Player performance in ice hockey

Patrick Lambrix and Niklas Carlsson

Linköping University, Sweden

Motivation for Niklas

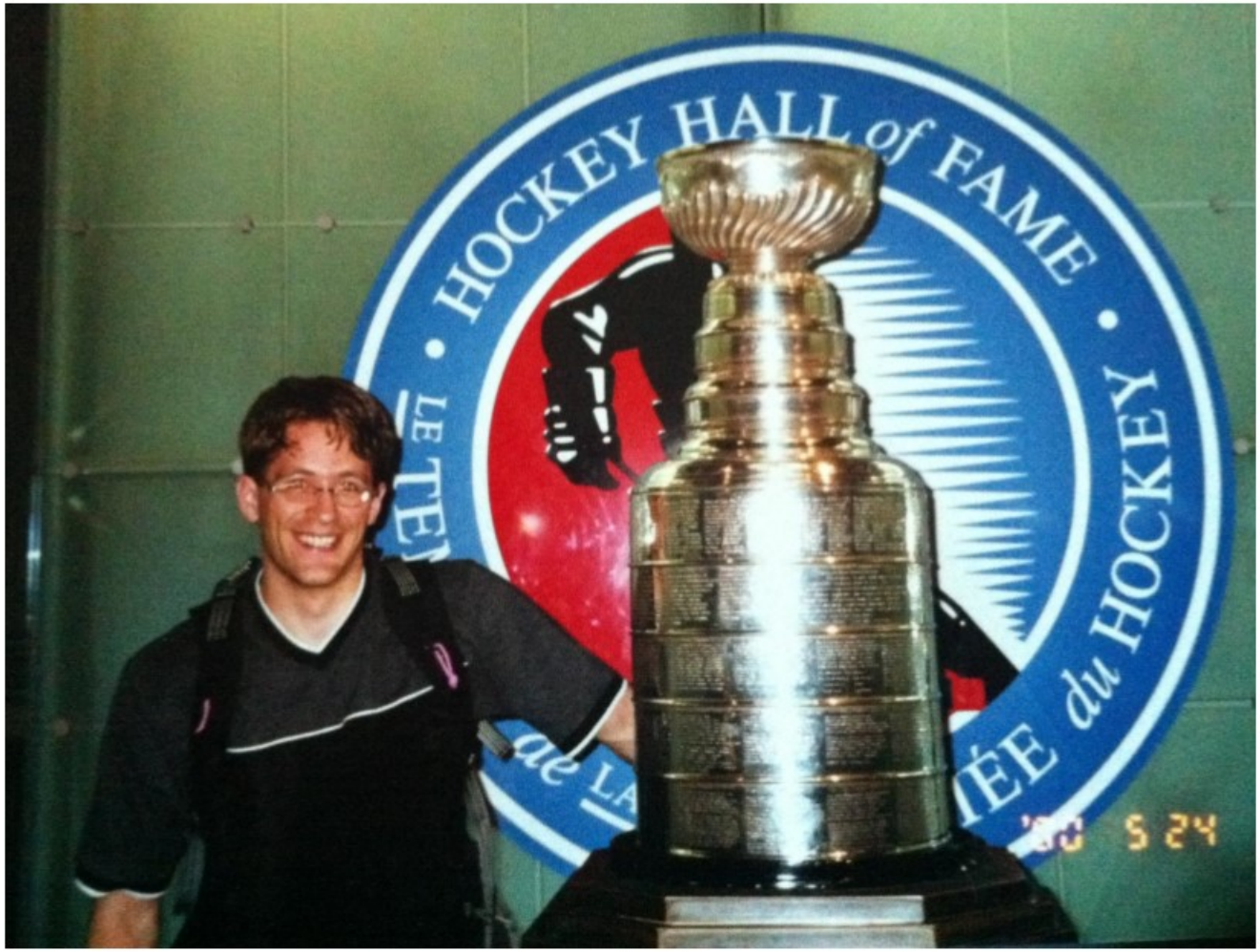
Niklas
(and many others)
dream of:

Motivation for Niklas

First try

Motivation for Niklas

A bit easier …



Motivation

Not completely given up first method …

Motivation for Patrick

Outline

- Motivation
- Methods and Results
- Conclusion

Performance metrics - traditional

RANK	SPELARE	NR	LAG	POS	GP	G	A	TP	PIM	GWG	PPG	SOG	HITS	BKS	+	-	+/-	TOI/GP
1	Ryan Lasch	81	Frölunda	F	37	10	25	35	16	3	2	66	9	2	23	31	-8	19:02
2	Joakim Lindström	10	Skellefteå	F	38	13	21	34	14	3	4	132	4	8	29	25	4	17:15
3	Derek Roy	9	Linköping	F	39	5	29	34	22	0	2	74	4	20	29	20	9	17:15

- *Offensive:* G: goals, A: assists, TP: points, GWG: game winning goals, PPG: powerplay goals, SOG: Shots on goal
- *Defensive:* HITS: hits, BKS: blocked shots
- +/-: plus-minus
- PIM: penalty minutes
- *Time:* GP: games played, TOI: time on ice

Performance metrics - advanced

- **Corsi:** shots

- **xG** (Expected Goals): assigns a value to each shot, based on the likelihood of the shot resulting in a goal.

- These metrics have made it into the ice hockey discourse

Performance metrics - advanced

Critique on advanced metrics: context

Some new approaches:

- Using Markov games
- THOR (Total Hockey Rating)

Motivation

Our goal: **Goal-based** metrics that take into account context

Variant 1

- Are variants on traditional metrics
- Are easy to understand for practitioners
- Take into account goal *importance*

Variant 2

- New advanced metrics
- Take into account actions leading to goals

Outline

- Motivation
- Methods and Results
- Conclusion

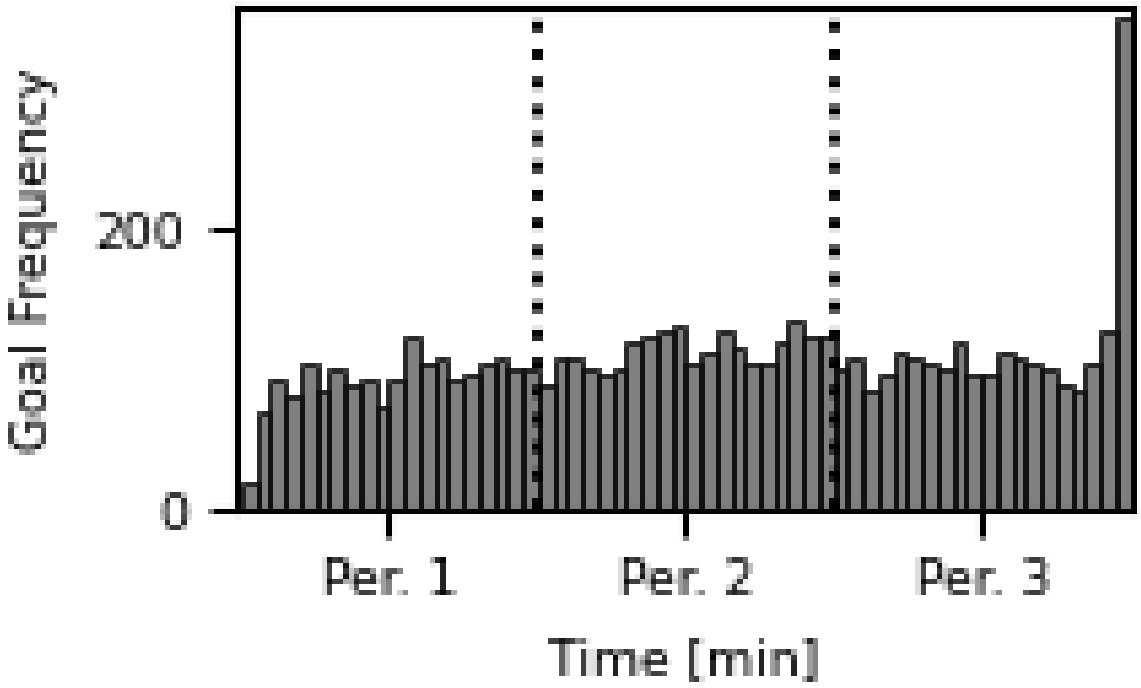
Data

- Play-by-play data from Sportlogiq
- Seasons 2007-2008 to 2013-2014
- Only regular season

Outline

- Motivation
- Methods and Results
  - Variant 1
- Conclusion

Goal frequency per minute



Game Points Importance Value for a goal

- Game points for NHL: 2pts for win, 1pt for tie and loss after overtime, 0pts for loss in regulation time.

Game Points Importance Value for a goal

- Outcome in regulation time: win, tie, loss

- Context:

<time, goal differential, manpower differential>

$$P(outcome \mid context) = \frac{Occ(context \mid outcome)}{Occ(context)}$$

Game Points Importance Value for a goal in a context

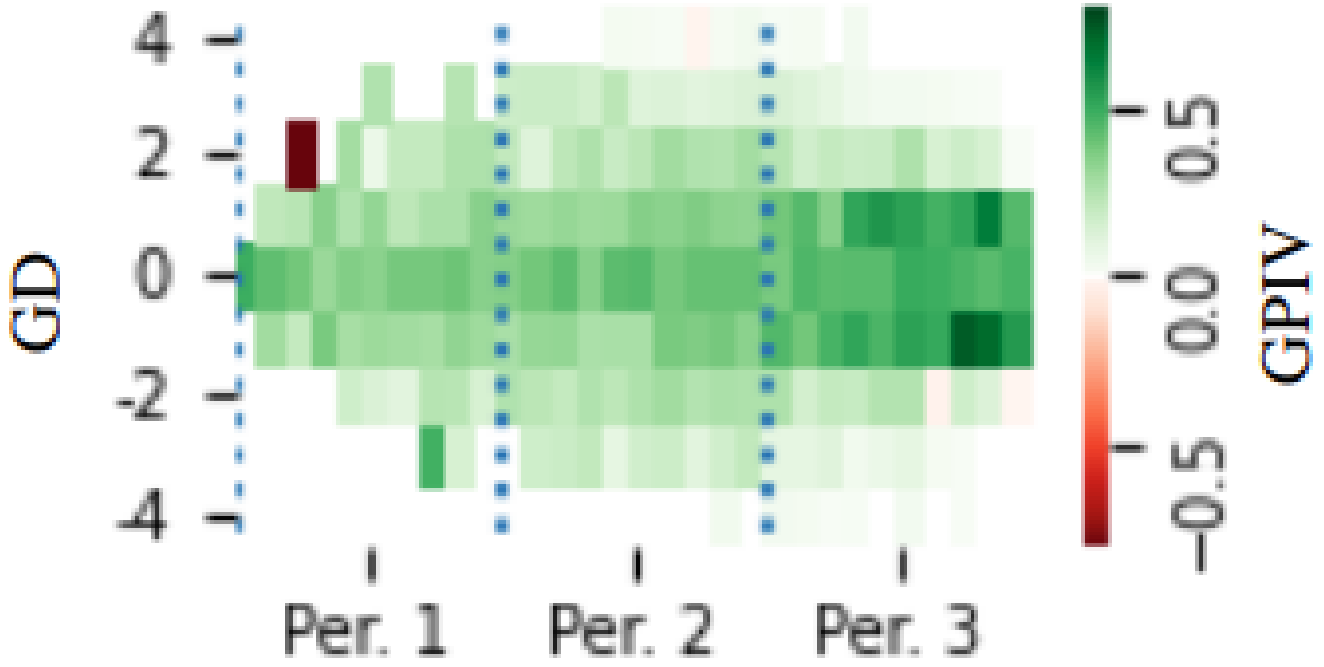
Change of probability of winning the game by scoring the goal

context = contextBG

$$GPIV(context) = 2 * [P(win \mid contextAG) - P(win \mid contextBG)] + 1 * [P(tie \mid contextAG) - P(tie \mid contextBG)]$$

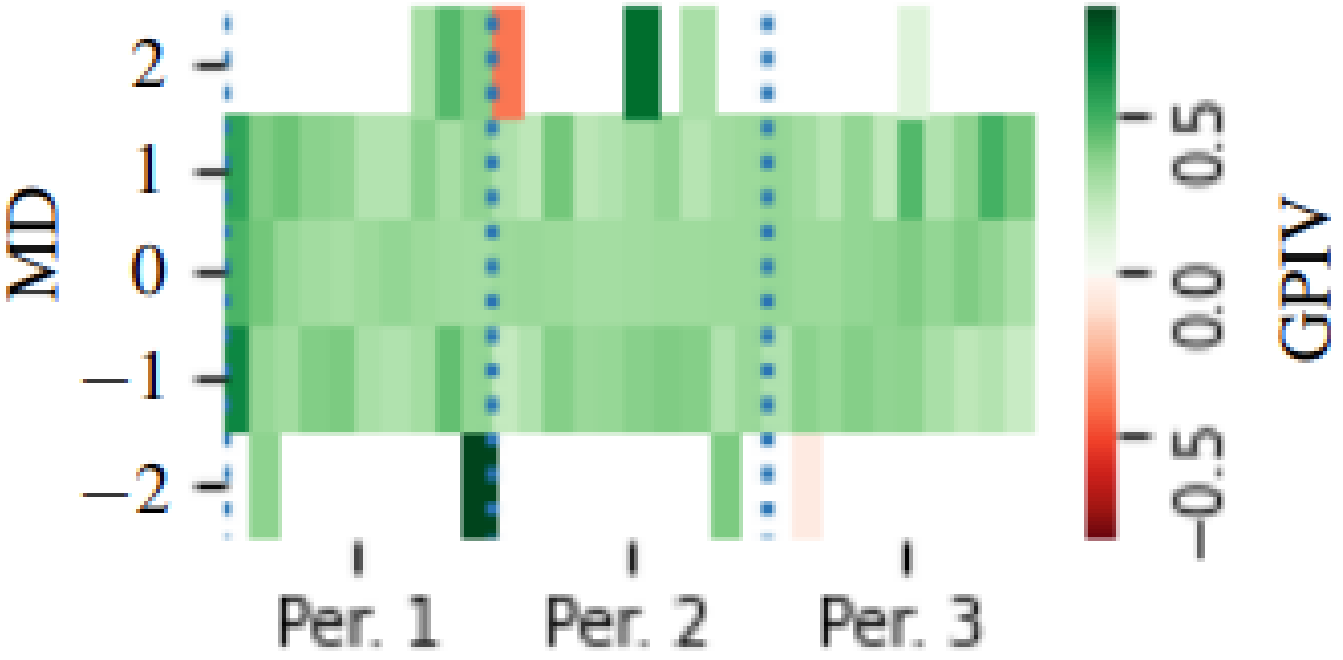
Change of probability of the game ending in a tie by scoring the goal

GPIV vs Goal Differential



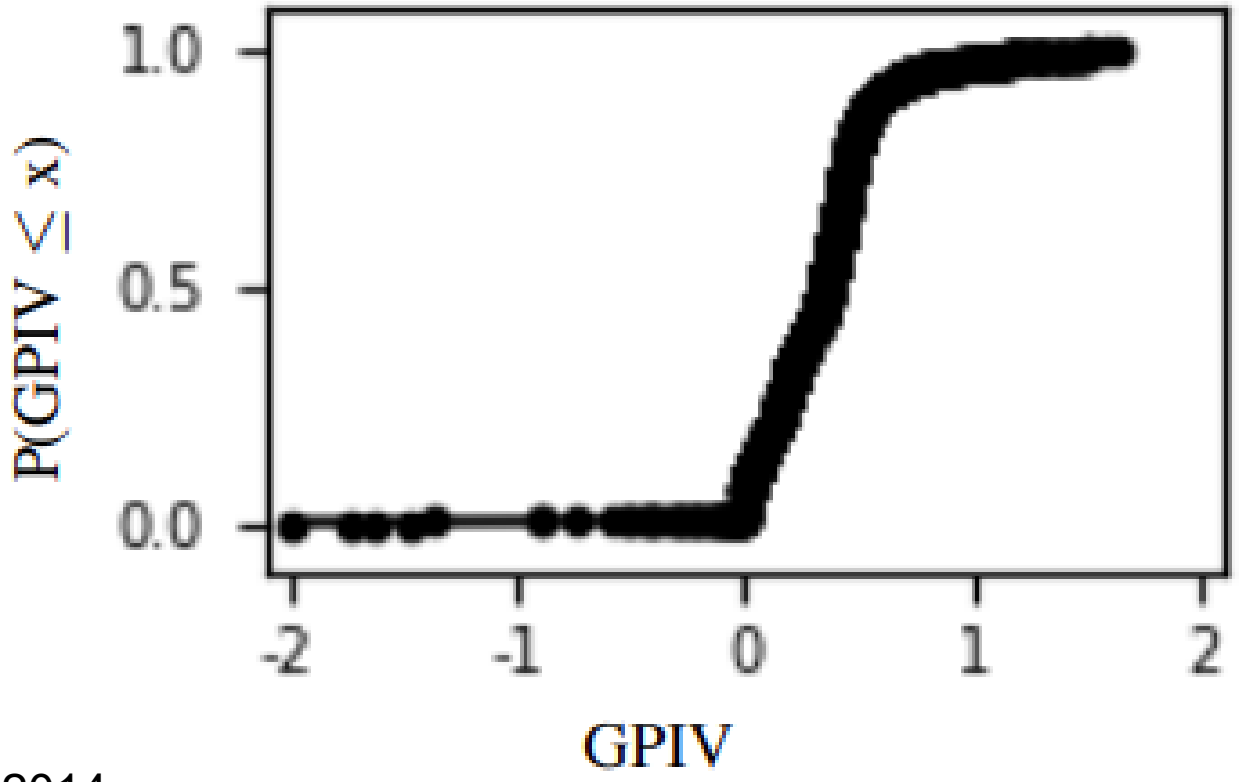
2013-2014 season

GPIV vs Manpower Differential



2013-2014 season

Cumulative Distribution Function for GPIV



2013-2014 season

New metrics

- Traditional metrics:

  Goal contributes 1 directly (Goals, +/-) or indirectly (Assists)

- Variants of Goals, Assists, Points, +/-:

  Goal contributes with its context-based GPIV

Top 10 players for GPIV-P

P-Rank	GPIV-P-Rank	Rank-diff	Player	Position	P	GPIV-P
2-3	1	1	Sidney Crosby	C	69	25.734
6-7	2	4	Alex Ovechkin	R	64	25.085
4	3	1	Joe Pavelski	C	67	23.467
1	4	-3	Tyler Seguin	C	70	22.259
5	5	0	Phil Kessel	R	66	22.006
6-7	6	0	Ryan Getzlaf	C	64	21.366
2-3	7	-5	Corey Perry	R	69	20.803
20-22	8	12	Blake Wheeler	R	51	20.295
20-22	9	11	Anze Kopitar	C	51	19.812
23-24	10	13	Eric Staal	C	50	19.791

2013-2014 season

Top 10 players for GPIV-G

G-Rank	GPIV-G-Rank	Rank-diff	Player	Position	G	GPIV-G
1	1	0	Alex Ovechkin	R	48	19.080
2	2	0	Corey Perry	R	42	14.145
3	3	0	Joe Pavelski	C	41	14.115
7-8	4	3	Sidney Crosby	C	35	13.674
4-6	5	-1	Phil Kessel	R	37	13.358
13	6	7	Jeff Skinner	L	32	12.666
31-40	7	24	Kyle Okposo	R	26	11.757
20-24	8	12	Marian Hossa	R	29	11.713
25	9	16	David Perron	L	28	11.638
4-6	10	-6	Tyler Seguin	C	37	11.527

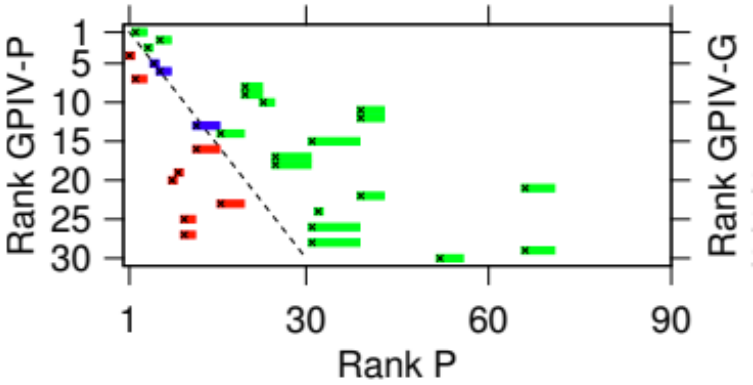
2013-2014 season

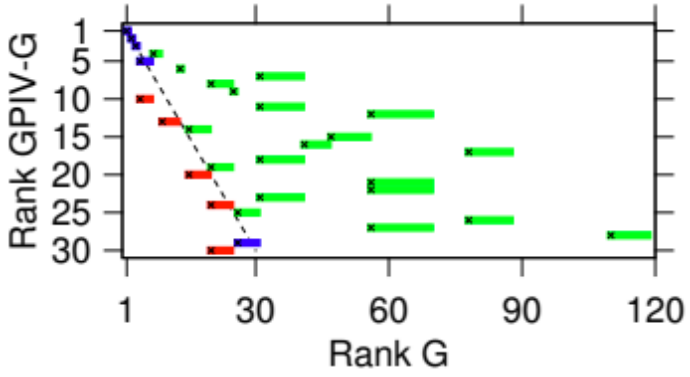
Top 10 players for GPIV-A (first)

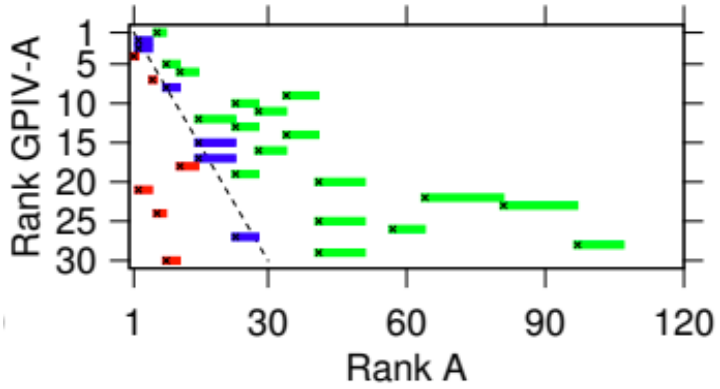
A-Rank	GPIV-A-Rank	Rank-diff	Player	Position	A	GPIV-A
6-7	1	5	Eric Staal	C	29	12.541
2-4	2	0	Ryan Getzlaf	C	34	12.376
2-4	3	-1	Sidney Crosby	C	34	12.060
1	4	-3	Joe Thornton	C	35	11.724
8-10	5	3	Matt Duchene	C	28	11.126
11-14	6	5	Keith Yandle	D	27	10.851
5	7	-2	Tyler Seguin	C	33	10.731
8-10	8	0	Gabriel Landeskog	L	28	10.461
34-40	9	25	Jiri Hudler	C	23	10.248
23-27	10	13	John Tavares	C	25	10.143

2013-2014 season

Rank changes P, G, A (first)

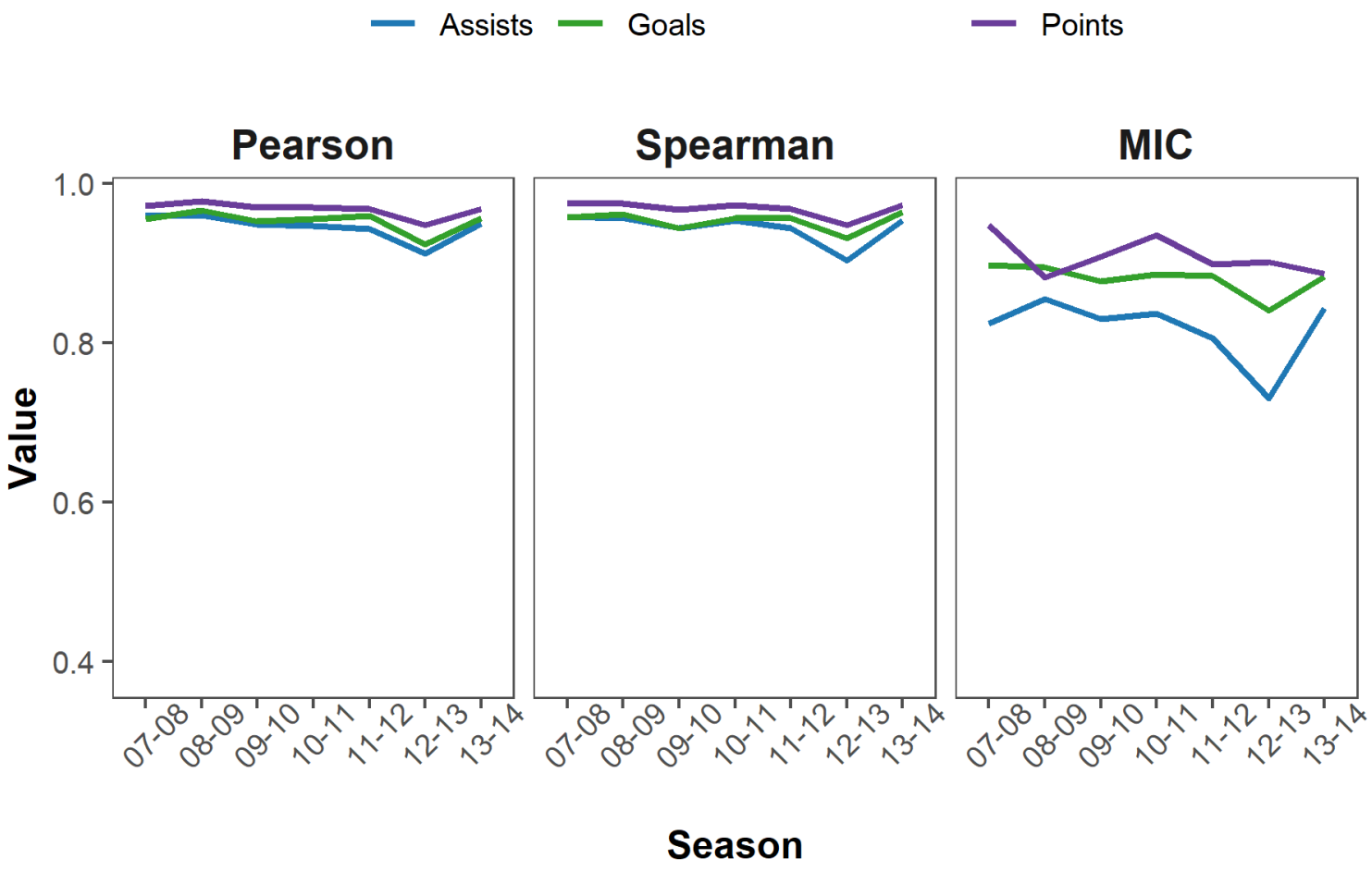




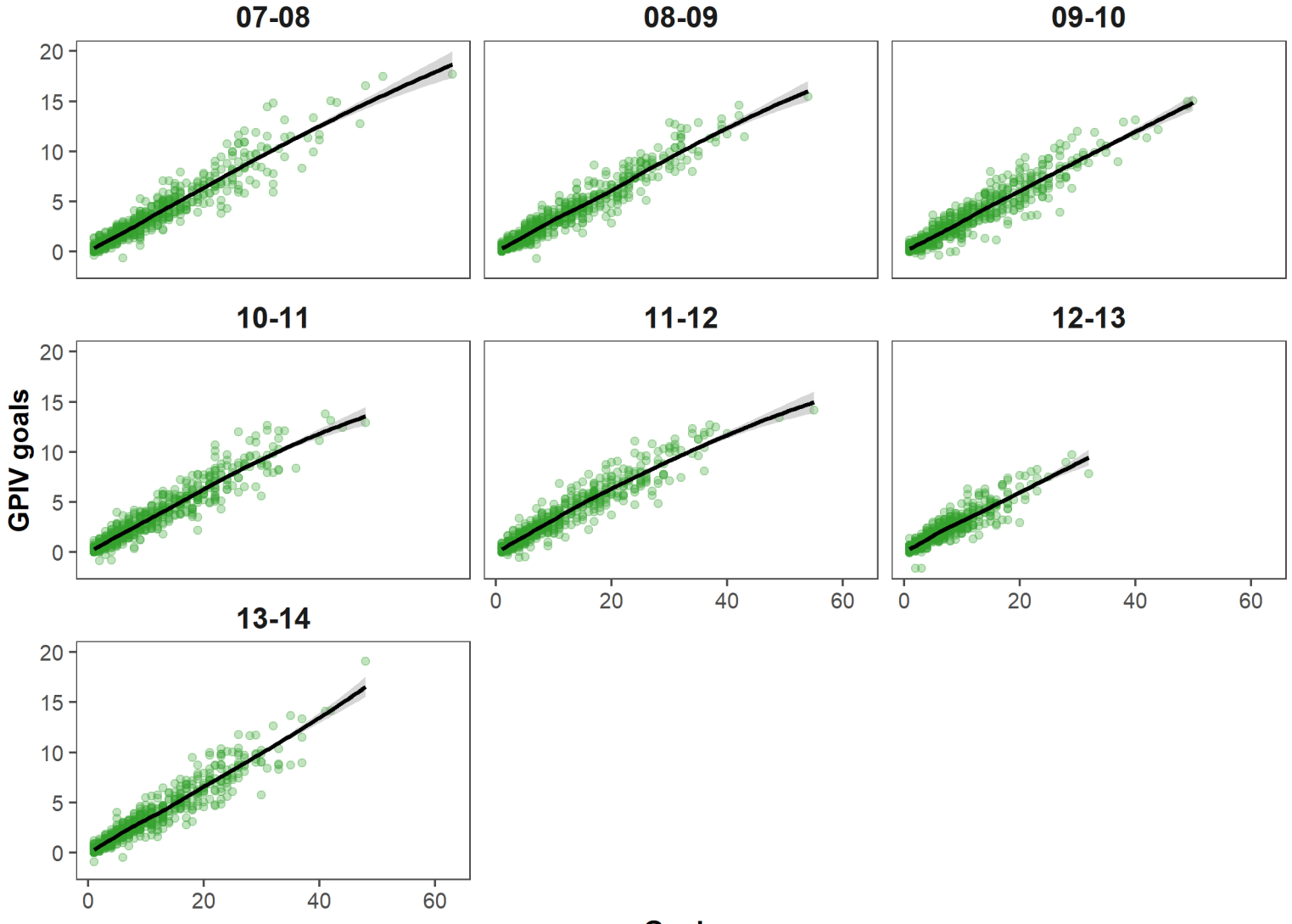


2013-2014 season

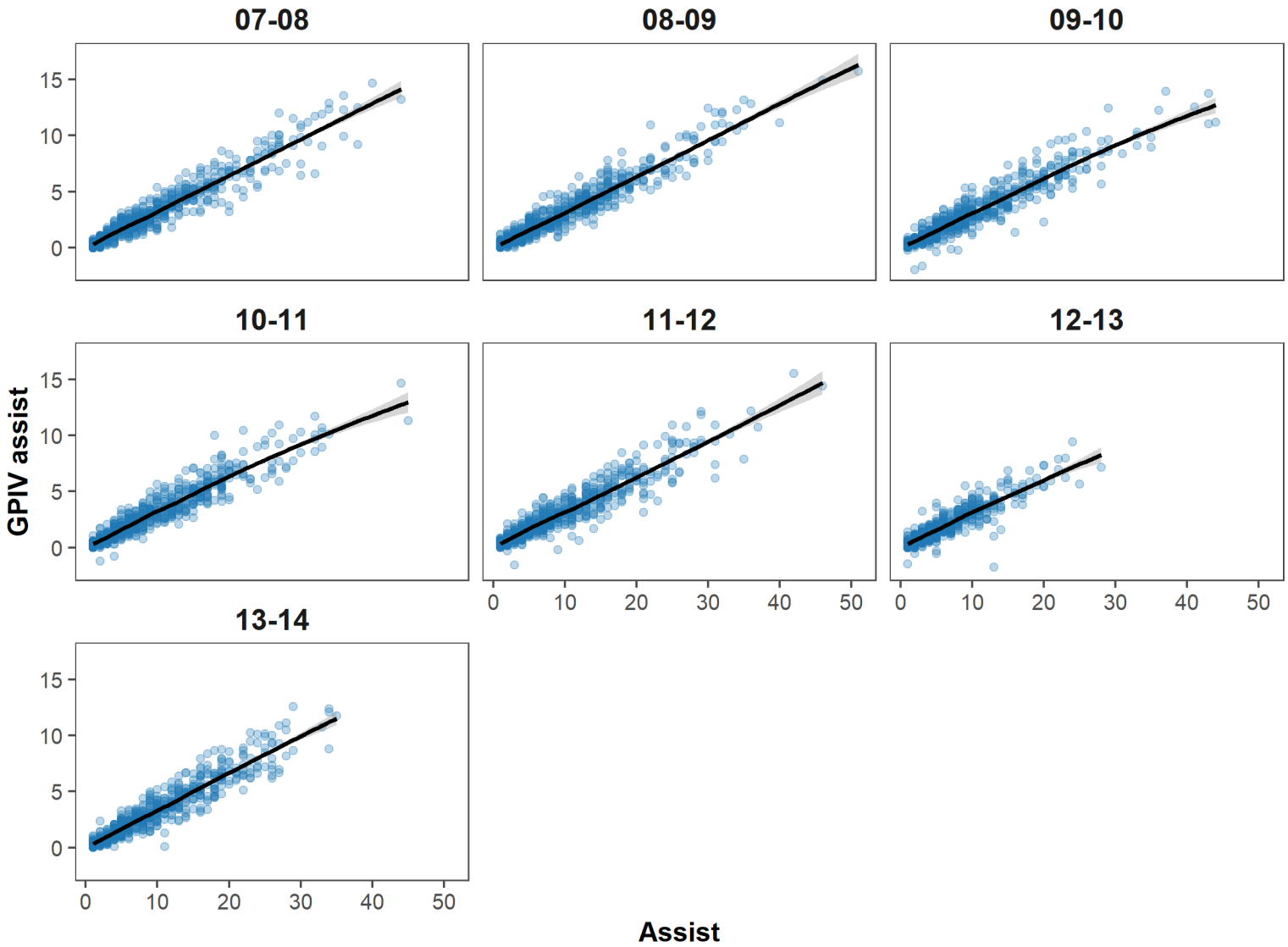
Correlations for 7 seasons



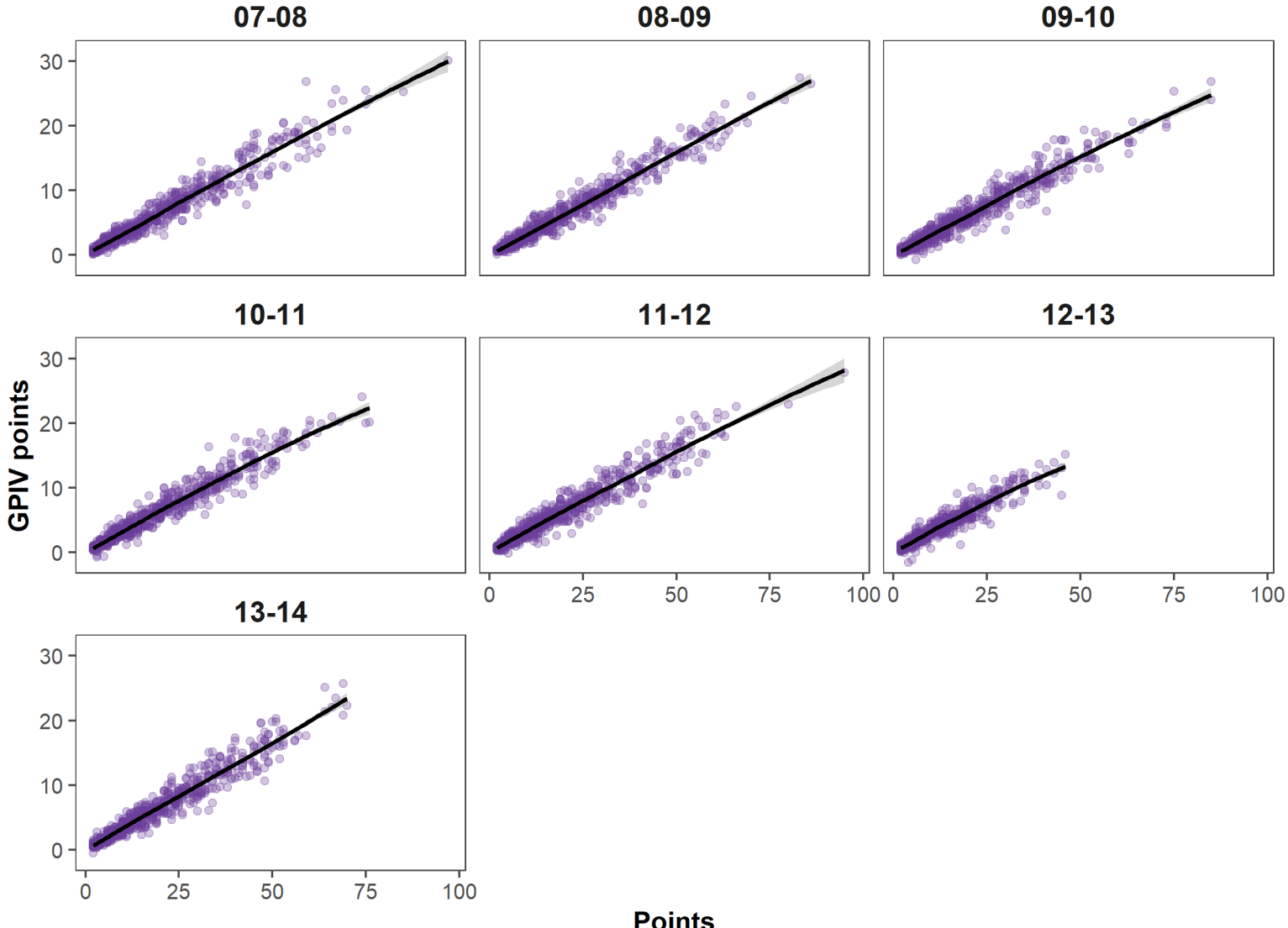
Goals vs GPIV-Goals



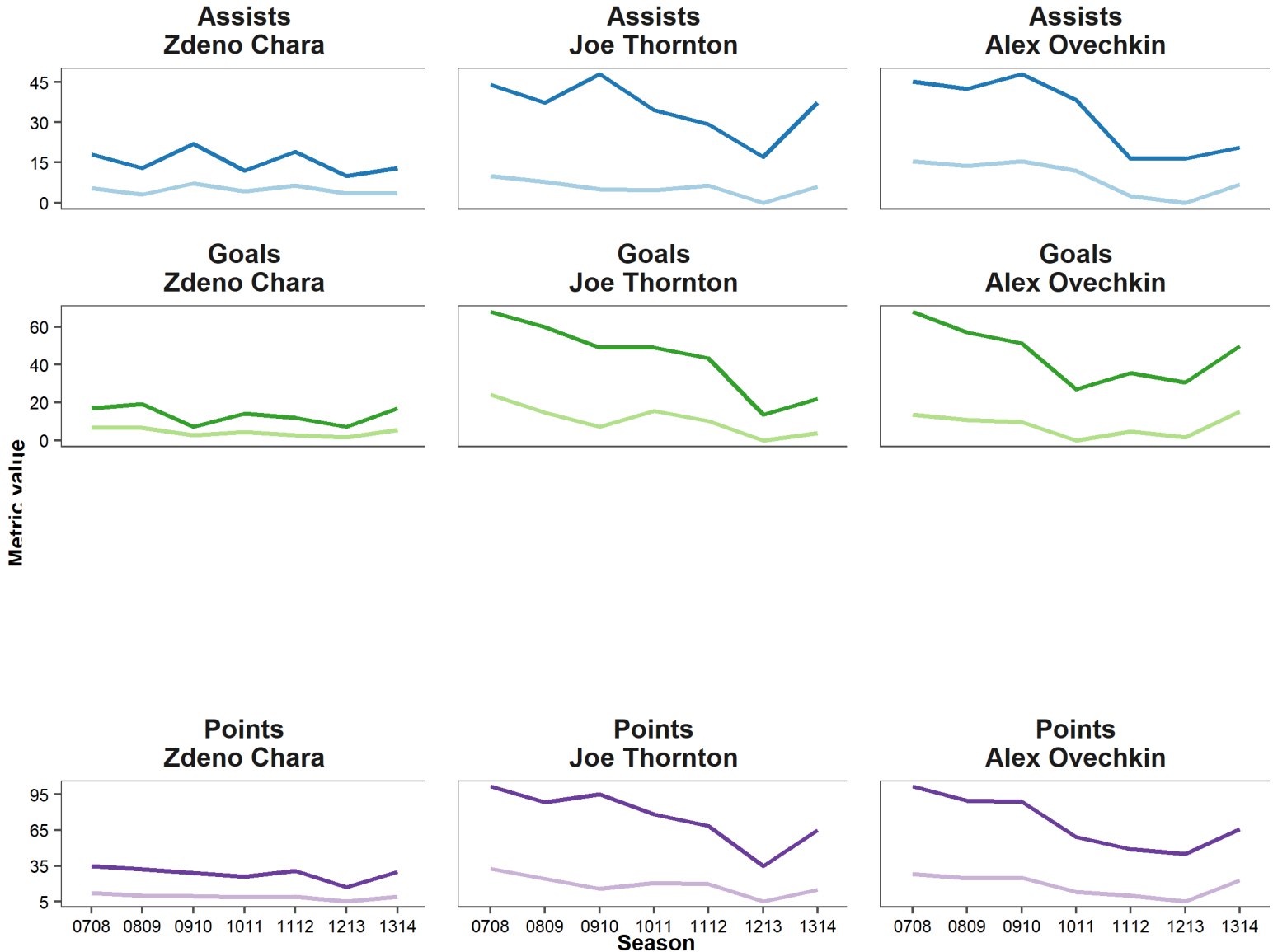
Assists vs GPIV-Assists (first)



Points vs GPIV-Points



Trends for some players



Currently working on / Future work

- Playoffs
- Overtime goals
- Change of GPIV during season
- GPIV over different seasons
- SHL

Outline

- Motivation
- Methods and Results
  - Variant 2
- Conclusion

Action Impact Model

- Based on the work by Routley and Schulte 2015*
- Idea:
  - Define state $s = < c, ps >$

    where c is a context and ps is a play sequence
  - Actions are performed in states
  - Define impact of action in a state
  - Define player impact based on action impacts

*Schulte's group presented a more extended model at IJCAI 2018.

Action Impact Model

Context

Notation	Name	Range
GD	Goal Differential	[-8,8]
MD	Manpower Differential	[-3,3]
P	Period	[1,7]

Action Impact Model

Events

Action Event	Start/End Event
Faceoff	Period Start
Shot	Period End
Missed Shot	Early Intermission Start
Blocked Shot	Penalty
Takeaway	Stoppage
Giveaway	Shootout Completed
Hit	Game End
Goal	Game Off
	Early Intermission End

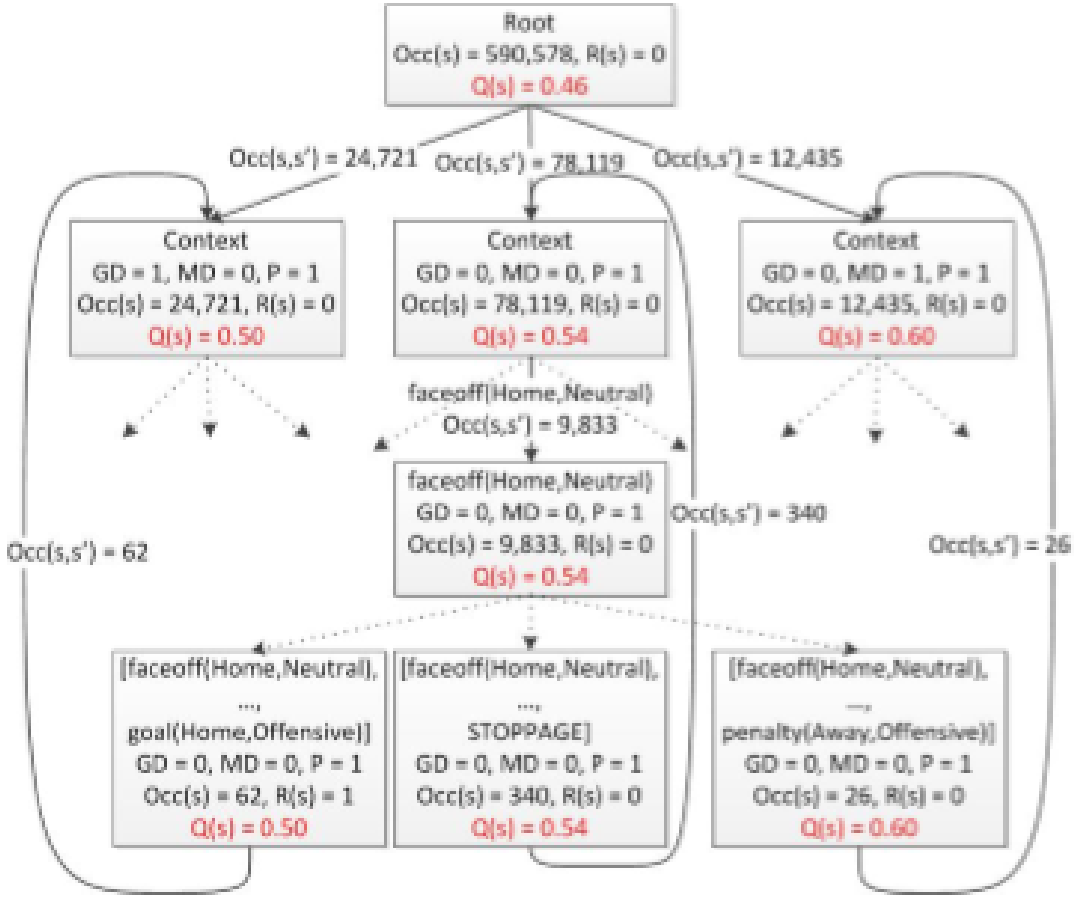
Action Impact Model

A play sequence is defined as

- the empty sequence or
- a sequence of events
  - first event: start marker
  - (possible) next events: action events
  - (possible) last event: end event

(→complete sequence)

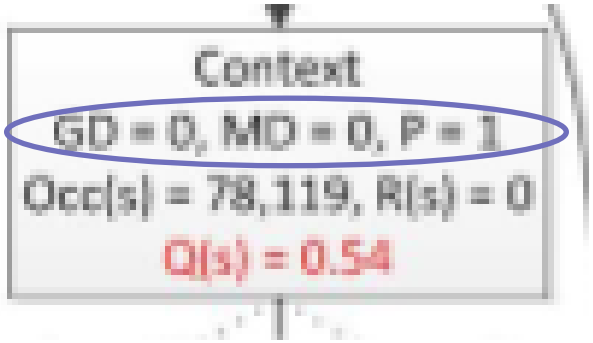
Action Impact Model

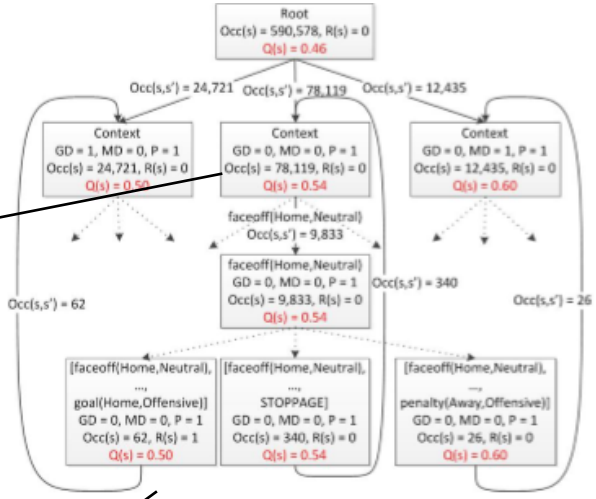


Routley and Schulte, 2015

Action Impact Model

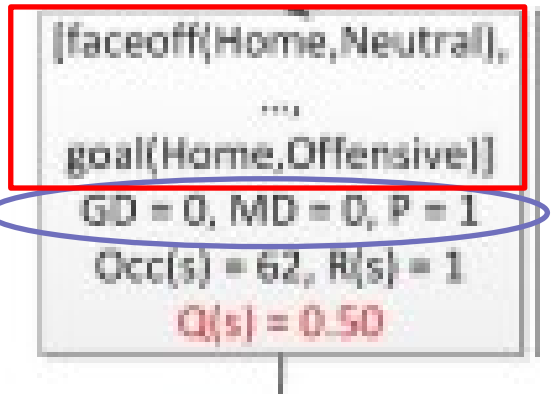
State $s = < c, ps >$





Context

Play sequence



Routley and Schulte, 2015

Action Impact Model

- Actions are performed in states

$< c, ps > * a =$

$< c, append(ps,a) >$ if state has no end event

(add action to play sequence, e.g., shot)

$< c', empty\text{-}set >$ if state has end event

(change context, e.g., after a goal)

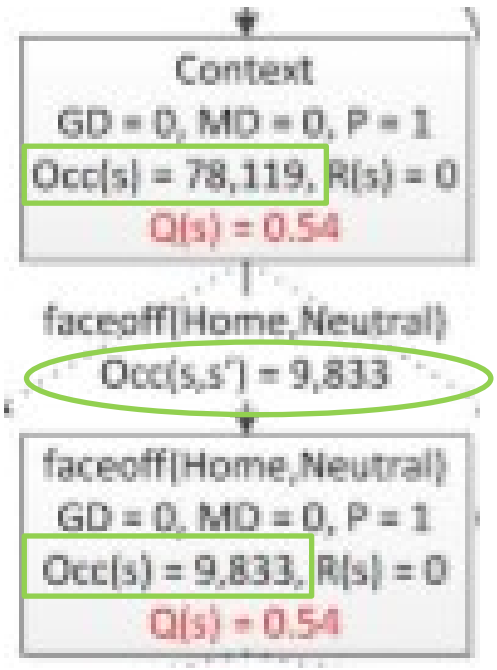
Action Impact Model

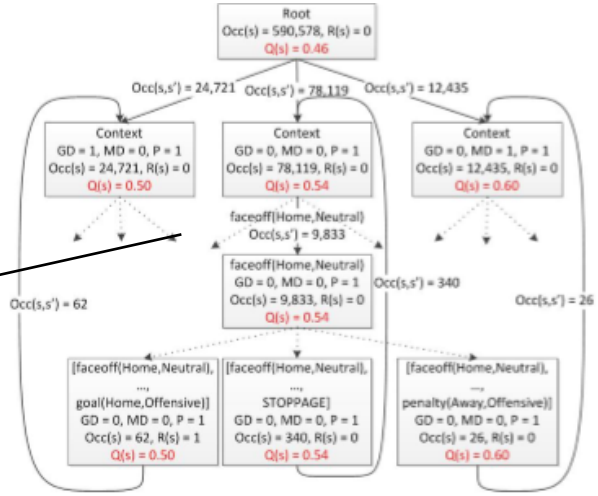
Based on play-by-play data:

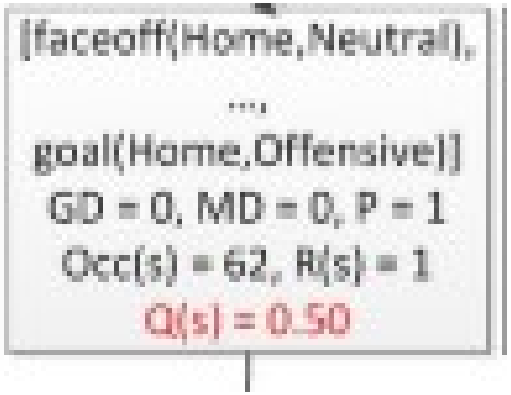
- Occurrences of state *s*: *Occ(s)*
- Occurrences of state *s* immediately followed by state *s'*: *Occ(s,s')*

- Transition probability $T(s,s') = Occ(s,s') / Occ(s)$

Action Impact Model







Occurrences

Occurrences

Routley and Schulte, 2015

Action Impact Model

Value iteration algorithm → Q-values

Reward function: goal states receive reward 1

(In single player experiments
also goal against reward -1)

- Impact of action *a* in state *s*: $Q_T(s * a) - Q_T(s)$

Action Impact Model

Algorithm 1 Dynamic Programming for Value Iteration

Require: Markov Game model, convergence criterion c, maximum number of iterations M

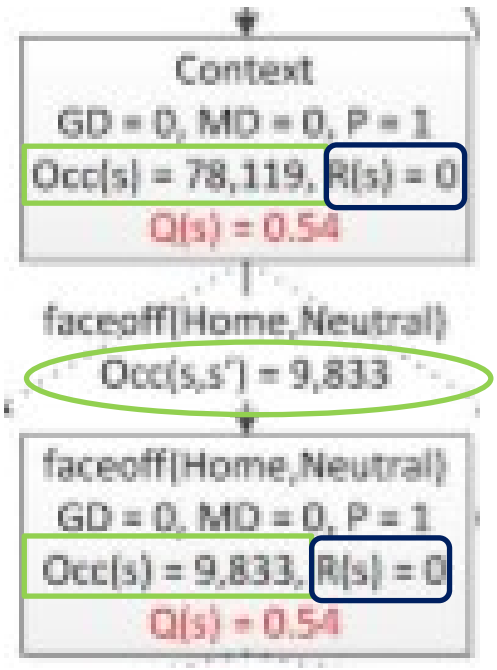
$lastValue = 0$
$currentValue = 0$
$converged = false$
for $i = 1; i \leq M; i \leftarrow i + 1$ **do**
  for all states s in the Markov Game model **do**
    if $converged == false$ **then**
      $Q_{i+1}(s) = R(s) + \frac{1}{Occ(s)} \sum_{(s,s') \in E} (Occ(s, s') \times Q_i(s'))$
      $currentValue = currentValue + |Q_{i+1}(s)|$
    end if
  end for
  if $converged == false$ **then**
    if $\frac{currentValue - lastValue}{currentValue} < c$ **then**
      $converged = true$
    end if
  end if
  $lastValue = currentValue$
  $currentValue = 0$
end for

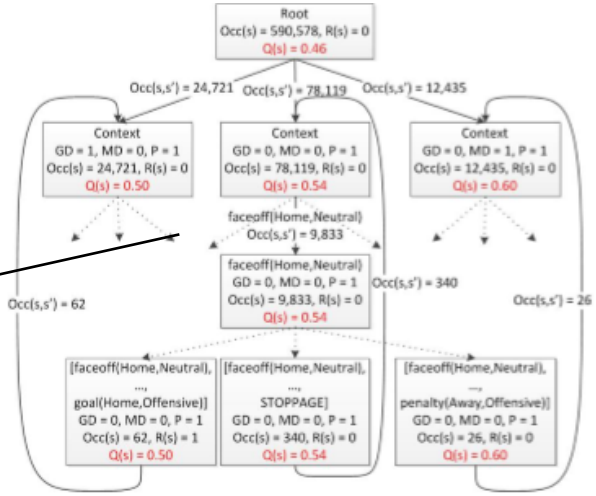
Action Impact Model

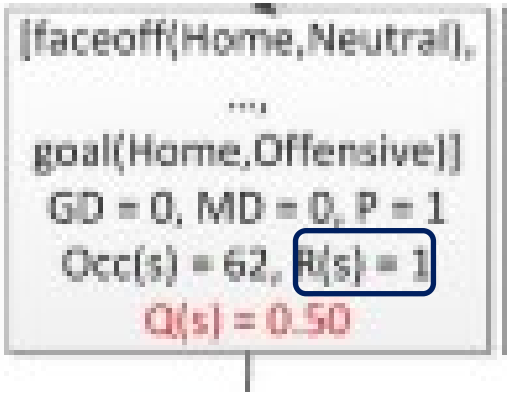
7: $$Q_{i+1}(s) = R(s) + \frac{1}{Occ(s)} \sum_{(s,s') \in E} (Occ(s,s') \times Q_i(s'))$$

Compute separate Q-values for Home and Away teams

Action Impact Model







Occurrences

Occurrences

Reward

Q-value

Routley and Schulte, 2015

Player Impact

Sum of action impacts

1. Based on all actions performed by the player (direct goal-based impact)
2. Based on actions when the player is on the ice (on-ice goal-based impact)

Variants normalized by time

Player Pair Impact

Sum of action impacts when both players are on the ice (on-ice goal-based impact)

Variants normalized by time

Definitions of impact

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)

Top players 2007-2008 and 2008-2009 for direct impact

PlayerName	Position	Age	Salary	GP	G	GA	PlusMin	Points	Direct	Directh	On-ice	On-iceh
2007												
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												
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

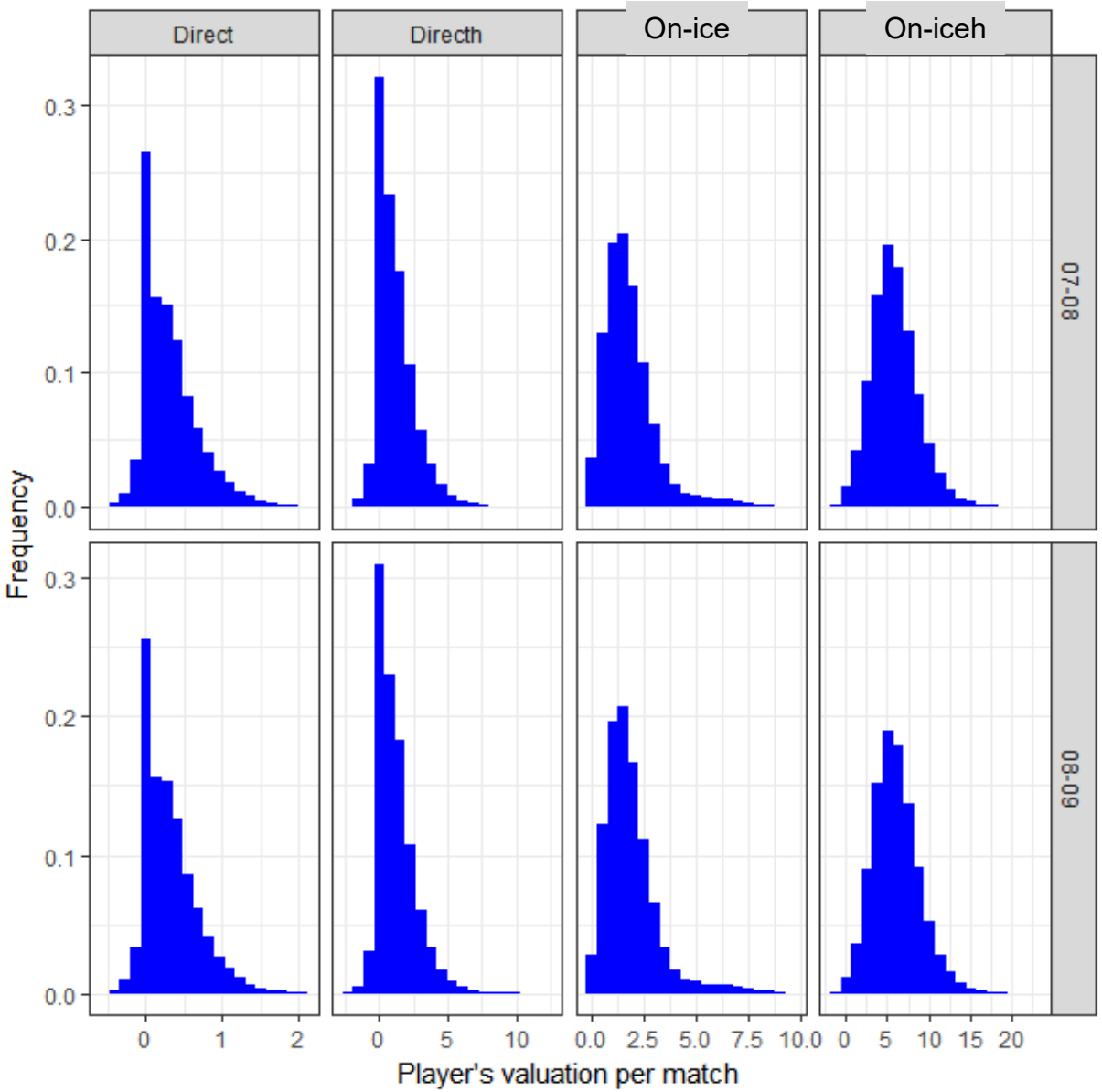
Table 5.1: Top 10 Players performance for 2007-2008 and 2008-2009 for the Direct metric.

Top players 2007-2008 and 2008-2009 for on-ice impact

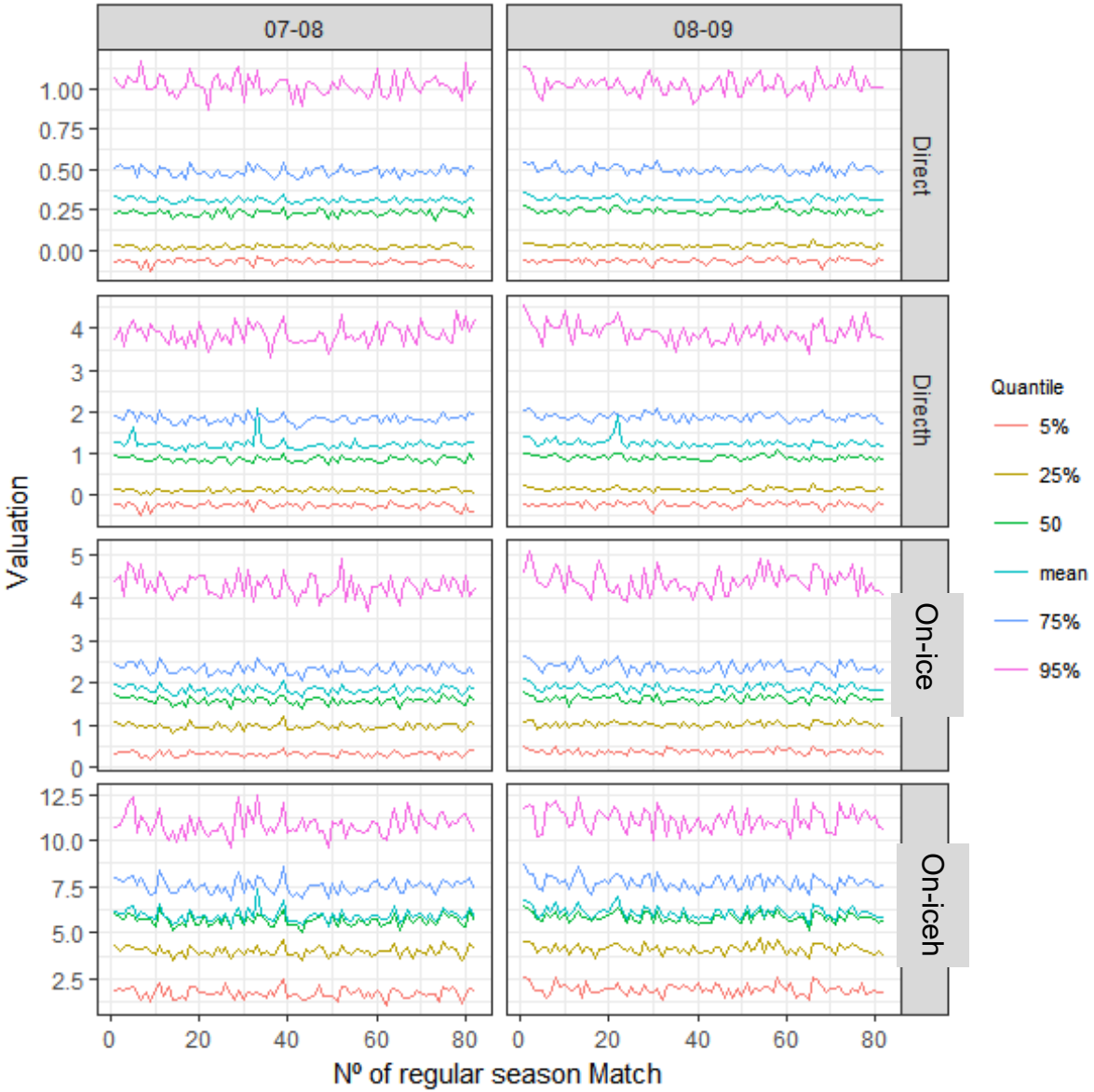
PlayerName	Position	Age	Salary	GP	G	GA	PlusMin	Points	Direct	Directh	On-ice	On-iceh
2007												
Dion Phaneuf	D	22	0.94	82	17	43	12	60	59.22	134.05	**246.12**	559.67
Alex Ovechkin	F	22	3.83	82	65	47	28	112	71.96	182.65	**232.56**	588.85
Tomas Kaberle	D	29	4.25	82	8	45	-8	53	38.32	93.36	**221.93**	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 Lidstrom	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 Backstrom	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

Table 5.4: Top 10 players performance for 2007-2008 and 2008-2009 for the (On-ice metric without goalkeeper positions

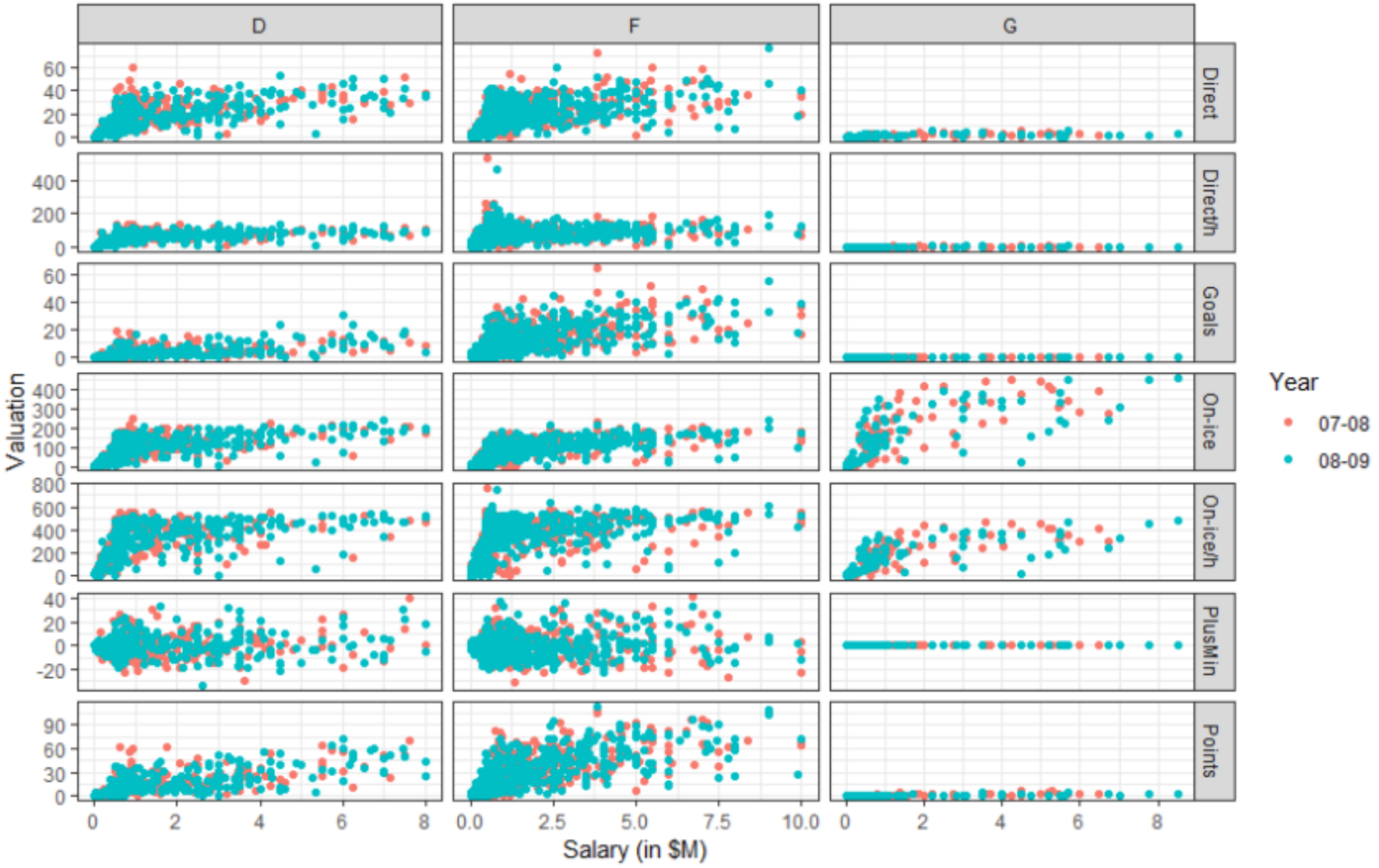
Distribution of impact values



Quantiles per game



Impact vs salary per position

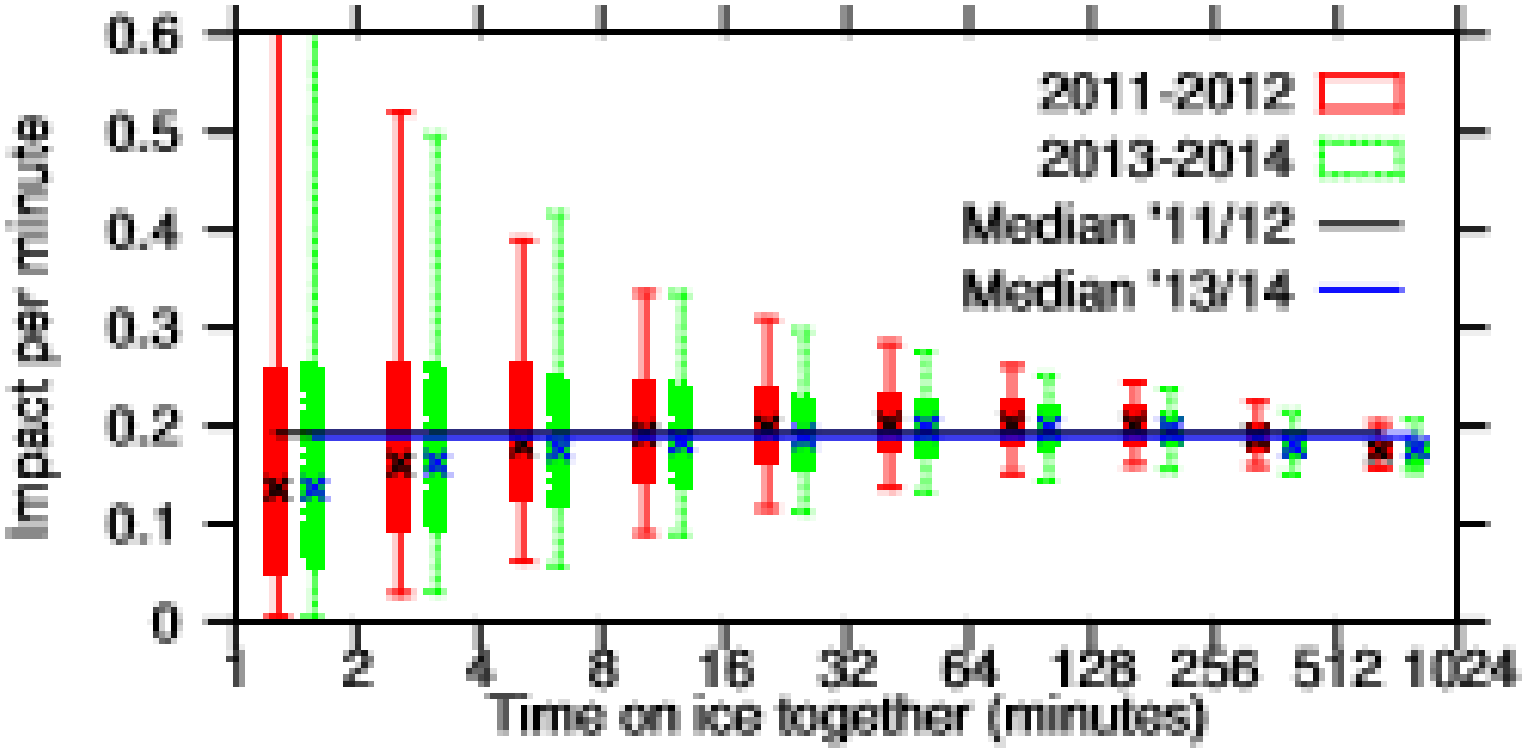


Top pairs 2011-2012

Table 3. Top pairs 2011-2012 according to total impact.

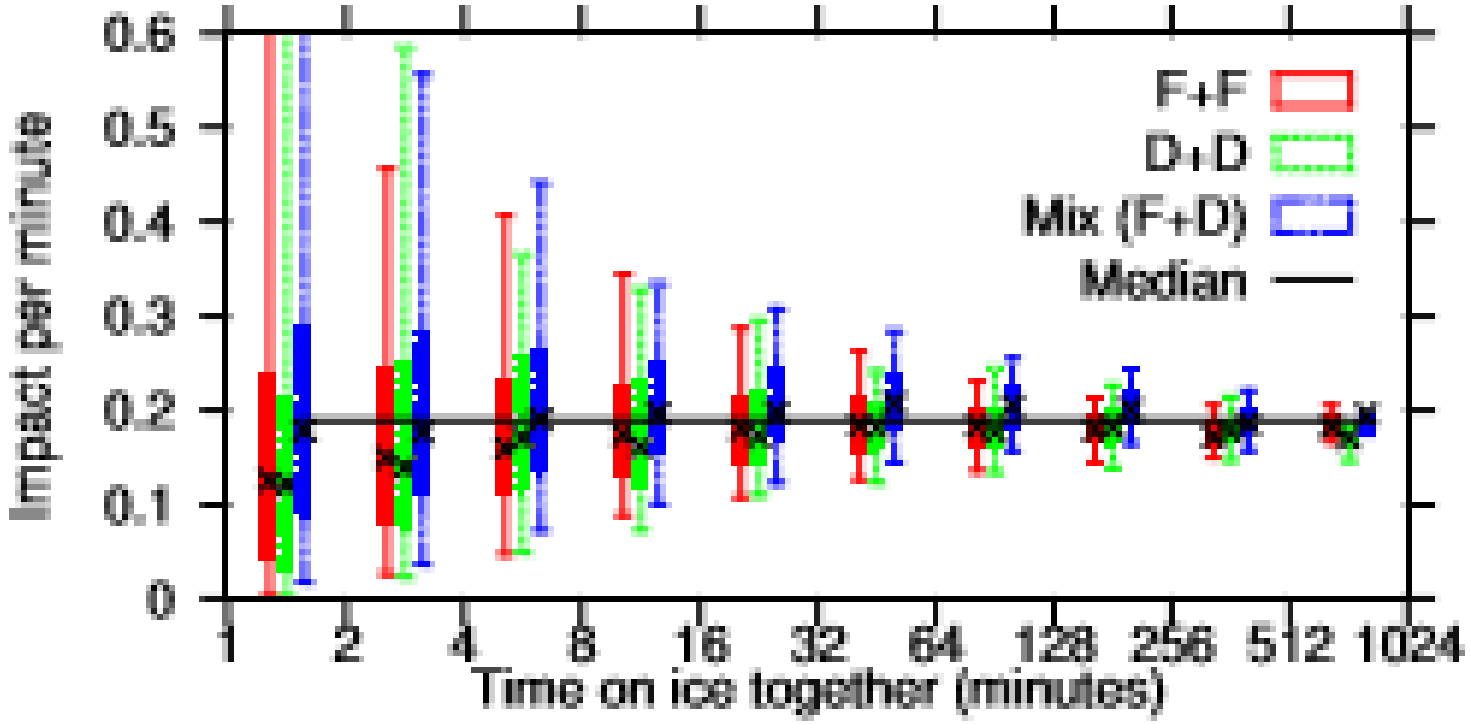
	Player 1					Player 2					Pair stats		
	Name	Pos	G	A	+/-	Name	Pos	G	A	+/-	Team	Impact	TOI
Forwards	Ilya Kovalchuk	R	37	46	-9	Zach Parise	L	31	38	-5	NJD	121.17	40,163
	Ryan O'Reilly	C	18	37	-1	Gabriel Landeskog	L	22	30	+20	COL	115.74	39,021
	Joe Pavelski	C	31	30	+18	Joe Thornton	C	18	59	+17	SJS	112.65	39,353
	Steven Stamkos	C	60	37	+7	Martin St. Louis	R	25	49	-3	TBL	111.77	35,941
	Milan Michalek	L	35	25	+4	Jason Spezza	C	34	50	+11	OTT	111.73	36,689
Defenders	Dan Girardi	D	5	24	+13	Ryan McDonagh	D	7	25	+25	NYR	155.28	55,911
	Filip Kuba	D	6	26	+26	Erik Karlsson	D	19	59	+16	OTT	134.74	47,985
	Francois Beauchemin	D	8	14	-14	Cam Fowler	D	5	24	-28	ANA	125.54	45,795
	Josh Gorges	D	2	14	+14	P.K. Subban	D	7	29	+9	MTL	125.16	44,390
	Carl Gunnarsson	D	4	15	-9	Dion Phaneuf	D	12	32	-10	TOR	123.06	36,181
Mixed	Jason Spezza	C	34	50	+11	Erik Karlsson	D	19	59	+16	OTT	110.58	35,990
	Joe Pavelski	C	31	30	+18	Dan Boyle	D	9	39	+10	SJS	106.04	35,612
	Joe Thornton	C	18	59	+17	Dan Boyle	D	9	39	+10	SJS	102.96	35,160
	Tomas Fleischmann	L	27	34	-7	Brian Campbell	D	4	49	-9	FLA	98.08	31,804
	Stephen Weiss	C	20	27	+5	Brian Campbell	D	4	49	-9	FLA	96.79	32,995

Impact per minute



Variation decreases when more joint TOI
Medians highest in 16-256 minutes joint TOI

Impact per minute



Mixed pairs may have higher impact

Outline

- Motivation
- Methods and Results
- Conclusion

Conclusions

- Introduced new goal-based performance metrics for ice hockey players
- Strong correlation for G, A, P between new and traditional metrics
- Pass the eye test

Future Work

- Combine variant 1 and 2 by using GPIV as the reward function in variant 2
- Alternative reward functions for the performance model
- Game prediction and season simulation

Thanks to our students

- Carles Sans Fuentes
- Dennis Ljung
- Jon Vik
- Min-Chun Shih
- Rabnawaz Jan-Sher
- Rasmus Säfvenberg
- Sofie Jörgensen
- Haris Kozlica
- Timmy Lehmus Persson

Further Reading

- Vik J, Shih M-C, Jansher, R, Carlsson N, Lambrix P, Not all goals are equally important - a study for the NHL, *MathSport International 2021 Conference*, 26-31, 2021.
- Lehmus Persson T, Kozlica H, Carlsson N, Lambrix P, Prediction of tiers in the ranking of ice hockey players, *7th Workshop on Machine Learning and Data Mining for Sports Analytics*, 89-100, 2020.
- Sans Fuentes C, Carlsson N, Lambrix P, Player impact measures for scoring in ice hockey, *MathSport International 2019 Conference*, 307-317, 2019.
- Ljung D, Carlsson N, Lambrix P, Player pairs valuation in ice hockey, *5th Workshop on Machine Learning and Data Mining for Sports Analytics*, 82-92, 2019.

- Papers available at the LiU Sports Analytics Group page:

https://www.ida.liu.se/research/sportsanalytics/