

Hockey Analytics

Tim B. Swartz *

Abstract

This paper provides a review of some of the key research topics in hockey analytics.

Keywords: Player evaluation, Match simulation, National Hockey League.

*Tim Swartz is Professor, Department of Statistics and Actuarial Science, Simon Fraser University, 8888 University Drive, Burnaby BC, Canada V5A1S6. Swartz has been partially supported by grants from the Natural Sciences and Engineering Research Council of Canada.

1 INTRODUCTION

There are various ways of identifying and ranking the “big” sports of the world. For example, participation is an obvious measure (Bialik 2012). Using financial criteria, one may consider variables such as the number of professional teams and leagues, player salaries, ticket prices, television viewership, team valuations, sponsorships, etc. Using financial criteria, there is no doubt that the six big team sports in the world are soccer, basketball, baseball, American football, hockey and cricket.

In the team sports listed above, there is immense pressure to perform at a high level. To help perform at a high level, analytics have infiltrated the sports world. Analytics problems in sport are vast and include important problems such as player evaluation and scouting, the management of salaries, player development, modification of tactics, training to improve performance, and the treatment and prevention of injuries.

Clearly, analytics in sport are an expensive business, and team sports have adopted analytics to various degrees. The sport with the greatest history in analytics is baseball where substantive contributions were made by Bill James with annual publications of his *Baseball Abstract* beginning in 1977. Gray (2006) provides a biography of James and his ideas. Baseball is particularly conducive to analytics as it may be described as a “discrete” game where only a countable number of outcomes can occur as the result of each pitch. Baseball also has an extensive history of data and record keeping which has been beneficial to the growth of baseball analytics. Baseball analytics has a large following. For example, the Society for American Baseball Research (www.sabr.org) promotes conferences and publications, and boasts a membership of over 6000 individuals interested in the nuances of the sport. It may be argued that baseball analytics (often referred to as sabermetrics) provided the inspiration for other sports to follow suit. The inspiration came through the widespread attention provided to the book “Moneyball” (Lewis 2003) which was developed into the popular 2011 Hollywood movie starring Brad Pitt. Moneyball chronicled the 2002 season of the Oakland Athletics, a small-market Major League Baseball team who sought efficiency gains through the acquisition of undervalued players.

In the academic literature, there has been longstanding activity in sports science which is concerned with topics such as exercise, health, medicine, physiology and psychology. In addition to sports science, there are other academic interests involving sport, and there now exist sports journals with a focus on topics such as economics, operations research, engineering and computer science. On the statistical side, the *Journal of Quantitative Analysis in Sports* (JQAS) became a

publication of the American Statistical Association in 2011, and the *Journal of Sports Analytics* was founded in 2015. [Albert and Bennett \(2014\)](#) provide a general review of statistics in sports.

However, despite the presence of academic journals having a focus on sport, there may be an even greater activity in industry where the activities and effort have been intentionally suppressed. Many teams in the big sports now have full-time analytics staff where their primary objective is to gain a competitive edge. These personnel are often prevented by their clubs to disclose the nature of their work.

Now, where does hockey fit in the above discussion? In terms of publications in peer reviewed journals, it seems that hockey analytics is middling amongst the big six sports. For example, in the five-year period 2012-2016, JQAS published 25, 18, 16, 10, 9, and 4 papers on basketball, soccer, American football, hockey, baseball and cricket, respectively. And we note that the number of baseball papers may not reflect the actual activity in the sport due to the availability of baseball-specific outlets (e.g. *Baseball Research Journal*).

[In terms of professional hockey, the National Hockey League \(NHL\) is the premier hockey league in the world consisting of 30 teams and many feeder leagues from Canada and the United States. The salaries in the NHL are the highest amongst professional leagues with a \\$73 million salary cap in the 2016-2017 season. Other top professional leagues exist in predominantly northern European countries including Sweden, Switzerland, Germany, Czech Republic and Finland. Along with Russia, these are also the countries where participation rates are highest and whose national teams play at the highest level. McFarlane \(1997\) provides a history of hockey.](#)

In terms of analytics staff in the NHL, the situation in 2010 indicated a general antipathy in hockey towards analytics. However, with the advent of Moneyball, the influence of analytics in other sports and the availability of data, the state of analytics has begun to change in the NHL (<http://www.sloansportsconference.com/content/changing-on-the-fly-the-state-of-advanced-analytics-in-the-nhl/>). For example, most NHL teams now have analytics staff. Some teams have even recruited academic statisticians to lead these efforts (e.g. Andrew Thomas - Minnesota Wild, Sam Ventura - Pittsburgh Penguins, Brian Macdonald - Florida Panthers).

In addition, conferences specializing in hockey analytics are now regular events. These include Edmonton (2014), Calgary (2014), Pittsburg (2014), Washington (2015), Fort Lauderdale (2016), Rochester (2015, 2016), Vancouver (2016, 2017) and Ottawa (2015, 2016, 2017). There are now perhaps more conferences than exist hockey research in need of dissemination.

This paper attempts to capture the current state of the rapidly changing world of hockey analytics. The emphasis is on problems that are more of a complex nature, the types of problems

that would be attractive to statisticians. There are lots of hockey analytics problems whose solutions are straightforward. For example, a team may want to know the proportion of faceoffs that are won by a particular player. As useful as this information may be to teams, we do not consider straightforward research investigations in this review paper. For a lively discussion of all types of research problems in hockey (with an emphasis on the non-technical), the book by Vollman, Awad and Fyffe (2016) is recommended. In the spirit of Bill James, Vollman has also self-published an annual *Hockey Abstract* beginning in 2013. In Section 2, we describe data that are available for hockey analytics. This, like everything in hockey analytics is currently in a state of flux. In Section 3, we describe the Holy Grail of analytics problems, player evaluation. A number of approaches have been proposed, and the main techniques are described. The topic of player evaluation is still very much an open area for research as all approaches suffer from multicollinearity issues. That is, players tend to play with common teammates and it is difficult to separate their respective contributions. In Section 4, the major problem of match simulation is described. The availability of a simulator allows the investigation of various questions of interest in a laboratory setting. The challenge involves the development of a realistic simulator that captures the characteristics of actual hockey games. In Section 5, some miscellaneous topics are explored. We conclude with a short discussion in Section 6.

2 DATA

In the early days of the NHL, and today in lower leagues, so-called “statisticians” would attend games and participate in data entry activities. Depending on the needs of their team, they might record events such as shots, saves, hits, zone entries and the results of face-offs. As the game modernized and matches were recorded, statisticians would watch film after the game, and could produce even more statistics.

Since the 1980’s (Kasan 2008), NHL data has been generated via the NHL’s Real Time Scoring System (RTSS). The procedures and the data that have been collected have evolved over the years. Today, at every NHL game, there are a crew of scorers for the home team who view matches and make decisions in real time. The data are uploaded to nhl.com and can be freely accessed. The data include events recorded in a play-by-play format. Although there are uniform standards that have been imposed on the crews, there has been some criticism over the accuracy of the RTSS data (<http://statsportsconsulting.com/2014/11/24/1357/>).

Notwithstanding the data integrity issues, Thomas and Ventura (2014) have provided a great

service by making the RTSS data easily accessible. They have created an R package *nhlscrapr* that provides detailed event information and processing for NHL games. The scraper retrieves play-by-play game data from the NHL’s RTSS database and stores the data in convenient files that can be handled by the R programming language. A typical match includes roughly 400 events per game which corresponds to an event roughly every 9 seconds. The *nhlscrapr* package can access NHL matches back to the 2002-2003 regular season.

The promise of player-tracking data in the NHL has been a much discussed topic amongst those involved in hockey analytics. A similar initiative has already taken place in the National Basketball Association (NBA) where the SportVU system has been in place since the 2013/2014 season. The NBA data has promoted a surge in research activities including previously difficult topics of investigation such as the evaluation of contributions to defense (Franks, Miller, Bornn and Goldsberry 2015).

Although some experimentation has taken place, as of the 2016-2017 season, player-tracking has not yet been fully implemented in the NHL in the sense that it is not freely available to everyone. Amongst the competing technologies, the company Sportvision has developed an approach that requires chips in both the puck and in player jerseys. An alternative technology promoted by the company SPORTLOGiQ is based on a single camera in each arena, machine learning and optical recognition software. There is great detail and accuracy in the SPORTLOGiQ database with events occurring every 1.2 seconds on average. The data also records the x and y coordinates for every player on the ice for each event and every frame; this is the player-tracking aspect of the data.

Similar to the NBA, we expect a surge of research activity in hockey analytics once player-tracking data becomes widely available. **With massive datasets expected, data mining techniques (Hand and Adams 2015) will play a larger role.**

3 PLAYER EVALUATION

The ultimate goal of any professional hockey team is to win, and assembling a “good” team is a component of winning. Although the makings of a good team in the NHL depend on many factors including team synergy, team depth and salary cap constraints, there is no doubt that the evaluation of individual talent is an important task in building a good team.

Of course, coaches and general managers typically have a strong sense of the contributions of players. However, sometimes there may be subtleties about player contributions that remain un-

detected. Therefore, objective measures of player evaluation form part of the evaluation process.

The *plus-minus statistic* has a long history for assessing player contributions in hockey. It is a simple and common statistic that purports to measure player impact. Excluding power-plays, the plus/minus statistic for a player is defined as the number of goals that were scored by the player's team while he was on the ice minus the number of goals that were scored against his team while he was on the ice. On powerplays, goals are only counted in situations when the short-handed team scores. The interpretation of the plus-minus statistic is that the larger the value, the greater the contribution of the player.

There are a number of difficulties with the plus-minus statistic. First, it has difficulty distinguishing between players who frequently play together. When a goal is scored, the plus-minus statistics for these players are adjusted equally (i.e. ± 1) even though the actual player contributions may be quite different. Second, the plus-minus statistic is dependent on ice-time. For example, if a player's ice-time is doubled, then his plus-minus will roughly double if his performance remains the same. As an example of how the plus-minus statistic can give misleading results, consider the case of Alex Ovechkin. In the 2013-2014 NHL regular season, Ovechkin's plus-minus statistic of -35 ranked him 785 out of 787 NHL players despite being selected to the second All-Star team and scoring the most goals (51) in the league. Part of the mystery is resolved by noting that Ovechkin played the most minutes of any of his teammates on the powerplay and rarely played during the penaltykill during the 2013-2014 season.

In response to the difficulties with the plus-minus statistic, many statistics have been proposed to help address player evaluation. We consider approaches that are grounded in statistical theory with an associated statistical model. Consequently, we will not discuss some popular statistics such as GVT (goals versus threshold) which rely on various component pieces with various weighting factors (Awad 2009). The statistics which we review are based on regression type procedures. Therefore, we first consider models of the form

$$y_i = x_i' \beta + \epsilon_i \tag{1}$$

where y_i is the i -th measurement on a continuous scale that describes the quality of the i -th event with respect to the home team. The covariate x_i is a known vector, β is an unknown parameter vector and ϵ_i is the corresponding random error. The key feature of (1) is that x_i contains an indicator variable for every player in the league such that a specific player's variable is set to 1/-1/0 according to whether he was a home team player who was on the ice when the event

occurred, whether he was a road team player who was on the ice when the event occurred, or whether he was not on the ice when the event occurred. It follows that the component of the β vector that corresponds to a player's indicator variable is a measure of his quality. Compared to plus-minus, the regression approach provides a partial effect which controls for the contributions of linemates and opponents.

Schuckers and Curro (2013) is one of several procedures for player evaluation based on (1). See also Macdonald (2012). Schuckers and Curro (2013) consider various events (e.g. shots, hits, takeaways, etc.) and estimate the probability that a goal arises within a 20-second window of the event. Therefore, events have value and the values determine the response variable y . For the covariate x , in addition to the indicator variables for the players, the covariate vector includes a home team indicator to account for the home team advantage and a zone start variable to account for the advantage of beginning a shift in the offensive zone. After fitting the model using ridge regression, and transforming the estimated player characteristic $\hat{\beta}_j$, they introduce an interpretable statistic referred to as THoR (Total Hockey Rating). The rankings which arise from the catchy acronym THoR appear to correspond to general intuition.

When the data y_i are bernoulli (i.e. $p_i = P(y_i = 1)$), logistic regression models have been proposed which take the form

$$\log\left(\frac{p_i}{1-p_i}\right) = x_i'\beta \tag{2}$$

where x_i and β have the same structure as in (1).

In the logistic regression context (2), Gramacy, Taddy and Tian (2017) considered goals (either for or against the home team) as the dependent variable y . For the covariate x , in addition to the indicator variables for the players, they also specified a home team effect, team-season effects, manpower effects, playoff effects and interaction terms. Their estimation methods were based on regularization which involves penalty terms in a classical framework. They also used these penalty terms to carry out full Bayesian analyses. Under regularization, many players are estimated as having no effect. This essentially reduces the parametrization of the problem and permits more accurate estimation of the remaining extreme players (i.e. those who are really good and those who are really bad). A drawback with the approach is that teams do not score many goals (roughly 5.5 total goals per match). Consequently, there is a sparsity in the dataset which is their motivation for using multiple seasons of data. We question the logic of the inclusion of the team-season effect which improves estimation. From our point of view, once the 10 skaters on

the ice are identified, there is no need for a team-season effect.

Although the various regression type procedures discussed above provide useful insight, they all suffer from the inability to distinguish between players who share most of their ice-time together.

3.1 Goaltender Evaluation

Although nothing prevents the analysis of goaltending using methods based on (1) or (2), such analyses ignore important information that is relevant to goaltending. And it is important to look at goaltending carefully because there exists a minority opinion that goaltenders in the NHL are essentially indistinguishable (Yost 2015). In fact, looking at simple save percentages and their associated standard errors suggests that most goalies are similar.

Schuckers (2017) provides a good review of goaltending statistics. In particular, the review emphasizes the need to distinguish between the difficulty of shots of various types (i.e. shot quality) and the need to assess the reliability of goaltending statistics. The latter is accomplished through observing correlations of goaltending statistics obtained from odd and even numbered shots.

Schuckers (2017) proposes the following general goaltender statistic:

$$\text{aSVP}_{\bar{S}} = \sum_{u=1}^U G(u)\bar{S}(u) \quad (3)$$

where there are U different shot types, the estimated probability (across the NHL) of facing a shot of type u is $\bar{S}(u)$ and the probability that the goaltender makes a save from a shot of type u is $G(u)$. The development of the $\text{aSVP}_{\bar{S}}$ statistic involves determining the shot types u , obtaining the league frequencies $\bar{S}(u)$ and the individual save probabilities $G(u)$. The determination of shot types is critical since they are potentially infinite, and the estimation of $G(u)$ is more difficult for large U . Shot type enumeration may involve many features including distance from the net, angle from the net, type of shot (e.g. slapshot, wristshot, backhand) and whether a shot is a consequence of a rebound. Schuckers (2017) outlines challenges and describes various approaches for the estimation of $G(u)$ including parametric and non-parametric methods.

The general statistic (3) strikes the author as a sound measure for distinguishing goaltenders. However, with the multitude of choices involved in the selection of shot types u and estimation procedures for $G(u)$, there appears to be potential for more development of $\text{aSVP}_{\bar{S}}$ statistics. Currently, the versions of $\text{aSVP}_{\bar{S}}$ as investigated by Schuckers (2017) do not provide reliability in

terms of strong season to season correlations.

4 MATCH SIMULATION

Simulation provides the capability to address questions of interest for which there may not be adequate data. In terms of sporting applications, simulation permits the investigation of new strategies without the risk of attempting a detrimental strategy in an actual game. The challenge involves the development of realistic simulators. [Asmussen \(2014\)](#) provides a review of stochastic simulation.

Suppose that a team has n possessions in a hockey match and that the probability of scoring on any possession is p . These beginning assumptions (although imperfect) suggest that the number of goals X scored by the team during the game may be modelled as a $\text{Binomial}(n, p)$ random variable. And with n large and p small (as is the case in hockey), it may be reasonable to assume that $X \sim \text{Poisson}(\theta)$ where $\theta = np$. The Poisson distribution and the related Exponential distribution (which models time until events) provide the natural starting points for hockey match simulators.

Marek, Sediva and Toupal (2014) fit Poisson-based models in the context of ice hockey data from the Czech Republic. Buttrey (2016) simulated NHL games in terms of goal scoring times and penalty durations using Exponential distributions. Parameters related to scoring and penalties were estimated by weighting recent matches more heavily. The utility of the simulator was demonstrated through betting when the simulated results disagree from sportsbook lines.

We provide a short review of the more sophisticated hockey match simulator developed by Thomas (2017). Rather than using boxscores, Thomas (2017) uses the detailed RTSS event data to capture greater realism with respect to the NHL. One of the first modifications introduced by Thomas (2017) is based on the recognition that Poisson scoring rates for both teams do not remain constant throughout a match. For example, the rates appear to depend on the current goal differential between the two teams. In particular, there is a higher probability of tied games than is predicted by independent Poissons. As a second example, Thomas (2017) observes that actual goal scoring times have a longer right tail in playoff overtime periods than during regular play when matches are tied. Thomas (2017) also provides the framework for modelling team (or even lineup) strength, and offensive and defensive capabilities.

Penalties can be also simulated exponentially with the rates depending on various scenarios including the home/road venue, the score and the time. The resulting manpower advantages

subsequently affects goal scoring rates.

The model can be extended beyond goals and penalties to include events as provided in the RTSS data. When doing so, added features can be provided such as increasing the probability of a goal in the near term following an offensive zone faceoff. The framework and the data are available to simulate matches with considerable realism.

5 MISCELLANEOUS TOPICS

5.1 Tactics

There are many aspects in hockey that are known to be beneficial such as skating fast, winning faceoffs (Schuckers, Pasquali and Curro 2012) and maintaining possession (Gentile 2014). However, we do not consider these to be tactics. Although you would like your team to skate fast, win faceoffs and maintain possession, as a coach, these are not practices over which you have direct control. If your team is not very good, you cannot simply tell your players to do these things.

We consider tactics as decisions that are under control of the coach. And because hockey is a game where it is difficult to maintain possession and dictate eventualities, there does not seem to be many opportunities for tactics.

There are formational and style issues that fall under tactics. For example, a coach decides how the puck ought to be brought out of their own zone, how many skaters should chase in the offensive zone, the extent to which teams should “pinch” at the blueline, etc. However, the value of decisions of this sort do not seem to have been carefully addressed through analytics. We consider these to be open problems.

One tactical area that has been explored is the decision to pull the goalie (Beaudoin and Swartz 2010). Historically, when a team has trailed late in a game, the team pulls its goalie in favour of an additional skater. The situation produces a tradeoff where the trailing team is more likely to score (with more skaters), and at the same time, the leading team is more likely to score (via an open net). Through simulation techniques and the analysis of scoring rates under different manpower situations, Beaudoin and Swartz (2010) argue that teams should pull their goalie with roughly three minutes remaining in a match when trailing by a single goal. Prior to the analysis, NHL coaches tended to pull their goalies with less than one minute remaining. It has been observed that there has been a recent trend to pull the goalie earlier (Davis and Lopez 2015).

Another opportunity for tactics is based on the tendencies of NHL officials. For example, Beaudoin, Schulte and Swartz (2016) analyzed NHL penalty calls and have noted that the road team, the team that is leading and the team with fewer penalties called against them are all more likely to have the next penalty called against them. Perhaps the recognition of referee biases can be used by teams to modify playing style and gain a competitive edge.

5.2 The NHL draft

Every year, the NHL has a draft where amateur players satisfying eligibility requirements are available for selection by NHL teams. The order in which teams select players is based primarily on the standings from the previous season where the poorer performing teams have earlier selections. The draft is therefore seen as a means of improving competitive balance.

The problem of draft selection is essentially a problem of player evaluation as discussed in Section 2. However, the draft problem is more difficult for two reasons: (1) the players that are under consideration are mostly young and will develop; therefore the problem of player evaluation involves the more complex problem of prediction, and (2) the players that are under consideration compete in different leagues; therefore league comparisons introduce additional uncertainty.

Tingling (2017) provides a review of many issues associated with the NHL draft. His primary thesis is that drafting well is difficult. He explores a number of topics including evaluation criteria for post-hoc draft analysis, the relative value of draft positions and the drafting prowess of teams and their scouts. Tingling (2017) also discusses the Central Scouting Service (CSS), the NHL's ranking system for draftees under consideration. The CSS provides baseline information to all teams, and its existence raises the question of the utility of team scouting.

5.3 Performance trajectories with age

Although this topic has been touched upon previously, its importance demands a separate section. Barring injuries, a typical developmental path for an athlete involves a period of improvement, followed by a period of peak performance followed by a period of decline. The reason why the recognition of these periods is so important for the NHL (and other leagues) is the combination of bad long term contracts and salary cap constraints can impair teams for many years.

Part of the dilemma facing teams is that established players want to be rewarded financially for their previous service whereas teams want to reward players for future performance. With young players, the degree to which they develop is unknown, and this causes uncertainty for

management. Further, with performance trajectory graphs (i.e. plots of player value versus age), there exists variability that ought to be taken into account; some players develop quicker/slower than average and some players decline quicker/slower than average. It is also worth asking whether a player's trajectory depends on his position? Another problem (previously mentioned) is the determination of value. For example, is it better to have an all-star level player for one year versus a solid contributing player for five years? Also, should the performance trajectory curve follow a parametric specification, and if so, what is the specification? Finally, a related statistical issue involves the data collection procedure. There exists a sampling bias; players who are no longer in the NHL and are not considered in a data analysis may be those players whose decline was rapid. Conversely, young players in the NHL and are considered in a data analysis may be those whose development was rapid.

With a focus on some of the challenges presented above, Brander, Eagan and Yeung (2014) provide a careful analysis of the performance trajectory problem. They conclude that the peak ages for forwards and defensemen are 26 years and 28 years, respectively. In addition, they argue that near peak performances for forwards and defencemen lasts for approximately 10 years and 12 years, respectively.

5.4 Pace of play

In sports, there are lots of questions of curiosity that do not directly affect the game itself. For example, who is the best player? Which is the best team? What is the probability of a given team winning the Stanley Cup? We have avoided such questions in this paper. However, there is one line of questioning that is a bit puzzling and we include it here as it may be related to tactics. We begin by asking what is meant by (i.e. how do we measure) playing at a fast pace? From this question follows many related questions. For example, does playing at a fast pace produce goals? Does playing at a fast pace increase goals against? Does playing with pace help teams win? Who are the pacy teams? Are there pacy players?

Silva, Davis and Swartz (2017) have attempted to tackle this problem using player-tracking data. With player tracking data, distance travelled lengthwise on the ice while in possession was defined as a measure of pace. Counter to their intuition, they found that this measure of pace did not correlate with statistics such as shots on goal by both teams or goals by both teams.

Pace is a topic that is of considerable importance in soccer. It is well known that strategies such as Jose Mourinho's "parking the bus" are meant to suppress pace and goal scoring. It is also known that goal scoring intensity in soccer increases as the game progresses, when players tire

and the game gets stretched. (Armatas, Yiannakos and Sileloglou 2007). Therefore, defining pace in an appropriate way and studying its consequences in hockey appears to be an open research problem.

6 DISCUSSION

Hockey is a difficult sport for analytics. Like soccer and basketball, it is a fluid game, a so-called “continuous” game where there are many moving pieces. Also, what happens “off the puck” is not typically measured; these difficult to detect actions can have a profound impact on the results of a game. Possession also makes hockey a difficult sport. Compared to soccer and basketball, it is more difficult to retain possession in hockey. When you cannot easily control possession, then it is more difficult to devise and implement offensive tactics.

On the other hand, data are more plentiful in hockey now than ever before, and more detailed data is on the imminent horizon. The new data will take the form of player-tracking data which will require data science skills on the part of analysts.

One way of looking at these difficulties and challenges is that there are opportunities in hockey analytics. The game is played in many countries, and hockey is one of the six wealthy big sports. This suggests that there should be future resources and interest for making meaningful contributions in hockey analytics.

7 REFERENCES

- Albert, J.H. and Bennett, J.M. (2014). Sports, Statistics in. *Wiley StatsRef: Statistics Reference Online*, DOI: 10.1002/9781118445112.stat00129.
- Armatas, V., Yiannakos, A. and Sileloglou, P. (2007). Relationship between time and goal scoring in soccer games: analysis of three World Cups. *International Journal of Performance Analysis in Sport*, 7(2), 48-58.
- Asmussen, S. (2014). Simulation of stochastic processes. *Wiley StatsRef: Statistics Reference Online*, DOI: 10.1002/9781118445112.stat04387.
- Awad, T. (2003). Understanding GVT, Part 3. *Hockey Prospectus, Numbers on Ice*, Accessed on February 6, 2017 at <http://www.hockeyprospectus.com/puck/article.php?articleid=236>
- Beaudoin, D., Schulte, O. and Swartz, T.B. (2016). Biased penalty calls in the National Hockey League. *The ASA Data Science Journal*, 9, 365-372.

- Beaudoin, D. and Swartz, T.B. (2010). Strategies for pulling the goalie in hockey. *The American Statistician*, 64, 197-204.
- Bialik, C. (2012). Figuring out sports participation. *Behind the Numbers, Wall Street Journal*, Online (June 8/12) at <http://blogs.wsj.com/numbers/figuring-out-sports-participation-1145/>.
- Buttrey, S.E. (2016). Beating the market betting on NHL hockey games. *Journal of Quantitative Analysis in Sports*, 12, 87-98.
- Brander, J.A., Egan, E.J. and Yeung, L. (2014). Estimating the effects of age on NHL player performance. *Journal of Quantitative Analysis in Sports*, 10, 241-259.
- Davis, N. and Lopez, M. (2015). NHL coaches are pulling goalies earlier than ever. Accessed on February 5, 2017 at <https://fivethirtyeight.com/features/nhl-coaches-are-pulling-goalies-earlier-than-ever/>
- Franks, A., Miller, A., Bornn, L. and Goldsberry, K. (2015). Counterpoints: advanced defensive metrics for NBA basketball. *Sloan Sports Analytics Conference, 2015*.
- Gentile, S. (2014). A reminder from Darryl Sutter: Don't mistake puck dominance for defense. Accessed on February 5, 2017 at <http://www.sportingnews.com/nhl/news/la-kings-news-darryl-sutter-quotes-corsi-for-definition-sidney-crosby/1ty2k6440ww8213spqckr4u09n>
- Gramacy, R., Taddy, M. and Tian, S. (2017). Hockey performance via regularized logistic regression. In *Handbook of Statistical Methods and Analyses in Sports*, J.H. Albert, M.E. Glickman, T.B. Swartz and R.H. Koning (editors), Chapman and Hall/CRC: Boca Raton, 287-306.
- Gray, S. (2006). *The Mind of Bill James: How a Complete Outsider Changed Baseball*, Doubleday, New York.
- Hand, D.J. and Adams, N.M. (2015). Data mining. *Wiley StatsRef: Statistics Reference Online*, DOI: 10.1002/9781118445112.stat06466.pub2.
- Kasan, S. (2008). Off-ice officials are a fourth team at every game. Accessed on February 5, 2017 at <https://www.nhl.com/news/off-ice-officials-are-a-fourth-team-at-every-game/c-388400>
- Lewis, M. (2003). *Moneyball: The Art of Winning an Unfair Game*, WW Norton, New York.
- Macdonald, B. (2012). Adjusted plus-minus for NHL players using ridge regression with goals, shots, Fenwick and Corsi. *Journal of Quantitative Analysis in Sports*, 8, Issue 3, Article 8.
- Marek, P., Sediva, B. and Toupal, T. (2014). Modelling and prediction of ice hockey match results. *Journal of Quantitative Analysis in Sports*, 10, 357-365.
- McFarlane, B. (1997). *Brian Mcfarlane's History of Hockey*, Sports Publishing Inc, Champaign IL.
- Schuckers, M. (2017). Statistical evaluation of ice hockey goaltending. In *Handbook of Statistical Methods and Analyses in Sports*, J.H. Albert, M.E. Glickman, T.B. Swartz and R.H. Koning (editors), Chapman and Hall/CRC: Boca Raton, 307-325.

- Schuckers, M. and Curro, J. (2013). Total hockey rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events. Proceeding of the 2013 MIT Sloan Sports Analytics Conference, Accessed on February 6, 2017 at <http://statsportsconsulting.com/thor/>
- Schuckers, M., Pasquali, T. and Curro, J. (2012). An analysis of NHL faceoffs. Accessed on February 5, 2017 at <http://statsportsconsulting.com/main/wp-content/uploads/FaceoffAnalysis12-12.pdf>
- Silva, R., Davis, J. and Swartz, T.B. (2017). The evaluation of pace of play in hockey. Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby BC, Canada.
- Thomas, A. (2017). Poisson/Exponential models for scoring in ice hockey. In *Handbook of Statistical Methods and Analyses in Sports*, J.H. Albert, M.E. Glickman, T.B. Swartz and R.H. Koning (editors), Chapman and Hall/CRC: Boca Raton, 271-285.
- Thomas, A.C. and Ventura, S.L. (2014). nhlsrapr: Compiling the NHL Real Time Scoring System Database for easy use in R. *R package version 1.8*, Accessed on February 6, 2017 at <http://CRAN.R-project.org/package=nhlsrapr>.
- Tingling, P. (2017). Educated guesswork - drafting in the National Hockey League. In *Handbook of Statistical Methods and Analyses in Sports*, J.H. Albert, M.E. Glickman, T.B. Swartz and R.H. Koning (editors), Chapman and Hall/CRC: Boca Raton, 327-339.
- Vollman, R., with T. Awad and I. Fyffe (2016). *Hockey Abstract Presents... Stat Shot: The Ultimate Guide to Hockey Analytics*. ECW Press: Toronto.
- Yost, T. (2015). Why analysis of goalies is so difficult. Accessed on February 6, 2017 at <http://www.tsn.ca/why-analysis-of-goalies-is-so-difficult-1.348286>

8 RELATED ARTICLES

Sports, Statistics in.