

# 11. Risiko, Belohnung und Verstärkungslernen in der Eishockey-Analytik

Sheng Xu, Oliver Schulte, Yudong Luo, Pascal Poupart, Guiliang Liu

## Abstrakt

Viele Entscheidungen im Sport sind deshalb so schwierig, weil sie eine Abwägung zwischen Risiko und Belohnung beinhalten. Aktionen wie ein Drei-Punkte-Schuss, das Tragen eines Pucks oder das Dribbeln mit einem Ball bergen ein höheres Risiko des Scheiterns und erfordern außergewöhnliche Fähigkeiten, bringen aber auch das Potential für eine höhere Belohnung mit sich. In diesem Beitrag werden computergestützte Tools für die *Risikoanalyse* beschrieben, um das Risiko zu modellieren, das mit den Entscheidungen von Teams und Sportlern verbunden ist. Wir nutzen *verteilungsbasiertes Verstärkungslernen* (RL) als Quelle von Konzepten und Techniken für die rechnergestützte Risikoanalyse. Verteilungs-RL-Techniken ermöglichen es uns, eine dynamische Verteilung der Ergebnisse von mehr als 1000 Spielen der National Hockey League zu modellieren. Wir finden starke Belege dafür, dass starke Mannschaften viele Risiken eingehen (0,90 Korrelation zwischen dem Saisonrang der Mannschaft und der Standardabweichung der Saison). Bei den Spielern finden wir gute Belege dafür, dass stärkere Spieler mehr Risiken eingehen (z.B. 0,86 Korrelation zwischen den Saisontoren eines Spielers und seiner Value-at-Risk-Kennzahl).

## 11.1 Einleitung: Risiken im Sport eingehen

Bei vielen Entscheidungen von Sportlern und Trainern geht es darum, ein höheres Risiko für eine potenziell höhere Belohnung in Kauf zu nehmen. Ein bekanntes Beispiel aus dem Basketball ist die Entscheidung zwischen einem Weitwurf, der potenziell drei Punkte bringt, und einem Schuss aus kürzerer Entfernung, der zwei Punkte bringt. Ein komplexeres Beispiel aus dem Eishockey ist das Abziehen des Torwarts, d. h. die Auswechslung eines Angreifers für den Torwart, wenn eine Mannschaft in Rückstand gerät. Wird der Torwart früher abgezogen, erhöht sich die Chance auf den Ausgleich, aber auch die Chance, dass die führende Mannschaft ein Tor erzielt, was das Spiel praktisch sofort entscheidet. Während die entscheidungstheoretisch optimale Wahl darin besteht, *den durchschnittlichen Erfolg*, d. h. die durchschnittliche Anzahl von Punkten und Siegen, zu *maximieren*, haben mehrere Sportanalysten beobachtet, dass Spieler und Trainer oft von einem sekundären Ziel beeinflusst werden, nämlich die *Wahrscheinlichkeit eines Misserfolgs zu minimieren*, oder allgemein die Wahrscheinlichkeit schlechter Ergebnisse. Zur Veranschaulichung dieses Punktes im Basketball-Szenario: Nehmen wir an, ein Spieler befindet sich in einer Situation, in der die Wahrscheinlichkeit, einen Dreipunktwurf zu erzielen, 20 % und die Wahrscheinlichkeit, einen Zweipunktwurf zu erzielen, 30 % beträgt. Dann ist die erwartete Punktzahl für beide Möglichkeiten gleich (nämlich 0,6). Die Wahrscheinlichkeit, dass der Dreipunktwurf misslingt, liegt jedoch bei 80 %, der Zweipunktwurf dagegen nur bei 70 %; die Wahrscheinlichkeit eines schlechten Ergebnisses ist beim Zweipunktwurf geringer. Betrachten wir nun eine andere Situation, in der der Weitschuss eine Erfolgswahrscheinlichkeit von 25 % hat. In diesem Fall beträgt die erwartete Anzahl der Punkte aus der Distanz 0,75 im Vergleich zu 0,6, und die optimale Entscheidung ist der Distanzwurf. Die Wahrscheinlichkeit eines Misserfolgs beträgt jedoch immer noch 75 % gegenüber nur 70 % für den Zweipunktwurf, so dass ein *risikoscheuer* Spieler immer noch den sichereren Zweipunktwurf vorziehen könnte.

Mehrere Sportanalysten haben argumentiert, dass Spieler und Mannschaften dazu neigen, risikoscheue Entscheidungen auf Kosten ihres Gesamterfolgs zu treffen, der über viele Spiele und Spielsituationen gemittelt wird. Pelechris (2016) liefert Belege dafür, dass American-Football-Trainer darauf abzielen, die

Varianz der erwarteten Punkte zu minimieren, und nicht die erwarteten Punkte direkt, vielleicht um öffentliche Kritik für Misserfolge zu vermeiden. Beaudoin und Swartz (2010) argumentierten, dass Eishockeymannschaften bei einem Rückstand ihre Torhüter früher abziehen sollten, um die Gewinnchancen zu maximieren. In der Tat haben NHL-Teams in letzter Zeit damit begonnen, ihre Torhüter früher abziehen.<sup>1</sup> Ein weiterer Beleg für einen Trend zu mehr Risikobereitschaft ist der Anstieg der Dreipunktversuche im Basketball, die von 22,2 % in der Saison 2010/11 auf 39,2 % in der Saison 2020-2021 steigen.<sup>2</sup> Der Trend zu risikoreicheren Aktionen im Laufe der Zeit stützt die These, dass Risikobereitschaft einen Vorteil gegenüber vorsichtigeren Taktiken bietet.

In diesem Beitrag untersuchen wir die Risikobereitschaft in der National Hockey League. Wir untersuchen verschiedene Möglichkeiten zur Quantifizierung des Risikos einer Entscheidung, darunter traditionelle Begriffe wie die Varianz/Standardabweichung der Ergebnisse und die Gini-Abweichung, ein alternatives Variabilitätskonzept, das sich gut für multimodale Verteilungen eignet (Luo et al., 2023). Aus technischer Sicht *erfordert die Modellierung des Risikos die Modellierung von Momenten höherer Ordnung der Verteilung der möglichen Handlungsergebnisse*. Um die Ergebnisverteilung über ihren Mittelwert hinaus zu modellieren, nutzen wir die jüngsten Arbeiten im Bereich des *verteilungsorientierten Verstärkungslernens* (VL). Das Verstärkungslernen ist ein Zweig des maschinellen Lernens, der untersucht, wie man sich in sequenziellen Entscheidungsszenarien verhält, wie wir sie in der Sportanalyse antreffen. Das Verstärkungslernen hat eine umfangreiche Reihe von Methoden zur Schätzung der erwarteten Ergebnisse von Handlungen entwickelt, die als Vorhersage- oder Strategiebewertungsmethoden bekannt sind (Sutton & Barto, 1998). Verteilungsbasiertes Verstärkungslernen ist eine neuere Entwicklung, die Methoden zur Modellierung der Verteilung von Handlungsergebnissen bietet. In einer kürzlich erschienenen Arbeit von Liu et al. (2022) wurde eine verteilungsbezogene RL-Methode zur Schätzung der Verteilung von Handlungsergebnissen für Play-by-Play-Daten (Ereignisse) entwickelt. Wir nutzen die Berechnungswerkzeuge aus ihrer Arbeit, um die Verteilungen von Handlungsergebnissen aus großen Play-by-Play-Ereignisdaten (1000+ Spiele, 1M+ Ereignisse) zu schätzen.

Wir wenden ihre Methodik an, um die Risikobereitschaft von Profispielern in der NHL zu quantifizieren und zu untersuchen. Die Hauptfrage ist, ob die Leistung mit der Risikobereitschaft zusammenhängt, und zwar sowohl für Teams als auch für Spieler. Wir untersuchen drei verschiedene Variabilitätskonzepte zur Quantifizierung des mit einer Ergebnisverteilung verbundenen Risikos: Standardabweichung, Gini-Abweichung und Value at Risk. Die Risikoauswirkung einer Aktion ist das Ausmaß, in dem sie die Variabilität der Spielergebnisse für das Team des handelnden Spielers erhöht/verringert.

Unsere wichtigsten Ergebnisse sind die folgenden: Für die *Mannschaftsleistung* zeigt das Gesamtrisiko der von einer Mannschaft in einer Saison ergriffenen Maßnahmen eine sehr hohe Korrelation mit der Mannschaftsleistung, gemessen an der Anzahl der gesamten Saisonpunkte. Bei Verwendung der Standardabweichung als Risikomaß beträgt die Korrelation 0,90, bei der Gini-Abweichung ebenfalls 0,90. Bei der zuvor von Liu et al. (2022) untersuchten Value-at-Risk-Kennzahl beträgt die Korrelation nur 0,51 bei einem Konfidenzniveau von 0,2, das die Risikobereitschaft widerspiegelt (siehe Abschnitt 11.1 für weitere Einzelheiten).

---

<sup>1</sup> <https://www.nhl.com/news/coaches-room-paul-maclean-nhl-pulling-goalies-trend-301099580>

<sup>2</sup> <https://www.nba.com/news/3-point-era-nba-75>



Abbildung 11.1: Systemkomponenten in unserer Risikoanalyse-Methodik

Für die *Spielerleistung* verwenden wir das Gesamtrisiko der von einem Spieler in einer Saison durchgeführten Aktionen als Maß für die Risikobereitschaft des Spielers. Alle Risikokennzahlen der Spieler weisen ein hohes Maß an zeitlicher Konsistenz auf, wobei ihre Risikowerte von Runde zu Runde schon nach weniger als zur Hälfte der Saison zum Gesamtwert von der Saison konvergieren. Die Metrik der Standardabweichung des Spielers erreicht eine substantielle Korrelation von 0,51 mit den Gesamttoren des Spielers in einer Saison. Die Value-at-Risk-Metrik erreicht eine noch höhere Torkorrelation von 0,86 (bei einem Konfidenzniveau von 0,2 für die Risikobereitschaft). Diese Ergebnisse belegen, dass die Variabilitäts-Risiko-Metriken sehr gut geeignet sind, um den Teamerfolg vorherzusagen, aber weniger geeignet, um den Spielerfolg vorherzusagen. Da es jedoch kein unabhängiges Maß für Spielerstärke gibt, sehen wir die Ergebnisse für Spieler als vorläufig an, und die Untersuchung der Risikobereitschaft von Spielern ist eine wertvolle Richtung für zukünftige Forschung. Unsere Studie konzentriert sich zwar auf Eishockey-Daten aus der National Hockey League (NHL), aber unsere Methoden lassen sich auf jeden Play-by-Play-Datensatz anwenden; siehe Liu et al. (2022) für eine Anwendung auf Fußballdaten.

**Gliederung des Papiers.** Unser Beitrag ist wie folgt aufgebaut. Wir beginnen mit einem Überblick über die Regeln des Eishockeys und unseren Play-by-Play-Datensatz. Dann gehen wir auf den Hintergrund des Reinforcement Learning (RL) ein, insbesondere auf die verteilungsorientierten RL-Techniken zum Lernen der Verteilung von Handlungsergebnissen aus Play-by-Play-Ereignisdaten. Unsere Diskussion konzentriert sich auf die wichtigsten Prinzipien und Intuitionen; Details zu unseren Lernmethoden finden Sie im Anhang und in den Referenzen. Ausgehend von einer dynamischen Verteilung der zukünftigen Handlungsergebnisse zu jedem Zeitpunkt eines Spiels definieren wir die Risikoauswirkung einer Handlung als die Erhöhung/Verringerung des Verteilungsrisikos aufgrund dieser Handlung. Die gesamte Risikoauswirkung aller Aktionen wird verwendet, um die Risikobereitschaft von Mannschaften und Spielern zu quantifizieren. Abbildung 11.1 fasst die Komponenten unseres Systems zusammen.

## 11.2 Eishockeyregeln und Eishockeydaten

**NHL-Regeln** Wir geben einen kurzen Überblick über die Spielregeln in der NHL (National Hockey League, 2014). NHL-Spiele bestehen aus drei Spielabschnitten, die jeweils 20 Minuten dauern. Ein Team muss innerhalb von drei Spielabschnitten mehr Tore als der Gegner erzielen, um das Spiel zu gewinnen. Bei ausgeglichener Spielstärke haben die Mannschaften fünf Schlittschuhläufer und einen Torwart auf dem Eis. Strafen führen dazu, dass ein Spieler auf der Strafbank sitzt. Der bestrafte Spieler wird für 2, 4, 5 oder 10 Minuten auf die Strafbank geschickt, und das bestrafte Team ist in Unterzahl, was zu einem Mannschaftsunterschied zwischen den beiden Teams führt. Die Periode, in der ein Team bestraft wird, wird als Powerplay für das gegnerische Team mit einem Mannschaftsvorteil bezeichnet. Ein Tor in

Unterzahl ist ein Tor, das von der bestraften Mannschaft erzielt wird, und ein Powerplay-Tor ist ein Tor, das von der Mannschaft im Powerplay erzielt wird.

**Datensatz** In diesem Beitrag verwenden wir einen proprietären Play-by-Play-Datensatz, der von Sportlogiq<sup>3</sup> erstellt wurde. Die Daten wurden durch eine Kombination aus Computer Vision und manueller Annotation erstellt. Der Datensatz enthält insgesamt 1196 Spiele, die zwischen dem 3. Oktober 2018 und dem 6. April 2019 gespielt wurden. Der Trainingsdatensatz für die Erstellung unseres Modells enthält 956 Spiele (vom 3. Oktober 2018 bis zum 24. Februar 2019). Tabelle 11.1 listet die in unserer Analyse verwendeten Merkmale auf (Liu et al., 2022) und Abbildung 11.2 zeigt die angepassten Koordinaten im Sportlogiq-Datensatz. Abbildung 11.2 zeigt ein schematisches Layout der Eishockeystrecke. Die Einheiten sind Fuß. Die angepassten Y-Koordinaten verlaufen von -42,5 am Boden bis 42,5. Die Torlinie liegt bei X = 89.

Tabelle 11.1: Die vollständige Liste der Spielmerkmale für den Eishockey-Datensatz. Die Tabelle verwendet angepasste räumliche Koordinaten, wobei negative Zahlen die defensive Zone des handelnden Spielers und positive Zahlen die offensive Zone bezeichnen.

Typ	Name	Bereich	
Eis Hockey	Räumliche Merkmale	X-Koordinate des Pucks Y-Koordinate des Pucks Geschwindigkeit des Pucks Winkel zwischen dem Puck und dem Tor	[-100, 100] [-42.5, 42.5] ( $-\infty$ , $+\infty$ ) [-3.14, 3.14]
	Zeitliche Eigenschaften	Verbleibende Spielzeit Dauer des Ereignisses (Aktion)	[0, 3,600] (0, $+\infty$ )
	In-Game- Funktionen	Punktedifferenz Personelle Situation Heim- oder Auswärtsteam Ergebnis der Aktion	( $-\infty$ , $+\infty$ ) {Gleiche Stärke, Unterzahl, Power Play} {Zu Hause, Auswärts} {Erfolgreich, Misserfolg}

Der Datensatz zeichnet Ereignisdaten auf, die als Play-by-Play-Daten bekannt sind. Play-by-Play-Daten spezifizieren den Zeitpunkt und den Ort von Aktionen, identifizieren den für die jeweilige Aktion verantwortlichen Spieler und enthalten die in Tabelle 11.1 beschriebenen **kontextbezogenen Ereignismerkmale**. In Tabelle 11.2 sind die häufigsten Aktionstypen in unserem Datensatz aufgeführt. Zur Veranschaulichung der Play-by-Play-Daten zeigt Tabelle 11.3 einen Ausschnitt.

<sup>3</sup> <https://sportlogiq.com>

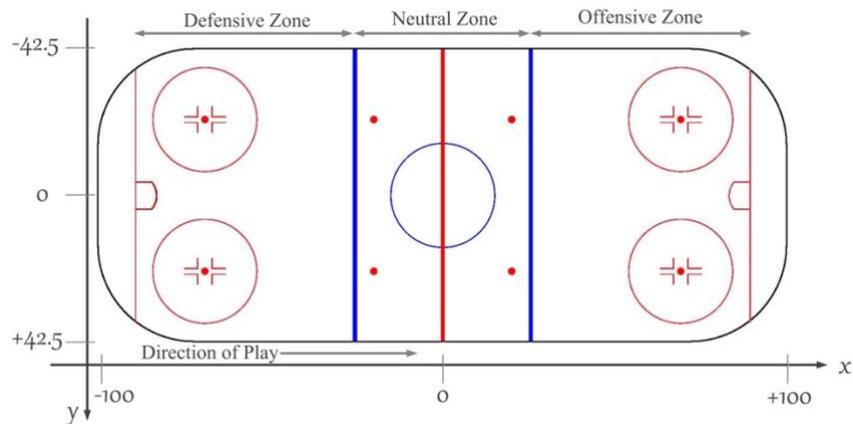


Abbildung 11.2: Spielfeldanordnung mit angepassten Koordinaten. Die Koordinaten werden so angepasst, dass die Offensivzone des Teams, das eine Aktion durchführt, rechts liegt.

Tabelle 11.2: Definition der häufigsten Aktionstypen im Datensatz

Aktion	Beschreibung
Block	Versuch, die Flugbahn des Pucks zu blockieren
Tragen	Kontrollierter Transport des Pucks über eine blaue Linie oder die rote Mittellinie
Check	Wenn ein Spieler versucht, seinen Körper einzusetzen, um einen Gegner vom Puck zu trennen
Abladen	Wenn ein Spieler den Puck in die Angriffszone schickt
Auskippen	Wenn ein verteidigender Spieler den Puck in die Bande spielt, ohne einen Mannschaftskameraden für einen Pass anzuvisieren
lpr	Loser Puck wird geborgen. Der Spieler hat den Puck erobert, als kein Spieler ihn kontrollierte
Abseits	Wenn ein Spieler jenseits der blauen Angriffslinie erwischt wird, bevor sein Mitspieler den Puck in die Angriffszone gebracht haben
Pass	Der Spieler versucht einen Pass zu einem Teamkollegen
Puck-Schutz	Wenn ein Spieler seinen Körper einsetzt, um den Puck entlang der Bande zu schützen
Annahme	Wenn ein Spieler einen Pass von einem Teamkollegen erhält
Schuss	Ein Spieler schießt auf das Tor
Schuss gegen	Ein Schuss wurde von der gegnerischen Mannschaft abgegeben

Tabelle 11.3: Beispielhafter Auszug aus den Play-By-Play-Daten.

gameld	playerId	Zeitraum	team	xKoordinate	yKoordi-	Arbeits	Aktion	Typ
--------	----------	----------	------	-------------	----------	---------	--------	-----

		ld		nate	kräfte		
849	402	1	15	-9.5	1.5	gleich	lpr
849	402	1	15	-24.5	-17	gleich	tragen
849	417	1	16	-75.5	-21.5	gleich	check
849	402	1	15	-79	-19.5	gleich	Puck-Schutz
849	413	1	16	-92	-32.5	gleich	lpr
849	413	1	16	-92	-32.5	gleich	Pass
849	389	1	15	-70	42	gleich	Block
849	389	1	15	-70	42	gleich	lpr
849	389	1	15	-70	42	gleich	Pass
849	425	1	16	-91	34	gleich	Block
849	395	1	15	-97	23.5	gleich	Annahme

### 11.3 Markov-Spielmodelle für den Sport

Das Verstärkungslernen bietet ein reichhaltiges Instrumentarium zur Abschätzung der zukünftigen Erfolgchancen strategischer Akteure. Schulte (2022) gibt eine kurze, leicht zugängliche Einführung in die Anwendung von RL in der Sportanalytik. Für Einzelagentenprobleme basiert RL auf dem grundlegenden Markov-Entscheidungsprozessmodell. Die Verallgemeinerung des Markov-Entscheidungsprozesses auf mehrere Entscheidungsträger führt zum Markov-Spiel-Modell. Markov-Spiel-Modelle wurden für verschiedene Sportarten wie Eishockey, Fußball und American Football entwickelt (Chan et al., 2020; Liu et al., 2020; Liu & Schulte, 2018). Wir verwenden das Eishockeymodell von Liu und Schulte (2018).

#### 11.3.1 Markov-Spielmodell für NHL-Eishockey

Ähnlich wie (Liu & Schulte, 2018) wenden wir das Markov Game Framework an, um die Spieldynamik für Sportspiele zu modellieren. Ein Markov-Spiel (Littman, 1994), das manchmal auch als stochastisches Spiel bezeichnet wird, ist definiert durch eine Menge von Zuständen  $\mathcal{S}$  und eine Sammlung von Aktionssätzen  $\mathcal{A}$ , eine für jeden Agenten in der Umgebung. Zustandsübergänge werden durch den aktuellen Zustand und eine Liste von Aktionen gesteuert, eine Aktion von jedem Agenten. Für jeden Agenten gibt es eine zugehörige Belohnungsfunktion, die einem Zustandsübergang eine Belohnung zuordnet. Ein Überblick darüber, wie ein Hockey-Markov-Spiel-Modell dieses Schema ausfüllt, ist wie folgt.

- Es gibt zwei Agenten, *Home* und *Away*, die ihre jeweiligen Teams vertreten.
- Die **Aktion**  $a_t$  bezeichnet die Bewegungen der Spieler, die den Puck kontrollieren. Unser Modell verwendet einen diskreten Aktionsvektor mit einer Ein-Schuss-Darstellung.
- Eine **Beobachtung** ist ein Merkmalsvektor  $x_t$  der einen Wert der in Tabelle 11.1 aufgeführten Merkmale zu einem diskreten Zeitschritt  $t$  angibt. Wir verwenden die vollständige Sequenz  $s_t \equiv (x_t, a_t, x_{t-1}, \dots, x_0)$  um den **Zustand** darzustellen (Mnih et al., 2015).



Abbildung 11. 3: Ein Eishockeyspiel wird in Torschussepisoden unterteilt

- Da wir zwei Agenten haben, haben wir zwei **Belohnungsfunktionen**, eine für das Heimteam und eine für das Auswärtsteam. Die Belohnung zum Zeitpunkt  $t$  ist 1 für eine Mannschaft, die ein Tor zum Zeitpunkt  $t$ , 0, wenn kein Tor erzielt wird; wir schreiben  $\text{goal}_{t,Home}$ ,  $\text{goal}_{t,Away}$ . Zum Beispiel,  $\text{goal}_{t,Home} = 1$  bedeutet, dass die Heimmannschaft ein Tor zum Zeitpunkt  $t$  erzielt. Es ist technisch nützlich, einen virtuellen "none"-Agenten für den Fall einzuführen, dass keine Mannschaft vom Zeitpunkt  $t$  bis zum Ende des Spiels ein Tor erzielt. Wenn keine der beiden Mannschaften am Ende des Spiels ein Tor schießt, schreiben wir  $\text{goal}_{T,Neither}$  wobei  $T$  der letzte Zeitschritt ist.

Das **Modell der erwarteten Tore**  $R_k(s_t, a_t) = P(\text{goal}_{t,k} = 1 | s_t, a_t)$  gibt die Wahrscheinlichkeit an, dass eine Mannschaft nach einer Aktion in einem bestimmten Spielzustand ein Tor erzielt. Dieses Modell geht von der **Markov-Annahme aus**, die besagt, dass die zum Zeitpunkt  $t$  verfügbare Zustandsinformation so informativ ist, dass die Torchancen nur auf der Grundlage des aktuellen Zustands geschätzt werden können, unabhängig vom aktuellen Zeitpunkt  $t$  und früheren Zuständen. Technisch gesehen ist das Markov-Spielmodell stationär und die Torwahrscheinlichkeit ist unabhängig vom aktuellen Zeitindex  $t$ . In ähnlicher Weise werden die **Übergangswahrscheinlichkeiten**  $P(s_{t+1}, a_{t+1} | s_t, a_t)$  als stationär angenommen und hängen nur vom aktuellen Spielzustand ab. Das bedeutet, dass für einen gegebenen Spielzustand  $s_t$  und Aktion  $a_t$  zum Zeitpunkt  $t$  die **Dynamik des NHL-Sports eine Verteilung über zukünftige Spielverläufe definiert, die nur vom aktuellen Zustand und der aktuellen Aktion abhängt**.

### 11.3.2 Die Erwartungswertfunktion

Wir unterteilen ein Sportspiel in **Torschussepisoden**, so dass jede Episode: 1) am Anfang des Spiels oder unmittelbar nach einem Tor beginnt und 2) mit einem Tor oder am Ende des Spiels endet ( $s_H$ ). Episoden erstrecken sich über Spielunterbrechungen.

Eine Schlüsselgröße beim Verstärkungslernen ist die *erwartete Belohnung* in Bezug auf künftige Trajektorien. Angesichts unserer binären Belohnung (Tor erzielen oder nicht) ist die erwartete Belohnung für ein Team die Chance, das nächste Tor zu erzielen, bezeichnet als  $Q_k(s_t, a_t)$ . Um den grundlegenden RL-Ansatz zum Lernen einer Q-Funktion zu erläutern, betrachten wir zunächst die erwartete Belohnung für einen begrenzten *Horizont*, d. h. eine feste Vorausschau-Länge  $H$ . Die Chance, innerhalb der nächsten  $H$  Schritte ein Tor zu erzielen, wird dann durch den Ausdruck definiert:

$$Q_k^H(s_t, a_t) = \sum_{h=0}^H \gamma^h P(\text{goal}_{t+h,k} = 1 | s_t, a_t) \quad (1)$$

Wie in Vorgängerstudien (Liu & Schulte, 2018; Liu et al., 2020), wählen wir  $\gamma = 1$ . In diesem Fall zählen alle Tore gleich, und der Wert  $Q_k^H$  ist einfach die Wahrscheinlichkeit, dass Team  $k$  innerhalb von  $H$  Schritten ein Tor erzielt.

Der Q-Wert erfüllt eine wichtige Rekursionsbeziehung, die als Update der dynamischen Programmierung bekannt ist:

$$Q_k^{H+1}(s_t, a_t) = R_k(s_t, a_t) + \gamma E_{s_{t+1}, a_{t+1} \sim P(s', a' | s_0 = s_t, a_0 = a_t)} [Q_k^H(s_{t+1}, a_{t+1})] \quad (2)$$

Das Prinzip der dynamischen Programmierung ist, dass eine Mannschaft nur dann ein Tor in  $H+1$  Schritten erzielt, wenn sie (1) sofort ein Tor erzielt oder (2) einen weiteren Schritt macht und dann innerhalb von  $H$  Schritten ein Tor erzielt. Der Wert-Iterationsalgorithmus verwendet die dynamische Programmierung zur

Schätzung der Q-Werte aus einem Datensatz  $\mathcal{D}$  von beobachteten Spielfolgen wie folgt.

1. Initialisieren Sie die Q-Werte für  $H = 0$  mit dem erwarteten Zielmodell  $R$ .
2. Gleichung (2) iterativ anwenden durch  $H = 1, H = 2, \dots$ , bis zur Konvergenz, wobei  $Q_k^{H+1}(s_t, a_t) \approx Q_k^H(s_t, a_t)$  für jedes Team und jedes Zustands-Aktions-Paar. Wir bezeichnen den **konvergenten Q-Wert** als  $Q_k(s_t, a_t)$ .

Für kontinuierliche Zustandsräume, wie wir sie beim Eishockey haben, kann der Erwartungswert  $E_{s_{t+1}, a_{t+1} \sim P(s', a' | s_0 = s_t, a_0 = a_t)}$  durch Mittelwertbildung über alle Übergänge geschätzt werden  $s_t, a_t; s_{t+1}, a_{t+1}$  die im Datensatz beobachtet werden. Im NHL-Modell von Schulte et al. (2017b) trat Konvergenz mit einer Vorlaufzeit von  $H = 13$  auf. Die Tatsache, dass Erwartungswerte in RL eine Vorlaufzeit beinhalten, bedeutet, dass sie die mittelfristigen Auswirkungen von Aktionen auf die Torausbeute erfassen können.

Mit einer unbeschränkten Vorausschau  $H \rightarrow \infty$  konvergiert die Wertiteration unter milden Bedingungen zu einem Fixpunkt, der die Bedingungen erfüllt:

$$Q_k(s_t, a_t) = R_k(s_t, a_t) + \gamma E_{s_{t+1}, a_{t+1} \sim P(s_{t+1}, a_{t+1} | s_t, a_t)}[Q_k(s_{t+1}, a_{t+1})] \quad (3)$$

die als **Bellman-Gleichung** für die Strategiebewertung bekannt ist. Im Anhang Abschnitt 8.1 wird erörtert, wie die Bellman-Gleichung zum Lernen eines neuronalen Netzmodells der Q-Funktion verwendet werden kann.

**Bemerkung.** Für Leser, die mit Reinforcement-Learning-Modellen vertraut sind, möchten wir kurz unser NHL-Modell mit anderen RL-Modellen vergleichen. Andere Leser können diesen Absatz überspringen, ohne Verlust der Kontinuität. Unser Lernsetting ist das Offline-Lernen, bei dem wir eine Wertfunktion aus einem Datensatz lernen, ohne Aktionen auszuführen; das heißt, unser Problem ist die Vorhersage von Ergebnissen und nicht die Kontrolle von Aktionen. In der Offline-Perspektive können die beobachteten Aktionen als ein weiteres Merkmal ähnlich wie Zustände behandelt werden. Formal gesehen ist das, was wir definiert haben, ein *Markov-Belohnungsprozess* in einem erweiterten Zustandsraum  $\mathcal{S} \times \mathcal{A}$  wobei ein erweiterter Zustand ein Paar ist  $(s, a)$  (vgl. (Sutton & Barto, 1998, Kap.6.4)). Die Q-Funktion, wie wir sie definiert haben, ist die Wertfunktion dieses Markov-Belohnungsprozesses. Wir haben die Q-Notation anstelle von V für Wertfunktion verwendet, weil ihre Bedeutung dieselbe ist wie bei der Politikbewertung: die erwartete kumulative Belohnung für einen Agenten bei einer aktuellen Aktion und einem aktuellen Zustand. Ein äquivalentes Modell würde darin bestehen, zunächst eine Strategie  $\pi_{Home}$  für die Heimmannschaft und eine andere Strategie  $\pi_{Away}$  für die Auswärtsmannschaft zu schätzen. Wenn zum Beispiel im Zustand  $s$  die Heimmannschaft den Puck in 30 % der Fälle abgibt, könnten wir schätzen  $\pi_{Home}(pass|s) = 30\%$ . Die Q-Funktion, wie wir sie definiert haben, stellt  $Q^{\pi_{Home}, \pi_{Away}}$ . Dies ist die Q-Funktion des NHL-Markov-Spiels, bei dem die Heimmannschaft und die Auswärtsmannschaft die in den Daten gezeigten Verhaltensregeln mit Aktionshäufigkeiten befolgen. Auch wenn das Markov-Belohnungsmodell für Offline-Daten vielleicht weniger vertraut ist als die Formulierung zur Bewertung der Politik, verwenden wir es, weil es konzeptionell einfacher ist und in der Tat gut zur Position des Sportanalysten passt, der die Spiele passiv beobachtet: Entscheidungen der Spieler sind Ereignisse, die der Sportanalyst analysieren muss, und nicht Entscheidungen, die er kontrollieren kann. Für weitere Diskussionen über die Reduzierung der Offline-Markov-Spielanalyse auf andere RL-Modelle siehe (Luo et al., 2020).

### 11.3.3 Lernen von Belohnungsverteilungen

Distributional RL lernt die Verteilung der Zufallsvariablen  $Z^k(s_t, a_t)$  die die Summe der (diskontierten) Belohnungen für zukünftige Episodentrajektorien liefert, die mit dem Zustand  $s_t$  und der Aktion  $a_t$  beginnen (Bellemare et al., 2017). Daher ist der Q-Wert die Erwartung der Z-Variablen:  $Q_k(s_t, a_t) = E(Z_k(s_t, a_t))$ . Ähnlich wie der Q-Wert folgt auch die Verteilung der Gesamtbelohnungen  $Z$  der **Verteilungs-Bellman-Gleichung**:

$$Z_k(s_t, a_t) \triangleq R_k(s_t, a_t) + \gamma E_{s_{t+1}, a_{t+1} \sim P(s_{t+1}, a_{t+1} | s_t, a_t)} [Z_k(s_{t+1}, a_{t+1})] \quad (4)$$

wobei  $X \stackrel{\Delta}{=} Y$  anzeigt, dass die Zufallsvariablen  $X$  und  $Y$  der gleichen Verteilung folgen. Bei einer Darstellung der  $Z_k$  Verteilungen die für einen Rechner angemessen ist, können wir iterativ die Bellman-Aktualisierungsgleichung (4) anwenden, um die Verteilung für verschiedene Vorausschau-Längen zu aktualisieren  $H = 1, H = 2, \dots$ , bis wir zu einer **konvergenten** Verteilung gelangen  $Z_k$ .

**Darstellung der Belohnungsverteilung.** Um die Verteilungs-Bellman-Gleichung anwendbar zu machen, stellt sich die Frage, wie man eine rechnerisch angemessene Darstellung wählt, die das Lernen unterstützt. Eine Möglichkeit besteht darin, eine parametrische Familie zu wählen, z. B. eine Gauß-Verteilung. Das Problem bei solchen parametrischen Familien ist, dass sie in der Regel unimodal sind. Unimodale Verteilungen sind für die komplexe Dynamik des Sports nicht geeignet, da Ereignisse in der Regel einen hohen Verzweigungsfaktor aufweisen, wobei verschiedenen Zweigen verschiedenen Modi entsprechen. Wenn beispielsweise ein Spieler einen Pass versucht, gibt es drei mögliche Ergebnisse: Der Pass wird abgefangen, erreicht den vorgesehenen Empfänger oder wird zu einem losen Puck. Für jedes dieser Ergebnisse gibt es alternative Folgeereignisse, usw. Die verschiedenen möglichen Ereignisfolgen bestimmen unterschiedliche Trefferwahrscheinlichkeiten, was zu einer sehr multimodalen Verteilung führt.

Bellemare et al. (2017) schlugen vor, Belohnungsverteilungen mithilfe von **Quantilsregression** zu modellieren; Liu et al. (2022) wandten Quantilsregression auf Eishockey- und Fußballdaten an.

Die Quantilsregressions (QR)-DQN-Methode stellt die bedingte Verteilung von  $Z$  dar durch eine gleichmäßige Mischung von  $N$  unterstützenden Quantilen als  $\hat{Z}_k(s_t, a_t) = \frac{1}{N} \sum_{i=1}^N \delta_{\theta_{k,i}(s_t, a_t)}$ , wobei  $\theta_{k,i}$  das Quantil auf dem Quantilniveau (oder Quantilindex) schätzt  $\tau_i = i/N$  für  $1 \leq i \leq N$  und  $\delta_{\theta_{k,i}}$  eine Dirac-Verteilung für  $\theta_{k,i}$  ist. Nehmen wir zum Beispiel an, dass  $N = 4$  und schätzen  $\hat{Z}_k(s_t, a_t)$  Quantile bei 25%, 50%, 75%, 100% als 0,1, 0,4, 0,7, 0,9. Dann hat die kumulative Dichtefunktion (cdf) von  $\hat{Z}_k(s_t, a_t)$  25 % der Werte bei 0,1 oder weniger, 50 % (der Median) der Werte bei 0,4 oder weniger, 75 % der Werte bei 0,7 oder weniger und 100 % der Werte bei 0,9 oder weniger (also keiner über 0,9). Innerhalb jedes Quantils wird die cdf als gleichmäßig über das Quantil approximiert (z. B. hat die cdf den Wert 0,1 zwischen 0 und 0,25).<sup>4</sup>

Bei einer festen Anzahl  $N$  von Zielquantilen ( $N = 4$  in unserem Beispiel) können wir ein neuronales Netz so trainieren, dass es als Eingabe ein Zustands-Aktionspaar  $(s, a)$  hat und als Ausgabe 4 Zahlen, die den Quantilen entsprechen. Durch Erhöhung der Anzahl  $N$  liefert dieses Verfahren eine nichtparametrische Annäherung an die kumulative Dichtefunktion von  $\hat{Z}$  und damit der Verteilung von  $\hat{Z}$ . Für weitere Einzelheiten siehe (Liu et al., 2022).

#### Definition der Belohnungsvariable: Erwartete Tore oder Tatsächliche Tore

---

<sup>4</sup> Wir haben die Notation im Vergleich zu Bellemare et al. (2017) leicht vereinfacht, da ein Quantil mit dem Mittelpunkt eines Bins assoziiert wird und nicht mit dem oberen Endpunkt des Bins.

Wir haben gute empirische Ergebnisse erzielt, indem wir tatsächliche Tore als Belohnungsgrößen verwendeten. Als weitere Verfeinerung folgen wir (Decroos et al. 2017) und zerlegen die Torerfolgswahrscheinlichkeiten in zwei Komponenten: die Wahrscheinlichkeit, einen Schuss zu verwalten, und die Wahrscheinlichkeit, dass ein Tor zu einem Schuss führt:

$$P(goal_{t,k}|s_{t_0}) = P(goal_{t,k}|s_t, shot) \times P(s_t, shot|s_{t_0})$$

Diese Gleichung besagt, dass die Wahrscheinlichkeit, ein Tor aus einem Anfangszustand ( $s_{t_0}$ ) zu erzielen, das Produkt ist aus der Wahrscheinlichkeit, einen Schuss zu erreichen, und der Wahrscheinlichkeit, dass der Schuss zu einem Tor führt. Im Eishockey gilt diese Gleichung, da der einzige Weg, ein Tor zu erzielen, darin besteht, zuerst einen Schuss abzugeben.

In unserer Anwendung des verteilungsbasierten Reinforcement Learning (RL) betrachten wir die Torerfolgswahrscheinlichkeiten  $P(goal_{t,k}|s_t, shot)$  als die Ergebnisse (virtuelle Belohnungen), deren Verteilung zu modellieren ist.

Die Motivation für die Verwendung der Schuss-Tor-Zerlegung ist folgende:

1. Ein Team kann weitgehend kontrollieren, ob es einen Schuss abgibt, während der Erfolg des Schusses von Faktoren wie der Fähigkeit des gegnerischen Torwarts abhängt, die weniger unter ihrer Kontrolle liegen. Daher sollte ein Modell der Team-/Spielerstärke Teams für das Erreichen von Schüssen belohnen.
2. Schüsse sind zwar selten, aber nicht so selten wie Tore.
3. Bei Verwendung von tatsächlichen Toren als Belohnungen ist unsere Ergebnisvariable ( $Z_k$ ) binär und definiert eine Bernoulli-Verteilung. Bei erwarteten Toren als Belohnungen erstreckt sich die Ergebnisvariable ( $Z_k$ ) über das Intervall  $[0,1]$ , und die Ergebnisverteilung ist eine informative Verteilung über Torerfolgswahrscheinlichkeiten.

Bisher haben wir beschrieben, wie man ein maschinelles Lernmodell erstellen kann, das die Verteilung der Ballbesitzergebnisse für eine bestimmte Mannschaft zu einem bestimmten Zeitpunkt und in einem bestimmten Kontext in einem Spiel schätzt. Wir zeigen nun, wie das Modell angewendet werden kann, um analytische Erkenntnisse für eine Sportart zu gewinnen. Insbesondere erörtern wir, wie die Risikobereitschaft von Mannschaften, die Risikobereitschaft von Aktionen und die Einstufung von Spielern nach der Risikobereitschaft ihrer Aktionen bewertet werden können.

#### 11.4 Risikomaße für eine Ergebnisverteilung

Luo et al. (2023) befassen sich eingehend mit den Eigenschaften verschiedener Risikomaße für eine Verteilung von Ergebnissen. In dieser Studie verwenden wir drei dieser Maße: Standardabweichung, Gini-Abweichung und Value-at-Risk (VaR). Die Varianz, das Quadrat der Standardabweichung und der VaR sind in der Portfolioanalyse häufig verwendete Risikomaße. Die Gini-Abweichung wird von Luo et al. (2023) für

multimodale Verteilungen empfohlen. Die formalen Definitionen lauten wie folgt.

**Definition 1** (Risikomaße). Für eine Zufallsvariable  $Z$  seien  $Z_1$  und  $Z_2$  zwei i.i.d. Kopien von  $Z$ , d.h.  $Z_1$  und  $Z_2$  sind unabhängig und folgen der gleichen Verteilung wie  $Z$ .

- Die Varianz ist definiert als  $\mathbb{V}[Z] = \frac{1}{2} \mathbb{E}[(Z_1 - Z_2)^2]$
- Die Standardabweichung ist die Quadratwurzel der Varianz  $\text{STD}[Z] = \sqrt{\mathbb{V}[Z]}$
- Die Gini-Abweichung ist definiert als  $\mathbb{D}[Z] = \frac{1}{2} \mathbb{E}[|Z_1 - Z_2|]$ .

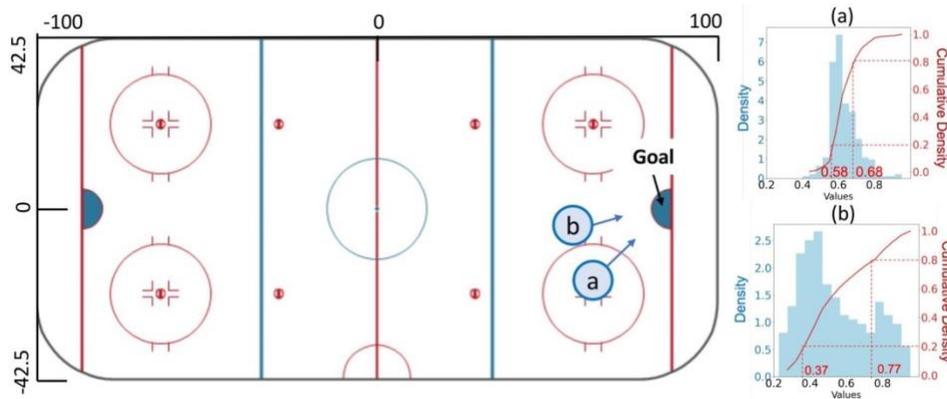


Abbildung 11.4: Die vorhergesagte Verteilung zukünftiger Tore in einem Eishockeyspiel zwischen Blues und Coyotes, NHL-Saison 2018-19. Die Schüsse werden an den Positionen (a) und (b) abgegeben. Die Verteilungen der nächsten Tore (a) und (b) haben *die gleiche Erwartung* (etwa 0,6), aber der erste Schuss hat eine viel geringere Varianz und Gini-Abweichung der Ergebnisse. Schuss (a) zeigt auch eine größere risikoaverse Schätzung (beim Konfidenzintervall 0,8 finden wir eine größere Chance auf das nächste Tor mit  $0,58 > 0,37$ ) und eine kleinere risikofreudige Schätzung (beim Konfidenzintervall 0,2 finden wir eine kleinere Chance auf das nächste Tor von  $0,68 < 0,77$ ).

Somit ersetzt die Gini-Abweichung die L2-Norm der Varianz durch die L1-Norm. Die Definition des **Value-at-Risk** (VaR) hängt von der Wahl eines **Konfidenzniveaus**  $\alpha \in (0,1]$ . Der VaR für das Niveau  $\alpha$  ist einfach das Quantil  $1 - \alpha$ . Im obigen hypothetischen Beispiel beträgt der Value-at-Risk für  $\alpha = 25\%$  also 0,7 und für  $\alpha = 50\%$  beträgt er 0,4. Intuitiv bietet der VaR eine Art Worst-Case-Analyse in Bezug auf ein benutzergesteuertes Risikoniveau. Wird beispielsweise  $\alpha = 0,8$  gewählt, entspricht dies einer *Risikoaversion*, weil das Gewicht auf schlechte Ergebnisse gelegt wird. Im Gegensatz dazu entspricht die Wahl von  $\alpha = 0,2$  einer *Risikobereitschaft*, weil sie einen größeren Bereich von positiven Ergebnissen berücksichtigt. Abbildung 11.4 illustriert die unterschiedlichen Maße von Risiko in verschiedenen Ergebnisverteilungen.

**Berechnung von Risikokennzahlen aus der Quantilsregression** Der VaR ist in Form von Quantilen definiert und kann daher auf natürliche Weise aus einem Quantilsregressionsmodell berechnet werden. Ausgehend von einer Quantilsdarstellung können wir die Varianz und die Gini-Abweichung wie folgt schätzen (Luo et al., 2023).

$$\mathbb{V}[\hat{Z}_k] \approx \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N (\theta_{k,i} - \theta_{k,j})^2 \quad (5)$$

$$\mathbb{D}[\hat{Z}_k] \approx \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N |\theta_{k,i} - \theta_{k,j}| \quad (6)$$

Abbildung 11.5 zeigt die Q-Werte und Standardabweichungen, die das trainierte Modell während eines Spiels zwischen den Flyers und den Maple Leafs zuordnet.

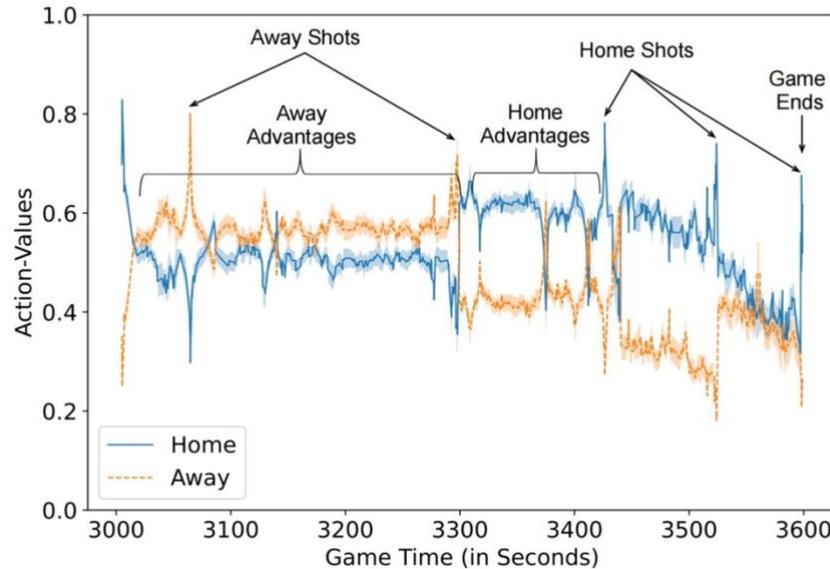


Abbildung 11.5: Veranschaulichung der dynamischen Verteilung der nächsten Torchancen durch Darstellung der entsprechenden Mittelwerte  $\pm$  Standardabweichung der Aktionswerte bei jedem Zeitschritt in einem Spiel zwischen den Flyers (Heimteam) und den Maple Leafs (Auswärtsteam) am 15. März 2019.

### 11.5 Messung des Risikos einer Aktion

Mit Hilfe der im vorigen Abschnitt beschriebenen Techniken können wir aus einer geschätzten Ergebnisverteilung Folgendes berechnen  $\hat{Z}_k(s_t, a_t)$  ein Risikomaß  $\rho_k(s_t, a_t)$ , wobei  $\rho$  eines der oben beschriebenen Risikomaße ist (Standard/Gini-Abweichung, VaR(c)). Ein einfacher Ansatz wäre, das Risiko einer Aktion in einem Spielzustand einfach durch das Risikomaß zu messen  $\rho_k(s_t, a_t)$ . Das Problem bei diesem Ansatz ist, dass er den allgemeinen Spielkontext der Aktion misst und nicht die spezifischen *Auswirkungen* der Aktion. Wenn ein Spieler beispielsweise einen Pass spielt, während seine Mannschaft ein leeres Tor hat, wird das Risikomaß hoch sein, unabhängig davon, wie riskant sein Pass ist. Intuitiv betrachtet findet der Pass an einem riskanten Ort und zu einem riskanten Zeitpunkt statt, trägt aber selbst nicht zum Risiko der Mannschaft bei.

Das gleiche Problem stellt sich in Bezug auf die erwarteten Aktionswerte (d. h. Q-Werte): Eine Mannschaft, die gegen ein leeres Netz spielt, hat eine hohe Chance, das nächste Tor zu erzielen, aber eine bestimmte Aktion eines Spielers erhöht die Torchancen seiner Mannschaft möglicherweise nicht über das Spielen gegen ein leeres Netz hinaus. Routley und Schulte (2015) schlugen vor, dieses Problem durch die Berechnung der **Auswirkung einer Aktion** zu lösen, die daran gemessen wird, wie sehr eine Aktion die Torchancen der Mannschaft im Ballbesitz *verändert*. Die **Auswirkung einer Aktion auf Torchancen** ist definiert für eine Aktion  $a_{t+1}$  und Zustand  $s_{t+1}$  wie folgt:

$$impact_k(s_{t+1}, a_{t+1}) = [Q_k(s_{t+1}, a_{t+1}) - Q_k(s_t, a_t)] \mathbb{1}_{p(s_{t+1}, a_{t+1}) \geq \epsilon} \quad (7)$$

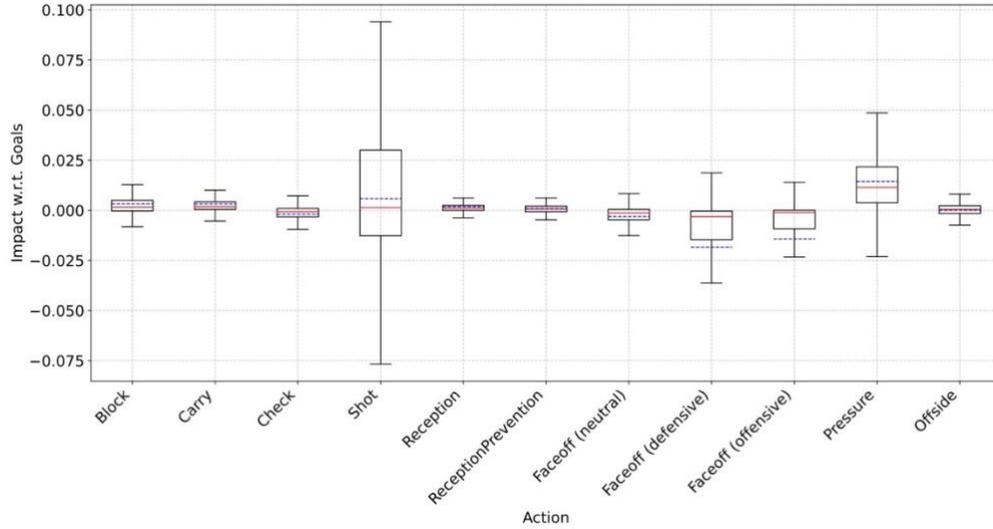


Abbildung 11.6: Box Plot für den Risikoeinfluss einer Aktion auf die Standardabweichung der Torchancen. Die rote Linie zeigt den Medianwert und die blaue Linie den Mittelwert an. Jeder Punkt stellt die Auswirkung der jeweiligen Aktion in einem Zustand dar (Ausreißer wurden entfernt).

Hier ist  $p$  ein Dichte-Schätzer für Zustands-Aktions-Paare, und  $\mathbb{1}_{p(\cdot) \geq \epsilon}$  ist ein Indikatorwert, der 1 ergibt, wenn die Wahrscheinlichkeit des Ereignisses unter einen Schwellenwert fällt  $\epsilon$ . Die Idee besteht darin, seltene Ereignisse herauszufiltern, da die Q-Wert-Schätzungen für seltene Ereignisse aufgrund des geringen Stichprobenumfangs häufig verzerrt sind. Bei einem diskreten Markov-Spielmodells haben Routley und Schulte (2015) alle Zustands-Aktions-Paare aus der Betrachtung entfernt, die weniger als 10 Mal in den Daten vorkommen. Für unseren kontinuierlichen Zustandsraum folgen wir Liu et al. (2022) und eliminieren anomale Ereignisse mit  $p < 20\%$ . Die Filterung von Anomalien ist kein wesentlicher Bestandteil unseres Risikoanalyse-Methodik.

Wir können den Ansatz für Torchancen auf die Risikoanalyse anwenden, indem wir messen, wie sehr eine Maßnahme das Risiko einer Maßnahme verändert, was wir als **Risikoauswirkung** bezeichnen:

$$Rimpact_k(s_{t+1}, a_{t+1}) = [\rho_k(s_{t+1}, a_{t+1}) - \rho_k(s_t, a_t)] \mathbb{1}_{p(s_{t+1}, a_{t+1}) \geq \epsilon} \quad (8)$$

wobei  $\rho$  eines unserer Risikomaße ist (Standard/Gini-Abweichung, Value at Risk).

Die Abbildungen 11.6 und 11.7 zeigen die Box-Plots für die Auswirkungen der Standard-/Gini-Abweichung für verschiedene Aktionen. Spielverändernde Ereignisse wie Schüsse, Assists und Strafen haben tendenziell einen großen Einfluss auf die Variabilität der Mannschaftsergebnisse. Defensive Faceoffs verringern tendenziell das Risiko einer Mannschaft, wahrscheinlich weil sie auf eine erfolgreiche Verteidigung folgen. Interessant ist, dass die Ausübung von Druck das Risiko einer Mannschaft tendenziell erhöht. Boxplots für den Value-at-Risk sind im Anhang zu finden.

## 11.6 Team-Performance und Team-Risiko-Tasking

Wir wenden unser Modell der Ergebnisverteilung auf Teams an, indem wir bewerten, wie viel Risiko ein Team insgesamt eingeht. Wir beginnen mit Teams, weil wir die

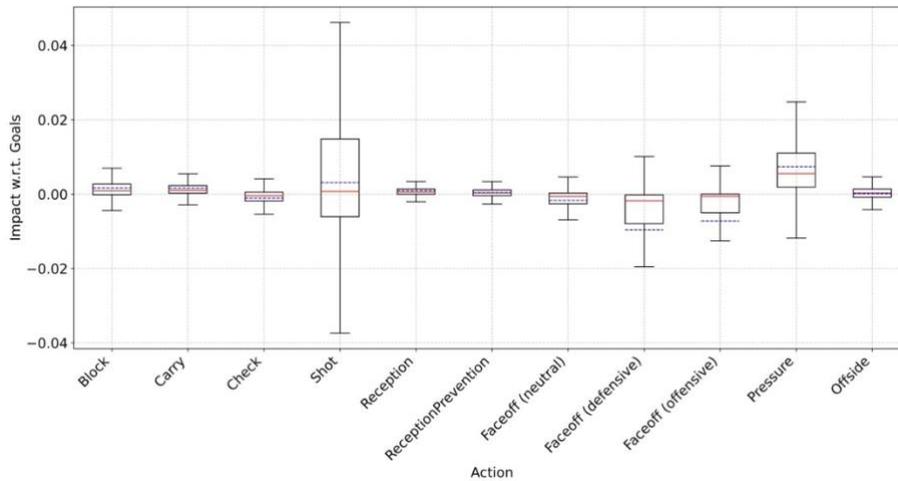


Abbildung 11.7: Box Plot für die Risikoauswirkung einer Aktion auf die Gini-Abweichung der Torchancen beim nächsten Tor. Die rote Linie zeigt den Medianwert und die blaue Linie den Mittelwert an. Jeder Punkt stellt die Auswirkung der jeweiligen Aktion in einem Zustand dar (Ausreißer wurden entfernt).

Gesamtpunktzahl einer Mannschaft während einer Saison als Maßstab für ihre Stärke verwenden können. *Die wichtigste Frage in diesem Abschnitt ist, ob die Risikobereitschaft eines Teams mit dem Teamerfolg korreliert.* Um die Risikobereitschaft von Teams zu quantifizieren, addieren wir die Risikoauswirkungen der gesamten Aktionen des Teams.

Für einen Datensatz  $\mathcal{D}$  sei  $g, t$  eine generische Instanz für die Ereignisnummer  $t$  im Spiel Nummer  $g$ . Außerdem sei  $\text{team}_{gt}$  die Mannschaft, die Aktion  $t$  im Spiel  $g$  ausführt, und ähnlich für den Zustand  $s_{gt}$  und die Aktion  $a_{gt}$ . Dann ist der gesamte Risikoeinfluss für die Mannschaft  $T$  im Datensatz gegeben durch

$$\text{RIM}_k = \sum_{g,t:\text{team}_{gt}=T} \text{R impact}_{k_t}(s_{gt}, a_{gt}) \quad (9)$$

wobei  $k_t$  die entsprechende Belastung zum Zeitpunkt  $t$  (Heim oder Auswärts). Gleichung (9) zeigt, wie eine Team-Risikoauswirkungsmetrik für jedes Risikomaß definiert werden kann, die wir wie folgt abkürzen: StdRIM = Risikoauswirkung für die Standardabweichung, GdRIM = Risikoauswirkung für die Gini-Abweichung, RIM(0,2) = Risikoauswirkung für den VaR mit risikobereitem Konfidenzniveau 0,2, RIM(0,8) = Risikoauswirkung für den VaR mit risikoscheuem Konfidenzniveau 0,8. RIM(c) wird von Liu et al. (2022) als RiGIM(c) bezeichnet.

Die Platzierung einer Mannschaft in der Saisonliga wird durch die Anzahl der Punkte bestimmt, die die Mannschaft in jedem Spiel erzielt. Daher messen wir die Korrelation zwischen der Risikoauswirkung einer Mannschaft in der Saison und der Gesamtpunktzahl der Saison als Maß für die Mannschaftsleistung. Wie die Abbildungen 11.8 und 11.9 zeigen, liefern sowohl die Standard- als auch die Gini-Abweichung Maßzahlen für die Risikobereitschaft, die sich hervorragend zur Vorhersage der Mannschaftsleistung eignen. Dies zeigt, dass *stärkere Teams mehr Risiken eingehen*. Tabelle 11.4 enthält die Pearson-Korrelationen zwischen der Gesamtpunktzahl des Teams und den Risikokennzahlen des Teams. Value-at-

Risikokennzahlen sind wesentlich weniger aussagekräftig in Bezug auf die Teamleistung, insbesondere bei einem hohen Konfidenzniveau von 0,8. Tabelle 11.5 zeigen die 10 besten Teams der Liga und ihre Risikokennzahlen.

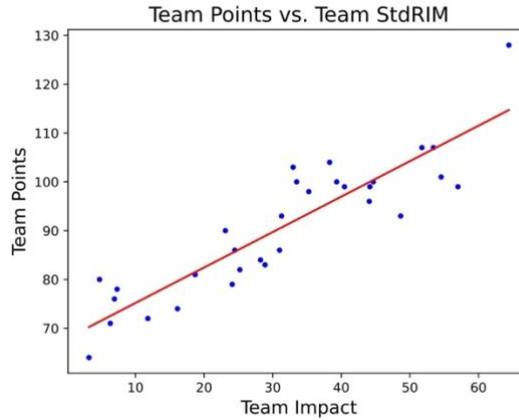


Abbildung 11.8: Mannschaftspunkte im Vergleich zu Team StdRIM.

Tabelle 11.4: Korrelationen zwischen der Risikoauswirkungskennzahl einer Mannschaft und ihren Saisonergebnissen. Die Standard/Gini-Abweichungen zeigen eine sehr hohe Vorhersagekraft für die Saisonleistung eines Teams, was zeigt, dass stärkere Teams mehr Risiken eingehen.

StdRIM	GdRIM	RIM(0.2)	RIM(0.8)
0.90	0.90	0.51	-0.24

### 11.7 Rangliste der Eishockeyspieler nach Risikobereitschaft

Wie bei Teams verwenden wir die Gesamtrisikoaussagen der Handlungen eines Spielers, um seine Risikobereitschaft zu bewerten. Sei  $pl_{gt}$  der Spieler, der zum Zeitpunkt  $t$  im Spiel  $g$  agiert. Dann ist die **gesamte Risikoauswirkung des Spielers** für den Spieler  $l$  im Datensatz ist gegeben durch

$$RIM_l = \sum_{g,t:pl_{gt}=l} Rimpact_{k_t}(s_{gt}, a_{gt}) \quad (10)$$

wobei wir die gleiche Notation wie in Abschnitt 6 verwenden. In den Tabellen 11.7 und 11.8 sind die 20 besten Spieler nach ihrem Risiko-Ranking aufgeführt. Bei Anwendung des "Sehtests" identifiziert der risikosuchende  $Var(0,2)$  viele Stars, wie Connor McDavid, Leon Draisaitl und Sidney Crosby. Die Standardabweichungskennzahl identifiziert ebenfalls einige Stars, wie Alex Ovechkin und Johnny Gaudreau. Aber sie hebt auch einige weniger bekannte Spieler hervor, wie Jason Zucker und Jaden Schwartz. Ein weiterer Unterschied besteht darin, dass die Standardabweichungsmetrik eine Tendenz zu Stürmern aufweist, während die  $Var(0,2)$ -Risikoauswirkungsmetrik mehr Center umfasst. Die Top-20-Tabellen für die anderen Risikokennzahlen befinden sich im Anhang.

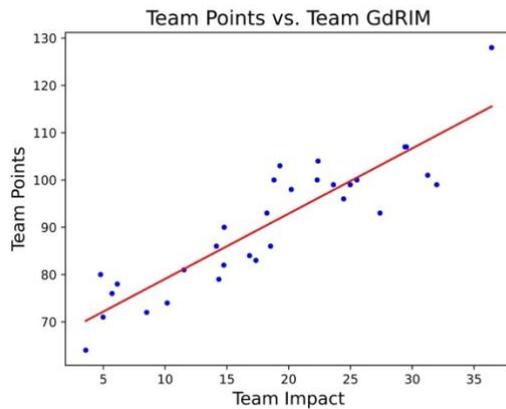


Abbildung 11.9: Team Points im Vergleich zum Team GdRIM.

Tabelle 11.5: Top 10 Teams mit StdRIM und GdRIM auf der Grundlage der gesamten Saison. Der tatsächliche Rang basiert auf der Gesamtpunktzahl. Der prognostizierte Rang basiert auf den Risikomaßen und misst, wieviel Risiko insgesamt eine Mannschaft in der Saison eingegangen ist (beide Maße ergeben dieselbe Rangfolge).

Name der Mannschaft	Rang nach Risikomodell	Tatsächlicher Rang	Punkte insgesamt	StdRIM	GdRIM
Lightning	1	1	128	64.43	36.42
Blues	2	12	99	57.01	31.98
Sharks	3	6	101	54.56	31.25
Flames	4	2	107	53.43	29.53
Bruins	5	3	107	51.76	29.41
Golden Knights	6	16	93	48.66	27.39
Maple Leafs	7	7	100	44.67	25.52
Hurricanes	8	11	99	44.17	24.98
Canadiens	9	14	96	44.09	24.45
Jets	10	10	99	40.47	23.62

Tabelle 11.6: Die 20 besten Spieler nach der Risikomessgröße Value-at-Risk mit einem risikobereiten Konfidenzniveau von 0,2 (d. h.  $RIM(0,2)$ ) auf der Grundlage der gesamten Saison.

Spieler Name	Position	Team	P	A	G	RIM(0.2)
Nikita Kucherov	RW	TBL	128	87	41	61.12
Mitchell Marner	RW	TOR	94	68	26	60.81
Johnny Gaudreau	LW	CGY	99	63	36	59.71
Patrick Kane	RW	CHI	110	66	44	56.55
Brad Marchand	LW	BOS	100	64	36	53.34
Mark Stone	RW	VGK	73	40	33	51.29
Connor McDavid	C	EDM	116	75	41	51.00
Leon Draisaitl	C	EDM	105	55	50	50.16
Timo Meier	RW	SJS	66	36	30	49.67
Blake Wheeler	RW	WPG	91	71	20	48.96
Sidney Crosby	C	PIT	100	65	35	48.56
Jonathan Huberdeau	LW	FLA	92	62	30	48.19
Kyle Connor	LW	WPG	66	32	34	47.85
Artemi Panarin	LW	CBJ	87	59	28	47.34
Evgenii Dadonov	RW	FLA	70	42	28	45.83
Cam Atkinson	RW	CBJ	69	28	41	45.70
Matthew Tkachuk	LW	CGY	77	43	34	45.32
Brendan Gallagher	RW	MTL	52	19	33	44.95
Jake Guentzel	LW	PIT	76	36	40	44.58
Brandon Saad	LW	CHI	47	24	23	43.69

Tabelle 11.7: Top 20 Spieler mit StdRIM auf Basis der gesamten Saison.

Spieler Name	Position	Team	P	A	G	StdRIM
Johnny Gaudreau	LW	CGY	99	63	36	13.07
Patrick Kane	RW	CHI	110	66	44	12.04
Nikita Kucherov	RW	TBL	128	87	41	11.68
Alex Ovechkin	LW	WSH	89	38	51	11.30
Mitchell Marner	RW	TOR	94	68	26	11.14
Cam Atkinson	RW	CBJ	69	28	41	11.05
Timo Meier	RW	SJS	66	36	30	10.77
Vladimir Tarasenko	RW	STL	68	35	33	9.31
Matthew Tkachuk	LW	CGY	77	43	34	9.30
Brad Marchand	LW	BOS	100	64	36	9.12
Jaden Schwartz	LW	STL	36	25	11	9.10
Brendan Gallagher	RW	MTL	52	19	33	9.04
David Pastrnak	RW	BOS	81	43	38	9.01
Kyle Connor	LW	WPG	66	32	34	8.95
Filip Forsberg	LW	NSH	50	22	28	8.72
Josh Anderson	RW	CBJ	47	20	27	8.30
Mikko Rantanen	RW	COL	87	56	31	8.28
Evgenii Dadonov	RW	FLA	70	42	28	7.91
Jason Zucker	LW	MIN	42	21	21	7.85
Jonathan Huberdeau	LW	FLA	92	62	30	7.73

Eine Schwierigkeit bei der Bewertung von Spielerrankings besteht darin, dass es, anders als bei Mannschaften, wenn überhaupt, dann nur wenige unabhängige Metriken für die Spielerleistung gibt (Franks et al., 2016)]. Wir folgen früheren Arbeiten (Liu & Schulte, 2018; Decroos et al., 2019) und betrachten die Korrelationen zwischen unseren Risikokennzahlen und anderen aussagekräftigen Spielerstatistiken wie Tore, Assists und Punkte (= Tore + Assists). Abbildung 11.10 zeigt die Korrelationen zwischen diesen Metriken und den Metriken für die Auswirkungen des Abweichungsrisikos Runde für Runde. Für jede Runde

in der Saison berechnen wir für jeden Spieler seinen bisherigen Gesamtrisikoeinfluss (z. B. Standardabweichungseinfluss über alle Spiele bis Runde 30) und setzen ihn in Beziehung zu seinen Statistiken (z. B. bis Runde 30 erzielte Gesamttore). Wir beobachten eine starke Korrelation für torbezogene Statistiken, die am Ende der Saison 0,51 erreicht für StdRIM and 0.56 for GdRIM. Man beachte, dass die Korrelation bereits nach etwa 30 Runden, also nach weniger als der Hälfte der Saison, relativ hoch ist. Dies bedeutet, dass *die Risikokennzahlen eine hohe Vorhersagekraft für die zukünftige Leistung der Spieler haben*. Die Autokorrelationskurve unten rechts bestätigt direkt die zeitliche Konsistenz der Risikokennzahlen. In dieser Darstellung wird der Wert der Risikokennzahl nach  $n$  Runden mit dem endgültigen Saisonwert korreliert. Wir sehen, dass bereits nach 25 Runden die bis dahin beobachtete Risikoauswirkung eines Spielers seine endgültige Risikoauswirkung mit einer Korrelation von **über 0,8** vorhersagt. Die starke zeitliche Autokorrelation ist ein Beweis dafür, dass Risikomessgrößen stabile Eigenschaften eines Spielers erfassen (Pettigrew, 2015).

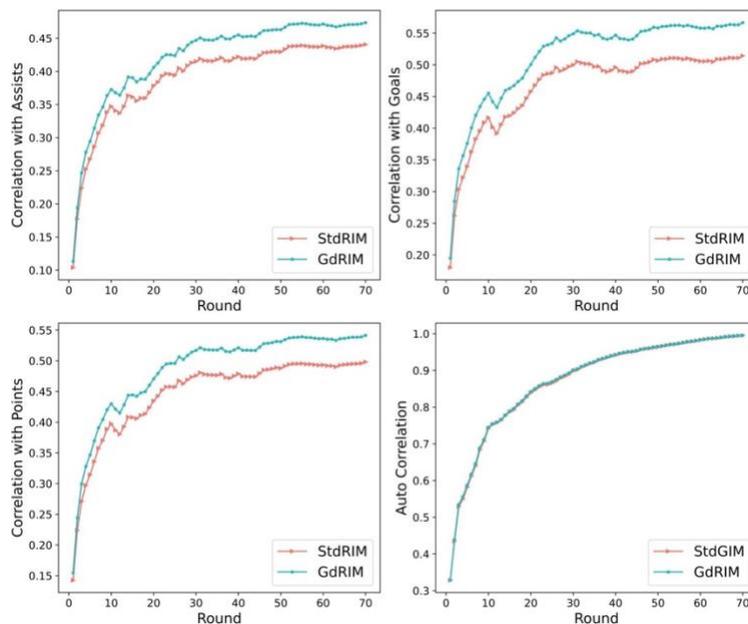


Abbildung 11.10: Korrelationen zwischen verschiedenen Spielerkennzahlen und Risikokennzahlen von Runde zu Runde. Die Abbildung zeigt die Korrelationen für die Risikokennzahlen Standardabweichung und Gini-Abweichung (StdRIM und GdRIM).

Abbildung 11.11 zeigt die Korrelationen zwischen Value-at-Risk-Kennzahlen und zielbezogenen Statistiken. Die Korrelationen sind sogar noch höher als bei den Abweichungsmetriken (0,76 vs. 0,5 bei Toren). Wie die Autokorrelation zeigt, ist die Risikobereitschaft eines Spielers, die durch den Value-at-Risk gemessen wird, während einer Saison stabil. Unsere Beobachtungen für Mannschafts- und Spielerrankings weisen daher in unterschiedliche Richtungen: Für Mannschaften gibt es deutliche

Hinweise darauf, dass die Standard/Gini-Abweichung die Risikobereitschaft misst, die auf die Stärke der Mannschaft hinweist, während für Spieler der Value-at-Risk-Wert mit einem niedrigen Konfidenzniveau besser zu korrelieren scheint. Korrelationen mit torbezogenen Statistiken sind nur ein oberflächliches Signal für die Stärke eines Spielers, da Tore selten vorkommen und nur einen kleinen Teil der relevanten Aktionen abdecken. Weitere Untersuchungen zur Risikobereitschaft der Spieler scheinen angezeigt.

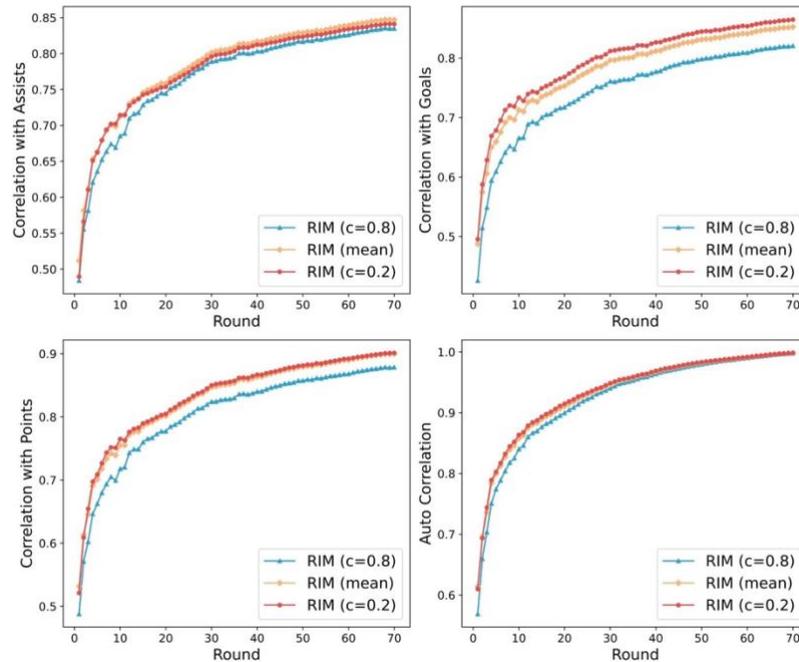


Abbildung 11.11: Korrelationen zwischen verschiedenen Spielerkennzahlen und Risikokennzahlen von Runde zu Runde. Die Abbildung zeigt die Korrelationen für die bedingte Value-at-Risk-Metrik (RIM(c)) und die Zieleinflussmetrik (basierend auf dem Erwartungswert). Wir stellen risikofreudige Konfidenzwerte (0,2), risikoscheue Konfidenzwerte (0,8) und einen neutralen Wert dar, bei dem das entsprechende Konfidenzniveau auf den Mittelwert der Ergebnisverteilung (seinen Q-Wert) festgelegt ist.

Eine mögliche Erklärung dafür, warum die abweichungsbasierten Metriken weniger mit Toren korrelieren und beim Augentest schlechter abschneiden, ist die Unterscheidung von zwei Möglichkeiten, wie die Handlungen eines Spielers die Ergebnisvariabilität erhöhen können: (1) *absichtliche Risiken* und (2) *ungezwungene Fehler*. Starke Spieler gehen kontrollierte Risiken ein, die auf dem Vertrauen in ihre Fähigkeiten beruhen. Beim Eishockey zum Beispiel wird ein starker Spieler den Puck oft in seine Angriffszone tragen und so die Verteidiger auf sich ziehen, was das Risiko erhöht, den Puck zu verlieren, aber auch die Chance auf einen erfolgreichen Angriff. Im Gegensatz dazu ist es sicherer, den Puck hinter dem Tor der verteidigenden Mannschaft zu versenken, führt aber tendenziell zu weniger Toren (Schulte et al., 2017a). Ein vergleichbares Szenario im Fußball wäre es, den Ball in Richtung des Tors der verteidigenden Mannschaft zu dribbeln, anstatt ihn an einen Mitspieler weiterzugeben. Ein Beispiel für einen ungezwungenen Fehler wäre ein Ballverlust während eines vielversprechenden Angriffs, der dazu

führt, dass sich die Bandbreite der wahrscheinlichen Ergebnisse von der Konzentration auf einen erfolgreichen Angriff bis hin zur Wahrscheinlichkeit, dass beide Mannschaften ein Tor erzielen, erweitert. Ein einfaches statistisches Modell für diese Unterscheidung ist ein Bernoulli-Modell für ein binäres Ereignis mit der Wahrscheinlichkeit  $p$ . Ein Modell für erwartete Tore könnte zum Beispiel die Wahrscheinlichkeit  $p$  für einen erfolgreichen Schuss zum Zeitpunkt  $t$  angeben. Die Varianz des Erfolgs ist gegeben durch  $p(1 - p)$  und ist maximal bei  $p = 50\%$ . Ein absichtliches Risiko kann die Erfolgchance von  $p < 50\%$  in Richtung  $p' > p$  erhöhen wobei  $p' < 50\%$ ; zum Beispiel kann ein absichtliches Risiko die Erfolgchance von 30% auf 40% erhöhen. *Ein solcher Schritt erhöht sowohl die Trefferchance als auch die Varianz.* Andererseits kann ein Fehler die Trefferchance von  $p$  in Richtung  $p' < p$  verringern mit  $p' < 50\%$ , zum Beispiel von 70% auf 60%. Eine solche Aktion *verringert die Gewinnchance und erhöht die Varianz.* Es ist möglich, dass die Risiken von starken Spielern hauptsächlich vom vorteilhaften absichtlichen Typ sind, während die Risiken von schwächeren Spielern vom schädlichen Fehlertyp sind, so dass beide ein hohes Risikomaß aufweisen können. Diese Analyse legt nahe, dass eine fruchtbare Richtung für zukünftige Forschung eine Spielermessung ist, die Risiko und Belohnung kombiniert, z.B. die Standardabweichung mit dem Mittelwert der Ergebnisverteilung, die mit der Aktion eines Spielers verbunden ist.

## 11.8 Fazit

Entscheidungssträger im Sport stehen oft vor einer *Abwägung zwischen Risiko und Erfolg*. Soll ein Basketballspieler einen 3-Punkte-Wurf aus der Ferne oder einen sichereren 2-Punkte-Wurf machen? Soll ein Eishockeyspieler den Puck mitnehmen, einem Mitspieler zuspelen oder ihn hinter dem Netz versenken? Um das Verhalten von Sportlern und Trainern bei Abwägungen zwischen Risiko und Nutzen zu untersuchen, sind Instrumente für die *Risikoanalyse* erforderlich. In diesem Beitrag wurden computergestützte Werkzeuge für die Risikoanalyse im Sport beschrieben, die auf Konzepten und Techniken des *distributiven Verstärkungslernens* basieren. Distributional RL zielt darauf ab, die Verteilung möglicher zukünftiger Ergebnisse zu modellieren, während traditionelles RL den Erwartungswert schätzt. Um die komplexen multimodalen Ergebnisverteilungen darzustellen, die sich aus der Sportdynamik ergeben, haben wir einen hochmodernen Ansatz für distributives RL adaptiert, der die Quantilsregression als aussagekräftigen nichtparametrischen Ansatz für die Modellierung von Verteilungen nutzt. Wir haben Verteilungs-RL-Techniken auf der Grundlage der Bellman-Gleichung angewandt, um eine dynamische Ergebnisverteilung in der National Hockey League für mehr als 1000 Spiele und mehr als 1 Million Ereignisse zu schätzen. In der Literatur zur Risikoanalyse werden verschiedene Methoden zur Quantifizierung des mit einer Ergebnisverteilung verbundenen Risikos vorgeschlagen. Wir haben einige der bekanntesten für Anwendungen in der Eishockey-Analyse bewertet, um die Fragen zu beantworten: Gehen stärkere Mannschaften mehr Risiken ein? Gehen stärkere Spieler mehr Risiken ein? Wir fanden heraus, dass die traditionelle Risikokennzahl Standardabweichung ein hervorragendes Maß ist, um den Erfolg einer Mannschaft vorherzusagen: Die Standardabweichung der Ergebnisverteilung eines Teams, aggregiert über eine Saison, zeigt eine Korrelation von 0,90 mit dem Tabellenplatz des Teams am Ende der Saison (bestimmt durch die Gesamtpunktzahl). Value-at-risk mit einem risikobereitem Konfidenzniveau von 0,2 zeigt eine geringere, aber immer noch starke Korrelation von 0,51.

Für das Spielerranking haben wir festgestellt, dass der Value-at-Risk mit einem Konfidenzniveau von 0,2, der über alle Aktionen eines Spielers in einer Saison aggregiert wird, eine sehr hohe Korrelation von 0,86 mit den gesamten Saisontoren des Spielers aufweist. Die Standardabweichung hingegen zeigt eine geringere, aber immer noch starke Korrelation von 0,51. Wir vermuten, dass die Modellierung des Risikoverhaltens von Spielern eine feinere Metrik erfordert, die unterscheiden kann zwischen absichtlichen Risiken, die

durch Handlungen entstehen, die ein hohes Maß an Geschicklichkeit erfordern, und Risiken, die sich aus Fehlern ergeben (z. B. Verlust des Pucks). Eine vielversprechende Quelle für solche feinkörnigen Messgrößen sind Maße aus der Portfoliotheorie, die Risiko und Erfolg (erwartetes Ergebnis) kombinieren. In der Sportanalytik könnten wir zum Beispiel eine Version der berühmten Sharpe-Ratio untersuchen, die den erwarteten Wert durch die Standardabweichung teilt.

Zusammenfassend lässt sich sagen, dass die Risikoanalyse ein vielversprechender neuer Ansatz für die Sportanalyse ist, der sich auf die schwierigen Abwägungen zwischen dem Eingehen von Risiken und der Maximierung der Erfolgchancen konzentriert, mit denen die Entscheidungsträger im Sport konfrontiert sind. Distributional Reinforcement Learning bietet die rechnerischen Werkzeuge, um sowohl Risiken als auch erwartete Belohnungen in großen Sportdatensätzen abzuschätzen. Unsere NHL-Studie bestätigt, dass starke Teams große Risiken eingehen und stützt die These, dass starke Spieler risikofreudig sind.

### 11.9. Danksagung

Diese Forschung wurde durch Forschungsmittel des kanadischen Natural Sciences and Engineering Research Council unterstützt, die an Schulte und Poupart vergeben wurden. Yudong Luo wurde außerdem durch ein David R. Cheriton Graduiertenstipendium und ein Graduiertenstipendium des Waterloo AI Institute unterstützt. Die in dieser Arbeit verwendeten Ressourcen wurden von der Provinz Ontario und der kanadischen Regierung über CIFAR sowie von Unternehmen, die das Vector Institute sponsern, zur Verfügung gestellt.<sup>5</sup> Der in dieser Arbeit analysierte Datensatz wurde großzügigerweise von Sportlogiq zur Verfügung gestellt.

### 11.10 Anhang

#### 11.10.1 Wertfunktionen und Wertverteilungen lernen

Abbildung 11.12 zeigt die Architektur unseres rekurrenten neuronalen Netzes zum Lernen von Q-Werten und Verteilungsquantilen. Bei einem Demonstrationsdatensatz von beobachteten Trajektorien (d.h.,  $\mathcal{D} = \{\tau\}$ ), kann eine Q-Funktion, die die Bellman-Gleichung erfüllt, durch Minimierung der 2-Norm des Fehlers der zeitlichen Differenz (TD) gelernt werden, die wie folgt definiert ist

$$\mathcal{L}(\theta) = \mathbb{E}_{\tau \sim \mathcal{D}} \left( \hat{Q}_k(s_{t+1}, a_{t+1}) + r_{k,t} - \hat{Q}_k(s_t, a_t) \right)^2 \quad (11)$$

wobei der Erwartungswert den Durchschnitt über alle Übergänge darstellt  $s_t, a_t, s_{t+1}, a_{t+1}$  die in dem Datensatz beobachtet wurden. Der Term  $\hat{Q}_k(s_t, a_t)$  steht für die aktuelle Wertschätzung und  $\hat{Q}_k(s_{t+1}, a_{t+1}) + r_{k,t}$  einen Vorausschau-Schritt. Die Minimierung ihrer quadratischen Differenz führt dazu, dass das neuronale Netz die Bellman-Gleichung Gleichung (3) erfüllt. Weitere Einzelheiten zur Umsetzung der zeitlichen Bellman-Gleichung mit der Quantilsregression finden Sie in Liu et al. (2022).

#### 11.10.2 Auswirkungen der Maßnahmen auf den Value-at-Risk

Abbildung 11.13 zeigt die Auswirkungen der Maßnahmen für Var(0,2) mit einem Konfidenzniveau von 0,2. Spielverändernde Ereignisse haben einen hohen Einfluss auf das Risiko, wie wir in Abschnitt 5 festgestellt

---

<sup>5</sup> <https://vectorinstitute.ai/partners/>

haben. Im Allgemeinen sind die Veränderungen des Risikos geringer als bei Standard/Gini-Abweichungen, da sich diese Metrik auf Ergebnisse mit geringer Wahrscheinlichkeit konzentriert.

Abbildung 11.14 zeigt die Auswirkungen von Maßnahmen auf das Risiko, gemessen durch  $\text{Var}(0,8)$  mit einem Konfidenzniveau von 0,8. Bei diesem Konfidenzniveau ist die Auswirkung des Risikos ähnlich wie die Auswirkung auf das Tor, d.h. sie misst eher die Zunahme der Torchancen als die Zunahme der Variabilität der Torchancen.

### 11.10.3 Die 20 besten Spieler-Tabellen für andere Risikokennzahlen

Tabelle 11.8 zeigt die 20 wichtigsten Akteure für die Risikokennzahl Gini-Auswirkung, die den 20 wichtigsten Akteuren für die Risikokennzahl Standardabweichung sehr ähnlich ist.

Für das risikoaverse Konfidenzniveau 0,8 identifiziert der bedingte Value-at-Risk viele Stars, wie Scheifele und Crosby. Diese Kennzahl zeigt eine Vorliebe für Center und schließt interessanterweise auch einige Verteidiger ein.

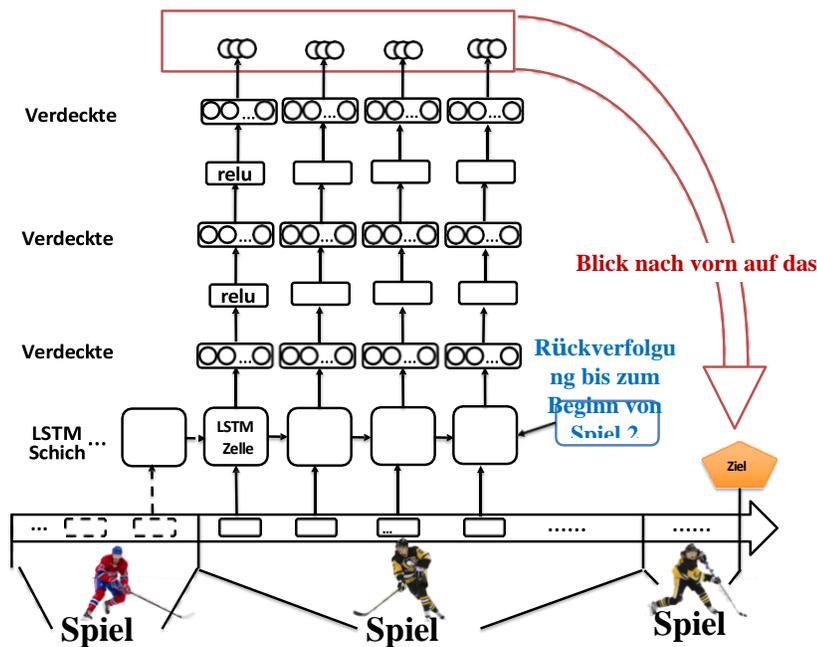


Abbildung 11.12: Unsere rekurrente neuronale Netzarchitektur zum Lernen von Q-Werten und einer Verteilung von Handlungsergebnissen. In jedem Zeitschritt erhält das RNN als Eingabe ein Paar  $s_t, a_t$  wobei der Zustand  $s_t$  ein Vektor der in Tabelle 1 dargestellten Merkmale ist. Es gibt eine Schätzung des Q-Wertes (Chance, das nächste Tor zu erzielen) oder eine Reihe von Quantilen aus, die die Verteilung der Handlungsergebnisse (erzielte Tore) darstellen. Wir verwenden eine LSTM-Architektur. Für weitere Einzelheiten siehe Liu et al. (2022).

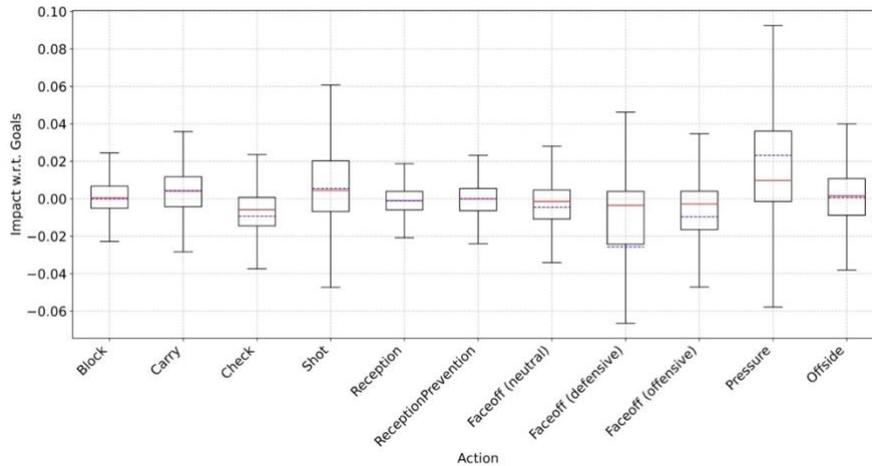


Abbildung 11.13: Box Plot mit Konfidenzniveau 0,2 (risikofreudig) Die rote Linie zeigt den Medianwert und die blaue Linie den Mittelwert an. Jeder Punkt stellt die Risikoauswirkung der jeweiligen Aktion in einem Zustand dar.

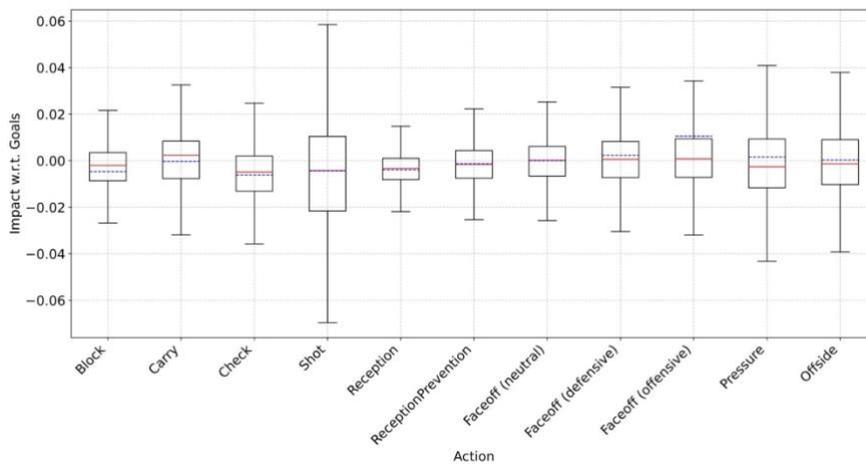


Abbildung 11.14: Box Plot für VaR mit Konfidenzniveau 0,8 (risikoavers) Die rote Linie zeigt den Medianwert und die blaue Linie den Mittelwert an. Jeder Punkt steht für die Risikoauswirkungen der jeweiligen Aktion in einem Zustand.

Tabelle 11.8: Die 20 besten Spieler mit GdRIM auf der Grundlage der gesamten Saison.

Spieler Name	Position	Team	P	A	G	GdRIM
Johnny Gaudreau	LW	CGY	99	63	36	7.15
Patrick Kane	RW	CHI	110	66	44	6.61
Nikita Kucherov	RW	TBL	128	87	41	6.53
Cam Atkinson	RW	CBJ	69	28	41	6.06
Alex Ovechkin	LW	WSH	89	38	51	6.00
Mitchell Marner	RW	TOR	94	68	26	5.98
Timo Meier	RW	SJS	66	36	30	5.97

Matthew Tkachuk	LW	CGY	77	43	34	5.18
Brad Marchand	LW	BOS	100	64	36	5.06
Kyle Connor	LW	WPG	66	32	34	5.03
Vladimir Tarasenko	RW	STL	68	35	33	4.96
Brendan Gallagher	RW	MTL	52	19	33	4.91
David Pastrnak	RW	BOS	81	43	38	4.87
Jaden Schwartz	LW	STL	36	25	11	4.87
Filip Forsberg	LW	NSH	50	22	28	4.64
Mikko Rantanen	RW	COL	87	56	31	4.62
Josh Anderson	RW	CBJ	47	20	27	4.52
Evgenii Dadonov	RW	FLA	70	42	28	4.44
Jason Zucker	LW	MIN	42	21	21	4.27
Jonathan Huberdeau	LW	FLA	92	62	30	4.22

Tabelle 11.9: Top 20 Spieler mit VaR bei Konfidenzniveau 0,8 auf Basis der gesamten Saison.

Spieler Name	Position	Team	P	A	G	RIM(0.8)
Aleksander Barkov	C	FLA	96	61	35	50.50
Leon Draisaitl	C	EDM	105	55	50	49.67
Mark Scheifele	C	WPG	84	46	38	48.73
Sidney Crosby	C	PIT	100	65	35	47.24
Jonathan Toews	C	CHI	81	46	35	45.22
Mitchell Marner	RW	TOR	94	68	26	42.71
Dylan Larkin	C	DET	73	41	32	41.82
Nikita Kucherov	RW	TBL	128	87	41	41.00
Max Domi	LW	MTL	72	44	28	40.47
Connor McDavid	C	EDM	116	75	41	40.45
Bo Horvat	C	VAN	61	34	27	39.92
Mika Zibanejad	C	NYR	74	44	30	39.29
Artemi Panarin	LW	CBJ	87	59	28	38.93
Sebastian Aho	C	CAR	83	53	30	38.55
Mark Stone	RW	VGK	73	40	33	38.38
Claude Giroux	C	PHI	85	63	22	37.98
Johnny Gaudreau	LW	CGY	99	63	36	37.86
Mathew Barzal	C	NYI	62	44	18	37.83
Nicklas Backstrom	C	WSH	74	52	22	37.78
Brad Marchand	LW	BOS	100	64	36	37.55

## Referenzen

- Beaudoin, D., & Swartz, T. B. (2010). Strategien zum Ziehen des Torwarts beim Eishockey. *The American Statistician*, 64(3), 197-204.
- Bellemare, M. G., Dabney, W., & Munos, R. (2017, Juli). A distributional perspective on reinforcement learning. In *International conference on machine learning* (pp. 449-458). PMLR.
- Chan, T. C., Fernandes, C., & Puterman, M. L. (2021). Punktgewinne im Fußball: Using Markov process-based value functions to assess team performance. *Operations Research*, 69(3), 877-894.
- Decroos, T., Bransen, L., Van Haaren, J., & Davis, J. (2019, Juli). Actions speak louder than goals: Valuing player actions in soccer. In *Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 1851-1861).
- Franks, A. M., D'Amour, A., Cervone, D., & Bornn, L. (2016). Meta-Analytik: Werkzeuge zum Verständnis der statistischen Eigenschaften von Sportmetriken. *Journal of Quantitative Analysis in Sports*,

12(4), 151-165.

- Littman, M. L. (1994). Markov games as a framework for multi-agent reinforcement learning. In *Machine learning proceedings 1994* (pp. 157-163). Morgan Kaufmann.
- Liu, G., & Schulte, O. (2018). Deep Reinforcement Learning in Ice Hockey for Context-aware Player Evaluation. arXiv preprint arXiv:1805.11088.
- Liu, G., Luo, Y., Schulte, O., & Kharrat, T. (2020). Deep soccer analytics: learning an action-value function for evaluating soccer players. *Data Mining and Knowledge Discovery*, 34, 1531-1559.
- Liu, G., Luo, Y., Schulte, O., & Poupart, P. (2022). Uncertainty-aware reinforcement learning for risk-sensitive player evaluation in sports game. *Advances in Neural Information Processing Systems*, 35, 20218-20231.
- Luo, Y., Schulte, O., & Poupart, P. (2020, Juli). Inverses Verstärkungslernen für Mannschaftssportarten: Valuing actions and players. In *IJCAI* (S. 3356-3363).
- Luo, Y., Liu, G., Poupart, P., & Pan, Y. (2023). Eine Alternative zur Varianz: Gini deviation for risk-averse policy gradient. *Advances in Neural Information Processing Systems*, 36, 60922-60946.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *nature*, 518(7540), 529-533.
- Nationale Eishockeyliga. Offizielle Regeln der Nationalen Eishockeyliga 2014-2015. <http://www.nhl.com/nhl/en/v3/ext/rules/2014-2015-rulebook.pdf>.
- Pelechrinis, K. (2016). Entscheidungsfindung im American Football: Evidence from 7 Years of NFL Data. In *MLSA@ PKDD/ECML*.
- Pettigrew, S. (2015). Assessing the offensive productivity of NHL players using in-game win probabilities. In *9th annual MIT sloan sports analytics conference* (Vol. 2, No. 3, S. 8).
- Routley, K. D. (2015). Ein Markov-Spielmodell zur Bewertung von Spieleraktionen im Eishockey.
- Schulte, O. (2022, September). Valuing Actions and Ranking Hockey Players With Machine Learning. In *Linköping Hockey Analytics Conference* (S. 2-9).
- Schulte, O., Khademi, M., Gholami, S., Zhao, Z., Javan, M., & Desaulniers, P. (2017). Ein Markov-Game-Modell zur Bewertung von Aktionen, Standorten und Teamleistung im Eishockey. *Data Mining and Knowledge Discovery*, 31, 1735-1757.
- Schulte, O., Zhao, Z., Javan, M., & Desaulniers, P. (2017, März). Apples-to-apples: Clustering and ranking NHL players using location information and scoring impact. In *Proceedings of the MIT Sloan Sports Analytics Conference*.
- Sutton, R. S., & Barto, A. G. (1998). Das Problem des Verstärkungslernens. *Reinforcement Learning: An introduction*, 51-85.

## Informationen über die Autoren

### **Oliver Schulte (korrespondierender Autor)**

School of Computer Science  
Simon Fraser University, Vancouver  
[oschulte@cs.sfu.ca](mailto:oschulte@cs.sfu.ca)

### **Sheng Xu**

School of Data Science,  
The Chinese University of Hong Kong, Shenzhen

Shenzhen, China  
[shengxu1@link.cuhk.edu.cn](mailto:shengxu1@link.cuhk.edu.cn)

**Yudong Luo**  
Cheriton School of Computer Science  
University of Waterloo  
Waterloo, Canada  
Vector Institute, Toronto, Canada  
[yudong.luo@uwaterloo.ca](mailto:yudong.luo@uwaterloo.ca)

**Pascal Poupart**  
Cheriton School of Computer Science  
University of Waterloo  
Waterloo, Canada  
Vector Institute, Toronto, Canada  
[ppoupart@uwaterloo.ca](mailto:ppoupart@uwaterloo.ca)

**Guiliang Liu**  
School of Data Science,  
The Chinese University of Hong Kong, Shenzhen  
Shenzhen, China  
[liuguiliang@cuhk.edu.cn](mailto:liuguiliang@cuhk.edu.cn)