Player impact measures for scoring in ice hockey

Carles Sans Fuentes, Niklas Carlsson, and Patrick Lambrix

Linköping University, Sweden

Abstract

A commonly used method to evaluate player performance is to attribute values to the different actions that players perform and sum up these values every time a player performs these actions. In ice hockey, such metrics include the number of goals, assists, points, plus-minus statistics and recently Corsi and Fenwick. However, these metrics do not capture the context of player actions and the impact they have on the outcome of later actions. Therefore, recent works have introduced more advanced metrics that take into account the context of the actions and perform look-ahead. The use of look-ahead is particularly valuable in low-scoring sports such as ice hockey. In this paper, we first extend a recent approach based on reinforcement learning for measuring a player's impact on a team's scoring. Second, using NHL play-by-play data for several regular seasons, we analyze and compare these and other traditional measures of player impact. Third, we introduce notions of streaks and show that these may provide information about good players, but do not provide a good predictor for the impact that a player will have the next game. Finally, streaks are compared for different player categories, highlighting differences between player positions and correlations with player salaries.

1 Introduction

In the field of sports analytics, many works focus on evaluating the performance of players. A commonly used method to do this is to attribute values to the different actions that players perform and sum up these values every time a player performs these actions. These summary statistics can be computed over, for instance, games or seasons. In ice hockey, common summary metrics include the number of goals, assists, points (assists + goals) and the plus-minus statistics (+/-), in which 1 is added when the player is on the ice when the player's team scores (during even strength play) and 1 is subtracted when the opposing team scores (during even strength). More advanced measures are, for instance, Corsi (sum of shots on goals, missed shots and blocked shots) and Fenwick (sum of shots on goals and missed shots)[1].

However, these metrics do not capture the context of player actions and the impact they have on the outcome of later actions. To address this shortcoming and to capture the ripple effect of actions (where one action increases/decreases the success of a later action, for example), recent works [9, 11, 4] have therefore introduced more advanced metrics that take into account the context of the actions and perform look-ahead. The use of look-ahead is particularly valuable in low-scoring sports such as ice hockey.

In this paper we use an existing approach for measuring player performance (based on actions performed by a player) as well as extend the approach (based on actions when a player is on the ice). Further, we introduce time-normalized versions of the approaches. The background to the existing approach is given in Sect. 3 and the performance metrics are defined in Sect. 4. In Sect. 5 we analyze these two metrics in different ways using data from the NHL 2007-2008 and 2008-2009 seasons. First, we look at the top 10 players for these metrics in the two seasons, and discuss performance distributions. Then, we compare these metrics with the traditional metrics goals, points and +/- and discuss the relation to salary. Third, we introduce two notions of streaks and show that information about these kinds of streaks may give indications on who is a good player but not for whether this particular player will contribute more or less in a game than on average over a season. Finally, we compare and contrast the streak durations observed for different player categories (e.g., based on player position and salary) and tie our findings to those in prior parts of the paper.

[1] See, e.g., https://en.wikipedia.org/wiki/Analytics_(ice_hockey).

This paper appears in *Proceedings of MathSport International 2019 Conference*, pp 307-317, Athens, Greece, 2019

2 Related work

Many of the models for evaluating player performance attribute a value to the actions the player performs and then compute a sum over all those actions. For instance, the goal measure attributes a value to goal-scoring actions, while the assists measure attributes a value to passes that lead to goals. This is also true for some newer performance measures such as Fenwick and Corsi that attribute value to shots. Several newer performance measures extend some of the traditional measures. For instance, several regression models have been proposed for dealing with the weaknesses of the +/- measure (e.g., [6, 7, 1]). Further, in [2] principal component analysis was performed based on 18 traditional measures and a performance measure based on the 4 most important components was proposed.

Some of the approaches take game context into account. Added goal value [8] is a measure that attributes value to goals, but the value of the goal is dependent on the situation in which it is scored, thereby taking some context into account. Another measure for player evaluation based on the events that happen when a player is on the ice is proposed in [10]. Event impacts are based on the probability that the event leads to a goal (for or against) in the next 20 seconds. Other works model the dynamics of an ice hockey game using Markov games where two opposing sides (e.g., the home team and the away team) try to reach states in which they are rewarded (e.g., scoring a goal). In [13] the scoring rate for each team is modeled as a semi-Markov process, with hazard functions for each process that depend on the players on the ice. A Markov win probability model given the goal and manpower differential state at any point in a hockey game is proposed in [3]. In [9, 11, 12, 4] action-value Q-functions are learned with respect to different targets. (See Section 3 for the model in [9].) Although the approaches use Markov-based approaches, the definitions of states and reward functions are different. The advantages of such approaches (e.g., [12]) are the ability to capture game context (goals have different values in a tie game than in a game where a team is leading with many goals), the ability to look ahead and thereby assigning values to all actions in the game, and the possibility to define a player's impact through the player's actions. In this paper we base our work on one of these approaches.

3 Background

We base our work on an initial model presented in [9], where action-value Q-functions are learned with respect to the next goal. The state space considers *action events* with three parameters: the action type (Faceoff, Shot, Missed Shot, Blocked Shot, Takeaway, Giveaway, Hit, Goal), the team that performs the action (home, away), and the zone (offensive, neutral, defensive). A *play sequence* is defined as the empty sequence or a sequence of events for which the first event is a start marker, the possible next events are action events, and the possible last event is an end event. If the play sequence ends with an end event, it is a *complete* play sequence. The start/end events are Period Start, Period End, Early Intermission Start, Penalty, Stoppage, Shootout Completed, Game End, Game Off, and Early Intermission End. Actions and play sequences occur in a context. In [9] a *context state* contains values for 3 context features. Goal Differential is the number of home goals minus the number of away goals. Manpower Differential is the number of home players on the ice minus the number of away players on the ice. Further, the Period of the game is recorded. A *state* is then a pair which contains a context state and a (not necessarily complete) play sequence.

Actions are performed in specific states. For action a and state s=$< c, ps >$, where c is the context state and ps is the play sequence, the resulting state of performing a in state s is denoted by $s * a$ and is defined as $< c, ps * a >$, where $ps * a$ is the play sequence obtained by appending action a to ps. For states with play sequences that are end events, the next state is a state of the form $< c', \emptyset >$ where c' is defined by the end event. For instance, a goal will change the goal differential and update the context.

Table 1: Basic action sets.

A is the set of all state-action-pairs $< s, a >$ where action a is performed in state s
$A_i(p_k)$ is the set of state-action-pairs when player p_k is on the ice
$A_p(p_k)$ is the set of state-action-pairs where the action is performed by player p_k $A_p(p_k) \subseteq A_i(p_k)$

Table 2: Player and player pair impact.

The direct goal-based impact of a player is   the sum of the goal-based impact values of the actions performed by the player: DGB-impact(p_k) = $\Sigma_{<s,a>\in A_p(p_k)}$ impact(s,a)
The on-ice goal-based impact of a player is   the sum of the goal-based impact values of the actions when the player is on the ice: OIGB-impact(p_k) = $\Sigma_{<s,a>\in A_i(p_k)}$ impact(s,a)

Transition probabilities between different states are based on play-by-play data. The transition probability TP(s,s') for a transition from state s to state s' is defined as Occ(s,s') / Occ(s) where Occ(s) is the number of occurrences of s in the play-by-play data and Occ(s,s') is the number of occurrences of s that are immediately followed by s' in the play-by-play data. Using a state transition graph with the computed transition probabilities, Q-values for states are learned using a value iteration algorithm. The **goal-based impact of an action** a **in a state** s, impact(s,a), is then defined as $Q_T(s * a) - Q_T(s)$ where T is the team performing the action.

The performance of a player is computed as the sum of the goal-based impacts of the actions the player performs (over a game or a season). This is equivalent to comparing the actions taken by a specific player to the actions of an average player.

For our work, we re-implemented the code available from [9] using Python and R. The reward for goals for is +1 and goals against -1. (In the original implementation only +1 for goals is used.) The resulting goal-based impact values for the actions were used as a base for our work on player performance.

Recently, in [4] an updated model was introduced (which we have not used) where more events as well as more context features are taken into account. The Q-function represents the probability that a team scores the next goal. A neural net representing the Q-function was learned. It was shown that the updated player performance measure based on the updated goal-based impact measure is different from other measures such as +/-, expected goal, win-above-replacement and goal-above-replacement. It also correlates better than the other measures with many standard success measures such as goals, assists, points, shots, and face-off win percentage as well as with salary.

4 Player metrics

We introduce two basic measures for computing goal-based performance (similarly to [5]) as well as variants that normalize the measure with respect to time on ice.

First, we define different sets of actions[2] for players (Table 1). We differentiate between actions performed by a player and actions performed (by the player or another player) when a player is on the ice. In Table 2 we define the **direct goal-based impact of a player** (DGB-impact) based on the actions the player performs (and this is essentially the impact as defined in [9]). Further, we define the **on-ice goal-based impact of a player** (OIGB-impact) using the actions when the player is on the ice. This

[2]In the remainder we use action as a shorthand for action in a particular state.

Table 3: Top 10 players for 2007-2008 and 2008-2009 for the direct impact.

Player Name	Position	Age	Salary	GP	G	A	+/-	Points	Direct	Direct/h	On-ice	On-ice/h
2007-2008												
Alex Ovechkin	F	22	3.83	82	65	47	28	112	71.96	182.65	232.56	588.85
Dion Phaneuf	D	22	0.94	82	17	43	12	60	59.22	134.05	246.12	559.67
Rick Nash	F	23	5.50	80	38	31	3	69	59.01	181.80	158.82	485.99
Jarome Iginla	F	30	7.00	82	50	48	27	98	58.94	161.92	204.12	560.88
Dustin Brown	F	23	1.18	78	33	27	-13	60	53.78	156.41	171.40	501.48
Brenden Morrow	F	28	4.10	82	32	42	23	74	51.15	146.62	171.59	504.57
Zdeno Chara	D	30	7.50	77	17	34	14	51	50.74	117.69	203.78	468.89
Trent Hunter	F	27	1.55	82	12	29	-17	41	50.31	167.65	153.36	508.27
Mike Green	D	22	0.85	82	18	38	6	56	48.26	122.63	219.72	545.08
Pavel Datsyuk	F	29	6.70	82	31	66	41	97	48.22	134.68	198.44	559.41
2008-2009												
Alex Ovechkin	F	23	9.00	79	56	54	8	110	75.93	194.34	239.89	612.23
Dustin Brown	F	24	2.60	80	24	29	-15	53	59.76	177.60	178.34	540.84
Shea Weber	D	23	4.50	81	23	30	1	53	53.14	136.10	201.19	511.36
Evgeni Malkin	F	22	3.83	82	35	78	17	113	50.76	134.92	220.41	591.75
Dion Phaneuf	D	23	7.00	79	11	36	-11	47	50.34	122.64	240.57	532.49
Vincent Lecavalier	F	28	7.17	77	29	38	-9	67	49.46	143.99	188.17	549.37
Sheldon Souray	D	32	6.25	81	23	30	1	53	49.38	125.86	203.08	514.73
Jeff Carter	F	24	4.50	82	46	38	23	84	48.88	141.78	189.35	548.30
Rick Nash	F	24	6.50	78	40	39	11	79	48.88	145.11	171.59	498.26
Martin St. Louis	F	33	5.00	82	30	50	4	80	47.82	135.55	204.19	569.06

allows for a measure that includes indirect impact on the game by being on the ice. Even when players do not perform registered actions, they can still influence the game; e.g., by opening up a path for a teammate who may score. For both measures we also define variants normalized by time on ice (TOI). In this paper the normalization uses 1 hour of TOI.

5 Data-driven analysis

In this paper we use the play-by-play data for the NHL regular season games for the 2007-2008 and 2008-2009 seasons made available by [9]. Traditional performance metrics are gathered from `www.nhl.com` while salary information was taken from `www.dropyourgloves.com`. For the 2007-2008 and 2008-2009 seasons this resulted in information about 944 and 979 players, respectively.

5.1 Goal-based impact

In Tables 3 and 4 we show the top 10 players for the direct and on-ice goal-based impacts for the 2007-2008 and 2008-2009 seasons. For the on-ice impact we removed the goalkeepers as these are much longer on the ice than the other players and therefore collect more impact. For both the direct and on-ice impact measures we see that both defenders and forwards appear in the top-10 lists. This is similar to the +/- measure where 4 defenders were in the top 10 for each of the seasons. The goals and points measures, however, are heavily dominated by forwards. In 2007-2008 and 2008-2009 the best defenders regarding points held places 38 (Nicklas Lidström) and 32 (Mike Green), respectively, while the best defenders regarding goals held places 104 (Dustin Byfuglien) and 80 (Shea Weber), respectively. Similarly to the +/- measure, defenders show up in the top lists for the on-ice impact to a larger degree than for the direct impact. One reason is that the on-ice impact allows for indirect contributions. Another reason may be that, in general, defenders often play more than forwards.

Fig. 1 shows relative impact frequencies for 99.8 % of the values of the different player impact measures for the regular season games in seasons 2007-2008 and 2008-2009. For this figure we have

Table 4: Top 10 players for 2007-2008 and 2008-2009 for the on-ice impact (goalkeepers removed).

Player Name	Position	Age	Salary	GP	G	A	+/-	Points	Direct	Direct/h	On-ice	On-ice/h
2007												
Dion Phaneuf	D	22	0.94	82	17	43	12	60	59.22	134.05	246.12	559.67
Alex Ovechkin	F	22	3.83	82	65	47	28	112	71.96	182.65	232.56	588.85
Tomas Kaberle	D	29	4.25	82	8	45	-8	53	38.32	93.36	221.93	551.72
Mike Green	D	22	0.85	82	18	38	6	56	48.26	122.63	219.72	545.08
Andrei Markov	D	29	5.75	82	16	42	1	58	42.37	105.18	213.81	530.37
Nicklas Lidström	D	37	7.60	76	10	60	40	70	29.04	66.41	205.68	480.18
Jarome Iginla	F	30	7.00	82	50	48	27	98	58.94	161.92	204.12	560.88
Zdeno Chara	D	30	7.50	77	17	34	14	51	50.74	117.69	203.78	468.89
Lubomir Visnovsky	D	31	2.05	82	8	33	-18	41	32.64	83.52	201.34	523.00
Roman Hamrlik	D	33	5.50	77	5	21	7	26	37.79	93.89	201.29	509.39
2008												
Dion Phaneuf	D	23	7.00	79	11	36	-11	47	50.34	122.64	240.57	532.49
Alex Ovechkin	F	23	9.00	79	56	54	8	110	75.93	194.34	239.89	612.23
Evgeni Malkin	F	22	3.83	82	35	78	17	113	50.76	134.92	220.41	591.75
Dan Boyle	D	32	6.67	77	16	41	6	57	36.11	88.65	219.94	539.81
Chris Pronger	D	34	6.25	82	11	37	0	48	43.40	99.89	217.92	503.72
Mike Green	D	23	6.00	68	31	42	24	73	46.41	106.62	214.33	493.09
Nicklas Backström	F	21	2.40	82	22	66	16	88	37.12	111.83	214.19	630.43
Braydon Coburn	D	23	1.20	80	7	21	7	28	40.78	100.10	211.64	516.12
Andrei Markov	D	30	5.75	78	12	52	-2	64	38.03	96.17	209.18	527.62
Mark Streit	D	31	4.10	74	16	40	6	56	39.38	97.60	206.59	504.31

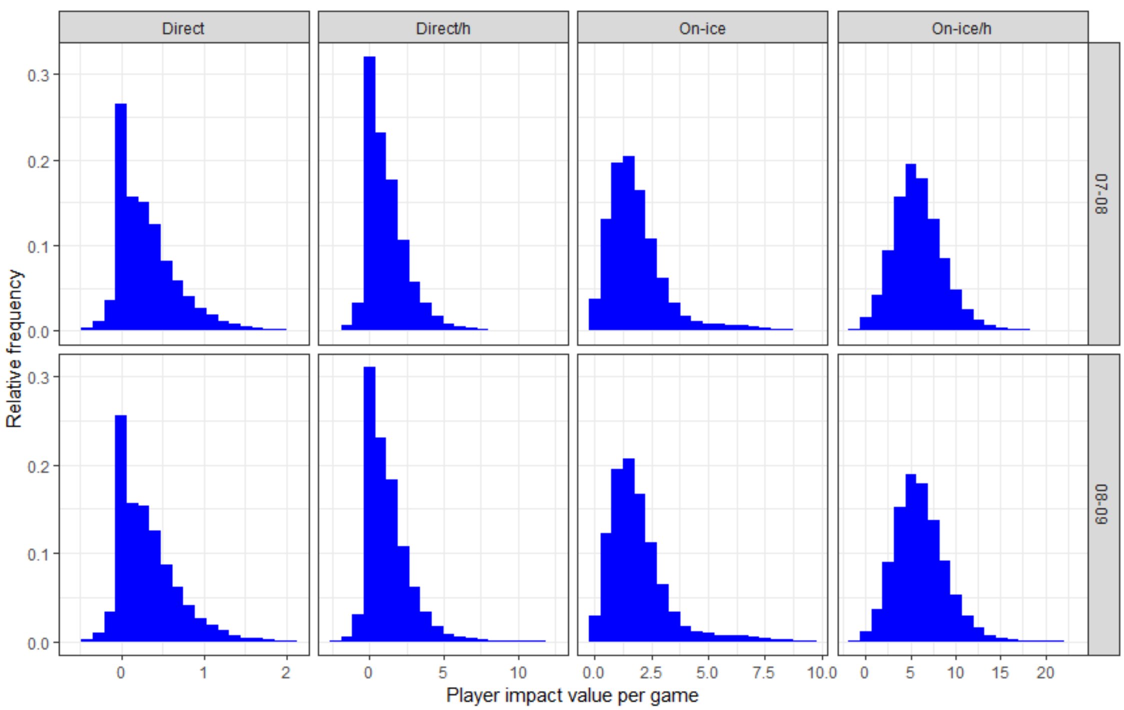


Figure 1: Player impact distributions for the 2007-2008 and 2008-2009 seasons.

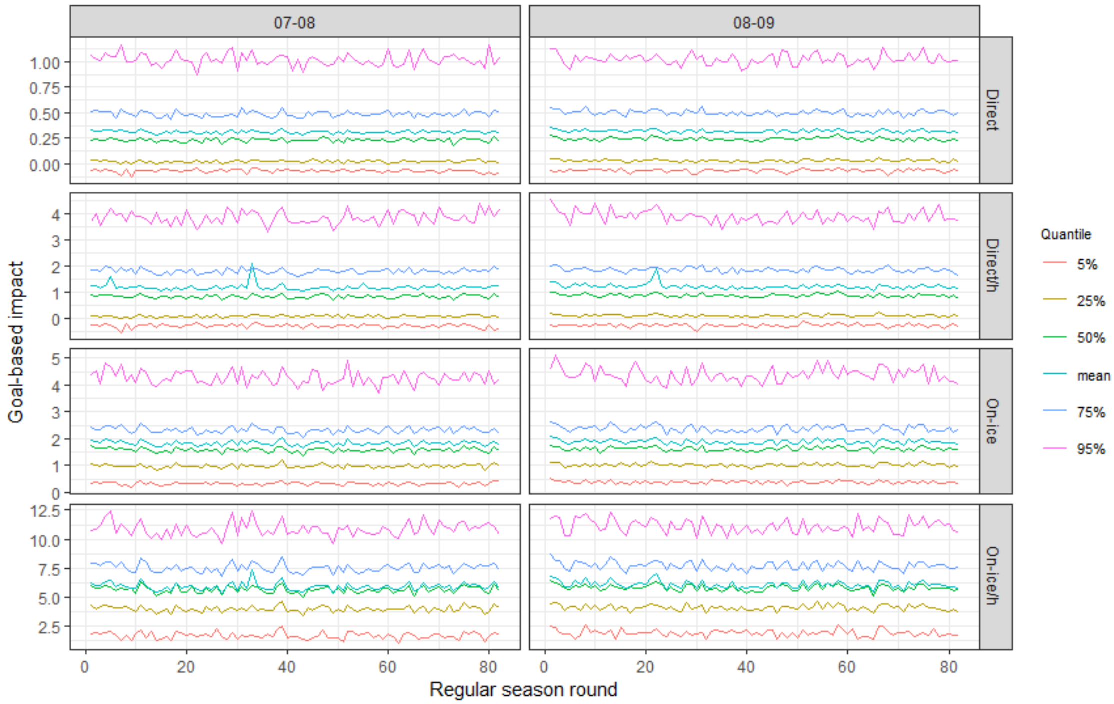
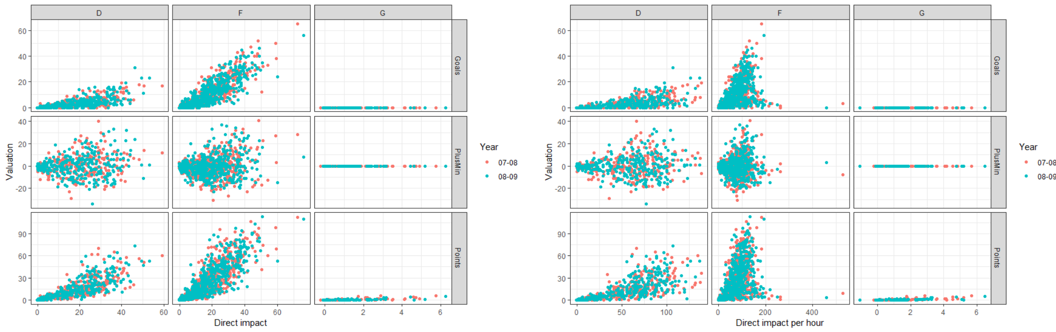


Figure 2: Player impact quantiles per game during 2007-2008 and 2008-2009 seasons.

excluded the two 0.01% tails. All measures are skewed towards the lower impacts. We also note that the distribution of the player impact is similar over the two seasons for each of the measures. In Fig. 2 we show the 5%, 25%, 75%, and 95% quantiles as well as the mean and median of all the goal-based player impacts per game for the regular season games in seasons 2007-2008 and 2008-2009. These are given for the direct and on-ice impacts and their normalized variants. The impact values for the quantile levels are rather stable during the season. Further, except for the normalized on-ice variant there is a clear separation between the mean and the median. We note that the levels of the player impact quantiles are similar over the two seasons for each of the measures. For the DGB-impact the values for the quantiles are about the following: 95% around 1, 75% around 0.5, mean around 0.36, 50% around 0.24, 25% around 0.03 and 5% around -0.06, while for the normalized DGB-impact the values for 95% around 3.9, 75% around 1.84, mean around 1.21, 50% around 0.86, 25% around 0.09 and 5% around -0.02. For the OIGB-impact these values are for 95% around 4.35, 75% around 0.23, mean around 1.8, 50% around 1.59, 25% around 1 and 5% around 0.34, while for the normalized OIGB-impact these values are for 95% around 11.95, 75% around 7.63, mean around 6, 50% around 0.57, 25% around 4 and 5% around 1.75. Further, we note that the gap between 95% and 75% is larger or much larger than between other quantiles of the same range. This may interpreted as that top players contribute much more than good players.

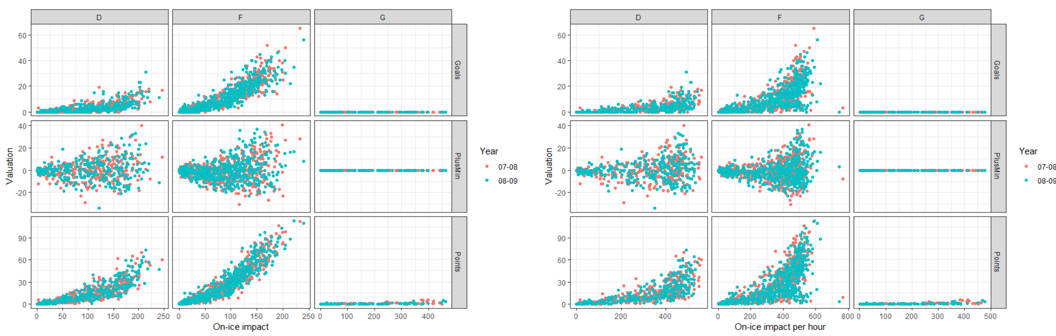
5.2 Goal-based impact versus other performance measures versus salary

In this section we compare the impact measures to the goal, points and +/- measures (Figs. 3 and 4). The impact measures follow the goals and points for forwards and defenders, but are not correlated with the +/- measure. Further, they seem to allow for a more fine-grained measure than the points and the



(a) (b)

Figure 3: Direct impact (a) and Direct impact per hour (b) versus Goals, +/- and Points for defenders, forwards and goal-keepers for the whole 2007-2008 and the whole 2008-2009 seasons.



(a) (b)

Figure 4: On-ice impact (a) and On-ice impact per hour (b) versus Goals, +/- and Points for defenders, forwards and goal-keepers for the whole 2007-2008 and the whole 2008-2009 seasons.

goals measures.

In Fig. 5 we plot the salary versus several performance measures. For defenders and forwards, the impact measures are similar to the goals and points in the sense that the higher the performance value, the higher the lowest salary for that value. However, the ranges for the salary for a particular performance value are quite large. For goalkeepers, as expected, the direct impact, goals and points have no correlation to the salary, but there is a similar trend as for forwards and defenders for the on-ice impact. We note that the on-ice impact for goalkeepers may be seen as a measure for the team when the goalkeeper is playing. The +/- measure does not seem to influence salary.

5.3 Streak durations

Over the duration of a season, player performance varies. Typically, point streaks (i.e., periods during which a player has points in consecutive games) are used to identify players that are "hot" or "cold". Furthermore, the points over the last few games (e.g., five games) are often reported, providing some idea of how a player (or team) is currently playing. However, due to the low-scoring nature of the game, long point streaks are becoming rarer[3], are seldom long lasting, and only assess offensive numbers. In this section we consider four alternative ways of identifying players currently on "hot streaks" and "cold streaks". First, instead of using points, we use one of the two metrics: (i) direct impact, and (ii) on-ice

[3] Longest point streak in a season (NHL). https://records.nhl.com/records/skater-records/scoring-streaks/longest-consecutive-point-streak

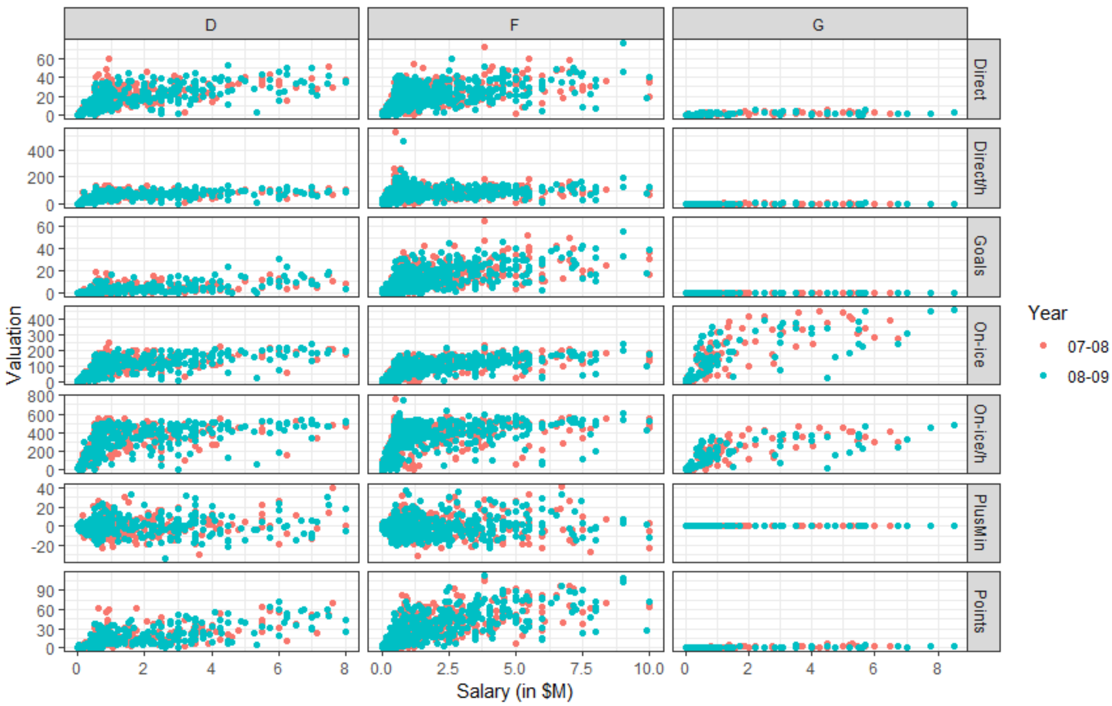
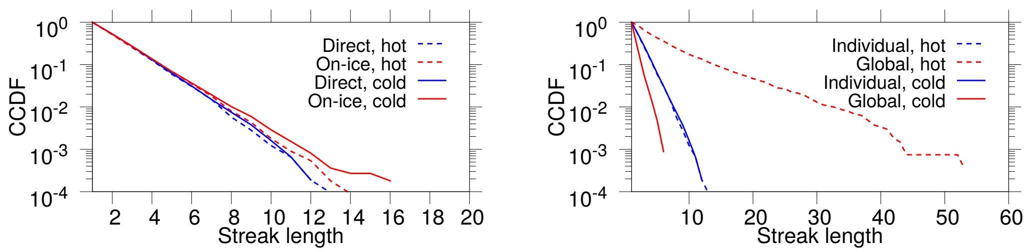


Figure 5: Salary versus different performance measures for defenders, forwards and goal-keepers for the whole 2007-2008 and the whole 2008-2009 seasons.



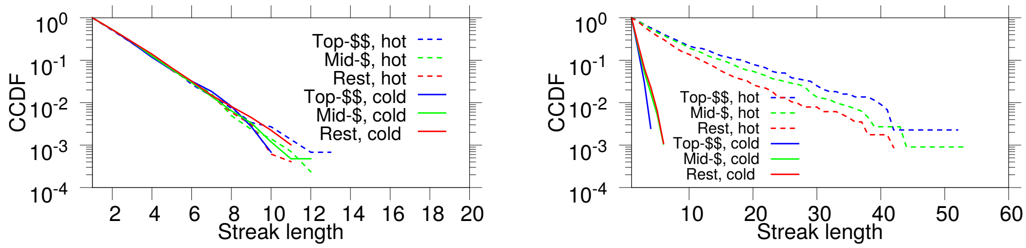
(a) Direct vs. on-ice (Individual)

(b) Individual vs Global (Direct)

Figure 6: Streak durations shown as empirical Complementary Cumulative Distribution Functions (CCDFs) over all layers during the 2007-2008 season.

impact. Second, we define “streaks” either based on how the player performs relative to an individual threshold (i.e., its median impact) or a global threshold (i.e., whether it has positive or negative impact). For the individual metric, a streak is defined as a sequence of games over which the player has an impact above (or below) the player’s median score (over the games the player plays in the season). To avoid assigning “hot streaks” to players that currently have zero impact, we only consider players that have a median impact above $\epsilon = 10^{-5}$. For the global measure, a hot/cold streak is defined as a sequence of games over which the player consistently has an impact strictly above ϵ or strictly below $-\epsilon$ for some small $\epsilon = 10^{-5}$, respectively. Throughout this analysis we only consider games the player participate in, not games missed by injury, being benched, or that the player for some other reason misses.

Fig. 6 shows the empirical Complementary Cumulative Distribution Functions (CCDFs) of the



(a) Individual (b) Global

Figure 7: Impact of player's salary range on the streak distributions (CCDFs). Results for 2007-2008 season when using the direct impact measure.

streak durations observed with these four methods across all players. Fig. 6(a) compare the two metrics (i.e., direct vs. on-ice) when using individual per-player thresholds (i.e., their median impact). While the on-ice metric results in slightly longer streaks, all four curves show clear straight-line behavior on lin-log scale suggesting that hot-streak durations when using an individual threshold is exponentially distributed. This itself suggests that hot streaks, when assessed relative to the players' average performance over a season, may actually be memoryless and recent performance history (including longer streaks) may not add value compared to just reporting the average performance over the entire season.

In contrast, as shown in Fig. 6(b), the "hot streak" durations when defined relative a global baseline (i.e., strictly positive/negative impact), have a somewhat heavier tail than suggested by a straight line. The reduced hazard rates observed here suggest that hot streaks when defined relative to such global baseline in fact carry memory. However, we note that part of this simply is due to these streaks often being due to good players (and teams) being more likely to be associated with these streaks. In summary, these results suggests that providing information about who is on a hot streak may primarily help indicate who is a good player and that this player is likely to contribute positively in the next game; however, it does not seem to be a good indicator whether this particular player will contribute more/less than on average over a season.

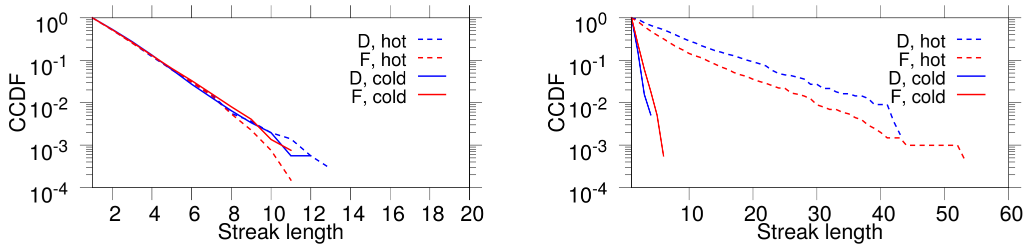
5.4 Streak durations for player groups

The above observations also hold when looking at individual player groups; e.g., based on salary range (Fig. 7) and player position (Fig. 8). For the salary ranges, we split the players into three categories: (i) the top-10% (with the highest salary), (ii) the mid-range players (with salaries in the top-40%, but not in the top-10%), and (iii) the rest (with salaries below those in the top-40%). Again, we note the typical straight-line behavior of an exponential distribution for the individual measures (Figs. 8(a) and 7(a)) and slightly heavier tails for the global measure (Figs. 8(b) and 7(b)).

Given our prior observation that good players are more likely to be associated with longer hot streaks (when using a global baseline), it is perhaps not surprising that long hot streaks are more frequently among the best paid players. Similarly, these players typically have shorter cold streaks than the less paid players.

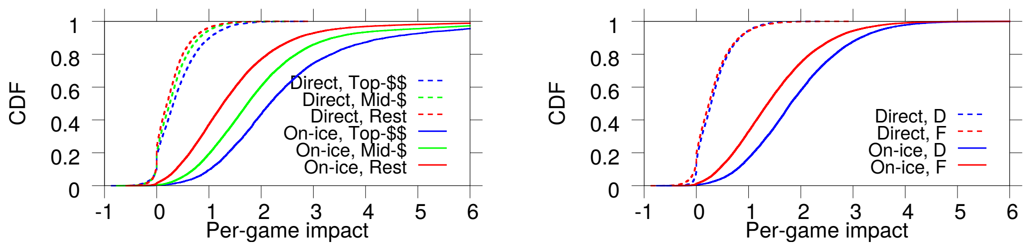
In general, we observe a strict ordering of the global CCDFs (shown Fig. 7(b)) based on salary range. This indicate that the better paid players in fact contribute more to the total impact of a team. To highlight these differences we also plotted CDFs of the per-game impact seen by players in the different salary categories (Fig. 9(a)) as a function of time (over the season) for different classes and example percentiles (Fig. 10(a)). And indeed, we did observed similar strict orderings here too.

Corresponding breakdowns based on player position highlight both differences and similarities. First, as shown in Fig. 9(b), although the direct impact is similar for the two player categories, de-



(a) Individual (b) Global

Figure 8: Impact of player position on the streak distributions (CCDFs). Results for 2007-2008 season when using the direct impact measure.



(a) Salary (b) Position

Figure 9: Empirical Cumulative Distribution Functions (CDFs) of the direct impact and on-ice impact for different player positions and salary ranges, as calculated on a per-game basis.

fenders typically have higher on-ice impact than forwards. Second, these observations are consistent across the season, as exemplified by the direct impact (median and 95%-ile scores) shown in Fig. 10(b). Third, although the longest hot streaks belong to forwards, the hot streaks in general are longer for defenders and shorter for forwards. Both the first and the last observation may in part be due to the top-defenders typically playing more minutes per game than forwards.

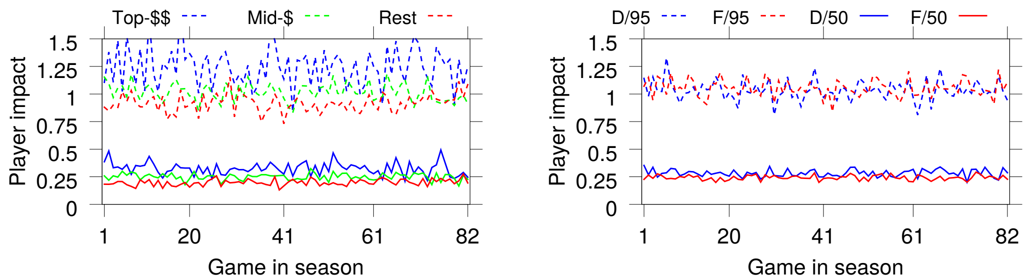
6 Conclusion

In this paper we introduced and analyzed approaches for measuring goal-based player impact in ice hockey. We showed that the measures, similar to the +/- measure, to a larger degree allow for defenders in the top rankings than goals and assists. There is a certain correlation (as expected) to goals and assists, but not to +/-. Further, we defined two notions of streaks that could be indicators of good players, but not for performance in the next game.

Regarding future work, one direction is to work with different reward functions in the Q-learning algorithm to investigate impact of player actions for different desirable outcomes (e.g., shots on goals, powerplays). Further, it would be interesting to extend the work in [5] and investigate alternative pair impact definitions.

References

[1] Robert B. Gramacy, Shane T. Jensen, and Matt Taddy. Estimating player contribution in hockey with regularized logistic regression. *Journal of Quantitative Analysis in Sports*, 9:97–111, 2013. doi: 10.1515/jqas-2012-



(a) Salary (b) Position

Figure 10: Comparison of the percentile values as seen for different player positions and salary ranges. Here, we show per-game impact values of the median and 95%-ile for each game during the 2007-2008 season.

0001.

[2] Wei Gu, Krista Foster, Jennifer Shang, and Lirong Wei. A game-predicting expert system using big data and machine learning. *Expert Systems with Applications*, 130:293–305, 2019. doi: 10.1016/j.eswa.2019.04.025.

[3] Edward H Kaplan, Kevin Mongeon, and John T. Ryan. A Markov Model for Hockey: Manpower Differential and Win Probability Added. *INFOR Information Systems and Operational Research*, 52(2):39–50, 2014. doi: 10.3138/infor.52.2.39.

[4] Guiliang Liu and Oliver Schulte. Deep reinforcement learning in ice hockey for context-aware player evaluation. In Jérôme Lang, editor, *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, pages 3442–3448, 2018. doi: 10.24963/ijcai.2018/478.

[5] Dennis Ljung, Niklas Carlsson, and Patrick Lambrix. Player pairs valuation in ice hockey. In Ulf Brefeld, Jesse Davis, Jan Van Haaren, and Albrecht Zimmermann, editors, *Machine Learning and Data Mining for Sports Analytics. MLSA 2018*, volume 11330 of *Lecture Notes in Computer Science*, pages 82–92, 2019. doi: 10.1007/978-3-030-17274-9_7.

[6] Brian Macdonald. A Regression-Based Adjusted Plus-Minus Statistic for NHL Players. *Journal of Quantitative Analysis in Sports*, 7(3), 2011. doi: 10.2202/1559-0410.1284.

[7] Brian Macdonald. An Improved Adjusted Plus-Minus Statistic for NHL Players. In *MIT Sloan Sports Analytics Conference*, 2011.

[8] Stephen Pettigrew. Assessing the offensive productivity of NHL players using in-game win probabilities. In *MIT Sloan Sports Analytics Conference*, 2015.

[9] Kurt Routley and Oliver Schulte. A Markov Game Model for Valuing Player Actions in Ice Hockey. In Marina Meila and Tom Heskes, editors, *Uncertainty in Artificial Intelligence*, pages 782–791, 2015.

[10] Michael Schuckers and James Curro. Total Hockey Rating (THoR): A comprehensive statistical rating of National Hockey League forwards and defensemen based upon all on-ice events. In *MIT Sloan Sports Analytics Conference*, 2013.

[11] Oliver Schulte, Mahmoud Khademi, Sajjad Gholami, Zeyu Zhao, Mehrsan Javan, and Philippe Desaulniers. A Markov Game model for valuing actions, locations, and team performance in ice hockey. *Data Mining and Knowledge Discovery*, 31(6):1735–1757, 2017. doi: 10.1007/s10618-017-0496-z.

[12] Oliver Schulte, Zeyu Zhao, Mehrsan Javan, and Philippe Desaulniers. Apples-to-apples: Clustering and Ranking NHL Players Using Location Information and Scoring Impact. In *MIT Sloan Sports Analytics Conference*, 2017.

[13] A.C. Thomas, Samuel L. Ventura, Shane Jensen, and Stephen Ma. Competing Process Hazard Function Models for Player Ratings in Ice Hockey. *The Annals of Applied Statistics*, 7(3):1497–1524, 2013.