

Die Vorteile von OBV gegenüber Packing, eine Analyse

Tore sind das oberste Spielziel im Fußball. Da ist es nur logisch, dass sich die ersten statistischen Modelle darauf konzentrierten, die Wahrscheinlichkeit eines Torerfolgs in bestimmten Situationen zu beurteilen. Dabei entstanden Modelle zur Einschätzung der Chancenqualität in Form zu erwartender Treffer (xG). Was aber, wenn wir noch einen Schritt weiter gehen wollen – indem wir den ganzen vorherigen Spielzug mit berücksichtigen? Wie lässt sich die Beteiligung der anderen Spieler am Tor analysieren? Welche Lösungsansätze gibt es dafür?

Aktuelle Ansätze

Bevor wir die StatsBomb-Lösung des On-Ball Value (OBV) vorstellen, möchten wir uns die verschiedenen aktuellen Methoden ansehen.

xG Chain: Diese Methode berücksichtigt bei der Berechnung der Wahrscheinlichkeit eines Torerfolgs die gesamten vorhergehenden Passstafetten. Der Ansatz honoriert die Spieler, die häufig an Spielzügen beteiligt sind, die zu hochwertigen Chancen führen. Das System hat gewisse Vorzüge. Allerdings muss am Ende der Passfolge immer ein Torschuss folgen. Andernfalls fallen alle vorherigen Pässe durchs Bewertungsraster. Ein weiteres Problem: Bei dieser Herangehensweise kann es passieren, dass der Beitrag eines Spielers überschätzt wird. So erhält etwa ein Teamkollege, der zuverlässig einen wichtigen Spielmacher mit einfachen Pässen bedient, in der Regel hohe Werte – selbst, wenn der Spielmacher den eigentlichen Beitrag zum Torerfolg leistet.

Progressive Pässe: Es gibt eine Möglichkeit, die Torvorbereitung weniger auf Schüsse zu reduzieren. Gewertet wird dann vielmehr die Anzahl an Pässen, die den Ball näher ans gegnerische Tor bringen. Hierbei wird ein bestimmter Schwellenwert für den Ballvortrag definiert. Das können die zurückgelegten Meter sein, oder die verkürzte Distanz zum gegnerischen Tor prozentual zur Ausgangsentfernung. Mit dieser Methode ist es möglich, erfolgreich diejenigen Spieler zu identifizieren, die am meisten zum Ballvortrag beitragen. Es ist aber immer noch ein relativ grobmaschiger Ansatz. So wird etwa ein Ballvortrag über 50 Meter ähnlich eingestuft wie ein Pass über 15 Meter. Außerdem ist in diesem System ein 15-Meter-Pass am Flügel genauso viel wert wie ein 15-Meter-Pass in den Strafraum.

Packing: Keine der oben genannten Methoden berücksichtigt den Standort der Spieler, die gerade nicht am Ball sind – anders als der Packing-Wert, der die Zahl der mit einem Pass überspielten Gegenspieler bewertet. Die Denkweise dahinter: Alle überspielten Gegner stehen nicht mehr hinterm Ball, sind also aus dem Spiel genommen. Die Bewertung der Pässe erfolgt anteilig zur Anzahl der überspielten Gegner.

Es gibt viele Abwandlungen dieses Messwertes. Keine von ihnen bezieht aber explizit den Passort oder die Spielsituation mit ein. Ob der Pass aus dem Spiel heraus, durch einen Freistoß oder in einer anderen Situation zustandekam, wird nicht berücksichtigt. Mit anderen Worten: Sobald dieselbe Anzahl an Gegenspielern überwunden wurden, sind die Pässe gleichwertig – ganz gleich, ob der Pass von der Mittellinie auf den Elfmeterpunkt kam, oder es sich um einen Kurzpass über fünf Meter am Mittelkreis handelte, mit dem nur die vorderste Pressing-Linie überspielt wurde. Auch der allgemeine Spielkontext fließt nur in geringem Maße mit in die Bewertung ein: Beim Konter gibt es deutlich weniger Gegenspieler, die mit einem Pass überspielt werden können, als bei einem kontrollierten Spielaufbau.

Berechnungsmodelle nach Ballbesitz: Eine neuere Methode zur Bewertung von Spielaktionen besteht darin, explizit die Änderung der Torwahrscheinlichkeit infolge eines jeden Ereignisses einzuschätzen. Hierbei fließen die Anfangs- und Endposition sowie der Spielkontext und manchmal auch der Verlauf des Ballbesitzes mit in die Berechnung ein. Diese Herangehensweise hat mehrere Vorteile, auf die wir unten noch genauer eingehen. Der Schwachpunkt der Methode liegt darin, dass sie Daten zur Position der Spieler abseits des Balles oft nur sehr begrenzt berücksichtigt.

Berechnungsmodelle nach Feldüberlegenheit: Manche diese Berechnungsmodelle berücksichtigen durchaus die Position der Spieler abseits des Balles auf Basis der Tracking-Daten. Es handelt sich dabei um wahrscheinlichkeitstheoretische Berechnungen der Feldüberlegenheit und der Chance auf Torerfolg bei einem erfolgreichen Pass in den betreffenden Bereich. Für ein solches Modell sind die kompletten Tracking-Daten erforderlich. Bislang gibt es nur eine sehr kleine Zahl an Teams, die ein System dieser Art entwickelt und genutzt haben. Die bekanntesten sind [Javier Fernandez](#) und Kollegen sowie [Will Spearman](#). Es ist sehr schwer, die dafür erforderlichen großen Mengen an Tracking-Daten zu beschaffen. Ein weiterer Knackpunkt ist das tiefgehende technische Fachwissen, das für die Berechnung erforderlich ist. Die Anwendung beschränkt sich damit auf die kleine Auswahl an Ligen, in denen Tracking-Daten verfügbar sind.

Der StatsBomb-Ansatz: On-Ball Value (OBV)

Aktuell bewerten wir den Ballvortrag mithilfe unseres eigenen Ballbesitz-Modells, dem On-Ball Value (OBV). Der OBV ist so angelegt, dass er den Ballvortrag danach bewertet, wie stark ein Ereignis die Tordifferenz beim aktuellen und nächsten Ballbesitz voraussichtlich verbessert oder verschlechtert. Die Ballbesitzphasen können wir dabei nach Offensiv- und Defensivbewegung unterteilen.

Wir sind der Meinung: Modelle zur Bewertung des Ballbesitzes wie der OBV sind anderen Methoden überlegen, die sich auf den Wert/Progressivität konzentrieren, den ein Spieler im Ballbesitz mit seinen Aktionen erreicht.

Intuitive und aussagekräftige Skala

Ein progressiver Pass erhöht normalerweise die Trefferchance – bei gleichzeitig sinkendem Risiko, im Falle des Ballverlusts einen Gegentreffer zu kassieren. Einer der Vorteile beim Einsatz des OBV zur Bewertung progressiver Pässe: Der Parameter ist voll auf den Nettowert der zu erwartenden Tordifferenz infolge einer Aktion geeicht und konzentriert sich auch explizit darauf. Damit gibt der OBV einen gut kalibrierten Faktor der wichtigsten Maßeinheit im Fußball wieder: die zu erwartenden Tore.

Das Ergebnis ist damit aussagekräftiger als die Resultate anderer Systeme, die etwa die Zahl der progressiven Bälle oder den Pacing-Wert überspielter Gegner über 90 Minuten ausgeben. Diese Ansätze versuchen, den progressiven Spielaufbau zu bewerten, indem sie die Torrate mit Ersatzparametern in Korrelation setzen – beispielsweise mit dem Abstand zum Tor oder der Zahl der Verteidiger zwischen Ball und Tor. Die Ergebnisse können eine grobe Richtung vorgeben. Was sie aber genau für den weiteren Spielverlauf bedeuten, bleibt ungewiss.

Ganzheitliche Bewertung von Aktionen (mit vergleichbaren Werten)

Wir können mit dem OBV den Progressionswert von Pässen berechnen, was in der Praxis auch geschieht. Das Modell hat aber eine noch deutlich größere Tragweite. Es verrät uns, mit welcher Wahrscheinlichkeit ein Team infolge einer bestimmten Aktion einen Treffer landet oder einen Gegentreffer hinnehmen muss. Das Prinzip können wir auf alle Pässe anwenden, um quantitativ deren Progressionswert zu bestimmen. Daneben erlaubt es uns, Werte über Ereignisse hinweg miteinander zu vergleichen – etwa den Nutzen von Ballvorträgen, abgefangenen Bällen etc. Schließlich beziehen sich alle Ereignisse auf eine gemeinsame Skala.

Der OBV beziffert den Wert einer jeden Aktion dahingehend, wie sie sich voraussichtlich auf die Tordifferenz eines Teams auswirkt. Dadurch können wir den Wert von Sportlern mit unterschiedlichen Spielstilen fürs Team einschätzen. Alle Analysen, die wir zusätzlich zum OBV durchführen, haben dasselbe Bezugssystem.

Nutzen-Risiko-Rechnung

Der OBV beurteilt den Wert eines Passes direkt dahingehend, wie er sich auf die Treffer- und Gegentreffer-Wahrscheinlichkeit des Teams auswirkt. Ist der Pass nicht erfolgreich, erhält der betreffende Spieler einen Abzug. Dafür werden zwei Parameter herangezogen: die Erfolgswahrscheinlichkeit, die das Team vor dem Ballverlust hatte, und die Wahrscheinlichkeit, dass der Gegner infolge des wiedererlangten Ballbesitzes einen Treffer erzielt.

Wir halten das für die komplettere Methode zur Beurteilung von Ballbewegungen. Die herkömmlichen Systeme bewerten den wesentlichen Faktor weniger direkt: den Einfluss des einzelnen Spielers auf die Fähigkeit seines Teams, Tore zu erzielen und Gegentreffer zu verhindern.

Ein wichtiger Vorteil unserer Methode besteht darin, erfolglose Aktionen in die Spielerbewertung mit einfließen zu lassen. Die erfolgreichen progressiven Pässe und die Zahl der mit einem erfolgreichen Pass überspielten Gegner (das Packing) greifen zu kurz. Beide Ansätze berücksichtigen nicht die fehlgeschlagenen Versuche eines Spielers, einen progressiven Pass zu spielen, die im Ballverlust münden. Es geht darum, herauszufinden, welchen Anteil ein einzelner Spieler an der Tordifferenz seines Teams hat. Dabei dürfen wir uns nicht allein auf die Anzahl positiver Aktionen beschränken. Anders gesagt: Wenn ein Spieler ein Risiko eingeht, wollen wir wissen, ob der Erfolg die Aktion rechtfertigt.

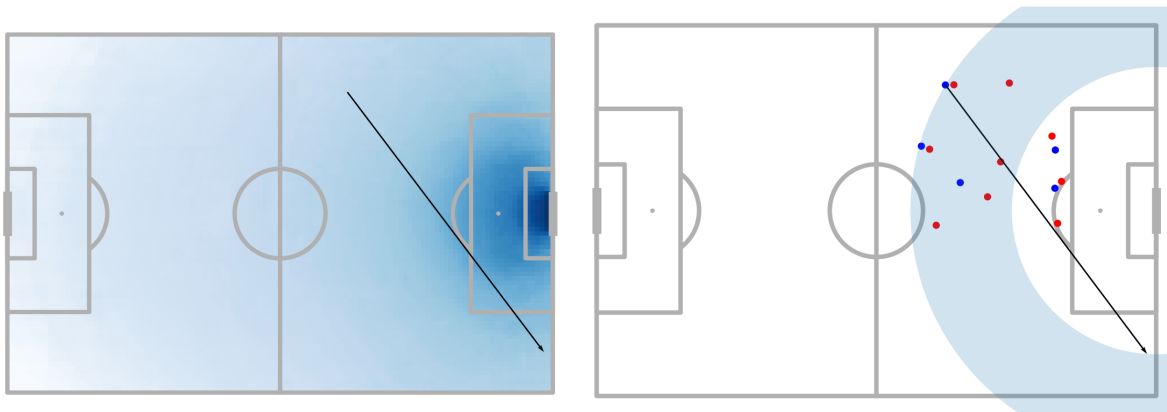
Konsistenz

Die Messwerte, die wir zur Bewertung des Ballvortrags verwenden, müssen direkt mit der Effektivität eines Teams verknüpft sein. Andernfalls können wir Spieler nicht verlässlich bewerten. Im Idealfall ergeben Aktionen mit identischer Beurteilung auch dieselben Werte, weil sie auf die gleiche Art und Weise zum Erfolg des Teams beitragen.

Nicht alle progressiven Pässe haben den gleichen Effekt auf die zu erwartende Tordifferenz eines Teams. Ein erfolgreicher Vertikalpass eines Innenverteidigers über 15 Meter wirkt sich anders auf die zu erwartende Tordifferenz aus als ein Pass über dieselbe Distanz in den Strafraum hinein. Ein anderes Beispiel: Ein Pass, mit dem vier Gegner überspielt werden, von denen drei auf der eigenen Spielfeldseite stehen, hat

einen anderen Wert für das Team, als ein Pass, der eine komplette kompakte Formation aus vier gegnerischen Mittelfeldspielern aus dem Spiel nimmt.

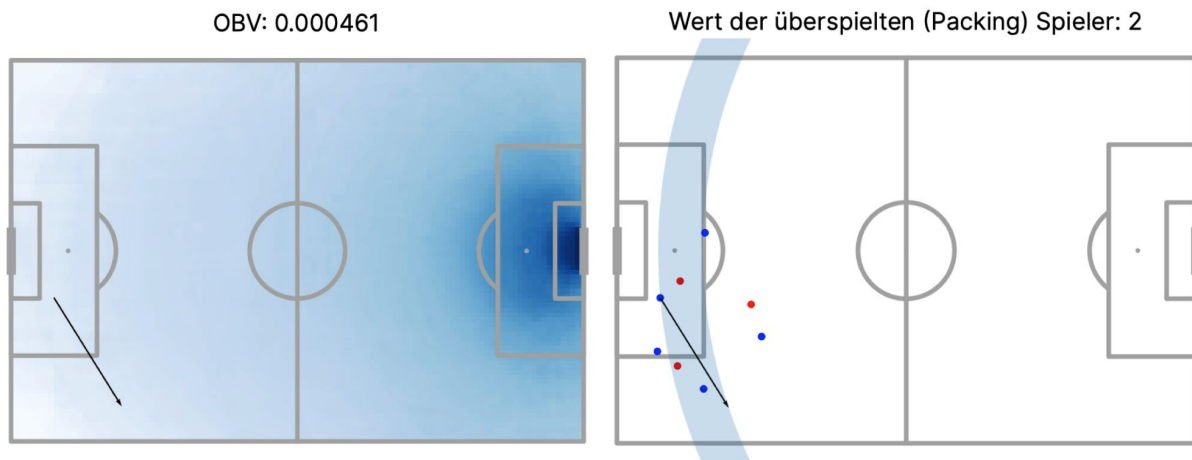
Die Darstellungen unten demonstrieren die Passbewertung anhand der beim OBV (links) und beim Packing (rechts) genutzten Daten. Die Heatmap in der Darstellung links zeigt die Ballbesitzfläche nach Einschätzung des OBV (auf einer Protokoll-Skala). Der markierte Bereich im rechten Bild ist die Zone, in der Gegner nach dem Packing-Prinzip überspielt wurden.



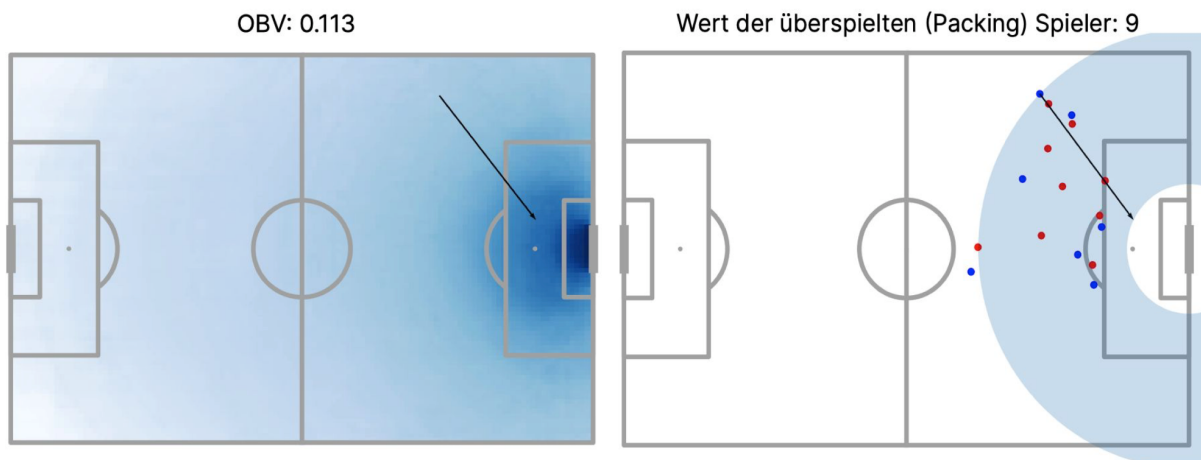
Rückpässe in den Strafraum sind ein prominentes Beispiel dafür, wie wichtig es ist, Aktionen dahingehend zu bewerten, wie sie die Trefferwahrscheinlichkeit beeinflussen. Im Beispiel unten verändert sich beim Pass die Entfernung zum Tor nicht groß. Auch werden nicht viele Gegner durch den Pass überspielt. Dennoch ergibt sich durch die Aktion eine deutlich erhöhte Trefferwahrscheinlichkeit.



Bei einer identischen Anzahl überspielter Gegner kann eine Aktion die Trefferwahrscheinlichkeit ganz unterschiedlich beeinflussen. Dazu ein Beispiel:



Wichtig ist, dass sich auch der umgekehrte Effekt ergeben kann: Zwei Pässe mit einem sehr ähnlichen OBV können sich in der Realität sehr unterschiedlich auf die Trefferwahrscheinlichkeit auswirken. In diesem Beispiel handelt es sich um einen schwierigen Pass durch die Abwehrreihen in den torgefährlichen Bereich. Die tatsächliche Trefferwahrscheinlichkeit vor der Aktion ist relativ gering.



Da der OBV die Position der Verteidigung nicht berücksichtigt, bewertet das System diesen viel einfacheren Pass sehr ähnlich:



Der OBV bringt zwar deutliche Vorteile mit sich. Es gibt aber keine objektiven und öffentlich verfügbaren Untersuchungen, die zur Passbewertung den relativen Effekt der Aktion mit Tracking-basierten Daten abgleichen. Wir haben daher eine quantitative Analyse durchgeführt, die einen solchen Abgleich ermöglicht und unser Verständnis unterschiedlicher Methoden erweitert.

Methodik

Das Ziel besteht letztendlich darin, den Wert einer bestimmten Aktion fürs Team einzuschätzen. Deshalb haben wir einige Systeme darauf trainiert, die Vorhersagegenauigkeit unterschiedlicher Informationen zu bewerten, auf denen der OBV und das Packing aufbauen. Dadurch wollten wir ermitteln, wie gut die Daten vorhersagen, ob ein Team im weiteren Verlauf des Ballbesitzes ein Tor erzielt.

Wir haben dafür verallgemeinerte additive Modelle (generalised additive models, kurz GAMs) verwendet. Diese erlauben es, nichtlineare Beziehungen zwischen den Merkmalen und dem Ziel herzustellen. Sie sind damit flexibler als eine logistische Regression, gleichzeitig aber weniger anfällig für gewählte Hyperparameter wie etwa Random Forests oder XGBoost-Modelle. Einfache Modelle wie GAMs sollten ausreichen, um sich einen allgemeinen Überblick über die Vorhersagegenauigkeit der einzelnen Merkmale zu verschaffen.

Wegen der eingeschränkten Verfügbarkeit großer Datensätze an Positions/Tracking-Daten war es notwendig, eine Packing-Version auf Basis von StatsBomb 360-Daten zu entwickeln. Dadurch berücksichtigen wir den Standort eines jeden Spielers im Ausschnitt zum Zeitpunkt der jeweiligen Aktion.

Der Datensatz enthält alle erfolgreichen Pässe aus dem Spiel heraus, mit StatsBomb 360-Ausschnitten der großen fünf europäischen Ligen aus der Saison 2020/2021 und etwa 1,1 Millionen Aktionen. Wir haben unterschiedliche Kombinationen an Merkmalen verwendet, entweder aus der Aktion oder den 360-Daten abgeleitet. Das erlaubte es uns, den kumulativen xG des Ballbesitzes zu berechnen, in den sich die einzelnen Pässe einfügen. Die Leistungswerte des Modells haben wir im System 100-mal verarbeitet. Dabei haben wir anhand des Ballbesitz-xG binäre Torergebnisse als Stichproben genommen. Bei mittleren Standardabweichungen von ± 2 über 100 Ergebnisse bei wiederholter Probenname erreichten wir ein Konfidenzintervall von 95 Prozent hinsichtlich der Vorhersagegenauigkeit eines Satzes an Merkmalen.

Ein wichtiger Hinweis: Die Videodaten enthielten nur einen Platzausschnitt, wodurch von Zeit zu Zeit nicht alle 22 Spieler in den 360-Daten Berücksichtigung fanden. Der Ansatz ist also nicht komplett gleichwertig zu einem Vergleich mit kompletten Packing-Daten. Die Packing-Zahlen auf Basis der 360-Daten weisen dennoch eine starke Korrelation zum vollständigen Packing-Verfahren auf. Der Vergleich liefert also immer noch aussagekräftige Ergebnisse zur relativen Vorhersagegenauigkeit der Faktoren, die dem OBV und dem Packing unterliegen – selbst, wenn die Ergebnisse nicht als absolut interpretiert werden dürfen.

Ergebnisse und Diskussion

Tabelle 1: Modell Performance-Metriken.

Modellbeschreibung	ROC AUC (höher ist besser)	Brier-Score (niedriger ist besser)
<i>Vorhersagekraft von Vorher-Nachher-Darstellungen</i>		
Nur Start- und Endpositionen (Hauptfaktor für OBV und simple Metriken für progressive Bälle)	0.638 \pm 0.004	0.0194 \pm 0.0002
Gegner näher am Tor als der Ball davor und danach	0.583 \pm 0.004	0.0196 \pm 0.0002
<i>Vorhersagekraft bei gegebenem Ausgangspunkt</i>		
Nur Ausgangspunkt	0.625 \pm 0.004	0.0195 \pm 0.0002
Ausgangspunkt & OBV für	0.639 \pm 0.004	0.0193 \pm 0.0002

Ausgangspunkt & Anzahl überspielter Spieler	0.634 ± 0.004	0.0194 ± 0.0002
Ausgangspunkt & Anzahl der überspielten Spieler & OBV für	0.640 ± 0.004	0.0193 ± 0.0002
<i>Nur Vorhersagekraft der Aktionswerte</i>		
OBV für	0.583 ± 0.004	0.0193 ± 0.0002
Anzahl der überspielten Spieler	0.556 ± 0.003	0.0196 ± 0.0002
<i>Referenzmodelle</i>		
Wert des Ballbesitz im Endpunkt	0.646 ± 0.004	0.0193 ± 0.0002
Wert des Ballbesitz im Endpunkt & Anzahl der Gegner hinter dem Ball danach	0.647 ± 0.004	0.0192 ± 0.0002

Der ROC AUC zeigt, wie gut ein Modell positive von negativen Fällen unterscheiden kann (in diesem Fall Torerfolg oder ausbleibender Torerfolg nach einer Ballbesitzphase). Bei einem ROC AUC von 0,5 wird jedem Pass dieselbe Wahrscheinlichkeit zugeschrieben, dass im weiteren Verlauf ein Tor folgt. Ein ROC AUC von 1 entspricht einer perfekten Anordnung der Vorhersagen. (Das heißt: Die Vorhersagewahrscheinlichkeit ist bei Pässen, die zum Treffer führen, immer höher, als bei den Pässen ohne anschließendem Torerfolg.)

Modelle, die Daten zum Standort der Aktion mit einschließen, haben eine höhere Vorhersagegenauigkeit. Das deutet darauf hin, dass der Standort einer Aktion hinsichtlich der Chancen eines Torerfolgs die Vorhersagegenauigkeit progressiver Ereignisse am stärksten beeinflusst. Wenn wir nur Ausgangs- und Endpunkte verwenden, ist die Vorhersagegenauigkeit geringer als beim kompletten OBV-Modell, das auch den weiteren Spielkontext mit einbezieht.

Ähnliche Vorhersagegenauigkeiten liefern bei Berücksichtigung des Pass-Ausgangspunktes der OBV des Passes und die Zahl der nach dem Packing-Prinzip überspielten Gegner.

Wenn wir die nach dem Packing-Prinzip überspielten Gegner beim OBV-Modell mit berücksichtigen, ist die dadurch gewonnene Vorhersagegenauigkeit relativ gering. Das deutet darauf hin, dass sowohl der OBV als auch das Packing ähnlich genau angeben, ob ein Treffer folgt oder nicht. Das bringt uns zur Vermutung, dass der OBV bereits einen

Großteil der Vorhersage-Ergebnisse enthält, zu denen wir mit zusätzlichen Packing-Daten auf 360-Basis gelangen würden.

Dieses Ergebnis scheint sich mit vorherigen Arbeiten zu decken, bei denen andere Datensätze zum Einsatz kamen. [Bobby Gardiner stellte fest](#), dass eine starke Korrelation zwischen dem Packing und den progressiven Pässen besteht. [Kuba Michalczyk](#) kam indes zu dem Schluss, dass reine Ereignisdaten bereits ein sehr präziser binärer Indikator für Pässe durch die Abwehrreihe sind.

Die Zahl der per Packing überspielten Gegner (auf 360-Basis) allein genügt, um eine Aussage zur Trefferwahrscheinlichkeit zu machen. Die Vorhersagegenauigkeit ist aber nicht so hoch wie beim OBV.

Mit der Einschränkung, dass unsere Packing-Daten nicht aus vollständigen Tracking-Daten stammen, können wir feststellen: Das Packing auf 360-Basis hat für sich genommen eine geringere Aussagekraft als positions- oder OBV-basierte Merkmale, wenn es darum geht, Treffer am Ende einer Ballbesitzphase vorherzusagen.

Schlussfolgerung

Wir haben eine Reihe qualitativer und quantitativer Argumente dazu vorgebracht, wie wir den Ballvortrag und dessen Nutzen fürs Team bewerten können.

Wir sind der Meinung: Bei der Bewertung des Ballvortrags ist es sinnvoll, wenn wir uns direkt ansehen, wie sich jede Aktion (voraussichtlich) auf die Trefferwahrscheinlichkeit auswirkt. Weiterhin haben wir festgestellt, dass Packing-Zahlen (auf 360-Basis) für sich genommen nicht so aussagekräftig für die nachfolgende Trefferwahrscheinlichkeit sind wie allgemein angenommen. Unterm Strich bietet das Packing keine großen Vorteile im Vergleich zu ereignisbasierten Ansätzen. (Zum selben Ergebnis sind auch schon frühere Untersuchungen zum Thema gelangt.)

Positionsdaten liefern durchaus wertvolle Zusatzinformationen. Das Packing allein greift aber wahrscheinlich zu kurz, wenn es darum geht, das Potenzial dieser Daten auszuschöpfen.