. If a game ends on an overtime goal we will have processed this event already. Otherwise we need to check for expiration one final time, since one or more penalties may have expired between the most recent event and the end of the game.

After applying this algorithm to every box score, we have a list of intervals for every game. These intervals are then combined into a single interval for each unique manpower situation encountered in the game. There may be as many as nine of these (since each team can have three, four or five skaters) or as few as one, in a game with no penalties. (In the 2008-2009 season there were two games which exhibited all nine possibilities, but there was no game with no penalties.) Each interval identifies the visiting and home teams, the numbers of skaters, the total time spent in that manpower configuration, and the number of goals scored by each of the teams during that configuration. It is also convenient to convert each of these records into two, one giving the number of goals scored by the home team and a second giving the number scored by the visitors; this is the format expected by standard statistics packages.

3.2. Overall Goal-Scoring Rates

Our final data set has 10,864 rows. Since (as we discuss below) our model assumes that goals arise like events from a Poisson process, it is reasonable to compute the league-wide scoring rate, by manpower situation. Table 1 shows the number of minutes spent in each situation (rounded to the nearest minute), the number of goals scored in that situation, and that situation's scoring rate, expressed in goals per sixty minutes. These numbers of goals are lower than the official totals for two reasons. First, they exclude empty-net goals – goals scored against a team which has removed its goalie in favor of an extra skater. Goals scored by the team removing its goalie are included. Second, a team that wins a shootout is awarded a goal in the official totals and those goals are excluded here as well. Of course the number of minutes teams collectively spent in 5-on-4 is exactly equal to the number spent in 4-on-5, and likewise for other i, j pairs with $i \neq j$. "Type" refers to even-strength (ES), power-play (PP), or short-handed (SH).

Manpower	Type	Goals	Time (Min.)	Rate
5.on.5	ES	4,508	113,957	2.37
4.on.4	ES	143	2,811	3.05
3.on.3	ES	11	186	3.55
5.on.4	PP	1,671	15,321	6.54
4.on.3	PP	55	506	6.52
5.on.3	PP	190	601	18.97
4.on.5	SH	208	15,321	0.81
3.on.5	SH	2	601	0.20
3.on.4	SH	11	506	1.30

Table 1: Number And Rate Per Sixty Minutes of Adjusted Goals By Manpower Situation, for All NHL teams in the 2008-2009 Regular Season

4. A MODEL FOR SCORING RATES AND GAME OUTCOMES

4.1. The Assumptions

In this section we specify our model. Under the assumptions of the model (and some other assumptions) we can then compute the probabilities of the different possible outcomes of any game. We assume that goals are scored according to a Poisson process whose rate depends on the teams in question and also on the manpower situation. Each team has two parameters, one for offense (denoted by, for example, α_{DET} for Detroit) and one for defense (denoted by, for example, β_{PIT} for Pittsburgh). There is, furthermore, one parameter common to all teams for each manpower situation (denoted by, for example, γ45 for the 4-on-5, with γ_{55} set to the baseline value of zero) and a single parameter (γ_H) common to all teams denoting the size of the home-ice advantage. The intercept is denoted by μ. We model the logs of the scoring rates by sums of these parameters.

So, for example, in a game between Detroit and Pittsburgh, with Detroit the home team, and with both teams at full strength, we would expect Detroit to score goals at a rate whose log is $\mu + \alpha_{DET} + \beta_{PIT} + \gamma_H$ and Pittsburgh to score goals at a rate whose log is $\mu + \alpha_{PIT} + \beta_{DET}$. When Detroit goes on the 5-on-4 power play, we expect the log of its rate to be $\mu + \alpha_{DET} + \beta_{PIT} + \gamma_{54} + \gamma_{H,}$ and the log of Pittsburgh's to be $\mu + \alpha_{PIT} + \beta_{DET} + \gamma_{45}$.

This model is, of course, naïve. It presumes that the rates are constant across all the minutes of every game, regardless of time of season, choice of goalie, injuries, or anything else. It presumes that all teams have the same home-

DOI: 10.2202/1559-0410.1334

ice advantage and the same advantage (disadvantage), in a multiplicative sense, from the power play (shorthandedness). All minutes in a particular situation are interchangeable under the model, and Detroit's offensive strength and Pittsburgh's defensive strength are taken to be additive on the log scale. Because of our success in producing this data, researchers now have the ability to examine lots of different models. However, we have not seen much improvement from more complicated models, and our model produces interesting results, as we see below.

To recap, the number of goals Detroit scores against Pittsburgh at 5-on-5 (call it $X_{\text{DET,PIT,5-on-5}}$) is, under the model, a Poisson random variable which depends on the parameters and also on the number of minutes during which the two teams were in that situation ($n_{\text{DET,PIT,5-on-5}}$). It is usual to think of λ as a rate per hour (since a game often lasts an hour) and n as a number of minutes. Under our model, then,

$$X_{\text{DET,PIT,5-on-5}} \sim \text{Poisson}\,(\lambda_{\text{DET,PIT,5-on-5}} \times n_{\text{DET,PIT,5-on-5}}/60)$$

and

$$\log\,(\lambda_{\text{DET,PIT,5-on-5}}) = \mu + \alpha_{\text{DET}} + \beta_{\text{PIT}} + \gamma_{\text{H}}$$

so that

$$\log\,(\text{E}(X_{\text{DET,PIT,5-on-5}})) = \mu + \alpha_{\text{DET}} + \beta_{\text{PIT}} + \gamma_{\text{H}} + \log\,(n_{\text{DET,PIT,5-on-5}}/60)$$

and similarly for all other situations. This is then a Poisson GLM (McCullagh and Nelder, 1989) and estimation of the parameters is straightforward. We use the statistical packages R (R Development Core Team, 2010) and S-Plus. The critical point is that we need to require that the coefficient of the log duration – that is, the coefficient of the last term on the right side – be 1; the `offset()` command in those two languages accomplishes this. This model produces 68 parameters: 29 for the offense of the thirty teams, with Anaheim (which comes first alphabetically) set to be the baseline with value 0; 29 for defense, eight for the nine manpower situations, and one for home-ice advantage, plus an intercept.

4.2. The Coefficients

Table 2 shows the offensive and defensive coefficients sorted from best (positive for offense, negative for defense). It also shows the actual numbers of (adjusted) goals scored by, and allowed by, each team, and the league rankings among those numbers. (Non-integer rankings reflect ties.) The rankings, not surprisingly, show general agreement between the coefficients of our model and the actual numbers of goals scored and allowed. Differences are attributable to differences in the numbers of penalty minutes accrued by the teams.

The coefficients in Table 2 can be interpreted as follows. When Detroit is on offense the log of the scoring rate is greater than that of Anaheim by 0.126. Since exp(0.126) = 1.13, we conclude that Detroit scores goals at a 13% higher rate than Anaheim does, other things being equal. Boston's defensive coefficient is –.178, and exp(–.178) = .837. So Boston gives up goals at 83.7% the rate of Anaheim. Table 3 shows the remaining parameters, together with their exponentiated values. We see that, according to the model, teams in a 5-on-4 score goals at a rate that is 175% greater than at 5-on-5, and teams in a 4-on-5 score at a rate only 34% of the full-strength rate. (The numbers in table 3 differ from the overall ones in table 1 since these former adjust for the team strength parameters).

4.3. Computing Probabilities of Victory

Predicting the outcome of a game is difficult, in part because the distribution of penalties is unpredictable. For example, penalties often come in bunches so they should not be treated as independent. Conditioning on the number and location of penalty minutes in a game would make for better estimates but seems unrealistic. For this model, we restrict ourselves to predicting the outcome of a game in which there are no penalties at all (which is not very different from the case in which penalties are charged equally).

Suppose that Detroit and Pittsburgh play and that their goal-scoring parameters are respectively λ_D and λ_P (for sixty minutes). We observe the two Poisson random variables X_D and X_P, denoting the numbers of goals, and we compute $X_\Delta = X_D - X_P$. Then Detroit will win if $X_\Delta > 0$, and Pittsburgh will win if $X_\Delta < 0$. If $X_\Delta = 0$, regulation time will end in a tie; we handle this case later.

The difference between two independent Poisson random variables is known to follow a Skellam distribution (Haight, 1967). The probability that X_Δ takes on value x is

$$S(x; \lambda_D, \lambda_P) = \Pr(X_\Delta = x \mid \lambda_D, \lambda_P) = \exp(-(\lambda_D + \lambda_P))\ (\lambda_D/\lambda_P)^{x/2}\ I_x(2\sqrt{[\lambda_D\lambda_P]}),$$

where $I_x()$ denotes the modified Bessel function of the first kind of order x. This distribution is not directly provided in R but is easy to compute because of that package's computation of the Bessel functions. The probability that Detroit wins in regulation is then just $\sum_{x>0} S(x)$. If a playoff game ends in a tie in regulation, play continues until a goal is scored. We envision that goals arise from one of two independent Poisson processes generating goals at rates λ_D and λ_P per sixty minutes, respectively. In this case the game ends if the Detroit process produces its first goal before the Pittsburgh process produces one, an event with probability $\lambda_D/(\lambda_D + \lambda_P)$. Finally, therefore, the probability that Detroit wins a playoff game is $\sum_{x>0} S(x) + S(0)\,\lambda_D/(\lambda_D + \lambda_P)$, versus Pittburgh's probability of $\sum_{x<0} S(x) + S(0)\,\lambda_P/(\lambda_D + \lambda_P)$.

DOI: 10.2202/1559-0410.1334

Team	Off (α)	GF	Rank	Team	Def (β)	GA	Rank
Detroit	0.126	280	1	Boston	–0.178	185	1
Boston	0.065	258	2	Minnesota	–0.121	191	2
Washington	0.058	254	3	New Jersey	–0.109	199	4
Philadelphia	0.045	249	4.5	NY Rangers	–0.098	207	6.5
Chicago	0.015	247	6	San Jose	–0.085	198	3
Pittsburgh	0.011	249	4.5	Chicago	–0.076	203	5
Toronto	0.009	242	9.5	Vancouver	–0.068	207	6.5
Anaheim	0.000	235	13	Ottawa	–0.052	218	13.5
Calgary	–0.001	243	7.5	Nashville	–0.051	215	10.5
Atlanta	–0.006	242	9.5	Florida	–0.042	212	9
San Jose	–0.023	243	7.5	Carolina	–0.040	210	8
New Jersey	–0.038	233	15	Columbus	–0.040	215	10.5
Buffalo	–0.039	241	11	Los Angeles	–0.039	217	12
Vancouver	–0.052	235	13	St. Louis	–0.024	218	13.5
Montreal	–0.058	235	13	Philadelphia	–0.015	229	18
St. Louis	–0.093	220	18	Buffalo	–0.010	222	15
Florida	–0.095	222	17	Anaheim	0.000	226	17
Carolina	–0.099	229	16	Pittsburgh	0.002	225	16
Edmonton	–0.108	218	20	Washington	0.029	235	19.5
Dallas	–0.112	219	19	Montreal	0.048	241	23
Columbus	–0.115	215	21	Calgary	0.070	240	22
Ottawa	–0.159	209	22	Detroit	0.087	235	19.5
Tampa Bay	–0.169	205	24	Edmonton	0.092	239	21
Minnesota	–0.170	207	23	Tampa Bay	0.100	259	27
Nashville	–0.176	201	25	Colorado	0.114	245	26
NY	–0.201	193	28.5	Dallas	0.115	244	25
Phoenix	–0.213	196	26	Phoenix	0.132	242	24
Colorado	–0.227	189	30	NY Islanders	0.154	267	28
NY Rangers	–0.245	195	27	Atlanta	0.194	272	29
Los Angeles	–0.250	193	28.5	Toronto	0.249	281	30

Table 2: Offensive and Defensive Coefficients, Plus Adjusted Goals For (GF) and Against (GA) And League GF/GA Rankings

Effect	Coefficient	Exp
(Intercept)	–3.216	0.040
5.on.5	0.000	1.000
4.on.4	0.249	1.282
3.on.3	0.409	1.505
5.on.4	1.012	2.752
4.on.3	0.989	2.688
5.on.3	2.070	7.924
4.on.5	–1.070	0.343
3.on.5	–2.476	0.084
3.on.4	–0.605	0.546
Home	0.084	1.088

Table 3: Other Coefficients and Their Exponentiated Values

For the regular season, a game goes to a shootout after five minutes of overtime. Detroit will win in overtime if its first goal precedes Pittsburgh's and is also within the first five minutes. This probability is

$$P_5(\lambda_D, \lambda_P) = [\lambda_D/(\lambda_D + \lambda_P)]\,(1 - \exp(-5(\lambda_D + \lambda_P)).$$

Pittsburgh will win in overtime with probability $P_5(\lambda_P, \lambda_D)$, and the game will go to a shootout with probability $\exp(-5(\lambda_D + \lambda_P))$. We assign the probability of winning a shootout to be 0.5 for each team, which is not unreasonable in light of the shootout data. (There is also no evidence of a home-ice advantage in the shootout.) So Detroit's overall probability of winning a regular season game, according to the model, is

$\sum_{x>0} S(x) + S(0)\, P_5(\lambda_D, \lambda_P) + 0.5 \exp(-5(\lambda_D + \lambda_P))$, while Pittsburgh's is

$\sum_{x<0} S(x) + S(0)\, P_5(\lambda_P, \lambda_D) + 0.5 \exp(-5(\lambda_D + \lambda_P))$.

Now we can use the coefficient values to get a numeric value for this estimated probability. The log of Detroit's (per-minute) scoring rate at home against Pittsburgh is $\mu + \alpha_{DET} + \beta_{PIT} + \gamma_H = -3.216 + 0.126 + 0.002 + .084 = -3.004$. Therefore Detroit's (sixty-minute) rate is 60 exp (–3.004) = 2.975. Meanwhile Pittsburgh's scoring rate has log equal to $\mu + \alpha_{PIT} + \beta_{DET} = -3.216 + 0.011 + .087 = -3.118$, leading to a sixty-minute scoring rate of 2.654. Using the computations above, we compute Detroit's predicted probability of winning a home, regular-season game in regulation to be .467, in overtime to be .034, and in a shootout to be .053, for a total of .554. Meanwhile Pittsburgh's probability of winning is .446 (.362 in regulation, .030 in overtime, and .053 in a shootout, after rounding). When Boston plays at home against the Islanders (to pick teams with

DOI: 10.2202/1559-0410.1334

very different scoring rates), we estimate a .772 probability of Boston winning, versus .228 for the Islanders.

Teams in hockey are ranked by points, with two points being awarded for each win (whether in regulation, overtime or by shootout), and one point being awarded for a loss in overtime or shootout. In our example above, the expected number of points awarded to Detroit is $2 \times .554 + (.030 + .053) = 1.19$, versus Pittsburgh's $2 \times .446 + (.034 + .053) = .979$.

4.4. Evaluating the Model

We evaluated this model's fit by comparing its predictions to the actual regular-season outcomes. We selected a cutoff date of February 27, 2009. In the 2008-2009 season, 922 games, roughly 62 games played by each team, were played up to and including that date.

Our predictions came in two forms, "static" and "dynamic." In static prediction, we update the model so that the coefficients reflect all the data up to and including February 27th and then predict the outcomes of all the subsequent games. In dynamic prediction, we updated the model after each day's games, so that on day d the coefficients reflect outcomes up to (but not including) d. We then used the resulting model to predict only the outcomes on day d.

In both cases we used data starting either at the beginning of the season, or (to reduce the effect of games played long ago) starting at day $d - n$, where the "days back" parameter n was set 10, 20 and 30. (It was convenient to count days because hockey teams play on different days. Notice that when n is 20, for example, models are built on the basis of 19 days.) We then tallied the numbers of wins and expected points for all teams across the set of days representing the final 308 games of the 2008-2009 season.

One drawback of this prediction scheme is that it uses only the overall (even-strength) scoring rates. Knowledge of a game's penalty structure should almost certainly lead to better predictions. Therefore, as a test of the model, we also performed prediction when the numbers of durations of the manpower situations were taken to be known. That is, we computed the expected numbers of goals for each team in a particular game, using the actual numbers of minutes of 5-on-4, 4-on-5 and so on situations observed in that game. (Obviously in practice these numbers are not known until the game is over.) We call these results "conditional" since they refer to numbers of goals scored, conditional on the number of minutes in each manpower situation.

Table 4 shows the correlation between the actual number of points observed by each team over the final 308 games and the numbers predicted under the different parameter settings.

	Dynamic		Static	
Days back	**Regular**	**Conditional**	**Regular**	**Conditional**
10 days	.682	.752	.314	.334
20 days	.688	.750	.298	.315
30 days	.636	.688	.283	.303
All season	.403	.425	.456	.475

Table 4: Correlations Between Actual and Predicted Points Under Different Settings

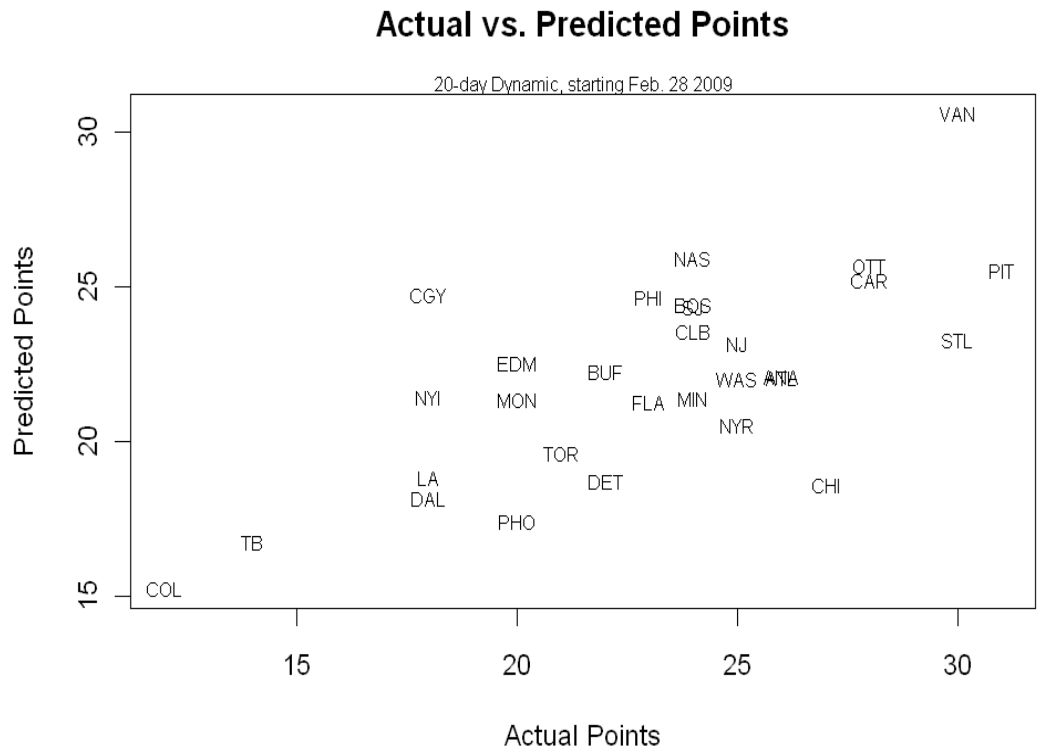
The correlation between the predicted and actual numbers of points is unmistakable, if not overwhelming. It is no surprise that the "conditional" correlations are uniformly better than their regular counterparts, although the differences are not always large. Perhaps more surprising is the difference in quality between the static and dynamic predictions; the static ones perform much worse in the situations where they have only a few days, but better when a lot of data is available. This suggests that (among other shortcomings in the model) the goal-scoring and goal-preventing rates may be changing fairly quickly.

By inspection of the table we have selected 20 days as the best number of days to use at the time the parameters are estimated. Figure 1 shows the plot of the actual number of points acquired by each team (noted by its two- or three-letter abbreviation) versus the number of points predicted with the regular, dynamic, 20-day-back model.

5. Conclusions

We have constructed a data set containing all the intervals, by manpower situation, for every NHL game in the 2008-2009 season. This might serve as an instructive exercise in collecting data from the internet. A simple probabilistic model for hockey outcomes is described and implemented in standard statistical software. Making predictions after about ¾ of the season, we find that dynamic predictions perform better than static ones, and that using data from the last twenty days improves on using the whole season's worth. Conditional predictions – ones that use knowledge of the number of penalties – perform better than unconditional ones, which demonstrates that the knowledge of penalty times can be important, but these predictions cannot be used in practice.

DOI: 10.2202/1559-0410.1334



In future work we hope to improve these models. It would be instructive to examine the rate at which the estimated goal-scoring and goal-preventing parameters change. We might consider looking for overdispersion in the GLM or other evidence of "streakiness." These estimates of scoring rates could be put to use handicapping games, or used in conjunction with Washburn's (1991) work to determine the optimal time at which one team should pull its goalie when playing another specific team.

References

Breiman, Leo (1986), *Probability and Stochastic Processes* (2nd Ed.), Palo Alto, CA, The Scientific Press

Haight, Frank A. (1967), *Handbook of the Poisson Distribution*, New York: John Wiley and Sons.

Insightful Corp. (2005), *S-Plus 7 for Windows User's Guide*, Seattle, WA, Insightful Corporation.

McCullagh, P. and Nelder, J.A. (1989), *Generalized Linear Models* (2nd Ed), London: Chapman and Hall.

R Development Core Team (2010), R: A Language And Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL `http://www.R-project.org`.

Washburn, A. (1991) "Still More on Pulling the Goalie," Interfaces Vol. 21, No. 2, March-April 1991, pp. 59-64

DOI: 10.2202/1559-0410.1334